Published in final edited form as:
Nat Microbiol. 2018 August ; 3(8): 932–938. doi:10.1038/s41564-018-0187-6.

African trypanosomes evade immune clearance by *O*-glycosylation of the VSG surface coat

Jason Pinger[#1], **Dragana Nešić**[#2,†], **Liaqat Ali**[#3], **Francisco Aresta-Branco**[4,5], **Mirjana Lilic**[2,‡], **Shanin Chowdhury**[1], **Hee-Sook Kim**[1], **Joseph Verdi**[6], **Jayne Raper**[6], **Michael A. J. Ferguson**[3,§], **F. Nina Papavasiliou**[4,§], and **C. Erec Stebbins**[5,§]

[1]The Rockefeller University, Laboratory of Lymphocyte Biology, New York, New York, USA

[2]The Rockefeller University, Laboratory of Structural Microbiology, New York, New York, USA

[3]Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dundee, UK

[4]Division of Immune Diversity, German Cancer Research Center, Heidelberg, Germany

[5]Division of Structural Biology of Infection and Immunity, German Cancer Research Center, Heidelberg, Germany

[6]Department of Biological Sciences, Hunter College, City University of New York, New York, USA

[#] These authors contributed equally to this work.

Abstract

The African trypanosome, *Trypanosoma brucei* spp., is a paradigm for antigenic variation, the orchestrated alteration of cell surface molecules to evade host immunity. The parasite elicits robust antibody-mediated immune responses to its Variant Surface Glycoprotein (VSG) coat, but evades immune clearance by repeatedly accessing a large genetic VSG repertoire and “switching” to antigenically distinct VSGs. This persistent immune evasion has been ascribed exclusively to amino acid variance on the VSG surface presented by a conserved underlying protein architecture. We establish here that this model does not account for the scope of VSG structural and biochemical diversity. The 1.4Å resolution crystal structure of variant VSG3 manifests heretofore

[§]Correspondence to: e.stebbins@dkfz-heidelberg.de (structural biology), n.papavasiliou@dkfz-heidelberg.de (trypanosome strains and infection assays), and m.a.j.ferguson@dundee.ac.uk (carbohydrate biochemistry).
[†]Current address: The Rockefeller University, Laboratory of Blood and Vascular Biology, New York, New York, United States of America.
[‡]Current address: The Rockefeller University, Laboratory of Molecular Biophysics, New York, New York, United States of America.

Data Availability

Coordinates and structure factors of VSG3 have been uploaded to the RCSB PDB (www.rcsb.org) with PDB ID: 6ELC.

Author Information

These authors contributed equally: JP, DN, and LA. Conceived and designed the experiments: JP, DN, CES, LA, MAJF, FNP. Protein purification: JP, DN, ML, FNP, FAB. Structural prediction analyses: CES, FAB. Crystallography: DN and CES. Mass spectrometry: LA, MAJF. Trypanosome genetics and growth, antibody assays, and mouse infection studies: JP, SC, FAB, FNP, JV, JR. Contributed reagents/materials/analysis tools: DN, LA, JP, ML, SC, FAB, H-SK, FNP, CES. Wrote the paper: JP, CES, FNP, DN, LA, MAJF. The authors declare that they have no competing financial interests. Correspondence should be addressed to: e.stebbins@dkfz-heidelberg.de (structural biology), n.papavasiliou@dkfz-heidelberg.de (trypanosome strains and infection/antibody assays), and m.a.j.ferguson@dundee.ac.uk (carbohydrate biochemistry).

unappreciated divergence in the tertiary fold and oligomeric state. The structure also reveals an *O*-linked carbohydrate on the top surface of VSG3, a modification previously unknown in African trypanosomes. Mass spectrometric analysis indicates that this *O*-glycosylation site is heterogeneously occupied in VSG3 by 0 to 3 hexose residues and is also present in other VSGs. We demonstrate that this *O*-glycosylation increases parasite virulence by impairing the generation of protective immunity. These data alter the paradigm of antigenic variation by the African trypanosome, expanding VSG variability beyond amino acid sequence to include surface post-translational modifications with immunomodulatory impact.

Antigenic variation, or the diversification of expressed surface antigens during the course of infection, is a microbial survival strategy exemplified by the African trypanosome (*Trypanosoma brucei* spp.), the causative agent of African sleeping sickness in humans and nagana in livestock. Central to this immune-evasion strategy is the **V**ariant **S**urface **G**lycoprotein (**VSG**) coat of ~10 million molecules that densely covers the organism in its mammalian-infectious forms. The VSG coat elicits a robust antibody response, which the parasite evades by "switching" to a new, antigenically distinct variant. The VSG-antibody interaction thus represents the central molecular interface between the *T. brucei* pathogen and its host, resulting in observable peaks and valleys of parasitemia that are the repeated outcome of cycles of VSG switching, antibody generation, and parasite killing1.

Despite the centrality of the VSG-antibody interface, key parameters of these interactions remain poorly characterized, such as how antibodies might bind the dense VSG surface array, and how the sequence diversity between *VSG* genes manifests structurally to allow continual evasion of accumulated antibody responses during infection2, 3. Long-held assumptions have ascribed the immune-evasive properties of new coats exclusively to divergence in the amino acid sequence of VSG antigenic surfaces displayed on a scaffold of conserved overall architecture. This view was first formed nearly three decades ago based on crystal structures of the N-terminal domains (NTDs) of two VSGs4, 5. These early structural analyses revealed three key features: (i) a general conservation of architecture consisting of a two-lobed, "dumbbell" arrangement, the "top" (facing away from the pathogen) and "bottom" regions separated by an elongated 3-helix bundle (Fig. 1a), (ii) the VSGs formed homodimers, and (iii) despite alignment in the three-dimensional folds of these molecules (particularly in the core helical bundle), their surface properties varied considerably, consistent with the generation of immunologically distinct entities.

Structural information and VSG sequence alignments have been used to allocate VSG N-terminal domains to different classes based on the presence of numerous analogous cysteine disulfides6. The two early VSG structures, VSG2 (also termed MITat1.2 and VSG221) and ILTat1.24, and a newer structure published during review of this manuscript (VSG M1.1, which is highly homologous to VSG27) all belong to "class A". Given the depth of the genomic archive and the substantial diversity within VSGs even at the level of primary sequence, it seemed unlikely that the three structures published to date would exhaustively cover variation in VSG protein space.

Here, we present the crystal structure of the NTD of a common "Class B" variant, VSG3 (also termed MITat1.3 and VSG224), which contains several distinct structural and

Europe PMC Funders Author Manuscripts

biochemical features as compared to previously published VSGs. Most notably, an unexpected *O*-linked glycan was identified in the VSG3 crystal structure at residue S317, located on the top surface of the NTD. Mass spectrometry analyses revealed heterogeneity in the number of hexoses linked to VSG3 S317, and also that analogous modifications are present in other Class B VSGs. Finally, mouse infection and antibody binding assays showed that the VSG3 *O*-glycan potently increases parasite virulence and impairs the host's ability to generate a protective antibody response against trypanosomes expressing a VSG3 surface coat.

Results

Structural and oligomeric divergence in the VSG3 NTD

Preliminary examinations of VSG3 showed that it not only differs markedly at the sequence level from VSG2, ILTat1.24 and VSG M1.1 but that it also scored poorly against these structures in structure-based prediction algorithms such as protein "threading"8. Intrigued by these analyses, we proceeded to crystallize and solve the structure of this variant (Methods and Supplementary Table 1).

