CHARACTERISATION OF THE UK ANTHRAX VACCINE AND HUMAN IMMUNOGENICITY

Tapasvi Modi[1], David Gervais[1], Stuart Smith[1], Julie Miller[1], Shaan Subramaniam[2], Konstantinos Thalassinos[2,3], Adrian Shepherd[3]

[1] Porton Biopharma Limited, Porton Down, Salisbury, Wiltshire, SP4 0JG, UK

[2] Institute of Structural and Molecular Biology, Division of Biosciences, Darwin Building Room 101A, University College London, Gower Street, London WC1E 6BT, UK

[3]Department of Biological Sciences and Institute of Structural and Molecular Biology, Birkbeck, University of London, Malet Street, London, WC1E 7HX, UK

ABSTRACT

The manufacture of the UK Anthrax vaccine (AVP) focuses on the production of Protective Antigen (PA) from the *Bacillus anthracis* Sterne strain. Although used for decades, several of AVP's fundamental properties are poorly understood, including its exact composition, the extent to which proteins other than PA may contribute to protection, and whether the degree of protection varies between individuals.

This study involved three innovative investigations. Firstly, the composition of AVP was analysed using liquid chromatography tandem mass-spectrometry (LC-MS/MS), requiring the development of a novel desorption method for releasing *B. anthracis* proteins from the vaccine's aluminium-containing adjuvant. Secondly, computational MHC-binding predictions using NetMHCIIpan were made for the eight most abundant proteins of AVP, for the commonest HLA alleles in multiple ethnic groups, and for multiple *B. anthracis* strains. Thirdly, antibody levels and toxin neutralising antibody (TNA) levels were measured in sera from AVP human vaccinees for both PA and Lethal Factor (LF).

It was demonstrated that AVP is composed of at least 138 *B. anthracis* proteins, including PA (65%), LF (8%) and Edema Factor (EF) (3%), using LC-MS/MS. NetMHCIIpan predicted that peptides from all eight abundant proteins are likely to be presented to T cells, a pre-requisite for protection; however, the number of such peptides varied considerably between different HLA alleles.

These analyses highlight two important properties of the AVP vaccine that have not been established previously. Firstly, the effectiveness of AVP within humans may not depend on PA alone; there is compelling evidence to suggest that LF has a protective role, with computational predictions suggesting that additional proteins may be important for individuals with specific HLA allele combinations. Secondly, in spite of differences in the sequences of key antigenic proteins from different *B. anthracis* strains, these are unlikely to affect the cross-strain protection afforded by AVP.

Keywords: Bacillus anthracis, Anthrax, Anthrax vaccine precipitated, Desorption, Proteomics, MHC-binding prediction

ABBREVIATIONS

AD – Alcohol Dehydrogenase, AVA – Anthrax Vaccine Adsorbed, AVP – Anthrax Vaccine Precipitated, Chap. – 60 kDa Chaperonin, CF – Culture Filtrate, CPC - Cetylpyridinium Chloride, CV – Coefficient of variation, DIA – Data Independent Acquisition, DTT – Dithiothreitol, EDTA - Ethylenediaminetetraacetic Acid, EF – Edema Factor, ET – Edema toxin, Eno – Enolase, FDR – False Discovery Rate, GFP - [Glu1]-fibrinopeptide B, HLA – Human Leukocyte Antigen, IgG – Immunoglobulin G, IAA – Idoacetamide, LC-MS/MS – Tandem Mass Spectrometry, LF – Lethal Factor, LT – Lethal Toxin, MHC – Major Histocompatibility Complex, MIG – Medical Interventions Group, MTT - methylthiazolyldiphenyl-tetrazolium bromide, NaCl – Sodium Chloride, NaOH – Sodium Hydroxide, NIBSC – National Institute of Biological and Control Standards, PA - Protective Antigen, PBL – Porton Biopharma Ltd., PBS – Phosphate Buffered Saline, PHE – Public Health England, PGK – Phosphoglycerate Kinase, PX01 – PX01-90, rPA – recombinant Protective Antigen, rLF – recombinant Lethal Factor, RT – room temperature, ToF – Time of Flight, TNA – Toxin Neutralisation Assay

1. Introduction

Bacillus anthracis is a highly virulent bacterium that is responsible for causing anthrax. Anthrax spores survive in the environment for a long time, are easily transmitted, and are associated with high rates of morbidity and mortality. For these reasons, anthrax has gained increasing attention as a potential bioterrorism agent. As a consequence, government agencies are interested in stockpiling anthrax vaccines that exhibit long-term stability and efficacy as a means to safeguard public health through mass immunisation, should the need arise.

There are two widely-used vaccines against anthrax: the US Anthrax Vaccine Adsorbed (AVA) vaccine, and the UK Anthrax Vaccine Precipitate (AVP) vaccine. AVP, which has been in production since the 1950's and is manufactured by Porton Biopharma Ltd (PBL), is the focus of this research. AVP is an alum precipitate of a sterile culture filtrate of the *B. anthracis* Sterne ($34F_2$) strain. Previous proteomic studies [1, 2] have shown that AVP contains at least 21 proteins including Protective Antigen (PA), Lethal Factor (LF) and Edema Factor (EF). However, the exact composition of AVP remains unknown, although – perhaps significantly – it is thought to contain more LF than AVA, based on antibody titres measured in sera from both animal and human studies [3-5].

Numerous studies have confirmed that PA is the principal immunogen of both AVP and AVA, with anti-PA antibody and Toxin Neutralising Antibody (TNA) levels generally accepted as correlates of protection when measuring vaccine efficacy [3, 6, 7]. However, several studies have also highlighted the additional protective role of LF, either because it enhances the PA-specific antibody response [3, 8, 9], or via the independent protective role of anti-LF antibodies [3, 10, 11]. Additionally, EF has been shown to protect against *B. anthracis* spore challenge in animal studies [12, 13], and it is known that anti-EF antibodies can neutralise Edema Toxin (ET) [14]. Other *B. anthracis* proteins such as cell wall proteins have also been shown to

trigger a protective immune response against anthrax in mice [15]. However, it is currently unknown whether AVP proteins other than PA have a significant protective role in humans.

In anthrax research, there has been a heavy reliance on animal studies, owing to the life-threatening nature of *B. anthracis* and the low rates of human infection. Large-scale studies of human AVA vaccinees are possible because of previous mandatory US military vaccination programmes, whereas comparable studies for AVP are infeasible, given the comparatively smaller number of AVP vaccinees. One AVA study involving 1000 vaccinated individuals concluded that African Americans have lower toxin neutralising antibodies than European Americans [16], raising the possibility that genetic differences play a role in the immune response to AVA and calling into question the relevance of non-human studies. Given their known associations with ethnicity and with differential responses to vaccination, HLA haplotypes are prime candidates as potential genetic factors underpinning the stratification of human responses to anthrax vaccines.

A multi-stranded investigation of AVP is presented here. Firstly, the composition of AVP was determined using a label-free quantitative proteomic liquid chromatography tandem mass-spectrometry (LC-MS/MS) approach. This required the development of a novel desorption method for releasing *B. anthracis* proteins from the vaccine's aluminium-containing adjuvant. Adjuvants are used to enhance the immune response and increase vaccine stability [17-21], with potassium aluminium sulphate (alum) used as an adjuvant in AVP. During the final steps of the AVP manufacturing process, the proteins in sterile Culture Filtrate (CF) are precipitated under gravity by adding aluminium potassium sulphate solution. The supernatant is discarded such that the bulk vaccine precipitate is 15x concentrated. The bulk vaccine concentrate is diluted with sterile saline to achieve 5x concentrated final product [22].

Secondly, a computational investigation of AVP immunogenicity was carried out. Computational MHC-binding predictions using NetMHCIIpan were made for the eight most

abundant proteins from AVP (as determined using LC-MS/MS studies), for the commonest HLA alleles in multiple ethnic groups, and for multiple *B. anthracis* strains. Thirdly, a small proof-of-concept *in vitro* study was designed to characterise the immune response from PA and LF in human AVP vaccinees. Antibody levels and toxin neutralising antibody (TNA) levels were measured in sera from AVP human vaccinees for both PA and LF.

