DOSCATs: Double Standards in Quantitative Proteomics

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Richard James William Bennett

September 2017

Acknowledgments

Firstly, I would like to thank my supervisor Professor Rob Beynon, whose expert guidance, advice, knowledge and support throughout my Ph.D. has been invaluable and allowed me to get to this stage. A big thank you also to my industrial supervisor Professor John Colyer at Badrilla for his continual support and encouragement.

Huge thanks to all past and present members of the Centre for Proteome Research who have made it such an enjoyable place to work and have been willing to help and offer advice whenever needed. I would particularly like to thank Stephen Holman and Philip Brownridge for their expert training in mass spectrometry, Guadalupe Gómez-Baena for her help with the meningitis work and Lynn McLean and Deborah Simpson for their help with protein expression and generally knowing things.

I would like to thank all my family and friends for humouring me when I talk to them about science and only occasionally asking when I am going to get a ‘proper job’. I would especially like to thank my parents, who have always encouraged me to pursue my ambitions and without whose sacrifices and support I would not be where I am today.

Finally, my greatest thanks go to my partner Laura, who has always been there for me without question and without whom this thesis would not be possible.

DOSCATs: Double Standards in Quantitative Proteomics

Richard J. W. Bennett

Abstract

Since its inception, the field of proteomics has shifted from being a qualitative discipline, generating long lists of proteins within a sample, to a quantitative one, where how much of a protein is reported. With the advent of systems biology, the routine analysis of biomarker levels, and the requirement for robust, reliable data comparable between different laboratories, the importance of absolute quantification, where proteins are quantified in absolute titre, is becoming increasingly important. There are two commonly used techniques for absolute protein quantification, based on either mass spectrometry (MS) or immunochemical techniques such as western blotting (WB). MS is generally considered the gold standard technique for quantification, but WB can offer greater sensitivity and is much more accessible to researchers. Neither are intrinsically quantitative techniques and so rely on standards; either isotope labelled peptides or recombinant proteins bearing an epitope are used for MS or WB respectively. To improve the robustness and reproducibility of quantitative data it would be advantageous to apply both techniques for orthogonal quantification, but due to the very different calibration standards, workflows rarely overlap. DOSCATs (Double Standard conCATamers) are novel calibration standards that can unite MS and WB workflows, allowing for the quantification of direct comparison of quantitative data between the two platforms. DOSCATs, based on QconCAT technology, combine a series of epitope sequences concatenated with peptides in a single artificial protein. Stable isotope labelled peptide for MS analysis are released upon digestion with an enzyme such as trypsin, and intact DOSCATs act to bear multiple epitopes for WB. Also included were restricted proteolysis sites that allow for a mobility shift within WB, lending greater flexibility to the standard. The aim of this thesis was to develop and optimise the use of DOSCAT technology so that they could be used to quantify target proteins in both quantitative platforms.

A DOSCAT protein was designed and constructed to quantify five proteins of the NF-κB pathway. The DOSCAT was expressed and purified and the 9/13 peptides and 3/5 epitopes included in the sequence were observed by MS and WB respectively, demonstrating the proof of concept. However, restricted proteases performed poorly and three antibodies were discontinued by the manufacturer, so a second iteration of the NF-κB DOSCAT was designed. This was used to calibrate quantification by selected reaction monitoring MS (SRM-MS) and automated capillary WB. For three target proteins, protein fold change and absolute copy per cell values measured by MS and WB were in excellent agreement. Building on this success, another DOSCAT was built for six proteins implicated to be indicative of paediatric *Streptococcus pneumoniae* meningitis infection. All six proteins were quantified by SRM-MS although QWB failed to quantify two targets as either DOSCAT or endogenous protein was not detected. SRM-MS data agreed very well with previous datasets generated for the same samples by label-free MS and QWB using full length standards, however, absolute values for DOSCAT calibrated QWB were inconsistent. This could be due to antibodies recognising DOSCAT and endogenous protein with different affinities.

This work demonstrates that DOSCATs can be used as multiplexed, dual purpose standards to unite MS and WB workflows. The DOSCAT approach has the potential to generate reliable quantitative information particularly relevant for systems biology studies and contribute to the desired increase in reproducibility of biological research.

Table of Contents

List of Figures

List of Tables

Abbreviations

Ab	Antibody
ABM	Acute bacterial meningitis
ACN	Acetonitrile
Ambic	Ammonium bicarbonate
AQUA	Absolute quantification peptide
Arg	Arginine
BSA	Bovine serum albumin
CCD	Charge coupled device
CID	Collision induced dissociation
CONSeQuence	Consensus predictor for Q-peptide sequence
cpc	Copies per cell
CSF	Cerebrospinal fluid
CV	Coefficient of variance
Cys	Cysteine
Da	Dalton
DDA	Data dependent acquisition
DIA	Data independent acquisition
DOSCAT	Double Standard Concatamer
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme linked immune sorbent assay
ESI	Electrospray ionisation
Glu-Fib	[Glu1]-Fibrinopeptide B
His	Histidine
HR/AM	High resolution/accurate mass
HRP	Horseradish peroxidase
IAM	Iodoacetamide
Ig	Immunoglobulin
IP	Intellectual property
IPTG	Isopropyl β-D-1-thiogalactopyranoside
iTRAQ	Isobaric tagging for relative and absolute quantification
IκB	Inhibitor of kappa B
LC	Liquid chromatography
Lys	Lysine
m/z	Mass-to-charge ratio
mAb	Monoclonal antibody

MALDI	Matrix Assisted Laser Desorption/Ionization
MASCOT	Mascot search engine, developed by Matrix Science
M-DOSCAT-i	Meningitis DOSCAT immunogen
M-DOSCAT-S	Meningitis DOSCAT standard
mM	Millimolar
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSIA	Mass spectrometric immunoassays
MW	Molecular weight
MWCO	Molecular weight cut-off
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NFκB-DOSCAT	NF-κB DOSCAT iteration 1
NFκB-DOSCAT -2	NF-κB DOSCAT iteration 2
OD	Optical density
pAb	Polyclonal antibody
PCR	Polymerase chain reaction
PEAKS	Peaks Studio software, developed by Bioinformatics Solutions Inc.
PrEST	Protein Epitope Signature Tag
PRM	Parallel reaction monitoring
PSAQ	Protein standard for absolute quantification
PVDF	Polyvinylidene fluoride
QconCAT	Quantification concatamer
Q-peptide	Quantotypic peptide
QqQ	Triple quadrupole
Q-TOF	Quadrupole time-of-flight
QWB	Quantitative western blotting
RT	Room temperature
RV3C	Human rhinovirus 3C protease
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SID	Stable isotope dilution
SIL	Stable isotope label
SILAC	Stable isotope labelling with amino acids in cell culture
SISCAPA	Stable isotope standards and capture by anti-peptide antibodies
SPP	*Streptococcus pneumoniae* positive
SRM	Selected reaction monitoring
SWATH	Sequential window acquisition of all theoretical mass spectra
TBS	Tris buffered saline
TCA	Trichloro acetic acid
TEV	Tobacco etch virus protease

TFA	Trifluoracetic acid
TMT	Tandem mass tags
TNFα	Tumour necrosis factor alpha
TOF	Time of flight
TVMV	Tobacco vein mottling protease C4
USP	Unique selling point
UV	Ultraviolet
WB	Western blotting

Chapter 1: Introduction

1.1. The rise of targeted absolute protein quantification

Proteomics is focussed on the study of all proteins within an organism or system at a particular time and under specific conditions. When the term proteomics was first coined in the mid-1990s, the burgeoning field was mostly concentrated on the qualitative analysis of proteins, generating lists of proteins contained in samples (Wilkins et al., 1996). Since this time proteomics has advanced considerably in terms of technology and focus. In the fields infancy, proteomic experiments were conducted using 2D-electrophersis gels (N. G. Anderson et al., 1996), which although permitted the large-scale cataloguing of proteins, was laborious, costly, and technically limited in terms of sensitivity. Subsequently, mass spectrometry (MS) and affinity based assays were developed and utilised to study the proteome of complex samples in greater depth.

As development of instrumentation and experimental workflows progressed, the focus of experiments shifted from the qualitative to the quantitative analysis of proteins. This gives an insight into not only what proteins are present in a sample but at what level they are present. Quantitative proteomics can be categorised into two outputs: relative and absolute quantification. In relative quantification experiments, protein abundance in a sample is reported relative to the amount of the same protein in another sample in terms of a fold-change increase or decrease. This type of experiment is commonly used to study global protein abundance changes between, for example, healthy control and disease state samples, or samples grown under different growth conditions. Whilst useful for the direct comparison of a limited number of samples in a single experiment, relative quantification is of limited value when it comes to, for instance, the determination of biomarkers protein levels, formation of mathematical models of protein interactions (an important facet of systems biology) and the elucidation of subunit stoichiometry in protein complexes. Additionally, quantitative values from relative quantification experiments are limited to the particular laboratory and samples they were derived from and so inter-laboratory comparisons cannot be performed. Therefore, interest in absolute quantification, in which proteins are quantified without reference to another sample in terms of an absolute titre or copies per cell, has massively increased in the past few years. Absolute quantification is commonly associated with a targeted approach, in which a group of specific proteins is preselected for analysis. This contrasts with the global proteomic methods that are used in qualitative or relative quantification experiments. For example, in biomarker discovery pipelines, relative quantification approaches might be used in the discovery phase to identify potential proteins based on the highest abundance change, but targeted absolute quantification of a selection of these proteins is required on a greater sample size for verification and validation. As will be discussed, targeted absolute quantification delivers greater sensitivity, selectivity and multiplexing capabilities required for biomarker validation as well as the myriad of other applications associated with the approach.

There are multiple strategies and workflows available to researchers to achieve both relative and absolute quantification, and these can be broadly split into those utilising either MS or protein binders to detect and quantify proteins. Both such approaches have advantages and disadvantages inherent to them that affects the quality and robustness of the quantitative data produced. This thesis is centred around the creation of a novel calibration material, DOuble Standard ConCATamers (DOSCATs), that can act as a bridge between MS and affinity assay methodologies, uniting the two disparate workflows to improve the accuracy, reproducibility and sensitivity of targeted absolute quantification experiments. To understand how this can achieved, MS and affinity methodologies will be thoroughly reviewed with regards to fundamental technological platforms, differing experimental workflows and reagents utilised. The motivations for uniting the two technologies will then be further explored, the principles behind DOSCATs explained and the aims of this thesis regarding the development of DOSCATs outlined.

1.2. Protein quantification using protein binders

For protein detection and quantification, assays based on molecules that bind to proteins with high selectivity are routinely used. Their use pre-dates mass spectrometry and their relative low cost and ease of use means that quantitative affinity-based assays are still much more prevalent than MS assays, despite arguably providing less robust and reproducible data.

This section will review the different types of protein binders that are now available and the key affinity-based techniques, collectively known as immunoassays, which rely on protein binders for protein quantification.

1.2.1. Types of protein binder

By far the most widely used protein binder in biochemical research is the antibody. Discovered in the late 19th century, their ability to bind almost any protein target with high specificity has led to their extensive use in molecular biology as a research tool, which extends as far back as the 1950s. More recently, antibodies have been employed as therapeutic agents due to possessing characteristics such as high specificity, activity and favourable pharmacokinetics (Wold et al., 2016).

The antibody is an integral part of an organism's immune response to foreign antigens due to the ability for antibodies to be produced to bind almost any target. This has come about through the evolution of a complex set of mechanisms that alter the DNA of individual B cells, the type of lymphocyte where antibodies are produced. Antibodies belong in the immunoglobulin (Ig) superfamily, of which there are five classes with the most common being immunoglobulin G (IgG) (Schroeder et al., 2010). Structurally, all Igs share similar properties in that they contain a constant (Fc) region that provides a scaffold for the antigen binding variable (Fab) regions (Figure 1.1). Igs consist of two heavy and two light chains, each of which are made up of at least one constant domain and one variable domain. Disulphide

bridges formed between conserved cysteine residues in the constant domain link together heavy and light chains. At the end of the variable domain resides a hypervariable region made up of variable loops linked together by β-strands; it is this that confers specificity to an antigen. Variability in amino acid sequence is achieved by the random rearrangement of gene segments that code for the variable region in a process known as V(D)J recombination (Tonegawa, 1983). A further genetic mechanism, somatic hypermutation, occurs after B cells are exposed to an antigen and stimulated to proliferate, during which areas of the Ig gene family that code for hypervariable regions are highly prone to mutations leading to amino acid sequence diversity (Schroeder et al., 2010).

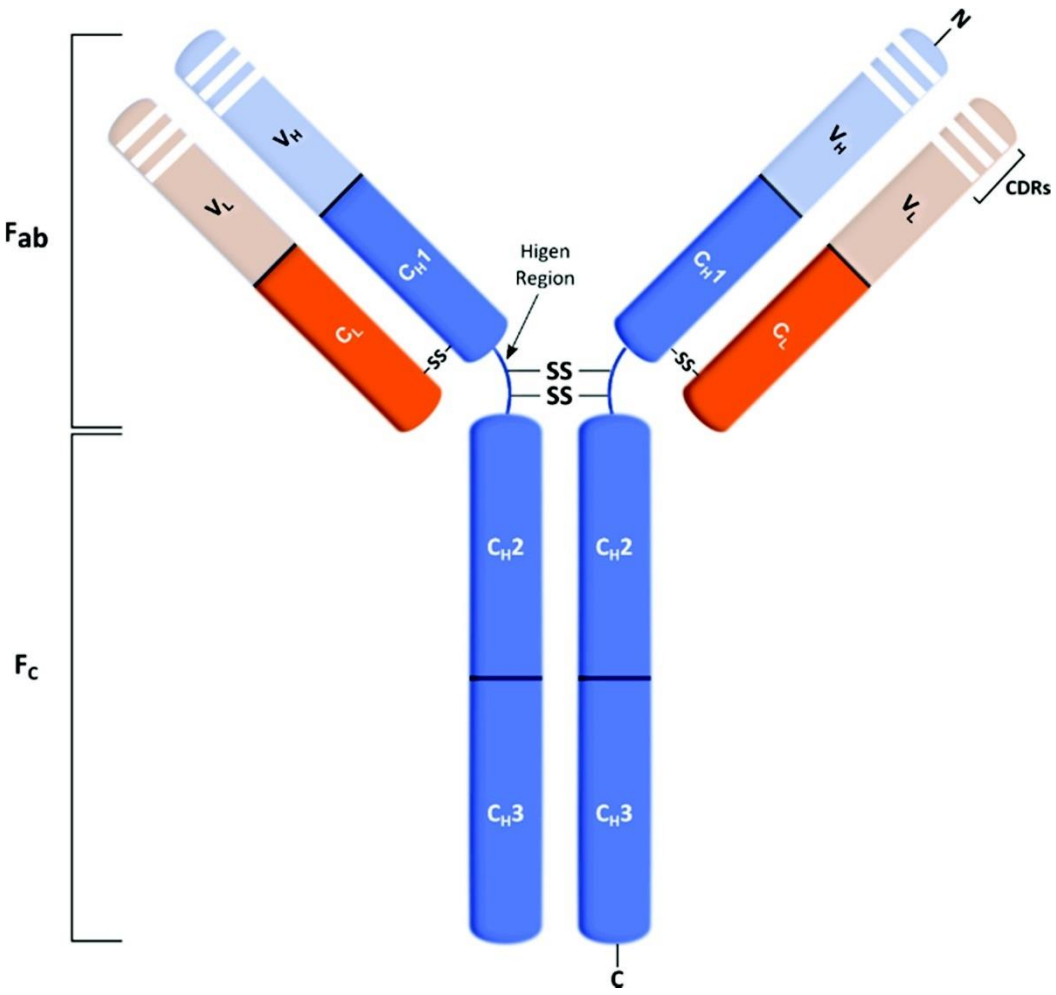


Figure 1.1 Structure of an immunoglobulin G (IgG). The molecule is composed of two light (L) and two heavy (H) chains, each comprising one variable (V) domain and at least one constant (C) domain. Chains are linked by disulphide bonds (SS). Figure adapted from Loureiro et al. (2015).

To create antibodies against protein targets for use as research tools, antigens relating to the target protein are introduced to an animal (most often rabbit or mice), whose B cells are stimulated to produce antibodies against the target. The antigen (or immunogen) can be the entire protein or a peptide fragment; when a peptide is used, it is usually coupled to a larger protein such as bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH) to ensure an immune response is elicited. Antibodies used in research can be divided into polyclonal and monoclonal, describing the process in which they were generated. Polyclonal antibodies (pAb) are derived from an animal's blood 2 – 3 months after immunisation, and are a heterogenous collection of antibodies from multiple B-cell lineages that recognise different epitopes on the same antigen. The crude immune-sera from an animal can be subjected to affinity purification to isolate specific antibodies, which can improve pAb specificity and minimise cross reactivity. pAbs are relatively cheap and quick to produce and through binding multiple epitopes they generally have a high overall affinity for their target and are more tolerant of minor epitopes changes such as glycosylation. However, pAbs produced from a single batch are finite; pAbs produced using the same immunogen in another animal will consist of an entirely different population of antibodies with a potentially very different affinity and specificity to the target. Moreover, the presence of many non-specific antibodies in anti-sera can lead to problems with cross reactivity and background signal. For monoclonal antibody (mAb) production, B cells from immunised animals are removed and fused with a myeloma cell to produce immortal hybridoma cells (Köhler et al., 1975). These are cultured, cloned and screened for the effectiveness of the antibodies they produce. Clones that produce the most effective antibody are cloned and cultured in isolation, providing a constant and consistent source of production. In contrast to pAbs, mAbs are homogenous and are highly specific to a single epitope. This massively reduces the probability of cross reactivity and improves reproducibility across experiments, which is advantageous for quantitative assays. Their method of production also means that mAbs are less prone to batch to batch variability and as such the consistency of performance can be more certain. However, due to the requirement to produce and clone hybridomas, mAbs are more expensive and time-consuming to produce than pAbs. They are also susceptible to small epitope changes preventing binding. Despite their widespread use, there are many shortcomings in using both types of antibodies as protein binders. They are inherently unstable and so must be used under physiological conditions and stored at low temperatures during their short shelf-life. They are slow and expensive to generate in animals and difficult to generate *in vitro*. The requirement of animals for Ab generation raises ethical and welfare issues and means that Abs can be raised only to antigens that are immunogenic, ruling out many potential targets and accounting for why Abs are only commercially available for a small fraction of the proteome.

The development of protein engineering technologies alongside the development of combinatorial libraries and *in vitro* selection techniques has led to the emergence of novel classes of protein binders generated without the need for animal immunisations that confer

significant advantages over antibodies (Uhlén et al., 2009). One such class are nucleic acid based binders known as aptamers (Stoltenburg et al., 2007). They rely on the principle that short DNA or RNA chains of under 100 nucleotides will fold in predictable three-dimensional structures capable of binding to target molecules with high affinity and specificity. Aptamers have been used in immunoassays (Q. Li et al., 2017; Shin et al., 2010), protein purification (Jia et al., 2016), point-of-care diagnostics (Dhiman et al., 2017) and as therapeutics (Ng et al., 2006). Aptamers are selected *in vitro* from large libraries of 10^{13} to 10^{15} random oligonucleotide sequences using systematic evolution of ligands by exponential enrichment (SELEX) (Tuerk et al., 1990). The target molecule is introduced to the library, which by chance will contain a population of oligonucleotides that have the necessary structure to bind the target. Unbound oligonucleotides are removed and bound oligonucleotides eluted and amplified by PCR using primers based on constant regions that flank each random sequence. The resultant enriched library is used for the next round of SELEX, and multiple rounds are performed until only nucleic acid species with the highest affinity and specificity for the target remain. These can be used as a heterogenous aptamer pool or individual aptamers can be isolated and sequenced to provide a homogenous population with rigid specificity to a particular epitope.

Protein based binders have also emerged due to the development of *in vitro* selection techniques such as phage display (G. Smith, 1985) and ribosome display (Hanes et al., 1997). Antibodies are difficult to express recombinantly as few traditional expression hosts contain the cellular machinery and environment required to fold and link subunits. However, smaller antibody fragments can be expressed successfully with common examples being the entire Fab region or single-chain variable fragment (scFv), which are variable regions from the heavy and light chains connected by a short flexible amino acid peptide linker (Frenzel et al., 2013). Other types of binders based on protein scaffolds containing a variable region with high affinity and specificity to a target have also been developed (Gebauer et al., 2009; Mouratou et al., 2015). Proteins that demonstrate very high stability, high solubility and high production yield in recombinant expression systems are selected as scaffolds. Libraries of scaffold proteins can be created, each with a random sequence of residues within the variable region that can have many different specificities. For recombinant antibodies, libraries contain the genetic information of variable regions of many different antibodies. These libraries can then be used with phage display (Clackson et al., 1991; Mouratou et al., 2015; Schirrmann et al., 2011) or ribosome display (Edwards et al., 2012; Grimm et al., 2011) techniques to select binders from the library with required specificity and affinity for a target. There have been numerous protein scaffolds structures developed for use as binders, some of which have become commercially available for numerous applications (Dias et al., 2017). The utility of some of these binders in immunoassays as either the primary or secondary detection molecule has been demonstrated (Jeong et al., 2016; Straw et al., 2013).

