Title: Correlating liposomal adjuvant characteristics to *in-vivo* cell mediated immunity using a novel mycobacterium tuberculosis fusion protein: A multivariate analysis study.

Authors: Elisabeth Kastner[a], M. Jubair Hussain[a], Vincent W Bramwell[a], Dennis Christensen[b] and Yvonne Perrie[a*]

[a]School of Life and Health Sciences, Aston University, Birmingham, UK.

[b]Statens Serum Institute, Copenhagen, Denmark.

Key Words: Liposome, subunit vaccine, multivariate analysis, in vivo correlation

***Corresponding author:**

Professor Yvonne Perrie
School of Life and Health Sciences,
Aston University,
Aston Triangle,
Birmingham,
B4 7ET.
Tel: +44 121 204 3991
Email address: y.perrie@aston.ac.uk

Abstract

Objective: In this study, we have used a chemometrics-based method to correlate key liposomal adjuvant attributes with *in vivo* immune responses based on multivariate analysis.

Methods: The liposomal adjuvant, composed of the cationic lipid dimethyldioctadecylammonium bromide and trehalose 6,6-dibehenate was modified with 1,2-distearoyl-*sn*-glycero-3-phosphocholine at a range of mol% ratios and the main liposomal characteristics (liposome size and zeta potential) was measured along with their immunological performance as an adjuvant for the novel, post exposure fusion tuberculosis vaccine, Ag85B–ESAT-6-Rv2660c (H56 vaccine). Partial least square regression analysis was applied to correlate and cluster liposomal adjuvants particle characteristics with *in-vivo* derived immunological performances (IgG, IgG1, IgG2b, spleen proliferation, IL-2, IL-5, IL-6, IL-10, INF-γ).

Key Findings: Whilst a range of factors varied in the formulations, decreasing the DSPC content (and subsequent zeta potential) together built the strongest variables in the model. Enhanced DDA and TDB content (and subsequent zeta potential) stimulated a response skewed towards a cell mediated immunity, with the model identifying correlations with INF-γ, IL-2 and IL-6.

Conclusion: This study demonstrates the application of chemometrics-based correlations and clustering, which can inform liposomal adjuvant design.

Introduction

For a vaccine to be regarded as effective, it must stimulate an adequate immune response, sustain safe administration and be patient friendly [1, 2]. Subunit vaccines contain selected purified antigens and potentially reduce side effects, eradicate reversion to virulence and the need for culturing harmful pathogens, whilst eliciting specific immune responses, ultimately generating a safer, more immunologically defined form of vaccination [2, 3]. As purified recombinant proteins generally induce low immunogenicity when administered alone, a suitable immunostimulatory adjuvant is required to produce a more potent vaccine [4, 5]. Liposomes are one of few immunological adjuvants approved for human administration and have been shown to be competent stimulators of an immune response [6]. In recent studies, key factors that influence the efficacy of liposomal adjuvant activity include vesicle charge, size and bilayer fluidity, as these affect interactions with immune system components [7]. For example, enhance antigen adsorption and retention, and an increased intensity in intracellular liposome presence, promoted by using cationic liposomal adjuvants is seen as a viable approach for effective vaccine delivery [[1, 8, 9,].

Despite potentially curative pharmacotherapies being readily available for many decades, tuberculosis (TB) is still the primary cause of preventable deaths worldwide [10]. The necessity of a host to inhibit Mycobacterium tuberculosis (MTB) infection is dependent upon the stimulation of cellular Th1 type immunity. Liposomal composition is a key variable that can influence the potency of such adjuvant delivery systems for TB vaccines. Cationic liposomes of dimethyldioctadecylammonium bromide (DDA) with an optimised incorporation of the glycolipid trehalose 6,6-dibehenate (TDB) forms an adjuvant system (CAF01) capable of stimulating powerful cell-mediated immunity against MTB, upon successful delivery of the recombinant TB fusion protein, Ag85B–ESAT-6 (H1 vaccine) [11].

With modern and high throughput analytical equipment, researchers often accumulate a large quantity of data, which necessitates the use of appropriate analytical tools for extraction of valuable information. Analysing such large data sets requires time and is a particular challenge for extracting the most useful information out of that data set. Computer-based methodologies are incorporated into the analysis of large data sets, in order to extract features within a reasonable timeframe. Often, the analysis of only one variable at a time is not sufficient and the simultaneous analysis of several variables is highly desirable. Multivariate analysis (MVA) is a flexible and multipurpose tool for data analysis. MVA can be used to provide an overview in a data set, for

classification and comparison between groups of data and for regression modelling between two sets of data, often referred as variables (X) and responses (Y). Opposed to multiple linear regression tools, MVA handles many variables and many observations at a time and deals with dimensionality problems. Furthermore, it can extrapolate using limited data sets and is relatively robust to noise in the variables, as well as the responses [12]. Principal components (PC) are computed through the multidimensional space to approximate the best data fit. In order to model the systematic variation in the data set, usually at least two PC are computed, orthogonal to each other, which aim to approximate the data as much as possible.

Principal component analysis (PCA) is the basis in a multivariate analysis, where a simple overview of the information in a dataset is required. Here, a large data set is grouped and trends and outliers are identified [13, 14]. PCA produces a summary, which identifies correlation between observations or groups. Furthermore, trends or sudden shifts in the dataset can be identified. PCA is used for identification of the relationship between the X-variables only and reduces the dimensionality of a multivariate data table into a lower-dimensional plane. Partial least square (PLS) analysis additionally deals with the Y-variables, the responses in a particular system or measurement. Here, the aim is to predict Y from X. The application of PLS determines how the responses are influenced by the factors and variables in a process, as well as identifying response correlations. Furthermore, we can use PLS to identify controlling factors responsible in achieving a desired response [14-16].

The application of relatively simple statistical analysis on experimentally obtained data is common practice. The use of more advanced statistical tools like Design of Experiment (DoE) studies and MVA studies are becoming more commonplace. Nevertheless, the combination of such theoretical multivariate models with experimentally obtained data or offline analysis may result in powerful systems providing extra information and confidence in a given research application. *In-vivo* testing of new pharmaceutical or biopharmaceutical compounds is time and cost intensive and currently indispensable during the development of new pharmaceutically active compounds. Whereas offline analytics are relatively simple and cost-effective, and if effective would be beneficial *in-vivo* predictions. This necessitates that the critical quality parameters of a given system are known and identified.

The goal of this study was to correlate and cluster *in-vivo* adjuvant activity from characteristics of a set of liposomal adjuvants containing the cationic lipid dimethyldioctadecylammonium bromide (DDA) and trehalose 6,6-dibehenate (TDB).

Liposomes formulated from DDA:TDB were chosen as the initial formulation as we have investigated and characterised its activity as an adjuvant [e.g. 7-9]. To generate a set of formulations based on DDA:TDB, we incorporated increasing levels of the saturated phosphatidylcholine, 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC), into DDA:TDB, where the ratio of DDA:TDB remained locked at a 8:1 molar ratio, resulting in 4 formulations with varying DDA,TDB and DSPC concentrations(Table 1). Using these formulations, we investigated the effect of liposomal composition and physical attributes on adjuvant action to identify key controlling features of the liposomes using MVA. MVA was used to both identify clusters of specific immune responses, and to verify a possible link to the physicochemical properties of an adjuvant, namely the size and zeta potential of the liposomes. A tuberculosis antigen vaccine candidate, known as H56, that combines the early secreted antigens of Ag85B–ESAT-6 with the latently expressed Rv2660c antigen, shown to provide protective immunity before and after exposure [17] was used in these studies. In this study, we combine the experimentally obtained data with a theoretical model that was based on PCA and PLS analysis in order to allow for prediction of liposomal adjuvant *in-vivo* performance.

