in vitro Characterisation of the Complement Cascade for Predicting Patient Outcome Post-operatively

Submitted by Peter Reader to the University of Exeter as a thesis for the degree of Doctor of Philosophy by Research in Medical Studies

May 2019

Signature: PR

Abstract

The identification of surgical patients at higher risk of infection enables targeted allocation of critical care resources to improve patient mortality. The Complement cascade of the innate immune system is known to increase risk of infection if compromised and can be tested *in vitro* as a potential method for stratification of high-risk patients.

Existing assays of Complement function are laboratory bound and require trained personnel to operate and interpret. This thesis describes the development of novel immunoassays for C3, C5a, TCC and TNFα, based on a multiplex biosensor platform with a duty cycle of <15 minutes. The assays for C3 and C5a were validated for use with serum samples with a CV of 23% and 21% respectively, within a dynamic range of 3.2-12.5 nM. When combined with automated data analysis presented here, the biosensor assays provide a step towards easy functional testing of Complement biomarkers at point-of-care. A new technique for assessing the monomeric purity of antibody samples is also presented for quality control in future immunoassay development.

A hypothesis-driven selection of candidate biomarker tests, denoting a compromised immune state, is possible with a mathematical model. Prerequisite techniques for optimising nonlinear models with unknown parameters are presented in a systems biology study of IgG binding to protein A/G to determine reaction kinetics and stoichiometry. The same techniques were used to fit a 187-parameter model of *in vitro* Complement activation to ELISA data from pooled human serum, with a mean absolute error of 14%. Sensitivity and flux analyses verified that the optimised model was consistent with existing knowledge of the system: regulation via Properdin and Factor H was predicted quantitatively.

The optimised model predicted the effect of Complement depletion on the concentration-time profiles of proteins following activation: new phenotypes of immune state. Model predictions for the variation in C5a and TCC formation rate phenotypes in healthy adults showed no significant difference (P>0.05) from the serum data of 22 volunteers. The model and cohort data provide an initial estimate of effect size for future clinical studies investigating the ability of these Complement activation phenotypes to identify high-risk surgical patients or identify the onset of infection.

Plain Language Summary

The risk of contracting an infection is increased when the body's natural defences are compromised in the presence of disease-causing microorganisms. As such, infection often occurs as an undesired result of surgery and is associated with increased patient discomfort and risk of death. The precise level of infection risk depends on both the type of surgery and the health of the patient, which is currently assessed with blood tests of immune system activity before surgery and during recovery. Patients perceived to be at an elevated risk of infection can be allocated additional health care resources, such as increased observation during recovery but infection remains a leading cause of death or delay in recovery in patients following surgery. This thesis reports the development of a new test of the immune system to help identify patients at higher risk of infection due to a compromised immune state from an underlying condition or debilitation.

Among the various components of the immune system that might be tested, the Complement cascade was chosen as the focus of the present study: a system of proteins which circulate in the blood to recognise and kill disease-causing microorganisms on a timescale of minutes. The Complement cascade is activated independently of cellular processes and involves the consumption of its components. Therefore, Complement protein levels can be measured to follow activation of the system over time and data from blood samples are largely representative of how the system behaves in the patient. As such, the Complement system of a patient can be sampled with existing blood collection methods and their Complement function tested quickly and accurately with minimal discomfort.

There are two existing tests of Complement system function used in current clinical care, both reporting the ability of patient blood serum samples to destroy animal cells. These tests are typically only used when a genetic deficiency is suspected and must be performed in the laboratory by trained staff. It was recognised that a new test that is fast and comparatively easy to operate might allow more frequent Complement function testing before surgery and at the bedside of recovering surgical patients. The present work proposes a test of Complement cascade function in a routine blood sample by measuring the consumption of proteins in response to a stimulus.

The concentrations of Complement proteins in blood samples can be tested with a number of methods, many of which use antibodies. Antibodies are proteins which bind specifically to other molecules and can be used in sensors to detect the presence and abundance of Complement proteins. The Shaw lab where the present study was based, has developed an in-house sensor platform, LiScAR, to observe the rate at which proteins in liquid samples bind to proteins on a sensor surface in real time. The first experimental chapter of this thesis describes the LiScAR in detail and how the technology can be used to determine the concentration of a protein in a sample in under 10 minutes.

Attempts to produce tests for Complement protein concentrations by immobilising antibodies on the sensor surface of the LiScAR proved more challenging than anticipated. Chapter 3 reports the results of testing a number of commercial antibodies, with many failing to produce working assays. The reasons behind failed assays were investigated and results point towards loss of antibody function following attachment to the sensor surfaces, as opposed to inadequate specificity or the orientation of antibodies. Fragmentation of antibodies during storage was also considered. The LiScAR was used to record the number of antigens bound by each antibody in a sample, with a theoretical limit of 1-2 (depending on antigen mass) and decreasing with fragmentation of the antibody structure. The new technique can be used for quality control of antibody samples in future work to improve experimental reproducibility. A number of new LiScAR tests worked well for samples of purified proteins but showed a lack of specificity or insufficient sensitivity when tested with human blood samples. Two new tests for the concentrations of Complement proteins C5a and C3 were validated for use with serum samples and, along with the test for antibody fragmentation mentioned previously, represent a significant contribution to the range of LiScAR tests and the larger field of blood testing.

Protein A/G (PAG) is a genetically engineered protein which binds to a common region on many antibodies and was used throughout the LiScAR studies to gauge instrument sensitivity and create antibody surfaces with directional control. PAG is known to have multiple binding sites for antibodies but a serendipitous result was recorded when mixing PAG and antibody in a known ratio: the initial reaction appeared to occur with 2 antibodies bound per PAG, contrary to the genetics literature which suggests 6 or 8. This observation prompted a detailed kinetic analysis of the reaction, reported in Chapter 4. The study used the LiScAR to rapidly sample antibody concentration after mixing with PAG, to track the progress of the reaction every 5 minutes. This work was used to learn the fundamental systems biology techniques of fitting complex models to limited experimental data.

The Complement cascade is a complex biological system and it was recognised that several variables would affect the ability of the aforementioned consumption test to identify Complement-depleted patients. These include the level of activation stimulus, the target proteins to measure and the delay between activation and measurement. It was also recognised that a mathematical model of the Complement system could be trained on a limited biological data set and then used to simulate a proposed Complement activation test for predicting the effect of Complement depletion on a test result. At the time of writing, a number of mathematical models for Complement are published in the literature but all either fail to consider all pathways of the system or lack relevant biological data for optimisation. Chapter 5 describes the construction and fitting of a new Complement system model to biological data from a pooled human serum sample following activation. The resulting model fitted well to the data and was verified to behave as expected according to the current understanding of the Complement cascade by simulating the effects of increasing the levels of proteins which are known to accelerate or inhibit activation.

The model was then used to simulate a number of potential Complement consumption tests: the concentrations of certain Complement proteins following a given level of stimulus and delay before testing. The results of the new tests were expected to vary in the healthy adult population due to natural variation in starting Complement protein concentrations in blood and this variation was modelled by repeating the test simulations 250 times with random combinations of protein concentrations within their normal ranges. The tests were then performed on a cohort of 22 adults in Chapter 6, to assess the accuracy of the model predictions and record the true variation in the new Complement consumption tests for the first time. The results suggested that the pooled serum and freshly collected serum samples displayed similar total Complement activation properties but some differences were apparent in the early reactions of the system. As such, the model failed to predict the results of some tests and was highly accurate for others, suggesting refinement of the model simulation is required. The data revealed that starting protein concentrations may not be entirely random as simulated in the model prediction. For example, individuals who had higher concentrations of a certain protein were likely to have higher concentrations of another protein.

The model was used to predict the Complement consumption test results of a Complement-depleted patient. Interestingly, there was no predicted correlation between the value of a patient’s Complement consumption test scores in healthy and depleted states. However, all tests were predicted to be sensitive to Complement depletion. Both of these predictions need to be validated with future research, performing these Complement depletion tests on a second cohort of patients who develop infections after surgery. The model and cohort data provide a first estimate of effect size for future clinical studies investigating the ability of these Complement tests to identify surgical patients at higher risk of infection, whilst the novel LiScAR tests provide a means of measuring C3 and C5 protein consumption with greater ease than existing methods. Performing these tests post-operatively might allow the onset of infection to be detected, enabling faster treatment and improved patient outcome.

Acknowledgements

I would like to thank my supervisors, Dr Chris Hyde and Dr Andrew Shaw, for their guidance throughout this endeavour and Dr Rouslan Olkhov for his detailed technical explanations of everything related to the LiScAR biosensor platform and a series of inspiring quotes which shaped my research journey. I must also acknowledge Dr Chris Scotton for providing the ethics and facilities for the clinical investigation and Dr Jacqueline Whatmore for performing all blood sampling.

A special thanks must go to my friends, you know who you are, for lifting me up when I was down and ensuring the good memories of my time at university outweigh the bad.

Finally, I would like to thank my parents for their unwavering support of my academic pursuits. It’s been a long time coming. Thank you for never doubting me.

List of Contents

List of Figures

List of Tables

List of Abbreviations

ADC	Antibody drug conjugate
aHUS	Atypical hemolytic uremic syndrome
AMD	Age-related macualr degenration
AP	Alternative Pathway
APACHE	Acute physiology, age, chronic health evaluation
AUC	Area under the curve
BNP	B-type natriuretic peptide
BSA	Bovine serum albumin
C(n)	Complement component n
C1-INH	C1 inhibitor
C4BP	C4-binding protein
C-ODE	Complement ordinary differential equation model
CP	Classical Pathway
CRP	C-reactive protein
CTAB	Cetyltrimethylammonium bromide
CV	Coefficient of variation
DAMP	Damage associated molecular pattern
DMF	Dimethylformamide
DSS	Disuccinimidyl suberate
ECDC	European Centre for Disease Prevention and Control
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbant assay
ESI	Electrospray ionisation
EWS	Early warning score
FB	Factor B
FD	Factor D
FH	Factor H
FI	Factor I
HAI	Healthcare associated infection
HF	High frequency
ICU	Intensive care unit
IL	Interleukin
IPA	Isopropanol
Ig	Immunoglobulin
k_a	Association rate constant
k_d	Dissociation rate constant
K_D	Equilibrium dissociation constant
LP	Lectin Pathway
LRT	Lower respiratory tract
LSP	Localised surface plasmon
LSPR	Localised surface plasmon resonance
LiScAR	Light scattering array reader
LoB	Limit of blank
LoD	Limit of detection
LoQ	Limit of quantification
mAb	Monoclonal antibody

MAC	Membrane attack complex
MAE	Mean absolute error
MASP	MBL-associated serine protease
MBL	Mannose binding lectin
MS	Mass spectrometry
NCEPOD	National confidential enquiry into patient outcome and death
NEWS	National early warning score
NHS	N-hydroxysuccinimide or National Health Service
NICE	National Institute for Health and Care Excellence
NIST	National Institute of Standards and Technology
NPV	Negative predictive value
ODE	Ordinary differential equation
PAG	Protein A/G
PAMP	Pathogen associated molecular pattern
PBS	Phosphate buffered saline
PCT	Procalcitonin
PPV	Positive predictive value
PRR	Pattern recognition receptor
QC	Quality control
QTOF	Quadrupole time of flight
RI	Refractive index
RIS	Refractive index sensitivity
RIU	Refractive index unit
RMSE	Root mean square error
ROC	Receiver operating characteristic
RSVF	Respiratory syncytial virus protein F
SAM	Self-assembled monolayer
sCR1	Soluble Complement receptor 1
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SOFA	Sequential organ failure assessment
SOP	Standard operating procedure
SPR	Surface plasmon resonance
TBG	Thyroxine-binding globulin
TCC	Terminal cascade complex
TLR	Toll-like receptor
TNFα	Tumour necrosis factor α
TNR	True negative rate
TPR	True positive rate
WBC	White blood cell
Z	Zymosan

1 Introduction

Nosocomial infections or healthcare associated infections (HAI) are those acquired in hospital (1), often resulting in prolonged patient hospitalisation, reoperation, readmission, and negatively impacting patient outcome (2, 3). HAI may also pose a risk to attending medical staff (4). The World Health Organisation estimates that 7% of hospitalised patients in First World countries develop a nosocomial infection (5). The pre-operative identification of patients at increased risk of post-operative complications has been identified as an effective method for guiding more effective care (6) and the early treatment of an initial post-operative complication is the best opportunity to limit patient death (failure to rescue) (7). The present study investigates *in vitro* activation of the Complement cascade in a blood sample, as a measure of immune state that might: 1) aid the identification of patients at high risk of HAI and 2) track patient recovery to rapidly reveal the onset of infection post-operatively.

The ability of a microorganism to cause disease is a function of host defence disruption, virulence and the immunologic competence of the host response (8). Invasive procedures which break the skin or enter a body cavity increase the likelihood of developing nosocomial infections (9) by providing a means of transmission for the microorganism to the host (10). Common types of nosocomial infection in patients include: central line-associated bloodstream infections, catheter-associated urinary tract infections, and ventilator-associated pneumonia (11). In Europe and the USA, the prevalence of nosocomial infections in patients admitted to the intensive care unit (ICU) is 9-37% (12) and is attributed to the immuno-compromised condition of these critically ill patients which increases their susceptibility to HAI (13). Figure 1.1 shows patient groups most at risk of HAI and the type of infections most prevalent in each group.

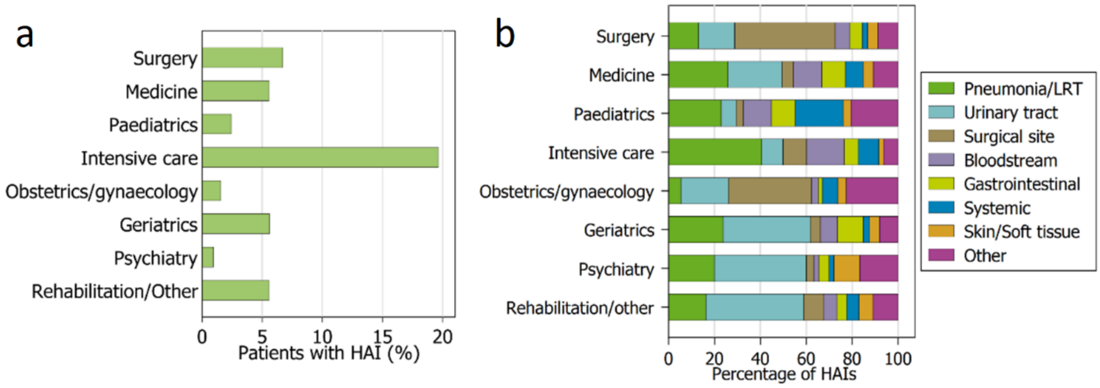


Figure 1.1 HAI statistics recorded by the European Centre for Disease Prevention and Control (ECDC). a) prevalence of HAI by patient group. b) prevalence of HAI types in each patient group. LRT: lower respiratory tract. Figure from ref (14).

The data indicate that surgical patients have the second highest risk of developing nosocomial infections after ICU patients which may be attributed to the effect of surgery on the immune system (15, 16). Over 200 million people worldwide have surgery requiring inpatient stay every year (17) and the most common type of nosocomial infection in surgical patients is surgical site infection (SSI) (3, 18, 19) predominantly caused by Gram positive bacteria (47%) and Gram negative *Enterobacteriaceae* (29%) (14). SSI rates depend largely on the type of surgical procedure and range from 0.04% for cataract surgery to 13.36% for high-risk abdominal hysterectomies (20). As with any type of infection, SSI can lead to sepsis as bacteria enter the bloodstream and the resulting inflammatory response causes damage to host tissues – a dysregulated host response (21). Sepsis is a major cause of death in surgical ICU (22, 23).

Sepsis is defined by The Society of Critical Care Medicine and the European Society of Intensive Care Medicine as life-threatening organ dysfunction caused by a dysregulated host response to infection (24). Patients with septic shock are a subset of sepsis patients which develop low blood pressure and exhibit a 40% risk of mortality (24). Each year, sepsis causes the deaths of ~31,000 patients and costs the NHS in England ~£2 billion (25). According a 2015 report by the National Confidential Enquiry into Patient Outcome And Death (NCEPOD), sepsis is first recorded within one week of surgery in the majority of cases (26) and any delays in detecting and treating infections such as sepsis increase the odds of patient mortality (27). The NHS estimates that 11,000 lives could be saved every year through better diagnosis and treatment of sepsis in the UK alone (25). Ultimately, the occurrence of post-operative complications of any kind within 30 days of surgery is the single most important determinant of long term patient survival (28) and therefore the prevention of post-operative infections that can lead to sepsis, such as HAI, is highly desirable.

The implementation of certain care pathways has been shown to reduce rates of SSI (29, 30). A care pathway is an outline of anticipated clinical care of appropriate duration and is used to aid a patient with a specific condition or set of symptoms move to positive outcomes (31, 32). Stratification of patients based on their predicted risk of post-operative infection is a patient-tailored approach to deciding the most effective care pathway to minimise risk of infection (33, 34). For example, the outcome of high-risk surgical patients may be improved by more intensive observation by medical staff (35) and pre-habilitation (personalised preparation for surgery) (36) - components of a care pathway which may be of little benefit to low-risk patients.

A determining factor of the patient’s care is the pre-operative evaluation. All patients presenting for surgery undergo pre-operative evaluation to identify risk factors that may lead to post-operative complications (infectious or non-infectious) (37). The pre-operative assessment includes a medical history and blood tests which are chosen based on the type of surgical procedure and the medical state of the patient. The National Institute for Health and Care Excellence (NICE) has published guidelines for the types of pre-operative tests recommended for different patients (38). For example, a full blood count tests the number of red blood cells, white blood cells and platelets and is not a recommended test for low-risk procedures such as excising a skin lesion but is recommended for comparatively high-risk procedures such as abdominal hysterectomy.

Within the surgical population, ~15% are considered high-risk because of certain biological characteristics such as age and comorbid disease (co-occurrence of one or more disorders) or the complexity of the surgical procedure and these patients account for the majority (~80%) of peri-operative deaths (39). Biological characteristics used to group or identify populations are called “biomarkers” or “phenotypes” (40) and an ideal identifier of high-risk patients would lead to the high-risk group accounting for closer to 100% of avoided complications. A number of biomarkers can be measured pre-operatively to assess patient health (41) such as the inflammatory markers C-reactive protein (CRP) and white blood cell (WBC) count which are routinely used to test for existing inflammatory conditions (42).

Tests for other plasma biomarkers may also be requested pre-operatively, although the Atlas of Variation in NHS Diagnostic Services in England (2017) highlights large geographical variation in the use of diagnostic tests (43). Glomerular filtration rate estimated from serum creatinine (a muscle waste product), and cardiac troponins (released by damaged heart muscle) may be used to reflect levels of pre-existing conditions of the kidneys and heart respectively, that predispose a surgical patient to non-infectious post-operative complications (39). Furthermore, the hormone B-type natriuretic peptide (BNP) regulates blood volume and raised pre-operative BNP has been show to predict 30-day cardiac death and major adverse cardiac events following vascular surgery (44).

When a post-operative infection does occur, early treatment is the best opportunity to prevent patient death (failure to rescue) (7). To facilitate early recognition and treatment of a patient’s deterioration post-operatively, patients’ observations charts are used to record measurements of vital signs and calculate some form of early warning score (EWS) (45). The current EWS used in the NHS is the National EWS (referred to as NEWS) and is used to assess the severity of acute illness based on six physiological measures: respiration rate, oxygen saturation, systolic blood pressure, pulse rate, level of consciousness and temperature (46). NICE Clinical Guideline 50, on the recognition and response to acute illness in adults in hospital states that the physiology of all patients in acute hospital settings should be measured at least every 12 hours and the frequency of monitoring should increase if abnormal physiology is detected such as the symptoms of sepsis. NCEPOD reported that 85% of patients with sepsis in 2015 were monitored using an EWS system performed at an appropriate frequency (26), suggesting improvements to the EWS would have a greater benefit to patient outcome than increased observation.

A blood culture test for the presence of bacteria in the blood is commonly used for sepsis diagnosis but may fail and requires hours to days before positivity of the culture is observed (47) – time during which a patient’s condition may worsen significantly without antibiotic intervention. Impaired tissue oxygenation occurs during sepsis, leading to an increase in serum lactate as a biproduct of anaerobic respiration (48). Lactate testing is increasingly used to assess the severity of sepsis and is a key component of the Surviving Sepsis Campaign (SSC)

Guidelines, although the majority of patients with sepsis are reportedly not tested (49). CRP correlates with mortality in critically ill patients (50) and can be used to indicate sepsis (51, 52). However, CRP does not allow discrimination of survivors from non-survivors in septic patients (53, 54). CRP is released by the liver in response to cytokines IL-1 and IL-6 to activate Complement and enhance phagocytosis (55). IL-6 and Procalcitonin (PCT) levels are significantly lower in sepsis survivors (54). CRP and PCT have become the most widely used biomarkers in the management of infection and sepsis in Europe (56).

Despite the existing measures presented here, post-operative recovery remains highly variable (57) and it is estimated that less than 30% of high-risk patients in the UK are admitted to critical care after surgery - highlighting difficulty in the timely identification high-risk patients despite EWS monitoring (58). Novel biomarkers of infection risk measured pre-operatively may guide cost-conscious approaches for patient stratification into high- and low-risk groups associated with care pathways appropriate for their level of post-operative infection risk (34). Furthermore, new and existing biomarkers might be measured post-operatively and at higher frequency in the care pathways of high-risk patients for timely identification of HAI to improve patient outcome

1.1 Biomarkers

Biomarkers are measurable indicators of medical state (59) and are used to indicate the current status of a patient or predict future risk of disease (60). Historic biomarkers included external indicators of physiological processes such as fever and swelling (61) and many tests in clinical practice are biomarkers: biochemical tests provide chemical or protein biomarkers, whereas physiological assessments provide anatomical biomarkers (62).

Biomarkers can be classified by the type of molecule and mechanism of action involved: genetic (63), epigenetic (64) or proteomic (65) for example. Alternatively biomarkers are often classified based on intended use: Types 0, 1, and 2 which may not be mutually exclusive (66, 67). Type 0 biomarkers are direct measures of disease state and correlate with clinical outcomes, such as CRP concentration for the assessment of inflammatory conditions. Type 1 biomarkers determine the biological effect of therapeutic intervention, such asHbA1c (a measure of glycated haemoglobin) used to monitor the treatment of diabetes. Type 2 biomarkers are referred to as ‘surrogate endpoints’; indirect measures of clinical endpoints such as how a patient feels or functions (68).

Surrogate endpoints are used when direct measurement of a clinical endpoint is more challenging: for example blood pressure is relatively easy to measure compared to direct measurement of heart function (with invasive echocardiography for example) (69). Surrogate biomarkers are not etiologic but are chosen based on their strong correlation with the clinical endpoint of interest and there are many examples where a favourable surrogate response to therapy is not indicative of a favourable clinical response. For example, although lowering blood pressure in patients with hypertension is a desirable surrogate endpoint to reduce risk of stroke, hypertensive patients who develop heart failure (an undesirable clinical endpoint) also show low blood pressure (69). Type 0 and Type 1 biomarkers are most appropriate for the present study as these would directly reveal pre-existing factors that influence patient risk of nosocomial infection and post-operative response to surgery or infection.

Many existing biomarkers were identified on the basis of understanding the underlying biological mechanisms of a the condition for which a biomarker is sought (a deductive method) (62). However, modern ‘Omics’ technologies (such as genomics, proteomics, metabolomics) can be used to generate thousands of molecular biomarker candidates (70) which may be compared in parallel for correlations with disease state (an inductive method) (62). Such technologies include protein microarrays and mass spectrometry, discussed later. Although high-throughput omics technologies are often used for biomarker discovery (71), the approach can suffer from the “curse of dimensionality” (72) where the value of testing additional biomarkers is disproportionately smaller than the effort required to statistically validate results (73). Hence, hypothesis-driven approaches are preferred because biological understanding of the biomarker under study (origin and function) allows better control of variables that might confound measurements of the biomarker during the validation process (74).

Ultimately, biomarkers chosen by either deductive or inductive approaches must be clinically validated by showing an acceptable discriminatory power between populations of cases and controls, as measured by a number of metrics from the contingency matrix in Table 1.1.

Table 1.1 Contingency matrix used to describe the possible outcomes of a biomarker test performed on a population of individuals with and without the condition of interest. The statistical performance metrics associated with the success rate of the biomarker are also shown.

Total population	Condition positive	Condition negative	Prevalence $= \frac{\text{condition positive}}{\text{total population}}$	Accuracy $= \frac{\text{true positive}+\text{true negative}}{\text{total population}}$
Test result positive	True positive	False positive, Type I error	Positive predictive value (PPV), precision $= \frac{\text{true positive}}{\text{test result positive}}$	
Test result negative	False negative, Type II error	True negative	Negative predictive value (NPV) $= \frac{\text{true negative}}{\text{test result negative}}$	
	True positive rate (TPR), recall, sensitivity $= \frac{\text{true positive}}{\text{condition positive}}$	True negative rate (TNR), specificity, selectivity $= \frac{\text{true negative}}{\text{condition negative}}$		

Sensitivity describes the probability of obtaining a positive test result when the condition is positive but the metric favours a test that always reports a positive result regardless of condition (ignores false positives). In contrast, PPV describes the probability of a given positive test result occurring due to a true condition but the metric favours a test which fails to identify many condition positive cases (ignores false negatives). Only by considering both sensitivity and specificity or PPV and NPV can a thorough understanding of biomarker performance be obtained, although when the consequences of Type I error are low, a test with higher sensitivity is preferred.

Results of biomarker testing are often described using a receiver operating characteristic (ROC) curve, in which the sensitivity (TPR) is plotted as a function of 1-specificity at different decision threshold values of the biomarker. A test decision made due to chance alone would display as a straight line with a gradient of 1 and an area under the curve (AUC) of 0.5. In contrast, an ideal biomarker would have ROC curve displaying 100% sensitivity (no false negatives) and 100% specificity (no false positives) with an area of 1. The Yuden Index, J, provides a single statistic which considers both sensitivity and specificity, J = sensitivity + specificity -1, the maximum vertical distance between the ROC curves of the biomarker and chance alone (75).

The ROC curve assessment allows the diagnostic performance of multiple biomarkers to be compared - a useful tool considering the plethora of published biomarkers. In 2017, a search of the PubMed database for publication titles containing the term “biomarker” revealed 3581 articles published that year and more than 100 different molecules have been suggested as useful biomarkers of sepsis (76). The increasing attention to biomarker research is evident in the growth in the number of yearly biomarker publications, Figure 1.2.

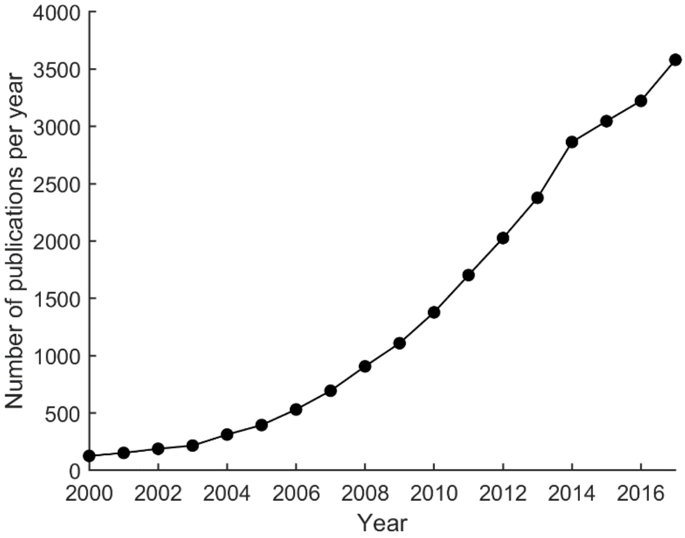


Figure 1.2 Number of publications per year returned by a PubMed query for publication titles containing the word “biomarker”, showing an increasing trend.

It has been estimated that at ~100,000 different biomarkers have been described in the literature but the number of biomarkers applied in clinical practice is significantly lower (77) and only 4 (9.1%) of 44 original proteomic biomarker tests published in 2006 made progress toward clinical application after 10 years (78).

The apparent bottleneck between biomarker discovery and adoption in clinical practice might be attributed to the challenges of biomarker validation (77). A clinical-grade assay must be available for the biomarker to be adopted in clinical practice (70) and biomarker validation testing is comprised of several steps (79): 1) assessment of assay performance characteristics (analytical validation); 2) assessment of assay performance regarding correlation with clinical endpoint of interest (clinical validation); and 3) assessment of assay impact on clinical outcome (clinical utility). A new biomarker must therefore reliably indicate a clinical endpoint and provide demonstrable benefit to the patient, either through improved patient management decision making or improved patient experience compared to existing practice (80).

Clinical validation studies can fail to show acceptable biomarker performance for a number of reasons (77). Firstly, discovery studies can overestimate the effect size of a given biomarker as a result of poor selection of case and control patients which should be matched for age, gender and other diseases. The use of different assays with different performance characteristics is also a challenge for biomarker validation in different laboratories. Consequently, follow-up studies in other laboratories may fail to replicate results, particularly if the function of the biomarker is incompletely understood. Poorly defined intended use of the biomarker can also lead to poor selection of case and control patients and new candidate biomarkers should be proposed based on their ability to support a well-defined clinical decision.

As such, a substantial effort is required to demonstrate the effectiveness of a biomarker. It has been suggested that the penetration of new biomarkers into clinical practice would be improved if discovery and validation studies considered the common reasons for biomarker validation failure discussed above and involved a single optimised assay used across multiple clinical studies (77). Furthermore, the ability of individual new biomarkers to outperform existing tests is not necessarily required, because the combination of several test results may provide more discriminatory power.

1.1.1 Biomarker Panels

For clinical conditions with multiple etiologies, such as a compromised immune system, combining the results from a panel of biomarkers may permit a better characterisation of patient condition. For example, multiple risk factors for cardiovascular disease chosen based on biological insight have been combined into equations which better predict the risk of developing heart conditions than any individual risk factor alone (81). A pre-hospital NEWS value (46) is associated with death within 48 hours of hospital admission (82). A further example is the Acute Physiology and Chronic Health Evaluation (APACHE) score, which is calculated at the beginning of admission to ICU to predict a patient’s mortality risk using a number of biomarkers such as WBC count, creatinine, sodium and potassium concentrations (83). Similarly, Apgar scores are used to track the health of newborn babies, using variables such as skin tone, heart rate and reflex irritability (84) and there are multiple other examples of scoring systems based on panels of biomarkers to maximise predictive value of patient outcome (85, 86). The severity of organ failure in sepsis may be quantified by a panel of tests including creatinine and bilirubin which make up the Sequential [Sepsis-Related] Organ Failure Assessment (SOFA) score (87). The change in SOFA score over time (ΔSOFA) has been demonstrated to evaluate treatment progress (88), effectively converting type 0 to type 1 biomarkers. A combination of SOFA score and additional biomarkers has been shown to predict the development of acute kidney injury in sepsis (89), highlighting the utility of combining existing scores with new biomarkers.

The use of biomarker panels may reduce the risk of incorrectly assessing a clinical endpoint compared to a single type 2 biomarker. For example, the APACHE score uses the worst values recorded for each biomarker in the first 24 hours after admission to the ICU (90), meaning that an underlying condition which confounds the result of one test (such as a surrogate endpoint) may still be detectable based on the other biomarkers which contribute to the total score. However, the choice of biomarkers to include in a panel is critical, as inclusion of insensitive markers can confound the predictions of the panel, particularly when a single value or score is calculated. The serum protein profiles of pigs have been screened with mass spectrometry to obtain large numbers of markers which show potential for the detection and identification of infections before the onset of symptoms, although specificity is low and could be attributed to the variation in irrelevant markers among individuals (91).

Biomarker candidates insensitive to the condition of interest can be avoided by deductively selecting type 0 biomarkers from the physiological system directly affected by said condition. Therefore, as dysfunction of the immune system can contribute to opportunistic disease (92), type 0 biomarkers indicating the efficacy with which the immune system protects the host can be deductively rationalised as candidates for a biomarker panel designed to predict nosocomial infection risk following surgery. The study in ref (34) supports the choice of immune system biomarkers and suggests that that pre-surgical immune state predicts the rate of post-operative recovery - potentially enabling patient stratification based on a predicted recovery profile. Patient immune state might also be monitored following surgery as a form of EWS to detect the onset of infection at the bedside

(93). Understanding the host response to infection offers insights into hypothesis-driven selection of candidate biomarker panels to predict the risk of nosocomial infection in surgical patients (94). An overview of the host immune response will now be presented.

1.2 The Host Immune Response

The host immune response is the defence mechanism which protects the host against invading pathogens and promotes wound healing and tissue repair (95). The immune system also responds to other immune challenges including medical devices and surgical trauma. The immune system may be divided into two components, each with a distinct protective role: the adaptive immune system provides highly specific host-protection from pathogens and is essential for an efficient immune response (96) but requires 3-5 days to respond to a previously unencountered pathogen (92) and hence the innate immune system is required as a first line of defence against infection (97). A critical attribute of both the innate and adaptive immune systems is the ability to distinguish self from non-self material.

The appearance of non-self biological patterns associated with pathogens (PAMPs) (98) is recognised rapidly by the innate immune system. Examples of PAMPs include bacterial peptidoglycan, lipopolysaccharide (LPS) (99) and surface flagellin proteins (100). Double stranded RNA is associated with viral infection (101), whilst complex carbohydrates such as zymosan and chitin are recognised from the cell walls of fungi (102). Parasitic pathogens are also identified by PAMPs in the form of glycolipids (103). PAMPs are bound by pattern recognition receptors (PRRs) of the innate immune system which, unlike the adaptive immune system, are genetically inherited rather than acquired. PRRs may be soluble or expressed on the surface of host cells such as phagocytes, which engulf and destroy pathogens, a process called phagocytosis. The Complement cascade represents an important source of soluble PRRs (104, 105), which directly induces cell lysis, opsonises pathogen surfaces to stimulate phagocytosis and recruits immune cells to the site of infection via chemoattraction. There are several families of cellular PRRs which differ in cellular location, structure and signal transduction. Cellular PRRs are located both inside the cell to detect intracellular pathogens such as viruses and on the external cell membrane to continuously monitor for PAMPs in extracellular fluids.

Toll-like receptors (TLRs) are a major PRR expressed on the surfaces of WBC's such as phagocytic cells. The variety of PRRs ensures the innate immune system responds rapidly to infection of the host by a range of pathogens.

The innate immune system can also be activated by trauma (accidental or surgical), which is accompanied by the extracellular release of damage associated molecular patterns (DAMPs) (106) also known as 'alarmins' (107). DAMPs include cellular components normally confined within the boundary of the cell membrane such as mitochondrial DNA (mtDNA) (108), high mobility group box-1 (HBGb1) (109), Cold-inducible RNA binding protein (CIRB) (110) and F-actin (111). Formyl peptides (112) may also be released by dying cells or may be secreted by pathogens. The innate immune system responds rapidly to the presence of both DAMPs and PAMPs by inducing inflammation. At the site of infection or injury, the Complement cascade releases inflammatory mediators whilst phagocytic cells release cytokines, cell signalling molecules which induce behavioural changes in nearby immune cells and attract further phagocytes to the site of infection. Cytokines also induce local vasodilation and increased vascular permeability, which is observed as redness, heat and swelling. Anti-inflammatory mediators are also released following inflammation in an attempt to restore homeostasis, as excessive inflammation causes damage to host-tissues (113). Impairment of the innate immune response is associated with increased susceptibility to infections, as is observed following sepsis (114), and hence biomarkers assessing innate immune system function could provide a measure of infection risk.

The adaptive immune system can be activated by the innate immune system and is essential for an efficient immune response (96). The adaptive immune system is comprised of two parts: humoral and cellular. The cellular immune response clears infected cells, whilst the humoral immune response clears circulating non-self material. In the humoral immune response, soluble immunoglobulin proteins (antibodies) bind to the surfaces of extracellular foreign molecules, marking them for destruction by WBCs. Antibodies are secreted by plasma cells and bind specifically to a small region (the epitope) on the surface of antigens. Antibodies contribute to host defence by binding to toxins and preventing their interaction with host cells. Antibodies opsonise pathogens for phagocytosis and activate the Complement cascade of the innate immune system. Critically, with the exception of natural antibodies (115), a previous antigenic challenge is required in order to express antibodies to a pathogen via the immunological memory of the adaptive immune system. Impairment of the cellular or humoral immune responses increases patient susceptibility to opportunistic infections, as is seen in acquired immunodeficiency syndrome (AIDS) when human immunodeficiency virus (HIV) infects and kills WBCs by inducing apoptosis (116).

Both innate and adaptive immune responses provide host-resistance to disease that is critical to survival. However, it is the innate immune system that provides the first line of defence against microbial pathogens (117) and biomarkers assessing the function of the innate immune response might therefore be used to predict risk of infection. Preoperative inflammation is known to correlate with postoperative infectious complications in surgical patients (118-120) and there is evidence to suggest that modulating preoperative inflammation could decrease the risk post-operative infectious complications (121). The suitability of the components from the innate immune response as candidate biomarkers of infection risk will now be discussed.

1.2.1 Candidate Immune System Biomarkers

The overview of the host immune response presented here reveals several immune system components which might be measured as biomarkers of a weakened immune system and might correlate with nosocomial infection risk or reveal onset of infection. To aid in the selection of candidate immune system biomarkers, each candidate can be compared to the characteristics of an ideal biomarker (122, 123):

1. specific to the condition of interest
2. sensitive to condition severity, correlates with prognosis
3. recovers to a baseline following effective therapy for “biomonitoring”
4. well defined intended use for “biomarker-guided therapy” decision making

In addition to these characteristics, the ease of biomarker testing should be considered. Less-invasive tests requiring smaller sample volumes are preferred in order to minimise patient discomfort.

The intended use of biomarker candidates in the present study can be clearly defined:

> The biomarker would be measured pre-operatively to determine the personal baseline of the patient and compared to the reference range of the biomarker in healthy adults at typical risk of HAI. A biomarker value outside of the reference range might indicate a high-risk patient and enable the patient to be stratified into an alternative care pathway, designed to improve the outcome of high-risk patients.

The biomarker might also be tested post-operatively as an EWS to profile patient recovery. Items 2 and 3 on the list of ideal biomarker characteristics require the biomarker to respond dynamically to a condition, ideally tracking condition severity with minimal lag as to avoid delays in clinical decision making.

Elevated cytokine concentrations indicate systemic inflammation and are attractive candidate biomarkers which might indicate patient risk of infection due to a compromised immune state. Systemic concentration variations in cytokines are observed after tissue injury (trauma or surgical) and infection (124), allowing cytokine testing to be performed on a variety of fluid samples such as serum (125) and tears (126), requiring minimal invasiveness. Furthermore, testing of cytokines can be performed using existing multiplex assays (such as Luminex bead-based assays) using small sample volumes (124). The human endotoxin model, an *in vivo* model of systemic inflammation in which LPS (a TLR agonist) is injected or infused intravenously in healthy volunteers, mimics the acute inflammatory response in sepsis and other acute inflammatory conditions (trauma, burns, surgery) (127). Within the first hour after LPS administration, pro-inflammatory cytokines TNFα, IFN and IL-6 appear in plasma and reach peak concentrations after 90, 120 and 120 minutes respectively (127). Anti-inflammatory cytokines IL-10 and IL-8 reach their maxima after approximately 3 hours and later (127). A continuous LPS infusion instead of a bolus intravenous injection extends the time to maximum in the cytokine response profiles (128), and the cytokine kinetics are dose dependent – markers of condition severity. Cytokines therefore display many characteristics of an ideal biomarker but some shortfalls are evident.

Firstly, the human endotoxin model is used as a direct measure of the cytokine response to infection but could not be used as a biomarker of pre-operative immune state because 1) the procedure causes significant discomfort to the subject and 2) immune state of the patient is altered as a consequence of the measurement – the observer effect. Instead, the standing concentrations of interleukins 6,8 and 10 as well as TNFα show greater promise for the diagnosis of inflammatory conditions such as sepsis (129). The persistence of TNFα and IL-6 in the serum following hypotension correlates with poor outcome (multiple system organ failure and death) in septic and non-septic, critically ill patients (130). IL-6 in synovial fluid has been demonstrated to identify periprosthetic joint infection after shoulder surgery with an optimal cut-off of 359.3 pg/mL (131) and statistically significant differences between IL-1 and TNFα have been used to identify patients with SSI following neck dissection (132). A cut-off value of an IL-6 level above 300 pg/mL in SIRS patients has been shown to correlate with larger risks of complications, including pneumonia and death (133). Serum cytokine cut-offs have also been demonstrated to identify neonatal sepsis (31 pg/mL IL-6, 17 pg/mL TNF-α and 1 pg/mL IL-1β) with accuracy improving when multiple markers were combined (134) – further highlighting the advantage of biomarker panels discussed previously. However, a more recent study reported an optimal cut-off value of 18 pg/mL for IL-6 and 10 pg/mL for CRP best identified neonatal sepsis (135) and yet another study derived a cut-off value of 10.85 pg/mL IL-6 (136), suggesting absolute cut-off values of cytokines are difficult to determine. The significant variation among individuals and the variety of factors influencing cytokine release such as menstruation (137) has hindered the definition on a ‘normal’ cytokine profile (129) (138).

Whilst the standing concentrations of cytokines may perform well for indicating an underlying condition, these biomarkers are released by immune cells only after the onset of a clinical condition or complication. Therefore, the standing concentrations of cytokines in healthy individuals are unlikely to indicate how effectively immune cells will respond to future infection. In contrast, antibody dependent cellular cytotoxicity by neutrophils is largely dependent on genetic variations in Fc receptor expression among patients (139) which influences the sensitivity of these cells. Testing for genetic biomarkers is becoming increasingly affordable with whole genome sequencing available as a service (140). However, a study of monozygotic twin pairs highlighted that whilst genetic testing is effective for identifying abnormal genes, the predictive value of whole genome sequencing is low for common conditions such as coronary heart disease, cancer, and stroke (141). Furthermore, DNA testing cannot reveal physical properties of an organism which vary over time due to behaviour and environmental effects on the immune system. Instead, epigenetic studies measuring changes to gene transcription as a direct result of the immune response are indicators of immune state and can be performed *in vitro* using mass cytometry to isolate immune cell types from patient blood.

Elevated transcription of the IL-1β gene (142, 143) can be used to monitor neutrophil priming, in which a priming agent such as TNFα induces enhanced phagocytosis and a more robust immune response to further stimulus. Correlations have also been found between signalling molecules controlling transcription of DNA in monocytes of post-operative patients and the speed of patient convalescence (144), suggesting that gene transcription in the cellular immune response to surgery could be used to predict recovery. Furthermore, approximately 10% of individuals infected with *Mycobacterium tuberculosis* develop active tuberculosis disease (TB) (145) and genetic testing of dendritic cells challenged with *Mycobacterium tuberculosis in vitro* has revealed differences in gene expression between resistant individuals diagnosed with latent TB and susceptible individuals that recovered from active TB (146). However, changes to gene transcription during sepsis are highly variable (147) and functional testing of a patient's immune cells to an *in vitro* challenge is a more direct test of a patient's potential ability to clear a future infection.

Functional testing of neutrophil performance characteristics *in vitro* such as phagocytic activity has been shown to predict the risk of infection, organ failure and mortality observed in patients with cirrhosis (148) and alcoholic hepatitis (149) (diseases associated with an increased incidence of microbial infection). Furthermore, reduced phagocytosis upon admission is a negative predictor of survival of patients with sepsis (150). Ultimately however, *in vitro* testing of cells does not capture significant effector functions of the cellular immune system such as vascular permeability and does not account for external impacts on cellular activity such as hormones. In contrast, the Complement system is almost entirely confined to the blood and is largely independent of external influences such as cellular activity. As such, the whole Complement system can be observed in serum and can be stimulated *in vitro* under conditions which closely approximate those *in vivo*.

The Complement cascade is a system of proteins which circulate the blood as inactive precursors and become activated upon recognition of PAMPs and DAMPs (151). Complement activation results in the consumption of circulating proteins after 30 minutes of activation *in vitro* (152), releasing protein fragments that opsonise pathogen surfaces or directly induce pathogen lysis. Measurement of Complement zymogens and activation products can be readily achieved using low volume serum samples and existing immunoassays. For a given stimulus, the dynamics of the Complement system depend on the relative concentrations of its interacting proteins according to the laws of enzyme kinetics and hence Complement activation dynamics should be sensitive to Complement depletion as is observed following sepsis (153). Complement therefore represents an ideal system of potential biomarkers which are fundamentally sensitive to infection and recover to a baseline following effective therapy - type 1 and type 2 biomarkers. Patients with a deficiency in one or more Complement proteins, due to an underlying inflammatory or genetic condition, could be at higher risk of infection (154, 155) and might be identified with an *in vitro* functional test of Complement activation dynamics – new risk biomarkers.

1.3 The Complement Cascade

The Complement system is primarily viewed as a first responder to microbial infections (156) and is comprised of plasma proteins which mostly circulate as inactive zymogens (157, 158). In addition to the protective role of Complement in innate immunity, it is increasingly apparent that the system is also responsible for the modulation of several complex tissue regeneration processes including: removal of obsolete synapses in the brain (159); inhibition of angiogenesis (160); mobilization of hematopoietic stem/progenitor cells during transplantation (161) and lipid metabolism (162). There is also crosstalk between Complement and the coagulation cascade (163, 164). For example, kallikrein cleaves C3, C5 and factor B (FB). Liver cells (hepatocytes) are the major source of Complement components in plasma but many other cell types synthesise Complement proteins (165). A specific example of extrahepatic biosynthesis of Complement is in the central nervous system, which utilises Complement for host immunity and obsolete synapse removal but is surrounded by the blood-brain or blood-spinal cord barrier and shielded from Complement proteins in the plasma (166). Whilst the diverse roles of Complement *in vivo* are still emerging, its mechanisms of activation as part of the immune response have been studied extensively since its discovery in the 19th century and are now well established (167).

Excluding extrinsic activation via crosstalk with the coagulation cascade, Complement activation occurs primarily via one or all of three possible initiation pathways: the Classical pathway (CP), Lectin pathway (LP) and Alternative pathway (AP) which lead to activation of the Terminal pathway (TP), Figure 1.3.

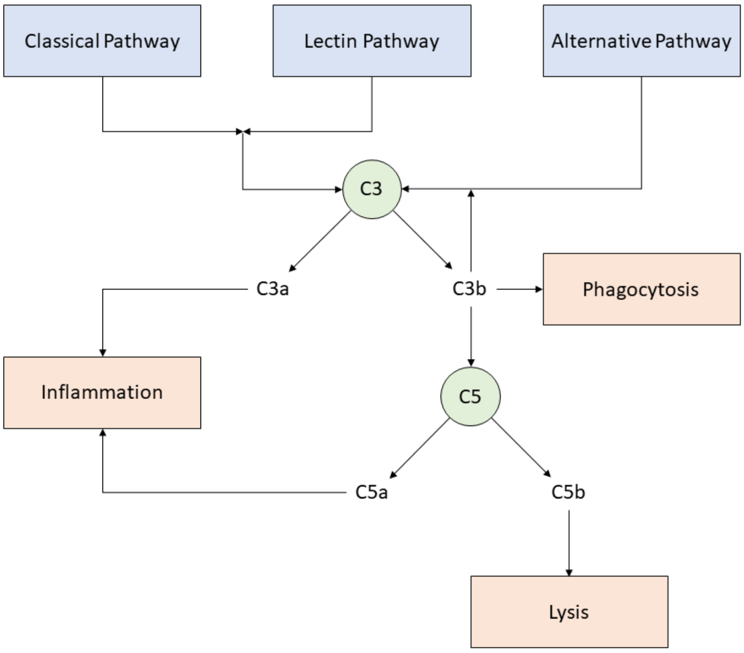


Figure 1.3 Simplified diagrammatic model of the Complement cascade. All three initiation pathways cleave C3 into the fragments C3a and C3b: inducing phagocytosis and inflammation. C3b further activates the AP in an amplification loop and initiates the TP which cleaves C5. C5a contributes to further inflammation whilst C5b results in the formation of the membrane attack complex to induce pathogen membrane lysis.

All three activation pathways converge with the cleavage of the proteins C3 and C5 into their constituent fragments. C3a and C5a are anaphylatoxins which attract and activate WBCs by binding to G-protein-coupled receptors (C3aR and C5aR) (156). C3b from the upper cascade is an opsonin (168) which covalently binds to pathogen surfaces via a thioester domain (169) and marks them for destruction by interacting with CR1 on phagocytic cells (156). C5b from the lower cascade is the first building block of the membrane attack complex (MAC), otherwise known as the terminal cascade complex (TCC) which forms pores pathogen membranes and induces lysis (156). When formed in solution, TCC is often referred to as sC5b-9. Anaphylatoxins, opsonins and MAC are important contributors to host defence against infection (156) and acquired or inherited Complement deficiency leads to increased risk of infection (154, 155). The Complement pathways will now be discussed in more detail.

1.3.1 Classical and Lectin Pathways

The CP and LP can be activated by specific pattern recognition molecules on exogenous material surfaces. C1q of the CP binds to CRP as well as immune complexes containing IgG and IgM (170), complementing the adaptive immune system (171). C1q also recognises PAMPs such as LPS (172) and non-immune self-ligands (173) such as membrane protrusions on apoptotic cells to aid in the clearance of dying cells (174). The equivalent pattern recognition molecule of the LP is mannose-binding lectin (MBL), which recognises sugars (including glucose, fructose and mannose) on pathogen surfaces (157, 175). Stable binding of C1q or MBL is achieved only when their ligands are clustered on a surface in a specific pattern.

Serine proteases are associated with C1q (C1r and C1s) and MBL (MASP1 and MASP2) in a Ca^{2+} dependent manner and hence the addition of chelating agents such as ethylenediaminetetraacetic acid (EDTA) to serum samples is used to prevent CP or LP activation (157). The serine proteases are zymogens which become activated by a conformational change in the structure of C1q or MBL following binding. C1q binding to immune complexes leads to the activation of C1r which in turn activates C1s (176) and similarly, upon MBL binding to its substrates, MASP1 activates MASP2. A notable difference between the two activation events is that a $C1r_2C1s_2$ tetramer is always associated with C1q, whilst 70% of MBL in blood is associated with a single homodimer of either MASP1 or MASP2 (177). The separation of MASP homodimers in different MBL molecules ensures that LP activation occurs only by colocalization and cross-activation of MBL containing different MASPs (178). Greater deposition of MBL therefore increases LP activation. MASP separation may also ensure the lack of spontaneous activation of the LP. Although the mechanism is unclear, there is evidence for spontaneous activity in the CP (179) in which its serine proteases are not separated.

C1s and MASP2 cleave C4 into fragments C4a and C4b which can bind locally via its thioester domain to immune complexes or cell surfaces (180). The locally bound C4b then binds to C2 which is cleaved in-situ by C1s, MASP1 or MASP2 to form a C3 convertase and a fragment of C2 (181). The nomenclature of the C2 fragments is inconsistent in the literature and should be addressed for clarity. Originally, C2a was used to indicate the fragment representing the activated form of the C2 proenzyme and proteolytic subunit of the C3 convertase (182). It was later discovered that this C2 fragment incorporated into the CP/LP C3 convertase is 70 kDa, whilst the other fragment is 30 kDa and the naming is inconsistent with the convention of denoting the smaller of two fragments with the earlier letter. The Complement Nomenclature Committee have published recommended names for many proteins (183) which will be used here, but C2b has not been universally adopted by the Complement community to denote the larger fragment of C2. The C3 convertase of the CP/LP is therefore referred to here as C4bC2a in agreement with the original work and the larger Complement community.

1.3.2 Alternative Pathway

Under normal physiological conditions, the AP is continuously active at a relatively low level, due to a process known as tick-over (184). Tick-over is the spontaneous hydrolysis of a labile thioester bond within C3 at a rate of 1×10^{-4} μM^{-1} min^{-1} (185) to produce small quantities of C3(H_2O). Gas bubbles and biomaterial surfaces are capable of increasing the rate of C3(H_2O) formation (186) such that tick-over is sensitive to different levels of stimulus. The structural change associated with the formation of C3(H_2O) exposes a binding site for factor B (FB), which binds in a Mg^{2+}-dependent manner. FB in the C3(H_2O)FB complex may be cleaved into fragments Ba and Bb by the serine protease factor D (FD). The result of tick-over is the continuous generation of the enzymatically active protein complex C3(H_2O)Bb, a C3 convertase with a half-life of 77 seconds (187) which cleaves C3 to produce low standing concentrations of C3a and C3b. The thioester domain of C3b enables the protein to opsonise any surface with exposed hydroxyl groups. However, the C3b thioester is relatively short-lived, with a half-life of 60 μs, ensuring that C3b binding occurs locally, within 60 nm of the C3 cleavage event (157). The low, sentinel concentration of short-lived C3b opsonin assures AP activation on any surface with exposed hydroxyl or amine residues. Notably, C3b has variable affinity for the hydroxyl groups of sugars (168), leading to variable AP activation by different surfaces. The term ‘Complement compatibility’, introduced by Tom Mollnes, describes the degree to which a material surface activates the complement system (188).

The AP is distinct from the CP and LP in that it lacks a well-defined recognition molecule. Instead, the AP serves to amplify any C3b produced by tick-over or the CP and LP. C3b typically binds in a Mg^{2+} dependent manner to FB, forming the C3 convertase C3bBb (189). There is evidence to suggest that FB binds to C3b via either Ba or Bb portions however only C3bBb and not C3bBa has the ability to cleave C3 (190), with consequences for negative feedback in the AP. The cleavage of C3 by C3bBb produces further C3b to form additional C3bBb: a positive feedback loop analogous to the chain reaction of neutrons in nuclear fission, Figure 1.4.

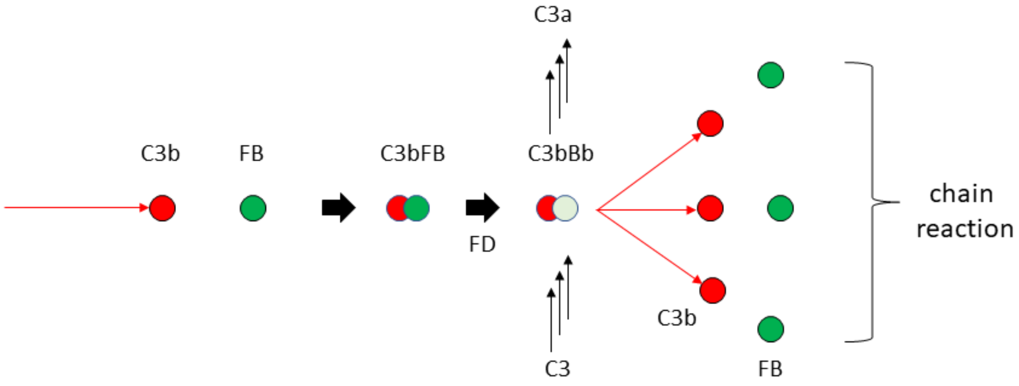


Figure 1.4 The amplification loop of the AP. C3b binds to FB, which in the presence of FD produces a C3 convertase capable of producing more C3b.

The AP is further enhanced by positive feedback using Properdin (P) (191), a protein released from WBCs during inflammation in response to C5a (192) and the only known positive regulator in the Complement system. P binds to the AP C3 convertase and increases its half-life by up to 10 fold (193), leading to enhanced AP amplification and increased C5a production and neutrophil activation. Surface plasmon resonance studies have shown that P can bind multiple C3b molecules (194) and provide a platform with additional, stabilised sites for C3 convertase assembly (195). Furthermore, there is evidence to suggest P is capable of binding to specific target surfaces including bacteria (196), providing the possibility of C3 convertase assembly being C3b-directed or P-directed (197). The role of P as a pattern recognition molecule is controversial as studies have shown that P binding to zymosan and *E. coli* is dependent on initial C3b deposition (198).

The formation of the C3 convertases by pattern recognition molecules sensitive to different molecular patterns provides the opportunity for differential activation of Complement according to the activating exogenous surface. For example, there is experimental evidence that non-encapsulated *Crypotococci* induce C3 cleavage rapidly via IgG binding and activation of the CP, followed by the AP with a lag of order minutes (199). In contrast, encapsulated *Cryptococci* have been shown to activate exclusively the AP, characterised by a 4-6 minute delay in C3 deposition following exposure to serum, attributed to the asynchronous and focal deposition of C3b (199). Measuring the dynamics of Complement activation might provide a means to identify the causative pathogen in sepsis – a differential diagnosis currently achieved using blood cultures which require hours to days of incubation and can fail. Irrespective of initiation pathway, Complement activation ultimately leads to activation of the TP.

1.3.3 Terminal Pathway

The bioactive C3b fragment is versatile. In addition to binding FB and hydroxyl groups on surfaces, C3b also binds to C4b and C3b of the CP/LP and AP C3 convertases. The resulting complexes, (C4bC2a)C3b and (C3bBb)C3b are serine proteases which lose the ability to cleave C3 and gain the ability to cleave C5 (157, 200) into its constituent fragments: C5a and C5b. The molecular details of C5 convertase formation remain poorly understood although experiments have shown that high C3b densities are required for C5 convertase formation on a surface (201). The dynamic equilibrium between C3 convertase and C5 convertase formation due to competition for C3b has been observed as a flux control point of Complement activation, in which the upper cascade appears to reach a threshold concentration of C3b before triggering the TP (152). C5 cleavage is the last enzymatic step in Complement activation and C5b sequentially binds to complement components C6, C7, C8 and C9 to form the cytotoxic MAC capable of lysing cell membranes (202).

The Complement system responds rapidly to a variety of stimuli via the three initiation pathways and amplification loop providing positive feedback. However, Complement is not cell-specific and activation may induce opsonisation and damage of host tissues unless regulated (203). The amplification of the stimulus must be controlled to prevent excessive activation causing harm to the host as is seen in systemic inflammatory response syndrome (SIRS) as a result of trauma (accidental or surgical) or infection (204-206).

1.3.4 Complement Regulation

Complement activation has been described as a 'double edged sword', as whilst the system is effective at protecting the host from infection, excessive activation in otherwise healthy individuals causes severe injury to host tissues. In addition to the rapid half-lives of many activated intermediates, the Complement system is regulated at multiple points to prevent excessive activation (157). Host cells are protected from Complement activation by expressing a collection of membrane regulators. These include CR1, membrane cofactor protein (MCP) and decay accelerating factor (DAF) which deactivate C3 and C4 convertases. In addition, MAC inhibitory protein (CD59) prevents MAC formation by blocking the binding of C9 to C8. Expression of cell surface Complement regulators is not always desired and confers protection of tumour cells against Complement-induced lysis (207).

Understanding Complement regulation *in vitro* is important to rationalise the choice of biomarkers. Soluble regulators of Complement include vitronectin (Vn) which prevents soluble sC5b-7 membrane insertion (208) and clusterin (Cl) which also prevents membrane insertion of TCC and its precursors (209). Furthermore, plasma carboxypeptidase enzymes cleave the C-terminal arginine from C3a to produce C3adesArg (210) which cannot bind to C3aR and activate WBC's (156). The same enzymatic cleavage occurs on C5a to produce C5adesArg which shows reduced affinity for C5aR. Assays for such Complement activation products should ensure the 'desArg' forms of these proteins are also measured.

Complement regulation significantly perturbs activation of the initiation pathways. In the CP, an average of 4 C2 molecules and 35 C4 molecules are cleaved per activated C1 complex due to rapid inhibition by the serpin, C1-inhibitor (C1-INH) (211). C1-INH also regulates MBL activation of the LP (212). C4 binding protein (C4BP) accelerates the natural decay of C4bC2a by binding to C4b and encouraging the dissociation of C2a (213). Subsequently, in the presence of C4BP in the fluid phase, the serine protease factor I (FI) cleaves C4b into inactive fragments C4c and C4d. The soluble form of Complement receptor type 1 (sCR1) (214, 215) also induces dissociation of C4bC2a and serves as a cofactor for the FI-mediated proteolysis of C4b. sCR1 has more diverse functions than C4BP and also plays a role in AP deactivation. In the presence of sCR1, FI cleaves C3b into iC3b which is unable to bind FB, preventing C3 convertase formation (193). However, the master regulator of the AP is factor H (FH), which directly blocks the binding of FB to C3b (216). FH also accelerates AP C3 convertase dissociation and acts as a cofactor for FI in deactivation of C3b. Critically, FH is able to bind C3(H_2O) to regulate both tick-over and AP C3 convertase formation but FH cannot bind to C3 as the binding site only becomes exposed after hydrolysis or cleavage. The impact of regulatory proteins on Complement activation is significant and deactivated proteins such as iC3b may provide sensitive measures of Complement activation in addition to their activated forms. The end products of regulation pathways in particular, may be good biomarkers of Complement activity due to their fundamentally long half-lives – these will not be consumed by further reactions.

Continual dysregulation of Complement activation can contribute to chronic inflammatory and degenerative conditions (217) including: systemic lupus erythematosus (SLE) (218), age-related macular degeneration (AMD) (219), atypical hemolytic uremic syndrome (aHUS) (219) and paroxysmal nocturnal hemoglobinuria (PNH) (220). A number of therapeutic inhibitors designed to control Complement activation are being developed (221, 222). Berinert (CSL Behring), Ruconest (Salix Pharmaceuticals) and Cinryze (Shire Pharmaceuticals) are preparations of C1-INH approved for treating hereditary angioedema (severe swelling) due to insufficient functional C1-INH (222). Lampalizumab (Roche/Genentech) is a monoclonal antibody (mAb) designed to inhibit FD activity, although recent phase III clinical trials in AMD patients showed no benefit (223). Cp40 (Amyndas), APL-1 and APL-2 (Apellis Pharmaceuticals) are preparations of Compstatin, a synthetic peptide which sterically hinders the binding of C3 to convertases (224); Comptstatin has been shown to improve the outcome of trauma-induced hemorrhagic shock in non-human primates (225).

Reduced Complement activation is sought during shock and certain autoimmune diseases (226). However, long-term suppression of the complete Complement system may increase risk of infection and additional antimicrobial prophylaxis is used to counter this risk (224). C5 is an attractive target for therapeutic inhibitors as negative regulation of the Complement system at this point preserves the chemoattractant and opsonisation functions of C3a and C3b, whilst preventing

C5 cleavage and the formation of TCC. Eculizumab (Soliris; Alexion Pharmaceuticals) (227) is a humanized monoclonal antibody that inhibits C5 binding and is approved for the treatment of PNH and aHUS. Despite a plethora of targets for inhibition, development of Complement drugs has been considered relatively slow (217), possibly due to the complexity of Complement activation and its diversity of functions which may confound *in vivo* studies. The development of therapeutics for Complement regulation may be aided by a systems level understanding of Complement activation, discussed later.

Complement activation is an important factor for effective clearance of infection and results in the consumption of circulating proteins. However, C3 depletion to a cut-off value of <0.578 mg/mL during sepsis strongly correlates with increased mortality (sensitivity 78.4% and specificity 99.8%) (153), exemplifying how Complement activation can be both protective and harmful to the host. C3 depletion relative to a personal healthy value pre-operatively (percentage consumption) might be used to better assess Complement activation in individuals with personal healthy protein concentrations at the extremes of the normal ranges. For example, a patient with a personal healthy C3 value in the 90th percentile would exhibit greater Complement activation, compared to a patient in the 10th percentile for C3 concentration, before achieving a given absolute cut-off value for C3 concentration. Patients considered to be Complement-depleted relative to their personal healthy values could be provided with antimicrobial prophylaxis if considered to be at a high risk of infection. To sustain C3 levels, exogenous supplementation of C3 has been demonstrated to enhance bacterial clearance and improve survival in septic mice (228). Sepsis survivors exhibit a higher mortality risk compared with age-matched non-septic critical care survivors (229) and individuals with chronic medical conditions are also at increased risk of future sepsis events (230). A recent review of Complement polymorphisms describes combinations of risk alleles whose aggregate effect on disease risk is significant: a ‘Complotype’ (231). These potentially high-risk patients are ideal candidates for stratification into the high-risk care pathway described previously.

The Complement response of a patient is a phenotype and likely an important factor in determining the course and severity of HAI (203). A Complement test might be used to identify patients with atypical Complement activation phenotypes which might indicate greater risk of developing infectious complications: type 0 biomarkers. As with the ΔSOFA score, a Complement activation test could be used repeatedly to assess Complement exhaustion over time and track condition severity with therapeutic intervention: type 1 biomarkers. Interestingly, the standing concentration of C3 does not correlate with the APACHE II or SOFA scores of septic patients (153), highlighting that Complement activation in sepsis is variable and relies on factors other than C3 concentrations. A better measure of Complement activation potential would include additional proteins of the system including activation products in a biomarker panel. Ideally, this panel of biomarkers would be measured on a rapid and multiplex assay platform to allow repeated evaluation of Complement activation phenotypes at point-of care. The trajectory of Complement activation phenotypes from pre-operative to post-operative immune states could reveal the impact of surgery on the immune system and provide evidence for the onset of nosocomial infection. Existing assays of Complement will now be discussed.

1.4 Existing Assays of Complement Activation

Assays for Complement proteins are increasingly important for initial diagnosis of disease, tracking progression of a condition and the patient response to complement-targeted therapies (203). There are four major classes of Complement diagnostic approaches: detection of opsonin deposition, measurement of zymogen concentrations, measurement of activation product concentrations and functional tests of Complement activity (203).

Detection of Complement proteins deposited on cells membranes can provide information about the type and localization of Complement activity in tissues (203). For example, localised Complement deposition has been observed on lesions in the brains of multiple sclerosis patients (232) and in joint tissues of those with rheumatoid arthritis (233). Immunohistochemistry is often used to observe antigen distribution over tissue sections and relies on Complement protein detection by antibodies conjugated to labels. The labels can be observed via a range of methods and are used to visualise the location of antibody binding on a tissue sample (234). Antibody labels include enzymes which induce colour changes, fluorescent markers (235), radioactive isotopes (236) and nanoparticles (237) which will be discussed later. Whilst immunohistochemistry is effective for showing Complement deposition, results cannot indicate the extent of

Complement depletion in healthy individuals with adequate surface regulators protecting the host tissues. In contrast, Complement proteins circulate in the blood of all individuals and are readily measured from routine blood samples.

The standing concentrations of inactive Complement precursors such as C1q, C1-INH, C4, C3, or FB are measured in clinical practice to detect Complement deficiency and so have validated assays and standards (238). Protein concentration is commonly measured by immunoprecipitation assays in which primary antibodies form complexes with proteins in the sample expressing specific epitope targets. Secondary antibodies, specific to the primary antibodies, may be added to increase the mass of the formed complexes which scatter light passed through the sample. The relationship between antibody-antigen complexes and the extent of light scattering is calibrated to determine the total concentration of antigen in the sample (239). However, immunoprecipitation gives no information about the conformation or activation state of the protein, and protein concentrations may appear normal in patients with dysfunctional proteins (240).

Activation products, such as protein fragments, directly indicate Complement activity may be more sensitive markers of dysregulation (241). For example, C3a has been measured alongside PCT to distinguish between SIRS and sepsis (242) – two conditions which present following surgery with similar symptoms of excessive inflammation (204). Complement activation products exhibit different properties to their parent proteins and these are exploited in assays for Complement activation. Firstly, many protein fragments from as a result of Complement activation are low molecular weight and their larger parent proteins may be filtered from solution (using polyethylene glycol for example) before immunoprecipitation analysis (243). Secondly, neo-epitopes on the surface of the protein fragments may be targeted by antibodies in immunoassays. Enzyme linked immunosorbent assays (ELISA) have been used extensively to quantify Complement activation in fluid samples (244) (245) but the epitope specificity of the antibody is important, highlighted by the recent finding that Eculizumab-C5 complexes express a C5a neoepitope resulting in false positives in some assays for C5a (246).

There are commercially available ELISA kits for Complement zymogens and fragments and several variations of the ELISA method. The most sensitive (by 2-

5 fold) is the sandwich ELISA (247). Antibodies are immobilised on a solid support, such as polystyrene or polypropylene, are used to capture antigens in a sample over time. Antigen concentration in the sample is proportional to the rate at which antigens bind to the capture antibodies. The assay surface is washed to remove unbound proteins and covered with detection antibodies which bind to epitopes on captured antigen. The epitopes of each antibody should ideally be on opposite sides of the antigen to avoid steric hinderance between antibodies binding to the antigen. The detection antibodies are conjugated to enzymes that produce a coloured product following interaction with a substrate which is added as the penultimate step in the procedure. A commonly conjugated enzyme is horseradish peroxidase (HRP) which oxidises chromogenic substrates such as tetramethylbenzidine (TMB) into coloured products. Peroxidase enzymes catalyse the reduction of hydrogen peroxide to water using two hydrogen atoms from TMB, which is oxidised from a colourless diamine to a blue diimine (248). The reaction is stopped in a final step via the addition of acid which reduces the pH of the solution outside the optimal working range of the enzyme. In a 'double-sandwich' ELISA, the detection antibody is not conjugated and a conjugated secondary antibody (specific to the Fc region of the detection antibody) is used to induce colouration. The extent of colouration is dependent on the number of antigens bound to the assay surface which is calibrated to derive the concentration of antigen in the sample.

Plasma protein concentrations vary among individuals (C3a concentration range = 29.7–173.8 ng / mL (249)) which confounds the use of strict cut-off points denoting Complement depletion. In addition, small activation fragments are removed by the kidneys such that C5a concentration measured in the plasma of healthy donors has been shown previously to be no different from septic patients (250). Higher mass biomarkers of Complement activation such as TCC would have low serum concentrations before elective surgery and rise post-operatively depending on the concentration of DAMPs produced by the procedure, further confounding the use of *in vivo* cut-off concentrations. Therefore, the best test of Complement activation potential is with a functional test: challenging patient serum *in vitro* and observing the Complement activation response. Such an approach was taken in a recent study of Complement activation after polytrauma (251).

Activation of the CP may be quantified using the CH50 test, which determines the dilution of serum required to lyse 50% of a known quantity of sheep erythrocytes (252). The sheep erythrocytes are sensitised to Classical Complement activation by pre-coating with anti-sheep antibodies. Serum for CH50 testing is diluted in buffer containing the divalent cations Mg^{2+} and Ca^{2+} required for Complement activation: the structure of the C1 complex is dependent on Calcium (253) and C4b binding to C2 is dependent on Magnesium (254). The test is an end-point assay and the protocol requires 50 minutes to complete (255).

Activation of the AP may be quantified via the AH50 test, a similar method to CH50. The binding of MBL to mannan requires Ca^{2+} (256) and the calcium-dependence of the CP and LP may be used to inhibit these pathways by reducing the bioavailability of Ca^{2+} with the chelating agent ethylene glycol tetraacetic acid (EGTA). The serum for AH50 testing is diluted using buffer containing only Mg^{2+} ions which ensures all haemolysis occurs as a result of AP activation, as the AP requires only magnesium for activation in the binding of C3b to FB (257-259).

The CH50 and AH50 tests are sensitive to several sources of error (255). Firstly, spontaneous lysis of erythrocytes may occur independently of Complement activity and the batch to batch variation of antibodies used to sensitise the cell membranes can lead to systematic errors. Secondly, a serum standard of known haemolytic activity is required to account for these errors. The CH50 results of the adult population exhibit a normal distribution with a factor of ~2 between minimum and maximum values (260, 261), highlighting significant variation in Complement activity in the healthy population.

Functional activity tests of all three initiation pathways in serum can be performed using ELISA, wherein the assay surface is coated with either antibody, mannan or LPS and the concentration of soluble TCC (sC5b-9) is recorded (262). Functional tests using TCC as an endpoint minimise the number of tests required to diagnose Complement deficiency (263); only if all Complement components are present in a serum sample and function correctly will the AH50 and CH50 (or equivalent ELISA assays) induce normal production of TCC (when compared to a standard reference range). An alternate functional activity test would involve activating all initiation pathways simultaneously and observing the response in each pathway with assays for specific activation products: biomarkers of the system response.

Owing to the reactivity of the Complement system *in vitro*, the assays of Complement discussed here may be influenced by pre-analytical factors such as sample preparation and handling. Activation markers can be rapidly produced in samples before analysis and confound results of activation studies. Freezing samples at -80 ^{o}C after mixing with the divalent cation chelating agent EDTA is often used to slow activation of all three initiation pathways, however, quantifiable activation still occurs (264). Furthermore, Complement can also be activated by sample thawing (241) and cation chelation only inhibits certain steps in the system which may induce flux bottlenecks and confound protein concentration studies. Ideally, Complement activity should be assessed immediately after sampling – achievable with minimal sample preparation and a rapid assay at point-of-care.

When sampling is performed correctly, the kinetics of Complement activation can be recorded accurately. EGTA serum and a synthetic pathway of purified FD, FB, FH, FI, C3, and P at physiological concentrations yield the same kinetics of C3 deposition on encapsulated *Cryptococci* (265) and *Bacteroides* (266). Furthermore, a synthetic pathway of 11 Alternative and Terminal cascade proteins (including negative regulatory protein FH) is capable of lysing sheep erythrocytes with a similar dose response curve to C4-depleted serum (267). The observation has two major implications. Firstly, extrinsic activation of the AP does not appear to contribute significantly to C3 opsonisation of bacterial surfaces. Secondly, the kinetics of Complement activation may be closely approximated using a well-defined set of proteins and equations for chemical kinetics in a mechanistic mathematical model of the system.

Complement is a complex system involving a balance of activating and deactivating reactions and many intermediates with short half-lives. The complexity of the system makes the hypothesis-driven selection of biomarkers challenging – which analytes should be measured following activation and when? Experimental data describing all Complement components following activation would provide a thorough characterisation of the system response *in vitro* but this would be a technical challenge due to the number of proteins involved. Mathematical models can be trained on limited data to help researchers to understand non-linear and emergent properties of biological systems, quantitatively predict the effects of perturbations, generate hypotheses and guide future studies (268). Non-linear features of systems can occur as a result of positive feedback such as the amplification loop in the Complement cascade for example. The emergent properties of a system are those which occur as a consequence of, but are not shared by, the interacting system components (269). For example, cells are self-assembled structures of molecules that when studied individually are considered ‘non-living’ but together display emergent, ‘living’ properties such as self-repair, reproduction and evolution via a complex system of interactions. Such a system is considered irreducible: unlikely to be fully understood studying each reaction in isolation (268). Systems biology models are used to understand and predict non-linear and emergent behaviour in complex systems such as the Complement cascade (270).

A mechanistically complete Complement activation model could be used to guide hypothesis driven selection of candidate biomarkers most sensitive to Complement depletion and patient risk of HAI. Complement is a suitable biological system for mechanistic modelling because the proteins and interactions of the system are largely well-defined as described previously. Complement activation can therefore be mathematically represented using differential equations, discussed in Chapters 2 and 4 to describe the concentration change of each species over time. Furthermore, the binding kinetics of some reactions in the Complement system have been derived by studying reactants in isolation. A fully mechanistic approach to Complement modelling is possible due to the existing biological understanding of the system.

Existing Complement models have been used to investigate therapeutic inhibition of certain interactions to guide drug design (271) and the effect of starting protein concentrations on bactericidal action (185). However, no existing Complement models are both fully mechanistic and consider all pathways of the system, preventing predictions of system behaviour based on any protein or interaction described in section 1.3. Inadequate biological data has been cited as the major limiting factor in the training and validation of complete Complement models (271).

In summary, the identification of surgical patients at higher risk of nosocomial infection is challenging but enables targeted allocation of critical care resources to improve high-risk patient mortality. The Complement cascade is a fundamental component of the innate immune system and responds rapidly to infection, which if compromised increases risk of HAI. The hypothesis of this thesis is that a direct measurement of Complement function might help identify surgical patients at higher risk of HAI. The best test of Complement activation potential is a functional test: challenging patient serum *in vitro* with an activation stimulus and observing the Complement activation response in each pathway. The *in vitro* response of Complement to a standard challenge can be profiled over time by measuring activation product concentrations (such as C5a) or the consumption of zymogens (such as C3) in serum with existing immunoassays. A depleted Complement cascade should exhibit a different activation response compared to a 'healthy' system because activation kinetics depend upon both the stimulus and relative concentrations of the circulating proteins. Properties of the concentration-time profiles of Complement proteins following activation such as initial rate and total change in concentration are patient phenotypes, observable physical properties where values outside of the healthy reference range might indicate that a patient is at higher risk of nosocomial infection. The proposed test of Complement activation potential is shown in Figure 1.5.

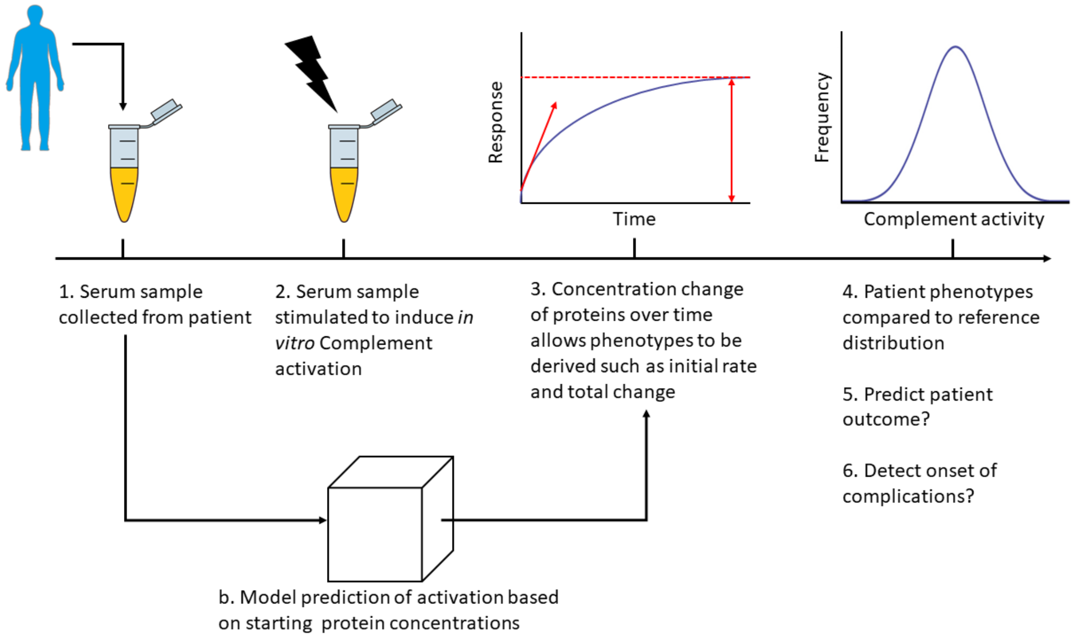


Figure 1.5 The proposed test of Complement activation potential: a possible measure of post-operative risk of nosocomial infection. A serum sample is collected from a patient either pre-operatively, with the intention of predicting infection risk, or post-operatively, with the intention of detecting the onset of Complications. All initiation pathways of the Complement system are activated in vitro and the response profiles of Complement activation products measured over time with immunoassays. Alternatively, Complement activation can be predicted using a mechanistic mathematical model which considers the starting concentrations of all proteins. Properties of the measured or predicted concentration-time profiles, such as initial rate and absolute change in analyte concentrations, are then compared to the reference range. Deviation from the reference range of patients at low risk of post-operative infection may suggest a patient is high-risk and inform clinical decision making such as the choice of care pathway.

Repeating the Complement function test after surgery would allow a patient's response to be compared to their pre-operative condition, wherein a significant relative decrease in Complement activation potential might indicate increased vulnerability to infection. Repeated testing over time as part of an EWS might enable early recognition of infection and track condition severity.

Studies of Complement in patients are currently limited by laboratory bound assays which require trained personnel to operate. Complement activation potential is occasionally assessed with functional tests such as CH50 and AH50 to identify genetic deficiencies but is not tested routinely in a clinical setting. The ease of studying Complement activation potential could be improved with a point-of-care biosensor for specific Complement protein concentrations. Ideally the biosensor would be multiplex, allowing a panel of biomarkers to be tested, and the assay would enable real time measurements of protein concentrations with minimal sample preparation to avoid the need for sample storage and allow the test to be performed with ease.

The protein concentration profiles most sensitive to a depleted Complement response are challenging to predict due to the highly non-linear change in protein concentrations during Complement activation (185). Complement activation *in vitro* should be fully characterised to enable a hypothesis-driven approach to identifying the most sensitive phenotypes of Complement depletion. A fully mechanistic model of Complement activation, describing the change in concentration of all species over time could be used to quantitatively predict phenotypes which are sensitive to Complement depletion. The identified phenotypes could then be tested in a clinical study comparing the phenotype distributions of healthy patients and those who develop nosocomial infection. An accurate *in vitro* Complement activation model might also be used to predict patient phenotypes based on a single measurement of the starting Complement protein concentrations in a serum sample. Use of a model prediction would reduce time for patient stratification into high- and low-risk groups. Such an approach would require both a multiplex assay of critical protein concentrations and a validated mathematical model of *in vitro* Complement activation – neither of which are currently available.

1.5 Thesis Aim and Objectives

This thesis will report the characterisation *in vitro* Complement activation phenotypes as potential biomarkers of immune state denoting a patient's risk of post-operative nosocomial infection. The same phenotypes should also reveal the onset of infection post-operatively and their measurement requires a quantitative assay platform capable of measuring Complement activation products over time. The phenotypes may also be predicted using a mechanistic model of the system considering the starting concentrations of proteins in a serum sample. Ultimately, the intended utility of the phenotypes is stratification of patients into care pathways to improve outcome.

The first objective was to establish a reproducible assay platform on which tests for Complement activation can be developed, with potential for multiplexing, miniaturisation and use at point-of-care – not possible with existing clinical assays of Complement function. Chapter 2 describes the development of a microarray biosensor platform, the Light Scattering Array Reader (LiScAR), inherited as an existing device for studying protein-protein interactions with the potential to be used as a multiplex immunoassay for serum proteins (272).

The second objective was to screen antibodies for sensitivity and specificity against analytes of interest using the LiScAR biosensor platform. Chapter 3 describes the development and performance assessment of novel multiplex immunoassays for Complement proteins A rapid test for antibody monomeric purity was developed to support the new assays, enabling quality control testing of antibody samples from different batches. A rapid and multiplex biosensor platform and validated assays for a number of Complement proteins allowed rapid kinetic data of cascade activation to be collected for the optimisation of a mathematical model of the Complement system.

The third objective was to produce a mathematical model of Complement activation. Chapter 4 describes a systems biology approach to developing kinetic models of a simple system. Chapter 5 describes the use of techniques discussed in Chapter 4 to train a fully mechanistic mathematical model of the Complement system. The optimised model was used to predict the response of the Complement system from different starting conditions. Quantitative Complement activation characteristics, system phenotypes, were explored using the model as metrics to indicate Complement depletion and possibly risk of nosocomial infection.

The fourth and final objective was to establish the distributions of *in vitro* Complement activation phenotypes for the healthy adult population. Chapter 6 presents the first measurements of *in vitro* Complement activation phenotype distributions in healthy adults and compares the results to modelled predictions for model validation. The phenotype variation was used to estimate the power of each phenotype to discriminate healthy from compromised Complement systems: a necessary prerequisite to future clinical studies investigated *in vitro* Complement activation phenotypes for patient stratification.

This thesis details the research journey from conceptualisation of a new biomarker, through development of a new assay and training of a new mathematical model, to recording biomarker values in a cohort of healthy individuals. This work therefore aims to provide all necessary prerequisites for further research into Complement response testing as a means to gauge immune state and infection risk.

2 Development and Performance Assessment of a Biophotonic Protein Biosensor: LiScAR

2.1 Introduction

The biosensor platform around which a significant portion of this thesis is based was inherited as an existing device designed for observing protein-protein interactions using light scattering (272). The biophotonic instrument, hereby referred to as the light scattering array reader (LiScAR), has been described previously in several publications and utilised for a broad range of experiments including: antibody affinity studies (273-275), serum antibody screening (276-278), lectin-fibrinogen affinity studies (279) and antibody epitope mapping (275). Previous research indicated that the LiScAR might be used as an immunoassay for serum proteins, such as those of Complement cascade, with significant potential for miniaturisation and eventually use at point-of-care. This chapter will first introduce the main properties of biosensors and show why the LiScAR is an appropriate platform on which to develop assays for studying the Complement cascade.

Biosensors are analytical instruments designed to produce a signal in response to the presence of a specific analyte or set of analytes (280). Conventional analytical techniques such as mass spectrometry (MS) and ELISA assays are the gold standard for quantitatively determining biomarker concentration in a sample but require trained personnel to operate and are laboratory bound – hindering their use for biomarker profiling of patients at point-of-care (281). In contrast, biosensors often surpass conventional analytical techniques with respect to ease of use and portability (282) by combining the specificity of biochemical interactions with a microprocessor into a single instrument with a small form factor (appropriate for use at point-of-care) where all data analysis is automated for the user. Biosensors may also enable more rapid sampling rates and lower running costs compared to MS or ELISA, making these devices attractive instruments for research use. Biosensors have many applications in fields such as food safety (283), medical diagnostics (284), drug discovery (285), and environmental monitoring (286) but regardless of application, all biosensors consist of the following common components (287):

- Bio-receptor – a biomolecule that generates a response upon interaction with analyte, in a process termed bio-recognition. The response may be observed as variations in temperature, colour, charge, etc. The part of the response that contains useful information is termed signal (288).
- Transducer – a device that converts the energy of the bio-recognition signal into another form, typically electrical energy.
- Electronics – the transduced signal is systematically processed and interpreted to make the data meaningful for the user.
- Display – the results of electronic processing are reported to the user graphically, typically by a computer screen.

The International Union of Pure and Applied Chemistry (IUPAC) recommends that a biosensor is classified according to the bio-receptor mechanism or mode of signal transduction (289).

2.1.1 Bio-receptors and signal transduction

Once a target analyte has been identified, a major task in developing a biosensor involves selection of a suitable bio-receptor. There are two major categories of bio-receptors: bio-ligands and bio-catalysts (290). Molecules capable of simply binding to a target analyte are considered bio-ligands and antibodies (282), nucleic acids (291) and lectins (292) are common examples. Bio-catalysts are also capable of binding analytes but have an additional activating role in biological reactions - converting bound substrates to products. Typical examples of bio-catalysts include enzymes (293) and whole cells (294). The range of available bio-receptors requires a variety of signal transduction technologies to convert the initial stimulus into electrical signals that can be analysed by a computer processor. The transducer element of a biosensor is not selective for the analyte but greatly influences the sensitivity of the instrument because signal transduction is a potential bottleneck in the transfer of information from bio-recognition stimulus to electrical output. Common modes of signal transduction include: electrochemical (295), thermometric (296), piezoelectric (297), optical (298) and magnetic (299).

The performance of biosensors, with varying bio-receptors and modes of signal transduction, may be compared using standardised instrument characterisation metrics. IUPAC strongly recommend that the performance criteria of a biosensor should be reported using the following metrics: calibration characteristics, selectivity, measurement duration, reproducibility, stability and lifetime (289).

2.1.2 Calibration Characteristics

An important analytical characteristic of a biosensor is the relationship between analyte concentration and the signal produced by the bio-recognition event: the calibration curve (300). Calibration curves are often plotted as signal with respect to analyte concentration on a log scale, and typically show deviation from linearity at the extremes of the concentration range, Figure 2.1.

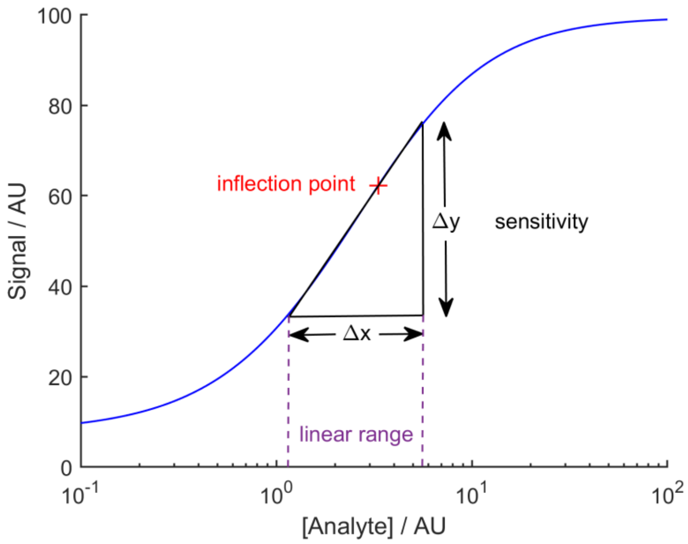


Figure 2.1 A typical calibration curve for a biosensor, showing the relationship between analyte concentration and signal. The curve is most sensitive to analyte concentration within a linear range (purple), over which the gradient is effectively constant and may be fitted to a straight line. The gradient of the linear portion of the curve (black) provides a measure of instrument sensitivity. The curve is typically sigmoidal and the inflection point (red) defines the position at which the sign of curvature changes and the gradient is steepest. The curve shown is asymmetrical and hence the inflection point is not in the centre of the linear range.

The sensitivity of a biosensor is in part determined by the slope of the calibration curve. All else being equal, a greater absolute change in signal for a given change in analyte concentration enables greater measurement resolution - the ability to distinguish between smaller differences in analyte concentration. A non-zero signal in the absence of analyte is common in analytical techniques and is termed background: a non-specific contributor to the analyte signal which must be included in the calibration curve to prevent artificially high readings at low analyte concentrations. Performing the standard instrument analysis procedure on a sample containing zero analyte is referred to as a blank measurement and provides an estimate of the background. Absolute signal magnitude (above the background) is not however, the sole contributor to instrument sensitivity and signal noise has an equally critical impact on calibration performance.

Noise is defined as any fluctuations in the instrument response that are independent of analyte concentration and hence greater noise reduces instrument performance (288). There are several sources of instrument noise which may or may not be avoidable. For example, environmental sources of noise, such as physical vibration from nearby electric motors or electromagnetic interference from power cables, may be avoided by isolating or insulating the analytical instrument from these signals. In contrast, thermal noise is unavoidable in electric circuits as this occurs due to the random motion of discrete charge carriers regardless of applied voltage. Similarly, the shot noise of a camera sensor occurs due to the discrete detection of photons emitted randomly from a light source.