The 1.4Å resolution crystal structure reveals that although VSG3 shares the overall two-lobe, three-helix bundle architecture with the Class A VSGs, there is substantial divergence in the fold as well as the topology, such as the connectivity of the lower-lobe subdomain (Fig. 1, Supplementary Fig. 1). Consistent with such variation, the oligomerization state of VSG3 is also different, existing in solution and the crystal asymmetric unit as a monomer rather than a dimer (Supplementary Fig. 2)4, 5, 7. Further divergence from previous VSG structures is found in the number and location of cysteine disulfide bonds in the top lobe of the protein (Supplementary Fig. 3), including one that anchors an antigenically critical loop (discussed below), as well as the position and structural context of an N-linked glycan in the lower lobe (Fig. 2a and Supplementary Fig. 4).

VSG3 is heterogeneously *O*-glycosylated

A striking feature of the VSG3 structure is an unexpected carbohydrate modification (Fig. 2a and b). Additional density attached to S317 revealed the presence an *O*-linked α-glucose moiety (Fig. 2b). This modification is located on the top surface of VSG3 in an exposed loop stabilized by flanking cysteines involved in a disulfide bridge (the peptide CTG**S**ASEGLC, residues 314-323, Fig. 3a), and is therefore likely to be highly accessible to the immune system.

Mass spectrometric analysis of intact VSG3 by ES-MS before and after digestion with PNGaseF validated the *O*-linked α-glucose moiety identified in the crystal structure, and additionally revealed heterogeneity in the number of *O*-linked glycans present (Supplementary Fig. 5). Subsequently, peptide mass analysis by LC-MS/MS confirmed that the peptide A311-K339 contained 0, 1, 2 or 3 hexose residues (Fig. 3b), and the MS/MS product ion spectra of several peptide glycoforms further defined the *O*-glycosylation site as S317 (Supplementary Fig. 6). Together, these data strongly suggest that there is a heterogeneous chain of 0-3 hexoses attached to S317. Based on the ES-MS spectra

(Supplementary Fig. 5a and b), we estimate the proportions of the 0, 1, 2 and 3 hexose glycoforms at 8%, 33%, 37%, and 22%, respectively. No evidence of additional glycosylation of serines or threonines was observed.

While the modification of serine and threonine residues by *O*- and phosphodiester-linked carbohydrate chains is known for other kinetoplastids9, such modifications have not been previously described in *T. brucei.* To our knowledge, this is also the first description of a Glcα1-*O*-Ser linkage in any organism, although Glcβ1-*O*-Ser linkages are known10. In addition, no carbohydrate of any kind has been previously shown to occur on the top surface of any VSG.

Other VSGs display analogous *O*-linked glycans

To assess the prevalence of surface *O*-linked glycosylation, we compiled a list of VSGs from diverse genomic locations11 with N-terminal domains that threaded to VSG3 and contained cysteine-flanked, serine- or threonine-containing loops analogous to the glycosylated loop in VSG3. A subset of this list is shown in Fig 3c. We then analyzed two of these: VSG11 and VSG615. Additionally, we examined VSG21, which fails to thread to VSG3. Mass spectrometric analysis showed that both VSG11 and VSG615 possessed heterogeneous *O*-linked hexose modifications analogous to those found in VSG3 (Fig. 3d and e), whereas VSG21 did not. These data suggest that surface *O*-linked glycosylation could be a widespread VSG modification present throughout trypanosome infections.

The VSG3 *O*-glycan enhances parasite virulence

Because of its central placement on the top surface of VSG3, we hypothesized that the *O*-linked glycan might constitute a critical antigenic determinant. To test this conjecture, we genetically engineered isogenic trypanosome strains that carried either a wild type VSG3 ($VSG3_{WT}$) coat or a VSG3 coat with a S317A point mutation ($VSG3_{S317A}$, which cannot be *O*-glycosylated). This was achieved by knocking these genes into a *VSG2* expression site, replacing the active *VSG2* gene (Methods and Supplementary Fig. 7). We then used these strains to infect naïve C57BL/6 mice.

The polyclonal antibodies that mediate *T. brucei* clearance are of IgM-class and directed solely against exposed epitopes of parasite-bound VSG12. To compare the ability of the two isogenic strains to evade this IgM response, we focused on the first peak of parasitemia (up to day 9 post infection). During this period, outgrowth of the $VSG3_{WT/S317A}$-expressing trypanosomes occurs, and mice either succumb to infection or clear the parasite population, leaving only those cells that have switched expression to a new VSG. We observed that all $VSG3_{WT}$-infected animals succumbed to infection during the first peak (Fig. 4a). In contrast, all but one of the $VSG3_{S317A}$-infected animals survived the first peak, having cleared the $VSG3_{S317A}$-expressing trypanosomes by day 8. Late mortality events (after day 10) in the $VSG3_{S317A}$ mouse group resulted from outgrowth of parasites which had switched expression to other VSGs. This robust differential in virulence was reproducible in two independent clone sets (Data from the two clone sets are combined in Fig. 4a and shown separately in Supplementary Fig. 8a and c), and in additional infections of a different mouse strain (CD-1; Supplementary Fig. 8e). We also observed lower peak parasitemia levels in the

$VSG3_{S317A}$ infections compared to the $VSG3_{WT}$ infections (Fig. 4b, Supplementary Fig. 8b, d, and f). However, growth curves between the wild type and mutant VSG3 trypanosomes were similar *in vitro* (Supplementary Fig. 7b), suggesting that the differences observed in infection were related to host factors rather than being clone-intrinsic. We therefore reasoned that our infection data might reflect a role for the VSG3 *O*-linked glycan chains in inhibiting the host's ability to generate a protective antibody response.

VSG3 *O*-glycosylation impairs host development of protective immunity

To address this hypothesis, we performed additional infection experiments in which mice were immunized against $VSG3_{WT}$ or $VSG3_{S317A}$ prior to infection (via inoculation with UV-irradiated parasites—see methods). This protocol was previously shown to yield an immune response that closely mimics the response to the infective parasite13. We then challenged these mice via injection of live cognate parasites at day 8 post-immunization and monitored mouse blood smears for the presence of trypanosomes on days 5-8 post-challenge. As shown in Fig. 4c, non-immunized mice (Fig. 4a and b) uniformly presented parasitemia by day 5. We found that pre-immunization with $VSG3_{S317A}$ was highly protective, suppressing detectable outgrowth of cognate $VSG3_{S317A}$ trypanosomes in 83.3% (10/12) of mice (Fig. 4c). In contrast, pre-immunization with $VSG3_{WT}$ was only marginally protective against cognate $VSG3_{WT}$ trypanosomes, suppressing outgrowth in 25% (3/12) of mice. All mice that presented parasitemia either succumbed to infection or cleared the parasites by day 8. The discrepancy in the protective effect of pre-immunization was also evident in survival differences between the two mouse groups (Supplementary Fig. 9a). These results provide evidence that the *O*-linked glycan chains displayed by VSG3 inhibit the ability of the host antibody response to mediate parasite clearance.

To directly examine the effect of the *O*-glycan chains on the properties of polyclonal IgM responses, we collected antisera elicited by $VSG3_{WT}$ or $VSG3_{S317A}$-coated parasites at day 8 post-infection. We then used flow cytometry (FACS) analysis to assess antisera binding in the context of the intact parasite surface coat (as surface binding is thought to be necessary and sufficient to predict clearance12). In accordance with our *in vivo* results, we found that $VSG3_{WT}$-elicited antisera exhibited markedly lower binding against $VSG3_{WT}$ trypanosomes, as compared to $VSG3_{S317A}$-elicited antisera binding $VSG3_{S317A}$ trypanosomes (Fig. 4d and e; Supplementary Fig. 9b-f).

