Contents lists available at ScienceDirect

The Cell Surface

Sorting of GPI-anchored proteins at the trypanosome surface is independent of GPI insertion signals

Thomas Henry Miller, Sabine Schiessler, Ella Maria Rogerson, Catarina Gadelha[*]

School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK

ABSTRACT

The segregation of glycosylphosphatidylinositol-anchored proteins (GPI-APs) to distinct domains on the plasma membrane of eukaryotic cells is important for their correct cellular function, but the mechanisms by which GPI-APs are sorted are yet to be fully resolved. An extreme example of this is in African trypanosomes, where the major surface glycoprotein floods the whole cell surface while most GPI-APs are retained in a specialised domain at the base of the flagellum. One possibility is that anchor attachment signals direct differential sorting of proteins. To investigate this, we fused a monomeric reporter to the GPI-anchor insertion signals of trypanosome proteins that are differentially sorted on the plasma membrane. Fusions were correctly anchored by GPI, post-translationally modified, and routed to the plasma membrane, but this delivery was independent of retained signals upstream of the ω site. Instead, ω−minus signal strength appears key to efficacy of GPI addition and to GPI-AP cellular level. Thus, at least in this system, sorting is not encoded at the time of GPI anchor addition or in the insertion sequence retained in processed proteins. We discuss these findings in the context of previously proposed models for sorting mechanisms in trypanosomes.

Introduction

Proteins destined to be anchored to the plasma membrane via glycosylphosphatidylinositol anchors are identified in the cell by the presence of a cleavable N-terminal signal peptide and C-terminal GPI-insertion signal (Lemus et al., 2023). These sequences direct newly synthesised proteins to the secretory pathway for terminal delivery at the plasma membrane. Cell surface polarity brings an additional level of complexity to this, whereby diffusion barriers define and regulate specialised membrane domains to which specific GPI-APs are delivered and retained (Trimble and Grinstein, 2015). The cellular machinery that decodes sorting signals to ensure correct delivery remains elusive, and the identity of the different signals themselves is also largely unknown.

In fungi, the primary signal for sorting appears to be the ω site environment: in yeast, 2 amino acids in the ω minus region are determinants of plasma membrane sorting and retention; whilst in *Aspergillus*, only 1 amino acid is needed (Caro et al., 1997; Frieman and Cormack, 2004; Ouyang et al., 2013). In mammalian cells, the GPI-anchor insertion sequence has been implicated in providing the signal for apical or basolateral sorting in polarised epithelia (Paladino et al., 2008; Bate et al., 2016; Puig et al., 2019), though this is not the sole

mechanism at play, with N-glycans being sufficient to target GPI-anchored and transmembrane proteins to the apical domain (Benting et al., 1999). Here we investigate if the rules guiding GPI protein sorting in fungi and animals are conserved more widely in eukaryotes by testing for their presence in trypanosomes, which are members of a clade that diverged from Opisthokonts close to the base of extant eukaryotes. *Trypanosoma* and *Plasmodium* (also an early diverging organism) are extreme examples of sorting, as they produce vast amounts of variant GPI-APs to evade host immune attack. Indeed, the GPI anchor structure was first solved in *T. brucei* (Ferguson et al., 1988), and this system remains one of the best studied outside of the Opisthokonta due to its polarised cell surface, good definition of a wide range of surface proteins, and the molecular tools with which hypotheses about membrane protein sorting and retention can be tested (Gadelha et al., 2009, Lacomble et al., 2009, Gadelha et al., 2015).

We took the availability of a large portfolio of validated surface proteins to test the influence of the GPI signal. We first asked whether different GPI anchoring sequences had an impact on the localisation of exogenous proteins. We then engineered a series of domain exchanges and deletions to dissect the influence of individual signal components. Our results show that signal sequences regulate efficacy of GPI addition

---

but does not constrain surface domain localisation.

Methods

Cell culture and transfection

Bloodstream-form *Trypanosoma brucei* Lister 427 single marker line (Wirtz et al., 1999) was grown in IMDM-based medium supplemented with 15 % foetal bovine serum and 0.1 µg ml^{-1} Puromycin. All experiments were carried out with logarithmically growing cells. 2.5×10^7 cells were transfected with 10 µg linearised plasmid DNA in Tb-BSF buffer (90 mM NaHPO4, 5 mM KCl, 0.15 mM CaCl$_2$, 50 mM HEPES pH 7.3; Burkard et al., 2007) using an Amaxa Nucleofector 2b device (Lonza) with program 'Z-001'. Selection was applied by addition of 5 µg ml^{-1} Hygromycin B.

Protein sequence prediction

GPI anchor insertion was predicted by PredGPI (Pierleoni et al., 2008). Proteins were considered only if they were a PredGPI hit with false-positive rate ≤ 0.01 and also had SignalP peptide prediction (Nielsen et al., 1997; Nielsen and Krogh, 1998) using the hidden Markov model methodology, "eukaryotic" settings and thresholds of $p \geq 0.9$) with $p \geq 0.7$ (because only proteins directed to the endoplasmic reticulum are processed for anchor addition). N-linked glycosylation was predicted by NetNGlyc (Gupta and Brunak, 2002).

Generation of fluorescent protein fusions

Fusion proteins were created using pSiG, a genetic toolkit specifically designed for tagging of GPI-anchored protein genes (Gadelha et al., 2015). The vector pSiG-HhsfG purposely includes processing signals (trypanosome signal peptide and GPI-anchor addition sequences derived from VSG-2, gene ID Tb427.BES40.22) flanking a superfolder GFP (sfGFP) with improved folding dynamics and greater resistance to changes in redox environments encountered in the secretory pathway, plus an epitope tag (3xHA) (Gadelha et al., 2015). A base construct for ectopic constitutive expression of membrane-bound fluorescent protein from the parasite β-tubulin locus was derived from pSiG-HhsfG by incorporating a fragment of *TUBB* ORF with an internal linearization site (NotI), to generate pSiG-HhsfG-Tub.

GPI insertion signal sequences from ESAG2 (Tb427.BES40.18), ESP5 (Tb927.5.291b), ESAG6 (Tb427.BES40.3) and GRESAG9 (Tb927.5.120) were amplified by PCR from genomic DNA (Supplementary Fig. 1A) and cloned into pSiG-HhsfG-Tub between XbaI and BamHI restriction sites to replace the VSG-2 signal. Further modifications were made to the ω−4 to ω−1 positions of pSiG^{ESAG2}-HhsfG-Tub, using annealing primers listed in Supplementary Table 1. All constructs were amplified in XL1 Blue *Escherichia coli* and sequence confirmed by Sanger sequencing. Integration into the correct genomic locus (the β-tubulin) was verified by multiplex PCR from gDNA using forward primer CTACCTGACAGCGTCTG and reverse primers GATGCAGATAGCCTCACG and CACTAGAGCTTATTTTATGGCAGC in a primer-limited manner.

To generate pSiG-HhmScaI-Tub, the mScarlet-I ORF was created by DNA synthesis (Integrated DNA Technologies IDT) and used to replace the sfGFP ORF in pSiG-HhsfG-Tub. Both fluorescent protein sequences were codon optimised to accommodate the trypanosome codon bias.

Live cell native fluorescence microscopy

For analysis of localisation of membrane-bound fluorescent proteins by native fluorescence, 1×10^6 live cells were harvested from mid-log phase cultures, washed in phosphate-buffered saline supplemented with 20 mM glucose (PSG; 137 mM NaCl, 3 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$), and resuspended in ~10 µL of PSG. 2 µL of this concentrated live cell suspension were transferred to glass slides and imaged using an Olympus BX51 microscope equipped with a 100x UPlanApo objective (1.35NA; Olympus) and CoolSnap-HQ CCD camera (6.45 µm pixel^{-1}, Photometrics) without binning. All images of fluorescent cells were captured at equal exposure settings without prior illumination (unless stated otherwise in figure legend). Images for level comparison were also processed in parallel with the same alterations to minimum and maximum display levels, except where stated. Image acquisition was controlled by µManager open source software (Edelstein et al., 2014). Processing and analysis were performed in ImageJ (Schneider et al., 2012).

Native fluorescence microscopy of chemically-fixed cells

For illustrative examples of cells with differential fluorescent protein localisation shown in Supplementary Fig. 2A, GG^{VSG2} or trypanosomes transfected with the construct pGad8-Tub (for inducible expression of GFP alone (no signal sequence for association with a membrane moiety) from the β-tubulin locus, as described in Wickstead et al. 2003) were harvested as above, allowed to adhere onto derivatized glass slides for 2 min (at density of 2×10^7 cells µL^{-1}), fixed for 10 min in 2.5 % w/v formaldehyde, and mounted in a solution containing DAPI and a photostabilizing agent (1 % w/v 1,4-Diazabicyclo(2.2.2)octane, 90 % v/v glycerol, 50 mM sodium phosphate pH 8.0, 0.25 mg/mL^{-1} 4′,6-diamidino-2-phenylindole).