Noise is a major source of variation between identical measurements on the same instrument, termed intra-assay variation or analytical noise. The Gaussian (or normal) distribution describes how repeated measurements cluster symmetrically around a mean value. The random variation among measurements in a normal distribution can be estimated using the sample standard deviation (SD) (301). The SD is the square root of the sample variance and a measure of the typical difference of a measurement from the mean of the sample:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \qquad 2.1$$

where x_i represents an individual measurement, $\bar{x}$ is their mean value and N is the number of repeats. The n-1 term is Bessel's correction, which improves the closeness of the estimated SD to the true standard deviation of the population (σ). One SD either side of the mean measurement value describes a region containing 68% of all individual measurements, whilst 95% of all measurements have values within an interval 2SD to either side of the mean. Not all distributions are normal and should be tested before attempting to calculate SD. Normality testing is discussed in Chapter 6.

A related statistic to SD is standard error (SE) (301), which estimates the standard deviation of a sampling distribution: a distribution of values where each is derived from multiple observations such as a distribution of mean values achieved in repeated experiments. In a distribution of mean values, the standard error of the mean (SEM) is used to estimate the average difference of each experiment's mean from the mean of the sampling distribution:

$$SE = \frac{\sigma}{\sqrt{n}} \tag{2.2}$$

where σ is the standard deviation of the population and n is the number of values in the sampling distribution. For infinite populations where the true value of σ is unknown, SD is used as an estimate – highlighting the importance of Bessel's correction. The 95% confidence interval (CI) provides the information about a range within which the 95% sample means will fall: 2SEM either side of the sampling distribution mean. Notably, a normality test is not required before calculating the SEM because according to the Central Limit Theorem: the distribution of mean values is closely approximated by a normal distribution, even if the distributions of the measurements in each sample are not normally distributed.

The low gradient regions of the calibration curve are most affected by analytical noise and have worse measurement resolution than the linear region. For example, when repeatedly sampled, the measured brightness of a light source (the number of photons emitted per unit time) will vary significantly if a relatively low number of photons are detected during the time frame of each measurement. In contrast, the brightness of a more intense light source (or the same light source recorded over a longer period of time) will produce more consistent estimates of brightness because the random variations in photon emission will be smaller relative to the recorded value: the measurement has a better signal to noise ratio.

The repeatability of an instrument is quantified by repeating the analytical procedure on a standard sample of analyte. The relative variability in apparent analyte concentrations (derived from reading off the standard curve) can be measured using the coefficient of variation (CV) (302) with the result often reported as a percentage:

$$CV = \frac{SD}{\mu} \times 100 \tag{2.3}$$

where µ is the sample mean value and SD is the sample standard deviation. It is evident that lower concentration samples will show higher CV for a constant SD: a lower signal to noise ratio.

During the design of biosensor platforms, it is advantageous to identify the sensitivity requirements of the instrument (303), which could be decided by the range of analyte concentration in samples of interest. For instance, a biosensor designed to measure C3 in serum would ideally be sensitive to the lowest concentration of the C3 distribution found in serum, else the instrument must simply report that the derived concentration is below its detection threshold. Furthermore, it is often advantageous for the dynamic range of the standard curve to encompass the analyte concentration range, as differential dilution of samples can lead to matrix effects (304), discussed later. The concentration ranges of Complement proteins should be considered for biosensor design and are shown in Figure 2.2.

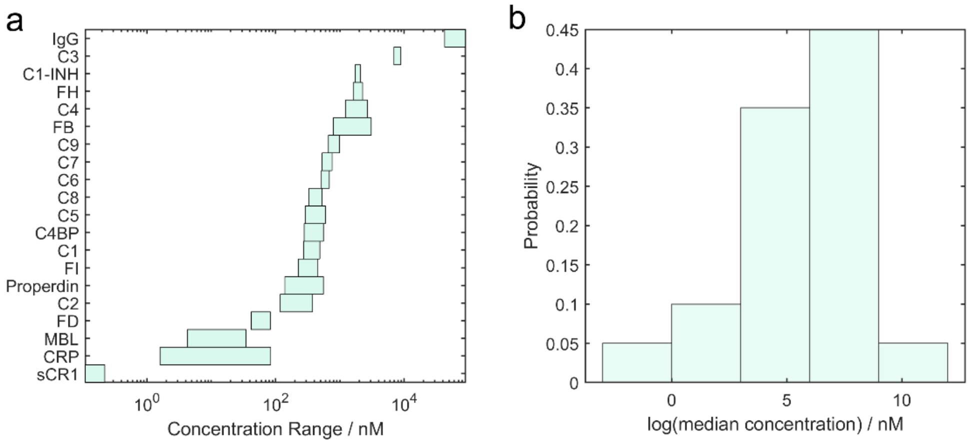


Figure 2.2 Concentrations of circulating Complement proteins in normal human serum. a) concentration ranges ordered by lower range limit. b) distribution of median Complement protein serum concentrations. Values from Table 6.1.

The majority of Complement protein concentrations in serum range between 100-3000 nM and their medians approximate a log normal distribution. The LiScAR should therefore have nanomolar sensitivity in order to be appropriate for studying the response of the Complement cascade.

To compare the lowest concentration of analyte that can be reliably measured on analytical instruments with varying analytical noise and calibration curve gradients, detection limits are defined, Figure 2.3. Detection limits are expressed in units of analyte concentration after converting the detection signal threshold to concentration using a calibration curve (288). Three types of detection limit are typically defined for analytical instruments and are designed to account for analytical noise which is random and normally distributed:

- Limit of blank (LoB) – the highest apparent analyte concentration when replicates of a blank sample are measured, defined mathematically as μ + xSD of the derived blank replicates. The value of x is chosen based on the required confidence in analyte detection; 1.645×SD is recommended and contains 95% of the blank values as described previously (305).
- Limit of detection (LoD) – the minimum concentration of analyte required to produce a statistically significant signal compared to the analytical noise derived from replicate blank samples. Traditionally LoD is defined as μ + 2×SD of the derived blank replicates. However, the method is overly optimistic and does not consider the analytical noise of the samples containing the LoD concentration of analyte. For example, a protein solution may induce more analytical noise than the blank solution containing only buffer. Therefore, a superior approach is to measure the analytical noise for a low concentration of analyte and define LoD as LoB + xSD of the low concentration replicates. The value of x is again chosen based on the required confidence and 1.645×SD is recommended to ensure 95% of samples at LoD concentration will produce values exceeding the LoB (305).
- Lower / Upper limit of quantitation (LLoQ / ULoQ) – the lowest or highest concentration of analyte which can be determined from the calibration curve with desired analytical precision, such as a certain CV value (305).

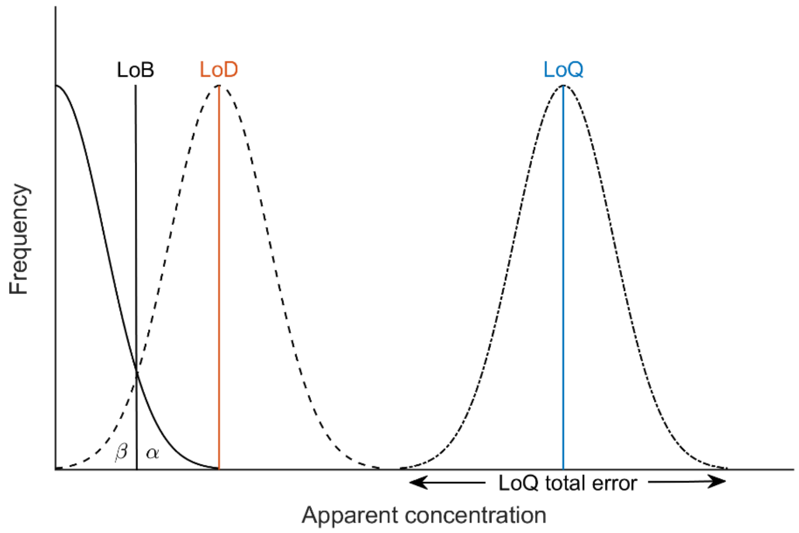


Figure 2.3 Graphical representation of LoB, LoD and LoQ metrics. The distribution of apparent concentrations of several blank replicates enables LoB to be defined (solid black). 5% of blank samples have apparent concentrations greater than the LoB (α), representing Type I error. The distribution of samples at a concentration of LoD (dashed black, red) is calculated from the SD of low concentration replicates. 5% of samples at LoD concentration have apparent concentrations below the LoB (β), representing Type II error. The apparent concentration of replicate samples at LoQ concentration (dot dashed black, blue) must show a required CV.

The overall analytical method is typically more complex than analysing pure, standard samples of analyte and the imprecision associated with additional steps such as dilution or mixing with reagents provides opportunities for error to accumulate. Therefore, in addition to the previous definition of LoD, the method detection limit (MDL) considers the preparation and matrix of the measured samples when calculating LoD (306). For instance, blood samples may need to be diluted before use with an assay to lower the analyte concentration in the sample into the dynamic range of the standard curve. The matrix of the sample may have a significant effect on LoD and should be controlled for by one of two ways in a standard operating procedure:

- The standard samples which are measured to produce the standard curve may be produced in the same matrix as the samples of interest. The standard curve can directly determine the analyte concentration in test samples, but preparation of the standards may be complicated.
- The standard samples measured to produce the standard curve may be produced in a different matrix to that of the samples of interest, simplifying standard production. A sample of known analyte concentration in the same matrix as the samples of interest (termed a spike) is then measured. The difference between the known concentration of the spike and the

biosensor reading defines a property called percentage recovery and can be used to adjust the readings of unknown samples.

Importantly, detection limits only measure the ability of a biosensor to distinguish between background and signal. Proving that biosensor signal is due to the presence a certain analyte is a question of selectivity, which requires additional experiments.

2.1.3 Selectivity

In theory, a biosensor is described as specific when a single analyte is responsible for the signal that is measured and no other (307). However, bio-receptors often fall short of being truly specific, instead simply exhibiting a degree of preference for the target analyte: selectivity (308). The phenomenon of selectivity is often due a high degree of similarity (chemical and structural) between analytes. For instance, there is a high degree of structural homology between IgG of sheep, goats and cows. Antisera raised against IgG from one species may contain antibodies capable of binding to the IgG of other species (309). Furthermore, selectivity for a protein is complicated if the target analyte can be incorporated into a larger structure. A monoclonal antibody specific for C3 of the Complement cascade may bind to C3, C3b and any other complexes containing C3b so long as the epitope remains and is sterically available. Herein lies a major challenge of developing antibody-based assays as the epitope targets are not commonly reported (310). Many biosensors in use rely on a single highly selective bio-recognition element, but an alternate approach is to incorporate several sensor elements into a single device: multiplex platforms.

Chemometrics is a data processing technique that uses mathematical and statistical methods to correlate patterns in multiplex or large analytical data sets with the physical properties of a sample. Published examples of chemometrics include determining the origin of medicines (311) and consumables (312) and a method for the detection of extra-terrestrial life (313) – all based on characteristic chemical ‘signatures’. Chemometric methods such as principal component analysis can combine data from arrays of sensor elements which react differentially with analytes of interest to create a ‘chemical fingerprint’, allowing samples to be compared in multidimensional parameter space. Chemometric analysis is therefore analogous to combining taste and smell to distinguish between foods and is sometimes referred to as a ‘chemical nose/tongue’ approach (308). The advantage of such an approach is that the variety of analytical data can make up for low selectivity of individual assay elements. Ideally all sensor elements of a multiplex biosensor will be highly selective and operate simultaneously -reducing both the sample volume required and the measurement duration compared to if the assays were operated individually. A disadvantage of chemometrics is the potential effect irrelevant data which may confound the analysis. As discussed in Chapter 1, a panel of validated, high-quality biomarkers is most desirable.

2.1.4 Sampling Rate and Measurement Reproducibility

The rate at which the concentration of analyte in a sample can be determined is a critical metric when reporting biosensor performance. For samples which are chemically stable, assays providing high accuracy but longer operation times are ideal. An example would be using an ELISA plate to measure the analyte concentration in 96 samples simultaneously over several hours. However, a significant challenge for biosensor development is providing a sampling rate that is sufficient to distinguish the kinetics of rapid biological processes.

High frequency sampling is particularly important for studying the Complement cascade as the system is activated on timescales of minutes rather than hours, and so sample collection, storage and measurement should be performed with minimal delay to achieve accurate results. As discussed in Chapter 1, activated Complement samples are typically stored with EDTA to halt activation by sequestering divalent cations required for certain protein-protein interactions. However, chemically halting a cascade at only certain points in the system is likely to cause bottlenecks of flux and provide misleading kinetic data. Therefore, Complement activation kinetics would ideally be studied using fresh serum sampled on a timescale capable of defining the rate of change in concentration of each protein of interest. Such an approach requires an instrument capable of measuring a sample rapidly and ideally repeatably at a high frequency to minimise cost - rather than using multiple single use assays such as ELISA plates. Not all biorecognition elements are appropriate for repeated use, often because the reaction with the target analyte is irreversible. Repeated measurements of stable samples must produce similar values, and the performance metric of the similarity is termed reproducibility and can be measured with CV.

As discussed, the variation in biosensor technology, sensitivity and sampling rate is significant. It is therefore unsurprising that biosensors also vary in terms of reproducibility over time: the ability of the biosensor to generate identical responses for a repeated experiment (287). A critical factor in determining the reproducibility of a biosensor is the stability of any biological components and reagents, which often degrade over time with repeated use and under storage conditions – hence the expiration dates accompanying commercial immunoassays. The half-life of any tethering chemistry involved in the construction of the biosensor is a further source of biosensor performance variation over time. Biosensor sensitivity loss due to bio-receptor degradation and tethering half-lives must be determined experimentally by repeated measurements of an analytical calibration standard. Understanding the lifetime of a biosensor ensures repeatable experiments and may enable signal loss to be corrected for, if well characterised.

2.1.5 Labelled and Label-free Techniques

Two classes of biosensors may be defined based on how the binding of analyte is detected by the sensing element: labelled where the detection event detects the presence of the label (assumed to be bound to the target with known stoichiometry) and label-free techniques where the native species is detected (314). Labelled techniques include surface-based assays such as ELISA and bead-based assays such as the Luminex platform, whilst label-free techniques include mass spectrometry (MS) and surface plasmon resonance (SPR) devices (315).

Detecting analyte directly according to physical properties such as mass or charge in a background of other species is often challenging. Therefore, many biological assays use a high affinity receptor such as a monoclonal antibody to bind analyte with high selectivity, before labelling the receptor with a signal-generating reagent that is more easily observable (316), Figure 2.4. Common labels include nanoparticles, fluorescent proteins and element isotopes (317), which are detected by measuring the light or radiation emitted by these labels when bound to the analyte of interest. A typical example of a labelled technique is ELISA, in which no signal is generated by the binding of analyte to the assay surface. Instead, a detection antibody conjugated to an enzyme is used to produce a coloured substrate and reveal the quantity of bound analyte via a secondary binding step in the assay protocol.

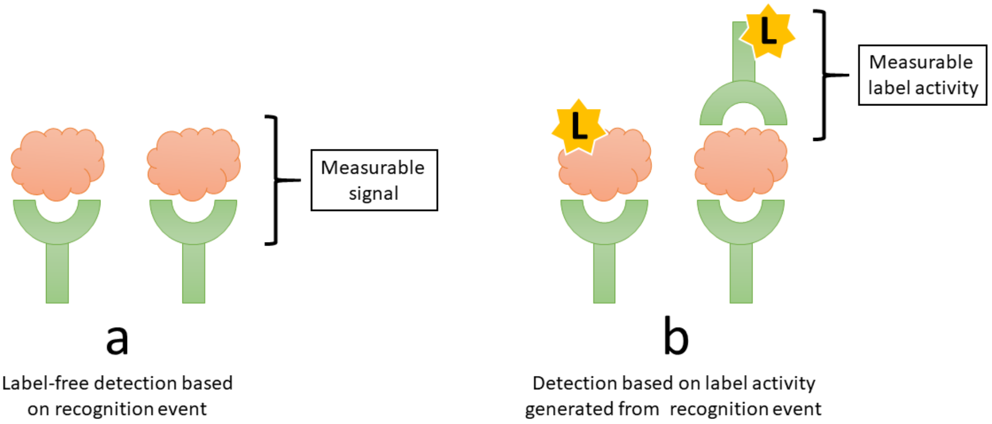


Figure 2.4 Origin of measurable signal in labelled and label-free techniques. Bio-recognition elements (green) bind to analyte (pink). a) Analyte may be detected based on the recognition event only in a label free biosensor. b) Analyte may be detected based on a readout of label activity (yellow). Figure adapted from ref (316).

Labelled assays such as ELISA typically exhibit LOD values two orders of magnitude lower than their label-free counterparts, with the most sensitive labelled techniques using conjugated secondary antibodies for signal amplification (318). However, the synthesis of labels may be complex or low yield and requires quality control to prevent sources of error leading to systematic inter assay variation. Furthermore, in the case of sandwich immunoassays, finding an antibody pair with compatible epitope locations and acceptable selectivity is challenging. Another disadvantage of using labels is that these techniques may have relatively low temporal resolution, particularly if the incubation time required for analyte detection with the labelling step is significant – highlighting the importance of high affinity antibodies as bio-recognition and detection reagents, discussed later. For consistency, label-based assays such as ELISA make a single measurement after a fixed time period and are known as end-point assays. In contrast, label-free biosensors enable the direct detection of many biological and chemical analytes (298) and may derive analyte concentration from the rate of binding observed in real time: kinetic assays.

Label-free biosensors often enable faster sampling speeds than labelled endpoint assays and may be broadly divided into two categories: optical (298) and non-optical (314). Non-optical techniques include electrochemistry (319), acoustic resonance (320), microcantilever (321), isothermal titration calorimetry (322) and quartz crystal microbalance technology (323). Most commercial analytical platforms use optical techniques (324) including SPR (325), optical waveguides (326) and disk resonators (327). Optical biosensors have benefited from progress in nano-technology which has enabled the production of cost-effective and small form factor devices with the potential to be used for biomarker measurements at point-of-care in a clinical setting (298, 328).

2.1.6 SPR-based Biosensors

The most extensively employed optical biosensors are those based on SPR (329). SPR is also one of the most prominent optical biosensor technologies used for pharmaceutical research (330) and a significant proportion of research papers report SPR-based assays for clinically relevant biomarkers with appropriate detection limits in a range of biological fluids (331). SPR-based sensors, such as the LiScAR, can detect the binding of biomolecules to a sensor surface functionalised with complementary proteins which can be reused. SPR-based sensors therefore offer rapid, repeatable, label-free measurements of binding kinetics from a single sample – a significant advantage over conventional clinical chemistry techniques such as ELISA.

SPR-based biosensors are not without disadvantages and the major drawback is that detection limits are fundamentally reliant on the mass of analyte. Labelled techniques such as ELISA assays can achieve picomolar LoD values (332), whilst the typical LoD for SPR methods is 1−10 nM for a 20-kDa protein (333). SPR detection limits are similar to those of ELISA for larger molecules (334) and labelling techniques are available for SPR to improve sensitivity (335), but at the cost of sampling speed. A further disadvantage of conventional SPR-based biosensors is the requirement for protein immobilisation on a sensor surface, which may limit mass transfer and has been attributed to the under-estimation of binding rate constants compared to the same reactions in solution (336, 337). The rate constants of a simple antibody-antigen interaction at a surface and in solution are compared using the LiScAR in Chapter 4.

In order to be adopted in a clinical setting, the advantages of SPR-based biosensors over conventional end-point assays must be demonstrated for the detection of biomolecules in clinical samples from patients. However, only 1% of research articles report the use of patient samples and the median number of patients in these studies is small (~10) (331). Instrument calibration using pooled biofluids is useful for technique development but may not be appropriate for actual clinical samples due to differences in protein adsorption profiles between patients (338).

For studying the Complement cascade, both labelled and label-free techniques have their place. Quantifying binding kinetics and rapidly sampling analyte concentrations are tasks most appropriate for a label-free biosensor platform, such as the LiScAR, whilst high precision estimates of low concentration proteins could be performed with labelled techniques such as ELISA. The characteristics of existing biosensors have been discussed and those most appropriate for measuring Complement activation in a clinical setting have been identified. The LiScAR platform described here is an SPR-based kinetic assay platform which when combined with a well-chosen bio-receptor and tethering chemistry should provide rapid, repeatable and low volume sampling of complex biological fluids with nanomolar sensitivity. The basic principles of SPR in relation to the signal generation in the LiScAR will now be described, before the construction and development of the device is presented in detail.

2.1.7 Surface Plasmon Resonance

Plasmons are oscillations of free electrons at the surface of metals, induced by the oscillating electric field of incident electromagnetic (EM) radiation (339). In particles smaller than the wavelength of incident EM radiation, such as metallic nanoparticles, all electrons in the volume of the particle experience a similar electric field and displace coherently from one side of the particle to the other: a localised surface plasmon (LSP) (339). The displacement of the negatively charged electrons from the positively charged metal lattice generates a restoring force that attracts the electrons back to the lattice – inducing an oscillation with direction parallel to the electric field of the incident EM radiation (340). The frequency at which a plasmon can oscillate is a function of particle size, shape, and composition. Furthermore, the lifetime of the oscillation is limited by energy loss (dephasing times for gold nanoparticles are of 2–50 fs depending on particle characteristics) but can be sustained by radiation with a wavelength matching the oscillation frequency (or resonance frequency) of the plasmon, in a process known as resonance, Figure 2.5.

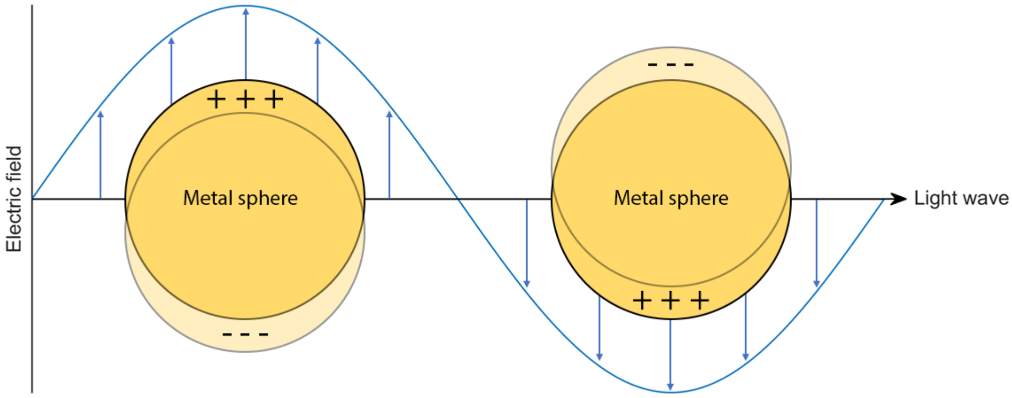


Figure 2.5 Illustration of a localised surface plasmon on a metal sphere with diameter smaller than the wavelength of incident EM radiation. Figure adapted from ref (341).

LSP resonance (LSPR) results in wavelength selective absorption at visible wavelengths for noble metal nanoparticles, which may be observed with a UV-visible spectrometer (315). Gold nanoparticles are commonly used in the construction of LSPR-based sensors (342) and the absorption spectrum of gold colloid reveals an absorbance peak corresponding to the resonance frequency, Figure 2.6.

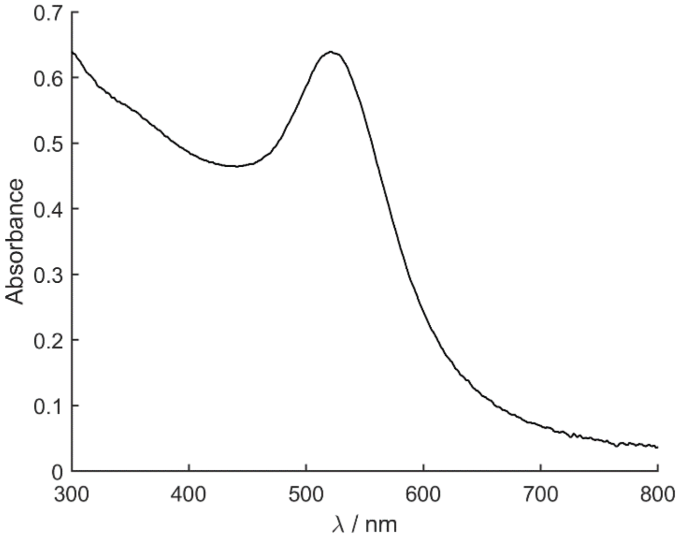


Figure 2.6 UV-visible absorption spectrum of 3–4 nm gold nanoparticles. The colloid was produced using the synthesis described in ref (343).

The spectrum suggests higher absorbance for shorter wavelengths but this is an artefact due to Rayleigh scattering of EM radiation by molecules and particles in the solution with smaller diameter than the wavelength of incident radiation. Rayleigh scattering is inversely proportional to wavelength and hence shorter wavelengths are scattered more and prevented from reaching the detector of the spectrometer (344). The width of the absorbance peak indicates the size distribution of the gold colloid whilst the position of the maximum shifts with particle size (larger particles red-shift) as well as particle shape and material (329). Furthermore, the amount of light scattered and absorbed (the extinction) of the nanoparticles changes with properties of the surrounding medium such as increasing refractive index (RI) (329).

RI is the ratio of the speed of light in a vacuum to the speed of light in a specified medium and is dependent on molecular structure, molecular weight and temperature (345). When proteins in solution bind to nanoparticles functionalised with complementary bio-receptors, the local refractive index of the nanoparticles increases and causes a red-shift and an increase in extinction at the resonance frequency of the LSP (329). The increase in scattering of monochromatic light by nanoparticles can be recorded to show the RI increase associated with analyte binding to the nanoparticles in real time: the basis of the LiScAR platform.

Single nanoparticles have been used to detect biotin with an LoD of 10 nM (346) and the high potential for miniaturisation is a key advantage of LSPR-related sensors as this reduces the required sample volume – an important issue for clinical testing. Miniaturisation also enables multiple sensors to be combined into a multiplex assay platform. The LiScAR is therefore an ideal platform for the development of a multiplex assay for Complement proteins in a clinical setting.

2.2 Aims and Objectives

The aim of this chapter is to describe and explain the principles behind the construction and operation of the LiScAR biosensor platform. This chapter will also present optimised methods to maximise performance of the biosensor, ahead of attempting antibody immobilisation for assays Complement proteins in Chapter 3.

The first objective was to study the fundamental methods behind assay surface fabrication, signal generation and instrument operation which were inherited from the research group (276).

The second objective was to investigate effect of changing several variables throughout the fabrication process and develop optimised standard operating procedures (SOPs) for assay construction. The operation the biosensor was made faster and more user-friendly by developing automated data processing routines.

The final objective was to demonstrate the procedure for kinetic analysis of analyte binding at the sensor surface and calibration of the instrument to derive analyte concentration.

2.3 Materials

The fundamental methods and SOP for the LiScAR were inherited from the research group and published in reference (276). The following materials are used for the construction of the LiScAR sensor surface. Auric chloride (#254169), sodium citrate (#W302600), glycerol (#G9012), hexadecyltrimethylammonium bromide (#H5882) and ascorbic acid (#A4403) were obtained from Simga Aldrich. The linker (HS-C11-EG6-OCH2-COOH) and spacer (HS-C11-EG3) self assembled monolayer components were purchased from Prochimia Surfaces. Pierce recombinant protein A/G (#21186) was obtained from Thermo Scientific. Human serum albumin and bovine serum albumin were obtained from Sigma-Aldrich. NIST RM 8671 was purchased from the National Institute of Standards and Technology. Standard instrument running buffer, also used in preparation and dilution of the samples, was phosphate-buffered saline (PBS) supplied in tablet form by Sigma-Aldrich. Ortho-phosphoric acid (85%) was obtained from Fluka and a 0.01 M aqueous solution used as regeneration buffer. Silane-prep slides were obtained from Sigma Aldrich and the COC, COP and amino microscope slides were obtained as bespoke materials from Scienion AG.

2.4 Gold Nanoparticle Array Production SOP

Gold nanoparticle colloid was synthesised following the methods described in ref (343). Auric chloride (200 µL, 50 mM) and sodium citrate (100 µL, 0.1 M) were added to 40 mL deionised water in a Duran bottle. $NaBH_4$ (0.151 g) was dissolved in 40 mL of deionised water to produce a 0.1 M solution and cooled in an ice bath to 0 ^{o}C. The cooled $NaBH_4$ (2.4 mL) was added quickly via pipette to the auric chloride solution with stirring. A colloid of gold nanoparticles was formed with an orange / red colour. The reaction mixture was stirred at room temperature for 10 minutes before incubation at 40 ^{o}C in a water bath with periodic stirring of <5 seconds every minute. The reaction mixture was allowed to stand at room temperature for 2 hours and gradually darkened to a ruby red colour – associated with a shift in the resonance frequency of the nanoparticles as they matured. A dark purple solution and the presence of black particulates would have indicated aggregation of nanoparticles requiring the synthesis to be repeated. The gold colloid is stable for several months at 4 ^oC and is referred to as a ‘seeds’ solution. The absorbance spectrum of the seeds solution was checked with spectrophotometer for quality control as described previously in Figure 2.6.

The seeds solution was diluted in a solution of sodium citrate and glycerol was added. The optimal seeds dilution factor and glycerol content were determined by previous optimisation experiments on the given substrate surface, as described in sections 2.7.2 and 2.7.3. Sodium citrate concentration was maintained during dilution of the seeds to prevent the colloid crashing out of solution. The diluted seeds solution was then printed as liquid droplets in an array format onto the chosen substrate surface using an inkjet printer (sciFLEXARRAYER S3 from Scienion AG, Berlin, Germany). Unless otherwise stated, all LiScAR experiments in the thesis were performed using arrays produced on an aminated glass microscope slide, referred to as the ‘amino’ substrate by the supplier (Scienion AG, Berlin, Germany) with best results achieved from printing a seeds solution at 1/10th the concentration of the synthesis solution and containing 20% glycerol (by volume). Printed arrays were incubated at 70% humidity for 2 hours to allow the gold nanoparticles to bind to the aminated surface. The printed arrays were washed under flow with deionised water from a tap or wash bottle and dried with N_2 gas.

Auric chloride (500 µL, 50 mM) was added to 50 mL of deionised water. Aqueous hexadecyltrimethylammonium bromide (CTAB) solution (0.2 M) was warmed to 30 ^0C with stirring until fully dissolved (the solution is above the limit of solubility at room temperature). The dissolved CTAB (50 mL, 0.2 M) was added to the auric chloride solution to produce a solution with a brown/orange colour. Ascorbic acid (450 µL, 0.1 M) was then added to the solution which turns colourless and is referred to as growth solution.

The printed arrays were incubated in the growth solution for 30 minutes at 28 ^oC - allowing the seed nanoparticles to grow into efficient light scattering particles. The nanoparticle arrays were washed immediately with deionised water under flow from a wash bottle or tap before standing for 2 days in deionised water. During this time, 3 water changes were performed to ensure all CTAB is removed from the array surfaces. The nanoparticle arrays were dried with N_2 gas and may be stored dry at 4^oC for several months.

2.5 LiScAR Instrument Configuration

The component layout of the LiScAR is shown in Figure 2.7. When installed into the LiScAR, the nanoparticle array was surrounded by a parafilm gasket and a glass cover window is placed on top. The glass cover window had two holes which housed microfluidic tubing with a push fit. The nanoparticle array, parafilm and cover window were assembled in layers to create a flow cell with the input and output holes of the cover window located at opposite sides of the nanoparticle array surface. The flow cell was clamped to a dove prism with index matching fluid between the rear of the nanoparticle array surface and the prism. A 100 mW LED illuminated the dove prism (and hence the nanoparticle array also) with 660 nm red light through a small aperture. Light leaving the prism was captured and dissipated by felt baffles. A monochrome video camera (Lumenera Lu13M) was positioned normal to the array surface to observe changes in the scattered light from the nanoparticles associated with local refractive index changes. The video camera had a ½ inch, progressive monochrome CCD sensor with 1392x1040 pixels (with 2×binning to 696x520) and a dynamic range of pixel brightness levels of 0-4095. The LiScAR components were contained in a blackout hardboard structure to insulate the components from external light sources that would contribute to background noise.

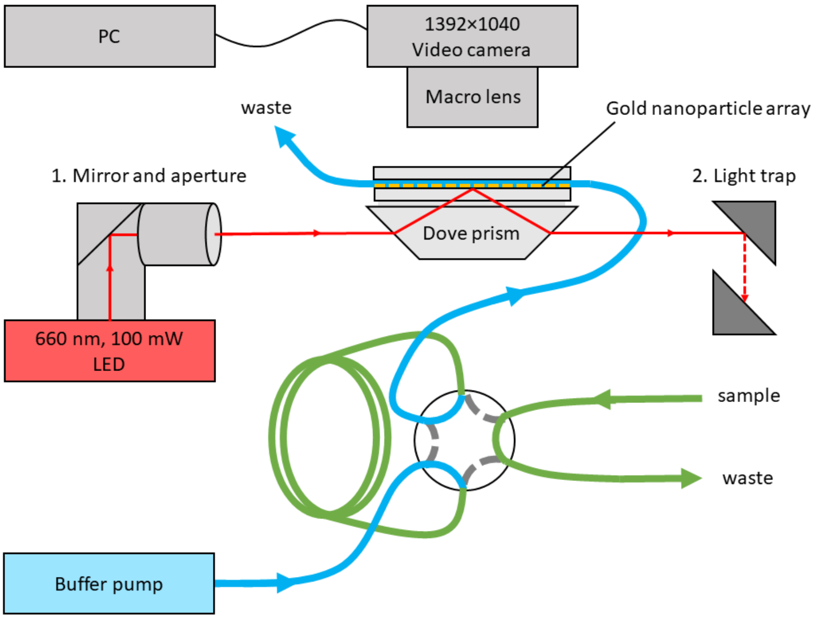


Figure 2.7 Component layout of the LiScAR biosensor platform. Components 1 and 2 were new additions to the inherited setup. Component 1 was used to reduce the form factor of the instrument and ensured the diameter of the light beam is within the cross section of the prism. Component 2 was a felt light trap. Both new components reduced the amount of reflected light returning to the prism which was a source of background noise. Figure adapted from ref (274).

The nanoparticle arrays could be installed in the LiScAR for observation and quality control. The array could be checked visually for signs of non-uniformity which would indicate a discontinuous substrate surface before printing or array damage after printing. The printed volume and hydrophobicity of the substrate surface are factors which, for a given seed solution, determined the size of the printed nanoparticle spot. A homogenous substrate surface should produce spots of a single diameter and brightness, although small variations were often observed. A typical nanoparticle array is shown in Figure 2.8.

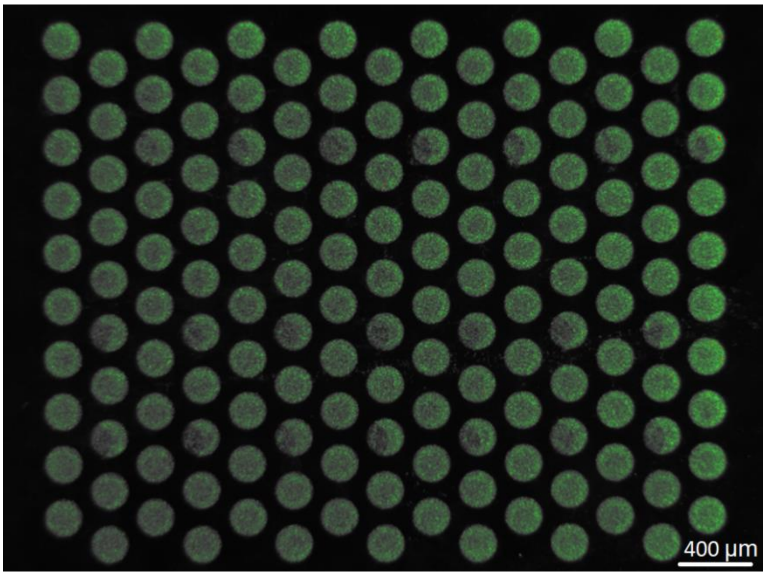


Figure 2.8 False colour image of gold nanoparticle spots as seen by the video camera of the LiScAR biosensor platform ((Lumenera Lu13M). Nanoparticle spots were developed on the amino substrate surface as described in section 2.4.

The brightness of the camera pixels focused on each assay spot were recorded at a frame rate of 30 frames per second (FPS). As discussed previously, the sensitivity of the instrument was improved by maximising the signal to noise ratio. The noise of the brightness signal from each assay spot was reduced by averaging video frames and can be estimated by the equation:

$$noise = \frac{y}{\sqrt{n}} \tag{2.4}$$

where n is the number of measurements and y is the noise without averaging. The effect of frame averaging on the magnitude of noise in the brightness signal of the nanoparticles spots was investigated and results are shown in Figure 2.9.

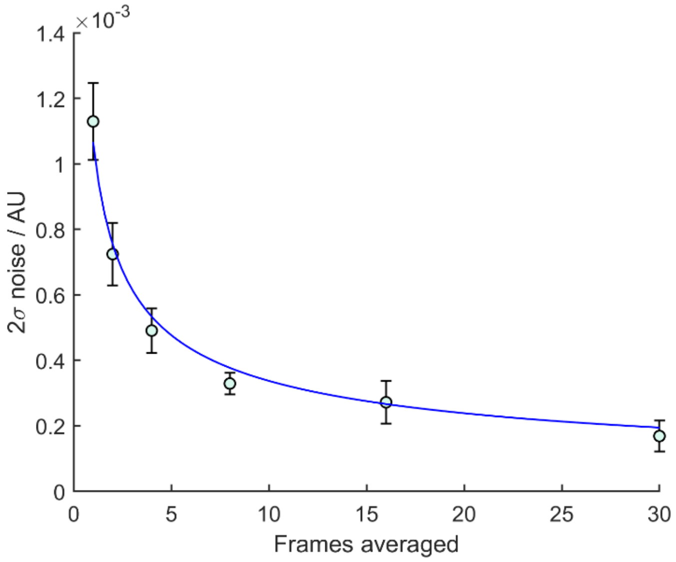


Figure 2.9 The effect of frame averaging on the noise of the nanoparticle spots signal recorded by the LiScAR biosensor platform. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. Data were collected at 30 FPS for 30 seconds at six frame averages: 1, 2, 4, 8, 16 and 30. The data are fitted to equation 2.4 with R^2*=0.99. Error bars are 2SD from n=3 technical replicates.*

The data fit well to equation 2.4, showing a 6-fold reduction in noise when 30 frames were averaged compared to no averaging. As anticipated, noise reduction showed limiting returns with further averaging – indicating that the inherited protocol averaging 30 frames every second is a good compromise between measurement duration and signal to noise ratio.

The established protocol for characterising array spot sensitivity was performed by measuring the brightness change associated with a known change in RI between PBS and double concentrated PBS (2×PBS). When the LiScAR was active, PBS buffer was continuously pumped into the flow cell. The sample loop (Figure 2.7) was loaded with sample and the buffer flow was redirected to push the contents of the sample loop across the array surface. After a period of time, the original flow configuration was restored to fill the flow cell with PBS once more. A sample injection of 2×PBS buffer can be seen as a step change in the brightness of each spot, Figure 2.10.

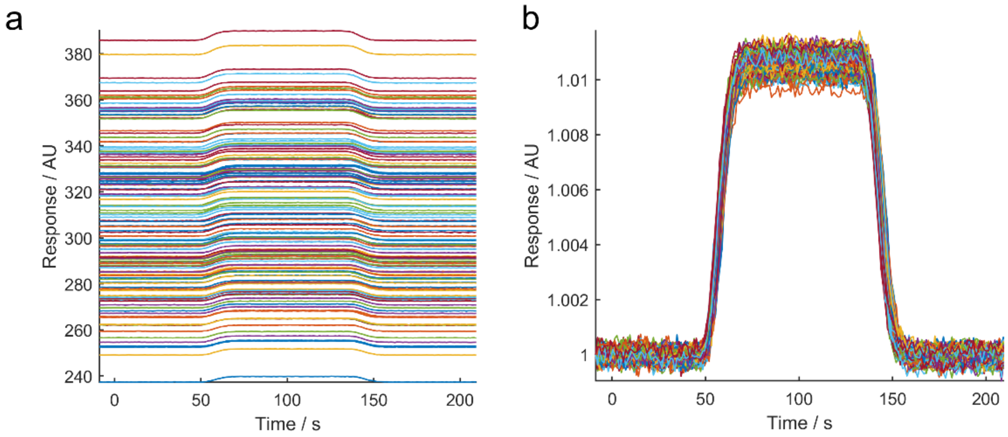


Figure 2.10 The change in spot brightness over time for a typical array of n=150 spots in response to a PBS to 2×PBS switch. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. a) raw data showing the variation in starting intensity of each individual spot. b) data processed by dividing each data point by the first so that each time course has a starting value of 1 (baseline normalisation). The data show that the signal change associated with the PBS to 2×PBS switch is typically ~1%.

The ΔRI associated with the PBS to 2×PBS switch can be measured with a refractometer and has a value of 1.6 mRIU at room temperature. One established measure of RI sensitivity for the LiScAR is nRIS, the percentage change in brightness per RIU, calculated using the equation:

$$nRIS = 100 \times \frac{\Delta\mathrm{signal}}{signal_{baseline} \times \Delta\mathrm{RI}} \qquad 2.5$$

where $\mathrm{signal}_{\mathrm{baseline}}$ is the mean signal under PBS and Δsignal is the change in brightness in response to the known ΔRI. The Δsignal was determined by subtracting $\mathrm{signal}_{\mathrm{baseline}}$ from the mean value of the plateau signal (array in 2×PBS). The PBS to 2×PBS switch was not instantaneous and data were also recorded from a transition period with duration dependent on the rate at which the volume of the flow cell is replaced with the sample of 2×PBS.

2.6 Signal Processing Routine for Refractive Index Sensitivity Calibration

To avoid user bias in deciding which regions should be averaged to determine baseline and plateau in the calculation of nRIS, the present study developed an automated signal processing routine to standardise the procedure. The workflow of the routine was applied to the response data of each and every assay spot and was comprised of several phases:

1. The first and last 10 data points were averaged and the gradient between the two was calculated by fitting the data to a linear equation.
2. The x values of each data point were substituted into the fitted linear equation (with the derived values of gradient and y axis intercept) to produce a theoretical linear baseline across the time range of the data.
3. The theoretical linear baseline was subtracted from the data to correct for signal drift over the time-course of the measurement, which can occur due to small changes in temperature or lamp intensity oscillations.
4. The response range of the data was determined and the data were split into two populations: those with brightness values above half of the response range (the plateau population) and those with brightness values below half of the response range (the baseline population).
5. The baseline population was split into a pre-plateau population and a post-plateau population – identified by a gap in the time values greater than 2 seconds (corresponding to the plateau population).
6. The pre-plateau and post-plateau data were trimmed one data point at a time, starting from the last and first time points respectively, until the data point to be removed had a brightness value lower than the mean of the remaining values.
7. The pre-plateau and post-plateau population were combined to form the overall baseline population.
8. The plateau population data was trimmed from both the start and end time points alternately, until the data point to be removed had a higher brightness value than the mean of the remaining plateau population.
9. The mean and standard deviation of the baseline and plateau populations was calculated.

The output of the automated analysis routine for an array of 150 spots is shown in Figure 2.11.

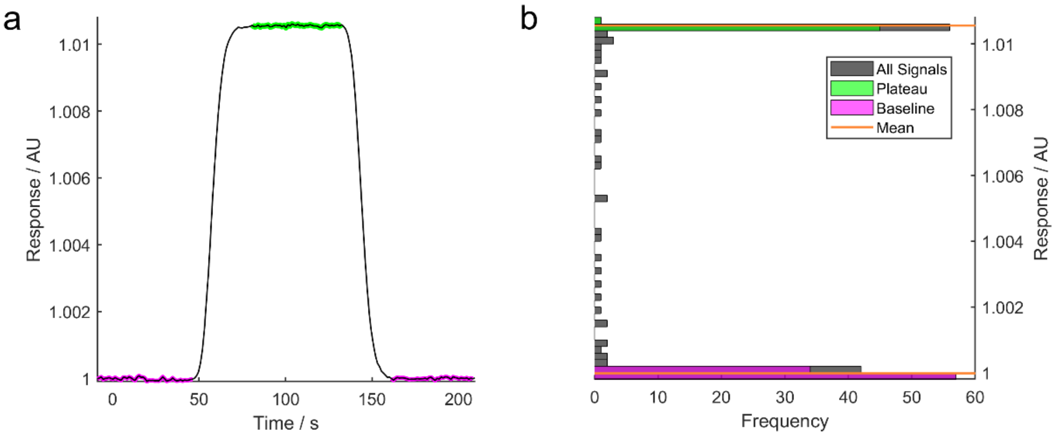


Figure 2.11 Output of the automated PBS switch analysis software performed on the mean signal from a typical array of n=150 spots. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. a) the change in brightness over time, in response to a PBS to 2×PBS switch. The baseline (purple) and plateau (green) regions are determined. b) a histogram of brightness values rotated to match the data of the PBS to 2×PBS switch in (a). The mean values (orange) are determined from data of the baseline (purple) and plateau (green) regions.

The nRIS metric is a noise-independent measure of assay signal but as mentioned previously, it is the signal to noise ratio that must be optimised to maximise assay sensitivity. The Δsignal was calculated by subtraction of $signal_{baseline}$ from $signal_{plateau}$ and hence the noise of Δsignal can be calculated using quadrature:

$$\sigma_{\Delta\text{signal}} = \sqrt{(\sigma_{baseline}{}^2 + \sigma_{plateau}{}^2)} \qquad 2.6$$

where $\sigma_{baseline}$ and $\sigma_{plateau}$ are the standard deviations of the baseline and plateau signals. The values of $\sigma_{baseline}$ and $\sigma_{plateau}$ are assumed to be the same and so equation 2.6 can be simplified:

$$\sigma_{\Delta\text{signal}} = \sqrt{2\sigma^2} = \sqrt{2} \times \sigma \qquad 2.7$$

where $\sigma_{\Delta\text{signal}}$ is the minimum Δsignal which can be observed with a statistical significance of 1 standard deviation. Converting the value of $\sigma_{\Delta\text{signal}}$ to RIU gives the LoB for RI change and is a noise-dependent measure of refractive index sensitivity, RIS:

$$RIS = \sqrt{2} \times \sigma \times \frac{\Delta\text{RI}}{\Delta\text{signal}} \qquad 2.8$$

A further output of the automated PBS switch analysis routine was a visualisation of relative array spot performance. The RIS of each array spot could be visualised in the dimensions of the printed array, to observe effects of spatial spot location on RI sensitivity such as heterogeneity in the aminated substrate. The RIS of each array spot in a typical array is visualised as a function of print order in Figure 2.12.

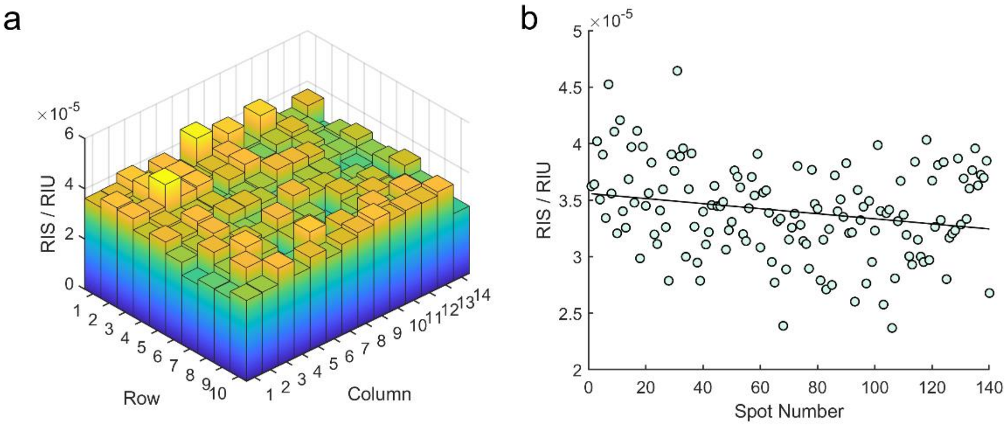


Figure 2.12 The RI sensitivity of each assay spot in a typical array of n=150 spots on a single array. a) RIS dependence on spot location on the array. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. Each bar represents a spot and highlights any variations due to the aminated substrate surface. b) RIS dependence on the order of printing with R^2=0.05 showing no effect of spot number on RIS.

The homogenous distribution of RIS values across the geometry of the array indicated a uniform substrate surface. Furthermore, a lack of correlation between spot number and RIS indicated no time-dependent change in the gold colloid whilst in the printer nozzle.

As will be discussed later, the signals from subsets of spots are typically averaged to produce a single assay channel with lower noise. Ideally all array spots would show identical RI sensitivity but in practice a distribution is observed, Figure 2.13.

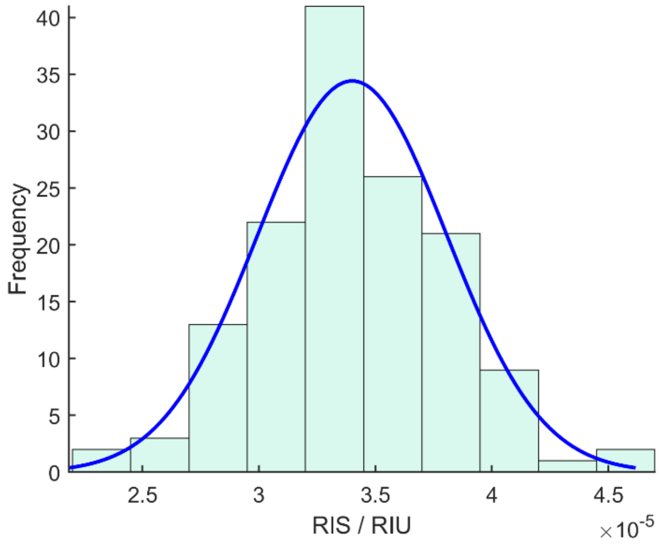


Figure 2.13 Histogram of RIS values for a typical array of n=150 spots on a single array. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. A normal distribution function (blue) fits the data with R^2=0.94.

The range of RIS values across an array was typically <35% above and below a mean of 1-10×10^{-5} RIU after optimising print conditions for the substrate material and batch number, discussed subsequently. For comparison, the commercially available SPR-based instruments from BIAcore may achieve detection limits below 3×10−7 RIU although the instruments typically use temperature controlled sensors to reduce noise (347).

Ideal arrays displayed a normal distribution of RIS values but many arrays showed skewed distributions with subsets of array spots exhibiting poor values of RIS, as discussed later. Spots with RIS values deviating significantly from the normal distribution (>3×SD) were routinely removed from further analyses by the automated analysis software to avoid reducing the response of an averaged signal from a group of spots.

2.7 Refractive Index Sensitivity Optimisation

Ultimately, the RIS of a given nanoparticle spot depends upon a number of variables during sensor surface manufacture which can be explored easily with the automated signal processing routine. A series of experiments were performed to assess the growth of the seed nanoparticles and optimise the established SOP using the RIS performance metric. Quantitative pass/fail limits for quality control (QC) of each array can then be derived.

2.7.1 Seeds Growth Time

The time over which the printed seed particles are grown determines the size of the resulting nanoparticles. The PBS switch analysis routine was used to assess the scattering properties of nanoparticle spots grown for different periods of time using the established SOP in section 2.4. The time dependence of several spot characteristics was investigated and results are shown in Figure 2.14.

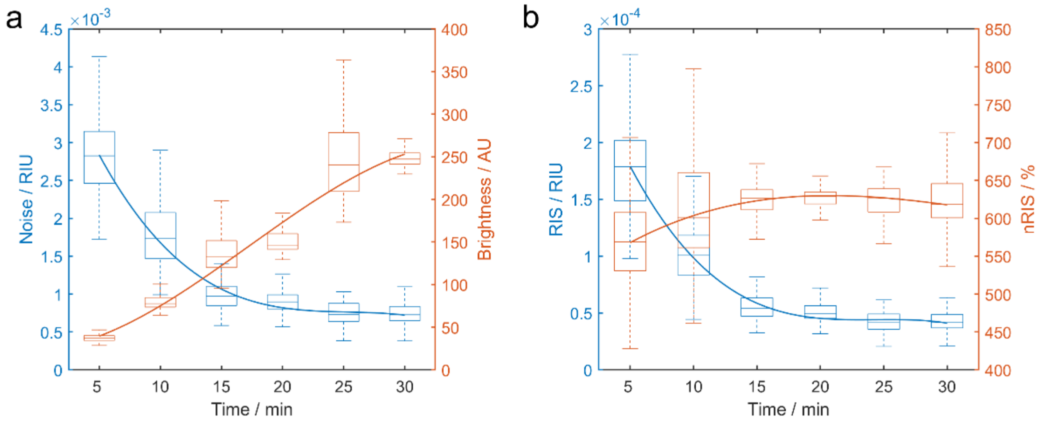


Figure 2.14 The effect of growth time on nanoparticle spot characteristics. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. a) noise and brightness. b) RIS and nRIS. Noise is defined here as 2SD of the background signal. The central mark of each box indicates the median of n=150 spots on a single array and the bottom and top edges of each box indicate the 25th and 75th percentiles respectively. The whiskers extend to the most extreme values. Best fit lines are shown to track the trends in the medians.

As the nanoparticles grew over time, the brightness of the printed spots increased. As shown previously, a negative relationship was observed between spot brightness and signal noise due to the number of photons collected by the camera sensor for a given exposure. The relatively unchanging nRIS and decreasing noise as the nanoparticles grew over time, lead to a proportional decrease in RIS with limiting returns after 25 minutes. The temperature of the growth solution was not investigated but is likely to correlate with growth time and should be investigated in future work for effects on nanoparticle shape. The nRIS did not appear to change significantly over time and may be determined predominantly based on characteristics of the printed seeds such as surface density which must also be optimised.

2.7.2 Nanoparticle Size and Surface Density

The seeds concentration and print viscosity should have a significant effect on the surface density of the nanoparticle spots. High concentrations of seeds should have an increased collision frequency with the substrate surface, increasing the rate of binding and hence the density of immobilised nanoparticles.

In contrast, viscosity is a measure of the friction between molecules of a fluid and higher concentrations of glycerol increase the viscosity of the solution and should reduce the rate of seeds binding to the substrate surface according to the Stokes-Einstein relation (348):

$$D\eta = \frac{k_B T}{6\pi r_H} \qquad 2.9$$

where k_B is the Boltzmann constant (1.38×10^{-23} J K^{-1}) and T is the temperature of the solution. The hydrodynamic radius, r_H, is the radius of an equivalent sphere which would diffuse at the same rate as the particle in solution. η is the viscosity of the solvent and D is the diffusion coefficient, a constant of proportionality between the movement of a particle due to molecular diffusion and movement due to a concentration gradient. Decreasing the temperature and increasing the viscosity or size of the seeds will slow the diffusion of the nanoparticles. Furthermore, the addition of glycerol to the seed solution prevents the printed droplet from evaporating, which would increase the seeds concentration in each printed droplet and the seeds would be deposited onto the substrate surface in their entirety upon complete drying.

The nanoparticle spots were characterised using scanning electron microscopy (SEM) images to observe the physical properties of the nanoparticle spots which may be responsible for variations in RIS. An SEM image of a typical nanoparticle spot on the amino surface is shown in Figure 2.26.

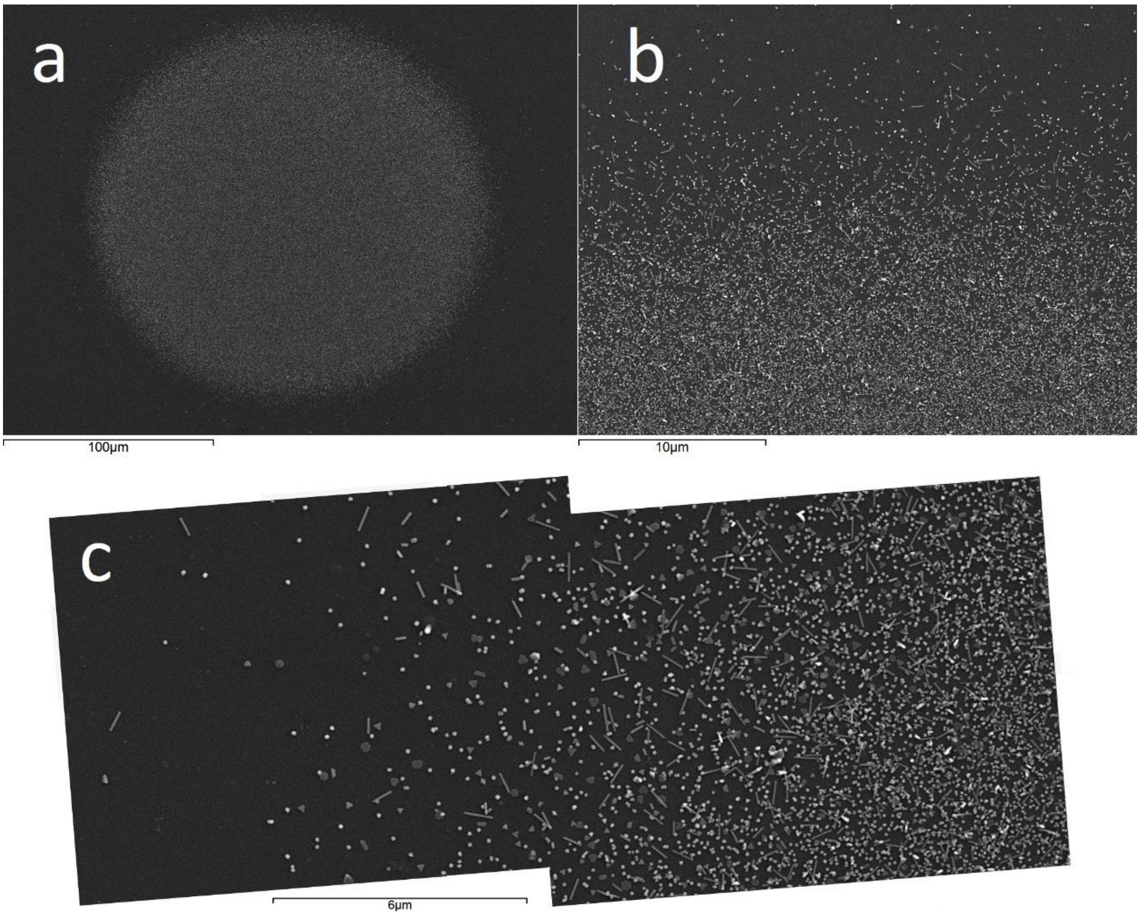


Figure 2.15 Scanning electron microscopy images of nanoparticles grown on an aminated substrate surface. The nanoparticles were developed on the amino substrate surface as described in section 2.4. a) image showing a typical array spot ~200 μm in diameter. b) image showing the edge of an array spot with the surface density of nanoparticles decreasing from centre to edge. c) high resolution image showing nanoparticles grow into a variety of shapes: plates, rods and spherical particles.

The seed nanoparticles grew into a variety of shapes including plates, rods and spherical nanoparticles. Some structures had sharp features (truncated polyhedral and the tips of rods) which may produce plasmons more sensitive to RI. The known magnification of the SEM images enables the dimensions of the nanoparticles to be derived from a given image.

SEM images of nanoparticle spots printed from solutions containing a variety of glycerol and seeds concentrations were analysed systematically with an automated image processing routine. The workflow of the routine was comprised of several stages:

1. A binary version of the original image file was created (all pixels either black or white) based on a brightness threshold set by the user.
2. A flood fill operation filled ‘holes’ in the image. Pixels designated as ‘holes’ were those that could not be reached by filling starting from the edge of the image.

3. Pixels were grouped into objects by an algorithm that looked for 8-connected pixels of the same colour. The pixel coordinates of each object were recorded.
4. The pixels of each object were counted and converted into an equivalent circular diameter (ECD) according to the equation:

   $$ECD = 2 \times \sqrt{\frac{A}{\pi}} \tag{2.10}$$

   where A is the area of the object in pixels.
5. Objects within the upper and lower ECD thresholds set by the user were recorded as a list of ECD values.

The routine was tested using an image of 1000 white squares on a black background and counted the correct number of squares. A typical analysis of a SEM nanoparticle image is shown in Figure 2.16. In the example shown, 136 objects were counted from a total of 63818 objects identified, highlighting the requirement for user-defined ECD thresholds for accurate results and a systematic brightness threshold for all images.

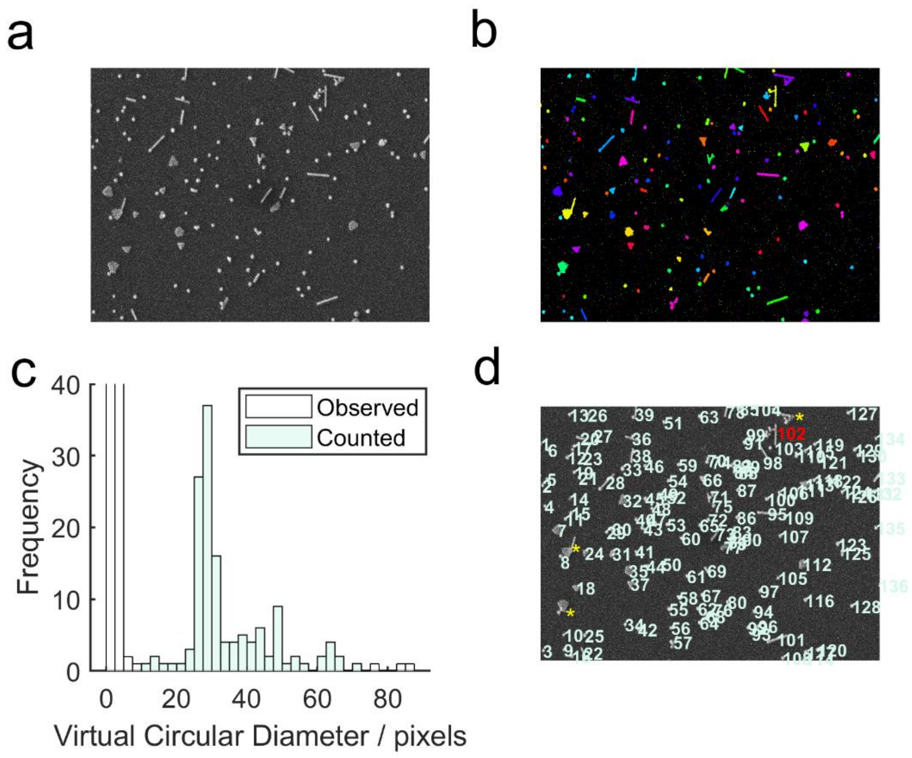


Figure 2.16 Analysis of nanoparticle size and density in a typical SEM image of an array spot. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. a) original SEM image of nanoparticles. b) individual objects in the image are highlighted in different colours. c) histogram of ECD values showing the sizes counted according to the ECD limits set by the user. d) original image with counted object numbers overlaid. Overlapping particles which produce ECD values higher than the upper ECD limit are discounted and shown with yellow asterisks. The object labelled 102 in red is below the upper ECD threshold and counted, despite being a cluster of nanoparticles – a source of error in the routine.

Overlapping particles were a source of error in the analysis routine and may be counted as a single large particle. However, large groups of indistinguishable nanoparticles were removed from the analysis by the upper ECD threshold. Clusters of small particles could be counted if the ECD of the cluster was within the ECD limits, but the original image was overlaid with the counted particle numbers which allowed visual inspection of the results. Nanoparticle clusters represent ~3% of the nanoparticle objects identified in Figure 2.16.

The analysis routine was used to assess images of array spots printed from solutions containing varying concentrations of glycerol and seeds, Figure 2.17. Due to the variation in nanoparticle density across the spot, images were captured from the centre and the edge of each spot.

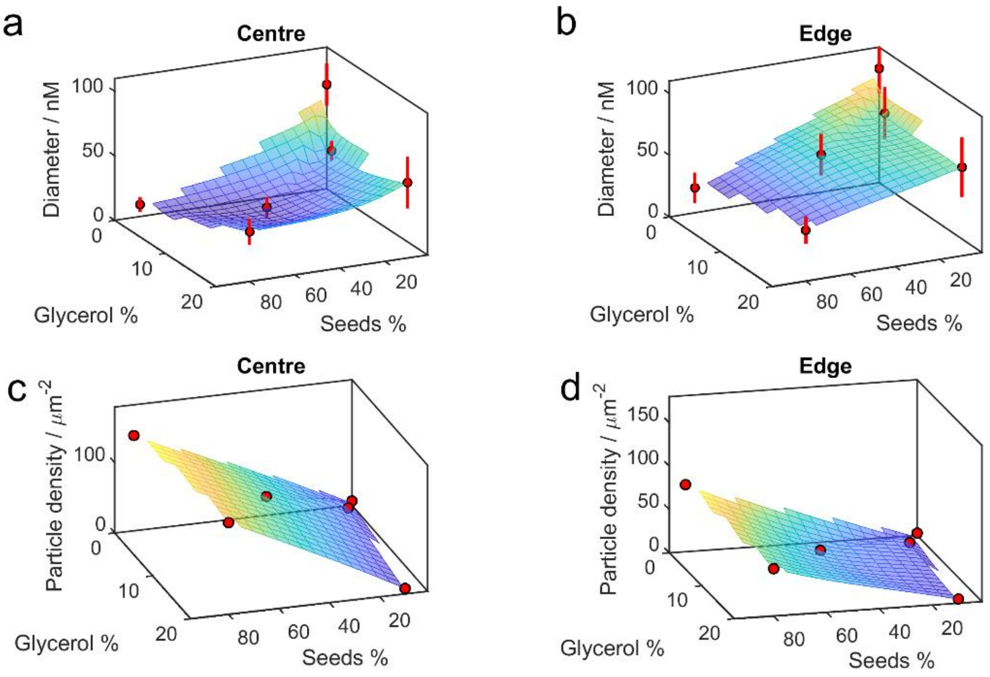


Figure 2.17 Effect of glycerol concentration and seeds solution concentration on the size and density of array spot nanoparticles. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. a) mean ECD of nanoparticles at the centre of an array spot. b) mean ECD of nanoparticles at the edge of an array spot. c) nanoparticle density at the centre of an array spot. d) nanoparticle density at the edge of and array spot. Glycerol concentration is expressed as percentage volume of the solution and seeds concentration is expressed as a percentage of the seeds solution described in section 2.4. Error bars are 2SD of nanoparticles counted (n>100) in images of array spots (n=1) developed on a single array.

In agreement with the Stokes-Einstein relation, nanoparticle density increased with seeds concentration - attributed to higher collision numbers and binding at higher concentrations. However, the effect of glycerol concentration on surface density was not obvious. Nanoparticle density was also greater in the centre of the spot compared to the edge, which may be attributed to droplet deformation after printing. The exposure of the substrate at the edge of the spot to printed seeds solution may be brief upon initial landing a flattening of the droplet and subsequent evaporation.

The ECD of the nanoparticles ranged from 24 to 93 nM, with surface densities of 6 to 170 µm^{-2}. The ECD of nanoparticles was visibly reduced by seeds concentration but glycerol concentration did not have a visible effect. The centre of each nanoparticle spot had a higher particle density and smaller mean ECD, which suggests nanoparticle size may be limited by the available surface area in high density regions. Figure 2.18 shows the effect of seeds and glycerol concentrations on RIS for a single assay spot.

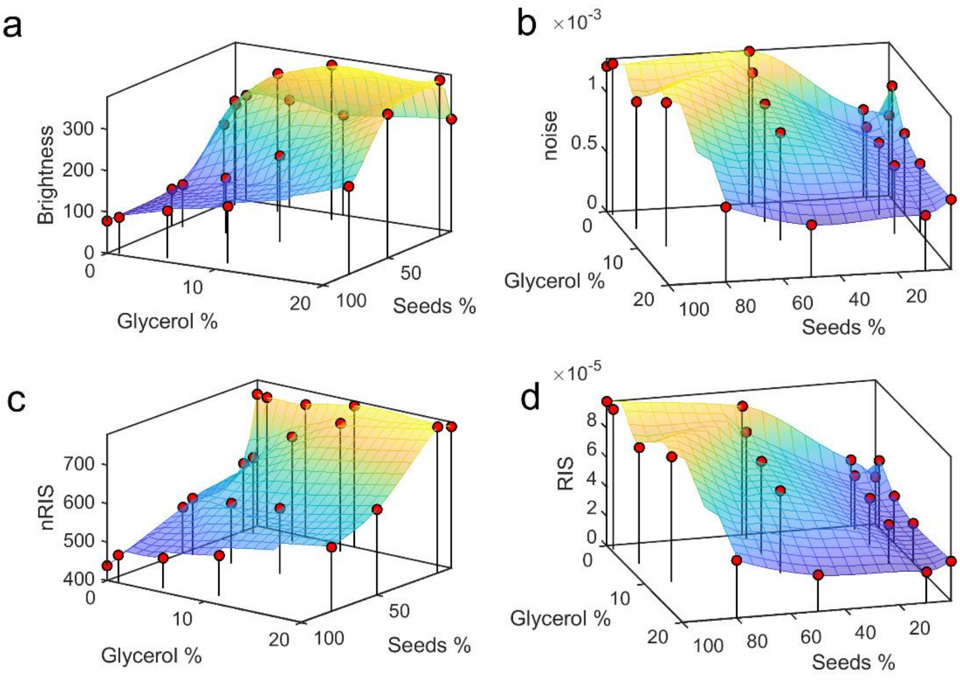


Figure 2.18 Effect of glycerol concentration and seeds solution concentration on nanoparticle spot characteristics. Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. Four spot parameters are reported by the signal processing routine: a) brightness, b) noise, c) nRIS and d) RIS. Data points are derived from the mean response of n=7 assay spots and 2 experimental repeats. Glycerol concentration is expressed as percentage volume of the solution and seeds concentration is expressed as a percentage of the seeds solution described in section 2.4.

As was observed in the nanoparticle diameter and density data, the effect of seeds concentration was greater than the effect of glycerol concentration within the tested ranges. The viscosity range covered by the solutions tested (0-20% glycerol) is approximately 1.01-1.88×10^{-6} m^2 s^{-1} according to the calculation in ref (349), based on initial work in refs (350, 351). Higher glycerol concentrations generally appeared to produce brighter, less noisy, and more sensitive spots, although a RIS minimum the in the box search surface was not identified. Glycerol concentrations above 20% were challenging to print due to the viscosity of the droplets and were not investigated. An optimum region for seeds concentration was found between 1-50%

The lowest RIS for the aminated substrate was achieved at 20% glycerol and 10% seeds with a value of (2.1±0.5)×10^{-5} RIU per spot. The RIS minimum corresponds to densities and ECD values in the ranges 4.8-7.9 μm^{-2} and 59-72 nm respectively (range between mean values for edge and off-centre particles).

The fabrication process in section 2.4 is designed to synthesise gold nanoparticles on any substrate surface and a variety of substrates were explored in an attempt to further improve the RI sensitivity of the LiScAR.

2.7.3 Substrate Material

Plastic substrates such as polypropylene (PP), Cyclic olefin copolymer (COC) and Cyclic olefin polymer (COP) are desirable as LiScAR assay surfaces because of the cost benefit over aminated glass, although an aminated glass substrate is available from Sigma Aldrich at ~10% the cost of the typically used amino substrate described in section 2.4. The viscosity and seeds concentration parameters were box-searched to optimise the printing conditions on the plastic and sigma substrates, Figure 2.19.

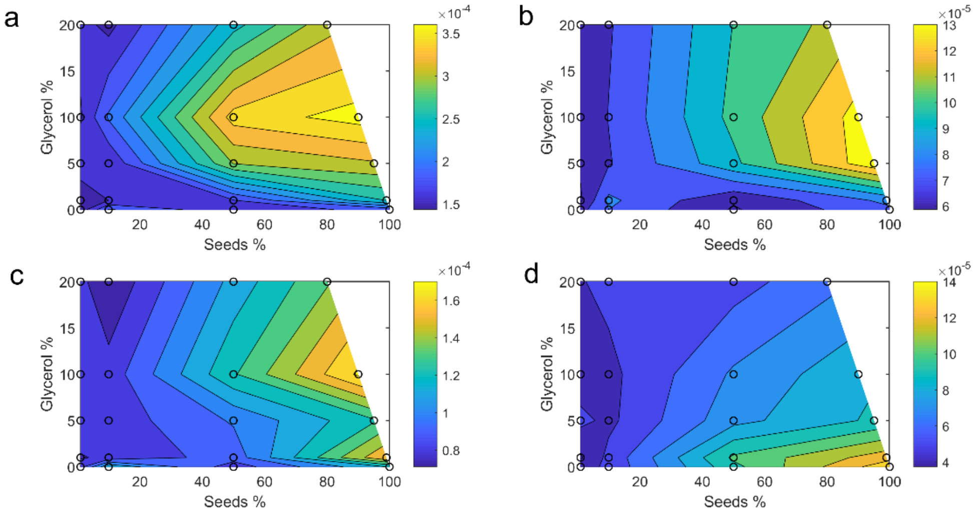


Figure 2.19 Box search of glycerol concentration and seeds solution concentration to optimise RIS on various substrate surfaces. Nanoparticle spots were developed 4 substrates as described in section 2.4: a) PP, b) COP, c) COC and d) Sigma. Glycerol concentration is expressed as percentage volume of the solution and seeds concentration is expressed as a percentage of the seeds solution described in section 2.4. Data points are derived from the mean response of n=7 assay spots.

The substrate onto which the nanoparticle array was printed had a demonstrable effect on the refractive index sensitivity of an array and the optimum combination of seeds and glycerol concentrations varied for different surfaces. The plastic and glass substrates are compared to the amino substrate in Figure 2.20.

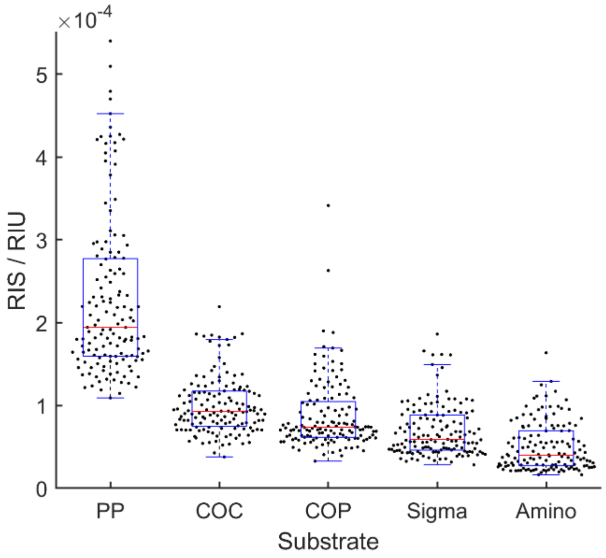


Figure 2.20 RIS values of nanoparticle spots developed on five substrates. Nanoparticle spots were developed using glycerol and seeds solution concentration variations to the printing protocol described in section 2.4 as shown in Figure 2.19. n=150 spots on a single array surface.

The plastic substrates produced arrays with higher variation and worse RI sensitivity compared to glass substrates. The PP substrate in particular displayed a large variation in RIS values, which could be attributed to nonlinear effects of glycerol and seeds concentration on baseline noise or nRIS of the nanoparticle spots. The mathematical relationship between RIS and nRIS can be derived by multiplying both sides of the nRIS equation by RIS:

$$nRIS \times RIS = 100 \times \frac{\Delta\text{signal}}{signal_{baseline} \times \Delta\text{RI}} \times \sqrt{2} \times \sigma \times \frac{\Delta\text{RI}}{\Delta\text{signal}}$$

Solving for RIS gives:

$$RIS = 100 \times \frac{\Delta\text{signal}}{signal_{baseline} \times \Delta\text{RI} \times \text{nRIS}} \times \sqrt{2} \times \sigma \times \frac{\Delta\text{RI}}{\Delta\text{signal}}$$

Simplifying by cancelling ΔRI and Δsignal terms gives:

$$RIS = 100 \times \frac{1}{signal_{baseline} \times \text{nRIS}} \times \sqrt{2} \times \sigma$$

Rearranging the equation gives:

$$RIS = 100 \times \frac{\sqrt{2}}{\text{nRIS}} \times \frac{\sigma}{signal_{baseline}} \qquad 2.11$$

Figure 2.21 shows equation 2.11 fitted to nRIS and RIS data from the nanoparticle spots grown on different substrates.

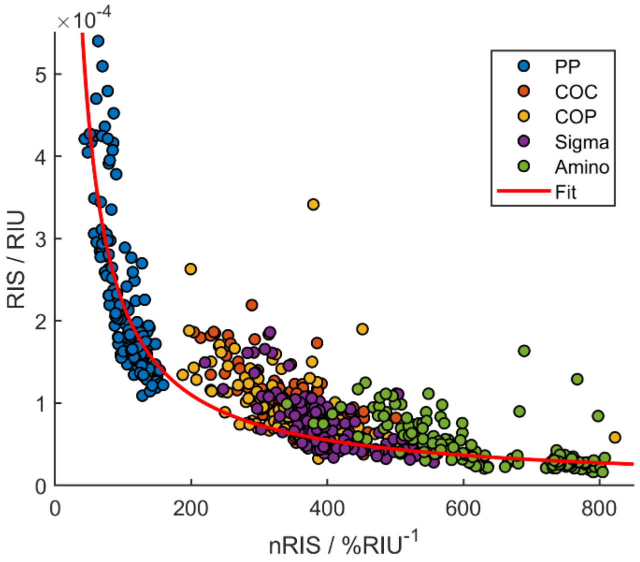


Figure 2.21 The relationship between RIS and nRIS. Nanoparticle spots were developed on a number of substrates using glycerol and seeds solution concentration variations to the printing protocol described in section 2.4 as shown in Figure 2.19. n=150 spots on a single array surface. The mathematical relationship fits the data (red line) with R^2=0.72.

The R^2 of the fit in Figure 2.21 suggests that equation 2.11 accounts for 72% of the variance in RIS. The estimated parameters of equation 2.11 reveal a fundamental property of the LiScAR setup: the ratio of baseline noise (σ) to spot brightness ($signal_{baseline}$) with a value of $(1.69\pm0.04)\times10^{-4}$ (95% confidence). When smaller, the ratio scales the nRIS vs RIS curve vertically such that smaller values of RIS are achieved for a given nRIS: the ratio is minimised when spots are brighter with lower baseline noise.

Plotting brightness against σ shows that spot noise decreased with the square root of sampling number (in this case the number of photons captured by the video camera, apparent brightness) as described previously in equation 2.4, Figure 2.22.

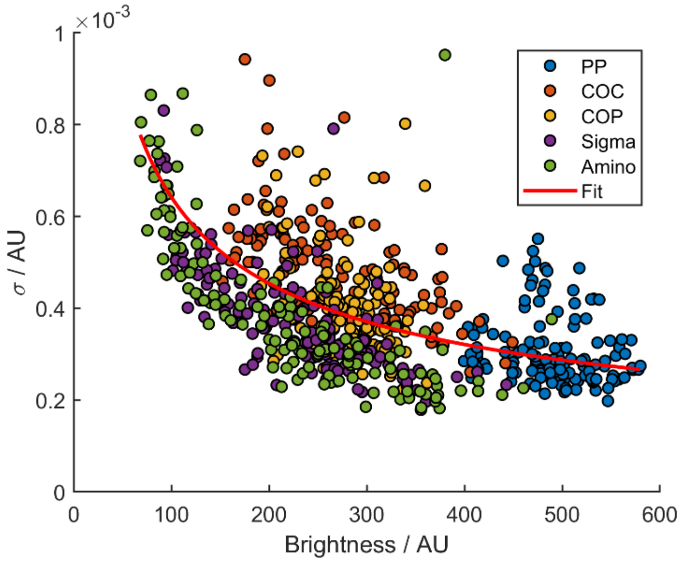


Figure 2.22 Effect of spot brightness on spot noise. Nanoparticle spots were developed on a number of substrates using glycerol and seeds solution concentration variations to the printing protocol described in section 2.4 as shown in Figure 2.19. Spot noise is 1SD for each of n=150 spots on a single array surface. The data fit to equation 2.4 with R^2=0.19 showing noise decreasing with the square root of brightness.

The R^2 of the fit in Figure 2.22 is relatively poor which may be due to the lack of brightness data between zero and 100 AU – the region of the curve where noise is expected to change most significantly. Importantly, Figure 2.22 does not predict which substrate produced the lowest RIS: the worst performing substrate in terms of RIS was PP, despite producing bright spots with low noise - highlighting the contribution of nRIS to RIS variation. nRIS is strongly influenced by the opacity of the substrate which if scratched or contains impurities can scatter light and contribute to absolute background signal - reducing the percentage brightness change caused by the nanoparticles.

In summary, the fabrication procedure for the nanoparticle surfaces was optimised for RI sensitivity on five substrate materials. Optimal results were achieved with a nanoparticle density and ECD of 4.8-7.9 μm^{-2} and 59-72 nm, using 20% glycerol and 10% seeds solution concentrations, the standard amino surface and a growth time of 30 minutes. The optimised SOP derived from the box search produced spots with nRIS = 770±26 % RIU^{-1} and RIS = $(2.1\pm0.5)\times10^{-5}$ RIU per spot. The RIS from the mean response signal of 7 spots (a single column on a typical 10×14 array) was 1.4×10^{-5} RIU. Variation between print batches on amino and sigma substrates lead to typical nRIS values in the range 650-850 and RIS values of $1\text{-}10\times10^{-5}$ RIU per spot.

The array format of the sensor surface and the automated data analysis routines improved the speed of data collection and sensitivity optimisation, allowing new batches of substrate surfaces to be optimised via box search before printing a new batch of gold array surfaces. Future work might investigate the ratio of different shaped nanoparticles on RIS and the temperature of the growth solution. Furthermore, the pH of the printed seeds solution might also be studied for its effect on nanoparticle surface density.

The optimised and calibrated gold nanoparticle array surface can be used to determine the RI of various solutions and this was demonstrated using binary mixtures of isopropanol (IPA) and water, Figure 2.23.

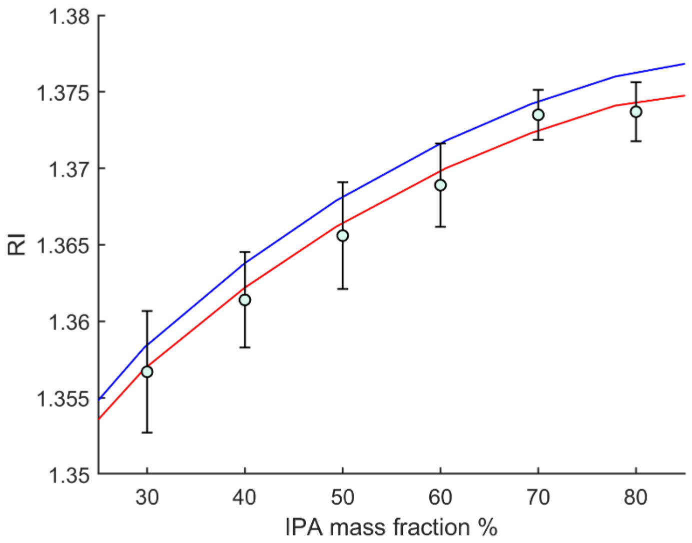


Figure 2.23 RI dependence on the mass fraction of isopropanol in a binary mixture with water. Literature values for the RI at 25 ºC (red) and 20ºC (blue) are in good agreement with the RI derived from the nanoparticle data (green). Nanoparticle spots were developed on the amino substrate surface as described in section 2.4. Error bars are 2SD from n=3 technical repeats.

The data are in good agreement with published values (352) for binary mixtures of IPA and water, validating the nanoparticle array as a reliable technique for measuring changes in RI. The transducer, electronics and display elements of the LiScAR have now been discussed and subsequent sections of the chapter will focus on the bio-receptor element – enabling the LiScAR to record the RI change associated with protein binding interactions.

2.8 LiScAR Bio-Receptor Tethering

The LiScAR is sensitive to RI changes associated with protein binding reactions within 93 ± 10 nm of spherical gold nanoparticles of diameter 90 ± 13 nm (353).

Therefore, analyte sensing in complex samples can be achieved by tethering of the bio-receptors to the gold nanoparticles.

Gold nanoparticles are negatively charged and will adsorb proteins with positively charged residues. The formation of protein corona around nanoparticles is a complex process (354), however it is widely accepted that surface adsorption can lead to structural changes in proteins, including denaturation (355). Proteins may have significantly reduced activity when adsorbed directly to the metal surface, leading to a reduction in assay sensitivity. The aim of protein immobilisation is therefore to maintain protein structure and accessibility of the relevant binding sites (356). Furthermore, many bio-sensors are limited by non-specific binding of background materials (fouling) when sampling complex mixtures such as human serum (357). The inherited LiScAR protocol uses a self-assembled monolayer (SAM) which insulates the gold surface from immobilised proteins and also serves as an anti-fouling layer.

SAMs are mixtures organic molecules which spontaneously form ordered layers on surfaces (358). SAMs on planar gold surfaces are widely used in the construction of a variety of surfaces for use in biosensors and nanomedicine (359). The SAM used in the studies reported here bound to the gold nanoparticles via thiol groups in a complex process, involving the dissociation of the S–H bond followed by the formation of the Au–S covalent bond (360). There were two components to the SAM, termed linker and spacer. The linker component terminated with a carboxylic acid group, which was used to immobilise proteins with a covalent peptide bond. The spacer component terminated with a hydroxyl group and formed an antifouling layer on the gold surface. The ratio of linker to spacer could be used to control the density of immobilised proteins. Repeated ethylene glycol (PEG) groups were common to both linker and spacer elements which prevented the non-specific adsorption of proteins (361). PEG is also an amphiphilic polymer which is surrounded by an organised structure of water molecules when hydrated, leading to hydrophilic surface passivation (362). A schematic of the SAM is shown in Figure 2.24.

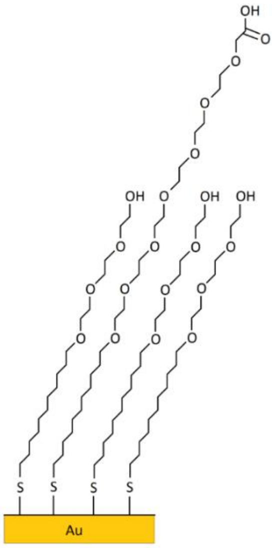


Figure 2.24 Schematic of the self-assembled monolayer bound to a gold surface via thiol groups. The spacer component is capped with a hydroxyl group and provides an antifouling layer over the gold surface. The linker component is capped with a carboxylic acid group which may be activated for tethering to proteins. The linker has a larger number of repeated PEG units than the spacer, exposing the carboxylic acid group above the SAM surface.

Most proteins contain an abundance of lysine amino acids with nucleophilic primary amines as R groups. Carbodiimide compounds are used for efficient crosslinking of these primary amines to the carboxylic acid terminus of the SAM linker component. Carbodiimides are considered zero-length crosslinkers because no portion of their chemical structure becomes part of the final bond, Figure 2.25.

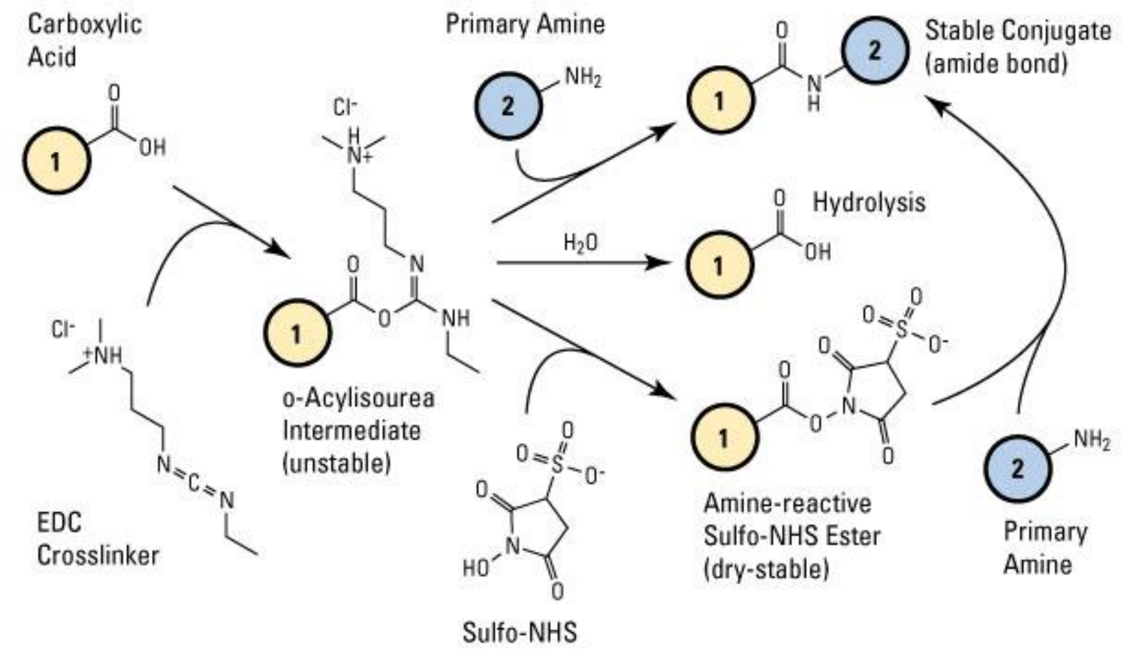


Figure 2.25 Schematic of EDC/NHS coupling chemistry. SAM linker (1) is tethered to the primary amine of a protein (2). The O-acylisourea intermediate is readily hydrolysed and the synthesis typically includes N-hydroxysuccinimide to extend the half-life of the intermediate. Image from ref (363).

In the present work, the carbodiimide compound 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) was used to activate carboxyl groups on the SAM linker, forming an active O-acylisourea intermediate that was easily displaced by nucleophilic attack from primary amino groups on the protein structure. The O-acylisourea intermediate is unstable and is readily hydrolysed by water. Therefore, N-hydroxysuccinimide (NHS) was included in the reaction mixture to form a dry-stable NHS ester intermediate that is less readily hydrolysed but still reactive towards primary amines (363).