Additional analyses of the cross-reactivities of our polyclonal antisera did not suggest a simple mechanistic explanation as to how the heterogeneous *O*-glycan chains cause elicitation of IgM responses that are poorly functional against cognate $VSG3_{WT}$ coats (Supplementary Fig. 9b-f). $VSG3_{WT}$-elicited antisera bind $VSG3_{S317A}$-expressing parasites well, which weakens the explanation that *O*-glycosylation simply causes elicitation of antibodies that are poorly functional. In the reverse combination, $VSG3_{S317A}$-elicited antisera also bind $VSG3_{WT}$-expressing parasites well, indicating that the *O*-glycosylation is not broadly able to inhibit antibody binding. These data therefore suggest that the poor binding of the WT/WT parasite/antisera combination may result from a combined effect of the heterogeneous *O*-glycan chains at both the antibody elicitation stage and in the subsequent antibody-parasite interactions (discussed further below). Despite the apparent

complexity of the underlying mechanism, the clear and accordant *O*-glycan-dependent discrepancies observed in parasite clearance following infection (Fig. 4a), generation of protective immunity from pre-immunization (Fig. 4c), and cognate antibody binding (Fig. 4e and f) altogether strongly indicate that the *O*-glycan chains that decorate VSG3 aid the parasite in evasion of the host's humoral immune response.

Discussion

The African trypanosome presents a fascinating paradigm for persistent immune evasion through antigenic variation. The discovery of the VSGs and of VSG coat switching dramatically altered our understanding of African trypanosomiasis and opened vistas for research into understanding how a pathogen can continuously avoid the ever-adapting mammalian immune system. Sequence analyses and some structural studies have provided many key insights into the bases of antigenic diversity between VSGs. However, the results we present in this study indicate that our understanding of VSG diversity at the protein level has been (and almost certainly remains) incomplete. We have demonstrated that VSG structural architecture is more variable than initially appreciated, expanding the VSG "diversity space" at the level of the protein fold and resulting molecular surfaces. Secondly and more unexpectedly, we have identified a VSG post translational modification (of a form previously unidentified in *T. brucei*) that changes the antigenic character of the VSG and exhibits potent immunomodulatory effects. The presence of an *O*-linked glycan also indicates the existence of an associated biochemical pathway, including an unidentified *O*-glycosyltransferase. Given the degree to which *O*-glycosylation of VSG3 increases parasite virulence and our identification of analogous modifications in other VSGs, further characterization of this pathway may provide important insights into *T. brucei* pathogenesis. Additionally, other unidentified types of VSG modifications could also modulate host-parasite interactions. The possibility of unidentified biochemical diversity highlights the need for continued examination of VSGs at the protein level, as such examinations may reveal additional immunomodulatory factors in *T. brucei*, with potential relevance to other pathogens or immune interactions.

Regarding the mechanistic basis of the poor functionality of antibodies elicited against wild type (*O*-glycosylated) VSG3 (Fig. 4c-e, Supplementary Fig. 9), it is important to first reiterate that these analyses represent overall functionality of polyclonal antisera. These antisera are comprised of a multitude of clonal antibodies (of IgM class), each exhibiting distinct properties. The factors influencing the overall reactivity of the polyclonal antisera thus include: (1) the binding epitope of each antibody, (2) whether the epitopes overlap with or are affected by the glycan, and if so, (3) whether the glycan affects each interaction enough to significantly alter antibody binding, and finally, (4) the abundance of each monoclonal IgM antibody within the population of antibodies that make up the polyclonal antisera. The subject of the effects of glycosylation on antibody binding is a poorly understood area, complicated by the diverse ways glycosylation can alter the antigenic character of the modified protein. Relevant literature demonstrates that glycosylation of proteins and peptides can have myriad effects on the binding profile of antibodies that target regions in the vicinity of the glycan(s), extending beyond simple exposure or occlusion of epitopes (reviewed in 14). An individual antibody raised against a glycoprotein may have

unaffected, increased, or decreased binding affinity based on the presence or absence of the glycan, the length and composition of the glycan chain, and “allosteric” effects of the glycan on the structural conformation of the surrounding peptidic regions (i.e. the affected epitope need not actually include the glycan). The impact of differential glycosylation can vary dramatically and unpredictably for individual antibodies 15.

The fact that VSG3 is variably modified with a chain of zero to three *O*-linked hexoses may also suggest that display of a heterogeneously glycosylated surface can impair immune function. Heterogeneous *O*-glycosylation of VSG3 would be expected to produce, minimally, four antigenically variable entities on the trypanosome surface (one for each hexose chain length, ignoring the possibility of multiple conformers for each chain). These multiple glycoforms would therefore present a more diverse set of epitopes for antibody elicitation and a less homogenous surface for antibody binding. Furthermore, IgM antibodies are particularly sensitive to epitope density because their binding is typically dependent on polyvalent interactions. For *T. brucei* specifically, effective IgM binding and trypanosome clearance have been demonstrated to require a critical density of cognate VSGs on the trypanosome surface13. The presence of various glycoforms on a $VSG3_{WT}$ coat could reduce the effective densities of any epitopes altered by glycosylation. In contrast, a $VSG3_{S317A}$ coat would be comparatively homogenous and potentially more amenable to binding for IgMs with certain specificities.

Overall, because of the complexity of the system and the quantity of unknown variables, it is impossible to predict *a priori* the prevalence or proportions of elicited antibodies in our antisera with given binding epitopes, glycoform specificities, affinities, and avidities, but these are some of the factors that underlie our binding results in aggregate. The mechanisms contributing to our results could be further examined by isolating a set of monoclonal antibodies from the polyclonal responses and testing their binding properties against different VSG3 glycoforms, or potentially VSG3-derived peptide glycoforms. Importantly however, the binding effects observed likely depend on multiple epitopes and the overall structure and packing of VSGs in the surface coat. Such studies could improve our understanding the impact of glycosylation on the antibody response, a topic with relevance far beyond *T. brucei* infection.

In summary, our data demonstrate unexpected antigenic variation in *T. brucei* VSG-coats that cannot be predicted by primary sequence analysis. Beyond significant additional variation in the protein fold, the high-resolution structure of VSG3 reveals an unanticipated (and potentially widespread) post-translational modification on the VSG surface that potently enhances parasite immune evasion. Antigenic variation in the African trypanosome therefore includes multiple overlapping immune evasion strategies: amino acid divergence, heterogeneous surface glycosylation, and perhaps additional variance yet to be uncovered.

Methods

Purification of VSGs

T. brucei expressing VSG2, VSG3, VSG3-S371A mutant, VSG11, VSG21, or VSG615 were cultured *in vitro* in HMI-916 media. Cells were pelleted and then lysed in 0.2mM

$ZnCl_2$, according to established protocols17. This lysis mixture was centrifuged, and the pellet containing the membrane material was re-suspended in pre-warmed (40°C) 10mM phosphate buffer. This enabled activation of endogenous lipases and resulted in the efficient release of surface VSG protein from the membrane. Following a second centrifugation, supernatant containing VSG3 protein was loaded onto an anion exchange column (Q-sepharose Fast-Flow, GE Healthcare), which had been equilibrated with 10mM phosphate buffer. Two passes through this column resulted in the C-terminal truncation of VSG3 (but not VSG2), presumably due to cleavage by endogenous proteases. This C-terminally truncated VSG3 was more stable than full length VSG3 and was used for crystallization. Further purification steps for VSG3 include one or, if necessary, two runs over a gel filtration column (Superdex 200, GE Healthcare) in 10mM HEPES pH 7.5 ((4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)), 150 mM sodium chloride (NaCl). Aliquots from the gel filtrations runs were subjected to SDS-PAGE analysis for visual inspection.