Immunoblotting

Lysates from 2×10^7 cells were resolved by reducing SDS-polyacrylamide gel electrophoresis and electro-transferred onto nitrocellulose membrane (GE healthcare) in 25 mM Tris, 192 mM glycine, 0.02 % SDS, 10 % methanol. Membranes were blocked with 5 % (w/v) skimmed milk in TBS-T (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 % (v/v) Tween-20) and protein detected by 800 ng ml^{-1} mixture of two anti-GFP monoclonal antibodies (7.1 and 13.1; Roche), 500 ng ml^{-1} anti-mCherry polyclonal antibody (Abcam), or 1:20,000 dilution anti-VSG-2 polyclonal antibody (Gadelha et al., 2015) followed by 160 ng ml^{-1} goat anti-mouse or 450 ng ml^{-1} goat anti-rabbit immunoglobulins conjugated to horseradish peroxidase (Sigma) in TBS-T containing 1 % (w/v) skimmed milk. Antibodies were detected using Western Lightning enhanced chemiluminescence reagent (GE healthcare) captured on X-ray films. For absolute fusion protein quantification, lysates were prepared, resolved and probed as above alongside a dilution series of recombinant fluorescent protein. Quantification-grade imaging of chemiluminescence was capture on a Fusion FX (Vilber).

Hypotonic cell lysis and cellular fractionation

3×10^7 cells were harvested, washed twice in cold (0°C) PSG, and resuspended in 90 µL cold H$_2$O with the following mixture of protease inhibitors: 1 mM EDTA, 5 µM E-64d, 7.5 µM pepstatin A, 50 µM leupeptin, 0.5 mM PMSF and 2 mM 1,10-phenanthroline. The cell suspension was left on ice for 15 min and then split into three equal samples (each containing ~1×10^7 cells). To one, hot (95°C) Laemmli buffer was added to prepare whole cell lysate. Another sample was fractionated into soluble (supernatant) and insoluble (pellet) fractions by centrifugation at 3400 g at 4°C for 5 min and hot Laemmli buffer added to each fraction. The third sample was transferred to 37°C for a further 15 min before being fractionated and solubilised as above. As a negative control to the activity of GPI-PLC, ESAG10 (a transmembrane protein at the parasite plasma membrane) endogenously tagged with sfGFP at its C-terminus (Gadelha et al., 2015) was included in the analysis.

Enzymatic removal of N-linked glycans

Peptide-N-glycosidase F (PNGase-F) was used to assess the glycosylation status of GPI-anchored fluorescent protein fusions. PNGase F

removes all *N*-linked glycans (with the exception of 1,3-core fucosylated species) by cleaving the bond between the innermost N-acetylglucosamine (GlcNAc) and the asparagine residue in glycoproteins. 2×10^7 cells were harvested and washed twice with PSG, resuspended in PSG at 2×10^6 cells μL^{-1} and deglycosylation was carried out according to manufacturer's instructions (NEB). Briefly, one volume of [2×] denaturing buffer (1 % SDS, 80 mM dithioreitol) was added, and samples heated at 95°C for 10 min. Each sample was then divided into 2 parts and treated with or without 1000 U PNGase-F plus 0.1 % Nonidet P40 in 50 mM sodium phosphate pH 7.5 (H_2O was added to the untreated part instead of enzyme) for 1 h at 37°C before solubilisation in Laemmli buffer at 95°C for 5 min. Note that under the latter conditions, required for PNGase F activity, some cellular protease activity may also occur, and a small amount of protein in all cases gets cleaved/degraded.

Results

GPI-anchored fluorescent protein fusions are stably expressed in vivo

To investigate the specific role of GPI insertion signals in the sorting of GPI-APs in African trypanosomes, C-terminal insertion sequences from membrane-associated proteins with different localisations were fused to monomeric superfolder GFP (sfGFP) targetted to the ER by a N-terminal signal peptide. Candidate insertion signals were selected from an infective-form *Trypanosoma brucei* cell surface protein atlas (also called 'surfeome', Gadelha et al., 2015) which constitutes the largest portfolio of validated surface membrane proteins to date in this system. The surfeome includes 51 predicted GPI-APs; their GPI insertion signals (from ω site to C-terminal end) vary in length between 21 and 32 amino acids, with cysteine, threonine, asparagine or serine (in order of predominance) most often as the predicted ω amino acid.

The trypanosome plasma membrane can be conceptually divided into three contiguous regions: the cell body, the flagellum, and a specialized region at the base of the flagellum called the flagellar pocket, which is the sole site of endocytosis and secretion in this organism. We selected 4 surfeome GPI-APs as representatives of distinct domains on the cell surface: ESAG2 (cell body membrane), ESP5 (endosomes), the heterodimeric transferrin receptor, of which the ESAG6 subunit is anchored by GPI (flagellar pocket and early endosomal compartments), and VSG-2 (entire plasma membrane; Fig. 1A,B). We also selected GRESAG9, a non-surface GPI-AP (Gadelha et al., 2015) which localises to the ER (Barnwell et al., 2010) (Fig. 1A,B).

To study exclusively the role of the GPI-anchor insertion sequence, we used "minimal" GPI insertion signals which include only the 4 amino acids upstream of ω (Fig. 1C,D). In all cases, this was sufficient to maintain the prediction for both GPI-anchor status and native ω site prediction in the resultant fusions. GPI-anchored sfGFP fusions (herein termed as "GG$^{<origin>}$", where < origin > refers to the source of the anchor insertion sequence) were stably expressed in vivo, as confirmed by immunoblotting of whole cell extracts (Fig. 1E). No impact on cellular growth, morphology or VSG synthesis was observed upon expression of fusion proteins (Supplementary Figure 3); GG lines grow at the same rate as parental cells (data not shown).

The predicted molecular weight of these fusions (32 kDa) is consistent with the apparent electrophoretic mobility of GG^{ESAG2}, although GG^{ESP5}/GG^{ESAG6} and GG^{VSG2}/GG^{GRESAG9} migrate slower (with apparent mass of ~34 and 36 kDa, respectively). Given that all but the insertion signal of these protein is identical, this heterogeneity in electrophoretic mobility could arise either from differences in the processed protein due to the 4 amino acids upstream of the ω site, or differences in the lipid and/or glycan content of GPI anchors. sfGFP has no predicted glycosylation sites within its ORF, and the sequence C-terminal of ω is removed on anchor addition. However, GG^{GRESAG9} contains the consensus sequon N-X-T/S upstream of the predicted ω site (Fig. 2A), whilst VSG-2 contains two N-glycosylation sites: N263 and N428 (Zitzmann et al., 2000), the latter being 2 residues passed the predicted ω

(Fig. 2A). For comparison, this study used in silico predictions to define the GPI anchor insertion signal for all proteins. As such, the VSG-2 ω site used here was N426, but the physically mapped ω site is S433 (Zitzmann et al., 2000). This would include the N428 in the retained sequence, allowing the post-translational modification of the resultant fusion. In agreement with this, migration of GG^{GRESAG9} and GG^{VSG2} is altered by PNGase-F treatment, demonstrating N-linked glycosylation of these two fusions (Fig. 2B). Likewise, the apparent size of deglycosylated GG^{GRESAG9} falls to the predicted 32 kDa, whereas GG^{VSG2} remains slightly higher than the other fusion proteins (Fig. 2B).

Fluorescent proteins are anchored by GPI to the plasma membrane

To test whether the sfGFP fusions were indeed anchored via a GPI moiety, we used a well-established assay of specific plasma membrane release by GPI-phospholipase C activity. *T. brucei* encodes a GPI-PLC that is expressed in bloodstream forms and localised to the flagellum membrane. At physiological temperature and when plasma membrane integrity is compromised, GPI-PLC-mediated hydrolysis of GPI anchors occurs, releasing GPI-APs into the soluble fraction (Cross, 1984). However, under cold conditions, GPI-PLC is inactive, meaning that GPI-APs remain associated to the plasma membrane and migrate with the insoluble fraction (Fig. 3A,B). To mimic the above conditions, cell lines were osmotically lysed by resuspension in water, and fractionated into soluble (supernatant) and insoluble (pellet) fractions. All five sfGFP-GPI fusions are sensitive to hydrolysis by GPI-PLC (Fig. 3C) and released from the plasma membrane in a manner similar to VSG (arrowhead, Fig. 3C), confirming that all fusions were modified by GPI anchor. GPI-anchored proteins found in intracellular locations (e.g. those in the secretory pathway/en route to the cell surface) are not substrates for the parasite GPI-PLC (Ferguson et al., 1985; Ferguson et al., 1986). This differential accessibility likely accounts for the small amount of protein remaining associated with the cell pellet at 37°C for GG^{ESP5}, GG^{ESAG6} and GG^{GRESAG9} (Fig. 3C).

