Abundance of tegument surface proteins in the human blood fluke *Schistosoma mansoni* determined by QconCAT proteomics

William Castro-Borges[a,*], Deborah M. Simpson[b], Adam Dowle[a,c], Rachel S. Curwen[a], Jane Thomas-Oates[c,d], Robert J. Beynon[b], R. Alan Wilson[a]

[a]Centre for Immunology & Infection, Department of Biology, University of York, Heslington, York, YO10 5DD, UK
[b]Protein Function Group, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool, L69 7ZB, UK
[c]Centre of Excellence in Mass Spectrometry, University of York, Heslington, York, YO10 5DD, UK
[d]Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK

ARTICLE INFO

ABSTRACT

The schistosome tegument provides a major interface with the host blood stream in which it resides. Our recent proteomic studies have identified a range of proteins present in the complex tegument structure, and two models of protective immunity have implicated surface proteins as mediating antigens. We have used the QconCAT technique to evaluate the relative and absolute amounts of tegument proteins identified previously. A concatamer comprising R- or K-terminated peptides was generated with $[^{13}C_6]$ lysine/arginine amino acids. Two tegument surface preparations were each spiked with the purified SmQconCAT as a standard, trypsin digested, and subjected to MALDI ToF-MS. The absolute amounts of protein in the biological samples were determined by comparing the areas under the pairs of peaks, separated by 6 *m/z* units, representing the light and heavy peptides derived from the biological sample and SmQconCAT, respectively. We report that aquaporin is the most abundant transmembrane protein, followed by two phosphohydrolases. Tetraspanin Tsp-2 and Annexin-2 are also abundant but transporters are scarce. Sm200 surface protein comprised the bulk of the GPI-anchored fraction and likely resides in the secreted membranocalyx. Two host IgGs were identified but in amounts much lower than their targets. The findings are interpreted in relation to the models of protective immunity.

1. Introduction

Schistosomes are platyhelminth parasites that can live for decades in the human portal vasculature, surrounded by humoral and cellular components of the immune system but apparently invulnerable to attack. Even more remarkable, these centimetre-long worms are covered by a naked syncytial layer of cytoplasm, called a tegument, which by rights should provide the perfect target. The outer surface of this layer, which forms the parasite-host interface, has an unusual multilaminate appearance [1] and its properties must be intimately linked with the worm's ability to evade attack. It is often referred to as a heptalaminate membrane [2], but the alternative interpretation, followed here, is that it represents a normal plasma membrane, overlain by a membrane-like secretion, termed a membranocalyx by analogy with the glycocalyx of conventional eukaryotic cells [3,4]. The cell bodies, containing the machinery for protein synthesis and export, lie below the musculature of the body wall and are connected to the tegument by narrow tubes of cytoplasm [5].

* Corresponding author at: Departamento de Ciências Biológicas, Núcleo de Pesquisas em Ciências Biológicas, Universidade Federal de Ouro Preto, Campus Morro do Cruzeiro s/n, Ouro Preto, Minas Gerais, Brazil. Tel.: +55 31 3559 1705.
E-mail address: williamcborges@hotmail.com (W. Castro-Borges).

Multilaminate vesicles originating there traffic to the syncytium to deliver their contents to the exterior at the base of tegument pits [1,3], thus contributing to a slow turnover of the membranocalyx into the blood stream [6,7]. The most plausible explanation for the properties and function of the membranocalyx is that it provides a relatively inert barrier to protect the underlying plasma membrane where conventional membrane enzymes and transporters reside. If this is correct, the membranocalyx must itself contain few macromolecules of parasite origin. However, its hydrophobic nature allows it to acquire lipophilic host molecules such as GPI-anchored membrane proteins [8] and erythrocyte glycolipids [9].

Methods have been developed to enrich the surface for biochemical analysis, with freeze/thaw/vortexing (F/T/V) being widely used to detach the membranes [10,11]. The advent of proteomic techniques permitted the composition of the total material released by F/T/V, to be investigated [12] but relatively few membrane-spanning proteins were identified. The application of density gradient centrifugation to enrich the surface membranes [11], followed by differential extraction with chaotropic agents of increasing strength yielded a final insoluble pellet containing primarily membrane proteins [13]. Subsequently, proteins on the surface of live worms accessible to biotinylation with sulfo-NHS-biotin were identified [14]. More recently, the surface proteins of live worms accessible to the enzymes trypsin and phosphatidyl-inositol phospholipase C (PiPLC) have also been documented [8]. The cumulative result of these studies is that we have a growing list of proteins which comprise the principal macromolecular constituents of the parasite interface. However, there is still much work required to place them in their correct context within the multilayered surface. In addition, and important for the identification of possible vaccine targets, we have no idea of the relative abundance of the various constituents.

The proteomic approach, linked to suitable fractionation procedures, is ideally suited to the qualitative identification of proteins in a complex mixture, but obtaining quantitative information has proved less tractable. The fractionation of soluble schistosome proteins by 2-dimensional electrophoresis (2-DE) before quantitation by staining and image analysis [15,16] provides a good indication of their relative abundance but the approach is not suitable for integral membrane proteins as ionic detergents are often required for their optimal extraction. The number of peptides identified per protein in tandem mass spectrometry has been used as a protein abundance index (PAI), normalised relative to the theoretical number of peptides [17]. The PAI was subsequently modified to an exponential form [18], and validated using synthetic peptides, giving a score roughly proportional to abundance. However, the response of a peptide in the mass spectrometer is a function of its amount, physiochemical properties and ability to ionise, which can all vary widely. The QconCAT approach has been developed to circumvent these limitations, allowing for the measurement of both relative and absolute abundance of multiple proteins within a mixture.

In the QconCAT approach, absolute quantification relies on the addition of isotopically labelled peptide internal standards (Q peptides) to the protein sample under investigation. The Q peptide has a different mass but otherwise identical physiochemical properties to the target peptide. These peptides therefore ionise similarly, meaning their relative abundances can be directly quantified from their responses in the mass spectrometer. Furthermore, since a known amount of Q peptide is added, the absolute amount of unmodified peptide can be calculated [19]. The key advance provided by the QconCAT technique is that, instead of chemically synthesising each Q peptide individually, it relies on the production of an artificial protein that is a concatamer of Q peptides (hence the term QconCAT) [20–22]. These are selected from proteins known to be present in the sample and facilitate the simultaneous quantitation of numerous target proteins. A gene encoding the requisite peptides is designed, and the concatamer is expressed heterologously in *Escherichia coli*. It can be labelled with stable isotopes by growth in the presence of *e.g.* [$^{13}C_6$]lysine and [$^{13}C_6$] arginine [22]. The labelled QconCAT protein is then purified, quantified and a known amount added to the complex protein sample under investigation. Tryptic digestion of the QconCAT-sample mix releases each of the QconCAT peptides in a strict 1:1 stoichiometry; subsequent MS analysis allows the quantification of each peptide present in the sample. We report here the design of a QconCAT gene encoding 33 signature peptides selected from different classes of protein identified in our previous proteomic studies of the schistosome tegument surface [13,14]. The labelled QconCAT standard was added to two tegument surface preparations, and the abundance of the selected peptides, and thus the proteins from which they were released, in each sample determined by mass spectrometry. Our results provide insights into the relative importance of processes occurring at the tegument surface of adult worms resident in the host bloodstream.

2. Materials and methods

2.1. Ethics statement

The procedures involving animals were carried out in accordance with the UK Animals (Scientific Procedures) Act 1986, and authorised on personal and project licences issued by the UK Home Office. The study protocol was approved by the Biology Department Ethical Review Committee at the University of York.

2.2. Selection of peptides, preparation and purification of the SmQconCAT construct

A total of 31 proteins was selected for quantification based on their previous identification in proteomic studies of the *S. mansoni* tegument [13,14]. Representative unique peptides were selected according to the criteria described by Pratt et al. [21]. Briefly, the majority of chosen peptides were in the 1000–2000 Da range, were either lysine or arginine terminated, and contained no cysteine or methionine residues. The artificial QconCAT gene [20] was synthesized *de novo* and cloned into pET21a by PolyQuant GmbH (Regensburg, Germany) as described [20]. The *E. coli* strain and media used were as follows: *E. coli* strain BL21(l)DE3 (*E. coli* B F⁻ dcm ompT hsdS$_B$(r$_B^-$ m$_B^-$) gal, l(DE3) was transformed with the vector and cultured in minimal media (1xM9 salts, 1 mM MgSO$_4$, 0.1 mM CaCl$_2$, 0.00005% (w/v) thiamine, 0.2% (w/v) glucose, unlabelled amino acids at 0.1 mg/mL or 0.2 mg/mL His, Tyr, Phe, Pro and Trp), supplemented either with unlabelled or

[$^{13}C_6$]arginine and [$^{13}C_6$]lysine at 0.1 mg/mL. The QconCAT was expressed in *E. coli* by induction with isopropyl-β-D-thiogalactopyranoside (IPTG) at mid-log phase and the cells were harvested after 5 h of induction by centrifugation at 1400 x *g* at 4 °C for 15 min. Inclusion bodies containing QconCAT were recovered by breaking cells using BugBuster Protein Extraction Reagent (Novagen, Nottingham, UK).

Inclusion bodies were redissolved in 20 mM phosphate buffer, 6 M guanidinium chloride, 0.5 M NaCl, 20 mM imidazole, pH 7.4. From this solution, [$^{13}C_6$]lysine/arginine-labelled or unlabelled QconCAT proteins were purified by affinity chromatography using a Ni-based resin (HisTrap HP Kit, Amersham Biosciences, UK). The column was pre-equilibrated with the same buffer used to solubilize the inclusion bodies. Following sample loading, HisTrap columns were washed with the same buffer prior to elution of the sample with the same buffer containing 500 mM imidazole and ten 1 mL fractions were collected. The purified QconCAT protein was desalted by three rounds of dialysis against 100 volumes 10 mM ammonium bicarbonate, pH 8.5, 1 mM dithiothreitol for 2 h using fresh buffer each time, exactly as described by Rivers et al., [22].