2.8.1 Assay Method

The method of immobilising proteins on the LiScAR sensor surfaces was inherited from the research group and is used for all LiScAR experiments in the thesis unless otherwise stated. Nanoparticle arrays were developed on the amino substrate as described in sections 2.4 before washing with ethanol. The SAM components were mixed in a 10:1 spacer to linker ratio at 2 mM in ethanol and the nanoparticle arrays were submerged in the solution for a minimum of 4 hours to allow uniform SAM formation. EDC (20 mg) and NHS (35 mg) were dissolved in 1 mL of dimethylformamide (DMF) before dilution with 20 mL IPA. SAM coated nanoparticle arrays were washed with IPA and left to stand in the EDC/NHS solution for 20 minutes. The nanoparticle arrays were washed rapidly with IPA followed by water under flow and then dried with N_2 gas. Protein solutions at >0.5 mg/mL in PBS were printed on top of specific groups of nanoparticle spots using an inkjet printer (sciFLEXARRAYER S3 from Scienion AG, Berlin, Germany) and become tethered to the activated SAM surface. The arrays were left in the printer at room temperature for 2 hours at 70% humidity before storage at 4ºC.

A functionalised array may be enclosed in the flow cell configuration discussed previously and stored in PBS running buffer. A flow rate of 6 mL/hr was typically used for binding experiments to eliminate mass transport effects (276). The SOP of the LiScAR is the present study was as follows:

1. The array was washed with 10 mM phosphoric acid solution to remove unbound proteins simply adsorbed onto the gold nanoparticle surface.
2. The array was blocked for 300 s with 1 mg/mL BSA solution in PBS to fill any spaces of the gold surface that would otherwise enable nonspecific binding of sample proteins.
3. The array was washed again with 10 mM phosphoric acid solution to remove BSA weakly adsorbed onto the gold nanoparticle surface.
4. A 2×PBS switch was performed, enabling calibration of sensor response to RI and RIS to be determined by the automated signal processing routine described previously.
5. A sample was injected to observe analyte association to the assay surface in real time: the association phase.
6. The flow was switched back to PBS running buffer to observe analyte dissociation: the dissociation phase.
7. A 200 nM sample of detection antibody (described in detail in 3.3.2) may be injected at this stage and the AUC may be calibrated for analyte concentration determination.
8. The array was washed with 10 mM phosphoric acid solution to remove bound analyte.

The array surface is reusable and steps 5-7 may be repeated in a duty cycle typically lasting 10 minutes in total. The brightness data were recorded for each spot and calibrated to RIU using the automated signal processing routine described in section 2.6. The response data from each spot were normalised to a starting value of zero and the data from multiple spots were averaged to produce a single kinetic trace of analyte binding with relatively low background noise compared to the individual spot responses. The number of spots used for averaging was typically 14, corresponding to a single row of a 14×10 array. The surfaces may be tested using a standard IgG reference material from NIST which binds to array spots functionalised with the Fc binding protein A/G (PAG), described in Chapter 3. Data from a typical analyte capture experiment are shown in Figure 2.26, comprising of an association phase followed by a dissociation phase.

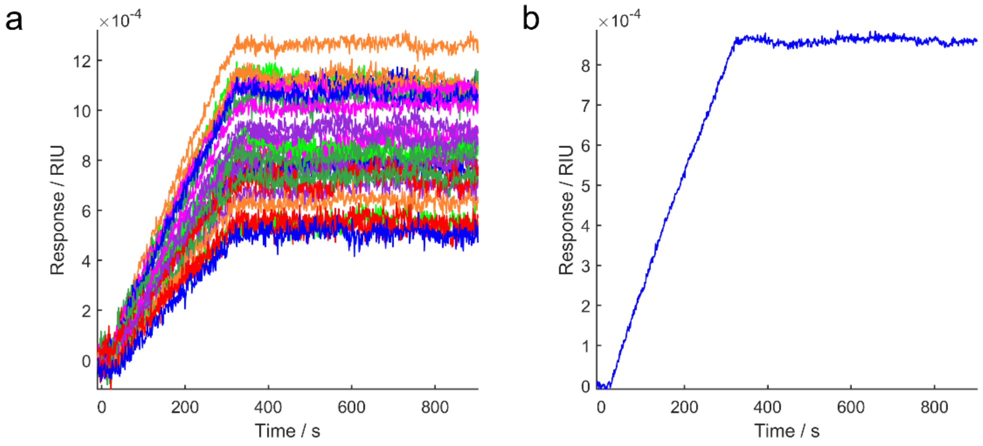


Figure 2.26 Kinetic response curves of NIST IgG (3.2 nM in PBS) binding to PAG on the LiScAR sensor surface. PAG was immobilised using the assay method described in 2.8.1. a) the signal from n=27 individual assay spots. b) the averaged signal from the same n=27 assay spots.

The data in Figure 2.26 show significant variation in the response of each assay spot. Inter-spot variation occurs as a result of variations in surface functionalisation which depends on the accuracy with which droplets of protein are printed onto the gold spots. Sub-optimal printing accuracy leads to gold nanoparticle regions with no functionalisation and hence no sensitivity to analyte, which results in a lower mean brightness change recorded by the LiScAR for each spot. Printing accuracy can be checked qualitatively by observing a newly printed array through a microscope lens. Alternatively, printing accuracy can be interrogated quantitatively by calculating the unresponsive pixels in a nanoparticle spot image as a percentage of the total spot pixels.

The raw kinetic trace is sensitive to changes in temperature, physical disturbances of the surface supporting the instrument and variations in illumination intensity of the LED. To observe these effects, a group of control spots are always included on each assay, printed from a 1 mg/mL BSA solution. BSA is an appropriate control protein for healthy human serum samples due to its stability and lack of specific binding affinity for human proteins. However, serum antibodies to BSA may develop as part of a patient's immune response (at nanomolar concentrations) following exposure to BSA which is used as a surgical sealant in procedures such as vascular reconstructions (364). For serum samples containing anti-BSA antibodies, a different control protein could be used such as human serum albumin. As with the assay spots, the signals of the control spots were normalised and averaged into a single channel which showed all changes in signal that were not due to the binding of analyte. The signal of the control spots was subtracted from the signal of the assay spots, producing a kinetic trace which did not drift and responded predominantly due to specific analyte binding, Figure 2.27.

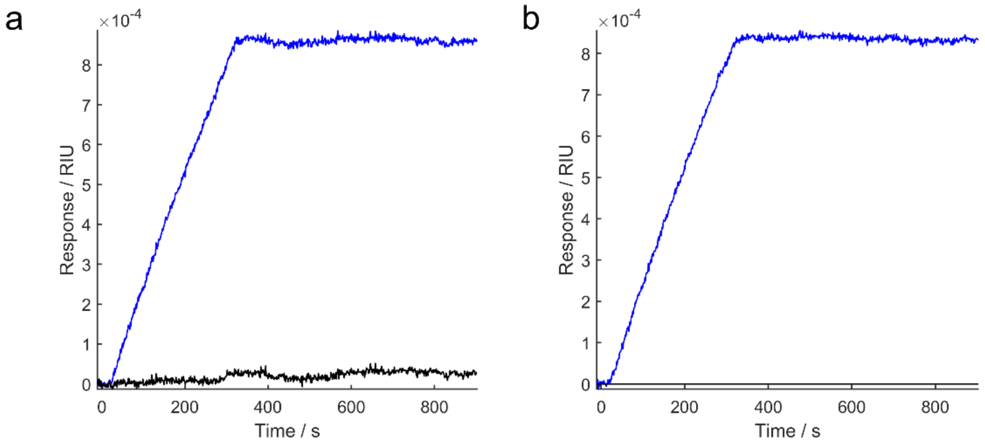


Figure 2.27 The effect of subtracting the assay response signal from a reference channel. PAG and BSA were immobilised using the assay method described in 2.8.1. a) raw assay signal of the NIST mAb (3.2 nM) binding to PAG assay channel (blue) and BSA reference channel (black). b) assay signal after reference signal subtraction. Signals are the averaged response of n=27 assay spots.

PAG spots and BSA spots were located randomly across the array to ensure these positive control and reference channels were representative of sensitivity and nonspecific signal changes respectively at all assay locations. The number of positive control and reference spots used for averaging was chosen to match the number of spots in each assay channel - ensuring the averaged signals from both assay and control spots displayed similar noise. The rate of analyte binding recorded by the assay spots can be analysed to derive the kinetics of the interaction and the maximum analyte load of the assay surface.

2.9 Kinetic analysis of sensor response

The competing processes of binding and dissociation of analyte at the assay surface are evident in the distinctive shape of the sensor response. The standard model for filling a biosensor surface is that of the Langmuirian adsorption isotherm (365), which describes a finite number of binding sites on the assay surface being filled by analyte over time. The assumptions of the Langmuirian adsorption isotherm are: 1) all sites are assumed to be identical; 2) each site can hold a maximum of 1 analyte; 3) there is no interaction between neighbouring binding sites and 4) the adsorption process must represent a dynamically reversible process. Initially, in what is termed the association phase, the rate of analyte binding to an unfilled assay surface is high and decreases as the number of available binding sites are filled up by analyte: leading to a horizontal asymptote. The initial rate of analyte association is therefore dependent on the concentration of analyte and available surface sites: a second order reaction with rate constant, k_a, having units of M^{-1} s^{-1}. The rate of analyte dissociation from the assay surface is independent of analyte concentration: a first order reaction with rate constant, k_d, having units of s^{-1}. Analyte dissociation occurs simultaneously to binding during the association phase and the overall rate of assay surface filling may be described by a differential equation showing the balance between the association and dissociation rates:

$$\frac{d\theta}{dt} = k_a[P](1 - \theta) - k_d\theta \qquad 2.12$$

where θ is the surface coverage and 1-θ is the remaining unfilled binding sites as a fraction of total binding sites; P is the concentration of protein analyte in solution and k_a and k_d are the association and dissociation rate constants respectively. The positive term describes association of analyte, whilst the negative term describes dissociation. The surface coverage at a specific time during association, $\theta(t)$, may be estimated by integrating equation 2.12 to give:

$$\theta(t) = \frac{\vartheta(t)}{\vartheta_m} = \frac{k_a[P]}{k_a[P] + k_d}(1 - e^{-(k_a[P]+k_d)t}) \qquad 2.13$$

where $\vartheta(t)$ is the experimentally observed change in signal at time t and ϑ_m is the signal that would be observed at surface saturation. ϑ_m is not attainable experimentally due to dissociation as shown in equation 2.13 and an equilibrium surface coverage is observed in practice. The relative values of k_a and k_d describe strength of the binding interaction and define the affinity constant, K_D, via the equation:

$$K_D = \frac{k_d}{k_a} \qquad 2.14$$

K_D is equal to the concentration of analyte required such that 50% surface coverage is maintained at equilibrium. Quantitative measures of protein-protein interactions are critical to understanding protein functions (366) and screening for new antibodies (367) for use as high affinity bio-receptors. Furthermore, investigating the relationship between binding affinity and protein structure is a fundamental component of drug design (366).

2.9.1 Parameter Estimation for the Langmuir model

The k_a, k_d and ϑ_m parameters of the Langmuir model can be estimated by fitting the model to RI data from the association and dissociation phases of a protein binding experiment on the LiScAR (368). The RI data are used to fit equation 2.15 for the association phase (equivalent to equation 2.13 scaled by ϑ_m) and equation 2.16 for the dissociation phase:

$$\vartheta(t) = \vartheta_m \frac{k_a[P]}{k_a[P] + k_d}\left(1 - e^{-(k_a[P]+k_d)t}\right) \qquad 2.15$$

$$\vartheta(t) = \vartheta_a e^{-k_d t} \qquad 2.16$$

where ϑ_a is the signal at the start of the dissociation phase. Model fitting techniques, parameter estimation and goodness of fit analysis are discussed in Chapter 4. Briefly, in order to estimate model parameters with good precision, the training data must represent the effects of all parameters of the model. For example, to best estimate the value of ϑ_m, ideal fitting data would show analyte binding until maximum surface coverage is achieved. The data would constrain the model such that a good fit would only be achieved over a small range of possible ϑ_m estimates. Similarly, a poor estimate of k_d will occur if the data used to the fit the model do not show significant analyte dissociation – a situation which can occur if the noise of the data exceeds the gradient of the dissociation phase. When the model is fitted to a single measurement of association and dissociation, the parameters of the model are poorly determined but the precision of parameter estimation is markedly improved by fitting the model to data from a range of analyte concentrations simultaneously: a global fit. A Langmuir model global fit to

LiScAR data of IgG binding to spots functionalised with PAG is shown in Figure 2.28.

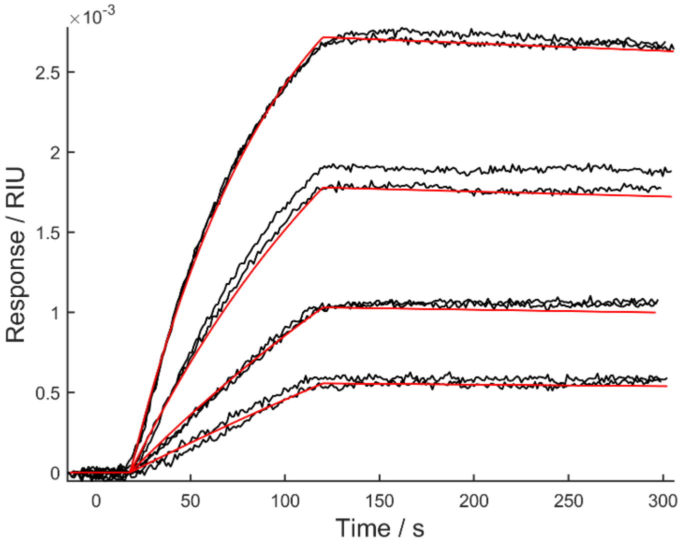


Figure 2.28 Global fit of the Langmuir model (red) to kinetic data from IgG binding to immobilised PAG (black). PAG was immobilised using the assay method described in 2.8.1. Signals are the averaged response of n=20 assay spots. Duplicate concentrations of IgG were sampled (6.25, 12.5, 35 and 50 nM) and higher concentrations induce greater response signal. The association phase shown here begins at 20 seconds and lasts 100 seconds in total. The dissociation phase shown here begins at 120 seconds and lasts 200 seconds. The fit estimates for the interaction are: $k_a=3.26 \pm 0.03\times10^5\ M^{-1}\ s^{-1}$ *and* $k_d=3.01 \pm 0.37\times10^{-4}\ s^{-1}$.

The precision of parameter estimation in the global fit comes from the increased constraints by the data on the model parameters. The higher concentrations provide a better estimate of ϑ_m, but increase the likelihood of collecting data which violates the assumptions of the Langmuir model such as cooperativity between binding sites and aggregation rather than 1:1 binding kinetics. In contrast, lower concentrations provide good quality kinetic data but may not show enough curvature during association to precisely define ϑ_m. The global fit enables all parameters to be estimated with improved confidence over a single kinetic trace.

Protein adsorption data often appear to be fit well by the Langmuir model (369). However, incorrect parameter estimates will be made if the data are collected from a process which violates any assumptions of the model. These limitations are important to understand before attempting to fit the Langmuir model to kinetic data.

2.9.2 Limitations of the Langmuir model

The assumption that all binding events on the surface are equal is easily violated by nonspecific binding. Whilst antibodies bind with affinity constants of

$10^{-10} - 10^{-11}$ M to specific epitopes (310), nonspecific binding of antibodies to low affinity targets is likely to occur at some binding sites at higher concentrations of analyte (~10^{-3} M) (370). When nonspecific binding occurs, the gradient of the association phase may be increased due to additional mass binding to the surface. The gradient of the dissociation phase is also typically increased initially as the nonspecific interactions often have rapid dissociation rates. Additional phenomena which occur at higher analyte concentrations are protein clustering and aggregation which will violate both the 1:1 binding assumption and the non-cooperativity assumption. Protein clustering and aggregation are discussed in detail in Chapter 3.

The effects of nonspecific binding, aggregation and clustering may or may not be evident in the shape of the kinetic data. Hence, to minimise nonspecific binding effects it is important to use the lowest concentrations of analyte that provide sufficient surface coverage to estimate all parameters of the Langmuir model. Nonspecific binding of analyte to the non-biological regions of the assay surface is also reduced by the incorporation of additional blocking steps during assay preparation. Step 2 of the standard operating procedure may be repeated to ensure hydrophobic and electrostatic interactions between analyte and the gold / SAM surface are effectively blocked before complex solutions are sampled (371).

Matrix effects are effects on an analytical technique caused by any elements of a sample that are not the analyte (304). Matrix effects account for the nonspecific effects mentioned previously and include the phenomenon of molecular crowding (372). The volume occupied by proteins in solution is unavailable to others because two molecules cannot occupy the same space at the same time. In crowded biological samples, the steric hinderance of large proteins effectively limits the volume available to smaller proteins, leading to discontinuous protein concentrations. Matrix effects may also affect binding rates by altering the rate of diffusion (mass transport) in the solution. For instance, a sample containing a relatively high concentration of albumin will exhibit a lower rate of diffusion onto the assay surface due to higher viscosity, as shown in 2.9. Matrix effects may be accounted for by diluting samples using a solution with a similar matrix to the sample - ensuring that any effect on the assay is consistent for all measurements regardless of dilution factor. Alternatively, matrix effects such as excluded volume and diffusion limitations will be minimised at relatively high sample dilutions in PBS buffer.

In theory, once the parameters of the Langmuir model are estimated using a global fit to kinetic data, the concentration of analyte in unknown samples may be estimated based on association and dissociation data. However, this approach is challenging due to errors in the estimated parameters of the Langmuir model and matrix effects of samples in complex mixtures. In practice, the area under the curve (AUC) of the association phase is calibrated to analyte concentration and used to construct a standard curve which is independent of the Langmuir assumptions.

Because the assay response cannot be measured at every concentration of analyte within the dynamic range of the assay, a mathematical function must be fit to the data which enables interpolation between the standard concentration response values. Selection of the standard curve function is critical to achieve accurate assay result. For example, a linear function can be applied to the linear range of the assay shown in Figure 2.1 but will not fit well to the sigmoidal relationship exhibited by a larger range analyte concentrations. In contrast, over-parameterised models may enable a perfect fit to the data but are subject to noise and may provide inaccurate interpolation between data points (known as overfitting) that does not represent the “true curve” (373). This is an example of the bias-variance trade-off commonly encountered in model fitting. The ideal standard curve function should fit well to the data with the smallest number of parameters and the 5-parameter logistic model is commonly applied to immunoassay data (373):

$$y = \mathrm{D} + \frac{\mathrm{A} - \mathrm{D}}{\left(1 + (x/C)^{B}\right)^{E}} \qquad 2.17$$

where A is the minimum asymptote (the response value at zero analyte), C is the inflection point and D is the maximum asymptote (the response value at infinite analyte concentration). Parameters B and E define the steepness and the asymmetry of the sigmoidal curve respectively. Figure 2.29 shows the calibration process for an IgG assay (based on nanoparticle spots functionalised with PAG) using the 5-parameter logistic equation.

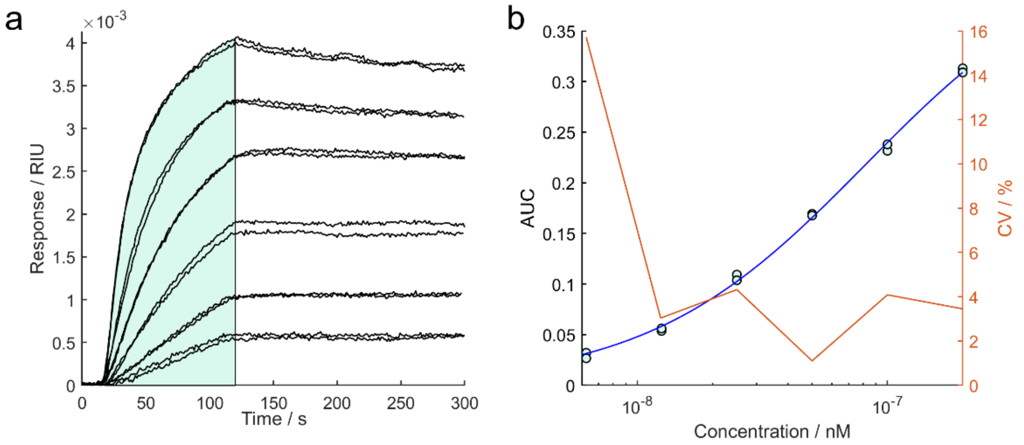


Figure 2.29 Calibration data of IgG binding to immobilised PAG. PAG was immobilised using the assay method described in 2.8.1. Signals are the averaged response of n=20 assay spots. Duplicate concentrations of IgG were sampled (6.25, 12.5, 25, 50, 100 and 200 nM) and higher concentrations induce greater response. a) kinetic response data (black) showing association and dissociation phases and the AUC region (green). b) standard curve of IgG standard concentration against AUC (green) which can be interpolated with a 5-parameter logistic standard curve (blue) to determine the concentration of IgG in an unknown sample. The CV at each concentration (orange) indicates the analytical precision across the curve.

The data show good reproducibility in the relationship between IgG concentration and AUC. The standard curve was used to derive the IgG concentration of the standards based on their AUC as a measure of assay performance. A maximum CV in derived concentration of 16% occured at the lowest standard concentration (6.25 nM). The CV plot allowed the LLoQ to be determined which for the example shown was 12.5 nM - based on a chosen CV threshold of 5%. For comparison, Enzo ELISA kits state that intra-assay CV should be less than 10% (374). The CV in derived concentration should increase as the IgG concentration approaches the ULoQ, which for the example shown was >200 nM IgG. In practice a new standard curve is performed for each LiScAR assay but in future batch calibration may be possible with highly reproducible assay surfaces.

2.10 Conclusions

These results show that the LiScAR is an appropriate biosensor platform on which to develop assays for studying the Complement cascade. The fabrication process of the LiScAR assay surfaces has been discussed and the inherited protocol was developed to improve the RI sensitivity of the nanoparticle spots. Furthermore, potential violations of the Langmuir model assumptions have been considered and solutions presented to ensure the validity of estimated kinetic parameters. The development of automated signal processing routines enabled rapid and standardised assay calibration and operation and is a first step towards enabling non-specialized technicians to operate the LiScAR - a major advantage for all biosensors. The optimised experimental methods and standard operating procedures for the LiScAR were used to produce reliable data for the studies presented in the subsequent chapters. Chapter 3 presents the results of antibody tethering to the LiScAR sensor surface to develop novel protein assays specific to biomarkers of Complement cascade activation and reports a method for antibody sample purity testing using the PAG surface presented here.

3 Antibody Properties in the Development of New Biophotonic Assays of Complement Activation

3.1 Introduction

Chapter 2 demonstrated that proteins can be printed onto the LiScAR assay sensor surface in an array format and tethered using EDC/NHS coupling chemistry, one of the most commonly used strategies for crosslinking proteins to gold nanoparticles (375). BSA and PAG were routinely printed to produce control spots and a positive control assay for IgG concentration, respectively. In theory, the same methods for protein immobilisation could be applied to immobilise any commercially available antibody to produce an assay for its antigen. The frequency of lysine amino acids in proteins is ~7% (376), suggesting that a number of amines are available for EDC/NHS tethering on most proteins. The structure of IgG is shown in Figure 3.1.

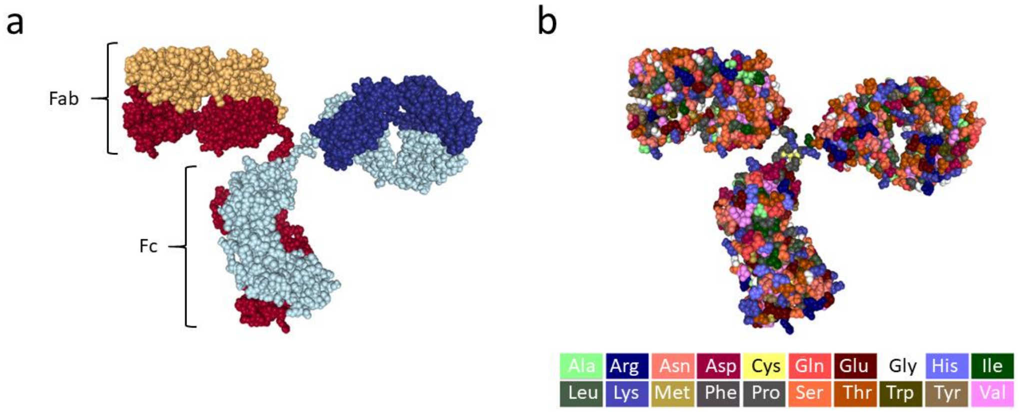


Figure 3.1 Structure of IgG. a) spacefill model coloured by chain showing heavy chains (light blue and red) and light chains (dark blue and orange); and b) spacefill model coloured by amino acid. Image created with NGL Viewer (377). Protein data bank ID: 1IGT.

The IgG structure is comprised of four major regions: two fragment antigen binding regions (Fab), the fragment crystallisable region (Fc) and the hinge region (378). The Fc region interacts with cell surface receptors and is connected via the hinge region to two Fab regions which bind to antigen. It is therefore the Fab regions which provide bio-recognition in immunoassays such as the LiScAR platform. There are typically 80 lysine residues present in IgG with ~50% accessible for tethering (379). Trastuzumab emtansine (trade name Kadcyla) and ibritumomab tiuxetan (trade name Zevalin) are examples antibody-drug conjugates (ADCs) in which antibody is conjugated via its lysine residues to ‘cytotoxic payloads’ (380) allowing targeted delivery to cancer cells (381).

The single flow cell of the LiScAR platform exposes all assay elements of the biophotonic sensor surface to a given sample, allowing multiple assays to be performed simultaneously as a multiplex biosensor. Fundamentally, the ϑ_m of the biophotonic assays is proportional to the mass of analyte bound to the surface and by extension the sensitivity of analyte capture is also dependent on analyte mass. The PAG-functionalised surface produces a strong signal during IgG capture but multiplex assays of Complement activation require proteins of different masses to be measured. Low-mass analytes such as C5a (10.4 kDa) and high-mass analytes such as TCC (1030 kDa) will exhibit different responses upon analyte capture. The sensitivity of assays may be improved with an additional (labelling) step in the SOP: detection.

Following analyte capture and preceding surface regeneration, an antibody sample may be injected over the assay surface which binds to epitopes on the captured analyte. A calibration curve is constructed from the AUC of the detection step only and somewhat standardises assay sensitivity because the masses of detection antibodies are similar. Notably, there will be some variation in detection step response due to varying binding ratios between analyte and detection antibody if polyclonal antibodies are used. The inclusion of a detection step increases the duty cycle time of the assay but can improve assay sensitivity if the total mass of detection antibody on the sensor surface is greater than the mass of captured analyte. Signal amplification with the detection step does not limit the biophotonic assays to analytes above a minimum mass threshold, allowing Complement activation biomarkers to be chosen without restriction by mass limits. Furthermore, the detection step occurs after any nonspecific binding effects between the sample and the sensor surface, allowing complex mixtures to be analysed that would otherwise perturb the signal of the capture step.

The first step in the development of a new assay is to identify analyte targets based on the information they provide. This study aims to develop assays of Complement activation which can be achieved by two strategies: 1) the consumption of circulating inactive proteins and 2) the release of activation products. C3 is consumed as a result of activation in all Complement pathways and patients with C3 glomerulopathy (a kidney disorder caused by dysregulated Complement activation) present with increased C3d and low C3 levels, reflecting ongoing Complement activation in these patients (382). Severely burned patients also show low C3 concentrations initially but C3 recovers to normal levels during treatment for survivors whilst non-survivors show diminished C3 recovery (383). C3 cleavage is therefore a good biomarker of Complement activation.

Proteins consumed or released by the CP and LP may also be used as measures of Complement activation. C4 is common to both pathways and C4d deposition on certain brain tumours has been associated with worsened patient outcome (384). C4 and C3 concentrations typically show a positive relationship and a C3/C4 ratio above 4.53 has been identified as a risk biomarker for cardiovascular events in acute coronary heart disease (385). A significant correlation has also been found between MBL concentration and MBL-associated C4 consumption in plasma from healthy individuals (386), potentially explaining why lower peri-operative MBL levels are associated with an increased occurrence of post-operative infections (387, 388). However, there is contrasting evidence suggesting MBL and MASP-2 levels in serum do not affect the clinical outcome in cystic fibrosis patients (389). The conflicting evidence in different patient groups further highlights that the best measure of Complement activation potential is with a functional test rather than inferring system properties from a given protein concentration.

The correlation between the concentration of AP or TP proteins and patient outcome has been investigated for many conditions, although the reasons for the correlation are often unclear. Patients with heart failure display significantly increased levels of FD and TCC, whilst low levels of FH and P are associated with adverse patient outcome (390). It has also been shown that sC5b-9 (fluid phase TCC) concentration predicts future cardiovascular events in type 2 diabetes patients with myocardial infarction (391). Furthermore, the relationship between C5a and sC5b-9 concentrations and patient death has been investigated during acute episodes of thrombotic thrombocytopenic purpura (TTP) – a blood clotting disorder (392) for which sC5b-9 concentration might also distinguish patients with TTP from those with aHUS (393). In addition, Bb and sC5b-9 concentrations in early pregnancy are strongly predictive of adverse outcomes in patients with SLE (394). These examples suggest that proteins consumed or released by the AP and TP are good biomarkers of Complement activation and are likely to show the greatest sensitivity due to the amplification loop of the AP.

Once analytes have been identified, the selection of antibodies with appropriate epitope targets is critical. Assays of Complement activation must be specific to either intact proteins or their activated fragments and two assay designs are proposed in Figure 3.2.

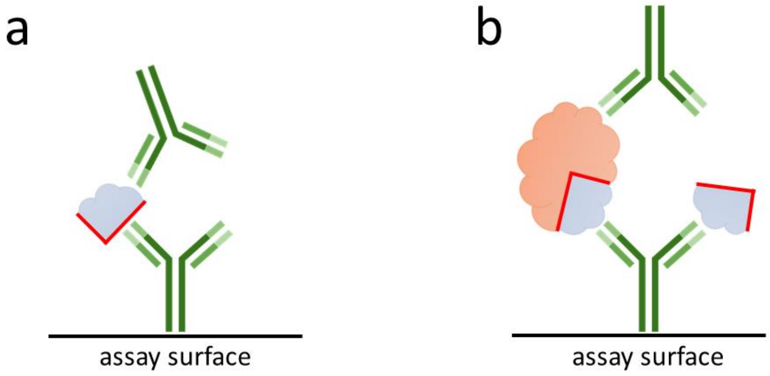


Figure 3.2 Types of sandwich immunoassay for measuring Complement activation. a) A protein fragment (blue) is captured by an antibody (green) via a neo-epitope (red) which is exposed only after cleavage. A detection antibody binds to an epitope elsewhere on the fragment. Complement activation is observed as an increase in assay signal due to the increase in neo-epitope concentration. b) The capture antibody binds to an epitope present on both the blue fragment and the intact protein. A detection antibody binds to an epitope on both the orange fragment and the intact protein. Complement activation is observed as a decrease in assay signal because the detection antibody cannot bind when the epitopes are separated.

The specificity of the assay in in Figure 3.2a is primarily due to the capture antibody which binds a neo-epitope exposed on the analyte of interest after cleavage from its parent protein. Assay response is subsequently enhanced by using a detection antibody which if polyclonal may enable multiple antibodies to detect a single captured analyte – maximising assay sensitivity.

The specificity of the assay in Figure 3.2b relies on both capture and detection antibodies binding to epitopes on the surface of the intact protein which are separated upon activation. A disadvantage of the assay in Figure 3.2b is that the captured epitope is not unique to the target analyte and cleaved protein fragments will compete with the intact protein for binding sites on the sensor surface. Competition for surface binding sites will be greatest when the ratio of intact protein concentration to fragment concentration is lowest. Therefore, this type of assay will overestimate protein consumption to a greater extent during *in vitro* experiments which cause greater consumption of Complement proteins. If the competing fragments are below the 30-50 kDa molecular mass filtration cut-off of the kidneys (395) they will be cleared from the plasma and any assay perturbed by these fragments *in vitro* may be better suited for assessing *in vivo* Complement activation.

Several antibodies have been reported in the literature to produce immunoassays for Complement activation products, some of which are commercially available. An ELISA for Ba has been used to profile the half-life of Ba in human plasma using anti-Ba mAb D22/3 (specific for a neo-epitope on Ba) and anti-factor B as an antibody pair (245) but these antibodies are no longer available to purchase. An ELISA for C5a has been used to profile Complement activation in septic baboons using anti-C5a mAb C17/1 (specific for a neo-epitope on C5a) and a polyclonal anti-C5 detection antibody (244). In addition, an ELISA for TCC has been used to quantify TCC in human plasma using mAb aE11 (specific for a neo-epitope on polymerised C9) (396) and polyclonal anti-C6 detection antibody (244).

Monoclonal antibodies specific to epitopes which are separated upon protein cleavage may be used for capture and detection antibodies in assays of protein consumption. The mAb K13/16 has been shown to bind both C3a and intact C3 (397), implying that the epitope of K13/16 is located on the outer surface of C3 which becomes part of the C3a fragment upon cleavage. Detection antibodies which bind to the C3b surface of intact C3 should produce a C3 activation assay where detection signal will decrease with C3 cleavage.

3.2 Aims and Objectives

This Chapter reports investigations into antibody properties in the development of novel assays of Complement activation.

The first objective was to test of a panel of candidate capture and detection antibodies for Complement proteins using the optimised assay production methods described in Chapter 2.

The second objective was to investigate the effects of changing the inherited protocol for surface functionalisation described in 2.8.1 in an attempt to optimise the sensitivity of the successful sandwich assays.

The third objective was to assess the performance of the novel Complement assays according to the metrics discussed in Chapter 2, such as sensitivity and specificity.

Owing to the variability in antibody properties encountered, the fourth objective was to develop a novel quality control test of antibody samples which could improve the reproducibility of future assay development on the LiScAR biosensor platform.

3.3 Antibody Screening

3.3.1 Capture Antibody Screening

A panel of candidate capture antibodies was immobilised on the amino substrate surface described in section 2.4 using the assay method described in 2.8.1. The antibodies tested are listed in Table 3.1. A polyclonal antibody for TNFα, a cytokine discussed in Chapter 1 as a biomarker of sepsis, was included in the panel. A polyclonal antibody for CRP was inherited from the research group as a working CRP capture antibody and was included in the antibody screen for reference. A monoclonal IgG from NIST was also included in the antibody screen as a well characterised reference material and will be discussed later in the chapter.

Table 3.1 Antibodies immobilised on the LiScAR sensor surface and tested for their ability to capture the listed antigens. Antibodies were immobilised and tested using the assay method described in 2.8.1.

Clone name	Product code	Supplier	Reactivity	Antigen tested	Product code	Supplier
unknown	A252	Pathway Diagnostics	Bb (neo)	Bb	A155	CompTech
K13/16	GAU 013-16-02	Thermo Fisher	C3a/C3	C3	A113c	CompTech
013III-1.16	MA1-82814	Thermo Fisher	C3b/C3	C3	A113c	CompTech
unknown	A209	Pathway Diagnostics	iC3b (neo)	iC3b	A115	CompTech
polyclonal	NBP1-31229	Novus Biologicals	C4b/C4	C4	A105c	CompTech
unknown	A251	Pathway Diagnostics	C4d/C4	C4	A105c	CompTech
C17/5	GAU 025-05-02	Thermo Fisher	C5a (neo)	C5a	A145	CompTech
unknown	A239	Pathway Diagnostics	TCC (neo)	TCC	A127	CompTech
polyclonal	AHP1212	Bio-Rad	TNFα	TNFα	PHP051	Bio-Rad
NIST mAb	8671	NIST	RSVF	RSVF	11049-V08B	Sino Biological

In each antibody screening experiment a minimum of 14 array spots were functionalised with BSA to provide a negative control channel for baseline subtraction. A number of spots were also functionalised with PAG as a positive control according to the standard operating procedure. Target analytes were diluted to 25 nM in PBS (with the exception of large molecular mass antigens C3 and TCC which were diluted to 15 nM) and injected into the LiScAR for 200 seconds, with a regeneration step of 100 seconds between samples.

Six antibodies were found to be successful in capturing their antigen from solution. The array format of the LiScAR assay surfaces allows cross-reactivity to be directly observed and the 6 capture antibodies were printed onto multiplex assay surfaces for specificity testing. Assays were considered adequately specific only if their array spots produced significantly more signal than the controls (3SD). Successful antigen capture experiments for all analytes are shown in Figure 3.3, compared to the polyclonal antibody for CRP which was inherited from the research group as a previously working antibody for CRP capture on the LiScAR sensor surface.

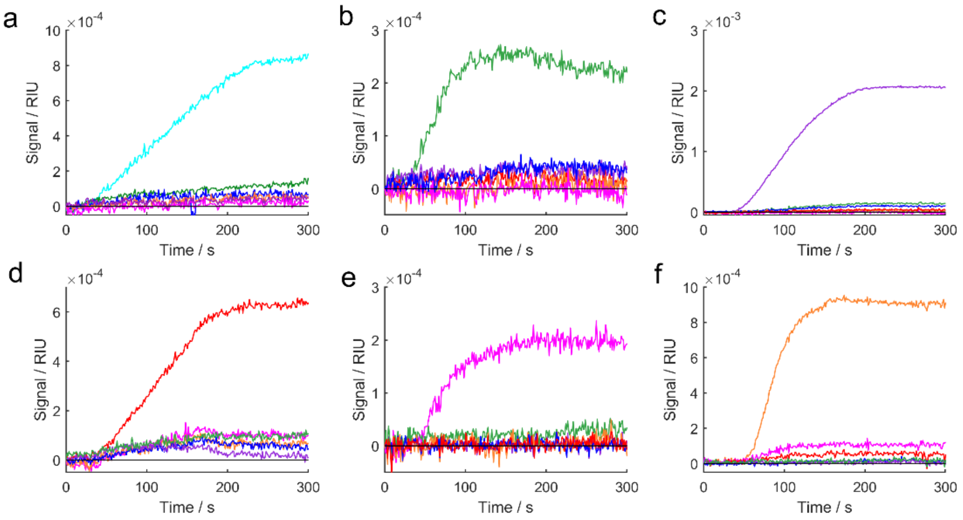


Figure 3.3 Signal response of LiScAR sensor surfaces to antigen samples. A number of proteins were immobilised using the assay method described in 2.8.1: PAG (blue), anti-C5a (green), anti-TCC (purple), NISTmAb (red), anti-TNFα (pink) and anti-CRP (orange). Each panel shows the array response to a 25 nM sample of antigen from Table 3.1: a) 15 nM C3; b) 25 nM C5a; c) 15 nM TCC; d) 25 nM RSVF; e) 25 nM TNFα and f) 25 nM CRP. Signals are the averaged response of n=15 assay spots.

Five novel assays were demonstrated for the detection of C3a, C5a, TCC, RSVF and TNFα in <5 minutes. All assays produced a specific response when exposed to purified solutions of their target antigen.

The capture antibody screening revealed that not all antibody samples tested were suitable for antigen capture after immobilisation on the LiScAR surface, with a success rate of <50%. The RI change associated with antigen capture correlated positively with antigen mass as expected but the success of assay production does not relate to a particular analyte mass range, Figure 3.4. For example, C5a (9 kDa) capture was observed whereas Ba (33 kDa) capture was not despite a three-fold increase in analyte mass.

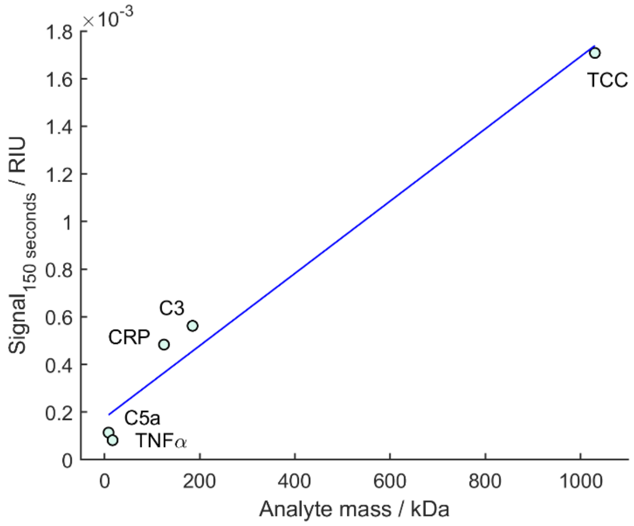


Figure 3.4 Effect of analyte mass on LiScAR signal following 150 seconds of association at 15 nM. Experiments were performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots. Analytes are labelled and fit to a straight line with R^2=0.97.

Deviation from the linear trend between analyte mass and assay response may occur as a result of differences in binding kinetics and the number of binding sites on the sensor surface. For example, a rapid association rate would enable maximum surface coverage to be achieved within a shorter period of time, whilst a lower number of binding sites on the assay surface would reduce ϑ_{m}.

3.3.2 Detection Antibody Screening

The use of a detection antibody can be used to improve the sensitivity of the biophotonic assays, as discussed previously. For a given assay, the number of available binding sites before analyte capture is constant. Within the dynamic range of the assay, the surface coverage achieved during analyte capture varies based on analyte concentration when the duration of association phase is fixed. The number of binding sites available for the detection antibody is determined by the surface coverage of captured analyte and hence precision of the detection step improves as ϑ_{m} is approached. Detection step association rate cannot be used to determine the concentration because analyte surface coverage is uncertain. Therefore, a relatively high concentration of detection antibody must be used, compared to the concentration of analyte standards, to maximise detection antibody surface coverage whilst minimising the duty cycle of the assay.

A panel of candidate detection antibodies were diluted to 200 nM in PBS for testing on the capture assays for C3, C5a, TCC and TNFα. The antibodies tested are listed in Table 3.2. The polyclonal antibody for CRP capture was inherited from the research group as a working CRP detection antibody and was included in the antibody screen for reference.

Table 3.2 Antibodies tested for their ability to detect the listed antigens captured by antibodies immobilised on the LiScAR sensor surface. Antibodies were immobilised and tested using the assay method described in 2.8.1.

Description	Product code	Supplier	Reactivity	Antigen tested
monoclonal	MA1-82814	Thermo Fisher	C3b/C3	C3
monoclonal	MCA2606	Bio-Rad	C3d/C3	C3
anti-sera	AHP1752	Bio-Rad	C3	C3
polyclonal	PA5-35000	Thermo Fisher	C5a	C5a
monoclonal	LF-MA0231	Thermo Fisher	C6	TCC
polyclonal	32034-05111	Assay Pro	C9	TCC
polyclonal	32556-05111	Assay Pro	TCC	TCC
monoclonal	ABIN112002	Antibodies Online	TCC	TCC
polyclonal	AHP1212	Bio-Rad	TNFα	TNFα

A capture step of 200 seconds was performed for each antigen at 15 nM. A detection step of 200 seconds was then performed using each candidate detection antibody at 200 nM (or anti-sera diluted 500-fold). 100 seconds of PBS running buffer between capture and detection is used as a reliable measure of the pre-detection signal for AUC analysis and could be used to estimate k_d of the capture step. Successful detection steps following analyte capture are shown in Figure 3.5 compared with the assay for CRP that was inherited from the research group. The anti-C3d/C3 monoclonal antibody MCA2606 was also able to bind captured C3 but the signal was weak, Appendix 1.

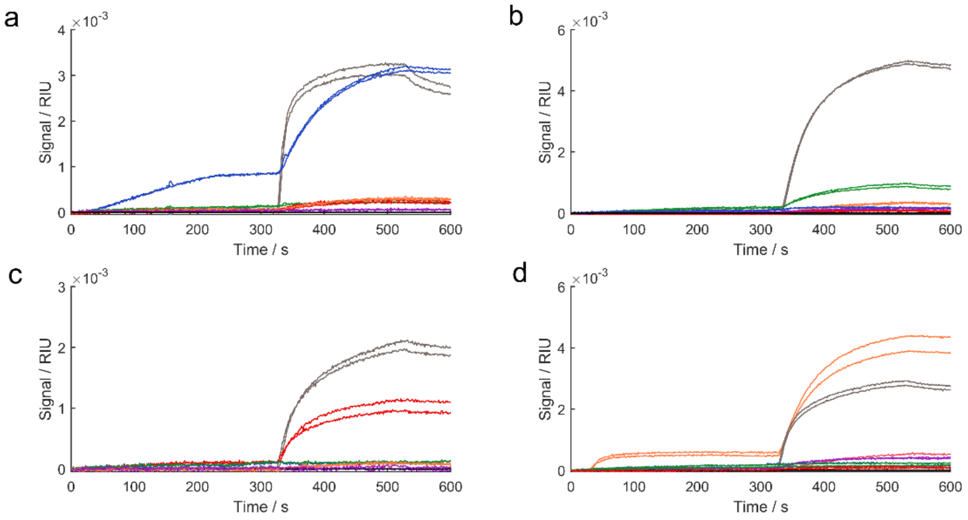


Figure 3.5 Signal response of LiScAR sensor surfaces to antigen capture and detection. A number of proteins were immobilised using the assay method described in 2.8.1: PAG (grey), anti-C5a (green), anti-TCC (purple), anti-TNFα (red) and anti-CRP (orange). Antigen capture (15 nM) and detection (200 nM) for a number of assays (Table 3.2) are shown: a) C3 and AHP1752; b) C5a and PA5-35000; c) TNFα and AHP1212; d) CRP and its detection antibody inherited from the research group. Signals are the averaged response of n=15 assay spots shown for n=2 duplicate experiments performed on the same assay surface.

Following the capture step, all assay channels exhibited a specific detection step response to their respective analytes only. TCC was the only analyte for which detection antibodies were sought but not found. Antigen detection was achieved for C3, C5a and TNFα. All detection steps showed significant curvature, suggesting maximum detection antibody coverage is approached within 200 seconds. Furthermore, the ΔRI of all detection steps was greater than the capture step – suggesting that the detection step increased assay sensitivity for these analytes in the duty cycle described here.

3.4 Optimisation of Biophotonic Assays

To achieve optimal analyte sensitivity, protein immobilisation should achieve a high surface density of antibodies in an orientation that does not hinder accessibility of the Fab region binding sites (398). Furthermore, covalent attachment of antibody to a solid support can lead to a change in the structural configuration of the protein resulting in reduced binding ability (399). A number of parameters of the assay fabrication process were investigated here for their effect on assay sensitivity: 1) concentration of printed antibody; 2) the delay between SAM activation and functionalisation; 3) the buffer in which the antibody is printed and 4) the orientation of immobilised antibody.

3.4.1 Print concentration

A large proportion of the cost associated with the fabrication of the LiScAR arrays is the protein or analyte used for functionalisation. The cost of each print is reduced by diluting the concentration of printed protein but may reduce the surface density and sensitivity of the resulting array spots. The effect of print concentration on assay performance was investigated by printing an anti-CRP polyclonal antibody sample (inherited from the research group) at a range of concentrations and recording the AUC of the CRP capture step, Figure 3.6.

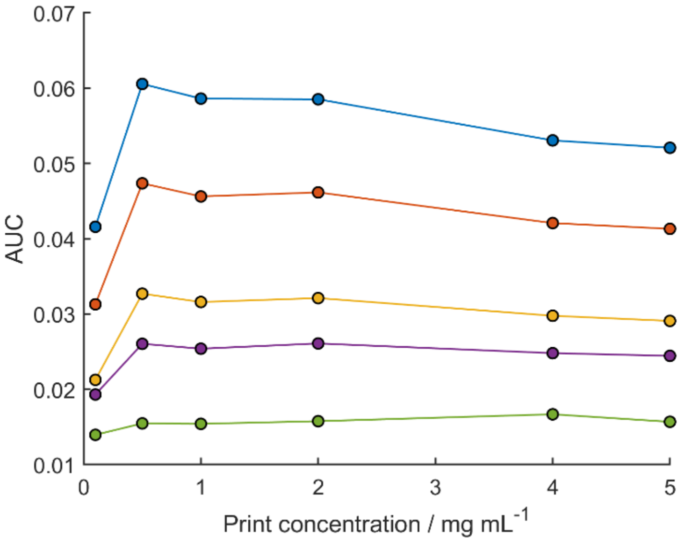


Figure 3.6 The effect of print concentration on response of the LiScAR CRP assay. Experiments were performed using the assay method described in 2.8.1. The AUC from several CRP injections is shown: 3.1nM (green) 6.3 nM (purple), 12.5 nM (yellow), 50 nM (orange) and 100 nM (blue). Data points were derived from the averaged response of n=15 assay spots on a single array surface.

The relationship between antibody print concentration and assay response was more significant at higher analyte concentrations. The standard printing protocol used 1 mg/mL print concentration and these data indicate that optimal sensitivity may be achieved with print concentrations in the range 0.5-2 mg/mL - offering a potential materials savings. At printed concentrations greater than 2 mg/mL, the response of the CRP assay did not improve. Other antibodies should be investigated to determine whether all proteins show a similar trend when printed under the same conditions.

3.4.2 NHS Ester Hydrolysis

A potential source of inter-assay sensitivity variation is the time delay between SAM drying and protein printing, which provides an opportunity for hydrolysis of the NHS ester intermediate. For best results, arrays were printed as soon as possible after removal from the EDC/NHS solution and dried with N_2 gas. The time required to prepare different numbers of arrays before printing may vary significantly if slides are dried with a gas hose individually as was the case in the present work: a single printed substrate could be printed in <5 minutes, whereas 20 substrates required up to 15 minutes to dry before printing. Water in the air is in vast excess to that of the NHS ester intermediates on the activated SAM surfaces before printing. Therefore, the rate of NHS ester hydrolysis is proportional only to the concentration of the ester and the rate equation becomes pseudo first order:

$$rate = k[A] \tag{3.1}$$

where k is the rate constant and [A] is the concentration of the NHS ester. First order reactions have a constant half-life:

$$t_{1/2} = \frac{\ln(2)}{k} \tag{3.2}$$

where $t_{1/2}$ is the half-life of the reaction and k is the rate constant. The relationship between assay response and printing delay may be written:

$$response_t = response_{start} \times 0.5^{\frac{t}{t_{1/2}}} \tag{3.3}$$

where $response_t$ is the response of the assay following a printing delay of t and $response_{start}$ is the response with no delay before printing.

The effect of NHS ester intermediate half-life on assay performance was investigated by delaying the time between SAM activation and printing of anti-CRP polyclonal antibody at 1 mg/mL onto an array surface. The detection step response of the anti-CRP array spots to CRP samples are shown as a fit to equation 3.3 in Figure 3.7. The fitted $response_{start}$ was allowed to vary for each analyte concentration whilst the half-life parameter was fitted globally.

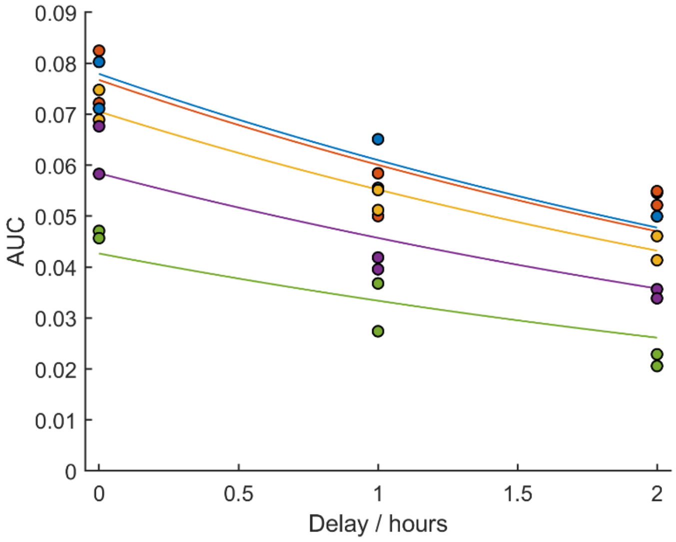


Figure 3.7 The effect of crosslinker half-life on LiScAR CRP assay performance. Experiments were performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots on a single assay surface. Duplicate determinations of analyte concentration are shown: 0.78 (green), 1.56 (purple), 3.13 (yellow), 6.25 (orange) and 12.5 (blue) nM. Equation 3.3 is fit globally to data from assay response data when capture antibodies are printed with a delay following SAM activation. The half-life of the crosslinker in air at 70% humidity and 20 ^{o}C was determined to be 2.8 ±0.7 hours.

The half-life of the NHS ester intermediate was estimated to be 170 ± 45 minutes. As expected, the half-life estimate in air is greater than the literature value of >120 min for the hydrolysis half-life of PEG-NHS in solution at pH 7.4 (400, 401). For the data shown, the first substrate was printed as fast as possible following removal from the EDC/NHS solution with a benchmark time of 2 minutes and the data indicate that the assay response decreases by ~10% over a subsequent 30 minutes delay to printing. In the present work, the time delay between SAM activation and protein printing was typically up to 10 minutes for a batch of <10 substrates and limited mainly by the time required to dry arrays removed from the EDC/NHS solution. The duration of the drying step was therefore not significantly detrimental to the sensitivity of small batches of functionalised arrays but should be considered for mass production, where the drying step should be automated to dry multiple arrays simultaneously.

3.4.3 Print Buffer

Antibody samples are typically stored in PBS with sodium azide and occasionally BSA. The BSA minimises loss of antibody from binding to the plastic container surface and the sodium azide is bactericidal. Ideally, antibodies would be printed for EDC/NHS coupling from purified solutions of PBS only because as the sodium azide and BSA compete with IgG for the activated SAM linkers: leading to lower surface functionalisation and reduced assay sensitivity.

BSA competition for SAM linkers was investigated for its effect on assay performance by printing solutions of anti-CRP polyclonal antibody containing BSA. The capture step response of the CRP assay to 50 nM CRP is shown in Figure 3.8. The effect of BSA competition at different ratios of linker to spacer SAM components was also tested.

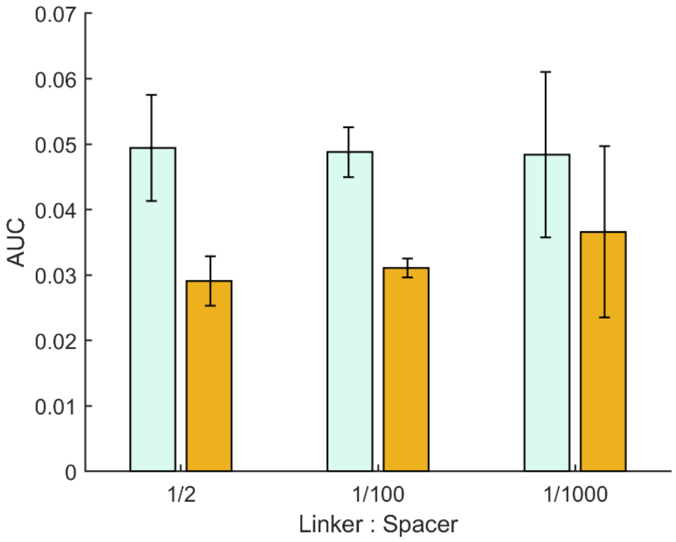


Figure 3.8 The effect of SAM linker availability on LiScAR CRP assay performance. Experiments were performed using the assay method described in 2.8.1. The response of the CRP assay to 50 nM CRP is shown for assay spots functionalised with 0 mg/mL BSA (green) and 1 mg/mL BSA (orange) at 3 linker : spacer ratios. Error bars are 2SD from n=3 experimental repeats recording the averaged response of n=15 assay spots.

The ratio of linker to spacer SAM components had no effect on the response of the CRP assay for the range of ratios shown in Figure 3.8. However, when competition for linker components was introduced, via the inclusion of BSA in the printed antibody solution, assay response was significantly reduced. The result highlights that for optimum results, commercially available antibodies supplied with BSA should be buffer exchanged to remove BSA before printing.

Additional properties of the printed antibody solution may affect assay sensitivity. The p*I* of an antibody is the pH at which the antibody will not migrate in an electric field due to the zwitterion structure having a net charge of zero. The isoelectric point (p*I*) of antibodies has been shown to vary significantly between samples (402), typically in the range 6-12 (402). Antibodies at pH<p*I* are positively charged and will be electrostatically attracted to the negatively charged gold nanoparticle surface (403) although their amine groups may react less readily with the SAM linker. Antibodies at pH>p*I* are negatively charged but their amine groups will be neutral and react more readily with the activated SAM linker. Print solution pH also has an effect on the hydrolysis half-life of PEG-NHS esters, which exceeds 120 min at pH 7.4, but is <9 min at pH 9 (400). The effects of electrostatic attraction and increased NHS ester half-life (pH<p*I*) are likely to give a greater improvement to surface coverage when the concentration of protein is low and limited by molecular collisions. Conversely, when protein concentration is not limiting, the effect of improved nucleophilicity (pH>p*I*) may be more advantageous to maximise surface coverage, despite reduced NHS ester half-life. pH may also affect antibody orientation on the assay surface (404) leading to variation in steric availability of binding sites, discussed later.

The effect of printed antibody pH on assay response was investigated experimentally using the CRP assay. The capture antibody was buffer exchanged using a dialysis spin column, from PBS (containing 0.05% sodium azide) into phosphate buffer (0.1 M) at a number of pH values and printed using the SOP. The capture and detection step responses of the CRP assay are shown in Figure 3.9.

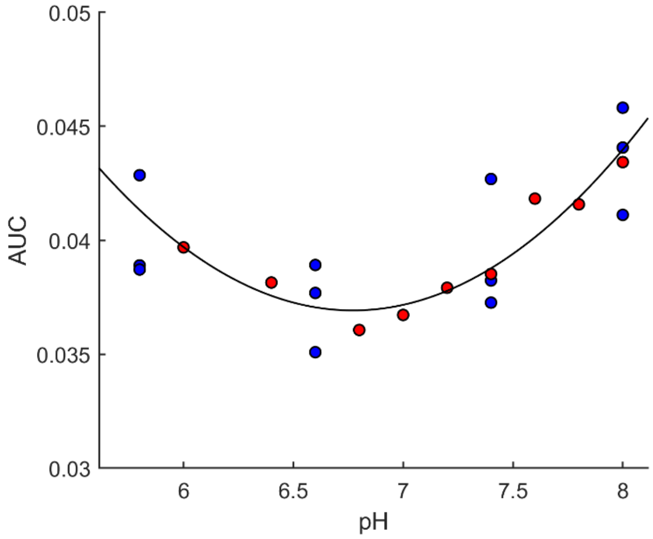


Figure 3.9 The effect of pH and buffer exchange on the LiScAR CRP assay capture step. Experiments were performed using the assay method described in 2.8.1. Buffer exchange prints (blue) are compared to 'pH on top' prints (red). Data points were derived from the averaged response of n=15 assay spots. A line of best fit (quadratic) is shown in black.

Assay response did not appear to vary significantly within the pH range tested – although pH 8 displayed the highest mean assay response. Buffer exchange of antibody samples using spin columns is not always possible due to the limited volumes of antibody samples. A second buffer printed immediately on top of printed assay spots in a 1:1 volume ratio was investigated for its ability to change the pH of antibody tethering without the need for buffer exchange - referred to as the ‘pH on top’ method. Phosphate buffer at a number of pH values was printed immediately following a standard print of anti-CRP from PBS at pH 7.4 and results are shown in Figure 3.9. Assay sensitivity was consistent between the buffer exchanged and ‘pH on top’ groups, suggesting that either method may be used to investigate the effect of solution pH on assay performance. Furthermore, assay spots functionalised with antibody after the buffer exchange (removing sodium azide) did not display a higher mean response to antigen than the ‘pH on top’ group. A limitation to the ‘pH on top’ method is the residual buffering capacity of the PBS antibody solution printed first, which will resist changes to pH associated with the second printed solution. A further limitation of the ‘pH on top’ method is the delay between printing the first and second buffers (~1 minute) during which time antibody tethering occurs as a function of the first buffer only. The error in estimated pH is minimised by printing 0.1 M solutions of the second buffer, which is an excess to the 0.01 M Na_2HPO_4 and 0.0018 M KH_2PO_4 found in PBS (the capture antibody buffer).

The capture antibodies for the CRP and C5a assays were printed using the ‘pH on top’ method using phosphate buffer and borate buffer at pH values 5-7.4 and 8-10 respectively. The capture and detection steps of the CRP and C5a assays were performed for a number of analyte concentrations to observe the effect of buffer and pH on assay sensitivity, Figure 3.10.

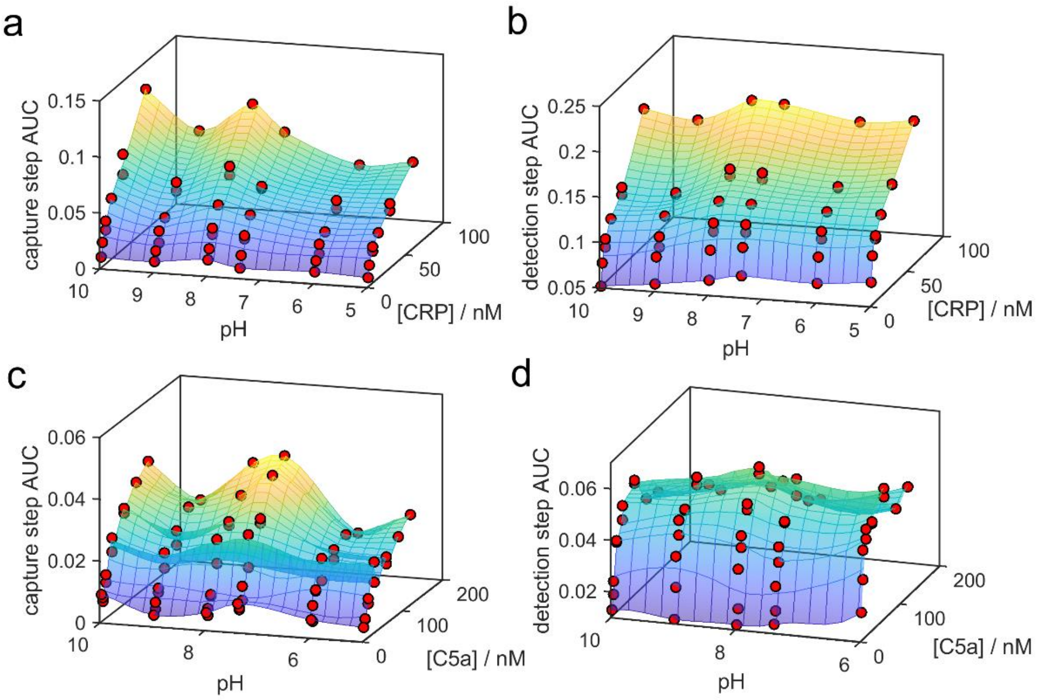


Figure 3.10 The effect of printing buffer and pH on the LiScAR CRP and C5a assay calibration curves. Experiments were performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots from two assays. a) CRP capture step. b) CRP detection step. c) C5a capture step. d) C5a detection step. Borate buffer was used for pH 8-10 and phosphate buffer was used for pH 5-7.4.

The data suggest that assay response was greatest when the antibodies are printed at pH 7.4-8. The results are in good agreement with literature showing reaction of NHS esters with amines is most efficient at pH 7.2–8.5 (405). Interestingly, assay performance appeared to rise from pH 9 to pH 10 in the capture step of both assays but not in the detection step. Experimental conditions such as pH can significantly affect the strength of individual thiol–gold contacts (360) and high pH values favour SAM dissociation from the gold surface (406) which reduces antifouling. The increase in capture signal relative to detection signal may therefore have occured as a result of nonspecific binding at higher pH values.

The optimised buffer conditions: 1 mg/mL antibody without BSA or sodium azide, borate buffer pH8 ‘on top’ and minimum delay following SAM activation did not generate assays when applied to antibodies which failed the initial screen.

3.4.4 Antibody Orientation

It is unlikely that assays failed due to lack of assay surface functionalisation, owing to the high abundance of available lysine residues for tethering. However, the covalent immobilisation of proteins has been shown to alter their primary structure and antibody activity has been shown to reduce with tethering proximity to the active site (407). The range of available tethering locations using EDC/NHS coupling typically leads to high heterogeneity of products formed during the synthesis of (ADCs) (381) and a significant portion of antibodies on the biophotonic surface could be oriented such that the binding of antigen to the Fab region is hindered (398).

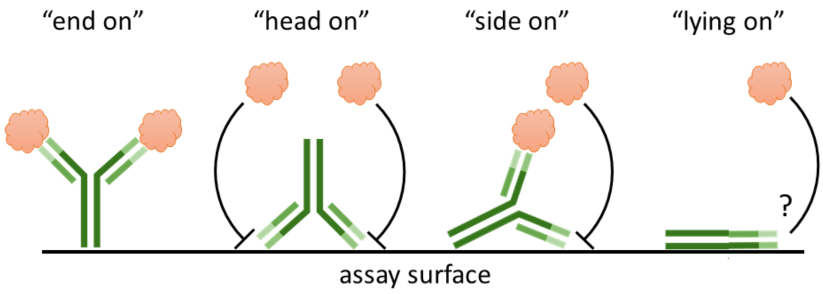


Figure 3.11 Possible orientations of antibodies immobilised on an assay surface via random deposition. The "end on" and "side on" orientations are expected to allow antigen binding, whereas the "head on" and "lying on" orientations may sterically hinder antigen binding. The relative abundance of each configuration will limit the mass of antigen which can be captured by the assay surface and by extension, the LiScAR assay response.

Tethering antibodies far from the Fab region minimises changes to the structural configuration of the Fab region which might reduce binding ability (399). Furthermore, the "end on" configuration is preferred due to minimal steric hinderance of antigen binding (398). Orientation control of antibodies on biosensor surfaces has been shown to improve assay sensitivity up to 200-fold compared to random deposition (408).

Site-specific immobilisation of antibodies on a sensor surface is achievable by a number of methods (398, 409). Firstly, the isoelectric point of the Fab region is often larger than that of the Fc region for most antibodies, leading to antibodies with dipoles which may be oriented on a surface using an electric field (410, 411). The unique properties of the Fc region may also be targeted for site specific immobilisation. Antibodies pre-treated with HCl buffer at pH 2 before immobilisation mainly adsorb "end on" to hydrophobic surfaces, due to an increase the hydrophobicity of the Fc region which is less stable than the Fab region (404). The Fc region is also glycosylated and periodate oxidation of diols in the sugars generates aldehydes which can tether to SAM capped with amines (via imine bonds) (412). The high abundance of histidine residues on the Fc region can be used to orient antibodies via coordination to surfaces with cations presented on chelates (413). Reduction of disulphide groups in the hinge region of IgG yields antibody fragments with exposed thiols which can be tethered to a maleimide surface (414). Biotin has been genetically incorporated into the Fc region of an antibody, leading to a 5-fold improvement in antigen detection limit when the antibody was immobilised on a streptavidin coated sensor surface compared to random deposition (415).

The most common method for controlled orientation of antibodies on a biosensor surface uses the molecular recognition of Fc binding proteins (411, 416-418). The PAG-functionalised assay surface of the IgG assay, discussed in Chapter 2, binds IgG via the Fc region and generates an IgG layer in which Fab regions are most likely to be oriented away from the assay surface (419). Following non-covalent binding of the antibody panel onto the PAG sensor surface, antigen capture was observed for antibodies which failed show antigen capture when crosslinked to SAM, Table 3.3.

Table 3.3 Comparison of antigen capture by two antibody immobilisation methods: covalent crosslinking to SAM and reversible binding to PAG. Antigen capture events are noted as successful if a specific signal is observed which is >3σ of the control channel.

	Antigen capture observed	
	Antibody covalently bound to SAM	Antibody non-covalently bound to PAG
Anti-Bb	no	yes
Anti-iC3b	no	yes
Anti-C4d/C4	no	no
Anti-C3b/C3	no	yes
Anti-C4b/C4	no	yes
Anti-TCC	yes	yes
Anti-C5a	yes	yes
Anti-C3a/C3	yes	yes

The result suggests that the anti-Bb, anti-iC3b, anti-C3b/C3 and anti-C4b exhibited reduced antigen binding ability when crosslinked to the SAM surface due to either unfavourable tethered orientations or structural changes associated with covalent tethering.

The disadvantage of noncovalent antibody immobilisation is that the capture antibody dissociates with elution of the antigen and must be reloaded onto the PAG surface before each measurement – increasing the duty cycle time, materials cost and sources of error. Furthermore, the PAG surface can capture detection antibodies non-specifically and increase the background signal of the detection step. In the production of affinity columns for affinity column chromatography, antibody is allowed to bind to the PAG within the column before covalent crosslinking of the PAG-IgG complexes to generate reusable surfaces. The same technique is possible on the PAG functionalised biophotonic surfaces using an appropriate chemical crosslinker.

Disuccinimidyl suberate (DSS) is a linear compound with NHS ester moieties at each end. As described in Chapter 2, NHS esters are susceptible to nucleophilic attack by primary amines to form stable amide bonds with lysine residues on the surface of proteins. In contrast to EDC, which is a zero length crosslinker, the two NHS esters of DSS are linked by a 8-carbon chain. DSS is therefore a ‘chemical stapler’ and an appropriate choice for crosslinking existing protein-protein interactions such as PAG-IgG. The structure of DSS and its mechanism of amine crosslinking are shown in Figure 3.12.

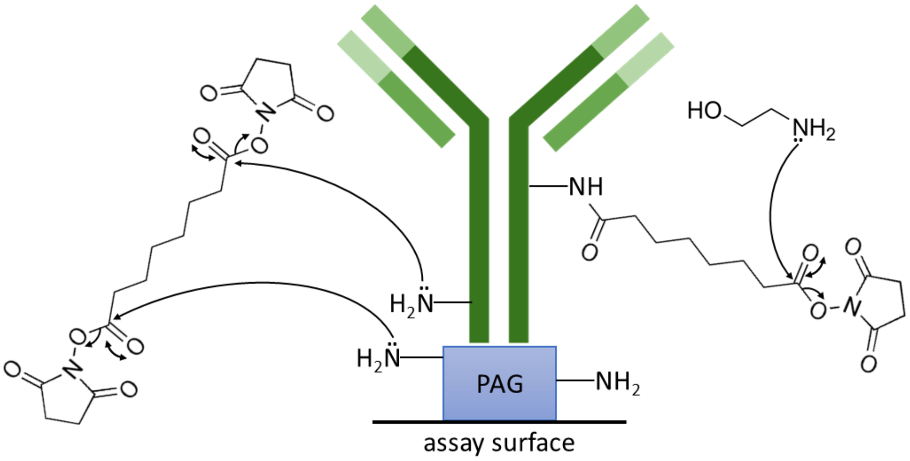


Figure 3.12 Chemical structure of DSS and amine crosslinking mechanism. The carbonyl group of DSS is susceptible to nucleophile attack by amine groups which leads to substitution of the NHS group and the formation of an amide bond. Both NHS esters of DSS can react with amine groups on proteins within the spacer arm length of 11.4 Å to form a chemical crosslink. Unreacted DSS can be quenched by the addition of ethanolamine.

The protocol developed for DSS crosslinking of IgG to PAG was as follows:

1. Biophotonic array of BSA and PAG was exposed to a high concentration of capture antibody until maximal surface coverage is achieved.
2. PBS running buffer was used to remove unbound IgG - ensuring a uniform 2-dimensional layer of IgG. Results were improved by repeating steps 1 and 2 several times to achieve a highly ordered and high density PAG-IgG surface.

3. 6.5 mg of DSS was dissolved in 0.5 mL DMF before mixing with PBS. The precipitate formed was removed with centrifugation for 5 minutes at 4.4 *g*. Failure to remove precipitate may block instrument tubing.
4. The DSS solution (0.014 M) was injected into the biophotonic instrument for a period of time.
5. An aqueous solution of ethanolamine (0.1 M) was used to quench the reaction for 10 minutes if the array was to be used immediately, otherwise the array was left in PBS overnight following step 6 to allow all NHS esters to hydrolyse. Use of an array expressing NHS esters leads to nonspecific and irreversible crosslinking of analytes.
6. The assay surface is washed for 100 seconds in H_3PO_4 (10 mM) to remove unbound IgG.

As with EDC/NHS coupling, a disadvantage of DSS crosslinking is the lack of specificity which may lead to biological activity loss when the crosslinker binds close to the Fab binding site. The duration of crosslinking is therefore a critical parameter in the protocol: assay sensitivity improves with antibody surface coverage but excessive crosslinking may lead to denaturing and a reduction in antigen binding ability. The fraction of PAG-IgG interactions which were successfully crosslinked was estimated by comparing the RI before and after the crosslinking and wash steps, Figure 3.13.

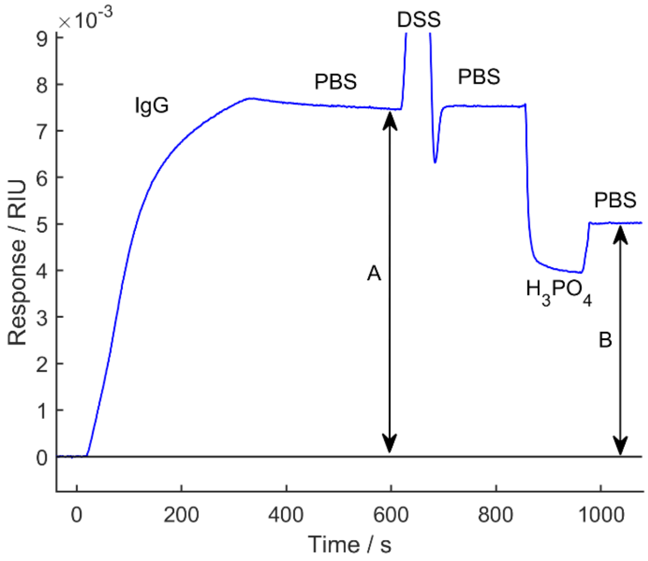


Figure 3.13 PAG channel response to DSS crosslinking of IgG. PAG was immobilised using the assay method described in 2.8.1. A and B represent the RI before and after crosslinking and are used to estimate the extent of IgG crosslinking. Signal is the averaged response of 15 assay spots.

The percentage coverage of IgG retained following crosslinking was estimated by expressing A as a percentage B in Figure 3.13. DSS crosslinking time was optimised by recording the ϑ_{m} of the CRP assay capture step for a 12.5 nM sample of CRP after different crosslinking times, Figure 3.14.

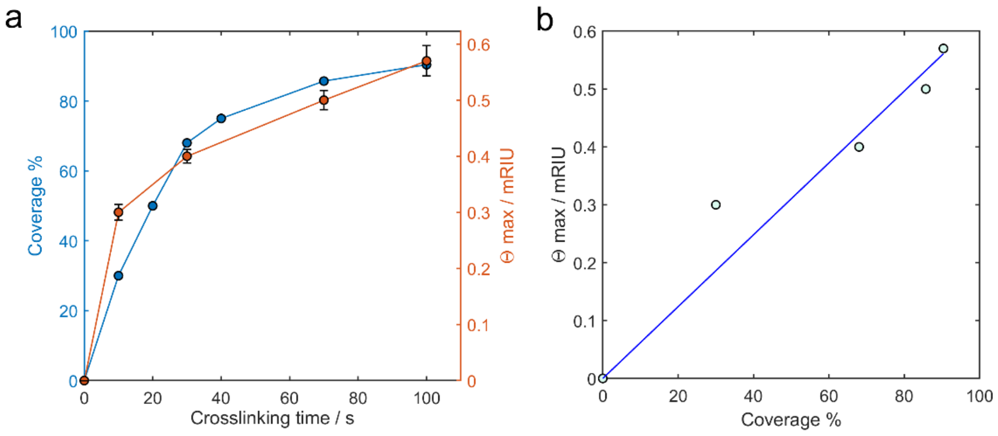


Figure 3.14 The relationship between DSS crosslinking time and ϑ_m of the LiScAR CRP assay capture step. PAG was immobilised using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots. a) time dependence of coverage % (blue) and ϑ_m (orange). Error bars are 95% CI from the Langmuir model estimate of ϑ_m from single experiments. b) the linear relationship between coverage % and ϑ_m with R^2=0.93.

The results show that that 90% surface coverage was retained after 100 seconds and antibody surface coverage is linearly correlated with assay response within range shown. Assay response is expected to show a decreasing trend with extensive crosslinking time but the data collected do not show the maximum.

The anti-Bb, anti-iC3b, anti-C3b/C3 and anti-C4b antibodies showed antigen capture when non-covalently bound to PAG (Table 3.3) but failed to show antigen capture when covalently immobilised on PAG or the SAM surface, indicating their lack of antigen binding ability is a result of covalent immobilisation rather than immobilised orientation. The anti-C4d/C4 failed to capture antigen in any experiment, whether covalently immobilised or not, suggesting an inherent lack of antigen binding ability which was confirmed by immobilising antigens on the SAM surface and observing a lack of antibody binding.

Antibodies which showed antigen capture when tethered directly to the SAM surface also captured antigen when crosslinked to PAG but the oriented surfaces displayed ϑ_{m} values for antigen capture that were 30% lower than when the same antibodies were directly tethered to the SAM surface. The inferior performance of the oriented surfaces may be attributed to denaturing of the capture antibody (due to DMF or crosslinking) and/or a more favourable IgG surface density on the SAM surface compared to the PAG surface. A lower surface density may limit the antigen binding capacity due to lack of immobilised Fab regions whilst high density packing of IgG may hinder the reaction between Fab regions and antigen (419). IgG loading times of 12 hours did not improve assay response. Future studies could attempt oriented crosslinking using water soluble crosslinkers such as sulfo-DSS to minimise denaturing of IgG and PAG.

3.5 Novel Assays of Complement Activation

The optimised assay fabrication processes, using EDC/NHS coupling, were used to produce four novel assays for C5a, TCC, C3 and TNFα. The assays use the antibody pairs shown in Table 3.4. Capture antibodies were immobilised on the amino substrate surface described in section 2.4 using the assay method described in 2.8.1.

Table 3.4 Antibody pairs used in the production of four novel sandwich immunoassays. Assays were performed using the assay method described in 2.8.1.

Target analyte	Capture antibody target	Supplier	Detection antibody target	Supplier
C5a	C5a neo-epitope	ThermoFisher #GAU 025-05-02	C5a/C5	ThermoFisher #PA5-35000
TCC	Polymerised C9 neo-epitope	Quidel #A239	N/A	N/A
C3	C3a/C3	ThermoFisher #GAU 013-16-02	C3	Bio-Rad #AHP1752
TNFα	TNFα	Bio-Rad AHP1212	TNFα	Bio-Rad #AHP1212

The specificity of the assay capture steps was established previously with mixtures of purified antigens in PBS (Figure 3.3) and the detection steps of each assay show a specific response following analyte capture from purified solutions of a single analyte (Figure 3.5). However, cross-reactivity was also tested in order to assess whether the assays could be performed simultaneously as a single multiplex assay.

The cross-reactivity of the detection steps was tested by immobilising all analytes on an assay surface and observing their response to the capture and detection steps of all assays. A capture step of 200 seconds was performed for each antigen at 15 nM. A detection step of 200 seconds was then performed using each detection antibody at 200 nM (or anti-sera diluted 500-fold). As seen in

Figure 3.15, 100 seconds of PBS running buffer between capture and detection provides a reliable measure of the pre-detection signal for AUC analysis and could also be used to estimate k_d of the capture step. Results of cross-reactivity testing are shown in Figure 3.15.

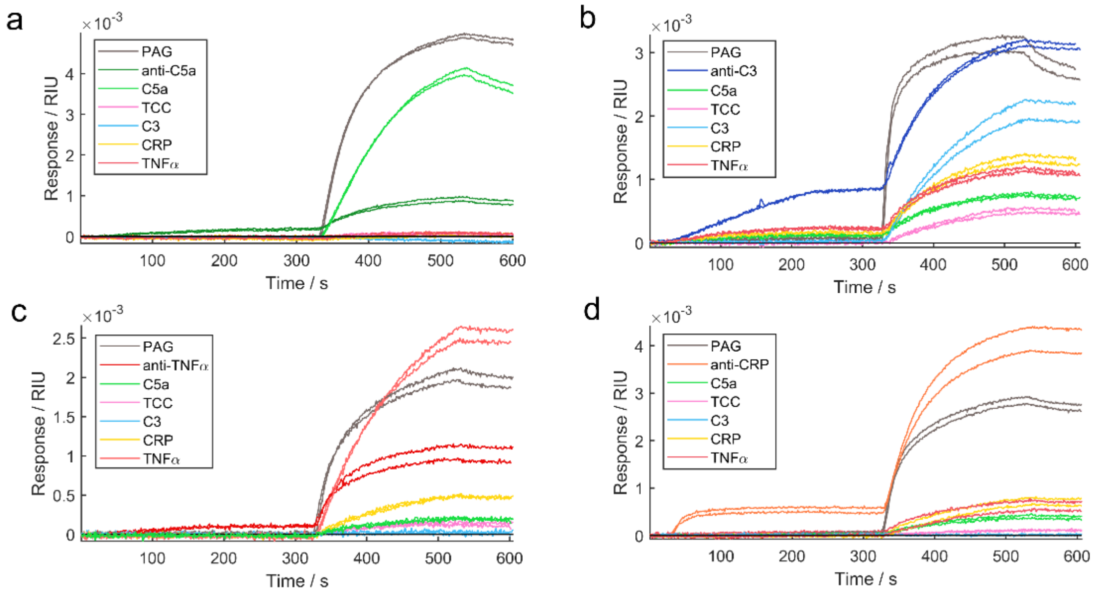


Figure 3.15 Signal response of LiScAR sensor surfaces, functionalised with analyte, to assay capture and detection steps. Experiments were performed using the assay method described in 2.8.1 for n=2 determinations of analyte: a) C5a, b) C3, c) TNFα and d) CRP. Signals are the averaged response of n=15 assay spots.

The ΔRI associated with detection antibody capture by immobilised PAG showed significant variation despite consistent concentrations. The C3 detection step had a maximum signal of 3 mRIU in the PAG channel whereas the C5a detection step exceeded 5 mRIU in the PAG channel. The varying RI change associated with different IgG samples binding to the PAG surface is explored in detail later in the Chapter.

All detection antibodies bound strongly to their respective immobilised analytes and the PAG positive control, indicating that few epitopes were denatured by surface tethering. Following analyte capture, the C5a detection step caused a response in the anti-C5a channel, the C5a channel and the PAG channel only – indicating that the detection antibody does not have affinity for the analytes of other assays (Figure 3.15a). The CRP channel appeared to respond to the TNFα detection step (Figure 3.15c) and the TNFα channel appeared to respond to the CRP detection step (Figure 3.15d) suggesting some cross-reactivity. However, the specific response of the anti-TNFα and anti-CRP channels to their respective analytes was significantly larger than the cross-reaction responses. The results suggest that the assays for C5a, TCC, TNFα and CRP may be performed simultaneously due to the specificity of their detection antibodies. The C3 detection step used antisera containing many different antibodies and appeared to exhibit affinity for all immobilised analytes (Figure 3.15b). Therefore, the C3 detection step must be performed separately and after the detection steps of the other assays. The anti-C3d/C3 monoclonal antibody MCA2606 was also able to bind captured C3 and showed an adequately specific response (Appendix 1) although the signal was significantly weaker than detection with antisera.

A multiplex array containing the assays for C5a, TCC, C3 and TNFα was used to assess the sensitivity of these new assays. The performance of the CRP assay is not presented here. Standard concentrations of each analyte were prepared at four-fold higher concentrations that desired for the standard curves and mixed equally, in random sequence before testing in duplicate. The detection steps of the C5a and TNFα assays were performed simultaneously first, followed by the detection step of the C3 assay. A 5-parameter logistic model was fit to the AUC data at each analyte concentration to produce standard curves. A CV analysis of the novel Complement assays is shown in Figure 3.16. Samples of inactivated, pooled human serum were diluted 30-fold and assayed for C5a and TCC as an initial test of assay susceptibility to nonspecific binding. The dissociation phase following analyte capture in serum was extended to 200 seconds to allow dissociation of non-specifically bound serum components until the signal baseline was stable.

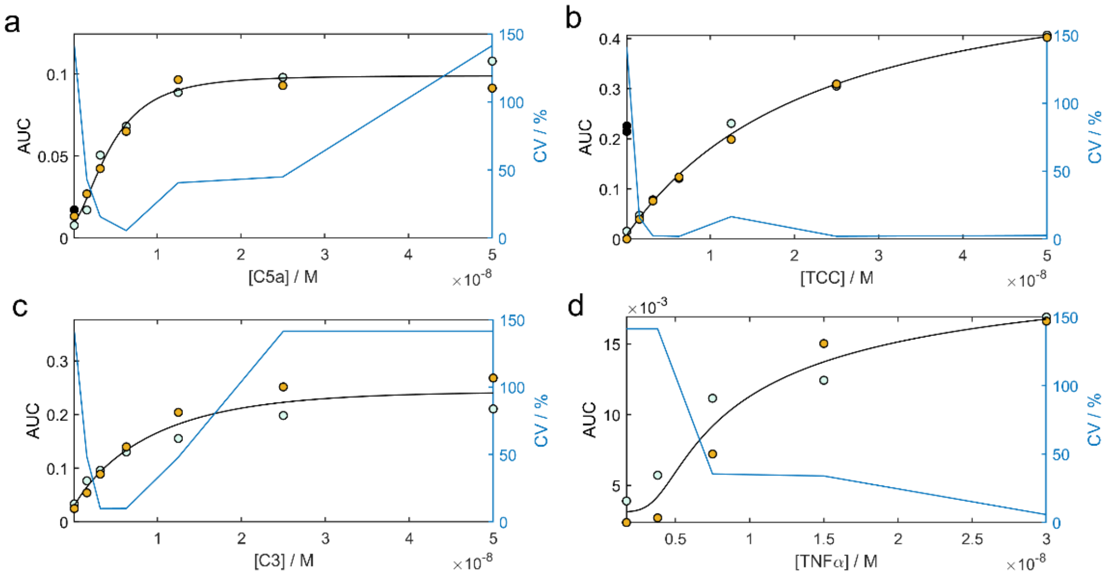


Figure 3.16 Calibration of AUC dependence on analyte concentration for a multiplex panel of LiScAR assays. Experiments were performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots. The standard curves of 4 assays are shown: a) C5a, b) TCC, c) C3 and d) TNFα. A 5-parameter logistic model (black) is fit to the data from n=2 determinations of purified standards in PBS (green and orange). A sample of pooled human serum (black) serves as a control in the C5a and TCC assays. Intra-assay CV at each analyte concentration is shown for each assay (blue).

The 5-parameter logistic model fits well to the data with $R^2>0.95$ for all assays, confirming that the analytes tested did not cross-react and demonstrating a novel 4-plex assay with a duty cycle of 15 minutes.

The C5a assay produced a reading of zero for pooled serum as expected and suggests that the TP of the pooled serum was not activated. The C5a assay was used to profile the increase in C5a concentration following activation of the pooled serum and results are shown in Chapter 7. The non-zero TCC reading for inactivated serum implies that the assay experienced significant matrix effects from the serum, or the epitope target of the capture antibody was not specific to TCC. The TCC assay antibody may have some affinity for C9 that is not incorporated into TCC, rendering the assay inappropriate for use with serum samples.

CV values were calculated across the range of the standard curves to show which analyte concentrations are derived most precisely by the assays. AUC values which exceed the limits of the standard curve model generated values of zero. Enzo ELISA kits state that intra-assay CV should be less than 10% (374) which is achieved for all three assays at an analyte concentration of 6.3 nM. The TCC assay exhibits the lowest CV values tested with a value of 2% at 6.3 nM and the performance is attributed to large AUC values associated with the mass of TCC which exceeds 1 megadalton. The dynamic ranges of each assay and associated CV values are shown Table 3.5.

Table 3.5 Performance assessment of the novel LiScAR Complement assays.

	C5a	TCC	C3	TNFα
LLoQ (CV)	3.2 nM (16%)	1.6 nM (16%)	3.2 nM (10%)	7.5 nM (35%)
ULoQ (CV)	12.5 nM (41%)	>50 nM (3%)	12.5 nM (48%)	>30 nM (6%)
Min CV (concentration)	5% (6.3 nM)	2% (6.3 nM)	10% (6.3 nM)	6% (30 nM)
Dynamic range CV %	21	7	23	25

The range of TNFα in healthy serum is 11 ± 7 pg/mL (420) and has a mass of 17 kDa (higher for dimers and trimers), corresponding to a concentration range of 0.66 ± 0.43 pM (lower for dimers and trimers) which is below the 7.5 nM LLoQ of the TNFα assay. As a result, the TNFα assay was not sufficiently sensitive for use with human serum and was not investigated further.

The dynamic ranges of the C5a and C3 assays were similar and smaller than the range of possible analyte values in serum. Consequently, samples should be assayed for C5a and C3 at a number of dilutions to ensure the analyte concentration falls within the dynamic range of the assay. The upper limit of quantitation of the TCC assay was not apparent in the range of standard concentrations tested and was not investigated in further experiments due to the prohibitively expensive quantities of purified protein required for calibration.

The specificities of the C5a, TCC, TNFα and C3 assays and their calibration characteristic were assessed here on a multiplex assay surface – demonstrating the ability of the LiScAR to record multiple biomarkers in a single test. However, before any assays can be used with confidence, the assays must be quantitatively validated with an existing analytical technique.

3.5.1 Assay validation

Validation of novel assays involves a comparison of the new technique to a gold standard (421) and the relationship between two variables (such as the measurements from two analytical techniques) is typically assessed with correlation and linear regression analysis (422). Correlation analysis quantifies the degree to which two variables are related by computing the Pearson correlation coefficient, r, based on the covariance of the results of both assays. Sample covariance is calculated using the equation:

$$cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1} \qquad 3.4$$

where n is the number of paired measurements and $\bar{x}$ and $\bar{y}$ represent the means of each measurement respectively. The sign of the sample covariance depends on the sign of the correlation (positive or negative). The magnitude of the covariance depends on the magnitude of the data. Correlation between data of different magnitudes can be easily compared using the Pearson correlation coefficient, which is calculated from the covariance using the equation:

$$r = \frac{cov(x,y)}{\sigma_x \sigma_y} \qquad 3.5$$

where σ_x and σ_y are the standard deviations of the two measures (described previously in Chapter 2). The Pearson correlation coefficient, r, has a value between -1 and +1 with a sign dependent on sign of the correlation (positive or negative) and a magnitude corresponding directly to the strength of the correlation (zero if no correlation). As such, a strong positive correlation will always have an r value close to 1, regardless of the magnitude of the data.

