Glycoconj J (2009) 26:915–921
DOI 10.1007/s10719-008-9116-x

Proteomic scale high-sensitivity analyses of GPI membrane anchors

Angela Mehlert · Michael A. J. Ferguson

Received: 6 December 2007 / Revised: 24 January 2008 / Accepted: 1 February 2008 / Published online: 11 March 2008

Abstract Glycosylphosphatidylinositol (GPI) anchored proteins are ubiquitous in eukaryotic cells. Earlier analysis methods required large amounts of purified protein to elucidate the structure of the GPI. This paper describes methods for analyzing GPIs on a 'proteomic' scale. Partially purified proteins may be run on sodium dodecyl sulphate polyacrylamide gel electrophoresis and then blotted onto a polyvinylidene difluoride (PVDF) membrane. Following identification of the protein the piece of PVDF may be subjected to various chemical treatments, which are specific for GPI structures. The first method uses gas chromatography–mass spectrometry and it enables the presence of a GPI anchor to be confirmed. The second method depends on the cleavage of phosphate bonds and permits the carbohydrate structure to be elucidated by electrospray or matrix assisted laser desorption ionization-time of flight mass spectrometry. The final method described uses deamination of the glucosamine residue to release the lipid moiety for analysis by mass spectrometry.

Keywords Glycosylphosphatidylinositol (GPI) · Mass spectrometry · Proteomic techniques

Abbreviations

CID	collision induced dissociation
DMSO	dimethylsuphoxide
ER	endoplasmic reticulum
ESAG	expression site associated gene
GC–MS	gas chromatography–mass spectrometry
GPI	glycosylphosphatidylinositol
HF	hydrofluoric acid
MALDI-Tof	Matrix assisted laser desorption ionization-time of flight spectrometry
Nmr	nuclear magnetic resonance
PI	phosphatidylinositol
PNGaseF	*N*-glycosidase F
PVDF	polyvinylidenedifluoride
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIM	selected ion monitoring
VSG	variant surface glycoprotein

A. Mehlert · M. A. J. Ferguson (✉)
Division of Biological Chemistry and Drug Discovery, College of Life Sciences, Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, UK
e-mail: m.a.j.ferguson@dundee.ac.uk

Introduction

Many eukaryotic cell surface proteins are anchored to the cell membrane using a glycosylphosphatidylinositol (GPI) anchor. The GPI anchor may be added to the C-terminus of a cell surface protein in the endoplasmic reticulum. A precursor glycolipid (in *Trypanosoma brucei* this is called glycolipid A and has the structure EthN-*P*-Man_3GlcN-phosphatidylinositol [PI]) is assembled in the endoplasmic reticulum and transferred to the nascent proteins by a transamidase enzyme, in exchange for the C-terminal signal peptide of the GPI anchored protein.

To elucidate fully the structure of a GPI anchor by conventional carbohydrate chemistry is a lengthy undertaking that requires large amounts of native protein and specialized techniques including nuclear magnetic resonance and radio-labelling. Therefore, we have been working on new methods, based on a 'proteomic' approach, that enable a GPI structure to be elucidated from a sample of the protein of interest electrophoresed on an sodium dodecyl sulphate (SDS)

polyacrylamide gel and blotted onto a polyvinylidenedifluoride (PVDF) membrane. These methods can facilitate the determination of both the lipid and carbohydrate structure of the GPI anchor and are dependent on sensitive mass spectrometric techniques. However, being based on mass spectrometry data, the method for carbohydrate determination cannot distinguish between isobaric carbohydrate residues. Consequently, certain assumptions about the GPI structure have to be made, this is assisted by the fact that the structure of all the GPI anchors so far analysed include the core structure:

$$\begin{aligned}&NH_2CH_2CH_2PO_4H-6Man\alpha1-2Man\alpha1-6Man\alpha1\\&\quad-4GlcN\alpha1-6D-myo-\text{inositol}-1-HPO_4\\&\quad-\text{lipid}[EthN-P-Man_3GlcN-PI].\end{aligned}$$