Materials and Methods

Materials

Dimethyldioctadecylammonium (DDA), trehalose 6,6-dibehenate (TDB) and 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) were purchased from Avanti Polar Lipids (Alabaster, Alabama, USA). The fusion protein Ag85B-ESAT-6-Rv2660c (H56 antigen), synthesised to a final concentration of 0.7 mg/mL, was obtained from the Statens Serum Institut (SSI, Copenhagen, Denmark). Tris-base (Ultra Pure), purchased from ICN Biomedicals (Aurora, OH) was used to make Tris buffer (adjusted to pH 7.4 with HCl). Phosphate Buffered Saline (PBS) tablets were purchased from Sigma-Aldrich Co. Ltd. (Dorset, UK). Chloroform and methanol (extra pure) were purchased from Fisher (UK). Double distilled water was used in preparation of all solutions.

Preparation of liposomes via lipid hydration

Liposome formulations were prepared by the long established method of lipid hydration [18]. Lipids were dissolved in a chloroform:methanol mixture (9:1 v/v), with DDA and TDB set to a 5:1 DDA:TDB weight ratio/8:1 molar ratio. Additional liposomal formulations were prepared where this DDA:TDB remained locked at this ratio but DDA:TDB was substituted with DSPC at ratios of 25, 50 and 75 % (Table 1). These lipid mixtures were added to a round bottomed flask and upon solvent extraction via

rotary evaporation and N_2 flushing, a dry film was produced. The remaining film was hydrated in Tris buffer (10 mM, pH 7.4) for 20 minutes at 10 °C above the main gel-to-liquid phase transition of DDA at ~47 °C [11, 19] or DSPC at 55 °C to completely hydrate the film and form liposomes. Addition of H56 was performed after liposome formation at final concentrations of 0.1 mg/mL. Antigen adsorption to liposomes was promoted by incubation for 30 minutes at room temperature.

Determination of particle size and zeta potential by dynamic light scattering

The z-average diameter and zeta potential was measured using via dynamic light scattering (DLS) (Malvern Zetasizer Nano-ZS, Malvern Instruments, Worcs., UK). Measurements took place at 25 °C in (1/10 dilution; 1 mM TRIS, pH 7.4). All measurements were carried out on triplicate batches of formulations.

Immunisation study

Vaccination of mice

All experiments were undertaken in accordance with the 1986 Scientific Procedures Act (UK). All protocols have been subject to local ethical review and were carried out in a designated establishment under the project license number PPL 30/2743. Female C57BL/6 mice, 6-8 weeks old were obtained from Charles River, UK. Vaccine preparations were prepared with the liposomes (Table 1) with the addition of Ag85B-ESAT-6-Rv2660 (H56) antigen to a final concentration of 0.1 mg/mL (5 µg/vaccine dose). All mice, with the exception of the naive group, were immunised intramuscularly (i.m.) with the proposed vaccines (0.05 mL/dose) three times, with two week intervals between each immunisation.

Sera collection

Five scheduled bleeds took place over the seven-week immunisation study with blood samples taken at regular intervals prior to termination. Blood drawn from the tail vein (50 µL) with micropipette capillary tubes coated in heparin solution (0.1% w/v in PBS), was added to 450 µL PBS (giving a final dilution of 1/10) and centrifuged using a Micro Centaur centrifuge at 13,000 RPM for 5 minutes. The supernatants of each mouse sample was collected and stored at -20 °C for future analysis.

In-vitro spleen cell culture

Spleen cell suspensions were produced into 10 mL RPMI 1640 cell culture medium (w/o Glutamine) supplemented with 10% (v/v) FBS and 1% (v/v) PSG (BioSera, East Sussex, UK). Cell suspensions were then centrifuged at 1000 RPM for 10 min at 15 °C and upon supernatant removal, the remaining pellet was resuspended in 10 mL

RPMI, before repeated centrifugation prior to pellet resuspension in 5 mL RPMI. Single cell suspensions were used to evaluate splenocyte proliferation and antigen specific cytokine responses. For splenocyte proliferation, H56 was added to sterile 96 well cell culture plates (Greiner Bio-One Ltd, Gloucestershire, UK) at various concentrations of 0-25 µg/mL with a positive control of concanavalin A (2 µg/mL). 100 µL of spleen cell suspensions were added and incubated at 37 °C, 5% CO_2, and upon 72 hours incubation, 40 µL of [^{3}H] thymidine at 0.5 (µCi) in supplemented RPMI was added per well and incubated for 24 hours. Well contents were harvested onto quartz filter mats (Skatron/Molecular Devices, Berkshire, UK) using a cell harvester (Titertek Instruments, Alabama, USA) and transferred to 20 mL scintillation vials (Sarstedt, Leciester, UK) containing 5 mL scintillation cocktail (Ultima Gold, PerkinElmer, Cambridgeshire, UK). Incorporation of [^{3}H] thymidine in cultured cells was measured with a scintillation counter.

Assessment of H56 specific antibody isotype titres

Serum samples were assessed for levels of IgG, IgG1 and IgG2b antibodies by the enzyme-linked immunosorbent assay (ELISA). The ELISA plates (96 well, flat bottomed, high binding, Greiner Bio-One Ltd, Gloucestershire, UK) were firstly coated with 3 µg/mL H56 antigen prior to overnight incubation at 4 °C. All plates were washed three times with PBST wash buffer (40 g NaCl, 1 g KCl, 1 g KH_2PO_4, 7.2 g Na_2HPO_4, ($2H_20$) per 5 litres of ddH_20, incorporating ~0.4 mL of Tween 20) (Microplate washer, MTX Lab Systems, INC., Virginia, USA). Plates were then blocked by coating each well with 100 µl of Marvel in PBS (dried skimmed milk powder, 4% W/V, Premier Foods, Hertfordshire, UK) and incubated for one hour at 37 °C before washing three times with PBST buffer. 140 µL of serum sample was serially diluted in PBS (70 µL sequentially) in dilution plates, added to the washed ELISA plates and incubated for one hour at 37 °C. Plates were then washed five times with PBST buffer before the addition of 60 µL/well of horseradish peroxidise (HRP) conjugated anti-mouse isotype specific immunoglobulins of IgG, IgG1 and IgG2b (AbD serotec, Oxfordshire, UK) diluted to 1/750, 1/4000 and 1/4000 in PBS respectively, to identify anti-H56 antibodies. Plates were washed a further five times with PBST buffer before adding 60 µL/well substrate solution (colouring agent: 6x 10 mg tablets of 2,2'-azino-bis (3-ethylbenzthiazoline-6-sulfonic acid) (ABTS; Sigma, Dorset, UK) in citrate buffer (0.92g Citric Acid + 1.956g NA_2 HPO_4 per 100 mL) incorporating 10 µL of hydrogen peroxide (30% H_2O_2/100 mL) and incubation for 30 min at 37 °C. Absorbance was read at 405 nm using a microplate reader (Bio-Rad Laboratories, model 680, Hertfordshire, UK). Known positive serum and pooled naïve mice sera were used as positive and negative controls respectively.