A further metric of correlation is linear regression, in which a straight-line model is fitted to a plot of one variable against the other. In contrast to covariance or Pearson correlation, linear regression enables the value of one variable to be predicted based on the value of the other. The regression line is obtained using the method of least squares (422), which finds the parameter values of a linear model that produce the smallest sum of the squared deviations between model and data, discussed further in Chapter 4. The goodness of fit (how closely the data match the fitted trend line) can be quantified with the coefficient of determination, R^2, representing the ratio of the sum of squared residuals between the data and the regression line to the data and the mean. Correlation and linear regression may be used compare the results obtained using a novel assay to the results obtained using a gold standard technique.

To assess the validity of the novel Complement assays, a number of analyte samples at a range of concentrations were measured on the novel biophotonic assays and the equivalent gold standard assays. The gold standard for C5a and TCC measurement is ELISA, whilst the standard for C3 measurement in the clinical laboratory is an immunoturbidimetric assay. The C5a and TCC assays were assessed using PBS samples of purified proteins. The C5a and C3 assays were assessed using the results of serum samples from the clinical investigation presented in Chapter 6. Activated serum samples were diluted 20-fold for the C5a assay, whilst inactivated pooled serum samples were diluted to between 40% and 90% to produce a range of C3 concentrations before testing by diluting samples 640-fold. The dissociation phase following analyte capture in serum was extended to 200 seconds to allow dissociation of non-specifically bound serum components until the signal baseline was stable. A linear regression analysis of all three assays compared to their gold standards is shown in Figure 3.17.

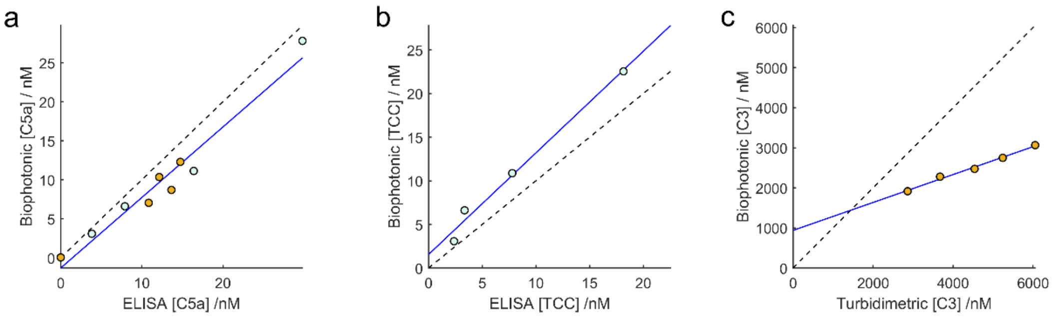


Figure 3.17 Correlation analysis of gold standard assays compared to the LiScAR Complement assays. Experiments were performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots. Three assays were assessed: a) C5a, b) TCC, and c) C3. PBS data (green) and serum data (gold) fit to a linear model with y intercept of zero (blue line) with R^2 values of: 0.96, 0.99 and 0.98 respectively. The ideal correlation, with a gradient of 1 and y intercept of zero, is shown for comparison (black dashed line).

The C5a and TCC assays showed a strong linear correlation with ELISA whilst the biophotonic C3 assay displayed a significant offset from the results obtained using the turbidimetric assay - indicated by the large non-zero y intercept and low gradient of the linear regression model. The quantitative results of correlation and regression analysis using the equations and methods mentioned previously are shown in Table 3.6.

Table 3.6 Results of correlation and regression analysis for the comparison of novel LiScAR assays to their gold standards.

	C5a	TCC	C3
Sample covariance	64.9	60.6	1.53×10^4
Correlation, r	0.98	0.99	0.99
Gradient of regression line	0.90	1.2	0.35
Intercept of regression line / nM	-1.3	1.6	940

The quantitative measurement of continuous variables always has some degree of associated error due to analytical imprecision and assessment of the agreement between two measures is therefore challenging. For example, the offset in the C3 assay is not reflected in the correlation and linear regression goodness-of-fit statistics as these analyses study the trend between two variables but do not value the differences in the pairs of results. The y-intercept of the regression line indicates poor absolute agreement but cannot distinguish between systematic error and nonlinear assay differences which is best assessed using a Bland-Altman analysis.

A Bland-Altman analysis uses the mean and standard deviation of the differences between measures to define limits of agreement (423). The test result is subtracted from the gold standard result and the difference is plotted as a percentage of the mean of the two readings such that large positive values of percentage difference indicate that the new assay under-estimates the concentration compared to the gold standard. A Bland-Altman analysis of the novel assays and their respective gold standards is shown in Figure 3.18.

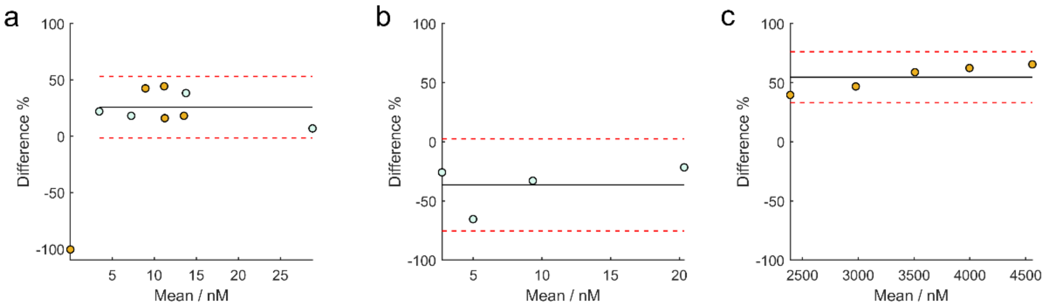


Figure 3.18 Bland-Altman analysis of the LiScAR Complement assays. LiScAR experiments were performed using the assay method described in 2.8.1 and data points were derived from the averaged response of n=15 assay spots. Three assays were assessed: a) C5a, b) TCC, and c) C3. PBS data (green) and serum data (gold) define the mean (black line) and 2SD (red dashed lines) of assay differences.

Bias between the novel assays and their gold standards is observed as a non-zero mean difference in the Bland-Altman plots and is apparent in all three assays to varying degrees. The bias is considered systematic if the distribution of the percentage difference is relatively constant across the range of the assay. The percentage difference is confounded by measurements approaching the LLoQ, where small absolute differences are perceived as large relative differences. For example, the blank serum reading of the C5a assay was excluded from the analysis as the value was lower than the LoQ of the assay.

The C5a and TCC assays showed errors with no dependence on the mean reading and the associated bias appears to be systematic: on average the C5a assay under-estimated by 26 ± 27 % and the TCC assay over-estimated by 36 ± 39 %. The bias could be accounted for in future assays to improve assay accuracy relative to the gold standard (for the analyte concentration range shown). The bias associated with the C3 assay was 55 ± 21 % and visibly worsened at greater analyte concentrations. The worsening error observed in the C3 assay is attributed to serum matrix effects induced by varying serum dilution. The turbidimetric assay for C3 suggests the pooled human serum had a C3 concentration of 1420 mg/L (7.68 µM) and the offset in the biophotonic C3 assay results may have been due to a number of factors. Firstly, error in the standard C3 concentrations may have occured as a result of concentration error in the stock which was supplied as a solution rather than reconstituted from a solid. Errors may also have occured due to the instability of the C3 protein during sample preparation. Calibration of the assays is critical and highlights the need for high quality standards of precisely known concentrations. Future work should use spiked C3 serum samples, providing a broader range of C3 concentrations and enabling a constant dilution factor to be used for Bland-Altman validation analysis. The C5a and C3 assays were used for rapid sampling of *in vitro* Complement activation, shown in Chapter 7.

The search strategy used here to find antibodies capable of capturing antigen on the LiScAR assay surface, involved printing of candidate capture antibodies and simultaneous testing with analyte samples. This approach is efficient for testing a panel of antibodies but is experimentally time consuming for individual antibodies as a new assay must be fabricated each time a new antibody becomes available for testing. Furthermore, the success rate of the search for capture antibodies was <50%. Future LiScAR assay development would benefit from rapid assessment of antibody sample characteristics ahead of printing or use for detection because kinetic parameters and specificity information is rarely available for commercial antibodies.

3.6 A Rapid and Quantitative Technique for Assessing IgG Monomeric Purity

In nontherapeutic antibodies used for research, a link has been made between poor reproducibility of results and variability in antibody quality (424-426). Antibody samples are known to undergo degradation over time during storage (427) and the proportion of intact antibodies in a sample (the monomeric purity) is tested to ensure potency of therapeutic antibodies (428). Antibodies for research should also be assessed for monomeric purity to ensure reproducible experiments (429) and a rapid technique for assessing the monomeric purity of antibody samples could test the extent of degradation in samples immediately before use.

Lot-to-lot comparisons of antibody samples should be performed with reference standards (430). NIST provide a monoclonal antibody, the NISTmAb, designed to support biopharmaceutical innovation by serving as reference standard for comparable evaluation of antibody-based techniques between different laboratories (431-433). The NISTmAb is a recombinant humanised IgG with a known sequence (434) specific to the respiratory syncytial virus protein F (RSVF) (435). The NISTmAb is well characterised with a known monomeric purity variation of 96.8-98.7% between batches, as characterised by comparing the relative peak areas of a non-denaturing size exclusion chromatogram (431, 432).

Non-enzymatic fragmentation in monoclonal antibody samples is frequently observed at the hinge region disulphide bonds. In this region, a single polypeptide connects the Fab and Fc fragments and hence cleavage is followed by separation of these fragments (436). The most abundant antibody fragments and their respective masses observed by non-denaturing size exclusion chromatography SEC (427) are shown in Figure 3.19.

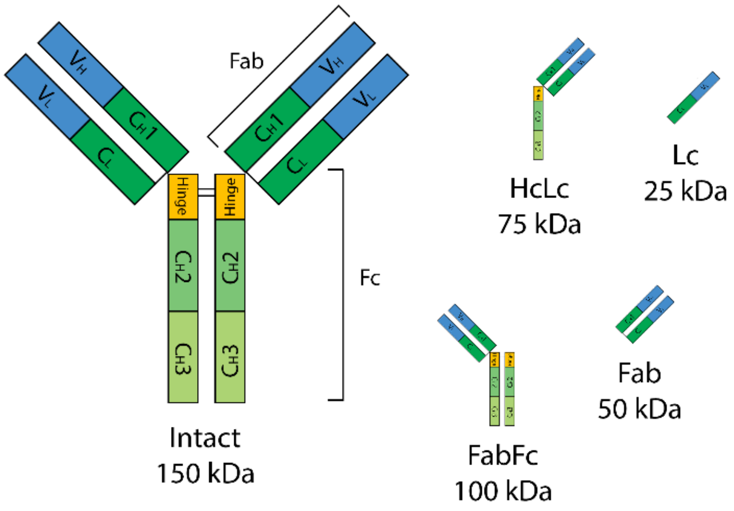


Figure 3.19 Fragmentation products typically observed in samples of IgG. Variable and constant domains are marked V and C respectively with subscripts for light and heavy chain marked L or H respectively. Heavy chain regions are numbered 1, 2, and 3 starting from the N-terminus and black lines indicate disulfide bonds. The most abundant antibody fragments recorded in IgG samples and their respective molecular masses are also shown. Heavy chain and Light chain fragments are labelled Hc and Lc respectively. The 50 kDa fragment may originate from the Fab region as shown or the Fc region. Similarly, the 25 kDa fragment may originate from the Lc as shown, the Hc of the Fab region, or Hc of the Fc region.

SEC, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry (MS) are the gold standard techniques used for characterisation and stability assessment of antibodies (427, 437, 438). Non-reducing SDS-PAGE analysis of a papain digestion of the NISTmAb shows bands for Fab and Fc fragments at 50 kDa, whilst under reducing conditions, the disulphide bonds between heavy and light chains are separated to produce two species of 25 kDa (439). The primary amino acid sequence and PTM structure of antibodies are readily confirmed by liquid chromatography and mass spectrometry (440) however there are no easy methods to establish the monomeric purity of a sample immediately before use. A plasmonic biosensor platform such as the LiScAR may enable rapid monitoring of monomeric purity in a research setting.

Plasmonic biosensors are typically used to observe antibody binding kinetics and are fundamentally mass sensors, as described in Chapter 2. Accurate kinetic parameters may be derived using antibody samples of high monomeric purity and fitting the Langmuir binding interaction model to kinetic data in a global fit (368, 369). However, kinetic analyses are fundamentally limited by unknown antibody concentration and monomeric purity because the presence of multiple binding interactions in fragmented antibody samples invalidates the 1:1 assumption of the Langmuir model (369).

The remainder of the present chapter describes a rapid technique to assess Fab specificity and monomeric purity using the LiScAR. Antigen or PAG were immobilised on the amino substrate surface described in section 2.4 using the assay method described in 2.8.1 to generate assays for Fab and Fc fragments. The technique described here is comprised of two steps, shown in Figure 3.20.

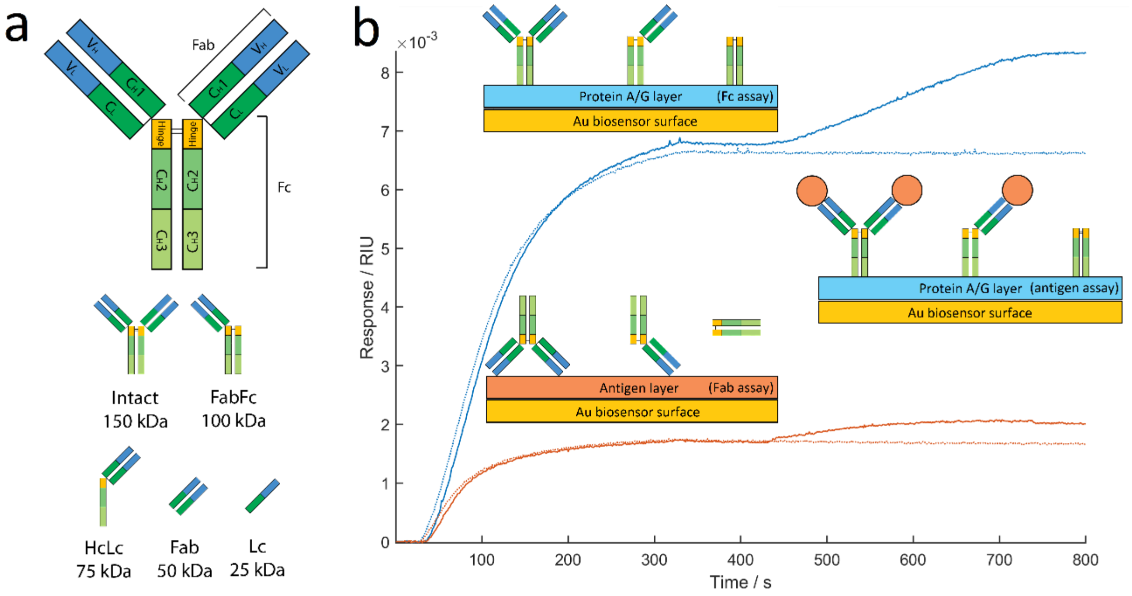


Figure 3.20 Biophotonic assays used to characterise the monomeric purity of an IgG sample. Proteins were immobilised on the amino substrate surface described in section 2.4 using the assay method described in 2.8.1. The Fab assay uses a sensor surface functionalised with antigen (orange) to confirm antibody specificity and derive antibody-antigen binding kinetics. The Fc assay uses a sensor surface functionalised with PAG (blue) to capture Fc fragments. Antibody captured by the Fc assay is used as an antigen assay to determine the antibody-antigen binding stoichiometry. Intact IgG may bind a maximum of 2 antigens (orange circles), whereas samples containing fragments lacking Fab regions will show a lack of antigen capture. The sensor response of the antigen surface (orange) and PAG surface (blue) to the NISTmAb is shown, with and without the inclusion of the antigen assay step (solid and dotted lines respectively) to highlight the signal of antigen capture. Signals are the averaged response of n=20 assay spots.

In the first step, antibody binding to the Fab assay provides data for kinetic analysis of the antibody-antigen interaction and the mass of antibody binding to the Fc assay is recorded simultaneously. For a constant number of binding sites on the Fc assay surface, an initial estimate of monomeric purity may be obtained by observing a mass deficit in degraded samples compared to the NISTmAb. In the second step, the mass of antigen which can bind to the antibodies immobilised on the PAG surface (an antigen assay) is recorded. The stoichiometry of the antibody-antigen binding reaction on the PAG surface is derived from the masses of antibody and antigen captured by the Fc and antigen assays. Samples with high monomeric purity should exhibit an average of 1-2 antigens bound per antibody, depending on antibody mass, whereas antibody fragments lacking Fab regions should display a lower mean stoichiometry. A panel of eight antibody samples was tested: the NISTmAb, a therapeutic Infliximab biosimilar and six commercially available monoclonal and polyclonal samples for research use. Results are validated with native ESI-QTOF mass spectrometry – a gold standard technique for assessing the relative abundance of antibody fragmentation products in a sample (441).

3.6.1 Materials and Methods

Pierce recombinant protein A/G (#21186) and anti-C5a C17/5 (#GAU025-05-02) were obtained from Thermo Scientific. Bovine serum albumin (BSA) (#A7030) was obtained from Sigma-Aldrich. NIST RM 8671 was purchased from the National Institute of Standards and Technology and respiratory syncytial virus protein F (RSVF, #11049-V08B) was obtained from Sino Biological Inc. C5a protein (#A145) was purchased from Complement Technologies Inc. Anti-CRP monoclonal, anti CRP polyclonal, anti-thyroxine-binding globulin (anti-TBG), TNFα, CRP and TBG were all obtained from a major research reagent provider whilst the Infliximab biosimilar was obtained from clinical storage, reflecting the conditions of therapeutic antibodies before use (both sources of antibody samples are made anonymous to enable publication of monomeric purity measures). Standard instrument running buffer and sample dilution buffer was phosphate-buffered saline (PBS) supplied in tablet form by Sigma-Aldrich. Ortho-phosphoric acid (85%) was obtained from Fluka and a 0.01 M aqueous solution used as regeneration buffer. The absorbance of the NISTmAb and other antibody samples was measured at 280 nm with a NanoDrop instrument and converted to total protein concentration using Beer’s Law and a molar extinction coefficient of 210,000 M^{-1} cm^{-1}, (as recommended by the manufacturer).

The binding kinetics and monomeric purity of 8 antibody samples were assessed using a multiplex array and the LiScAR platform. BSA, PAG and antigen were printed with from 1 mg/mL solutions and the array surface was blocked with BSA (1 mg/mL, 300 seconds) before use to minimise non-specific binding. The response of each assay to refractive index (RI) change was calibrated as described in Chapter 2. The mean limit of detection per individual sensor element was $6.6 \pm 2.9 \times 10^{-2}$ mRIU.

The NISTmAb was injected into the LiScAR at concentrations of 1.56, 3.23 and 6.25 nM to provide NIST-RSVF and NIST-PAG binding data for kinetic global fitting as shown in Chapter 2. For monomeric purity testing, all antibody samples were diluted to 100 nM (based on manufacturer stated concentration) and injected over the assay surface for a fixed period of 300 seconds to monitor association kinetics to the Fab assay and Fc assay. The assays were washed in PBS running buffer for 500 s to monitor dissociation kinetics. Each antigen sample was prepared at 100 nM in PBS and injected over the assay surfaces for 300 s to observe antigen capture by IgG previously loaded on the PAG surface.

The sensor response to antibody or antigen binding at maximum surface coverage, ϑ_{m}, was derived for all experiments with kinetic fitting of the Langmuir adsorption isotherm. For clarity, the ϑ_{m}value of the Fab assay is written $\vartheta_{\mathrm{m\ Fab}}$ and the ϑ_{m}value of the Fc assay is written $\vartheta_{\mathrm{m\ Fc}}$ and the ϑ_{m}value of the antigen assay is written $\vartheta_{\mathrm{m\ antigen}}$. To minimise experimental variation, all antibody samples were tested in duplicate and in random sequence on a single multiplex array.

ESI-QTOF experiments were performed by the University of Bath to estimate monomeric purity using a gold standard technique, following 13.a.ii of the protocol list in ref (442). The MS experiments analysed native, intact forms of the antibodies. The degree of antibody degradation was expressed as a relative area ratio of fragments compared to intact protein in the deconvoluted mass spectra.

3.6.2 Results

A series of kinetic traces were recorded for each antibody sample with the Fc, Fab and antigen assays. The Langmuir model was fit to the data as described in Chapter 2. The ϑ_{m} of the NISTmAb sample binding via the Fab region to its antigen (RSVF) immobilised on the sensor surface was determined as 1.73 (±0.03) mRIU by fitting the Langmuir model to data from 100 nM antibody samples. The kinetic parameters for the NISTmAb–RSVF interaction were derived as $k_a = (8.9 \pm 0.2) \times 10^4\ \mathrm{M}^{-1}\mathrm{s}^{-1}$ and $k_d = (2.4 \pm 0.6) \times 10^{-4}\ \mathrm{s}^{-1}$ by fitting the Langmuir model to 3 low concentration samples, Figure 3.21. The equilibrium constant was estimated to be $K_D = 2.7\ (\pm 0.7)$ nM. The errors were derived from the covariance matrix of the fit and are stated at 95% confidence limits.

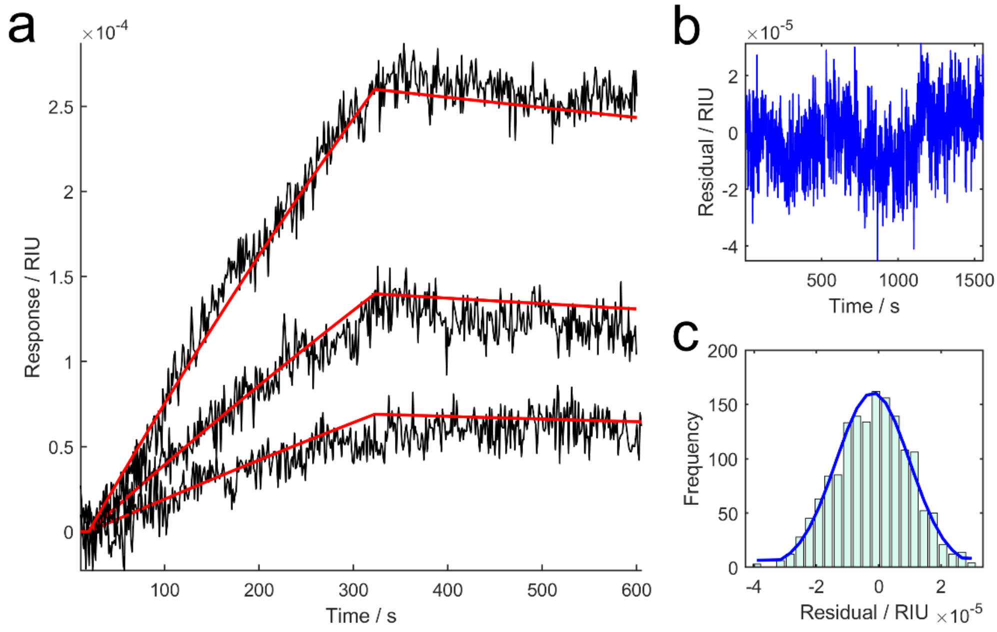


Figure 3.21 RI time dependence for the NISTmAb binding to its antigen, RVSF, immobilised on the sensor surface. RSVF was immobilised using the assay method described in 2.8.1. a) kinetic response curves obtained from 1.56, 3.23 and 6.25 nM NISTmAb samples (black), and global fit to the Langmuir adsorption model (red). Signals are the averaged response of n=20 assay spots and higher concentrations induce a greater response. b) residuals over the time course of the fit, showing consistent error between model fit and data. c) histogram of the residuals indicating the error between the model fit and data is normally distributed.

From four repeated experiments at 100 nM, the ϑ_{m} of the NISTmAb sample binding via the Fc region to its PAG immobilised on the sensor surface was determined as 6.66 (±0.96) mRIU. Complex binding kinetics and deviations from 1:1 Langmuir kinetics are evident in the data however a global fit to four low concentrations, whilst enabling the ϑ_{m} value to float, enabled the model to fit well and estimates the kinetics of the interaction as $k_{\mathrm{a}} = (5.9 \pm 1.1) \times 10^{4}$ $\mathrm{M}^{-1}\mathrm{s}^{-1}$ and $k_{\mathrm{d}} = (1.96 \pm 1.85) \times 10^{-5}$ s^{-1}. The error associated with the dissociation rate constant is large due to the apparently slow dissociation rate – a common source of error highlighted in Chapter 2.

The responses of the Fab, Fc and antigen assays were recorded for all antibody samples at 100 nM, Figure 3.22.

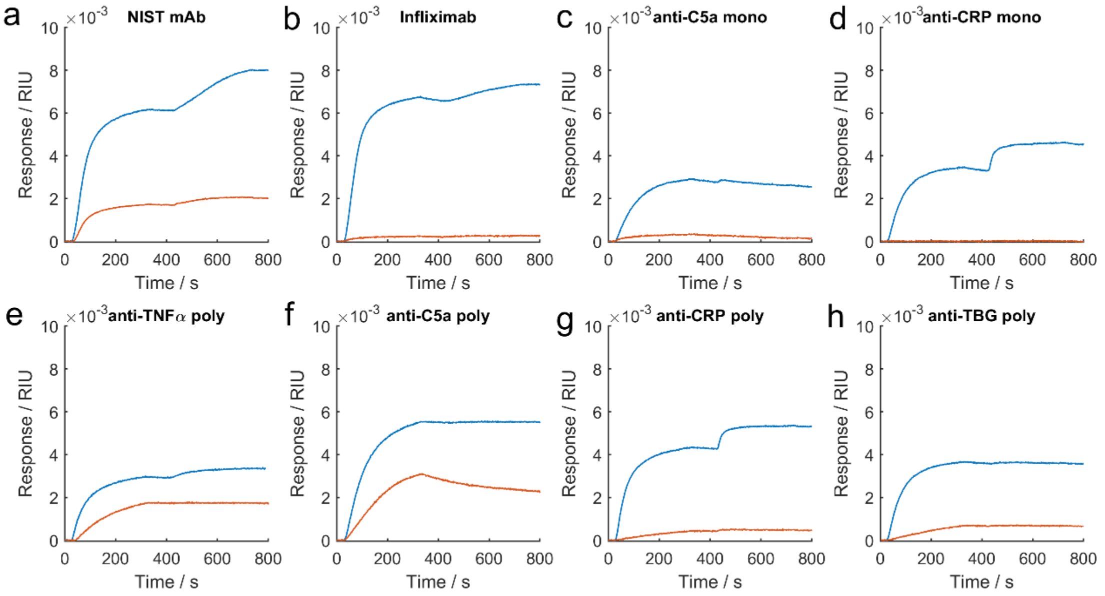


Figure 3.22 Raw kinetic response curves of the PAG surface (blue) and antigen surface (orange) to antibody binding (0-300 seconds) and antigen binding (400-700 seconds). PAG and antigen surfaces were constructed Experiments were performed using the assay method described in 2.8.1. Signals are the averaged response of n=20 assay spots. Two types of antibody were tested: a-e) monoclonal antibody samples and e-h) polyclonal antibody samples.

The $\vartheta_{\mathrm{m\ Fab}}$, $\vartheta_{\mathrm{m\ Fc}}$ and $\vartheta_{\mathrm{m\ antigen}}$ values of all antibody samples were derived by kinetic fitting and are shown in Table 3.7. All values assume antibody concentrations stated in the product data sheets which are confirmed by absorbance at A280 nm. The CV of $\vartheta_{\mathrm{m\ Fc}}$ for the NISTmAb was 7%.

Table 3.7 The $\vartheta_{m\ Fab}$, $\vartheta_{m\ Fc}$ and $\vartheta_{m\ antigen}$ values derived from fitting the Langmuir model to kinetic data from all antibody samples in the panel. The mass of TNFα is shown as a monomer, but there is evidence for dimers and trimers in solution. Errors are 95% CI.

Sample	$\vartheta_{\mathrm{m\ Fab}}$ / mRIU	$\vartheta_{\mathrm{m\ Fc}}$ / mRIU	$\vartheta_{\mathrm{m\ antigen}}$ / mRIU	Antigen Mass / kDa
NISTmAb	1.73±0.03	6.66±0.96	3.37±0.77	53
Infliximab biosimilar	0.22±0.06	6.82±0.21	1.32±0.01	17*
Anti-C5a mono	0.42±0.13	3.07±0.28	0.11±0.04	10.4
Anti-CRP mono	0.02±0.02	3.57±0.13	1.44±0.68	125
Anti-TNFα poly	2.34±0.13	3.23±0.88	0.42±0.06	17
Anti-C5a poly	5.82±1.48	5.36±0.60	0.06±0.04	10.4
Anti-CRP poly	0.76±0.30	4.33±0.09	1.18±0.45	125
Anti-TBG poly	1.60±0.70	3.74±0.24	0.08±0.05	54

The $\vartheta_{\mathrm{m\ Fab}}$ was consistently lower than the $\vartheta_{\mathrm{m\ Fc}}$ for all antibody samples tested, indicating that for a given surface area, more antibodies were adsorbed when bound via the Fc region to the PAG surface compared binding via the Fab region to the antigen surface. The therapeutic Infliximab biosimilar antibody had a particularly low $\vartheta_{\mathrm{m\,Fab}}$ value indicating a low surface density or epitope presentation.

An estimate of the relative binding site density differences between PAG surfaces and the antigen surfaces was obtained by comparing the $\vartheta_{\mathrm{m\,Fab}}$ / $\vartheta_{\mathrm{m\,Fc}}$ ratio of the NISTmAb (0.26 ± 0.04) to that of the panel of antibodies with a range = (0.004 – 1.1) × 10^{-3}. Furthermore, an estimate of antibody monomeric purity was obtained by comparing the $\vartheta_{\mathrm{m\,antigen}}$ / $\vartheta_{\mathrm{m\,Fc}}$ ratio of the antibody samples, after normalising for antigen binding mass: dividing $\vartheta_{\mathrm{m\,antigen}}$ and $\vartheta_{\mathrm{m\,Fc}}$ by the antigen masses and intact IgG masses respectively, Figure 3.23.

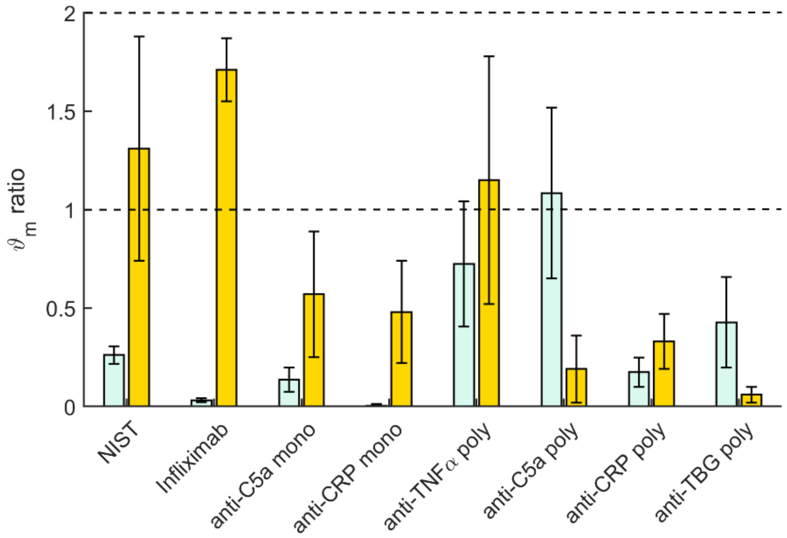


Figure 3.23 ϑ_m *analysis showing the* $\vartheta_{m\,Fab}$ / $\vartheta_{m\,Fc}$ *ratio (green) and the* $\vartheta_{m\,antigen}/\vartheta_{m\,Fc}$ *ratio (orange). The* ϑ_m *values of the* $\vartheta_{m\,antigen}/\vartheta_{m\,Fc}$ *ratio are normalised for the mass of the antigen and intact antibody. The stoichiometric limit of the* $\vartheta_{m\,antigen}/\vartheta_{m\,Fc}$ *ratio is between 1 and 2 (dashed lines). Error bars are 2SD from n=2 determinations.*

To validate the $\vartheta_{\mathrm{m\,antigen}}$ / $\vartheta_{\mathrm{m\,Fc}}$ ratio as a measure of antibody monomeric purity, the biosensor results were compared with results from the ESI mass spectrum recorded for each antibody, Figure 3.24.

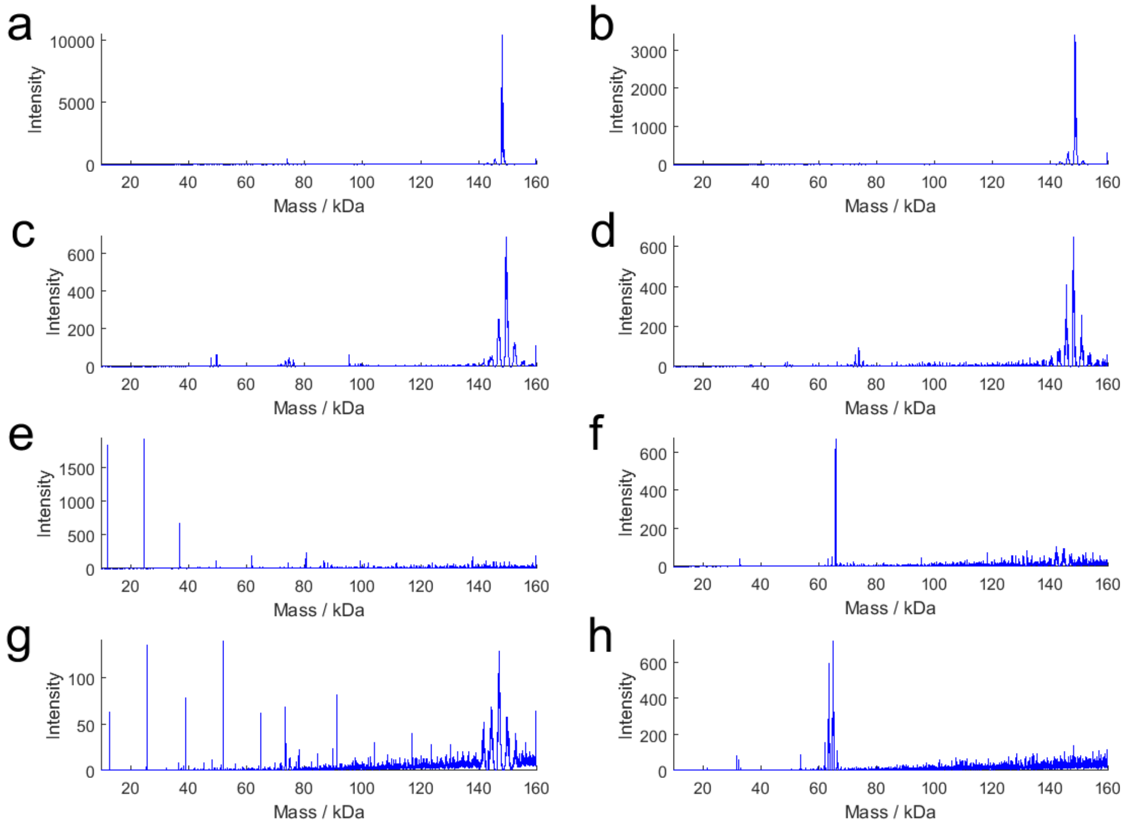


Figure 3.24 The deconvolved mass spectra of antibody samples. Eight samples are shown: a) the NISTmAb; b) Infliximab biosimilar, c) anti-C5a monoclonal, d) anti-CRP monoclonal, e) anti-TNFα polyclonal, f) anti-C5a polyclonal, g) anti-CRP polyclonal and h) anti-TBG polyclonal. The data are the averaged signal from n=3 determinations. More information on these antibody samples is provided in section 3.6.1.

The deconvolved mass spectrum of the NISTmAb is dominated by a species at 148.2 kDa attributed to the pure, non-fragmented IgG structure. All samples showed varying intensities of intact IgG at ~150 kDa, often with multiple glycoforms and many antibody samples show additional lower mass proteins at varying abundances. The NISTmAb Fab fragment has a reported mass of 47 (±5) kDa (439) and the mass spectrum shows a low intensity mass peak at 48 kDa (not visible at the scale of Figure 3.24).

The relative abundances of fragments relative to the intact antibody were determined from the AUC of regions in which masses are typically present: 140-160 kDa, 90-110 kDa, 65-85 kDa, 40-60 kDa and 15-35 kDa, Figure 3.25. These regions of the mass spectrum correspond to the predicted fragmentation products shown in Figure 3.19.

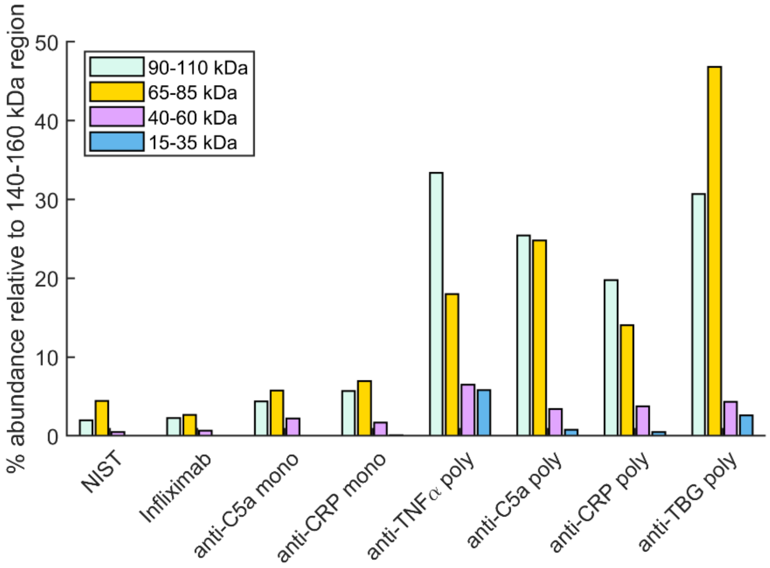


Figure 3.25 The abundances of fragments in each antibody sample evaluated by taking the AUC of specific regions of the mass spectra corresponding to IgG fragmentation products. The percentage abundances of the most abundant peaks are plotted relative to the intact region (140-160 kDa): 90-110 kDa (light blue), 65-85 kDa (gold), 40-60 kDa (pink) and 15-35 kDa (blue). All regions may contain Fc moieties.

The HcLc fragment, expected at 75kDa, was the most abundant fragmentation product in all monoclonal antibody samples, whilst the Fc-Fab fragment, expected at 100 kDa, was most abundant in the polyclonal samples with the exception of anti-TBG. The AUC of impurities relative to the intact IgG AUC was <20% for monoclonal antibodies, whilst all polyclonal samples displayed impurities present at >20% relative to the intact IgG structure. The monomeric purity estimated by ESI was the peak area of the intact region expressed as a percentage of total peak area and is tabulated in Table 3.8. The monomeric purity of the intact NISTmAb was determined by ESI to be 94% compared with the reported value of 96.5% by NIST (431).

3.6.3 Discussion

Antibody sample monomeric purity is critical for immuno-kinetic assay development, binding kinetics studies and effective immunotherapy. Two potential assays of antibody sample fragmentation were assessed here. The first was based on the mass deficit observed for fragmented samples binding via the Fc region to PAG on a the LiScAR sensor surface, the Fc assay. The second assay assessed monomeric purity using the binding ratio of antigen to antibody on a PAG surface, the antigen assay. Affinity of antibody for antigen was confirmed with an antigen surface which captures antibody via the Fab region, the Fab assay. Monomeric purity estimates from the LiScAR assays were validated using native ESI-QTOF mass spectrometry.

ESI data suggest the most abundant glycoform of the intact NISTmAb had a mean molecular mass of 148.2 kDa and the sample displays a distinctive glycosylation pattern. The masses of NISTmAb glycoforms are in good agreement with published values (434), Figure 3.26.

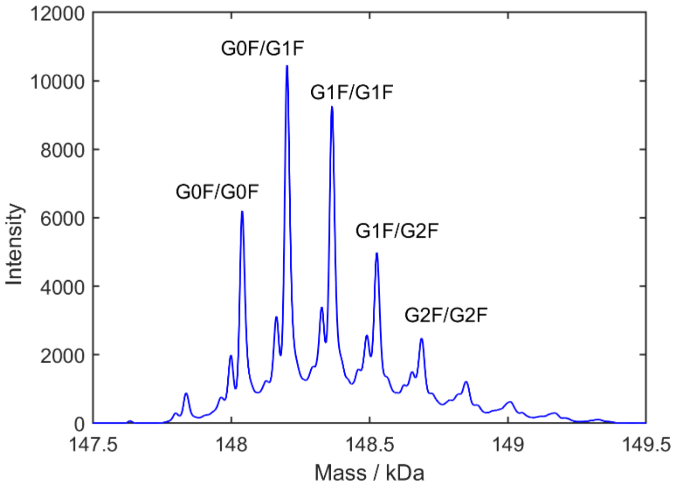


Figure 3.26 The deconvoluted mass spectrum of the NISTmAb intact IgG region. The highest intensity peak has a mass of 148.2 kDa and the glycosylation pattern is in good agreement with previously identified glycoforms in mass spectrometry studies. The data are the averaged signal from n=3 determinations.

The homogeneity of the NISTmAb contrasts with the heterogeneity of intact IgG masses in the anti-C5a monoclonal, anti-CRP monoclonal and anti-CRP polyclonal samples, attributed to multiple N-deglycolsylated glycoforms (443). The ESI data validate the NISTmAb as a good standard antibody of high monomeric purity.

From the ESI data, an effective mass of antibody, M_{ESI}, can be derived from the mass spectrum with the equation:

$$M_{ESI} = \frac{aA + bB + cC + dD}{a + b + \mathrm{c} + \mathrm{d}} \qquad 3.6$$

where lowercase letters represent the percentage AUC of the mass spectrum relative to the intact peak area (140-160 kDa) from the regions: 90-110 kDa (predicted Fab-Fc complex), 65-85 kDa (predicted HcLc complex), 40-60 kDa (predicted Hc or Lc) and 15-35 kDa (predicted Lc or Hc monomer of Fc). Uppercase letters represent the mean mass of the integration region: 150 kDa (intact), 100 kDa (predicted FabFc complex), 75 kDa (predicted HcLc complex), 50 kDa (predicted Hc or Lc) and 25 kDa (predicted Lc or Hc monomer of Fc). The derived M_{ESI} for each of the antibodies in the panel are shown in Table 3.8.

The monomeric purity of the NISTmAb (and reassuringly that of the therapeutic infliximab biosimilar) was high, which was not the case for the other commercially available samples that would be used for research applications. The reliability of experiments involving Western blots, immunohistochemistry, flow cytometry and immunoprecipitation is negatively impacted by variations in antibody purity as these techniques fundamentally rely on a connection between Fc and Fab regions of the antibody. The consequence of purity on kinetic analysis needs to be considered with the fundamental errors of the 1:1 binding model or Langmuirian kinetics conventionally used to analyse antibody binding events. Fitting a set of low concentrations in the global fit methodology for the NISTmAb produces and error of 15% for k_a and 10% for k_d due to the uncertainties in the fit over the timescale of the experiments (368). Model fitting errors would dominate the analysis of the NISTmAb and Infliximab biosimilar samples but this would not be true for all members of the antibody panel tested here. For the case of the anti-TBG sample, any reported antibody kinetics analyses without a sample purity assessment may have an error of up to 46% due to low monomeric purity.

The packing density of IgG around a protein is a property of epitope density and steric packing. Antibodies with identical masses should therefore achieve identical ϑ_{m} values for a given antibody packing density, albeit after differing periods of time to allow variations in association kinetics. There was significant variation in the $\vartheta_{\mathrm{m\ Fab}}$ (Table 3.7) which may be explained by variation in availability of epitopes for antibody binding as evidenced by the larger $\vartheta_{\mathrm{m\ Fab}}$ of polyclonal anti-CRP compared to the monoclonal anti-CRP sample on the same CRP surface. In contrast, the $\vartheta_{\mathrm{m\ Fc}}$ showed less variation than $\vartheta_{\mathrm{m\ Fab}}$, attributed to the use of a common PAG surface for all antibody samples with a consistent number of binding sites.

Variation in $\vartheta_{\mathrm{m\ Fc}}$ may occur due to differing mean IgG mass differences (monomeric purity differences) between IgG samples as observed with ESI. The Fc assay presented here will bind with high affinity to both intact IgG and the most abundant fragmentation products which contain the Fc region - a concern for PAG based affinity purification of antibodies but a potential antibody fragmentation assay when combined with surface plasmon technology. Any given ϑ_{m} derived from fitting of the Langmuir equation may be related to the adsorbed mass on the surface, *M*, by de Feijter's formula (444):

$$M_{adsorbed} = d_A \frac{\Delta RI}{dn/_{dc}} \qquad 3.7$$

where d_A is the thickness of the adsorbed layer and *dn/dc* is the rate of change of refractive index with surface concentration (the refractive index increment) and ΔRI is the change in the refractive index. The value of *dn/dc* has been measured for protein as 0.182 cm^3 g^{-1} and shown to vary with buffer conditions (445) but the effect is negligible (444).

The largest uncertainty in the estimate of M is the value of d_A. IgG layers adsorbed onto a protein A and silica surfaces have been observed previously using ellipsometry with thicknesses of 4 nm to 16 nm respectively (446, 447). Neutron reflectivity measurements of IgG4 binding to a protein A / BSA surface report three sequential layers leading to total thicknesses between 6.1 nm and 25 nm depending on washing and blocking of the surface (448) because nonspecific 3D aggregation upon adsorption at high IgG concentrations can enable adsorbed layers to exceed the maximum length of a single IgG molecule, 16 nm (449). The mean cross section dimension of the IgG crystal structure (450) suggests a d_A of 10 nm, assumed constant for all antibody samples. The surface adsorbed mass density depends only on the difference in the refractive index (ΔRI) as a result of adsorbed antibody.

Substituting $\vartheta_{\mathrm{m\,Fc}}$ of NISTmAb for ΔRI in equation 3.7 predicts the areal density of the NISTmAb when all Fc binding domains on the PAG binding sites are filled, $M_{\mathrm{adsorbed}} = (3.66 \pm 0.53) \times 10^{-8}$ g cm^{-2}. By comparison, electrostatic adsorption of monoclonal antibodies onto a variety of surfaces has been reported at densities in the range 5×10^{-7} to 5×10^{-6} g cm^{-2} (451, 452). Electrostatic adsorption is expected to achieve a higher surface mass density compared with the ordered two-dimensional layer of IgG bound to a PAG surface. The M_{adsorbed} values of each antibody studied are presented in Table 3.8.

The ESI data indicate that effectively all protein in the NISTmAb sample was antibody monomer (93.5%) with a monomeric mass, M_{monomer}, of 148.2 kDa. The NISTmAb data and the Avogadro constant, N_{a}, can be used to calibrate the surface density of intact antibodies bound at saturation, $N_{\mathrm{Ab}} = (M_{\mathrm{adsorbed}} / M_{\mathrm{monomer}}) \times N_{\mathrm{a}} = (1.49 \pm 0.21) \times 10^{11}$ cm^{-2}. Comparatively, the maximum number of monoclonal antibodies in a monolayer has been determined by adsorption onto a silicon nitride surface as 1.2×10^{12} cm^{-2}, although it was estimated that up to 75% of the protein may be non-specifically adsorbed rather than covalently bound (453). The unknown purity of antibody samples in the literature is a potentially large source of error for published N_{Ab} values derived from protein mass density measurements whereas the N_{Ab} figure presented here is of known accuracy due to the monomeric purity estimate from ESI.

The number of PAG binding sites on the surface was constant for the NISTmAb and if all antibodies bound to the surface with the same packing density then an estimate of effective mass at each binding site, M_{Fc}, may be estimated from asserting a simple ratio:

$$\frac{\vartheta_m^{Fc}}{M_{Fc}} = \frac{\vartheta_m^{monomer}}{M_{monomer}} \qquad 3.8$$

where $\vartheta_m^{monomer}$ and $M_{monomer}$ are the $\vartheta_{m\ Fc}$ and mass of the monomeric reference material, which in the present study is the NISTmAb with $\vartheta_{m\ Fc}$ = 6.66±0.96 mRIU and a mass of 148.2 kDa. When a sample has degraded, a fraction of Fc binding domains on the PAG will be occupied by fragments with a lower mean molecular mass than the intact IgG: a mass deficit producing a lower $\vartheta_{m\ Fc}$ for constant number of occupied binding sites. The estimated monomeric purity from the Fc assay can be derived by expressing M_{Fc} as a percentage of $M_{monomer}$, Table 3.8.

Table 3.8 Experimental properties derived for the antibody panel.

Sample	ESI mass, M_{ESI} / kDa	ESI Monomeric purity / %	Fc areal density, $M_{adsorbed}$ / g cm^{-2} $\times10^{-8}$	Plasmonic mass, M_{Fc} / kDa	Plasmonic monomeric purity / %
NISTmAb	145	93.5	3.7±0.5	148±21	100±14
Infliximab biosimilar	146	94.7	3.8±0.1	152±5	103±3
Anti-C5a mono	142	89.0	1.7±0.2	68±6	46±4
Anti-CRP mono	141	87.4	1.96±0.07	80±3	54±2
Anti-TNFα poly	123	61.1	1.8±0.4	72±19	49±13
Anti-C5a poly	127	64.7	3.0±0.3	119±12	80±8
Anti-CRP poly	132	72.4	2.38±0.05	96±2	65±1
Anti-TBG poly	119	54.2	2.1±0.1	83±5	56±3

A weak positive correlation is observed between M_{ESI} and M_{Fc}, the estimates of mean antibody mass in a sample according to ESI or the LiScAR Fc assay, Figure 3.27.

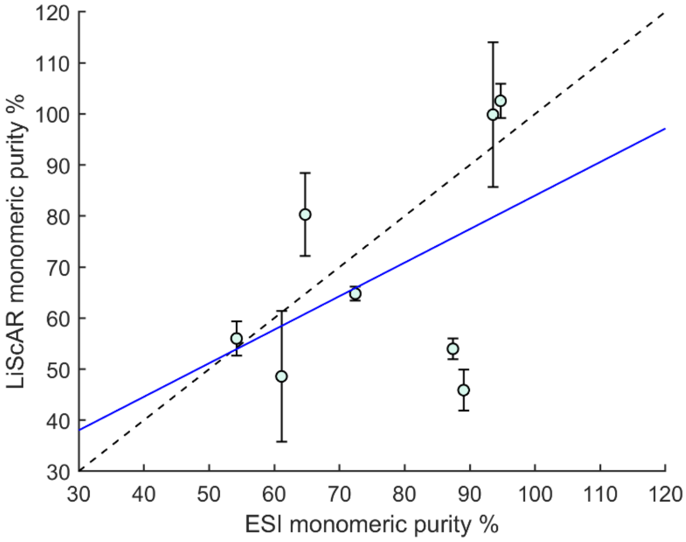


Figure 3.27 Correlation between the monomeric purity of antibody samples estimated by the LiScAR Fc assay and ESI. A linear fit to the data (blue) and perfect correlation (dashed black) are shown. R^2=0.22. Error bars are 2SD from n=2 determinations.

Several phenomena may account for lack of correlation observed. Firstly, N_{Ab} at $\vartheta_{\mathrm{m\ Fc}}$ may be a fundamental property of a given antibody due to glycan variation and the locations of PAG binding domains on the structure, leading to some antibody monomers binding in orientations which favour closer packing on the surface (454). The $\vartheta_{\mathrm{m\ Fc}}$ will also be sensitive to nonspecific aggregation and clustering of IgG (448) although aggregates of the NISTmAb IgG are only observed following deliberate physical agitation for days at room temperature (455) and the antibody samples tested here are not expected to contain significant aggregates due to appropriate storage.

The lack of correlation between $\vartheta_{\mathrm{m\ Fc}}$ and ESI monomeric purity may best be explained by the orientation and higher packing density of low mass antibody degradation products, which may have access to a greater number of binding sites on the PAG surface than monomeric IgG. Therefore, it is likely that the $\vartheta_{\mathrm{m\ Fc}}$ of the Fc assay is only sensitive to severe antibody degradation.

A more sensitive measure of monomeric purity is to support antibodies on the PAG surface and measure the antigen binding ratio, $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$, which after normalising for molecular mass, shows a strong dependence on monomeric purity estimated by ESI, Figure 3.28.

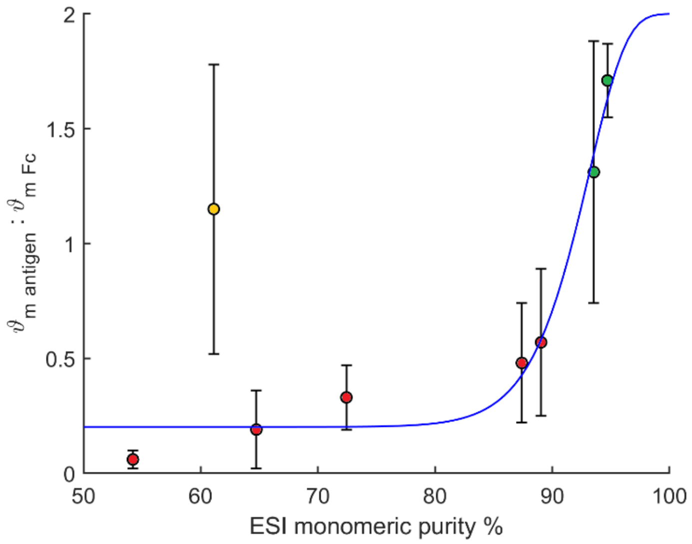


Figure 3.28 The relationship between monomeric purity estimated from the ESI data and the mass normalised $\vartheta_{m\ antigen}$ / $\vartheta_{m\ Fc}$ *biosensor measurement. The NISTmAb and Infliximab biosimilar (green) are >90% pure according to ESI data and exhibit* $\vartheta_{m\ antigen}$ / $\vartheta_{m\ Fc}$ *ratios within the expected limits of 1-2 antigens per antibody. Other commercially available antibodies (red) are <90% pure and show less than 1 antigen bound per antibody. A 5-parameter logistic curve fits well to the data (blue) is constrained by a maximum asymptote of 2 (the upper limit of the antigen-antibody interaction stoichiometry). The anti-TNFα polyclonal (orange) does not fit the trend line, likely due to a unique fragmentation pattern resulting in a majority of 100 kDa HcLc fragments which may pack tightly on the PAG surface and remain capable of binding 1 antigen. Error bars are 2SD from n=2 determinations.*

The $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio measure of antibody integrity shows a strongly rising trend with the pure antibodies showing a ratio that falls within the expected stochiometric limits of 1 or 2 antigens binding to per antibody. The Infliximab biosimilar antibody binds to a relatively low mass antigen (TNFα with a monomer mass of 17 kDa) and displayed an antigen-antibody binding ratio of 1.71 ± 0.16. The NISTmAb binds to a higher-mass antigen (53 kDa) and displayed a lower antigen-antibody binding ratio of 1.31 ± 0.57 attributed to greater steric hindrance of antigens. The $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio is antigen-dependent for larger antigens but would always exceed 1 with adequate antibody spacing on a plasmonic surface. The NISTmAb performed well as a standard material for the antibodies shown, but calibration of the assay would likely be improved by comparing each tested sample to a high-purity reference standard of the same antibody.

Of the commercially available research antibodies, anti-TBG antibody showed a relatively high $\vartheta_{\mathrm{m\ Fab}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio but displayed the weakest $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio. These two ratios indicate that both Fab and Fc regions of TBG were present in the sample, but a significant proportion were separated due to degradation. The high fragmentation suspected in the anti-TBG antibody is supported by the lack of intact antibody in the ESI mass spectrum, Figure 3.24.

The $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio is a measure of the biological activity of an antibody sample, whereas the ESI monomeric purity estimated from mean mass penalises any form of antibody fragmentation. The ESI data predicted a low monomeric purity in the anti-TNFα polyclonal but the sample displayed a relatively high $\vartheta_{\mathrm{m\ antigen}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio. The result may be explained by the unique degradation pattern shown in Figure 3.25 which indicates that the most abundant species was the 100 kDa HcLc fragment which is capable of binding to PAG and one antigen. The fragmentation in the anti-TNFα polyclonal remains identifiable by comparison to the Infliximab biosimilar which shares the same antigen but exhibited a higher antigen-antibody binding ratio.

Using the $\vartheta_{\mathrm{m\ Fab}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio recorded by the LiScAR, the antibody panel readily divided into two types of samples: the high purity (>90%) samples (NISTmAb and Infliximab biosimilar) and the remainder of the panel with lower purity (<90%). The $\vartheta_{\mathrm{m\ Fab}}$ / $\vartheta_{\mathrm{m\ Fc}}$ ratio was significantly different for these two material types as evidenced by a two-sample *t*-test and two sample Kolmogorov-Smirnov (P=0.01 and P=0.03 respectively). These statistical tests are discussed in Chapter 6. The assays developed here show the potential to test the specification of antibodies immediately before use as assay materials and other antibody-dependent techniques to prevent variability in experimental results caused by sample degradation.

3.7 Conclusions

The NISTmAb provides an important new benchmark material which enables the calibration of biosensors assessing monomeric purity of antibody samples with assays for Fab, Fc and antigen. Analysis of a panel of commercially available research grade antibodies using ESI-QTOF and the LiScAR highlighted significant variation in monomeric purity and raises accuracy concerns for previously reported k_a, k_d and K_D measurements of similar materials in the literature. Without sample monomeric purity analysis (of both the antibody and the antigen) these measurements may have errors greater than 46% in highly fragmented samples. The antigen assay technique to estimate antibody monomeric purity can be performed in 15 minutes and may be combined with additional rapid tests for IgG aggregates as described for the NISTmAb in ref (455).

Four novel assays have also been presented here for the quantitative concentration determination of C5a, TCC, C3 and TNFα. The calibration curves of the four new assays are reproducible and can be performed simultaneously with purified PBS standards in a 4-plex assay. Of the new assays, the C5a and C3 assays were suitably sensitive and selective for use with serum samples and can be combined into a 2-plex assay with a duty cycle of <15 minutes. The Complement assays represent a proof-of-concept for point-of-care Complement testing and are used in Chapter 7 to collect high frequency kinetic data of Complement activation *in vitro*.

Chapters 2 and 3 have demonstrated how immobilisation of proteins and antibodies on the LiScAR surface enables binding kinetics to be studied. Chapter 4 introduces fundamental systems biology modelling techniques and considers whether the binding kinetics measured for surface interactions under PBS conditions reflect the binding interactions that occur in complex mixtures such as the serum. The transferability of kinetic models is a critical consideration when constructing a complex systems biology model such as that of Complement activation.

4 Kinetic Modelling of Biological Systems

4.1 Introduction

The mathematical modelling approach of systems biology is becoming an established strategy for understanding complex biological systems (456). Computer modelling tools such as Systems Biology Markup Language (SBML) (457) and COPASI (458) are readily available and allow computer models to be constructed and shared between research groups. Four main types of models can be identified (459): verbal, diagrammatic, physical and formal. Verbal models are used to qualitatively represent a system in words and any ambiguity is then clarified by showing the system structure or interactions in diagrammatic form. Physical models may be used when the scale of a system is not relevant such as spatial models of proteins. Formal models represent quantitative knowledge of the system using mathematical equations and can be interrogated rapidly using a computer. Formal models must be constructed with care, with consideration given to the isolation of certain objects (compartmentalisation), the treatment of time as a continuous or discrete variable and whether similar objects are represented as individuals or by populations. To further increase the variety of models, the interactions of species in the model may rigidly follow a set of rules (deterministic) or occur via chance (stochastic).

Systems biology asserts that since genes and proteins are molecules and exert their functions through other molecules, a thorough explanation of a biological system may be achieved by combining knowledge of all molecular interactions involved (460). Biochemical systems obey the same physical laws as physics and chemistry, such as mass balance and energy transfer. Biochemical interactions can therefore be represented by chemical reactions, which can be used to construct deterministic differential equation models describing each species in system (461). Differential equations concerned with a single independent variable, such as time, are referred to as ordinary differential equations (ODEs). A reversible bimolecular binding reaction between reactants A and B is shown in equation 4.1 as a molecular equation, whilst equation 4.2 describes the same reaction as an ODE for the rate of concentration change of A over time:

$$A + B \underset{k_{d1}}{\overset{k_{a1}}{\rightleftarrows}} AB \qquad 4.1$$

$$\frac{d[A]}{dt} = k_{d1}[AB] - k_{a1}[A][B] \qquad 4.2$$

where A and B could represent distinct proteins and AB is the complex formed when A and B bind together. As seen in previous chapters: k_a and k_d represent the association and dissociation rate constants respectively. The ODE describes the change in the concentration of A with respect to time, based on the concentrations of A, B and AB. The ODE has a positive term which describes the formation of A from the breakdown of AB and a negative term which describes the removal of A by the reaction with B. In a larger system where A might also bind to proteins other than B, additional negative terms would be included in the ODE for the consumption of A. By extension, all reactions in a larger network can be written as rates of change of concentration in time based on the formation and destruction of complexes. The ODE shows that a critical distinction is made between the concentration of A as part of the AB complex and unbound A which is free to react with B. Stoichiometry is conserved and the total quantity of a species is calculated by the sum of its unbound and bound forms: $A + AB = A_{total}$.

ODE models are widely used in systems biology to quantitatively predict the kinetic behaviour of complex systems because they are able to consider both structure and reaction kinetics of a system (461, 462). A systems biology approach could be used to test the hypothesised reaction mechanism between A and B shown above by substituting the starting concentrations (of A, B and AB) and rate constants (k_a and k_d) into an ODE describing each species in the system. An ODE solver iteratively calculates the concentration of each species in the system after distinct time steps and returns the corresponding concentrations at each time step. The model simulation can then be compared to experimental time series data where a good agreement between the model and the data provides evidence that the structure of the model, its kinetic constants and starting conditions may be correct.

The process of predicting system outputs based on knowledge of the system interactions is known as the ‘forward problem’ (463). Such an approach is possible if the starting concentrations of all species in the model and the rate constants are known, but this is often not the case in biological systems. Some rate constants may not have been measured experimentally, and the concentrations of certain species may be unknown. Furthermore, such integrated models may not fit well to biological data (464, 465).

The ‘inverse problem’ is the systems biology approach used to estimate model structure or unknown parameters such as rate constants by fitting model outputs to data (466). Data fitting (also termed optimisation) is typically achieved with some variation of the routine shown in Figure 4.1.

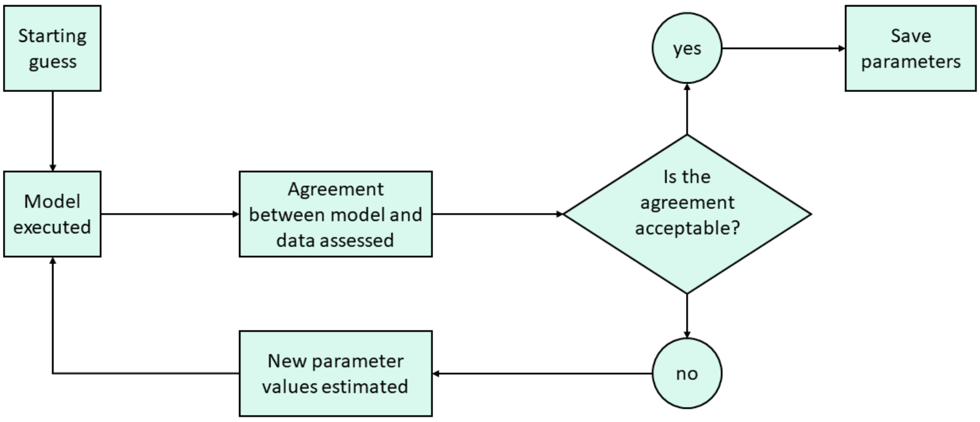


Figure 4.1 Process diagram of parameter estimation by model fitting to data. Agreement threshold to exit the loop is predefined before the process begins. The agreement is typically assessed with an objective function such as root mean square or relative absolute error. An escape from the loop may also be included to terminate the routine in the event that acceptable agreement is not reached - a maximum number of attempts for example.

The parameters to be estimated (or floated) are given initial guess values which are iteratively changed until the agreement between model and data reaches a predefined threshold: an optimised model. The goodness-of-fit is evaluated for each new set of parameters with an objective function (also termed cost or merit function) whose minimum or maximum value is sought (depending on objective function) by an optimisation procedure (467). Typically used objective functions include root mean square error (RMSE) or mean absolute error (MAE) which indicate a better fit when their values decrease and hence the task of the fitting routine is to minimise these objective functions.

The decision to terminate the fitting routine is performed based on predefined criteria. For example, the routine may stop upon reaching a parameter combination which produces an objective function below some threshold such as the confidence intervals of the data. More commonly, fitting routines may end regardless of the absolute objective function value when improvements to the fit are only achieved with very small changes to the parameters (a parameter precision tolerance). Finally, the fit may end after a certain number of parameter combinations have been evaluated. Finding the optimal combination of parameters may be computationally expensive (requiring many iterations), especially when estimating large numbers of parameters which produce nonlinear trends in the objective function (468). It is very inefficient to grid search the parameter space of large models with many floated parameters and more efficient parameter estimation algorithms have been developed (462, 465).

Many parameter estimation algorithms are available but may be categorised into two groups: local and global (469). Nonlinear functions may contain many local minima: parameter estimates which produce better objective function values than similar estimates but do not produce the single best solution. If several local minima exist, the solution which produces the smallest objective function value is referred to as the global minimum. Local parameter estimation algorithms find the lowest objective function value within the basin of attraction of the starting point. A basin of attraction is a set of starting parameter values which lead to a local minimum in the objective function when each parameter is varied in the direction associated with the favourable change in objective function value. The basins of attraction for a local and global minimum, due to the single floated parameter x, are shown in Figure 4.2.

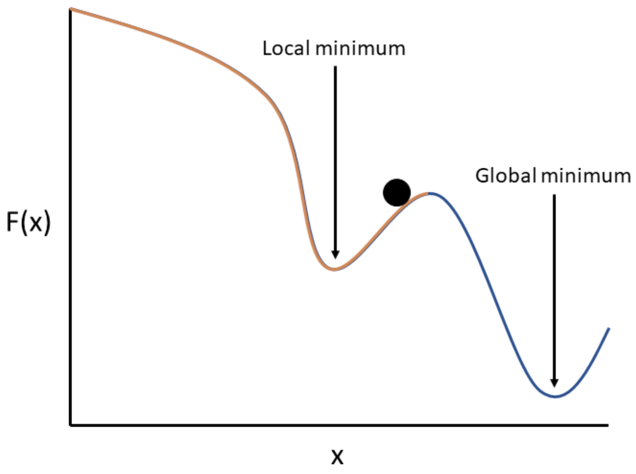
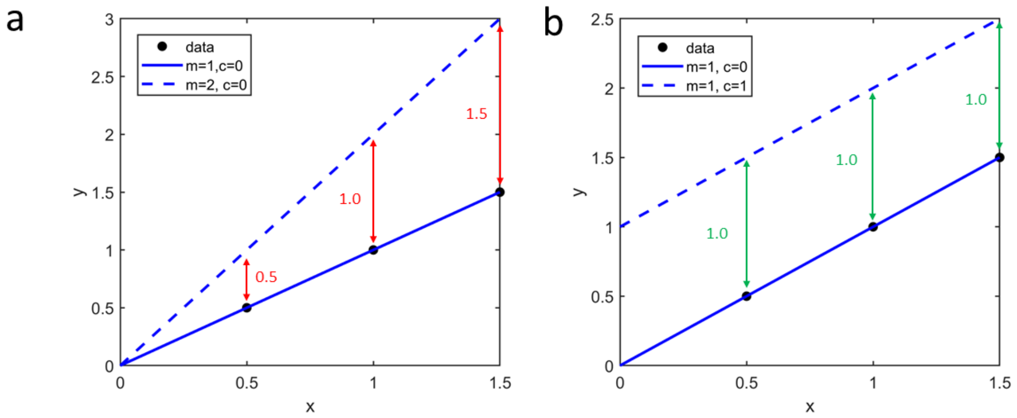


Figure 4.2 Sketch of a two-dimensional objective function surface, showing the objective function value, F(x), at different values of x. A starting value (black circle) and the basins of attraction for the local minimum (orange) and the global minimum (blue) are shown.

Most systems biology data fitting problems are multi-dimensional and their objective function surfaces cannot be visualised graphically. However, basins of attraction are still present in higher dimensions and Figure 4.2 highlights the importance of a good initial guess when using local solvers. If the initial guess for x was within the local basin of attraction, then a local solver would not find the global solution as this would involve increasing the objective function value before reaching the global minimum. Global solvers attempt to search more than one basin of attraction to improve the probability of finding the global solution. The search efficiency and confidence that the fit solution is global can be improved by the inclusion of *a priori* knowledge of the system under study. Many fitting algorithms use bounds to constrain the parameter values within theoretically plausible ranges, although such *a priori* information should be used with caution to avoid bias in the model.

Given that the equations of the model are constant during parameter estimation, confidence in the modelled reaction network is essential. Mechanistic knowledge of the system under study is therefore required to produce an accurate model. Incorrect assumptions of stoichiometry and excluded system inputs / outputs will lead to estimated parameters with little or no physical relevance to the true process. If the structure of the model is uncertain, the modelling process often includes the construction of many alternative model structures which can be compared after data fitting (470). The structures of models which fit poorly to the data can be rejected, providing some mechanistic insight, however the true mechanism may still be unclear as a potentially infinite number of incorrect model structures could produce equally good fits to the data (471). For example, the model $y=mx+c$ will fit well to data showing a linear relationship between x and y. The model $y=mx+(d/e)$ will fit equally well to the data when $d/e=c$ but an infinite number of equally well-fitting solutions for parameters d and e are possible. Parameters d and e are therefore referred to as being highly correlated and as such their values are poorly defined by the data because as one increases the other decreases proportionally to achieve the same fit result. A recent review of 17 systems biology models showed that many of the obtained model parameters were ‘sloppy’, meaning not well-defined (472). At the fit solution, several techniques are available to assess how confidently the parameter values are determined by the data.

Insensitivity of the cost function to variation in a parameter value leads to large confidence limits on the derived value of a parameter – low precision estimates. A sensitivity analysis is used to estimate the change of the cost function with respect to a given model parameter. The sensitivity of a function with respect to each variable can be evaluated at each observation in a data set to generate a Jacobian matrix (473), Figure 4.3. The Jacobian matrix indicates the relative influence of each parameter on the goodness-of-fit at different regions of the data set and can be used to assess the correlation of model parameters .



c

	Δ function / Δ parameter	
	m	c
observation 1	0.5	1
observation 2	1	1
observation 3	1.5	1

Figure 4.3 Estimation of the Jacobian matrix for the function y=mx+c. The function fits perfectly to the data (residuals = 0) when m=1 and c=0 (solid blue). Two variations of the fitted function are shown: a) parameter m is changed by 1; b) parameter c is changed by 1. The change in function value at each observation (red) is recorded to produce c) the Jacobian matrix.

The columns of the Jacobian matrix shown in Figure 4.3 do not show any correlation: a change to the gradient parameter, m, increasingly affects the goodness-of-fit at larger values of x, whereas a change to the intercept parameter, c, affects the goodness of fit equally at all observations. The lack of correlation (positive or negative) between columns in the Jacobian indicates that the estimated parameters are well-defined by the model fit to the data. In contrast, a Jacobian matrix from fitting parameters d and e in the model *y=(d/e)x+c* would show a perfect negative relationship between the columns for *d* and *e* parameters, indicating no unique solution to the optimisation problem due to the high correlation in these parameters.

The relationship between two columns in the Jacobian can be assessed quantitatively using covariance. Performing a covariance analysis on all pairwise combinations of columns in the Jacobian matrix is used to generate the variance-covariance matrix, where the off-diagonal elements are the covariance of two parameters and the diagonal elements are simply the variances of each parameter. As described in Chapter 3, parameter relationships are more easily compared by normalising covariance to correlation. The process of converting of the variance-covariance matrix to the correlation matrix is shown as an example in Figure 4.4.

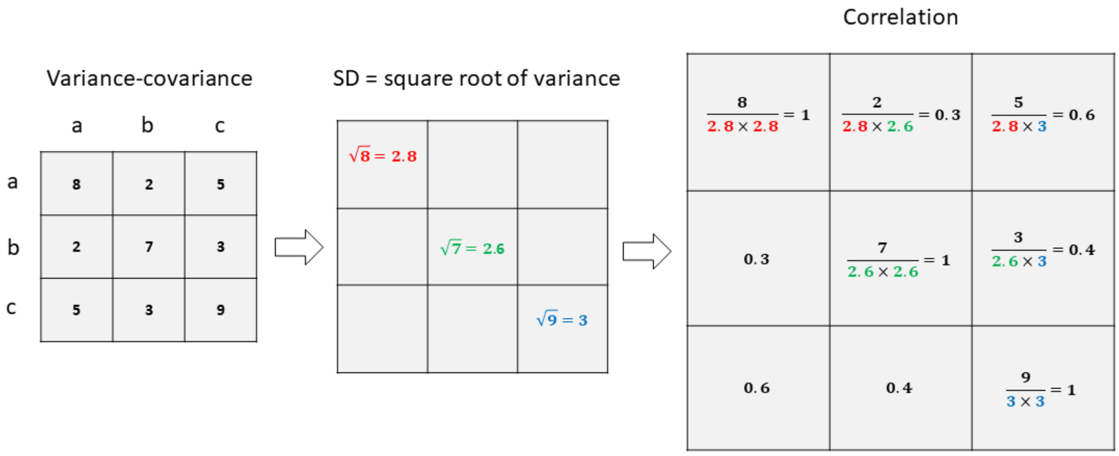


Figure 4.4 Flow chart showing the generation of the correlation matrix from the covariance matrix. The variance-covariance matrix of parameters a, b, and c is used to generate the standard deviation of each parameter. The correlation matrix is generated by dividing the variance-covariance matrix by the product of the standard deviations for each parameter. The correlation matrix indicates that parameters a and c are more correlated with each other than with b.

Parameter correlation during model fitting may be avoided by several strategies. Firstly, additional data may be included in the fit which is disproportionately sensitive to one parameter in a correlated pair. Secondly, the number of floated parameters may be reduced. This may be achieved by simplification of the model to include fewer parameters, fixing certain parameters or asserting that multiple parameters should hold the same value based on *a priori* knowledge (474). The first approach may prevent the deduction of physically representative parameters such as rate constants if the model structure does not represent the true mechanisms in the system under study. The latter approaches require mechanistic understanding of the system.

The requirement for mechanistic understanding before parameters can be usefully estimated is a major concern for critics of systems biology who claim that the inverse problem is impossible to solve for such complex systems with unclear mechanistic pathways (475). A rebuttal to the sceptics of systems biology is the emergent field of synthetic biology in which systems biology has provided the knowledge necessary to develop effective synthetic biological systems (476). Examples include the prediction of genetic modifications for optimal succinate production in *E. coli* (477) and control of metabolic flux with synthetic protein scaffolds based on systems level knowledge of the metabolic pathways (478). A systems level understanding must precede the construction of optimised synthetic systems and demonstrating a systems level understanding is possible with a validated kinetic model.

Model validation is an essential step in the systems biology approach to modelling. A model must be able to predict the result of observations, but successful predictions cannot guarantee that a model is correct but rather fail to to provide evidence for rejection (479). Published systems biology studies appear to use a common strategy for the development of kinetic models (461):

1. The model structure is defined, often by ODEs, based on existing knowledge of the system under study.
2. Unknown parameters such as rate constants are estimated by fitting the model to experimental data (training data), typically in the form of time series concentration data of one or more species in the model.
3. The agreement between model and data is assessed for both training data and independent data under different experimental conditions which were not used for parameter estimation: model validation.

Models intended for use as quantitative predictive tools should always be validated and predicting the behaviour of the system in a new experimental setting is the ultimate test of predictive kinetic models (461). Partitioning of the existing data into training and validation data sets is common practice but, as with model structure design, may be subject to bias (discussed in Chapter 6). One approach is to fit the model to all available data before repeated validation with new data as it is made available with successive studies.

There are several reasons for the inability of a model to predict the outcomes described by a new data set despite fitting well to the training data set. The model structure may not be applicable to the conditions of the validation data and model parameters may be poorly defined during optimisation or incorrectly defined due to multiple local minima during the fitting process. For qualitative investigations, the precise value of a parameter is not always required to draw meaningful biological conclusions of system dynamics (480), but quantitative predictions require accurate and transferable model parameters. A reasonable modelling approach is therefore to include relevant published rate constants where available (481) and fit the remaining parameters to experimental data (469). Ideally, the included published rate constants should be derived from a similar experimental setting to the data for model fitting and such a setting should be representative of the intended system setting for model predictions. For example, if predicting system behaviour *in vivo* is the intended use of a model, care should be taken so that the *in vitro* conditions used for model optimisation reflect the *in vivo* conditions as much as possible (482). Association rate and equilibrium constants have been shown to vary by an order of magnitude due to nonspecific interactions and molecular crowing in actual intracellular environments (483). Kinetic studies performed *in vitro* have the advantage that the experimental conditions can be more strictly controlled compared to *in vivo* and perturbations can be made to estimate parameter values even for detailed kinetic mechanisms (484).

Equation 4.2 describes a reaction in a standard system setting such as PBS but the active or effective concentration of a species available in a different system setting, such as a surface, may be different. Models of protein binding reactions may be complicated by properties such as avidity, steric effects and interface activity (485, 486) which likely differ between system settings. Furthermore, the excluded volume effect of macromolecular crowding increases the effective concentration of macromolecules in a solution (487). The effective concentration of a species can be modelled as the product of the known concentration and an activity coefficient:

$$a_1 = \gamma_1[A] \qquad 4.3$$

where γ_1 is the activity coefficient of the species A and depends on the difference in protein activity between the observed system setting and the standard setting used for kinetic modelling - the difference between PBS and cytoplasm for example. Significant differences in the effective activity of a species observed in varying system settings may fundamentally limit the predictive usefulness of a deterministic model derived from a single setting (488). These transferability uncertainties of γ highlight the difficulty in applying a model to different system settings.

Models of protein binding reactions are particularly important in the field of synthetic multivalent ligand binding which studies next-generation therapeutic agents and biomaterials (489). The architecture of synthetic protein scaffolds may be genetically designed to vary in size and valency, which control reaction kinetics and mechanism of action (490). The genetically-modified protein PAG has multiple binding sites with high affinity for mammalian immunoglobulins: a model multivalent binding system (491). There is genetic evidence for 8 binding domains on PAG which are derived from the binding domains of protein A and protein G (492). However, when purchased, PAG is described as having 6 binding sites despite the genetic evidence for 8 and there is clearly an unknown fraction of binding sites available sterically, Figure 4.5.

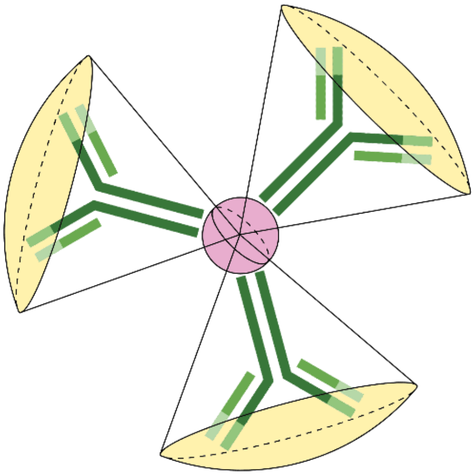


Figure 4.5 Structure of the PAG-IgG complex. PAG (purple) is shown binding to IgG (green). Each bound IgG will occupy a fraction of the available spherical surface area (yellow) and must reduce the solid angle available to further IgG molecules.

PAG is widely used for ligand capture in affinity chromatography (493) where there is an assumption that the protein activity on a separation surface is the same as in solution and in large excess due to the number of binding sites and mass on the surface. The number of IgG binding sites on the PAG is unclear, with conflict between the value quoted by suppliers and the genetic modification data. Biochemical titration data and a kinetic model of the system may enable the binding ratio to be observed directly and provide data for kinetic modelling of this relatively simple system.

4.2 Aims and Objectives

The aim of this chapter is to demonstrate a systems biology approach to mathematically modelling a multivalent ligand binding system (IgG binding to PAG with unknown stoichiometry) using the techniques described previously, ahead of the more challenging task of modelling the Complement cascade in Chapter 5.

The first objective was to record experimental kinetic data of the system under study. The biophotonic assay for IgG, described in Chapter 2, was used to collect high-frequency data of IgG concentration following mixing with PAG.

The second objective was to construct a mechanistic model of the PAG-IgG system. The model structure and rate constants were derived in the standard system setting of PBS by fitting the model to the experimental data. Parameter optimisation methods were considered to ensure a global minimum is found and parameter correlation is assessed and minimised based on the methods discussed previously.

The final objective was to compare the reaction kinetics of the system in PBS to those in an alternate system setting where PAG is immobilised on a sensor surface, from which activity coefficients can be derived. The transferability of model parameters and structure between system settings is assessed.

4.3 Materials and Methods

Pierce recombinant protein A/G (5 mg) was obtained from Thermo Scientific (#21186). IgG antibody used was rabbit anti-sheep (1 mg/mL) obtained from Bio-Rad (#5184-2304). Human serum albumin and bovine serum albumin were obtained from Sigma-Aldrich. Standard instrument running buffer, also used in preparation and dilution of the samples, was phosphate-buffered saline (PBS) supplied in tablet form by Sigma-Aldrich. Ortho-phosphoric acid (85%) was obtained from Fluka and a 0.1 M aqueous solution used as regeneration buffer. GAPSII aminated glass prep slides were obtained from Corning (#40006).

The IgG test array was constructed using a different printer to that described in the SOP in Chapter 2 (AJ1000 from ArrayJet, Edinburgh, UK). Printing accuracy was improved by increasing the viscosity of printed solutions and hence glycerol was routinely included in the samples. The assay was constructed using 24 assay elements functionalised with PAG printed from a 2 mg/mL solution in PBS and

50% glycerol. 16 light intensity controls were printed from a 2 mg/mL solution of BSA in PBS. The control elements were used to remove the effect of refractive index variation with temperature, intensity changes in the lamp illumination and non-specific protein binding. Nonspecific binding sites on the array surface were blocked with human serum albumin (5 μM, 300 seconds) before assay operation.

Standard solutions of IgG (2.5 nM, 25 nM, 50 nM, 100 nM, 200 nM) were injected over the assay surface for a fixed period of 100 seconds to monitor association kinetics to PAG and washed in PBS for 200s to monitor dissociation kinetics. Association and dissociation timescales were chosen to provide typical errors in estimated association and dissociation rate constants of <15% and <10% respectively (368), whilst allowing rapid sampling of 5 minutes per sample. The resulting immune-kinetic assay traces were fitted simultaneously for all concentrations to the Langmuir model, discussed in Chapter 2, to produce a global fit and derive the kinetic constants of the interaction.

PAG-IgG reaction mixtures were prepared with PAG mole fractions of: 0.09, 0.14, 0.2, 0.25, 0.3, 0.35, 0.43 and 0.5. The starting IgG concentration was 100 nM for each mole fraction. The IgG concentration in a 500 μL sample of each reaction mixture was measured at 50±30s after mixing (variation due to speed of operator) and every 5 minutes thereafter for 1 hour.

4.4 IgG Assay Calibration and PAG-IgG Binding Kinetics at the LiScAR Sensor Surface

The fitted Langmuir model, for the PAG-IgG binding interaction at the interface of the sensor surface, gave estimates of k_a and k_d of (3.26 ±0.03) $\times 10^5$ M^{-1} s^{-1} and (3.01 ±0.37) $\times 10^{-4}$ s^{-1} respectively, Figure 4.6a. The estimated kinetic constants compare favourably with literature values (494).

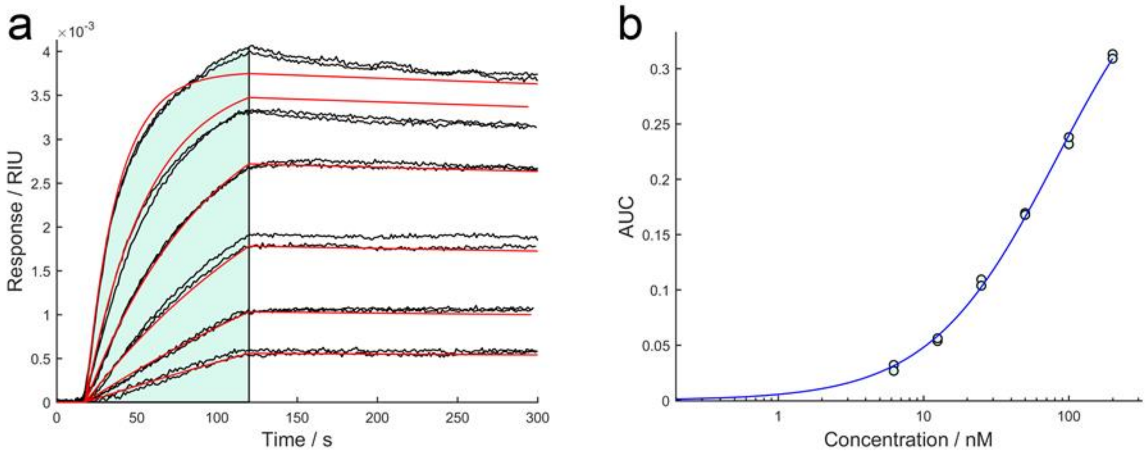


Figure 4.6 IgG capture by PAG immobilised on the LiScAR sensor surface. PAG was immobilised on the amino substrate surface using the assay method described in 2.8.1. Signals are the averaged response of n=20 assay spots. a) Duplicate kinetic responses (grey) of the sensor surface to varying concentrations of IgG (6.25, 12.5, 35, 50, 100,200 nM). Higher concentrations induce greater response and data deviated significantly from the globally fitted 1:1 Langmuir model (red). AUC of the kinetic response curve was recorded over 100 seconds of IgG association (green). b) Standard curve of n=2 IgG concentration determinations against AUC. The data fit to a logistic model (blue) with R^2=0.997.

The IgG immune-kinetic assay was calibrated using the AUC of the association step, plotted against IgG concentration and fitted to a logistic standard curve as described in Chapter 2, Figure 4.6b. The IgG assay Limit of Detection (derived here as 3σ of the noise distribution of the PBS blank) was 4.7×10-5 RIU corresponding to 0.4 nM. The Limit of Quantitation, (derived here as 10σ of the noise distribution of the PBS blank) was 1.7 nM. The typical assay error, estimated from 9 sample repeats at 100 nM, was 5 nM (2σ). The duty cycle of the IgG assay was 5 minutes, allowing for a surface regeneration step after each sample measurement which removes previously captured IgG in preparation for the next sample.

4.5 PAG-IgG Binding Kinetics in Solution

Binary solutions of IgG/PAG were prepared with different PAG mole fractions (xPAG) in the range 0 – 0.5 (corresponding to starting ratios in the range 1:0 to 1:1 IgG:PAG). The solutions were sampled for unbound IgG using the rapid biophotonic assay over a period of 1 hour, Figure 4.7. From an initial IgG concentration of 100 nM, the antibody concentration is seen to fall rapidly within the first 50 seconds followed by further, slower decrease over the proceeding hour.

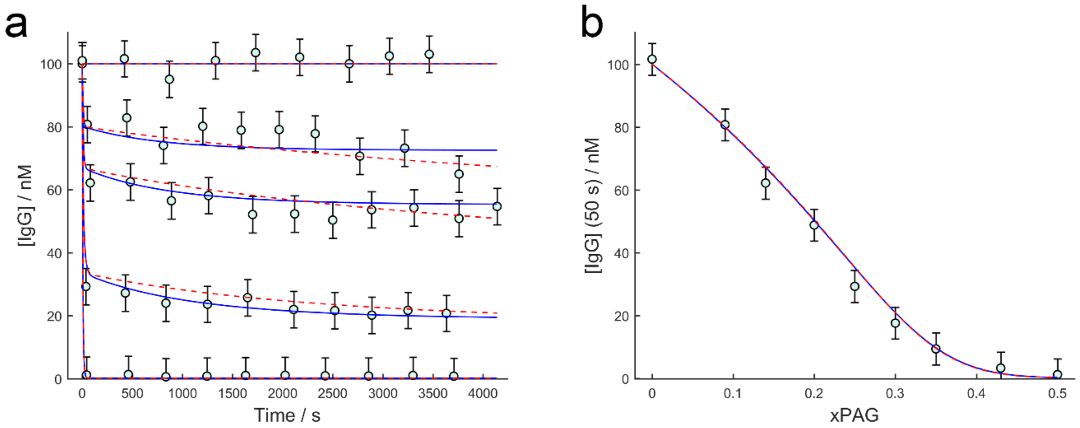


Figure 4.7 Time course analysis of IgG binding to solution phase PAG. a) Unbound IgG concentration is shown for varying ratios of antibody and PAG: control, 10:1, 6:1, 3:1, 1:1. The data were fitted to a kinetically distinct 3-binding site model (solid blue) and an 8-binding site model (dashed red). b) Titration of the free IgG against the mole fraction of PAG, sampled at the end of the first kinetic process 50±30 seconds after mixing. All data points represent single measurements with error bars of 5nM, the typical assay error (±2σ). A 3-site model (solid blue) and an 8-site model (dashed red) are overlaid for comparison.