The unusual chemistry of GPI anchors allows exquisitely specific reactions to be carried out, enabling detection of the GPI structures in a background of many other more common cellular glycolipids. The ability of aqueous hydrofluoric acid (aq. HF) to cleave phosphomono- and di-ester bonds enables the purification of the carbohydrate portion of the GPI from the protein and lipid components. When the reaction is carried out on PVDF membrane, the protein and the lipid remain on the membrane, whereas the GPI carbohydrate core structure may be washed off in water, permitting its removal for analysis. In addition, the deamination of the glucosamine residue using nitrous acid is specific for this free amino sugar, which is very unusual in nature but common to all GPIs. In this case the protein, still attached to the carbohydrate residue of the GPI, remains on the PVDF membrane and the phosphatidylinositol lipid which is released by the deamination reaction is extracted using organic solvent, allowing its analysis. The application of these techniques on a proteomic scale has enabled the structural elucidation of several GPI structures

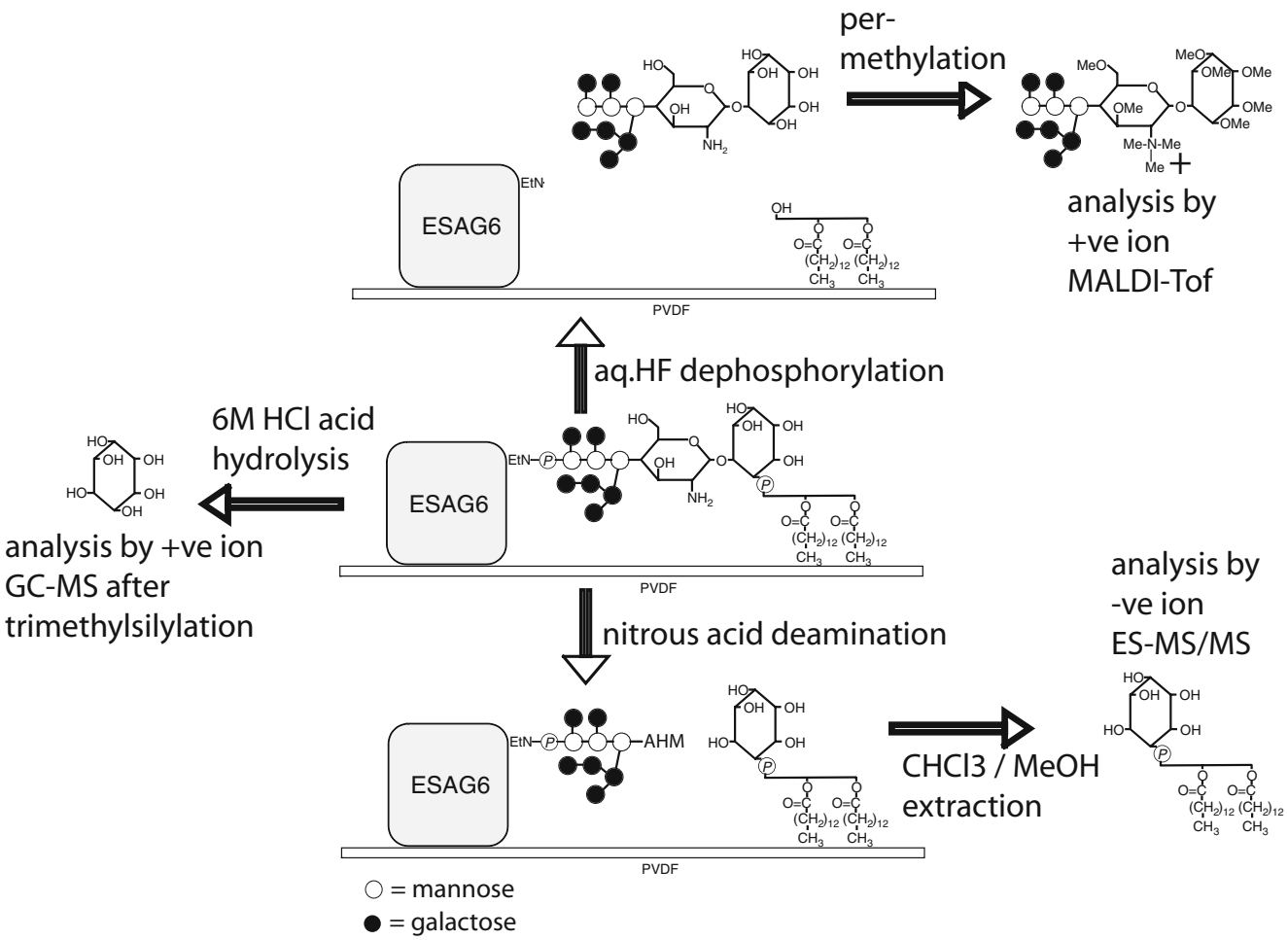


Fig. 1 Reaction scheme for the analysis of the *T. brucei* TfR GPI anchor. This example is adapted from [5]. The ESAG6 component of *T. brucei* TfR is immobilized on PVDF by Western blotting. Initial inositol analysis may be performed