2.3. Molecular characterisation of the SmQconCAT

The homogeneity of the SmQconCAT preparation was evaluated by separation of a 3 μg aliquot by 1-DE using a pre-cast NUPAGE 4-12% Bis-Tris gel (Invitrogen) after a 45 min run at 200 V. The gel was then fixed in 40% methanol, 10% acetic acid for 30 min, stained with SYPRO Ruby (Invitrogen) for 2 h in the dark and imaged using a Molecular Imager FX (Bio-Rad, Bath, UK). A protein band corresponding to the expected molecular mass for SmQconCAT was excised from the gel and processed for in-gel digestion exactly as described by Castro-Borges et al. [23]. An aliquot of 1–2 μL of the digestion supernatant containing the peptides was spotted on a MALDI plate and dried before the addition of 0.6 μL of a saturated solution of α-cyano-4-hydroxycinamic acid (CHCA; Sigma, Poole, UK) in 50% (v:v) acetonitrile (ACN)/0.1% (v:v) trifluoroacetic acid (TFA). Positive-ion MALDI mass spectra (MS) were obtained using a 4700 Proteomics Analyzer with TOF/TOF Optics (Applied Biosystems, Framingham, USA) in reflector mode, over the *m/z* range 800–4000, and monoisotopic masses obtained from centroids of raw, unsmoothed data.

In order to optimize the concentration of SmQconCAT to be used for quantitation and to fragment each of the precursor ions to determine their amino acid sequence, two aliquots of 2.5 and 10 μg SmQconCAT (in 10 mM ammonium bicarbonate) were digested with trypsin in solution at 1:20 (enzyme:substrate) at 37 ° C over 24 h. A 3 μL aliquot of each digested sample was then injected onto a reversed-phase PS-DVB monolithic column (200 μm i.d.×5 cm, LC Packings, Amsterdam, Netherlands). Peptides were separated using a two-step linear gradient of 2–50% (v:v) ACN in 0.1% aqueous heptafluorobutyric acid over 30 min, followed by 50–80% (v:v) ACN in the same solvent over 5 min, at a flow rate of 3 μL/min; UV absorbance at 214 nm was monitored. Fractions were collected onto a MALDI target plate using a Probot (Dionex, Bannockburn, USA) with simultaneous addition of 6 mg/mL CHCA matrix solution in 60% (v:v) ACN. Positive-ion MALDI mass spectra (MS) were obtained essentially as described above. The precursor mass window was set to a relative

resolution of 50, and the metastable suppressor was enabled. The default calibration was used for product ion spectra, which were baseline-subtracted (peak width 50) and smoothed (Savitsky-Golay with three points across a peak and polynomial order 4); peak detection used a minimum S/N of 5, local noise window of 50 *m/z*, and minimum peak width of 2.9 bins. Product ion spectra were assigned from their respective peptide sequences.

2.4. Freeze/thaw removal of the tegument from adult *S. mansoni*

The tegument of adult worms was removed according to the freeze/thaw method of Roberts, S., et al. [11] with slight modifications. Briefly, the adult worm burden of 20 infected mice (approximately 1600 worms) that had been stored frozen at –80 °C in ~10×1 mL plastic microfuge tubes was slowly defrosted on ice after addition of 1 mL RPMI-1640 medium (Gibco Invitrogen, Paisley, UK) containing 20 μL protease inhibitor cocktail (PIC; Sigma, Poole, UK). After a single, gentle inversion, the thaw supernatant (S1) was removed from each tube and the adult worms were transferred to 10 mL plastic round bottom tubes, each filled to a depth of ~1 cm. Three millilitres of ice-cold RPMI supplemented with PIC were added before each tube was vortexed for 10, one second bursts. The vortex supernatant (S2), containing the detached tegument, was removed. S2 was spun at 120×*g*, at 4 °C for 30 min in a round bottomed tube in a GPR benchtop centrifuge (Beckman-Coulter, High Wycombe, UK). The supernatant (S2S) was again removed. The pellet (S2P), containing the tegument, was resuspended in ~100 μL RPMI-1640 before loading onto a sucrose gradient (20–70% w:v sucrose in 10 mM Tris-HCl, pH 7.4, 7 cm total depth in a round bottomed, clear polyethylene centrifuge tube). The gradient was spun at 112,000 ×*g* in an L8 centrifuge (Beckman-Coulter) using a SW41 swing out rotor at 4 °C for 45 min.

After centrifugation, the gradient was assessed for the presence of a milky band, approximately one third from the top of the tube. One millilitre fractions were collected from the top of the tube and each assayed for the presence of alkaline phosphatase (an apical membrane marker [11]) using BluePhos substrate (Kirkegaard & Perry Laboratories, Washington, USA) in a microtitre format. Samples containing the highest enzymatic activity were pooled and diluted at least 10 times in 10 mM Tris-HCl pH 7.4 before spinning for 40 min at 74,000×*g*, 4 °C in the L8 centrifuge with the SW41 rotor. The final pellet resulting from this centrifugation step was termed the Gradient Pellet (GP) and contained the enriched tegument membranes. The GP was resuspended in 0.5 mL 40 mM Tris-HCl pH 7.4 and stored frozen at –80 °C until use. Three separate biological replicate GPs were prepared, each divided into two technical replicates, making a total of six QconCAT analyses.

2.5. In-gel digestion of the GP fraction and LC-MS

The protein content of the membranous GP fraction was determined by separation of a 5 μL aliquot by 1-DE along with different concentrations of a soluble worm antigen preparation (SWAP) on the same gel. The GP fraction and reference standards were stained with SYPRO Ruby and analysed by densitometry. An 80 μg aliquot of the GP fraction (20 μL) was then reduced with

10 mM DTT and denatured for 10 min at 65 °C in the presence of 2× lithium dodecyl sulfate (LDS)–protein loading buffer (Invitrogen). The sample was then loaded onto a 1.0 mm 4–12% NuPAGE Bis-Tris Gel (Invitrogen) and allowed to migrate for 5 min at 200 V using 1×2-[N-Morpholino]ethanesulfonic acid (MES) running buffer (Invitrogen). Two aliquots of non-isotopically labeled QconCAT were run in parallel lanes to assess for co-migration with GP proteins. After electrophoresis, the gel was fixed in 40% (v:v) methanol/10% (v:v) acetic acid for 20 min and stained in Bio-Safe Coomassie overnight. Following a de-staining step in fixative solution, the Coomassie-stained smear of GP proteins was sliced into 5 gel pieces. In-gel digestion was then performed as previously described [23]. The recovered supernatant containing the GP peptides was concentrated in a SpeedVac (Thermo Lifescience, Basingstoke, UK) to dryness and peptides finally resuspended in 20 μL 0.1% (v:v) TFA. A 3 μL aliquot of the preparation was injected onto a reversed-phase PS-DVB monolithic column (200 μm i.d.×5 cm, LC Packings, Amsterdam, Netherlands). Fractions were collected onto a MALDI target plate using a Probot (Dionex, Bannockburn, USA) with simultaneous addition of 6 mg/mL CHCA matrix solution in 60% (v:v) ACN. Positive-ion MALDI mass spectra were obtained essentially as described above.

2.6. In-solution digestion of the GP fraction mixed with the SmQconCAT protein

An 80 μg aliquot of total GP fraction was submitted to in-solution digestion along with 2.5 μg of isotopically-labeled SmQconCAT. Briefly, the GP fraction was first solubilised in SDS, with final concentration of the detergent kept at 0.1%. Proteins were then reduced using 6 mM tris(2-carboxyethyl) phosphine for 1 h at 60° C. After cooling at room temperature, proteins were alkylated in the presence of 12 mM methyl methanethiosulfonate for 10 min in the dark. MS grade trypsin (Promega, Southampton, UK), previously dissolved in 0.5 M triethylammonium bicarbonate was added to a final ratio of 1:80 (trypsin:substrate) for the first experiment and 1:20 for the second and third biological replicate experiments. Trypsin digestion was allowed to proceed for 24 h at 37 °C. A stock 10% (v:v) TFA solution was then added to a final concentration of 0.1% before recovery of peptides by passage through a solid phase Strata C18-E extraction cartridge (55 μm, Phenomenex, Macclesfield, UK), followed by several column washes in 0.1% (v:v) TFA and final elution in 750 μL of 50% (v:v) ACN / 0.1% (v:v) TFA. The eluted fraction was concentrated under vacuum to dryness and peptides resuspended in 20 μL 0.1% (v:v) TFA. A 3 μL aliquot of the preparation was injected onto the reversed-phase PS-DVB monolithic column and peptides separated essentially as described above under a 2–50% (v:v) ACN linear gradient over 30 or 60 min.

2.7. MS spectrum and the search for light and heavy ions for quantification of tegumental proteins

An MS spectrum was acquired for each individual MALDI spot (180 in total per plate) over the m/z range 800–4000, using a 4700 Proteomics Analyser (Applied Biosystems) mass spectrometer. Each heavy peptide, originating from the SmQconCAT protein and its respective peptide light ion, derived by digestion of the GP fraction, were manually located from extracted ion chro-matograms, allowing for a 0.1 m/z unit tolerance. These were created by plotting the signal/noise ratio against spot position for each expected mass, using 4000 Series Explorer software (version 3.6, Applied Biosystems). Peptide pairs present in the same spectrum were selected for product ion analysis under the conditions described above. Product ion spectra were compared to those obtained following digestion of the SmQconCAT alone and assigned using their respective peptide sequences.

2.8. Monitoring of proteolyis of SmQconCAT and analyte proteins

In order to determine completeness of digestion of analyte and SmQconCAT proteins, the in-solution digestion protocol was performed essentially as described above. However a 10 μL aliquot of the digestion mix was removed and inactivated by addition of 10 μL of 10% (v:v) formic acid at the following time points: 0, 5, 10, 15, 20, 25, 30, 60, 120 min and 24 h. Each peptide fraction was then loaded onto C18 tips (Millipore) and the peptides eluted in 10 μL of 50% (v:v) ACN/0.1% (v:v) TFA. One microliter of each time point sample was spotted onto a MALDI plate, the peak area for each precursor ion recorded using the analysis tool available in the 4000 Series Explorer software (version 3.6, Applied Biosystems), and a graph plotted to determine the time for the reaction to reach its asymptote.