2. Materials and methods

2.1 Desorption methods

Proteins bound to alum are not suitable for liquid chromatography tandem mass-spectrometry (LC-MS/MS) proteomic analysis, as alum would interfere and suppress the signal observed during the MS analysis. Hence, a desorption method was developed to release or desorb proteins in AVP from alum. AVP was obtained from PBL, nine desorption methods using salts and surfactants were assessed for recovery of proteins using the Micro BCA assay (Thermo Fisher, UK) and a size-based separation using 1D gel electrophoresis. Sodium hydroxide (NaOH), sodium citrate, succinic acid, sodium phosphate dibasic, guanidine hydrochloride, urea, ammonium sulphate, cetylpyridinium chloride (CPC) and ethylenediaminetetraacetic acid (EDTA) were purchased from Sigma Aldrich, UK. RapiGest™ SF surfactant was purchased from Waters, UK and ProteaseMAX™ surfactant was purchased from Promega, UK. Regenerated cellulose centrifugal concentrators were purchased from Merck Millipore, UK. Protein estimation Micro BCA kit was purchased from ThermoFisher, UK.

An identical volume of AVP (600 µL) was used to investigating the following nine desorption methods:

1) 5µL of 10M sodium hydroxide was added. The solution was vortexed for 30 seconds until it turned clear. In order to neutralise the solution, 10µL of 3M sodium citrate was added immediately; 2) 500µL of 250mM succinic acid, pH 3.5 was added and incubated for 30 min

at room temperature (RT) shaking [20]; 3) 500µL of 0.66M sodium phosphate dibasic, 3mM EDTA, pH 7.0 was added and incubated for 3 hours at 37°C, shaking [23]; 4) 1mg of RapiGest™ SF surfactant was dissolved in AVP and incubated at 37°C for 24 hours; 5) 1mg of ProteaseMAX™ surfactant was dissolved in AVP and incubated at 37°C for 24 hours; 6) 500µL of 4M guanidine hydrochloride, pH 7.0 was added and incubated for 24 hours at RT, shaking; 7) 500µL of 8M urea, pH 7.0 was added and incubated for 24 hours at RT, shaking; 8) 500µL of 1M ammonium sulphate, 27mM CPC, pH 7.0 was added and incubated for 24 hours at RT, shaking [21]; 9) 500µL of 0.66M sodium phosphate dibasic, 3mM EDTA, pH 7.0 and 1mg of RapiGest™ SF surfactant was added and incubated for 3 hours at RT, shaking [23].

After the desorption process was completed, each solution was centrifuged at 14,000xg for 2 min. The supernatants were transferred into regenerated cellulose centrifugal filter units and concentrated to ~100µL by centrifuging at 14,000xg. The solutions were buffer exchanged into phosphate buffer saline (PBS - 100 mM Sodium Phosphate and 150 mM NaCl, pH 7.2) five times by adding 400µL of PBS each time. The final volume of each sample was ~100µL. The protein concentration of each sample was determined by Micro BCA kit using the manufacturer's recommended procedure. Desorbed proteins were analysed by 1D gel electrophoresis. The samples for 1D gel electrophoresis were treated as follows: 10µL of 4x LDS was added to 30µL of sample and incubated at 70°C for 10 min. 20µL of each sample was loaded on 4-12% bis-tris gel and run for 45 minutes at 200V using MES running buffer. The gel was stained using Coomassie blue stain.

Recovery of desorbed proteins was calculated by subtracting the discarded supernatant concentration from culture filtrate giving a theoretical estimate of the concentration of AVP precipitate. The percentage of recovery of desorbed proteins was calculated by measuring the difference between the estimated and quantified protein concentrations.

2.2 Proteomics

2.2.1 Sample preparation

The composition of Culture Filtrate (CF) and desorbed AVP was determined using a label-free quantitative proteomic LC-MS/MS approach. CF and AVP samples were obtained from PBL. Figure 1A shows the workflow of sample preparation and LC-MS/MS analysis for CF and AVP. In summary, AVP was desorbed using 0.66M Sodium Phosphate, 1 mM EDTA, pH 7.0 buffer for 3 hours at 37°C. The samples were incubated with 8M urea in 50 mM Ammonium Bicarbonate (Sigma Aldrich, UK), at 40°C for 10 minutes shaking. Subsequently, the samples were treated with 10mM DTT (Sigma Aldrich, UK) at 56°C for 30 minutes, shaking; and 20mM Iodoacetamide (IAA) (Sigma Aldrich, UK) in the dark at RT for 30 minutes. The samples were diluted in 50mM Ammonium Bicarbonate solution such that final urea concentration reduced from 8M to 1M. The solubilised samples were digested by incubating with sequencing-grade modified trypsin (Promega, UK) at 1:50 trypsin: protein concentration at 37°C for 16 hours, with shaking. The samples were dried in SpeedVac system at 35°C. The samples were then desalted using Empore SPE Disks C_{18}, diam. 47 mm (Sigma, UK), then dried in SpeedVac system at 35°C. The dried samples were resuspended in LC-MS running buffer (3% acetonitrile, 0.1% formic acid in LC-MS grade water; all from Thermo Fisher, UK). An internal standard of trypsin digested BSA (125 fmoles) (ThermoFisher, UK) was spiked into the samples.

2.2.2 Liquid chromatography–tandem mass spectrometry

Separation of peptides was performed using a Waters NanoAcquity Ultra-Performance Liquid Chromatography system and data acquired in MS^E based Data Independent Acquisition (DIA) mode [24]. In summary, the samples were desalted using a reverse-phase SYMMETRY C_{18} trap column (180µm internal diameter, 20mm length, 5µm particle size, Waters, UK) at a flow

rate of 8µL/min for 2 minutes. Peptides were separated by a linear gradient (0.3µL/min, 35°C column temperature; 97-60% Buffer A over 60 minutes) using a custom-made Acquity UPLC M-Class Peptide BEH C_{18} column (130Å pore size, 75µm internal diameter, 400mm length, 1.7µm particle size, Waters, UK). [Glu1]-fibrinopeptide B (GFP, Waters, UK) was used as lockmass at 100fmol/µL. Lockmass solution was delivered from an auxiliary pump operating at 0.5µL/min to a reference sprayer sampled every 60 seconds.

The nanoLC was coupled online through a nanoflow sprayer to a Q-ToF hybrid mass spectrometer (HDMS Synapt G2-Si; Waters, UK). The instrument was operated in positive ion mode and tuned to a mass resolution of ~20,000 (full width at half maximum). The ToF analyser was externally calibrated with fragment ions of [Glu1]-fibrinopeptide B (GFP, Waters, UK) for m/z range of 175.11 to 1285.54.

Data were lockmass-corrected with the monoisotopic mass of the doubly-charged precursor of GFP (785.8426 *m/z*), post-acquisition. Accurate mass measurements were made using a data-independent mode of acquisition [25]. Briefly, energy in the collision cell was alternated between low energy (4 eV) and high energy (energy ramp from 16-38 eV) modes every 0.6 seconds to acquire precursor and fragment ion spectra for retention time alignment and peptide sequencing during database processing. Measurements were made over a *m/z* range of 50-2000 Da. Each sample was analysed in technical triplicates.

2.2.3 Database processing

Raw data were searched using PLGS v3.0.2 (Waters, UK). The raw data was lockmass-corrected, smoothed, background subtracted and deisotoped. The peptide and fragment ion retention times were aligned [26]. Data were searched against Uniprot complete protein database for *Bacillus anthracis* Sterne $34F_2$ strain. Carbamidomethyl-C and oxidation were specified as fixed and variable modifications, respectively. A maximum of two missed

cleavages of the protease were allowed for semi-tryptic peptide identification. For peptide identification, three corresponding fragment ions were set as a minimum criterion whereas for protein identification a minimum of two corresponding peptide identification and seven fragment ions were required. Protein level FDR rate was maintained at 1% estimated based upon the number of proteins identified from a decoy database. The proteins were quantified using the Hi3 quantification method [27].

2.3 Computational studies

2.3.1 MHC class II epitope prediction

MHC class II predictions were carried out to identify 15-mer epitopes in eight proteins in AVP, using NetMHCIIpan 3.2 [28]. The alleles for MHC II predictions were selected based on the data published by Wang et al., [29]; the 25 alleles result in >99% human population coverage. Python scripts were written to enable automation (supplemental information). Binding affinities of peptides with IC_{50} ≤50 nM cut off and ≤500 nM cut off were used to select strong and medium binding epitopes, respectively [30].

