THE
EMBO
JOURNAL

Manuscript EMBO-2009-70981

Distinct donor and acceptor specificities of Trypanosoma brucei oligosaccharyltransferases

Luis izquierdo, Benjamin Schulz, Joao Rodrigues, Maria Lucia Guther, James Procter, Markus Aebi

Corresponding author: Michael Ferguson, Dundee, University of

Review timeline:

Submission date:	30 March 2009
Editorial Decision:	24 April 2009
Revision received:	27 May 2009
Editorial Decision:	10 June 2009
Revision received:	17 June 2009
Accepted:	18 June 2009

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 24 April 2009

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are enclosed. As you will see, all three reviewers show significant interest in your work and appreciate the high quality of the data presented. However, all three - and referee 3 in particular - raise a number of concerns that would first need to be addressed before we could consider publication of your manuscript. In particular, referee 3 is not fully convinced by your evidence that the different STT3 isoforms have different acceptor specificities, primarily based on an inaccurate calculation of isoelectric points. Clearly, it is critical that this issue is resolved. In addition, I would draw your attention to the comments of referee 1 regarding the need to make your manuscript more accessible to non-specialists, and to consider alternative explanations for the data you present.

In the light of the referees' positive recommendations, I would therefore like to invite you to submit a revised version of the manuscript, addressing all the comments of all three reviewers (particularly, but not restricted to, those highlighted above). I should add that it is EMBO Journal policy to allow only a single round of revision. Therefore, acceptance of your paper will depend on your ability to fully answer the points raised by the referees.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE REPORTS

Referee #1 (Remarks to the Author):

This paper is a major investigation of glycosylation specificity at the deep level in trypanosomes, and as such, is at a very high impact. Trypanosomes represent the best model outside of classical systems for such studies, and the work is thorough, well documented and overall convincing.
My major issue is with the writing - this is very dense and compact, and there is a major assumption that the reader is on the same wavelength as the authors - this is not always true and this reviewer found that in parts she had to dig to fully appreciate the context. Further, there is a rather inflexible current running through the piece, and the possibility of alternate explanations does not seem to be well entertained, considered or mentioned in many places. This is important as EMBO is a general journal and the work needs to be made more accessible.

The potential that the STT specificity is modified by heterologous expression is not discussed - the peptide analysis in yeast is very smart and detailed, but the system there is very different, including the absence of complex class processing, the overall environment in the ER and major differences in growth rates and protein turnovers for example when compared against trypanosomes. The analysis therefore ought to be qualified a little.

The possibility that differential glycosylation is also environmental needs to be considered - i.e. distance from membrane, influenced by interactions with chaperones, and also quaternary structure?

A dot is used as opposed to delta for ko, which is confusing. Also, there is a lot of data in the yeast section - what is point of the FOA for example in the present context? Explain a little more for accessibility.

I am not sure I totally buy the argument that the differences in the glycopeptide profiles cannot be due to expression levels - presumably there is a detection limit here and a cutoff that is used for the MS analysis, which is of course appropriate, but the over-expressed trypanosome STTs could be acting more efficiently due to greater activity in general. This is not necessarily the same as total protein level.

Explain reason for doing the rnai in a heterozygote - is this for increased efficiency or other reasons?

Need to make clear that the biochemical analysis by endoH is on new VSG and the MS is on bulk material.

Could the data not also indicate that the dual oligoman form is processed, or could be processed to the mature form (wild type is not really correct here as it is a genetic term; it is VSG as made in a wold type trypanosome). the statement as read suggests that the dual oligoman form is somehow excluded from further action, and possibly degraded. Is this what the authors intend, and if so, is there other evidence to support this or is it a new model at this point?

I do not find the differences to be very clear between 3C and 3D - I may be missing something but overall these do not seem so radically different. Perhaps better annotation here of the mw involved or rearranging the panels so that the corresponding surface and ER pools are adjacent would help?

In addition, the material that is analysed here is ghost and sVSG rather than, sensu stricto, ER and surface.

STTA:The authors need to also entertain the possibility here that the phenotype is secondary - for example a GTase or processing enzyme that is critical for progression to complex glycans may have been strongly affected by the altered glycosylation environment, and leads to the very huge difference inthe overall profile of glycans that are detected. I accept totally that the authors explanation is reasonable, but there are other possibilities.

STTB: Again, I agree that the proposal is a good one, but I also think the same possibility, that an additional factor that is critical, could have been very strongly affected. I am not that comfortable with the idea of such exclusive function, as this is rare in such complex cases where there are so many players.

Is the pI correlation also supported by work in other organisms?

Mouse data: Why no kinetic count here for growth in mice? What time point is the count and why? Is there recovery of growth for example consistent with the rnai machinery breaking? I think this is important as the analysis presented indicates that the cells have reverted, but there is the trivial explanation that something else has affected growth here, rather than the rnai itself and the authors are just looking at revertants throughout the experiment.

No validation of the knockdown - qRTPCR? This is a critical control.

It is not true that this is the first example of essentiality in vivo and not in vitro. Why is this an issue anyway?

Referee #2 (Remarks to the Author):

This is an extremely complete and thorough study of role played by different STT3 proteins of T. Brucei and tells us a lot about what they actually do. The work is of a very high technical standard and a large amount of compelling and well controlled data is presented. A key finding is the difference in the N-linked glycan that is transferred to substrate proteins by different STT3s. The context of the N-glycosylation site in influencing subsequent modification is also very clearly shown. The work is all pulled together in a very complementary and biologically relevant fashion culminating in the use of a mouse model for T. Brucei infection.

Minor comments.