2.9. SmQconCAT and PiPLC-released proteins from live parasites

Adult parasites were obtained by liver perfusion of BALB/c mice previously infected with 180 cercariae. The recovered parasites were thoroughly and gently washed in RPMI-1640 medium, with cell debris and damaged parasites removed using a light microscope, essentially as described [8]. The remaining worms were transferred to a 30 mL flask and incubated for 1 h at 37 °C / 4% CO_2 in 5 mL of HEPES-buffered RPMI-1640 containing PiPLC (from *Bacillus cereus,* Sigma) at 1.25 U/mL. At the end of this procedure, the supernatant was removed and concentrated at 4 °C using a 5000 Da cut-off centrifugation device (Vivaspin 6, Sartorius Stedim Lab Ltd, Lincoln UK). Ten micrograms of this preparation was mixed with 2.5 μg of SmQconCAT and in-solution digestion was performed as described above. LC-MS was performed as described above and the target peptides quantified for comparison with the results obtained with the GP fraction. This experiment was performed three times.

2.10. Statistical analyses

The reproducibility of the QconCAT method was determined by comparing the means of abundance estimates (n=6) for the two peptides from alkaline phosphatase and Sm200 present in the construct, using the Student's t-test.

3. Results

3.1. Characterization and verification of the SmQconCAT

The sequences of the peptides selected for quantification of the 31 tegument proteins, the m/z values of their protonated

molecules and the accession numbers of the parent proteins are presented in Table 1. To assess the reproducibility of the method, alkaline phosphatase and Sm200 were each represented by two peptides in the protein construct, (Fig. 1A, C). In order to assess the homogeneity of the recombinant SmQconCAT after purification by nickel-affinity chromatography, a 3 μg aliquot was separated on a 4–12% NuPAGE gel under denaturing conditions. Coomassie staining of the gel revealed a protein of approximately 50 kDa, consistent with the predicted mass (Fig. 1B). The isotopically-labelled version of the SmQconCAT used for quantitation differs in mass by <200 Da, and produced a virtually identical electrophoretic profile (data not shown).

The protein band was then excised from the gel and subjected to in-gel digestion. The digest was analyzed by MALDI-TOF mass spectrometry (Fig. 1A). In this single-band analysis, the 33 constituent GP signature peptides were detected and their identities confirmed from their product ion spectra. However, at this initial stage we observed very low signals at m/z 1275, 1685 and 2661 for the peptides expected from the Na+/K+ATPase alpha subunit, Sm25 and Dynein

light chain, respectively. Although peptides in the construct are in equimolar proportions, and were selected precisely because they had been detected in previous analyses, we nevertheless observed up to a 20 fold difference in peak intensity between the individual peptides signals in this unfractionated mixture.

3.2. Protein content of the GP preparation

For quantification of tegument proteins we chose the gradient pellet (GP) fraction, (Fig. 2A), generated using the Freeze/Thaw/Vortex protocol to enrich for membrane proteins. As the majority of membrane protein assays are incompatible with detergents, we determined the protein content of the GP fraction by 1-DE, Coomassie staining and densitometric analysis, using defined amounts of soluble proteins from adult worm homogenate (SWAP) as the standards (Fig. 2B). A graphical representation of the linear relationship between optical density and the concentration of SWAP standards (Fig. 2C; $r^2=0.99$) provided a good estimate of the amount of GP used for tegument protein quantitation.

Table 1 – SmQconCAT signature peptides for quantitation of GP and PiPLC-released proteins.					
Peptide	Sequence	[M+H]$^+$	[M+H]$^+$ on SmQconCAT	Protein ID	Accession [a]
T2	VFDEIQQK	1006.5	1012.6	Tretraspanin, Tsp-2	Smp_181530
T3	VIGEVDEYPDYIK	1539.8	1545.8	200 kDa GPI-anchored surface protein — Sm200_b	Smp_017730
T4	GFLLLVEGGR	1060.6	1066.7	Alkaline phosphatase — AlkPhos_a	Smp_155890
T5	GHSFTTTAER	1106.5	1112.6	Actin	Sm01276
T6	LASIGFVLYR	1138.7	1144.7	Calpain	Smp_137410
T7	VFVLEDFER	1153.6	1159.6	Dysferlin	Smp_141010
T8	QNDSSNLYEITIR	1552.8	1558.8	EF-hand calcium binding protein — Sm20.8	Smp_074460
T9	SQIHDIVLVGGSTR	1481.8	1487.9	Heat shock protein 70 (Hsp70)	Smp_106930.1
T10	EVHTAWTQPR	1224.6	1230.7	Immunoglobulin gamma-3 heavy chain (*Mus musculus*)	BAA03476
T11	VNSAAFPAPIEK	1243.7	1249.7	Immunoglublulin gamma-1 heavy chain (*Mus musculus*)	AAA75163
T12	ILEIPFNSTNK	1275.7	1281.7	Na+/K+ATPase alpha subunit	Smp_015020
T13	LLHEVTPGESHIPR	1584.8	1590.9	Fimbrin	Sm13240, AAA29882
T14	YHNTIVEVISK	1302.7	1308.8	Annexin 2	Smp_077720
T15	QGNNEGPASESLLYPR	1731.8	1737.9	Glucose transport protein type 1 — GluTP1	Smp_012440
T16	DGIPINFPNENR	1385.7	1391.7	Sm29	Smp_072190
T17	AGHIHGAASDEVFK	1438.7	1444.8	Phosphodiesterase	Smp_153390.1
T18	FISLLPDVEIGEK	1459.8	1465.9	Gelsolin	Smp_008660.1
T19	SLLDNTYGR	1038.5	1044.6	EF-hand calcium binding protein - Sm22.6	Smp_045200
T20	NFSFGFLNFEFK	1496.7	1502.8	Voltage-dependent anion-selective channel (VDAC)	Smp_091240.1
T21	GYLNPVESNEEVR	1505.7	1511.8	High-affinity copper uptake protein (HACUP)	Smp_048230
T22	HHYIETIEFTVK	1516.8	1522.8	Alpha-2 macroglobulin	Smp_089670
T23	HPNYENQVDNLQR	1626.8	1632.8	Tetraspanin, TE736	Sm04463
T24	TIHGQIIQFK	1184.7	1190.7	200 kDa GPI-anchored surface protein - Sm200_a	Smp_017730
T25	SDFVVDVDYDDSHR	1668.7	1674.8	Aquaporin-3	Smp_005740
T26	GPFNTVGITGTVK	1290.7	1296.8	Alkaline phosphatase — AlkPhos_b	Smp_155890
T27	FSQYGDNLHK	1208.6	1214.6	Tetraspanin, Sm23	Smp_017430
T28	GFSPADISVQWLQR	1603.8	1609.9	Immunoglobulin M heavy chain (*Mus musculus*)	AAH96667
T29	LFIIYISLDLFSLK	1685.0	1691.0	GP18-22 (Sm25)	Smp_195180
T30	NFPIDQPVGSQR	1357.7	1363.7	Amino Acid transporter	ABQ02403
T31	QVIDAIETILPYEVR	1759.0	1765.0	Phospholipid scramblase	Smp_008860
T32	LIEDPLGSLDLFSVIR	1787.0	1793.0	ATP-diphosphohydrolase 1	Smp_042020.1
T33	EADVSLTAFVLIALQEAR	1946.1	1952.1	Complement component 3 (*Mus musculus*)	AAH43338
T34	HFGSYVTHETHNFIYFYLDDR	2661.2	2667.3	Dynein light chain	Smp_095520