Aptamers, recombinant antibodies and protein scaffold binders hold many potential advantages over antibodies (Stoltenburg et al., 2007). *In vitro* generation removes the need for animals and permits the use of small non-immunogenic or toxic molecules as immunogens. Once selected for, they can be chemically synthesised or recombinantly expressed very cheaply to provide a consistent source of material, which, due the inherent stability the material used, has a prolonged shelf life and can be stored at ambient temperature. Unlike antibodies, they can all function in non-physiological conditions and their relatively small size allows access to epitopes that larger antibodies may be sterically blocked from binding.

Despite these clear advantages, the uptake of these novel binders by the research community has been slow and they are not yet widely used (Rozenblum et al., 2016). In the case of aptamers this is in part due to technical deficiencies in the SELEX process, which requires considerable time- and labour-consuming optimisation to produce aptamers with affinity and specificity comparable to antibodies (Kanagawa, 2003; Musheev et al., 2006). The use of next generation sequencing to better characterise aptamers and improvements in the SELEX procedure are hoped to lead to the consistent generation of high-affinity in an efficient manner (Gotrik et al., 2016; Rozenblum et al., 2016). Moreover, they are all still relatively young technologies and the standardisation of their production, optimisation and utility in immunoassays has not yet transpired. As this occurs and their value becomes apparent it is likely their uptake by the research community will accelerate, but for now and despite their limitations antibodies remain the first choice for protein detection.

1.2.2. Quantitative immunoassays

Immunoassays simply describe assays that use protein binders to select a target antigen from a complex sample with a high degree of specificity. Immunoassays can be made quantitative through the use of calibration standards, which generally take the form of recombinant proteins that resemble or are identical the target analyte. The two most widely used assays for targeted absolute protein quantification, western blotting (WB) and Enzyme Linked Immunosorbent Assay (ELISA) will be discussed here. In the following descriptions antibodies are defined as the primary binder, although any of the protein binders described above could and have been used instead.

Western blotting was first described in the late 1970s (Towbin et al., 1979) and the basics have not changed much since. Proteins are separated based on their molecular weight by SDS-PAGE and transferred to a protein binding membrane such as nitrocellulose or PVDF by electrotransfer. Non-specific sites on the membrane are blocked followed by incubation with a primary Ab specific to the target and a horseradish peroxidase (HRP) conjugated secondary Ab specific to the Fc region of the primary antibody. Exposure of HRP to a peroxide substrate yields an enhanced chemiluminescent (ECL) signal that is detected by X-ray film. In its original form, WB is a semi-quantitative rather than a quantitative technique due to the limited dynamic range afforded by film detection, the high variability of the protein electrotransfer step and the

poor throughput and reproducibility resultant from the many manual handling steps required (Koller et al., 2005). Since first described, there have been several technological improvements that have aimed to improve reproducibility and make the technique truly quantitative.

Improving in the throughput and reproducibility of protein transfer from gel to membrane have come from the development of semi-dry transfer methods such as iBlot, which can reduce the transfer time from gel to membrane to under 10 minutes (J. M. Silva et al., 2014). Reversible stains have been developed that allow for the assessment of protein transfer completeness before Ab probing takes places (Antharavally et al., 2004; Colella et al., 2012). Other advances have been directed towards improving detection technologies. Cooled charged coupled device (CCD) cameras have been used rather than film for hugely improved linear dynamic range, although this approach can still be limited by signal saturation. Rather than chemiluminescence, detection based on fluorescence has been developed by conjugating near infrared dyes to secondary antibodies (Zellner et al., 2008). This massively increases the dynamic linear range of detection, although sensitivity can be limited. Fluorescent signal is stable for long periods of time, improving reproducibility. Moreover, different fluorescent dyes can be used together for multiplexed detection; for example, target protein and a loading control such as GAPDH can be analysed together, overcoming the need to chemically strip and re-probe the membrane. Fluorescent nanoparticles known as quantum dots have been also been used as a multiplexed detection method (Gilroy et al., 2010; Scholl et al., 2009). Their surface can be functionalised to bind primary Abs, removing the need for secondary antibodies and giving excellent sensitivity. These technological advances have made quantitative western blotting (QWB) somewhat routine, however, care must be taken in the experimental design, optimisation and data stages to ensure quantitative data are reliable (Gassmann et al., 2009; Mollica et al., 2009). Quantification using WB is usually relative, with loading controls such as GAPDH or β-actin used to normalise variations in protein load across different lanes. Absolute quantification of target proteins by WB using recombinant protein standards for calibration has been reported using detection methods based on chemiluminescence and CCD camera (Murphy et al., 2009, 2011) or fluorescence (Bromage et al., 2009; Y. V Wang et al., 2007). WB has also been employed for proteome-wide quantification through the creation of a *Saccharomyces cerevisiae* fusion library, in which each open reading frame was tagged with a tandem affinity purification (TAP) tag and protein detection was achieved through the use of a single anti-TAP tag antibody (Ghaemmaghami et al., 2003).

Other systems have attempted to redesign the western blot process to remove the SDS-PAGE and transfer steps as well as automate the entire procedure (Figure 1.2). In one approach known as µWesterns, microfluidic systems and microchip electrophoresis are used to perform the entire process on tiny glass slides (Hughes et al., 2012). Samples are separated by electrophoresis over a short distance (3 mm) in benzophenone-derivatized polyacrylamide

gel, which upon irradiation with UV light captures the proteins *in situ*. Antibody reagents are introduced and fluorescent signal is imaged by a CCD camera. Another approach, marketed as 'Simple Westerns' by ProteinSimple, uses capillaries pre-coated with a benzophenone derivative to separate proteins by electrophoresis, which are then crosslinked to capillary walls by UV irradiation for *in situ* analysis. The separating gel is washed away and antibodies are flowed through the capillary before luminol and peroxide are used to generated chemiluminescent signal, which is detected by a CCD camera (Nguyen et al., 2011; O'Neill et al., 2006; Rustandi et al., 2012).

Compared to 'classic' WB these processes are fast (µWestern 10 – 60 m, Simple Western 3 h), high throughput, use tiny amounts of samples and Ab, have high sensitivity and dynamic linear range and are more reproducible due to the minimisation of manual sample handling steps and the elimination of the protein transfer step. Both platforms have demonstrated their utility since becoming commercially available; µWesterns have recently been used quantify proteins in single cells loaded in individual microwells (Sinkala et al., 2017) and Simple Western technology has been used for absolute quantification of proteins with a high degree of accuracy and reproducibility at sub picogram per nanogram of cell lysate loaded (J. Q. Chen et al., 2013; Hamm et al., 2015; Loughney et al., 2014; Rustandi et al., 2015; Xu et al., 2017).

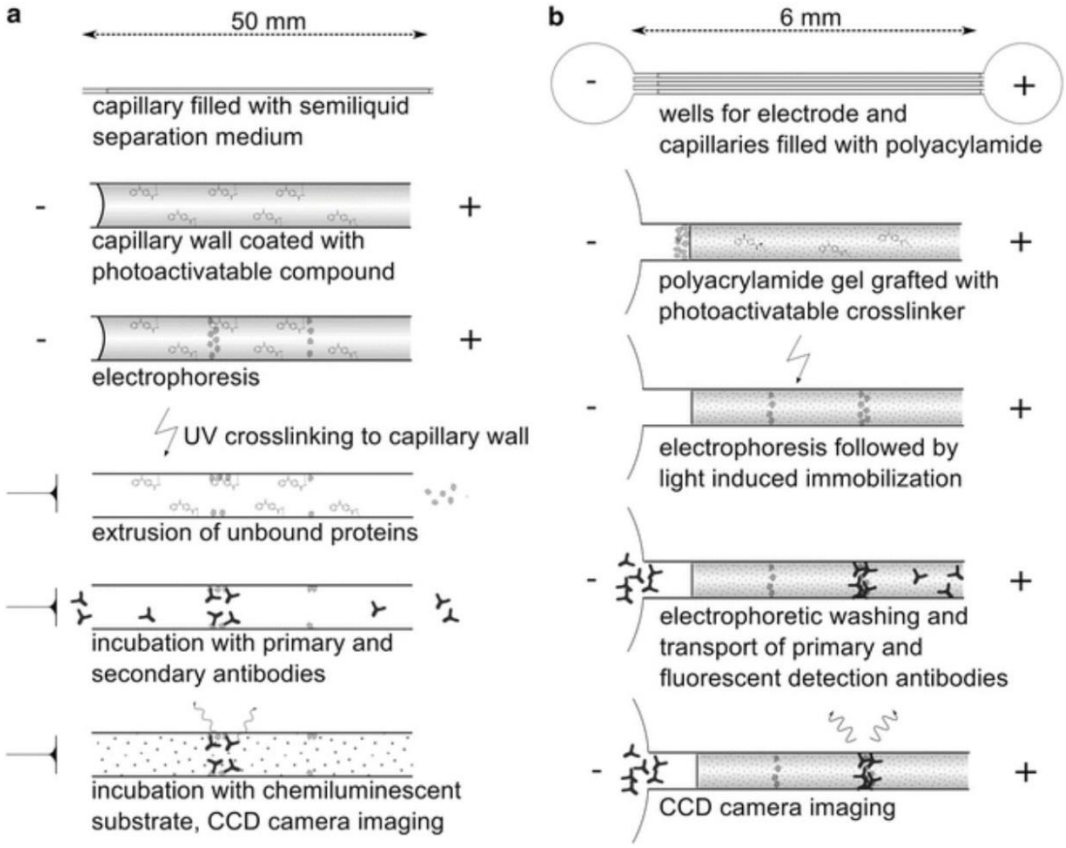


Figure 1.2 Schematics for automated western blotting systems. a) Capillary automated WB system developed by ProteinSimple and marketed as Simple Westerns (Nguyen et al., 2011). Protein separation and detection by antibodies takes place within distinct capillaries, and instruments can process up to 96 capillaries in a single automated run. b) Microfluidic western blotting system on microchips, marketed as µWesterns (Hughes et al., 2012). Top panel shows three separation channels and other panels close-ups of a single channel; each microchip can hold 48 separation channels. Both devices can separate proteins by size or isoelectric focusing. Figure reproduced from Kurien and Towbin (2015).

The ELISA, first developed in 1971 (Engvall et al., 1971; Van Weemen et al., 1971), has become the gold standard for protein quantification using protein binders due to the sensitivity, accuracy and throughput it affords. As such, diagnostic grade ELISAs against many targets have been developed and approved for clinical use. ELISAs are performed on 96-well polystyrene plates that can passively bind antigens or antibodies. Detection is achieved using a label conjugated antibody, with enzyme labels such as HRP or alkaline phosphatase the most common to produce a colorimetric, chemiluminescent or fluorescent signal.

Different formats of ELISAs have been developed termed direct, indirect and capture (or sandwich) assays (Wild et al., 2013). In both direct and indirect assays, the antigen is bound to the plate surface and detected by either a labelled primary antibody (direct assay) or a primary antibody and labelled secondary antibody (indirect assay). For sandwich assays, a capture antibody is bound to the plate surface and specifically binds the target antigen when a sample is passed over. Unbound molecules are then washed away and a primary and labelled secondary antibody are used to detect the captured antigen. There is an additional form of ELISA that the above detection formats can be adapted to, known as competitive ELISA. Here, an unlabelled antibody is incubated with a sample and antibody/antigen complexes are formed, which are then added to a well pre-coated with the target antigen. Unbound antibodies (i.e. those that have already bound antigen in the sample) are washed away and bound antibodies are detected. This means that the signal output in inversely proportional the amount of target analyte in the sample. Competitive ELISA is especially useful for use with crude or complex samples as pre-purification of the antigen is not required.

Direct ELISA is the fastest technique with the fewest steps, but specific labelled primary Abs are required for every target and the lack of signal amplification reduces assay sensitivity. Indirect assays offer this signal amplification to improve sensitivity but the addition of a secondary antibody raises the chance of cross-reactivity, resulting in non-specific signal. Sandwich assays offer the highest sensitivity and specificity as two distinct Abs are used for capture and detection. However, this can also be a weakness as these Ab pairs need to be identified and very well validated to minimise non-specific binding, which would result in inaccurate quantification. Such prior Ab validation is vitally important as unlike WB there is no sample separation step to check that Abs are binding to a protein at the expected MW. As such, high-grade ELISAs are very time-consuming and expensive to set up and usually limited to commercially available kits to clinically relevant targets.

The classic ELISAs formats are suited to high throughput analysis against a single target protein. More recently, multiplexed ELISA systems have been developed that allow the detection and quantification of multiple targets at once. These systems can split into planar and suspension multiplex assays (Tighe et al., 2015). In suspension assays antibodies are immobilised onto fluorescently activated plastic or magnetic microbeads, with the fluorescent signature of a bead directly relating to the antibody bound to it. The microbeads are then mixed with the sample in the liquid phase, where analyte is captured by the immobilised antibodies.

Detection antibodies labelled with a reporter dye are then added. Using flow cytometric methods, each bead is analysed individually by lasers which excite the reporters so that the type of bead (and so specificity of antibody immobilised to it) and the signal is ascertained. Planar assays are analogous to protein microarrays, which have their origins in the fields of genomics and transcriptomics, where they have been extensively used for gene expression analysis. Such arrays allow for the global quantitative analysis of proteomes. They are much like a sandwich ELISA assays; high affinity capture ligands to multiple different targets are immobilised onto a solid phase such as a microtitre plate or functionalised glass slide, over which a sample is passed. Detection antibodies bind captured analytes and signal generated though through direct or indirect methods.

Both types of multiplex approaches offer advantages over singleplex assays not only in terms of increased throughput but also lower sample input and increased dynamic range. Multiplex assays have been commercialised (Gupta et al., 2016; Tighe et al., 2015); available planar assays on the market include MULTI-ARRAY technology (Meso Scale Discovery) and ImmunoCAP (Thermo Scientific), and commercially available suspension assays include xMAP (Luminex), Cytometric Bead Array (BD Bioscience) and Multi Bead (Enzo Life Science). However, these products all require expensive instrumentation and suffer the same limitations as singleplex ELISA in that they require highly validated antibodies for specific detection. Furthermore, problems with antibody cross-reactivity becomes more prevalent as multiplexing is increased, limiting the amount of proteins that can be detected without compromising accuracy. Novel techniques to mitigate the effect of cross reactivity are currently being explored (Juncker et al., 2014). For example, ‘DNA barcoding’ techniques such as the proximity ligation assay (Lundberg, Thorsen, et al., 2011) or the proximity elongation assay (Lundberg, Eriksson, et al., 2011) use DNA molecules as reporters for specific binding. Capture and detection Abs in a sandwich assay are tagged with DNA strands that upon both Abs binding to the same target molecule are brought into proximity, permitting their ligation or hybridization by complementary sequences. These newly formed DNA strands can act as a reporter signal upon amplification by PCR. Even with these advances, compared to classic ELISA techniques, multiplex technologies are not as mature meaning that the regulatory guidelines for the validation of multiplex biomarker assays are not in place and as such there are few clinically available multiplex assays (Tighe et al., 2015).

Both WB and ELISAs can be employed for targeted protein quantification to great effect, but it should be noted that the key to accurate and reproducible immunoassays in any format is a well validated protein binder. Poor reproducibility of published results has become a major issue in research (Begley et al., 2012; Prinz et al., 2011), and this in part has been attributed to poor quality antibodies, with studies finding that many commercially available antibodies or assay kits lack specificity or otherwise do not meet minimal required standards (Berglund et al., 2008; Michel et al., 2009; Rifai et al., 2013). Additionally, some antibodies are application specific and validation for one application does not guarantee success on another; for

example, an antibody that works in a western blot where proteins are denatured might not work in an ELISA where proteins are in their native state and so epitopes possibly sterically obscured. The research community has started to tackle this problem by developing standard operating practices for antibody validation (Acharya et al., 2017; Uhlen et al., 2016) and creating tools to rank antibodies that have been validated (Björling et al., 2008) or previously cited in peer reviewed papers (Helsby et al., 2014). Such efforts coupled with the development of antibody alternatives should increase the reliability of protein binders and ensure that immunoassays continue to be considered for target protein quantification.

1.3. Protein quantification by mass spectrometry

Mass spectrometry has emerged as the key technology in proteomics, due to its ability to, amongst other applications, rapidly identify and quantify vast number of biomolecules in complex samples, identify sites of post translational modifications and elucidate the structure of protein complexes (Aebersold et al., 2016). This section will give a brief overview of the fundamentals of biomolecular MS, detail the strategies and workflows developed to enable protein quantification by MS and finally describe the different data acquisition strategies that have been established to enable MS-based targeted quantitative assays.

1.3.1. Biomolecular mass spectrometry

At the most basic level, mass spectrometers consist of an ion source, a mass analyser that measures the mass to charge ratio (*m/z*) of ionised molecules and a detector that registers the number of ions at each *m/z* value. The development of two so-called soft ionisation techniques, matrix-assisted laser desorption/ionisation (MALDI) (Karas et al., 1987; Tanaka et al., 1988) and electrospray ionisation (ESI) (Nohmi et al., 1992; Whitehouse et al., 1985; M Yamashita et al., 1984; Masamichi Yamashita et al., 1984) facilitated the routine use of MS for biomolecular analysis. This is because mass spectrometers can only detect and measure the mass of a molecule after it has been converted into the gas phase; these techniques enabled the ionisation and transfer of biomolecules into the gas phase without fragmentation so that relatively large molecules such as proteins and peptides could be analysed by MS.

MALDI uses energy from the laser excitation of a matrix to produce singly charged ions from large molecules. A sample is co-crystallised with a matrix solution that strongly absorbs at the wavelength of the laser beam. Irradiation of the sample mixture with a laser pulse desorbs and ionises the matrix material, transferring matrix and analyte into the gas phase. At this stage protons are transferred to analyte molecules, and analyte ions progress through the mass spectrometer. MALDI has a rapid sampling speed and is tolerant to sample contamination, meaning that it has been used for multiple applications (Chang et al., 2007). However, it is a notoriously variable technique (Szájli et al., 2008) and the non-continuous nature of sampling makes it incompatible with in-line liquid chromatography (LC) separation, meaning that MALDI is not routinely used in quantitative proteomic analysis.

The mechanism of ESI can be broken down into three major stages (Kebarle et al., 2009) (Figure 1.3). Firstly, the sample solution is passed through a capillary tip, through which a high voltage (2.5 – 4.0 kV) is applied. This causes the formation of charged droplets in an elongated cone shape known as a Taylor Cone. Secondly, solvent evaporation (caused by elevated source temperature and the introduction of nitrogen drying gas) causes shrinkage of the charged droplets, which in turn increases repulsion between the charges at the droplet surface. When the droplet is reduced to a certain size the Rayleigh limit is met, where the surface charge repulsion overcomes the surface tension of the droplet. The resultant instability causes the droplet to disperse into many smaller charged droplets in what is known as Coulomb fission. This process of droplet disruption continues leading to very small, highly charged droplets from which the final stage of ESI takes place, the production of gas phase ions that subsequently move into the mass spectrometer. Unlike MALDI, ESI results in the formation of multiply charged ions so that large biomolecules come into a *m/z* range that can be measured by MS. Moreover, its mechanism makes it highly amenable to be coupled to liquid flow techniques such as LC, making it the preferred ionisation choice in proteomic analysis.