2.3.2 B. anthracis strain data

The PA, LF and EF protein sequences from known 33 *B. anthracis* strains were analysed for identifying mutations using MegAlign software [31]. The list of all *B. anthracis* strains evaluated in the study are detailed in supplemental information, the genomic sequence data was obtained from NCBI database [32]. Substitutions were then analysed in the context of the MHC class II to see whether they changed the immunogenic properties of these proteins.

2.4 In vitro studies

2.4.1 Blood collection

Eight AVP vaccinated volunteers and two non-vaccinated control volunteers were recruited for the study. A total of 29 ml of blood was collected from each volunteer. Although the volunteers were not recruited based on their vaccination dates, details of their AVP vaccination history were taken.

2.4.2 Ethics statement

Human AVP vaccinees and healthy control volunteers based at PBL, Porton Down, participated in the context of a study protocol (Ref: R&D 325) approved by the PHE Independent ethics committee, UK; the subjects were all adults (both male and female) aged over 18 years and all provided written, informed consent.

2.4.3 HLA tissue typing

HLA Tissue typing analysis was contracted to Proimmune Ltd, UK. Blood (4 ml) was collected in K_2EDTA tubes (Midmeds, UK) and stored at -80°C. The MHC II alleles for each donor for the 6x loci (2 x DRB1, 2 x DQB1 and 2 x DPB1) were reported.

2.4.4 Anti-PA and anti-LF IgG ELISA assay

Blood (5 ml) that was collected in clot activator coated tubes (Midmeds, UK) was centrifuged at 2,000 x g for 10 minutes. The supernatant was recovered and stored at -80°C for Anti-PA and Anti-LF IgG ELISA and TNA assay.

These tests were performed by Medical Interventions Group (MIG) at Public Health England (PHE), Porton Down. Briefly, 96-well plates (NUNC flat bottomed wells, ThermoFisher, UK) were coated overnight with 0.5µg/mL of either purified rPA (*E.coli* derived, PHE, Porton Down) or rLF (*B. anthracis* derived, PHE, Porton Down), before addition of serial diluted human serum samples and reference (PHE, Porton Down). The reference serum was prepared by conversion of plasma collected from AVP vaccinated individuals (not originating from this study). Anti-human IgG Fcγ specific antibody conjugated to Alkaline Phosphatase (Jackson Immunoresearch, UK) was used to produce a colorimetric response proportional to the amount of PA or LF specific antibody, when substrate (AP Yellow, BioFX & surmodics, UK) was added. Plates were read using a Versamax plate reader with SoftMax Pro 5.2 analysis software (Molecular Devices, UK). Each sample was assigned a titre against a five-parameter logistic human sera reference curve. The reference sera was assigned heuristic values of 960U/ml (PA) and 500U/ml (LF) based on the mean ED_{50} value (median effective dose) of multiple runs on previous occasions.

2.4.5 Toxin neutralisation assay

Sera were serially diluted and incubated with Lethal Toxin (LT - formed by the association of PA and LF) (PHE, Porton Down) at a controlled concentration. This was transferred to 96-well plates seeded with a mouse macrophage cell line (J774A.1) known to be sensitive to anthrax toxin-mediated cytotoxicity. Cell survival was assessed through uptake of methylthiazolyldiphenyl-tetrazolium bromide (MTT, Sigma, UK) by surviving cells; this provided a colorimetric readout of survival. Plates were read using a Versamax plate reader with SoftMax Pro 5.2 analysis software (Molecular devices, US). Sample ED_{50} values (the dilution of serum required for a 50% reduction in cytotoxicity) were compared to a reference serum (PHE, Porton Down), which allows an NF_{50} (50% Neutralisation Factor) value to be calculated for each sample. The reference serum was prepared by conversion of plasma collected from AVP vaccinated individuals (not originating from this study).

3. Results

3.1 Desorption methods

Several different salts and surfactants were investigated for the desorption of proteins from alum in AVP. Figure 1B shows the size-based separation of desorbed AVP proteins on 1D gel electrophoresis, using nine different desorption methods. The desorption methods with NaOH (Lane 3, Figure 1B), EDTA (Lane 5, Figure 1B), ammonium sulphate and CPC (Lane 11, Figure 1B) and combination of RapiGest™ SF surfactant and EDTA (Lane 12, Figure 1B) gave good recovery of desorbed proteins; the measured protein concentration of recovered proteins was 27.1, 18.6, 11.6 and 18.1 µg/mL using Micro BCA assay, respectively. The desorption methods with succinic acid (Lane 4, Figure 1B), RapiGest™ SF surfactant (Lane 6, Figure 1B), ProteaseMAX™ surfactant (Lane 7, Figure 1B), guanidine hydrochloride (Lane 8, Figure 1B) and urea (Lane 10, Figure 1B) did not desorb AVP proteins from alum. This corroborated with the Micro BCA assay results.

Although the maximum desorption of proteins from alum in AVP was using the NaOH method, the higher molecular weight bands on the 1D gel (Lane 3, Figure 1B) were faint in comparison to other methods. This confirms that the harsh conditions associated with NaOH degrades proteins. Hence, this method was not taken forward. The EDTA desorption reagent (Lane 5, Figure 1B) gave good recovery of proteins. EDTA is probably responsible for chelating aluminium ions, thus enabling desorption of proteins from alum. The ammonium sulphate and CPC method produced a pellet in the sample after the desorption process. Although, the recovery of proteins was good from this method (11.6 µg/mL using Micro BCA method) and the protein bands are comparably dark on the gel (Lane 11, Figure 1B); the method was thought to create repeatability issues due to pellet formation, hence this method was not taken forward. The combination of RapiGest™ SF surfactant and EDTA (Lane 12, Figure 1B) did

not enhance the recovery of proteins in comparison to EDTA alone (Lane 5, Figure 1B), hence this method was also not taken forward.

Based on size-based separation of desorbed proteins (Figure 1B) and Micro BCA assay results, 0.66M Sodium Phosphate Dibasic, 3mM EDTA, pH 7.0 was found to be the optimum desorption reagent for proteins in AVP. Nevertheless, further studies (comparing protein concentration of CF, discarded supernatant and desorbed AVP) revealed that the recovery of desorbed proteins from AVP using this method was between 40-60%.

3.2 Proteomics studies

3.2.1 Identification of proteins in CF and AVP

Based on the analysis of LC-MS/MS data using the *B. anthracis* proteome, a total of 163 and 261 proteins were identified in CF and AVP, respectively (Figure 2A). The complete list of proteins identified in CF and AVP is given in the supplemental information. 138 proteins were found to be common to both CF and AVP. Two biological replicates were prepared from two batches of CF and AVP, with each sample analysed by LC-MS/MS in triplicate. A maximum precision of 11% and 19% coefficient of variation (CV) was measured for the identification of proteins in the two batches of CF and AVP respectively, including the biological replicates. More proteins were identified in AVP in comparison to CF, because AVP is five times more concentrated than CF, due to the alum precipitation step during AVP manufacture. Hence, low abundance proteins were identified in AVP that were not identified in CF. AVP samples had more variability in replicates, potentially due to the variability in the desorption process.

3.2.2 Relative quantitation of PA, LF and EF in CF and AVP

PA was found to be the most abundant protein in CF and AVP, followed by LF and EF (Figure 2B). For CF and AVP, PA accounted for 65% and 64% of total protein respectively, LF accounted for 6% and 8% respectively, and EF accounted for 3% of total protein in both. Repeatability for relative quantitation of PA, LF and EF proteins in two batches of CF and AVP was good, a maximum CV of 18% was measured, including two biological replicates and triplicate LC-MS/MS analysis (Figure 2C).

3.2.3 Absolute quantitation of proteins in CF and AVP

PA was found to be the most abundant protein in CF and AVP, followed by LF, enolase, PX01-90, EF, 60kD chaperonin, alcohol dehydrogenase and phosphoglycerate kinase. PA was measured to be 615 and 2831 ng/mL in CF and AVP; LF was measured to be 52 and 345 ng/mL in CF and AVP; EF was measured to be 30 and 119 ng/mL in CF and AVP, respectively (Figure 2D). A maximum 25% CV was measured for the quantification of abundant proteins in two batches of CF, including two biological replicates. The repeatability of AVP biological replicates was poor and a CV of 15% was measured for PA (the most abundant protein in AVP), and approx. 40% for less abundant proteins. AVP samples had more variability in biological replicates, possibly due to the variability in the desorption process. The repeatability of the triplicate analysis of each AVP sample was <20%.