Quantification of cytokine production by the sandwich ELISA

Isolation of splenocyte cell suspensions and plating onto 96 well cell culture plates was conducted as summarised above. The cells were subsequently incubated for 48 hours at 37 °C (5% CO_2), prior to supernatant removal and storage at -70 °C for future analysis. Quantification of the cytokines, IL-2, IL-5, IL-6, IL-10 and IFN-γ within cell culture supernatants took place using each specific DuoSet ELISA development kit (R&D Systems, Oxfordshire, UK). The plates were firstly coated with 100 µL capture antibody per well and incubated at room temperature overnight. The plates were then washed three times with PBST buffer before blocking. The plates were subsequently incubated at room temperature for a minimum of one hour before washing a further three times. 100 µL/well of sample or standards was then added to each well and incubated for two hours at room temperature. The plates were washed three times before adding 100 µL of cytokine specific detection antibody per well and incubation for two hours at room temperature. Upon washing three times, 100 µL of Streptavidin-horseradish peroxidise (HRP) was added per well (diluted 1/200). The plates were then covered to avoid exposure to direct light and incubated at room temperature for 20 minutes. After three more washes, 100 µL substrate solution was added to each well (1:1 mixture of colour reagent A and B: stabilised hydrogen peroxide and stabilised tetramethylbenzidine respectively). The plates were then covered and incubated at room temperature for 20 minutes. The experimental reaction was halted by adding 50 µL stop solution (2N H_2SO_4) per well. The optical density was immediately determined using a microplate reader at 450 nm (Bio-Rad Laboratories, model 680, Hertfordshire, UK).

Statistical tests

Data was analysed by one-way analysis of variance (ANOVA) followed by the Tukey test to compare mean values of different groups. Differences were considered to be statistically significant at $p < 0.05$.

Multivariate Data Analysis

Principal Component Analysis (PCA) and Partial Least Square (PLS) regression analysis was performed (SIMCA version 13.0, Umetrics) in order to analyse more than one variable at a time. The relationship between the variables DDA concentration, liposome size and zeta potential and the immunological responses (IgG, IgG1, IgG2b, INF-γ, IL-2, IL-5, IL-6, IL-10, spleen proliferation) was displayed in a loading plot, using all experimentally obtained raw data in this study. Model fit was interpreted by goodness of fit (R^2) and goodness of prediction (Q^2) and regarded as good for $R^2>0.5$

Weights were selected to maximize the correlation. The loading scatter plot was used for identifying relationships between the variables and the responses, as well as the relationships between the variables themselves and the responses themselves. For interpretation, a line from a selected variable was drawn though the origin of the loading scatter plot and X- and Y-variables were projected on the line. Variables opposite to each other were determined as negatively correlated, positive correlation was determined with variables adjacent to each other. The specific regression coefficients plots are used to evaluate the X-Y relations in the here computed PLS model. Correlated responses demonstrate similar coefficient profiles, whereas uncorrelated responses would show a different profile. The model was validated using a permutations plot with 40 permutations for each Y-response.

Results and Discussion

Liposomal adjuvants characteristics

Upon vesicle production, dynamic light scattering was used to determine the particle size, polydispersity and zeta potential of the liposomes before and after H56 antigen addition (0.1 mg/mL: 5 µg/vaccine dose). In the present study, DDA-TDB remained locked at a molar ratio of 8:1, as previous studies found this ratio to be most beneficial in immunological performance [11]. This formulation was modified by the incorporation of DSPC in substitution for DDA-TDB at various molar % ratios, therefore the concentrations of DSPC, DDA and TDB were each varied but the DDA and TDB concentrations were linked (Table 1). From the results, it can be seen that varying the composition of the liposomes resulted in changes in both vesicle size and zeta potential (Table 2). The particle size of DDA-TDB liposomes in Tris buffer prior to substitution was ~500 nm, with a polydispersity of 0.3 and a strong cationic surface charge of ~50 mV (Table 2), in accordance with previous results [11, 19, 20]. Incorporation ofDSPC generated significantly larger vesicles ($P < 0.05$) but with no clear trend of DSPC concentration to vesicle size and all remained in a sub micrometer size range of 650-850 nm. In contrast the zeta potential decreasing with increasing DSPC, as would be expected (Table 2). Upon surface adsorption of H56 antigen the particle size of all formulations increased significantly ($P < 0.05$) to 850 -1300 nm depending on the formulation, whilst cationic zeta potential decreased (Table 2). For all 4 formulations tested antigen loading was > 85 % (results not shown), with no significant difference, presumably due to the high cationic lipid content/anionic antigen content even with the 75 % DSPC formulation. For MVA analysis the liposome characteristics post-addition of antigen were used.

Immunological characterization for H56 specific antibody isotypes

When considering the antibody responses in mice immunised, by day 37 all four of the liposome formulations induced significantly higher ($P < 0.05$) IgG immune responses in mice compared to mice immunised with antigen alone, with no significant difference between the formulations (Fig. 1A). A similar trend was noted with IgG1 responses in the vaccinated mice (Fig. 1B). In the case of IgG2b (Fig. 1C), liposomal adjuvants composed of 75 mol% DSPC generated significantly lower ($P < 0.05$) levels of antibody titres at all time points tested compared to DDA-TDB, and IgG2b responses were not significantly different to responses in mice immunised with non-adjuvanted H56 (Fig. 1C). This suggests that up to 50 % DSPC within the liposome formulation did not compromise the immunogenic effect of the DDA-TDB adjuvant, which is capable of inducing protective cellular immunity against TB when administered with a model vaccine antigen [21]. This data is in line with previous studies conducted within our group, where DDA was directly replaced with DSPC but the TDB concentrations were not changed (and hence the 8:1 molar of DDA-TDB was not maintained) [22]. This suggests that IgG1 antibody responses remain high over a wider range of DDA and TDB concentrations and liposome characteristics whilst IgG2b decreased with decreasing DDA content, irrespective of the DDA-TDB ratio.

Immunological characterization for H56 specific spleen proliferation rates

Antigen specific splenocyte proliferation in mice previously vaccinated with the liposomal systems and upon re-stimulation with H56 vaccine at increasing concentrations from 0-25 µg/mL was assessed. DDA-TDB liposomal adjuvants generated the strongest cell proliferation (Fig. 2). However, cell proliferation was seen to be dependent on DDA-TDB concentration as there is a notable trend of decreasing responses from cells harvested from mice immunised with liposomes containing increasing DSPC levels (and corresponding decreasing levels of DDA-TDB) within the liposome formulation (Fig. 2). Indeed, liposomal adjuvants containing 75 mol% DSPC were consistently low even upon re-stimulation at higher H56 concentrations (Fig. 2).

Spleen cell cytokine responses

Spleen cell cytokine responses from mice immunised with the various liposomal formulations show variable correlation to the DSPC content (Fig. 3). In general, IFN-γ, IL-2 and IL-6 levels were shown to decrease with increasing DSPC content (Fig. 3A, B and D). Whilst IL-5 production was low for all groups (Fig. 3C), with mice which received antigen alone having similar levels to those mice which received liposomal adjuvants. In contrast, the presence of DSPC in the liposomal adjuvant tended to increase IL-10 responses (Fig 3E).

With increasing replacement of DDA-TDB with DSPC in the formulation, the zeta potential decreases and the strength of immune response tends to skew towards a Th2 type response, even with the small decreases in zeta potential noted in these formulations (Table 1). The effect of liposomal charge has been studied previously for the quality of immunity stimulated with Ag85B-EAST-6 antigen [20] in which it was noted that production of IFN-γ was strongly dependent upon the liposomal adjuvants being positively charged. In contrast, DDA-TDB substituted with 75 mol% DSPC displayed a weak cellular immune response. The resultant Th2 type immune response observed can be considered to be independent of the surface charge of the system, corresponding with previous studies [20] stating that a Th2 type elicited response was not significantly affected by liposomal adjuvant charge.