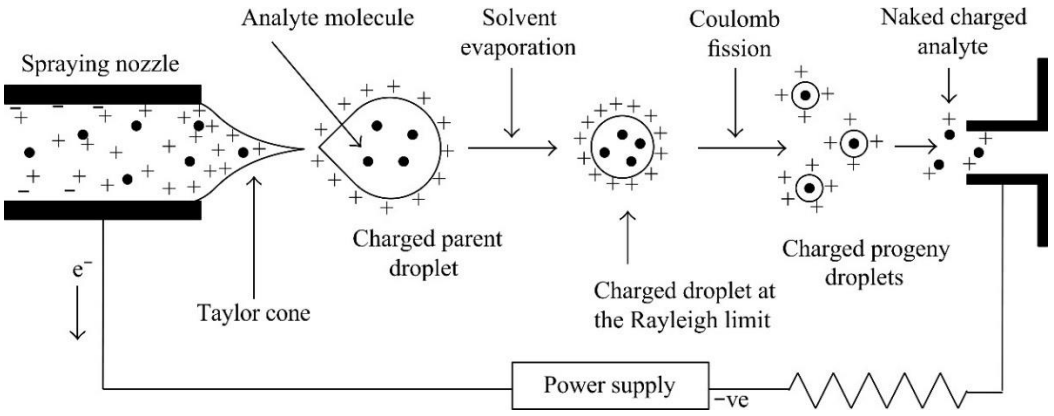


Figure 1.3 Schematic of electrospray ionisation mechanism in positive ion mode. Electric charge at the capillary tip causes the formation of droplets, which through solvent evaporation and Coulomb fission are reduced to charged gas phase ions that move in to the mass spectrometer. Figure reproduced from Banerjee and Mazumdar (2012).

Once in the gas phase, ions are separated based on their *m/z* in a mass analyser. Types of mass analyser commonly used in proteomics include quadrupole, time of flight (TOF) and ion traps. It is beyond the scope of this introduction to describe the detailed workings of these mass analysers, although they have been thoroughly reviewed elsewhere (Aebersold et al., 2003; Boesl, 2017; El-Aneed et al., 2009; Glish et al., 2003) and an overview of each will be given here.

First described in the 1950s, quadrupoles are arguably the simplest type of mass analyser; they are made up of four cylindrical rods arranged parallel to one another through which DC and oscillating RF electric fields are passed though. The ratio of these DC and AC electric fields dictates the motion of ions through the quadrupole; only certain ions of specific *m/z* values will have a sufficiently stable trajectory to pass through the quadrupole to the detector at a particular field strength. By systematically altering field strengths specific *m/z* values can be selected for. In comparison to other mass analysers quadrupoles have low mass resolution and accuracy, but due to their mode of operation can stay tuned to a particular *m/z* value for a period of time, making them very useful for analysis in targeted experiments.

In TOF mass analysers, ions are separated based on their velocities in a low-pressure flight tube. Groups of ions are accelerated by an electric field of known strength into the flight tube. As all ions share the same kinetic energy, those with lower *m/z* values have a greater velocity through the tube and so reach the detector first. Through knowledge of the time taken for an ion to hit the detector after acceleration and the distance of the flight tube, the *m/z* value of the ion can be calculated. TOF analysers offer high resolution, sensitivity and mass accuracy, but have a low duty cycle (the time the system is actively measuring ions) as sampled packets of ions must be cleared of the flight tube before the next packet can arrive (Boesl, 2017).

There are multiple types of ion trap that all operate on the common principle of capturing charged particles and subjecting them to MS analysis. The Fourier-transform ion-cyclotron-resonance mass spectrometer (FTICR-MS) captures ions in a high vacuum using a magnetic field where frequencies of trapped ions can be measured and converted into *m/z* values (Aebersold et al., 2003). FTICR-MS gives very high mass resolution and very high mass accuracy, but FTICR instruments are very expensive and require space-consuming high field superconducting magnets. A more recently developed type of ion trap, the Orbitrap (Makarov, 2000), can deliver mass accuracy and resolution comparable with an FTICR instrument without the obtrusive hardware requirements (Hu et al., 2005).

Most modern mass spectrometers used for proteomics are hybrid instruments in that have more than one mass analyser, for example quadrupole time of flight (Q-TOF), triple quadrupole (QqQ) or Q-Orbitrap (Aebersold et al., 2003). Such instrument design allows for the implementation of tandem MS (MS/MS), in which precursor ions of a specific *m/z* are selected (MS1) and then fragmented to produce product ions that are subsequently detected by another mass analyser (MS2). Collision induced dissociation (CID) is the method most commonly used for the fragmentation of peptides. In CID, ions are accelerated and collided with an inert neutral gas such as helium or argon. The kinetic energy generated by these collisions is converted into internal energy that results in the dissociation of the most labile bonds in the structure; these being the amide bonds along the peptide backbone to produce *b*- and *y*- ions (Johnson et al., 2015; Mitchell Wells et al., 2005). Higher-energy collisional dissociation (HCD) is similar in principle to CID but specifically refers to beam-type CID performed in an octopole collision cell. HCD is associated with higher activation energy and

shorter activation times that can result in additional fragmentation of *b*- and *y*- ions. Another type of fragmentation known as electron transfer dissociation (ETD) occurs by the transfer of electrons from radical anions to multiple charged ions. This forms an unstable positive radical cation, which then fragments at the N–Cα bond, producing *c*- and *z*- fragment ions (Coon et al., 2005; Qi et al., 2017). Unlike CID, ETD does not result in the removal of labile post translational modifications (PTMs) such as phosphoryl groups, making it especially useful for studying PTMs in proteomes.

The final component of the mass spectrometer is the detector, which monitors the charge or current produced by an ion when passes or hits a surface. The signal is amplified and converted into the form of a mass spectrum. The choice of detector will depend on the type of mass analyser; common detectors include the photomultiplier, the electron multiplier and micro-channel plate detectors.

1.3.2. Experimental approaches for quantitative mass spectrometry

MS based proteomics can be divided into top-down and bottom-up proteomics, in which analysis is performed on intact proteins or peptides acting as surrogates for proteins respectively. Although quantification strategies for top-down proteomics have been developed (Du et al., 2006; Hung et al., 2012; Ntai et al., 2014), approaches are less advanced than bottom-up proteomics and indeed the majority of quantitative proteomics workflows are carried out at the peptide level. A typical bottom-up proteomics workflow begins with the enzymatic digestion of a protein into peptides using an enzyme of known specificity. Trypsin is usually used due to its specific C-terminal cleavage of arginine and lysine residues, which typically occur in protein sequences at a frequency to produce peptides of ~ 7-35 residues (Swaney et al., 2010). Such peptides are an optimal length for LC-MS analysis as they have a mass between ~ 600-4000 Da resulting in charged (2+) species between *m/z* 300 and 2000, which are easily detected by MS. Other enzymes used in proteomics experiments include ArgC, LysC, AspN, GluC, LysN and chymotrypsin, which can be used individually or in parallel to improve sequence coverage of proteins (Giansanti et al., 2016). Following digestion, samples are separated in order reduce the complexity of sample entering the mass spectrometer, thus increasing proteome coverage and permitting quantification of less abundant proteins. For peptide analysis LC is usually employed in-line with the mass spectrometer, so as peptides elute they are instantly subjected to ESI and MS/MS analysis. In some instances additional sample fractionation is performed beforehand using techniques such as strong anion exchange, capillary isoelectric focussing or size exclusion chromatography (Zhang et al., 2014). Such pre-fractionation can dramatically increase proteome coverage, but the additional handling steps can result in a reduction in reproducibility and accuracy. Tandem mass spectrometry is commonly performed in data-dependent acquisition (DDA) mode, in which a fixed number of precursor ions (for example the top 10 most intense ions in an MS1 scan) are selected for fragmentation and analysis by MS2. The fragment ion spectra and *m/z* value of each peptide ion are used to enable confident identification of peptides in the sample, which

are then mapped back to proteins. The vast number of fragment ion spectra generated in the MS/MS of a complex sample would make such analysis by an individual impossible. The routine performance of these data processing steps has been enabled by the development of an array of bioinformatic software and databases such as MASCOT and PEAKS (Bruce et al., 2013).

The quantification of peptides and proteins identified by MS is complicated by the fact that MS is not an inherently quantitative technique. This is due to differences in ionisation efficiency between peptides, so the signal intensity of a peptide ion does not relate to its abundance (Oss et al., 2010). Moreover, the presence of interfering compounds in a sample matrix can also alter ionisation efficiency, reducing reproducibility and limiting the ability to compare analytes in different samples (Sarkar et al., 2009). Therefore, several labelling strategies based on the use of stable isotope labels (SIL) have been developed to facilitate protein quantification by MS (Figure 1.4). These work on the basis that SIL peptides will have the same physiochemical properties as analyte peptides, so they will ionise with equal efficiency and the MS intensity can be directly compared.

One group of approaches involves the chemical derivatization of reactive groups in proteins or peptides with SILs. Lysine and cysteine side chains as well as the N-terminus of the sequence are most commonly used, due the well-characterised conjugation chemistry associated with these groups. In the first reported such approach, known as isotope-coded affinity tagging (ICAT) (Gygi, 1999), cysteine residues in two samples to be compared are derivatized with a heavy or light version of a reagent comprising biotin and a linker region capable of incorporating deuterium atoms. After sample combination and enzymatic cleavage, derivatised peptides can be isolated by avidin affinity chromatography and analysed by LC-MS/MS, revealing the identity and quantity of the peptide and which sample it originated from. Although initially successful and followed up by a number of studies aiming to improve properties of the tag (Hansen et al., 2003; Yi et al., 2005), ICAT is ultimately limited in that it cannot detect proteins and peptides that don't contain cysteine residues. Subsequent labelling strategies have targeted the ε-amino group on lysine residues or the N-terminus through N-hydroxysuccinimide (NHS) chemistry. Such strategies include isobaric tags for relative and absolute quantification (iTRAQ) (Ross et al., 2004), tandem mass tagging (TMT) (Thompson et al., 2003) and stable-isotope dimethyl labelling (Raijmakers et al. 2008; Boersema et al. 2009). iTRAQ and TMT utilise isobaric tags that are of the same mass but produce different fragment ions upon MS/MS analysis, thus acting as reporter ions for peptides in a specific sample. Peptides from different samples are tagged with different isobaric tags, pooled and analysed by LC-MS/MS. This determines the peptide sequence and abundances of the reporter ions are used to quantify peptides relative to other samples. Multiple samples can be analysed at once by this approach; iTRAQ is available in 4- or 8-plex and TMT in 10-plex. Such approaches, though, are relatively expensive and time-consuming, and quantitative accuracy can be compromised by interfering ions, especially in complex samples (H. Li et al.,

2017). Dimethyl labelling offers a fast and inexpensive alternative by converting primary amines in all peptides to dimethylamines using isotopomers of formaldehyde and cyanoborohydride. This creates peptides of slightly different masses in up to three samples, which can then be mixed and differentiated by MS/MS. The low-cost and ease of the labelling reaction is a major advantage, although the use of deuterium in the label can lead to retention time shifts in differentially labelled peptides (S. Wang et al., 2007).

As well as labelling peptides through chemical means, SILs can be introduced metabolically during cell growth. In the most popular metabolic labelling method, stable isotope labelled amino acids in culture (SILAC) (Ong et al., 2002), cells are cultured with [$^{13}C_6$] or [$^{15}N_7$] labelled arginine and lysine, which are directly incorporated into proteins as they are synthesised. Labelled and unlabelled samples are mixed and peptides from each sample discriminated by the mass shift resultant from the label. SILAC ensures that every peptide is labelled at the earliest point in the workflow, minimising sample handling errors leading to improved reproducibility. However, multiplexing capabilities are limited and the workflow is time-consuming and expensive.

The approaches described above all result in relative quantification. For absolute quantification, a different strategy based on stable isotope dilution (SID) is required. A SIL peptide standard of known concentration is added to the sample, acting as an internal calibrant so that analyte concentration can be determined based on the difference in signal intensity. Peptides chosen to act as standards are known as quantotypic; such peptides are unique to the proteome of interest, do not contain residues that are likely to be subject to post-translational or chemical modifications and are readably detectable by LC-MS/MS. Multiple approaches to deliver SIL peptides into the sample have been developed (Figure 1.5). Absolute quantification (AQUA) peptides are chemically synthesised with isotope labels and can be spiked into samples before or after proteolytic digestion (Kettenbach et al., 2011). There is no limit as to the number of AQUA peptides that can be spiked into a single sample, although costs can become prohibitive when large numbers of proteins are to be analysed. Additionally, some peptides can be difficult to chemically synthesise, which could be an issue when quantifying small proteins containing a limited number of Q-peptides. A relatively low-cost strategy to produce many SIL peptides for multiple proteins is to use a quantification concatamer (QconCAT) protein (Beynon et al., 2005; Julie M Pratt et al., 2006). QconCATs are artificial recombinant proteins that concatenate several Q-peptides into a single sequence that is expressed heterologously in *E.coli* grown in isotopically enriched media. The QconCAT protein can then be affinity purified, quantified, and added to a sample prior to digestion so that the SIL peptide standards are released during proteolysis. This approach facilitates a large degree of multiplexing using a single standard and the utilisation of biosynthesis permits the generation of peptide sequences that were unavailable by chemical synthesis. The utility of multiplexing using QconCATs has been demonstrated through the absolute quantification of the entire *Saccharomyces cerevisiae* proteome (Lawless et al., 2016) and various protein

pathways and networks (P. Brownridge et al., 2013; Carroll et al., 2011; Messiha et al., 2014), amongst others. A drawback of the QconCAT approach is the potential for differential proteolysis between the QconCAT and analyte proteins due to different local environments around cleavage sites (P. Brownridge & Beynon, 2011). Missed cleavages would lead to inaccurate protein measurements, however this can be mitigated by adding natural flanking sequences to either side of tryptic peptides in the QconCAT sequence, maximising the possibility of equivalent digestion (Cheung et al., 2015).

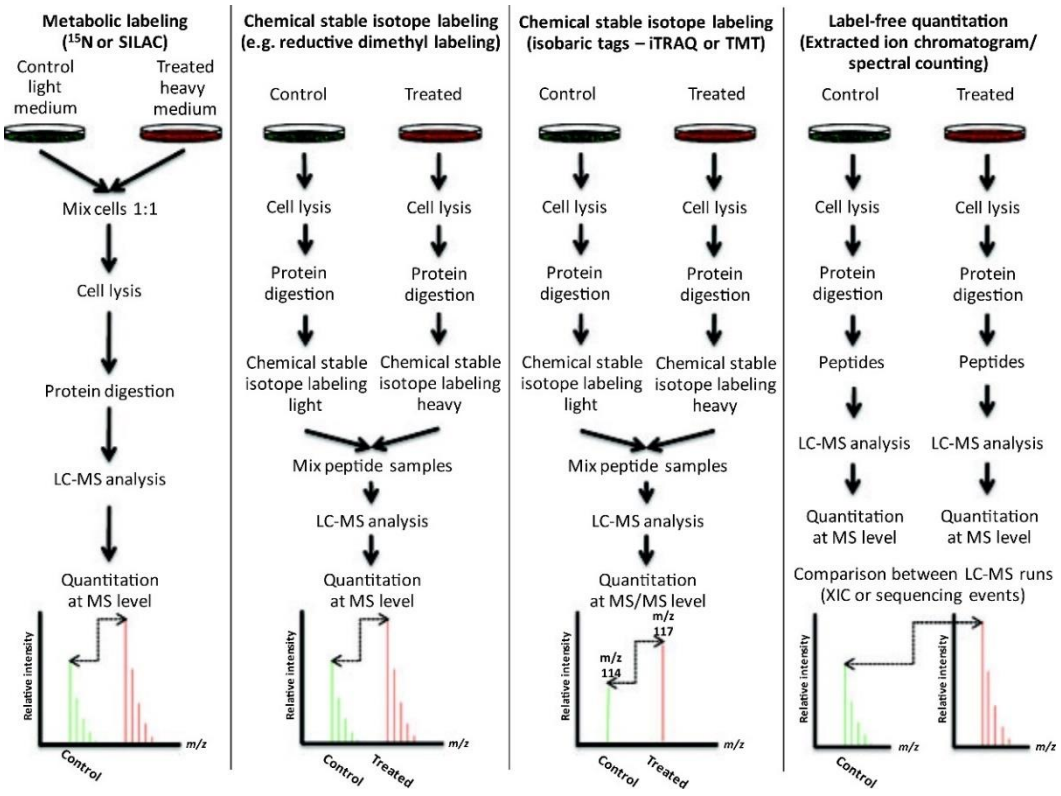


Figure 1.4 Workflows for relative quantitative proteomics. In metabolic labelling techniques such as SILAC, labels are introduced by growing cells in stable isotope enriched media so samples can be combined very early in the workflow, thus reducing sample handling variations. In chemical labelling techniques samples are combined are protein digestion and peptide labelling. Quantification is performed at the MS/MS level for iTRAQ compared to the MS level for other techniques. Label-free strategies use precursor ions intensities or spectral counting to determine peptide abundances between runs. Figure reproduced from Engholm-Keller and Larsen (2013).

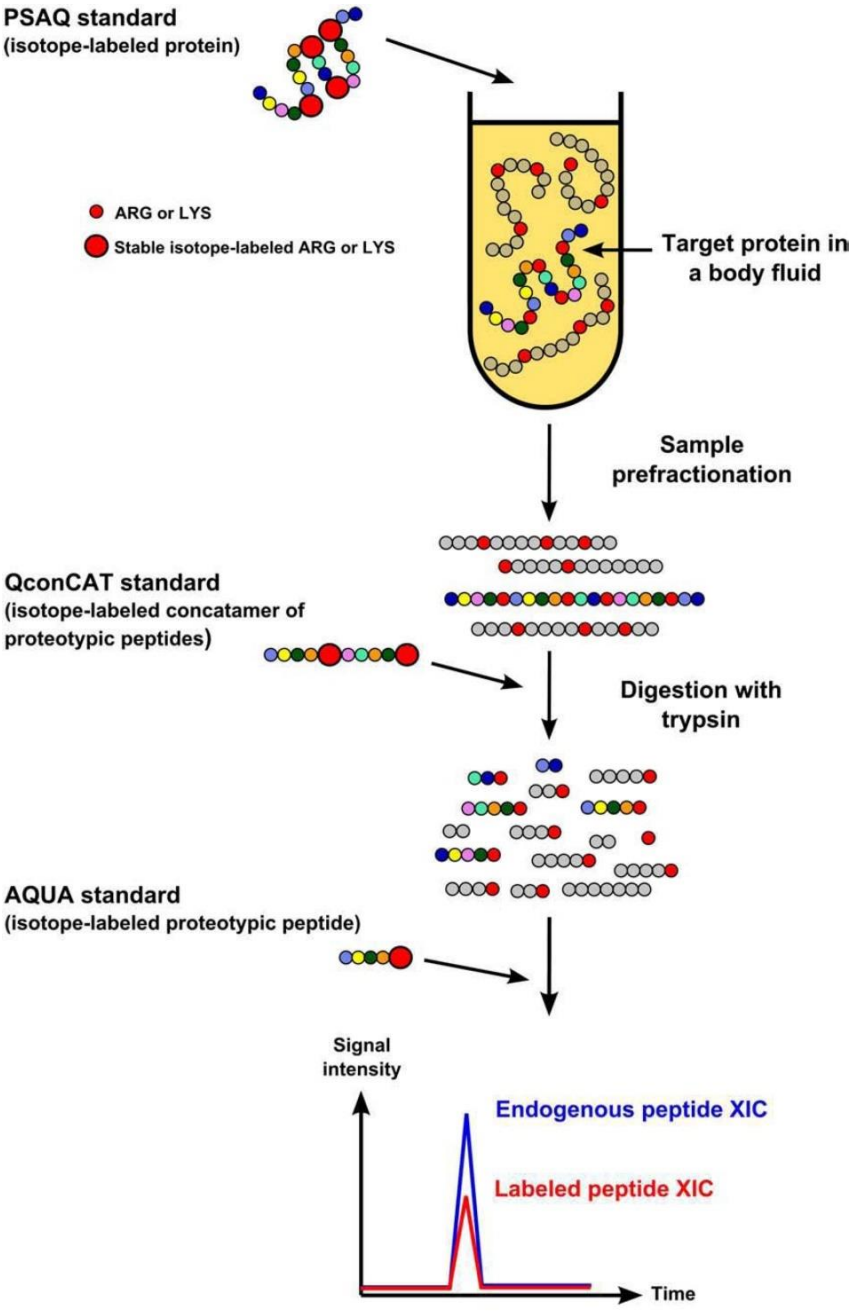


Figure 1.5 Isotope dilution strategies for absolute quantitative proteomics. Schematic of the workflow for absolute protein quantification detailing different types of stable isotope labelled standard and where in the workflow they are introduced to the sample. Protein standards such as PSAQ are introduced pre-fractionation (if carried out) and pre-digestion. Quantification concatamers (QconCATs) are introduced just prior to digestion and peptide standards just before MS analysis. Quantification is achieved by comparing the peak areas of extracted ion chromatograms (XIC) from unlabelled endogenous and labelled standard peptides. Figure reproduced from Brun et al. (2009).