3.3 Computational studies

3.3.1 MHC II epitope prediction

MHC II epitopes were predicted using NetMHCIIpan for the eight most abundant proteins in AVP (identified by MS studies) and for 25 class II HLA alleles in order to assess the impact of

allelic differences in human populations. Given the reasonable assumptions that a) epitope binding affinity is correlated with the strength of the immune response, and b) a larger proportion of peptides predicted to bind with moderate to high affinity (≤500 nM) are likely to be true epitopes than those predicted to bind with low affinity (≤5000 nM), only predicted epitopes with high and intermediate binding affinity were taken forward for analysis. Two IC_{50} binding thresholds advocated by the Immune Epitope Database (IEDB) [30], ≤50 nM and ≤500 nM, were adopted for predicted epitopes with high binding affinities (strong binders) and intermediate binding affinities (medium binders), respectively. The number of strong and medium-plus-strong binders for eight AVP proteins and 25 HLA alleles are shown in Tables 1 and 2 respectively.

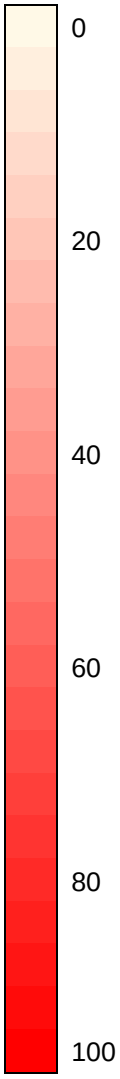
With respect to the eight AVP proteins and 25 HLA alleles investigated, LF and EF had the highest number of predicted class II strong binding epitopes overall, including the highest numbers for 8 and 6 of the 25 alleles respectively (Table 1). PA, LF and EF proteins were predicted to have >5 strong binding epitopes for between 7 and 10 of the HLA alleles. The other five abundant proteins were predicted to have >5 strong binding epitopes for between 3 and 7 of the HLA alleles. This data suggests that different proteins in AVP may be contributing to a protective T cell response in different HLA alleles. For example, in individuals having HLA-DQA1*0501-DQB1*0301 allele, only four proteins (60 kDa chaperonin, alcohol dehydrogenase, enolase and phosphoglycerate kinase) have >5 strong binding epitopes (Table 1).

Additionally, PA, LF and EF were predicted to have >5 medium-plus-strong epitopes for all HLA alleles investigated, except alleles HLA-DQA1*0301-DQB1*0302 and HLA-DQA1*0401-DQB1*0402. It is notable that the 60 kDa chaperonin protein had a much higher number of predicted epitopes for both of these alleles (15 and 29 epitopes respectively) (Table 2). Once again, LF and EF had the highest number of predicted class II strong-plus-medium binding

epitopes overall, including the highest numbers for 13 and 6 of the 25 alleles respectively (Table 1).

Table 1 - Predicted number of Strong binding MHC II epitopes, derived using NetMHCIIpan (IC_{50} cut-off of ≤50 nM)

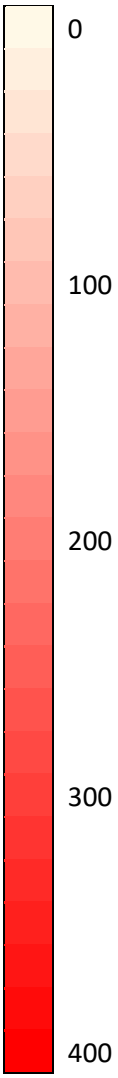
HLA II Alleles	PA	LF	EF	PX01	Chap.	AD	Eno	PGK
HLA-DPA1*0103-DPB1*0201	3	6	12	9	0	0	0	0
HLA-DPA1*0103-DPB1*0401	0	5	4	9	0	0	0	0
HLA-DPA1*0201-DPB1*0501	0	0	0	0	0	0	0	0
HLA-DPA1*0301-DPB1*0402	0	5	0	0	0	0	0	0
HLA-DQA1*0101-DQB1*0501	0	0	0	0	0	0	0	0
HLA-DQA1*0102-DQB1*0602	0	0	0	0	3	0	5	0
HLA-DPA1*0201-DPB1*0101	0	5	0	0	0	0	0	0
HLA-DQA1*0301-DQB1*0302	0	0	0	0	0	0	0	0
HLA-DQA1*0401-DQB1*0402	0	0	0	0	0	0	0	0
HLA-DQA1*0501-DQB1*0201	0	0	0	0	0	0	0	0
HLA-DQA1*0501-DQB1*0301	5	0	0	0	12	11	7	19
HLA-DRB1*0101	44	100	101	27	42	20	43	49
HLA-DRB1*0301	7	21	0	0	4	2	0	0
HLA-DRB1*0401	1	3	10	0	0	0	3	0
HLA-DRB1*0404	5	0	5	4	0	1	0	5
HLA-DRB1*0405	0	0	12	0	0	0	0	0
HLA-DRB1*0701	15	33	28	7	0	2	4	12
HLA-DRB1*0802	0	0	0	0	0	0	0	0
HLA-DRB1*0901	0	18	14	4	3	1	4	11
HLA-DRB1*1101	6	5	20	2	0	0	6	4
HLA-DRB1*1302	32	36	30	9	10	10	8	11
HLA-DRB1*1501	5	30	7	7	0	0	0	0
HLA-DRB3*0101	4	20	0	0	0	2	0	0
HLA-DRB4*0101	6	2	0	0	0	0	0	0
HLA-DRB5*0101	7	25	28	11	0	0	14	14
Total No. of Epitopes	140	314	271	89	74	49	94	125



MHC class II epitope binding predictions were carried out to identify 15-mer epitopes in eight most abundant proteins in AVP, using NetMHCIIpan 3.2 [28], across 25 HLA alleles covering >99% human population [29]. To select strong binding epitopes, a binding affinity cut-off of IC_{50} ≤50 nM was applied [30]. The numbers on the heatmap scale refer to the absolute number of strong binding epitopes.

Table 2 - Predicted number of Medium-plus-Strong binding MHC II epitopes, derived using NetMHCIIpan (IC_{50} cut-off of ≤500 nM)

HLA II Alleles	PA	LF	EF	PX01	Chap.	AD	Eno	PGK
HLA-DPA1*0103-DPB1*0201	55	118	146	67	13	2	34	60
HLA-DPA1*0103-DPB1*0401	41	88	68	52	5	0	23	25
HLA-DPA1*0201-DPB1*0501	14	32	50	26	0	0	4	3
HLA-DPA1*0301-DPB1*0402	33	61	48	44	3	0	17	14
HLA-DQA1*0101-DQB1*0501	22	28	40	27	0	0	11	13
HLA-DQA1*0102-DQB1*0602	58	40	34	34	137	63	63	62
HLA-DPA1*0201-DPB1*0101	50	121	130	66	12	0	35	57
HLA-DQA1*0301-DQB1*0302	5	2	0	0	15	0	0	4
HLA-DQA1*0401-DQB1*0402	5	4	3	0	29	0	7	4
HLA-DQA1*0501-DQB1*0201	29	56	54	20	75	23	65	35
HLA-DQA1*0501-DQB1*0301	89	48	58	39	197	114	121	111
HLA-DRB1*0101	346	398	358	263	299	165	188	225
HLA-DRB1*0301	107	124	53	23	45	49	30	41
HLA-DRB1*0401	139	211	172	86	64	33	67	71
HLA-DRB1*0404	171	223	189	121	124	57	89	105
HLA-DRB1*0405	112	207	172	89	39	24	51	69
HLA-DRB1*0701	195	235	205	105	113	90	100	106
HLA-DRB1*0802	75	111	112	34	75	28	56	43
HLA-DRB1*0901	169	214	195	74	134	97	121	115
HLA-DRB1*1101	128	202	221	114	89	33	79	89
HLA-DRB1*1302	244	277	224	147	122	84	92	102
HLA-DRB1*1501	150	214	190	112	52	41	73	96
HLA-DRB3*0101	82	130	86	31	31	32	27	34
HLA-DRB4*0101	156	236	195	158	123	45	72	100
HLA-DRB5*0101	176	216	258	134	115	55	86	108
Total No. of Epitopes	**2651**	**3596**	**3261**	**1866**	**1911**	**1035**	**1511**	**1692**



MHC class II epitope binding predictions were carried out to identify 15-mer epitopes in eight most abundant proteins in AVP, using NetMHCIIpan 3.2 [28], across 25 HLA alleles covering >99% human population [29]. To select strong-plus-medium binding epitopes, a binding affinity cut-off of IC_{50} ≤500 nM was applied [30]. The numbers on the heatmap scale refer to the absolute number of strong-plus-medium binding epitopes.

