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

Jason Pinger1,2, Dragana Nešić2,7,9, Liaqat Ali3,9, Francisco Aresta-Branco4,5, Mirjana Lilic2,8, Shanim Chowdhury1, Hee-Sook Kim1, Joseph Verdi9, Jayne Raper6, Michael A. J. Ferguson3* and C. Erec Stebbins5*

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 1.4-Å-resolution crystal structure of the variant VSG3 manifests convergence in the tertiary fold and oligomeric state. The structure also reveals an O-linked carbohydrate on the top surface of VSG3. Mass spectrometric analysis indicates that this O-glycosylation site is heterogeneous occupied in VSG3 by zero to three 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.

1The Rockefeller University, Laboratory of Lymphocyte Biology, New York, NY, USA. 2The Rockefeller University, Laboratory of Structural Microbiology, New York, NY, USA. 3Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, UK. 4Division of Immune Diversity, German Cancer Research Center, Heidelberg, Germany. 5Division of Structural Biology of Infection and Immunity, German Cancer Research Center, Heidelberg, Germany. 6Department of Biological Sciences, Hunter College, City University of New York, New York, NY, USA. 7Present address: The Rockefeller University, Allen and Frances Adler Laboratory of Blood and Vascular Biology, New York, NY, USA. 8Present address: The Rockefeller University, Laboratory of Molecular Biophysics, New York, NY, USA. 9These authors contributed equally: Jason Pinger, Dragana Nešić, Liaqat Ali. *e-mail: m.a.j.ferguson@dundee.ac.uk; n.papavasiliou@dkfz-heidelberg.de; e.stebbins@dkfz-heidelberg.de

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 1.4-Å-resolution crystal structure of the variant VSG3 manifests convergence in the tertiary fold and oligomeric state. The structure also reveals an O-linked carbohydrate on the top surface of VSG3. Mass spectrometric analysis indicates that this O-glycosylation site is heterogeneous occupied in VSG3 by zero to three 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 variant surface glycoprotein (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 parasitaemia that are the repeated outcome of cycles of VSG switching, antibody generation and parasite killing.

Despite the centrality of the VSG–antibody interaction, the key parameters involved 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 infection. 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 on the basis of crystal structures of the amino (N)-terminal domains (NTDs) of two VSGs. These early structural analyses revealed three key features: 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 three-helix bundle (Fig. 1a); the VSGs formed homodimers; and 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 NTDs to different classes based on the presence of numerous analogous cysteine disulfides. 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 VSG2) 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 biochemical features as compared to previously published VSGs. Most notably, an unexpected O-linked glycan was identified in the VSG3 crystal structure at residue Ser 317, located on the top surface of the NTD. Mass spectrometry (MS) analyses revealed heterogeneity in the number of hexoses linked to VSG3 Ser 317, and also that analogous modifications are present in other class B VSGs. Finally, mouse infection and...
by these analyses, we proceeded to crystallize and solve the structure of the VSG3 NTD. 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 and Supplementary Fig. 1). Consistent with such variation, the oligomerization state of the VSG3 NTD is also different, existing in the number of oligomers present (Supplementary Fig. 5).

**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’.

The 1.4-Å-resolution crystal structure reveals that although VSG3 is heterogeneous O-glycosylated, a striking feature of the VSG3 structure is an unexpected carbohydrate modification (Fig. 2a,b). Additional density attached to Ser 317 revealed the presence of an α-linked glucose moiety (Fig. 2b). This modification is located on the top surface of VSG3 in an exposed loop stabilized by flanking cysteine residues involved in a disulfide bridge (the peptide CTS2SEGLC, residues 314–323, Fig. 3a), and is therefore likely to be highly accessible to the immune system.