GPI insertion signal is insufficient to restrict surface domain localisation

All 5 fusion proteins were clearly present at the cell periphery, the flagellar pocket and endosomal compartments when assessed by native fluorescence (Fig. 4, Supplementary Fig. 2), strongly indicating their localisation at the plasma membrane. This is in agreement with accessibility of GPI-PLC, as is the localisation of a proportion of the proteins to the endocytic pathway. The amount of fusion protein that reaches the parasite plasma membrane is very similar between individual GGs, although their intracellular amounts differ (Fig. 4). Significantly, the cellular localisations of the fusion proteins did not recapitulate those of the endogenous proteins from which the GPI anchor insertion signals were derived (see Fig. 1), indicating that although the N-terminus signal peptide and the GPI insertion signal are sufficient to sort an exogenous fluorescent protein to the plasma membrane, they are insufficient to direct any other domain restriction in trypanosomes. Given the small amount of sequence retained, it is perhaps unsurprising that this does not alter the fate of fusion proteins, but these data also strongly suggest that, unlike in other systems, there is no signal in the type of anchor attached to specific sequences in trypanosomes or determination of a particular processing pathway set by the GPI anchor signal.

Surface domain localisation for all proteins here is not the result of formation of fluorescent protein dimers. The superfolder GFP used in this study was originally engineered to be monomeric, and indeed crystallizes as a monomer (Pédelacq et al., 2006), although a weak ability to dimerise has been reported for sfGFP fused onto an endoplasmic reticulum membrane protein (Cranfill et al., 2016). This is particularly important because dimerisation and oligomerisation status of GPI-APs are known to have an effect on their apical fate in epithelial cells (Catino et al., 2008; Lebreton et al., 2021). However, unrestricted surface localisation was unchanged on recreation of a fusion with the
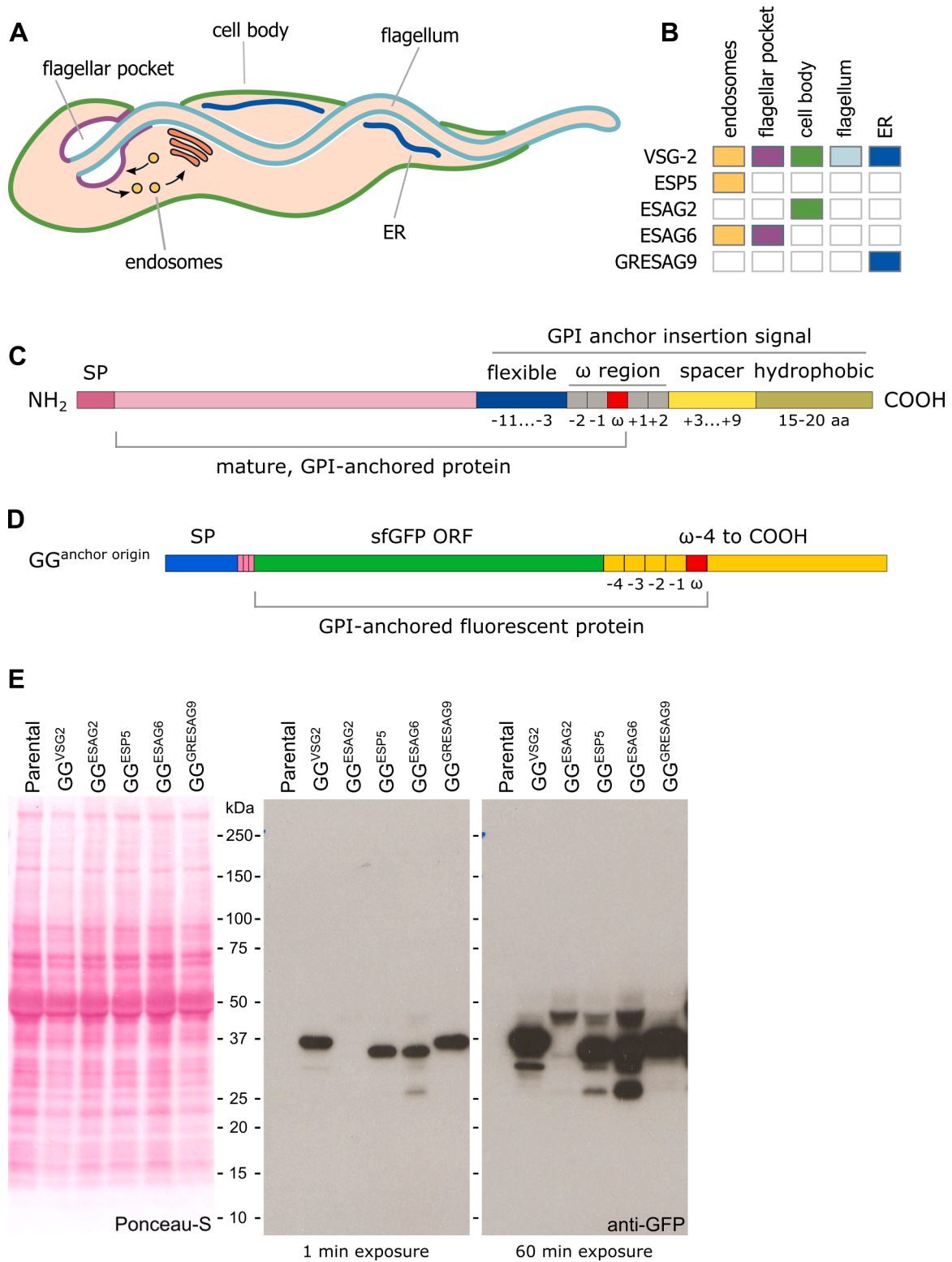
Fig. 1. The cell surface landscape of *Trypanosoma brucei*. A) Cartoon illustrating the parasite's compartmentalised surface membrane and secretory pathway. B) Cellular localisation of trypanosome GPI-APs used in this study (colours according to compartments in A). Data from native fluorescence signal distribution observed for each endogenous-locus tagged cell line from Gadelha et al., 2015. C) Schematic representation of a GPI anchor signal architecture. D) Experimental setup used in this study: representation of superfolder GFP (sfGFP, green) fusion proteins (pre-processing) bearing the VSG-2 N-terminal signal peptide (blue), 3 HA epitopes (pink) and a candidate C-terminal GPI-anchoring signal sequence (yellow, derived from VSG-2, ESAG2, ESP5, ESAG6 or GRESAG9). Cartoons not to scale. E) Ponceau S-stained nitrocellulose membrane of whole cell lysates shows protein loading (left). Immunoblot of sfGFP-GPI fusions using mouse immunoglobulins α-GFP (middle, right). GG^{ESAG2} signal is detected upon long exposure.
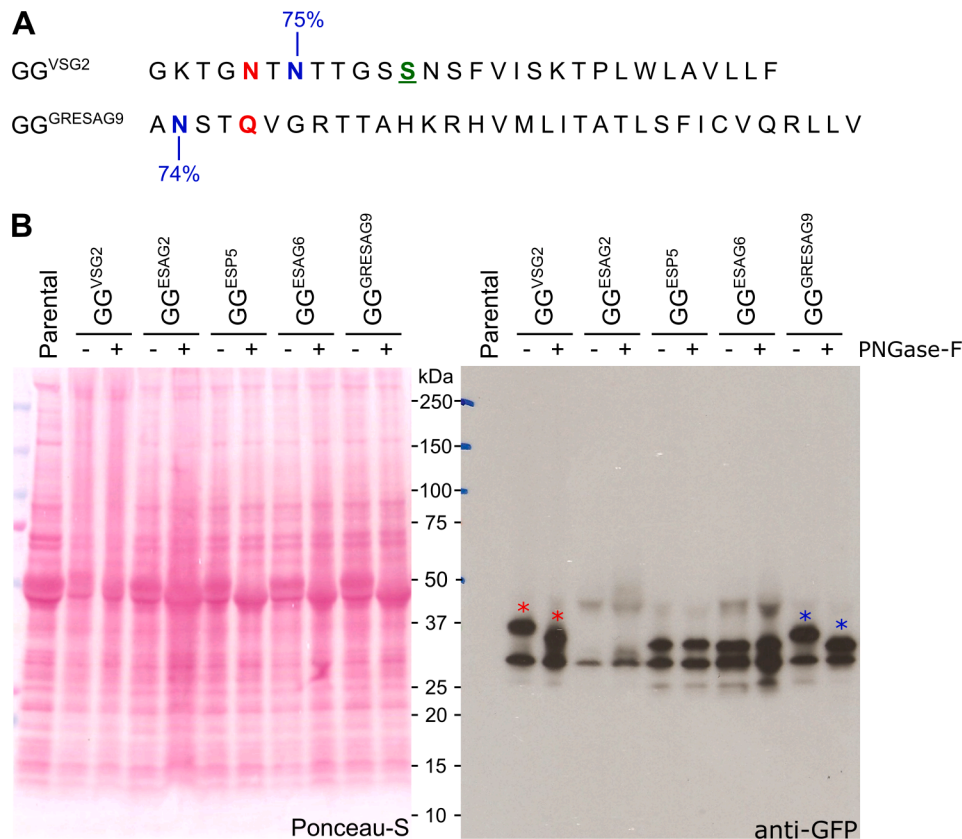
Fig. 2. Differential effect of glycosidase on GPI-anchored fluorescent proteins reveals GG^{VSG2} and GG^{GRESAG9} to be N-glycosylated. **A)** Amino acid sequence from $\omega-4$ to the C-terminus end of GG^{VSG2} and GG^{GRESAG9}. Predicted ω shown in red; predicted N-glycosylation site (and prediction probability) shown in blue. Likely GG^{VSG2} ω site, based on previous physical mapping and presence of N-glycosylation, shown in green underlined. **B)** Immunoblot of whole cell lysates, treated (+) or untreated (−) with the glycosidase PNGase F (cleaves N-linked glycans) reveals an electrophoretic shift for GG^{VSG2} (red star) and GG^{GRESAG9} (blue star) after deglycosylation only. The lower MW band present in all samples (with or without PNGase F) reflects the small amount of degradation/cleavage affecting the fluorescent protein, which is intrinsic to a protease assay such as this.

monomeric red fluorescent protein mScarlet-I (Bindels et al., 2017) (Fig. 4) demonstrating that the effect here is very likely unrelated to oligomerisation of the test proteins.