3.3.2 Efficacy of AVP against different B. anthracis strains

Multiple sequence alignments of PA, LF and EF protein sequences from 33 known *B. anthracis* strains were generated. These alignments revealed 3, 4 and 6 single amino acid differences for PA, LF and EF respectively between the vaccine (Sterne) strain and the other strains. NetMHCIIpan was used to make medium-plus-strong 15-mers epitope predictions spanning each of these substitutions. These predictions indicate that specific epitopes in the Sterne strain may be absent (i.e. the corresponding peptides occur as non-binders) or changed (i.e. the corresponding peptides contain a modified set of TCR-facing amino-acid residues) in other *B. anthracis* strains; and other *B. anthracis* strains may contain epitopes not present in the Sterne strain (i.e. with the corresponding Sterne strain peptides occurring as non-binders). The results are summarised in Table 3.

Table 3 - Predicted medium-plus-strong binding MHC II epitope differences between the Sterne (vaccine) strain and other *B. anthracis* strains, derived using NetMHCIIpan (IC_{50} cut-off of ≤500 nM)

Amino acid change	*B. anthracis* strains	Number of missing epitopes[1]	Number of additional epitopes[2]	Number of changed epitopes[3]
PA I433V	HYU01	9	0	5
PA P565S	CDC 684, SK-102, Vollum 1B, Vollum	10	12	9
PA A600V	BA1015, Canadian Bison, CDC 684, isolate IT Carb1-6241, isolate IT Carb3-6254, PAK-1, RA3, SK-102, Turkey32, V770-NP-1R, Vollum 1B, Vollum, Pollino, $P.NO_2$, Larissa, HYU01, H9401, A1144	18	5	11
LFK155X	$P.NO_2$	4	9	13
LF S299A	1C3, 4NS, A16, A16R, A0248, A1144, A2012, Ames 0462, Ames BA1004, BA1015, Canadian Bison, CDC 684, H9401, Larissa, Ohio, $P.NO_2$, Pak-1, Pollino, Shikan, SK-102, Stendal, Turkey 32, V770-NP-1R, VCM1168, Vollum 1B, Vollum	0	4	16
LF S299T	BA1035, HYU01, RA3, SVA11	3	0	12
LF Q346E	H9401	0	5	1
LF E709G	BA1035, HYU01, $P.NO_2$, RA3, SVA11	0	0	0
EF D84G	A16R	0	40	1
EF D180G	BA1035, HYU01, RA3, SVA11	22	16	19
EF I318T	BA1035, HYU01, RA3, SVA11	10	0	0
EF G352V	A16R	0	0	0
EF E443D	Canadian Bison	0	0	0
EF E467G	Canadian Bison	5	18	1

Differences in predicted medium-plus-strong binding MHC class II epitopes due to single amino acid differences in PA, LF and EF proteins from 33 known *B. anthracis* strains compared with the AVP vaccine (Sterne) strain are detailed above. MHC class II epitope binding predictions were carried out to identify 15-mer epitopes, using NetMHCIIpan 3.2 [28], across 25 HLA alleles covering >99% human population [29]. To select medium-plus-strong binding epitopes, a binding affinity cut-off of IC_{50} ≤500 nM was applied [30].

[1]The number of HLA-II alleles for which a Sterne strain epitope in Table 2 is predicted to be a non-binder in a non-Sterne strains

[2]The number of HLA-II alleles for which an epitope is predicted with a non-Sterne strains that is not predicted to be an epitope with the Sterne strain

[3]The number of HLA-II alleles for which a Sterne strain epitope is predicted to present a different epitope in non-Sterne strains (i.e. with a different TCR-facing amino-acid residue).

3.4 In vitro studies

The HLA types and immunisation history of patients are shown in Table 4, and anti-PA antibody, anti-LF antibody and TNA levels are shown in Table 5.

Table 4 - Patient HLA types immunisation history

Sample No.	Immunisation History (Primary Immunisation Year, Boosters Years)	HLA-DRB1		HLA-DQB1		HLA-DPB1	
1	2013, 2015, 2016	*01:01:01	*15:01:01	*06:02:01	*05:01	*04:02:01	*04:01:01
2	2012, 2014, 2016	*04:01/35/63 /145/179	*11:01/11:08 /11:37/11:17 5/13:14	*03:02:01	*03:01:01	*03:01:01	*03:01:01
3	2014, 2016	*15:01:01	*04:01:01	*06:02:01	*03:01:01	*04:01:01	*04:01:01
4	1999, 2016	*04:01:01	*11:02:01	*03:01:01	*03:19:01	*04:01:01	*04:01:01
5[A]	N/A	*15:01:01	*15:01:01	*06:02:01	*06:02:01	*02:01	*04:01:01
6	2015, 2016	*15:01:01	*15:01:01	*06:02:01	*06:02:01	*02:01	*04:01:01
7	2007, 2009, 2011, 2013, 2016	*15:01:01	*11:04:01	*06:02:01	*03:01:01	*04:01:01	*11:01:01
8	2014, 2016	*04:08:01	*07:01/79	*03:03:02	*03:01:01	*04:01:01	*04:01:01
9	2016, N/A	*03:01:01	*04:01:01	*02:01:01	*03:01:01	*01:01:01	*20:01:01
10[A]	N/A	*07:01/79	*07:01/79	*02:02:01	*02:02:01	*17:01:01	*17:01:01

Eight AVP vaccinated volunteers and two non-vaccinated control volunteers (denoted by suffix A) were recruited for the study. Although the volunteers were not recruited based on their vaccination dates, details of their AVP vaccination history were taken. The subjects were all adults (both male and female) aged over 18 years. HLA tissue typing analysis was performed to determine the MHC II alleles for each donor for the 6x loci (2 x DRB1, 2 x DQB1 and 2 x DPB1).

1

2

Table 5 – Average Anti-PA and Anti-LF antibody titres, and TNA levels in blood sera of AVP vaccinees

Sample No.	Anti-PA Antibody Titre (U/mL)	Anti-LF-Antibody Titre (U/mL)	TNA (NF_{50})*1000
1	263	561	66
2	261	139	38
3	784	734	493
4	724	128	166
5[A]	0	0	0
6	533	470	318
7	306	118	91
8	538	149	89
9	706	429	75
10[A]	0	0	0

End-point titers of serum IgG to recombinant Protective Antigen (PA) and recombinant Lethal Factor (LF) and 50% neutralisation factor (NF_{50}) values for serum neutralization of Lethal Toxin (LT) determined in a J774A.1 macrophage-based Lethal Toxin neutralization assay (TNA) in AVP vaccinees. Measurement of antibody levels was performed with at least four replicates; TNA levels were measured in duplicates. [A]Samples from control volunteers.

Anti-PA and anti-LF antibody titres and TNA levels measured in AVP vaccinees are highly variable. As expected, the antibody titre and TNA levels in the control samples were below the detection limit (Table 5). An analysis of the vaccinee data in Tables 4 and 5 shows that there is no clear correlation between vaccinee immunisation history (whether the number and/or timing of vaccinations) and antibody titre (whether anti-PA and/or anti-LF antibody titre), nor between anti-PA and/or anti-LF antibody titre and TNA level. Linear regression analysis showed positive correlation between PA, LF and PA+LF antibody titres and TNA levels in sera of AVP vaccinees - 0.482, 0.548 and 0.639 respectively (Figure 3B).