[a] Acession numbers can be found at http://www.genedb.org/Homepage for parasite proteins and at http://www.ncbi.nlm.nih.gov/ for host proteins.
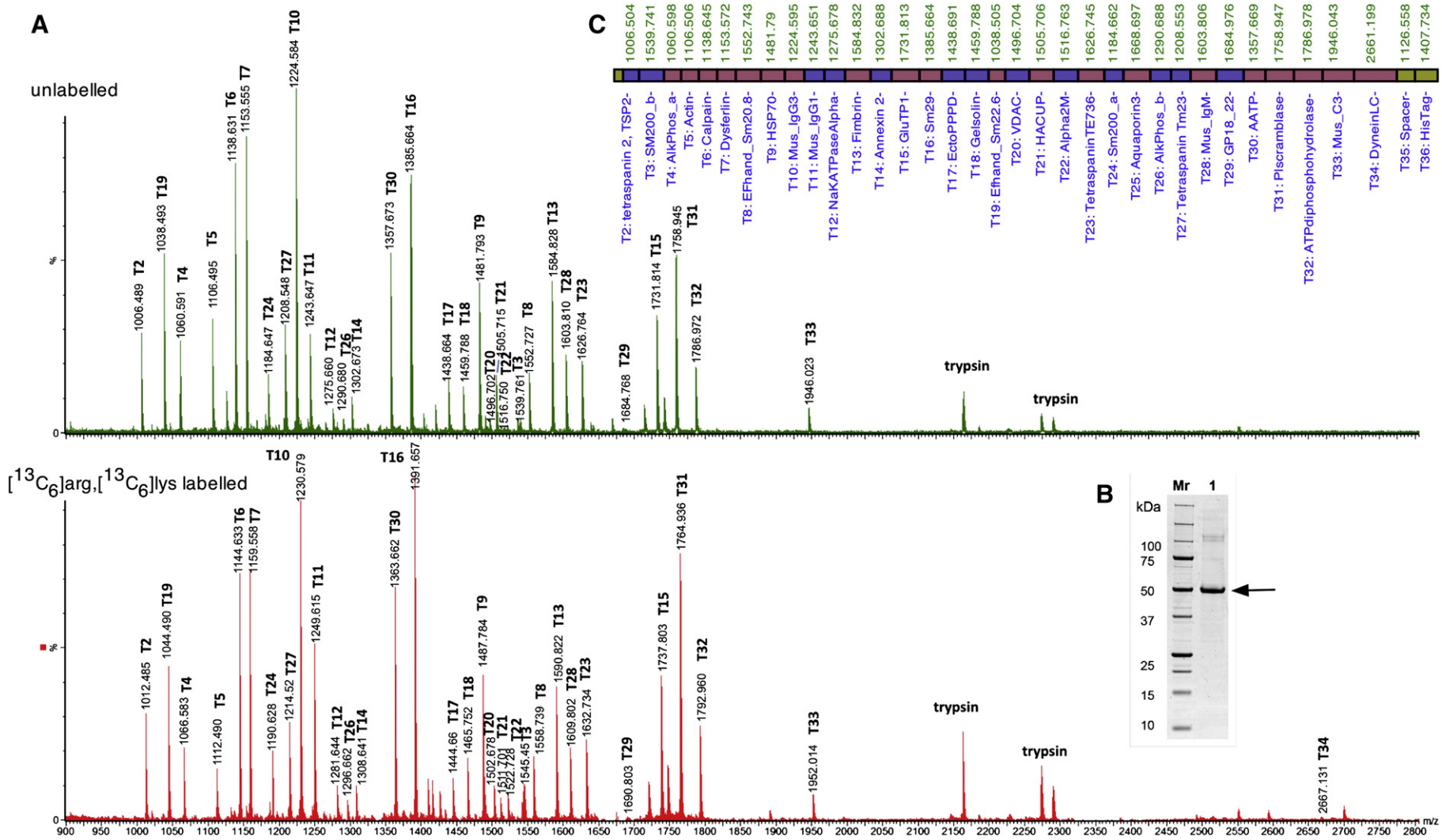
Fig. 1 – Molecular characterisation of the SmQconCAT construct. (A) Peptide mass fingerprint analysis of the tryptic peptides derived from in-gel digestion of the SmQconCAT (top and bottom spectra represent respectively the unlabelled and [$^{13}C_6$]lys/[$^{13}C_6$]arg labelled SmQconCAT). (B) Electrophoretic profile of a 3.0 μg aliquot of unlabelled SmQconCAT separated under denaturing conditions and stained with SYPRO Ruby. Note the expected Mr of the SmQconCAT at approximately 50 kDa (arrowed). (C) A schematic representation of the arrangement of lysine – (K, blue) and arginine – (R, red–brown) terminated tryptic (T) peptides in the SmQconCAT construct. Trypsin autolysis peaks are also indicated in the spectra.
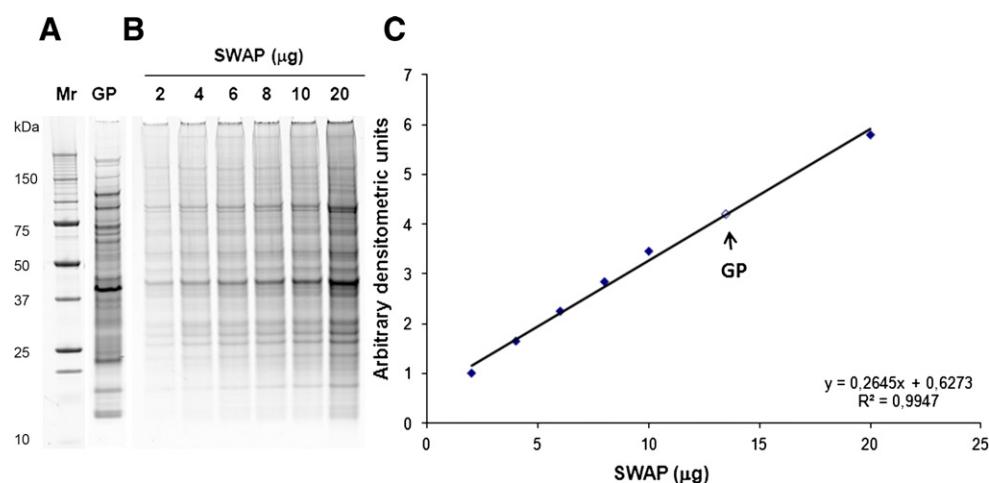
Fig. 2 – Electrophoretic profile and quantitation of the GP preparation. (A) A 5 μL aliquot of the GP preparation was separated alongside increasing amounts of SWAP as reference standards (B) under denaturing conditions and the gel stained with SYPRO Ruby. (C) Graphic representation of the linear relationship between optical density and the concentration of SWAP standards. The amount of protein contained in the 5 μL aliquot of GP shown on the gel, after densitometric analysis, equates to 13.5 μg protein (arrowed).

3.3. Optimization of tryptic digestion

We next determined whether in-gel or in-solution digestion would provide the most appropriate protocol for recovery of both analyte and isotopically-labeled standard peptides for quantification. MS fingerprinting revealed that up to 17 peptides could be detected following in-solution digestion whereas only 11 peptides were identified following in-gel digestion (Supplementary Fig. 1). Furthermore, after in-gel digestion we failed to detect analyte "signature peptides" known to be present in reasonable amounts e.g. for aquaporin, tetraspanin-Tsp2 and cooper ion transporter. This conclusion was based on our current knowledge of the protein composition of the GP using multi-dimensional liquid chromatography followed by tandem mass spectrometry (unpublished data). Among the peptides detected, a marked difference in signal intensity could be seen when the two digestion protocols were compared. Presumably, the rate of diffusion out of the gel depends on the size and physicochemical properties of the different peptides, which could in turn explain the range of responses from the various peptides in the different digests. Consequently, we adopted in-solution digestion for method optimization and quantitation purposes.

The QconCAT approach requires addition of a known amount of the standard to the analyte prior to proteolysis. We therefore optimized the amount of SmQconCAT added in order to maximise mass spectrometric responses and facilitate interpretation of the spectra. Two samples of SmQconCAT differing fourfold in concentration were digested and submitted to nano HPLC fractionation, followed by MS fingerprinting of the constituent peptides. A total of 30 out of 33 peptides could be detected in the 2.5 μg aliquot of digested SmQconCAT, the majority displaying a suitable signal to noise ratio (Supplementary Fig. 2). In contrast, 10 μg of the SmQconCAT digest, fractionated and analysed under the same conditions, did not produce a proportional increase in peptide ion intensities. In fact, in some cases peptide signals were detected at equivalent intensities regardless of whether 2.5 or 10 μg was analysed. Samples spiked with 2.5 μg of SmQcon-CAT proved adequate for all future analyses.

Finally, the digestion time that would result in optimal enzymatic cleavage of the proteins to be quantified was established. Tryptic digestion of GP plus 2.5 μg (50 pmol) SmQconCAT was monitored over a 24 h period. The area of each of the selected peptide peaks containing the light ion from the GP and/or its corresponding heavy ion from the SmQconCAT, was extracted using the 4700 Proteomics Analyzer operating software. The time course for digestion of four representative proteins is illustrated in Fig. 3. Although it was not possible to record the peak areas for all the peptide pairs, as we did not employ nano HPLC fractionation of the resulting sample, our data provide strong evidence for optimal tryptic digestion of the GP components and the SmQconCAT by 24 h. In fact, after digestion for 120 min the recorded peptide areas had reached a plateau, demonstrating the end of proteolysis during the initial phase of digestion.

3.4. Relative abundance of proteins in the GP preparation

Once the digestion time and the ratio of GP and SmQconCAT were optimised, determination of the relative abundance of the GP components depended exclusively on comparing the ratios obtained when each pair of heavy and light peptide ions were present in the spectra. The signature peptide for ATP-diphosphohydrolase (m/z 1787.0) present in the GP fraction separated by 6 m/z units from the corresponding $[^{13}C_6]$Arg-labelled version (m/z 1793.0) derived from the SmQconCAT, is representative of the data obtained (Fig. 4). The data can be displayed in two complementary ways. The absolute mass of proteins in the GP can be estimated using the formula Light
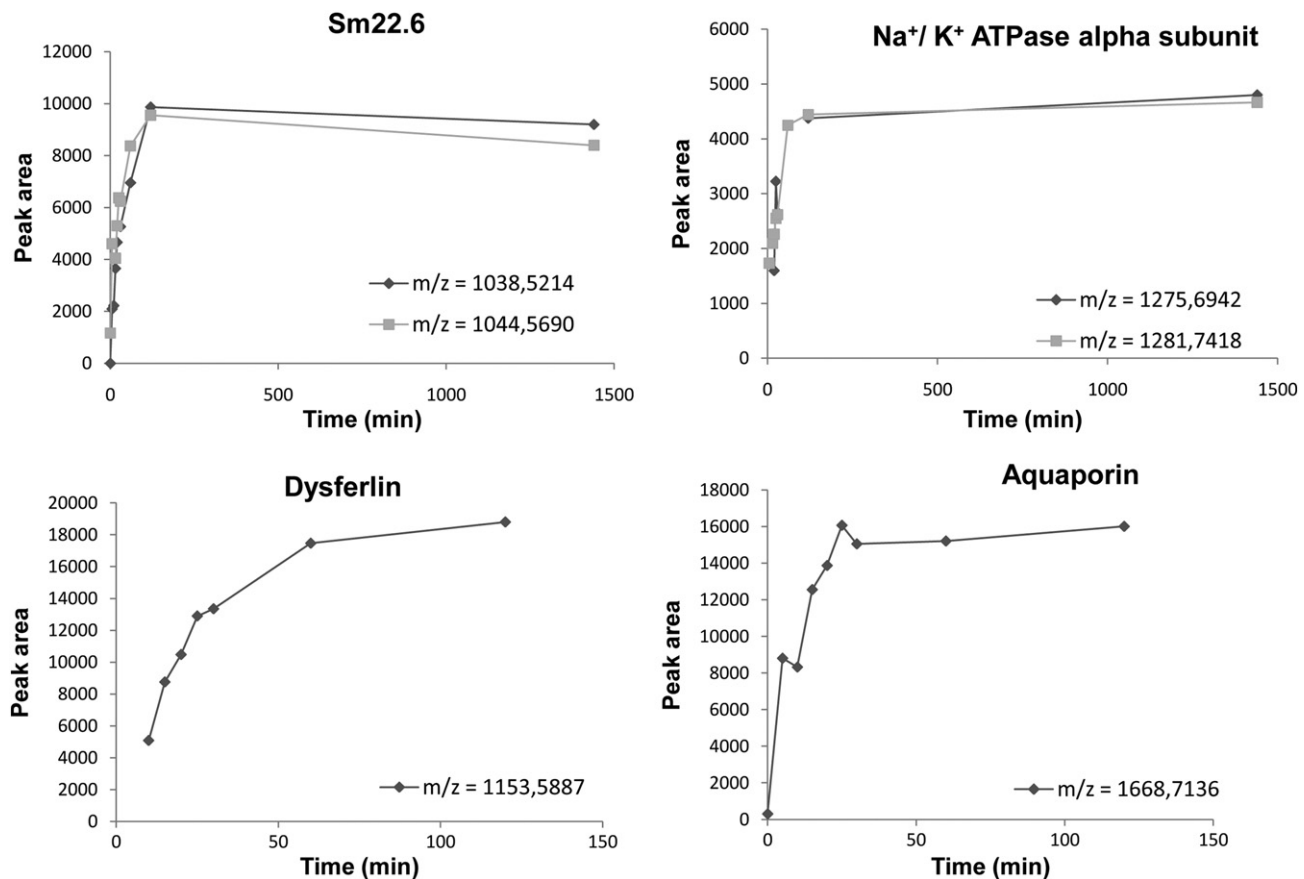
Fig. 3 – **Time course for tryptic digestion of the GP preparation spiked with SmQconCAT. Trypsin activity on GP and SmQconCAT proteins was monitored over 24 h at 37 °C at an enzyme-substrate ratio of 1:20. The peptide peak area profile of four representative peptides is shown, two (Sm22.6 & Na+/K+ ATPase) with light and heavy peptides from GP and SmQconCAT respectively (upper panels) and two (dysferlin and aquaporin) with light peptides only (lower panels). Note that tryptic digestion had virtually reached an asymptote in the first 2 h of the experiment attesting to optimal digestion of the GP-QconCAT mixture over the 24 h period.**

