A next-generation proteome array for *Schistosoma mansoni*

Rafael Ramiro de Assis\(^a,c\), Fernanda Ludolf\(^b\), Rie Nakajima\(^c\), Al Jasinskasc, Guilherme C. Oliveira\(^d\), Philip L. Felgner\(^c\), Soraya T. Gaze\(^a\), Alex Loukas\(^e\), Philip T. LoVerde\(^f\), Jeffrey M. Bethony\(^a,g,*\), Rodrigo Correa-Oliveira\(^a\), Carlos E. Calzavara-Silva\(^a,c\)

\(^a\) Laboratório de Imunologia Celular e Molecular, Instituto Rene Rachou, Fiocruz, Belo Horizonte, MG, Brazil
\(^b\) Infectología e Medicina Tropical, Facultade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
\(^c\) Protein Microarray Laboratory, Division of Infectious Disease, School of Medicine, University of California, Irvine, CA, USA
\(^d\) Vale Technology Institute of Sustainable Development, Belém, PA, Brazil
\(^e\) Centre for Biodiscovery and Molecular Development of Therapeutics, Australian Institute of Tropical Health & Medicine, James Cook University, Cairns, QLD, Australia
\(^f\) Departments of Biochemistry and Pathology, School of Medicine, University of Texas Health Science Center, 7703 Floyd Curl Dr., San Antonio, TX 78229, USA
\(^g\) Department of Microbiology, Immunology, and Tropical Medicine, School of Medicine and Health Science, The George Washington University, Washington, DC, USA

**Abstract**

A proteome microarray consisting of 992 *Schistosoma mansoni* proteins was produced and screened with sera to determine antibody signatures indicative of the clinical stages of schistosomiasis and the identification of subunit vaccine candidates. Herein, we describe the methods used to derive the gene list for this array (representing approximately 10% of the predicted *S. mansoni* proteome). We also probed a pilot version of the microarray with sera from individuals either acutely or chronically infected with *S. mansoni* from endemic areas in Brazil and sera from individuals resident outside the endemic area (USA) to determine if the array is functional and informative.

© 2016 Published by Elsevier Ltd on behalf of Australian Society for Parasitology.

Schistosomiasis remains among the most important of the neglected tropical diseases (NTDs), with some estimates indicating that chronic schistosomiasis may cause as many as 28 million Disability Adjusted Life Years (DALYs) lost and 280,000 deaths annually (King, 2010). Based on such estimates, schistosomiasis rivals malaria as one of the world’s most important tropical infections (King, 2010). *Schistosoma mansoni* infection usually peaks in early adolescence and declines in adulthood, an age-related pattern that has led to the hypothesis that individuals in endemic areas acquire resistance to infection (Woolhouse and Hagan, 1999). While several host age-related factors have been hypothesised (e.g., alterations in hormone level, thickening of the skin, host behaviour, etc. (Fulford et al., 1998)), the most compelling explanation for resistance remains the acquisition of immunity coincident with the natural death of the worms in the host, which occurs approximately 10–15 years after initial infection (Woolhouse and Hagan, 1999). It is further hypothesised that dying adult worms release previously inaccessible “surface” and “subsurface” schistosome antigens, exposing those to the host immune system (Woolhouse and Hagan, 1999), which expand until a repertoire of antibodies evolves into a protective immune response. An accelerated version of this naturally acquired immunity occurs when individuals undergo repeated rounds of treatment with the anthelmintic drug praziquantel (PZQ), which increases the permeability of the membranes of schistosome cells to calcium ions, inducing a contraction of the parasite, paralysis and the release of antigens.

Both natural and drug-induced resistance have the following characteristics as discussed previously (Woolhouse and Hagan, 1999): antibodies are generated to a repertoire of surface and subsurface antigens released by worm death or PZQ, which gradually expands until resistance is achieved. Due to the number and the diversity of the antigens released by parasite death or PZQ, traditional serological technologies such as a simplex indirect ELISA (one antigen at a time) are unable to capture the “repertoire” of antibodies generated during the progression to natural and drug-induced immunity. Recent advances in high order multiplexing, such as protein microarrays, provide a high-throughput technology that can comprehensively capture the "immunome" associated with both of these forms of resistance, i.e., the "repertoire" of antibodies against the antigens or epitopes that interface with the host immune system during natural and drug-induced immunity (Gaze et al., 2014; Driguez et al., 2016). Protein array
technology takes advantage of genomics to select and design open reading frames (ORFs) for each annotated gene, with the protein products expressed in a cell-free system and printed in an array format (Gaze et al., 2014; Driguez et al., 2016). The arrays can then be probed efficiently and economically with small volumes of host sera or plasma. The high-throughput nature of the proteome microarray approach to antigen discovery is the key to determining the repertoire of antigens that can elicit defined immune responses in schistosomiasis and determine patient-to-patient and species-dependent differences in the response during natural and drug-induced immunity.