Multivariate analysis for clustering Th1 and Th2 type immune responses to adjuvant characteristics

Multivariate model evaluation

Whilst the above *in-vivo* results are in line with previous studies, it is difficult to investigate the multifactorial changes in liposome attributes that occur when the lipid composition is modified, therefore the principle aim of this work was to analyse this *in-vivo* data set using MVA. Initially, the correlation of two fitted principal components (PC1 and PC2) for the overall model fit was determined as loadings and weights. The model type was PLS with 12 observations. Initially we selected the liposome size and the DDA concentration as x-variables, (2 X-variables and 10 Y-variables). This data was chosen in order to assess whether the size of the liposome or the DDA concentration (which was linked to the TDB concentration) is the most contributing factor in the vaccine immune response. The fraction of the X-variation modelled in PC1 was 62 % (eigenvalue 1.24) and 100 % in PC2 (eigenvalue 0.764). The fraction of the Y-variation modelled with the first PC was 46 %, and 13 % in the second PC. The cumulative goodness of fit was 0.59 and the cumulative goodness of prediction was 0.37.

In the second analysis study, we selected the liposome size, zeta potential and DDA concentration as variables (3 X-variables and 9 Y-variables). Obviously, given that zeta potential measurements are be directly linked to the amount and type of lipid used (as well as the aqueous media the liposomes are suspended in), the zeta potential represents a response towards lipid composition. However, this set allowed us to verify how the model predicted the influence of zeta potential on immune responses *in-vivo*. Here, the cumulative goodness of fit was 0.97 and the goodness of prediction was 0.52, with two PC fitted (PC1 with 64% of the fraction in the X-variation modelled

(eigenvalue 1.92), 97% respectively in the PC2 (eigenvalue of 0.98); Y-variation modelled in first PC was 44%, 53% in PC2). Unfortunately, including the TDB concentration as a variable resulted in a non-statistically valid model.

PC1 and PC2 in both model setups were regarded to comprise satisfactory information to construct a predictive model on the data set. Furthermore, we analysed the cumulated R^2 and Q^2 values for each Y-variable (Fig. 4), in both model setups. R^2 represents a goodness of the model fit and describes how well the variation of the respective variable is explained; Q^2 indicates how well the respective variable can be predicted. A threshold value for $R^2 > 0.5$ was chosen for valid models; values below 0.5 indicated noise present. IgG and IgG1 responses were shown to be insignificant in both designs chosen, due to negative Q^2 value (Fig. 4 A and B). Spleen proliferation, INF-γ, IL-2, and IL-6 showed good model fit above 0.5, with respective good prediction power indicated by a relatively low level of noise in the data set (Fig. 4 A and B). Goodness of prediction for the responses IgG2b, IL-5 and IL-10 was at or below 0.5, indicating a higher amount of noise present for these responses.

PLS regression to cluster H56 specific antibody isotypes

Modelling of the data revealed no strong or moderate outliers present (evaluated in the PCA analysis; data not shown). Due to insignificance in the model for the antibody subtypes IgG and IgG1, these were removed from further analysis, with IgG2b remaining, but at a low confidence level. This is in line with the basic statistical analysis in Figure 1 that revealed no significant difference between the formulations for IgG and IgG1, confirming that these antibody subtypes are not an ideal measure for vaccine efficacy in these systems, indicated by statistical insignificance in the PLS analysis.

PLS regression to analyse specific spleen proliferation rates

The liposomal adjuvants were shown to promote splenocyte proliferation upon restimulation with H56 antigen, demonstrated by the strong correlation between the variables DDA (Fig. 4C) (and zeta potential; Fig. 4D) to splenocyte proliferation in the coefficient plot, with size not shown to correlate with responses. DDA concentration is the most influential variable for the response spleen proliferation, visible by the high coefficient value (close to 1) as well as a small confidence interval. The loading scatter plots shows a close correlation of splenocyte proliferation response to the variable DDA (Fig. 4E) and zeta potential (Fig. 4F), identifying their strong correlation. This confirms that the biggest effect to spleen proliferation rates is the increase in DDA content, which is strongly linked to the zeta potential of a vaccine. As indicated in Figure 3, the peak of proliferation correlates with DDA-TDB liposomes, which have the strongest zeta potential (Table 2).

PLS regression to cluster cytokines responses

The PLS analysis revealed a statistical significance for the responses INF-γ, IL-2, IL-6 and IL-10 for the variables DDA (Fig. 4 C), as well as for the variable zeta potential (Fig. 4D) again as would be expected due to their link. Overall, the DDA content as well as the zeta potential showed a positive correlation to INF-γ, IL-2 and IL-6, and an inverse correlation to the response IL-10. The increase in DDA in a vaccine adjuvant formulation gave no notable correlation in size but does result in a higher zeta potential, which is predicted to increase the specific INF-γ, IL-2 and IL-6 production *in-vivo*. The corresponding peak in INF-γ production (Fig. 3A) was detected for the DDA-TDB liposomes, which also provided the strongest cationic zeta potential (Table 2). Here, the model predictions are in line with the previous reported results that showed increasing cationic charge (but with constant TDB concentrations across the formulations) enhanced INF- γ as well as IL-6 [22]. However in addition to this, the model suggests no impact of DDA and zeta potential content on IL-5, but an inverse correlation between the response IL-5 and the liposome size (Fig. 4C and D), indicating that a smaller liposome size is predicted to increase the specific IL-5 production. Nevertheless, initial model evaluation of the response IL-5 indicated a level of noise present in the data set, which should be considered in any predictions made until model validation is verified.

The specific regression coefficients (Fig. 4E and F) represent the X-Y relations in the computed PLS model; which simplifies the model overview. Correlated responses demonstrate similar coefficient profiles. Similar coefficient profiles for the responses INF-γ , IL-2 and IL-6 for the variables DDA and zeta potential suggesting a grouping and relation between those cell mediated responses, which can be clustered together as Th1-specific immune responses driven by the DDA content. This cluster is visible in both loading scatter plots (Fig. 4E and F) and not influenced by zeta potential being included in the model as a variable or a response, with a strong cluster of the responses INF-γ, IL-2, IL-6 and IgG2b, all which are linked with Th1 specific immune responses. When several Y-variables need to be modelled and analysed together, PLS offers the ability to generate a simpler depiction of data sets, rather than generating separate models for each response. It is recommended to analyse strongly correlated Y-variables together and group them, as their correlation stabilizes the model [12]. However, this only applies for dependent responses that measure and incorporate similar measurements.

Model summary

We see that the DDA (and the linked TDB) concentration in a vaccine formulation is a crucial variable and most importantly more influential to the immunological response than the actual liposome size (across the range considered). Generally, a strong link between the DDA concentration and zeta potential could be identified; for selecting the zeta potential as a Y-response (Fig. 4E) as well as a X-variable (Fig. 4F), its close link to DDA as a variable confirms the significance of the zeta potential to initiating a Th1 mediated immune response *in-vivo*. Overall, the model developed was statistically valid for the variables, DDA and zeta potential (spleen proliferation, IFN- γ, IL-2, IL-6, IL-10, IgG2b), and to limited extent liposome size (in the case of IL-5), as summarized by the importance of the x-variables. The variable influence on projection plot (VIP) (Fig. 5), which summarizes all components and y-variables [23], indicated that the variable DDA content (Fig. 5A) and zeta potential (Fig. 5A&B) were ranked as the variables with the highest impact in the PLS models. However, whilst the zeta potential is shown to strongly influence the immune responses *in-vivo* and thus could be taken as a controlling factor, it is directly linked to the DDA content. Furthermore, we have previously shown that liposomes of the same DDA content, and hence same zeta potential, gave different immunological profiles depending on the TDB content [11]. This demonstrates that controlling factors between the formulation and the physico-chemical characteristics must be identified when applying MVA to avoid incorrect interpretation.

Model validation

To assess the validity of the predictions made by the PLS analysis, the model was validated using respective permutations plots for each specific Y-response (Fig. 6). The permutation plots helped to assess the validity of the PLS model by assessing the risk of invalidity and verifying that the model does not only fit the current data set, but also predicts Y from new observations.