Signal strength modulates efficacy of GPI addition and protein level

Given that all fusion constructs integrate at the same highly transcribed genomic region (the parasite β-tubulin locus), and have the same 5′ and 3′-UTR sequences, it was expected that protein amounts would be similar. Consistent with this, different fusion proteins have mostly similar levels, ~20,000 molecules per cell (Fig. 5). The exception is GG^{ESAG2} which is lower in abundance, ~800 copies per cell (Fig. 5), in spite of correct integration being verified by diagnostic multiplex PCR (Supplementary Fig. 1B,C).

Differential structural features at and around the GPI insertion site are known to dictate protein-specific processing (Mehlert et al., 1998; Zitzmann et al., 2000). We therefore tested if the surrounding sequence microenvironment to the GG^{ESAG2} anchor insertion signal could influence fusion protein regulation. Sequential removal of $\omega-4$ to $\omega-1$ from GG^{ESAG2} had no impact on the amount of the resultant fusions (Fig. 6), indicating the correct processing of GPI-anchor insertions despite truncations to the signal. These data agree with and expand previous work on the ω−plus region of VSG sequences (Böhme and Cross, 2002). The GPI-anchor insertion signals of VSGs are generally well conserved (being composed of 17 or 23 amino acids, in which ω is Ser, Asp or Asn, $\omega+2$ is serine, and $\omega+7$ almost always lysine) but substantial modification of ω−plus residues of VSG-4 did not affect its cellular levels, GPI anchoring, or targeting to the plasma membrane (Böhme and Cross, 2002).

It is noticeable that the ω−minus region of GG^{ESAG2} contains two charged residues (Fig. 6B), whilst the other fusions (which are more abundant in protein level) have one or none. The successive deletions above did not markedly alter the region's chemical nature (Fig. 6B), at least in terms of amino acid charge. If this feature plays a role in regulating the amount of GG^{ESAG2} protein, then significantly altering that environment may proportionally affect cellular levels. To test this, the 4 amino acids upstream of GG^{ESAG2} ω site were replaced with those from VSG-2, ESP5, ESAG6 and GRESAG9, or for 4 small neutral residues (AAAG) which should act as a flexible linker with minimal interference with the anchor (Fig. 7A,B). These 5 modifications to GG^{ESAG2} were analysed by GPI anchor prediction algorithms to ensure that the ω site would remain the one originally predicted (Fig. 7B). The chimeric anchor signals caused an increase in cellular levels of GG^{ESAG2} (Fig. 7C). The highest level was observed for the artificial ω−minus region created by the 4 neutral residues, followed by the chimeras created with GRESAG9, ESP5, VSG-2 and then ESAG6 ω−minus domains (in order of protein abundance, Fig. 7C). This indicates that, even with an ESAG2 C-terminal end, changing only 4 residues upstream of the ω site improves the strength of the overall GPI insertion signal and, consequently, increases the abundance of the GPI-AP.

Re-engineering of GG^{ESAG2} ω−minus region did not affect its routing to the secretory pathway or ultimate residence at the cell surface (Fig. 8). All modifications made to the ESAG2 $\omega-4$ to $\omega-1$ residues generated localisations not dissimilar to its native signal (Fig. 4). This suggests that, although cellular levels of GG^{ESAG2} are dependent on efficiency of anchor addition alone, cellular location is not: the ω-minus region can be extensively altered by truncations or domain exchange,
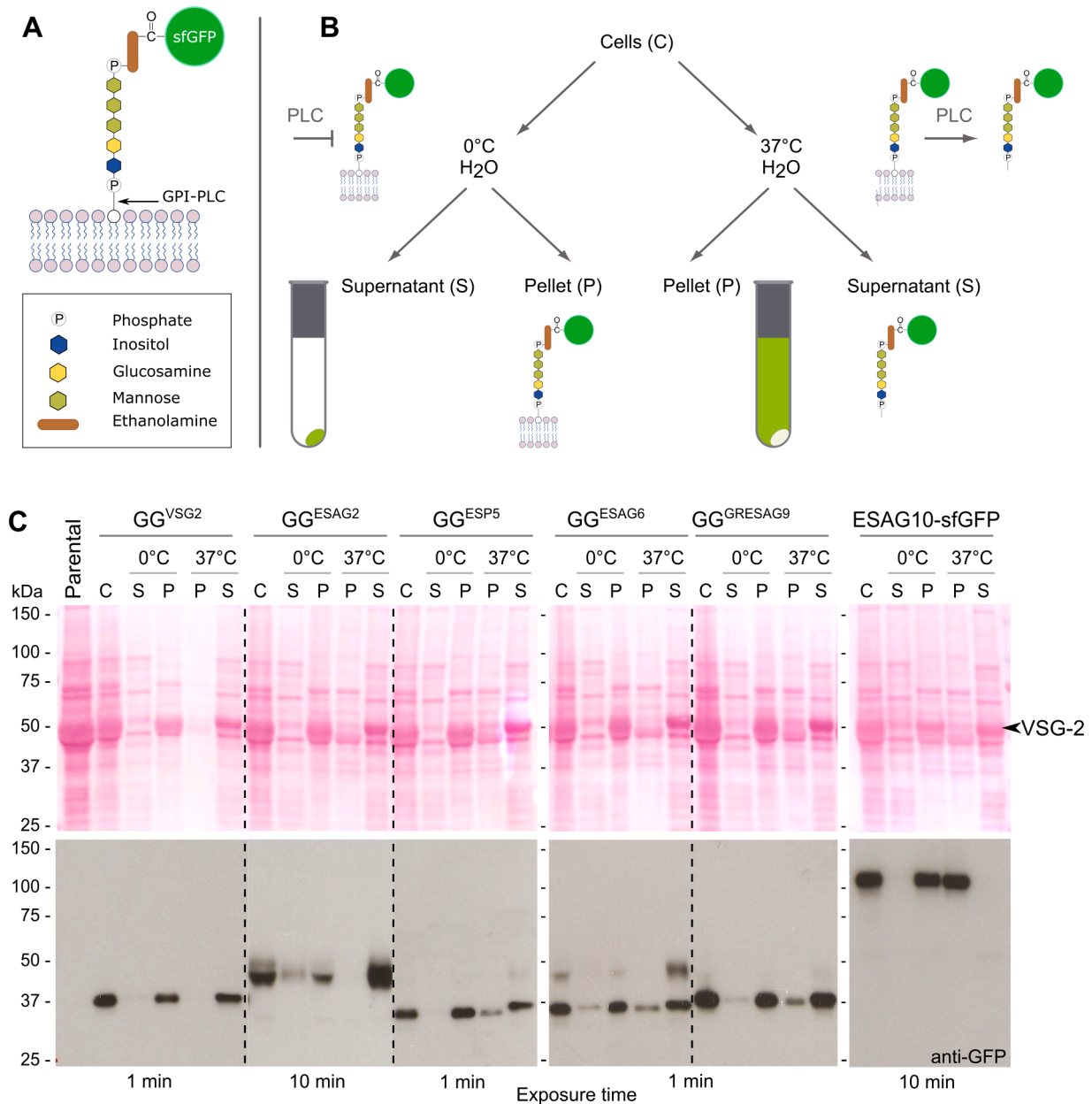
Fig. 3. sfGFP-GPI fusions are efficiently anchored to the plasma membrane via GPI. **A)** *T. brucei* encodes a phospholipase C (PLC) which cleaves GPI anchors upon access to the cell surface (by hypotonic lysis of live cells at 37°C). A simplified structure of the GPI core covalently linked to sfGFP depict an inositol ring in glycosidic linkage to a glucosamine and three mannose residues are linked to ethanolamine by a phosphodiester bond. Cartoon not to scale. **B)** This endogenous PLC can be used to remove the GPI moiety of sfGFP, rendering it soluble upon fractionation. Flow diagram of hypotonic lysis of trypanosomes for endogenous PLC to gain access to the parasite surface membrane. Cells are lysed either warm or cold, and whole cell lysates fractioned into soluble and insoluble fractions. At cold, the PLC is inactive, anchors are not cleaved, and GPI-APs are retained on the parasite plasma membrane (pellet). At 37°C, PLC cleaves the anchor and GPI-APs migrate to the supernatant. **C)** Hypotonic lysis fractions resolved by SDS-PAGE. The visible shift of the VSG-2 band (arrowhead) from the cold pellet to the warm supernatant acts as an internal control. Immunoblot of hypotonic lysis samples resolved by SDS-PAGE above. A shift from pellet to supernatant is seen for all sfGFP-GPI fusions, demonstrating those to be efficiently anchored by GPI at the cell surface. ESAG10 (a transmembrane domain protein at the plasma membrane) serves as control.

with or without change to the overall charge, and levels of processed protein be affected, but the resultant GPI-AP still reaches the parasite cell surface in all cases.