protein MW/Heavy QconCAT MW x light peak area/heavy peak area x 2.5 μg. For ATP-diphosphohydrolase, using the ratios in Fig. 4, $61{,}350/55{,}746.17 \times 40{,}488/14{,}243 \times 2.5\ \mu g = 7.82\ \mu g$ total protein = 9.8% of the GP protein. The alternative approach is to normalize the data relative to a chosen indicator protein. Since the proteins in the biological sample are each represented by a single peptide in the concatamer (except for Sm200 and alkaline phosphatase, from which two peptides were included) the relative plot provides an indication of the number of molecules of each in the preparation, independent of their mass.

By spiking 80 μg of GP with 2.5 μg of SmQconCAT, we were able to determine the abundance of 21 proteins in the GP fraction. These proteins fulfilled the criterion of their signature peptides being detected in at least three out of the six technical replicates. The compiled results are expressed as fold enrichment relative to the membrane-spanning anion channel as unity (Fig. 5; note the log scale of the y-axis). It is of note that the two peptides for Sm200 and alkaline phosphatase in the concatamer, showed very good agreement in their values indicating high reproducibility between the individual experiments (the means for each pair of peptides were not significantly different P > 0.05). In addition, the dynamic range of the SmQconCAT technique at the chosen concentration was > 200, and capable of estimating the concentration of aquaporin at 5.47x and fimbrin at 0.025x enrichment, relative to the anion channel in a single GP sample. A further five parasite and one host proteins (Sm20.8, Sm23, glucose transporter SGTP1, alpha-2 macroglobulin, Na+/K+ ATPase alpha subunit and complement C3) were detected in one or two replicates, while three proteins (calpain, dynein light chain and IgM heavy chain) were not detected in any sample.

The proteins identified are grouped according to their putative cellular functions. The relative abundance plot reveals that the solute transporter aquaporin is the dominant protein detected, in terms of number of molecules. In comparison, the amino acid and the copper ion transporters were 0.7× and 0.12× less abundant than the anion channel norm. The glucose transporter SGTP1 was not detected while SGTP4 was not in the concatamer. Of the three enzymes that hydrolyse organic phosphates, alkaline phosphatase (1.7×) and ATP-diphosphohydrolase (2.1×) were relatively abundant
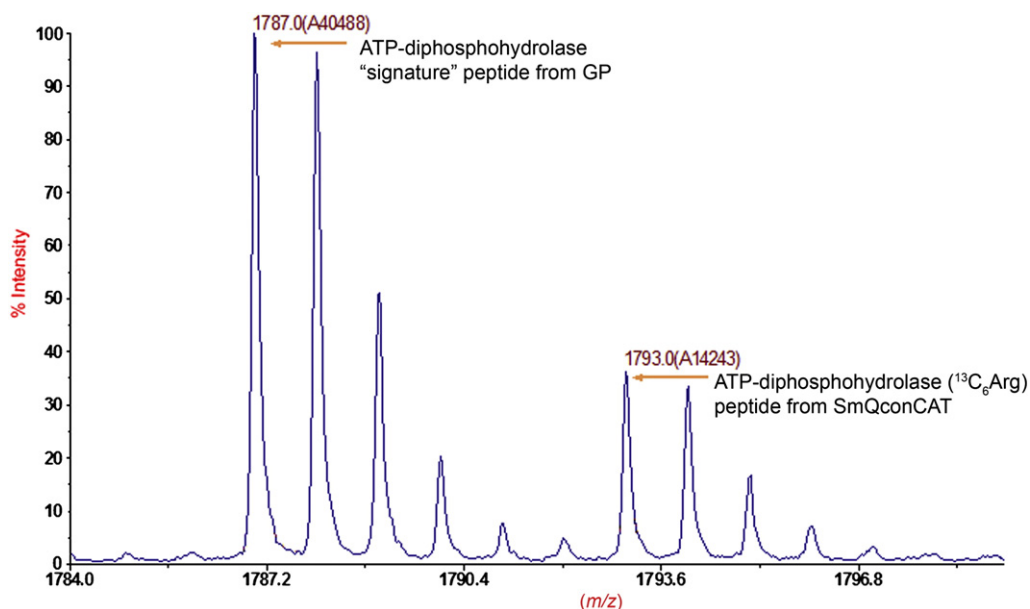
Fig. 4 – Specimen calculation of protein abundance A representative spectrum for quantitation of ATP-diphosphohydrolase showing the light peptide (L; m/z 1787.0) derived from tryptic digestion of GP separated by 6 m/z units from its corresponding heavy peptide (H; m/z 1793.0) originating from the added SmQconCAT. The area under the L (40488) and H (14243) peptide peaks is used to calculate the L/H ratio (2.84) from which absolute and relative amounts of the proteins in the 80 μg of GP is determined by reference to the 2.5 μg SmQconCAT added.

compared to phosphodiesterase (0.7×); the flippase (phospholipid scramblase, 0.6×) was the least detected. Dysferlin was the most abundant protein contributing to membrane structure (2.7×) while two tetraspanins were present in different amounts (1.5x and 0.4x) and a third (tetraspanin Sm23) was detected in only one replicate. The membrane-associated
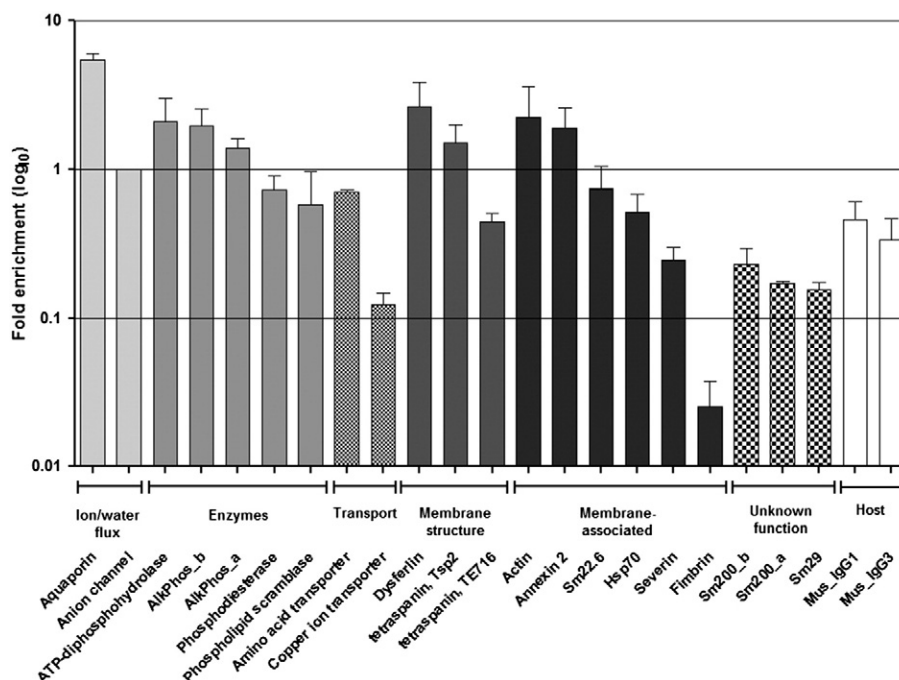
Fig. 5 – Quantitation of *S. mansoni* tegument proteins in the GP using the QconCAT approach. Relative abundance for 21 tegument proteins using the SmQconCAT construct. The mean fold enrichment ± S.D is expressed on a log scale relative to the anion channel as unity, providing information about abundance on a molar basis, i.e. the height of the bars is proportional to the number of molecules of each protein. Proteins are grouped according to their molecular function. Note the good agreement obtained for alkaline phosphatase and Sm200, each represented by two peptides in the SmQconCAT, to assess reproducibility of the technique.

Table 2 – Absolute quantification of GP and PiPLC-released proteins based on the SmQconCAT approach.