Crystallization and Structural Determination

For crystallization, the purified VSG3 N-terminal domain was concentrated to 2mg/ml in a buffer containing 10mM HEPES pH 7.5, 150mM NaCl. Native crystals were grown by vapor diffusion using hanging drops formed from mixing a 1:1 volume ratio of the protein with an equilibration buffer consisting of 16-22% polyethylene glycol (PEG) molecular weight 3350 Da, 200-325mM NaCl, and 100mM Tris pH 8.2 (tris(hydroxymethyl)aminomethane). For cryoprotection, crystals were transferred directly into a buffer with a 25% PEG 3350 Da, 300mM NaCl, 100m Tris pH 8.2, 10% glycerol, and flash-cooled immediately afterward to 100K (-173.15°C). Crystals formed in the space group I23 with a single VSG3 monomer in the asymmetric unit. For phase determination, crystals were grown in similar conditions in which NaCl was replaced with sodium bromide (18-22% PEG 3350 Da, 275-400mM sodium bromide, 100mM Tris pH 8.2). Prior to flash-freezing, those crystals were transferred into cryoprotection buffers (25% PEG 3350 Da, 50mM Tris pH 8.2, 5% glycerol) containing stepwise increasing concentration of sodium bromide from 0.5M to 2M.

Data were collected at Advanced Photon Source (APS) at Argonne National Laboratory at beamline 24-ID-C and processed onsite through the RAPD software pipeline18, 19, 20, 21, 22. Bromine soaked crystals were collected at a wavelength of 0.9194Å and the structure solved by SAD (Single-wavelength Anomalous Dispersion) from the anomalous signal due to 28 partially ordered bromine ions using the SHELX suite23. A partial model of the bromine-soaked protein was built by the PHENIX suite24. This model was then used in arp/wARP25, 26 as a starting model for automated building against the native data to 1.4Å resolution. Several cycles of automated building with arp/wARP, refinement with REFMAC27, 28, and manual inspection and model building led to a final model with and R/Rfree of 13.0%/16.9%. Two carbohydrates are present in the model – an N-linked $Man_2GlcNAc_2$ to N67 and an *O*-linked Glc to S317.

Digestion, de-*N*-Glycosylation and Mass Spectrometry

Preparation of *N*-glycopeptides—Purified VSG3 (400 μg) was denatured and reduced (200 μl 10 mM sodium phosphate, pH 8.0, 8 M urea, 40 mM dithiothreitol (DTT), 2 h, room

temperature (RT)) and S-alkylated by adjusting to 100 mM iodoacetamide (IAA), 1 h, RT, in the dark. Unreacted IAA was quenched by the addition of 4 µl of 1M DTT and sample adjusted to 400 µl containing 2 mM calcium acetate 80 µg Pronase (Sigma-Aldrich). After digestion for 24 h, RT the Pronase glycopeptides were enriched using cotton wool hydrophilic interaction liquid chromatography prior to reverse phase liquid chromatography29 tandem mass spectrometry (LC-MS/MS).

Preparation of *O*-glycopeptides—Purified VSG3 (50 µg in 25 µl 10 mM sodium phosphate, pH 8) was reduced (10 mM DTT, 85 °C, 20 min), S-alkylated (25 mM IAA, 1h, RT in the dark), diluted with an equal volume of 2 x GluC buffer and digested with 1:25 (w/w) endoproteinase GluC (New England BioLabs) for 24 h, 37 °C, with shaking. The GluC fragments were separated using NuPAGE 4-12% Bis-Tris protein gels (Invitrogen), stained with Coomassie Brilliant Blue (Thermo Scientific) and the 17 kDa fragment was excised and subjected to in-gel trypsin digestion for LC-MS/MS analysis.

LC-MS/MS analysis of glycopeptides—The LC was an Ultimate U3000 Nano LC System (Dionex) fitted with a C18 trap (PepMap nanoViper, Thermo Scientific) and resolving column (PepMap RSLC) with inner diameters of 100 and 75 µm and lengths of 2 and 50 cm, respectively. Mobile phases were 0.1% formic acid in 2% acetonitrile (solvent A) and 0.08% formic acid in 80% acetonitrile (solvent B). Samples were loaded in solvent A and eluted as follows: 0% B for 5 min, gradient to 40% B in 122 min and to 98% B in 10 min held at 98% B for 20 min followed by 20 min re-equilibration with 2% B. The LC system was coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific) equipped with an Easy Spray ion source and operated in positive ion mode. The spray voltage was 2 kV and the full scans were acquired in a Fourier transform MS mass analyser over *m/z* 335-1800 at a resolution of 60,000. The MS/MS analyses were performed under data-dependent mode to fragment the top 15 precursors using collision induced dissociation (CID). The raw files, converted to mgf format by MSConvert (proteowizard.sourceforge.net), were searched against the VSG3 protein sequence using the Mascot software (v.2.4.0, Matrix Science Inc., Boston, MA) with the parameters set as follows: peptide tolerance, 5 ppm; MS/MS tolerance, 0.5 Da; enzyme, trypsin; one missed cleavage allowed; and fixed carbamidomethyl modifications of cysteines. Oxidation of methionine and serine conversion to de-hydro alanine were used as variable modifications.

LC-MS analysis of intact and de-N-glycosylated VSG3—VSG3 (50 µg) was denatured in 1% sodium dodecyl sulphate for 10 min at 95 °C and then digested with 2500 units of glycerol free N-glycosidase F (PNGaseF, NEB) in 1% n-octyl-β-D-glucopyranoside for 24 h at 37 °C. The de-*N*-glycosylated product was precipitated in 10% trichloroacetic acid for 24 h at 4 °C, collected by centrifugation (16 x 000 g, 20 min, 4 °C), washed with cold acetone, air dried and redissolved in 1% formic acid. A control (non-treated) VSG3 sample was similarly TCA-precipitated and redissolved in 1% formic acid. The samples were analysed by LC-MS using an Agilent 1200 System fitted with a C8 column (Agilent), 25µm ID, 360 mm long. The mobile phases were the same as above. Samples were loaded in solvent A and eluted as follows: 7% B for 8 min, gradient to 25% B in 25 min and to 60% B in 48 min held at 60% B for 10 min followed by 10 min re-equilibration with 7% B. The LC

was coupled to an Agilent 6520 Q-Tof mass spectrometer. The spray voltage was 1.7 kV and the ion transfer tube was set at 360 °C. MS scans were collected over *m/z* 335-3200 and spectra were deconvoluted using the Maximum Entropy software with the settings mass range 35-45k; mass step, 1.0 Da; S/N threshold, 30; adduct, proton; average mass, 90% peak height; maximum consecutive charge state, 5; and minimum protein fit score of 8.

Cloning of VSG3 and VSG3-S317A knock-in vectors pKI224 and pKI224-S317A

Plasmid cloning was performed as follows: first, the wild type *VSG3* gene was amplified from trypanosomal genomic DNA (*Lister427*) into a pGEM subcloning vector (Promega A1360), and the S317A mutation was introduced by QuickChange PCR. Wild type and mutant *VSG3* genes were then cloned into the vector pUB39 30 (replacing *VSG117* in the original vector via HindIII and BamHI sites) to create pSUN7 and pSUN8. *VSG3* and *VSG3-S317A* were amplified from pSUN7 and pSUN8, respectively, using primers 224_BsiWI_For (GCGACACGTACGCGGCATGCAAGCGGCAGCA) and 224_AvrII_Rev (AATTAACCTAGGGGGCAATTCCAGCTGTC) (these PCR fragments contained the WT/S317A *VSG3* coding regions along with the surrounding *VSG117* 5' and 3' UTRs from the pUB39 vector). The backbone vector pSY37F1D-CTR-BSD was digested with BsiWI and SpeI, and the larger fragment was isolated by gel purification. *VSG3* and *VSG3-S317A* PCR fragments were digested with BsiWI and AvrII and ligated into the pSY37F1D-CTR-BSD fragment, resulting in final plasmids pKI224 and pKI224-S317A, respectively.