Discussion

Differentiation of cell surface composition is ubiquitous across eukaryotes. It enables structural and functional specialisation that is a fundamental requirement to localising different activities to distinct regions of the plasma membrane. A central question is whether generic or conserved rules govern the delivery and retention of proteins to specialised domains on the surface. For surface proteins anchored by GPI, a number of mechanisms could act to direct protein localisation: 1) the biosynthesis and insertion of membrane domain-specific anchors could enable different proteins to be directed to their respective specialised domain (provided that a cellular machinery is in place to decode and sort these distinct anchors). 2) Alternatively, different GPI-insertion signal sequences could drive proteins into distinct processing routes, each of which resulting in a different endpoint at the plasma membrane. 3) For some/most organisms, there could be no role in specific localisation for GPI-anchor type or insertion sequence, with sub-surface localisation instead linked to generic features of proteins themselves,
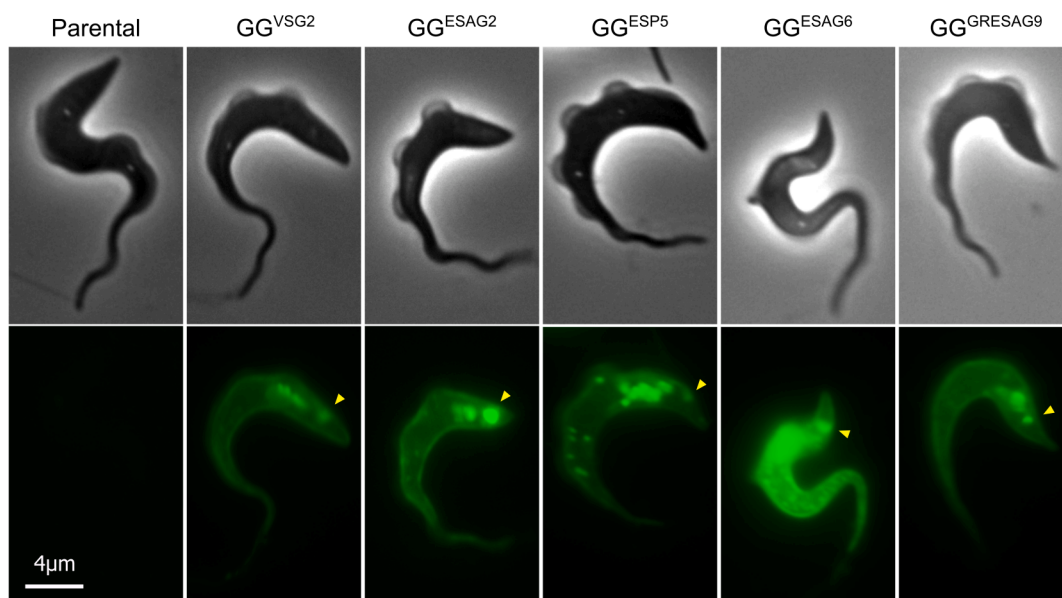
Fig. 4. GPI-anchor insertion signals are sufficient to sort fluorescent proteins to the cell surface, but are not the determinants of surface domain specialisation. Live microscopy of trypanosome lines expressing sfGFP-GPI fusions shows native fluorescent signal on the cell periphery, indicative of parasite cell surface (cell body, flagellum) as well as flagellar pocket membrane (indicated by yellow arrowhead) and endosomes. Signal from superfolder GFP is shown in green. Each image is representative of the distribution observed for the respective cell line. All images captured and processed under equal settings.
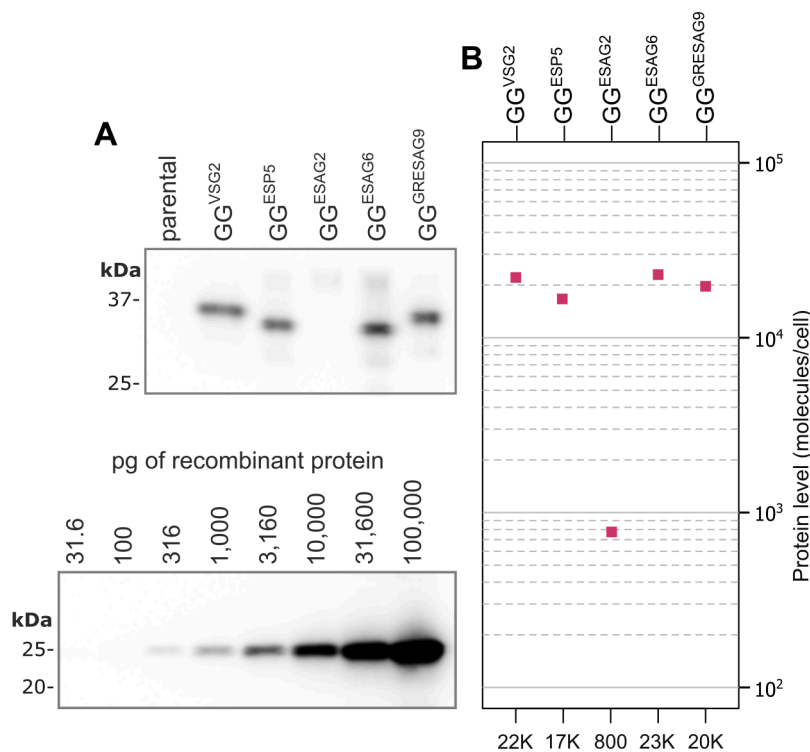

Fig. 5. Fluorescent protein quantification. **A)** Quantitative immunoblotting of whole cell lysates against known amounts of recombinant fluorescent protein, to estimate number of molecules per cell **(B)**.

such as oligomerisation status, abundance, or post-translational modifications. 4) Finally, no generic mechanisms may exist for an organism, with localisation resulting from a potentially complex interplay of multiple effects specific to each individual protein. Below we discuss our findings in African trypanosomes in the context of models previously proposed.

Multi-anchor model

To date, trypanosome GPI structures have only been elucidated for VSG, which represents the vast majority of anchors in a trypanosome cell (VSG constitutes ~10 % of the total cellular proteome and > 90 % of proteins found on the cell surface; reviewed in Borges et al., 2021). Although all characterised GPI anchors share a common core structure
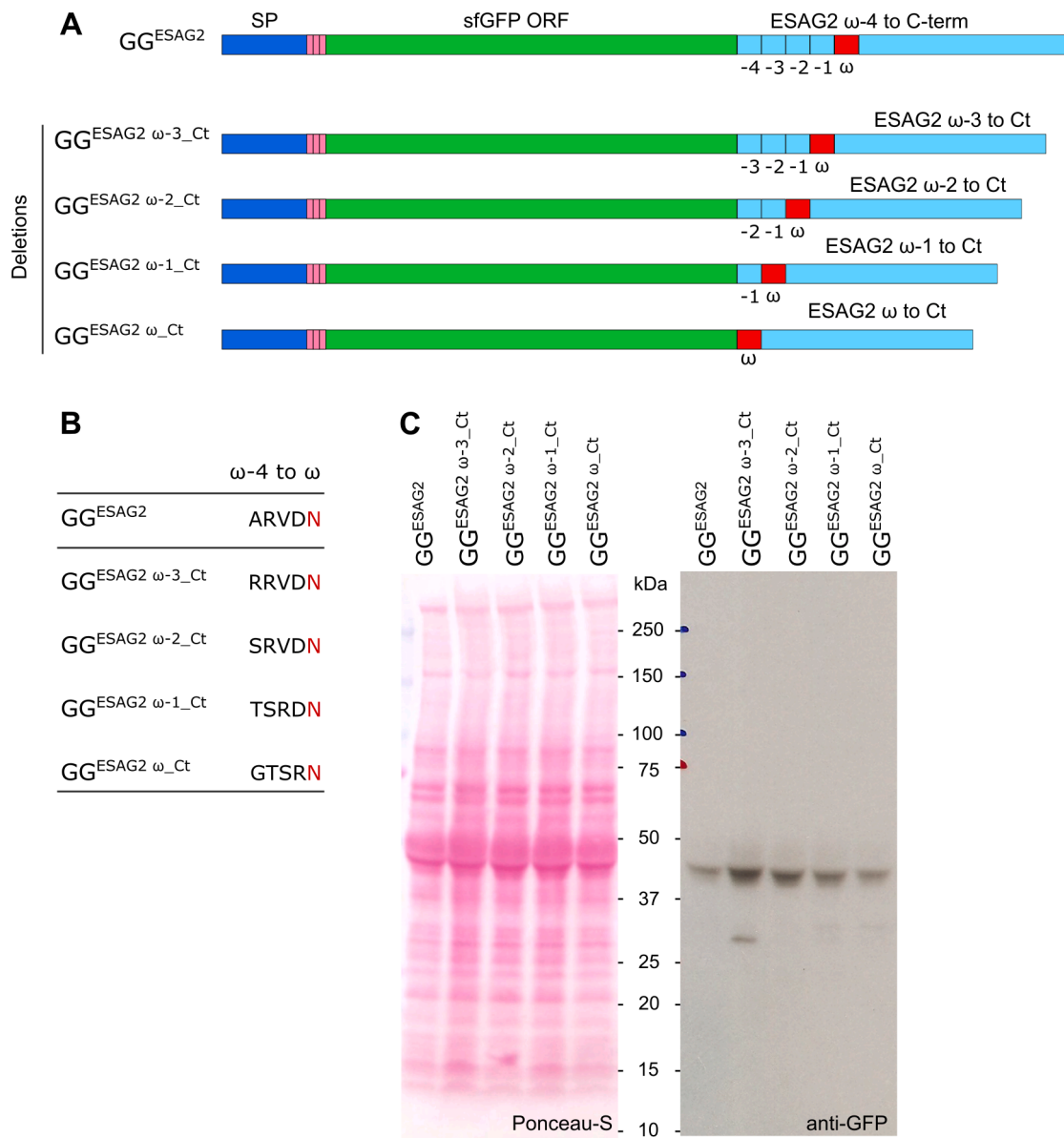
Fig. 6. Truncations to the GPI insertion signal do not alter protein processing. **A)** Schematic representation of modifications generated onto the ESAG2 GPI insertion signal fused to sfGFP. GG[ESAG2] was re-engineered by sequential deletion of ω−4 to ω−1. **B)** Amino acids in positions ω−4 to ω−1 resultant of re-engineering. The ω residue is shown in red. **C)** Effect of deletions to the ESAG2 GPI insertion sequence on protein level and anchoring efficiency.