Protein microarrays corresponding to ≥80% of the Plasmodium falciparum proteome have been used to screen for immunogenicity of malaria proteins with known as well as unknown functions (Dent et al., 2015). More recently, Arama et al. (2015) associated the overall antigen specificity of IgM and IgG responses, influenced by the host genetic background, with protection from malaria (Arama et al., 2015). By assessing genome-wide serological data of S. mansoni infections, we have attempted to uncover mechanisms driving natural and drug-induced immunity such as host age, sex, residence, behaviour (e.g., water contact) and chemotherapy (e.g., number of times treated with PZQ). A first generation schistosome protein microarray consisted of 230 Schistosoma japonicum sequences and 50 S. mansoni sequences (Driguez et al., 2016). The S. mansoni full-length cDNA sequences were selected based on the identification of the corresponding proteins in the tegument by proteomic studies, with a focus on proteins with transmembrane domains and signal peptides (Gaze et al., 2014; Pearson et al., 2015; Driguez et al., 2016). The first generation protein array was then screened with sera from a cohort of individuals from Brazil, some of whom were “putatively resistant” to S. mansoni infection and some of whom were chronically infected with schistosomiasis. These two cohorts showed discrete antigen recognition profiles with distinct isotype preferences for different antigens (Gaze et al., 2014). The results of this screen provided the proof-of-principal that a schistosome protein array could be a useful tool by which to analyse the protective immune response in schistosomiasis and determine patient-to-patient and species-dependent differences in the response during natural and drug-induced immunity.

Herein, we describe the production of a ‘next generation’ S. mansoni protein microarray, with over 900 S. mansoni antigens (Fig. 1). While numerous proteome arrays have been against viral (e.g. Dengue virus, Vaccinia virus), bacterial (e.g. Leptospira interrogans, Brucella melitensis) and protozoan pathogens (e.g. P. falciparum) (Cannella et al., 2014; Lessa-Aquino et al., 2015), few arrays have been made for multicellular pathogens; indeed the only ones we are aware of are the pilot schistosome arrays (Chen et al., 2014; Gaze et al., 2014; Pearson et al., 2015) and a 623 protein Necator americanus (hookworm) array (Tang et al., 2014). The next generation S. mansoni proteome array we described herein consisting of 992 recombinant proteins is the largest array for a helminth parasite.

One of the challenges for construction of eukaryotic protein microarrays is the sheer number of proteins encoded in the genomes of these complex organisms. ORFs need to be amplified from cDNA due to the presence of introns, and specific filtering methods are required to deal with stage-specific gene expression patterns and secondary structure of proteins (Berriman et al., 2009). The proteins comprising the next generation S. mansoni protein microarray were selected using four criteria. The primary criterion was the presence of the protein on the surface of the tegument of S. mansoni, as determined by enzymatic shaving studies using trypsin or phosphatidylinositol-specific phospholipase C (PiPLC) treatment (Castro-Borges et al., 2011a,b), as well as biotinylation of tegument proteins (Brischi and Wilson, 2006; Sotillo et al., 2016). A second criterion was the selection by seropositivity of the protein when S. mansoni adult worm antigen extracts were screened with sera from individuals resident in endemic areas using two-dimensional gel electrophoresis (Ludolf et al., 2014). The third criterion was a bioinformatics analysis of the predicted proteome by selecting proteins with signal peptides and/or transmembrane motifs as predictors of cellular localisation and epitope prediction using TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) and BepiPred 1.0 (http://www.cbs.dtu.dk/services/BepiPred/).

The selected proteins were further classified as having higher expression in the cercariae, schistosomula or adult stages, as well
Fig. 2. Recognition profile of infected and non-infected individuals to schistosome antigens printed on the pilot microarray. The 92 proteins randomly selected to compose the pilot array were probed (1:50) using sera of two negative individuals from the USA, two chronically infected individuals and two Schistosoma mansoni-infected individuals collected during the acute phase of the infection, all from Brazil. Anti-IgG (1:200) was used as the secondary antibody. Proteins were also positive in the first generation S. mansoni array (Gaze et al., 2014). (A) Representative heatmap of selected S. mansoni proteins from sera of infected and non-infected individuals. From 92 proteins tested in the pilot array, we selected 52 positive proteins to generate the representative heatmap, ordered based on chronic sample results. Data from duplicates of negative control sera and S. mansoni-infected sera were extracted and the means of signal intensities of each group were used to construct the representative heatmap. (B) Recognition of each tested protein. From the 92 probed proteins, eight proteins were differentially recognised between: A, negative controls and chronically infected individuals; B, negative controls and acute phase infected individuals; and C, chronically infected individuals and acute phase infected individuals. A Kruskal–Wallis multiple comparison test with Dunn’s correction was used to perform statistical comparisons. (C) Overall recognition of the 92 tested proteins from sera of negative controls, chronically infected patients and individuals during the acute phase of infection. P values were retrieved using the two-tailed Mann–Whitney statistical test.
Comparison of antigens recognised by IgG in sera from individuals with chronic and acute *Schistosoma mansoni* between the expanded 992 antigen *S. mansoni* array and the previous 37 antigen *S. mansoni* array (Gaze et al., 2014).