1. The "deltas" have been lost throughout the text in relation to STT3 deleted yeast strains.
2. What is the definition of an efficient or poor glycosylation site as used for Figure 1?
3. The basis for identifying specific oligosaccharides could be better stated in the legend to Figure 2.
4. There seems to be something odd about Figures 3C and 4C in that the mobilities of the TbSTT3A,B,C and TbSTT3B RNAi products (-Tet) are not quite identical (assuming the samples were run on the same gel). Fig. 3C, lanes 1 and 7 should be identical but are not? Likewise Fig. 4C, lanes 1 and 7 should be that same but have apparently different mobility. The EndoH resistance of the doubly glycosylated form of ER associated VSG is also not convincing (Fig. 4C, lane 5).
5. The proportion of EndoH sensitive and insensitive VSG in siTbSTT3A cells cannot be readilly estimated from the image since provided it the product is overexposed (Fig. 3C).
6. What is the consequence of siTbSTT3C on VSG processing?
7. Supplementary Fig. S1 could be better labelled.
8. Page 5 refers to the "in vivo" analysis of OST function by RNAi. On page 8, this analysis becomes "in vitro". These terms can become confusing when dealing with different systems and organisms in the same manuscript. It may be better to avoid their use unless absolutely necessary. Alternatively please be entirely consistent.

Referee #3 (Remarks to the Author):

The manuscript from Izquierdo et al addresses the substrate specificity of the multiple STT3 homologues in T. brucei. The authors have discovered that 2 of 3 STT3 genes are expressed in vitro, and they present strong evidence that the two STT3 proteins use different donor oligosaccharides. The also propose that the two STT3 proteins recognize acceptor sites with different properties. The evidence here is less convincing due to errors in the calculation of protein isoelectric points. After the manuscript is revised to improve clarity and to correct the errors, it could be reconsidered.

1. The authors suggest that the difference in ER (Fig. 3C) and cell surface (Fig. 3D) glycan composition of VSG221 in STT3ARNAi (+Tet) Fig. 3D is explained by more efficient cell surface expression of VSG221 that has a complex glycan on N263. I am puzzled by the low ion intensity of N263 glycopeptides that carry the high mannose glcyan in the STT3ARNAi (+Tet) cells (Table SIV). Are glycopeptides in Table SIV and Fig. S3 derived from total VSG221 or cell surface VSG221? If cell surface VSG221 was used for all glycan analysis (Fig. 2, Table S3, Table S4, Fig. S3), the authors should clearly state this. If this is the case, glycan analysis utilized a pool that was enriched for he wild type rather than mutant glycan structures in the TbSTT3A RNAi.

2. I'm not sure how the authors calculated isoelectric points below 2.0 for a 14 residue peptide, since this is simply impossible. Since these peptides are embedded in a protein, the calculation should only be dependent upon the side chains (Nα-amino group on residue 1, no C-terminal carboxyl group on residue 13). Side chains (DERKHC and Y) are the only contributing factors. The most acidic side chains (D and E have pKs of ~4). Poly-aspartate (~50 D residues) has a calculated isoelectric point of about 2.5, while polyarginine has a calculated isoelectric point of 13.5, so this defines the possible range of IEPs for peptides. For this reason I am extremely skeptical that the mean pI of Tb. STT3A sites is 2.5, since that corresponds to the poly-acidic extreme of the possible range of IEP. I also don't see how one can have a variance in isoelectric point (Table SV, variance for + category =14.1) that exceeds the possible range of peptide isoelectric points. I looked at several of the peptides in the PRIDE depository, and found that the authors calculated isoelectric points are off by as much as 3 pH units. For example, a 13 residue glycopeptide (DVGALNDTAVLSE) in the file Tb05030H13.470 is listed as having an isoelectric point of 0!!, when in fact it is closer to 3.2-3.3.

3. The authors present evidence that a single TbSTT3 is sufficient in vitro (Fig. 3B and 4B) and indicate that at least one STT3 is essential based upon Fig. S1F. It should be noted that the cells seem to be adapting to the double knockdown after about 120 hours. Perhaps this is due to inactivation of the RNAi construct, but this possibility has not been tested. The authors also present convincing evidence that both STT3s are required for efficient infection of animals. In this case they suggest that the surviving cells (<10% of wild type) are explained by a reduced effect of the RNAi. This hypothesis is not well supported by the digestion data (Fig. 6C and 6D). In the case of the STT3A RNAi, the EndoH sensitivity of cell surface VSG211 looks remarkably similar to that shown in Fig. 3D for the in vitro experiment (where there is no selection pressure to lose the knockdown construct). In the case of the STT3B RNAi (Fig. 6D), the experiment is somewhat flawed by the observation that -Dox sample resembles the +Dox sample, rather than the wild type. One wonders whether this is an experimental error. Secondly, the possibility that there are EndoH sensitive glycans in Fig. 6D (+Dox) would need to be confirmed by ES-MS analysis. Rather than embark upon such an analysis, I suggest that the authors remove the suggestion that the surviving cells are escaping the RNAi, since they have not tested this by RT-PCR. It seems more likely that the TbSTT3s do not have 100% specificity for the donor oligosaccharide and acceptor sites, and will transfer a non-preferred donor (GlcNAc2Man5 for STT3B) at a certain frequency to the non-preferred site (N263 for TbSTT3B).