Model validation is a crucial diagnostic function of MVA. Here, the X-data is left unmodified, whilst the Y-data is permuted and arranged in a different order after which a PLS model is fitted to the permuted data set. The derived models are cross-validated by computing R^2 and Q^2. This random shuffling of the Y-data allows comparing the permuted values with the real R^2 and Q^2 values of the model. This permutation procedure is repeated for a certain number, mostly between 25 and 100, (here, we chose 40), which leads to the generation of parallel PLS models thus establishing reference distributions based on random data. These references are used to assess the statistical significance in the initial PLS model [24].

Here, the goodness of fit and prediction (R^2 and Q^2) of the current model were compared with the R^2 and Q^2 of randomly permuted Y-observations while the X-variables were maintained constant. For each Y-variable, 40 permutations were selected. The R^2 and Q^2 values from the original model were shown on the far right end of the respective graphs, whereas the Y-permuted models were shown on the left side. The correlation between permuted Y-vector to the original X-vector was depicted by the horizontal correlation axis. The criteria for model validity have been selected as the intercept of the Q^2 regression line at or below zero. Furthermore, the validity was assessed by depiction of all permuted R^2 values below the R^2 of the original model.

The initial model that evaluated the zeta potential as a response, showed an excellent model validity with its respective permutation plot (Fig. 6A), confirming that the response zeta potential can be modelled and described by PLS methods. Models for the responses spleen proliferation, INF-γ, IL-2 and IL-6 showed excellent permutation plots (Fig. 6 B, C, D, E), confirming the validity of the PLS model and predictions made from selected responses. Validation for the variables IgG, IgG1, IgG2b, IL-5 and IL10 failed (plots not shown), confirming the previous invalidity of the models as already seen in initial model evaluation (Figure 4 A and B). Furthermore, this confirms that the initially detected higher level of noise present for IL-5, IL-10 and IgG2b resulted in a non-valid model, exemplifying that any predictions made using MVA depend on verifying the validity of the models by the permutation testing.

Nevertheless, interpretations should be made in consideration of the assay accuracy, which might lead to a higher level of noise in the data set, as seen for the variables IgG2b, IL-5 and IL-10. Although clear trends and clusters were visible, interpretation always depends on the accuracy of the assay. Furthermore, wider formulation profiling is required to challenge this use of MVA in more complex vaccine adjuvant studies. However, results here emphasize the use of multivariate analysis as a new tool for *in-vivo* vaccine efficacy correlations and cluster analysis for Th1 specific immune responses.

This study shows that useful clustering, trends and predictions can be made using MVA tools when a range of factors are varied (in this case DDA, TDB and DSPC content which results in variations in vesicle size and zeta potential). Correlating *in-vivo* data may be a cost effective way for initial information about vaccine efficiency. Information extracted from MVA may speed up the drug and process development process, as desired *in-vivo* immune response targets might be predicted and are dictated by the characteristics of the adjuvant or delivery system. From the present

study evaluations, the extraction of information from *in-vivo* data by partial least square regression models gives a powerful tool to further characterize a vaccine formulation. It can be used for initial clustering of *in-vivo* specific immune responses and help to allow for future predictions of vaccine efficiency; overall, a new and useful method to speed up the development process of a vaccine candidate.

MVA is a useful tool for not only summarizing and visualizing data sets, it also allows for classification and identification of quantitative relationships between variables [12]. Matrices can be of alterable amounts of variables and observations, allowing for flexibility in generating the data set. The application of those mathematical and statistical tools is highly applicable for determination of relationships between various measurements derived from a system or process [25]. We define the relationship between two properties, where the effect of one property that can easily be measured in the laboratory is related to the second property, which is more difficult to measure. Initially, data of both property measurements are obtained, which are then built into a model using multivariate regression, linking the dependent and independent variables.

The most significant advantage of using multivariate tools is the ability to analyse multiple variables simultaneously, along with the reduction of the dimensionality of the data set by projecting the data into a lower dimension thus improving data interpretation and presentation [26]. Visualization and simplification of complex pharmaceutical data is one of the main advantages of using MVA tools, and it is highly applicable in pharmaceutical research and process or product development [27]. MVA is furthermore often applied in diagnostics tools, where the identification of the major contributing variables leads to the isolation of the deviation, frequently applied in industrial processes for product quality control [26].

Conclusion

In conclusion, models were developed to cluster and predict Th1 immune responses to the vaccine formulation dependent on liposomal adjuvant characteristics. Substitution of DDA:TDB with DSPC reduced the cationic zeta potential and resulted in variations in vesicle size. The extent of DSPC incorporation correlated to polarised immune responses with a combination of cellular and humoral immunity. We have shown that the use of multivariate tools allows for clustering and predictions from key liposome characteristics to specific *in-vivo* immune responses. The reliability of derived PLS models suggests its general usefulness for predicting *in-vivo* specific immune responses from offline measurements. Such multivariate approaches may be useful in correlating key characteristics to critical quality attributes of a vaccine formulation.

Specific variable-dependences and independences support the selection of key variables that need to be further optimized in a development process. Such methods may be particularly useful for screening many variables at a time, especially in early stage development processes.

Acknowledgements

This work was part funded by the EPSRC Centre for Innovative Manufacturing in Emergent Macromolecular Therapies (E Kastner), NewTBVAC (contract no.HEALTHF3-2009-241745) and Aston University.

References

1. Perrie, Y., et al., *Vaccine adjuvant systems: enhancing the efficacy of sub-unit protein antigens.* International journal of pharmaceutics, 2008. **364**(2): p. 272-280.
2. Black, M., et al., *Advances in the design and delivery of peptide subunit vaccines with a focus on Toll-like receptor agonists.* Expert review of vaccines, 2010. **9**(2): p. 157-173.
3. Mohammed, A.R., et al., *Increased potential of a cationic liposome-based delivery system: enhancing stability and sustained immunological activity in pre-clinical development.* European journal of pharmaceutics and biopharmaceutics, 2010. **76**(3): p. 404-412.
4. O'Hagan, D.T. and E. De Gregorio, *The path to a successful vaccine adjuvant–'the long and winding road'.* Drug discovery today, 2009. **14**(11): p. 541-551.
5. Holten-Andersen, L., et al., *Combination of the cationic surfactant dimethyl dioctadecyl ammonium bromide and synthetic mycobacterial cord factor as an efficient adjuvant for tuberculosis subunit vaccines.* Infection and immunity, 2004. **72**(3): p. 1608-1617.
6. Bramwell, V.W. and Y. Perrie, *Particulate delivery systems for vaccines.* Critical Reviews™ in Therapeutic Drug Carrier Systems, 2005. **22**(2).
7. Henriksen-Lacey, M., et al., *Liposomal vaccine delivery systems.* Expert opinion on drug delivery, 2011. **8**(4): p. 505-519.
8. Christensen, D., et al., *Cationic liposomes as vaccine adjuvants.* 2007.
9. Smith Korsholm, K., et al., *The adjuvant mechanism of cationic dimethyldioctadecylammonium liposomes.* Immunology, 2007. **121**(2): p. 216-226.
10. Sosnik, A., et al., *New old challenges in tuberculosis: potentially effective nanotechnologies in drug delivery.* Advanced drug delivery reviews, 2010. **62**(4): p. 547-559.
11. Davidsen, J., et al., *Characterization of cationic liposomes based on dimethyldioctadecylammonium and synthetic cord factor from< i> M. tuberculosis</i>(trehalose 6, 6'-dibehenate)—A novel adjuvant inducing both strong CMI and antibody responses.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2005. **1718**(1): p. 22-31.
12. Eriksson, L., *Multi-and megavariate data analysis*. 2006: MKS Umetrics AB.
13. Jackson, J.E., *A user's guide to principal components*. Vol. 587. 2005: John Wiley & Sons.
14. Wold, S., et al., *Multivariate data analysis in chemistry*, in *Chemometrics*. 1984, Springer. p. 17-95.
15. Wold, S., M. Sjöström, and L. Eriksson, *PLS-regression: a basic tool of chemometrics.* Chemometrics and intelligent laboratory systems, 2001. **58**(2): p. 109-130.
16. Wold, S., et al., *Some recent developments in PLS modeling.* Chemometrics and intelligent laboratory systems, 2001. **58**(2): p. 131-150.
17. Aagaard, C., et al., *A multistage tuberculosis vaccine that confers efficient protection before and after exposure.* Nature medicine, 2011. **17**(2): p. 189-194.
18. Bangham, A., M.M. Standish, and J. Watkins, *Diffusion of univalent ions across the lamellae of swollen phospholipids.* Journal of molecular biology, 1965. **13**(1): p. 238-IN27.
19. Christensen, D., et al., *Trehalose preserves DDA/TDB liposomes and their adjuvant effect during freeze-drying.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2007. **1768**(9): p. 2120-2129.
20. Henriksen-Lacey, M., et al., *Liposomal cationic charge and antigen adsorption are important properties for the efficient deposition of antigen at the injection site and ability of the vaccine to induce a CMI response.* Journal of controlled release, 2010. **145**(2): p. 102-108.