Protein identity	Protein MW (Da)	GP			PiPLC-released fraction		
		Ratio L/H	Amount in 80 μg	%	Ratio L/H	Amount in 10 μg	%
Aquaporin	32,890	9.518	14.05	18.1			
ATP-diphosphohydrolase	61,350	4.093	11.26	14.1			
AlkPhos_a (GFLLLVEGGR)	59,338	3.553	9.46	9.6	0.175	0.466	6.8
AlkPhos_b (GPFNTVGITGTVK)	59,338	2.225	5.92		0.336	0.894	
Dysferlin[a]	21,9170[a]	5.914	58.12*				
Tetraspanin Tsp-2	26,181	2.07	2.43	3			
Actin	41,953	4.142	7.79	9.7			
Annexin 2	39,295	4.019	7.08	8.9			
Sm200_a (TIHGQIIQFK)	189,192	0.403	3.42	3.75	0.827	7.02	87
Sm200_b (VIGEVDEYPDYIK)	189,192	0.302	2.56		1.22	10.35	
Sm29	21,200	0.267	0.25	0.31	0.046	0.044	0.44
Mus_IgG1 heavy chain	50,836	0.8	1.82	2.28			
Mus_IgG3 heavy chain	51,633	0.276	0.64	0.8			
			Σ 56.01 μg[b]	Σ 70.54		Σ 9.414 μg	Σ 94.2

[a] There are two gene models for Dysferlin, one with a MW of 9 kDa that would greatly reduce its contribution.
[b] excluding Dysferlin. Total amounts in GP and PiPLC-released fraction were calculated using the mean values obtained for Alkaline phosphatase and Sm200 peptides. Only the most abundant components of the GP and PiPLC-released fraction were selected for absolute quantification.

proteins comprise a diverse group that is not detached from the plasma membrane when the GP is prepared and washed in Tris buffer. Actin (2.24×) and annexin 2 (1.9×) are the most prominent, HSP70, Sm22.6 and severin (0.5, 0.74 and 0.24×, respectively) moderately so, but fimbrin is decidedly scarce in this group. Most surprising is the apparent paucity of the Sm200 surface protein (mean 0.2×, but see below). The other schistosome-specific protein, Sm29, was even less abundant (0.16×). The values (0.34×and 0.46×) obtained for the two IgG subclasses confirm the presence of host antibody firmly bound to the tegument surface but not especially abundant, while the IgM heavy chain was not detected and complement C3 found only in two technical replicates.

The absolute mass contribution of identified proteins to the 80 μg total protein of the GP sample can also be calculated (Table 2). Aquaporin is again the dominant protein representing 17% of the total. The ATP-diphosphohydrolase and alkaline phosphatase collectively comprise 24% of total proteins. Among the structural proteins, actin and annexin 2 account for 19% of total proteins, whilst the vaccine candidate tetraspanin Tsp-2, the most abundant of the three represented in the concatamer, is 3% of the total; the Sm200 surface protein, because of its high MW accounts for 3.75% of the GP protein content, while another vaccine candidate, Sm29, is only 0.3%. The dysferlin presents a problem because there are two gene models in the *S. mansoni* genome database (http://www.genedb.org/Homepage/Smansoni) both containing the signature peptide, the second with a MW of ~9000 Da. Using this value to calculate the dysferlin contribution reduces it to 0.95 μg (1.2% of the total).

3.5. PiPLC shaving of the tegument indicates Sm200 abundance

Our research has revealed that three tegument surface proteins represented in the SmQconCAT (Sm200, alkaline phosphatase, Sm29) were both accessible to biotinylation on live worms and could be removed by incubation with PiPLC. We therefore investigated the abundance of proteins represented in SmQconCAT, in the PiPLC-released fraction.
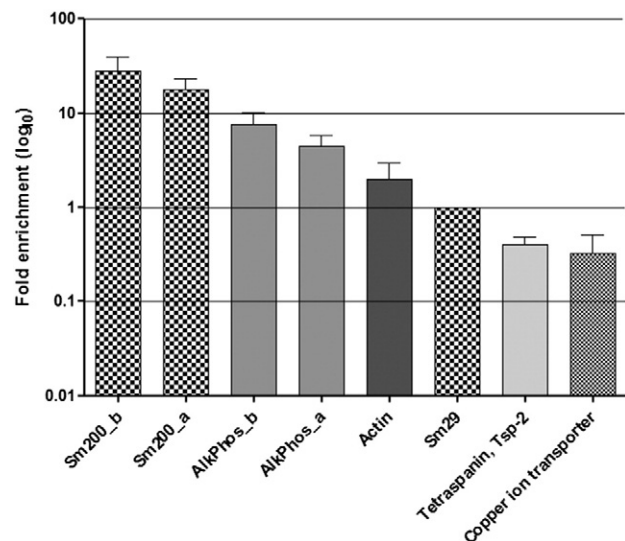
Fig. 6 – Quantitation of S. mansoni tegument proteins released after PiPLC shaving of live worms. Live worms were submitted to gentle shaving of the tegument surface using PiPLC and the recovered supernatant spiked with SmQconCAT for quantitation of GPI-anchored proteins. The mean fold enrichment±S.D for proteins found in the three analyses is expressed on a log scale relative to Sm29 as unity. Note the striking difference in molar abundance of Sm200 compared with its amount in the GP preparation. Sm200_a=TIHGQIIQFK; Sm200_b=VIGEVDEYPDYIK; AlkPhos_a=GFLLLVEGGR; AlkPhos_b=GPFNTVGITGTVK.

Replicate samples comprising 10 μg of PiPLC-released protein were obtained, much less than for the GP, but highly enriched in a small number of proteins. We were able to identify the two peptides representative of Sm200, the two peptides for alkaline phosphatase and the Sm29 signature peptide (Fig. 6). We also identified the copper ion transporter, tetraspanin-Tsp2 and actin as minor constituents. Normalizing the data relative to Sm29 as unity reveals that Sm200 is 23 times and alkaline phosphatase 5.7 times more abundant than Sm29 on a molar basis. Actin is also more abundant (2.0×) while tetraspanin (0.4×) and copper ion transporter (0.32×) are less abundant than Sm29.

When the absolute mass of three known GPI-anchored schistosome proteins identified in the PiPLC sample is calculated, we can see that Sm200 is by far the most abundant constituent (Table 2). It accounts for an estimated 87% of the protein in the sample, while the contribution of alkaline phosphatase is 6.8% and the Sm29 a mere 0.44%.

4. Discussion

The QconCAT technique was developed to quantitate known protein constituents in a biological sample, using signature peptides selected on the basis that they were found in previous MS/MS analyses. This means that they should have a good chance of generating a mass spectrometric response and being detected. However, there are constraints that do limit choice. Because the quantification was performed by MALDI-TOF MS, individual peptides in the concatamer must differ sufficiently in molecular weight to be distinguished in the peptide mass fingerprint, not forgetting that each will be represented as heavy and light versions of differing mass, in our case differing in mass by 6 Da [22]. It is also possible, when the QconCAT is added to a complex mixture of proteins like the GP, there will be unknown peptides present with a mass, close or identical to the target peptide pairs of interest, which interferes with quantitation. However, it proved possible to synthesise an artificial gene encoding >30 schistosome peptides that fulfilled the fundamental criteria. This enabled the simultaneous quantification of at least 21 proteins in the tegument surface membrane complex of S. *mansoni* accounting for approximately 80% of the total protein content of the GP. However, three peptides were never identified, even analysing the concatamer alone for fingerprinting analysis, among them the parasite dynein light chain, at 2660 Da the largest peptide in the construct.

The QconCAT concept is simple and elegant. The biological sample being spiked with an optimum quantity of the heavy QconCAT and then processed for MS in the normal way. In our case, the ratio of light to heavy peaks, 6 *m/z* units apart, permitted both absolute and relative quantities of the represented proteins to be determined. The one drawback is that a decision about what peptides to include has to be made at the outset; as work progresses the experimenter might want information about proteins not represented. Three of the seven GPI-anchored surface proteins released by PiPLC treatment of live worms [8] were included but not parasite carbonic anhydrase, ADP-ribosyl cyclase (SARC) and two CD59 orthologues, nor on the tegument-associated host proteins

CD44, CD48 and CD90 that we identified later. The reproducibility of the method was demonstrated by the lack of significant difference between the two estimates of abundance for Sm200 and alkaline phosphatase. We opted to analyse 80 μg aliquots of the GP containing hydrophobic membrane proteins, quantified by densitometry of hydrophilic proteins (SWAP) and this could affect the accuracy of the absolute value obtained.

A small number of proteins make up the major part of the gradient pellet, their preponderance presumably reflecting the importance to the parasite of the processes they mediate. Aquaporin is the chief constituent of the plasma membrane, in terms both of number of molecules and mass. It was identified by compositional analysis of the GP [13] and by biotinylation of live parasites [14]; its presence at the tegument surface of adult worms was subsequently confirmed by immunocytochemistry [24]. It appears as a pore controlling water influx, as a portal for drug entry [25] and it may also regulate the efflux of lactic acid, generated by adult worms in large quantities [24]. Aquaporins occur as tetramers in plasma membranes, each monomer serving as a separate pore with six membrane-spanning helices connected by five inter-helix loops, three external to the plasma membrane. It is thus similar to tetraspanin Tsp-2, the inter-helix loops of which have been developed as a vaccine candidate [26] and given the abundance of aquaporin, a similar approach may be worth pursuing.

The second major group are the phosphohydrolases; in number of molecules and mass, ATP-diphospohydrolase and alkaline phosphatase are each approximately half the abundance of aquaporin. In contrast only about one third the number of phosphodiesterase molecules is present, compared to the other two enzymes. All three were identified by compositional analysis [13] and biotinylation of live worms [14]; they have been demonstrated at the tegument surface by immunocytochemistry [27] [28,29]. Alkaline phosphatase is GPI-anchored and released by PiPLC treatment [8] whereas the other two phosphohydrolases possess membrane-spanning domains. Clearly, with their large ecto-domains these phosphohydrolases are capable of breaking down a range of organic phosphate substrates immediately external to the parasite. Free ATP in particular can act as a danger signal to promote inflammation [30] and it has been suggested that the surface ATP-diphosphohydrolase may remove ATP released by host cells as a result of damage inflicted by schistosomes migrating in the vasculature [29]. Whether the phosphohydrolases also serve the more prosaic function of breaking down phosphorylated metabolites into organic moiety and phosphate ions for separate uptake, or play a central role in immune evasion has yet to be confirmed experimentally but their abundance attests to their importance.