Generation of $VSG3_{WT}$ and $VSG3_{S317A}$ -expressing *T. brucei*

The wild type *VSG3* and *VSG3-S371A* point mutant genes were transfected into trypanosomes by means of the previously established pSY37F1D-CTR-BSD vector13.This vector contains a blasticidin-resistance gene (BSD) cassette and allows a "knock-in" transfection in which the *VSG* gene included in the vector replaces the endogenous *VSG2* gene in the telomeric Bloodstream-form Expression Site 1 (BES1) of the *Lister427* genome (see diagram, Supplementary Fig. 7a). First, the wild type *VSG3* and *VSG-S317A* point mutant genes were cloned into the PSY37F1D-CTR-BSD vector backbone to create plasmids pKI224 and pKI224-S317A, respectively. These plasmids were then linearized with XhoI and transfected into a parental clone expressing VSG2 via the Amaxa nucleofector protocol as previously described 31. BSD resistant clones were selected based on expression of VSG3 (by FACS) and counterselected for dual expressors (which co-expressed VSG2; also by FACS), using monoclonal antibodies to VSG2 and VSG3. VSG2 counterselection was used to eliminate a subset of clones which integrated the VSG cassette ectopically. Two sets of clones were generated. Transfection of *VSG3* and *VSG3-S317A* into the *Lister 427* "single marker" cell line32 produced clones KI.$VSG3_{WT}$-1 and KI.$VSG3_{S317A}$-1, respectively. Transfection of *VSG3* and *VSG3-S317A* into a different VSG2-expressing *Lister 427* cell line (GPIPLC null, 30) produced clones KI.$VSG3_{WT}$-2 and KI.$VSG3_{S317A}$-2, respectively. KI.$VSG3_{WT}$-2 and KI.$VSG3_{S317A}$-2 were generated on a GPIPLC-/- background. The lack of a GPI-PLC gene does not affect parasite virulence or the course of parasitemia 30 and consequently should not affect interpretation of the mouse infection results shown. The presence of the single point mutation within $VSG3_{S317A}$ was verified in the KI.$VSG3_{S317A}$ clones through RT-PCR amplification and sequencing.

Mouse Infection Assays

For Fig. 4a and b, groups of C57BL/6J or CD-1 mice were infected via i.p. injection of 100 live trypanosomes, expressing either $VSG3_{WT}$ or $VSG3_{S317A}$. For Fig. 4c and Supplementary Fig. 9a, immunization-challenge experiments were performed as previously described 13 using C57BL/6J mice. For immunization, mice received an i.p. injection with 2-3 x 10^6 UV-irradiated trypanosomes (KI.$VSG3_{WT}$-2 or KI.$VSG3_{S317A}$-2; these clones lack the endogenous lipase GPI-PLC, and therefore retain intact VSG coats following irradiation30) in 100ul HMI-9, and an identical repeat injection 3 days later. The subsequent challenge injection was administered 8 days after the initial immunization, and consisted of 100 live trypanosomes (as in the Fig. 4a and b experiments). For all infection experiments, mouse survival was assessed daily, and parasitemia was monitored starting on day 4 or 5 by blood-smear (blood from mouse tail clippings examined under a microscope for the presence of trypanosomes) or by counting parasite numbers using a hemocytometer. For Fig 4c, mice were permanently scored as “confirmed parasitemia” when parasites were first identified by blood smear, regardless of whether parasitemia subsequently fell to undetectable levels. All mice were wild type males or females, aged 6-9 weeks at experiment start (Jackson Laboratory). We did not have an a priori, pre-specified effect size from which to preemptively determine an appropriate sample size for our mouse infection experiments, but the sample sizes were determined to be large enough to ensure proper statistical analyses and reproducibility. Mice were assumed to be sufficiently randomized by Jackson Laboratory and no additional randomization was performed. Experimenters were not blinded to mouse group allocation, but measurements (survival assessment, parasitemia counts, FACS MFI) were not subjective. All animal experiments were approved by Rockefeller University’s institutional animal care and use committee under protocol #16894.

Elicitation of $VSG3_{WT}$ and $VSG3_{S317A}$-specific polyclonal IgM antisera

C57BL/6J mice were injected i.p. with 1 x 10^3 $VSG3_{WT}$ or $VSG3_{S317A}$-expressing trypanosomes in HMI-9 (three mice each). After 4 days, infections were cleared with 250ug berenil/mouse injected i.p., and berenil treatment was repeated after 24 hours. Clearance was necessary to prevent $VSG3_{WT}$-infected animals from succumbing to infection prior to developing significant antibody titers. On day 8 post-injection, blood was collected via cardiac puncture, and serum was separated from whole blood using Microtainer serum collection tubes (BD 365967).

FACS analyses to determine IgM binding

All samples, solutions, and incubation steps were at 4°C or on ice to prohibit antibody internalization during sample preparation. 10^6 $VSG3_{WT}$ or $VSG3_{S317A}$-expressing trypanosomes were isolated from culture, resuspended in HMI-9 containing dilutions of $VSG3_{WT}$ or $VSG3_{S317A}$-specific polyclonal IgM antisera (see above), and incubated for 10 min. Cells were washed once with HMI-9, then incubated with 1:1000 dilution of goat anti-mouse IgM-FITC (SouthernBiotech 1021-02) in HMI-9 for 10 min. Cells were washed once again with HMI-9, resuspended in HMI-9 containing 500ng/ml Propidium Iodide (BD 556463) (for exclusion of dead cells from analysis), and analyzed using a BD-FACSCalibur (Fig. 4c) or BD-LSRII flow cytometer (Fig. 4d) and FlowJo software (Treestar).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank George Cross (Rockefeller University) and Hedda Wardemann (DKFZ) for critical reading of the manuscript and for general advice, Margarida Sanches-Vaz and Luisa Figueiredo (IMM, Lisbon) for help with mouse experiments and Monica Chandra (DKFZ) for providing us with purified VSG615. We also thank the staff at Argonne National Laboratories (NE-CAT) for beamline support. NE-CAT is funded by an NIH/NIGMS grant (P41 GM103403) and the Pilatus 6M detector on 24-ID-C beam line is funded by an NIH-ORIP HEI grant (S10 RR029205). The Advanced Photon Source, within which NE-CAT is located, is a U.S. Department of Energy (DOE) User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This work was also supported by funds to C.E.S. and F.N.P. from the German Cancer Research Center (DKFZ, Heidelberg) and Rockefeller University, by NIH/NIAID (AI085973) to FNP, and by a Wellcome Trust Senior Investigator Award (101842) to M.A.J.F. The Dundee mass spectrometry facility is supported by Wellcome Trust grant 097045.