4. The authors would like to conclude that TbSTT3C and TbSTT3B also have different acceptor site specificities based upon the data presented in Fig. 1. It would be nice if the authors could use the same size window (4 residues on each side of NXT/S in Fig. 1, and 5 residues on each side in Fig.5) to consider the impact of adjacent residues on glycosylation efficiency. It would be preferable if the method of peptide comparison was the same in both cases (correctly calculated isoelectric point). The weakness of the Fig. 1D method is that the reader can deduce very little about the net charge of the sequences flanking poorly and efficiently modified glycosylation sites. In the case of STT3C, the number of efficiently used sites appears to be 12, while inefficient sites number 17, so any conclusions are somewhat blunted by the small number of peptides. This reviewer also wonders whether the method of analysis (use of cell surface protein) biases against the detection of certain hypoglycosylated variants, if lack of a glycan at a certain site has a negative impact on protein folding in the ER, a phenomena which is not unheard of. Efficient and inefficient sites both contain 0-3 acidic residues within this region, but it is not clear whether the poorly glycosylated outliers (2 or 3 acidic residues) have neutral PI's due to the presence of one or more basic residues. I can't help wondering whether there is a more effective and convincing way to present the data than box diagrams. For example can the authors plot glycosylation efficiency vs IEP for STT3C? The observation that the TOS1_417 site is glycosylated, albeit not with 100% efficiency, by Tb SXTT3B, but not by yeast OST is interesting. There is some statistical evidence from B. Imperiali's lab that lysine at +4 relative to N in the sequon is underrepresented in sequon databases.

Supplemental Section

1. Table S1.

(a) An asparagine in the glycopeptide CCW14_87 (m/z = 1105.4781) is marked by an asterisk, presumably indicating that this site is glycosylated. This seems to be the only asparagine marked by an asterisk; N residues in other glycopeptides appear to be in a bold font. Is there some reason that this peptide is marked differently?

(b) The significance of the Rank column is unclear since all peptides are entries are ranked 1.

© European Molecular Biology Organization

(c) #Previously identified15. The meaning of the superscript 15 is not clear.
(d) The authors should add a sentence beneath the table to define the score column. Why does the peptide GAS1_253* have a score of (68) (i.e. score within parentheses)
(e) The expect column should also be defined. It is not clear why the exponent is capitalized in some cases (i.e. 8.4E-04), but not in others (i.e. 8.2e-08).
(f) The glycopeptide EXG2_50 should have an asterisk after it.

2. Figure S3 and Table SIV. Each of these supplemental items contains information on the glycan structure of pronase peptides derived from VSG221 produced by T. brucei cells that express TbSTT3A, TbSTT3B or TbSTT3A+TbSTT3B. It is not clear to me why some intensity data presented in Figure S3 is in conflict with tabulated data in Table SIV. As an example, consider FigS3, panel A, N263 glycopeptide NET (Hex3HexNAc2). Figure S3 shows a 60% intensity value for this glycopeptide in both the -Tet and +Tet conditions. However, ion intensity values in Table SIV for this peptide ({plus minus}Tet in STTA siRNA) are 439.4 (-Tet) and 173.4 (+Tet). The S3 figure legend does not explain how % intensity values were calculated, but it seems likely that the Table SIV values were normalized to the intensity of another glycopeptide. At best, Fig. S3A is confusing, because it appears that siRNA knockdown of STT3A does not reduce transfer of the GlcNAc2Man5 to the N263, but instead increases transfer of GclcNAc2Man9 to the N428 site.

3. Figure S4. This figure could be expanded to show the source of glycoproteins (total glycoprotiens or cell surface glycoproteins?) and to indicate whether the samples used for digestion were enriched by sequential lectin chromatography (ConA and then ricin) or split sample chromatography (ConA or ricin).

Minor points:

1. The Schulz and Aebi reference is listed twice in the References list, once in 2008 (no Vol and page #) and once in 2009.
2. Legend Fig. 1(C), the sentence "Data is shown in Table S1" should read "Data is shown in Table S2. "
3. Define symbols used in in Fig. 2 (GlcNac, Man, etc.)
4. The manuscript should mention that results of the proteomics analysis of T. brucei glycopeptides is available in the PRIDE database (list accession #). Hopefully they will correct the annotated IEPs.

1st Revision - authors' response 27 May 2009

Referee #1 (Remarks to the Author):
This paper is a major investigation of glycosylation specificity at the deep level in trypanosomes, and as such, is at a very high impact. Trypanosomes represent the best model outside of classical systems for such studies, and the work is thorough, well documented and overall convincing. My major issue is with the writing - this is very dense and compact, and there is a major assumption that the reader is on the same wavelength as the authors - this is not always true and this reviewer found that in parts she had to dig to fully appreciate the context. Further, there is a rather inflexible current running through the piece, and the possibility of alternate explanations does not seem to be well entertained, considered or mentioned in many places. This is important as EMBO is a general journal and the work needs to be made more accessible.

>>>> We accept this criticism and have written a more expansive introduction and discussion, including providing alternative explanations, to address it.

The potential that the STT specificity is modified by heterologous expression is not discussed - the peptide analysis in yeast is very smart and detailed, but the system there is very different, including the absence of complex class processing, the overall environment in the ER and major differences in growth rates and protein turnovers for example when compared against trypanosomes. The analysis

therefore ought to be qualified a little.

>>>> We accept this criticism. The heterologous expression system using yeast has been used successfully before to analyze the activity of T.cruzi and Leishmania STT3s (Castro et al. 2006; Nasab et al. 2008; Hese et al. 2009) and this is now stated clearly in the revised introduction. In addition, a statement acknowledging the potential limitation raised by the reviewer is now made in the discussion as follows: "Although we can not formally exclude the possibility that heterologous expression of TbSTT3s in yeast might affect their specificities, or that STT3 protein levels necessarily correlate directly with enzymatic activity, the heterologous complementation of Stt3p mutant yeast strains with STT3 proteins from protozoan organisms appears, thus far, to be a good experimental system to study STT3 function and specificity (Shams-Eldin et al. 2005; Castro et al. 2006; Nasab et al. 2008; Hese et al. 2009).

The possibility that differential glycosylation is also environmental needs to be considered - i.e. distance from membrane, influenced by interactions with chaperones, and also quaternary structure?