Compared with the above enzymes, the anion, amino acid and copper transporters are much less prevalent in the GP preparation. Note that only the amino acid transporter has been localised to the surface by immunocytochemistry [31]. All three were identified by proteomic analysis of the GP [13] but only the anion channel was biotinylated on live worms [14]. The indubitable surface location of the anion and copper transporters requires confirmation using specific antibodies. Of the three proteins involved in membrane structure, dysferlin

presents an anomaly. There are two gene models of widely different molecular weight, both containing the signature peptide, the smaller of which would place dysferlin as a minor component contributing<1 μg (1%) to the total GP. Alternatively dysferlin could represent fragments of contaminating material derived from the underlying muscle after F/T/V treatment. Indeed, in the mammal dysferlin is a muscle protein, mutations of which lead to one form of muscular dystrophy. It is associated with the plasma membrane and contains domains that may play a role in calcium-mediated membrane fusion events; it can also bind to caveolin present in lipid rafts. However, the biotinylation of dysferlin on live worms [14] argues against this contamination. An updated version of the *S. mansoni* genome, together with immunocytochemical localisation, should help resolve these issues.

Of the three tetraspanins represented in the QconCAT, two (Sm23 and Tsp-2, [32]; [26]) have been advocated as vaccine candidates. Tsp-2 is by far the most abundant tetraspanin constituent and also appears to be a very peripheral protein. Its surface location has been confirmed by immunocytochemistry [26]. It was biotinylated on live worms [14] and detected among the proteins released by enzymatic shaving of live worms with PiPLC (although not itself enriched; [8]), an observation confirmed here. We conclude that if any tetraspanin provides structure to the membranocalyx, then it is likely to be Tsp-2. The actin, severin and fimbrin in the GP preparation appear firmly attached to the plasma membrane. Only the actin, a known component of the tegumental spines [33] that are tightly apposed to the plasma membrane, makes a significant contribution (~10%) to the composition of the GP. It is also believed to form part of the dense cytoskeleton that early ultrastructural studies revealed lies beneath the plasma membrane [3]. The abundance of actin is notable in view of its interaction with BAR domain proteins that control membrane curvature, and with Rho-family GTPases [34]. The superficial location of the former proteins was revealed by trypsin treatment of live worms [8] whilst a number of small GTPases are also associated with the tegument surface (unpublished data). The severin and fimbrin, present in small amounts, may also be part of the spines and/or cytoskeleton via their ability to interact with actin. Finally, in the structural category, Annexin 2 is also abundant in the GP (~9%). Its surface location has been confirmed by immunocytochemistry [35] and it was biotinylated on live worms [14]. Indeed we have suggested that, given the known phospholipid-binding properties of this group of proteins, Annexin 2 could promote adhesion of the membranocalyx to the plasma membrane by acting like a molecular 'velcro' via the four binding domains it possesses [8].

Sm200 and Sm29, the two schistosome-specific proteins of unknown function quantified in the GP, are apparently only minor constituents (3.75% and 0.3% respectively in absolute terms but approximately the same number of molecules, due to the tenfold difference in MW). Sm200 protein has orthologues in the related trematode parasites *Fasciola hepatica, Clonorchis sinensis* and *Opisthorchis viverrini* (sequence information available at the Gasser Laboratory, Faculty of Veterinary science, University of Melbourne, Australia, http://bioinfosecond.vet.unimelb.edu.au/index.html). It would be instructive to know if these orthologues were similarly expressed solely on the tegument surface of those three flukes as it would imply a Class-wide conservation of function. Sm200, first cloned by Hall et al. [36] was already known to be GPI-anchored and located in the tegument surface [37,38]. Surprisingly, it was not found in initial compositional analysis of the GP [13] but was accessible to biotinylation in live worms [14] and was enriched by enzymatic shaving of live worms with both PiPLC and trypsin [8]. These observations indicate that it is readily lost during the processing of tegument membranes by differential extraction for MS analysis. Furthermore, its turnover into the vascular environment has been confirmed by its identification in circulating lipoprotein particles from the blood of schistosome-infected humans [39]. Sm29 has been demonstrated as localised to the tegument surface membranes [40] and it has been developed as a vaccine candidate [41]. It was identified by compositional analysis of the GP [13] and was biotinylated on live worms [14]. It was also the most highly enriched protein following PiPLC treatment of live worms [8], Sm200 was the second and alkaline phosphatase the third most enriched; these are values relative to untreated control worms and tell us nothing about abundance.

Relative and absolute quantitation of these three proteins in the PiPLC preparation versus GP preparations provides a dramatically divergent picture of their abundance. Displaying the PiPLC data relative to Sm29, reveals 23x as many Sm200 and 5.7x as many alkaline phosphatase molecules. In terms of mass, that means 198x as much Sm200 and 16x as much alkaline phosphatase. These observations accord with the results of 2D electrophoretic fractionation of PiPLC-released proteins (Fig. 1 in ref [8]) where Sm200 is dominant, alkaline phosphatase detectable and Sm29 not identified. Yet if we look at the amount of these three proteins in the GP there is only 14x as much Sm200 and 36x as much alkaline phosphatase, relative to Sm29. Thus, whilst the Sm29 and the alkaline phosphatase are in approximately the same proportions in the two preparations, Sm200 is in a great excess in the PiPLC fraction. How can these discrepancies be reconciled? It is possible that due to its large size and GPI anchor, Sm200 is easily detached from the membranes into solution during the thawing, vortexing, washing and centrifugation steps that lead to GP preparation. However, the obvious explanation is that the Sm200 segregates in a different 'compartment' to Sm29 and alkaline phosphatase, the membranocalyx, which separates from the plasma membrane during GP preparation. Indeed, during development of the F/T/V technique for tegument membrane isolation [11], multilamellar membrane aggregates were observed in the transmission electron microscope and it was speculated that these were accumulations of the outer bilayer, i.e. the membranocalyx. A sandwich ELISA to detect the presence of Sm200 in consecutive density gradient fractions, might finally provide a route to direct membranocalyx enrichment for full compositional analysis, exactly as alkaline phosphatase was used as the marker for the plasma membrane [11].

Host immunoglobulin heavy chains for IgM, IgG1 and IgG3 plus complement factor C3 were identified by biotinylation of live worms [13]. Complement factors C3 and C4 were released from live adults by trypsin shaving [8] and C4 has also been identified by compositional analysis of GP (unpublished data) but no members of the membrane attack complex C5–C9 have

been found. The first four of these host proteins were represented in the SmQconCAT. The amounts of IgG1 and IgG3 detected were small both in molar (one third to one half that of the anion channel) and absolute terms (2.3 and 0.8%, respectively of the GP). For comparison, the two C3 spectra obtained indicate the number of molecules is one tenth that of the anion channel, i.e. less than a third that of the two IgGs, amounting to 1% of the GP mass. This means that the bound immunoglobulins and complement factors are present at the tegument surface in small amounts, relative to their potential targets and the scarcity of C3 suggests that no amplification of the complement cascade is occurring. If the detected antibodies could be recovered intact, then they could be used in Western blotting assays to pinpoint their antigenic targets [42]. It would also be instructive to determine whether worms under immunological pressure, such as that applied by the rhesus macaque [43], had much more antibody bound at the tegument surface. The proposed mechanism of self cure in this primate host would predict that was the case.

Two animal models convincingly demonstrate acquired immunity to *S. mansoni* and so could provide a useful paradigm for the vaccine development. These are the radiation-attenuated (RA) cercarial vaccine, extensively characterised in laboratory mice [44] and to a lesser extent in primates [45,46], and the self cure phenomenon in rhesus macaques [43]. In the mouse model we have shown that challenge parasite elimination occurs in the lungs of vaccinated animals and is dependent on the recruitment of CD4+ T cells with Th1 characteristics. Independent work comparing the tetraspanin Sm23 with the RA vaccine supports this paradigm, revealing that failure to evoke pulmonary inflammatory responses following challenge elicited lower protection [47]. During priming of the immune response by RA larvae in the skin-draining lymph nodes, dendritic cells appeared to graze the intact tegument, suggesting the involvement of loosely attached surface antigens [48]. The current and associated studies [14],[8] give a clearer idea of what those mediating antigens might be, with the proviso that they may need to be detached from the surface for processing by accessory cells to trigger the recall immune response. Sm200 is pre-eminent in that list, because of its abundance, peripheral location and release to the external environment. Tetraspanin Tsp-2, though less abundant, is similarly prominent, especially if it proves to be located in the membranocalyx, and Annexin 2 may also be detachable. Aquaporin and the three phosphohydrolases are worth considering, especially alkaline phosphatase because of its GPI-anchor. The newly documented location of carbonic anhydrase, SARC and the two CD59 orthologues [8] makes them potential candidates, but there is no information about their abundance and location in the membranocalyx versus plasma membrane. It is noteworthy that almost all of the above-mentioned candidates are expressed in the lung stage schistosomulum target of protection in the RA vaccine model [8]. The obvious way forward to assess the importance of any or all of these proteins is to use recombinants to probe the skin-draining lymph node and pulmonary responses of mice exposed to the RA vaccine.

Self cure in the rhesus macaque is a protracted process, over weeks rather than hours, not an acute lethal killing of worms. It is mediated by a different mechanism whereby antibodies must reach a crucial threshold to inhibit vital worm functions [48]. For the tegument we have suggested they act by blocking the function of enzymes or transporters via their active sites. The current study highlights the suitability of aquaporin in this context, and the three phosphohydrolases are also potential targets with carbonic anhydrase and SARC also worth considering. Conversely, minor plasma membrane components could prove better targets because fewer molecules would need to be blocked. The anion and copper transporters (provided their localisation is confirmed), the amino acid transporter and SGTP-4 ([49]; not included in this study) are all valid minor constituents. All this presupposes that antibodies can reach the targets in sufficient quantities to exert an effect. Indeed, it seems likely that the active sites on key proteins will be protected from antibody binding, for example by the judicious positioning of glycan epitopes to exert steric hindrance. Thus the induction of antibodies with blocking activity may not be a trivial task and it would be instructive to determine whether self-curing rhesus macaques are better able to do this than permissive human or murine hosts. On a final note of caution about both models, we have suggested that the host response to a single protein is unlikely to be sufficient to disable the larval or adult schistosome [48].