References

1. Matthews KR, McCulloch R, Morrison LJ. The within-host dynamics of African trypanosome infections. Philos Trans R Soc Lond B Biol Sci. 2015; 370
2. Hsia R, Beals T, Boothroyd JC. Use of chimeric recombinant polypeptides to analyse conformational, surface epitopes on trypanosome variant surface glycoproteins. Mol Microbiol. 1996; 19:53–63. [PubMed: 8821936]
3. Schwede A, Macleod OJ, MacGregor P, Carrington M. How Does the VSG Coat of Bloodstream Form African Trypanosomes Interact with External Proteins? PLoS Pathog. 2015; 11:e1005259. [PubMed: 26719972]
4. Metcalf P, Blum M, Freymann D, Turner M, Wiley DC. Two variant surface glycoproteins of Trypanosoma brucei of different sequence classes have similar 6 A resolution X-ray structures. Nature. 1987; 325:84–86. [PubMed: 2432433]
5. Blum ML, et al. A structural motif in the variant surface glycoproteins of Trypanosoma brucei. Nature. 1993; 362:603–609. [PubMed: 8464512]
6. Carrington M, et al. Variant specific glycoprotein of Trypanosoma brucei consists of two domains each having an independently conserved pattern of cysteine residues. J Mol Biol. 1991; 221:823–835. [PubMed: 1942032]
7. Bartossek T, et al. Structural basis for the shielding function of the dynamic trypanosome variant surface glycoprotein coat. Nat Microbiol. 2017; 2:1523–1532. [PubMed: 28894098]
8. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015; 10:845–858. [PubMed: 25950237]
9. Mendonca-Previato L, Todeschini AR, Heise N, Previato JO. Protozoan parasite-specific carbohydrate structures. Curr Opin Struct Biol. 2005; 15:499–505. [PubMed: 16154349]
10. Takeuchi H, Kantharia J, Sethi MK, Bakker H, Haltiwanger RS. Site-specific O-glucosylation of the epidermal growth factor-like (EGF) repeats of notch: efficiency of glycosylation is affected by proper folding and amino acid sequence of individual EGF repeats. J Biol Chem. 2012; 287:33934–33944. [PubMed: 22872643]
11. Cross GA, Kim HS, Wickstead B. Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427. Mol Biochem Parasitol. 2014; 195:59–73. [PubMed: 24992042]
12. Black SJ, et al. Regulation of parasitaemia in mice infected with Trypanosoma brucei. Curr Top Microbiol Immunol. 1985; 117:93–118. [PubMed: 3896678]
13. Pinger J, Chowdhury S, Papavasiliou FN. Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. Nat Commun. 2017; 8:828. [PubMed: 29018220]

14. Lisowska E. The role of glycosylation in protein antigenic properties. Cell Mol Life Sci. 2002; 59:445–455. [PubMed: 11964123]
15. Rangappa S, et al. Effects of the multiple O-glycosylation states on antibody recognition of the immunodominant motif in MUC1 extracellular tandem repeats. Med Chem Commun. 2016; 7:1102–1122.
16. Hirumi H, Hirumi K. Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. J Parasitol. 1989; 75:985–989. [PubMed: 2614608]
17. Cross GA. Release and purification of Trypanosoma brucei variant surface glycoprotein. J Cell Biochem. 1984; 24:79–90. [PubMed: 6725422]
18. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]
19. Evans P. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr. 2006; 62:72–82. [PubMed: 16369096]
20. Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011; 67:282–292. [PubMed: 21460446]
21. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013; 69:1204–1214. [PubMed: 23793146]
22. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. [PubMed: 15299374]
23. Sheldrick GM. A short history of SHELX. Acta Crystallogr A. 2008; 64:112–122. [PubMed: 18156677]
24. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
25. Langer G, Cohen SX, Lamzin VS, Perrakis A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc. 2008; 3:1171–1179. [PubMed: 18600222]
26. Winn MD, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011; 67:235–242. [PubMed: 21460441]
27. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
28. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011; 67:355–367. [PubMed: 21460454]
29. Ali L, et al. The O-glycomap of lubricin, a novel mucin responsible for joint lubrication, identified by site-specific glycopeptide analysis. Mol Cell Proteomics. 2014; 13:3396–3409. [PubMed: 25187573]
30. Leal S, et al. Virulence of Trypanosoma brucei strain 427 is not affected by the absence of glycosylphosphatidylinositol phospholipase C. Mol Biochem Parasitol. 2001; 114:245–247. [PubMed: 11378204]
31. Burkard G, Fragoso CM, Roditi I. Highly efficient stable transformation of bloodstream forms of Trypanosoma brucei. Mol Biochem Parasitol. 2007; 153:220–223. [PubMed: 17408766]
32. Wirtz E, Leal S, Ochatt C, Cross GA. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol Biochem Parasitol. 1999; 99:89–101. [PubMed: 10215027]

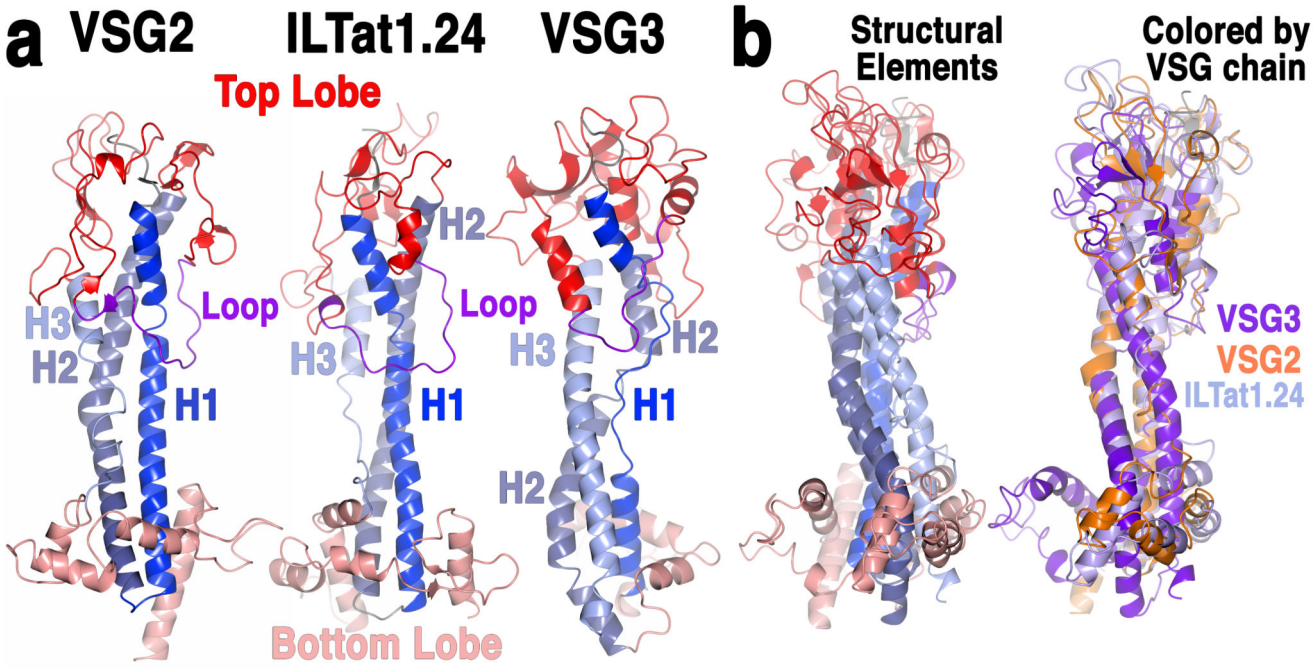

Fig. 1. Substantial structural divergence between VSGs.