21. Agger, E.M., et al., *Cationic liposomes formulated with synthetic mycobacterial cordfactor (CAF01): a versatile adjuvant for vaccines with different immunological requirements.* PloS one, 2008. **3**(9): p. e3116.
22. Hussain, M.J., et al., *Th1 immune responses can be modulated by varying dimethyldioctadecylammonium and distearoyl-sn-glycero-3-phosphocholine content in liposomal adjuvants.* Journal of Pharmacy and Pharmacology, 2014. **66**(3): p. 358-366.
23. Kubinyi, H., *3D Qsar in Drug Design: Volume 1: Theory Methods and Applications*. Vol. 1. 1993: Springer.
24. van der Voet, H., *Comparing the predictive accuracy of models using a simple randomization test.* Chemometrics and Intelligent Laboratory Systems, 1994. **25**(2): p. 313-323.
25. Lopes, J.A., et al., *Chemometrics in bioprocess engineering: process analytical technology (PAT) applications.* Chemometrics and Intelligent Laboratory Systems, 2004. **74**(2): p. 269-275.
26. Kourti, T., J. Lee, and J.F. Macgregor, *Experiences with industrial applications of projection methods for multivariate statistical process control.* Computers & chemical engineering, 1996. **20**: p. S745-S750.
27. Rajalahti, T. and O.M. Kvalheim, *Multivariate data analysis in pharmaceutics: a tutorial review.* International journal of pharmaceutics, 2011. **417**(1): p. 280-290.

Tables.

Table 1: Incorporation of DSPC into DDA-TDB formulations at 25, 50 and 75 mol%.

	Weight µg per dose		
Formulation (mol%)	DDA	TDB	DSPC
DDA:TDB	250	50	0
+ 25% DSPC	188	36	88
+ 50% DSPC	125	25	175
+ 75% DSPC	63	14	264

Values of weight and µmoles in the various liposome formulations where DDA:TDB was locked at a 5:1 wt ratio/8:1 molar ratio and increasingly replaced with DSPC in a 50 µL dose.

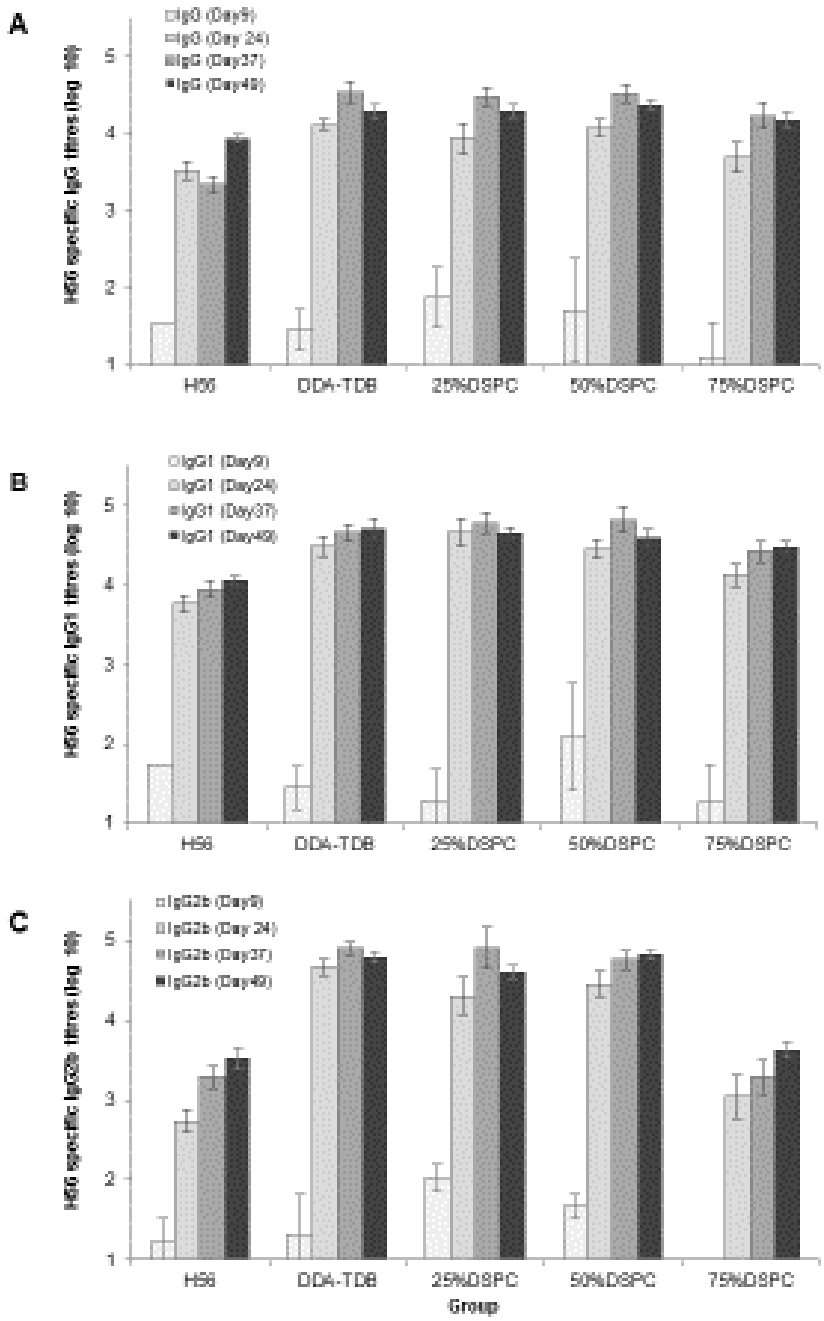
Table 2: Particle size, polydispersity and zeta potential liposomal adjuvants prior to and post H56 antigen adsorption.

Formulation	Antigen	DDA/TDB	25% DSPC	50% DSPC	75% DSPC
Vesicle size (nm)		517 ± 29	640 ± 24	856 ± 114	734 ± 67
	+ H56	981 ± 198	1266 ± 151	1036 ± 92	852 ± 52
Polydispersity		0.32 ± 0.01	0.34 ± 0.01	0.32 ± 0.01	0.33 ± 0.02
	+ H56	0.42 ± 0.02	0.46 ± 0.06	0.54 ± 0.14	0.42 ± 0.1
ZP (mV)		45.7 ± 0.7	42.7 ± 1.9	35.4 ± 3.6	33.2 ± 0.5
	+H56	47.4 ± 6.1	41.4 ± 3.7	31.7 ± 6.4	28.7± 5.3

The liposomes were produced by lipid hydration in Tris buffer (10 mM, pH 7.4) and with H56 vaccine antigen added at 0.1 mg/mL. Characterisation used a Malvern Nanosizer ZS. Results denote the mean ± s.d. for three independent experiments.