of ethanolamine-P6Manα1-2Manα1-6Manα1-4GlcNα1-6PI (EtN-P-Man$_3$GlcN-PI), diversity is derived from lipid remodelling, glycan addition, and various substitutions to the core assembly (Ji et al., 2023). In some examples, such as VSG-4, as many as four distinct glycan structure variants were found (Ferguson et al., 1988). This indicates the cellular potential for different anchors being used for the same protein (at least for the VSG protein family). What influence, if any, do these modifications have on the operational sorting of the resultant protein? In the case of VSG-4, probably very little, as the protein can access all domains on the cell surface irrespective of its GPI anchor chemical diversity. The findings reported here suggest that surface protein residence does not seem encoded in the C-terminus of the anchored protein, whereby the molecular nature of the GPI insertion signal sequence either has no effect on the type of anchor being added (and, as such, all GPI-APs are associated with a generic 'cell surface' anchor), or that an anchor-location relationship does not exist (i.e. multiple anchors occur but they bare no influence on where a GPI-AP ultimately resides).

Oligomerisation status

An alternative suggestion was made based on two very closely-related proteins: VSG-2 and the VSG-related transferrin receptor (TfR). VSG-2 is associated to the plasma membrane by 2 GPI anchors (GPI$_2$), whereas TfR is anchored by a single GPI (GPI$_1$). On this basis, and several mutagenesis experiments, the number of anchors in a GPI-AP was proposed as the determinant of intracellular sorting in trypanosomes: GPI$_1$ molecular complexes localise to the flagellar pocket, but GPI$_2$ ones escape pocket retention and localise to the entire cell surface (Schwartz et al., 2005).

As tantalisingly simple and attractive as it is, this model does not account for the scope of VSG structural and biochemical diversity – a family containing >3000 gene copies in the genome of *T. brucei* (Cross et al., 2014). For example, VSG-3, VSG-9, VSG-11, VSG-615 and mVSG-1954 have been shown to exist in solution and in crystal form (and possibly on the membrane) as monomers and trimers (Pinger et al., 2018; Umaer et al., 2021; Chandra et al., 2023; Đaković et al., 2023).
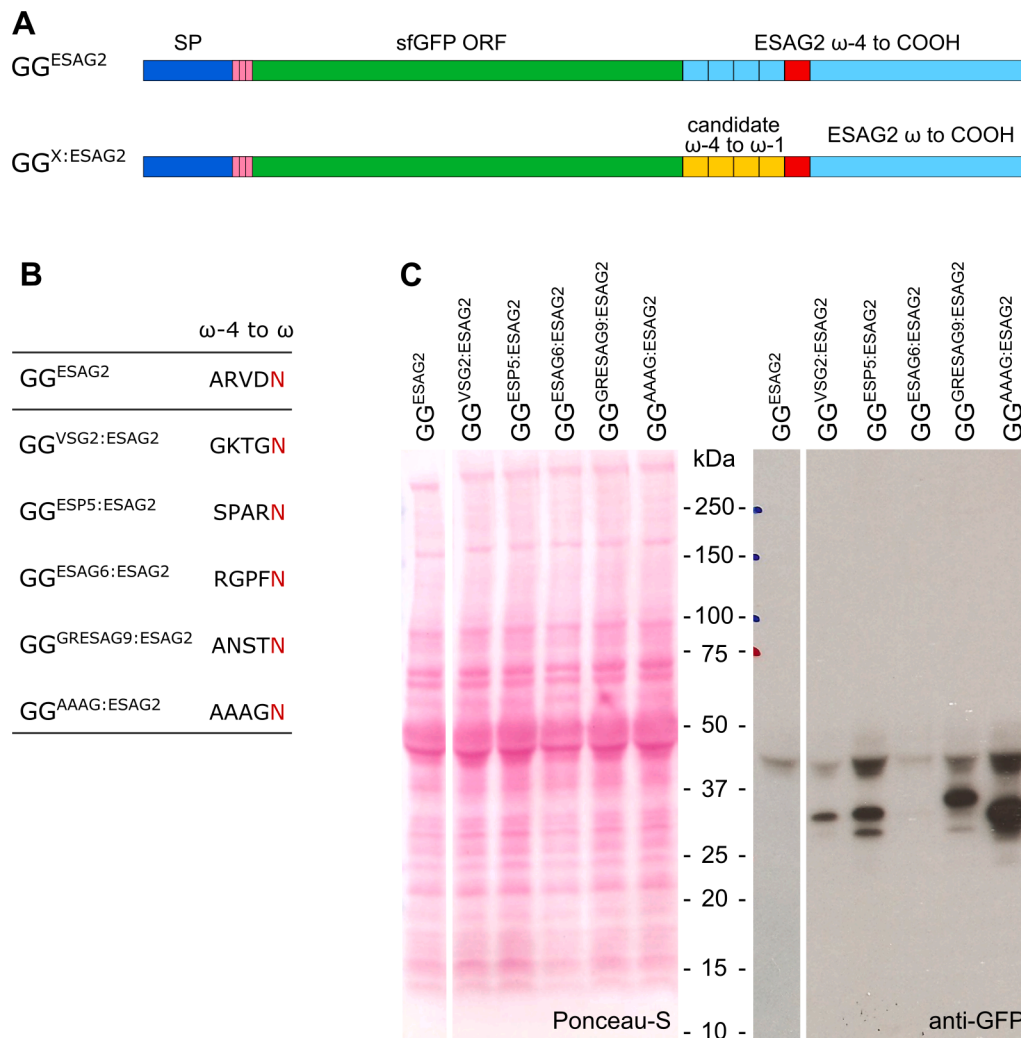
Fig. 7. Domain exchange affects GPI anchor addition. **A)** Schematic representation of artificial ESAG2 GPI insertion signals generated by replacement of native ω−4 to ω−1 residues with the respective positional ones from the other candidate proteins, or four small neutral residues. **B)** Amino acids in positions ω−4 to ω−1 resultant of re-engineering. The ESAG2 ω residue is shown in red. **C)** Effect of modifications to the ESAG2 GPI insertion sequence on protein level and anchoring efficiency. Chimeric signals made with the ω−minus residues from GRESAG9 > ESP5 > VSG-2 > ESAG6 (in order of abundance) greatly increase GPI anchoring and processing of fusion proteins, with the highest level observed for ω−4 to ω−1 composed of AAAG. The GRESAG9:ESAG2 chimeric signal causes a slightly higher band due to the N-glycosylation site within the ω−minus region (see Fig. 2).

Neither does it explain the cell surface localisation observed for two other VSG-related GPI-APs: ESAG2 (excluded from the flagellar pocket and flagellum membranes) and ESAG11 (excluded from the flagellar pocket) (Gadelha et al., 2015).

As for non-VSG GPI-APs, they too do not seem to conform to the GPI valence rule, because the parasite's monomeric haptoglobin-haemoglobin receptor (HpHbR) resides in the flagellar pocket membrane (Vanhollebeke et al., 2008) but forms a dimer with high avidity for dimeric HpHb (Lane-Serff et al., 2014). Importantly, our use of monomeric fluorescent proteins reinforces the notion that a GPI_1 feature is not sufficient for differential cellular sorting and flagellar pocket residence. Thus, whilst there could be a sorting mechanism based on oligomerisation status for particular subclasses of proteins, it is clearly not a generic tenet.