after digestion of the membrane with 6 M HCl (*left*). The membrane can be treated with aq. HF to release the hydrophilic GPI glycans into solution, for subsequent permethylation and positive ion MALDI-Tof mass spectrometry (*top*), or treated with nitrous acid that separates the protein from the PI component of the GPI anchor. In the latter case, the hydrophobic PI moiety remains attached to the PVDF membrane and can be recovered by solvent extraction of the membrane for negative ion ES-MS/MS mass spectrometry (*bottom*)

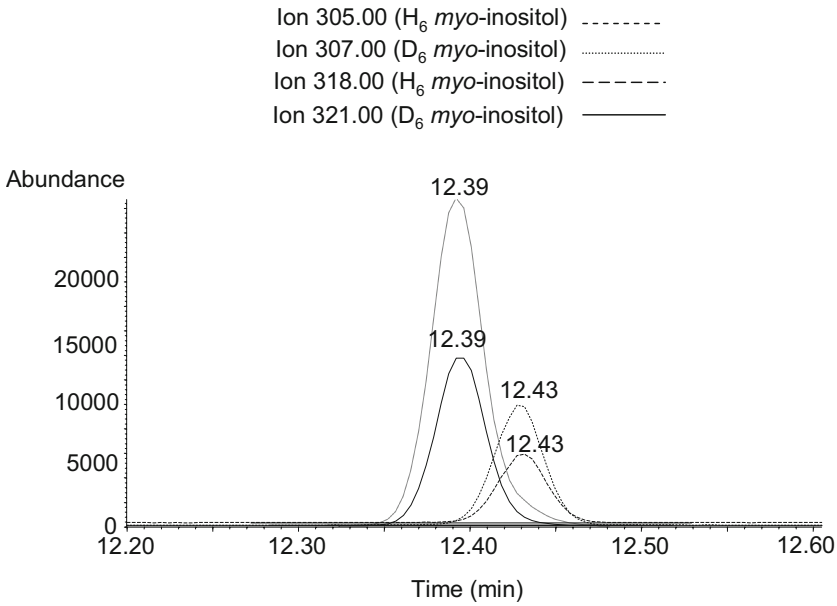
from proteins of low abundance [4, 5–9]. However, the first problem which has to be faced is whether the protein of interest is actually carrying a GPI anchor. There are predictive algorithms for the C- terminal GPI signal sequences, but they are not infallible. We have been able to use a simple acid digest and gas chromatography–mass spectrometry (GC–MS) approach to determine whether the protein of interest on a PVDF blot contains *myo*-inositol. If the protein band of interest contains this polyol, then further analysis to determine the GPI structure is warranted. This initial strategy is shown in Fig. 1 along with the methods for determining the lipid and carbohydrate components of the GPI.

Materials and methods

All reagents were supplied by Sigma Aldrich Chemical Co., Poole, UK or VWR International, Lutterworth, UK, unless otherwise stated.