>>>> We accept this criticism and have incorporated the following statement in the discussion: "While we conclude that local pI around the glycosylation sequon correlates with sequon usage by the different TbSTT3s, we acknowledge that other factors, such as position in the polypeptide and, thus, local secondary and tertiary structure are also likely to play significant roles in differential N-glycosylation in T.brucei. These aspects are currently under investigation."

A dot is used as opposed to delta for ko, which is confusing. Also, there is a lot of data in the yeast section - what is point of the FOA for example in the present context?
Explain a little more for accessibility.

>>>> The conversion of deltas to dots was a pdf conversion artefact - we apologise to the reviewers for not picking it up. The FOA experiment indicates the complementation of the lethality of yeast stt3 cells by TbSTT3. The following sentence has been added to the 'Yeast manipulation' section in Material and Methods: "Addition of 5-FOA to the media allowed for selection of cells which had lost the URA plasmid encoding the wild type yeast STT3. These cells could only survive if expression of a plasmid-borne T. brucei STT3 paralogue complemented the lack of yeast STT3."

I am not sure I totally buy the argument that the differences in the glycopeptide profiles cannot be due to expression levels - presumably there is a detection limit here and a cutoff that is used for the MS analysis, which is of course appropriate, but the over-expressed trypanosome STTs could be acting more efficiently due to greater activity in general. This is not necessarily the same as total protein level.

>>>>We absolutely agree that the TbSTT3p paralogues are expressed at a much higher level than wild type yeast OTase, and that this difference in dosage probably effects the glycosylation efficiency of certain sites. This increased dosage effect could explain the partial glycosylation of some sites that are never observed to be glycosylated by yeast OTase (TOS1_417 and CCW14_87). However, it is not at all expected that the glycosylation of different sites is increased with TbSTT3Bp (TOS1_417) and TbSTT3Cp (CCW14_87). This implies that factors in addition to dosage levels influence the site-specific activites of these enzymes. For clarity, we have incorporated a statement to that effect in the discussion (the underlined section within the following): "Although we can not formally exclude the possibility that heterologous expression of TbSTT3s in yeast might affect their specificities, or that STT3 protein levels necessarily correlate directly with enzymatic activity, the heterologous complementation of Stt3p mutant yeast strains with STT3 proteins from protozoan organisms appears, thus far, to be a good experimental system to study STT3 function and specificity (Shams-Eldin et al. 2005; Castro et al. 2006; Nasab et al. 2008; Hese et al. 2009).

Explain reason for doing the rnai in a heterozygote - is this for increased efficiency or other reasons?

>>>> Yes, it was to increase RNAi efficiency. The following has been added to the relevant section

of the Results: "This deletion did not affect the glycosylation profile of the VSG221 reporter glycoprotein (Fig. 2A and B) but it provided us with a heterozygote cell line where subsequent selective RNAi knockdowns of TbSTT3s would be maximised."

Need to make clear that the biochemical analysis by endoH is on new VSG and the MS is on bulk material.

>>>> The following new text now makes this clear: "Cell surface VSG can be purified in a soluble form (sVSG221) after osmotic shock, a process that releases VSG from the parasite surface by the action of endogenous GPI-specific phospholipase C that cleaves the dimyristoylglycerol lipid component of the VSG GPI anchor (Cross, 1975, 1984; Ferguson et al, 1985). This cell surface-derived sVSG221 is amenable to glycoform analysis as an intact glycoprotein by ES-MS and as Pronase glycopeptides by ES-MS and ES-MS/MS (Jones et al, 2005; Urbaniak et al, 2006; Manthri et al, 2008; Stokes et al 2008). In addition, the N-glycosylation status of both cell surface-derived sVSG221 and of newly-synthesized (presumably ER-associated) VSG221, i.e., that which is not released by osmotic shock (Ferguson et al 1985), can be assessed by EndoH and PNGaseF digestion and analysis by SDSPAGE and Coomassie blue staining or Western blot, respectively."

Could the data not also indicate that the dual oligoman form is processed, or could be processed to the mature form (wild type is not really correct here as it is a genetic term; it is VSG as made in a wild type trypanosome). The statement as read suggests that the dual oligoman form is somehow excluded from further action, and possibly degraded. Is this what the authors intend, and if so, is there other evidence to support this or is it a new model at this point?

>>>> The reviewer is correct to point out this out. The revised and extended version explains the situation much more clearly and explains the data in our previous paper, derived from both bioinformatics analyses and experimental analyses using alphamannosidease inhibitors, that indicate that: "As noted previously (Manthri et al. 2008), the strict demarcation between these different routes to complex and oligomannose N-glycans appears to be due to the absence of the Golgi mannosidase II activity that permits the conversion of triantennary Man5GlcNAc2 to Man3GlcNAc2 in other eukaryotes." This statement is now in the discussion.

I do not find the differences to be very clear between 3C and 3D - I may be missing something but overall these do not seem so radically different. Perhaps better annotation here of the mw involved or rearranging the panels so that the corresponding surface and ER pools are adjacent would help?

We understand the reviewer's point but this is exactly why the data in Fig. 3D are also obtained backed up by ES-MS of the whole VSG and ES-MS and ES-MS/MS of their Pronase glycopeptides.

In addition, the material that is analysed here is ghost and sVSG rather than, sensu stricto, ER and surface.