Supplementary materials related to this article can be found online at doi:10.1016/j.jprot.2011.06.011.

REFERENCES

[1] Hockley DJ, McLaren DJ. *Schistosoma mansoni*: changes in the outer membrane of the tegument during development from cercaria to adult worm. Int J Parasitol 1973;3:13–25.

[2] Van Hellemond JJ, Retra K, Brouwers JF, van Balkom BW, Yazdanbakhsh M, Shoemaker CB, et al. Functions of the tegument of schistosomes: clues from the proteome and lipidome. Int J Parasitol 2006;36:691–9.

[3] Wilson RA, Barnes PE. The tegument of *Schistosoma mansoni*: observations on the formation, structure and composition of cytoplasmic inclusions in relation to tegument function. Parasitology 1974;68:239–58.

[4] Skelly PJ, Wilson RA. Making sense of the schistosome surface. Adv Parasitol 2006;63:185–284.

[5] Morris GP, Threadgold LT. Ultrastructure of the tegument of adult *Schistosoma mansoni*. J Parasitol 1968;54:15–27.

[6] Saunders N, Wilson RA, Coulson PS. The outer bilayer of the adult schistosome tegument surface has a low turnover rate in vitro and in vivo. Mol Biochem Parasitol 1987;25:123–31.

[7] Wilson RA, Barnes PE. The formation and turnover of the membranocalyx on the tegument of *Schistosoma mansoni*. Parasitology 1977;74:61–71.

[8] Castro-Borges W, Dowle A, Curwen RS, Thomas-Oates J, Wilson RA. Mass spectrometric identification of exposed proteins on the surface of the schistosome tegument released

by enzymatic shaving: a rational approach for selection of vaccine candidates. PLoS Negl Trop Dis 2011;5:e993.

[9] Goldring OL, Clegg JA, Smithers SR, Terry RJ. Acquisition of human blood group antigens by *Schistosoma mansoni*. Clin Exp Immunol 1976;26:181–7.

[10] Kusel JR. Protein composition and protein synthesis in the surface membranes of *Schistosoma mansoni*. Parasitology 1972;65:55–69.

[11] Roberts SM, MacGregor AN, Vojvodic M, Wells E, Crabtree JE, Wilson RA. Tegument surface membranes of adult *Schistosoma mansoni*: development of a method for their isolation. Mol Biochem Parasitol 1983;9:105–27.

[12] van Balkom BW, van Gestel RA, Brouwers JF, Krijgsveld J, Tielens AG, Heck AJ, et al. Mass spectrometric analysis of the *Schistosoma mansoni* tegumental sub-proteome. J Proteome Res 2005;4:958–66.

[13] Braschi S, Curwen RS, Ashton PD, Verjovski-Almeida S, Wilson A. The tegument surface membranes of the human blood parasite *Schistosoma mansoni*: a proteomic analysis after differential extraction. Proteomics 2006;6:1471–82.

[14] Braschi S, Wilson RA. Proteins exposed at the adult schistosome surface revealed by biotinylation. Mol Cell Proteomics 2006;5:347–56.

[15] Curwen RS, Ashton PD, Johnston DA, Wilson RA. The *Schistosoma mansoni* soluble proteome: a comparison across four life-cycle stages. Mol Biochem Parasitol 2004;138:57–66.

[16] Mathieson W, Wilson RA. A comparative proteomic study of the undeveloped and developed *Schistosoma mansoni* egg and its contents: the miracidium, hatch fluid and secretions. Int J Parasitol 2010;40:617–28.

[17] Rappsilber J, Ryder U, Lamond AI, Mann M. Large-scale proteomic analysis of the human spliceosome. Genome Res 2002;12:1231–45.

[18] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 2005;4:1265–72.

[19] Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 2003;100:6940–5.

[20] Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods 2005;2:587–9.

[21] Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. Nat Protoc 2006;1:1029–43.

[22] Rivers J, Simpson DM, Robertson DH, Gaskell SJ, Beynon RJ. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. Mol Cell Proteomics 2007;6:1416–27.

[23] Castro-Borges W, Cartwright J, Ashton PD, Braschi S, Guerra Sa R, Rodrigues V, et al. The 20S proteasome of *Schistosoma mansoni*: a proteomic analysis. Proteomics 2007;7:1065–75.

[24] Faghiri Z, Camargo SM, Huggel K, Forster IC, Ndegwa D, Verrey F, et al. The tegument of the human parasitic worm *Schistosoma mansoni* as an excretory organ: the surface aquaporin SmAQP is a lactate transporter. PLoS One 2010;5:e10451.

[25] Faghiri Z, Skelly PJ. The role of tegumental aquaporin from the human parasitic worm, *Schistosoma mansoni*, in osmoregulation and drug uptake. FASEB J 2009;23:2780–9.

[26] Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, et al. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. Nat Med 2006;12:835–40.

[27] DeMarco R, Kowaltowski AT, Mortara RA, Verjovski-Almeida S. Molecular characterization and immunolocalization of *Schistosoma mansoni* ATP-diphosphohydrolase. Biochem Biophys Res Commun 2003;307:831–8.

[28] Rofatto HK, Tararam CA, Borges WC, Wilson RA, Leite LC, Farias LP. Characterization of phosphodiesterase-5 as a surface protein in the tegument of *Schistosoma mansoni*. Mol Biochem Parasitol 2009;166:32–41.

[29] Bhardwaj R, Skelly PJ. Purinergic signaling and immune modulation at the schistosome surface? Trends Parasitol 2009;25:256–60.

[30] Bours MJ, Swennen EL, Di Virgilio F, Cronstein BN, Dagnelie PC. Adenosine 5′-triphosphate and adenosine as endogenous signaling molecules in immunity and inflammation. Pharmacol Ther 2006;112:358–404.

[31] Krautz-Peterson G, Camargo S, Huggel K, Verrey F, Shoemaker CB, Skelly PJ. Amino acid transport in schistosomes: characterization of the permeaseheavy chain SPRM1hc. J Biol Chem 2007;282:21767–75.

[32] Da'dara AA, Skelly PJ, Wang MM, Harn DA. Immunization with plasmid DNA encoding the integral membrane protein, Sm23, elicits a protective immune response against schistosome infection in mice. Vaccine 2001;20:359–69.

[33] Zhou Y, Podesta RB. Surface spines of human blood flukes (*Schistosoma mansoni*) contain bundles of actin filaments having identical polarity. Eur J Cell Biol 1989;48:150–3.

[34] Saarikangas J, Zhao H, Lappalainen P. Regulation of the actin cytoskeleton-plasma membrane interplay by phosphoinositides. Physiol Rev 2010;90:259–89.

[35] Tararam CA, Farias LP, Wilson RA, Leite LC. *Schistosoma mansoni* Annexin 2: molecular characterization and immunolocalization. Exp Parasitol 2010;126:146–55.

[36] Hall TM, Joseph GT, Strand M. *Schistosoma mansoni*: molecular cloning and sequencing of the 200-kDa chemotherapeutic target antigen. Exp Parasitol 1995;80:242–9.

[37] Sauma SY, Strand M. Identification and characterization of glycosylphosphatidylinositol-linked *Schistosoma mansoni* adult worm immunogens. Mol Biochem Parasitol 1990;38:199–209.

[38] Sauma SY, Tanaka TM, Strand M. Selective release of a glycosylphosphatidylinositol-anchored antigen from the surface of *Schistosoma mansoni*. Mol Biochem Parasitol 1991;46:73–80.

[39] Sprong H, Suchanek M, van Dijk SM, van Remoortere A, Klumperman J, Avram D, et al. Aberrant receptor-mediated endocytosis of *Schistosoma mansoni* glycoproteins on host lipoproteins. PLoS Med 2006;3:e253.

[40] Cardoso FC, Pacifico RN, Mortara RA, Oliveira SC. Human antibody responses of patients living in endemic areas for schistosomiasis to the tegumental protein Sm29 identified through genomic studies. Clin Exp Immunol 2006;144:382–91.

[41] Cardoso FC, Macedo GC, Gava E, Kitten GT, Mati VL, de Melo AL, et al. *Schistosoma mansoni* tegument protein Sm29 is able to induce a Th1-type of immune response and protection against parasite infection. PLoS Negl Trop Dis 2008;2:e308.

[42] Cutts L, Wilson RA. The protein antigens secreted in vivo by adult male *Schistosoma mansoni*. Parasitology 1997;114(Pt 3):245–55.

[43] Wilson RA, Langermans JA, van Dam GJ, Vervenne RA, Hall SL, Borges WC, et al. Elimination of *Schistosoma mansoni* adult worms by rhesus macaques: basis for a therapeutic vaccine? PLoS Negl Trop Dis 2008;2:e290.

[44] Coulson PS. The radiation-attenuated vaccine against schistosomes in animal models: paradigm for a human vaccine? Adv Parasitol 1997;39:271–336.

[45] Kariuki TM, Farah IO, Yole DS, Mwenda JM, Van Dam GJ, Deelder AM, et al. Parameters of the attenuated schistosome vaccine evaluated in the olive baboon. Infect Immun 2004;72:5526–9.

[46] Eberl M, Langermans JA, Frost PA, Vervenne RA, van Dam GJ, Deelder AM, et al. Cellular and humoral immune responses

and protection against schistosomes induced by a radiation-attenuated vaccine in chimpanzees. Infect Immun 2001;69:5352–62.

[47] Riengrojpitak S, Anderson S, Wilson RA. Induction of immunity to Schistosoma mansoni: interaction of schistosomula with accessory leucocytes in murine skin and draining lymph nodes. Parasitology 1998;117(Pt 4):301–9.

[48] Wilson RA, Coulson PS. Immune effector mechanisms against schistosomiasis: looking for a chink in the parasite's armour. Trends Parasitol 2009;25:423–31.

[49] Skelly PJ, Shoemaker CB. Rapid appearance and asymmetric distribution of glucose transporter SGTP4 at the apical surface of intramammalian-stage Schistosoma mansoni. Proc Natl Acad Sci USA 1996;93:3642–6.