**Viability of antibody binding assays showed that the VSG3 O-glycan potently increases parasite virulence.** As a result 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 VSG3 variants expressing site, replacing the VSG2 gene with a S317A mutation which cannot be O-glycosylated. This was achieved by knocking the genes for these coats 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 defined the O-glycosylation site as Ser 317 (Supplementary Fig. 6). Together, these data strongly suggest that there is a heterogeneous chain of 0–3 hexoses attached to Ser 317. On the basis of the electrospray MS spectra (Supplementary Fig. 5a,b), we estimate the proportions of the 0, 1, 2 and 3 hexose glycoformas 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 kinetoplastids, such modifications have not previously been described in T. brucei. To our knowledge, this is also the first description of a Glc1- O-Ser linkage in any organism, although Glcβ1- O-Ser linkages are known. In addition, no carbohydrate of any kind has previously been 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 locations with NTDs 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 analysed two of these: VSG11 and VSG615. We also examined VSG21, which fails to thread to VSG3. MS analysis showed that both VSG11 and VSG615 possessed heterogeneous O-linked hexose modifications analogous to those found in VSG3 (Fig. 3d,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. As a result 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 VSG (VSG3wt) or a VSG3 coat with a S317A point mutation (VSG3S317A, which cannot be O-glycosylated). This was achieved by knocking the genes for these coats 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 VSG. To compare the ability of the two isogenic strains to evade this IgM response, we focused on the first peak of parasitaemia (up to day 9 post-infection). During this period, outgrowth of parasites that had switched expression to other VSGs.

VSG3 O-glycosylation impairs host development of protective immunity. To address this hypothesis, we performed additional infection experiments in which mice were immunized against VSG3WT or VSG3S317A before infection (via inoculation with ultraviolet-irradiated parasites—see Methods). This protocol was previously shown to yield an immune response that closely mimics the response to the infective parasite. 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,b) uniformly presented parasitaemia by day 5. We found that pre-immunization with VSG3S317A was highly protective, suppressing detectable outgrowth of cognate VSG3WT trypanosomes in 83.3% (10/12) of mice (Fig. 4c). In contrast, pre-immunization with VSG3WT was only marginally protective against cognate VSG3WT trypanosomes, suppressing outgrowth in 25% (3/12) of mice. All mice that presented parasitaemia 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 VSG3WT or VSG3S317A-coated parasites at day 8 post-infection.
We then used flow cytometry (fluorescence-activated cell sorting (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 clearance\(^1\)). 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,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,f) together 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. Unexpectedly, we have also identified a VSG...
post-translational modification 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. Moreover, other unidentified types of VSG modification 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. 3a–e and 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: the binding epitope of each antibody; whether the epitopes overlap with or are affected by the glycan, and if so, whether the glycan affects each interaction enough to significantly alter antibody binding; and, finally, 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 previously9). 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 (that is, the affected epitope need not actually include the glycan). The impact of differential glycosylation can vary dramatically and unpredictably for individual antibodies.

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 homogeneous 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 surface9. The presence of various glycoforms on a VSG3NT coat could reduce the effective densities of any epitopes altered by glycosylation. In contrast, a VSG3NTA coat would be comparatively homogeneous 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 probably depend on multiple epitopes and the overall structure and packing of VSGs in the surface coat. Such studies could improve our understanding of 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-9® media. Cells were pelleted and then lysed in 0.2 mM ZnCl2, according to established protocols32. The VSG-encoding subtilisin and culture was centrifuged, and the pellet containing the membrane material was resuspended in pre-warmed (40 °C) 10 mM 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 10 mM phosphate buffer. Two passes through this column resulted in the carboxy-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 crystalization. Further purification steps for VSG3 include one or, if necessary, two runs over a gel filtration column (Superdex 200, GE Healthcare) in 10 mM HEPES pH 7.5, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 150 mM sodium chloride (NaCl). Aliquots from the gel filtration 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 2 mg ml−1 in a buffer containing 10 mM HEPES pH 7.5, 150 mM NaCl. Native crystals were grown by vapour 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 3,350 Da, 200–325 mM NaCl and 100 mM Tris pH 8.2. The VSG3NTA crystal structure was obtained by soaking in 2 M PEG 3350 Da, 300 mM NaCl, 100 mM Tris pH 8.2, 10% glycerol and flash-cooled immediately after 100 K (~173.15 °C). Crystals formed in the space group I223 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 3,350 Da, 275–400 mM sodium bromide, 100 mM Tris pH 8.2). Before flash-freezing, those crystals were transferred into cryoprotection buffers (25% PEG 3,350 Da, 50% M Tris pH 8.2, 5% glycerol) containing a stepwise increasing concentration of sodium bromide from 0.5 M to 2 M.