The decrease in available IgG after the first 50 seconds is shown in Figure 4.7b and has a zero asymptote (LoQ 1.7 nM) at mole fractions of PAG in the range 0.35 - 0.43, corresponding to 1.6 ± 0.3 IgG molecules bound to PAG. The effective activity of PAG during the initial 50 seconds of association is the ratio of bound IgG to the genetic sequence maximum of 8 sites, $\gamma_{PAG} = \frac{1.6 \pm 0.3}{8} = 0.20 \pm 0.04$. The apparent cooperative binding effect observed in Figure 4.7b has been previously reported for protein A (495) and may be explained by fewer antibodies bridging two PAG proteins (by binding via two Fc epitopes) at higher mole fractions of antibody (454), although this is not modelled.

The number of bound IgG molecules per individual PAG molecule after 1 hour is shown in Figure 4.8. The maximum IgG : PAG binding ratio was 2.8 ± 0.5, markedly similar to the published ratio of 2.5 - 3.1 for protein A (496) but significantly below the 8 binding sites suggested by the sequence: corresponding to an activity, $\gamma_{PAG} = \frac{2.8 \pm 0.5}{8} = 0.35 \pm 0.06$.

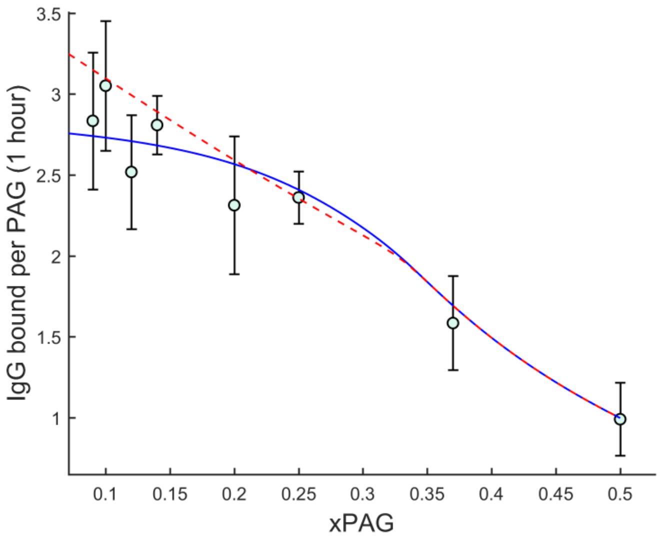


Figure 4.8 Effective stoichiometry of antibody (IgG) bound per PAG at varying mole fractions of PAG. Data were derived from the mean of n=3 determinations at 55 minutes ±5 minutes after mixing with PAG. Error bars are ±2σ. A 3-binding site model (solid blue) and an 8-binding site model (dashed red) are overlaid.

The steric blocking of neighbouring IgG binding sites on the genetically modified PAG appeared to be no better than for native protein A. Despite reports that the sum of weaker contacts in a flexible context around a central scaffold can be more efficient than precise design and strong individual interactions (497), steric hindrance between binding domains must be considered in ligand design. Future synthetic Fc binding proteins should include contacts around platforms with considered geometries able to better accommodate IgG in all binding sites.

4.6 PAG-IgG System Modelling

An ordinary differential equation (ODE) model is presented here for the sequential filling of 8 binding sites (the number suggested by the genetics literature) on the PAG molecule assuming sequential non-cooperative binding. The model is described by molecular equations 1a-8a and mathematically by differential equations 1b-8b with k_a and k_d representing the forward and reverse reaction rate constants respectively. All effective activity coefficients were assumed to be 1.

$$IgG + PAG \underset{k_{d1}}{\overset{k_{a1}}{\rightleftarrows}} IgG_1PAG \tag{1a}$$

$$\frac{d[IgG]}{dt} = k_{d1}[IgG_1PAG] - k_{a1}[IgG][PAG] \tag{1b}$$

$$IgG + IgG_1PAG \underset{k_{d2}}{\overset{k_{a2}}{\rightleftarrows}} IgG_2PAG \tag{2a}$$

$$\frac{d[IgG]}{dt} = k_{d2}[IgG_2PAG] - k_{a2}[IgG][IgG_1PAG] \tag{2b}$$

$$IgG + IgG_2PAG \underset{k_{d3}}{\overset{k_{a3}}{\rightleftarrows}} IgG_3PAG \tag{3a}$$

$$\frac{d[IgG]}{dt} = k_{d3}[IgG_3PAG] - k_{a3}[IgG][IgG_2PAG] \tag{3b}$$

$$IgG + IgG_3PAG \underset{k_{d4}}{\overset{k_{a4}}{\rightleftarrows}} IgG_4PAG \tag{4a}$$

$$\frac{d[IgG]}{dt} = k_{d4}[IgG_4PAG] - k_{a4}[IgG][IgG_3PAG] \tag{4b}$$

$$IgG + IgG_4PAG \underset{k_{d5}}{\overset{k_{a5}}{\rightleftarrows}} IgG_5PAG \tag{5a}$$

$$\frac{d[IgG]}{dt} = k_{d5}[IgG_5PAG] - k_{a5}[IgG][IgG_4PAG] \tag{5b}$$

$$IgG + IgG_5PAG \underset{k_{d6}}{\overset{k_{a6}}{\rightleftarrows}} IgG_6PAG \tag{6a}$$

$$\frac{d[IgG]}{dt} = k_{d6}[IgG_6PAG] - k_{a6}[IgG][IgG_5PAG] \tag{6b}$$

$$IgG + IgG_6PAG \underset{k_{d7}}{\overset{k_{a7}}{\rightleftarrows}} IgG_7PAG \tag{7a}$$

$$\frac{d[IgG]}{dt} = k_{d7}[IgG_7PAG] - k_{a7}[IgG][IgG_6PAG] \tag{7b}$$

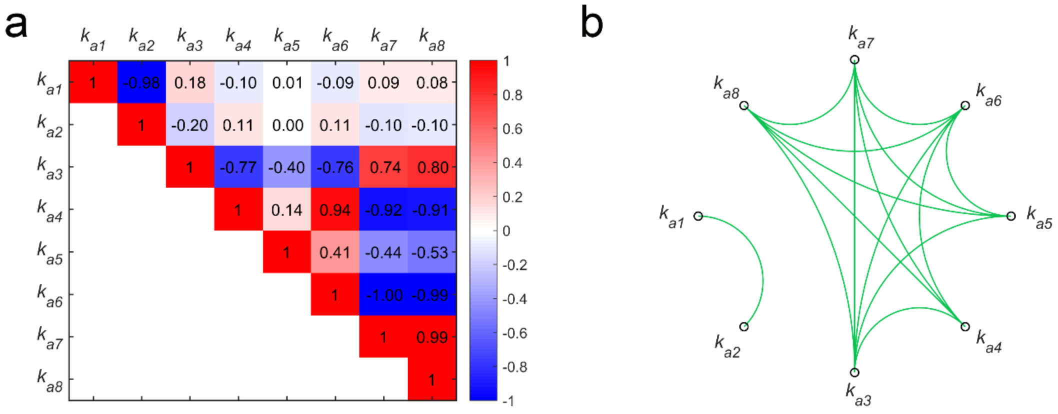
$$IgG + IgG_7PAG \underset{k_{d8}}{\overset{k_{a8}}{\rightleftarrows}} IgG_8PAG \tag{8a}$$

$$\frac{d[IgG]}{dt} = k_{d8}[IgG_8PAG] - k_{a8}[IgG][IgG_7PAG] \tag{8b}$$

The ODE solutions were propagated numerically in MATLAB with the function ‘ode15s’ and fitted simultaneously to all data shown in Figure 4.7 with equal weighting, using a non-linear least squares method (498, 499). The initial guess values for the association and dissociation rate constants of the model were taken from the surface data shown in Figure 4.6. The model fit did not converge to a solution when all 16 rate constants were allowed to float, owing to the high correlations between the parameters.

4.6.1 Parameter Reduction

The first parameter reduction technique was to fix the PAG-IgG dissociation rate at the rate observed on the surface. The dissociation rate reflects how well bound the IgG is to the PAG and is least likely to vary between settings. There was little opportunity for surface avidity for the Fc region binding and constant half-life is a reasonable assumption. The dissociation rates for all PAG-IgG$_n$ complexes were set at the dissociation rate observed on the sensor surface, 3.01 × 10^{-4} s^{-1}, and the association rates of all reactions were fitted to the data. The 8 association rate constants were determined with poor confidence and remained highly correlated. The correlation matrix and a node map of parameter correlation are shown in Figure 4.9.



	k_{a1}	k_{a2}	k_{a3}	k_{a4}	k_{a5}	k_{a6}	k_{a7}	k_{a8}
k_{a1}	1	-0.98	0.18	-0.10	0.01	-0.09	0.09	0.08
k_{a2}		1	-0.20	0.11	0.00	0.11	-0.10	-0.10
k_{a3}			1	-0.77	-0.40	-0.76	0.74	0.80
k_{a4}				1	0.14	0.94	-0.92	-0.91
k_{a5}					1	0.41	-0.44	-0.53
k_{a6}						1	-1.00	-0.99
k_{a7}							1	0.99
k_{a8}								1

Figure 4.9 Parameter correlation for the 8-binding site, 8-parameter fit of the association rate constant parameters. The initial guess values for the association rate constants of the fitting routine were taken from the Langmuir model fit to the LiScAR surface data. a) Heat map of the correlation matrix. b) Node map of absolute pairwise parameter correlation, showing two distinct groups of parameters. Every model parameter is represented by a node and is linked to another parameter node if their absolute pairwise correlation is greater than 0.2. The figure highlights the rationale for grouping k_{a1-2} into k-fast and k_{a3-8} into k-slow.

The heat map of the correlation matrix shows a striking division into two parameter groups and the node map shows the connectivity of the rate constants. The parameters in the model may be grouped based on the correlation connectivity into a single fast association reaction (combining k_{a1} - k_{a2}) and a single slow-kinetic association (combining k_{a3}-k_{a8}), which allowed the fit to converge with good root-mean-square-error (RMSE), reduced parameter correlation and improved parameter confidence.

The 8-binding site ODE model may also be reduced to describe the sequential filling of 3 binding sites only: the number determined experimentally by titration, Figure 4.8, using the same parameter grouping as the 8-site model (*k*-fast = k_{a1-2} and *k*-slow = k_{a3}). The results of fitting the reduced models are shown in Table 4.1.

Table 4.1 Results of fitting 8 and 3 binding site models to IgG-PAG binding time course data.

Model	Fit Property	Biosensor Interface Value	Least squares local minimum estimate	Pattern search global minimum estimate
8 binding site model	*k*-fast / M^{-1} s^{-1}	3.26×10^5	$1.90 \pm 0.09 \times 10^6$	1.89×10^6
	k-slow / M^{-1} s^{-1}	3.26×10^5	$6.10 \pm 0.01 \times 10^3$	6.10×10^3
	k_d / s^{-1}	3.01×10^{-4}	-	-
	RMSE / nM	27.0	3.46	3.46
3 binding site model	*k*-fast / M^{-1} s^{-1}	3.26×10^5	$1.86 \pm 0.08 \times 10^6$	1.85×10^6
	k-slow / M^{-1} s^{-1}	3.26×10^5	$1.24 \pm 0.05 \times 10^4$	1.24×10^4
	k_d / s^{-1}	3.01×10^{-4}	-	-
	RMSE / nM	12.6	3.31	3.31

The RMSE for the 8 site and 3 site models were not significantly different: 3.25 nM and 3.31 nM respectively (typical assay error was 5 nM). Considering both goodness-of-fit and titration data, the reaction mechanism in solution is best described by 3 antibody binding steps, rather than 8 as predicted by the genetic evidence. The correlation coefficients between the *k*-fast and *k*-slow parameters for the 8 and 3 binding site models are -0.03 and -0.05 respectively. The low value indicates that the *k*-fast and *k*-slow parameter groups are kinetically distinct in the data.

The *k*-fast and *k*-slow rate constants reduction demonstrates a persistent challenge for systems biology modelling. Kinetically distinct processes in a deterministic model must have significantly different rates to be determined experimentally, highlighting that kinetic data need to be recorded at a sufficiently rapid sampling rate. Similar fast or slow kinetic processes may never be distinguishable and may only be studied in isolation. Furthermore, even with high quality data, there must be a global minimum in the fitted model if the kinetic parameters are to decouple sufficiently to allow the model to produce precise parameter estimates. The *k*-fast-*k*-slow parameter space in the current model can be explored in a box search and plotted in 3D with RMSE as the cost function of the fit, Figure 4.10.

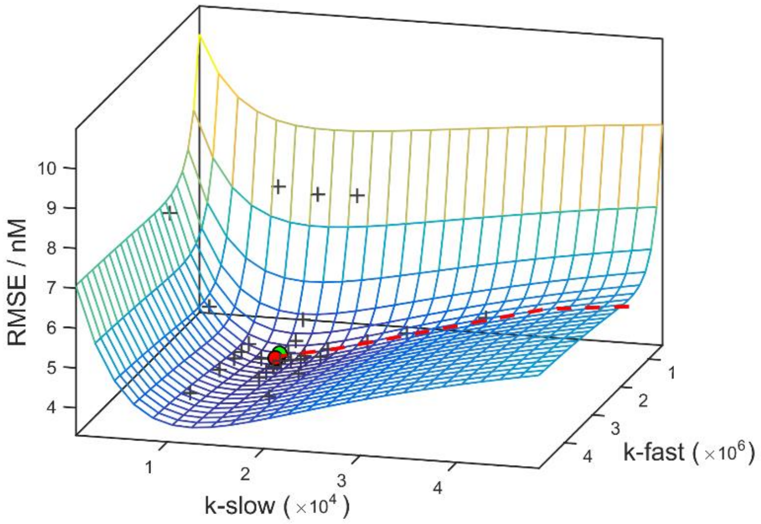


Figure 4.10 Box-search parameter space of root mean square error at various combinations of k-fast and k-slow for the 3-binding site model fit. The global minimum estimated by the pattern search global solver (green sphere) is within 95% confidence limits of the least squares fit result (red sphere). The least squares fit search path (red dashed line) and pattern search model evaluations (black crosses) originate from k-fast and k-slow initial guesses set at the surface studies value of 3.26×10^5 *(outside plot range).*

The parameter space exploration allowed the depth of the global minimum from the starting point to be determined: 27.0 – 3.46 nM (RMSE) for the 8-site model and 12.6 – 3.31 nM (RMSE) for the 3-site model. Critically, two fitting routines, least squares and a pattern search method, find the same minimum and show good starting-point invariance, Table 4.1. Starting point invariance is a good test of model accuracy but is rarely checked so easily. There can be reasonable certainty in the purity of the two grouped rate constants, *k*-fast and *k*-slow, and that their resolution into kinetically distinct process is limited by the lack of data in the first 50 seconds of the titration time courses. Hence with some confidence, the PAG-IgG interaction is accurately described by two processes: an initial rapid binding of two IgG ligands followed more slowly by at most one further IgG.

4.7 Discussion

There are two observations in the results with consequences for spin-column separation science: 1) the surface excess of binding sites for IgG on the PAG surface may be significantly lower than expected and 2) loading times for the column may be as short as 50 seconds, rather than 10 minutes as indicated in some protocols. Both binding site models have the same fitted *k*-fast rate constant and predict that after 50 seconds, unbound IgG concentration is reduced by 50% and 99% at PAG mole fractions of 0.20 and 0.43 respectively, Figure 4.7. The antibody titration shows that $k_{a1\text{-}2}$ dominate the reaction in early time and complete antibody binding may be achieved rapidly with a sufficient PAG mole fraction. Over-loading of the column is probably the major concern if reliance on 6-fold binding site excess (as sold) or 8-fold (as indicated by the genetic sequence) were critical in the adoption or modification of a protocol.

The association rate of the IgG-PAG reaction differs significantly between solution and surface tethered PAG. The fitted 8-site and 3-site models have mean association rate constants that are 47 ± 8 % and 282 ± 20 % greater than k_a derived from biosensor data of the IgG association to tethered PAG at a plasmon surface (3.26 ± 0.03) $\times 10^5$ M^{-1} s^{-1}. The difference between the interface and solution-phase reaction rates may be attributed to several factors including differences in diffusion coefficients (limiting rate constants), reduced effective activity at the interface and limited transport across any interface that forms between the tethered proteins and the bulk solution. However, there is also a clear steric effect with the acceptance angle of the surface collision (Figure 4.5) likely limited, which will limit PAG binding activity when the reaction is confined to a surface. The consequences of steric hinderance include reduced surface packing density (454) as seen in Chapter 3, enhanced cooperative binding (495) and potentially different surface complexes due to surface avidity (500). These latter effects are apparent as departures from the Langmuir model that are visible at high IgG concentrations in Figure 4.6. All these effects can be grouped into an 'effective activity' coefficient for a given experimental setting. Only if the effective activity of the experimental setting is known can a "pure" rate constant be determined which is transferable to other system settings.

To derive the activity of a new experimental setting, the concept of a standard setting must first be considered for comparison. For protein-protein interactions the standard setting must capture the fundamental kinetics and thermodynamics of the interaction - minimising the effects of concentration perturbations such as hydrodynamic effects in the crowded environments of cell cytoplasm (501) which can be explored later. The rate constants determined in the standard setting provide the best understanding of association, dissociation and affinity of the ligand-protein interaction and define the standard number of binding sites. PAG-IgG does not have 8 simultaneously functional sites as predicted by the genetic sequence. In this study, the number of kinetically active sites, $\gamma_{PAG} = \frac{2.8 \pm 0.5}{8} = 0.35 \pm 0.06$, defines an effective activity for the PAG in PBS.

A model which fits well to *in* vitro data may not fully account for the added complexity of *in vivo* conditions (482). In the cytoplasm, as much as 40% of the total volume may be occupied by proteins (502), and crowding effects will generate several perturbations to protein-protein interactions and ligand binding (372). Reduced diffusion rate also changes the collision number within the environment. Interactions may be sterically favoured or hindered and concentrations of reactants may be enhanced or depleted by association with different components. The description of a second order reaction with the activity, γ, captures the effective concentration variations between the simple system settings observed but may not account for the addition of new pathways. Model transfer between system settings requires the contributing factors of the new setting to be understood or measured empirically. Models of the immune system may require the flexibility to be adapted for healthy and sick patients, where the patient's response changes the viscosity and protein composition of the blood. Amplification of transferability uncertainties is likely to occur with larger reaction networks such as that of the Complement cascade and if not accounted for, could lead to a loss of any model predictability.

4.8 Conclusions

The issues of model transferability presented here, highlight the need for a model to be validated in every new system setting for which predictions are made. Model transfer to new system settings is possible using activity coefficients as presented here but such an approach may be challenging for large models such as those of Complement cascade activation. For example, research has shown that the assembly rate of the surface-bound AP C3 convertase is 80-fold greater in the presence of physiological Mg^{2+} (0.5 mM) than in 2 mM EDTA (a magnesium ion chelating agent), whilst in the fluid-phase the same reaction is estimated to differ less than three-fold under the same conditions (257). Ultimately, fitting a model to data in the most relevant system setting avoids the challenges of transferability. Complement activation studies are typically performed in serum and phenotype testing would likely occur *in vitro* for control over the conditions of activation. It is therefore logical that a Complement activation model for predicting patient phenotypes should also be developed using pooled human serum *in vitro*. Existing kinetic parameters of Complement protein interactions should be taken from experiments performed in serum to develop a mechanistic model of Complement activation using the systems biology techniques presented here.

5 Kinetic Modelling of *in vitro* Complement Cascade Activation

5.1 Introduction

Chapter 4 showed that there are three major scientific applications of kinetic models of biological systems. Firstly, model construction enables the system response to be simulated by knowing the starting conditions and objects involved – challenging with diagrammatic models alone. Secondly, comparison of a kinetic model to data can provide evidence for system structure and model fitting can provide kinetic constants. Thirdly, the model predictions of system behaviour under different conditions can be used to generate hypotheses and guide new experiments to confirm any simulated effects. Ultimately, the goal of systems modelling is to achieve a refined model which accurately reproduces the relationship between any given starting conditions and the system response recorded in experimental data (459). It follows that a predictive model of Complement activation could be used to predict the Complement activation dynamics of patient serum and guide a hypothesis-driven approach to the selection of potential risk biomarkers of nosocomial infection.

The Complement cascade is characterised by a series of rapid, consecutive kinetic processes with catalytic activity both positively and negatively regulated. C3 and C5 convertases are negatively regulated by a number of proteins in solution (FH, FI, C1-INH, C4BP and sCR1) and on the surface of host cells (157, 158) as discussed in Chapter 1. The lifetime and catalytic activity of the C3 and C5 convertases controls the extent of Complement activation which can become dysregulated in conditions such as sepsis (503). The structures of most protein complexes and their interactions in the Complement system are well-known, allowing a mechanistic representation for Complement activation to be constructed following the recommended nomenclature for each protein (183), Figure 5.1.

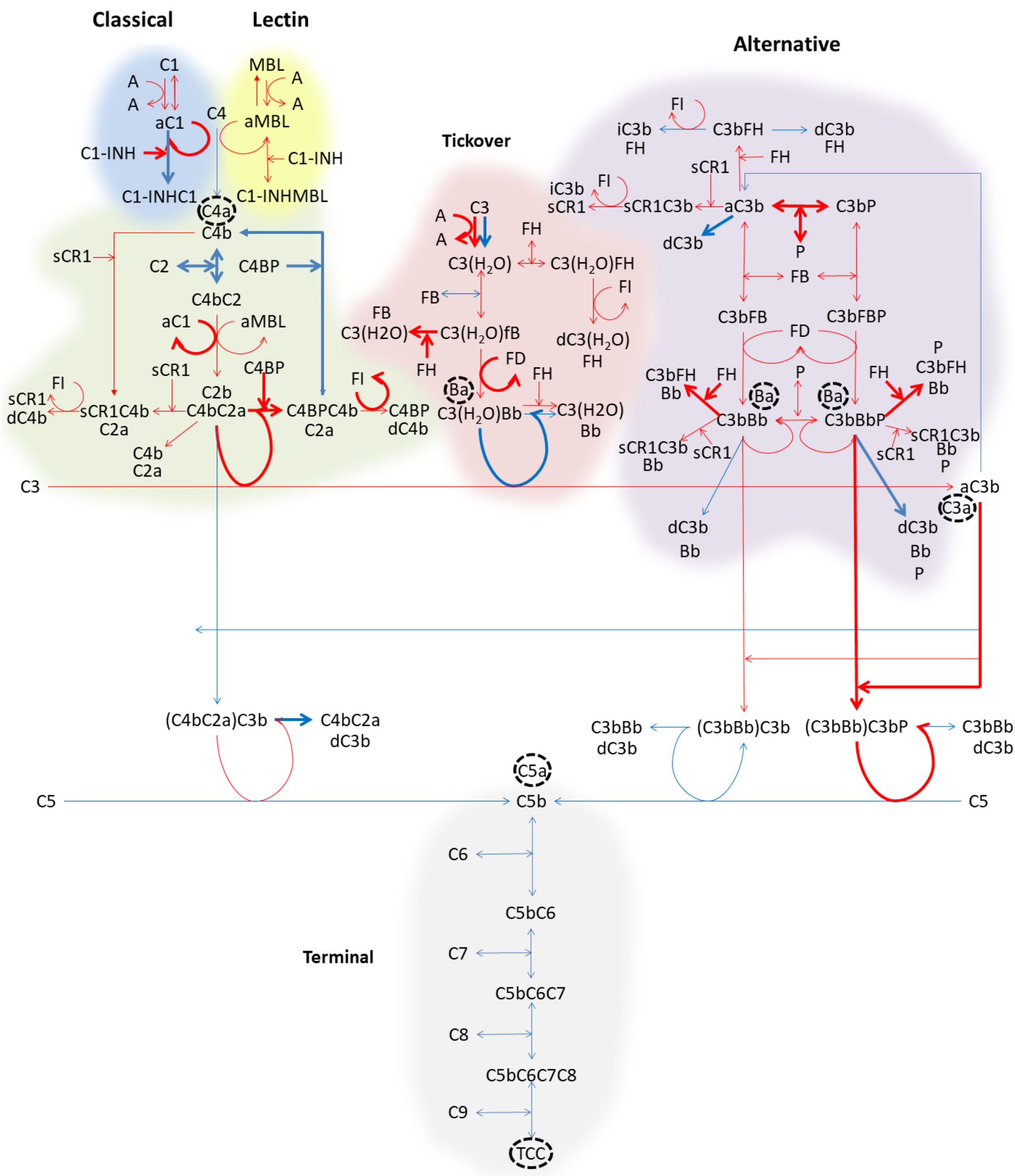


Figure 5.1 Schematic of the mechanisms of Complement activation and inhibition. All complement pathways are highlighted: CP (blue), LP (yellow), CP/LP (green), tick-over (red), AP (purple) and TP (grey). Reversible reactions are indicated by double headed arrows. Reactions with known rate constants are shown in blue, whilst reactions with unknown rate constants are shown in red (all rate constants are tabulated in Appendix 3 and Appendix 5). The protein species for which data were collected for model fitting are circled with dashed lines. The most sensitive reactions (accounting for 95% of model sensitivity) from the sensitivity analysis in section 5.3.4 are emboldened to show that system dynamics is controlled by reactions located throughout the system.

Despite extensive study (167), some reactions of the Complement system are incompletely understood or remain contentious. For example, there is evidence to suggest that FB binds to C3b via either Ba or Bb portions (190) but the interaction is rarely mentioned in studies of the AP. A more recent study showed that the Ba portion of FB binds in a Mg^{2+} independent process followed by the slower binding of the Bb portion (504). Furthermore, it is well reported that

Complement activation occurs *ex vivo* and complicates studies of Complement in serum and plasma samples, but the effect is typically attributed to tick-over in the AP, despite reports of spontaneous formation of C3 fragments in human plasma showing C1q dependence (505). Experimental evidence of spontaneous C1 activation is conflicting: one study estimated a first order rate constant of 0.173 min^{-1} in purified solutions of C1 components which was suppressed in human serum (506), whilst other studies have shown that IgG oligomers (507) or immune complexes are required to activate C1 (508). A further study concluded that C1 activation could be a result of unidentified proteolytic enzymes and can be both an intra- and intermolecular process impacted by experimental conditions (179) . LP tick-over has not been reported. Uncertainty in the mechanism of certain intermolecular binding reactions is a challenge when constructing of the Complement reaction network and previous studies have studied Complement activation dynamics using Complement models which focus on the well-known reactions. A list of existing Complement models and their properties is shown in Table 5.1.

Table 5.1 Existing Complement models published in the literature.

Model	Year	Pathways modelled	Optimisation method	Activation stimulus
Hirayama et al. (509)	1996	CP, TP	Minimal protein consumption	Spontaneous C1-C4 complex
Korotaevskiy et al. (185)	2009	CP, AP, TP	Minimal protein consumption; *in vitro* bactericidal data	Population of cells
Liu et al. (510)	2011	CP, LP, TP, CRP crosstalk	Fitting to CRP, C4, C3 and C4BP time series data; *in vitro* bactericidal data	CRP-C1 complex and PAMPs
Zewde et al. (511)	2016	AP, TP	None	Surface for C3b deposition
Sagar et al. (271)	2017	LP, AP	Fitting to C3a and C5a time series data	Zymosan
Zewde et al. (512)	2018	CP, AP, TP	None	Spontaneous C1 activation and surface for C3b deposition
Present study	2019	CP, LP, AP, TP	Fitting to C3a, C5a, C4a, Ba and TCC time series data	Zymosan

The present study requires a model that can predict the effects of starting protein conditions (such as depletion) on activation dynamics on the complete Complement system to deductively identify potential phenotypes of an immune state which is more vulnerable to post-operative nosocomial infection. The dogma of systems biology is that the qualitative mechanisms of Complement activation reported in the literature and represented in Figure 5.1, can only be investigated quantitatively through a mechanistic model of the system as a whole rather than isolated or simplified components (513). A thorough Complement model should therefore include each of the bimolecular reactions and consider the catalytic activity of each resulting enzyme and its inhibitors.

All previous models of the Complement system have involved the use of ODEs but none have included mechanistic reactions for all pathways of the Complement system (none consider the LP for example), instead concentrating either on a subset of pathways or reducing model complexity by combining reactions into non-mechanistic equations. As discussed in Chapter 4, model simplification is driven by the major challenges facing an ODE model of a complex system: 1) insufficient quality data; 2) large numbers of kinetic parameters that are highly correlated and need to be optimised; 3) locating a global minimum in the parameter space and 4) data set invariance. Inadequate biological data has been cited as the major limiting factor in the optimisation of previous Complement models (271). Bimolecular rate constants of the Michaelis-Menten reaction scheme for enzymes (514) may not be experimentally accessible for all short-lived intermediates and hence are often ignored in favour of reducing the parameters of the model. Non-mechanistic models are inherently unable to predict system behaviour beyond the context of the optimised model, such as the effect of additional inhibitory processes targeting an enzyme-substrate complex or different activation stimuli. As such, all previous models are inappropriate for use in the present study which aims to predict the effects of starting protein conditions on activation dynamics on the complete Complement system.

A new model of Complement activation must therefore be developed and optimised specifically to predict system dynamics in the system setting studied here: *in vitro* activation of human serum. A previous model of Complement activation is optimised for *in vitro* zymosan activation of human serum (271) and used C3a and C5a time series data from the Shaw lab to fit model outputs. The system setting of *in vitro* activation by a single stimulus is better suited to the present study than the system settings of other previous models which either consider activation via a growing population of cells, spontaneous activation or are simply not optimised for a specific stimulus at all (see Table 5.1). As described in Chapter 4, ODEs are widely used in biological modelling to capture the quantitative and dynamic nature of complex systems over time (462) and using ODEs in the present study allows the published rate constants compiled from previous models to be reused, although estimating the missing parameters remains challenging.

Notwithstanding the modelling challenges, quantitative and mechanistic modelling of Complement activation should reveal the dynamics of the cascade and enable predictions to be made of therapeutic relevance, such as the effect of protein concentrations (deficiency) or system disruption by a drug on Complement activation dynamics. Of the 109 reactions shown in Figure 5.1, the kinetic constants of 34 have been reported previously (Appendix 3) but there is no prior information on the kinetics of the remaining reactions. As shown in Chapter 4, for a model to be quantitatively accurate, the model parameters should be known or estimated from fitting the model to experimental data. Model fitting is complicated by parameter correlation and the presence of multiple local minima - challenges which were overcome in Chapter 4 by grouping similar parameters using prior knowledge of their values and fitting using global optimisation algorithms. A starting value for parameter estimation within or close to the global minimum is beneficial.

Despite a large range of recorded enzyme kinetics, a recent review of the enzyme kinetic parameters: K_m, k_{cat} and k_{cat}/K_m for ~2500 enzymes reveals that many are "moderately efficient" (515). For example, the relatively small range of k_{cat} values within 1 – 100 s^{-1} accounts for 60% of all recorded enzymes. Median kinetic constants derived from the distribution of all known enzymes can be used as initial guess values for unknown kinetic parameters and a further reduction in the number of floated parameters can be achieved by assuming that all constitutive protein concentrations in pooled serum take the median values of their respective reference ranges. These two prior assumptions provide a biologically appropriate starting point for optimisation of a complete and mechanistic ODE model of Complement activation: referred to here as the C-ODE model.

5.2 Aims and Objectives

The aim of this chapter is to describe the development of a fully mechanistic Complement model, based on the accepted reaction mechanisms in Figure 5.1. The purpose of the model is predicting Complement activation dynamics from specific starting conditions such as patient serum protein concentrations and the level of activation stimulus.

The first objective was to study Complement activation markers (Ba, C4a and TCC) with ELISA, over 24 hours following the *in vitro* activation of pooled human serum with zymosan at several concentrations. The new data set was combined with data for C3a and C5a from previous activation studies using the same protocol (152) to produce a consistent data set of activation markers from all Complement pathways.

The second objective was to construct an ODE model of Complement activation considering all bimolecular reactions and enzyme catalysis and incorporating available published rate constants. The optimisation strategies discussed previously were used to fit the model to the ELISA data and estimate unknown rate constants.

Finally, parameter sensitivity and flux analyses were performed on the model to verify that the dynamics of Complement activation align with current understanding of the system. The C-ODE model predictions of new system phenotypes may then be validated in future studies.

5.3 Materials and Methods

5.3.1 *in vitro* serum activation

Pooled Normal Human AB Serum was obtained from Innovative Research (IPLA-SERAB-OTC). Phosphate buffered saline (PBS), EDTA and zymosan were obtained from Sigma Aldrich. ELISA assays were obtained from Quidel for Ba, C4a and TCC (#A035, #A033, #A020). All serum experiments were performed in a Level 2 laboratory following relevant risk assessments.

Suspensions of zymosan in PBS (1 mL) were homogenised by vortex mixing and added to pooled normal human serum (24 mL) to give final zymosan concentrations of 0, 0.01, 0.1 and 1 mg/mL. The serum was maintained at 37 °C in an incubator with orbital shaking at 190 rpm during the experiment. Serum samples of 1 mL were collected by pipette from the reaction mixtures at time points: −0.5, 0.5, 1, 2, 4, 8, 16 and 24 hours relative to the addition of zymosan. Samples were immediately mixed with aqueous EDTA (50 µL, 10 mM) to give a final EDTA concentration of 0.5mM and centrifuged for 30 minutes at 3000 *g*. The supernatant was removed, avoiding sampling fat deposits or the zymosan pellet, before freezing with liquid nitrogen and storage at -20 ^{0}C. The ELISA assay protocols were followed as detailed by the manufacturer. Before ELISA analysis of all samples, the first and final samples in the time courses were analysed at several dilution factors to ensure all samples fell within the dynamic range of each assay. Samples were diluted according to Table 5.2. ELISA calibration curves are shown in Appendix 2 with control samples exhibiting errors of <4%.

Table 5.2 Dilution factors used with samples from each zymosan time series for use with ELISA. Dilution factors include the dilution due to the addition of EDTA and zymosan solutions.

Assay	0 mg / mL zymosan samples dilution factor	0.01 mg / mL zymosan samples dilution factor	0.1 mg / mL zymosan samples dilution factor	1 mg / mL zymosan samples dilution factor
C4a	272.5	272.5	381.5	381.5
Ba	4375	6562.5	6562.5	13125
TCC	87.5	164	246	820

The ELISA data collected here were merged with the previously published data for C3a and C5a data (152) producing a dataset containing five markers of Complement activation sampled at eight time points over four different activation experiments of the same pooled serum sample. The final dataset (Figure 5.2) contains 160 data points with a mean ±2SD derived from at least two statistical repeats (C3a and C5a data from the previous work show biological repeats). The epitope decay observed previously in the C3a assay (152, 516) was corrected using the half-life equation 3.3 with a negative power: $corrected = data \times 0.5^{-\frac{t}{t_{1/2}}}$ where the half-life of the epitope decay was measured previously to be 28 hours (152).

5.3.2 C-ODE Model Construction

The C-ODE model is comprised of 85 distinct reaction mechanisms parameterised by 109 rate constants allowing for reversible reactions and 78 ODE’s describe the concentration of each species over time. All reactions and species included in the C-ODE model are presented in Appendix 3 and Appendix 4, respectively. The reaction network, shown in Figure 5.1, was developed into a fully mechanistic model divided into three types of reactions. Firstly, simple bimolecular reactions and enzyme catalysis were constructed using the familiar Michaelis-Menten reaction scheme. Secondly, a new class of activation reactions was used to represent the complex activation of C1, MBL and C3 by zymosan, for which there was no kinetic data or mechanism. Lastly, a set of sink reactions was used to represent decay of activated species within the *in vitro* reaction mixture, removing these species permanently from the system.

The Michaelis-Menten reaction type is exemplified in equation 5.1 with the formation of the tick-over C3 convertase:

$$C3(H_2O)FB + FD \underset{k_d}{\overset{k_a}{\rightleftarrows}} C3(H_2O)FBFD \xrightarrow{k_{cat}} C3(H_2O)Bb + FD \tag{5.1}$$

where k_a is the affinity constant, k_d is the dissociation constant and k_{cat} is the enzyme turnover number. Two spontaneous reactions were included in the model: C1 -> aC1 and C3 -> C3(H_2O). The rate of spontaneous C3 hydrolysis is known and was assigned a value of 1×10^{-4} min^{-1} (184, 185). The rate constant of spontaneous C1 activation was considered to be unknown in the present study due to the contradictions in the literature discussed previously (179). The rate constants of the Michaelis-Menten reactions can be converted to enzyme kinetic constants: the affinity constant, $k_d / k_a = K_D$; the Michaelis constant, $(k_{cat} + k_d) / k_a = K_m$; and a measure of catalytic efficiency, k_{cat}/K_m. The enzyme kinetic constants may then be compared to the data in ref (515).

Zymosan is reported to activate all Complement pathways but the mechanisms are complex with unknown rate constants. For example, the sequential activation processes of C1 and MBL by their associated serine proteases involves multiple proteins reacting simultaneously as described in Chapter 1. Furthermore, a recent review of the AP states that an active process rather than tick-over should be shown as the initiator of AP activation (517), with gas bubbles and biomaterial surfaces all known to increase the rate of C3(H2O) formation with an unknown mechanism (186). Considering the uncertain mechanisms of Complement activation, the model contains a minimal set of activation reactions which provide a simple means to express the complex Complement initiation processes in each pathway, Table 5.3.

Table 5.3 Activation reactions of the C-ODE model. Species 'A' represents the activating process, which in the present study is zymosan.

Reaction name	Reaction mechanism
A1	A + C1 -> A + aC1
A2	A + MBL -> A + aMBL
A3	A + C3 -> A + C3(H2O)

The zymosan in the reaction mixture had an unknown surface area but was prepared as a fine homogenous suspension which was well mixed to prevent settling during activation and is expected to behave as a solution phase reactant rather than a surface. Hence, the activation reactions were written as simple bimolecular reactions where zymosan is considered as having a constant concentration and any surface filling effects are ignored. The activation reactions enabled the familiar Michaelis Menten reaction scheme to be used for modelling the subsequent activity of the activated proteins: aC1, aMBL and $C3(H_2O)$.

A final set of reactions are sink reactions, Michaelis-Menten reactions invoked to remove species from the reaction mixture when the activation products are inert to further reactions. The sink reactions therefore reduce model complexity by simplifying products of "dead end" metabolic pathways for which there is no data for model fitting (518). The full set of sink reactions is shown in Table 5.4.

Table 5.4 Sink reactions of the C-ODE model.

Reaction name	Reaction mechanism
d1	aC1 + C1inh -> C1inhC1
d2	aMBL + C1inh -> C1inhMBL
d3	C3(H2O)FHFI -> dC3(H2O) + FH + FI
d4	C4BPC4bFI -> C4BP + dC4b + FI
d5	sCR1C4bFI -> sCR1 + dC4b + FI
d6	sCR1C3bFI -> sCR1 + iC3b + FI
d7	C3bFHFI -> iC3b + FH + FI

Sink reactions d1 and d2 permanently remove C1 and MBL from the serum respectively. C1-INH (also termed C1inh for modelling syntax) is a serpin and suicide inhibitor which irreversibly binds and inactivates C1r and C1s proteases in the C1 complex of the CP as well as the MASP-2 protease in the LP (519). C1inh therefore prevents the proteolytic cleavage of later complement components C4 and C2 by the activated forms of C1 and MBL. Reaction d3 inactivates the spontaneously formed C3(H2O) using negative regulation from

FH and FI. Reactions d4 and d5 describe C4 deactivation mediated by C4BP, soluble CR1 and FI.

Following dissociation from complexes or due to its natural half-life, C3b is converted to dC3b - considered to be inert to further reactions with FB and existing convertases. However, dC3b is not a sink species as it is further converted to iC3b by FI in the presence of a cofactor in sink reactions d6 and d7. Species for which there is no deactivated form (such as P, FB and Bb) are regenerated following dissociation from a complex, which simulates the natural equilibrium of these species in the serum.

Of the 109 rate constants in the model, 34 were assigned published literature values, Appendix 3. Unknown rate constants of Michaelis-Menten reactions were given starting values equal to median enzyme values (515) of: $k_a = 6\ \mu M^{-1}\ min^{-1}$; $k_d = 0.06\ min^{-1}$; $k_D = 10$ nM and turnover number, $k_{cat} = 840\ min^{-1}$. The activation reactions in Table 5.3 are simplifications of complex processes and were assigned initial values of $1\ \mu M\ min^{-1}$. The starting concentrations of the 78 constituent proteins were set at their median physiological values in serum (the pooled-serum approximation) and all intermediate enzyme species formed as part of the cascade were assumed to have negligible concentrations before activation and are given starting concentrations of zero (Appendix 4). The ODEs were solved using a stiff solver, ‘ode15s’ in MATLAB R2017a software (520) with relative and absolute tolerances of $1\times10^{-6}\ \mu M$ to produce a time series simulation of all 78 species in the model for a given zymosan concentration.

5.3.3 C-ODE Model Parameter Optimisation

The C-ODE model parameters were optimised by fitting the model outputs for C3a, Ba, C4a, C5a and TCC to the ELISA assay data at all zymosan concentrations: a global fit to 160 data points. The goodness-of-fit cost function was the square of the residuals (least squares) between the model and the data: ensuring outlying points had a disproportionate effect on the fit. The concentrations of each species in the dataset varied by two orders of magnitude and hence the data were normalised to values of 100 to ensure residuals for all data points were weighted equally by the cost function. Similarly, all floated parameters were normalised to values of 100 to ensure all parameter deviations were explored to equal extents by the step tolerance of the global optimisation routine.

The 56 unknown association rate constants and turnover numbers were allowed to vary in the fitting process whilst the 19 unknown dissociation rate constants, 34 literature rate constants and 78 constituent protein concentrations were fixed. The unknown dissociation rate constants were fixed at a median k_d value to remove high correlation with the k_a and k_cat parameters. The k_a and k_cat parameters of the Michaelis-Menten reactions were constrained in the fitting process by 15-fold above and below their starting values. The parameter search space encompassed 67%, 86% and 64% of the published values for K_m, $k_\mathrm{cat}/K_\mathrm{m}$ and k_cat in reference (515) respectively. The activation reactions (Table 5.3) were allowed to vary over a larger range of 1000-fold above and below their assigned initial values of 1 $\mu M^{-1}\ min^{-1}$, as no prior information was known for these abstract reactions.

The fit was performed with a gradient descent method implemented in MATLAB as ‘lsqcurvefit’ (498, 499). The fitting routine was stopped when the cost function changed by less than a tolerance of 1×10^{-3} relative to the previous value (0.1%). The converged fit at this tolerance defined a local minimum. Excursions away from the local minimum were generated by a global optimisation routine to initiate 9 additional starting parameter values sampled randomly and uniformly from the fit bounds of the parameters before the gradient descent solver was executed. The fit solutions from all starting points were ranked in order of goodness-of-fit as judged by root mean square error (RMSE) and the best solution is presented here as the C-ODE model. RMSE serves as a multivariable function to combine the magnitudes of the residuals for all sample points into a single value measure of fit accuracy and penalises larger errors with the same bias as the least squares fitting routine.

The gradient fitting routine provided the variance-covariance matrix, discussed in Chapter 4, was used to derive the correlation matrix and estimate 95% confidence limits of the fitted parameter values. The optimised C-ODE model was transferred to freely available COPASI software (458, 521) and software invariance demonstrated using a Bland-Altman comparison analysis as discussed in Chapter 3.

5.3.4 Local Sensitivity and Flux analysis

A major challenge when assessing large ODE models is to distinguish model features which are fundamental from those which are accidental or irrelevant

(522). Two key tests were used to assess whether the C-ODE model was consistent with published properties of Complement activation: 1) a local sensitivity analysis of model reactions to identify which kinetic parameters have the most effect on model output and 2) a flux analysis of model species to assess the activity of a particular pathway as a consequence of all related catalytic interactions (523).

The local sensitivity analysis of the C-ODE model parameters was performed in MATLAB in the following sequence: 1) all model parameters were set to their values at the fit solution; 2) a single model parameter was perturbed by +1%; 3) the model was executed; 4) goodness-of-fit to the data was evaluated using the RMSE cost function; and 5) steps 1 to 4 were repeated for every parameter and ranked according to magnitude of change in RMSE.

Flux represents the total turnover number for a given reaction, considering the concentration of the reactants and their respective rate constants. The C-ODE model was transferred to COPASI software (version 4.20.158) which was used to simulate the flux in events per minute through each reaction of the model using the following simulation settings: compartment volume = 1 L, duration = 1440 min, interval size = 1 min, relative tolerance = 1×10^{-6}, absolute tolerance = 1×10^{-12}, maximum internal steps = 1×10^{4}. The model is described entirely by Appendix 3, Appendix 4 and Appendix 5.

5.4 Results

The activation of pooled human serum by different concentrations of zymosan was profiled using Ba, C4a, C3a, C5a and TCC concentrations over 24 hours, Figure 5.2. The analytes Ba and C4a are markers of AP and CP/LP activation respectively, whilst C3a and C5a / TCC are measures of upper cascade and TC activity respectively.

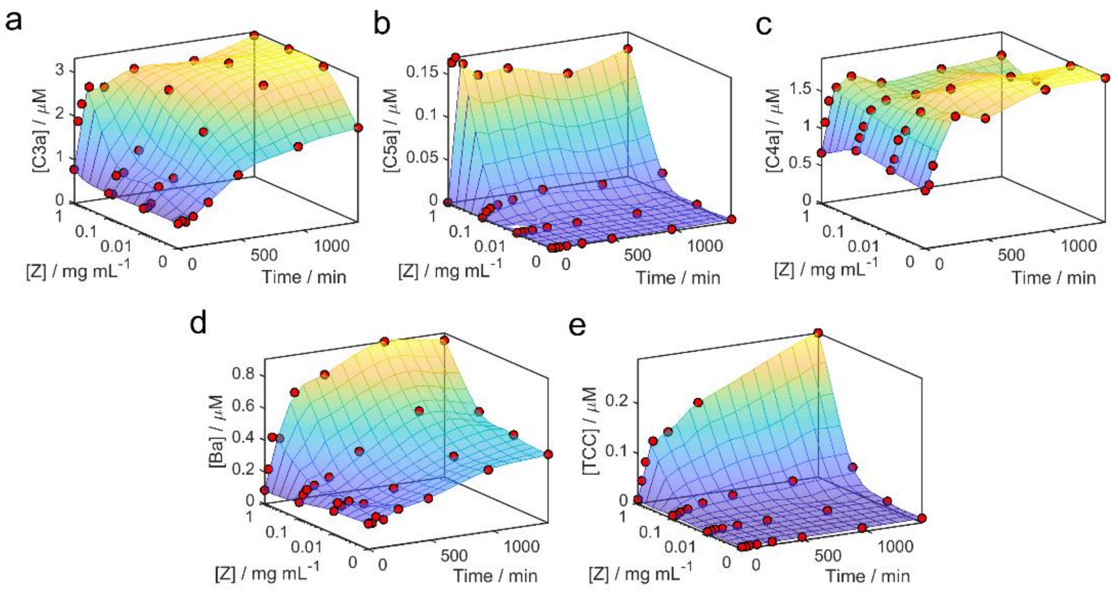


Figure 5.2 Complement activation response to varying zymosan concentrations over time. A number of analytes are shown: a) C3a, b) C5a, c) C4a, d) Ba, and e) TCC. ELISA data (red) are interpolated with a 3D mesh to aid visualisation. Data represent the mean of n=2 technical replicates.

All Complement activation markers measured in the pooled human serum increased in concentration over the 24-hour time course and the trend of a given analyte over time was similar for all zymosan concentrations, whilst the kinetics of each analyte were distinctly different. All markers of Complement activation appeared to reach a plateau after ~4 hours for all experiments, except for TCC which continued to rise gradually over 24 hours. The effect of zymosan on the CP/LP and AP activation is shown in Figure 5.3.

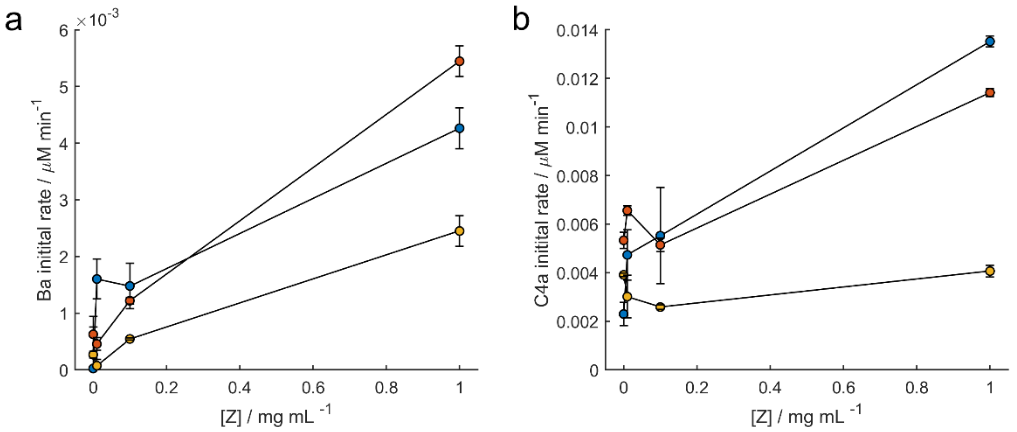


Figure 5.3 Zymosan dependence of AP and CP/LP activation in pooled serum. The gradient of analyte concentration change between 0-30 (blue), 0-60 (red) and 0-240 (yellow) minutes following zymosan activation is shown. a) rate of Ba concentration increase. b) rate of C4a concentration increase. Error bars are 2SD from n=2 technical replicates.

The rate of C4a and Ba production showed zymosan dependence when judged by the gradient of concentration increase. Furthermore, there was evidence for a Z-independent production of C4a: in the absence of zymosan, the rate of C4a production over the 4 hours of serum incubation was (3.91±0.07)×10^{-3} µM min^{-1}. This value for CP/LP activation may be compared to (2.7±0.4)×10^{-4} µM min^{-1} for Ba and (4.7±3.2)×10^{-4} µM min^{-1} for C3a in AP tick-over.

C3a can be used a measure of total upper cascade activation and the C3a : C5a ratio has been used previously to identify a flux control threshold at 0.01 mg/mL zymosan (152), above which significant activation of the terminal cascade occurs. The C3a : TCC ratio is also a measure of flux between the upper and lower cascades and includes the effect of terminal cascade inhibition, Figure 5.4.

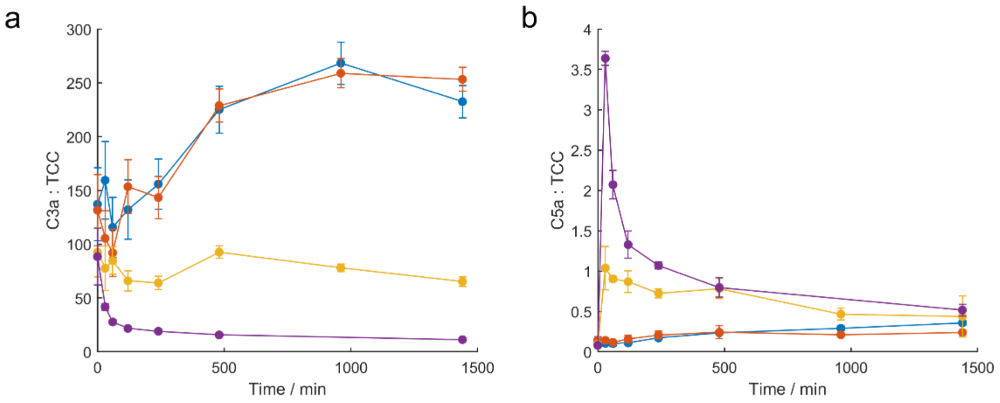


Figure 5.4 Relative Complement pathway flux over time following zymosan activation. A number of zymosan activations are shown: control (blue), 0.01 mg/mL (orange), 0.1 mg/mL (yellow) and 1 mg/mL (purple). a) upper and lower cascade (C3a : TCC) and b) early and late lower cascade (C5a : TCC). Error bars are 2SD from n=2 technical replicates.

The C3a : TCC ratio started at a value of 112±51 and evolved in time with strong dependence on zymosan concentration. During tick-over and at the lowest zymosan dose (0.01 mg/mL), the C3a : TCC ratio rose to a value of 243±29 at 24 hours. In contrast, at 0.1 mg/mL zymosan the C3a : TCC ratio remained relatively constant at 78±24 and at the highest zymosan dose the C3a : TCC ratio fell to 11±1 after 24 hours.

The C5a : TCC ratio is a measure of lower cascade flux and started at a value of 0.11±0.07, followed by a lag phase where the transient value correlated positively with zymosan dose. In contrast to the C3a : TCC ratio which was zymosan dependent after 24 hours, the C5a : TCC ratio converged at a value of 0.4±0.2 for all zymosan activations after 24 hours, attributed to the linear system of reactions between C5 and TCC. The expected ratio was 1 TCC produced per C5 cleavage and may have been influenced here by the rising TCC assay response.

Before any optimisation, the model fitted to the experimental data with a RMSE of 53% and mean absolute error (MAE) of 41%. The parameters in the model were then optimised using the global fitting strategy (Appendix 5) which reduced the RMSE to 19% and the MAE to 14%, approaching the experimental error in the ELISA repeats (21% RMSE and 12% MAE) Figure 5.5.

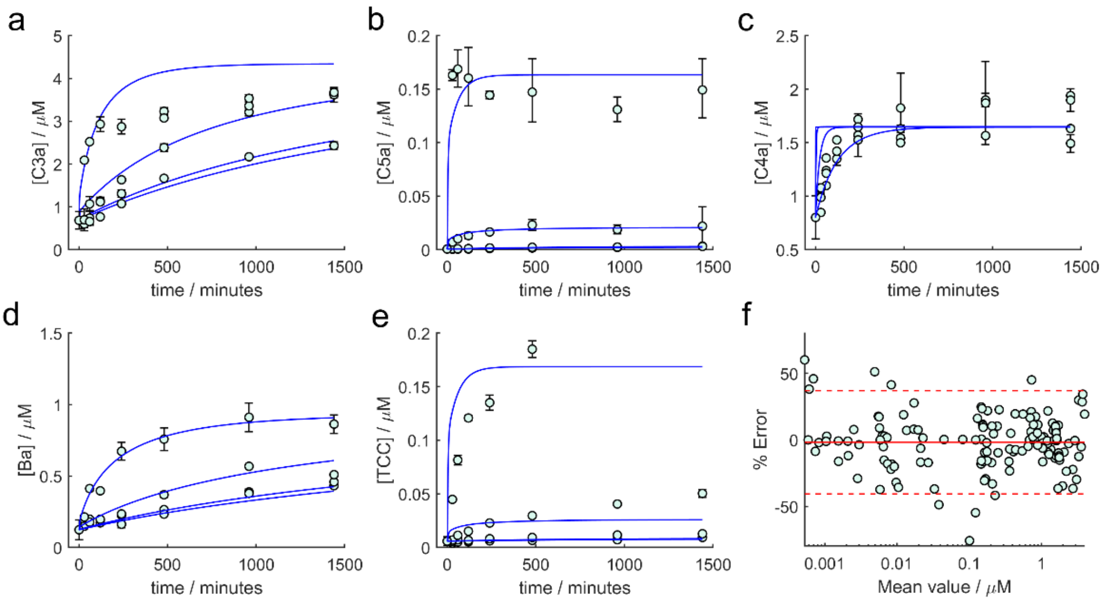


Figure 5.5 C-ODE model fit to zymosan activation of pooled human serum over time, profiled using five Complement activation markers: a) C3a, b) C5a, c) C4a, d) Ba and e) TCC. Error bars are 2SD from n=2 technical replicates. The global minimum fit to the data is shown in blue with a MAE of 14%. f) Bland-Altman analysis of the model fit residuals expressed as percentage error, with a mean of -2% (solid red) and 2SD = 37% (dashed red). The model fits the data with R^2=0.93.

There was a good general correlation between the data and the fit with $R^2 = 0.93$. The residuals of the model fit (excluding those at time = 0) may be expressed as percentage error and were distributed normally around a mean of -2% and SD = 19.4%. The Bland-Altman plot in Figure 5.5f shows there was no apparent concentration bias in the C-ODE global minimum as the residuals were distributed evenly across the complete range of analyte concentrations. The fit residuals are explored in more detail in Figure 5.6.

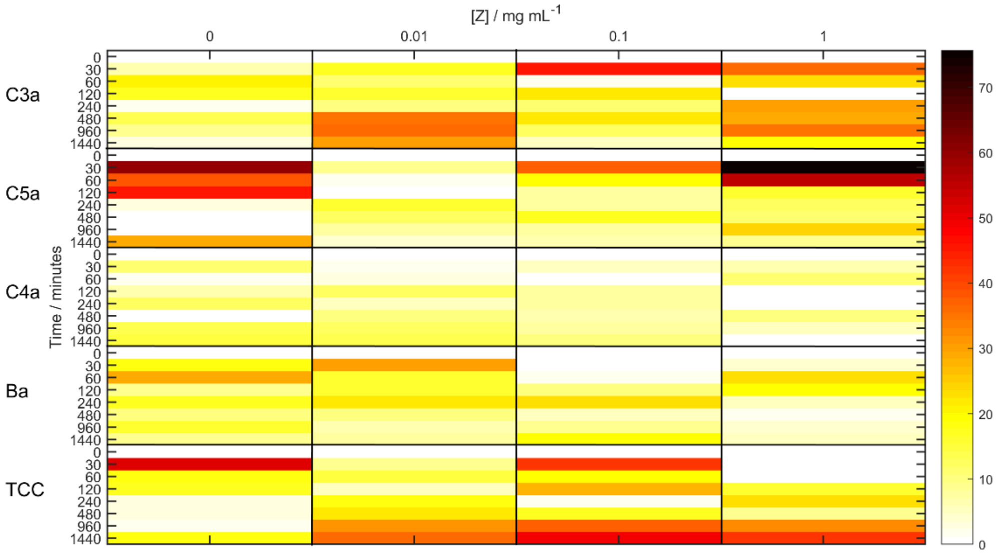


Figure 5.6 Heat map of C-ODE model fit residuals as a percentage of each data point. The figure shows where the model fits best (white and yellow) and worst (red and black) in the data set.

The fit deviated most from the data in early time, suggesting more data should be collected 0-2 hours following activation to better understand initial dynamics. The largest discrepancy between the model and data was for C5a at 30 minutes following activation, with an error of 70% and this region was explored further with the LiScAR assay for C5a in Chapter 7.

The optimised model was transferred to COPASI software and the concentration-time outputs were checked for agreement with the MATLAB output using a Bland-Altman analysis, Figure 5.7.

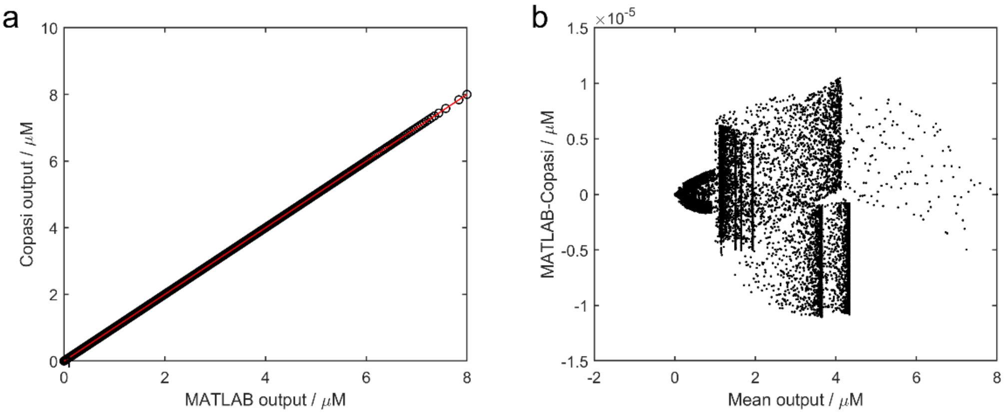


Figure 5.7 MATLAB solver output compared to COPASI for the C-ODE model fit. a) regression analysis with R^2=1, intercept=0 and slope=1. b) Bland-Altman analysis showing relative differences between outputs. The relative tolerance of both ODE solvers was 1×10^{-6} relative to the state value (0.001%).

The COPASI model output agreed with the MATLAB output within the tolerance of the ODE solvers, giving confidence to the use of COPASI for further analysis of the model.

Four basins of attraction (discussed in Chapter 4) were inferred from the fitting process by common minima in RMSE (19-20%, 26-29%, 37% and 50-53%) found by gradient descent from different starting parameter values, Figure 5.8.

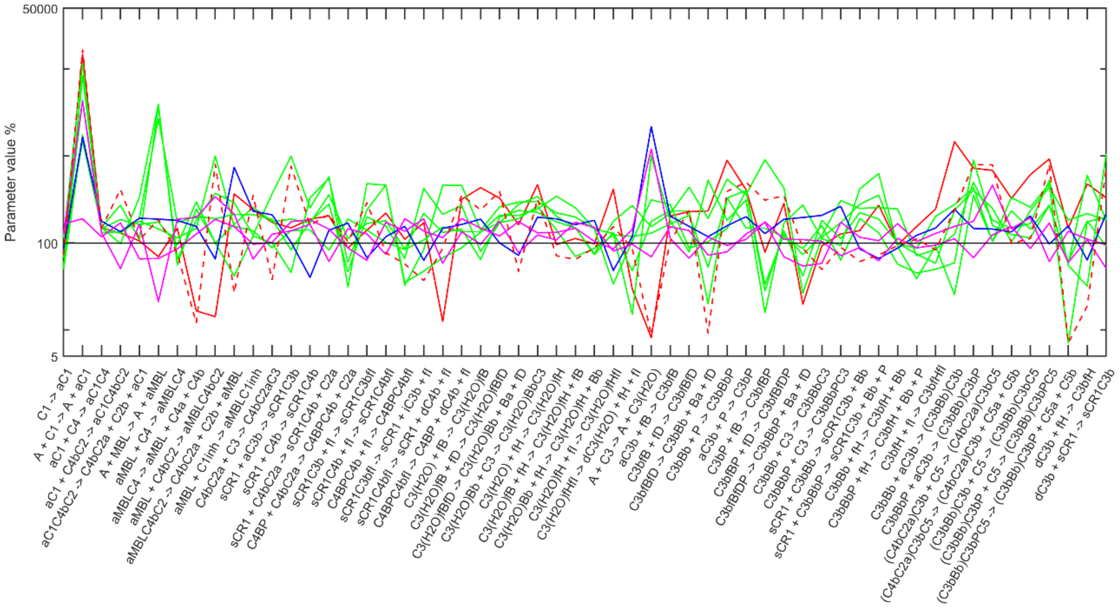


Figure 5.8 Local minima found by the global parameter optimisation routine for the C-ODE model fit. The initial guess is shown in black and the solutions are coloured by RMSE: 19-20% (red), 26-29% (green), 37% (blue) and 50-53% (purple). The global solution is also shown (dashed red).

The starting point based on median enzyme kinetics produced a solution within the second-best basin of attraction with a RMSE of 27%, highlighting the importance of global optimisation when searching the parameter space of large non-linear models for a global minimum.

The parameter estimates were constrained by bounds during the fitting process to reduce the parameter search space and the deviation of each parameter from its starting point is shown for the global solution in Figure 5.9.

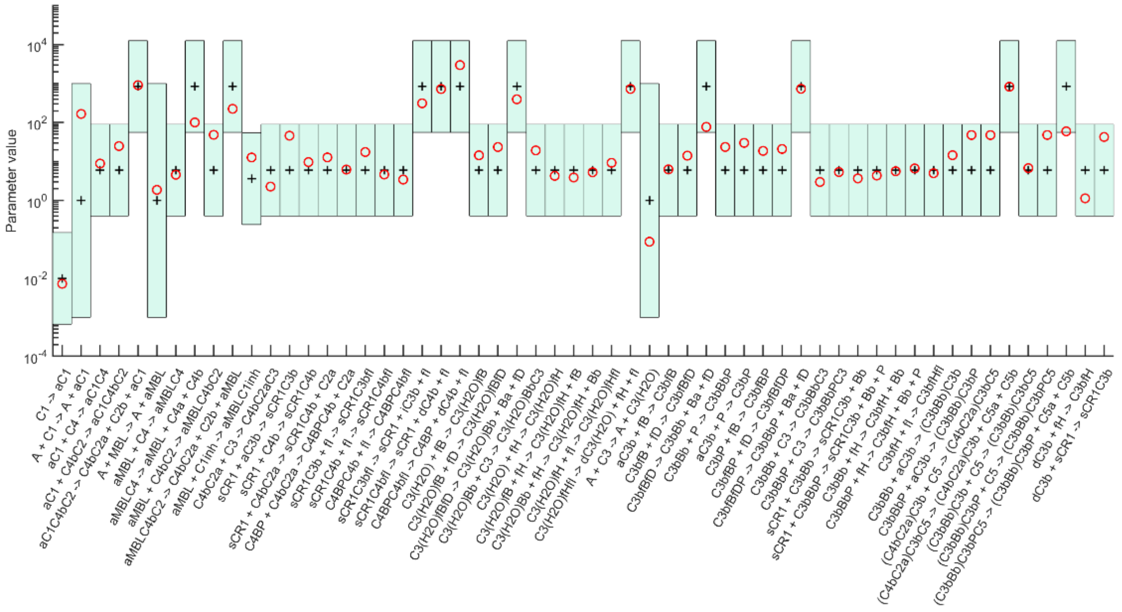


Figure 5.9 Parameter estimates of the C-ODE model fit compared to the fit bounds. Floated model parameters were given canonical starting guess values (black crosses) and were allowed to float within the fit bounds (green bars). The global solution of the fitting routine (red circles) is also shown.

The estimated parameter values at the fit solution were not at the limits of their fit bounds, indicating that the global minimum was wholly within the fit bounds and a different minimum would be required to improve the goodness-of-ft. The activation reactions detailed in Table 5.3 were not informed by median enzyme kinetics and were given larger fit bounds than the Michaelis-Menten reactions. The AP and LP activation reactions were changed by the fitting process from an initial guess of 1 to 0.1 and 2 µM^{-1} min^{-1} respectively, whilst the CP activation reaction was increased to a greater extent from an initial guess of 1 to 166 µM^{-1} min^{-1}. The change in rate constants by the fitting routine gives some indication as to the activity of each pathway although all have larger confidence intervals than their values due to over-parameterisation of the model fit to limited data.

The distributions of the optimised C-ODE parameters are shown in Figure 5.10, compared to the distributions of kinetic constants for Complement reactions (adopted in the model) and all enzymes in ref (515).

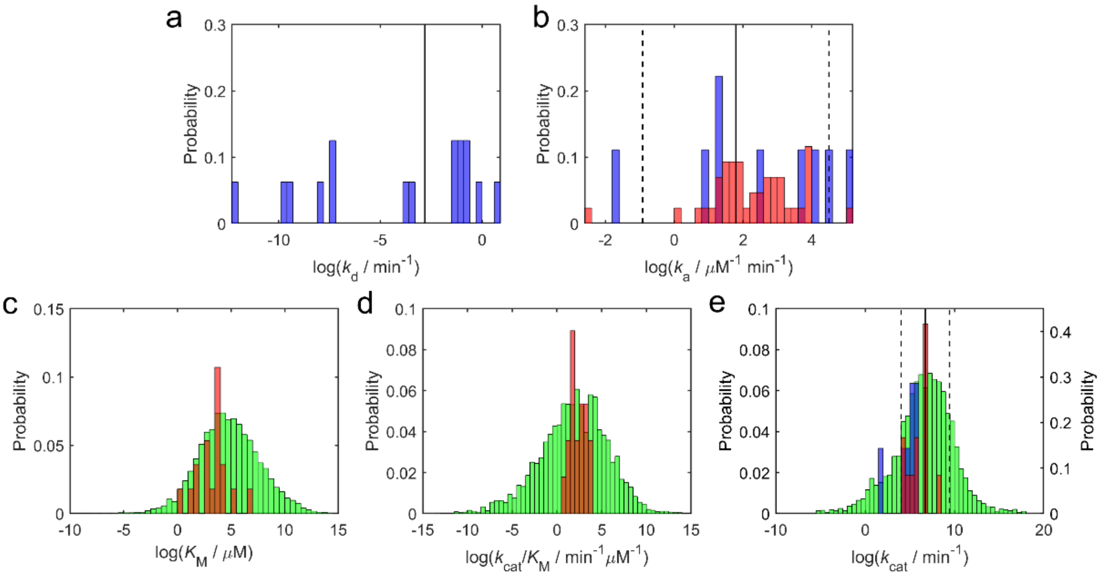


Figure 5.10 Distributions of C-ODE model parameters: a) k_d values, b) K_a values, c) K_M values, d) k_{cat}/K_M values, and e) k_{cat} values. Literature values in the model (blue) are compared to estimated values in the model (red) and the median values of all enzymatic reactions from supplementary data supplied by ref (515) (green). The value of the initial estimate before fitting (black line) and the fit bounds of the Michaelis-Menten reactions (dashed black lines) are also shown.

Critically, the fixed dissociation constant for unknown reactions was within the range of the values published for complexes in the Complement cascade (Figure 5.10a). The fit bounds for the fitted association constants covered the range of all published values for the Complement system (Figure 5.10b) with two exceptions: the formation of TCC (k_a=170 µM^{-1} min^{-1}) and the binding of FH to aC3b (k_a=0.18 µM^{-1} min^{-1}). A single published turnover number for the reaction C3bFHFI -> iC3b + FH + FI (k_{cat}=5.8 min^{-1}) was outside the parameter search region but within the range for all known enzymes (Figure 5.10e).

As described in Chapter 4, poorly defined parameters can be identified by a correlation analysis. Pearson correlation coefficients for the derived parameters varied between 1 (high positively correlated) and -1 (highly negatively correlated) and the distribution of correlation coefficients is shown in Figure 5.11.

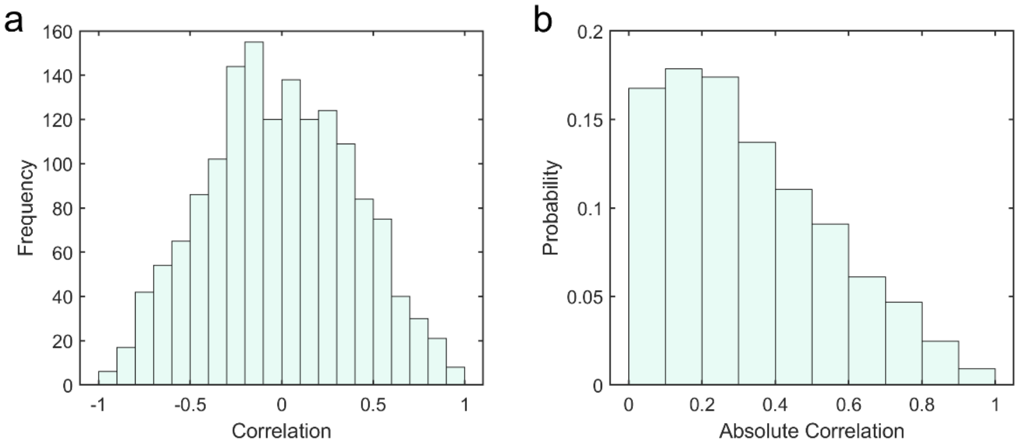


Figure 5.11 Distribution of parameter correlation from the correlation matrix of the C-ODE model fit. a) raw data showing a near-symmetrical distribution. b) absolute correlation values showing the majority of parameters are correlated <0.5.

The correlation analysis shows a nearly symmetrical distribution around zero, with a mean of -9.9×10^{-3} and SD = 0.4. The median absolute correlation was 0.29. Of the 1540 off-diagonal-pairwise parameter combinations, 258 (17%) and 1193 (77%) had correlations with values below 0.1 and 0.5 respectively. 52 (3%) had parameter combinations with values greater than 0.8. As discussed in Chapter 4, the correlation matrix could be assessed in future work to group similar parameters and reduce the error in the parameter estimates, but such an approach is challenging for large models.

Despite the low confidence in estimated parameters, the significance of each reaction on the model output can be explored with a local sensitivity analysis, Figure 5.12.

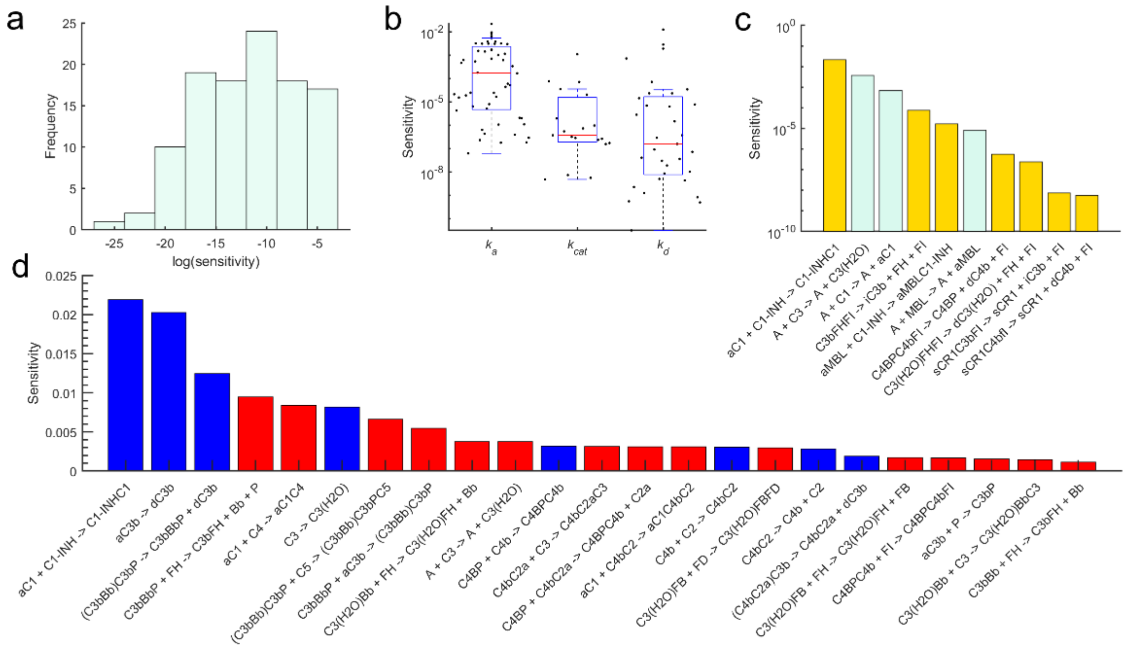


Figure 5.12 Local sensitivity analysis of C-ODE model rate constants. a) histogram of parameter sensitivity. b) sensitivity of k_a, k_{cat} and k_d kinetic parameters. c) sensitivity of the activation (green) and sink reactions (gold) at the fit solution. d) sensitivity of all rate constants, including known kinetic values and all dissociation rates which are fixed in the fitting process (blue) and floated parameters (red). For clarity, the 23 most sensitive rate constants (accounting for 95% of total sensitivity) are shown from a total of 109.

The distribution of sensitivity values for all model rate constants is shown in Figure 5.12a and is skewed towards higher values, suggesting a small number of reactions have little effect on the model fit. Figure 5.12b shows the relative sensitivities of the k_a, k_d and k_{cat} parameters. Three k_d values were highly sensitive outliers: the half-lives of the P-stabilised AP C5 convertase, C4bC2 and the CP/LP C5 convertase – deactivating reactions with literature values. The sensitivity of the activation reactions (Table 5.3) and sink reactions (Table 5.4) are shown in Figure 5.12c and indicate that CP and AP reactions were more sensitive that LP reactions.

The 23 most sensitive rate constants (accounting for 95% of total rate constant sensitivity) are plotted in order of sensitivity in Figure 5.12d which shows the CP/LP convertase binding to C3 was more sensitive than the equivalent reactions of the tick-over and AP convertases. In contrast, the AP (P stabilised) C5 convertase binding to C5 was more sensitive than the equivalent reactions of the CP/LP C5 convertase. The starting protein concentrations (fixed at their median serum values), known rate constants (fixed at literature values) and the unknown k_d parameters (fixed at the median enzyme value) account for 70% of the parameters in the C-ODE model and induced 86% of the total change in RMSE recorded during the analysis. The sensitivity analysis results for all model reactions and species are shown in Appendix 6 and Appendix 7 respectively.

An inspection of the overall C-ODE model performance and the activity of reactions in the cascade was tested with a flux analysis. The time-dependent flux through the C3 and C5 convertases of all Complement pathways is shown in Figure 5.13.

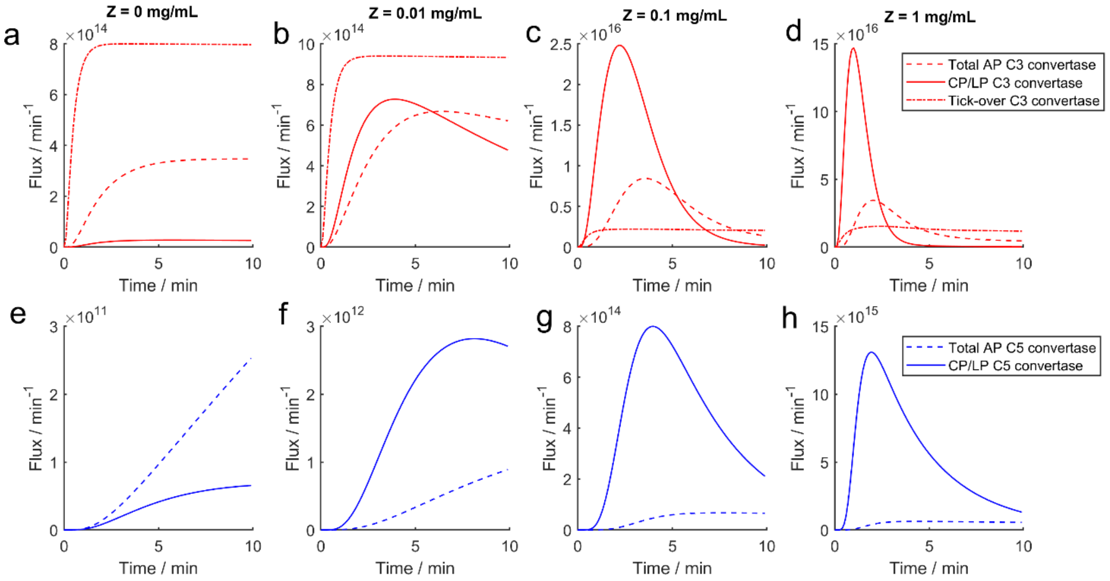


Figure 5.13 Time-dependent flux of substrates (in a 1 L compartment of serum) for the tick-over C3 convertase (dot-dashed lines), AP convertases (dashed lines) and the CP/LP convertases (solid lines) at four concentrations of zymosan. The simulation was performed using the COPASI settings described in section 5.3.4. Plots a-d show the C3 convertases (red) and plots e-h show the C5 convertases (blue). The flux shown for the AP is from both forms of the convertase (associated with P or not).

The maximum convertase activity in the C-ODE model occurs within 10 minutes of Complement activation by zymosan. The C3 convertases reach peak flux before the C5 convertases as anticipated, although the lag between upper and lower cascade decreases at higher zymosan concentrations. Furthermore, the C3 convertases display higher peak flux than the C5 convertases.

Integration of the flux profiles provides an estimate of the substrate turnover in each Complement pathway for a given time period. The effect of zymosan on 1) flux maximum, 2) time to flux maximum and 3) substrate turnover is shown quantitatively in Figure 5.14.

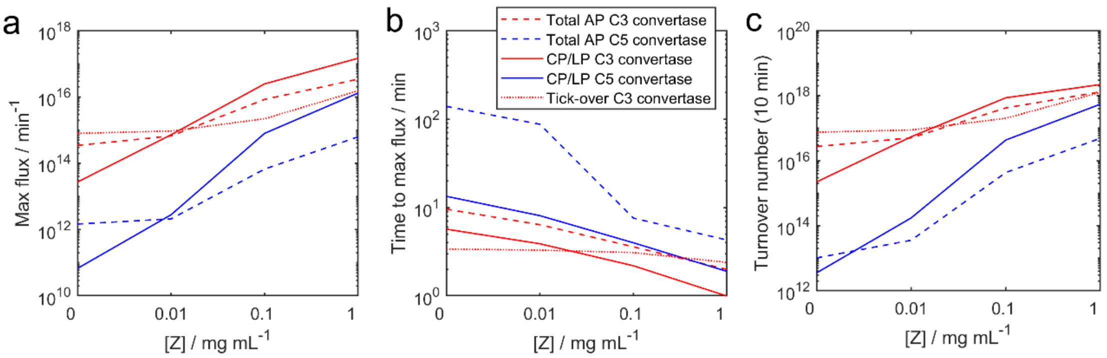


Figure 5.14 Zymosan dependence of the flux of substrates (in a 1 L compartment of serum) through the convertases in each pathway: a) flux maximum; b) time to flux maximum; and c) the number of substrates converted over 10 minutes. The simulation was performed using the COPASI settings described in section 5.3.4.

The CP/LP C3 and C5 convertases display higher peak flux, earlier time to maximum and higher substrate turnover than the AP and tick-over convertases at higher concentrations of zymosan, whilst the opposite is true without zymosan. The transition in C3 and C5 flux from majority AP to majority CP/LP occurs at 0.01-0.1 mg/mL zymosan in agreement with the data in Figure 5.4.

5.5 Discussion

The present study profiled Complement cascade activation by zymosan in pooled serum from the activation markers: Ba (AP), C4a (CP/LP) and TCC (TP). These data were added to existing activation data for C3a (total upper cascade) and C5a (early TP) to produce a data set totalling 160 data points from five analytes in four pathways. The experimental data are consistent with the existing understanding of Complement activation (such as linear flux in the TP, Figure 5.1) and provide a quantitative insight into the stoichiometry of the response and pathway activation. The data were used to optimise the C-ODE model of Complement activation containing 109 rate constants and 78 species. The C-ODE model was optimised based on two starting assumptions: 1) that the median values of all known enzyme parameters are good approximations for unknown kinetic constants and 2) that the respective median serum protein concentrations are good approximations for the values of the constitutive proteins in pooled serum. The optimised C-ODE model fitted the data within experimental RMSE and enabled pathway flux balance to be simulated under specific starting conditions, such as zymosan dose, to compare the activity in different Complement pathways.

5.5.1 Model Development

The C-ODE model describes Complement using exclusively bimolecular chemical equations as have the previous Complement models shown in Table 5.1, despite the current understanding that Complement initiation occurs on an activating surface, with zymosan providing a protected microenvironment for C3b (524). The kinetics of binding reactions confined to surfaces exhibit different kinetics to those in solution (525), as discussed in Chapter 4. The model might be made more accurate in future work by including equations such as the Langmuir equation in Chapter 2, for the interaction of proteins with a finite surface. Nevertheless, uniform activation of Complement throughout the reaction mixture is encouraged in the present study by continuous mixing of a fine suspension of zymosan which in previous work has been measured at a particle diameter of ~3 µm (526). Uniform Complement activation should more closely approximate bimolecular solution-phase kinetics than a continuous zymosan surface under serum without mixing, by minimising limitations of diffusion. A uniform reaction mixture also ensures that samples provide reliable estimates of concentration variations and the analytes chosen are not surface bound.

There is no experimental data for the concentration variation over time for most species in the C-ODE model, which is a major source of high parameter correlation and starting point dependence of model fitting. Chapter 4 showed that parameter correlation can be reduced by fixing unknown k_d values and a similar approach was taken here. The kinetic data from ∼2500 unique enzymes in ref (515) was used to fix the unknown k_d values at the most highly probable median value, allowing the fitting process to control catalytic activity with k_a and k_{cat} parameters only. The median sensitivity of the k_d parameters was lower than the median sensitivity of the k_a and k_{cat} parameters (Figure 5.12b) rationalising the fitting strategy of floating k_a and k_{cat} whilst fixing the k_d parameters because errors in the assumed k_d values will have comparatively lower impact on the model output. The fitting strategy resulted in only 3% of parameters being correlated with a Pearson correlation coefficient > 0.8 and a median value of 0.29.

The parameter search space was limited by fit bounds to ensure estimated parameter values were within a biologically plausible range with the highest probability (Figure 5.10). The chosen fit bounds are somewhat supported by observing that the 31 of the 34 published kinetic constants for Complement are wholly within the fit bounds and suggests that Complement is a system of “moderately efficient” enzymes. To ensure the best fit to the data was achieved within the fit bounds, multiple basins of attraction were explored in the parameter search space by running a gradient descent solver at multiple starting points (Figure 5.8). The best solution was chosen based on RMSE (Figure 5.9) and ensured that the specific combination of estimated rate constants best predicted the response of all analytes to zymosan, simultaneously (Figure 5.5).

Of all parameters in the model, the optimised k_a and k_{cat} parameters accounted for only 14% of model sensitivity at the fit solution, highlighting the extent to which the fixed parameters shape the model. Despite the *a priori* biological constraints of the fixed parameters, the optimised model fitted well to the experimental data – evidence supporting the model structure. The starting point of the fit based on median biological values led to a local solution with the second best minimum and suggests that median kinetic constants of enzymes may be used a good starting point for constructing other systems biology models with many unknown parameters – although global optimisation is still essential to explore the parameter space for the best solution.

5.5.2 Complement Activation Characteristics and Model Verification

In the present study, zymosan was used to activate the Complement system in serum. Zymosan is a polysaccharide found in the cell walls of yeast and has been shown to directly activate the AP, inducing rapid C3 cleavage in mixtures of purified AP proteins by providing a surface for C3b binding and convertase formation (524). Zymosan activation of the LP has also been demonstrated (527) and CP activation occurs in serum due to natural antibodies against yeast cell wall components (199). The data in Figure 5.3 show Ba and C4a rates of formation were zymosan-dependent, an observation which is consistent with anticipated activation of the AP and CP/LP.

Zymosan independent activation was observed in Ba as expected but also in C4a, supporting the previous reports of spontaneous activation in the CP/LP. Activation of the CP and LP was not distinguishable using the analytes shown because C4a is common to both pathways (Figure 5.1). However, C4BP reportedly regulates the assembly of LP convertases more strongly than the CP convertases because of a ~7 to 13-fold greater affinity for C4b produced via the LP than the CP (528), suggesting the majority of C4bC2a production may have been due to CP activity. Normal human serum has been shown to reduce spontaneous C1 activation, observed in purified solutions of proteins, to a negligible effect (506), suggesting IgG oligomers (507) may have contributed to C1 activity in the pooled serum used here. The spontaneous C1 activation rate recorded here for pooled serum may be different to the rate observed in other serum samples and may limit the transferability of the C-ODE model to a different system setting, such as fresh serum or PBS.

Before model validation with independent data, an intuitive test of the C-ODE model was to verify that model behaviour conformed to the existing understanding of the system. An equation for the spontaneous activation of C1 allowed the C-ODE model to fit well to the data (Figure 5.5) and the fitted rate constant was 0.01 min^{-1} compared to 0.173 min^{-1} measured in purified solutions of C1 components (506). The C-ODE model does not distinguish between C4b produced by CP or LP but the sensitivity analysis indicates that CP activation was more sensitive than the LP activation reaction by 800-fold (Figure 5.12c), in agreement with evidence for CP dominance over LP.

The sensitivity analysis suggests the CP/LP C3 convertase substrate binding and catalysis reactions were more sensitive than those of the AP C3 convertases, whereas the opposite is true for the C5 convertases. The peak flux values shown in Figure 5.13 are in agreement with the sensitivity analysis and show that a 100-fold zymosan increase from 0.01 to 1 mg/mL triggers a 200-fold increase in the (peak) flux of C4bC2a compared to a 50-fold increase in C3Bb(P). The turnover analysis in Figure 5.14 suggests a more complex mechanism: the CP/LP dominates at higher zymosan concentrations, whilst the AP is more active at lower zymosan concentrations. The switch occurs between 0 and 0.1 mg/mL zymosan - identified previously as a flux control point in ref (152) and the data in Figure 5.4. Comparing CP/LP and AP activity following infection might provide a novel method for identifying pathogen load.

The relative turnover in each pathway is also time-dependent. The maximum flux in the C-ODE model occurs at 0.2 seconds following 1 mg mL^{-1} zymosan activation in the cleavage of C4 by the activated C1 complex. However, CP/LP activity is short-lived and the total flux from upper cascade to lower cascade after 20 minutes is predominantly due to the AP C5 convertases, which display higher turnover than the CP/LP C5 convertases over the 24 hours simulation at all zymosan concentrations except 1 mg/mL where CP/LP and AP C5 turnover is similar. The result of minimal CP/LP activity after 20 minutes is consistent with evidence that AP amplification is responsible for 80% of TCC formation whether Complement is activated via the CP or LP (529, 530).

The flux analysis shows that Complement activation of pooled serum by zymosan displays a characteristic pattern of pathway flux. Figure 5.14b reveals that the AP convertases reach peak flux 1-4 minutes later than those of the CP/LP and the lag decreases with zymosan concentration. The result is consistent with experimental evidence that non-encapsulated *crypotococci* induce C3 cleavage rapidly via the CP followed by the AP with a lag of order minutes (199). In contrast, uncapsulated *cryptococci* have been shown to exclusively activate the AP, characterised by a 4-6 minute lag in C3 deposition following activation due to the asynchronous and focal deposition of C3b (199). More early time data should be collected to observe the initial rate of C3 cleavage more precisely and this is attempted with the LiScAR assay for C3 in Chapter 7. Comparing CP/LP and AP activity following infection might provide a novel method for identifying pathogens.

The deviation of the floated parameters from their (equal) starting values is shown in Figure 5.9 and can be compared to qualitative observations of reactions. The literature states that FB has a higher affinity for C3(H2O) than to C3b (157) and this is reflected in the fitted values for these association constants in the C-ODE model: 14.44 μM^{-1} min^{-1} versus 6.29 μM^{-1} min^{-1}, leading to estimated affinity constants of 37 nM and 9.5 nM respectively. Furthermore, experimental evidence has shown the convertase activity of $C3(H_2O)Bb$ to be 50% lower than the C3bBb complex, as modelled here, whilst being more resistant to inhibition and remaining active significantly longer (186). The flux profiles are in good agreement, with $C3(H_2O)Bb$ dominating C3 turnover at low zymosan concentrations and persisting for longer but showing the lowest peak flux at higher concentrations of zymosan.