>>>> We will stick to the existing descriptions but have added the following explanation in the text: "Cell surface VSG can be purified in a soluble form (sVSG221) after osmotic shock, a process that releases VSG from the parasite surface by the action of endogenous GPI-specific phospholipase C that cleaves the dimyristoylglycerol lipid component of the VSG GPI anchor (Cross, 1975, 1984; Ferguson et al, 1985). This cell surface-derived sVSG221 is amenable to glycoform analysis as an intact glycoprotein by ES-MS and as Pronase glycopeptides by ES-MS and ES-MS/MS (Jones et al, 2005; Urbaniak et al, 2006; Manthri et al, 2008; Stokes et al 2008). In addition, the N-glycosylation status of both cell surface-derived sVSG221 and of newly-synthesized (presumably ER-associated) VSG221, i.e., that which is not released by osmotic shock (Ferguson et al 1985), can be assessed by EndoH and PNGaseF digestion and analysis by SDS-PAGE and Coomassie blue staining or Western blot, respectively."

STTA:The authors need to also entertain the possibility here that the phenotype is secondary - for example a GTase or processing enzyme that is critical for progression to complex glycans may have

been strongly affected by the altered glycosylation environment, and leads to the very huge difference inthe overall profile of glycans that are detected. I accept totally that the authors explanation is reasonable, but there are other possibilities.
STTB: Again, I agree that the proposal is a good one, but I also think the same possibility, that an additional factor that is critical, could have been very strongly affected. I am not that comfortable with the idea of such exclusive function, as this is rare in such complex cases where there are so many players.

>>>> *We think that the revised version provides more balanced arguments than the original and we thank the reviewer for asking us to do this.*

Mouse data: Why no kinetic count here for growth in mice? What time point is the count and why? Is there recovery of growth for example consistent with the rnai machinery breaking? I think this is important as the analysis presented indicates that the cells have reverted, but there is the trivial explanation that something else has affected growth here, rather than the rnai itself and the authors are just looking at revertants throughout the experiment.

>>>> *As I am sure the reviewer appreciates, there are many ways to perform in vivo gene essentiality for T.brucei and that there is some debate about the "best" protocol. The protocol we favour because, in our opinion, of its simplicity and clarity is that performed here ñ i.e., we perform an acute infection in groups of 5 mice and measure parasitaemia 3 days later (apologies for not stating this previously). Thus, the control animals show robust parasitaemias (mean levels around 5 x 108 parasites per ml of blood) with which to compare the doxycline-induced knockdown parasitaemia. Our experience with a number of cell lines has shown that measurement before 3 days will give unmeasurably low parasitaemias if the gene is essential while taking measurements beyond 3 days simply selects for the inevitable RNAi escape mutants*.*
At the 3 day time point, RNAi knockdown of essential genes gives the result seen here, i.e., (a) an obvious reduction on parasitaemia (to <0.5 x 108 per ml) if the gene is essential and (b) sufficient cells to do some biochemical phenotyping (in this case the Western blots ± glycosidase treatments) to ascertain whether these few surviving cells are tending to wild-type biochemistry through incomplete knockdown and/or escape "reversion". This, in itself, is a further indication of gene essentiality.

**see, for example: Chen Y, Hung CH, Burderer T, Lee GS. (2003) Development of RNA interference revertants in Trypanosoma brucei cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. Mol Biochem Parasitol. (2003) Feb;126(2):275-9.*

No validation of the knockdown - qRTPCR? This is a critical control.

>>>> *We believe that Northern blot, as well as qRTPCR, is a good validation assay for the knockdowns of TbSTT33A and TbSTT3B (Fig. 3A and 4A).*

It is not true that this is the first example of essentiality in vivo and not in vitro. Why is this an issue anyway?

>>>> *We have removed the claim about it not being the first such example and we have reduced and modified the discussion to simply read: "This result emphasises that gene functionality should be tested in vivo, i.e., in an animal model of infection, before making final conclusions about gene essentiality."*

Referee #2 (Remarks to the Author):

This is an extremely complete and thorough study of role played by different STT3 proteins of T. Brucei and tells us a lot about what they actually do. The work is of a very high technical standard and a large amount of compelling and well controlled data is presented. A key finding is the difference in the N-linked glycan that is transferred to substrate proteins by different STT3s. The context of the Nglycosylation site in influencing subsequent modification is also very clearly shown.

© European Molecular Biology Organization

The work is all pulled together in a very complementary and biologically relevant fashion culminating in the use of a mouse model for T. Brucei infection.

Minor comments.

1. The "deltas" have been lost throughout the text in relation to STT3 deleted yeast strains.

>>>> The conversion of deltas to dots was a pdf conversion artefact - we apologise to the reviewers for not picking it up.

2. What is the definition of an efficient or poor glycosylation site as used for Figure 1?

>>>> In material and methods, just at the end of 'Site-specific glycosylation occupancy determination' we have added the sentence: "Glycosylation sites with more than 95% occupancy were defined as being "efficiently" glycosylated."

3. The basis for identifying specific oligosaccharides could be better stated in the legend to Figure 2.

>>>> We agree and have re-written the figure legend as follows:

"Fig. 2. Mass spectrometric analyses of intact mature sVSG221 glycoforms before and after selective knockdown of TbSTT3A and TbSTT3B expression. Samples of sVSG221 were analyzed by ES-MS. The spectra show the masses of the various glycoforms of the intact mature sVSG221 molecules. The inset cartoons represent our interpretation of those glycoform masses in terms of the ranges of Nglycans present at each of the two N-glycosylation sites. These assignments combine additional data from the ES-MS and ES-MS/MS analyses of Pronase glycopeptides from the same sVSG221 preparations (Fig. S3 and Table SIV). In the inset cartoons, endo-H-resistant N-glycans are in red and endo-H-sensitive N-glycans are in blue. Filled circles and squares (red or blue) represent mannose and N-acetylglucosmaine residues, respectively, and open circles represent galactose residues. "

4. There seems to be something odd about Figures 3C and 4C in that the mobilities of the TbSTT3A,B,C and TbSTT3B RNAi products (-Tet) are not quite identical (assuming the samples were run on the same gel). Fig. 3C, lanes 1 and 7 should be identical but are not? Likewise Fig. 4C, lanes 1 and 7 should be that same but have apparently different mobility. The EndoH resistance of the doubly glycosylated form of ER associated VSG is also not convincing (Fig. 4C, lane 5).