Other approaches use protein rather than peptides as standards. The protein standards for absolute quantification (PSAQ) methodology utilises full length SIL proteins expressed and added to the sample prior to digestion (Brun et al., 2007). The Full-Length Expressed Stable Isotope-labeled Proteins for Quantification (FLEXIQuant) workflow builds on this by expressing protein grown in wheat germ extract, allowing the quantification of PTMs (Singh et al., 2009). Methods using protein standards are considered the most accurate for quantification as digestion between standard and analyte is highly likely to be equivalent. Additionally, fractionation of standard and analyte prior to digestion is permitted allowing greater sensitivity to be achieved. However, there is limited potential for multiplexing due to the great cost and time required to prepare a unique protein standards for each protein.

Although label-based methods allow for accurate quantification of proteins, they require costly reagents and the requirement for peptide standards means that absolute quantification cannot be performed on a proteome level. To overcome this, label-free approaches have been developed for relative and absolute quantification of the proteome. Label-free workflows are based on counting the number of spectra that are matched to a peptide (spectral counting), or comparing precursor ion intensities (Figure 1.4). Spectral counting methods are based on the observation that the number of peptide spectral matches (PSMs), the number of distinct peptides identified and the sequence coverage for a protein all correlate with protein quantity (H. Liu et al., 2004; Rappsilber Ryder, U., Lamond, A. I., Mann, M., 2002). Therefore, quantification of a protein is achieved by summing the number of MS2 spectra that are matched against its peptides. Popular spectral counting methods include Exponentially Modified Protein Abundance Index (EMPAI) (Ishihama et al., 2005) and Absolute Protein Expression (APEX) (Braisted et al., 2008; Lu et al., 2007). However, spectral counting is associated with poor linear dynamic range due to saturation effects, poor reproducibility due to DDA sampling, misassignment of peptide ions that are shared between proteins and an intrinsic bias towards more abundant proteins (Bantscheff et al., 2012; Grossmann et al., 2010). Ion intensity methods are based on the observation that there is excellent correlation between peptide concentration and MS1 signal intensity over a linear range greater than four orders of magnitude (Bondarenko et al., 2002; W. Wang et al., 2003). For quantification, all or a subset of peptides for a protein are chosen and their MS1 ion intensities aggregated through summation or averaging peptide fold changes. Widely used ion intensity methods include Hi3 (or Top3) (J. C. Silva et al., 2006) and Intensity Based Absolute Quantification (iBAQ) (Schwanhäusser et al., 2011). For Hi3, quantification is based the observation that the average signal intensity of the top 3 most intense peptide ions is correlated with the total amount of a protein in a sample (Ahrné et al., 2013; Grossmann et al., 2010; J. C. Silva et al., 2006). Absolute quantification using Hi3 is made possible by comparing analyte signal to spiked-in reference peptides or the total amount of protein injected into the instrument (Ahrné et al., 2013). Protein quantification using iBAQ is achieved by summing intensities of all peptides for a specific protein, which is then divided by the number of theoretically observable peptides to estimate protein level (Schwanhäusser et al., 2011). Compared to spectral

counting, ion intensity methods have been shown to be more accurate and reproducible (Ahrné et al., 2013; Arike et al., 2012). Both Hi3 and iBAQ have been employed for global proteome quantification across a range of samples, demonstrating their utility for quantifying a large number of proteins in a single experiment at a relatively low cost (Bantscheff et al., 2012; Megger et al., 2013).

Label-free quantification is considered less accurate and precise than quantification achieved using stable isotope labels, especially for low abundance proteins (P. Brownridge, Holman, et al., 2011; Krey et al., 2014). However, the nature of labelling experiments limits quantitative throughput and multiplexing and so the quantification of an entire proteome in a single LC-MS/MS would be very difficult using a labelling approach. Label-free methods certainly have their place in the proteomics toolkit, especially in discovery experiments where high sample throughput is more important than accuracy. However, for absolute protein quantification where accuracy, precision and a low limit of detection is key, such as biomarker verification studies, targeted approaches using SIL standards are still considered the gold standard approach.

1.3.3. Acquisition strategies for quantitative mass spectrometry

Although highly useful for protein discovery, shotgun MS proteomic workflows based on DDA do not have the high sensitivity, quantitative accuracy and reproducibility required for the absolute quantification of targeted proteins (Holman, Sims, and Eyers 2012). Mass spectrometers used in DDA mode for quantitative analysis generally have a low duty cycle due to continuous ionisation and the process of precursor ion selection; whilst one packet of product ions is being analysed peptides eluted from the column are lost, leading to the loss of information and thus a low number of points across the chromatographic peak. This translates into decreased accuracy and poor sensitivity. Additionally, the stochastic nature of precursor ion selection can lead to different peptides being selected for fragmentation for similar samples, reducing reproducibility. As peptide ions with the highest intensity are selected for fragmentation and MS2 analysis, there is an inherent bias towards the analysis and quantification of the most abundant proteins, limiting the dynamic range of analysis. Due to the many limitations of DDA, targeted MS techniques have been developed that when used in conjunction with SIL strategies can quantify proteins with exquisite selectivity and high reproducibility (Figure 1.6).

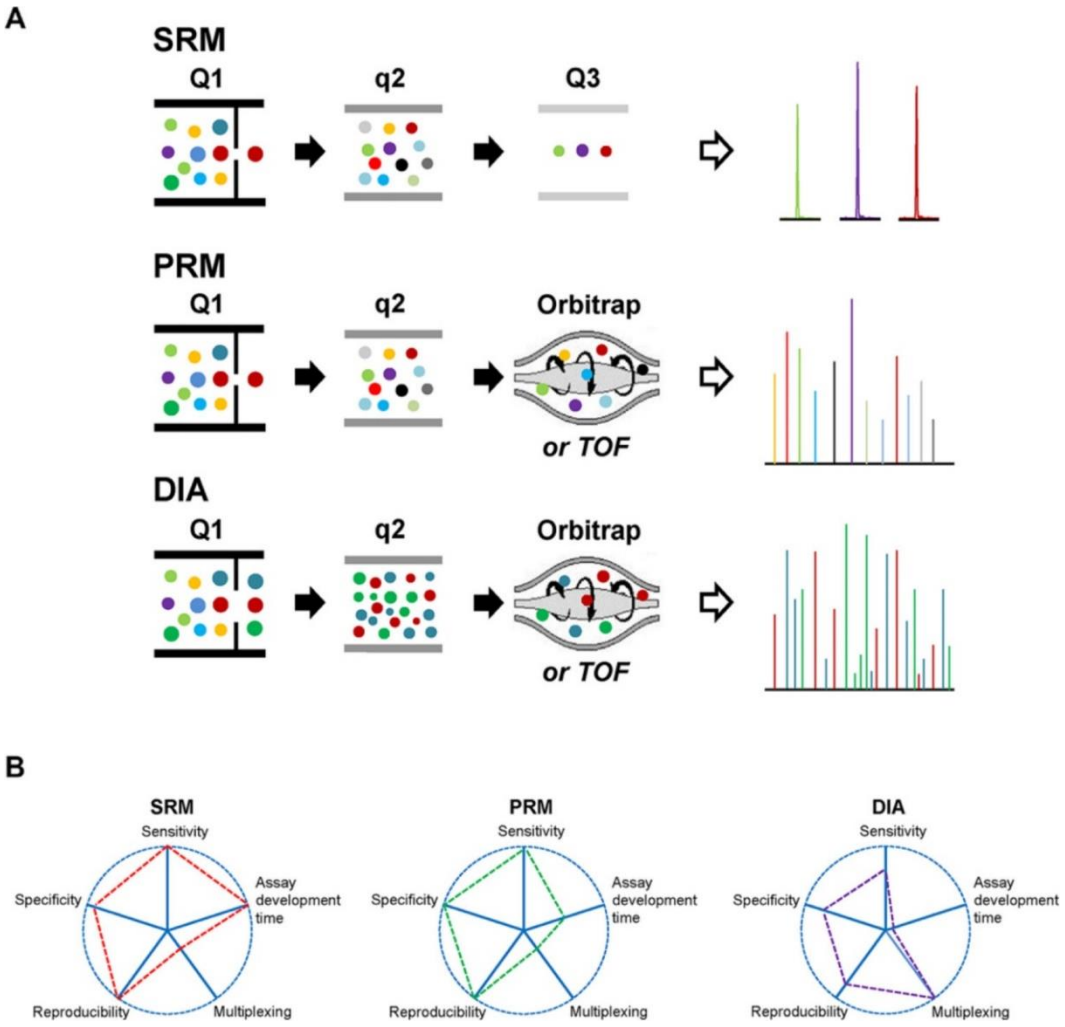


Figure 1.6 Acquisition strategies for targeted MS proteomics. a) Schematic diagrams of MS configurations for SRM, PRM and DIA. In SRM, a precursor ion for a specific peptide is selection in Q1 and transmitted to q2 for fragmentation by CID; specific product ions resultant from target peptide fragmentation are selected in Q3 for detection. PRM follows the same first two steps as SRM but all product ions are monitored by a HR/AM mass analyser. In DIA, all precursor ions within a defined *m/z* window are selected in Q1, fragmented in q2 and all product ions are monitored by a HR/AM mass analyser. The process is repeated through the full MS1 *m/z* range in a stepwise fashion. b) Comparison of SRM, PRM and DIA for targeted quantification. DIA is the least time-consuming technique to set up and the best for multiplexing, but lacks sensitivity, specificity and reproducibility. SRM and PRM perform as well as one another, other than PRM being more selective due to the use of a HR/AM mass analyser. The sensitivity comparison between SRM and PRM is based on quantification of the relatively small number of target peptides (e.g. ~50 peptides) in a single analysis. Figure reproduced from Tujin Shi et al. (2016).

Selected reaction monitoring (SRM) is such an approach that has been used extensively to study small molecules (Kondrat et al., 1978), but it has only recently been adopted for proteomics where its qualities have been swiftly recognised (Marx, 2012; “Method of the Year 2012,” 2013). SRM is typically performed on a triple quadrupole mass spectrometer, with two quadrupoles acting as selective mass filters and the other as a collision cell for ion fragmentation. In an SRM experiment, ions are delivered to the first quadrupole through which only precursor ions of a specific *m/z* value can pass through. These precursor peptide ions are then subjected to low energy CID in the second quadrupole. The resultant product ions are transferred into the third quadrupole, which only permits a specific product ion of a single *m/z* value through to the detector. This combination of precursor and product ion *m/z* values (known as a transition) alongside knowledge the chromatographic retention time of the peptides makes SRM an exquisitely selective technique. Moreover, multiple transitions can be monitored in a single run, which allows for a high degree of multiplexing, especially when SRM is scheduled so that peptides are monitored in a specific time window relating to when they elute from the LC column. SRM has become a mature technology with clarity on its background, implementation and limitations, and as such has been used with SIL peptides in for targeted quantification in many studies (Vidova and Spacil 2017; Holman, Sims, and Eyers 2012).

SRM is capable of quantifying proteins down to 50 copies per cell in *S. cerevisiae* (Picotti et al., 2009), however, this is a function of being able to load many cells onto a column. In human plasma or serum, commonly used for the detection of biomarkers as it is easily obtained, SRM can routinely detect proteins in the µg/mL range. As most biomarkers are in the pg/mL range, this is clearly insufficient. To improve sensitivity, sample enrichment using affinity reagents has been developed in what has become known as ‘immuno-SRM’. Affinity enrichment can be performed at the protein or peptide level. The development of Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA) methodology relies on the development of anti-peptide antibodies to enrich analyte and SIL standard peptides post-digestion. Anti-peptide Abs can be immobilised to a column or magnetic beads, sample passed over and the unbound fraction washed away, and bound peptides eluted for MS analysis (N. L. Anderson et al., 2004, 2009). SISCAPA has been used to detect proteins down to the low ng/mL level, for example FGF15 at a LOD of 0.1 ng/mL in mouse plasma (Katafuchi et al., 2015) and a TIMP1 glycoform at 0.8 ng/mL in human serum (Yeong et al., 2009). The protein level immunoenrichment of samples, termed Mass Spectrometric Immunoassay (MSIA) uses antibodies against a protein or particular protein isoform for enrichment before digestion and has been used to detect protein down to similar levels (Gauthier et al., 2015; Krastins et al., 2013). Sensitivity has been further improved through tandem affinity enrichment approaches, in which samples are enriched first at the protein level, and then again at the peptide level post-digestions. Using this approach β-NGF was quantified in human serum down to 0.7 pg/mL (Neubert et al., 2013), representing a 10-fold increase in sensitivity compared to regular immuno-SRM techniques. Immuno-SRM techniques have great potential to improve

sensitivity of targeted SRM experiments, but are limited in their multiplexing ability as exposing a sample to multiple Abs at one time increases the possibly of non-specific binding, thus increasing noise and decreasing sensitivity (Tujin Shi et al., 2016). They also require the rapid generation of affinity regents against protein and peptide targets; approaches using recombinant antibody fragments (Whiteaker et al., 2014) and aptamers (Zhao et al., 2011) have been proposed to fulfil this. There remains the fact, though, that affinity enrichment incurs additional expense, increases experimental time, and raises the potential for variability through extra sample handling steps. As an alternative, antibody-independent methods for improving SRM sensitivity have been developed. In an approach known as PRISM (high-pressure, high-resolution separations with intelligent selection and multiplexing), samples are separated by high-resolution reversed-phase capillary LC and analysed by multiplexing SRM, which when combined with immunoaffinity depletion of abundant proteins detected plasma proteins at 50 – 100 pg/mL levels (T. Shi et al., 2012). Building on this, two-dimensional (2D) high-resolution reversed-phase LC using low and high pH RP-LC in succession to fractionate first samples and then resultant fractions has led to the quantification of proteins at 10 pg/mL levels in non-depleted serum (Nie et al., 2017). This matches and even outperforms immunoassays in terms of sensitivity, but the method requires highly complex sample preparation and fractionation that limits sample throughput, and unless carefully controlled through automation would not be highly reproducible (Figure 1.6b).

A major limitation with SRM is the development time required to define optimal peptides, assay parameters and assay characteristics for a high quality SRM assay. Tools and software have been developed to aid with the selection of peptides and assay parameters such as CONSeQuence (Eyers et al., 2011), Peptide Atlas (Deutsch et al., 2008), SRMAtlas (Kusebauch et al., 2016) and Skyline (MacLean et al., 2010), but the process of assay validation to reach the highest standards set by the community (Carr et al., 2014) is still a time-intensive task.

There have been technological developments that aim to maintain the selectivity and sensitivity of SRM whilst improving development time and throughput. The development of hybrid instruments with high-resolution and accurate-mass (HR/AM) mass analysers, such as quadrupole-Orbitrap and quadrupole-TOF mass spectrometers, has facilitated the development of new targeted MS modes. Parallel reaction monitoring (PRM) (Gallien et al., 2012; Peterson et al., 2012; Schilling et al., 2015) is similar to SRM in that precursor ions of a specific *m/z* value are isolated by a quadrupole and fragmented by CID or HCD. The full MS/MS spectra (rather than a specific fragment ion) for the precursor ion is then acquired in the HR/AM mass analyser. Fragmenting every product ion improves assay selectivity and removes the need to select transitions for each peptide, making the assay less time-consuming to set up. Studies using PRM on Orbitrap and TOF instruments have shown it to have equivalent accuracy, sensitivities, signal linearity, dynamic range reproducibility and repeatability compared to SRM (Kockmann et al., 2016; Ronsein et al., 2015; Schiffmann et

al., 2014; Schilling et al., 2015). However, these results were obtained using relatively simple samples with few co-eluting peptides, so MS acquisition parameters could be set towards their optimum (i.e. high maximum fill time and high orbitrap resolving power). In more complex samples where peptide co-elution is increased or where a greater degree of multiplexing is required, these acquisition parameters must be adjusted, which ultimately affects sensitivity, accuracy and reproducibility (Bourmaud et al., 2016; Gallien et al., 2014; Tujin Shi et al., 2016). Novel dynamic acquisition strategies that rely on the real-time adjustment of PRM acquisition parameters based on the detection of spiked-in SIL peptides have been used to quantify up to 600 peptides in complex samples (Gallien et al., 2015), although the accuracy of reproducibility of such methods remains to be seen (Tujin Shi et al., 2016). Although currently less-established than SRM, PRM offers great promise due to its superior assay performance, and may eventually replace SRM as the gold standard for targeted quantification experiments if future studies show that recognised performance metrics for quantitative measurements can be met (Carr et al., 2014).

Both SRM and PRM are limited in their multiplexing ability, making global proteome quantification in a single run impossible. To overcome this, targeted analysis based on data independent acquisition (DIA) modes have also been developed. In DIA, all peptides ions within a defined *m/z* window are isolated in a quadrupole and fragmented together by CID or HCD. All product ions are then analysed by HR/AM mass analysers and the process is repeated throughout the full *m/z* range in a stepwise fashion. Since first being described (Venable et al., 2004), significant advances in DIA have been made through the development of new generation HR/AM mass spectrometers and different acquisition techniques to take advantage of the improved MS performance (H. Li et al., 2017; Tujin Shi et al., 2016). Once such technique, known as Sequential Windowed Acquisition of all THeoretical mass spectra (SWATH), has particularly garnered attention as a method for proteome-wide quantification of target proteins (Gillet et al., 2012). In SWATH, a wide, slightly overlapping precursor mass range is used, resulting in fast scan times for the entire mass range so that peptides are measured 8 – 10 times across their chromatographic elution profile. The resultant highly complex MS/MS spectra (referred to as digital maps) are deconvoluted post-acquisition using spectral libraries built from fragment ion spectra. For each target peptide these contain *m/z* values for the precursor and selected fragment ions, the relative intensities of fragment ions and retention time data for the peptide (Schubert et al., 2015). SWATH-MS has been used with label-free (Y. Liu et al., 2015; Mcqueen et al., 2015), chemical labelling (Russell et al., 2016) and SIL label standards (Y. Liu et al., 2013; Nakamura et al., 2016) to perform relative and absolute quantification in a wide range of samples (Anjo et al., 2017). When compared with SRM and PRM, protein fold changes and dynamic range have been shown to be similar using SWATH, but sensitivity up to 10-fold lower (Gillet et al., 2012; Y. Liu et al., 2013). Another study using a limited number of peptide analytes and SIL standards showed that SRM and PRM outperformed SWATH in terms of accuracy and precision, especially for low abundance proteins (Kockmann et al., 2016).

SWATH and DIA techniques in general have great potential for proteome-wide targeted quantification. For targeted analysis of a limited number of proteins, however, DIA is currently inferior to the gold standard SRM in terms of accuracy, sensitivity and reproducibility. These metrics could be improved with sample fractionation or depletion of the most abundant proteins, but these introduce additional sample handling steps that increase variability. Future development in instrumentation such as increased scan speeds, improved resolution and employing ion mobility to add an extra dimension of sample separation may improve DIA for targeted analysis, but this remains to be seen.

1.4. DOSCATs as double standards

Out of all the reviewed methods for targeted protein quantification using mass spectrometry and immunochemistry, SRM analysis with SIL standards is recognised as the gold standard for absolute quantification. Immunochemical methods are practised much more prevalently, though, with WB extensively used in a semi-quantitative manner and increasingly used for absolute quantification with the aid of improving technology. Whilst ELISAs are superior for quantification, the development time and expense associated with developing new well validated ELISAs means that the range of potential targets does not match that of WB.

Western blotting, although a much-maligned method for protein quantification, does hold some advantages over SRM (Table 1.1). It is generally seen as more sensitive and is much more accessible to the general research community due to its relatively low cost and ease of use. Additionally, WB is often requested by reviewers for orthogonal validation of MS-generated proteomics datasets without any evidence for the analytical equivalence of both methods (Aebersold et al., 2013). Therefore, there is a motivation to explore the performance of quantitative WB compared to MS. However, directly comparing WB and MS outputs within a single workflow is difficult due to the different calibration standards used (recombinant proteins and SIL peptides for WB and MS respectively). DOSCAT technology aims to bridge the gap between the two techniques by acting as a double standard for WB and MS analysis.

DOSCATs continue the trend started by QconCATs of using protein engineering to create artificial proteins to improve MS-based analyses, other examples being QCAL (Eyers et al., 2008), QCAL-IM (Chawner et al., 2012) and RePLiCAL (Stephen W Holman et al., 2016) proteins as standards for assessing and normalising instrument conditions across repeat runs. DOSCATs are artificial protein designed *in silico* and expressed recombinantly in *E.coli*, and function by concatenating epitope sequences for WB analysis from multiple proteins so that the standard is recognised by multiple antibodies. They also embed peptides from the same proteins for MS analysis, thus acting as a single multiplexed standard for WB and MS quantification. They can be spiked into the sample pre-digestion, so the sample can be taken down both MS and WB workflows, facilitating the seamless integration of the two workflows and the direct comparison of outputs from both techniques.