The C-ODE model was used to assess species limiting Complement activation after 24 hours of activation with 1 mg/mL zymosan, Figure 5.15.

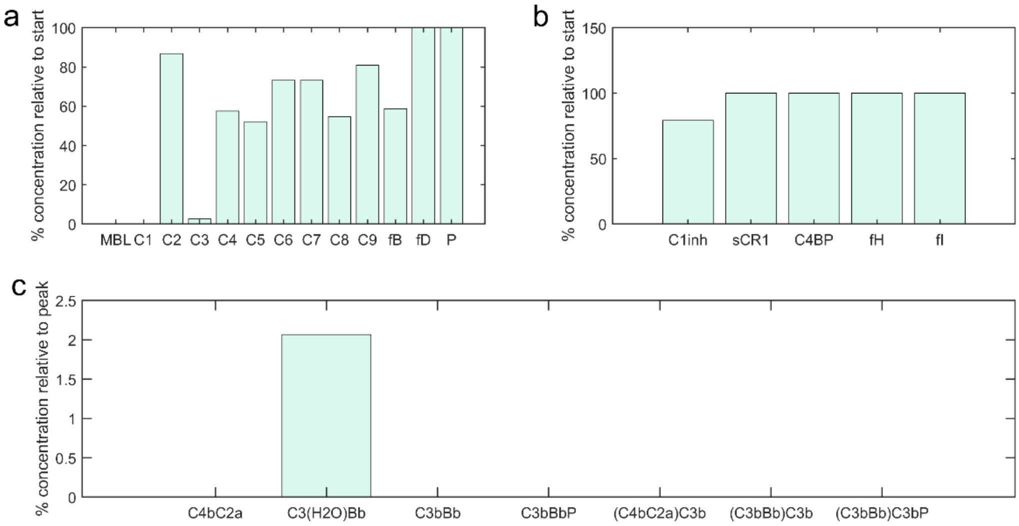


Figure 5.15 Model analysis of limiting species after 24 hours of Complement activation with 1 mg/mL zymosan. The simulation was performed using the COPASI settings described in section 5.3.4. The concentrations of model species after activation relative to their starting values are shown for: a) circulating activation proteins, b) circulating deactivation proteins and c) convertase complexes. MBL and C1 are completely depleted and C3 is depleted by >90%, whilst all other constituent protein concentrations remain at >50% of their initial values. The cascade is limited by the removal of C3 and C5 convertases by regenerated regulatory proteins. The only convertase remaining after 24 hours is the tick-over C3 convertase (at 2% of peak concentration).

MBL and C1 are completely depleted and C3 is depleted by 97%, whilst all other constituent protein concentrations remain at >50% of their initial values. Before C3 is fully consumed, the C-ODE model suggests the cascade is predominantly limited by the removal of C3 and C5 convertases by regenerated regulatory proteins. The only convertase remaining after 24 hours is the spontaneously generated tick-over C3 convertase (at 2% of peak concentration).

The duration of tick-over is predicted to exceed the 24-hour time-course of the data presented here, as although C1 is completely depleted, all other constituent protein concentrations including MBL, C3 and C4 remain at >50% of their initial values, Figure 5.16.

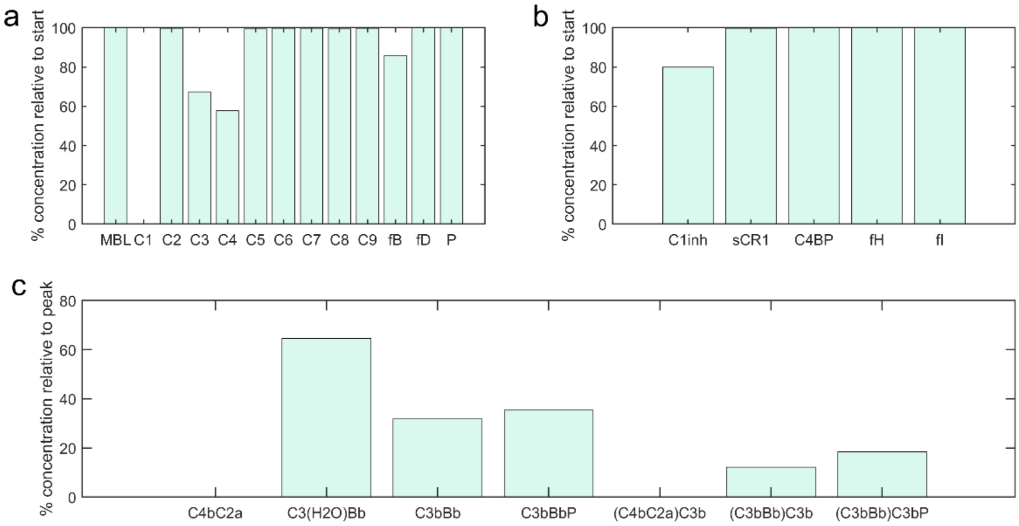


Figure 5.16 Model analysis of limiting species after 24 hours of Complement activation via tick-over. The simulation was performed using the COPASI settings described in section 5.3.4. The concentrations of model species after activation relative to their starting values are shown for: a) circulating activation proteins, b) circulating deactivation proteins and c) convertase complexes. As for activation with zymosan, C1 is completely depleted, although all other constituent protein concentrations, including MBL and C3, remain at >50% of their initial values. C4 is more depleted than C3. The CP/LP convertases are completely depleted whilst all other convertases remain at >10% of their peak values indicating that tick-over in vitro *continues after 24 hours at 37^{o}C.*

After 24 hours of tick-over all convertases remain at >10% of their peak values except for the CP/LP C3 and C5 convertases. Not until 25 days of tick-over is the tick-over C3 convertase reduced to <1% of its peak value and C3 reduced by >99%. All other convertases have negligible concentrations below 0.01% of their peak values, reaffirming the longevity and dominance of the tick-over convertase at low levels of stimulus.

The C-ODE model indicates that the most sensitive reactions of the system are inhibitory rather than activating in agreement with a previous Complement model (185) and conforming to existing knowledge that Complement is highly suppressed under normal physiological conditions to prevent over-activation and damage to host cells. As discussed in Chapter 1, Complement inhibiting drugs are typically designed to target proteins responsible for activation: such as C5 targeted by Eculizumab; FD targeted by Lampalizumab and C1 targeted by Berinert, Ruconest and Cynrise (222). Figure 5.12 shows the top three most sensitive rate constants are regulatory and had literature values: C1 inhibition by C1inh; the C3b thioester half-life and the half-life of the P-stabilised AP C5 convertase. The sensitivity analysis suggests that enhancement of the existing regulatory mechanisms of the cascade may be a powerful strategy to treat patients with overactive Complement systems, whilst drugs designed to suppress these inhibitory reactions might enhance the Complement response in immunocompromised patients. The sensitivity analysis in Figure 5.12 is also consistent with previous models of Complement activation in the reactions identified as being most sensitive. One model (510) identified C3b binding, C2 cleavage, C4 cleavage and C4bC2a decay as the most sensitive reactions, whilst another model (511) identified FH binding to C3(H_2O) and P binding as being the most sensitive reactions. All these reactions appear in the 23 most sensitive reactions of the C-ODE model in Figure 5.12.

P is the only known positive regulator of AP activity and is secreted by neutrophils following stimulation by *N*-formyl-methionine-leucine-phenylalanine (fMLP) or tumour necrosis factor alpha (TNF-α) during infection (531) and endothelial cells produce P when induced by shear stress (532). Notably, the concentration of P is significantly decreased in sepsis patients on admission to intensive care but increased after clinical recovery to exceed levels observed in healthy volunteers (533). The effect of P concentration on model behaviour can be observed in the sensitivity analysis: the rate constants of C3 and C5 binding to P-stabilised convertases are more sensitive than their equivalent reactions in the absence of P (Figure 5.12). Furthermore, COPASI was used to quantitatively estimate the effect of P concentration on the flux of the AP C3 and C5 convertases according to the C-ODE model, Figure 5.17.

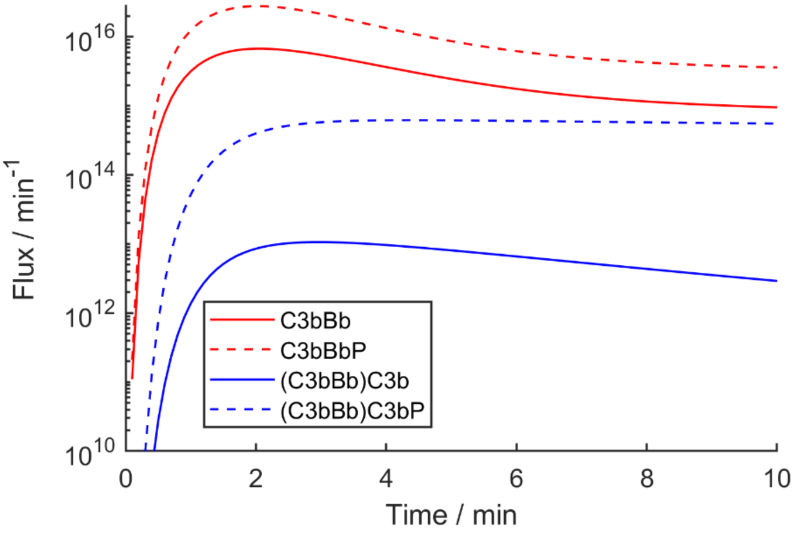


Figure 5.17 Model prediction of time-dependent flux of substrates (in a 1 L compartment of serum) through the C3 convertases (red) and C5 convertases (blue) following 1 mg/mL zymosan activation. The simulation was performed using the COPASI settings described in section 5.3.4. The effect of P on the Alternative convertase flux is shown for un-stabilised convertase (solid lines) and P-stabilised (dashed lines).

The flux analysis is consistent with the current understanding of P function and the C3 and C5 convertases in the C-ODE model display significantly higher flux when stabilised by P. Notably, the relative flux increase due to P is greater for the C5 convertase than the C3 convertase, suggesting that P may have a greater effect on terminal pathway activation.

Therapeutic recombinant P has been shown to confer increased resistance to infection (534) and the C-ODE model was used to predict the effect of doubling the concentration of P in the pooled serum, Figure 5.18.

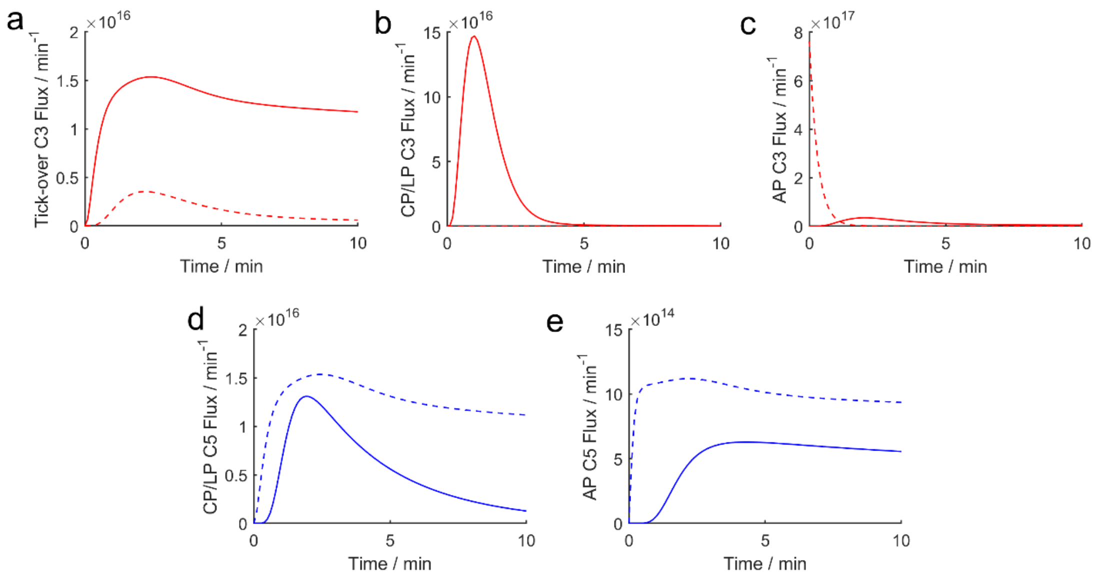


Figure 5.18 Model prediction of convertase flux (in a 1 L compartment of serum) following 1 mg/mL zymosan activation at median P concentration (solid lines) and twice-median P concentration (dashed lines) for C3 (red) and C5 (blue) convertases. The simulation was performed using the COPASI settings described in section 5.3.4.

Doubling P concentration is predicted to increase peak flux through the AP C3 and C5 convertases by a factor of 10 and 2 respectively. The peak flux through the CP/LP C5 convertase was also increased by ~10%. Perhaps most interestingly, flux is decreased in the convertases that do not contain C3b (tick-over and CP/LP C3 convertases), attributed to competition for the C3 substrate. The result highlights the utility of the model for predicting the complex behaviour of the Complement system to therapeutic interventions. In principle the prediction may be personalised for a given patient, by matching the profile of starting protein conditions which may be greater or lower than the median values studied here.

Complement FH deficiency is a rare disorder which results in dysfunctional or undetectable concentrations of circulating FH, resulting in activation of the AP and secondary depletion of C3 and other proteins of the AP (535). The plasma levels of FH are positively correlated with plasma C3 in mice (228) and mice with FH deficiency have 87-94% less C3 than healthy controls (536). Human FH levels influence the degree of C3 opsonization and killing of *Phenuomococci* with an optimal concentration of 300 µg/mL (1.9 µM) - the value adopted in the C-ODE model (537). FH-deficient patients are therefore more susceptible to infections, C3 glomerulopathy, and HUS (535). The C-ODE model was used to quantitatively estimate the effect of complete FH deficiency on the consumption of C3 and C5 by observing the differences in C3 and C5 convertase flux, Figure 5.19.

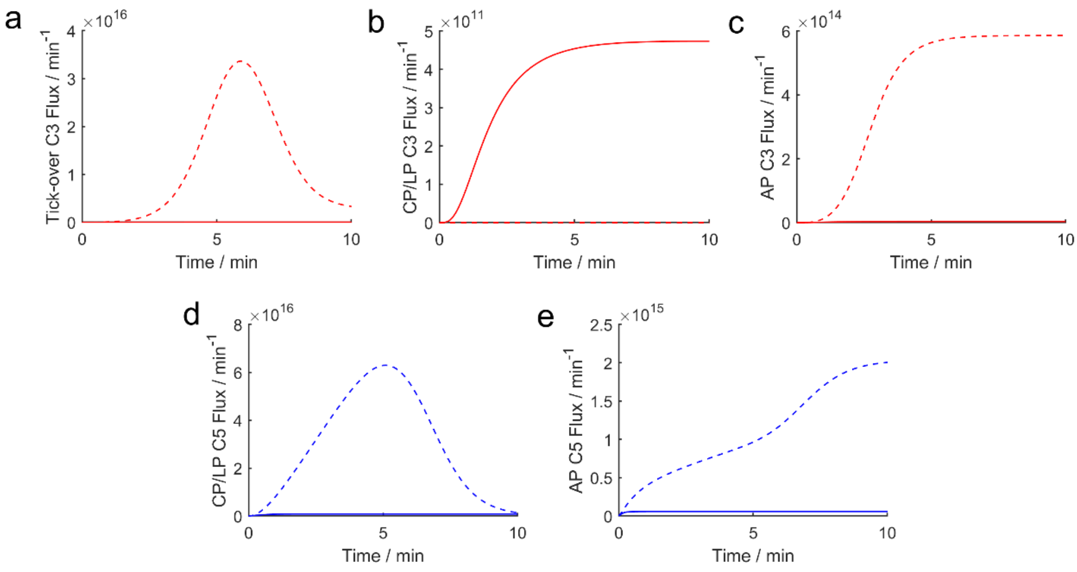


Figure 5.19 Model prediction of convertase flux (in a 1 L compartment of serum) following 0 mg/mL zymosan activation at median FH concentration (solid lines) and FH-deficient serum (dashed lines) for C3 (red) and C5 (blue) convertases. The simulation was performed using the COPASI settings described in section 5.3.4.

The C-ODE model predicts that C3 and C5 consumption is increased significantly at 0 mg/mL zymosan in the absence of FH. Peak C3 convertase flux is increased ~1000-fold, whilst peak AP C3 flux and CP/LP and AP C5 flux is increased ~100-fold. Interestingly, the CP/LP C3 convertase flux is reduced to zero in the absence of FH - attributed to increased competition for the C3 substrate as seen with the P supplementation simulation in Figure 5.18. The CP/LP C5 convertase flux profile peaks at 5 minutes in the absence of FH which causes the rate of increasing AP C5 convertase flux to slow at 5 minutes before continuing to rise. The patterns of convertase flux are a new insight into the dynamics of Complement activation that would be difficult to measure experimentally due to the challenges associated with distinguishing between protein consumption due to each pathway. The results further indicate that inhibiting one pathway in order to measure flux through the other would not capture the effects of substrate competition between convertases of different pathways.

5.6 Conclusions

The C-ODE model is the most comprehensive model of Complement activation to date: incorporating published rate constants and optimised with biological data from all pathways of the system at varying degrees of stimulus. The use of median enzyme kinetics and median serum concentrations produced a good starting point for model fitting and these values could be used as effective priors when constructing other systems biology models with unknown parameters. Additional priors might also be considered, such as reactions associated with pathways under higher selection pressure exhibiting higher values of k_{cat} (515). There is evidence for the CP evolving from the LP to complement the adaptive immune system (216, 538), and the AP is hypothesised to be the evolutionarily oldest component (although this is contested (539)). The mean k_{cat} values published for the AP are 50% higher than those of the CP/LP, although the standard deviations indicate the difference is not statistically significant. The effect of substrate mass on K_m (515) was not investigated but may also be considered to improve initial parameter estimates.

The sensitivity and flux analysis of the C-ODE model revealed characteristics of the Complement cascade which have been previously reported, including: disproportionate upper and lower cascade activation (152); zymosan activation proceeding via CP followed by AP after a lag of several minutes (199); the importance of P for increasing AP convertase turnover and the importance of FH to limit activation (Figure 5.1). The model also highlighted the sensitivity of the system to its many inhibitory mechanisms, which may inspire novel strategies for therapeutic regulation of Complement. A logical progression of the present study would be to collect data to define the most sensitive unknown rate constants estimated by the sensitivity analysis. The ultimate validation of the model is to test predictions based on different starting conditions.

The C-ODE model has the potential to be personalised for individual cases and substituting the starting protein concentrations for values across the physiological range may be used to predict the variation of *in vitro* Complement activation by zymosan in the healthy population. The sensitivity of the C-ODE model to different surfaces could also be adjusted via the activation reactions which ultimately control the flux through each pathway but are currently optimised for a suspension of zymosan in PBS. Chapter 6 reports the results of *in vitro* Complement testing of healthy adult serum using the same zymosan activation protocol described here. The results provide a first look at the variation of Complement activation in the healthy adult population and provide independent data from the training data set with which to validate predictions of the C-ODE model.

6 *in vitro* Complement Activation Phenotype Variation in a Healthy Adult Population

6.1 Introduction

Complement activation dynamics in the fluid phase depend on the stimulus and the relative concentrations of Complement proteins and can be profiled over time using immunoassays for activation products such as TCC. The data shown in Chapter 5 indicate two characteristics of Complement activation that can be observed with minimal sampling: 1) the initial rate of activation marker concentration change and 2) the total activation marker concentration change (or capacity) following activation. These measures of Complement activation dynamics are phenotypes of immune state which may be useful in characterising host response to infection (34) as discussed in section 1.1 and are relatively simple to record using immunoassays as shown in section 5.4. Serum can be challenged *in vitro* with a standard concentration of zymosan and phenotypes measured after specific times – coined a ‘Z-test’.

The C-ODE model, developed in Chapter 5, was optimised using data from pooled human serum and predicts the response of the Complement system in pooled serum to varying degrees of zymosan stimulus. As stated in Chapters 4 and 5, predictive models must fit well to both the data used for parameter estimation and the results of independent experiments in a process known as model validation (540). Chapter 4 described an ODE-based model which fitted well to data in a solution setting but required adjustment via activity coefficients to be usefully predictive when the same interaction occurred at a surface. The intended utility of the C-ODE model is to accurately predict the Complement response to zymosan in patient serum samples with varying starting protein concentrations. Whether or not the C-ODE model can usefully be applied to different starting conditions must be tested (541).

The importance of model validation is such that some studies publish model construction (structure selection and parameter estimation) and model validation independently, allowing the model to be repeatedly validated in new system settings (484, 542). As mentioned in Chapter 4, a common methodology applied by the modelling community for the analysis of biological models is partitioning the available data into two parts: a training dataset for model construction and a ‘hold-out’ validation dataset which is used to test the predictive power of the model (543). However, partitioning of a single dataset into training and validation sets can lead to incorrect conclusions about the validity of a model. For example, a validation dataset using the same dependent variables under very similar conditions to the training dataset is a poor test of model predictions. Similarly, a training data set which is very limited (failing to define key features of the fitted model) is unlikely to predict hold-out data of untrained dependent variables under extremely different conditions to the training data. Therefore, accurate biological knowledge of the system structure and parameter values is required in order to partition the data such that construction and validation is not biased – a paradoxical situation for complex models with uncertain structures and unknown parameters. To minimise bias in the development of models with uncertain structures, random and repeated partitioning of available data into training and validation sets (cross-validation) can be used (461). However, the structure of the C-DOE model is relatively well-known and is based on the accepted reaction mechanisms of 78 proteins. Of 109 rate constants, 34 have published values and unknown rate constants were constrained to physiological ranges and estimated during model fitting to experimental data of 5 activation products following zymosan activation of pooled serum. Therefore, the C-ODE model structure and a significant proportion of its parameters were not biased by the fitting process, making the model appropriate for validation with independent data without refitting or cross validation.

The Complement phenotypes discussed here provide independent data for model validation and are functional tests of immune state which could be investigated for their ability to identify hospital patients at higher risk of nosocomial infection. As discussed in Chapter 1, prognostic models are commonly used combine two or more risk biomarkers to predict clinical outcomes but relatively few published prognostic models are adopted for clinical practice due to insufficient validation (544). The choice of sample size, statistical methods and inclusion criteria are important decisions in the design of clinical investigations and will be discussed briefly here.

6.1.1 Sample Size

Determining the optimal number of units of analysis to be included in a study is an essential component of study design. Larger sample sizes may require more time and resources to collect the data, whilst smaller sample sizes reduce the power of the clinical study and risk that real differences will not be identified (545, 546). Sample size planning ensures that the study is efficient: large enough to confidently answer the main research question but not excessively large to waste resources (547). The main research question of the present study is: *are the distributions of Complement activation phenotypes observed in the healthy adult population significantly different from those predicted by the C-ODE model?*

The present study is a preliminary investigation which will be used to determine the distributions of Complement activation phenotypes. Preliminary studies are used to estimate a given effect size and facilitate the planning and conduct of pilot studies which precede full trials (548). The present study requires sufficient data to estimate the range and shape of the phenotype distributions which can then be compared to C-ODE model predictions. A sample size of 10-75 is recommended for pilot studies depending on the effect size identified in preliminary studies (549).

6.1.2 Statistical Methods

A number of statistical tests are available to determine whether two distributions are different but the choice of statistical test is important and depends upon the type of data being compared. Analytical data may be discrete or continuous and may be from a known or unknown distribution. Parametric tests assume data have come from a probability distribution which can be defined using an equation, such as a normal distribution (550). Parametric tests will therefore lead to erroneous conclusions when used to compare samples from distributions which violate the assumptions of the test. The present study will derive rate and capacity phenotype measurements from analyte concentration data on a continuous scale with no prior knowledge of the distributions. A normality test should be performed before choosing an appropriate statistical test to compare the datasets.

The normality of data can be assessed visually using quantiles (also known as percentiles) (551). Quantiles are ‘cut points’ which split a distribution into equally likely ranges: the median is the ‘0.5 quantile’ (50th percentile) and splits a distribution into two groups with 50% of the data falling below the value of the median. A quantile-quantile plot is generated by plotting the *i*th smallest value of the test data against the $\frac{i-0.5}{n}$ quantile of the standard normal distribution (mean

= 0, σ = 1). For example, the smallest value (*i*=1) in a sample of 10 values would be plotted against the value of the 0.05 quantile of the standard normal distribution, -1.64, below which 5% of the data fall. A quantile-quantile plot is shown in Figure 6.1.

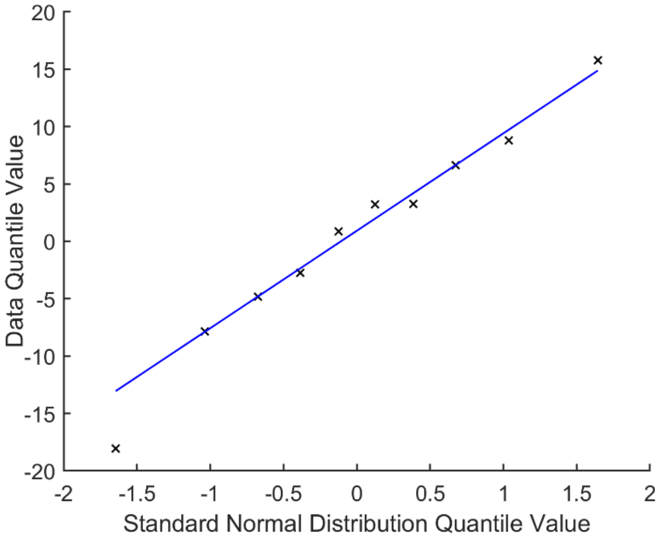


Figure 6.1 Quantile-quantile plot of 10 data points from a random normal distribution with a mean of zero and σ of 10. For any sample of n=10, the values of the x axis are unchanged. The data quantile values show a linear relationship with the standard normal distribution quantile values, indicating the data is normally distributed.

The data in the quantile-quantile plot will fit well to a linear model if normally distributed. The quantile-quantile plot is valuable for observing where deviations from normality occur in the distribution. For example, in Figure 6.1 the data quantile value for the 0.05 quantile (x=-1.64) is lower than expected – although the result is known to be due to chance because the data were generated randomly.

The Shapiro-Wilk (SW) test is a strongly recommended method for quantitatively testing the normality of a distribution (551) to produce the SW test statistic, W. The SW test is related to the quartile-quartile plot as the data are ordered by size and multiplied by coefficients derived from the expected values of the standard normal distribution quantiles (the x axis in Figure 6.1) during the calculation of W (552). The null-hypothesis of the SW test is that the population is normally distributed, with a test statistic of 1 for a perfectly normal distribution.

The value of the test statistic will vary due to chance and by a larger amount for small sample sizes. Therefore, a measure of confidence in the test result considering the sample size is useful. The P value of a statistical test is the probability of recording test statistic larger than a critical value expected for a given sample size when the null hypothesis is true (incorrectly rejecting the null hypothesis due to chance). As such, if the P value is lower than the desired significance level then there is evidence against the null hypothesis and a value of $P<0.05$ is often chosen (553) due to convention rather than any mathematical basis and no single index should substitute for scientific reasoning (554). Once normality of the data has been evaluated, additional comparative statistics must be applied to test whether the clinical and simulated data could be drawn from the same population.

If the distributions being compared are normally distributed, the mean values may be compared using the two-sample t-test (555). The two-sample t-test is analogous to the signal to noise ratio discussed in Chapter 2: the difference between the sample means is divided by a measure of variation in the distributions. Hence, the value of the test statistic is maximised when the sample means are well determined by low variance and the distributions are well-separated. There are two main variations of the two-sample t-test: the Student's t-test is used when both samples have equal variances and Welch's t-test is used when both samples have unequal variances (556). Statistically significant differences in variance may be assessed using the two-sample F-test (557). The null hypothesis of the two-sample F-test is that both samples come from normal distributions with the same variance whilst the null hypothesis of the two-sample t-test is that the data in the two distributions have equal means. The null hypothesis of each test may be rejected if the P value obtained for the test statistic is lower than a predefined significance level.

If the distributions being compared are not both normally distributed, they may be compared using a cumulative density function (CDF) and the two-sample Kolmogorov-Smirnov (KS) test (558). A CDF evaluated at a given value in a distribution, is the probability of that value being less than or equal to the given value. The two-sample KS test simply calculates the largest difference in CDF value between two distributions and is therefore sensitive to any shape or magnitude differences between distributions – reminiscent of the Yuden Index statistic discussed in Chapter 1. The null hypotheses of the KS test is that both samples come from the same distributions and is rejected if the P value obtained for the test statistic is lower than a predefined significance level. An example of different CDF plots and a graphical representation of the two-sample KS test statistic are shown in Figure 6.2.

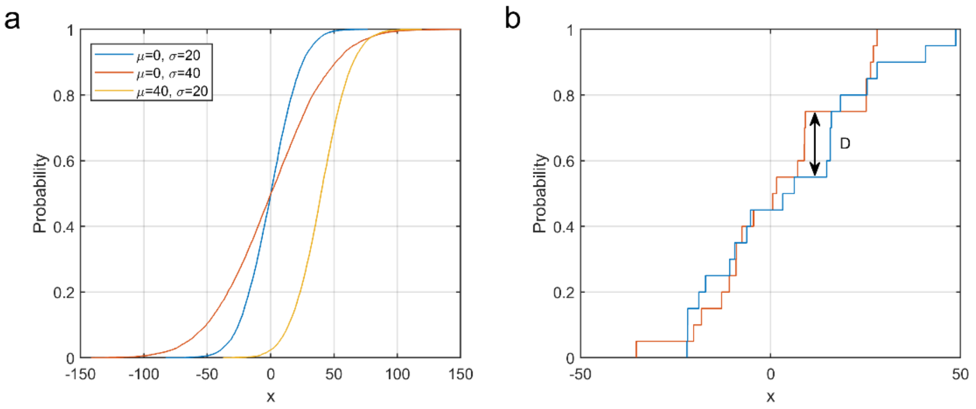


Figure 6.2 Empirical CDF plots. a) simulation using 10,000 samples of 3 random normal populations. Blue and orange distributions have equal means and unequal variances. Blue and yellow distributions have unequal means and equal variances; b) simulation using 20 values sampled randomly from 2 populations with equal means and variances. The largest difference in CDF value is the KS test statistic, labelled D, with a value of 0.2. The value of D is due to chance and will decrease with larger and more representative sample sizes.

The Complement phenotype data collected in the present study were first assessed for normality before the most appropriate test for distribution differences was chosen: either the parametric t-test or the non-parametric KS test. The P values of any statistical test be reported so that the precision of model predictions can be judged on a continuous scale rather than arbitrary classification of results as “significant” or “non-significant” (559).

6.1.3 Inclusion and Exclusion Criteria

The power of a study may be increased by reducing the heterogeneity with strict inclusion / exclusion criteria. For example, serum samples could be excluded from the study before Complement phenotype testing if the concentrations of any proteins are outside the expected healthy ranges simulated by the C-ODE model. However, excessively strict inclusion criteria risk that conclusions drawn from the study do not represent the population to which the findings might be applied.

There is no statistical association between the gender or age of individuals and the levels of C3 opsonisation (560) but other differences in Complement activity have been observed and significant variation has been reported (561). The aim of the present study is to establish the total variation in Complement activation between individuals and hence strict inclusion criteria would bias these results. A good recruitment criterion for the present investigation was self-certified healthy adults with no chronic inflammatory conditions that might influence Complement activation due to illness. CRP can be screened in each sample by turbidimetric assay in the clinical chemistry laboratory to check for excessive inflammation as a post-recruitment exclusion criterion. In time, CRP testing could be performed on the LiScAR biosensor platform in parallel with the Complement assays as a point-of-care Z-test.

6.2 Aims and Objectives

The aim of this chapter is to report the variation of Complement activation phenotypes in the healthy adult population and test the accuracy of the C-ODE model in predicting the phenotype distributions.

The first objective was to collect *in vitro* Complement activation data from healthy volunteers which provided reference distributions of certain Complement activation phenotypes. Two phenotypes of Complement activation were proposed here, based on zymosan activation *in vitro*: initial rate of analyte concentration increase (rate phenotype) and absolute change in analyte concentration (capacity phenotype).

The second objective was to predict the variation in Complement activation predicted in the healthy population using the C-ODE model from Chapter 5. The clinical data and model prediction were then compared using an appropriate statistical test. The agreement between the measured phenotype distributions and those predicted by the model indicates the extent to which the model is predictive in the new system setting of fresh serum samples: model validation.

The final objective was to predict how Complement depletion might affect the measured phenotypes, demonstrating the use of the C-ODE model and data to estimate an effect size with which future clinical investigations could be planned. Complement depletion is expected to be a risk factor for infection as discussed in Chapter 1.

6.3 Materials and Methods

The recruitment criteria for the study were adult volunteers self-reporting as healthy with no chronic inflammatory conditions that might impact their Complement protein concentrations and Z-test results. Volunteers were recruited through personal contact. The collection of blood samples was approved by the local University of Exeter Medical School Research Ethics Committee (UEMS REC reference number: May18/B/152Δ1). Informed consent was obtained from the volunteers who were required to complete the consent form shown in Appendix 8 before the study could begin. All serum experiments were performed in a Level 2 laboratory following relevant risk assessments. CRP was screened in each sample by turbidimetric assay in a Clinical Chemistry Laboratory to assess for inflammation and was used as a post-recruitment exclusion criterion. C3, C4 and IgG concentrations were also tested to indicate whether the study cohort was well matched the reference ranges reported for these proteins – a potential source of error for model predictions.

6.3.1 Modelling Predictions

The C-ODE model and *in vitro* data from Chapter 5 were used to choose common sampling time points that would capture the initial rate and capacity of Complement protein concentration increase for a number of activation products. ELISA data for Ba, C3a, C4a, C5a and TCC concentration increase in response to varying zymosan dose were used to train the model and hence there is higher confidence in the predictions of these analytes than those without data for validation. A comparison of the C-ODE model fit to pooled serum data from a 1 mg/mL zymosan activation and published data using a similar protocol (516) with 10 mg/mL zymosan and heat-aggregated IgG are shown in Figure 6.3.

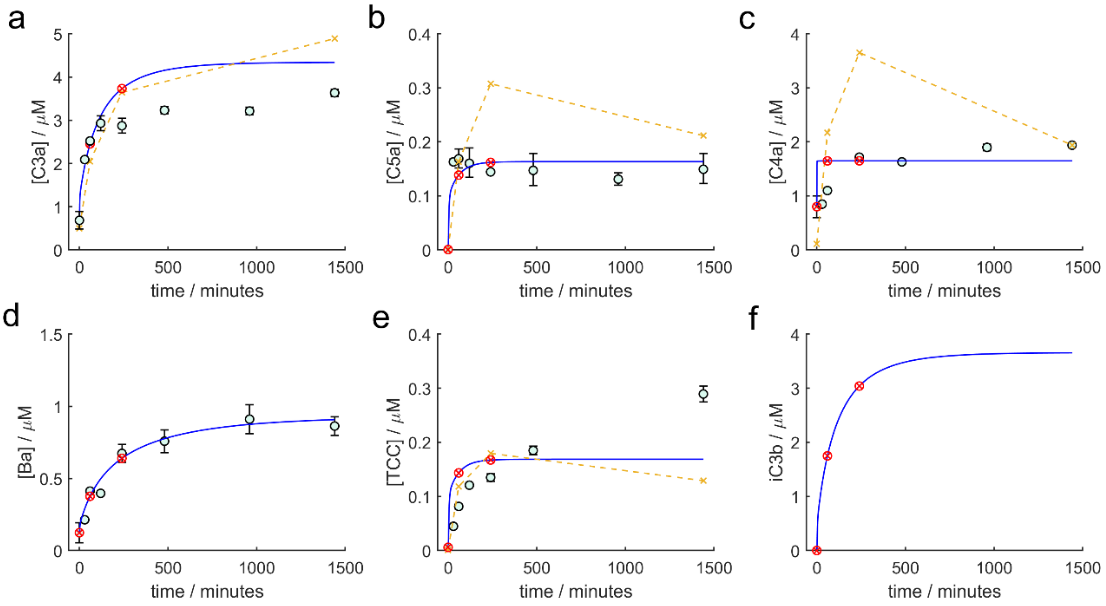


Figure 6.3 Complement activation model fit (blue) to ELISA data (green) from section 5.4 showing sample points for the clinical investigation and cohort simulations (red targets). Sampling times were chosen based on their ability to determine initial rate and total capacity of concentration increase. Data from the literature following a similar protocol with heat aggregated IgG (516) are shown for comparison (gold). Error bars are 2SD from duplicate determinations.

The timescale of analyte concentration increase was similar for all previously measured analytes and the sampling time points chosen to measure initial rate and capacity of concentration increase were: zero hours, 1 hour and 4 hours. Incidentally, the sampling times were the same as chosen by Bergseth et al. (516) in defining their human serum standard for Complement activation studies. The rate phenotypes were derived in units of concentration per minute using the formula: rate = (A_1 - A_0) / 60, where (A_1 - A_0) was the difference between the analyte concentrations at zero and 1 hour following zymosan activation. The capacity phenotype of each analyte was derived using the formula: capacity = A_4 - A_0, the difference between the analyte concentrations at zero and 4 hours following zymosan activation.

The rate and capacity phenotype distributions of each analyte were predicted by the C-ODE model using 250 starting conditions. Each starting condition was generated by sampling randomly from the published serum reference ranges (Table 6.1) of the 18 pre-activation proteins with circulating concentrations. Sampling was performed in MATLAB using the ‘betarnd’ function to generate a random sample from a beta distribution (shape parameters = 5,5) scaled to the limits of each reference range. The beta distribution is applied to model random variables in a number of fields including genetics to model allele frequency (562). The resulting sample distribution approximated normal random sampling (Shapairo Wilk, $P<0.05$). Beta random sampling has the benefit of being non-zero strictly within the set limits – unlike a continuous normal distribution, there is zero probability of randomly generating a protein concentration outside of its reference range.

6.3.2 Clinical Investigation Protocol

ELISA assays for Ba, C4a, C5a, iC3b and TCC were purchased from Quidel (#A033, #A035, #A021, #A006, #A020). Turbidimetric assays for IgG, C3 and C4 were performed using the Cobas platform (Roche) at Royal Devon and Exeter Hospital Clinical Chemistry Laboratory. Phosphate buffered saline (PBS), EDTA and zymosan were obtained from Sigma Aldrich. Serum collection tubes (#367895), 21 guage needles (#360213) and adaptors (#364815) were purchased from Medisave.

A new venepuncture performed by a phlebotomist was used to draw blood (10 mL) from the volunteers into sterile serum collection tubes (Red BD vacutainer with silica clot activator coating). The serum collection tubes were inverted several times and allowed to stand at room temperature for 30 minutes to allow clotting before centrifugation at 1,300 *g* for 10 minutes at room temperature, after which the serum supernatant was removed from the collection tubes for further centrifugation at 1,300 *g* for 5 minutes at 4^{o}C as recommended elsewhere (563). Serum supernatant (3 mL) was removed and added to 7 mL Bijou flasks.

Serum (1 mL) was sampled from the Bijou flasks into Eppendorf tubes (1.5 mL) containing EDTA (50 µL, 10 mM) to give a final EDTA concentration of 0.5 mM. The serum samples were centrifuged at 5,000*g* for 30 minutes (a necessary step to remove zymosan in subsequent samples and performed here for consistency). The supernatants were transferred to Eppendorf tubes (1.5 mL) and frozen with dry ice. The resulting time zero samples were stored at -80 ^{0}C.

Zymosan (25 mg) was suspended in PBS (1 mL) by shaking overnight in a 1.5 mL Eppendorf tube to ensure a fine suspension is achieved. The suspended zymosan solution (80 µL, 25 mg/mL) was added to the serum remaining in the Bijou flasks to give a total zymosan concentration of 1 mg/mL. The reaction mixtures were incubated at 37 ^{0}C with shaking at 190 RMP. After 1 hour and 4 hours, 1 mL of serum was sampled from the Bijou flask and prepared for storage using the same protocol described for the time zero samples. Serum samples were thawed on ice as recommended elsewhere (241) before use with ELISA for Complement activation products. The turbidimetric assays for C3, C4, CRP and IgG were performed by the Clinical Chemistry Laboratory using standard operating procedures.

6.3.3 Statistical Analysis

The measured and predicted phenotype distributions are presented as histograms with 5 bins (quintiles) and normalised by probability from which a number of distribution properties were derived. Firstly, the percentage difference between the median predicted and measured phenotype values of each analyte is reported as a test of relative model accuracy. Secondly, the overlap (intersection) of the distributions is reported as a percentage of the total histogram area as an initial test of distribution shape differences. Finally, all phenotype distributions were tested for normality using a Shapiro-Wilk test with a significance level of P=0.05 as described previously. Normal distributions were tested for equal variance using a two-sample F-test implemented using the ‘vartest2’ function in MATLAB. Normally distributed samples with equal variances were compared using a two tailed Student’s t-test using the ‘ttest2’ function in MATLAB assuming equal variances. Normally distributed samples with unequal variances were compared using a two tailed Welch’s t-test using the ‘ttest2’ function in MATLAB assuming unequal variances. Non-normal distributions were compared using a Kolmogorov–Smirnov test using the ‘kstest2’ function in MATLAB. All test decisions were made using a significance level of P=0.05.

The Complement activation model was constructed using differential equations and molar concentrations. For consistency, in the present study all assay results are also expressed in molar concentrations, rather than mass per unit volume as is commonplace for reporting protein concentrations in the literature. Correct conversion from mass to moles requires knowledge of the molecular mass of the protein used to calibrate each assay which is complicated by multimeric protein complexes. For example, the C1 complex is composed of a recognition subcomponent, C1q (460 kDa), and serine protease subcomponents, C1r (90 kDa) and C1s (80 kDa) in a 1:2:2 ratio (564). As there are several serine protease subunits per C1 complex, using the total molecular mass of C1 (800 kDa) in the conversion of mass to molar concentration would under-estimate the molar concentration of purified C1q standards, highlighting the importance of understanding the true molecular mass of the assay calibration standards. The molecular masses of analytes used in all calculations are shown in Table 6.1.

6.4 Results

A cohort of 22 individuals was recruited self-reporting no chronic inflammatory conditions. The CRP concentrations were recorded in all patient samples: 15 samples had CRP concentrations below the detection limit of the turbidimetric assay (1 mg/L) and the maximum recorded concentration was 4 mg/L. All CRP concentrations were below the upper healthy reference range limit of 10.5 mg/L (565). Consequently, no samples were rejected from the cohort due to chronic inflammation. The C3, C4 and IgG concentration distributions at time-zero are shown in Table 6.1 compared to the reference ranges from the literature.

*Table 6.1 Analyte molecular masses used in concentration calculations, reference ranges used for model predictions and standing concentrations of serum proteins measured in cohort serum samples (n=22). *In serum the basic unit of MBL may oligomerise to higher-order molecules of varying mass.*

Analyte	Molecular mass / kDa	Modelled concentration range / µM	Measured concentration median (range) / µM
CRP	125	1.6-84×10^{-3} (565)	8.0(8.0-32)×10^{-3}
C3	185 (566)	6.9-8.9 (249)	5.8(4.5-7.5)
C4	205 (566)	1.2-2.7 (567)	0.93(0.59-1.4)
IgG	150	43-90 (568)	67(47-92)
Ba	33	-	3.2(1.7-5.6)×10^{-2}
C4a	8.8	-	3.9(3.6-4.6)×10^{-3}
C5a	10.4	-	6.4(0.6-16)×10^{-4}
iC3b	176	-	8.0(3.6-14)×10^{-2}
TCC	1030	-	2.7(0.7-6.1)×10^{-4}
C1q	460 (566)	0.27-0.48 (569)	-
C2	93 (570)	0.12-0.38 (571)	-
C5	190 (566)	0.30-0.60 (572)	-
C6	105 (573)	0.51-0.69 (573)	-
C7	92 (574)	0.53-0.76 (574)	-
C8	151 (575)	0.33-0.53 (575)	-
C9	71 (576)	0.66-0.99 (576)	-
FB	93 (566)	0.80-3.1 (571)	-
FD	24 (566)	4.2-8.3×10^{-2} (577)	-
FI	88 (578)	0.23-0.45 (579)	-
FH	155 (580)	1.6-2.3 (579)	-
P	53 (581)	0.14-0.55 (582)	-
C4BP	540 (583)	0.28-0.56 (583)	
C1-INH	110 (584)	1.7-2.1 (249)	-
MBL	100* (585)	4.3-35×10^{-3} (586)	-
sCR1	200 (587)	1.1-2.2×10^{-4} (588)	-

The measured serum range for IgG was in good agreement with the reference range. The serum ranges for C3 and C4 varied by a factor 1.6 and 2.4 from minimum to maximum respectively, in good agreement with the modelled ranges of C3 and C4 which varied by a factor 1.3 and 2.3 from minimum to maximum respectively. The median C3 and C4 values measured in the cohort sera (5.8 µM and 0.93 µM) were 28% and 54% lower than the median values used for the optimisation of the C-ODE model (8.0 µM and 2.0 µM).

The Complement system response to zymosan was profiled by ELISA over 4 hours following activation of the fresh serum samples and is shown in Figure 6.4 compared with the pooled serum data used for model fitting.

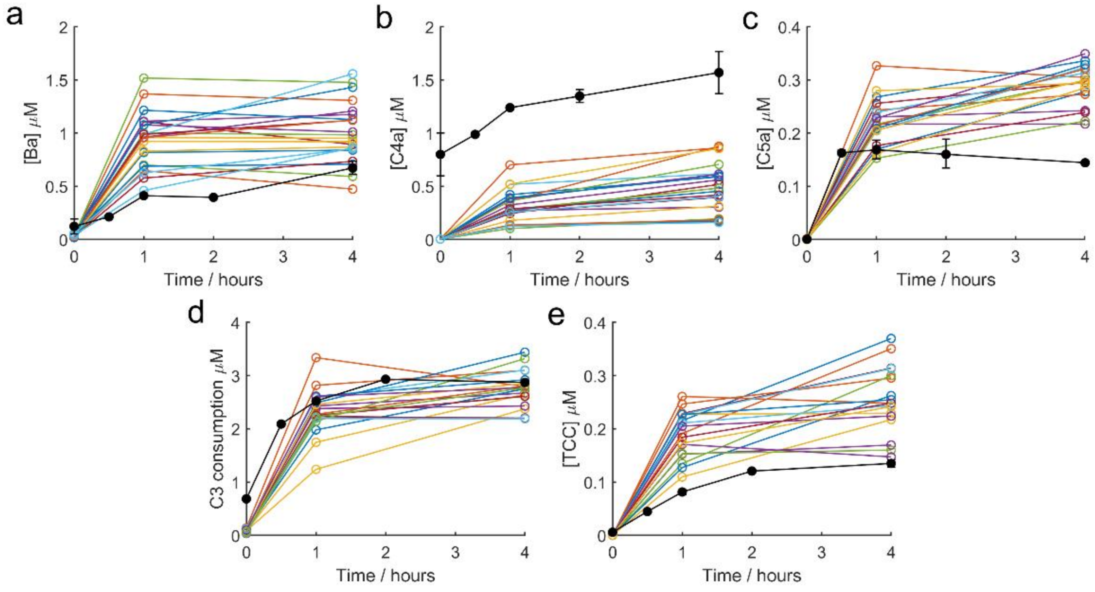


Figure 6.4 in vitro Complement activation in serum by 1 mg/mL zymosan. Z-test results from n=22 healthy volunteers (coloured) are compared to pooled human serum (black) profiled using 5 activation markers: a) Ba, b) C4a, c) C5a, d) C3 consumption and e) TCC. Error bars are 2SD from n=2 technical repeats. C3 consumption is assessed by comparing iC3b data from the cohort serum to C3a data from pooled human serum used to fit the C-ODE model (species which are produced in a 1:1 ratio with each molecule of cleaved C3).

All analytes showed an initial rise from their $t = 0$ concentrations and reached a plateau between 1 and 4 hours - similar concentration profiles to those observed for the pooled serum used to fit the C-ODE model in Chapter 5. The starting concentrations of most analytes were similar between pooled and cohort sera, although a significantly higher C4a concentration in the pooled serum suggests pre-analytical CP/LP activation. The mean rate of concentration increase for all analytes between 0 and 1 hours following activation (rate phenotype) was similar for pooled and cohort serum samples, although Ba and TCC rates appeared lower in the pooled serum. Furthermore, the total change in analyte concentration after 4 hours (capacity phenotype) appeared to be lower in the pooled serum for C5a and TCC.

Excluding CRP (which was largely below the detection limit of the assay) all other measured analytes were normally distributed (Shapiro-Wilk, $P > 0.05$) except C4a which was not normally distributed (P=0.002). The normal distributions measured in the experimental phenotypes support the choice of normal random sampling from the of constituent protein reference ranges to generate the predicted phenotype distributions from the C-ODE model.

The C-ODE model was unmodified throughout the present study and the parameters were not re-fitted to the phenotype data measured here. The medians and ranges of the rate and capacity phenotypes, both experimental and theoretical, are tabulated in in Table 6.2 and histograms of the distributions are shown in Figure 6.5.

Table 6.2 Comparison of phenotypes measured in the cohort serum (n=22) and predicted by the C-ODE model (n=250).

Species	Cohort Median Rate (range) / nM min^{-1}	Cohort Median Capacity (range) / µM	Predicted Median Rate (range) / nM min^{-1}	Predicted Median Capacity (range) / µM
Ba	15.4 (7.0-24.8)	1.01 (0.62-1.53)	4.25 (2.82-6.77)	0.526 (0.370-0.715)
C4a	4.59 (1.66-11.63)	0.480 (0.157-0.873)	13.9 (9.6-18.8)	0.833 (0.575-1.129)
C5a	2.97 (2.10-4.48)	0.243 (0.178-0.288)	2.17 (0.77-6.40)	0.197 (0.067-0.400)
iC3b	36.8 (19.2-53.3)	2.68 (2.14-3.38)	29.2 (20.1-42.2)	3.07 (2.33-3.89)
TCC	3.30 (1.82-4.34)	0.251 (0.160-0.370)	2.71 (0.77-6.40)	0.197 (0.067-0.400)

The rate and capacity phenotypes of Ba and iC3b were significantly larger by a factor of 3-10 than the phenotype of C4a from the CP/LP and C5a or TCC phenotypes of the TP. The response profile of iC3b displayed the fastest rate and largest capacity of the analytes measured. Further, the C4a rate phenotype displayed the largest range and hence patient variability, varying by ~7-fold. The distributions of rate and capacity phenotypes for the cohort serum data are shown for each analyte in Figure 6.5.

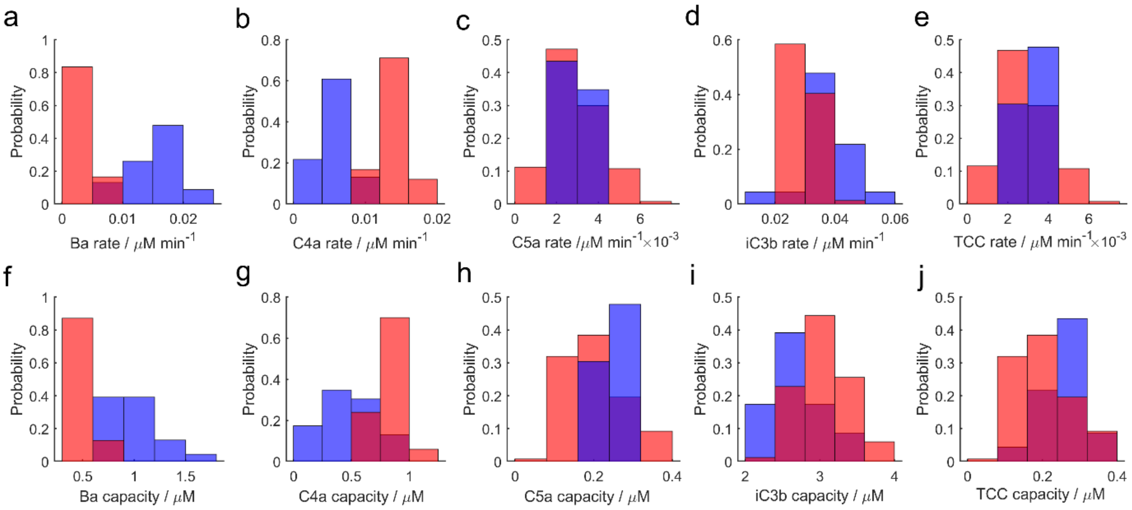


Figure 6.5 Distributions of Complement system biomarkers: a-e) rate phenotype distributions for Ba, C4a, C5a, iC3b and TCC and f-g) capacity phenotype distributions for the same analytes. Biological data from n=22 healthy volunteers (blue) are compared to n=250 simulation data (red).

The distribution comparison between the cohort phenotypes and the C-ODE predictions was tested using the percentage difference in the medians, the percentage histogram overlap and either a two tailed t-test or a KS test (depending on normality), Table 6.3. The median percentage difference between the phenotype medians of the prediction and the data was 21% and the median histogram intersection was 36%. The best agreement between measured and predicted phenotype distributions was for the rate phenotypes of C5a (KS test P=0.05) and TCC (KS test P=0.09). Notably, the iC3b rate and capacity phenotypes were also reasonably well predicted by the model with median percentage differences of -20% and 15% respectively.

Table 6.3 Results of statistical tests comparing rate and capacity phenotype distributions measured in the cohort serum (n=22) to the model predictions (n=250). Median percentage difference is difference between measured and predicted values as a percentage of the measured value. Histogram intersection is the overlapping histogram area as a percentage of the total histogram area. Distributions are tested for normality with the SW test (α=0.05) an appropriate agreement test is performed (t-test for normal distributions, KS test for non-normal distributions). The F-test determines if population variances are significantly different. P values > 0.05 indicate failure to reject the null hypotheses and are emboldened.

Analyte	Median percentage difference	Distribution percentage intersection	SW test P value	F-test P value	Agreement test	P value
Cohort Ba rate	-72	7	**0.84**		KS	3.9×10^{-19}
Predicted Ba rate			0.016			
Cohort C4a rate	202	7	**0.11**	4.1×10^{-2}	Welch's t-test	2.8×10^{-14}
Predicted C4a rate			**0.73**			
Cohort C5a rate	-9	70	**0.55**		KS	**0.053**
Predicted C5a rate			3.6×10^{-7}			
Cohort iC3b rate	-20	34	0.044		KS	5.5×10^{-9}
Predicted iC3b rate			0.013			
Cohort TCC rate	-28	51	**0.63**		KS	**0.090**
Predicted TCC rate			3.6×10^{-7}			
Cohort Ba capacity	-48	7	**0.37**	1.2×10^{-35}	Welch's t-test	2.7×10^{-8}
Predicted Ba capacity			**0.32**			
Cohort C4a capacity	74	23	**0.30**	7.1×10^{-8}	Welch's t-test	2.8×10^{-7}
Predicted C4a capacity			**0.73**			
Cohort C5a capacity	-20	39	**0.27**		KS	0.020
Predicted C5a capacity			8.9×10^{-6}			
Cohort iC3b capacity	15	38	**0.69**	**0.57**	Student's t-test	9.2×10^{-6}
Predicted iC3b capacity			**0.079**			
Cohort TCC capacity	-22	44	**0.71**		KS	8.2×10^{-4}
Predicted TCC capacity			8.90E-06			

All measured phenotype distributions appeared normally distributed as judged by the SW test with P>0.05, with the exception of iC3b rate which had a marginal P value of 0.044. In contrast, many predicted distributions from the model were not normally distributed according to the SW test, despite normal random sampling of starting protein concentrations for each simulation.

6.5 Discussion

The variation of two new Complement phenotypes was investigated for the first time: 1) the initial rate of activation marker increase over 1 hour (rate phenotype) and 2) the total change in activation marker concentration over 4 hours (capacity phenotype) following a standard zymosan challenge of serum - the Z-test. The experimental results from the cohort serum samples were used as an independent data set to test the accuracy and precision of the C-ODE model predictions of phenotype variation based on the concentration ranges of the 18 circulating proteins and zymogens in the adult population. The comparison of measured and predicted phenotype ranges tests whether the C-ODE model optimisation captured the concentration-dependence of each pathway in the cascade without re-optimisation or fitting to the new data.

6.5.1 Measured Phenotypes

The response profiles of 5 Complement activation marker proteins to the Z-test were successfully measured *in vitro,* under the ethics protocol, in the serum of 22 adult volunteers. The experimental data for the cohort in Table 6.2 point to significant variation in the C4a rate and capacity phenotypes which varied by 7-fold and 5-fold respectively. The other rate and capacity phenotypes in the cascade indicated smaller variations. The Ba and iC3b rate phenotypes of the AP varied equally in range by 3-fold but AP activation is tightly regulated via degradation of C3b (measured with the iC3b phenotype), a process which appeared to respond twice as fast as AP activation (measured with the Ba phenotype). The iC3b capacity phenotype was also twice that of Ba, further highlighting the extent to which C3b is regulated. By comparison, the median C4a rate phenotype of the CP/LP was only 25% that of Ba and 13% that of iC3b, showing similar kinetics to the lower cascade phenotypes of C5a and TCC. For the first time, the variations in differential flux across the cascade pathways have been quantified by the Z-test protocol in a cohort of healthy individuals.

To minimise pre-analytical activation of Complement in the Z-test protocol, samples were immediately mixed with EDTA to slow activation of all three initiation pathways during centrifugation before storage. Quantifiable Complement activation still occurs following addition of EDTA (264) and hence all samples were snap frozen and stored at -80^{o}C. To minimise Complement by sample thawing (241) all samples were thawed slowly on ice before testing with ELISA. The Z-test protocol in the present study used the same sample preparation methods as the study pooled serum data, used to train the C-ODE model in Chapter 5. Consequently, any pre-analytical Complement activation should have been systematically consistent between experiments. The C3 and C4 concentrations in of the cohort sera exhibited similar relative ranges but different absolute values to the reference ranges in Table 6.1. There are significant variations in reported reference ranges for C3 and C4 in serum which confounds the choice of a reference range: C3=660-1850 mg/L and C4=150-520 mg/L in ref (589); C3=750-1650 mg/L and C4=200-650 mg/L in ref (590); C3=800-2100 mg/L and C4=150-500 mg/L in ref (591); C3=830-1770 mg/L C3 and C4=150-450 mg/L C4 (592); C3=1280-1640 mg/L in ref (249) and C4=150-450 mg/L in ref (593). Ultimately, the C3 and C4 concentrations measured in our cohort C3=840-1380 mg/L (4.5-7.5 µM) and C4=120-290 mg/L (0.6-1.4 µM) were within the total ranges reported in the referenced studies, 660-2100 mg/L (3.6-11.5 µM) C3 and 150-650 mg/L (0.7-3.2 µM) C4, suggesting minimal pre-analytical Complement activation and accurate phenotype data.

Confidence in the value of a given protein concentration or phenotype is a function of assay error, sample dilution error and potentially variations in the mass and homogeneity of zymosan samples. A correlation analysis reveals that despite these potential sources of error, strong correlations were evident between protein concentrations and measured activation phenotypes, Figure 6.6.

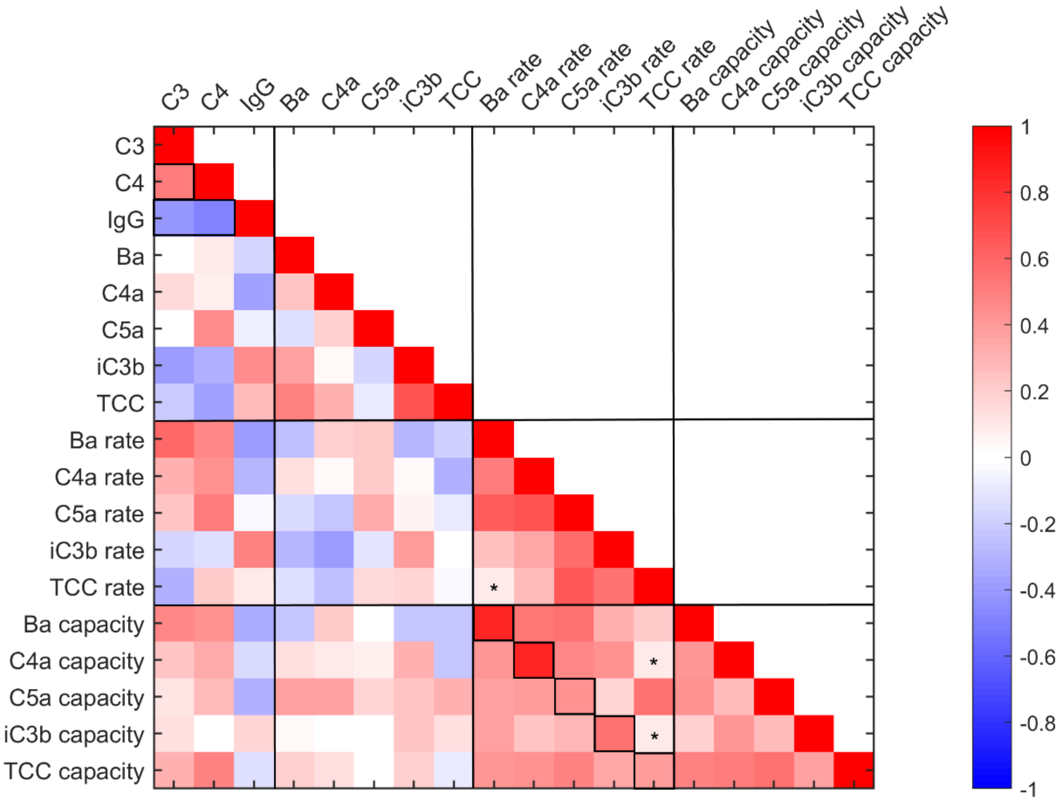


Figure 6.6 Heat map of correlation matrix showing Complement analyte standing concentrations and phenotypes (n=22). The plot is split according to parameter type (intact proteins, activation products, rate and capacity phenotypes). Boxes show notable high correlations and stars represent phenotypes with Pearson correlation coefficients <0.1.

The correlation analysis shows C3 and C4 concentrations were strongly correlated (Pearson correlation = 0.63) which is consistent with previous reports (594, 595) including a correlation with body mass index (596). IgG was found to be negatively correlated with C3 and C4 in the cohort serum (Pearson correlation = -0.45 and -0.49 respectively) which is also consistent with previous research (597). Correlations between observed protein concentrations might be considered to improve the accuracy of future models reliant on modelling protein concentration variations. Co-expression of genes could be considered also: the genes for C2, C4 and FB are coded for by genes on chromosome 6 (598) and the genes for C5 and C8 are located on chromosome 9 (595).

The correlation coefficients of Complement activation products appear to show some dependence on molecular weight, Figure 6.7. High molecular weight Complement activation products Ba (33 kDa) and iC3b (176 kDa) show a strong positive correlation with TCC (>1 MD), whilst C4a (8.7 kDa) and C5a (10 kDa) below the mass cut-off of the kidneys (30-50 kDa) (395) do not.

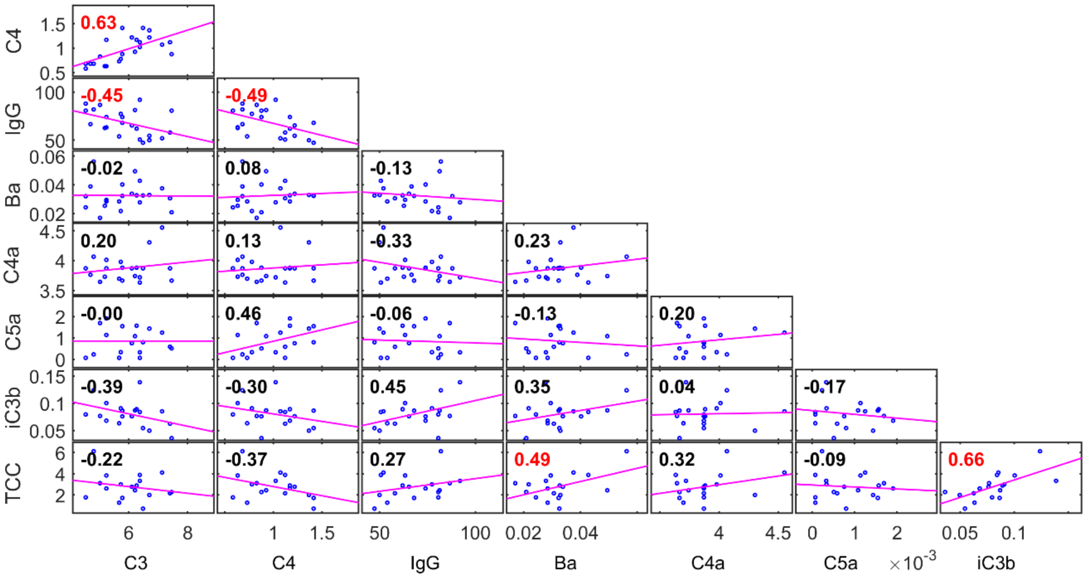


Figure 6.7 A matrix of plots showing correlations among pairs of analytes in µM (n=22). Lines of best fit are shown for each scatter plot and Pearson Coefficients displayed in red indicate significant correlation with P values <0.05.

The correlation between the standing concentrations of high molecular weight analytes with TCC highlights the importance of using larger biomarkers to observe Complement activation *in vivo*. Indeed, the C5a concentration measured in the plasma of healthy donors has been shown previously to be identical to septic patients (250), whereas all markers of Complement activation are seen to respond rapidly in the *in vitro* activations shown here, regardless of molecular weight.

Considering the consistency of these results with previous observations of serum protein correlations, the correlation between serum proteins and the novel Complement activation phenotypes can be investigated with confidence. All Complement activation phenotypes were positively correlated as expected given that all are measures of the strength of the Complement response to activation. Furthermore, a strong positive correlation was evident between the rate and capacity phenotype pairs for each analyte, Figure 6.8.

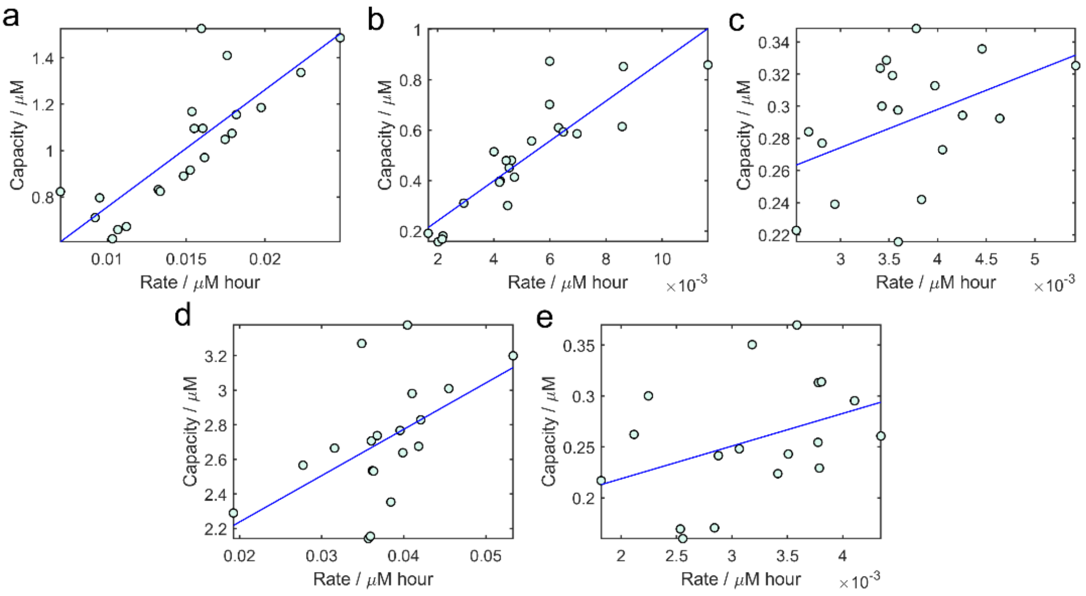


Figure 6.8 Linear regression analysis of the correlation between rate and capacity phenotypes in a cohort of n=22 healthy volunteers, measured for a number of activation products: a) Ba, b) C4a, C) C5a, d) iC3b and e) TCC. R^2=0.67, 0.77, 019, 0.28, 0.15 respectively. Pearson correlation coefficients=0.82, 0.87, 0.44, 0.53, 0.39, respectively with Ba, C4a and iC3b exhibiting P-values <0.05 indicating significant correlation.

The linear regression analysis implies that the capacity phenotype may be predictable based on initial rate: analytes Ba and C4a capacity phenotypes show the strongest dependence on initial rate (linear R^2=0.67 and 0.77 respectively). The TCC rate phenotype was least correlated with the other phenotypes (marked with stars in Figure 6.6), suggesting that a good estimate of Complement activation dynamics could be obtained by observing the minimal set of phenotypes: TCC rate, Ba rate, C4a capacity and iC3b capacity.

The standing concentrations of C3 and C4 were positively correlated with all capacity phenotypes as expected – an intuitive result considering the role of these proteins as a source of components for the C3 and C5 convertases. Interestingly, no single analyte was highly correlated with all the measured phenotypes. C3 and C4 showed low mean absolute correlations with all phenotypes with Pearson correlation coefficients = 0.22 and 0.32 respectively, indicating that individual protein concentrations are poor predictors of Complement activation potential. Other proteins such as those used for positive and negative regulation of the system are highly sensitive reactants in the system according to the sensitivity analysis in Chapter 5 and a full profile of the constituent proteins for each patient would be required to classify protein combinations that are import to system performance. Indeed, only with a validated mathematical model and all starting protein concentrations can Complement function in serum be precisely predicted.

6.5.2 Model Predictions

A comparison of measured and predicted phenotypes, based on a measured set of starting protein concentrations, would be the ultimate model validation test but requires assays for all Complement proteins in all serum samples. In this study, the predictive accuracy of the model is tested not on the reproducibility of fitting data but on the variation of novel system properties which are the patient response phenotypes. The C-ODE model was used to predict the distributions of phenotypes based on varying the concentration combinations of the 18 circulating proteins and zymogens that make up the cascade model.

The complete profiling of the 18 constitutive proteins is a technical challenge and the C-ODE model was optimised previously based on published median serum values rather than measured concentrations. Optimisation based on median concentrations was a reasonable assumption for pooled serum but any errors between the concentrations used for model optimisation and their true values will have induced bias in the fitted rate constants. Considering the potential for model bias, the protein ranges used in the phenotype predictions were not matched to the C3 and C4 concentrations measured in the cohort serum. Because the model fits well to pooled serum data, an accurate model prediction for the mean phenotype values (of proteins in the training data set) would only occur if the true mean starting protein concentrations in the training and cohort sera were similar. If the true mean starting protein concentrations in the training and cohort sera were different, the C-ODE model predictions will fail, regardless of differences in the protein concentration values used in the model. Matching the numerical values of C3 and C4 concentrations in the model to those measured in the cohort sera would risk producing a good median phenotype prediction due to chance. The validity of the model was therefore tested by comparing the variation in phenotype values which indicated whether the model reproduced the sensitivity of the Complement system to relative rather than absolute protein concentration variations. Table 6.1 shows that the relative variation in C3 and C4 concentrations measured in the cohort sera were in good agreement with the modelled variation. Furthermore, the consumption of C3 shown in Figure 6.4 was similar for both pooled and cohort sera (2.5-3 µM as measured by C3a and iC3b respectively), suggesting the C3 activation dynamics of the pooled serum were largely representative of the cohort serum for total cascade activity.

A quantitative analysis comparing the model predictions to the cohort data (Table 6.3 and Figure 6.5) revealed that the median percentage difference between measured and predicted phenotypes values was 21%, confirming that the pooled serum used for model training was reasonably representative of the cohort sera studied here. The measured and predicted phenotype distributions had a median histogram overlap of 32% for all phenotypes measured, suggesting reasonable precision considering high uncertainty in both the model parameters and model sensitivity to protein concentrations. Furthermore, iC3b was not a species present in the training data set, yet it's rate and capacity phenotypes were reasonably well predicted by the C-ODE model. These results suggest that the data used for model training, particularly C3a, was somewhat effective in capturing the dynamics of other model species.

Many of the predicted phenotype distributions exhibited SW test P values <0.05 (evidence against normality) whereas all measured cohort phenotypes appeared normally distributed. Evidence for non-normality may be due to the large sample size of the simulation distribution which could allow trivial differences from normality to have been perceived as significant by the SW test. To investigate the bias associated with the sample size of the modelled phenotypes, all statistical tests were performed for different numbers of model simulations. The 250 values predicted by the model were sampled uniformly at random, with replacement, and compared to the experimental data for each phenotype. The procedure was repeated 50 times for each number of observations to determine the range and median P values, Figure 6.9.

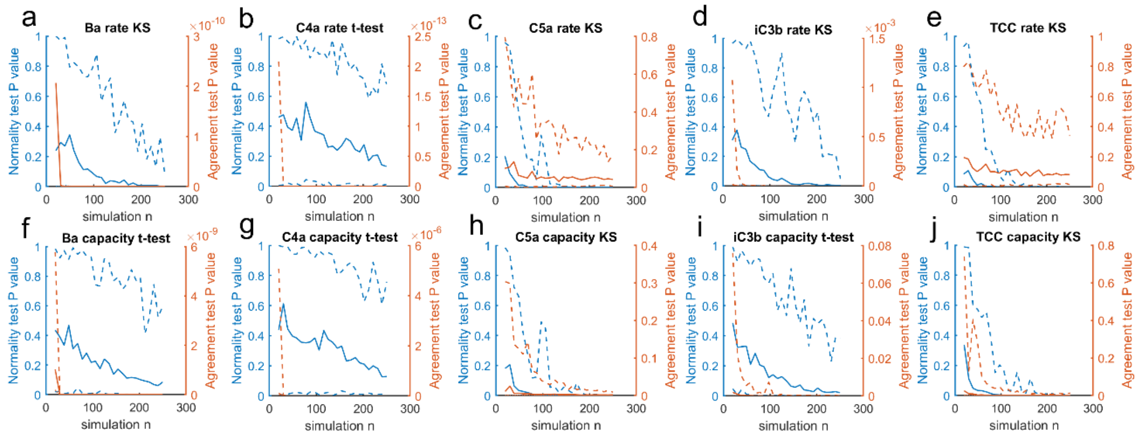


Figure 6.9 Effect of modelled phenotype sample size on the median (solid lines) and range (dotted lines) of P values from the SW test for normality (blue) and the t-test or KS comparison tests (orange): a-e) rate phenotype distributions for Ba, C4a, C5a, iC3b and TCC; and f-g) capacity phenotype distributions for the same analytes.

As expected, the median P value of the SW test for normality invariably decreased with sample size such that at n=20 all phenotype distributions appeared normal with P>0.05. The modelled phenotype distributions for Ba rate, Ba capacity, C4a capacity and iC3b capacity also appeared to be normally distributed at n=250 and as such were compared to their experimental distributions with a two-sample t-test. The phenotype distributions which for which there was evidence against normality at n=250 were compared to their experimental distributions with the KS test.

The P values from the t-test and KS tests exhibited a larger range at n=20 as expected due to the effects of randomness on small samples. As such, the C5a capacity, iC3b capacity and TCC capacity phenotypes occasionally exhibited P values >0.05 at n=20. However, the median P values of the comparison tests did not vary sufficiently with n to alter the conclusions drawn from testing at n=250 in Table 6.2: the C5a rate (Figure 6.9c) and TCC rate (Figure 6.9e) phenotypes were the only phenotype predictions with median P>0.05 at all simulation sample sizes tested here.

A further source of bias in the assessment of the C-ODE model’s ability to predict the Complement activation phenotypes was the choice of concentration ranges for modelling predictions, which show significant variation in the literature as discussed previously. To investigate the effect of starting protein concentration ranges on the C-ODE model predictions, the simulation was performed on 3 variations of the reference ranges listed in Table 6.1. To capture the effect of using larger reference ranges, the limits of the concentration sampling ranges in the model were extended by 20% of the median value. To capture the effect of using systematically different reference ranges the concentration sampling ranges in the model were translated by -20% and +20% of the median value. The minimum and maximum phenotype values, percentage difference of medians and percentage histogram overlap was recorded for all variations and is shown in Figure 6.10.

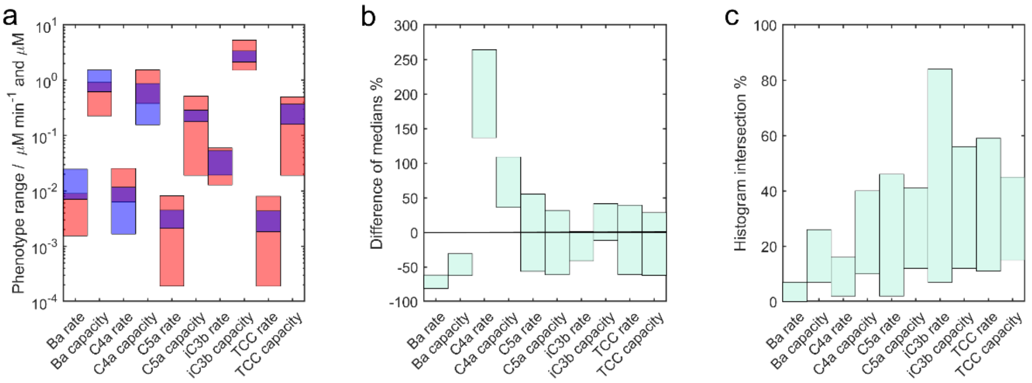


Figure 6.10 Variation of C-ODE model phenotype predictions (n=25) and their agreement with the cohort data (n=22) with varying concentration sampling concentration ranges. Ranges were extended by 20% of the median, translated by -20% and +20% of the median and the most extreme values are shown here: a) phenotype ranges recorded from model predictions (red) and experimental data ranges (blue); b) percentage differences between predicted and recorded phenotype median values and c) histogram intersection between predicted and recorded phenotype distributions.

With different concentration sampling ranges, the ranges of measured C5a, iC3b and TCC phenotypes were fully within the predicted ranges. Furthermore, the median phenotype values predicted for C5a, iC3b and TCC appeared robust and remained accurate to within 60% of the cohort data regardless of the reference range used for the simulation. The accuracy of the C-ODE model predictions for the iC3b, C5a and TCC phenotypes (common to all initiation pathways) indicates that model training using Ba, C3a, C4a, C5a and TCC concentration data from the pooled serum was effective for capturing total Complement activity. The KS test revealed no significant difference between measured and predicted distributions of the C5a and TCC rate phenotypes (P=0.05 and P=0.09 respectively) using the sampling ranges in Table 6.1.

The measured ranges of Ba and C4a phenotypes remained only partially within the predicted ranges, regardless of the sampling ranges used for the simulation and the poor agreement is attributed to the differences observed between pooled serum and cohort serum for Ba and C4a despite similar C3 consumption. Figure 6.5 shows that the AP (profiled using Ba) was more active in the cohort serum than predicted by the C-ODE model, whilst the CP/LP was less active than predicted. Consistency between pooled and cohort data for late cascade markers indicates that regardless of the relative contributions of the AP and CP/LP, zymosan activation of Complement yielded consistent flux through the lower cascade. In future work, the C-ODE model should be re-optimised using data from serum where all concentrations are known and preferably using fresh serum due to differences is relative CP/LP and AP dynamics. The resulting model might yield precise predictions of Complement activation using specified starting concentrations and should be validated using pairwise comparisons of data and model prediction for each starting condition.

Model validation is fundamentally performed to test whether a model is accurate enough to fulfil an intended purpose and a major goal of systems biology modelling is to develop a model that has sufficient accuracy to predict the response of a system to perturbation. The accuracy with which the C-ODE model predicted the distributions of iC3b, C5a and TCC phenotypes suggests that the model might also be used to predict the effect size of Complement depletion on these phenotypes. An optimal cut-off value of 0.578 mg/mL C3 has been identified as a good predictor of mortality in sepsis patients despite no correlation with APACHE II or SOFA scores (153) and corresponds to a consumption of 30-70% from the C3 reference ranges discussed previously. 1 hour of 1 mg/mL zymosan activation was simulated by the C-ODE model, resulting in 31 ± 6 % C3 depletion (2SD), to generate depleted starting conditions. The phenotype predictions from depleted starting conditions were compared to those of the healthy starting conditions to estimate the sensitivity of the phenotypes to pre-Z-test Complement depletion, Figure 6.11. The model predictions suggest the C4a phenotypes are most sensitive to Complement depletion compared to the other activation markers.

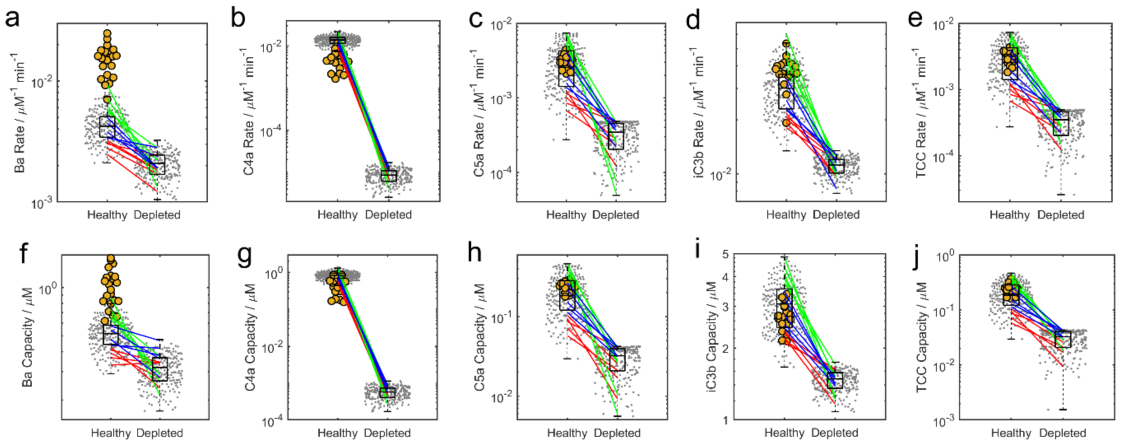


Figure 6.11 Comparison of Complement activation phenotypes for cascade activation markers Ba, C4a, C5a, iC3b and TCC in healthy and Complement-depleted sera: a-c) rate phenotype and f-h) capacity phenotype. Box and whisker plots are shown for the n=250 model simulation values (grey) and suggest all phenotypes are sensitive to Complement depletion (pre-test deficit of 31 ± 6 % C3). The data from n=22 healthy volunteers (orange) are shown for comparison. Selected simulations (n=5) with healthy phenotype values from above the upper quartile (green), below the lower quartile (red) and the interquartile range (blue) are tracked between healthy and depleted states showing no correlation.

The position of each simulated phenotype value within the healthy and depleted phenotype value distributions was tracked and reveals that there is no correlation between healthy phenotype value and depleted phenotype values. The result suggests individuals may lose Complement activation potential at different rates, independent of their healthy phenotype value. The phenotype trajectories between healthy and depleted states suggest that some cascades make rapid transitions from super-median to sub-median phenotype values, which in principle would be predictable with a full cascade profile of all constituent proteins in a patient. Correlation between any of the 18 starting protein concentrations and the gradient between healthy and depleted phenotypes could be investigated in further studies to identify combinations of proteins which are most sensitive to Complement depletion.

The simulation demonstrates the utility of the model in predicting patient response phenotypes and enables hypothesis-driven selection of biomarkers to test in future studies. The logical next step for investigations into these phenotypes is to collect data to validate the model prediction for the phenotypes in Complement-compromised patients with a pilot study. The sample sizes required to assure adequate power to detect statistical significance can be estimated using Complement activation models such as the C-ODE model discussed here.

6.6 Conclusions

The present study recorded the variation of 10 phenotypes of immune state in the healthy adult population and provided an independent data set with which to test the accuracy of C-ODE model. The model correctly predicted the variation for some phenotypes but the cohort data show that the pooled serum may have pre-analytical activation in CP/LP with a consequent bias in the model training which was observed in the validation process for Ba and C4a phenotypes. The accuracy of the analyte measurements might be improved by performing concentration analysis immediately after sampling with a rapid assay platform to minimise pre-analytical Complement activation. In time, the phenotypes presented here might be measured using the LiScAR biosensor platform at point-of-care. The correlations between standing protein concentrations in the cohort serum reveal that the model predictions might be further improved by considering co-expression of proteins when simulating starting protein concentrations.

The difference in phenotype values between the data for healthy individuals and the simulation of Complement depleted individuals was significant and highlights the utility of the model as a tool for power calculations in the design of future Complement activation studies. The C4a rate phenotype showed the largest variation in the cohort, varying by ~7-fold and is predicted to be the most sensitive phenotype to Complement depletion studied here. Identification of patients at high-risk of infection based on C3 and C4 concentrations may not be adequate as evidenced by the large reference ranges and lack of correlation between protein concentrations and patient phenotypes. The phenotypes are direct measures of Complement activation potential and are better suited to predicting patient response to infection.

7 Conclusions and Future Work

This thesis began with the impact of nosocomial infections, particularly SSI, on the outcome of elective surgical patients and highlighted research showing that timely identification high-risk patients is beneficial but challenging (58). The use of ‘risk-biomarkers’ or ‘phenotypes’ measured pre- and/or post-operatively was identified as one method which could help healthcare professionals to stratify patients into high and low-risk groups and guide the allocation of additional peri-operative care resources to high-risk patients and improve outcome (34). Several risk biomarkers are already in use, including CRP and WBC to gauge the immune states of patients (42). This thesis proposed that Complement activation phenotypes could provide a new source of risk-biomarkers and could be measured by activating Complement in patient serum *in vitro* and profiling the system response. Patients with Complement activation phenotypes, such as initial rate or absolute change (capacity) of analyte concentration, at the extremes of the distribution expected in healthy adults might be at higher risk of HAI and could be stratified into an appropriate care pathway based on this rapid assessment of immune state. Repeating the phenotype measurements post-operatively might also be used to detect the onset of infection and improve patient outcome via timely therapeutic intervention.

7.1 Summary of Research Findings

Existing assays of Complement activation products, such as ELISA, are confined to the laboratory and require extensive training to operate. Therefore, it was recognised that a biosensor platform capable of rapidly measuring Complement activation products in serum samples with relative ease would be beneficial to the study of Complement phenotypes in a research or clinical setting. The LiScAR biosensor platform, developed by the Shaw Lab, can perform rapid and multiplex measurements of protein-sensor interactions and its performance was reviewed in Chapter 2. Automated signal processing routines were developed to standardise the calibration of the instrument as a protein assay and ultimately improve the ease of data collection. Chapter 3 reported the development and optimisation of LiScAR assays for C3, C5a, TCC and TNFα which could be calibrated simultaneously as a 4-plex assay with purified protein samples. The new LiScAR assays for C3 and C5a were validated with ELISA and turbidimetric assays and can be performed simultaneously and repeatedly in serum samples, with a duty cycle of 15 minutes and a CV of 23% and 21% respectively, within a common dynamic range of 3.2-12.5 nM. It may be concluded that the LiScAR is a good platform for the development of assays required to track the progress of reactions with rapid and repeated measurements.

A novel LiScAR assay for the monomeric purity of IgG samples was developed to accompany the new assays for quality control of assay materials. Both the LiScAR method and ESI-MS revealed that the commercial antibodies tested varied significantly in monomeric purity, with some as low as 54% intact. It may be concluded that careful quality control screening of antibody samples is required before use in reproducible research and the new monomeric purity assay provides a convenient new test for this purpose. The LiScAR assays could now be used to reliably measure *in vitro* Complement activation phenotypes in patient serum samples with a less complex and faster method than ELISA – a step towards Complement testing at point of care. The antigen array presented in Chapter 3 might also be used to screen for the presence of antibodies in serum against specific antigens to test the efficacy of vaccines, whilst the PAG-based IgG assay in Chapter 2 could quantify total IgG concentration. All assays may be included on a single multiplex array according to the research needs.

An alternate approach to profiling *in vitro* Complement activation phenotypes with multiple measurements on multiplex assays is to predict the Complement response profile of a patient with a mathematical model, considering the starting protein concentrations measured in a single serum sample. Chapter 5 reported the development of a new mechanistic ODE model of Complement cascade activation, C-ODE. The C-ODE model is the most recent of 7 Complement models and is the only model optimised using biological data from all pathways of the system. The C-ODE model was optimised using the model fitting techniques described in Chapter 4 and ELISA data of Complement activation products in pooled human serum following zymosan activation. The C-ODE model was first verified by simulating the effect of FH deficiency and P supplementation on Complement activation, with results that were consistent with expectations of system behaviour in each condition. The mechanistic C-ODE model demonstrated its utility as a research tool by predicting significant competition for substrate among the initiation pathways in the FH-deficient and P-supplemented simulations – system behaviour that could not be quantitatively predicted with a diagrammatic model alone. These simulations may benefit future investigations into Complement modulation therapy for drug design and the model is available as a COPASI file that can be shared among the research community.

Following verification, the C-ODE model was used to estimate the variation in Complement activation phenotypes in the healthy population. The predictions considered the reference ranges of all 18 constituent protein concentrations in the model and results were compared with serum data from a cohort of 22 healthy volunteers in Chapter 6. The predicted distributions of C5a and TCC rate phenotypes were validated by observing no statistically significant difference from the experimental data, despite the challenges associated with model transfer to new system settings highlighted in Chapter 4. These challenges included increased pre-analytical CP/LP activity in the pooled serum data set as evidenced by a higher initial C4a concentration. It may be concluded that the mechanistic approach to modelling is useful for predicting system behaviour given sufficient training data and refinement of such a model is possible with iterative fitting to new experimental data as it becomes available. Ultimately, a limited experimental dataset and a fully mechanistic mathematical model allowed the complex behaviour of the Complement system to be studied, even the behaviour of species that were not in the training data set. The same approaches described here, such as the assumptions of median enzyme kinetics as a starting point for model optimisation, could be usefully applied to other systems such as coagulation – a possible source of new phenotypes to predict infection.