4. Discussion

4.1 Desorption of proteins from AVP

Nine different salt- and surfactant-based desorption methods were investigated for desorption of proteins from alum in AVP. The desorption method using 0.66M Sodium Phosphate Dibasic,

3mM EDTA, pH 7.0 was found to be the optimum desorption reagent for AVP. However, even this method only recovered between 40-60% of proteins from AVP. The low recovery of proteins from AVP is corroborated by similar findings by other groups investigating desorption of proteins from the closely-related aluminium based adjuvant Alhydrogel® (Brenntag Biosector, Denmark) (aluminium hydroxide). Alhydrogel adjuvants are more widely used; hence several studies have reported the stability profile of Alhydrogel based vaccines. Vassely et al., [20] have reported that due to chemical and physical changes in proteins adsorbed to Alhydrogel, the desorption of proteins from Alhydrogel is difficult. Another study has shown that the strength of the protein bound to Alhydrogel increases with time, and harsh desorption buffer conditions are required to recover proteins [21]. Our study also highlights the difficulty of analysing alum-adsorbed vaccines.

4.2 Proteomic LC-MS/MS studies

Proteomic LC-MS/MS studies demonstrated that AVP is composed of 261 proteins, including PA (65%), LF (8%) and EF (3%). This method had showed good repeatability for relative quantitation of PA, LF and EF in complex AVP samples. A maximum CV of 18% was measured in AVP, including two biological replicates and triplicate LC-MS/MS analysis. Several other Hi3 quantitation studies have reported similar CVs for complex samples [27, 33, 34]. Further, out of the 21 proteins previously identified in AVP using 2D DIGE and MS studies by the NIBSC group [1], 17 proteins were common to proteins identified in this study.

In this study, PA was measured to be 615 and 2831 ng/mL in CF and AVP; LF was measured to be 52 and 345 ng/mL in CF and AVP, respectively. Previously, the average concentration of PA and LF was reported to be 3710 and 990 ng/ml, respectively, in culture supernatant, determined by ELISA [22]. The quantitation data for PA and LF using ELISA is significantly different from LC-MS/MS data in the present study. Although Hi3 label-free quantitation is deemed reliable for absolute quantitation of proteins [27, 33], sample preparation steps

involving desorption, sample concentration using centrifugal filters and solid phase extraction steps using C_{18} disks resulted in loss of proteins. Calculations showed that only 20% of the sample was recovered from LC-MS/MS experiment, based on the total protein estimation by Micro BCA assay. Sample manipulations in standard polypropylene eppendorfs also resulted in loss of protein, hence Eppendorf LoBind microcentrifuge tubes were used. These losses explain the significant differences in absolute quantitation of CF using ELISA and LC-MS/MS methods. Ideally, sample preparation steps without desorption, concentration step and solid phase extraction should have been used to minimise the losses; however, this proved to be infeasible, as CF contains interfering substances that needed to be removed in order to achieve a reproducible LC-MS/MS analysis.

4.3 Computational Studies

4.3.1 MHC II epitope prediction

In this study, we undertook a broad analysis of MHC class II binding spanning multiple *B. anthracis* strains and proteins, and multiple HLA alleles. It is important to be cautious when interpreting this data, as MHC-peptide binding affinity is just one of the many factors contributing to the T cell response; whereas MHC-peptide binding can be predicted with reasonable accuracy, in the present context there is no information about other important factors such as T cell precursor frequency [35-37] and the breadth of T cell response [38]. The emerging picture from many independent research studies is a highly complex one; hence within a given individual, specific epitopes may be protective whereas others may have a negative impact, for example by blocking or slowing down the T cell response [39, 40], or by inducing autoimmunity [35, 40, 41]. In principle and presumably in practice, the same epitope may lead to different outcomes in different individuals.

Further, NetMHCIIpan may be prone to either over- and under-prediction in specific cases. It has been observed, for example, that standard computational tools predict a subset of experimentally-verified immunodominant peptides to bind too weakly to form epitopes [42]. However, in the context of this study, computational methods are of sufficient accuracy (with an area under the receiver operating characteristics curve [AUC] commonly greater that 0.8 [28]) to provide insights into the protective potential of distinct combinations of human HLA alleles and *B. anthracis* proteins – a combination that poses an unsolved challenge to experimental approaches.

Bearing these points in mind, there are nevertheless cautious but potentially important conclusions that can be drawn from the MHC II-peptide binding prediction. In interpreting these results, the confidence that a given antigen is likely to be protective with respect to a given HLA allele depends on the number of predicted epitopes, and in particular the number of epitopes predicted to bind with at least moderate strength. If the number of medium-plus-strong epitopes is low, there is a greater possibility that an individual will lack TCRs capable of binding to any of the peptide-MHC complexes associated with that combination of antigen and HLA allele.

NetMHCIIpan predicted that peptides from all eight proteins are likely to be presented to T cells, however the number of such peptides varied considerably between different proteins and different HLA alleles (Tables 1 and 2).

The “core” vaccine components PA, LF and EF were associated with >5 medium-plus-strong class II epitopes for all but two alleles, with LF having the largest number of epitopes overall (Table 1). The two notable exceptions were common HLA-DQ alleles HLA-DQA1*0301-DQB1*0302 and HLA-DQA1*0401-DQB1*0402; in both these cases, a much higher number of medium-plus-strong epitopes was associated with the 60 kDa Chaperonin protein (15 and 29 epitopes respectively).

Given that individuals have multiple class II HLA alleles, the data in Table 2 suggests that most individuals vaccinated with AVP have the potential to undergo a protective T cell response, although the proteins involved may vary between individuals. At the same, it appears that the efficacy of PA alone is not guaranteed for all individuals and that the presence of additional proteins may enhance the prospects that AVP affords broad protection. These results indicate the potential value of large-scale computational studies: even though individual predictions require validation; such studies are faster, cheaper and potentially more relevant than traditional animal studies.

4.3.2 Efficacy of AVP against different B. anthracis strains

Amino-acid differences in PA, LF and EF proteins between 33 *B. anthracis* strains and the vaccine (Sterne) strain were identified, and predictions were made using NetMHCIIpan to access their potential impact on MHC class II epitopes. Our analysis shows that a number of medium-plus-strong epitopes are missing, added or changed as a result of these differences (Table 3); however, the number of these predicted epitope transformations is very small in comparison to the total number of predicted Sterne strain epitopes for each of these proteins (Table 2). Consequently, we conclude that the differences between *B. anthracis* strains are unlikely to impact the cross-strain efficacy of the T cell response induced by the vaccine, although there remains a slim possibility that an individual's immunodominant vaccine-induced response may be comparatively ineffective against a different strain if one or more critical epitopes are absent or changed within that strain.

136 *4.4 In vitro studies*

137

138 The anti-PA and anti-LF antibody titres and TNA levels measured in AVP vaccinees were
139 highly variable. Positive correlation was measured between PA, LF and PA+LF antibody titres
140 and TNA levels in sera of AVP vaccinees - 0.482, 0.548 and 0.639, respectively (Figure 3B).

141

142 Although, there is evidence that anti-PA antibodies and TNA levels have good correlation in
143 human studies [7, 43, 44], strong correlation was not observed in this study, potentially due to
144 the small sample size. Previous studies have shown that antibody and TNA levels can be
145 highly variable in AVA and AVP vaccinees [3, 16, 44]. Hence, it is likely that antibody and TNA
146 levels are modulated by a range of other factors, such as age, gender, T cell and B cell
147 memory, and genetic differences (including HLA allelic differences) in humans could be
148 responsible for variable antibody levels [14, 45, 46]. Pajewski et al., [47] had investigated the
149 impact of HLA polymorphisms on anti-PA antibody response in AVA vaccinees and reported
150 that DRB1–DQA1–DQB1 haplotypes *1501–*0102–*0602, *0101–*0101–*0501 and *0102–
151 *0101–*0501 were associated with significantly lower anti-PA antibody levels. However, it was
152 not possible to identify such a correlation in this study.

153

154 Nevertheless, it is apparent that, whereas most vaccinees have higher anti-PA antibody titres
155 than anti-LF titres, none of the vaccinees have low levels of anti-LF antibodies. Moreover, the
156 TNA levels do not suggest that any vaccinees have negligible capacity to neutralise the
157 anthrax toxin, although, in the absence of vaccinated humans becoming infected with *B.*
158 *anthracis*, it is unclear what TNA levels are necessary to afford protection in humans.