Saturable membrane barriers

All fusion proteins engineered here were transcribed from the same genomic locus. Their translation, at least across most of the ORF, should be equal for the various cell lines analysed. And yet, there is a noticeable difference in cellular levels of GG^{ESAG2}, possibly as a result of differential

processing of its ω−minus region. Differences in protein level is important in the context of possible sorting mechanisms. HpHbR has been estimated at 200–400 copies per trypanosome cell (Vanhollebeke et al., 2008) and its localisation is restricted to the flagellar pocket. Whilst GG^{ESAG2} is present at ∼800 copies per cell, and clearly visible over the entire cell surface. If the membrane barriers that compartmentalise the flagellar pocket away from the rest of the cell surface operate in a saturable manner as it has been previously proposed (Mussmann et al., 2003), then saturation must kick in between 400 and 800 molecules for any given surface protein, which would seem tricky.

The situation gets more complicated with the TfR (present at 3000–4000 copies per cell). In excess of normal levels (by ectopic over-expression or iron starvation) TfR is no longer retained in flagellar pocket and endosomes, and escapes to the entire cell surface (Mussmann et al., 2003). The relevance of such artificial gene up-regulation to endogenous protein sorting remains unclear, but the discovery of surface membrane proteins with localisations specific to each individual domain (and combinations thereof) on the parasite surface (Gadelha et al., 2015) indicates that sorting to surface domains in trypanosomes must be more complex than just a saturable mechanism of flagellar pocket retention.
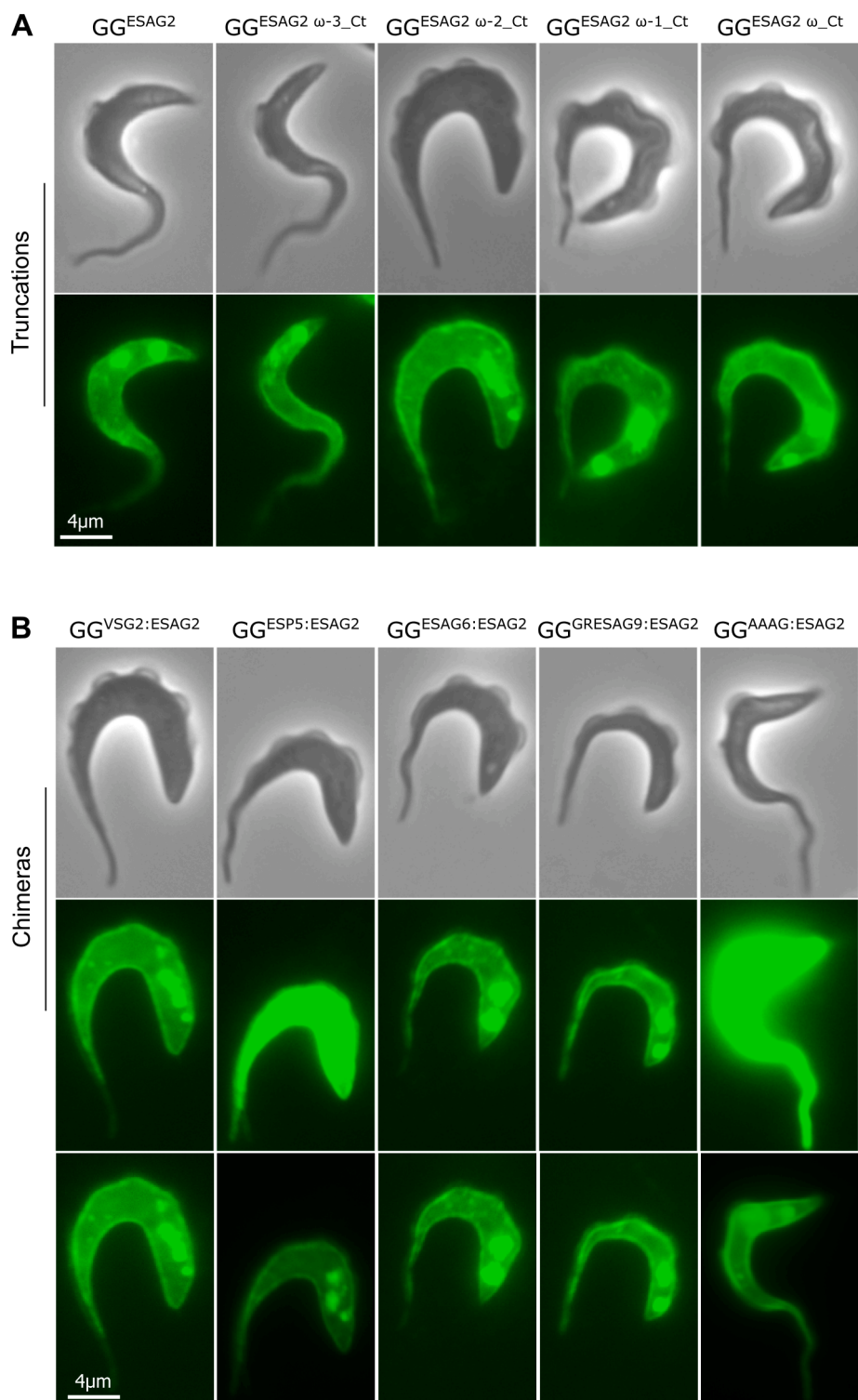
Fig. 8. GPI anchoring to the cell surface membrane persists upon modifications to the ω environment. Live microscopy of cell lines expressing modified GPI anchor signals. Native fluorescence is detected over the entire cell surface plus endosomes in truncated **(A)** and chimeric **(B)** ESAG2 anchor-expressing cells, similar to the signal observed from the native ESAG2 anchor (Fig. 4). All images captured and processed equally, except for the bottom panel in B, which has been differentially processed for clarity.

Protein-specific signals

Interactions via the GPI anchor are not the only connections available to a GPI-AP. Following lateral diffusion across the plasma membrane, local chaperones could recognise and retain specific proteins at their final destination. Distinct proteins may also interact differently with the multiple membrane barriers that define specialised domains on the cell surface. These interactions still need to be signalled from somewhere other than the ω region. We did consider that motifs within the primary sequence could be used to sort GPI-APs to their respective domains or enable them to cross specific domain boundaries. But even sensitive search methods were not able to find simple common motifs

among known surface proteins of same location.

Conclusion

Here we created the tools to dissect the ω region and its relationship with GPI-anchored protein sorting at the cell surface. We used 5 GPI-APs with distinct features (from a validated set of > 50) in search for common signals and mechanisms. We found that an N-terminal signal peptide and a C-terminal insertion sequence are sufficient for directing GPI-APs to the secretory pathway, for correct insertion of a GPI anchor, and for delivery to the plasma membrane. But we note that ultimate surface membrane specialisation is not determined by individual GPI insertion sequences, whereby the domain compartmentalisation seeing in the polarised membrane of African trypanosomes is not recapitulated by that signal alone. In doing so, we re-examine previous notions regarding protein sorting and retention which were proposed when there was insufficient knowledge of GPI-AP diversity.

The molecular signal(s) which impact on the ultimate distribution of GPI proteins in trypanosomes, alongside the cellular machineries that decode and interpret such signals, remain unknown. Nonetheless, here we uncovered an unexpected independence of known signals for this model systems. Validated sets of distinct GPI-APs will help us study the generic or diverse rules that govern protein segregation and create unique functional compartmentalisation of the cell surface.

CRediT authorship contribution statement

Thomas Henry Miller: Writing – review & editing, Investigation, Formal analysis. **Sabine Schiessler:** Investigation, Formal analysis. **Ella Maria Rogerson:** Writing – review & editing, Investigation, Formal analysis. **Catarina Gadelha:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

Given her role as an Associate Editor of the journal, Catarina Gadelha had no involvement in the peer review of this article, and has no access to information regarding its peer review. Full responsability for the editorial process for this article was delegated to Neil Gow.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tcsw.2024.100131.

References

Barnwell, E.M., van Deursen, F.J., Jeacock, L., Smith, K.A., Maizels, R.M., Acosta-Serrano, A., Matthews, K., 2010. Developmental regulation and extracellular release of a VSG expression-site-associated gene product from *Trypanosoma brucei* bloodstream forms. J. Cell Sci. 123, 3401–3411.

Bate, C., Nolan, W., McHale-Owen, H., Williams, A., 2016. Sialic acid within the glycosylphosphatidylinositol anchor targets the cellular prion protein to synapses. J. Biol. Chem. 291, 17093–17101.

Benting, J.H., Rietveld, A.G., Simons, K., 1999. N-Glycans mediate the apical sorting of a GPI-anchored, raft-associated protein in Madin-Darby canine kidney cells. J. Cell Biol. 146, 313–320.

Bindels, D.S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K.E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M.A., Gadella Jr., T.W., 2017. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat. Methods 14, 53–56.

Böhme, U., Cross, G.A.M., 2002. Mutational analysis of the variant surface glycoprotein GPI-anchor signal sequence in *Trypanosoma brucei*. J. Cell Sci. 115, 805–816.

Borges, A.R., Link, F., Engstler, M., Jones, N.G., 2021. The glycosylphosphatidylinositol anchor: a linchpin for cell surface versatility of trypanosomatids. Front. Cell Dev. Biol. 9, 720536.