This will allow for a fair assessment of the quality generated by WB compared to a MS gold standard. Moreover, orthogonal quantification of target proteins in a unified workflow will improve quantitative accuracy and utilise the best attributes of the two platforms. Techniques orthogonal to WB, such as MS, has been proposed to validate antibody specificity, and DOSCAT technology would be well placed to aid in this process.

Table 1.1 Comparison of western blotting and mass spectrometry for target absolute protein quantification.

Parameter	Western blotting	Mass spectrometry (SRM)
Sensitivity	Can routinely detect pg/mL protein in complex samples	Capable of detecting µg/mL protein in complex samples, although this can be improved to pg/mL by sample enrichment/depletion/fractionation
Reproducibility	Highly variable technique with many manual handling steps, although new automated techniques drastically improve this	Highly reproducible; typical %CV < 20%
Selectivity	Dependent on specificity of the antibody	Use numerous parameters: peptide elution time, multiple transitions
Quantification	Classic methodology semi-quantitative at best although technological advances and automation make the technique quantitative	Gold standard for quantification
Ease of use	Assays requires optimisation but technically simple methodology	Requires experienced operators to optimise and execute assay
Cost	Low equipment and reagent costs, although some instrumentation can be expensive e.g. CCD camera	Significant upfront and ongoing investment in instrumentation, maintenance and associated expertise

1.5. Aim and objectives

The overarching objective of this thesis is to develop DOSCAT technology to demonstrate its utility in quantifying a set of target proteins in a complex sample. To achieve this, there are several landmark aims that must be met. Firstly, the principles of DOSCAT design must be established and refined; for example, the rationale behind the selection of epitopes and peptides and how they are assembled within the protein sequence. Secondly, the DOSCAT must be shown to express recombinantly in *E.coli* and that the epitopes and peptides contained within it can be detected in WB and MS respectively. Finally, the DOSCAT can be deployed in a dual-quantitative workflow and data from both platforms compared.

The first results chapter is concerned with the initial design and proof of principle of a DOSCAT in a model system. The next two result chapters' detail how two separate DOSCATs can be deployed to quantify target proteins in complex samples. Throughout the results chapters unforeseen challenges that arise during the deployment of DOSCAT are addressed and solutions discussed. Finally, chapter 6 discusses the potential for commercialising of DOSCAT technology with an emphasis on intellectual property and bringing the technology to market.

Chapter 2: Materials and methods

2.1. Harvesting, counting and sonication of SK-N-AS cells

SK-N-AS cells between passages 7-14 were grown to 80% confluency in 75 cm^2 flasks in Minimum Essential Media (37 °C, 5% CO_2/humidity). In each set of experiments a set of three cultures were stimulated with 10 ng/mL TNFα (Calbiochem, UK) for a period of 24 h. At the same time media was replaced on the unstimulated cells without added TNFα. To harvest cells, media was aspirated and adherent cells were washed three times with sterile PBS (5 mL). Cell dissociation buffer (Sigma, UK) (1 mL) was added and the cells placed in the incubator at 37°C for 5 min. PBS (137 mM NaCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, pH 7.4) (4 mL) was added to each flask and the contents of each flask transferred to individual 15 mL Falcon tubes. Cells were dispersed by repeated uptake and aspiration from a 5 mL pipette and 50 µL of cell suspension was removed and added to an equal volume of 0.4% (w/v) trypan blue in PBS. Immediately after mixing, dye suspension (10 µL) of cells was pipetted into each of the two chambers of a counting slide and cell numbers counted in a TC10 cell counter (Bio-RAD, Hemel Hempstead, UK). The counter gave a direct reading of the total cells/mL and the viable cells/mL of suspension. Cells were harvested by centrifugation at 160 *x g* (Eppendorf, Cambridge, UK), supernatant removed and cell pellets stored at -20°C. Cell pellets were re-suspended in 25 mM ammonium bicarbonate (Sigma) (not pH adjusted) at 100 µL/1 x 10^6 cells and sonicated on ice using three 10 sec pulses at 30% amplitude delivered from a 3 mm probe of a Sonics Vibra Cell™ (Jencons Scientific Ltd, UK). Benzonase nuclease (Merck Millipore, UK, #70746) (2.5 U/100 µL cell lysate) was added and the cell lysate was held on ice and not fractionated further.

2.2. Collection and processing of CSF

CSF was collected with ethical approval and with informed consent as described by Gómez-Baena et al. (2017).

2.3. Preparation of competent *E. coli* cells

BL21 *E.Coli* cells were streaked onto a culture plate containing LB agar (Merck, Germany) and incubated at 37 °C overnight. A single colony from this culture was used to inoculate 10 mL LB broth (Merck, Germany), which was incubated overnight at 37 °C. LB broth pre-warmed at 37 °C (100 mL) was inoculated with 1 mL of the overnight culture and grown at 37 °C with shaking until an absorbance at 600 nm of 0.5 was reached. The culture was incubated on ice for 10 min and 50 mL of the culture centrifuged at 4,000 rpm, 5 min, 4 °C. Supernatant was discarded and cell pellets resuspended in 40 mL ice cold 0.1 M $CaCl_2$, incubated on ice for 20 min, and centrifuged 4,000 rpm, 5 min, 4 °C. Supernatant was discarded and the cell pellet resuspended in 1 mL 0.1 M $CaCl_2$. Cells were then used for transformation immediately or stored with 15% glycerol at -80 °C.

2.4. Transformation of *E.coli* cells

Genes for DOSCAT protein were optimised for expression in *Escherichia coli*, synthesised and ligated into a pET21a plasmid vector (Eurofins Genomics, Ebersberg, Germany). The DNA was diluted to 1 ng/µL in TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 8.0) and 5 ng (5 µL) DNA added to 150 µL competent *E.coli* cells and the mixture incubated on ice for 45 min. Cells were heat-shocked by placing them in a 42 °C water bath for 30 sec followed by incubation on ice for 2 min. A 1 mL volume of LB broth pre-warmed at 37 °C was added to the cells and this was incubated at 37 °C for 60 mins. Cells were centrifuged at 5,000 rpm, 5 mins and the cell pellet resuspended in 200 µL LB broth. Transformed cells (100 µL) were spread onto culture plates containing LB agar with 50g/mL ampicillin (Sigma, A9518) and incubated at 37 °C overnight. If not being used immediately for expression, glycerol stocks were created by inoculating 10 mL LB broth, 50 µg/µL ampicillin with a single transformed colony and incubating at 37 °C overnight with shaking. In a sterile cryovial, 1 mL of the overnight culture was added to 0.5 mL sterile 60% glycerol, gently mixed and stored at -80 °C.

2.5. Expression and purification of stable isotope labelled DOSCATs

A single colony of BL21 DE3 *E.coli* transformed with plasmid DNA was used to inoculate 10 mL LB broth, 50 µg/µL ampicillin and incubated for 6 h at 37 °C with shaking. A 100 µL volume of this culture was added to 10 mL minimal medium (48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 19 mM NH_4Cl, 1 M $MgSO_4$, 0.1 M $CaCl_2$, 20% glucose, 0.5% (*w/v*) thiamine) containing 50 µg/µL ampicillin and incubated overnight at 37 °C with shaking. Minimal medium with amino acids was prepared with 0.1 mg/mL [$^{13}C_6$]Arg/[$^{13}C_6$]Lys (Sigma), 0.1 mg/mL hydrophilic amino acids and 0.2 mg/mL hydrophobic amino acids. A 4 mL volume of this overnight culture was used to inoculate 200 mL minimal media with amino acids containing 50 µg/µL ampicillin, which was incubated at 37 °C. When the culture OD at 600 nm reached 0.6, expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Alfa Aesar, B21149) and the cells were grown for a further 3 h. *E.coli* cells were separated from culture media by centrifuging at 3500 *x g*, 15 min, 4°C. Cell pellets from 50 mL culture were resuspended in 2.5 mL sonication buffer (50 mM $NaPO_4$, 25 U/mL benzonase nuclease, 1 x Complete EDTA-free protease inhibitor tablet (Roche, #11836170001), pH 8.0). Cells were sonicated on ice using 10 sec pulses at 30% amplitude delivered from a 3 mm probe of a Sonics Vibra Cell™ until 130 joules was reached.

For purification, the *E.coli* lysate was centrifuged at 6000 *g*, 8 min, 4°C to separate insoluble and soluble fractions. DOSCATs were present in the insoluble inclusion body pellets, which were solubilised by incubating for 30 min in 4 mL 20 mM $NaPO_4$, 0.5 M NaCl, 10 mM imidazole, 6 M guanidine hydrochloride, pH 7.4. Inclusion body samples were filtered using a 1.20 µm syringe filter (Milliex GP, Merck Millipore, UK) before purification of the His-tagged DOSCAT on a 1 mL His-trap HP column (equilibrated in the solubilisation buffer, above) using the ÅKTA start system (GE Healthcare, USA). Bound proteins were eluted by applying a linear gradient

of 0-100% elution buffer (20 mM $NaPO_4$, 0.5 M NaCl, 0.5 M imidazole, 6 M guanidine hydrochloride, pH 7.4) over 20 min at a flow rate of 1 mL/min. Eluted fractions containing DOSCAT were pooled and dialysed against 50 mM ammonium bicarbonate, 1mM DTT (Melford Laboratories Ltd, MB1015), pH 8.5. RapiGest *SF* (Waters, UK) was added to the storage buffer at a final concentration of 0.1% (w/v) to reduce DOSCAT adsorption to plastic surfaces. DOSCAT solution was aliquoted and stored in low bind tubes (Corning, USA) at -20°C.

2.6. Expression of M-DOSCAT-i

The immunogen M-DOSCAT-i did not express using the protocol described in section 2.5, and so alternative cell lines were used: SL BL21, C43 (DE3), BL21(DE3)-pLysS (pLysS), Rosetta-pLysS (Rosetta) and Rosetta-gami-pLysS (Rosetta-gami). The M-DOSCAT-i plasmid was transformed into all cells as previously described. Agar and LB broth (both Merck Millipore) contained 1% glucose and 50 µg/mL ampicillin (Sigma, A9518) for all cell strains, with the following additions for some cell strains. For Rosetta and pLysS cells 35 µg/mL chloramphenicol (made up in 100% ethanol) was added and for Rosetta-gami 35 µg/mL chloramphenicol, 10 µg/mL tetracycline (made up in 75% ethanol) and 15 µg/mL leucine (made up in MilliQ H_2O) was added. After transformation cells were cultured in the appropriate agar overnight at 37°C and a single colony used to inoculate 10 mL of the appropriate LB media, which was cultured overnight at 37°C with shaking. This overnight culture (1 mL) was used to inoculate 100 mL of the appropriate LB media, which was cultured at 37°C with shaking and expression induced by 0.5 mM IPTG when the culture OD at 600 nm reached 0.6.

2.7. Protein concentration determination

Protein was assayed using a modified Bradford assay. SK-N-AS cell lysate and DOSCAT were diluted 1:50 and 1:100 in Milli-Q (18Ω) water and 100 µL of each sample added to a microtitre plate in duplicate followed by the addition of 200 µL protein assay reagent (Thermo Fisher Scientific, Cramlington, UK). The plate was read at 600 nm on a microplate reader (Multiscan) using Ascent software (Thermo Fisher Scientific) and protein concentrations interpolated from a BSA standard curve.

2.8. Restricted proteolysis of DOSCAT

The DOSCAT protein was incubated with either TEV protease (Invitrogen), TVMV protease (Biomol, Germany), human RV3C protease (GE Healthcare), factor Xa, thrombin (Sigma) or enteropeptidase (Novagen) using the following reaction buffers and conditions specific to each protease. TEV and TVMV proteolysis was carried out in 50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 1 mM DTT at 30°C. RV3C proteolysis was carried out in 50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT at 4°C. Enteropeptidase proteolysis was carried out in 20 mM Tris-HCl pH 7.4, 50 mM NaCl, 2 mM $CaCl_2$ at RT. Factor Xa proteolysis was carried out

in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM $CaCl_2$ at RT. Thrombin proteolysis was carried out in 20 mM Tris-HCl, 150mM NaCl, 25 mM $CaCl_2$, pH 8.0 at RT. All incubations were carried out overnight, with time points taken for SDS-PAGE analysis as described in the text.

2.9. In-gel trypsin digestion

Bands were excised from SDS-PAGE gels stained with Comassie Brilliant Blue (PhastGel Blue, GE Healthcare, 17051801). Gel pieces were reduced in 100 µL 100 mM DTT for 15 mins at 60°C and then alkylated by the addition of 7 µL 250 mM IAM (Sigma, #1002483822) for 30 mins at RT in the dark. The reaction was quenched by the addition of 3 µL 100 mM DTT and incubation at RT for 5 mins. The supernatant was removed and the gel piece washed twice in 500 µL ammonium bicarbonate with 10 mins shaking at RT each time. The gel piece was then dehydrated in 100 µL 100% ACN with 15 mins shaking at RT. This was then removed, residual ACN evaporated through air drying and 30-50 µL 0.02 µg/µL trypsin in 40 mM ammonium bicarbonate added to the gel piece. After 15 mins incubation at RT 50-100 µL 40 mM ammonium bicarbonate in 5% ACN was added so as to completely cover the gel piece. The sample was incubated at 37°C overnight (16-18 h) with gentle shaking. Following this an equal volume of 100% ACN was added to the solution and the incubation continued for 20 mins. The gel pieces were briefly spun down and the supernatant collected and speed-vacced to remove ACN.

2.10. In-solution trypsin digestion

Samples were diluted in 25 mM ammonium bicarbonate and proteins were denatured by the addition of Rapigest (0.1% w/v) and heating at 80°C for 10 min. Samples were reduced (addition of 10 µL 60 mM DTT and heating at 60°C for 10 min) and alkylated (addition of 180 mM iodoacetamide and incubation at RT for 30 min in the dark). In digests containing DOSCAT, up to 2 pmol of [Glu1]-Fibrinopeptide B (Waters) was added. Mass spectrometry grade trypsin (Promega, USA) was reconstituted in 50 mM acetic acid to 0.2 µg/µL and 10 µL added to digests followed by incubation at 37°C for 4.5 h. At this stage an additional 2 µg trypsin was added and the sample incubated at 37°C overnight. The digestion was ended and RapiGest *SF* removed by acidification (1.5 µL trifluoroacetic acid followed by incubation at 37°C for 45 min). Digests were made up to 225 µL by the addition of acetonitrile:water (2:1) and precipitate (resultant from the breakdown of RapiGest *SF*) removed by centrifugation (13,000 *g*, 30 min, 4°C). Supernatant was removed from the precipitate pellet and carried forward for use as sample.

2.11. SDS-PAGE

SDS-PAGE was performed as described by Laemmli (1970), unless otherwise stated in the text. Samples were mixed 1:1 with 2x reducing sample buffer (100 mM Tris-HCl pH 6.8, 4% (*w/v*) SDS, 0.2% (*w/v*) broophenol blue, 20% (*v/v*) glycerol, 200 mM DTT) and heated at 95°C

for 5 min. Samples were run on 12% or 15% (specified in the text) SDS-PAGE gels at 200V for 45 min. Protein bands were visualised with Coomassie Brilliant Blue.

2.12. Western blotting

Samples separated by SDS-PAGE were transferred electrophoretically to nitrocellulose or PVDF membrane using Towbin transfer buffer (25mM Tris, 192mM Glycine, 10% methanol) and either a constant voltage of 100 V for 1 h at RT or a constant current of 150 mA, 4 °C overnight (referred to as overnight transfer). Membranes were blocked in 5% milk in TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.4) solution for 1 h, incubated with primary antibody in TBS-T (1x TBS, 0.01% Tween 20) for 3 h at RT and then washed 3 times for 5 min in TBS-T. For chemiluminescent detection membranes were incubated with a HRP conjugated anti-mouse or anti-rabbit secondary antibody in TBS-T for 1 h at RT, washed 4 times for 5 min in TBS-T and following a final 5 min TBS wash incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for 5 min. Bands were imaged using film (Fuji) or CCD camera (Fujifilm LAS-3000). For fluorescent detection, membranes were incubated in the dark with a near infra-red fluorophore conjugated secondary antibody (IRDye® 680RD, Li-COR Biosciences) for 1 h at RT, washed as described above and imaged using an Odyssey SA scanner (Li-COR Biosciences) detecting in the 700 nm channel.

2.13. Automated capillary western blotting and data analysis

Automated capillary western blotting was performed on Wes instrumentation (ProteinSimple, CA). Samples were prepared for analysis per the ProteinSimple user manual. All reagents were provided by ProteinSimple other than the primary antibody. Samples were mixed at a 4:1 ratio with a 5 x master mix containing SDS (exact concentration unknown as master mix composition proprietary), 40 mM DTT and fluorescent molecular weight standards and heated at 95 °C for 5 min. Samples plus biotinylated molecular weight standards (ProteinSimple) were loaded along with blocking solution, wash buffers, primary antibodies, horseradish-peroxidase conjugated secondary antibodies and chemiluminescent substrate into a plate prefilled with stacking and separation matrices. Fully automated western blotting was performed using the Wes system; proteins were separated by electrophoreses at 375 V for 25 min, immobilised to the capillary by proprietary UV crosslinking and incubated with primary and secondary antibodies for 30 min each. Chemiluminescent signal was captured by a charge-coupled device camera and the resulting image was analysed by Compass software (ProteinSimple).

2.14. Quantification of labelled DOSCAT using Glu-Fib standard

Samples were analysed using nanoAcquity UPLC™ system (Waters) coupled to a Xevo™ TQS triple quadrupole mass spectrometer (Waters) in SRM mode with Q1 and Q3 operating at unit mass resolution. Digested sample (1 µL) was loaded onto a trapping column (C18, 180 µm x 20 mm, Waters) using partial loop injection for 3 min at a flow rate of 5 µL/min with 0.1% (v/v) formic acid. The sample was resolved on the analytical column (nanoACQUITY UPLC

HSS T3, C18, 75 µm x 150 mm x 1.8 µm column, Waters) using a gradient of 97% A (0.1% (v/v) formic acid) 3% B (99.9% (v/v) acetonitrile, 0.1% (v/v) formic acid) to 60% A 40% B over 30 min at a flow rate of 300 nL/min followed by washing with buffer B and re-equilibration. SRM analysis was performed using an electrospray voltage of 3000 V and a source temperature of 80°C. [Glu1]-Fibrinopeptide B was detected by measuring three transitions of both isotopic variants over the entire course of the chromatographic run. Quantification was performed by integrating extracted ion chromatograms in heavy and light channels and comparing the two.

2.15. Quantification of peptides by SRM and data analysis

Quantification was performed on a Xevo™ TQS triple quadrupole mass spectrometer coupled to a nanoAcquity UPLC™ system (Waters) (using the parameters described above). Analysis was performed using the same settings as previously described in scheduled SRM mode in which 15 data points were acquired over a 30 sec chromatographic peak within a 4 min window. Collision energy was optimised for each peptide (Supplementary Table 1).

Samples were prepared so that 1 µg cell lysate with either 1 fmol or 0.1 fmol DOSCAT were analysed by SRM methodology. Samples were prepared by serial dilution of the master mix sample with the cell lysate only digest. Both isotopic variants of each target peptide were analysed by three transitions. The final transition list was divided in two runs to achieve a minimum dwell time of 30 msec with each sample being analysed with both transition lists.

Data were analysed by Skyline (MacLean et al., 2010) and absolute quantification values were calculated from the standard:analyte ratio and the known concentration of internal stable isotope labelled standard. Quantification as copies per cell was derived from knowledge of the number of cells loaded onto the column and the loading of the accurately quantified DOSCAT standard.

2.16. Tandem Mass Spectrometry and data analysis

LC-MS/MS was performed on a nanoAcquity chromatography system coupled to either Synapt G2 (Waters) or LTQ-Orbitrap Velos (Thermo Fisher Scientific) instruments, or a Ultimate 3000 nano system (Dionex/Thermo Fisher Scientifc) coupled to a QExactive (Thermo Fisher Scientific) instrument.