Data were collected at the Advanced Photon Source at Argonne National Laboratory at beamline 24-ID-C and processed on-site through the RAPD software pipeline26–29. Bromine-soaked crystals were collected at a wavelength of 0.9194 Å and the structure was solved by single-wavelength anomalous dispersion from the anomalous signal due to 28 partially ordered bromine ions using the SHELEX suite30. A partial model of the bromine-soaked protein was built by the PHENIX suite31. This model was then used in arp/wARP32 as a starting model for automated building against the native data to 1.4 Å resolution. Several cycles of automated building with arp/wARP, refinement with REFMAC7/21 and manual inspection and model building led to a final model with an Rmerge of 13.0%/16.9%. Two carbohydrates are present in the model—an N-linked Man,GlcNAc, to Asn 67 and an O-linked Glc to Ser 317.

Digestion, de-N-glycosylation and MS. 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), for 2 h, at room temperature) and S-alkylated by adjusting to 100 mM iodoacetamide (IAA), for 1 h, at room temperature. Labeled IAA was quenched by the addition of 4 µl of 1 M DTT and the sample was adjusted to 400 µl containing 2 mM calcium acetate 80 µg Pronase (Sigma-Aldrich). After digestion for 24 h at room temperature, the
Promisc glycopeptides were enriched using cotton wool hydrophilic interaction LC before reverse-phase 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), d-alkylated (25 mM IAA, for 1 h, at room temperature, in the dark), diluted with an equal volume of 2 µl GluC buffer and digested with 1.25 w/w endoproteinase GluC (New England BioLabs) for 24 h, at 37 °C, with shaking. The GluC fragments were separated using Nanosil SCX 40 μm, 40 nL, His tag protein gels (Invitrogen brilliant blue (Thermo Scientific) and the 17kDa fragment was subjected to in-gel trypsin digestion for LC–MS/MS analysis.

LC–MS/MS analysis of glycopeptides. The LC instrument was an Ultimate 3000 Nano LC System (Dionex) fitted with a C18 trap (PepMap nanoViper, Thermo Scientific) and a resolving column (PepMap RSLC) with inner diameters of 100 and 75 µm and lengths of 2 and 50 cm, respectively. The mobile phases were 0.1% formic acid in 2% acetonitrile (solvent A) and 0.08% formic acid in 80% acetonitrile (solvent B). The 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 35–1800 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. The raw files, converted to mgf format by MConvert (http://proteowizard.sourceforge.net), were searched against the VSG3 protein sequence using Mascot software with the parameters set as follows: peptide tolerance, 5 ppm; MS/MS tolerance, 0.5 Da; enzyme, trypsin; one missed cleavage allowed; and fixed carboximidoethyl 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 2,500 units of glycosyl-free neuraminidase F (PNGaseF; NEB) in 1% w/v octyl-p-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 (16,000 g, for 20 min, at 4 °C), washed with cold acetone, air-dried and redissolved in 1% formic acid. A control (non-treated) VSG3 sample was similarly trichloroacetic acid-precipitated and redissolved in 1% formic acid. The samples were analysed by LC–MS using an Agilent 1200 System for 20 min (Agilent), with an internal mass range of 25 µm and a length of 360 mm. The mobile phases were as the same. The 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 system was coupled to an Agilent 6520 Q-ToF mass spectrometer. The spray voltage was 1.7 kV and the ion transfer was set at 360 °C. MS scans were collected over m/z 335–3,200 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 VSG 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 supervector (Promega A1360), and the S317A mutation was introduced by QuikChange PCR. Wild-type and mutant VSG3 genes were then cloned into the vector pUB9 (replacing VSG17 in the original vector via HindIII and BamHI sites) to create pSU7 and pSU8. VSG3 and VSG3–S317A were amplified from pSU7 and pSU8, respectively, using primers 224_BsiWI_For and 224_BsiWI_Rev (see above) (these PCR fragments were amplified from pSUN7 and pSUN8, respectively, using primers 224_BsiWI_For and 224_BsiWI_Rev (see above)). These plasmids were then linearized with XhoI and transfected into a parental 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 Amaza nucleofector protocol as previously described. BSD resistant clones were selected on the basis of expression of VSG3 (by FACS) and counterselected for dual expressors (which co-expressed VSG2; also by FACS), using monoclonal antibodies against VSG2 and VSG3. VSG2 counterselection was used to eliminate a subset of clones that integrated the expression cassette ectopically. Two sets of clones were generated. Transfection of VSG3 and VSG3–S317A into the Lister427 single marker cell line produced clones KL.VSG3WT–1 and KL.VSG3S317A–1, respectively. Transfection of VSG3 and VSG3–S317A into a different VSG2-expressing Lister427 cell line (GPI-PLC null) produced clones KL.VSG3WT–2 and KL.VSG3S317A–2, respectively. KL.VSG3WT–2 and KL.VSG3S317A–2 were subsequently stained with a GPI-PLC antibody. The loss of a GPI-PLC gene does not affect parasite virulence or the course of parasitaemia and should not affect interpretation of the mouse infection results shown. The presence of the single point mutation within VSG3_S317A was verified in the KL.VSG3_S317A clones through real-time PCR amplification and sequencing.