159

160 Given the current lack of knowledge about the TNA levels needed to provide protection against
161 *B. anthracis* in humans, anthrax vaccine studies have inevitable limitations, which are
162 compounded here by the small-scale nature of the *in vitro* component of this study. Indeed,
163 this represents the major limitation of this work; it is hoped that future studies will incorporate

many more vaccinees and will explicitly evaluate the potential importance of HLA and protein specificity highlighted by the computational results presented here.

Notwithstanding these limitations, there are several important and novel conclusions. Although direct comparison with AVA is not possible, notably because of the different cell lines and different reference standard used for the TNA assay, it is interesting to contrast the results here with the large AVA study conducted by James and co-workers [16]. In that study, 69% of vaccinees had no detectable anti-LF antibodies, whereas all vaccinees in this study had moderate to high anti-LF titres. Although, the composition of AVA is not published, AVA is thought to contain negligible amounts of LF, based on anti-LF antibody levels measured in several animal and human studies [3-5]. Additionally, over 40% of the AVA vaccinees were deemed to have low TNA levels; although it is hard to calibrate the TNA levels in this study, there is no evidence of very low TNA activity, with only a single individual having a TNA (NF_{50})*1000 below 50. Taken together, these observations suggest that anti-LF antibody response from AVP potentially enhances protection, and this is broadly consistent with previous observations about the efficacy of anti-LF antibodies in neutralising lethal toxin (formed by the association of PA and LF) [10, 11] and about the speed and extent of the anti-LF antibody response in comparison to the anti-PA response in naturally-acquired cutaneous anthrax patients [45].

Finally, this work demonstrates that AVP contains many protein components that have not previously been identified and suggests that several proteins not normally considered relevant – notably LF, EF, PX01-90 and 60KD Chaperonin – are reasonably abundant within AVP. Computational studies have shown that all of these proteins have the potential to afford protection for individuals with HLA allele combinations that are predicted to have relatively few PA epitopes. Further work is needed to validate this data experimentally using *in vitro* MHC molecules and antigen binding studies.

This work also shows the potential importance of considering inter-strain differences and identifies specific epitopes that are modified or absent in a subset of other *B. anthracis* strains. Further work is needed to identify whether these epitopes are important targets of the protective T cell response induced by the vaccine in some recipients.

ACKNOWLEDGEMENTS

We would like to thank Roger Hinton, Managing Director of PBL, for making facilities and funds available for this work. The G2-Si ion mobility mass spectrometer was purchased with a grant from the Wellcome Trust (104913/Z/14/ZBM). We would like to thank Diane Williamson at DSTL, Porton Down, for helpful discussions and guidance for the *in vitro* studies. Further thanks go to MIG Group in Public Health England for performing ELISA and TNA assays and to the entire Development team at PBL.

REFERENCES

[1] Whiting G, Wheeler JX, Rijpkema S. Identification of peptide sequences as a measure of Anthrax vaccine stability during storage. Human vaccines & immunotherapeutics. 2014;10:1669-81. DOI: 10.4161/hv.28443.

[2] Whiting GC, Rijpkema S, Adams T, Corbel MJ. Characterisation of adsorbed anthrax vaccine by two-dimensional gel electrophoresis. Vaccine. 2004;22:4245-51. DOI: 10.1016/j.vaccine.2004.04.036.

[3] Dumas EK, Garman L, Cuthbertson H, Charlton S, Hallis B, Engler RJM, et al. Lethal factor antibodies contribute to lethal toxin neutralization in recipients of anthrax vaccine precipitated. Vaccine. 2017;35:3416-22. DOI: 10.1016/j.vaccine.2017.05.006.

[4] Turnbull PC. Anthrax vaccines: past, present and future. Vaccine. 1991;9:533-9. DOI: 10.1016/0264-410x(91)90237-z.

[5] Turnbull PC, Leppla SH, Broster MG, Quinn CP, Melling J. Antibodies to anthrax toxin in humans and guinea pigs and their relevance to protective immunity. Med Microbiol Immunol 1988;177:293-303.

[6] Chen L, Schiffer JM, Dalton S, Sabourin CL, Niemuth NA, Plikaytis BD, et al. Comprehensive analysis and selection of anthrax vaccine adsorbed immune correlates of protection in rhesus macaques. Clinical and vaccine immunology : CVI. 2014;21:1512-20. DOI: 10.1128/CVI.00469-14.

[7] Reuveny S, White MD, Adar YY, Kafri Y, Altboum Z, Gozes Y, et al. Search for correlates of protective immunity conferred by anthrax vaccine. Infect Immun. 2001;69:2888-93. DOI: 10.1128/IAI.69.5.2888-2893.2001.

[8] Chen Z, Moayeri M, Crown D, Emerson S, Gorshkova I, Schuck P, et al. Novel chimpanzee/human monoclonal antibodies that neutralize anthrax lethal factor: evidence for possible synergy with anti-pa antibody. Infect Immun 2009;77:3902-8.

[9] Pezard C, Weber M, Sirard JC, Berche P, Mock M. Protective immunity induced by Bacillus anthracis toxin-deficient strains. Infect Immun. 1995;63:1369-72.
[10] Albrecht MT, Li H, Williamson ED, LeButt CS, Flick-Smith HC, Quinn CP, et al. Human monoclonal antibodies against anthrax lethal factor and protective antigen act independently to protect against Bacillus anthracis infection and enhance endogenous immunity to anthrax. Infect Immun. 2007;75:5425-33. DOI: 10.1128/IAI.00261-07.
[11] Staats HF, Alam SM, Scearce RM, Kirwan SM, Zhang JX, Gwinn WM, et al. In vitro and in vivo characterization of anthrax anti-protective antigen and anti-lethal factor monoclonal antibodies after passive transfer in a mouse lethal toxin challenge model to define correlates of immunity. Infect Immun. 2007;75:5443-52. DOI: 10.1128/IAI.00529-07.
[12] Winterroth L, Rivera J, Nakouzi AS, Dadachova E, Casadevall A. Neutralizing monoclonal antibody to edema toxin and its effect on murine anthrax. Infect Immun. 2010;78:2890-8. DOI: 10.1128/IAI.01101-09.
[13] Leysath CE, Chen KH, Moayeri M, Crown D, Fattah R, Chen Z, et al. Mouse monoclonal antibodies to anthrax edema factor protect against infection. Infect Immun. 2011;79:4609-16. DOI: 10.1128/IAI.05314-11.
[14] Dumas EK, Gross T, Larabee J, Pate L, Cuthbertson H, Charlton S, et al. Anthrax Vaccine Precipitated Induces Edema Toxin-Neutralizing, Edema Factor-Specific Antibodies in Human Recipients. Clinical and vaccine immunology : CVI. 2017;24. DOI: 10.1128/CVI.00165-17.
[15] Uchida M, Harada T, Enkhtuya J, Kusumoto A, Kobayashi Y, Chiba S, et al. Protective effect of Bacillus anthracis surface protein EA1 against anthrax in mice. Biochemical and biophysical research communications. 2012;421:323-8. DOI: 10.1016/j.bbrc.2012.04.007.
[16] Crowe SR, Garman L, Engler RJ, Farris AD, Ballard JD, Harley JB, et al. Anthrax vaccination induced anti-lethal factor IgG: fine specificity and neutralizing capacity. Vaccine. 2011;29:3670-8.
[17] Marrack P, McKee AS, Munks MW. Towards an understanding of the adjuvant action of aluminium. Nature reviews Immunology. 2009;9:287-93. DOI: 10.1038/nri2510.
[18] Wen Y, Shi Y. Alum: an old dog with new tricks. Emerg Microbes Infect. 2016;5:e25. DOI: 10.1038/emi.2016.40.
[19] Glenny AT PC, Waddington H, Wallace U. The antigenic value of toxoid precipitated by potassium alum. J Pathol Bacteriol. 1926;3:8-45.
[20] Vessely C, Estey T, Randolph TW, Henderson I, Cooper J, Nayar R, et al. Stability of a trivalent recombinant protein vaccine formulation against botulinum neurotoxin during storage in aqueous solution. Journal of pharmaceutical sciences. 2009;98:2970-93. DOI: 10.1002/jps.21498.
[21] Zhu D, Huang S, McClellan H, Dai W, Syed NR, Gebregeorgis E, et al. Efficient extraction of vaccines formulated in aluminum hydroxide gel by including surfactants in the extraction buffer. Vaccine. 2012;30:189-94. DOI: 10.1016/j.vaccine.2011.11.025.
[22] PHE. Anthrax Vaccine Precipitated. 1 ed: Public Health England; 2014.
[23] Chittineni SPaM, S. C. Improved method for Hepatitis B vaccine in-vitro potency. International Journal of Pharmaceutical Science Invention. 2014;3:39-42.
[24] Patel VJ, Thalassinos K, Slade SE, Connolly JB, Crombie A, Murrell JC, et al. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. Journal of proteome research. 2009;8:3752-9. DOI: 10.1021/pr900080y.
[25] Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, et al. Quantitative proteomic analysis by accurate mass retention time pairs. Analytical chemistry. 2005;77:2187-200. DOI: 10.1021/ac048455k.
[26] Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ. Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics. 2009;9:1696-719. DOI: 10.1002/pmic.200800564.
[27] Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. Molecular & cellular proteomics : MCP. 2006;5:144-56. DOI: 10.1074/mcp.M500230-MCP200.