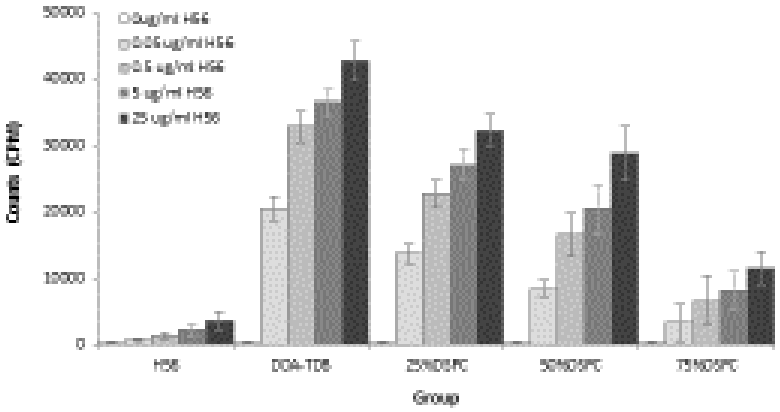
675 **Figure Legends**

676



677 **Figure 1**

678 **Figure 1**: Mean serum H56 specific antibody isotype titres stimulated by DDA-TDB
679 and substitution with 25-75 mol% DSPC (n=5, +/- standard error) for A: IgG, B: IgG1
680 and C: IgG2b subsets. Values display the positive reciprocal end point dilution
681 (log10). Sera was collected prior to the first immunisation and on days 9, 24, 37 and
682 49 respectively thereafter. Serum samples obtained across various time intervals
683 upon immunisation were analysed for the presence of anti-H56 specific antibodies by
684 the enzyme-linked immunosorbent assay (ELISA).



685

686 **Figure 2**. Spleen cell proliferation stimulated by H56 vaccine antigen (at 0, 0.05, 0.5,
687 5 and 25 μg/mL; n=5, mean of replicates ± standard error) for DDA-TDB and
688 substitution with 25-75 mol% DSPC. The level of H56 antigen specific splenocyte
689 proliferation was indicated by the extent of [^{3}H] labelled Thymidine incorporation into
690 cultured splenocytes.

691

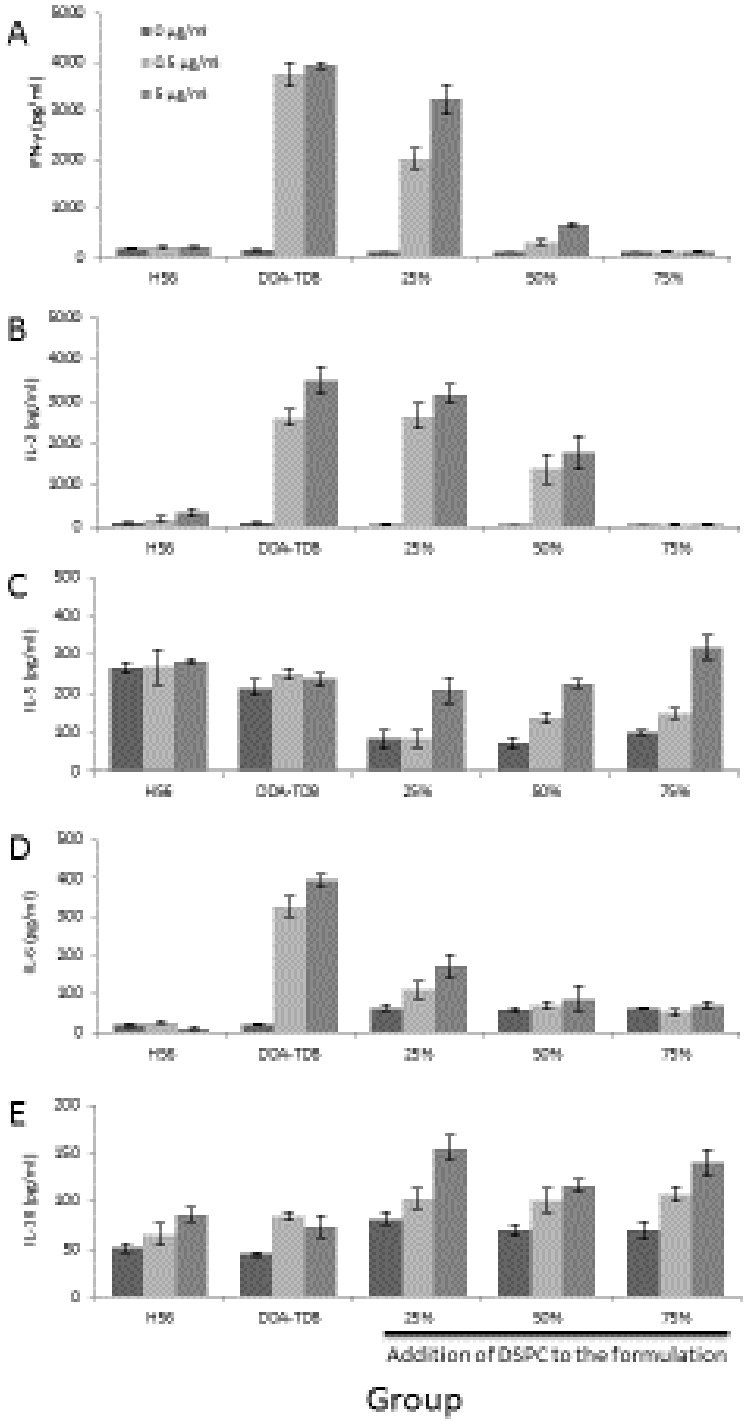


Figure 3

Figure 3. Spleen cell cytokine production in response to re-stimulation with H56 antigen at 0, 0.5 and 5 µg/mL, quantified for A: IFN-γ, B: IL-2, C: IL-5, D: IL-10 and E: IL-6. Results represent mean average cytokine production of five spleens per vaccination group +/- standard error.