Mouse infection assays. For Fig. 4a,b, groups of C57BL/6 or CD-1 mice were infected via intraperitoneal (i.p.) injection of 100 live trypanosomes, expressing either VSG3WT or VSG3S317A. For Fig. 4c and Supplementary Fig. 9a, immunization–challenge experiments were performed as previously described using C57BL/6 mice. For immunization, mice received an i.p. injection with 2×10^6 ultraviolet-irradiated trypanosomes (KI.VSG3−2 or KI.VSG3S317A−2; these clones lack the endogenous lipase GPI-PLC, and therefore retain intact VSG coats following irradiation) in 100 µl HMI-9, and an identical repeat injection 3 days later. The subsequent challenge infection was administered 8 days after the initial immunization, and consisted of 100 live trypanosomes (a final challenge dose of 10^7 trypanosomes). All animal experiments were approved by Rockefeller University’s institutional animal care and use committee under protocol no. 16984.

Elicitation of VSG3WT and VSG3S317A-specific polyclonal IgM antisera. C57BL/6 mice were infected i.p. with 1×10^7 VSG3WT or VSG3S317A-expressing trypanosomes in HMI-9 (three mice each). After 4 days, infections were cleared with 250 mg berenil per mouse i.p., and berenil treatment was repeated after 24 h. Clearance was necessary to prevent VSG3WT-infected animals from succumbing to infection before developing significant antibody titres. On day 8 post-infection, 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. One million VSG3WT or VSG3S317A-expressing trypanosomes were isolated from culture, resuspended in HMI-9 containing dilutions of VSG3WT or VSG3S317A-specific polyclonal IgM antisera (see above), and incubated for 10 min. Cells were washed once with HMI-9, and then incubated with a 1:10,000 dilution of goat anti-mouse IgM–fluorescein isothiocyanate (FITC) (SouthernBiotech 1021-02) (for exclusion of dead cells from analysis) and analysed using a BD-FACSCalibur (Fig. 4c) or a BD-LSRII flow cytometer (Fig. 4d) and FlowJo software (Treestar).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Coordinates and structure factors of VSG3 have been uploaded to the RCSB PDB with PDB ID: 6ELC.