Burkard, G., Fragoso, C.M., Roditi, I., 2007. Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. Mol. Biochem. Parasitol. 153, 220–223.

Caro, L.H., Tettelin, H., Vossen, J.H., Ram, A.F., van den Ende, H., Klis, F.M., 1997. In silico identification of glycosylphosphatidylinositol-anchored plasma membrane and cell wall proteins of *Saccharomyces cerevisiae*. Yeast. 13, 1477–1489.

Catino, M.A., Paladino, S., Tivodar, S., Pocard, T., Zurzolo, C., 2008. N- and O-glycans are not directly involved in the oligomerization and apical sorting of GPI proteins. Traffic 9, 2141–2150.

Chandra, M., Đaković, S., Foti, K., Zeelen, J.P., van Straaten, M., Aresta-Branco, F., Tihon, E., Lübbehusen, N., Ruppert, T., Glover, L., Papavasiliou, F.N., Stebbins, C.E., 2023. Structural similarities between the metacyclic and bloodstream form variant surface glycoproteins of the African trypanosome. PLoS Negl. Trop. Dis. 17, e0011093.

Cranfill, P.J., Sell, B.R., Baird, M.A., Allen, J.R., Lavagnino, Z., de Gruiter, H.M., Kremers, G.-J., Davidson, M.W., Ustione, A., Piston, D.W., 2016. Quantitative assessment of fluorescent proteins. Nat. Methods 13, 557–562.

Cross, G.A., 1984. Release and purification of *Trypanosoma brucei* variant surface glycoprotein. J. Cell Biochem. 24, 79–90.

Cross, G.A.M., Kim, H.S., Wickstead, B., 2014. Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. Mol. Biochem. Parasitol. 195, 59–73.

Đaković, S., Zeelen, J.P., Gkeka, A., Chandra, M., van Straaten, M., Foti, K., Zhong, J., Vlachou, E.P., Aresta-Branco, F., Verdi, J.P., Papavasiliou, F.N., Stebbins, C.E., 2023. A structural classification of the variant surface glycoproteins of the African trypanosome. PLoS Negl. Trop. Dis. 17, e0011621.

Edelstein, A.D., Tsuchida, M.A., Amodaj, N., Pinkard, H., Vale, R.D., Stuurman, N., 2014. Advanced methods of microscope control using μManager software. J. Biol. Methods 1, e10.

Ferguson, M.A.J., Low, M.G., Cross, G.A.M., 1985. Glycosyl-sn-1,2-dimyristylphosphatidylinositol is covalently linked to *Trypanosoma brucei* variant surface glycoprotein. J. Biol. Chem. 260, 14547–14555.

Ferguson, M.A.J., Duszenko, M., Lamont, G.S., Overath, P., Cross, G.A.M., 1986. Biosynthesis of *Trypanosoma brucei* variant surface glycoproteins. N-glycosylation and addition of a phosphatidylinositol membrane anchor. J. Biol. Chem. 261, 356–362.

Ferguson, M.A.J., Homans, S.W., Dwek, R.A., Rademacher, T.W., 1988. Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. Science 239, 753–759.

Frieman, M.B., Cormack, B.P., 2004. Multiple sequence signals determine the distribution of glycosylphosphatidylinositol proteins between the plasma membrane and cell wall in *Saccharomyces cerevisiae*. Microbiol. 150, 3105–3114.

Gadelha, C., Rothery, S., Morphew, M.K., McIntosh, J.R., Severs, N.J., Gull, K., 2009. Membrane domains and flagellar pocket boundaries are influenced by the cytoskeleton in African trypanosomes. Proc. Natl. Acad. Sci. U.S.A. 106, 17425–17430.

Gadelha, C., Zhang, W., Chamberlain, J.W., Chait, B.T., Wickstead, B., Field, M.F., 2015. Architecture of a parasite surface: complex targeting mechanisms revealed through proteomics. Mol. Cell. Proteomics. 14, 1911–1926.

Gupta, R., Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac. Symp. Biocomput. 310–322.

Ji, Z., Nagar, R., Duncan, S.M., Sampaio Guther, M.L., Ferguson, M.A.J., 2023. Identification of the glycosylphosphatidylinositol-specific phospholipase A2 (GPI-PLA2) that mediates GPI fatty acid remodeling in *Trypanosoma brucei*. J. Biol. Chem. 299, 105016.

Lacomble, S., Vaughan, S., Gadelha, C., Morphew, M.K., Shaw, M.K., McIntosh, J.R., Gull, K., 2009. Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in trypanosomes revealed by electron microscope tomography. J. Cell Sci. 122, 1081–1090.

Lane-Serff, H., MacGregor, P., Lowe, E.D., Carrington, M., Higgins, M.K., 2014. Structural basis for ligand and innate immunity factor uptake by the trypanosome haptoglobin-haemoglobin receptor. Elife 3, e05553.

Lebreton, S., Paladino, S., Liu, D., Nitti, M., von Blume, J., Pinton, P., Zurzolo, C., 2021. Calcium levels in the Golgi complex regulate clustering and apical sorting of GPI-APs in polarized epithelial cells. Proc. Natl. Acad. Sci. U.S.A. 118, e2014709118.

Lemus, L., Hegde, R.S., Goder, V., 2023. New frontiers in quality control: the case of GPI-anchored proteins. Nat. Rev. Mol. Cell Biol. 24, 599–600.

Mehlert, A., Richardson, J.M., Ferguson, M.A.J., 1998. Structure of the glycosylphosphatidylinositol membrane anchor glycan of a class-2 variant surface glycoprotein from *Trypanosoma brucei*. J. Mol. Biol. 277, 379–392.

Mussmann, R., Janssen, H., Calafat, J., Engstler, M., Ansorge, I., Clayton, C., Borst, P., 2003. The expression level determines the surface distribution of the transferrin receptor in *Trypanosoma brucei*. Mol Microbiol. 47, 23–35.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.

Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 122–130.

Ouyang, H., Chen, X., Lü, Y., Wilson, I.B., Tang, G., Wang, A., Jin, C., 2013. One single basic amino acid at the w-1 or w-2 site is a signal that retains glycosylphosphatidylinositol-anchored protein in the plasma membrane of *Aspergillus fumigatus*. Eukaryot. Cell 12, 889–899.

Paladino, S., Lebreton, S., Tivodar, S., Campana, V., Tempre, R., Zurzolo, C., 2008. Different GPI-attachment signals affect the oligomerisation of GPI-anchored proteins and their apical sorting. J. Cell Sci. 121, 4001–4007.

Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., Waldo, G.S., 2006. Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 24, 79–88.

Pierleoni, A., Martelli, P.L., Casadio, R., 2008. PredGPI: a GPI-anchor predictor. BMC Bioinformatics 9, 392.

Pinger, J., Nešić, D., Ali, L., Aresta-Branco, F., Lilic, M., Chowdhury, S., Kim, H.S., Verdi, J., Raper, J., Ferguson, M.A.J., Papavasiliou, F.N., Stebbins, C.E., 2018. African trypanosomes evade immune clearance by O-glycosylation of the VSG surface coat. Nat. Microbiol. 3, 932–938.

Puig, B., Altmeppen, H.C., Linsenmeier, L., Chakroun, K., Wegwitz, F., Piontek, U.K., Tatzelt, J., Bate, C., Magnus, T., Glatzel, M., 2019. GPI-anchor signal sequence influences PrPC sorting, shedding and signalling, and impacts on different pathomechanistic aspects of prion disease in mice. PLoS Pathog. 15, e1007520.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675.

Schwartz, K.J., Peck, R.F., Tazeh, N.N., Bangs, J.D., 2005. GPI valence and the fate of secretory membrane proteins in African trypanosomes. J. Cell Sci. 118, 5499–5511.

Trimble, W., Grinstein, S., 2015. Barriers to the free diffusion of proteins and lipids in the plasma membrane. J. Cell Biol. 208, 259–271.

Umaer, K., Aresta-Branco, F., Chandra, M., van Straaten, M., Zeelen, J., Lapouge, K., Waxman, B., Stebbins, C.E., Bangs, J.D., 2021. Dynamic, variable oligomerization and the trafficking of variant surface glycoproteins of *Trypanosoma brucei*. Traffic 22, 274–283.

Vanhollebeke, B., De Muylder, G., Nielsen, M.J., Pays, A., Tebabi, P., Dieu, M., Raes, M., Moestrup, S.K., Pays, E., 2008. A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans. Science 320, 677–681.

Wickstead, B., Ersfeld, K., Gull, K., 2003. The frequency of gene targeting in *Trypanosoma brucei* is independent of target site copy number. Nucleic Acids Res. 31, 3993–4000.

Wirtz, E., Leal, S., Ochatt, C., Cross, G.A.M., 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. Mol. Biochem. Parasitol. 99, 89–101.

Zitzmann, N., Mehlert, A., Carrouée, S., Rudd, P.M., Ferguson, M.A.J., 2000. Protein structure controls the processing of the N-linked oligosaccharides and glycosylphosphatidylinositol glycans of variant surface glycoproteins expressed in bloodstream form *Trypanosoma brucei*. Glycobiology 10, 243–249.