(a) Comparison of monomeric VSG structures. VSG2 (MITat1.2), ILTat1.24 and VSG3 are shown as ribbon diagrams colored by six structurally conserved elements: the three helices of the stalk-like bundle (three shades of blue for each individual helix, H1-H3), the top and bottom lobes (red and salmon, respectively), and a conserved, elongated loop in the top lobe (purple). **(b)** Superposition of the three monomers, colored on the left by the schema described in (a) and on the right by individual protein as indicated in the labeling.

Europe PMC Funders Author Manuscripts

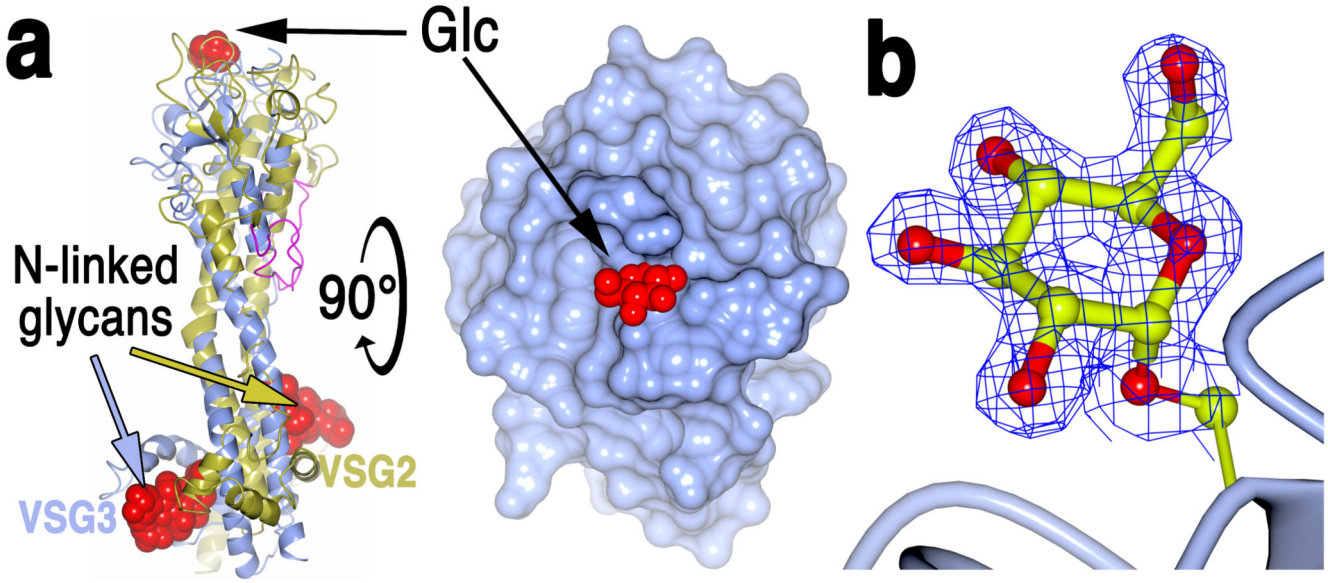

Fig. 2. Structural identification of *O*-linked Carbohydrate on VSG3 Surface.

(a) Left is a superposition of VSG3 (blue) and VSG2 (gold) with conserved purple "top loop." The $GlcNAc_2$ cores of the lower lobe N-linked glycans of the two VSGs are displayed as red space-filling atoms (coloring of arrows pointing to the glycans indicates to which VSG they belong). On the top surface of the upper lobe is a red space-filling representation of the α-glucose moiety attached to S317 of VSG3. A 90-degree rotation of the left model is rendered with molecular surfaces (blue for VSG3 protein, red for a space-filling representation of the glucose). **(b)** 1.4 angstrom resolution electron density map from the refined model of VSG3 clipped to the volume around the modeled glucose (Glc). S317 is shown connected to a ribbon diagram representation of the protein main chain (blue).

Europe PMC Funders Author Manuscripts

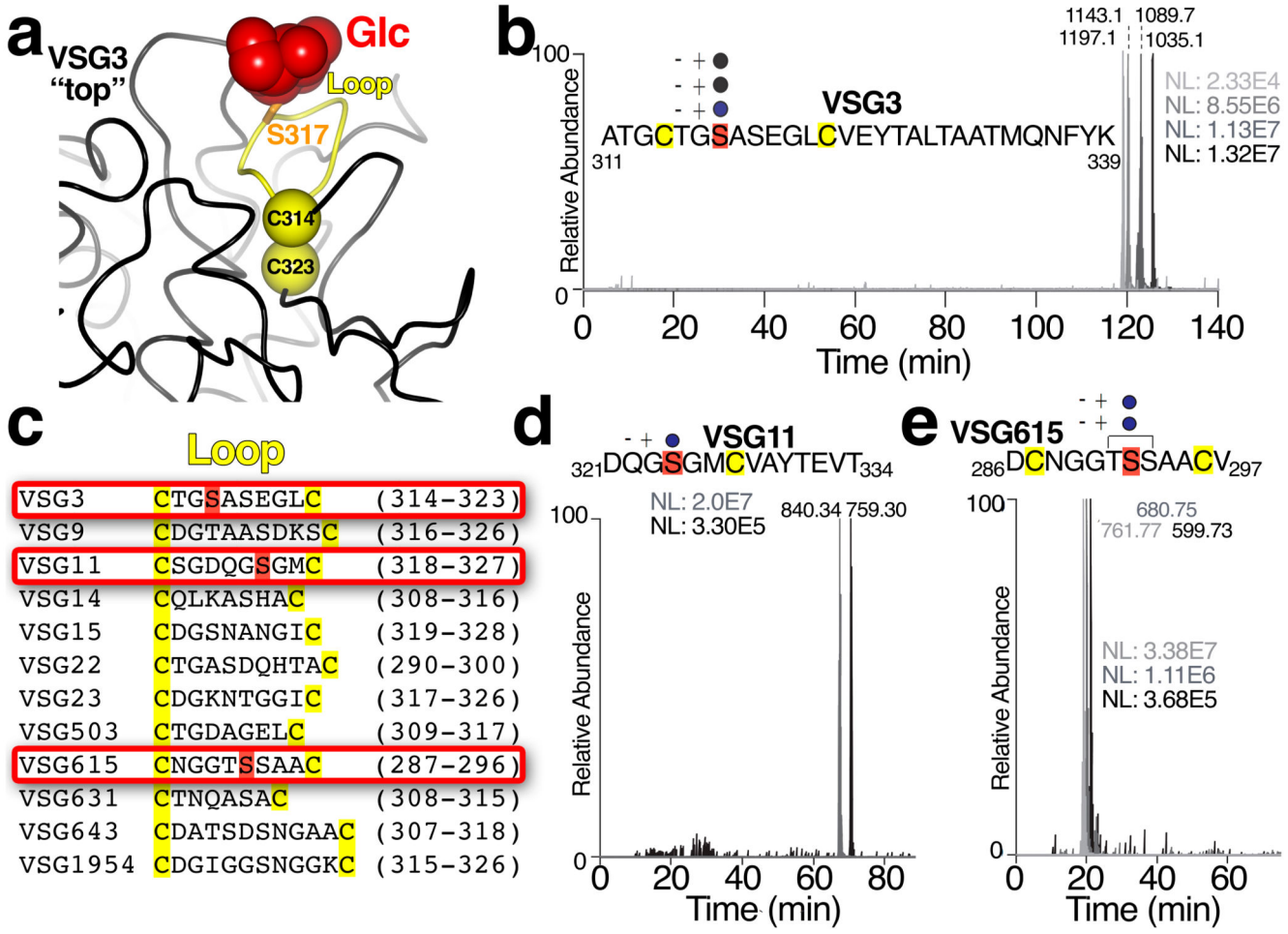

Fig. 3. Identification of heterogeneous *O*-linked glycans in surface loops of multiple VSGs.