Received: 14 May 2018; Accepted: 4 June 2018; Published online: 9 July 2018

References
1. Matthews, K. R., McCulloch, R. & Morrison, L. J. The within-host dynamics of African trypanosome infections. Phil. Trans. R. Soc. B https://doi.org/10.1098/rstb.2014.0288 (2015).
2. Hsie, R., Beals, T. & Boothroyd, J. C. Use of chimeric recombinant polypeptides to analyse conformational, surface epitopes on trypanosome variant surface glycoproteins. Mol. Microbiol. 19, 53–63 (1996).
3. Schwede, A., Macleod, O. J., MacGregor, P. & Carrington, M. How does the VSG coat of bloodstream form African trypanosomes interact with external proteins? PLoS Pathog. 11, e1005259 (2015).
4. Metcalfe, P., Blum, M., Freymann, D., Turner, M. & Wiley, D. C. Two variant surface glycoproteins of *Trypanosoma brucei* of different sequence classes have similar 6 Å resolution X-ray structures. *Nature* 325, 84–86 (1987).
5. Blum, M. L. et al. A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature* 362, 603–609 (1993).
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.* 221, 823–835 (1991).
7. Bartossek, T. et al. Structural basis for the shielding function of the dynamic trypanosome variant surface glycoprotein coat. *Nat. Microbiol.* 2, 1523–1532 (2017).
8. Kelley, L. A., Meulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858 (2015).
9. Mendonca-Previato, L., Todeschini, A. R., Heise, N. & Previato, J. O. Prototozan parasite-specific carbohydrate structures. *Curr. Opin. Struct. Biol.* 15, 499–505 (2005).
10. Takeuchi, H., Kathandra, J., Sethi, M. K., Bakker, H. & Haltwanger, R. S. Site-specific O-glycosylation 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.* 287, 33934–33944 (2012).
11. Kabsch, W. Crystallography of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 72, 799–819 (2006).
12. Cross, G. A., Kim, H. S. & Wickstead, B. Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.* 195, 59–73 (2014).
13. Javiran, S. J., Chowdhury, S. & Papavasiliou, F. N. Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. *Nat. Commun.* 8, 828 (2017).
14. Lisowska, E. The role of glycosylation in protein antigenic properties. *Cell Mol. Life Sci.* 59, 445–455 (2002).
15. Pinger, J., Chowdhury, S. & Papavasiliou, F. N. Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. *Nat. Commun.* 8, 828 (2017).
16. E. The role of glycosylation in protein antigenic properties. *Cell Mol. Life Sci.* 59, 445–455 (2002).
17. Cross, G. A., stocks, 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.* 7, 1102–1122 (2016).
18. Hirumi, H. & Hirumi, K. Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum albumin. *J. Mol. Biol.* 24, 79–90 (1984).
19. Cross, G. A. Release and purification of *Trypanosoma brucei* variant surface glycoprotein. *J. Cell. Biochem.* 24, 79–90 (1984).
20. Cross, G. A. Release and purification of *Trypanosoma brucei* variant surface glycoprotein. *J. Cell. Biochem.* 66, 125–132 (2010).
21. Evans, P. A. Scaling and assessment of data quality. *Acta Crystallogr D* 62, 72–82 (2006).
22. Evans, P. A. Introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D* 67, 282–292 (2011).
23. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D* 69, 1204–1214 (2013).
24. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 50, 760–763 (1994).
25. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr A* 64, 112–122 (2008).
26. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66, 213–221 (2010).
27. Langer, G., Cohen, S. X., Lanzlin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* 3, 1171–1179 (2008).
28. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53, 240–245 (1997).
29. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* 67, 355–367 (2011).
30. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. L. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567 (1999).
31. Böhme, U. & Cross, G. A. M. Mutational analysis of the variant surface glycoprotein GPI-anchor signal sequence in *Trypanosoma brucei*. *J. Cell Sci.* 115, 805–816 (2002).
32. Burkard, G., Fragoso, C. M. & Roditi, I. Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 153, 220–223 (2007).
33. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. A tightly regulated inductive expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 99, 89–101 (1999).
34. Leal, S. et al. Virulence of *Trypanosoma brucei* strain 427 is not affected by the absence of glycosylphosphatidylinositol phospholipase C. *Mol. Biochem. Parasitol.* 114, 245–247 (2001).

Acknowledgements
We thank G. Cross (Rockefeller University) and H. Wardemann (DKFZ) for critical reading of the manuscript and for general advice, M. Sanches-Vaz and L. Figueiredo (IMM, Lisbon) for help with mouse experiments and M. Chandra (DKFZ) for providing us with purified VSG615. We also thank the staff at Argonne National Laboratories (NE-CAT) for beamline support, and D. Ore at the Structural Biology and the High-Throughput Sequencing and Spectroscopy Resource Centers at Rockefeller University. 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 US 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 F.N.P. and by a Wellcome Trust Senior Investigator Award (101842) to M.A.J.F. The University of Dundee MS facility is supported by Wellcome Trust grant 097045.

Author contributions
J.P., D.N., C.E.S., L.A., M.A.J.F. and F.N.P. conceived and designed the experiments. J.P., D.N., M.L., F.N.P. and F.A.B. carried out the protein purification. C.E.S. and F.A.B. performed the structural prediction analyses. D.N. and C.E.S. carried out the crystallography analyses. L.A. and M.A.J.F. performed the MS analyses. J.P., S.C., F.A.B., F.N.P., J.V. and J.R. carried out the trypanosome genetics and growth analyses, antibody assays and mouse infection studies. D.N., L.A., J.P., M.L., S.C., F.A.B., H.-S.K., F.N.P. and C.E.S. contributed reagents, materials and analysis tools. J.P., C.E.S., F.N.P., D.N., L.A. and M.A.J.F wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41564-018-0187-6.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to M.A.J.F or F.N.P or C.E.S.
Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed |
|-----|-----------|
|   | ☑  | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
|   | ☑  | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
|   | ☑  | The statistical test(s) used AND whether they are one- or two-sided |
|   | ☑  | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
|   | ☑  | A description of all covariates tested |
|   | ☑  | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
|   | ☑  | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
|   | ☑  | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
|   | ☑  | Give P values as exact values whenever suitable. |
|   | ☑  | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
|   | ☑  | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
|   | ☑  | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |
|   | ☑  | Clearly defined error bars |
|   | ☑  | State explicitly what error bars represent (e.g. SD, SE, CI) |

Software and code

Policy information about availability of computer code

Data collection
Crystalllographic data collected at Advanced Photon Source (APS) at Argonne National Laboratory at beamline 24-ID-C and processed onsite through the RAPD software pipeline. Flow cytometry data was collected using FACSDiva software (BD Biosciences).