Inositol analysis

The protein to be analysed should be resolved on an SDS polyacrylamide gel and transferred to a PVDF blot (GE Healthcare, Chalfont St. Giles, UK). The protein of interest may be visualized by staining with amido black or Ponceau S if it is sufficiently intense, alternatively a parallel lane may be stained using an appropriate antibody. The regions of interest are cut out, along with equivalent areas from above and below the protein band, which are used as controls for the experiment. The strips of PVDF must be washed in methanol and again in water and then placed in a small (2 ml) Reactivial (Perbio Science UK Ltd, Cramlington, UK). Fifty picomole of D_6-*myo*-inositol (*myo*-Inositol-1,2,3,4,5,6- d6 from QMX, Labs, Thaxted, UK) should be included in each vial as an internal standard and enough 6 M HCl (constant boiling from Perbio, as above) to cover the strip added, usually 200 μl is sufficient. The vial is sealed with a Teflon lined cap and the contents hydrolysed for 16 h at 110°C. The hydrolysates are removed from the strips and the strips are then washed with 200 μl water. The hydrolysate and washings are combined and dried in a Speedvac (Thermo Fisher Scientific, Basingstoke, UK), redried from 50 μl methanol and derivatised with trimethylsilyl reagent as described in Ferguson 1992 [2]. Briefly, 500 μl anhydrous pyridine is mixed with 50 μl trimethylsilazane and 150 μl hexamethylsilazane in an acid washed tube. The mixture is briefly centrifuged or allowed to settle and 30 μl of the reagent added to each sample 30 min prior to running on an Agilent (Agilent technologies, South Queensferry, UK) 6890 gas chromatography system linked to a 5973 MSD-mass spectrometer (GC–MS), with an Agilent HP-5, or HP-5ms column (30 m, 0.25 mm i.d., film thickness 0.25 μm). A selected ion monitoring (SIM) programme is used. The GC programme has an injector temperature of 280°C and the interface between the GC and MS is held at 280°C, the He carrier gas flow is 0.7 ml/min with a column pressure of 5 psi and a splitless injector is used. The initial oven temperature is 80°C and is held for 1 min, then a gradient of 60°C/min to 140°C, followed by a



Ion m/z	Time of peak	Area ion m/z
305	12.43	203738
307	12.39	596193
318	12.43	116938
321	12.39	297046

Fig. 2 Inositol analysis using a selected ion monitoring programme by GC–MS. The *m*/*z* 305 and 307 ions represent the native (H_6) *myo*-inositol in the sample, the *m*/*z* 307 and 321 ions represent the internal standard (D_6) *myo*-inositol. The amount of *myo*-inositol in the sample may be calculated using the formula: (area of peak a or c area of peak b or d)×amount of (D_6) *myo*-inositol internal standard

gradient of 15°C/min up to a final temperature of 250°C. This temperature is held for 5 min and the oven returned to the initial temperature of 80°C. *Myo*-inositol is detected by the characteristic ions at *m/z* 305 and 318 and the D_6 labelled internal standard *myo*-inositol is detected using the ions at *m/z* 307 and 321, see example of ion chromatograms in Fig. 2. The amount of *myo*-inositol from the sample is calculated using the formula: (area $m/z318 \div$ area $m/z321) \times$ amount of D_6 *myo*-inositol internal standard.

GPIs have only been reported to occur at the C-terminus of a protein so an estimate of the molar quantity of *myo*-inositol in a pure protein must equal the molar quantity of the protein itself.

Carbohydrate structure of GPI

To reduce heterogeneity due to N-linked glycans, the GPI anchored protein to be analysed is usually treated with *N*-glycosidase F (PNGaseF; New England Biolabs, Hitchin, UK) before running on a precast gradient gel (4–12% bis tris from Invitrogen, Paisley, UK) and the gel is blotted onto a PVDF membrane, as above. It is helpful to run a known standard GPI anchored protein in an adjacent lane to the protein of interest. We generally use one of the well characterized variant surface glycoproteins (VSGs) from *T. brucei* [2, 6]. The blot is then stained with amido black to reveal the location of the protein of interest and the protein