[28] Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology. 2018;154:394-406. DOI: 10.1111/imm.12889.
[29] Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC bioinformatics. 2010;11:568. DOI: 10.1186/1471-2105-11-568.
[30] IEDB. http://wwwiedborg/Accessed June 2017.
[31] MegAlign. https://www.dnastar.com/t-megalign.aspx. Sep 2016.
[32] NCBI. https://www.ncbi.nlm.nih.gov/. Jul 2016.
[33] Kramer G, Woolerton Y, van Straalen JP, Vissers JP, Dekker N, Langridge JI, et al. Accuracy and Reproducibility in Quantification of Plasma Protein Concentrations by Mass Spectrometry without the Use of Isotopic Standards. PLoS One. 2015;10:e0140097. DOI: 10.1371/journal.pone.0140097.
[34] Chevreux G, Tilly N, Bihoreau N. Quantification of proteins by data independent acquisition: Performance assessment of the Hi3 methodology. Anal Biochem. 2018;549:184-7. DOI: 10.1016/j.ab.2018.03.019.
[35] Jenkins MK, Moon JJ. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. Journal of immunology. 2012;188:4135-40. DOI: 10.4049/jimmunol.1102661.
[36] Ford ML, Koehn BH, Wagener ME, Jiang W, Gangappa S, Pearson TC, et al. Antigen-specific precursor frequency impacts T cell proliferation, differentiation, and requirement for costimulation. The Journal of experimental medicine. 2007;204:299-309. DOI: 10.1084/jem.20062319.
[37] Martinez RJ, Andargachew R, Martinez HA, Evavold BD. Low-affinity CD4+ T cells are major responders in the primary immune response. Nat Commun. 2016;7:13848. DOI: 10.1038/ncomms13848.
[38] Edwards LJ, Evavold BD. T cell recognition of weak ligands: roles of signaling, receptor number, and affinity. Immunol Res. 2011;50:39-48. DOI: 10.1007/s12026-011-8204-3.
[39] Musson JA, Walker N, Flick-Smith H, Williamson ED, Robinson JH. Differential processing of CD4 T-cell epitopes from the protective antigen of Bacillus anthracis. The Journal of biological chemistry. 2003;278:52425-31. DOI: 10.1074/jbc.M309034200.
[40] Sadegh-Nasseri S, Kim A. Selection of immunodominant epitopes during antigen processing is hierarchical. Molecular immunology. 2018. DOI: 10.1016/j.molimm.2018.08.011.
[41] Fontaine M, Vogel I, Van Eycke YR, Galuppo A, Ajouaou Y, Decaestecker C, et al. Regulatory T cells constrain the TCR repertoire of antigen-stimulated conventional CD4 T cells. The EMBO journal. 2018;37:398-412. DOI: 10.15252/embj.201796881.
[42] Gfeller D, Bassani-Sternberg M. Predicting Antigen Presentation-What Could We Learn From a Million Peptides? Frontiers in immunology. 2018;9:1716. DOI: 10.3389/fimmu.2018.01716.
[43] Pittman PR, Leitman SF, Oro JG, Norris SL, Marano NM, Ranadive MV, et al. Protective antigen and toxin neutralization antibody patterns in anthrax vaccinees undergoing serial plasmapheresis. Clinical and diagnostic laboratory immunology. 2005;12:713-21. DOI: 10.1128/CDLI.12.6.713-721.2005.
[44] Quinn CP, Sabourin CL, Schiffer JM, Niemuth NA, Semenova VA, Li H, et al. Humoral and Cell-Mediated Immune Responses to Alternate Booster Schedules of Anthrax Vaccine Adsorbed in Humans. Clinical and vaccine immunology : CVI. 2016;23:326-38. DOI: 10.1128/CVI.00696-15.
[45] Ingram RJ, Metan G, Maillere B, Doganay M, Ozkul Y, Kim LU, et al. Natural exposure to cutaneous anthrax gives long-lasting T cell immunity encompassing infection-specific epitopes. Journal of immunology. 2010;184:3814-21. DOI: 10.4049/jimmunol.0901581.
[46] Ovsyannikova IG, Pankratz VS, Vierkant RA, Pajewski NM, Quinn CP, Kaslow RA, et al. Human leukocyte antigens and cellular immune responses to anthrax vaccine adsorbed. Infect Immun. 2013;81:2584-91. DOI: 10.1128/IAI.00269-13.
[47] Pajewski NM, Parker SD, Poland GA, Ovsyannikova IG, Song W, Zhang K, et al. The role of HLA-DR-DQ haplotypes in variable antibody responses to anthrax vaccine adsorbed. Genes Immun. 2011;12:457-65. DOI: 10.1038/gene.2011.15.

Figures:

Figure 1 – **A** - Workflow of sample preparation and LC-MS/MS analysis for CF and AVP; **B** - Comparison of Desorption methods - size-based separation of desorbed AVP proteins on 1D gel electrophoresis (Lane 1 – Molecular Weight Std, Lane 2 – Blank, Lane 3 – Sodium Hydroxide and Sodium Citrate method, Lane 4 – Succinic Acid method, Lane 5 - Sodium phosphate dibasic, EDTA method, Lane 6 - RapiGest™ SF surfactant method, Lane 7 - ProteaseMAX™ surfactant method, Lane 8 - Guanidine hydrochloride method, Lane 9 – Blank, Lane 10 – Urea method, Lane 11 - Ammonium sulphate, CPC method, Lane 12 - RapiGest™ SF surfactant, EDTA method).

Figure 2 – **A** - Comparison of the number of proteins identified in two batches of AVP and CF using LC-MS/MS. 261 proteins were found in AVP, 163 proteins were found in CF, 138 proteins were found common in CF and AVP (Two biological replicates and three analytical replicates were performed); **B** – Composition of AVP (PA was the principle component of the vaccine (64%), LF was found to be 8% and EF was found to be 3%, 258 proteins were found in lower abundances, comprising the other 25%) (Venn 2019); **C** – Relative Quantitation of PA, LF and EF was found similar in CF and AVP (Two biological replicates and three analytical replicates were performed). Error bars represent ± 1 Standard Deviation about the mean; **D** - Top 8 most abundant proteins in CF and AVP (PA – Protective Antigen, LF – Lethal Factor, Eno – Enolase, PX01-90, EF – Edema Factor, Chap – Chaperonin 60, AD – Alcohol Dehydrogenae, PGK – Phosphoglycerate Kinase). Error bars represent ± 1 Standard Deviation about the mean.

Figure 3 - A - End-point titers of serum IgG to recombinant Protective Antigen (PA) and recombinant Lethal Factor (LF) measured against a five-parameter logistic human sera reference curve and 50% neutralisation factor (NF_{50}) values for serum neutralization of Lethal Toxin compared to the reference serum determined in a J774A.1 macrophage-based Lethal

Toxin neutralization assay (TNA) in AVP vaccinees (n=10). Measurement of antibody levels was performed with at least four replicates, TNA levels were measured in duplicates. Data has been plotted with 95% confidence interval of the mean. Sample 5 and 10 are from control volunteers; **B** – Linear regression showing correlation between serum PA, LF and PA+LF Antibody titres and TNA levels in AVP vaccinees was 0.482, 0.548 and 0.639 respectively.