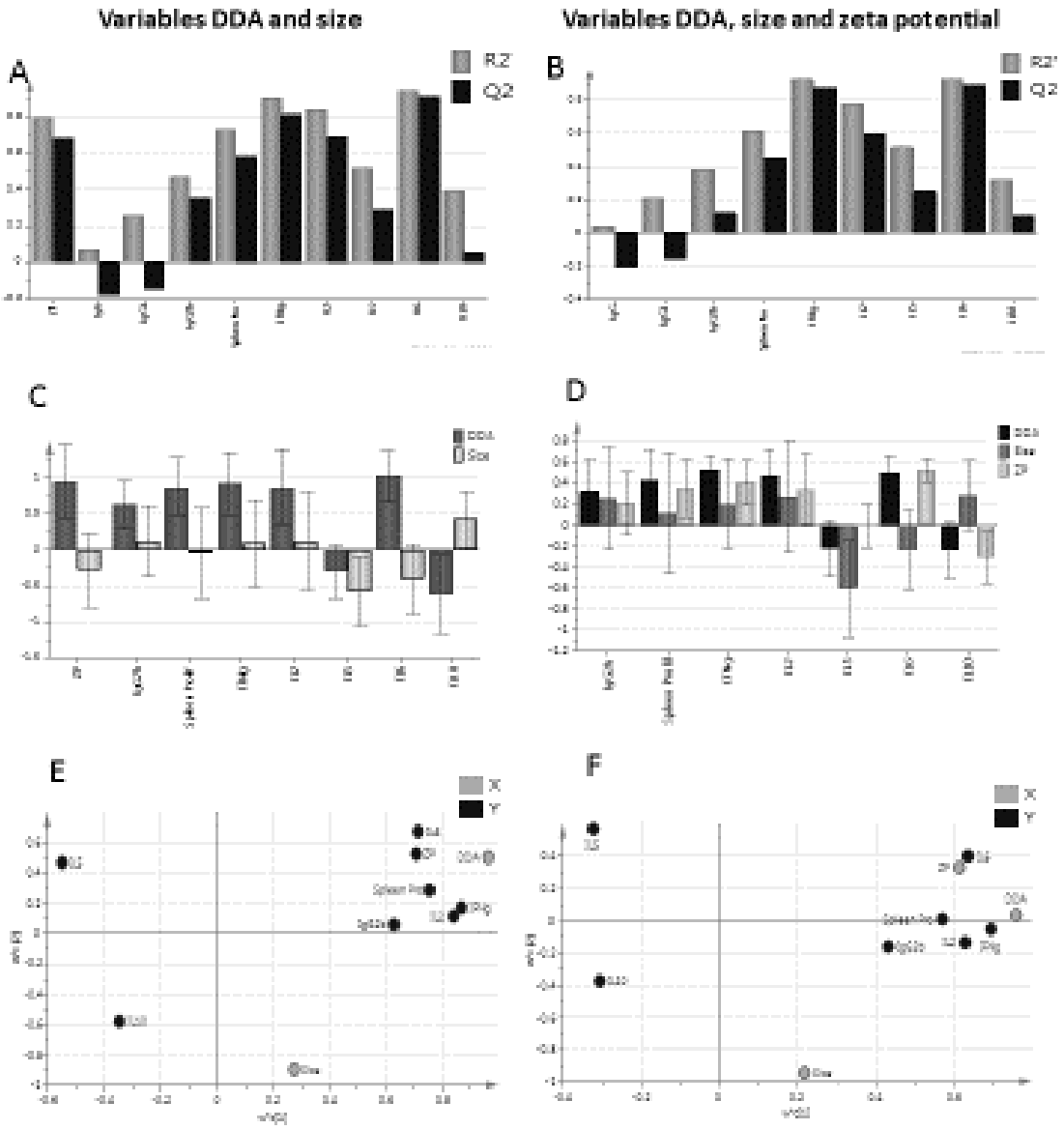


Figure 4. X/Y overview plot indicating the cumulated R^2 and Q^2 values for each response for A) DDA and size and B) DDA, size and zeta potential. Well modelled responses show a R^2 and Q^2 value above 0.5 IgG and IgG1 responses show poor model fit (negative Q^2), that indicates noise and no correlation between the X and the Y variables for those responses (statistical insignificance). PLS analysis results with Coefficient overview, displaying the coefficients for all responses to interpret how the X-variables affect the Y-variables for C) DDA and size and D) DDA, size and zeta potential. Loading scatter plot, where the relation between X and Y- variables are displayed for E) DDA and size and F) DDA, size and zeta potential.

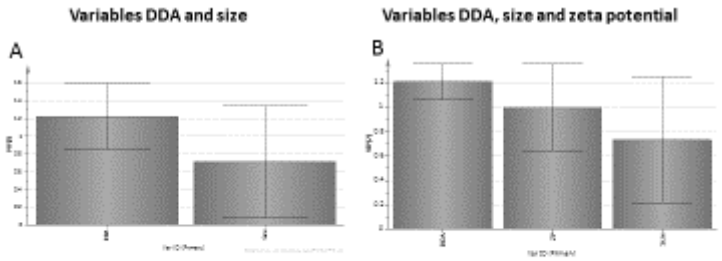
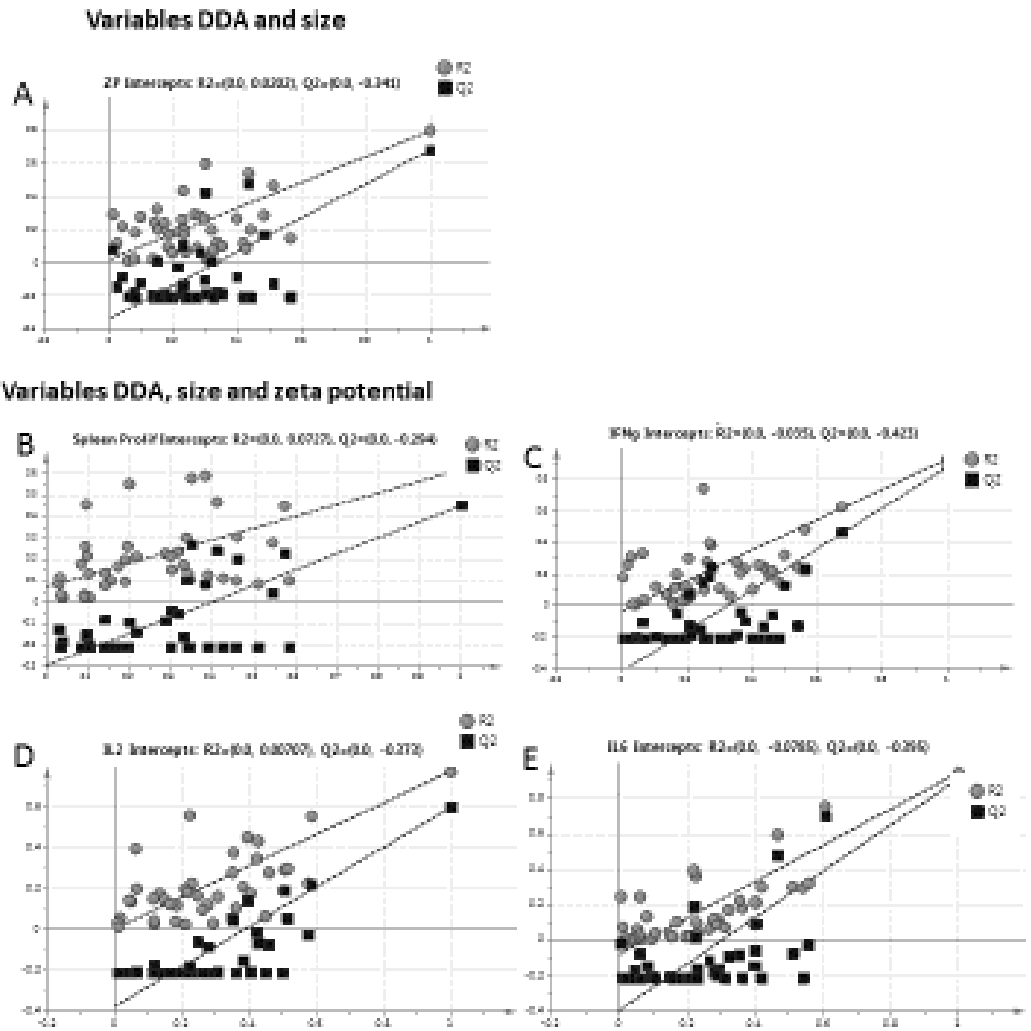


Figure 5

Figure 5. VIP plot (variable importance for projection) summarizing the importance of the variables liposome size and zeta potential. The VIP plot is sorted from high to low and indicates the value of the variable zeta potential as the most important X-variable in the PLS model for A) DDA and size and B) DDA, size and zeta potential.



719

720 **Figure 6**. Permutations plot for A: zeta potential, B: spleen proliferation, C: IFN- γ , D:
721 IL-2, E: IL-6. Model validity was assessed for 40 permutations. The correlation
722 between permuted Y-vector to the original X-vector is depicted by the horizontal
723 correlation axis. The criteria for model validity have been selected as the intercept of
724 the Q^2 regression line at or below zero.

725