For the nanoAcquiity system, samples were trapped onto a Symmetry C18 precolumn (180 µm id, 20 mm long, 5 µm particles) (Waters Corporation) over 3 min, at a flow rate of 25 µL/min in 2% (v/v) ACN /0.1% (v/v) formic acid. Bound peptides were resolved on a nanoAcquity UPLC C18 column (75 µm id, 150 mm long, 3 µm particles) at 300 nL/min over a 60 min linear gradient from 3 to 85% (v/v) ACN in 0.1% v/v formic acid. For the ultimate 3000 nano system, samples were loaded onto a trap column (Acclaim PepMap 100, 2 cm x 75 µm inner diameter, C18, 3 µm particle size, 100Å pore) at 5 µL/min with an

aqueous solution containing 0.1% (v/v) TFA and 2% (v/v) ACN. After 3 min, the trap column was set in-line with an analytical column (Easy-Spray PepMap® RSLC 15 cm × 75 µm inner diameter, C18, 2 µm, 100Å) (Dionex/Termo Fisher). Peptides were eluted by using an appropriate mixture of solvents A and B. Solvent A was 0.1% (v/v) formic acid in HPLC-grade water, and solvent B was 0.1% (v/v) formic acid in HPLC grade acetonitrile 80% (v/v). Separations were performed by applying a linear gradient of 3.8% to 50% solvent B over 35 min at 300 nL/min followed by a washing step (5 min at 99% solvent B) and an equilibration step (15 min at 3.8% solvent B).

The LTQ-Orbitrap Velos was operated in data-dependent acquisition mode. The 20 most intense multiply charged ions were sequentially fragmented at 35% of normalized collision energy. Precursors selected were dynamically excluded for 20 s. The QExactive was operated in data-dependent acquisition mode. Following a full MS scan between m/z 350 to 2000 (mass resolution of 60,000 FWHM at m/z 200), a data-dependent top-16 method MS2 analysis was performed with a target value of 1 x 10^5 ions determined with automatic gain control. Precursor ions were isolated with an isolation window of m/z 1.2, with scans acquired at a mass resolution of 30,000 FWHM at m/z 200 and dynamic exclusion of 20 s. The Synapt G2 instrument was operated in data-independent acquisition mode (MS^E).

Raw data for all instruments were converted into a *.mgf format peaklist file by Proteome Discoverer 1.1 (Thermo Fisher Scientific, Waltham, MA) using default parameters. Independent *.mgf files for each sample were searched against a database containing sequences of all DOSCAT iterations with MASCOT search engine (version 2.5.1, Matrix Science), using trypsin as specific enzyme, carbamidomethylation of cysteine as fixed modification, methionine oxidation as variable modification and one trypsin missed cleavage, a mass tolerance of 10 ppm for precursors and 0.6 Da for fragment ions. The false discovery rate (FDR) was calculated using the decoy database tool in MASCOT.

2.17. MALDI-TOF mass spectrometry

Samples were mixed 1:1 (*v/v*) with a 10 mg/mL solution of α-cyano-4-hydroxycinnamic acid in 50% ACN, spotted onto the MALDI target plate and air-dried. MALDI-TOF analysis was performed on an Ultraflex instrument (Bruker) operated in positive ion mode. Spectra were acquired at a laser energy of 40-70% of the maximum energy with 4000 shots per spectrum and a laser repetition rate of 1000 Hz. Spectra were acquired between *m/z* values of 700 and 4000. Data were analysed using FlexAnalysis to generated peak lists, which were searched against an in-house DOSCAT sequence database using Mascot.

Chapter 3: Proof of principle of a DOSCAT approach

3.1. Introduction

As with any novel technology, feasibility of the approach must be established to show that it has practical potential. Therefore, the first stage of the work is to design and use a DOSCAT to quantify a set of proteins in a well-defined model system. This includes the establishment and refinement of DOSCAT design principles, optimisation of protein expression and purification, and finally, a demonstration that the standard can be used to successfully calibrate both quantitative WB and SRM-MS experiments.

If western blotting using DOSCAT is to be routinely used for absolute protein quantification, it is important to understand how the multiple variables inherent to the technique contribute to quantitative accuracy and reproducibility. Therefore, another component of the proof of principle work will be determine the best method to execute a quantitative WB. Different WB methodologies and their advantages and disadvantages have been reviewed extensively in Chapter 1. Western blotting using ECL or fluorescent detection and, more recently, automated capillary WB, are the most routinely used methods for protein quantification. These three platforms will be evaluated in terms of their characteristics to deliver the best quantitative results, permitting an informed choice of the optimal technique to use with DOSCAT for future quantitative experiments.

As part of a model system to test the DOSCAT principle, proteins involved in NF-κB signaling were selected. The signaling pathways governing NF-κB activation are well-studied (Figure 3.1) and through the regulation of a wide spectrum of genes they are implicated in inflammation, immune response and disease (Hoesel et al., 2013; Karin, 2009; Lawrence, 2009). The NF-κB transcription factors are made up of Rel-family proteins including p50, RelA/p65, p52 and RelB (Hayden et al., 2012). In unstimulated cells, these proteins are inactivated by the binding of the inhibitory kappa B family of proteins, including IκBα, IκBβ and IκBε (Karin et al., 2000), or by forming a complex with p100. In the canonical pathway, cell stimulation (from, for example, growth factors, proinflammatory cytokines and antigen receptor binding) leads the activation of the IκB kinase (IKK) complex. This phosphorylates IκBs, leading to their poly-ubiquitination and subsequent degradation in the proteasome (Palombella et al., 1994; Scherer et al., 1995). This has the effect of activating the NF-κB transcription factors p65 and p50 through their translocation to the nucleus where they bind to DNA and influence gene expression (Karin et al., 2000). In the non-canonical pathway, the activation of different receptor classes causes the activation of the NF-κB inducing kinase (NIK), which phosphorylates and activates IKKα. This goes on to phosphorylate p100 leading to its ubiquitination and partial degradation to p52, which in complex with RelB activates gene expression.

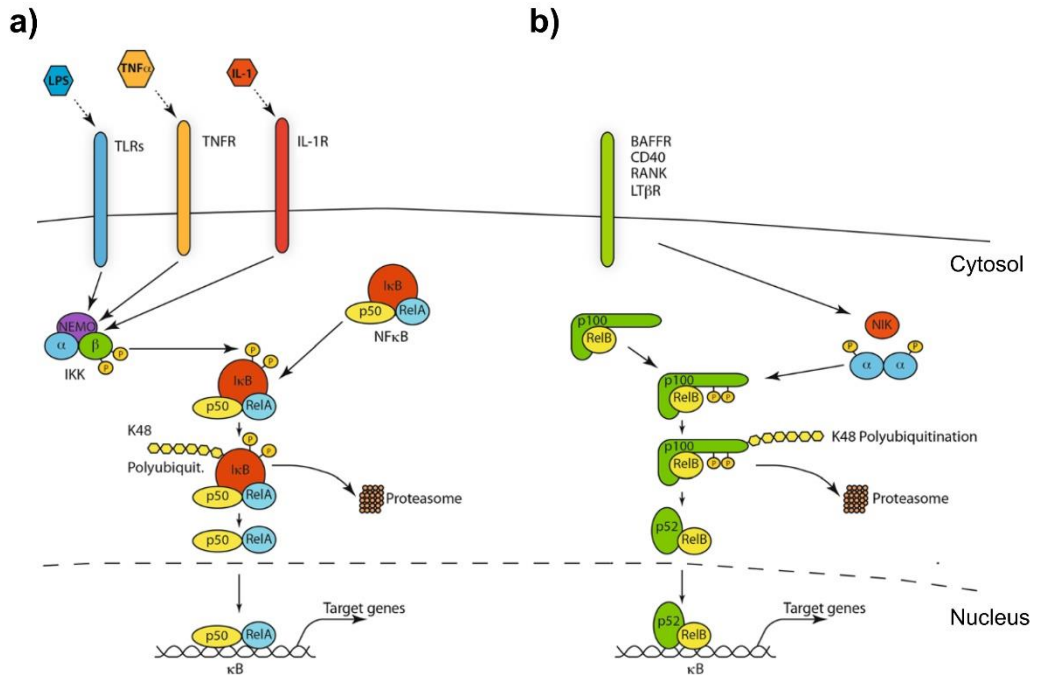


Figure 3.1 The canonical and non-canonical NF-κB signalling pathway. a) In the canonical pathway lipopolysaccharides (LPS), tumour necrosis factor α (TNFα) or interleukin-1 (IL-1) activate Toll-like receptors (TLRs), tumour necrosis factor receptor (TNFR) and interleukin-1 receptor (IL-1R) respectively. This leads to IKKβ activation in the IKK complex, which then phosphorylates IκBα leading to its polyubiquitination and degradation in the proteasome. The NF-κB complex is then free to translocate to the nucleus and activate gene transcription. b) The non-canonical pathway; the activation of B-cell activation factor (BAFFR), CD40, receptor activator for nuclear factor kappa B (RANK) or lymphtoxin β-receptor (LTβR) leads to the NIK-mediated activation of IKKα, which phosphorylates p100 leading to its degradation to p52, allowing p52-RelB heterodimers to active gene transcription. Figure adapted from Hoesel and Schmid (2013).

Designing a DOSCAT that targets proteins contained in the NF-κB pathway has several advantages. As the molecular and regulatory mechanisms involved in NF-κB mediated signaling are so well understood, it lends itself as an ideal model system for the proof of principle quantification experiments. The change in abundance of target proteins upon perturbations by specific signaling events is well profiled, so there is a benchmark to test the results from DOSCAT mediated quantitation against. Moreover, no studies to date have quantified in copy per cell terms the NF-κB proteins in a cell, so there will be additional value in generating quantitative data using this system for systems biology and mathematical modelling projects. An additional advantage to using a well-studied system is the abundance of analytical reagents (specifically antibodies) to detect proteins, which is critical when designing a DOSCAT.

The aim of this chapter is to complete the early design and optimisation work required to demonstrate that the DOSCAT approach is viable and potentially a useful technology. Using five target proteins from the NF-κB pathway, DOSCAT design principles will be outlined and

expression and purification protocols optimised. The utility of a DOSCAT approach will be determined by testing whether peptides contained within DOSCAT are detectable by MS and whether epitopes contained within its sequence can be recognised by target antibodies. Furthermore, three different WB methods most appropriate for quantitative work will be assessed to determine which has the best analytical properties for protein quantification.

3.2. Results and discussion

3.2.1. Analysis of quantitative western blot techniques

For absolute protein quantification by immunoblotting, it is important to understand the characteristics of different immunoblotting platforms so an assessment can be made as to which is be most suitable for quantification. Three different western blot platforms were tested. Two of these were classic WB, where following electrotransfer to PVDF membrane the detection step was performed through either an HRP linked secondary antibody and electro-chemiluminescence captured by a CCD camera, or a near infra-red fluorophore linked secondary antibody that can be detected by digital imaging (LI-COR). The third technique was based on automated capillary blotting using the Wes system (Protein Simple). Each platform was evaluated in terms of inter- and intra- assay variability, linear range and sensitivity.

Three QconCAT proteins of differing molecular weights were selected as a training set (Table 3.1). Each protein contained a C-terminal His tag, allowing for all proteins to be detected using the same His-tag antibody. The three proteins were quantified by a Bradford assay and combined at an equimolar ratio into a single master mix, which was serially diluted and loaded in replicates across the same gel (ECL, n = 4; LI-COR, n = 3, Wes, n = 3). An extended dilution series was run for the LI-COR and Wes systems as an increased dynamic range compared to classic ECL was expected. For both classic WB approaches densitometry on the bands was performed using Odyssey software (LI-COR) and for Wes the associated Compass software (Protein Simple) was used for analysis. This experiment was replicated twice for each platform. Exemplar blots for each platform are displayed in Figure 3.2. It should be noted that for capillary WB the Compass analysis software generates a pseudo-gel image, in which the intensity of each band is relative to the most intense signal on the gel image. For this reason, relatively weak signals are not always visible on the gel without adjusting the image contrast, which can distort the more intense bands. Therefore, bands are not observable for lower concentrations of master mix even though a signal was measured.

Table 3.1 QconCAT proteins used as a training set for western blot analysis.

Protein	Molecular weight (kDa)	Stock concentration (pmol/µL)
DBK1	87.8	2.3
CC02	60.0	3.4
LS1	30.7	15.2

The signal for each concentration point was first averaged across technical repeats within a single blot, and this value was averaged for all blot replicates (Figure 3.3). The response between concentration and signal was not identical between proteins (see difference in slope gradients), suggesting that either the protein concentrations within the master mix or the signal response was inaccurate. To assess this, a sample of the master mix was analysed by SDS-PAGE alongside each individual protein (Figure 3.4). The relative signal for the individual QconCAT proteins agreed with their measured concentrations, with LS1 being the most intense followed by DBK1 and CC02. However, the master mix clearly did not contain an equimolar mix of each QconCAT as intended. The signal for CC02 was the most intense, with the signal for DBK1 and LS1 about 50% and 20% of this respectively. This would account for the observed differences in signal response; the gradient of the slopes generated by automated WB analysis closely followed this distribution. This was not the case for ECL and LI-COR results, though. This could be explained by signal saturation effects limiting the maximum signal so that even though more protein is present, the measured signal and response over a concentration range is similar.

The results from the SDS-PAGE gel suggests that protein concentrations in the master mix were inaccurate. This makes an accurate assessment of assay sensitivity in terms of an absolute value difficult. However, linearity and dynamic range can be directly compared between the three platforms. As mentioned, signal saturation at higher protein concentrations is evident for classic ECL, with r^2 values of 0.93 and below, representative of a small linear dynamic range. An improved dynamic linear range was obtained by fluorescent detection for DBK1 and LS1 with $r^2 > 0.97$, but for CC02 the fluorescent signal saturated when more than 0.5 pmol protein was loaded. The linearity of response for automated WB was very good across all three test proteins ($r^2 > 0.97$), however, the linear signal falls away at lower concentrations of proteins, which indicates a lack of sensitivity compared to fluorescent and ECL detection.

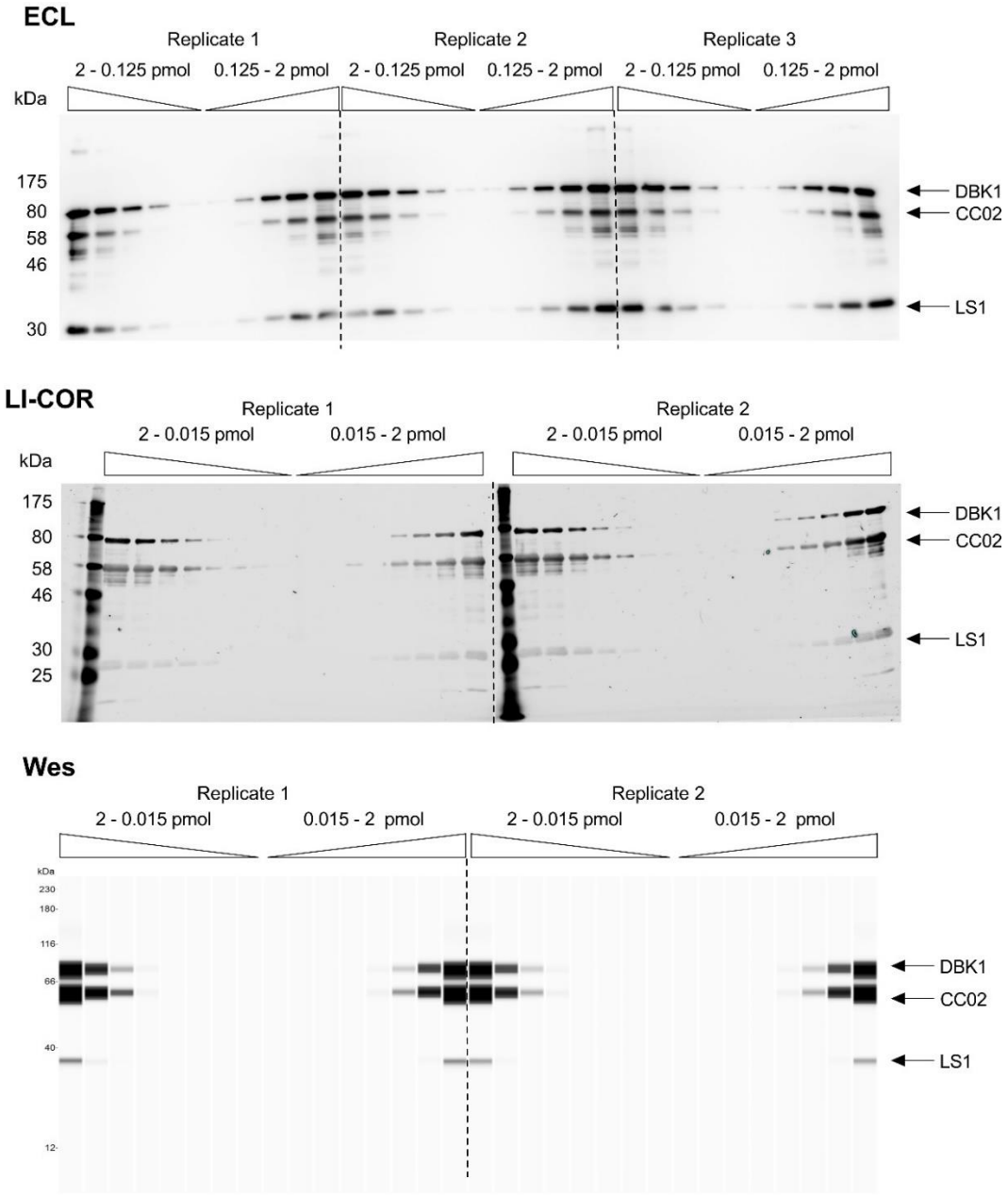


Figure 3.2 Analysis of QconCAT master mix using ECL, fluorescent (LI-COR) and automated capillary western blotting (Wes). A master mix of the QconCATs DBK1, CC02 and LS1 was prepared and a dilution series created in sample buffer. For ECL and fluorescent WBs the master mix was separated by SDS-PAGE, transferred onto PVDF membrane, probed with anti-His tag antibody at a 1:1000 dilution and detected using ECL or fluorescence (using the LI-COR system) as described in methods. Automated capillary WB was performed on Wes instrumentation using the same anti His-tag primary Ab at a 1:100 dilution as described in methods.

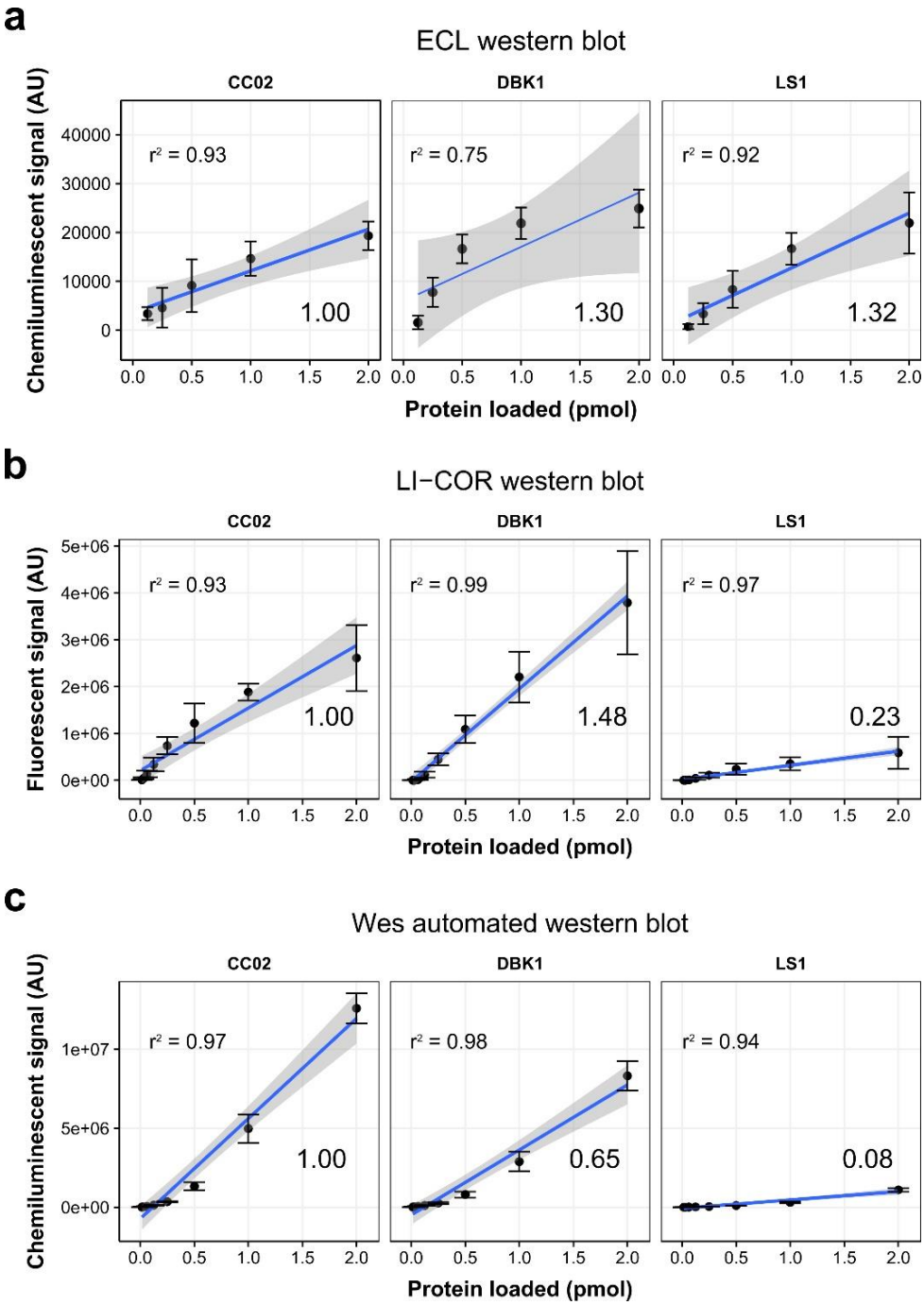


Figure 3.3 Signal response in different WB platforms for QconCAT proteins. Signal at each concentration point was determined by densitometry (for ECL and LI-COR data) or by Compass software (for Wes data) and averaged across technical repeats within a single blot, with this value being averaged for all blot replicates (n = 2). Error bars represent standard error, n = 2.