>>>> The samples for each panel were run on the same gel. We have rotated the image in Fig 4C by minus 3° to compensate for gel "smiling". The other differences (eg. Fig. 3C lanes 1 and 7) are due to slightly different loadings and/or transference during Western blotting. This is the best we can do, bearing in mind we are looking at the differences in migration of 50 kDa glycoproteins ± 1.5 kDa glycans. For information, the images shown are representative of several (>4) independent experiments.

5. The proportion of EndoH sensitive and insensitive VSG in siTbSTT3A cells cannot be readilly estimated from the image since provided it the product is overexposed (Fig. 3C).

>>>> We do not agree. Eg. all of the signal shifts in Fig 3C lane 5 even though the blot is well-developed. If it was underdeveloped perhaps some percentage might not have shifted and we would not realised it.

6. What is the consequence of siTbSTT3C on VSG processing?

>>>> We would certainly have liked to address this as well but since we did not find TbSTT3C expression at the mRNA level in either bloodstream or procyclic forms of the parasite, we could not

© European Molecular Biology Organization

assess the effect of its lack in the T. brucei cells. In the discussion we suggest that this protein might be playing a role in other T. brucei life stages that cannot be cultured in the lab. In addition it would be technically very difficult to try to specifically shut down TbSTT3B and C, since they are almost identical.

7. Supplementary Fig. S1 could be better labelled.

>>>> We agree and this has been done and the figure legend improved as well.

8. Page 5 refers to the "in vivo" analysis of OST function by RNAi. On page 8, this nalysis becomes "in vitro". These terms can become confusing when dealing with different systems and organisms in the same manuscript. It may be better to avoid their use unless absolutely necessary. Alternatively please be entirely consistent.

>>>> This is a good point and we have replaced "in vivo" with "in trypanosomes" and "in vitro" with "in culture" whenever these are the meanings we wish to convey.

Referee #3 (Remarks to the Author):

The manuscript from Izquierdo et al addresses the substrate specificity of the multiple STT3 homologues in T. brucei. The authors have discovered that 2 of 3 STT3 genes are expressed in vitro, and they present strong evidence that the two STT3 proteins use different donor oligosaccharides. The also propose that the two STT3 proteins recognize acceptor sites with different properties. The evidence here is less convincing due to errors in the calculation of protein isoelectric points. After the manuscript is revised to improve clarity and to correct the errors, it could be reconsidered.

1. The authors suggest that the difference in ER (Fig. 3C) and cell surface (Fig. 3D) glycan composition of VSG221 in STT3ARNAi (+Tet) Fig. 3D is explained by more efficient cell surface expression of VSG221 that has a complex glycan on N263. I am puzzled by the low ion intensity of N263 glycopeptides that carry the high mannose glcyan in the STT3ARNAi (+Tet) cells (Table SIV). Are glycopeptides in Table SIV and Fig. S3 derived from total VSG221 or cell surface VSG221? If cell surface VSG221 was used for all glycan analysis (Fig. 2, Table S3, Table S4, Fig. S3), the authors should clearly state this. If this is the case, glycan analysis utilized a pool that was enriched for he wild type rather than mutant glycan structures in the TbSTT3A RNAi.

>>>> All MS glycopeptide analyses, i.e., those in Table SIV and Fig. S3 as well as in Fig. 2 and Table SIII, are derived from cell surface sVSG221 and this is more clearly stated now in the revised paper. This explains the low ion intensity of N263 highmannose carrying glycopeptides because the cell-surface pool used for the analysis is enriched for the wild type glycoform, as we state on page 7 of the manuscript: "The much greater proportion of modified versus wild type VSG221 in the newlysynthesised VSG221 versus mature sVSG221 (compare Fig. 3C and Fig. 3D) suggests that the small proportion of wild type VSG221 present in the ER is enriched during ER exit and/or transit to the cell surface, which is consistent with the importance of the correct glycosylation of the Asn263 site in VSG221 (Blum et al, 1993; Izquierdo et al, 2009)."

2. I'm not sure how the authors calculated isoelectric points below 2.0 for a 14 residue peptide, since this is simply impossible. Since these peptides are embedded in a protein, the calculation should only be dependent upon the side chains ((Noα-amino group on residue 1, no C-terminal carboxyl group on residue 13). Side chains (DERKHC and Y) are the only contributing factors. The most acidic side chains (D and E have pKs of ~4). Poly-aspartate (~50 D residues) has a calculated isoelectric point of about 2.5, while polyarginine has a calculated isoelectric point of 13.5, so this defines the possible range of IEPs for peptides. For this reason I am extremely skeptical that the mean pI of Tb. STT3A sites is 2.5, since that corresponds to the poly-acidic extreme of the possible range of IEP. I also don't see how one can have a variance in isoelectric point (Table SV, variance

for + category =14.1) that exceeds the possible range of peptide isoelectric points. I looked at several of the peptides in the PRIDE depository, and found that the authors calculated isoelectric points are off by as much as 3 pH units. For example, a 13 residue glycopeptide (DVGALNDTAVLSE) in the file Tb05030H13.470 is listed as having an isoelectric point of 0!!, when in fact it is closer to 3.2-3.3.

>>>> We are extremely grateful to the reviewer for pointing out this serious error. All pIs have been recalculated and the relevant Figures and Tables corrected.