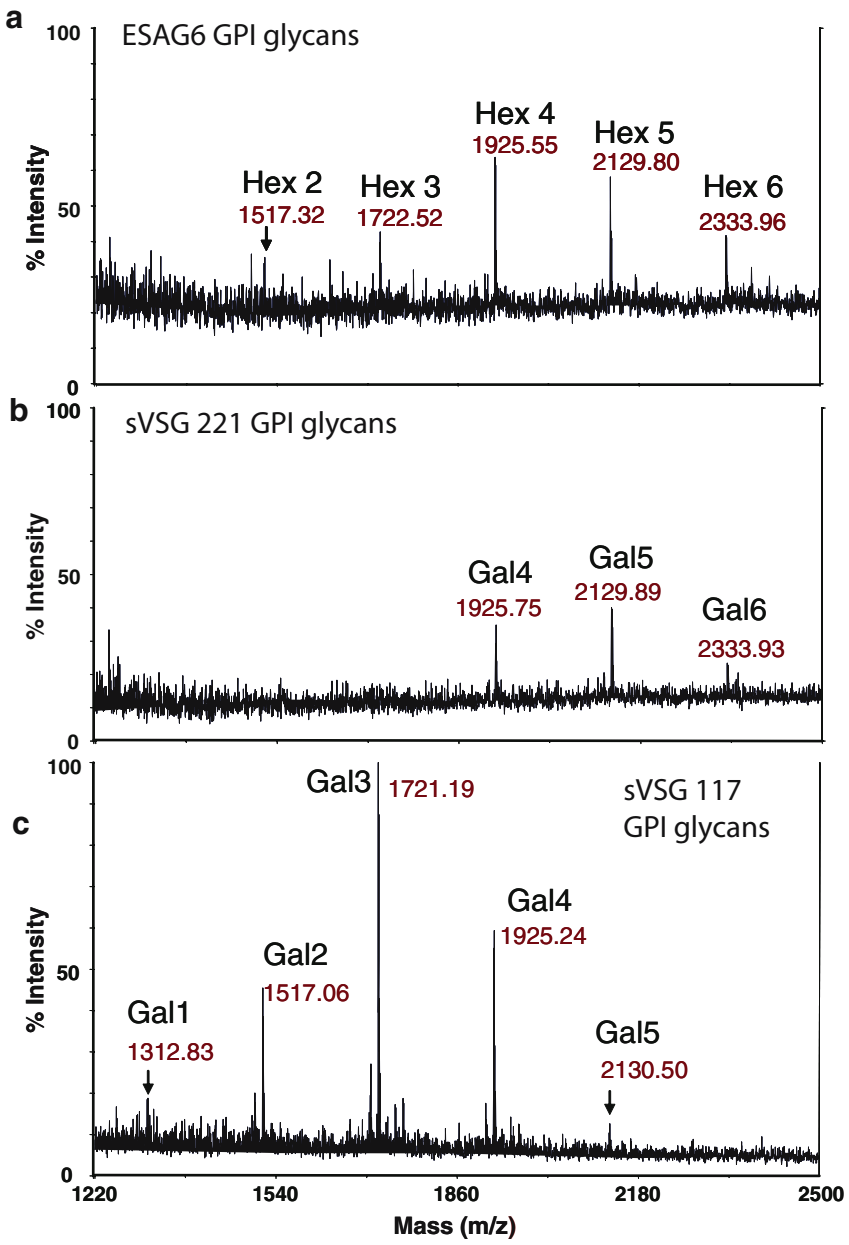


Fig. 3 Mass spectra of permethylated GPI glycans. Positive ion MALDI-Tof mass spectra of permethylated GPI glycans from **a** the ESAG6 component of TfR, **b** sVSG221 and **c** sVSG117. The annotated ions correspond to the permethylated GPI glycan core of Man_3-$GlcNMe_3$-inositol with up to six side chain Gal residues, as indicated

band is excised and placed in a screw top Eppendorf tube. The PVDF strip is wetted with methanol and aq. HF is added to cover the PVDF pieces, usually 40–50 μl is sufficient. The aq. HF digest is continued for approximately 60 h on ice/water. The aq. HF (which should contain the dephosphorylated glycans) is then removed from the piece of membrane which is washed with 2×100 μl water, the washings being added to the aq. HF digest. This diluted aq. HF is then lyophilised several times to remove the HF. The GPI glycans are dissolved in a small volume, usually 50 μl of water, transferred to an acid washed 2 ml Reactivial (Perbio) and dried in a Speedvac concentrator to remove the water. The samples are then redried from a small volume (usually 20 μl) of methanol and permethylated according to the protocol of Ferguson [2].