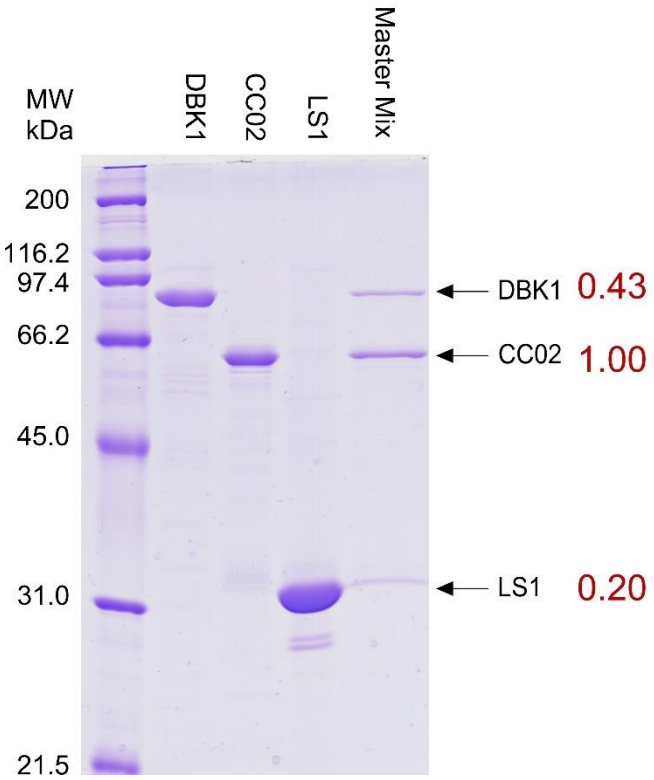


Figure 3.4 SDS-PAGE analysis of test QconCAT proteins and master mix. An equal volume of each QconCAT stock (DBK1, CCC02 and LS1) was mixed 1:1 with SDS sample buffer and analysed by SDS-PAGE, as was the QconCAT master mix containing a supposedly equimolar amount of each protein. The numbers on the right denote the densitometric ratio of the proteins in the master mix normalised to the signal of the CC02 band in the master mix lane only.

To find the intra assay variability, the coefficient of variance of technical repeats at each concentration was calculated. These CVs were averaged across the multiple gel replicates to give an average CV for each concentration point (Figure 3.5). Generally, intra-assay variability increased as less protein was loaded and the resultant signal approached the noise. The LI-COR system performed substantially worse than the other two systems tested, with CV > 20% in almost all cases. Classic ECL and the Wes system performed equivalently at higher concentrations of protein, but the performance on Wes was maintained down to the smallest amount of protein loaded. The mean signal of all technical repeats was calculated, and the variance in terms of % CV across different gel replicates was used to assess inter-assay variability at each concertation point (Figure 3.6). Again, variability increased as protein load decreased and out of the three platforms LI-COR performed the worst. Across the full range of protein loaded, lowest levels of inter assay variability were observed on the Wes system.

This work has compared classic WB using either ECL or fluorescent detection with automated WB, showing that the automated platform Wes offers improved inter-assay and intra-assay variability and good signal linearity and dynamic range.

For two proteins LI-COR demonstrated the best signal linearity across the range of protein concentrations tested. The lack of linear response for CC02 when measured by fluorescence is unusual, and there does not seem to be an obvious explanation. Both of the other platforms rely on electrochemiluminescene for detection, which is known for its limited linear dynamic range due to signal saturation (Elbaggari et al., 2008; Taylor et al., 2013). Fluorescent detection does not have this limitation and has shown to vastly improve linearity compared to classic ECL (Zellner et al., 2008). The data presented here (other than for CC02) would agree with this, although the difference with automated WB is much subtler. A caveat to this is the very low signal at the lowest protein concentrations detected by automated WB. These are essentially below the sensitivity limit and so affect the coefficient of determination value. If these data points were removed Wes would perform as well as LI-COR detection, albeit not with equivalent sensitivity. It may well be possible that the linear response would continue above the amount of protein loaded in these experiments, so further experimentation would be required to determine the full dynamic linear range of both systems.

The intra- and inter- assay CVs measured in this experiment agree well with previous studies that have assessed variability in automated WB and classic WB using ECL detection (J. Q. Chen et al., 2013; Loughney et al., 2014; Rustandi et al., 2012). It has been demonstrated that the majority of error in electrophoresis and blotting experiments is introduced by the operator (Koller et al., 2005), so it is no surprise that the Wes system, which automates all of the process other than sample preparation, performs the best out of the three platforms. What is less clear is why the LI-COR system performs so poorly in this experiment. The LI-COR system would be expected to perform as well or better as ECL detection, but this was not the observed result. The LI-COR and ECL experiments were performed at different times, so it is of course possible that multiple factors (reagent and buffers stability, protein transfer, antibody binding etc.) combined to give poor reproducibility on the given day. Error may also have been introduced in the densitometric analysis; fluorescent detection resulted in a higher background and non-specific bands (Figure 3.2) that may have interfered with accurate measurement of signal intensity. A more accurate assessment of variability between ECL and fluorescent detection methods could be obtained performing electrophoresis, transfer and immunoblotting together, and splitting the membrane prior to the detection step (secondary antibody incubation onwards).

Although this study gives a good insight into the best platform for quantitative WB is it limited by the relatively small number of proteins, antibodies and replicates used. Different antibodies and proteins may engender different results, and whilst the number of replicates performed here was limited by the available resources, increasing them would allow for a proper statistical assessment of differences between the techniques. Additionally, there are many more variables in a western blot that went untested in this experiment, such as primary and secondary antibody concentration, incubation and development times and electrotransfer settings.

Despite these limitations, there is a clear suggestion from the data that out of the three platforms tested automated WB using Wes offers the best properties for quantitative work. Moreover, compared to classic western blot approaches, automated WB is much faster (3 h vs ~ 2 days) and more economical. This raises the prospect of reproducible, high-throughput, quantitative assay that could be calibrated using DOSCAT technology.

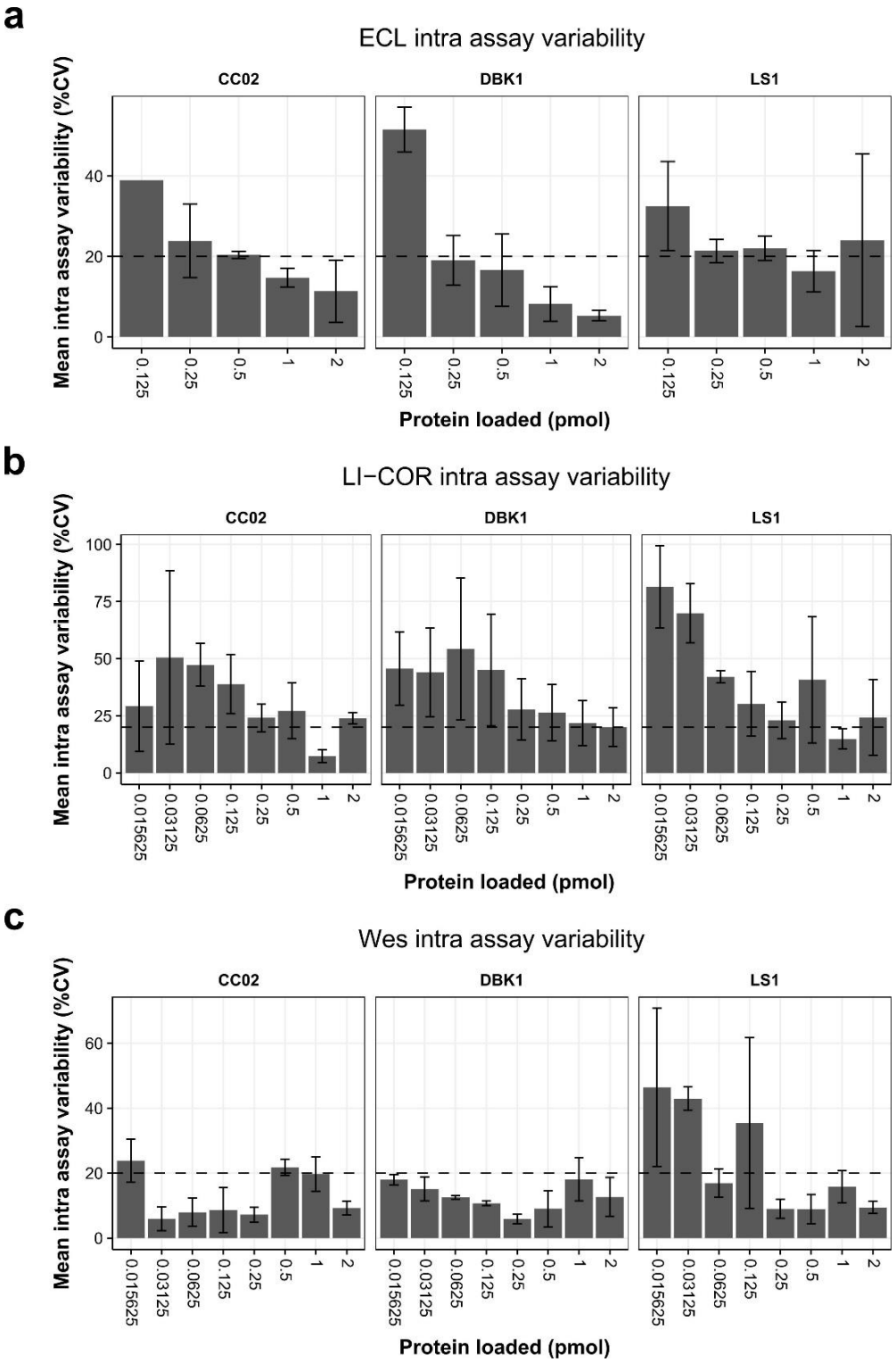


Figure 3.5 Intra assay variability for each western blot platform. Variability in terms of % CV for each concentration point was calculated from technical repeats on each gel replicate (ECL n = 4; LI-COR, Wes n = 3). Error bars represent standard error between gel replicates, n = 2. Note that CC02 at 0.125 pmol was not detected on one of the ECL blots, so a standard error calculation was not possible.

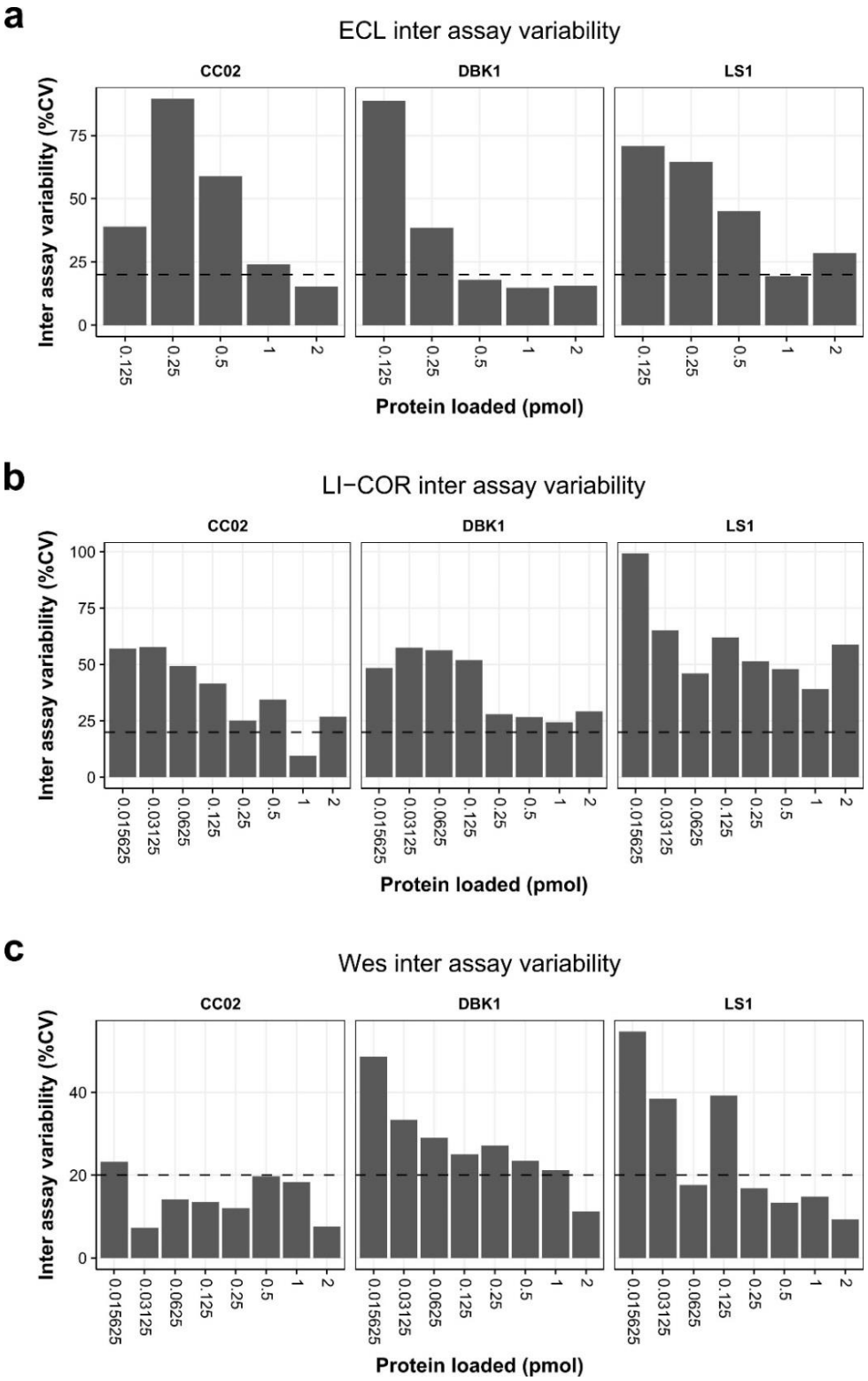


Figure 3.6 Inter assay variability for each western blot platform. Technical repeats for each concentration point across a single gel were averaged and the variance between gel replicates (n = 2) was calculated in terms of % CV.

3.2.2. NFκB-DOSCAT design, expression and validation

Five target proteins involved in NF-κB signaling were selected; p65, RelB, IκBα, IκBβ and IκBε. For each protein, an epitope sequence for a specific antibody and at least two quantotypic peptides were selected and concatenated into a single sequence to create what will be referred to as the NFκB-DOSCAT.

The selected antibodies and the epitopes used in the NFκB-DOSCAT are summarised in Table 3.2). The selection of epitopes to include in the NFκB-DOSCAT was driven by the commercial availability of antibodies for which the epitope was known and well characterised. This can be difficult to obtain from antibody manufacturers as they are not always willing to disclose peptide immunogen sequences or epitope mapping results due to commercial sensitivity. Some manufacturers would, however, disclose the amino acid around which the epitope was centered. Based on this information, the epitope region for DOSCAT was defined as the 25 amino acids either side of the disclosed central residue in the endogenous protein sequence. Since the average immunogen sequence is approximately 15 - 20 amino acids (Hancock et al., 2005), there was confidence that this approach would allow for the inclusion of the true epitope. As well as having epitope sequence data, the antibody itself must be well validated, although this is not always possible, especially on less well studied proteins. For each selected antibody, there was accompanying western blot data (provided by the manufacturer) to demonstrate the antibody detected a single protein at the expected molecular weight within a complex sample (e.g. cell lysate).

Established criteria for selecting Q-peptides, developed through multiple quantitative proteomics studies employing stable isotope labelled peptides, were used to choose Q-peptides to be included in the NFκB-DOSCAT. Firstly, the data repositories PeptideAtlas and Global Proteome Machine were accessed and peptides that met the criteria laid in Chapter 1 were shortlisted for inclusion. Using the data repository approach peptides could not be identified for IκBα and RelB, so other peptides were selected based on their PeptideSieve score and predicted observability. Since the design of the NFκB-DOSCAT, other tools have become available for the selection of Q-peptides. The consensus prediction system CONSeQuence uses four independent machine learning algorithms to select Q-peptides that, based on their physiochemical properties, are predicted to have the best detectability in an ESI-MS experiment (Eyers et al., 2011). More recently developed is SRM ATLAS, a database of validated SRM assays for 99.7% of the human proteome (Kusebauch et al., 2016). SRM coordinates for up to five peptides per protein are contained in the database, allowing for rapid identification of optimal Q-peptides. For the design of future DOSCATs, these tools could be employed together to maximise the possibility that a peptide will ionize efficiently and quantitative data will be obtained. To each selected Q-peptide a flanking region of 3 amino acids from the endogenous sequence were added to either end of the peptides. This strategy has been shown to improve quantitative accuracy by equalising digestion efficiency between standard and analyte (Cheung et al., 2015).

In addition to the peptide and epitopes to enable quantification, several restricted specificity endopeptidase sites were included in the NFκB-DOSCAT design. This allowed for the possibility of shifting the electrophoretic mobility of NFκB-DOSCAT in QWB, conferring greater flexibility if standard and analyte had overlapping motilities, or introducing the possibility of multiplexing to measure two analytes within a single lane. Four endopeptidase were selected (Figure 3.7); each being frequently applied to cleave fusion tags from recombinant proteins and so widely available commercially as well as with a well-defined cleavage site that was not present in any target analytes.

Q-peptides, epitopes and restricted specificity endopeptidase sites were arranged *in silico* (Figure 3.7). The exact positions of Q-peptides and epitopes was arbitrary, with peptides for the same protein mostly grouped together and epitopes spaced out evenly throughout the NFκB-DOSCAT sequence. The endopeptidase sites were inserted between epitopes in such a way so that the two proteolytic fragments generated would not have a similar electrophoretic mobility shift as a target analyte when detected by the same antibody. At the N-terminus, a methionine initiator residue followed by short sacrificial peptide to protect true Q-peptides from exoproteolytic activity were added. Following this, the sequence of [Glu1]-Fibrinopeptide B (Glu-Fib, EGVNDNEEGFFSAR) was added to enable quantification of the NFκB-DOSCAT itself against a commercially available Glu-Fib standard. A hexa-histidine tag was added to the C terminus to enable purification.

Table 3.2 Antibodies selected for use with NFκB-DOSCAT.