3. The authors present evidence that a single TbSTT3 is sufficient in vitro (Fig. 3B and 4B) and indicate that at least one STT3 is essential based upon Fig. S1F. It should be noted that the cells seem to be adapting to the double knockdown after about 120 hours. Perhaps this is due to inactivation of the RNAi construct, but this possibility has not been tested. The authors also present convincing evidence that both STT3s are required for efficient infection of animals. In this case they suggest that the surviving cells (<10% of wild type) are explained by a reduced effect of the RNAi. This hypothesis is not well supported by the digestion data (Fig. 6C and 6D). In the case of the STT3A RNAi, the EndoH sensitivity of cell surface VSG211 looks remarkably similar to that shown in Fig. 3D for the in vitro experiment (where there is no selection pressure to lose the knockdown construct). In the case of the STT3B RNAi (Fig. 6D), the experiment is somewhat flawed by the observation that -Dox sample resembles the +Dox sample, rather than the wild type. One wonders whether this is an experimental error. Secondly, the possibility that there are EndoH sensitive glycans in Fig. 6D (+Dox) would need to be confirmed by ES-MS analysis. Rather than embark upon such an analysis, I suggest that the authors remove the suggestion that the surviving cells are escaping the RNAi, since they have not tested this by RT-PCR. It seems more likely that the TbSTT3s do not have 100% specificity for the donor oligosaccharide and acceptor sites, and will transfer a non-preferred donor (GlcNAc2Man5 for STT3B) at a certain frequency to the non-preferred site (N263 for TbSTT3B).

>>>> We thank the reviewer for these thoughtful comments. Regarding to the RNAi experiment of both STT3A+B proteins, Fig. S1F, the possibility of loss of RNAi effect is indeed well known in trypanosomes (see Chen Y, Hung CH, Burderer T, Lee GS. (2003) Development of RNA interference revertants in Trypanosoma brucei cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. Mol Biochem Parasitol. (2003) 126: 275-9.)

We also acknowledge the complexity and "imperfection" of Figure 6D (lanes 7-9) where it is clear that the minus Dox VSG is showing signs of partial RNAi effects, presumably due to the leakiness of the Tet regulation system (as observed previously by others, see Mol Biochem Parasitol. (2005) 144:142-8). This would explain why we observe some EndoH resistant VSG in lane 8, although high levels of wild type VSG are still present. Nevertheless, as correctly pointed out by the reviewer, we can not exclude the possibility that the TbSTT3s do not have 100% specificity for the donor oligosaccharide and acceptor sites, and will transfer a non-preferred donor at a certain frequency to their non-preferred sites. This is now included in the discussion.

4. The authors would like to conclude that TbSTT3C and TbSTT3B also have different acceptor site specificities based upon the data presented in Fig. 1. It would be nice if the authors could use the same size window (4 residues on each side of NXT/S in Fig. 1, and 5 residues on each side in Fig.5) to consider the impact of adjacent residues on glycosylation efficiency. It would be preferable if the method of peptide comparison was the same in both cases (correctly calculated isoelectric point). The weakness of the Fig. 1D method is that the reader can deduce very little about the net charge of the sequences flanking poorly and efficiently modified glycosylation sites. In the case of STT3C, the number of efficiently used sites appears to be 12, while inefficient sites number 17, so any conclusions are somewhat blunted by the small number of peptides. This reviewer also wonders whether the method of analysis (use of cell surface protein) biases against the detection of certain hypoglycosylated variants, if lack of a glycan at a certain site has a negative impact on protein folding in the ER, a phenomena which is not unheard of. Efficient and inefficient sites both contain 0-3 acidic residues within this region, but it is not clear whether the poorly glycosylated outliers (2 or 3 acidic residues) have neutral PI's due to the presence of one or more basic residues. I can't help wondering whether there is a more effective and convincing way to present the data than box

diagrams. For example can the authors plot glycosylation efficiency vs IEP for STT3C? The observation that the TOS1_417 site is glycosylated, albeit not with 100% efficiency, by Tb SXTT3B, but not by yeast OST is interesting. There is some statistical evidence from B. Imperiali's lab that lysine at +4 relative to N in the sequon is underrepresented in sequon databases.

>>>> We have recalculated all of the pIs (as described above) and applied a similar method of analysis to both the yeast and trypanosome date, as suggested ñ this is found in the new Figure 6 . We could not find published data on the frequency lysine and +4 in the literature and so have not discussed this particular aspect.

Supplemental Section

1. Table S1.
(a) An asparagine in the glycopeptide CCW14_87 (m/z = 1105.4781) is marked by an asterisk, presumably indicating that this site is glycosylated. This seems to be the only asparagine marked by an asterisk; N residues in other glycopeptides appear to be in a bold font. Is there some reason that this peptide is marked differently?

>>>> This was meant to indicate that asparagines in glycosylation sequons were also indicated by a bold N. This has been added to the table footnotes. (b) The significance of the Rank column is unclear since all peptides are entries are ranked 1.

>>>> The rank column has been deleted for clarity.

(c) #Previously identified15. The meaning of the superscript 15 is not clear.

>>>> This was an error and it has been corrected. It is a call-out for Schulz BL, Aebi M (2009) Analysis of glycosylation site occupancy reveals a role for Ost3p and Ost6p in site-specific N-glycosylation efficiency. Mol Cell Proteomics 8: 357-364.

(d) The authors should add a sentence beneath the table to define the score column. Why does the peptide GAS1_253* have a score of (68) (i.e. score within parentheses)

(e) The expect column should also be defined. It is not clear why the exponent is capitalized in some cases (i.e. 8.4E-04), but not in others (i.e. 8.2e-08).

>>>>A summary of the methods used to generate the data in the table has been added. The 'Score' and 'Expect' columns contain equivalent data, and the 'Score' column has been removed for clarity. There is no difference between the meaning of the 'e' or 'E' nomenclatures. This has been standardized.