Briefly, the samples are processed alongside a positive carbohydrate control (*e.g.* lacto-*N*-tetraose, Sigma). The dried samples are dissolved in 50 μl of dimethylsuphoxide (DMSO) allowing at least 20 min for dissolution. A slurry 120 mg/ml NaOH in DMSO is prepared by quickly grinding two or three NaOH pellets in an appropriate volume of DMSO with an acid washed pestle and mortar. Fifty microlitre of this slurry is immediately added to each sample and they are left for a further 20 min. This is followed by two additions of 10 μl of methyl iodide (iodomethane), followed by 10 min incubations. A further 20 μl of methyl iodide is added and the samples left for 20 min.

Note: Methyl iodide is a powerful carcinogen and so these additions must be done in a fume hood.

After the permethylation reaction, the permethylated glycans must be recovered from the excess reagents in order for the glycans to be analysed by mass spectrometry. This is achieved by adding 250 μl of chloroform followed by 1 ml 0.2 g/ml sodium thiosulphate in water to each sample. The thiosulphate removes any molecular iodine and the permethylated glycans partition into the chloroform phase. The vials are vortexed for at least a minute and the contents are allowed to settle. The water phase is discarded and the chloroform phase is subjected to five further washes with 1 ml water with 1 min vortexing each time. By the final wash the water phase should look completely clear. The chloroform phase may be recovered from under the final water wash using a Pasteur pipette and transferred to another acid washed glass tube and then dried using a stream of nitrogen.

The permethylated GPI glycans, which contain a quaternary ammonium group on the permethylated glucosamine residue ($GlcNMe_3^+$), may be analysed in several ways. Following dilution in 50% acetonitrile with 1% formic acid they can be introduced to a electrospray mass spectrometer, the ions in positive mode will usually run as $[M+H]^{2+}$ or $[M+Na]^{2+}$ ions, [1, 6]. Alternatively, the sample may be diluted in methanol and an aliquot mixed with dihydroxybenzoic acid matrix, spotted onto a matrix assisted laser desorption ionization (MALDI) plate and analysed on a MALDI-time of flight (Tof) mass spectrometer again in positive mode, the major ion seen corresponds to $[M]^+$, [5], for examples see Fig. 3. The ion seen in MALDI-Tof which corresponds to the $Man\alpha1-2Man\alpha1-6Man\alpha1-4GlcNMe_3^+\alpha1-6D-myo-inositol$ core structure is *m/z* 1109.

Note: The quaternary amine formed by the permethylation of the GlcN residue makes positive ion mass spectrometry of these derivatives extremely sensitive.

Lipid structure of the GPI

The protein to be analysed should be digested with PNGaseF, then electrophoresed and blotted onto PVDF alongside a known standard, as in the previous methods. Similarly, the blot is stained with amido black and the bands of interest and a control blank piece cut out. The