(a) Cysteine-flanked glycan loop of VSG3 with the loop in yellow and the disulfide-bridged cysteines shown as yellow spheres and labeled, S317 shown as a stick model in orange, and the glucose atoms in red. **(b)** Extracted ion chromatogram showing the elution profiles of $[M+3H]^{3+}$ ions (*m/z* 1035.1, 1089.1, 1143.1 and 1197.1) for the A311 to K339 peptide of VSG3 containing 0, 1, 2 and 3 hexose residues, respectively. These elute by reversed-phase HPLC from most hydrophilic to least hydrophilic with decreasing numbers of sugar residues attached. This experiment was performed once. **(c)** Comparison of hypothesized "*O*-linked glycan" loops in additional VSGs. *O*-glycosylation was examined and confirmed for three VSGs (boxed in red): VSG3, VSG11, and VSG615. **(d)** and **(e)** are similar to (**b**) showing the extracted ion chromatogram results for VSG11 and VSG615, respectively, with the determined (VSG11) and the proposed (VSG615) modified serine residues indicated within the relevant peptide (the MS/MS data of the VSG615 glycopeptides did not allow unambiguous assignment of S292 as the unique *O*-glycosylation site, and thus T291 and S297 could also potentially be sites of modification). These experiments were performed twice each with similar results.

Europe PMC Funders Author Manuscripts

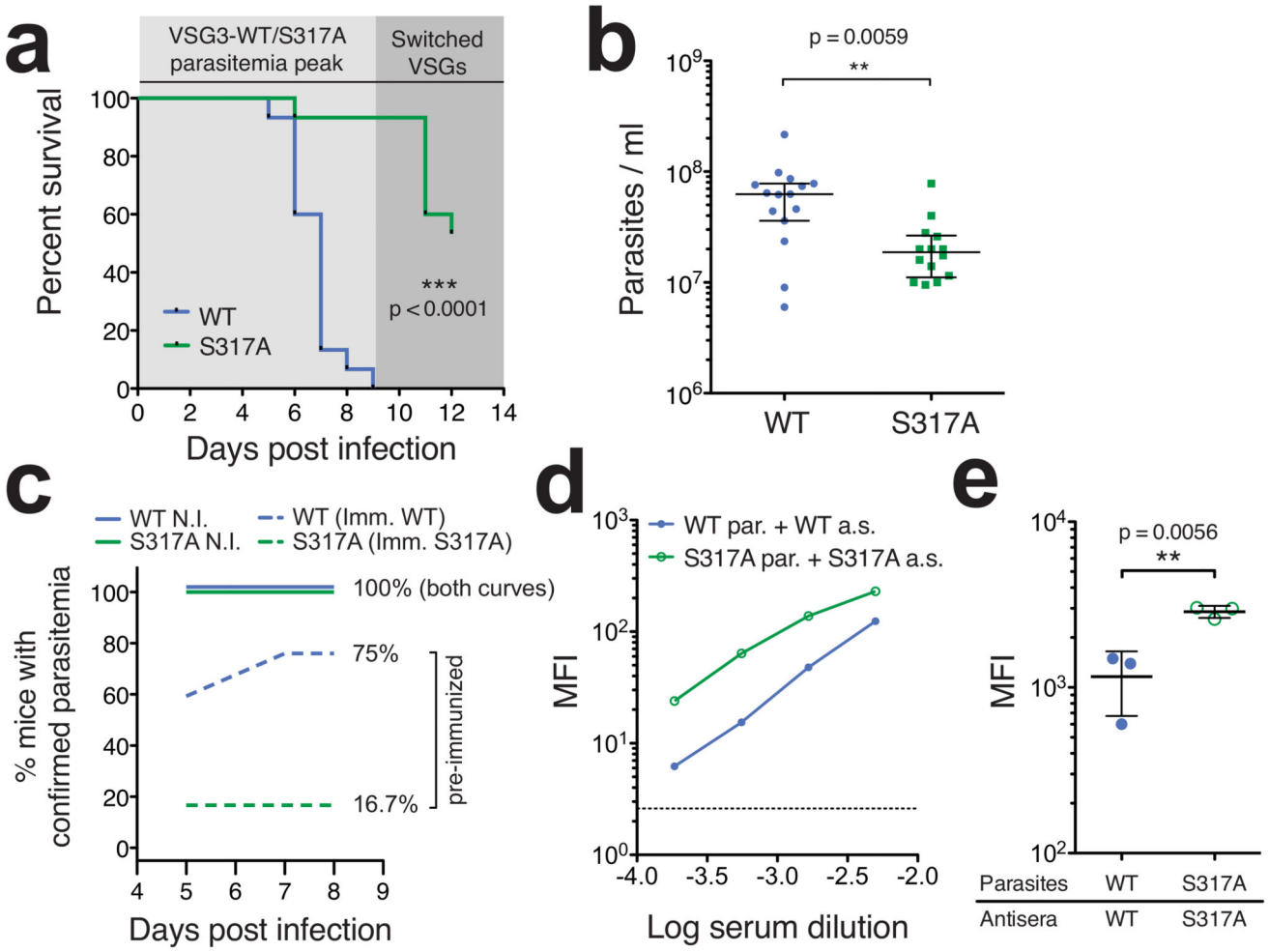

Fig. 4. Presence of *O*-linked glycan impairs immune function.

(a) Survival of mice infected with trypanosomes expressing $VSG3_{WT}$ or non-glycoslyated mutant $VSG3_{S317A}$ (n = 15 mice/group). Shaded regions represent sequential "peaks" of parasitemia. Data are combined results from two experiments with independent clone sets and representative of results within each set (n = 5 or 10 mice/group, see Supplementary Fig. 8). Statistical analysis was performed using a two-sided log-rank test (Mantel-Cox). **(b)** Parasitemia levels at day 5 post-infection. Data are combined from two independent clone sets as in (a) (separated in Supplementary Fig. 8). Statistical values were calculated by two-sided Mann-Whitney test; bars represent median and interquartile range **(c)** Emergence of parasitemia in non-immunized (N.I.) and pre-immunized mice. N.I. data are from infections shown in (a) and (b) (n = 15 mice/group). Pre-immunized mice (n = 12 mice/group) were inoculated with UV-irradiated trypanosomes expressing $VSG3_{WT}$ or $VSG3_{S317A}$ (Imm. WT or S317A, respectively) on days -8 and -5 relative to challenge. Challenge injection contained live parasites expressing the same VSG to which the mice were immunized. Mice were permanently scored as "parasitemia confirmed" when parasites were first detected via blood smear. **(d)** Comparison of $VSG3_{WT}$ or $VSG3_{S317A}$-elicited IgM antisera (a.s.) binding to live $VSG3_{WT}$ or $VSG3_{S317A}$-expressing parasites (par.). Binding was measured by FACS

analysis using dilutions of primary antisera, visualized after counterstaining with FITC-labeled anti-mouse-IgM, and reported as mean fluorescence intensity (MFI). Data are shown for one antiserum sample per VSG type. This experiment was performed twice with similar results. **(e)** Similar to (d) but with a single dilution (1:2000) of distinct antisera, elicited by infection of multiple mice (n = 3 independent antisera/VSG type; error bars represent mean ± SD). Statistical values were calculated by two-sided *t*-test, assuming normal distribution, and similarity of variances was confirmed by F test. For additional statistical analyses see Supplementary Fig. 9. For (d) the dotted line represents background MFI. This experiment was performed once.

Europe PMC Funders Author Manuscripts