(f) The glycopeptide EXG2_50 should have an asterisk after it.

>>>> This has been added.

(f) The glycopeptide EXG2_50 should have an asterisk after it.

>>>> Corrected.

2. Figure S3 and Table SIV. Each of these supplemental items contains information on the glycan structure of Pronase peptides derived from VSG221 produced by T. brucei cells that express TbSTT3A, TbSTT3B or TbSTT3A+TbSTT3B. It is not clear to me why some intensity data presented in Figure S3 is in conflict with tabulated data in Table SIV. As an example, consider FigS3, panel A, N263 glycopeptide NET (Hex3HexNAc2). Figure S3 shows a 60% intensity value

for this glycopeptide in both the -Tet and +Tet conditions. However, ion intensity values in Table SIV for this peptide ({plus minus}Tet in STTA siRNA) are 439.4 (-Tet) and 173.4 (+Tet). The S3 figure legend does not explain how % intensity values were calculated, but it seems likely that the Table SIV values were normalized to the intensity of another glycopeptide. At best, Fig. S3A is confusing, because it appears that siRNA knockdown of STT3A does not reduce transfer of the GlcNAc2Man5 to the N263, but instead increases transfer of GclcNAc2Man9 to the N428 site.

>>>> We apologise for the lack of clarity (now improved in the legends to Table SIV and legend to Fig. S3). The figures in Table SIV are the raw data (ion counts) for each glycopeptide ion and the % intensities in Figure S3 are normalized to the intensity of the most intense glycopeptide ion in each analysis. Thus, using the refereeís example, for N263 glycopeptide NET(Hex3HexNAc2) the intensities of 439.4 (-Tet) and 173.4 (+Tet) are normalized to 724.3 (-Tet, intensity of GPI anchor+5Gal) and to 294.7 (+Tet, NTT+NaHex9HexNAc2), respectively. The purpose of retaining both of these somewhat redundant data sets is (a) to not loose sight of the raw data (Table SIV) and (b) to provide a graphic image (Fig. S3) that allows one to observe clearly how the NET site gets Man8 and Man9 glycans in the STT3A siRNA +Tet cells and not in the -Tet cells and how the NTT site gets abundant complex glycans in the STT3B siRNA +Tet cells and few in the -Tet cells.

3. Figure S4. This figure could be expanded to show the source of glycoproteins (total glycoprotiens or cell surface glycoproteins?) and to indicate whether the samples used for digestion were enriched by sequential lectin chromatography (ConA and then ricin) or split sample chromatography (ConA or ricin).

>>>> *The figure has been improved and also the following has been inserted into the results section to help the reader understand the process: "Briefly, the glycoproteins were captured from total trypanosome lysates by sequential affinity chromatography on immobilised ricin followed by immobilised ConA. The ricin- and ConA-binding fractions were eluted with appropriate sugars and individually processed by digestion with endoH followed by PNGaseF. Thus, endoH-sensitive glycopeptides appear 203 Da heavier by mass spectrometry due to the GlcNAc residue left attached to the relevant Asn residues by endoH while the remaining endoH-resistant, but PNGaseFsensitive, sites appear 1 Da heavier by mass spectrometry because of the conversion of the relevant Asn residues to Asp by PNGaseF."*

Minor points:

1. The Schulz and Aebi reference is listed twice in the References list, once in 2008 (no Vol and page #) and once in 2009.

>>>> *Corrected.*

2. Legend Fig. 1(C), the sentence "Data is shown in Table S1" should read "Data is shown in Table S2. "

>>>> *Corrected.*

3. Define symbols used in in Fig. 2 (GlcNac, Man, etc.)

>>>> *Done.*

4. The manuscript should mention that results of the proteomics analysis of T. brucei glycopeptides is available in the PRIDE database (list accession #). Hopefully they will correct the annotated IEPs.

>>>> *The PRIDE entries have been corrected for IEPs and the accession number is now quoted in the paper*

2nd Editorial Decision 10 June 2009

Thank you for submitting the revised version of your manuscript EMBOJ-2009-70981R. It has now been seen again by referee 3, who now finds it acceptable for publication - his/her comments are attached below.

However, before we can accept your paper, both we and the referee have noticed a number of issues - primarily with the figures - that need to be fixed:

1. The font size you use for the text labels, particularly in Figure 1 (but also in some of the other figures) is very small, and I worry that some of the labels will not be clearly legible in the final version. Please could you increase the font size where appropriate?

2. As referee 3 points out, there is a problem with the conversion of Figure 5. This should not be a problem, since we use the original files. However, I also notice that there is no x-axis line for this chart, which looks a little odd!

3. As the referee points out, your reference in the text to Figure 6 (on page 9) is incorrect - you refer to 6C&D rather than B&C.

I would therefore ask you to make these changes and submit the revised version of your manuscript and figures through our online system. Once we have the amended version, we will then be able to formally accept your manuscript.

Many thanks,

REFEREE REPORT

Referee #3 (Remarks to the Author):

The manuscript from Izquierdo et al. has been carefully revised to address the points I raised during the previous review. In my opinion, the revised manuscript reports interesting and novel results concerning the substrate specificity of the T. brucei Stt3 proteins. The authors have made a serious and successful effort to make the revised manuscript more easily read by non-specialists.

Minor points

1. Figure 5 is truncated in the middle of pI = 11.5 (a single bar, ~10% complex oligosaccharide is shown). I assume that this is a problem that occurred during pdf conversion, but if not, the figure needs to be fixed.

2. Figure 6 is a big improvement relative to the previous version (Fig. 1D). The text refers to panels 6C and 6D, instead of 6B and 6C.

2nd Revision - authors' response 17 June 2009

All of the remaining pints have been attended to. Many thanks.