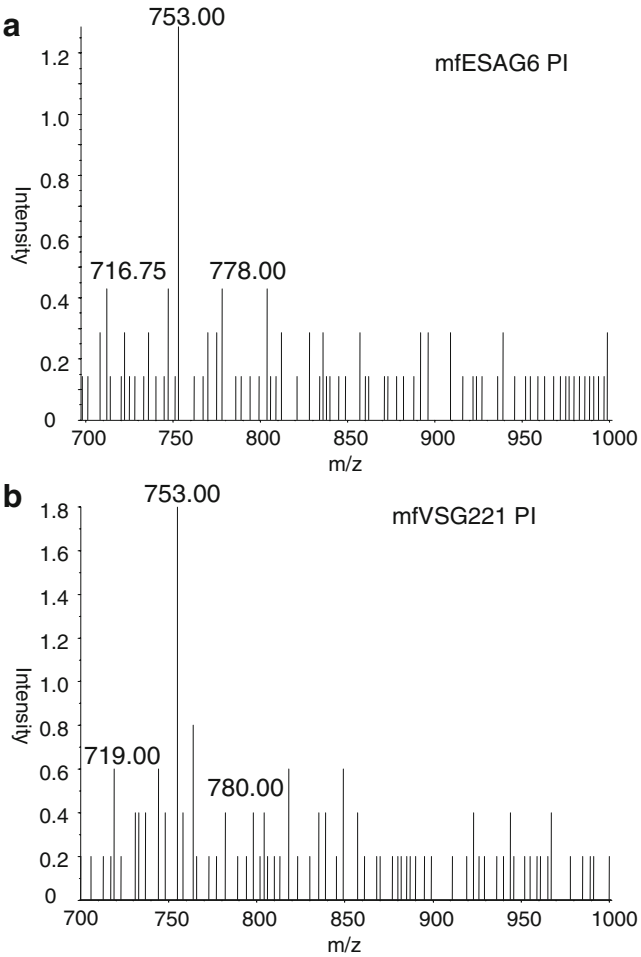


Fig. 4 Mass spectra of nitrous acid-released PI moieties. Negative precursor ion spectra (precursors of *m/z* 241) of the PI fractions from *T. brucei* ESAG6 (**a**) and mfVSG221 (**b**). The ions at *m/z* 753 correspond to dimyristoyl-PI

PVDF strips were washed in methanol three times, then in chloroform/methanol 1:1 (*v*/*v*) three times, then again with methanol three times. Any GPIs on the blots are then deaminated with nitrous acid using freshly prepared 250 mM $NaNO_2$ in 150 mM sodium acetate buffer pH 4.0 (37°C, 3 h). The PVDF strips are then washed with water three times, dried and may be stored indefinitely in a glass vessel at 4°C under argon, until analysis.

The PI moieties, which have been released from the GPI may be extracted from the PVDF by adding 50 μl of chloroform/methanol (1:1 *v*/*v*) to the strip of blot and the solution loaded into a nanospray gold-coated glass tip (type F from Waters, Manchester, UK.). Analysis is carried out on a Q-StarXL electrospray tandem-MS/MS system (Applied Biosystems, Warrington, UK.). Precursor ion scans of the fragment *m*/*z* 241, corresponding to the inositol-1,2-cyclic phosphate, are collected, see Fig. 4. All the GPI anchors from bloodstream form trypanosomes which have been elucidated so far, have contained *sn*-1,2-dimyristoyl-PIs which give the precursor ion *m*/*z* 753 [3]. Our results mean that the expression site associated gene 6 (ESAG6) GPI has the same lipid. Where more material is available, the PI moiety of a GPI anchor may be identified by electrospray or Q-tof mass spectrometry and the structure of its lipids confirmed by tandem (collision induced dissociation) mass spectrometry as in [9].

Results

Inositol analysis

The graphs in Fig. 2 show the result of an inositol analysis as described in "Inositol analysis" of the "Materials and methods." Using the equation shown earlier and assuming that 50 pmol of D_6 *myo*-inositol internal standard had been used, then the amount of *myo*-inositol in the sample would be 19.7 pmol. To be sure of an accurate estimate of inositol content these analyses should be run in triplicate.

Carbohydrate structure

Figure 3 shows the MALDI-Tof spectra obtained using the proteomic technique described in "Materials and methods". The spectra in Fig. 3b, c show that the GPI structures of the two different soluble VSG (sVSG) variants used as controls are consistent with the literature [3, 6]. Part a of Fig. 3 shows that the ESAG6 GPI glycans are similar to those of sVSG 221, although perhaps slightly larger. The smallest moiety found corresponds to a side chain containing three hexoses, the largest corresponding to a six hexose containing side chain.

Lipid structure

The spectra in Fig. 4 show the precursor ion scans of the fragment *m*/*z* 241, which is the inositol-1,2-cyclic phosphate from a, the membrane form of ESAG6 and b, membrane form (mf) VSG 221. The same precursor ion (*m*/*z* 753) is found in each sample so we concluded that the PI part of the ESAG6 GPI anchor has the same structure as that of all VSGs so far analysed *i.e.*, *sn*-1,2-dimyristoyl-PI. Unfortunately, due to the nature of precursor ion scanning experimental data, the spectra are of low resolution. We have published other data produced in a similar way where more material was available [8]. In this case the fragmentation data showed that several different lipids were found attached to the GPIs of the low molecular weight glycolipid antigen from *Cryptosporidium parvum*, whereas our data for the ESAG6 is consistent with other lipid data for GPIs from *T. brucei*.