Data analysis
Crystalllographic data collected at Advanced Photon Source (APS) at Argonne National Laboratory at beamline 24-ID-C and processed onsite through the RAPD software pipeline. Bromine SAD phasing was calculated using the SHELX suite. A partial model of the bromine-soaked protein was built by the PHENIX suite. This model was then used in arp/wARP 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 REFMAC, and manual inspection and model building led to a final model with and R/Rfree of 13.0%/16.9%.

LC-MS/MS analysis of glycopeptide 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). LC-MS analysis of intact and de-N-glycosylated VSG3 scans were collected over m/z 335-3200 and spectra were deconvoluted using the Maximum Entropy software.

Flow cytometry data was analyzed using FlowJo software. Data plotting and statistical analyses were performed using Prism (Graphpad) and R.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

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

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

We did not have an a priori prespecified effect size from which to preemptively determine an appropriate sample size for our mouse infection experiments (Fig. 4 and Supplementary Figs. 8 and 9) but the sample sizes were large enough to detect the effect observed and to verify reproducibility. Also see statement in Methods, page 19.

Data exclusions

No data were excluded.

Replication

Non-immunized mouse infection findings were reproduced for a total of n = 15 mice/experimental group (Figure 4a and b and Supplementary Fig. 8a-d) as well as n = 6 mice/experimental group for a different mouse strain (Extended Data Fig8e-f). Pre-immunized mouse infection findings were reproduced for n = 12 mice/experimental group (Fig. 4c and Supplementary Fig. 9a). Antibody binding assays were reproduced with 3 independent mouse-derived antiserum samples for each experimental group. All results were found to be reproducible.

Randomization

Mice were assumed to be sufficiently randomized by Jackson Laboratory and no additional randomization was performed.

Blinding

Experimenter were not blinded to mouse group allocation, but measurements (survival assessment, parasitemia counts, FACS MFI) were not subjective.

Reporting for specific materials, systems and methods

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ✓   | Unique biological materials |
| ✓   | Antibodies |
| ✓   | Eukaryotic cell lines |
| ✓   | Palaeontology |
| ✓   | Animals and other organisms |
| ✓   | Human research participants |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ✓   | ChIP-seq |
| ✓   | Flow cytometry |
| ✓   | MRI-based neuroimaging |

Antibodies

Antibodies used

We have used our own monoclonals for screening VSG3 and VSG3 mutant knock in clones (reported in PMID: 22952449 and made freely available to the community since then, through the monoclonal antibody facility of Rockefeller University).

Validation

These antibodies were initially validated in PMID: 22952449, and have been used in multiple subsequent studies (e.g. PMCID: PMC5635023).
Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) | Trypanosoma brucei Lister 427 strain; George Cross Laboratory, Rockefeller University. See also Methods, page 18.

Authentication | RNAseq.

Mycoplasma contamination | Tested and found negative.

Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals, ARRIVE guidelines recommended for reporting animal research

Laboratory animals | C57BL/6J and CD-1 mice were used. All mice were wild type males or females, aged 6-9 weeks at experiment start (Jackson Laboratory) [see methods, p. 18-20]. All experiments described were approved by the Rockefeller University Institutional Animal Care and Use Committee under protocol #16894 and comply with all NIH mandated ethical regulations regarding the use of vertebrate animals. (Statement also included in methods).

Wild animals | This study did not involve wild animals.

Field-collected samples | This study did not involve field-collected samples.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation | See section "FACS analyses to determine IgM binding" in Methods, page 20

Instrument | BD-FACSCalibur and BD-LSRII

Software | FACSDiva software was used for sample collection, FlowJo software was used for analysis.

Cell population abundance | n/a. Samples are not mixed cell types

Gating strategy | Samples are gated via forward and side scatter to eliminate debris and cell aggregates, and dead cells are excluded via elimination of cells which stain positive in the propidium iodide channel (see methods, page 20).

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.