The cohort serum data revealed that the standing concentrations of individual proteins were not correlated with any of the measured rate and capacity phenotypes, highlighting the necessity for a direct test of Complement activation potential such as the standard Z-test described here. The cohort phenotype data provide a first estimate of the variation in Complement activation in healthy adults – a benchmark against which the serum of high-risk patients could now be judged in future studies. The model was used to predict the effect size of Complement depletion on the phenotype distributions, which supported the hypothesis that these tests would be sensitive measures of Complement depletion in patients. The model also predicts that there is no correlation between healthy and Complement-depleted phenotype values in terms of position in the simulated distributions, suggesting highly personal transitions between these states. It may be concluded that Complement activation is a highly personal measure of immune state that should be investigated further for predicting patient risk of infection and detecting the onset of infection. The LiScAR shows potential to perform these tests at the bedside.

7.2 Strengths and Weaknesses

The C-ODE model was initially constructed using a bottom-up approach (599) by combining kinetic parameters measured in isolation in an established, complex reaction structure to predict systemic behaviour. This approach is distinct from top-down systems biology which looks for correlations between system behaviour and proteins or genes, in order to understand the effects of input on output (600). Top-down models often use ‘black-box’ models, such as artificial neural networks (601), which do not require mechanistic understanding of the system (602). Black-box models, are therefore an abstraction of the underlying physical processes of the system and can be optimised to fit well to experimental data and make predictions using a minimal set of parameters (602). However, their lack of physical interpretation is also the main disadvantage of black-box models, which often cannot investigate specific physical perturbations on the system without re-training – the benefit of mechanistic kinetic models such as C-ODE. For example, the C-ODE model can be used to predict the effect of inhibitors at any reaction in the system and this is not possible using non-mechanistic models.

Kinetic models are typically constructed for relatively small systems and of the 462 models in the 2014 Biomodels database (603), the median number of reactions and species was 12 and 11 respectively (604). The C-ODE model is therefore significantly larger than most kinetic models and many of the 187 parameters (concentrations and rate constants) were unknown, representing a significant optimisation challenge. The assumptions of median dissociation constants and starting concentrations allowed the model to be optimised by fitting 56 unknown association rate constants and turnover numbers – a reasonable reduction in the dimensionality of the optimisation problem. However, the individual fitted rate constants of the fitted C-ODE model were poorly defined, in part due to a lack of experimental data and could not be used to construct accurate subsystem models containing the same reactions in an alternate structure - a model of the AP only for example.

In contrast, the Silicon Cell approach to modelling (464) aims to produce highly transferable models combining smaller models constructed based on experimentally determined structures and rate constants studied in isolation. Unlike the optimised C-ODE model, Silicon Cell models are not fit to datasets of system behaviour and any differences between model predictions and experimental data are instead used to identify new effectors of the system (464). As a result, Silicon Cell models do not necessarily describe systemic behaviour accurately but ensure all model parameters can be transferred between systems which is useful when simulating modular synthetic metabolic pathways. A Silicon Cell approach to modelling the complete Complement system could be considered in future work as more kinetic parameters become available.

Metabolic networks containing easily isolated enzymes are ideal for studying using the Silicon Cell approach and a model of yeast glycolysis based entirely on *in vitro* data from the isolated enzymes has been used to describe *in vivo* data, albeit with additional system branches (482). Notably however, Silicon Cell model parameters are limited to models of similar system settings, as discussed in Chapter 4, and hence the conditions of study should be carefully considered to allow parameter transfer to future systems. Future work could study the isolated pathways of the Complement cascade, as discussed in Chapter 1, to provide missing rate constants in the C-ODE model. A protocol for such an experiment is presented in at the end of this chapter.

“All models are wrong, but some are useful” is a common expression in the modelling community. The phrase captures the idea that no model can fully capture the complexity of physical phenomena and a model should be judged on its ability to fulfil a purpose. In the present study, the C-ODE model was developed to predict the concentration-time profiles of all species in the Complement system following activation by zymosan *in vitro*. The C-ODE model therefore allowed the hypothesis driven selection of candidate biomarkers that might predict patient risk of infection. The model predictions of C5a and TCC rate phenotype variation were somewhat validated in Chapter 6 and further refinement of the model is required to improve predictions of the other phenotypes measured. Because the C-ODE model is mechanistic and contains all species in the system, refitting would be relatively straightforward if new time course data became available.

Many of the Complement activation products observed with the pooled human serum time course study in Chapters 5 and 6 were produced rapidly following activation. Accurate measures of reaction rates were therefore challenging to derive with assays such as ELISA which require samples to be stored before testing – introducing sources of error as described in Chapter 6. These sources of error might be overcome with real-time measurements of Complement activation. The present study demonstrated automated analysis of data from 4 LiScAR assays performed simultaneously on a single sample in 15 minutes – eliminating the need for sample storage. Operating the LiScAR is also relatively simple and enables significantly faster data collection compared to four ELISA's or the multiplex Complement assays offered by Merck's Luminex platform. The LiScAR could now be used to collect data to refine the C-ODE model, particularly in early time, although a number of improvements to the technology can be identified and are discussed here.

Firstly, to be adopted as a commercial instrument, the LiScAR instrument would need to be redesigned to pass the European Union health and environmental safety requirements (CE marking) required for all *in vitro* diagnostic devices sold in Europe. This would be a significant design challenge considering the microfluidic requirements of the instrument. A disadvantage of the LiScAR Complement assays developed here is the volume and concentration of detection antibodies required - which lead to a higher cost of each test compared to ELISA. The C5a assay, for example, used 450µL of 200 nM detection antibody to ensure a detection step of 200 seconds occured without dilution of the antibody in the 1 mL sample loop at the 6 mL hour^{-1} (1.7 µL s^{-1}) flow rate used in the LiScAR setup. At the time of writing, the detection antibody retails at £576 per 200 µL (1.07 mg mL^{-1}) - enough for 13 tests at a resulting cost of £44 per test or £88 for an estimate of either rate or capacity phenotype, discussed in Chapter 6. For comparison, the 96-well C5a ELISA used in Chapter 6 retails at £500 and can perform 80 tests (losing 2 columns to duplicate standard curve measurements and controls/blanks) at a resulting cost of £6 per test. Future work could minimise assay costs by re-designing the LiScAR platform to reduce the sample volumes required. This could be done by minimising the volume of the flow cell and the diameter of all tubing.

A further limitation of the LiScAR assay platform is the complex fabrication process of the assay surfaces, requiring tight control over a number of processes which impact assay sensitivity. A common issue encountered was variability in the synthesis of the gold nanoparticle seeds due to the nature of the method itself which was performed by hand. Cooled $NaBH_4$ was injected via syringe into the auric chloride / sodium citrate solution and the reaction proceeded rapidly. Thorough mixing was required to ensure a uniform population of seed nanoparticles was produced but mixing must not be so vigorous as to cause aggregation of the nanoparticles in the colloid – referred to colloquially as “crashing out” and observed as black / purple flakes in the solution. The use of magnetic bars to stir the mixture was attempted to reduce experimental variation but appeared to encourage nanoparticle aggregation. The seeds solution is stable at 4 ^{o}C and a well-performing stock of nanoparticle seeds was used for all experiments in the later stages of this project to minimise variations in seeds synthesis. In time, synthesis of the gold nanoparticle seeds could be automated with a liquid handling device for the highest possible reproducibility.

A further issue in the array fabrication process is variability in substrate surfaces which, as shown in Chapter 2, significantly impacts the growth and RI sensitivity of the printed gold nanoparticles. As such, the protocol for printing and growing the seed nanoparticles required re-optimisation with each new batch of substrates from the same manufacturer. The accuracy of printing also varied significantly between experiments. The printing technology is outside the scope of this thesis but is highly sensitive to sample viscosity and required extensive re-optimisation (target practice) for each protein or antibody sample - involving significant waste of antibody samples and substrates. Both substrate quality control and printing accuracy are engineering challenges that should be addressed before large-scale manufacture of the LiScAR assay surfaces but careful experimental design allowed the complex fabrication process of the LiScAR assay surfaces to be performed with good reproducibility and minimal waste. In the current project, nanoparticle arrays on new batches of substrate were tested for sensitivity (as judged by the metrics RIS and nRIS) before tethering of proteins to prevent wasted batches. Other properties of the substrate might also be tested ahead of nanoparticle printing, such as hydrophobicity, to quickly estimate the quality of substrate batches without the need to test nanoparticle arrays. Printing accuracy variations can be minimised by diluting samples in a common diluent such as PBS.

The Bland-Altman analysis in Chapter 3 revealed errors in the calibration of the C5a, TCC and C3 assays. The bias associated with the C3 assay was 55 ± 21 % and not systematic but visibly worsens at greater analyte concentrations. These errors could be accounted for in future assays to improve assay accuracy relative to the gold standard (at least for the analyte concentration range and conditions tested) but ultimately must be addressed. Appendix 1 shows the results of testing an alternative assay for C3, using monoclonal rather than polyclonal detection antibody, although the use of a neo-epitope capture antibody is preferable.

On average, the C5a assay systematically under-estimated by 26 ± 27 % and the TCC assay systematically over-estimated by 36 ± 39 %. The systematic calibration errors of all assays may be improved with the use of common protein standards of precisely known purity and concentration, such as the IgG reference material from NIST used to calibrate the monomeric purity assay in Chapter 3. The Bland-Altman analysis in the present study allows systematic calibration errors to be corrected for, providing a better agreement with the Gold Standard tests as demonstrated for the C5a assay later in this chapter.

A visual representation of all factors influencing measurement uncertainty is shown in Figure 7.1. To maximise reliability of measured Complement activation phenotypes, additional experiments might be performed in future work to understand the effects of factors not investigated here.

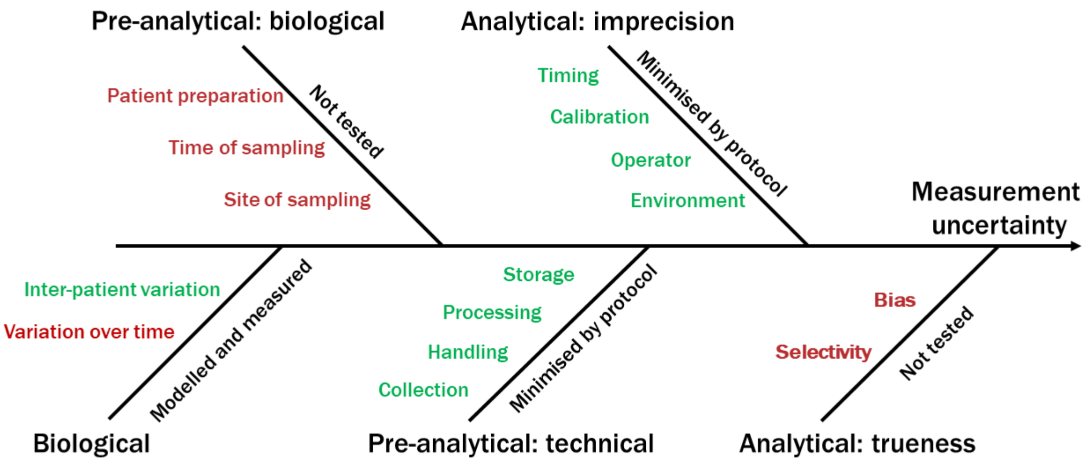


Figure 7.1 Feather diagram of the contributing factors for Complement phenotype measurement uncertainty, adapted from ref (605). Factors in green are controlled or measured in the present study, whereas factors in red should be investigated further.

Biological uncertainty stemming from variation among individuals was explicitly investigated in the present study of Complement activation phenotypes, using both model predictions and data from a cohort of healthy individuals. However, diurnal or circadian variations in plasma protein concentrations are a concern for the reliability of measurements and are observed for several proteins (e.g. bilirubin) (606) and hormones (e.g. cortisol) (607). Circadian variations in Complemented-mediated solubilisation of immune complexes have been reported in rheumatoid arthritis patients, suggesting blood could be sampled at the same time of day for increased reliability of Complement testing (608). However, the influence of diurnal variation on Complement activation has been studied using C3dg, showing no significant variation in plasma samples (609). Variation for an individual over time could be investigated in future work to understand the lifetime over which Complement phenotype test results are representative of immune state and inform the ideal measurement frequency during recovery. The model predictions in Chapter 6 predict that the studied activation phenotypes could change rapidly during infection, suggesting that these could be used to track condition severity and response to emerging therapeutic strategies such as supplementation with P (534) or C3 (228) – type 1 and type 2 biomarkers.

Analytical and pre-analytical factors were not specifically investigated in the present study of Complement activation phenotypes but were minimised with the use of commercial serum collection tubes and ELISA assay kits, as well as a consistent and established sample preparation protocol from ref (152). Many procedures in the experimental protocol were time dependent, meaning that results had the potential to be affected by variations in incubation and sampling times. In the Complement activation phenotypes study, samples were collected over a period of several weeks and labelled in chronological order. Sample aliquots were also defrosted and diluted in preparation for analysis in chronological order, which presents a potential source of non-systematic error. However, the concentrations of Complement activation markers showed no dependence on sample number, as judged by the data at 1 hour following the zymosan challenge in Figure 7.2. The samples measured at 1 hour following the zymosan challenge were used to measure the rate of analyte increase, making these data the most sensitive to variations in analytical timing. It may be concluded that the method of serum sample preparation described here had minimal effect on the phenotype results.

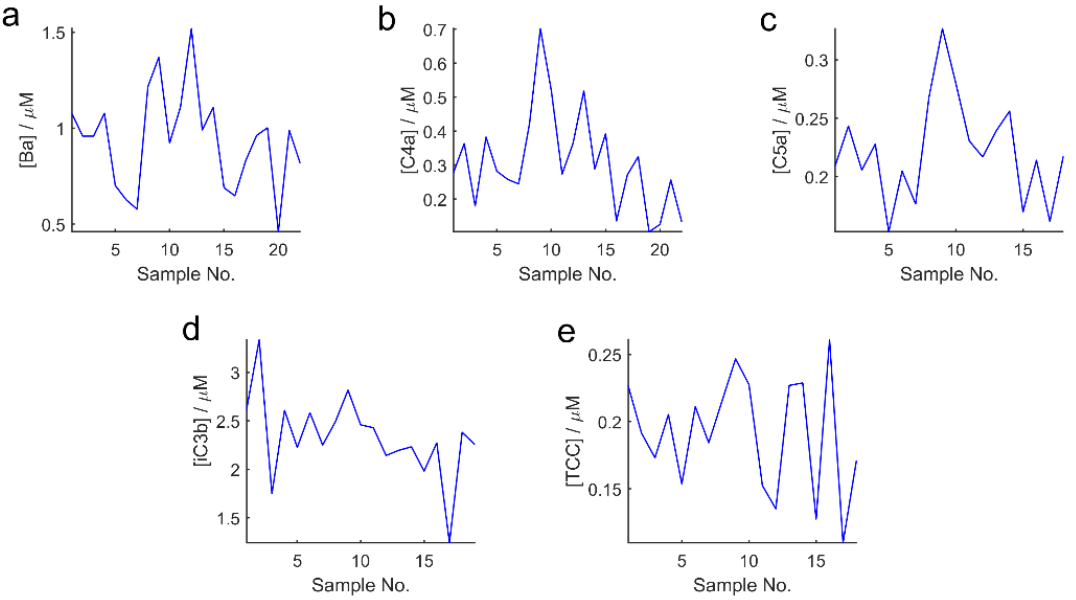


Figure 7.2 Dependence of analyte concentration (at 1 hour following zymosan challenge) on sample number. Sample number represents multiple effects including the age of experimental materials such as zymosan and the inverse of defrost time before dilution and analysis. No common trends are observed, indicating that the assay procedure did not influence the relative results.

The choice of sampling time was a source of bias in deriving the Complement activation phenotypes. The protocol shown in Chapter 6 used the same sampling times for all analytes but the model indicates that more accurate measures of rate and capacity could be achieved with different sampling times for each analyte, due to different times to maximum. For example, Figure 6.3 shows that sampling at 0, 1 and 4 hours provided a good estimate of initial rate and total change in C5a and TCC but C4a reached a maximum before 1 hour, whilst Ba and iC3b reached their maxima after 4 hours. Using different sampling times for each analyte would complicate the method of data collection for the rate and capacity phenotypes, although the correlation between these phenotypes in Figure 6.8 suggests the rate phenotypes alone may be a sufficient measure of Complement activity. Automated sampling and a rapid assay such as the LiScAR could measure different activation products at different times, from a single sample, following activation in real-time. Lack of sample preparation and storage would significantly reduce experimental complexity.

An additional factor which was not investigated in the present study of Complement activation phenotypes was the effect of sampling location. However, a new venepuncture of a vein in the arm was used throughout the study to collect all samples to minimise any effects. Furthermore, no significant difference has been found between protein biomarker concentrations in venous and arterial blood (610) suggesting sampling location is unlikely to affect results of systemic serum markers in the current work.

The effects of gender, age or sex on Complement activation phenotypes were not explored in the present study due to ethical restrictions enforcing data anonymity but are known to impact Complement activity (561). Many volunteers were young adults which may go some way to explain why the recorded C3 and C4 concentration distributions cover a comparatively narrow region of the total ranges reported in the literature. A broader range of test-subjects could be explored in future work to achieve a data set more representative of the general population.

Ultimately, although the ability of the Complement activation phenotypes to identify high-risk patients was not tested experimentally, the C-ODE model suggests a significant difference in phenotype values between healthy and Complement-depleted individuals. The C-ODE model accurately predicted the distribution of C5a and TCC rate phenotypes giving some confidence to the predictions. The C-ODE model predicts no correlation between the phenotype value before and after Complement-depletion – an interesting result supporting the initial hypothesis that repeated monitoring of these phenotypes would be useful to track patient immune state post-operatively to detect the onset of infection. The healthy cohort phenotype data and C-ODE model provide good preliminary data of an effect size, which could be used to plan a future clinical investigation of phenotype values from patients who develop infections following surgery.

7.3 Future Work

The present study developed two new assays for Complement activation, which in time could enable *in vitro* Complement testing in a clinical setting and measure Complement function with a less complex protocol than ELISA or the CH50 or AH50 tests mentioned previously. The new assays also allow rapid sampling of Complement activation to provide additional data for refinement of the C-ODE model. Preliminary data from the new assays in section 3.5 are presented here, complete with protocols, to inspire future studies using these assays as research tools.

7.3.1 High-Frequency Sampling of C5a in Human Serum

Preliminary data were collected to demonstrate use of the LiScAR platform for high-frequency (HF) sampling of C5a in serum following Complement activation. Pooled human serum was activated with 1 mg/mL zymosan as described in the protocol as described in section 5.3.1. 50 µL of serum was sampled from the reaction mixture, diluted 10-fold with 450 µL PBS and injected into the LiScAR following centrifugation at 1,300 *g* for 2 minutes. The response of the LiScAR to the measurement sequence (described in Chapter 3) is shown in Figure 7.3.

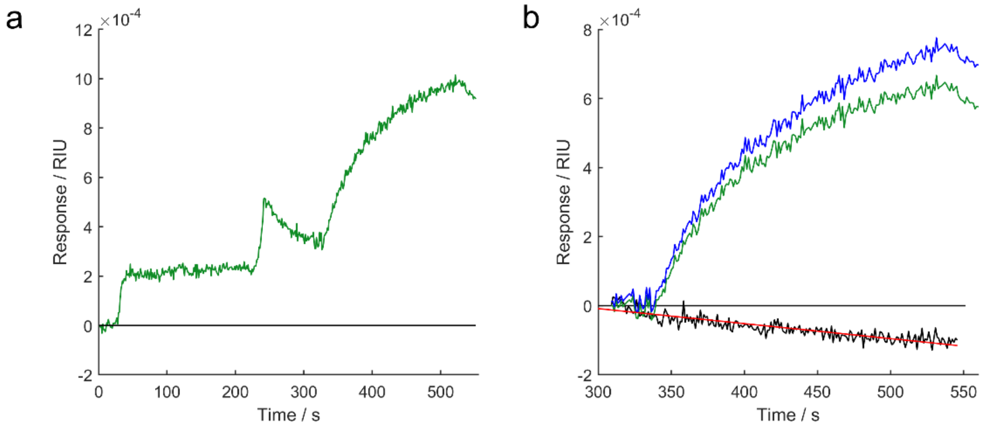


Figure 7.3 Response of anti-C5a assay spots (green) during the measurement sequence of an activated serum sample. The C5a assay is described in section 3.5 and was performed using the assay method described in 2.8.1. Signals are the averaged response of n=15 assay spots. a) The capture step at 20-220 seconds, followed by a PBS wash at 220-320 seconds and a detection step at 320-520 seconds with 200 nM polyclonal anti-C5a. b) The original detection step (green) is corrected for drift (blue) by subtracting a linear model (red) fit to signal drift data recorded by testing a serum sample and omitting the detection step (black).

The kinetics of the association phase, during which C5a capture occurs, were masked by non-specific binding of 10-fold diluted serum (20-200 s). Non-specific binding to the BSA reference spots was revealed by a PBS wash step which shows rapid dissociation from the reference spots: causing a rapid increase in the apparent response of the C5a assay spots at 220 s - an artefact of reference subtraction. The PBS wash step described in Chapter 3 was shortened from 200 s to 100 s to minimise the assay duty cycle and increase sampling frequency. 100 s of PBS washing allowed the majority of non-specifically bound material from the assay surface but performing the measurement sequence on a serum sample without performing the detection step revealed that the baseline continues to drift after 100 s – a potential source of error considering that the calibration curve used PBS samples which did not suffer from this effect. The signal drift in the HF data was corrected for by fitting a linear model to the drifting baseline and subtracting the linear model from the detection step data as shown in Figure 7.3. An alternate approach would be to calibrate the assay using C5a standards diluted using a matrix of pooled human serum.

The C5a concentration in the serum sample was derived as per the standard operating procedure in section 2.8.1 and comparing the AUC of the corrected detection step data to the standard curve using purified human C5a in PBS, as described in Chapters 2 and 3. The HF sampling method and the standard sampling method described in Chapter 3 (with PBS washing between capture and detection steps for 200 s until baseline stability) were used to profile pooled serum activated by 1 mg/mL zymosan, Figure 7.4. All serum experiments were performed in a Level 2 laboratory following relevant risk assessments.

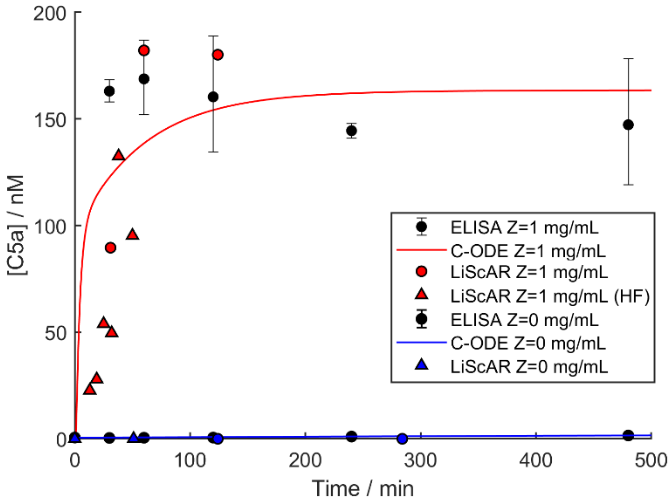


Figure 7.4 C5a concentration profile in pooled human serum following incubation with 0 mg/mL zymosan (blue) and 1 mg/mL zymosan (red). Data from the novel C5a assay developed in section 3.5 using a PBS wash step of 200 seconds (circles) are in good agreement with HF data from the same assay using a shorter wash step of 100 s and drift correction (triangles). LiScAR data points were derived from the averaged response of n=15 assay spots. The Bland-Altman analysis in section 3.5 suggests that, on average, the C5a assay systematically under-estimates the C5a concentration compared to ELISA by 26% (likely due to errors in the concentrations of the calibration standards) and this was corrected for in the data shown. The C-ODE model fit (solid curves) from Chapter 5 joins the ELISA data (black circles) and the LiScAR data in a continuous trend. Error bars are 2SD from n=2 technical repeats.

The LiScAR data for C5a collected using the standard and HF duty cycles show a similar trend in C5a production rate, although a thorough Bland-Altman analysis should be performed to determine the accuracy of the HF assay compared to the standard sampling routine. After correction for the 26% systematic error observed for the LiScAR C5a assay in Chapter 3, the LiScAR data closely match the ELISA data used to fit the C-ODE model at 60 minutes and 120 minutes following activation. The HF LiScAR data provide a first look at the C5a response profile within 30 minutes of activation and the real-time sampling allows Complement activation to be studied without using EDTA and freeze-thawing of samples - avoiding pre-analytical activation. The real-time LiScAR data indicate that the model over-estimates C5a production rate in early time. Future work could repeat the serum activation experiments in Chapter 5 with LiScAR assays to obtain HF C5a data for all zymosan activations for further refinement of the C-ODE model in early time.

7.3.2 Isolated Complement Pathways

Complex fluids such as human serum present a number of challenges for quantitative assays of analytes, such as matrix effects of excluded volume and diffusion limitations, as discussed in Chapter 2. These effects may also impact the transferability of model parameters derived in different solutions, as discussed in Chapter 4. Matrix effects can be eliminated entirely by reconstructing the system under study with purified proteins in the same matrix used for assay calibration. Furthermore, subsystems of isolated proteins can be studied to provide missing rate constants of complex systems, as mentioned previously in the context of the Silicone Cell approach to modelling.

Preliminary data were collected to demonstrate the use of the LiScAR to measure C3 consumption in an isolated AP of the constituent purified proteins. The fluid-phase association of FB to C3b is known to involve Mg^{2+} ions, but research using excess EDTA has shown that the requirement for Mg^{2+} is not absolute and in the absence of divalent cations, the fluid-phase reaction rate is reduced 3-fold (257) allowing AP activation to proceed in PBS. An isolated AP was produced according to Table 7.1 in PBS, containing proteins at 1/10th of their physiological concentrations in serum.

Table 7.1 Concentrations and volumes of AP components used to produce an isolated system in PBS.

	C3	FB	FD	FH	P
stock concentration / mg mL^{-1}	5	1	1.56×10^{-4}	1	8.04×10^{-2}
volume added to reaction mixture / µL	14.8	9	450	15	11.2
concentration in reaction mixture / mg mL^{-1}	0.148	1.80×10^{-2}	1.4×10^{-4}	3.00×10^{-2}	1.80×10^{-3}
physiological concentration / mg mL^{-1}	1.48	0.180	1.40×10^{-3}	0.300	1.8×10^{-2}

The reaction mixture was activated using zymosan as described in section 5.3.1 and profiled with the LiScAR assay for C3, calibrated with C3 standards in PBS as described in section 3.5. There is a possibility of hydroxyl groups on the SAM spacer binding to C3b in activated serum samples and the effect on assay performance should be evaluated in future work. Furthermore, the capture antibody tethered to the sensor surface in the C3 assay is monoclonal and known to bind an epitope on the C3a portion of C3. The effect of C3a competition for C3 on the LiScAR C3 assay was not tested and must be performed in future work to achieve accurate quantitative results. However, C3a has a mass of 9 kDa which should accommodate the binding of fewer polyclonal C3 detection antibodies than C3 (185 kDa) to produce a visible decrease in assay response. The aim of this experiment was to demonstrate a functional isolated AP and the C3 consumption for the isolated AP is shown in Figure 7.5, in response to zero and 1 mg/mL zymosan.

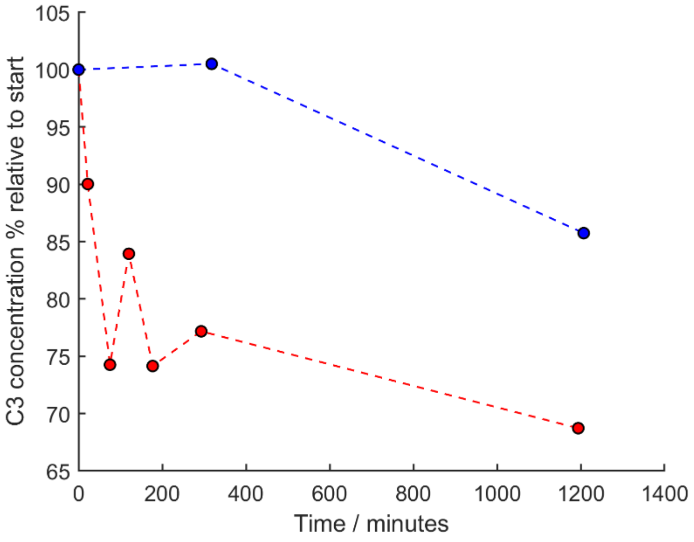


Figure 7.5 C3 consumption following activation of isolated AP proteins (in PBS at 1/10th of their serum concentrations) with zero (blue) and 1 mg/mL zymosan (red). C3 concentration data was recorded using the LiScAR biosensor platform and presented as a percentage relative to the starting value. The C3 assay is described in section 3.5 and was performed using the assay method described in 2.8.1. Data points were derived from the averaged response of n=15 assay spots.

Despite the unknown contribution of C3a to the assay response, 1 mg/mL zymosan activation showed greater C3 consumption compared to the control experiment, suggesting the isolated AP was functional and activated in PBS. Future work on the isolated AP described here could include Mg^{2+} and major serum proteins (albumins, globulins and fibrinogen) in the reaction mixture to better approximate the conditions in serum and provide accurate data for refinement of unknown AP parameters C-ODE model. Alternatively, the isolated AP demonstrated here could be used to quantitatively investigate the effect of different constituent protein concentrations on the system response – the effect of doubling Properdin or removing FH for example, as simulated in the C-ODE model.

The LiScAR biosensor platform is well suited to studying the isolated AP: the assays are quantitative, fast (<15 minutes), repeatable, multiplex and require small sample volumes (<1 mL). The selectivity of any novel assays should be thoroughly investigated in order to achieve physically relevant kinetic parameters and the anti-C3d/C3 antibody in Chapter 3 shows the ability to detect C3 captured by the anti-C3a/C3 capture antibody used for the existing assay, Appendix 1. The use of this antibody should provide a more selective C3 activation assay due to the specificity of the detection step, although the sensitivity and specificity of this assay requires further development and validation.

7.4 Conclusions

This present work has progressed from conceptualisation of a new biomarker, through development of new assays and training a new mathematical model, to recording the biomarker values in a cohort of healthy individuals. Furthermore, preliminary experiments have been proposed, using the new assays to provide more data for refinement of the model and hypothesis testing. The combination of new assays and model provides a powerful new tool for studying Complement activation in a clinical setting: the assays enable rapid profiling of Complement proteins in patient serum samples outside of the clinical chemistry laboratory with minimal user training, whilst the C-ODE model can be continuously refined with new data to improve predictions of Complement response phenotypes, based on starting protein concentrations in patient serum. Phenotype data from Complement-depleted patients could now be collected in a clinical investigation to validate the effect size of Complement depletion predicted by the cohort data and C-ODE model. The present study has made demonstrable progress towards the pre-operative identification of patients at higher risk of nosocomial infection and rapid detection of infection post-operatively.

Publications

Reader PP, Shaw AM. Kinetic Analysis of the Multivalent Ligand Binding Interaction between Protein A/G and IgG: A Standard System Setting. The Journal of Physical Chemistry B. 2017;121(38):8919-25.

Reader PP, Olkhov RV, Reeksting S, Lubben A, Hyde CJ, Shaw AM. A rapid and quantitative technique for assessing IgG monomeric purity, calibrated with the NISTmAb reference material. Analytical and Bioanalytical Chemistry. 2019;411(24):6487-96.

Bibliography

1. Khan HA, Baig FK, Mehboob R. Nosocomial infections: Epidemiology, prevention, control and surveillance. Asian Pacific Journal of Tropical Biomedicine. 2017;7(5):478-82.
2. Rebollo MH, Bernal JM, Llorca J, Rabasa JM, Revuelta JM. Nosocomial infections in patients having cardiovascular operations: A multivariate analysis of risk factors. The Journal of Thoracic and Cardiovascular Surgery. 1996;112(4):908-13.
3. Badia JM, Casey AL, Petrosillo N, Hudson PM, Mitchell SA, Crosby C. Impact of surgical site infection on healthcare costs and patient outcomes: a systematic review in six European countries. Journal of Hospital Infection. 2017;96(1):1-15.
4. WHO. The burden of health care-associated infection worldwide 2016 [Available from: http://www.who.int/gpsc/country_work/burden_hcai/en/
5. WHO. Health care-associated infections fact sheet 2014 [Available from: http://www.who.int/gpsc/country_work/gpsc_ccisc_fact_sheet_en.pdf.
6. Edwards M, Whittle J, Ackland GL. Biomarkers to guide perioperative management. Postgraduate Medical Journal. 2011;87(1030):542.
7. Ferraris VA, Bolanos M, Martin JT, Mahan A, Saha SP. Identification of Patients With Postoperative Complications Who Are at Risk for Failure to RescueComplications and Risk for Failure to RescueComplications and Risk for Failure to Rescue. JAMA Surgery. 2014;149(11):1103-8.
8. Machetti M, Viscoli C, Varnier O. Infections in Patients with Febrile Neutropenia: Epidemiology, Microbiology, and Risk Stratification. Clinical Infectious Diseases. 2005;40(Supplement_4):S240-S5.
9. Dave J, Pittard A, Inweregbu K. Nosocomial infections. Continuing Education in Anaesthesia Critical Care & Pain. 2005;5(1):14-7.
10. AS C. Patient Safety and Quality: An Evidence-Based Handbook for Nurses 2008 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK2683/?report=classic.
11. CDC. Types of Healthcare-associated Infections 2014 [Available from: https://www.cdc.gov/HAI/infectionTypes.html.
12. Trubiano JA, Padiglione AA. Nosocomial infections in the intensive care unit. Anaesthesia & Intensive Care Medicine. 2015;16(12):598-602.
13. Haque M, Sartelli M, McKimm J, Abu Bakar M. Health care-associated infections - an overview. Infection and drug resistance. 2018;11:2321-33.
14. ECDC. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals. 2012.
15. Faist E, Wichmann M, Kim C. Immunosuppression and immunomodulation in surgical patients. Current Opinion in Critical Care. 1997;3(4):293-8.
16. Dąbrowska AM, Słotwiński R. The immune response to surgery and infection. Central-European journal of immunology. 2014;39(4):532-7.
17. Minto G, Biccard B. Assessment of the high-risk perioperative patient. Continuing Education in Anaesthesia Critical Care & Pain. 2013;14(1):12-7.
18. Tang R, Chen HH, Wang YL, Changchien CR, Chen J-S, Hsu K-C, et al. Risk Factors For Surgical Site Infection After Elective Resection of the Colon and Rectum: A Single-Center Prospective Study of 2,809 Consecutive Patients. Annals of Surgery. 2001;234(2):181-9.

19. System LUH. Surgical site infections are the most common and costly of hospital infections 2017 [Available from: https://www.sciencedaily.com/releases/2017/01/170119161551.htm.
20. Ng WK, Kaur MN, Thoma A. Plastic surgeons' self-reported operative infection rates at a Canadian academic hospital. Plastic surgery (Oakville, Ont). 2014;22(4):237-40.
21. WHO. Improving the prevention, diagnosis and clinical management of sepsis 2018 [Available from: http://www.who.int/sepsis/en/.
22. Kaml GJ, Davis KA. Surgical Critical Care for the Patient with Sepsis and Multiple Organ Dysfunction. Anesthesiology Clinics. 2016;34(4):681-96.
23. Sannini A, Tison A, Martin C, Mokart D, Houvenaeghel G, Blache JL, et al. Predictive perioperative factors for developing severe sepsis after major surgery. BJA: British Journal of Anaesthesia. 2005;95(6):776-81.
24. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):801-10.
25. Care DoHaS. New action to reduce sepsis gov.uk2015 [Available from: https://www.gov.uk/government/news/new-action-to-reduce-sepsis.
26. NCEPOD. Just Say Sepsis! A review of the process of care received by patients with sepsis 2015 [Available from: https://www.ncepod.org.uk/2015report2/downloads/JustSaySepsis_FullReport.pdf.
27. Ma X-Y, Tian L-X, Liang H-P. Early prevention of trauma-related infection/sepsis. Military Medical Research. 2016;3(1):33.
28. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. Ann Surg. 2005;242(3):326-41; discussion 41-3.
29. Trussell J, Gerkin R, Coates B, Brandenberger J, Tibi P, Keuth J, et al. Impact of a patient care pathway protocol on surgical site infection rates in cardiothoracic surgery patients. American journal of surgery. 2008;196(6):883-9; discussion 9.
30. Glotzbecker M, Troy M, Miller P, Berry J, Cohen L, Gryzwna A, et al. Implementing a Multidisciplinary Clinical Pathway Can Reduce the Deep Surgical Site Infection Rate After Posterior Spinal Fusion in High-Risk Patients. Spine deformity. 2019;7(1):33-9.
31. Campbell H, Hotchkiss R, Bradshaw N, Porteous M. Integrated care pathways. BMJ. 1998;316(7125):133.
32. Schrijvers G, van Hoorn A, Huiskes N. The care pathway: concepts and theories: an introduction. International journal of integrated care. 2012;12(Spec Ed Integrated Care Pathways):e192-e.
33. Levett DZH, Grocott MPW. Cardiopulmonary Exercise Testing for Risk Prediction in Major Abdominal Surgery. Anesthesiology Clinics. 2015;33(1):1-16.
34. Fragiadakis GK, Gaudillière B, Ganio EA, Aghaeepour N, Tingle M, Nolan GP, et al. Patient-specific Immune States before Surgery are Strong Correlates of Surgical Recovery. Anesthesiology. 2015;123(6):1241-55.
35. Jhanji S, Thomas B, Ely A, Watson D, Hinds CJ, Pearse RM. Mortality and utilisation of critical care resources amongst high-risk surgical patients in a large NHS trust. Anaesthesia. 2008;63(7):695-700.
36. Gillis C, Li C, Lee L, Awasthi R, Augustin B, Gamsa A, et al. Prehabilitation versus rehabilitation: a randomized control trial in patients

undergoing colorectal resection for cancer. Anesthesiology. 2014;121(5):937-47.
37. Zambouri A. Preoperative evaluation and preparation for anesthesia and surgery. Hippokratia. 2007;11(1):13-21.
38. NICE. Clinical guideline NG45 Preoperative tests (update) Routine preoperative tests for elective surgery 2016 [Available from: https://www.nice.org.uk/guidance/ng45.
39. Pearse RM, Holt PJE, Grocott MPW. Managing perioperative risk in patients undergoing elective non-cardiac surgery. BMJ. 2011;343.
40. Trugenberger CA, Wälti C, Peregrim D, Sharp ME, Bureeva S. Discovery of novel biomarkers and phenotypes by semantic technologies. BMC bioinformatics. 2013;14:51-.
41. Mayeux R. Biomarkers: potential uses and limitations. NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics. 2004;1(2):182-8.
42. Gans SL, Atema JJ, Stoker J, Toorenvliet BR, Laurell H, Boermeester MA. C-Reactive Protein and White Blood Cell Count as Triage Test Between Urgent and Nonurgent Conditions in 2961 Patients With Acute Abdominal Pain. Medicine. 2015;94(9):e569.
43. England PH. The 2nd Atlas of Variation in NHS Diagnostic Services in England: NHS; 2017 [Available from: https://fingertips.phe.org.uk/documents/DiagnosticAtlas_FINAL.pdf.
44. Rodseth RN, Padayachee L, Biccard BM. A meta-analysis of the utility of pre-operative brain natriuretic peptide in predicting early and intermediate-term mortality and major adverse cardiac events in vascular surgical patients. Anaesthesia. 2008;63(11):1226-33.
45. Kyriacos U, Jelsma J, Jordan S. Monitoring vital signs using early warning scoring systems: a review of the literature. Journal of Nursing Management. 2011;19(3):311-30.
46. Physicians RCo. National Early Warning Score (NEWS) 2 - Standardising the assessment of acute illness severity in the NHS 2017 [Available from: https://www.ombudsman.org.uk/sites/default/files/National%20Early%20Warning%20Score%20%28NEWS%29%20-%20Standardising%20the%20assessment%20of%20acute-illness%20severity%20in%20the%20NHS_0.pdf.
47. Ning Y, Hu R, Yao G, Bo S. Time to positivity of blood culture and its prognostic value in bloodstream infection. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2016;35(4):619-24.
48. Shapiro NI, Howell MD, Talmor D, Nathanson LA, Lisbon A, Wolfe RE, et al. Serum Lactate as a Predictor of Mortality in Emergency Department Patients with Infection. Annals of Emergency Medicine. 2005;45(5):524-8.
49. Rhee C, Murphy MV, Li L, Platt R, Klompas M, Centers for Disease C, et al. Lactate Testing in Suspected Sepsis: Trends and Predictors of Failure to Measure Levels. Critical care medicine. 2015;43(8):1669-76.
50. Lobo SM, Lobo FR, Bota DP, Lopes-Ferreira F, Soliman HM, Melot C, et al. C-reactive protein levels correlate with mortality and organ failure in critically ill patients. Chest. 2003;123(6):2043-9.
51. Povoa P. C-reactive protein: a valuable marker of sepsis. Intensive care medicine. 2002;28(3):235-43.

52. Orati JA, Almeida P, Santos V, Ciorla G, Lobo SM. Serum C-reactive protein concentrations in early abdominal and pulmonary sepsis. Revista Brasileira de terapia intensiva. 2013;25(1):6-11.
53. Miguel-Bayarri V, Casanoves-Laparra EB, Pallas-Beneyto L, Sancho-Chinesta S, Martin-Osorio LF, Tormo-Calandin C, et al. Prognostic value of the biomarkers procalcitonin, interleukin-6 and C-reactive protein in severe sepsis. Medicina intensiva. 2012;36(8):556-62.
54. Tschaikowsky K, Hedwig-Geissing M, Braun GG, Radespiel-Troeger M. Predictive value of procalcitonin, interleukin-6, and C-reactive protein for survival in postoperative patients with severe sepsis. Journal of Critical Care. 2011;26(1):54-64.
55. Sproston NR, Ashworth JJ. Role of C-Reactive Protein at Sites of Inflammation and Infection. Frontiers in immunology. 2018;9:754-.
56. Reinhart K, Bauer M, Riedemann NC, Hartog CS. New approaches to sepsis: molecular diagnostics and biomarkers. Clinical microbiology reviews. 2012;25(4):609-34.
57. Kehlet H, Dahl JB. Anaesthesia, surgery, and challenges in postoperative recovery. Lancet. 2003;362(9399):1921-8.
58. Pearse RM, Harrison DA, James P, Watson D, Hinds C, Rhodes A, et al. Identification and characterisation of the high-risk surgical population in the United Kingdom. Critical care (London, England). 2006;10(3):R81.
59. Strimbu K, Tavel JA. What are Biomarkers? Current opinion in HIV and AIDS. 2010;5(6):463-6.
60. Pletcher MJ, Pignone M. Evaluating the Clinical Utility of a Biomarker: A Review of Methods for Estimating Health Impact. Circulation. 2011;123(10):1116-24.
61. McDermott JE, Wang J, Mitchell H, Webb-Robertson B-J, Hafen R, Ramey J, et al. Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data. Expert opinion on medical diagnostics. 2013;7(1):37-51.
62. Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. Postgraduate Medical Journal. 2009;85(1008):538.
63. Novelli G, Ciccacci C, Borgiani P, Papaluca Amati M, Abadie E. Genetic tests and genomic biomarkers: regulation, qualification and validation. Clinical cases in mineral and bone metabolism : the official journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases. 2008;5(2):149-54.
64. Garcia-Gimenez JL, Seco-Cervera M, Tollefsbol TO, Roma-Mateo C, Peiro-Chova L, Lapunzina P, et al. Epigenetic biomarkers: Current strategies and future challenges for their use in the clinical laboratory. Critical reviews in clinical laboratory sciences. 2017;54(7-8):529-50.
65. Li D, Chan DW. Proteomic cancer biomarkers from discovery to approval: it's worth the effort. Expert review of proteomics. 2014;11(2):135-6.
66. Frank R, Hargreaves R. Clinical biomarkers in drug discovery and development. Nature Reviews Drug Discovery. 2003;2:566.
67. Firestein GS. A biomarker by any other name. Nature Clinical Practice Rheumatology. 2006;2:635.
68. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clinical Pharmacology & Therapeutics. 2001;69(3):89-95.
69. Weintraub WS, Lüscher TF, Pocock S. The perils of surrogate endpoints. European Heart Journal. 2015;36(33):2212-8.

70. Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P. The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline. PROTEOMICS – Clinical Applications. 2008;2(10-11):1386-402.
71. Quezada H, Guzman-Ortiz AL, Diaz-Sanchez H, Valle-Rios R, Aguirre-Hernandez J. Omics-based biomarkers: current status and potential use in the clinic. Boletin medico del Hospital Infantil de Mexico. 2017;74(3):219-26.
72. Forshed J. Experimental Design in Clinical 'Omics Biomarker Discovery. Journal of Proteome Research. 2017;16(11):3954-60.
73. Tebani A, Afonso C, Marret S, Bekri S. Omics-Based Strategies in Precision Medicine: Toward a Paradigm Shift in Inborn Errors of Metabolism Investigations. International journal of molecular sciences. 2016;17(9):1555.
74. Scatena R. Advances in Cancer Biomarkers: From biochemistry to clinic for a critical revision: Springer; 2015.
75. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32-5.
76. Biron BM, Ayala A, Lomas-Neira JL. Biomarkers for Sepsis: What is and What Might Be? Biomarker Insights. 2015;10s4:BMI.S29519.
77. Holland R. What makes a good biomarker? 2016 [Available from: http://ojs.whioce.com/index.php/apm/article/download/7/7.
78. Parker LA, Chilet-Rosell E, Hernández-Aguado I, Pastor-Valero M, Gea S, Lumbreras B. Diagnostic Biomarkers: Are We Moving from Discovery to Clinical Application? Clinical Chemistry. 2018:clinchem.2018.292854.
79. Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, et al. Fit-for-Purpose Method Development and Validation for Successful Biomarker Measurement. Pharmaceutical Research. 2006;23(2):312-28.
80. Masucci GV, Cesano A, Hawtin R, Janetzki S, Zhang J, Kirsch I, et al. Validation of biomarkers to predict response to immunotherapy in cancer: Volume I - pre-analytical and analytical validation. Journal for immunotherapy of cancer. 2016;4:76-.
81. Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular disease risk profiles. American Heart Journal. 1991;121(1, Part 2):293-8.
82. Abbott TEF, Cron N, Vaid N, Ip D, Torrance HDT, Emmanuel J. Pre-hospital National Early Warning Score (NEWS) is associated with in-hospital mortality and critical care unit admission: A cohort study. Annals of medicine and surgery (2012). 2018;27:17-21.
83. Bouch DC, Thompson JP. Severity scoring systems in the critically ill. Continuing Education in Anaesthesia Critical Care & Pain. 2008;8(5):181-5.
84. The Apgar Score. Pediatrics. 2015;136(4):819.
85. Rapsang AG, Shyam DC. Scoring systems in the intensive care unit: A compendium. Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine. 2014;18(4):220-8.
86. Curtis JR, van der Helm-van Mil AH, Knevel R, Huizinga TW, Haney DJ, Shen Y, et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. Arthritis care & research. 2012;64(12):1794-803.
87. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive care medicine. 1996;22(7):707-10.
88. Jones AE, Trzeciak S, Kline JA. The Sequential Organ Failure Assessment score for predicting outcome in patients with severe sepsis and

evidence of hypoperfusion at the time of emergency department presentation. Critical care medicine. 2009;37(5):1649-54.
89. Lee CW, Kou HW, Chou HS, Chou HH, Huang SF, Chang CH, et al. A combination of SOFA score and biomarkers gives a better prediction of septic AKI and in-hospital mortality in critically ill surgical patients: a pilot study. World journal of emergency surgery : WJES. 2018;13:41.
90. Ho KM, Dobb GJ, Knuiman M, Finn J, Lee KY, Webb SAR. A comparison of admission and worst 24-hour Acute Physiology and Chronic Health Evaluation II scores in predicting hospital mortality: a retrospective cohort study. Critical Care. 2005;10(1):R4.
91. Koene MGJ, Mulder HA, Stockhofe-Zurwieden N, Kruijt L, Smits MA. Serum protein profiles as potential biomarkers for infectious disease status in pigs. BMC Veterinary Research. 2012;8:32-.
92. Medzhitov R, Janeway C. Innate Immunity. New England Journal of Medicine. 2000;343(5):338-44.
93. Freeman J, McGowan JE, Jr. Risk factors for nosocomial infection. The Journal of infectious diseases. 1978;138(6):811-9.
94. Cui P, Fang X. Pathogenesis of infection in surgical patients. Current Opinion in Critical Care. 2015;21(4):343-50.
95. Pavlov VA, Ulloa L, Metz CN. Chapter 22 - Infection and Sepsis. In: Arnason BG, editor. NeuroImmune Biology. 9: Elsevier; 2010. p. 309-20.
96. Alberts B JA, Lewis J, et al. Chapter 24, The Adaptive Immune System. Molecular Biology of the Cell 4th edition. New York: Garland Science; 2002.
97. Alberts B JA, Lewis J, et al. Chapter 25, Innate Immunity. Molecular Biology of the Cell 4th edition. New York: Garland Science; 2002.
98. Matzinger P. The Danger Model: A Renewed Sense of Self. Science. 2002;296(5566):301.
99. Kusumoto S, Fukase K, Shiba T. Key structures of bacterial peptidoglycan and lipopolysaccharide triggering the innate immune system of higher animals: Chemical synthesis and functional studies. Proceedings of the Japan Academy Series B, Physical and Biological Sciences. 2010;86(4):322-37.
100. Ali GS, Reddy A. PAMP-triggered immunity: Early events in the activation of FLAGELLIN SENSITIVE2. Plant signaling & behavior. 2008;3(6):423-6.
101. Choi JP, Kim YS, Tae YM, Choi EJ, Hong BS, Jeon SG, et al. A viral PAMP double-stranded RNA induces allergen-specific Th17 cell response in the airways which is dependent on VEGF and IL-6. Allergy. 2010;65(10):1322-30.
102. Sorrell TC, Chen SC. Fungal-derived immune modulating molecules. Advances in experimental medicine and biology. 2009;666:108-20.
103. Teixeira MM, Almeida IC, Gazzinelli RT. Introduction: innate recognition of bacteria and protozoan parasites. Microbes and Infection. 2002;4(9):883-6.
104. Bidula S, Schelenz S. A Sweet Response to a Sour Situation: The Role of Soluble Pattern Recognition Receptors in the Innate Immune Response to Invasive Aspergillus fumigatus Infections. PLoS Pathogens. 2016;12(7):e1005637.
105. Roozendaal R, Carroll MC. Emerging Patterns in Complement-Mediated Pathogen Recognition. Cell. 2006;125(1):29-32.
106. Lord JM, Midwinter MJ, Chen Y-F, Belli A, Brohi K, Kovacs EJ, et al. The systemic immune response to trauma: an overview of pathophysiology and treatment. Lancet (London, England). 2014;384(9952):1455-65.

107. Manson J, Thiemermann C, Brohi K. Trauma alarmins as activators of damage-induced inflammation. The British journal of surgery. 2012;99 Suppl 1:12-20.
108. Mathew A, Lindsley TA, Sheridan A, Bhoiwala DL, Hushmendy SF, Yager EJ, et al. Degraded mitochondrial DNA is a newly identified subtype of the damage associated molecular pattern (DAMP) family and possible trigger of neurodegeneration. Journal of Alzheimer's disease : JAD. 2012;30(3):617-27.
109. VanPatten S, Al-Abed Y. High Mobility Group Box-1 (HMGb1): Current Wisdom and Advancement as a Potential Drug Target. Journal of medicinal chemistry. 2018;61(12):5093-107.
110. Zhang F, Brenner M, Yang W-L, Wang P. A cold-inducible RNA-binding protein (CIRP)-derived peptide attenuates inflammation and organ injury in septic mice. Scientific Reports. 2018;8(1):3052.
111. Ahrens S, Zelenay S, Sancho D, Hanc P, Kjaer S, Feest C, et al. F-actin is an evolutionarily conserved damage-associated molecular pattern recognized by DNGR-1, a receptor for dead cells. Immunity. 2012;36(4):635-45.
112. Dorward DA, Lucas CD, Chapman GB, Haslett C, Dhaliwal K, Rossi AG. The Role of Formylated Peptides and Formyl Peptide Receptor 1 in Governing Neutrophil Function during Acute Inflammation. The American Journal of Pathology. 2015;185(5):1172-84.
113. Binkowska AM, Michalak G, Słotwiński R. Current views on the mechanisms of immune responses to trauma and infection. Central-European Journal of Immunology. 2015;40(2):206-16.
114. Delano MJ, Thayer T, Gabrilovich S, Kelly-Scumpia KM, Winfield RD, Scumpia PO, et al. Sepsis Induces Early Alterations in Innate Immunity That Impact Mortality to Secondary Infection. The Journal of Immunology. 2011;186(1):195.
115. Berghof TVL, van der Klein SAS, Arts JAJ, Parmentier HK, van der Poel JJ, Bovenhuis H. Genetic and Non-Genetic Inheritance of Natural Antibodies Binding Keyhole Limpet Hemocyanin in a Purebred Layer Chicken Line. PLOS ONE. 2015;10(6):e0131088.
116. Alimonti JB, Ball TB, Fowke KR. Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS. Journal of General Virology. 2003;84(7):1649-61.
117. Krishna S, Miller LS. Innate and adaptive immune responses against Staphylococcus aureus skin infections. Seminars in immunopathology. 2012;34(2):261-80.
118. Moyes LH, Leitch EF, McKee RF, Anderson JH, Horgan PG, McMillan DC. Preoperative systemic inflammation predicts postoperative infectious complications in patients undergoing curative resection for colorectal cancer. British journal of cancer. 2009;100(8):1236-9.
119. Kubo T, Ono S, Ueno H, Shinto E, Yamamoto J, Hase K. Elevated preoperative C-reactive protein levels are a risk factor for the development of postoperative infectious complications following elective colorectal surgery. Langenbeck's archives of surgery. 2013;398(7):965-71.
120. Gonzalez-Martinez S, Olona Tabuena N, Martin Baranera M, Marti-Sauri I, Moll JL, Morales Garcia MA, et al. Inflammatory markers as predictors of postoperative adverse outcome in octogenarian surgical patients: an observational prospective study. Cirugia espanola. 2015;93(3):166-73.
121. De Magistris L, Paquette B, Orry D, Facy O, Di Giacomo G, Rat P, et al. Preoperative inflammation increases the risk of infection after elective colorectal

surgery: results from a prospective cohort. International journal of colorectal disease. 2016;31(9):1611-7.
122. Morrow David A, de Lemos James A. Benchmarks for the Assessment of Novel Cardiovascular Biomarkers. Circulation. 2007;115(8):949-52.
123. Maisel A. Biomonitoring and Biomarker-Guided Therapy: The Next Step in Heart Failure and Biomarker Research. Journal of the American College of Cardiology. 2011;58(18):1890-2.
124. Aziz N. Measurement of Circulating Cytokines and Immune-Activation Markers by Multiplex Technology in the Clinical Setting: What Are We Really Measuring? Forum on immunopathological diseases and therapeutics. 2015;6(1-2):19-22.
125. Schiegnitz E, Kammerer PW, Schon H, Blatt S, Berres M, Sagheb K, et al. Proinflammatory cytokines as serum biomarker in oral carcinoma-A prospective multi-biomarker approach. Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology. 2018;47(3):268-74.
126. Wei Y, Gadaria-Rathod N, Epstein S, Asbell P. Tear cytokine profile as a noninvasive biomarker of inflammation for ocular surface diseases: standard operating procedures. Investigative ophthalmology & visual science. 2013;54(13):8327-36.
127. Andreasen AS, Krabbe KS, Krogh-Madsen R, Taudorf S, Pedersen BK, Moller K. Human Endotoxemia as a Model of Systemic Inflammation. Current Medicinal Chemistry. 2008;15(17):1697-705.
128. Taudorf S, Krabbe KS, Berg RMG, Pedersen BK, Møller K. Human Models of Low-Grade Inflammation: Bolus versus Continuous Infusion of Endotoxin. Clinical and Vaccine Immunology. 2007;14(3):250.
129. Monastero RN, Pentyala S. Cytokines as Biomarkers and Their Respective Clinical Cutoff Levels. International Journal of Inflammation. 2017;2017:11.
130. Pinsky MR, Vincent J-L, Deviere J, Alegre M, Kahn RJ, Dupont E. Serum Cytokine Levels in Human Septic Shock: Relation to Multiple-System Organ Failure and Mortality. Chest. 1993;103(2):565-75.
131. Frangiamore SJ, Saleh A, Farias Kovac M, Grosso MJ, Zhang X, Bauer TW, et al. Synovial Fluid Interleukin-6 as a Predictor of Periprosthetic Shoulder Infection. JBJS. 2015;97(1):63-70.
132. Candau-Alvarez A, Gil-Campos M, De la Torre-Aguilar MJ, Llorente-Cantarero F, Lopez-Miranda J, Perez-Navero JL. Early Modification in Drainage of Interleukin-1β and Tumor Necrosis Factor-α Best Predicts Surgical-Site Infection After Cervical Neck Dissection for Oral Cancer. Journal of Oral and Maxillofacial Surgery. 2015;73(6):1189-98.
133. Giannoudis PV, Harwood PJ, Loughenbury P, Van Griensven M, Krettek C, Pape H-C. Correlation Between IL-6 Levels and the Systemic Inflammatory Response Score: Can an IL-6 Cutoff Predict a SIRS State? Journal of Trauma and Acute Care Surgery. 2008;65(3):646-52.
134. Ng PC, Cheng SH, Chui KM, Fok TF, Wong MY, Wong W, et al. Diagnosis of late onset neonatal sepsis with cytokines, adhesion molecule, and C-reactive protein in preterm very low birthweight infants. Archives of Disease in Childhood - Fetal and Neonatal Edition. 1997;77(3):F221.
135. Laborada G, Rego M, Jain A, Guliano M, Stavola J, Ballabh P, et al. Diagnostic Value of Cytokines and C-reactive Protein in the First 24 Hours of Neonatal Sepsis. Amer J Perinatol. 2003;20(08):491-502.

136. Boskabadi H, Maamouri G, Tavakol Afshari J, Mafinejad S, Hosseini G, Mostafavi-Toroghi H, et al. Evaluation of serum interleukins-6, 8 and 10 levels as diagnostic markers of neonatal infection and possibility of mortality. Iranian journal of basic medical sciences. 2013;16(12):1232-7.
137. O'Brien SM, Fitzgerald P, Scully P, Landers A, Scott LV, Dinan TG. Impact of gender and menstrual cycle phase on plasma cytokine concentrations. Neuroimmunomodulation. 2007;14(2):84-90.
138. Lyon D, Walter J, Munro CL, Schubert CM, McCain NL. Challenges in Interpreting Cytokine Biomarkers in Biobehavioral Research: A Breast Cancer Exemplar. Biological research for nursing. 2011;13(1):25-31.
139. Treffers LW, Zhao XW, van der Heijden J, Nagelkerke SQ, van Rees DJ, Gonzalez P, et al. Genetic variation of human neutrophil Fcgamma receptors and SIRPalpha in antibody-dependent cellular cytotoxicity towards cancer cells. Eur J Immunol. 2018;48(2):344-54.
140. Institute NHGR. The Cost of Sequencing a Human Genome 2016 [Available from: https://www.genome.gov/27565109/the-cost-of-sequencing-a-human-genome/.
141. Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The Predictive Capacity of Personal Genome Sequencing. Science Translational Medicine. 2012;4(133):133ra58.
142. Nguyen GT, Green ER, Mecsas J. Neutrophils to the ROScue: Mechanisms of NADPH Oxidase Activation and Bacterial Resistance. Frontiers in cellular and infection microbiology. 2017;7:373-.
143. Yao Y, Matsushima H, Ohtola JA, Geng S, Lu R, Takashima A. Neutrophil priming occurs in a sequential manner and can be visualized in living animals by monitoring IL-1β promoter activation. Journal of immunology (Baltimore, Md : 1950). 2015;194(3):1211-24.
144. Gaudilliere B, Fragiadakis GK, Bruggner RV, Nicolau M, Finck R, Tingle M, et al. Coordinated Surgical Immune Signatures Contain Correlates of Clinical Recovery. Science translational medicine. 2014;6(255):255ra131-255ra131.
145. O'Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ, Berry MP. The immune response in tuberculosis. Annu Rev Immunol. 2013;31:475-527.
146. Blischak JD, Tailleux L, Myrthil M, Charlois C, Bergot E, Dinh A, et al. Predicting susceptibility to tuberculosis based on gene expression profiling in dendritic cells. Scientific Reports. 2017;7(1):5702.
147. Tang BM, Huang SJ, McLean AS. Genome-wide transcription profiling of human sepsis: a systematic review. Critical care (London, England). 2010;14(6):R237-R.
148. Taylor NJ, Manakkat Vijay GK, Abeles RD, Auzinger G, Bernal W, Ma Y, et al. The severity of circulating neutrophil dysfunction in patients with cirrhosis is associated with 90-day and 1-year mortality. Alimentary Pharmacology & Therapeutics. 2014;40(6):705-15.
149. Mookerjee RP, Stadlbauer V, Lidder S, Wright GA, Hodges SJ, Davies NA, et al. Neutrophil dysfunction in alcoholic hepatitis superimposed on cirrhosis is reversible and predicts the outcome. Hepatology (Baltimore, Md). 2007;46(3):831-40.
150. Danikas DD, Karakantza M, Theodorou GL, Sakellaropoulos GC, Gogos CA. Prognostic value of phagocytic activity of neutrophils and monocytes in sepsis. Correlation to CD64 and CD14 antigen expression. Clinical and experimental immunology. 2008;154(1):87-97.

151. Ricklin D, Hajishengallis G, Yang K, Lambris JD. Complement - a key system for immune surveillance and homeostasis. Nature immunology. 2010;11(9):785-97.
152. Morad HOJ, Belete SC, Read T, Shaw AM. Time-course analysis of C3a and C5a quantifies the coupling between the upper and terminal Complement pathways in vitro. Journal of Immunological Methods. 2015;427:13-8.
153. Ren J, Zhao Y, Yuan Y, Han G, Li W, Huang Q, et al. Complement Depletion Deteriorates Clinical Outcomes of Severe Abdominal Sepsis: A Conspirator of Infection and Coagulopathy in Crime? PLOS ONE. 2012;7(10):e47095.
154. Prodeus AP, Zhou X, Maurer M, Galli SJ, Carroll MC. Impaired mast cell-dependent natural immunity in complement C3-deficient mice. Nature. 1997;390:172.
155. Botto M, Kirschfink M, Macor P, Pickering MC, Wurzner R, Tedesco F. Complement in human diseases: Lessons from complement deficiencies. Mol Immunol. 2009;46(14):2774-83.
156. Noris M, Remuzzi G. Overview of complement activation and regulation. Seminars in nephrology. 2013;33(6):479-92.
157. Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement System Part I – Molecular Mechanisms of Activation and Regulation. Frontiers in Immunology. 2015;6:262.
158. Merle NS, Noe R, Halbwachs-Mecarelli L, Fremeaux-Bacchi V, Roumenina LT. Complement System Part II: Role in Immunity. Frontiers in Immunology. 2015;6:257.
159. Stephan AH, Barres BA, Stevens B. The Complement System: An Unexpected Role in Synaptic Pruning During Development and Disease. Annual Review of Neuroscience. 2012;35(1):369-89.
160. Langer HF, Chung K-J, Orlova VV, Choi EY, Kaul S, Kruhlak MJ, et al. Complement-mediated inhibition of neovascularization reveals a point of convergence between innate immunity and angiogenesis. Blood. 2010;116(22):4395.
161. Janowska-Wieczorek A, Marquez-Curtis LA, Shirvaikar N, Ratajczak MZ. The Role of Complement in the Trafficking of Hematopoietic Stem/Progenitor Cells. Transfusion. 2012;52(12):2706-16.
162. Gauvreau D, Roy C, Tom F-Q, Lu H, Miegueu P, Richard D, et al. A new effector of lipid metabolism: Complement factor properdin. Molecular Immunology. 2012;51(1):73-81.
163. Conway EM. Reincarnation of ancient links between coagulation and complement. Journal of Thrombosis and Haemostasis. 2015;13(S1):S121-S32.
164. Amara U, Flierl MA, Rittirsch D, Klos A, Chen H, Acker B, et al. Molecular intercommunication between the complement and coagulation systems. Journal of immunology (Baltimore, Md : 1950). 2010;185(9):5628-36.
165. Morgan BP, Gasque P. Extrahepatic complement biosynthesis: where, when and why? Clinical and Experimental Immunology. 1997;107(1):1-7.
166. Veerhuis R, Nielsen HM, Tenner AJ. Complement in the Brain. Molecular immunology. 2011;48(14):1592-603.
167. Nesargikar PN, Spiller B, Chavez R. The complement system: history, pathways, cascade and inhibitors. European Journal of Microbiology & Immunology. 2012;2(2):103-11.
168. Sahu A, Kozel TR, Pangburn MK. Specificity of the thioester-containing reactive site of human C3 and its significance to complement activation. Biochemical Journal. 1994;302(Pt 2):429-36.

169. Barnum SR. Complement: A primer for the coming therapeutic revolution. Pharmacology & Therapeutics. 2017;172:63-72.
170. McGrath FDG, Brouwer MC, Arlaud GJ, Daha MR, Hack CE, Roos A. Evidence That Complement Protein C1q Interacts with C-Reactive Protein through Its Globular Head Region. The Journal of Immunology. 2006;176(5):2950.
171. Kishore U, Ghai R, Greenhough TJ, Shrive AK, Bonifati DM, Gadjeva MG, et al. Structural and functional anatomy of the globular domain of complement protein C1q. Immunology letters. 2004;95(2):10.1016/j.imlet.2004.06.015.
172. Roumenina LT, Popov KT, Bureeva SV, Kojouharova M, Gadjeva M, Rabheru S, et al. Interaction of the globular domain of human C1q with Salmonella typhimurium lipopolysaccharide. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 2008;1784(9):1271-6.
173. Gaboriaud C, Frachet P, Thielens NM, Arlaud GJ. The Human C1q Globular Domain: Structure and Recognition of Non-Immune Self Ligands. Frontiers in Immunology. 2011;2:92.
174. Navratil JS, Watkins SC, Wisnieski JJ, Ahearn JM. The globular heads of C1q specifically recognize surface blebs of apoptotic vascular endothelial cells. Journal of immunology (Baltimore, Md : 1950). 2001;166(5):3231-9.
175. Turner MW. The role of mannose-binding lectin in health and disease. Molecular Immunology. 2003;40(7):423-9.
176. Kishore U, Ghai R, Greenhough TJ, Shrive AK, Bonifati DM, Gadjeva MG, et al. Structural and functional anatomy of the globular domain of complement protein C1q. Immunology Letters. 2004;95(2):113-28.
177. Héja D, Kocsis A, Dobó J, Szilágyi K, Szász R, Závodszky P, et al. Revised mechanism of complement lectin-pathway activation revealing the role of serine protease MASP-1 as the exclusive activator of MASP-2. Proceedings of the National Academy of Sciences. 2012;109(26):10498.
178. Degn SE, Kjaer TR, Kidmose RT, Jensen L, Hansen AG, Tekin M, et al. Complement activation by ligand-driven juxtaposition of discrete pattern recognition complexes. Proceedings of the National Academy of Sciences. 2014;111(37):13445.
179. Hosoi S, Circolo A, Borsos T. Activation of human C1: analysis with Western blotting reveals slow self-activation. The Journal of Immunology. 1987;139(5):1602.
180. Yokoyama I, Waxman F. Clonal variations in complement activation and deposition of C3b and C4b on model immune complexes. Immunology. 1993;80(2):168-76.
181. Krishnan V, Xu Y, Macon K, Volanakis JE, Narayana SVL. The structure of C2b, a fragment of complement component C2 produced during C3 convertase formation. Acta Crystallographica Section D: Biological Crystallography. 2009;65(Pt 3):266-74.
182. Rapp HJ, Borsos T. Complement and Hemolysis. Science. 1963;141(3582):738.
183. Kemper C, Pangburn MK, Fishelson Z. Complement Nomenclature 2014. Molecular Immunology. 2014;61(2):56-8.
184. Pangburn MK, Schreiber RD, Muller-Eberhard HJ. Formation of the initial C3 convertase of the alternative complement pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3. J Exp Med. 1981;154(3):856-67.

185. Korotaevskiy AA, Hanin LG, Khanin MA. Non-linear dynamics of the complement system activation. Mathematical Biosciences. 2009;222(2):127-43.
186. Bexborn F, Andersson PO, Chen H, Nilsson B, Ekdahl KN. The Tick-Over Theory Revisited: Formation and Regulation of the Soluble Alternative Complement C3 Convertase (C3(H(2)O)Bb). Molecular immunology. 2008;45(8):2370-9.
187. Fishelson Z, Pangburn MK, Müller-Eberhard HJ. Characterization of the initial C3 convertase of the alternative pathway of human complement. The Journal of Immunology. 1984;132(3):1430.
188. Mollnes TE. Complement and Biocompatibility. Vox Sanguinis. 1998;74(S2):303-7.
189. Forneris F, Ricklin D, Wu J, Tzekou A, Wallace RS, Lambris JD, et al. Structures of C3b in Complex with Factors B and D Give Insight into Complement Convertase Formation. Science (New York, NY). 2010;330(6012):1816-20.
190. Pryzdial EL, Isenman DE. Alternative complement pathway activation fragment Ba binds to C3b. Evidence that formation of the factor B-C3b complex involves two discrete points of contact. Journal of Biological Chemistry. 1987;262(4):1519-25.
191. Schwaeble WJ, Reid KBM. Does properdin crosslink the cellular and the humoral immune response? Immunology Today. 1999;20(1):17-21.
192. Blatt AZ, Saggu G, Kulkarni KV, Cortes C, Thurman JM, Ricklin D, et al. Properdin-Mediated C5a Production Enhances Stable Binding of Platelets to Granulocytes in Human Whole Blood. Journal of immunology (Baltimore, Md : 1950). 2016;196(11):4671-80.
193. Lambris JD, Lao Z, Oglesby TJ, Atkinson JP, Hack CE, Becherer JD. Dissection of CR1, factor H, membrane cofactor protein, and factor B binding and functional sites in the third complement component. The Journal of Immunology. 1996;156(12):4821.
194. Hourcade DE. The Role of Properdin in the Assembly of the Alternative Pathway C3 Convertases of Complement. Journal of Biological Chemistry. 2006;281(4):2128-32.
195. Fearon DT, Austen KF. Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase. The Journal of Experimental Medicine. 1975;142(4):856.
196. Spitzer D, Mitchell LM, Atkinson JP, Hourcade DE. Properdin Can Initiate Complement Activation by Binding Specific Target Surfaces and Providing a Platform for De Novo Convertase Assembly. The Journal of Immunology. 2007;179(4):2600.
197. Kemper C, Atkinson JP, Hourcade DE. Properdin: Emerging Roles of a Pattern-Recognition Molecule. Annual Review of Immunology. 2010;28(1):131-55.
198. Harboe M, Johnson C, Nymo S, Ekholt K, Schjalm C, Lindstad JK, et al. Properdin binding to complement activating surfaces depends on initial C3b deposition. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(4):E534-e9.
199. Wilson MA, Kozel TR. Contribution of antibody in normal human serum to early deposition of C3 onto encapsulated and nonencapsulated Cryptococcus neoformans. Infection and Immunity. 1992;60(3):754-61.
200. Berends ET, Gorham RD, Jr., Ruyken M, Soppe JA, Orhan H, Aerts PC, et al. Molecular insights into the surface-specific arrangement of complement C5 convertase enzymes. BMC biology. 2015;13:93.

201. Zwarthoff SA, Berends ETM, Mol S, Ruyken M, Aerts PC, Józsi M, et al. Functional Characterization of Alternative and Classical Pathway C3/C5 Convertase Activity and Inhibition Using Purified Models. Frontiers in immunology. 2018;9:1691-.
202. Bubeck D. The Making of a Macromolecular Machine: Assembly of the Membrane Attack Complex. Biochemistry. 2014;53(12):1908-15.
203. Ricklin D, Barratt-Due A, Mollnes TE. Complement in clinical medicine: Clinical trials, case reports and therapy monitoring. Molecular Immunology. 2017;89:10-21.
204. Balk RA. Systemic inflammatory response syndrome (SIRS): where did it come from and is it still relevant today? Virulence. 2014;5(1):20-6.
205. Huber-Lang MS, Vidya Sarma J, Zetoune FS, Ward PA. Role of Complement in Multi-Organ Dysfunction Syndrome. In: Szebeni J, editor. The Complement System: Novel Roles in Health and Disease. Boston, MA: Springer US; 2004. p. 465-80.
206. Charchaflieh J, Wei J, Labaze G, Hou YJ, Babarsh B, Stutz H, et al. The Role of Complement System in Septic Shock. Clinical and Developmental Immunology. 2012;2012:8.
207. Jurianz K, Ziegler S, Garcia-Schuler H, Kraus S, Bohana-Kashtan O, Fishelson Z, et al. Complement resistance of tumor cells: basal and induced mechanisms. Mol Immunol. 1999;36(13-14):929-39.
208. Milis L, Morris CA, Sheehan MC, Charlesworth JA, Pussell BA. Vitronectin-mediated inhibition of complement: evidence for different binding sites for C5b-7 and C9. Clinical and experimental immunology. 1993;92(1):114-9.
209. Tschopp J, Chonn A, Hertig S, French LE. Clusterin, the human apolipoprotein and complement inhibitor, binds to complement C7, C8 beta, and the b domain of C9. Journal of immunology (Baltimore, Md : 1950). 1993;151(4):2159-65.
210. Campbell WD, Lazoura E, Okada N, Okada H. Inactivation of C3a and C5a octapeptides by carboxypeptidase R and carboxypeptidase N. Microbiology and immunology. 2002;46(2):131-4.
211. Ziccardi RJ. Activation of the early components of the classical complement pathway under physiologic conditions. The Journal of Immunology. 1981;126(5):1769.
212. Davis AE, 3rd. Biological effects of C1 inhibitor. Drug news & perspectives. 2004;17(7):439-46.
213. Blom AM, Webb J, Villoutreix BO, Dahlbäck B. A Cluster of Positively Charged Amino Acids in the C4BP α-Chain Is Crucial for C4b Binding and Factor I Cofactor Function. Journal of Biological Chemistry. 1999;274(27):19237-45.
214. Carina D, Manuel P, Lars F, Gertraud S, Jürg-Alfred S. Soluble complement receptor type 1 (CD35) is released from leukocytes by surface cleavage. European Journal of Immunology. 1994;24(11):2725-31.
215. Di Bona D, Montalto G, Clemenza L, Bascone F, Accardo P, Bellavia D, et al. Soluble complement receptor type 1 (sCR1) in chronic liver diseases: serum levels at different stages of liver diseases. Clinical and Experimental Immunology. 1998;114(1):102-5.
216. Identification of distinct C3b and C4b recognition sites in the human C3b/C4b receptor (CR1, CD35) by deletion mutagenesis. The Journal of Experimental Medicine. 1988;168(5):1699-717.

217. Ricklin D, Reis ES, Lambris JD. Complement in disease: a defence system turning offensive. Nature Reviews Nephrology. 2016;12:383.
218. Manderson AP, Botto M, Walport MJ. The Role of Complement in the Development of Systemic Lupus Erythematosus. Annual Review of Immunology. 2004;22(1):431-56.
219. Zipfel PF, Heinen S, Józsi M, Skerka C. Complement and diseases: Defective alternative pathway control results in kidney and eye diseases. Molecular Immunology. 2006;43(1):97-106.
220. Hill A, Richards Stephen J, Hillmen P. Recent developments in the understanding and management of paroxysmal nocturnal haemoglobinuria. British Journal of Haematology. 2007;137(3):181-92.
221. Schrezenmeier H, Höchsmann B. Drugs that inhibit complement. Transfusion and Apheresis Science. 2012;46(1):87-92.
222. Morgan BP, Harris CL. Complement, a target for therapy in inflammatory and degenerative diseases. Nature Reviews Drug Discovery. 2015;14:857.
223. Holz FG, Sadda SR, Busbee B, Chew EY, Mitchell P, Tufail A, et al. Efficacy and Safety of Lampalizumab for Geographic Atrophy Due to Age-Related Macular Degeneration: Chroma and Spectri Phase 3 Randomized Clinical TrialsEfficacy and Safety of Lampalizumab for Geographic Atrophy Due to Age-Related Macular DegenerationEfficacy and Safety of Lampalizumab for Geographic Atrophy Due to Age-Related Macular Degeneration. JAMA Ophthalmology. 2018;136(6):666-77.
224. Mastellos DC, Yancopoulou D, Kokkinos P, Huber-Lang M, Hajishengallis G, Biglarnia AR, et al. Compstatin: a C3-targeted complement inhibitor reaching its prime for bedside intervention. European journal of clinical investigation. 2015;45(4):423-40.
225. van Griensven M, Ricklin D, Denk S, Halbgebauer R, Braun CK, Schultze A, et al. Protective Effects of the Complement Inhibitor Compstatin CP40 in Hemorrhagic Shock. Shock. 2019;51(1):78-87.
226. Ballanti E, Perricone C, Greco E, Ballanti M, Di Muzio G, Chimenti MS, et al. Complement and autoimmunity. Immunologic research. 2013;56(2-3):477-91.
227. Legendre CM, Licht C, Muus P, Greenbaum LA, Babu S, Bedrosian C, et al. Terminal Complement Inhibitor Eculizumab in Atypical Hemolytic–Uremic Syndrome. New England Journal of Medicine. 2013;368(23):2169-81.
228. Yuan Y, Ren J, Wu X, Cao S, Li J. Exogenous C3 Postpones Complement Exhaustion and Confers Organ Protection in Murine Sepsis. Journal of Surgical Research. 2011;168(1):e87-e94.
229. Dreiher J, Almog Y, Sprung CL, Codish S, Klein M, Einav S, et al. Temporal trends in patient characteristics and survival of intensive care admissions with sepsis: A multicenter analysis*. Critical Care Medicine. 2012;40(3):855-60.
230. Wang HE, Shapiro NI, Griffin R, Safford MM, Judd S, Howard G. Chronic Medical Conditions and Risk of Sepsis. PLOS ONE. 2012;7(10):e48307.
231. Harris CL, Heurich M, Cordoba SRd, Morgan BP. The complotype: dictating risk for inflammation and infection. Trends in Immunology. 2012;33(10):513-21.
232. Ingram G, Loveless S, Howell OW, Hakobyan S, Dancey B, Harris CL, et al. Complement activation in multiple sclerosis plaques: an immunohistochemical analysis. Acta neuropathologica communications. 2014;2:53.

233. Kemp PA, Spragg JH, Brown JC, Morgan BP, Gunn CA, Taylor PW. Immunohistochemical determination of complement activation in joint tissues of patients with rheumatoid arthritis and osteoarthritis using neoantigen-specific monoclonal antibodies. Journal of clinical & laboratory immunology. 1992;37(4):147-62.
234. Duraiyan J, Govindarajan R, Kaliyappan K, Palanisamy M. Applications of immunohistochemistry. Journal of pharmacy & bioallied sciences. 2012;4(Suppl 2):S307-S9.
235. Coons AH, Kaplan MH. Localization of antigen in tissue cells; improvements in a method for the detection of antigen by means of fluorescent antibody. J Exp Med. 1950;91(1):1-13.
236. Dubois L, Andersson K, Asplund A, Björkelund H. Evaluating real-time immunohistochemistry on multiple tissue samples, multiple targets and multiple antibody labeling methods. BMC Research Notes. 2013;6(1):542.
237. Page Faulk W, Malcolm Taylor G. Communication to the editors: An immunocolloid method for the electron microscope. Immunochemistry. 1971;8(11):1081-3.
238. Shih AR, Murali MR. Laboratory tests for disorders of complement and complement regulatory proteins. American Journal of Hematology. 2015;90(12):1180-6.
239. Corthell JT. Chapter 8 - Immunoprecipitation. In: Corthell JT, editor. Basic Molecular Protocols in Neuroscience: Tips, Tricks, and Pitfalls. San Diego: Academic Press; 2014. p. 77-81.
240. Ariceta G, Besbas N, Johnson S, Karpman D, Landau D, Licht C, et al. Guideline for the investigation and initial therapy of diarrhea-negative hemolytic uremic syndrome. Pediatric Nephrology. 2008;24(4):687.
241. Yang S, McGookey M, Wang Y, Cataland SR, Wu HM. Effect of Blood Sampling, Processing, and Storage on the Measurement of Complement Activation Biomarkers. American Journal of Clinical Pathology. 2015;143(4):558-65.
242. Selberg O, Hecker H, Martin M, Klos A, Bautsch W, Köhl J. Discrimination of sepsis and systemic inflammatory response syndrome by determination of circulating plasma concentrations of procalcitonin, protein complement 3a, and interleukin-6. Critical Care Medicine. 2000;28(8):2793-8.
243. Ingham KC. Protein precipitation with polyethylene glycol. Methods in enzymology. 1984;104:351-6.
244. Mollnes TE, Redl H, HØGÅSen K, Bengtsson A, Garrhd P, Speilberg L, et al. Complement activation in septic baboons detected by neoepitope-specific assays for C3b/iC3b/C3c, C5a and the terminal C5b-9 complement complex (TCC). Clinical & Experimental Immunology. 1993;91(2):295-300.
245. Oppermann M, Gotze O. Characterization of physiologic breakdown products of the complement fragment Ba. Mol Immunol. 1994;31(4):307-14.
246. Nilsson PH, Thomas AM, Bergseth G, Gustavsen A, Volokhina EB, van den Heuvel LP, et al. Eculizumab-C5 complexes express a C5a neoepitope in vivo: Consequences for interpretation of patient complement analyses. Mol Immunol. 2017;89:111-4.
247. Aydin S. A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA. Peptides. 2015;72:4-15.
248. Bally RW, Gribnau TCJ. Some Aspects of the Chromogen 3,3´,5,5´-Tetramethylbenzidine as Hydrogen Donor in a Horseradish Peroxidase Assay1989.

249. Derzsy Z, Prohászka Z, Rigó J, Füst G, Molvarec A. Activation of the complement system in normal pregnancy and preeclampsia. Molecular Immunology. 2010;47(7):1500-6.
250. Stove S, Welte T, Wagner TO, Kola A, Klos A, Bautsch W, et al. Circulating complement proteins in patients with sepsis or systemic inflammatory response syndrome. Clinical and diagnostic laboratory immunology. 1996;3(2):175-83.
251. Burk AM, Martin M, Flierl MA, Rittirsch D, Helm M, Lampl L, et al. Early complementopathy after multiple injuries in humans. Shock. 2012;37(4):348-54.
252. Harboe M, Thorgersen EB, Mollnes TE. Advances in assay of complement function and activation. Advanced Drug Delivery Reviews. 2011;63(12):976-87.
253. Major B, Kardos J, Kékesi KA, Lőrincz Z, Závodszky P, Gál P. Calcium-dependent Conformational Flexibility of a CUB Domain Controls Activation of the Complement Serine Protease C1r. Journal of Biological Chemistry. 2010;285(16):11863-9.
254. Mortensen S, Jensen JK, Andersen GR. Solution Structures of Complement C2 and Its C4 Complexes Propose Pathway-specific Mechanisms for Control and Activation of the Complement Proconvertases. Journal of Biological Chemistry. 2016;291(32):16494-507.
255. Costabile M. Measuring the 50% haemolytic complement (CH50) activity of serum. Journal of visualized experiments : JoVE. 2010(37):1923.
256. Zhang Y, Suankratay C, Zhang X, Jones DR, Lint TF, Gewurz H. Calcium-independent haemolysis via the lectin pathway of complement activation in the guinea-pig and other species*. Immunology. 1999;97(4):686-92.
257. Pryzdial EL, Isenman DE. A reexamination of the role of magnesium in the human alternative pathway of complement. Mol Immunol. 1986;23(1):87-96.
258. Torreira E, Tortajada A, Montes T, de Córdoba SR, Llorca O. 3D structure of the C3bB complex provides insights into the activation and regulation of the complement alternative pathway convertase. Proceedings of the National Academy of Sciences. 2009;106(3):882.
259. Pangburn MK, Muller-Eberhard HJ. The C3 convertase of the alternative pathway of human complement. Enzymic properties of the bimolecular proteinase. The Biochemical journal. 1986;235(3):723-30.
260. Total Serum Hemolytic Complement (CH50) Activity in Normal Korean Adults. J Pathol Transl Med. 1983;17(3):271-4.
261. Lee M. Basic skills in interpreting laboratory data: ASHP; 2009.
262. Seelen MA, Roos A, Wieslander J, Mollnes TE, Sjöholm AG, Wurzner R, et al. Functional analysis of the classical, alternative, and MBL pathways of the complement system: standardization and validation of a simple ELISA. Journal of Immunological Methods. 2005;296(1):187-98.
263. Wen L, Atkinson JP, Giclas PC. Clinical and laboratory evaluation of complement deficiency. Journal of Allergy and Clinical Immunology. 2004;113(4):585-93.
264. Morgan AR, O'Hagan C, Touchard S, Lovestone S, Morgan BP. Effects of freezer storage time on levels of complement biomarkers. BMC research notes. 2017;10(1):559-.
265. Kozel TR, Wilson MA, Pfrommer GS, Schlageter AM. Activation and binding of opsonic fragments of C3 on encapsulated Cryptococcus neoformans by using an alternative complement pathway reconstituted from six isolated proteins. Infection and immunity. 1989;57(7):1922-7.

266. Bjornson AB, Magnafichi PI, Schreiber RD, Bjornson HS. Opsonization of bacteroides by the alternative complement pathway reconstructed from isolated plasma proteins. The Journal of experimental medicine. 1987;165(3):777-98.
267. Schreiber RD, Muller-Eberhard HJ. Assembly of the cytolytic alternative pathway of complement from 11 isolated plasma proteins. J Exp Med. 1978;148(6):1722-7.
268. Fischer HP. Mathematical modeling of complex biological systems: from parts lists to understanding systems behavior. Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism. 2008;31(1):49-59.
269. Polack F, Stepney S. Emergent Properties Do Not Refine. Electronic Notes in Theoretical Computer Science. 2005;137(2):163-81.
270. Symons J. Computational Models of Emergent Properties. Minds and Machines. 2008;18(4):475-91.
271. Sagar A, Dai W, Minot M, LeCover R, Varner JD. Reduced order modeling and analysis of the human complement system. PLOS ONE. 2017;12(11):e0187373.
272. Olkhov RV, Shaw AM. Label-free antibody-antigen binding detection by optical sensor array based on surface-synthesized gold nanoparticles. Biosensors & bioelectronics. 2008;23(8):1298-302.
273. Olkhov RV, Shaw AM. Quantitative label-free screening for antibodies using scattering biophotonic microarray imaging. Anal Biochem. 2010;396(1):30-5.
274. van Vuuren BJ, Read T, Olkhov RV, Shaw AM. Human serum albumin interference on plasmon-based immunokinetic assay for antibody screening in model blood sera. Anal Biochem. 2010;405(1):114-20.
275. Read T, Olkhov RV, Williamson ED, Shaw AM. Kinetic epitope mapping of monoclonal antibodies raised against the Yersinia pestis virulence factor LcrV. Biosensors & bioelectronics. 2015;65:47-53.
276. Olkhov RV, Fowke JD, Shaw AM. Whole serum BSA antibody screening using a label-free biophotonic nanoparticle array. Anal Biochem. 2009;385(2):234-41.
277. Olkhov RV, Parker R, Shaw AM. Whole blood screening of antibodies using label-free nanoparticle biophotonic array platform. Biosensors & bioelectronics. 2012;36(1):1-5.
278. Olkhov RV, Kaminski ER, Shaw AM. Differential immuno-kinetic assays of allergen-specific binding for peanut allergy serum analysis. Anal Bioanal Chem. 2012;404(8):2241-7.
279. Olkhov RV, Weissenborn MJ, Flitsch SL, Shaw AM. Glycosylation characterization of human and porcine fibrinogen proteins by lectin-binding biophotonic microarray imaging. Anal Chem. 2014;86(1):621-8.
280. Turner A, Karube I, Wilson GS, editors. Biosensors : Fundamentals and Applications. 1 ed. Oxford, New York: Oxford University Press; 1987.
281. Dennison MJ, Turner APF. Biosensors for environmental monitoring. Biotechnology Advances. 1995;13(1):1-12.
282. Sharma S, Byrne H, O'Kennedy Richard J. Antibodies and antibody-derived analytical biosensors. Essays in Biochemistry. 2016;60(1):9-18.
283. Lavecchia T, Tibuzzi A, Giardi MT. Biosensors for functional food safety and analysis. Advances in experimental medicine and biology. 2010;698:267-81.
284. Mascini M, Tombelli S. Biosensors for biomarkers in medical diagnostics. Biomarkers. 2008;13(7-8):637-57.