Conclusions

Using the methods described above, we have been able to determine several GPI structures from proteins, which were difficult to purify in large amounts. The isolation of the protein of interest by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gives an extra dimension of purification, which ensures that the GPI analysis is of that protein and not any other contaminating material. SDS-PAGE is a particularly useful way of removing contaminating lipids, which are common in protein preparations.

Sadly, these proteomic methods do have disadvantages as well as their obvious advantages of scale. The GC column used for the inositol analysis after acid hydrolysis will become contaminated and make it useless for non-SIM work. The same contamination causes the MS ion source to require frequent cleaning. The use of aq. HF will remove any other groups (*e.g.* ethanolamine) attached to the GPI through phosphate bonds in addition to those phosphate linkages between the glycan and protein and the inositol and lipid. Therefore, it has to be assumed that the ethanolamine phosphate linkage between the C-terminus of the protein is to the third mannose of the glycan, however all the structures analysed by more conventional techniques have repeatedly demonstrated this general arrangement. The permethylated glycan experiment provides only the mass of the GPI glycan, any further information about the glycan structure must be achieved by other methods such as labeling with a radioactive or fluorescent probe at the reducing end of the molecule and exo-glycosidase sequencing by chromatography or high performance thin layer chromatography. This technique

also uses aq. HF which, despite the small quantities used, requires careful handling. The removal of the HF by freeze drying also has long-term deleterious consequences for the machinery used.

Acknowledgements The authors would like to acknowledge the help of Jeff Priest, Thierry Fontaine, Lynn Thomson and Maria-Agata Catino for their valuable contributions towards these methods.

This work was supported by a Wellcome Trust programme grant (071463) and an Medical Research Council programme grant.

References

1. Baldwyn, M.A.: Analysis of glycosylphosphatidylinositol protein anchors: the prion protein. Meth. Enzymol. **205**, 172–187 (2005)
2. Ferguson, M.A.J.: Chemical and enzymic analysis of GPI anchors. In: Turner, A.J., Hooper, N. (eds.) Lipid Modification of Proteins: A Practical Approach, pp. 191–230. IRL, Oxford (1994)
3. Ferguson, M.A., Homans, S.W., Dwek, R.A., Rademacher, T.W.: Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. Science **239**, 753–759 (1988)
4. Fontaine, T., Magnin, T., Melhert, A., Lamont, D., Latgé, J.-P., Ferguson, M.A.J.: Structures of the GPI membrane anchors from *Aspergillus fumigatus* membrane proteins. Glycobiology **13**, 169–177 (2003)
5. Mehlert, A., Ferguson, M.A.J.: Structure of the GPI anchor of the *T. brucei* transferrin receptor. Mol. Biochem. Parasitol. **151**, 220–223 (2007)
6. Mehlert, A., Richardson, J.M., Ferguson, M.A.J.: Structure of the glycosylphosphatidylinositol membrane anchor glycan of a class-2 variant surface glycoprotein from *Trypanosoma brucei*. J. Mol. Biol. **277**, 379–392 (1998)
7. Priest, J.W., Xie, L.-T., Arrowood, M.J., Lammie, P.J.: The immunodominant 17-kDa antigen from *Cryptosporidium parvum* is glycosylphosphatidylinositol-anchored. Mol. Biochem. Parasitol. **113**, 117–126 (2001)
8. Priest, J.W., Mehlert, A., Arrowood, M.J., Riggs, M.W., Fergsuon, M.A.J.: Characterization of a low molecular weight glycolipid antigen from *Cryptosporidium parvum*. J. Biol. Chem. **278**, 52212–52222 (2003)
9. Thomson, L.M., Lamont, D.J., Mehlert, A., Barry, J.D., Ferguson, M.A.J.: Partial structure of glutamic acid and alanine-rich protein, a major surface glycoprotein of the insect stages of *Trypanosoma congolense*. J. Biol. Chem. **277**, 48899–48904 (2002)