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Description</th>
<th>Selection method</th>
<th>Immunoreactivity by infection status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SMP_171190.1</strong></td>
<td>MEG-8 family</td>
<td>Bioinformatic</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_014570.2</strong></td>
<td>Saposin, JPR008139 Saposin</td>
<td>Bioinformatic</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_040790.1</strong></td>
<td>Peptidyl prolyl cis trans isomerase B</td>
<td>Bioinformatic</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_048230.1</strong></td>
<td>Solute carrier family 31 (copper transporters)</td>
<td>Surface</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_002740.2</strong></td>
<td>Hypothetical protein</td>
<td>Expression</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_030370.1</strong></td>
<td>Caleculin</td>
<td>Serologic</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_149590.1</strong></td>
<td>Gamma aminobutyric acid receptor subunit</td>
<td>Bioinformatic</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_194970.1</strong></td>
<td>25 kDa integral membrane protein</td>
<td>Surface</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_194960</strong></td>
<td>Similar to tetraspanin</td>
<td>Expression</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>SMP_045200.1</strong></td>
<td>Tegument-allergen-like protein</td>
<td>Surface</td>
<td>-</td>
</tr>
<tr>
<td><strong>SMP_072190.1</strong></td>
<td>Membrane associated protein 29</td>
<td>Surface</td>
<td>-</td>
</tr>
<tr>
<td><strong>SMP_005740.1</strong></td>
<td>Aquaporin 9 (Small solute channel 1)</td>
<td>Bioinformatic</td>
<td>-</td>
</tr>
</tbody>
</table>

*Proteins also used on the previous described array by Gaze et al. (2014).*

As sex-biased gene expression (Protasio et al., 2012; Rofatto et al., 2013), which constituted the fourth criterion. Two hundred and sixty-one proteins were discovered by biotinylation or enzymatic shaving of the tegument, 59 proteins were selected due to seropositivity, 219 proteins were selected by bioinformatics analysis and 453 proteins were selected based on stage or sex-biased gene expression. The final list of proteins comprised close to 10% (n = 992) of the entire predicted 10,852 proteins of *S. mansoni* (Protasio et al., 2012) as shown in Fig. 1 and are available in Mendelian Data (https://data.mendeley.com, doi: dx.doi.org/10.17632/jm8ygfd9hm.1). Approximately 10% of the 992 proteins on the array were randomly sampled and printed on a pilot array (Fig. 1 and Supplementary Table S1).

PCR amplification of predicted ORFs was carried out using cDNA from both cercariae and adult worms. Any ORFs longer than 3000 bp were amplified and cloned as overlapping fragments (at least 100 bp overlap). cDNA was synthesised by reverse transcription from total RNA extracted from *S. mansoni* cercariae and adult worms using oligoDT and random hexamers as primers and the Improm-II reverse transcriptase system (Promega, WI, USA) according to the manufacturer’s protocol, with the exception of annealing which was performed at 25 °C for 5 min. Primers were designed to contain both 20 nucleotides from the start or end of target sequence and 20 nucleotides from the recombinant site of the vector (ACGACACACATATGTCGTAG on the forward primer and TCCGGACATCTGCTGGA on the reverse primer) in order to allow homologous recombination with the pXT7 plasmid (Davies et al., 2005). A list of the primers is available in Mendelian Data (dx.doi.org/10.17632/jm8ygfd9hm.1). PCR amplification of each ORF was carried out using the HotMaster Mix (5 Prime, MD, USA) and 10 ng of *S. mansoni* cDNA per 25 μl reaction.