285. Yu D, Blankert B, Viré JC, Kauffmann JM. Biosensors in Drug Discovery and Drug Analysis. Analytical Letters. 2005;38(11):1687-701.
286. Bahadır EB, Sezgintürk MK. Applications of commercial biosensors in clinical, food, environmental, and biothreat/biowarfare analyses. Analytical Biochemistry. 2015;478:107-20.
287. Bhalla N, Jolly P, Formisano N, Estrela P. Introduction to biosensors. Essays in Biochemistry. 2016;60(1):1-8.
288. Biermann H. Important factors that inluence the determination fo detection limits Environmental Hazards Assessment Program State of California Department of Food and Agriculture 1989 [Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.639.9973&rep=rep1&type=pdf.
289. Thevenot DR, Tóth K, Durst RA, Wilson GS. Electrochemical Biosensors: Recommended Definitions and Classification. Pure and Applied Chemistry1999. p. 2333.
290. Gabriela P, Otilia B, Vlad-Oros B. Sol-gel technology in enzymatic electrochemical biosensors for clinical analysis2011.
291. Du Y, Dong S. Nucleic Acid Biosensors: Recent Advances and Perspectives. Analytical Chemistry. 2017;89(1):189-215.
292. Wang B, Anzai J-i. Recent Progress in Lectin-Based Biosensors. Materials. 2015;8(12):8590-607.
293. Rocchitta G, Spanu A, Babudieri S, Latte G, Madeddu G, Galleri G, et al. Enzyme Biosensors for Biomedical Applications: Strategies for Safeguarding Analytical Performances in Biological Fluids. Sensors (Basel, Switzerland). 2016;16(6):780.
294. Gui Q, Lawson T, Shan S, Yan L, Liu Y. The Application of Whole Cell-Based Biosensors for Use in Environmental Analysis and in Medical Diagnostics. Sensors (Basel, Switzerland). 2017;17(7):1623.
295. Ronkainen NJ, Halsall HB, Heineman WR. Electrochemical biosensors. Chemical Society Reviews. 2010;39(5):1747-63.
296. Ramanathan K, Danielsson B. Principles and applications of thermal biosensors. Biosensors and Bioelectronics. 2001;16(6):417-23.
297. Skládal P. Piezoelectric biosensors. TrAC Trends in Analytical Chemistry. 2016;79:127-33.
298. Damborský P, Švitel J, Katrlík J. Optical biosensors. Essays in Biochemistry. 2016;60(1):91-100.
299. Nabaei V, Chandrawati R, Heidari H. Magnetic biosensors: Modelling and simulation. Biosensors and Bioelectronics. 2018;103:69-86.
300. Borisov IA, Lobanov AV, Reshetilov AN, Kurganov BI. Quantitative analysis of the calibration dependences for biosensors. Applied Biochemistry and Microbiology. 2000;36(3):215-20.
301. Lee DK, In J, Lee S. Standard deviation and standard error of the mean. Korean journal of anesthesiology. 2015;68(3):220-3.
302. Reed GF, Lynn F, Meade BD. Use of coefficient of variation in assessing variability of quantitative assays. Clinical and diagnostic laboratory immunology. 2002;9(6):1235-9.
303. Ward WK. How to design a biosensor. Journal of diabetes science and technology. 2007;1(2):201-4.
304. Smeraglia J, Baldrey SF, Watson D. Matrix effects and selectivity issues in LC-MS-MS. Chromatographia. 2002;55(1):S95-S9.
305. Armbruster DA, Pry T. Limit of Blank, Limit of Detection and Limit of Quantitation. The Clinical Biochemist Reviews. 2008;29(Suppl 1):S49-S52.

306. Greg Wells HP, Charles William Russ IV. Signal, Noise and Detection Limits in Mass Spctrometry, Thechnical Note.
307. Vessman J. Selectivity or specificity? Validation of analytical methods from the perspective of an analytical chemist in the pharmaceutical industry. Journal of Pharmaceutical and Biomedical Analysis. 1996;14(8):867-9.
308. Peveler WJ, Yazdani M, Rotello VM. Selectivity and Specificity: Pros and Cons in Sensing. ACS Sensors. 2016;1(11):1282-5.
309. Esteves M-B, Binaghi RA. Antigenic similarities among mammalian immunoglobulins. Immunology. 1972;23(2):137-45.
310. Saper CB. A Guide to the Perplexed on the Specificity of Antibodies. Journal of Histochemistry & Cytochemistry. 2008;57(1):1-5.
311. Custers D, Cauwenbergh T, Bothy JL, Courselle P, De Beer JO, Apers S, et al. ATR-FTIR spectroscopy and chemometrics: An interesting tool to discriminate and characterize counterfeit medicines. J Pharm Biomed Anal. 2015;112:181-9.
312. Devi A, Jangir J, K AA. Chemical characterization complemented with chemometrics for the botanical origin identification of unifloral and multifloral honeys from India. Food research international (Ottawa, Ont). 2018;107:216-26.
313. Johnson SS, Anslyn EV, Graham HV, Mahaffy PR, Ellington AD. Fingerprinting Non-Terran Biosignatures. Astrobiology. 2018.
314. Sang S, Wang Y, Feng Q, Wei Y, Ji J, Zhang W. Progress of new label-free techniques for biosensors: a review. Critical Reviews in Biotechnology. 2016;36(3):465-81.
315. Li M, Cushing SK, Wu N. Plasmon-Enhanced Optical Sensors: A Review. The Analyst. 2015;140(2):386-406.
316. Rhouati A, Catanante G, Nunes G, Hayat A, Marty J-L. Label-Free Aptasensors for the Detection of Mycotoxins. Sensors. 2016;16(12).
317. Syahir A, Usui K, Tomizaki K-y, Kajikawa K, Mihara H. Label and Label-Free Detection Techniques for Protein Microarrays. Microarrays. 2015;4(2):228-44.
318. Varma M. Limitations of Label-Free Sensors in Serum Based Molecular Diagnostics 2015 [Available from: https://arxiv.org/ftp/arxiv/papers/1505/1505.01032.pdf.
319. Yang H. Enzyme-based ultrasensitive electrochemical biosensors. Current Opinion in Chemical Biology. 2012;16(3):422-8.
320. Länge K, Rapp BE, Rapp M. Surface acoustic wave biosensors: a review. Analytical and Bioanalytical Chemistry. 2008;391(5):1509-19.
321. Hansen KM, Thundat T. Microcantilever biosensors. Methods. 2005;37(1):57-64.
322. Recht MI, Nienaber V, Torres FE. Chapter Three - Fragment-Based Screening for Enzyme Inhibitors Using Calorimetry. In: Feig AL, editor. Methods in enzymology. 567: Academic Press; 2016. p. 47-69.
323. Marx KA. Quartz Crystal Microbalance: A Useful Tool for Studying Thin Polymer Films and Complex Biomolecular Systems at the Solution−Surface Interface. Biomacromolecules. 2003;4(5):1099-120.
324. Cooper MA. Non-optical screening platforms: the next wave in label-free screening? Drug Discovery Today. 2006;11(23):1068-74.
325. Wijaya E, Lenaerts C, Maricot S, Hastanin J, Habraken S, Vilcot J-P, et al. Surface plasmon resonance-based biosensors: From the development of different SPR structures to novel surface functionalization strategies. Current Opinion in Solid State and Materials Science. 2011;15(5):208-24.

326. Yuan G, Gao L, Chen Y, Liu X, Wang J, Wang Z. Improvement of optical sensing performances of a double-slot-waveguide-based ring resonator sensor on silicon-on-insulator platform. Optik. 2014;125(2):850-4.
327. Grist SM, Schmidt SA, Flueckiger J, Donzella V, Shi W, Talebi Fard S, et al. Silicon photonic micro-disk resonators for label-free biosensing. Optics Express. 2013;21(7):7994-8006.
328. Dey D, Goswami T. Optical biosensors: a revolution towards quantum nanoscale electronics device fabrication. Journal of biomedicine & biotechnology. 2011;2011:348218.
329. Sepúlveda B, Angelomé PC, Lechuga LM, Liz-Marzán LM. LSPR-based nanobiosensors. Nano Today. 2009;4(3):244-51.
330. Olaru A, Bala C, Jaffrezic-Renault N, Aboul-Enein HY. Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. Critical reviews in analytical chemistry. 2015;45(2):97-105.
331. Masson J-F. Surface Plasmon Resonance Clinical Biosensors for Medical Diagnostics. ACS Sensors. 2017;2(1):16-30.
332. Zhang S, Garcia-D'Angeli A, Brennan JP, Huo Q. Predicting detection limits of enzyme-linked immunosorbent assay (ELISA) and bioanalytical techniques in general. Analyst. 2014;139(2):439-45.
333. Wink T, van Zuilen SJ, Bult A, van Bennekom WP. Liposome-Mediated Enhancement of the Sensitivity in Immunoassays of Proteins and Peptides in Surface Plasmon Resonance Spectrometry. Analytical Chemistry. 1998;70(5):827-32.
334. Breault-Turcot J, Poirier-Richard HP, Couture M, Pelechacz D, Masson JF. Single chip SPR and fluorescent ELISA assay of prostate specific antigen. Lab on a Chip. 2015;15(23):4433-40.
335. Aubé A, Charbonneau DM, Pelletier JN, Masson J-F. Response Monitoring of Acute Lymphoblastic Leukemia Patients Undergoing l-Asparaginase Therapy: Successes and Challenges Associated with Clinical Sample Analysis in Plasmonic Sensing. ACS Sensors. 2016;1(11):1358-65.
336. Yarmush ML, Patankar DB, Yarmush DM. An analysis of transport resistances in the operation of BIAcore; implications for kinetic studies of biospecific interactions. Molecular immunology. 1996;33(15):1203-14.
337. Nieba L, Krebber A, Pluckthun A. Competition BIAcore for measuring true affinities: large differences from values determined from binding kinetics. Anal Biochem. 1996;234(2):155-65.
338. Rodriguez Emmenegger C, Brynda E, Riedel T, Sedlakova Z, Houska M, Alles AB. Interaction of Blood Plasma with Antifouling Surfaces. Langmuir. 2009;25(11):6328-33.
339. Mayer KM, Hafner JH. Localized Surface Plasmon Resonance Sensors. Chemical Reviews. 2011;111(6):3828-57.
340. Vincenzo A, Roberto P, Marco F, Onofrio MM, Maria Antonia I. Surface plasmon resonance in gold nanoparticles: a review. Journal of Physics: Condensed Matter. 2017;29(20):203002.
341. Hong Y, Huh Y-M, Yoon D, Yang J. Review Article Nanobiosensors Based on Localized Surface Plasmon Resonance for Biomarker Detection2013.
342. Bedford EE, Spadavecchia J, Pradier CM, Gu FX. Surface plasmon resonance biosensors incorporating gold nanoparticles. Macromolecular bioscience. 2012;12(6):724-39.
343. Jana NR, Gearheart L, Murphy CJ. Wet Chemical Synthesis of High Aspect Ratio Cylindrical Gold Nanorods. The Journal of Physical Chemistry B. 2001;105(19):4065-7.

344. Richard BM, Walter RL, Joseph NF. Laser Rayleigh scattering. Measurement Science and Technology. 2001;12(5):R33.
345. Fundamental Principles Governing Solvents Use. In: Wypych G, editor. Handbook of Solvents (Second Edition). Oxford: ChemTec Publishing; 2014. p. 11-72.
346. Nusz GJ, Marinakos SM, Curry AC, Dahlin A, Höök F, Wax A, et al. Label-Free Plasmonic Detection of Biomolecular Binding by a Single Gold Nanorod. Analytical Chemistry. 2008;80(4):984-9.
347. Leonard P, Hearty S, Brennan J, Dunne L, Quinn J, Chakraborty T, et al. Advances in biosensors for detection of pathogens in food and water. Enzyme and Microbial Technology. 2003;32(1):3-13.
348. A.Einstein. Investigation on the Theory of the Brownian Movement. New York: Dover Publications; 1926.
349. Reading Uo. Calculate density and viscosity of glycerol/water mixtures 2018 [Available from: http://www.met.reading.ac.uk/~sws04cdw/viscosity_calc.html.
350. Cheng N-S. Formula for the Viscosity of a Glycerol−Water Mixture. Industrial & Engineering Chemistry Research. 2008;47(9):3285-8.
351. Volk A, Kähler CJ. Density model for aqueous glycerol solutions. Experiments in Fluids. 2018;59(5):75.
352. Chu K-Y, Thompson AR. Densities and Refractive Indices of Alcohol-Water Solutions of n-Propyl, Isopropyl, and Methyl Alcohols. Journal of Chemical & Engineering Data. 1962;7(3):358-60.
353. Read T, Olkhov RV, Shaw AM. Measurement of the localised plasmon penetration depth for gold nanoparticles using a non-invasive bio-stacking method. Physical Chemistry Chemical Physics. 2013;15(16):6122-7.
354. Khan S, Gupta A, Verma NC, Nandi CK. Kinetics of protein adsorption on gold nanoparticle with variable protein structure and nanoparticle size. The Journal of Chemical Physics. 2015;143(16):164709.
355. Pan H, Qin M, Meng W, Cao Y, Wang W. How Do Proteins Unfold upon Adsorption on Nanoparticle Surfaces? Langmuir. 2012;28(35):12779-87.
356. Boujday S, Bantegnie A, Briand E, Marnet P-G, Salmain M, Pradier C-M. In-Depth Investigation of Protein Adsorption on Gold Surfaces: Correlating the Structure and Density to the Efficiency of the Sensing Layer. The Journal of Physical Chemistry B. 2008;112(21):6708-15.
357. Miodek A, Regan ME, Bhalla N, Hopkins AN, Goodchild AS, Estrela P. Optimisation and Characterisation of Anti-Fouling Ternary SAM Layers for Impedance-Based Aptasensors. Sensors. 2015;15(10).
358. Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM. Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. Chemical Reviews. 2005;105(4):1103-70.
359. Pensa E, Cortés E, Corthey G, Carro P, Vericat C, Fonticelli MH, et al. The Chemistry of the Sulfur–Gold Interface: In Search of a Unified Model. Accounts of Chemical Research. 2012;45(8):1183-92.
360. Xue Y, Li X, Li H, Zhang W. Quantifying thiol–gold interactions towards the efficient strength control. Nature Communications. 2014;5:4348.
361. Mrksich M, Whitesides GM. Using Self-Assembled Monolayers That Present Oligo(ethylene glycol) Groups To Control the Interactions of Proteins with Surfaces. Poly(ethylene glycol). ACS Symposium Series. 680: American Chemical Society; 1997. p. 361-73.

362. Cerruti M, Fissolo S, Carraro C, Ricciardi C, Majumdar A, Maboudian R. Poly(ethylene glycol) Monolayer Formation and Stability on Gold and Silicon Nitride Substrates. Langmuir. 2008;24(19):10646-53.
363. Fisher T. Carbodiimide Crosslinker Chemistry 2018 [Available from: https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/carbodiimide-crosslinker-chemistry.html.
364. Mogues T, Li J, Coburn J, Kuter DJ. IgG antibodies against bovine serum albumin in humans--their prevalence and response to exposure to bovine serum albumin. J Immunol Methods. 2005;300(1-2):1-11.
365. Langmuir I. The Adsorption of Gaes on Plane Surfaces of Glass, Mica and Platinum. Journal of the American Chemical Society. 1918;40(9):1361-403.
366. Kastritis PL, Bonvin AMJJ. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. Journal of the Royal Society Interface. 2013;10(79):20120835.
367. Ylera F, Harth S, Waldherr D, Frisch C, Knappik A. Off-rate screening for selection of high-affinity anti-drug antibodies. Analytical Biochemistry. 2013;441(2):208-13.
368. Read T, Olkhov RV, Williamson ED, Shaw AM. Label-free Fab and Fc affinity/avidity profiling of the antibody complex half-life for polyclonal and monoclonal efficacy screening. Analytical and Bioanalytical Chemistry. 2015;407(24):7349-57.
369. Latour RA. The langmuir isotherm: A commonly applied but misleading approach for the analysis of protein adsorption behavior. Journal of Biomedical Materials Research Part A. 2014;103(3):949-58.
370. Amin S, Barnett GV, Pathak JA, Roberts CJ, Sarangapani PS. Protein aggregation, particle formation, characterization & rheology. Current Opinion in Colloid & Interface Science. 2014;19(5):438-49.
371. Buchwalow I, Samoilova V, Boecker W, Tiemann M. Non-specific binding of antibodies in immunohistochemistry: fallacies and facts. Scientific Reports. 2011;1:28.
372. Minton AP. The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media. Journal of Biological Chemistry. 2001;276(14):10577-80.
373. Gottschalk PG, Dunn JR. The five-parameter logistic: A characterization and comparison with the four-parameter logistic. Analytical Biochemistry. 2005;343(1):54-65.
374. Conrad B. %CV in ELISA: How to Reduce Them and Why They’re Important 2018 [Available from: http://www.enzolifesciences.com/science-center/technotes/2018/january/cv-in-elisa-how-to-reduce-them-and-why-they-re-important/.
375. Bartczak D, Kanaras AG. Preparation of Peptide-Functionalized Gold Nanoparticles Using One Pot EDC/Sulfo-NHS Coupling. Langmuir. 2011;27(16):10119-23.
376. Tian L, Liu S, Wang S, Wang L. Ligand-binding specificity and promiscuity of the main lignocellulolytic enzyme families as revealed by active-site architecture analysis. Scientific Reports. 2016;6:23605.
377. Prlić A, Bradley AR, Duarte JM, Rose PW, Rose AS, Valasatava Y. NGL viewer: web-based molecular graphics for large complexes. Bioinformatics. 2018;34(21):3755-8.

378. 3 - Antibody structure–function relationships. In: Strohl WR, Strohl LM, editors. Therapeutic Antibody Engineering: Woodhead Publishing; 2012. p. 37-595.
379. Wang L, Amphlett G, Blättler WA, Lambert JM, Zhang W. Structural characterization of the maytansinoid–monoclonal antibody immunoconjugate, huN901–DM1, by mass spectrometry. Protein Science : A Publication of the Protein Society. 2005;14(9):2436-46.
380. Diamantis N, Banerji U. Antibody-drug conjugates--an emerging class of cancer treatment. British journal of cancer. 2016;114(4):362-7.
381. Dennler P, Fischer E, Schibli R. Antibody Conjugates: From Heterogeneous Populations to Defined Reagents. Antibodies. 2015;4(3).
382. Wehling C, Amon O, Bommer M, Hoppe B, Kentouche K, Schalk G, et al. Monitoring of complement activation biomarkers and eculizumab in complement-mediated renal disorders. Clinical and Experimental Immunology. 2017;187(2):304-15.
383. Kang HJ, Kim JH, Lee EH, Lee YK, Hur M, Lee KM. Change of complement system predicts the outcome of patients with severe thermal injury. The Journal of burn care & rehabilitation. 2003;24(3):148-53.
384. Makela K, Helen P, Haapasalo H, Paavonen T. Complement activation in astrocytomas: deposition of C4d and patient outcome. BMC cancer. 2012;12:565.
385. Palikhe A, Sinisalo J, Seppänen M, Haario H, Meri S, Valtonen V, et al. Serum Complement C3/C4 Ratio, a Novel Marker for Recurrent Cardiovascular Events. American Journal of Cardiology. 2007;99(7):890-5.
386. Zimmermann-Nielsen E, Baatrup G, Thorlacius-Ussing O, Agnholt J, Svehag SE. Complement Activation Mediated by Mannan-Binding Lectin in Plasma from Healthy Individuals and from Patients with SLE, Crohn's Disease and Colorectal Cancer. Suppressed Activation by SLE Plasma. Scandinavian Journal of Immunology. 2002;55(1):105-10.
387. Siassi M, Hohenberger W, Riese J. Mannan-binding lectin (MBL) serum levels and post-operative infections. Biochemical Society transactions. 2003;31(Pt 4):774-5.
388. Ytting H, Christensen IJ, Jensenius JC, Thiel S, Nielsen HJ. Preoperative mannan-binding lectin pathway and prognosis in colorectal cancer. Cancer Immunology, Immunotherapy. 2005;54(3):265-72.
389. Carlsson M, Sjoholm AG, Eriksson L, Thiel S, Jensenius JC, Segelmark M, et al. Deficiency of the mannan-binding lectin pathway of complement and poor outcome in cystic fibrosis: bacterial colonization may be decisive for a relationship. Clin Exp Immunol. 2005;139(2):306-13.
390. Shahini N, Michelsen AE, Nilsson PH, Ekholt K, Gullestad L, Broch K, et al. The alternative complement pathway is dysregulated in patients with chronic heart failure. Scientific Reports. 2017;7:42532.
391. Mellbin LG, Bjerre M, Thiel S, Hansen TK. Complement Activation and Prognosis in Patients With Type 2 Diabetes and Myocardial Infarction. Diabetes Care. 2012;35(4):911.
392. Wu TC, Yang S, Haven S, Holers VM, Lundberg AS, Wu H, et al. Complement activation and mortality during an acute episode of thrombotic thrombocytopenic purpura. Journal of Thrombosis and Haemostasis. 2013;11(10):1925-7.
393. Cataland SR, Holers VM, Geyer S, Yang S, Wu HM. Biomarkers of terminal complement activation confirm the diagnosis of aHUS and differentiate aHUS from TTP. Blood. 2014;123(24):3733.

394. Kim MY, Guerra MM, Kaplowitz E, Laskin CA, Petri M, Branch DW, et al. Complement activation predicts adverse pregnancy outcome in patients with systemic lupus erythematosus and/or antiphospholipid antibodies. Annals of the rheumatic diseases. 2018;77(4):549-55.
395. Ruggiero A, Villa CH, Bander E, Rey DA, Bergkvist M, Batt CA, et al. Paradoxical glomerular filtration of carbon nanotubes. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(27):12369-74.
396. Mollnes TE, Lea T, FrØLand SS, Harboe M. Quantification of the Terminal Complement Complex in Human Plasma by an Enzyme-Linked Immunosorbent Assay Based on Monoclonal Antibodies against a Neoantigen of the Complex. Scandinavian Journal of Immunology. 1985;22(2):197-202.
397. Thomas SN, van der Vlies AJ, O'Neil CP, Reddy ST, Yu SS, Giorgio TD, et al. Engineering complement activation on polypropylene sulfide vaccine nanoparticles. Biomaterials. 2011;32(8):2194-203.
398. Welch NG, Scoble JA, Muir BW, Pigram PJ. Orientation and characterization of immobilized antibodies for improved immunoassays (Review). Biointerphases. 2017;12(2):02D301.
399. Rao SV, Anderson KW, Bachas LG. Oriented immobilization of proteins. Microchimica Acta. 1998;128(3):127-43.
400. Nojima Y, Iguchi K, Suzuki Y, Sato A. The pH-Dependent Formation of PEGylated Bovine Lactoferrin by Branched Polyethylene Glycol (PEG)-N-Hydroxysuccinimide (NHS) Active Esters. Biological and Pharmaceutical Bulletin. 2009;32(3):523-6.
401. Hermanson GT. Chapter 3 - The Reactions of Bioconjugation. In: Hermanson GT, editor. Bioconjugate Techniques (Third Edition). Boston: Academic Press; 2013. p. 229-58.
402. Goyon A, Excoffier M, Janin-Bussat M-C, Bobaly B, Fekete S, Guillarme D, et al. Determination of isoelectric points and relative charge variants of 23 therapeutic monoclonal antibodies. Journal of Chromatography B. 2017;1065-1066:119-28.
403. Pei Z, Anderson H, Myrskog A, Dunér G, Ingemarsson B, Aastrup T. Optimizing immobilization on two-dimensional carboxyl surface: pH dependence of antibody orientation and antigen binding capacity. Analytical Biochemistry. 2010;398(2):161-8.
404. van Erp R, Linders YEM, van Sommeren APG, Gribnau TCJ. Characterization of monoclonal antibodies physically adsorbed onto polystyrene latex particles. Journal of Immunological Methods. 1992;152(2):191-9.
405. Vashist SK. Comparison of 1-Ethyl-3-(3-Dimethylaminopropyl) Carbodiimide Based Strategies to Crosslink Antibodies on Amine-Functionalized Platforms for Immunodiagnostic Applications. Diagnostics. 2012;2(3):23-33.
406. Bhatt N, Huang P-JJ, Dave N, Liu J. Dissociation and Degradation of Thiol-Modified DNA on Gold Nanoparticles in Aqueous and Organic Solvents. Langmuir. 2011;27(10):6132-7.
407. Grawe RW, Knotts TA. The effects of tether placement on antibody stability on surfaces. The Journal of Chemical Physics. 2017;146(21):215102.
408. Trilling AK, Harmsen MM, Ruigrok VJB, Zuilhof H, Beekwilder J. The effect of uniform capture molecule orientation on biosensor sensitivity: Dependence on analyte properties. Biosensors and Bioelectronics. 2013;40(1):219-26.

409. Trilling AK, Beekwilder J, Zuilhof H. Antibody orientation on biosensor surfaces: a minireview. Analyst. 2013;138(6):1619-27.
410. Chen H, Huang J, Lee J, Hwang S, Koh K. Surface plasmon resonance spectroscopic characterization of antibody orientation and activity on the calixarene monolayer. Sensors and Actuators B: Chemical. 2010;147(2):548-53.
411. Sun Y, Du H, Feng C, Lan Y. Oriented immobilization of antibody through carbodiimide reaction and controlling electric field. Journal of Solid State Electrochemistry. 2015;19(10):3035-43.
412. Yuan Y, Yin M, Qian J, Liu C. Site-directed immobilization of antibodies onto blood contacting grafts for enhanced endothelial cell adhesion and proliferation. Soft Matter. 2011;7(16):7207-16.
413. Batalla P, Bolívar JM, Lopez-Gallego F, Guisan JM. Oriented covalent immobilization of antibodies onto heterofunctional agarose supports: A highly efficient immuno-affinity chromatography platform. Journal of Chromatography A. 2012;1262:56-63.
414. Billah MM, Hodges CS, Hays HCW, Millner PA. Directed immobilization of reduced antibody fragments onto a novel SAM on gold for myoglobin impedance immunosensing. Bioelectrochemistry. 2010;80(1):49-54.
415. Yang H-M, Bao R-M, Yu C-M, Lv Y-N, Zhang W-F, Tang J-B. Fc-specific biotinylation of antibody using an engineered photoactivatable Z–Biotin and its biosensing application. Analytica Chimica Acta. 2017;949:76-82.
416. Lee Y, Jeong J, Lee G, Moon JH, Lee MK. Covalent and Oriented Surface Immobilization of Antibody Using Photoactivatable Antibody Fc-Binding Protein Expressed in Escherichia coli. Anal Chem. 2016;88(19):9503-9.
417. Guo S-L, Chen P-C, Chen M-S, Cheng Y-C, Lin J-M, Lee H-C, et al. A Fast Universal Immobilization of Immunoglobulin G at 4°C for the Development of Array-based Immunoassays. PLOS ONE. 2012;7(12):e51370.
418. Verma SK, Amoah A, Schellhaas U, Winterhalter M, Springer S, Kolesnikova TA. “To Catch or Not to Catch”: Microcapsule-Based Sandwich Assay for Detection of Proteins and Nucleic Acids. Advanced Functional Materials. 2016;26(33):6015-24.
419. Bae YM, Oh B-K, Lee W, Lee WH, Choi J-W. Study on orientation of immunoglobulin G on protein G layer. Biosensors and Bioelectronics. 2005;21(1):103-10.
420. Arican O, Aral M, Sasmaz S, Ciragil P. Serum levels of TNF-alpha, IFN-gamma, IL-6, IL-8, IL-12, IL-17, and IL-18 in patients with active psoriasis and correlation with disease severity. Mediators of inflammation. 2005;2005(5):273-9.
421. Streiner DL, Norman GR. “Precision” and “Accuracy”: Two Terms That Are Neither. Journal of Clinical Epidemiology. 2006;59(4):327-30.
422. Bewick V, Cheek L, Ball J. Statistics review 7: Correlation and regression. Critical Care. 2003;7(6):451-9.
423. Giavarina D. Understanding Bland Altman analysis. Biochemia Medica. 2015;25(2):141-51.
424. Baker M. Reproducibility crisis: Blame it on the antibodies. Nature. 2015;521(7552):274-6.
425. Weller MG. Quality Issues of Research Antibodies. Analytical Chemistry Insights. 2016;11:ACI.S31614.
426. Baker M. Antibody anarchy: A call to order. Nature. 2015;527:545.
427. Vlasak J, Ionescu R. Fragmentation of monoclonal antibodies. mAbs. 2011;3(3):253-63.

428. Correia IR. Stability of IgG isotypes in serum. mAbs. 2010;2(3):221-32.
429. Voskuil JLA. The challenges with the validation of research antibodies. F1000Research. 2017;6:161-.
430. Points to consider in the manufacture and testing of monoclonal antibody products for human use (1997). U.S. Food and Drug Administration Center for Biologics Evaluation and Research. Journal of immunotherapy (Hagerstown, Md : 1997). 1997;20(3):214-43.
431. Schiel JE, Turner A, Mouchahoir T, Yandrofski K, Telikepalli S, King J, et al. The NISTmAb Reference Material 8671 value assignment, homogeneity, and stability. Analytical and Bioanalytical Chemistry. 2018;410(8):2127-39.
432. Turner A, Yandrofski K, Telikepalli S, King J, Heckert A, Filliben J, et al. Development of orthogonal NISTmAb size heterogeneity control methods. Analytical and bioanalytical chemistry. 2018;410(8):2095-110.
433. Schiel JE, Mire-Sluis A, Davis D. Monoclonal Antibody Therapeutics: The Need for Biopharmaceutical Reference Materials. State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 1 Monoclonal Antibody Therapeutics: Structure, Function, and Regulatory Space. ACS Symposium Series. 1176: American Chemical Society; 2014. p. 1-34.
434. Formolo T, Ly M, Levy M, Kilpatrick L, Lute S, Phinney K, et al. Determination of the NISTmAb Primary Structure. State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 2 Biopharmaceutical Characterization: The NISTmAb Case Study. ACS Symposium Series. 1201: American Chemical Society; 2015. p. 1-62.
435. McLellan JS, Chen M, Kim A, Yang Y, Graham BS, Kwong PD. Structural basis of respiratory syncytial virus neutralization by motavizumab. Nature structural & molecular biology. 2010;17(2):248-50.
436. Cohen SL, Price C, Vlasak J. β-Elimination and Peptide Bond Hydrolysis: Two Distinct Mechanisms of Human IgG1 Hinge Fragmentation upon Storage. Journal of the American Chemical Society. 2007;129(22):6976-7.
437. Liu H, Gaza-Bulseco G, Sun J. Characterization of the stability of a fully human monoclonal IgG after prolonged incubation at elevated temperature. Journal of chromatography B, Analytical technologies in the biomedical and life sciences. 2006;837(1-2):35-43.
438. Zhang H, Cui W, Gross ML. Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. FEBS Letters. 2014;588(2):308-17.
439. Karageorgos I, Gallagher ES, Galvin C, Gallagher DT, Hudgens JW. Biophysical characterization and structure of the Fab fragment from the NIST reference antibody, RM 8671. Biologicals. 2017;50:27-34.
440. Resemann A, Jabs W, Wiechmann A, Wagner E, Colas O, Evers W, et al. Full validation of therapeutic antibody sequences by middle-up mass measurements and middle-down protein sequencing. mAbs. 2016;8(2):318-30.
441. Beck A, Sanglier-Cianférani S, Van Dorsselaer A. Biosimilar, Biobetter, and Next Generation Antibody Characterization by Mass Spectrometry. Analytical Chemistry. 2012;84(11):4637-46.
442. Rosati S, Yang Y, Barendregt A, Heck AJR. Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. Nature Protocols. 2014;9:967.
443. Higel F, Seidl A, Sorgel F, Friess W. N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. European journal of pharmaceutics and

biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik eV. 2016;100:94-100.
444. Vörös J. The Density and Refractive Index of Adsorbing Protein Layers. Biophysical Journal. 2004;87(1):553-61.
445. Ball V, Ramsden Jeremy J. Buffer dependence of refractive index increments of protein solutions. Biopolymers. 1998;46(7):489-92.
446. Lassen B, Malmsten M. Structure of Protein Layers during Competitive Adsorption. Journal of Colloid and Interface Science. 1996;180(2):339-49.
447. Wang D-S, Chang C-C, Shih S-C, Lin C-W. An ellipsometric study on the density and functionality of antibody layers immobilized by a randomly covalent method and a protein A-oriented method2009. 303-10 p.
448. Mazzer AR, Clifton LA, Perevozchikova T, Butler PD, Roberts CJ, Bracewell DG. Neutron reflectivity measurement of protein A–antibody complex at the solid-liquid interface. Journal of Chromatography A. 2017;1499:118-31.
449. Clark NJ, Zhang H, Krueger S, Lee HJ, Ketchem RR, Kerwin B, et al. Small-Angle Neutron Scattering Study of a Monoclonal Antibody Using Free-Energy Constraints. The Journal of Physical Chemistry B. 2013;117(45):14029-38.
450. Tan YH, Liu M, Nolting B, Go JG, Gervay-Hague J, Liu G-y. A Nanoengineering Approach for Investigation and Regulation of Protein Immobilization. ACS nano. 2008;2(11):2374-84.
451. Schlatter D, Barner R, Fattinger C, Huber W, Hübscher J, Hurst J, et al. The difference interferometer: application as a direct affinity sensor. Biosensors and Bioelectronics. 1993;8(2):109-16.
452. Stenberg E, Persson B, Roos H, Urbaniczky C. Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. Journal of Colloid and Interface Science. 1991;143(2):513-26.
453. Williams RA, Blanch HW. Covalent immobilization of protein monolayers for biosensor applications. Biosensors and Bioelectronics. 1994;9(2):159-67.
454. Hanson DC, Schumaker VN. A model for the formation and interconversion of protein A-immunoglobulin G soluble complexes. The Journal of Immunology. 1984;132(3):1397.
455. Cheung CSF, Anderson KW, Patel PM, Cade KL, Phinney KW, Turko IV. A new approach to quantification of mAb aggregates using peptide affinity probes. Scientific Reports. 2017;7:42497.
456. Cedersund G, Roll J. Systems biology: model based evaluation and comparison of potential explanations for given biological data. The FEBS Journal. 2009;276(4):903-22.
457. Hucka M, and the rest of the SF, Finney A, and the rest of the SF, Sauro HM, and the rest of the SF, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics. 2003;19(4):524-31.
458. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI—a COmplex PAthway SImulator. Bioinformatics. 2006;22(24):3067-74.
459. Motta S, Pappalardo F. Mathematical modeling of biological systems. Briefings in Bioinformatics. 2013;14(4):411-22.
460. Breitling R. What is Systems Biology? Frontiers in Physiology. 2010;1:9.
461. Hasdemir D, Hoefsloot HCJ, Smilde AK. Validation and selection of ODE based systems biology models: how to arrive at more reliable decisions. BMC Systems Biology. 2015;9:32.

462. Degasperi A, Fey D, Kholodenko BN. Performance of objective functions and optimisation procedures for parameter estimation in system biology models. npj Systems Biology and Applications. 2017;3(1):20.
463. Engl HW, Flamm C, Kügler P, Lu J, Müller S, Schuster P. Inverse problems in systems biology. Inverse Problems. 2009;25(12):123014.
464. Snoep JL. The Silicon Cell initiative: working towards a detailed kinetic description at the cellular level. Current Opinion in Biotechnology. 2005;16(3):336-43.
465. Chou IC, Voit EO. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. Mathematical Biosciences. 2009;219(2):57-83.
466. Tarantola A. Popper, Bayes and the inverse problem. Nature Physics. 2006;2:492.
467. Jármai K, Farkas J. Mechanics and design of tubular structures: Springer; 2014.
468. Liu Y, Gunawan R. Parameter estimation of dynamic biological network models using integrated fluxes. BMC Systems Biology. 2014;8(1):127.
469. Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J, Jirstrand M. Kinetic models in industrial biotechnology – Improving cell factory performance. Metabolic Engineering. 2014;24:38-60.
470. Link H, Kochanowski K, Sauer U. Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. Nature Biotechnology. 2013;31:357.
471. Prinz AA, Bucher D, Marder E. Similar network activity from disparate circuit parameters. Nature Neuroscience. 2004;7:1345.
472. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP. Universally Sloppy Parameter Sensitivities in Systems Biology Models. PLoS computational biology. 2007;3(10):e189.
473. Weisstein EW. Jacobian MathWorld--A Wolfram Web Resource [Available from: http://mathworld.wolfram.com/Jacobian.html.
474. Brown KS, Sethna JP. Statistical mechanical approaches to models with many poorly known parameters. Physical Review E. 2003;68(2):021904.
475. Brenner S. Sequences and consequences. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010;365(1537):207.
476. Liu D, Hoynes-O'Connor A, Zhang F. Bridging the gap between systems biology and synthetic biology. Frontiers in microbiology. 2013;4:211.
477. Ranganathan S, Suthers PF, Maranas CD. OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLoS computational biology. 2010;6(4):e1000744.
478. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, et al. Synthetic protein scaffolds provide modular control over metabolic flux. Nat Biotechnol. 2009;27(8):753-9.
479. Popper K. The logic of scientific discovery: Routledge; 2005.
480. Ashyraliyev M, Jaeger J, Blom JG. Parameter estimation and determinability analysis applied to Drosophila gap gene circuits. BMC Systems Biology. 2008;2(1):83.
481. Alvarez-Vasquez F, Sims KJ, Hannun YA, Voit EO. Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models. Journal of Theoretical Biology. 2004;226(3):265-91.
482. Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, et al. Can yeast glycolysis be understood in terms of in vitro

kinetics of the constituent enzymes? Testing biochemistry. European Journal of Biochemistry. 2000;267(17):5313-29.
483. Minton AP. How can biochemical reactions within cells differ from those in test tubes? Journal of Cell Science. 2006;119(14):2863.
484. du Preez FB, van Niekerk DD, Kooi B, Rohwer JM, Snoep JL. From steady-state to synchronized yeast glycolytic oscillations I: model construction. Febs j. 2012;279(16):2810-22.
485. De Crescenzo G, Pham PL, Durocher Y, O'Connor-McCourt MD. Transforming Growth Factor-beta (TGF-β) Binding to the Extracellular Domain of the Type II TGF-β Receptor: Receptor Capture on a Biosensor Surface Using a New Coiled-coil Capture System Demonstrates that Avidity Contributes Significantly to High Affinity Binding. Journal of Molecular Biology. 2003;328(5):1173-83.
486. Fan X, White IM, Shopova SI, Zhu H, Suter JD, Sun Y. Sensitive optical biosensors for unlabeled targets: A review. Analytica Chimica Acta. 2008;620(1–2):8-26.
487. Angeles-Martinez L, Theodoropoulos C. The Influence of Crowding Conditions on the Thermodynamic Feasibility of Metabolic Pathways. Biophysical Journal. 2015;109(11):2394-405.
488. Kirk PDW, Babtie AC, Stumpf MPH. Systems biology (un)certainties. Science. 2015;350(6259):386-8.
489. Kiessling LL, Strong LE, Gestwicki JE. Chapter 29. Principles for multivalent ligand design. Annual Reports in Medicinal Chemistry. Volume 35: Academic Press; 2000. p. 321-30.
490. Kiessling LL, Gestwicki JE, Strong LE. Synthetic Multivalent Ligands as Probes of Signal Transduction. Angewandte Chemie International Edition. 2006;45(15):2348-68.
491. Eliasson M, Andersson R, Olsson A, Wigzell H, Uhlén M. Differential IgG-binding characteristics of staphylococcal protein A, streptococcal protein G, and a chimeric protein AG. The Journal of Immunology. 1989;142(2):575-81.
492. Eliasson M, Olsson A, Palmcrantz E, Wiberg K, Inganäs M, Guss B, et al. Chimeric IgG-binding receptors engineered from staphylococcal protein A and streptococcal protein G. Journal of Biological Chemistry. 1988;263(9):4323-7.
493. Huse K, Böhme H-J, Scholz GH. Purification of antibodies by affinity chromatography. Journal of Biochemical and Biophysical Methods. 2002;51(3):217-31.
494. Bronner V, Tabul, M. and Bravman, T. Protein interaction analysis tech note 5820: rapid screening and selection of optimal antibody capturing system. : Bio-Rad Laboratories, Inc.; 2009 [Available from: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_5820A.pdf.
495. O'Keefe E, Vennett V. Use of immunoglobulin-loaded protein A-bearing staphylococci as a primary solid phase immunoadsorbent in radioimmunoassay. Journal of Biological Chemistry. 1980;255(2):561-8.
496. Ghose S, Hubbard B, Cramer SM. Binding capacity differences for antibodies and Fc-fusion proteins on protein A chromatographic materials. Biotechnology and Bioengineering. 2007;96(4):768-79.
497. Martos V, Castreño P, Valero J, de Mendoza J. Binding to protein surfaces by supramolecular multivalent scaffolds. Current Opinion in Chemical Biology. 2008;12(6):698-706.

498. Coleman TF, Li Y. On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. Mathematical Programming. 1994;67(1):189-224.
499. Coleman TF, Li Y. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. SIAM Journal on Optimization. 1996;6(2):418-45.
500. Ogi H, Motohisa K, Hatanaka K, Ohmori T, Hirao M, Nishiyama M. Concentration dependence of IgG–protein A affinity studied by wireless-electrodeless QCM. Biosensors and Bioelectronics. 2007;22(12):3238-42.
501. Ando T, Skolnick J. Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. Proceedings of the National Academy of Sciences. 2010;107(43):18457-62.
502. Smith S, Cianci C, Grima R. Macromolecular crowding directs the motion of small molecules inside cells. Journal of the Royal Society Interface. 2017;14(131):20170047.
503. Markiewski MM, DeAngelis RA, Lambris JD. Complexity of complement activation in sepsis. Journal of Cellular and Molecular Medicine. 2008;12(6a):2245-54.
504. Rooijakkers SHM, Wu J, Ruyken M, van Domselaar R, Planken KL, Tzekou A, et al. Structural and functional implications of the alternative complement pathway C3 convertase stabilized by a staphylococcal inhibitor. Nature immunology. 2009;10(7):721-7.
505. Manderson AP, Pickering MC, Botto M, Walport MJ, Parish CR. Continual Low-Level Activation of the Classical Complement Pathway. The Journal of Experimental Medicine. 2001;194(6):747-56.
506. Ziccardi RJ. Spontaneous activation of the first component of human complement (C1) by an intramolecular autocatalytic mechanism. Journal of immunology (Baltimore, Md : 1950). 1982;128(6):2500-4.
507. Tschopp J. Kinetics of activation of the first component of complement (C1) by IgG oligomers. Mol Immunol. 1982;19(5):651-7.
508. Dodds AW, Sim RB, Porter RR, Kerr MA. Activation of the first component of human complement (C1) by antibody—antigen aggregates. Biochemical Journal. 1978;175(2):383.
509. Hirayama H, Yoshii K, Ojima H, Kawai N, Gotoh S, Fukuyama Y. Linear systems analysis of activating processes of complement system as a defense mechanism. Biosystems. 1996;39(3):173-85.
510. Liu B, Zhang J, Tan PY, Hsu D, Blom AM, Leong B, et al. A Computational and Experimental Study of the Regulatory Mechanisms of the Complement System. PLoS computational biology. 2011;7(1):e1001059.
511. Zewde N, Gorham RD, Jr., Dorado A, Morikis D. Quantitative Modeling of the Alternative Pathway of the Complement System. PLOS ONE. 2016;11(3):e0152337.
512. Zewde N, Morikis D. A computational model for the evaluation of complement system regulation under homeostasis, disease, and drug intervention. PLOS ONE. 2018;13(6):e0198644.
513. Kitano H. Systems Biology: A Brief Overview. Science. 2002;295(5560):1662.
514. A Sexton DJ, A Kuzmič, Petr. A ‘ hit-and-run ’ kinetic mechanism for the analysis enzyme progress curves under low substrate concentrations. 2010.
515. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, et al. The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. Biochemistry. 2011;50(21):4402-10.

516. Bergseth G, Ludviksen JK, Kirschfink M, Giclas PC, Nilsson B, Mollnes TE. An international serum standard for application in assays to detect human complement activation products. Molecular Immunology. 2013;56(3):232-9.
517. Lachmann PJ. Looking back on the alternative complement pathway. Immunobiology. 2018;223(8):519-23.
518. Mackie A, Keseler IM, Nolan L, Karp PD, Paulsen IT. Dead End Metabolites - Defining the Known Unknowns of the E. coli Metabolic Network. PLOS ONE. 2013;8(9):e75210.
519. Singer M, Jones AM. Bench-to-bedside review: The role of C1-esterase inhibitor in sepsis and other critical illnesses. Critical Care. 2011;15(1):203-.
520. Shampine L, Reichelt M. The MATLAB ODE Suite. SIAM Journal on Scientific Computing. 1997;18(1):1-22.
521. Bergmann FT, Hoops S, Klahn B, Kummer U, Mendes P, Pahle J, et al. COPASI and its applications in biotechnology. Journal of Biotechnology. 2017;261:215-20.
522. Domijan M, Rand DA. Using constraints and their value for optimization of large ODE systems. Journal of The Royal Society Interface. 2015;12(104).
523. Lee SY, Park JM, Kim TY. Chapter four - Application of Metabolic Flux Analysis in Metabolic Engineering. In: Voigt C, editor. Methods in enzymology. 498: Academic Press; 2011. p. 67-93.
524. Fearon DT, Austen KF. Activation of the alternative complement pathway due to resistance of zymosan-bound amplification convertase to endogenous regulatory mechanisms. Proceedings of the National Academy of Sciences of the United States of America. 1977;74(4):1683-7.
525. Reader PP, Shaw AM. Kinetic Analysis of the Multivalent Ligand Binding Interaction between Protein A/G and IgG: A Standard System Setting. The Journal of Physical Chemistry B. 2017;121(38):8919-25.
526. Faro-Trindade I, Willment JA, Kerrigan AM, Redelinghuys P, Hadebe S, Reid DM, et al. Characterisation of Innate Fungal Recognition in the Lung. PLOS ONE. 2012;7(4):e35675.
527. Brouwer N, Dolman KM, van Zwieten R, Nieuwenhuys E, Hart M, Aarden LA, et al. Mannan-binding lectin (MBL)-mediated opsonization is enhanced by the alternative pathway amplification loop. Molecular Immunology. 2006;43(13):2051-60.
528. Rawal N, Rajagopalan R, Salvi VP. Stringent Regulation of Complement Lectin Pathway C3/C5 Convertase By C4b-Binding Protein (C4bp). Molecular immunology. 2009;46(15):2902-10.
529. Harboe M, Garred P, Karlstrøm E, Lindstad JK, Stahl GL, Mollnes TE. The down-stream effects of mannan-induced lectin complement pathway activation depend quantitatively on alternative pathway amplification. Molecular Immunology. 2009;47(2):373-80.
530. Harboe M, Ulvund G, Vien L, Fung M, Mollnes TE. The quantitative role of alternative pathway amplification in classical pathway induced terminal complement activation. Clinical & Experimental Immunology. 2004;138(3):439-46.
531. Kouser L, Abdul-Aziz M, Nayak A, Stover CM, Sim RB, Kishore U. Properdin and Factor H: Opposing Players on the Alternative Complement Pathway “See-Saw”. Frontiers in Immunology. 2013;4:93.
532. Fisher AB, Al-Mehdi AB, Manevich Y. Shear stress and endothelial cell activation. Critical Care Medicine. 2002;30(5):S192-S7.
533. Stover CM, McDonald J, Byrne S, Lambert DG, Thompson JP. Properdin Levels in Human Sepsis. Frontiers in Immunology. 2015;6:24.

534. Ali YM, Hayat A, Saeed BM, Haleem KS, Alshamrani S, Kenawy HI, et al. Low-dose recombinant properdin provides substantial protection against Streptococcus pneumoniae and Neisseria meningitidis infection. Proceedings of the National Academy of Sciences. 2014;111(14):5301.
535. Fijen CAP, Kuijper EJ, Te Bulte MT, Van De Heuvel MM, Holdrinet ACJM, Sim RB, et al. Heterozygous and homozygous factor H deficiency states in a Dutch family. Clinical & Experimental Immunology. 1996;105(3):511-6.
536. Bao L, Haas M, Quigg RJ. Complement Factor H Deficiency Accelerates Development of Lupus Nephritis. Journal of the American Society of Nephrology. 2011;22(2):285.
537. van Opzeeland FJH, Bootsma HJ, Langereis JD, de Jonge MI, de Groot R, van Selm S, et al. Complement Factor H Serum Levels Determine Resistance to Pneumococcal Invasive Disease. The Journal of infectious diseases. 2016;213(11):1820-7.
538. Fujita T. Evolution of the lectin–complement pathway and its role in innate immunity. Nature Reviews Immunology. 2002;2:346.
539. Dodds AW. Which Came First, the Lectin/Classical Pathway or the Alternative Pathway of Complement? Immunobiology. 2002;205(4):340-54.
540. Edda Klipp WL, Christoph Wierling, Axel Kowald, Hans Lehrach, Ralf Herwig. Systems Biology: A Textbook: Wiley; 2009.
541. Wade A. Derivation versus validation. Archives of Disease in Childhood. 2000;83(6):459.
542. du Preez FB, van Niekerk DD, Snoep JL. From steady-state to synchronized yeast glycolytic oscillations II: model validation. Febs j. 2012;279(16):2823-36.
543. Hasdemir D, Hoefsloot HCJ, Smilde AK. Validation and selection of ODE based systems biology models: how to arrive at more reliable decisions. BMC systems biology. 2015;9:32-.
544. Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? BMJ. 1995;311(7019):1539.
545. Suresh KP, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. Journal of Human Reproductive Sciences. 2012;5(1):7-13.
546. Halpern SD, Karlawish JT, Berlin JA. The continuing unethical conduct of underpowered clinical trials. JAMA. 2002;288(3):358-62.
547. Röhrig B, du Prel J-B, Blettner M. Study Design in Medical Research: Part 2 of a Series on the Evaluation of Scientific Publications. Deutsches Ärzteblatt International. 2009;106(11):184-9.
548. Peter G. Smith RHMaDAR. Field Trials of Health Interventions: A Toolbox. Chapter 13, Preliminary studies and pilot testing. Oxford (UK) 2015.
549. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Statistical methods in medical research. 2016;25(3):1057-73.
550. Kim TK. T test as a parametric statistic. Korean journal of anesthesiology. 2015;68(6):540-6.
551. Ghasemi A, Zahediasl S. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. International Journal of Endocrinology and Metabolism. 2012;10(2):486-9.
552. NIST. Anderson-Darling and Shapiro-Wilk tests 2019 [Available from: https://www.itl.nist.gov/div898/handbook/prc/section2/prc213.htm.

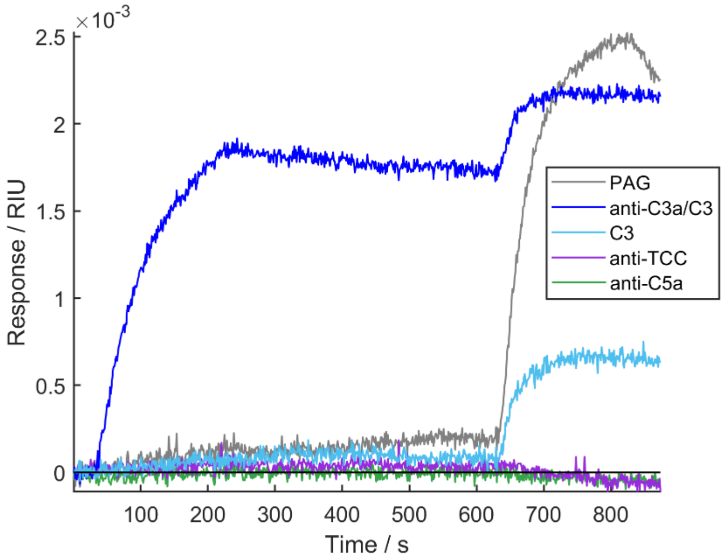
553. Fisher RA. Statistical Methods for Research Workers, 14. Aufl., Oliver & Boyd, Edinburgh, London 1970. XIII, 362 S., 12 Abb., 74 Tab., 40 s. Biometrische Zeitschrift. 1971;13(6):429-30.
554. Bangdiwala SI. Understanding Significance and P-Values. Nepal Journal of Epidemiology. 2016;6(1):522-4.
555. NIST. Two-Sample t-Test for Equal Means 2019 [Available from: https://www.itl.nist.gov/div898/handbook/eda/section3/eda353.htm.
556. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology. 2006;17(4):688-90.
557. NIST. F-Test for Equality of Two Variances 2019 [Available from: https://www.itl.nist.gov/div898/handbook/eda/section3/eda359.htm.
558. NIST. Kolmogorov-Smirnov Goodness-of-Fit Test 2019 [Available from: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm.
559. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. European journal of epidemiology. 2016;31(4):337-50.
560. Benasutti H, Wang G, Vu VP, Scheinman R, Groman E, Saba L, et al. Variability of Complement Response toward Preclinical and Clinical Nanocarriers in the General Population. Bioconjugate Chemistry. 2017;28(11):2747-55.
561. Gaya da Costa M, Poppelaars F, van Kooten C, Mollnes TE, Tedesco F, Würzner R, et al. Age and Sex-Associated Changes of Complement Activity and Complement Levels in a Healthy Caucasian Population. Frontiers in Immunology. 2018;9(2664).
562. Reviews C. Methods and Applications of Statistics in Clinical Trials2017.
563. Lachmann PJ. Preparing serum for functional complement assays. Journal of Immunological Methods. 2010;352(1):195-7.
564. Venkatraman Girija U, Gingras AR, Marshall JE, Panchal R, Sheikh MA, Harper JAJ, et al. Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(34):13916-20.
565. Erlandsen EJ, Randers E. Reference interval for serum C-reactive protein in healthy blood donors using the Dade Behring N Latex CRP mono assay. Scandinavian journal of clinical and laboratory investigation. 2000;60(1):37-43.
566. Morgan BP. The complement system: an overview. Methods in molecular biology (Clifton, NJ). 2000;150:1-13.
567. CompTech. C4 product description 2019 [Available from: https://www.complementtech.com/uploads/catalog/description/A105_C4.pdf.
568. Agarwal S, Cunningham-Rundles C. Assessment and clinical interpretation of reduced IgG values. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology. 2007;99(3):281-3.
569. Horak P, Hermanova Z, Zadrazil J, Ciferska H, Ordeltova M, Kusa L, et al. C1q complement component and -antibodies reflect SLE activity and kidney involvement. Clinical rheumatology. 2006;25(4):532-6.
570. Comptech. C2 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533bcac7490.

571. Oglesby TJ, Ueda A, Volanakis JE. Radioassays for quantitation of intact complement proteins C2 and B in human serum. Journal of Immunological Methods. 1988;110(1):55-62.
572. CompTech. C5 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/description/A120_C5.pdf.
573. CompTech. C6 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533df54aeb6.
574. CompTech. C7 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533e0f013b0.
575. CompTech. C8 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533e2c3a440.
576. CompTech. C9 Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/description/A126_C9.pdf.
577. CompTech. FD Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/description/A136_Factor_D.pdf.
578. Tsiftsoglou SA, Arnold JN, Roversi P, Crispin MD, Radcliffe C, Lea SM, et al. Human complement factor I glycosylation: Structural and functional characterisation of the N-linked oligosaccharides. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 2006;1764(11):1757-66.
579. Geerlings MJ, Kremlitzka M, Bakker B, Nilsson SC, Saksens NT, Lechanteur YT, et al. The Functional Effect of Rare Variants in Complement Genes on C3b Degradation in Patients With Age-Related Macular DegenerationRare Complement Gene Variants and C3b Degradation in AMDRare Complement Gene Variants and C3b Degradation in AMD. JAMA Ophthalmology. 2017;135(1):39-46.
580. Schmidt CQ, Slingsby FC, Richards A, Barlow PN. Production of biologically active complement factor H in therapeutically useful quantities. Protein Expression and Purification. 2011;76(2):254-63.
581. Hartmann S, Hofsteenge J. Properdin, the positive regulator of complement, is highly C-mannosylated. The Journal of biological chemistry. 2000;275(37):28569-74.
582. Stover CM, McDonald J, Byrne S, Lambert DG, Thompson JP. Properdin levels in human sepsis. Frontiers in immunology. 2015;6:24-.
583. CompTech. C4 Binding Protein Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533d2de3b7a.
584. CompTech. C1 Esterase Inhibitor Product Description 2019 [Available from: https://www.complementtech.com/uploads/catalog/5c533b5793692.
585. Keizer MP, Kamp A, van Mierlo G, Kuijpers TW, Wouters D. Substitution of Mannan-Binding Lectin (MBL)-Deficient Serum With Recombinant MBL Results in the Formation of New MBL/MBL-Associated Serine Protease Complexes. Frontiers in immunology. 2018;9:1406-.
586. de Benedetti F, Auriti C, D'Urbano LE, Ronchetti MP, Rava L, Tozzi A, et al. Low serum levels of mannose binding lectin are a risk factor for neonatal sepsis. Pediatric research. 2007;61(3):325-8.
587. Hara T, Fukuda H, Masaoka T, Matsumoto M, Seya T. Development of an ELISA assay for soluble CD35 (C3b/C4b receptor): high levels of soluble CD35 in LE-positive patients with hematological malignancies. Immunol Lett. 1994;41(2-3):249-53.
588. Di Bona D, Montalto G, Clemenza L, Bascone F, Accardo P, Bellavia D, et al. Soluble complement receptor type 1 (sCR1) in chronic liver diseases:

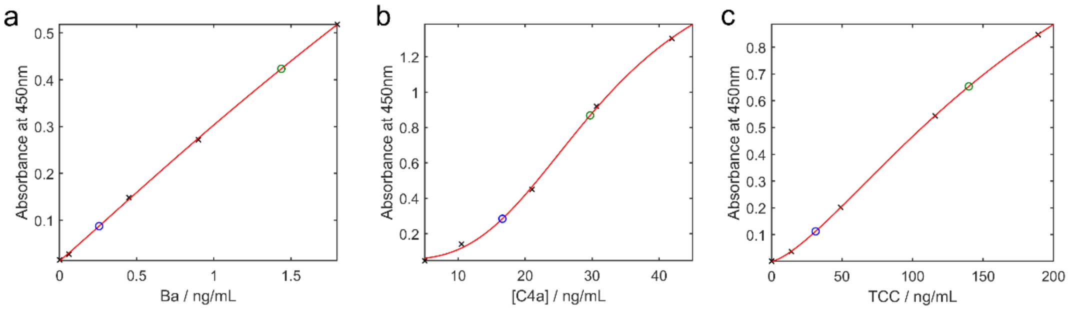
serum levels at different stages of liver diseases. Clin Exp Immunol. 1998;114(1):102-5.
589. Gomo R, Sibanda EN. Determination of normal ranges for serum C3 and C4 levels in an adult Zimbabwean population. The Central African journal of medicine. 2002;48(11-12):137-41.
590. Hospital ITJCU. Complement C3 and C4: NHS; 2019 [cited 2019 8th January]. Available from: https://www.southtees.nhs.uk/services/pathology/tests/complement-c3-c4/.
591. Bristol NN. Complement C3 + C4 2019 [Available from: https://www.nbt.nhs.uk/severn-pathology/requesting/test-information/complement-c3-c4.
592. Marks V, Cantor T, Mesko D, Pullmann R, Nosalova G. Differential Diagnosis by Laboratory Medicine A Quick Reference for Physicians: Springer-Verlag Berlin Heidelberg; 2002.
593. Lhotta K, Würzner R, Kronenberg F, Oppermann M, König P. Rapid activation of the complement system by cuprophane depends on complement component C4. Kidney International. 1998;53(4):1044-51.
594. Yilmazer M, Fenkci V, Fenkci S, Aktepe O, Sonmezer M, Kurtay G. Association of serum complement (C3, C4) and immunoglobulin (IgG, IgM) levels with hormone replacement therapy in healthy post-menopausal women. Human Reproduction. 2003;18(7):1531-5.
595. Gaya da Costa M, Poppelaars F, van Kooten C, Mollnes TE, Tedesco F, Würzner R, et al. Age and Sex-Associated Changes of Complement Activity and Complement Levels in a Healthy Caucasian Population. Frontiers in immunology. 2018;9:2664-.
596. Yang Y, Chung EK, Zhou B, Blanchong CA, Yu CY, Füst G, et al. Diversity in Intrinsic Strengths of the Human Complement System: Serum C4 Protein Concentrations Correlate with C4 Gene Size and Polygenic Variations, Hemolytic Activities, and Body Mass Index. The Journal of Immunology. 2003;171(5):2734.
597. Yamada K, Yamamoto M, Saeki T, Mizushima I, Matsui S, Fujisawa Y, et al. New clues to the nature of immunoglobulin G4-related disease: a retrospective Japanese multicenter study of baseline clinical features of 334 cases. Arthritis Research & Therapy. 2017;19(1):262.
598. Colten HR. Expression of the MHC class III genes. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 1984;306(1129):355-66.
599. Du B, Zielinski DC, Kavvas ES, Dräger A, Tan J, Zhang Z, et al. Evaluation of rate law approximations in bottom-up kinetic models of metabolism. BMC Systems Biology. 2016;10(1):40.
600. Bruggeman FJ, Hornberg JJ, Boogerd FC, Westerhoff HV. Introduction to systems biology. Exs. 2007;97:1-19.
601. Ben Kröse BKaPvdS, Patrick Smagt. An introduction to Neural Networks: University of Amsterdam; 1993.
602. Dalchau N. Understanding biological timing using mechanistic and black-box models. New Phytologist. 2012;193(4):852-8.
603. Le Novere N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, et al. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic acids research. 2006;34(Database issue):D689-91.

604. Hasdemir D, Hoefsloot HCJ, Westerhuis JA, Smilde AK. How informative is your kinetic model?: using resampling methods for model invalidation. BMC Systems Biology. 2014;8(1):61.
605. Smith AF, Messenger M, Hall P, Hulme C. The Role of Measurement Uncertainty in Health Technology Assessments (HTAs) of In Vitro Tests. PharmacoEconomics. 2018;36(7):823-35.
606. Pocock SJ, Ashby D, Shaper AG, Walker M, Broughton PM. Diurnal variations in serum biochemical and haematological measurements. Journal of clinical pathology. 1989;42(2):172-9.
607. Kage A, Fenner A, Weber B, Schöneshöfer M. Diurnal and ultradian variations of plasma concentrations of eleven adrenal steroid hormones in human males. Klinische Wochenschrift. 1982;60(13):659-66.
608. Petersen I, Baatrup G, Brandslund I, Teisner B, Rasmussen GG, Svehag SE. Circadian and diurnal variation of circulating immune complexes, complement-mediated solubilization, and the complement split product C3d in rheumatoid arthritis. Scandinavian journal of rheumatology. 1986;15(2):113-8.
609. Troldborg A, Jensen L, Deleuran B, Stengaard-Pedersen K, Thiel S, Jensenius JC. The C3dg Fragment of Complement Is Superior to Conventional C3 as a Diagnostic Biomarker in Systemic Lupus Erythematosus. Frontiers in Immunology. 2018;9(581).
610. Kelly E, Owen CA, Abraham A, Knowlton DL, Celli BR, Pinto-Plata V. Comparison of arterial and venous blood biomarker levels in chronic obstructive pulmonary disease. F1000Research. 2013;2:114.
611. Zadura AF, Theander E, Blom AM, Trouw LA. Complement Inhibitor C4b-Binding Protein in Primary Sjögren’s Syndrome and its Association With Other Disease Markers. Scandinavian Journal of Immunology. 2009;69(4):374-80.

Appendices



Appendix 1 Signal response of array surfaces to 100 nM C3 (capture step) and 100 nM anti-C3d/C3 (detection step). The assay is described in section 3.3.2 and was performed using the assay method described in 2.8.1. This alternate detection antibody for intact C3 shows affinity for captured and immobilised C3 and exhibits no cross-reactivity with the assays for C5a and TCC. The detection step should not detect captured C3a. The detection step signal is significantly weaker than when C3 antiserum is used.



Appendix 2 Standard curves of ELISA assays. Three assays are shown: a) Ba, b) C4a and c) TCC. Data are the mean of duplicate determinations (black). The data fit to a 5-parameter logistic curve with R^2>0.999 for all assays. High and low controls are shown in green and blue, respectively, displaying errors of <4%.

Appendix 3 Table of published rate constants adopted in the C-ODE model. The majority of rate constants are taken from a publication describing a previous kinetic model of Complement activation in ref (185). Units for first order and second order rate constants are min^{-1} and μM^{-1} min^{-1}respectively.

Reaction	Rate constant	Reference / Notes
'aC1C4 -> aC1 + C4a + C4b	300	(185)
C4b + C2 -> C4bC2	2.5	(185)
C4bC2 -> C4b + C2	2	(185)
aC1 + C1inh -> C1inhC1	3.6	(185)
C4bC2a -> C4b + C2a	0.35	(185)
C4bC2aC3 -> C4bC2a + C3a + aC3b	190	(185)
C3 -> C3(H2O)	0.0001	(185)

C3(H2O)FB -> C3(H2O) + FB	0.54	assumed same as Bb
C3(H2O)BbC3 -> C3(H2O)Bb + C3a + aC3b	55	(186) half the rate of alternative convertase
C3(H2O)Bb -> C3(H2O) + Bb	0.54	(185)
C3(H2O)FH -> C3(H2O) + FH	0.00069	(185) Inferred from C3bFH -> dC3b + FH
C3bBbC3 -> C3bBb + C3a + aC3b	110	(185)
C3bBbPC3 -> C3bBbP + C3a + aC3b	190	(185)
C3bBb -> dC3b + Bb	0.37	(185)
C3bBbP -> dC3b + Bb + P	0.028	(185)
aC3b + FH -> C3bFH	0.18	(185)
C3bFH -> dC3b + FH	0.00069	(185)
C3bFHFI -> iC3b + FH + FI	5.8	(185)
aC3b -> dC3b	192.5	(157) t1/2=60 µs
C4bC2a + aC3b -> (C4bC2a)C3b	96	(185)
(C4bC2a)C3b -> C4bC2a + dC3b	0.28	(185)
(C3bBb)C3b -> C3bBb + dC3b	0.23	(157) t1/2=3 min
(C3bBb)C3bP -> C3bBbP + dC3b	0.023	(157) t1/2=30 min
(C3bBb)C3bC5 -> (C3bBb)C3b + C5a + C5b	290	(185)
C5b + C6 -> C5bC6	3.6	(185)
C5bC6 -> C5b + C6	5.40E-06	(185)
C5bC6 + C7 -> C5bC6C7	44	(185)
C5bC6C7 -> C5bC6 + C7	7.00E-05	(185)
C5bC6C7 + C8 -> C5bC6C7C8	66	(185)
C5bC6C7C8 -> C5bC6C7 + C8	0.000315	(185)
C5bC6C7C8 + C9 -> TCC	170	(185)
TCC -> C5bC6C7C8 + C9	5.20E-05	(185)
C4BP + C4b -> C4BPC4b	12	(185)
C4BPC4b -> C4BP + C4b	0.96	(185)

Appendix 4 Table of starting concentrations for all species in the C-ODE model. Activation products were given starting values of zero unless measured in section 5.4.

Species name	Starting concentration / uM	Reference / Notes
(C3bBb)C3b	0	activation product
(C3bBb)C3bC5	0	activation product
(C3bBb)C3bP	0	activation product
(C3bBb)C3bPC5	0	activation product
(C4bC2a)C3b	0	activation product
(C4bC2a)C3bC5	0	activation product
A	0-0.02	activation stimulus
Ba	0.124	measured
Bb	0	activation product

C1	0.38	(569)
C1inh	1.9	(249)
C1inhC1	0	activation product
C2	0.25	(571)
C2a	0	activation product
C2b	0	activation product
C3	8	(249)
C3(H2O)	0	activation product
C3(H2O)Bb	0	activation product
C3(H2O)BbC3	0	activation product
C3(H2O)FB	0	activation product
C3(H2O)FBFD	0	activation product
C3(H2O)FH	0	activation product
C3(H2O)FHFI	0	activation product
C3a	0.685	measured
C3bBb	0	activation product
C3bBbC3	0	activation product
C3bBbP	0	activation product
C3bBbPC3	0	activation product
C3bP	0	activation product
C3bFB	0	activation product
C3bFBP	0	activation product
C3bFBFD	0	activation product
C3bFBFDP	0	activation product
C3bFH	0	activation product
C3bFHFI	0	activation product
C4	2	(593)
C4BP	0.42	(611)
C4BPC4b	0	activation product
C4BPC4bFI	0	activation product
C4a	0.801	measured
C4b	0	activation product
C4bC2	0	activation product
C4bC2a	0	activation product
C4bC2aC3	0	activation product
C5	0.34	(566)
C5a	0.001	measured
C5b	0	activation product
C5bC6	0	activation product
C5bC6C7	0	activation product
C5bC6C7C8	0	activation product
C6	0.61	(566)
C7	0.61	(566)
C8	0.36	(566)
C9	0.85	(566)
MBL	0.014	(586)
P	0.35	(533)
TCC	0.006	measured

aC1	0	activation product
aC1C4	0	activation product
aC1C4bC2	0	activation product
aC3b	0	activation product
aMBL	0	activation product
aMBLC1inh	0	activation product
aMBLC4	0	activation product
aMBLC4bC2	0	activation product
dC1	0	activation product
dC3(H2O)	0	activation product
dC3b	0	activation product
dC4b	0	activation product
dMBL	0	activation product
FB	1.9	(571)
FD	0.058	(566)
FH	1.94	(579)
FI	0.34	(579)
iC3b	0	activation product
sCR1	0.00017	(588)
sCR1C3b	0	activation product
sCR1C3bfI	0	activation product
sCR1C4b	0	activation product
sCR1C4bfI	0	activation product

Appendix 5 Table of reactions in the C-ODE model with unknown kinetic constants. The results of model fitting are shown, constrained within upper and lower bounds and starting from an initial guess value corresponding to median enzyme kinetics. Units for first order and second order rate constants are min^{-1} *and* μM^{-1} min^{-1}*respectively.*

Reaction	Initial guess	Lower bound	Upper bound	Fitted value
C1 -> aC1	0.01	6.67E-04	0.15	7.26E-03
A + C1 -> A + aC1	1	0.001	1000	166.42
aC1 -> C1	0.06	na	na	0.06
aC1 + C4 -> aC1C4	6	0.4	90	8.85
aC1C4 -> aC1 + C4	0.06	na	na	0.06
aC1 + C4bC2 -> aC1C4bC2	6	0.4	90	24.86
aC1C4bC2 -> aC1 + C4bC2	0.06	na	na	0.06
aC1C4bC2 -> C4bC2a + C2b + aC1	840	56	12600	911.04
A + MBL -> A + aMBL	1	0.001	1000	1.87
aMBL -> MBL	0.06	na	na	0.06
aMBL + C4 -> aMBLC4	6	0.4	90	4.57
aMBLC4 -> aMBL + C4	0.06	na	na	0.06
aMBLC4 -> aMBL + C4a + C4b	840	56	12600	100.76
aMBL + C4bC2 -> aMBLC4bC2	6	0.4	90	48.46

aMBLC4bC2 -> aMBL + C4bC2	0.06	na	na	0.06
aMBLC4bC2 -> C4bC2a + C2b + aMBL	840	56	12600	225.34
aMBL + C1inh -> aMBLC1inh	3.6	0.24	54	12.69
C4bC2a + C3 -> C4bC2aC3	6	0.4	90	2.27
C4bC2aC3 -> C4bC2a + C3	0.06	na	na	0.06
sCR1 + aC3b -> sCR1C3b	6	0.4	90	45.99
sCR1 + C4b -> sCR1C4b	6	0.4	90	9.65
sCR1 + C4bC2a -> sCR1C4b + C2a	6	0.4	90	12.74
C4BP + C4bC2a -> C4BPC4b + C2a	6	0.4	90	6.21
sCR1C3b + fI -> sCR1C3bfI	6	0.4	90	17.48
sCR1C4b + fI -> sCR1C4bfI	6	0.4	90	4.64
C4BPC4b + fI -> C4BPC4bfI	6	0.4	90	3.44
sCR1C3bfI -> sCR1 + iC3b + fI	840	56	12600	311.72
sCR1C4bfI -> sCR1 + dC4b + fI	840	56	12600	727.20
C4BPC4bfI -> C4BP + dC4b + fI	840	56	12600	2988.82
C3(H2O) + FB -> C3(H2O)FB	6	0.4	90	14.44
C3(H2O)FB + FD -> C3(H2O)FBFD	6	0.4	90	23.48
C3(H2O)FBFD -> C3(H2O)FB + FD	0.06	na	na	0.06
C3(H2O)FBFD -> C3(H2O)Bb + Ba + FD	840	56	12600	395.52
C3(H2O)Bb + C3 -> C3(H2O)BbC3	6	0.4	90	19.29
C3(H2O)BbC3 -> C3(H2O)Bb + C3	0.06	na	na	0.06
C3(H2O) + FH -> C3(H2O)FH	6	0.4	90	4.28
C3(H2O)FB + FH -> C3(H2O)FH + FB	6	0.4	90	3.89
C3(H2O)Bb + FH -> C3(H2O)FH + Bb	6	0.4	90	5.29
C3(H2O)FH + FI -> C3(H2O)FHFI	6	0.4	90	9.26
C3(H2O)FHFI -> dC3(H2O) + FH + FI	840	56	12600	730.03
A + C3 -> A + C3(H2O)	1	0.001	1000	0.09
aC3b + FB -> C3bFB	6	0.4	90	6.29

C3bFB -> dC3b + FB	0.06	na	na	0.06
C3bFB + FD -> C3bFBFD	6	0.4	90	14.15
C3bFBFD -> C3bFB + FD	0.06	na	na	0.06
C3bFBFD -> C3bBb + Ba + FD	840	56	12600	77.35
C3bBb + P -> C3bBbP	6	0.4	90	23.59
aC3b + P -> C3bP	6	0.4	90	30.01
C3bP -> dC3b + P	0.06	na	na	0.06
C3bP + FB -> C3bFBP	6	0.4	90	18.71
C3bFBP -> C3bP + FB	0.06	na	na	0.06
C3bFBP + FD -> C3bFBFDP	6	0.4	90	20.99
C3bFBFDP -> C3bFBP + FD	0.06	na	na	0.06
C3bFBFDP -> C3bBbP + Ba + FD	840	56	12600	735.42
C3bBb + C3 -> C3bBbC3	6	0.4	90	2.98
C3bBbC3 -> C3bBb + C3	0.06	na	na	0.06
C3bBbP + C3 -> C3bBbPC3	6	0.4	90	5.37
C3bBbPC3 -> C3bBbP + C3	0.06	na	na	0.06
sCR1 + C3bBb -> sCR1C3b + Bb	6	0.4	90	3.68
sCR1 + C3bBbP -> sCR1C3b + Bb + P	6	0.4	90	4.37
C3bBb + FH -> C3bFH + Bb	6	0.4	90	5.62
C3bBbP + FH -> C3bFH + Bb + P	6	0.4	90	6.65
C3bFH + FI -> C3bFHFI	6	0.4	90	5.01
C3bBb + aC3b -> (C3bBb)C3b	6	0.4	90	14.52
C3bBbP + aC3b -> (C3bBb)C3bP	6	0.4	90	47.79
(C4bC2a)C3b + C5 -> (C4bC2a)C3bC5	6	0.4	90	47.60
(C4bC2a)C3bC5 -> (C4bC2a)C3b + C5	0.06	na	na	0.06
(C4bC2a)C3bC5 -> (C4bC2a)C3b + C5a + C5b	840	56	12600	831.45
(C3bBb)C3b + C5 -> (C3bBb)C3bC5	6	0.4	90	6.76
(C3bBb)C3bC5 -> (C3bBb)C3b + C5	0.06	na	na	0.06
(C3bBb)C3bP + C5 -> (C3bBb)C3bPC5	6	0.4	90	47.97
(C3bBb)C3bPC5 -> (C3bBb)C3bP + C5	0.06	na	na	0.06

(C3bBb)C3bPC5 -> (C3bBb)C3bP + C5a + C5b	840	56	12600	59.50
dC3b + FH -> C3bFH	6	0.4	90	1.13
dC3b + sCR1 -> sCR1C3b	6	0.4	90	42.72

Appendix 6 Table of reactions in the C-ODE model ordered by sensitivity (the effect on the RMSE of the model fit).

Reaction	Sensitivity (∆ RMSE per 1% increase)
aC1 + C1inh -> C1inhC1	2.19E-02
aC3b -> dC3b	2.02E-02
(C3bBb)C3bP -> C3bBbP + dC3b	1.24E-02
C3bBbP + FH -> C3bFH + Bb + P	9.47E-03
aC1 + C4 -> aC1C4	8.39E-03
C3 -> C3(H2O)	8.18E-03
(C3bBb)C3bP + C5 -> (C3bBb)C3bPC5	6.60E-03
C3bBbP + aC3b -> (C3bBb)C3bP	5.44E-03
C3(H2O)Bb + FH -> C3(H2O)FH + Bb	3.78E-03
A + C3 -> A + C3(H2O)	3.77E-03
C4BP + C4b -> C4BPC4b	3.20E-03
C4bC2a + C3 -> C4bC2aC3	3.16E-03
C4BP + C4bC2a -> C4BPC4b + C2a	3.10E-03
aC1 + C4bC2 -> aC1C4bC2	3.10E-03
C4b + C2 -> C4bC2	3.05E-03
C3(H2O)FB + FD -> C3(H2O)FBFD	2.98E-03
C4bC2 -> C4b + C2	2.84E-03
(C4bC2a)C3b -> C4bC2a + dC3b	1.90E-03
C3(H2O)FB + FH -> C3(H2O)FH + FB	1.73E-03
C4BPC4b + FI -> C4BPC4bfI	1.67E-03
aC3b + P -> C3bP	1.52E-03
C3(H2O)Bb + C3 -> C3(H2O)BbC3	1.43E-03
C3bBb + FH -> C3bFH + Bb	1.14E-03
(C3bBb)C3bPC5 -> (C3bBb)C3bP + C5a + C5b	1.10E-03
C3bBb + P -> C3bBbP	1.02E-03
C4BPC4b -> C4BP + C4b	7.25E-04
A + C1 -> A + aC1	7.05E-04
C4bC2a + aC3b -> (C4bC2a)C3b	6.23E-04
C3(H2O) + FB -> C3(H2O)FB	6.12E-04
aC3b + FB -> C3bFB	5.37E-04
(C4bC2a)C3b + C5 -> (C4bC2a)C3bC5	4.64E-04
C3bBbP + C3 -> C3bBbPC3	3.38E-04

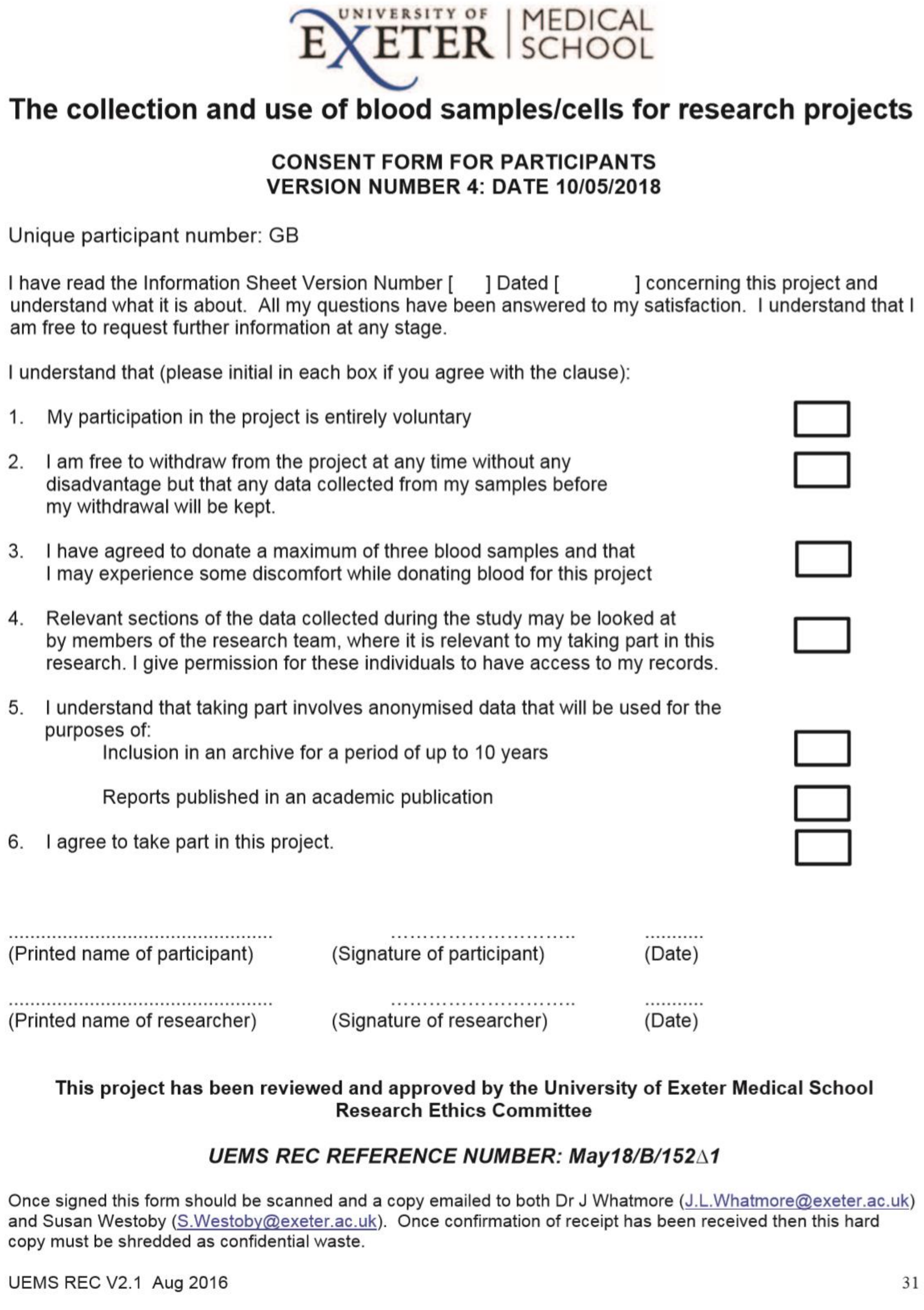
C3(H2O) + FH -> C3(H2O)FH	1.78E-04
C3bFH + FI -> C3bFHFI	1.67E-04
C1 -> aC1	1.44E-04
C3bFHFI -> iC3b + FH + FI	7.67E-05
C3(H2O)Bb -> C3(H2O) + Bb	7.20E-05
aC1C4 -> aC1 + C4a + C4b	7.13E-05
C3bFB + FD -> C3bFBFD	5.98E-05
C3bFBP + FD -> C3bFBFDP	5.20E-05
C3(H2O)FH + FI -> C3(H2O)FHFI	4.40E-05
aC1 -> C1	3.86E-05
C3(H2O)BbC3 -> C3(H2O)Bb + C3a + aC3b	3.55E-05
C3bBb -> dC3b + Bb	3.46E-05
C3bBb + C3 -> C3bBbC3	2.48E-05
aMBL + C4 -> aMBLC4	2.46E-05
C3bFB -> dC3b + FB	2.33E-05
C3(H2O)FB -> C3(H2O) + FB	2.33E-05
dC3b + FH -> C3bFH	2.13E-05
C4bC2aC3 -> C4bC2a + C3a + aC3b	2.04E-05
aC3b + FH -> C3bFH	1.90E-05
aMBL + C1inh -> aMBLC1inh	1.67E-05
C5b + C6 -> C5bC6	1.54E-05
C3bBbP -> dC3b + Bb + P	1.45E-05
A + MBL -> A + aMBL	8.08E-06
(C3bBb)C3bPC5 -> (C3bBb)C3bP + C5	7.76E-06
C3bP -> dC3b + P	6.95E-06
C3bP + FB -> C3bFBP	6.51E-06
C4bC2a -> C4b + C2a	6.36E-06
aMBL + C4bC2 -> aMBLC4bC2	5.11E-06
C5bC6C7 + C8 -> C5bC6C7C8	4.29E-06
(C3bBb)C3b -> C3bBb + dC3b	2.83E-06
C3bBb + aC3b -> (C3bBb)C3b	2.08E-06
(C3bBb)C3b + C5 -> (C3bBb)C3bC5	2.01E-06
C3bFBFD -> C3bBb + Ba + FD	1.96E-06
C3(H2O)BbC3 -> C3(H2O)Bb + C3	1.48E-06
C5bC6 + C7 -> C5bC6C7	1.11E-06
sCR1 + C4b -> sCR1C4b	1.07E-06
C3(H2O)FBFD -> C3(H2O)Bb + Ba + FD	9.76E-07
aMBLC4 -> aMBL + C4a + C4b	7.35E-07
sCR1 + C4bC2a -> sCR1C4b + C2a	6.41E-07
C4BPC4bfI -> C4BP + dC4b + fI	5.49E-07
sCR1 + aC3b -> sCR1C3b	4.36E-07

sCR1 + C3bBbP -> sCR1C3b + Bb + P	3.98E-07
C3bBbC3 -> C3bBb + C3a + aC3b	3.75E-07
C3bFBP -> C3bP + FB	3.75E-07
aMBLC4bC2 -> C4bC2a + C2b + aMBL	3.19E-07
dC3b + sCR1 -> sCR1C3b	2.93E-07
C4bC2aC3 -> C4bC2a + C3	2.93E-07
aC1C4bC2 -> C4bC2a + C2b + aC1	2.87E-07
C3bFBFDP -> C3bBbP + Ba + FD	2.65E-07
C3(H2O)FHFI -> dC3(H2O) + FH + FI	2.38E-07
sCR1C4b + FI -> sCR1C4bFI	2.34E-07
C5bC6C7C8 + C9 -> TCC	1.88E-07
sCR1C3b + FI -> sCR1C3bfI	1.81E-07
(C4bC2a)C3bC5 -> (C4bC2a)C3b + C5a + C5b	1.74E-07
C3(H2O)FH -> C3(H2O) + FH	1.60E-07
C3bBbPC3 -> C3bBbP + C3a + aC3b	1.57E-07
C3(H2O)FBFD -> C3(H2O)FB + FD	9.23E-08
(C4bC2a)C3bC5 -> (C4bC2a)C3b + C5	7.27E-08
aMBL -> MBL	7.10E-08
sCR1 + C3bBb -> sCR1C3b + Bb	6.29E-08
C3bBbPC3 -> C3bBbP + C3	4.09E-08
C3bFBFD -> C3bFB + FD	3.64E-08
aC1C4bC2 -> aC1 + C4bC2	3.04E-08
aMBLC4 -> aMBL + C4	1.91E-08
C3bFH -> dC3b + FH	1.02E-08
aC1C4 -> aC1 + C4	8.67E-09
sCR1C3bFI -> sCR1 + iC3b + FI	7.44E-09
sCR1C4bFI -> sCR1 + dC4b + FI	5.64E-09
aMBLC4bC2 -> aMBL + C4bC2	5.06E-09
(C3bBb)C3bC5 -> (C3bBb)C3b + C5a + C5b	4.86E-09
C3bBbC3 -> C3bBb + C3	4.39E-09
(C3bBb)C3bC5 -> (C3bBb)C3b + C5	2.29E-09
C3bFBFDP -> C3bFBP + FD	1.28E-09
C5bC6C7 -> C5bC6 + C7	8.38E-10
C5bC6C7C8 -> C5bC6C7 + C8	5.62E-10
TCC -> C5bC6C7C8 + C9	5.02E-10
C5bC6 -> C5b + C6	3.18E-11

Appendix 7 Table of species in the C-ODE model ordered by sensitivity (the effect on the RMSE of the model fit).

Species	Sensitivity (Δ RMSE per 1% increase)
C5a	0.05161
C3	0.038944
FH	0.035733
C3a	0.03158
TCC	0.030827
C1	0.027729
C1inh	0.025781
Ba	0.017575
C4BP	0.016905
C4	0.013029
C5	0.004588
C2	0.004214
FD	0.003385
FB	0.002252
P	0.002102
FI	0.001433
C4a	0.001145
MBL	2.87E-05
C6	2.85E-05
C8	3.14E-06
sCR1	1.68E-06
C7	1.11E-06
C9	1.97E-07
(C3bBb)C3b	0
(C3bBb)C3bC5	0
(C3bBb)C3bP	0
(C3bBb)C3bPC5	0
(C4bC2a)C3b	0
(C4bC2a)C3bC5	0
Bb	0
C1inhC1	0
C2a	0
C2b	0
C3(H2O)	0
C3(H2O)Bb	0
C3(H2O)BbC3	0
C3(H2O)FB	0
C3(H2O)FBFD	0
C3(H2O)FH	0
C3(H2O)FHFI	0
C3bBb	0
C3bBbC3	0
C3bBbP	0
C3bBbPC3	0
C3bP	0

C3bFB	0
C3bFBP	0
C3bFBFD	0
C3bFBFDP	0
C3bFH	0
C3bFHFI	0
C4BPC4b	0
C4BPC4bFI	0
C4b	0
C4bC2	0
C4bC2a	0
C4bC2aC3	0
C5b	0
C5bC6	0
C5bC6C7	0
C5bC6C7C8	0
aC1	0
aC1C4	0
aC1C4bC2	0
aC3b	0
aMBL	0
aMBLC1inh	0
aMBLC4	0
aMBLC4bC2	0
dC3(H2O)	0
dC3b	0
dC4b	0
iC3b	0
sCR1C3b	0
sCR1C3bFI	0
sCR1C4b	0
sCR1C4bFI	0



UNIVERSITY OF EXETER | MEDICAL SCHOOL

The collection and use of blood samples/cells for research projects

CONSENT FORM FOR PARTICIPANTS
VERSION NUMBER 4: DATE 10/05/2018

Unique participant number: GB

I have read the Information Sheet Version Number [ ] Dated [ ] concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I understand that (please initial in each box if you agree with the clause):

1. My participation in the project is entirely voluntary ☐
2. I am free to withdraw from the project at any time without any disadvantage but that any data collected from my samples before my withdrawal will be kept. ☐
3. I have agreed to donate a maximum of three blood samples and that I may experience some discomfort while donating blood for this project ☐
4. Relevant sections of the data collected during the study may be looked at by members of the research team, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
5. I understand that taking part involves anonymised data that will be used for the purposes of:
    Inclusion in an archive for a period of up to 10 years ☐
    Reports published in an academic publication ☐
6. I agree to take part in this project. ☐

.... (Printed name of participant) ………………………….. (Signature of participant) .... (Date)

.... (Printed name of researcher) ………………………….. (Signature of researcher) .... (Date)

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee

UEMS REC REFERENCE NUMBER: May18/B/152∆1

Once signed this form should be scanned and a copy emailed to both Dr J Whatmore (J.L.Whatmore@exeter.ac.uk) and Susan Westoby (S.Westoby@exeter.ac.uk). Once confirmation of receipt has been received then this hard copy must be shredded as confidential waste.

UEMS REC V2.1 Aug 2016 31

Appendix 8 Consent form for participants in the clinical investigation described in Chapter 6.