The *S. mansoni* proteome array was produced using a high-throughput approach in which competent DH5α *Escherichia coli* cells were incubated with a mixture of the pXT7 plasmid and target amplicons in an in vivo recombination procedure (Davies et al., 2005). Cloning success was verified by agarose gel electrophoresis of purified plasmids and quality control PCR. As an additional quality control measure, a representative portion (20% of the ORFs of all size ranges) of the clones was sequenced. Full-length proteins encoded by each purified plasmid were in vitro expressed according to the manufacturer’s protocol (RTS 100 Escherichia coli HY kit – 5 Prime, MD, USA). Proteins were printed onto eight pad nitrocellulose-coated AVID glass slides (Grace Bios, Oregon, USA) with an OmniGrid 100 microarray printer (Genomic Solutions, Ann Arbor, MI, USA). The success of expression was assessed by the detection of a histidine tag. Expression success was over 95% (data not shown). Each pad contained, in addition to the expressed proteins, the products of the empty vectors, purified human IgG, as well as purified anti-human IgG and empty spots. Thus, more than 200 slides, each with eight pads containing 992 proteins, were printed and stored at room temperature, protected from light and humidity.

To assure that probing and data extraction procedures were optimised, we probed a pilot microarray for human IgG with sera from two control individuals (i.e., not resident in an *S. mansoni* endemic area), two sera from individuals with chronic *S. mansoni* infection, and sera from two individuals in the acute phase of *S. mansoni* infection, in order to detect substantial differences in antibody profiles between individuals chronically infected with either chronic or acute schistosomiasis relative to the negative controls (Fig. 2). Of the 92 proteins printed on the pilot array, over 50 were recognised by sera from individuals with either chronic or acute schistosomiasis. Overall, differences in IgG profiles were most apparent between non-endemic controls and *S. mansoni* chronically infected individuals; we also detected substantial differences in antibody profiles between individuals with chronic and acute schistosomiasis (Fig. 2C). Among the immunoreactive proteins, eight showed differential recognition profiles between negative control sera and sera from individuals with chronic or acute schistosomiasis. Four proteins were also differentially recognised between sera from individuals with chronic and acute schistosomiasis (Fig. 2A and B, Table 1).

A number of proteins that were differentially recognised between individuals chronically infected with *S. mansoni* compared with controls are noteworthy (Table 1). Smp_171190.1 was included in the gene list for the array after a search for signal peptides and/or transmembrane motifs, as a predictor of cellular localisation. This protein belongs to the micro-exon gene (MEG) family, which could suggest a molecular system for creating protein variation through alternative splicing and is implicated in immune evasion by schistosomes, which presents a significant obstacle to vaccine development (DeMarco et al., 2010). Smp_030370.1, a calectulin-like protein, was spotted on the array, due its immunoreactivity with chronically infected *S. mansoni* subjects, using two-dimensional western blotting (Ludolf et al., 2014); this protein was previously suggested as a target for the generation of new vaccines and drugs against schistosomiasis (Sotillo et al., 2015). TE736, a 25 kDa tetraspanin integral membrane protein

Table 1
Comparison of antigens recognised by IgG in sera from individuals with chronic and acute *Schistosoma mansoni* between the expanded 992 antigen *S. mansoni* array and the previous 37 antigen *S. mansoni* array (Gaze et al., 2014).
(Smp_1949701) was recognised by IgG from sera of acutely infected individuals. This protein was also a target of IgG3 and IgE in chronically infected subjects reported by Gaze et al. (2014). Other tetraspanins have shown promise as vaccine antigens for *S. mansoni* infection (Tran et al., 2006) and TET36 needs to be further studied.

Comprehensive proteome arrays that are enriched for surface proteins such as that described herein provide an ideal platform by which to reveal new candidate antigens for schistosomiasis. In summary, we have described the selection, printing and probing of a next generation *S. mansoni* protein array, which is comprised of more than 900 proteins, the largest proteome array produced for a multicellular pathogen. Proteins were selected on the basis of the four methods described above, which allows for a broad range of immunogenic and vaccine candidates. Probing of a pilot version of the array with sera from subjects with acute or chronic schistosomiasis revealed substantial immunoreactivity and paves the way for extensive profiling of schistosomiasis systems immunology. Future work will utilise this protein array to address outstanding questions on the acquisition of human immunity and immunopathology to schistosomiasis, including: (i) the antibody signatures of the acute phase of *S. mansoni* infection compared with chronically infected individuals; (ii) the antibody signatures of individuals with hepatosplenic schistosomiasis compared with chronically infected individuals without hepatosplenomegaly; and (iii) antibody signatures of subjects who have developed immunological resistance to *S. mansoni* after PZQ therapy compared with those who are susceptible to reinfection after treatment.

Acknowledgements

Funds for this project were from National Institutes of Health, USA, grant numbers U01 R01-A1095916-04, U19-A108968-05 and P50AI098507, and from National Council for Scientific and Technological Development, Brazil, (CNPq) 478131/2012-7, Rafael Ramiro de Assis was funded by CNPq (168260/2014-0).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpath.2014.06.001.

References