Protein	Antibody	Clonality	Immunogen data
p65	CST #3034	Rabbit polyclonal	Centred around Ser276
RelB	CST #4954	Rabbit polyclonal	Centred around Ser424
IκBα	CST #9242	Rabbit polyclonal	Centred around Arg29
IκBβ	CST #9248	Rabbit polyclonal	Centred around Arg155
IκBε	CST #9249	Rabbit polyclonal	Centred around Gly205

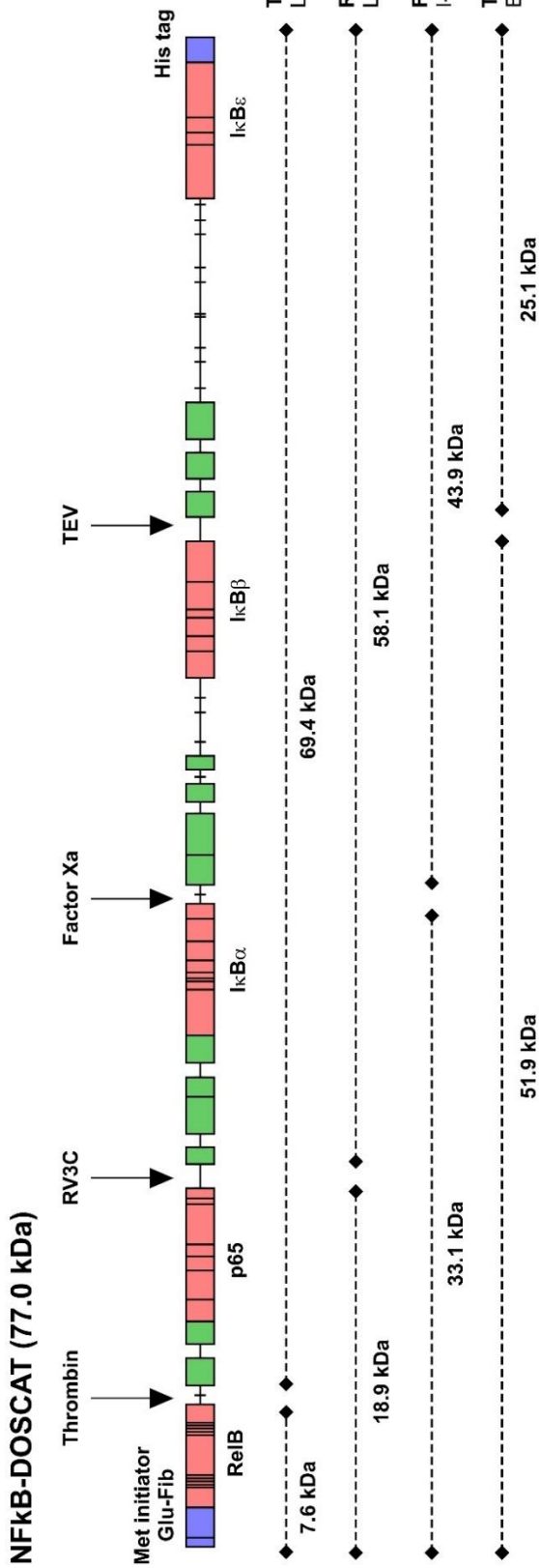


Figure 3.7 Protein map of the NFκB--DOSCAT. Green boxes represent quantotypic peptides, red boxes define the extent of antibody binding epitopes and purple boxes the [Glu1]-Fibrinopeptide B calibration peptide at the N-terminus and the hexa-His purification tag at the C-terminus. Arrows indicate the location of cleavage sites for each of the specific proteases and dotted lines the molecular weight of the fragments they generate upon cleavage (RV3C, human rhinovirus 3 C protease; TEV, tobacco etch virus protease).

The genes for the NFκB-DOSCAT were codon optimised for *E.coli* and cloned into pET21a vectors (Eurofins Genomics), which were transformed into BL21(DE3) *E.coli* cells. Cells were grown in minimal media containing [$^{13}C_6$]Arg/[$^{13}C_6$]Lys (Figure 3.8a) and expression was induced during mid-log phase by the addition of 1 mM IPTG (Figure 3.8b). The cell pellets were lysed by sonication and fractionated into soluble and insoluble fractions by centrifugation. Analysis of both fractions by SDS-PAGE revealed that the NFκB-DOSCAT was contained in the insoluble fraction. The insoluble fraction was solubilised in a strong chaotrope (6 M guanidine hydrochloride) and purified by Ni-NTA affinity chromatography (Figure 3.9a). The NFκB-DOSCAT was eluted over four fractions, which were pooled and dialysed in a 50 mM ammonium bicarbonate, 1 mM DTT storage buffer. Western blots using an anti his-tag antibody were used to demonstrate NFκB-DOSCAT purity (Figure 3.9b). Degradation products were visible at high (100 ng) loadings only, indicating that the degradation rate was low. As most of the NFκB-DOSCAT was intact and any degradation products could be approximately quantified as a fraction of total protein, degradation was not deemed a problem that would affect future quantitative experiments. Purified NFκB-DOSCAT was subjected to trypsin digestion and the peptides analysed by MALDI-TOF mass spectrometry (Figure 3.10). MS1 ions were searched against a database containing the NFκB-DOSCAT sequence using Mascot, with 44% sequence coverage. Out of the 13 Q-peptides built in to DOSCAT, 9 were observed by MALDI-TOF MS, with additional peptides from epitope and endopeptidase cleavage regions also visible. Furthermore, when searched against human and *E.coli* databases there were no significant protein matches were found, indicating that were no contaminants present.

For the first time a DOSCAT has been successfully expressed and purified, and subsequent analysis has demonstrated that the protein is intact with minimal degradation and that the expected sequence is correct. As they are designed on similar principles, it was predicted that DOSCATs would behave in a similar manner to QconCATs, and this is what was observed. The NFκB-DOSCAT accumulates in inclusion bodies, which is consonant with previous observations of over 10 years of expression with QconCAT constructs (P. J. Brownridge et al., 2012). As the NFκB-DOSCAT is an artificial protein designed *de novo*, there is not the same evolutionary pressure as there is on 'real' proteins to be soluble; therefore, upon expression in *E.coli* it accumulates into insoluble inclusion bodies (Fahnert et al., 2004; Kane et al., 1988). This does not present a problem in that DOSCATs do not need to be folded and it can in fact be advantageous as it acts as a pre-purification step. The expression and purification protocols developed for QconCATs can be directly transferred to DOSCATs with minimal changes, raising the prospect for fast, simple expression for future DOSCATs.

Not all Q-peptides were observed by MALDI-TOF analysis, which may be explainable by the limitations of using MALDI for ionisation. It is known that MALDI is a highly variable ionisation method that is heavily influenced by the matrix used, sample complexity and amount and types of peptides that are present within the sample (Szájli et al., 2008). MALDI displays a

preference for arginine-containing peptides (Krause et al., 1999), which is reflected in the observed results. It is possible that when analysed with LC-ESI-MS the peptides are readily observed for their use for quantification is permitted.

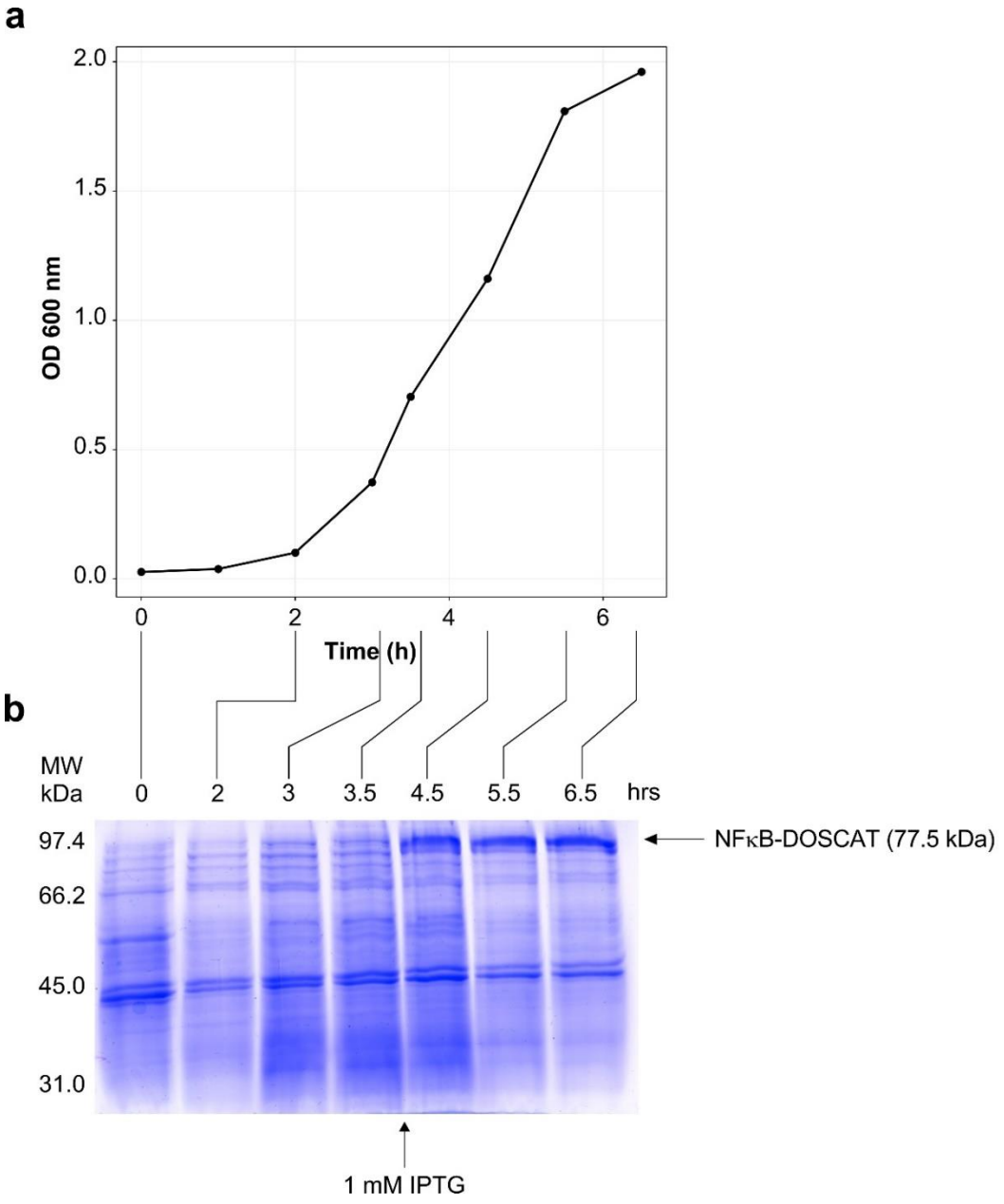


Figure 3.8 Expresion of NFκB-DOSCAT. a) Growth of transformed *E.coli* cells as measured by optical density at 600 nm. b) SDS-PAGE analysis of time points from *E. coli* culture containing NFκB-DOSCAT plasmid, grown in minimal media containing [$^{13}C_6$]Arg/[$^{13}C_6$]Lys for 3 h after inoculation, with expression induced by 1 mM IPTG at indicated time point.

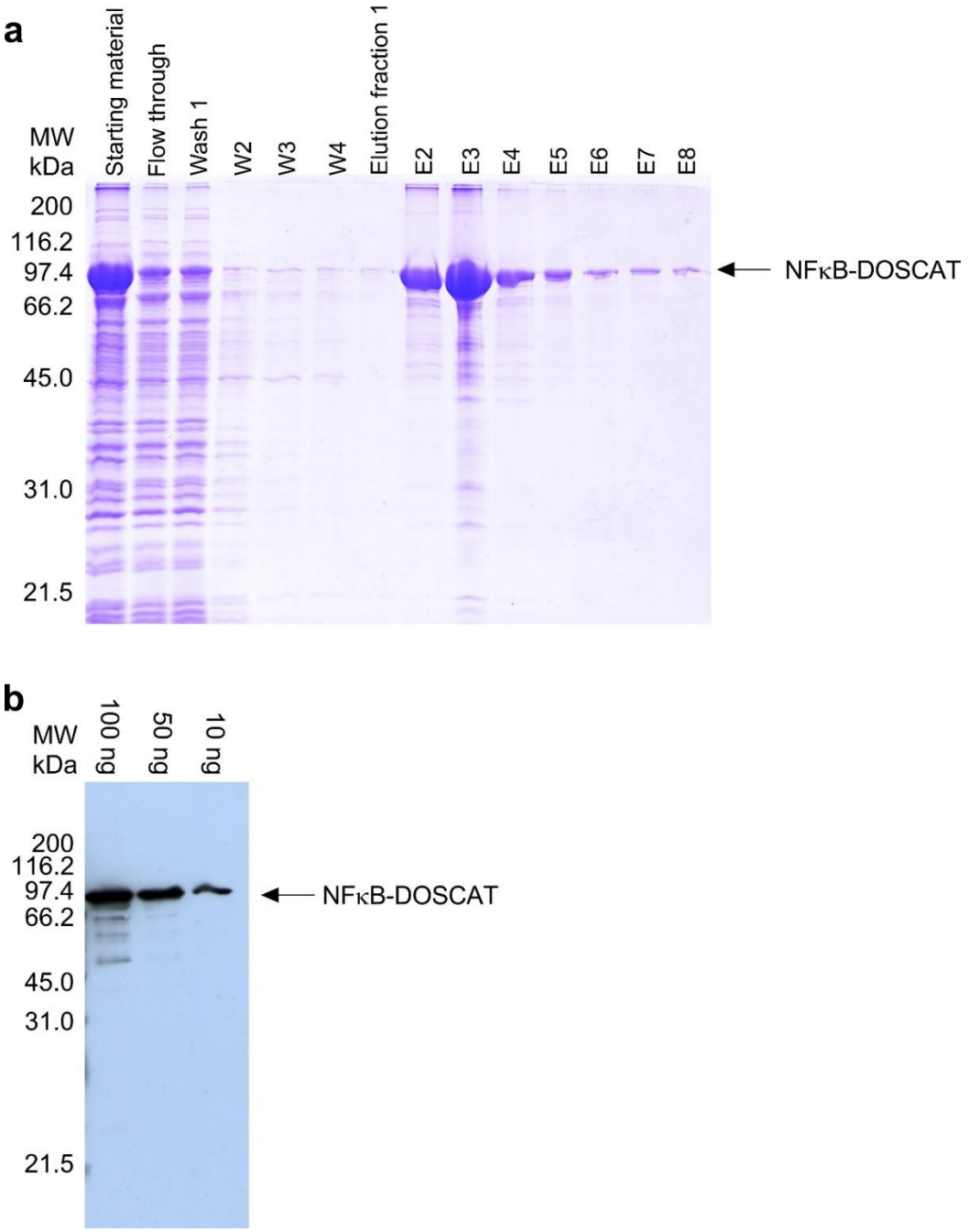


Figure 3.9 NFκB-DOSCAT purification. a) SDS-PAGE analysis of starting matrial, flow through, washes (W) and eluted fractions (E) resultant from NFκB-DOSCAT purification on a Histalon column; b) Western blot of purified NFκB-DOSCAT; a dilution series of purified NFκB-DOSCAT was separated by SDS-PAGE, transferred onto nitrocellulose membrane and probed with an anti His-tag antibody.

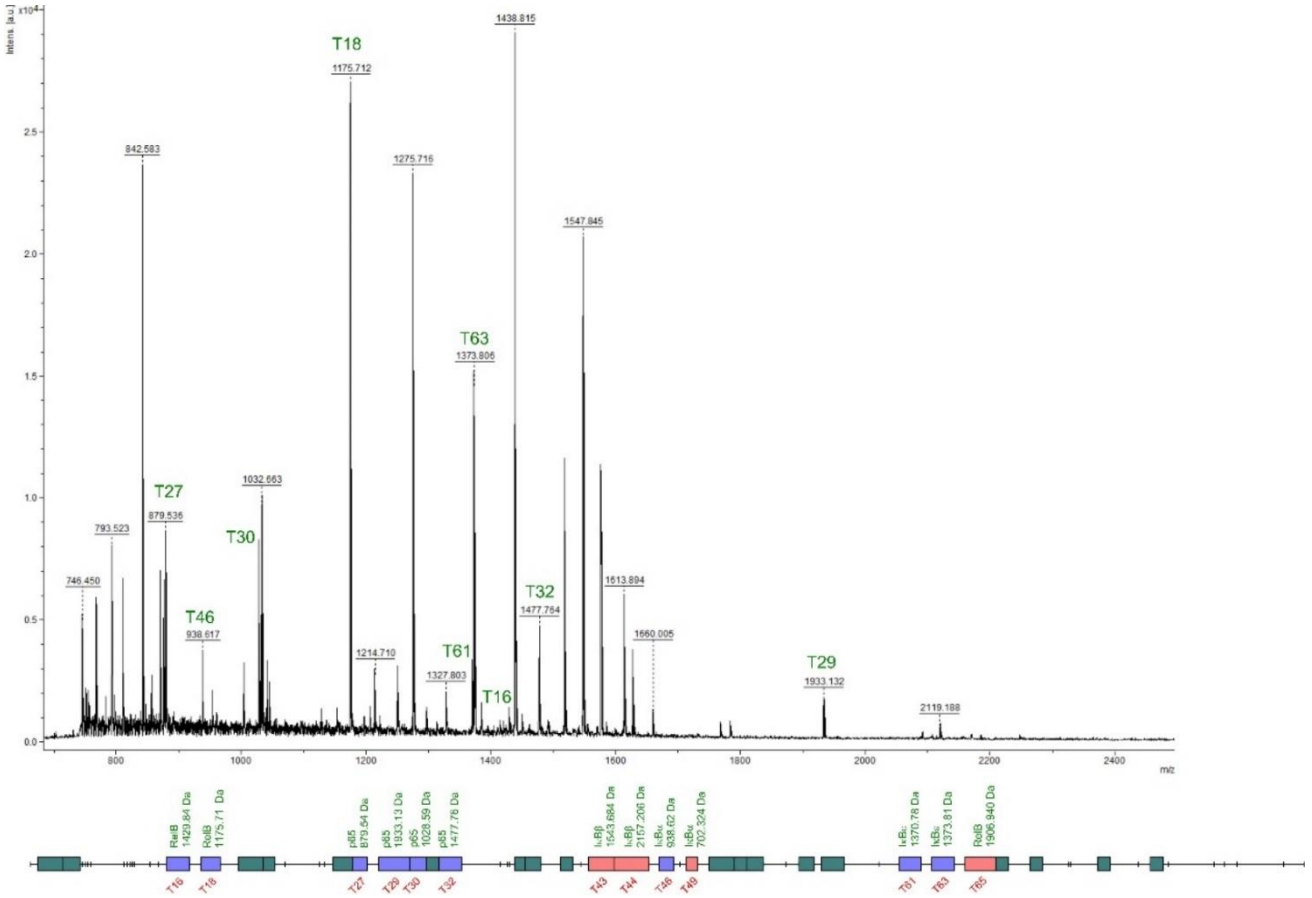


Figure 3.10 MALDI-TOF analysis of NFκB-DOSCAT. NFκB-DOSCAT was digested with trypsin and analysed by MALDI-TOF mass spectrometry. Top: mass spectrum annotated with observed Q-peptides resultant from the analysis; bottom: peptide map displaying obversed Q-peptides (green), other observed peptides (purple) and Q-peptides not observed (red).

3.2.3. NFκB-DOSCAT utility in western blotting

Once the NFκB-DOSCAT had been expressed and validated, the next step was to test if it could be detected in a western blot by the chosen NF-κB antibodies. Using a classic western blot approach, the NFκB-DOSCAT was detected by p65 and RelB antibodies but not by the antibodies to IκBβ or IκBε (Figure 3.11). There were additional non-specific bands above and below the expected molecular weight of the NFκB-DOSCAT that could be attributed to aggregation and degradation products, which would be noticeable when high amounts of protein are loaded. It should be noted that at this point stocks of the IκBβ antibody had been depleted, and the product line had in the interim been discontinued by the manufacturer. For this reason, it could not be used in subsequent experiments.

As future quantitative experiments were to be performed using automated western blotting, the remaining antibodies were tested in this platform using NFκB-DOSCAT and SK-N-AS cell lysate. The Wes system has different performance characteristics to a classic western blot, so initial experiments were performed to optimise methodologies and protein loading. Dilution series of NFκB-DOSCAT and SK-N-AS were loaded and probed with the RelB antibody. A clean signal was obtained for SK-N-AS cell lysate (Figure 3.12a), but not for NFκB-DOSCAT when loaded on its own (Figure 3.12b). At higher loadings of NFκB-DOSCAT (2 and 0.2 ng/µL) there were ‘dips’ in the electropherogram trace that were characteristic of protein burn out. This is where there are such high levels of protein loaded in the capillaries the ECL substrate is used up almost instantly before an image can be captured. Peaks present at 37 and 55 kDa may represent degradation products. In all traces, there was a constant signal above 100 kDa that scaled with the amount of NFκB-DOSCAT loaded. It was hypothesised that the observed signal was the NFκB-DOSCAT forming non-specific interactions with the capillary wall during electrophoresis rather than resolving to a clean peak. To test this, NFκB-DOSCAT was made up in buffers containing SDS, urea and SK-N-AS cell lysate. SDS and urea can act to disrupt protein-surface interactions, whereas the cell lysate would act as a complex background to minimise any interactions taking place. The addition of 2% SDS or 1 M urea had no effect on protein adsorption (Figure 3.12c), however, when the NFκB-DOSCAT was combined with cell lysate, a clear band for NFκB-DOSCAT and endogenous protein was observed (Figure 3.12d). Despite having a molecular weight of 77.5 kDa, the electrophoretic mobility of NFκB-DOSCAT in automated capillary WB was observed to be 100 kDa.

Based on this result, a dilution series of NFκB-DOSCAT was spiked into SK-N-AS cell lysate and probed with each antibody (Figure 3.13). Using p65, RelB and IκBα antibodies, distinct bands for the NFκB-DOSCAT and the endogenous protein were observed. However, only endogenous protein was detected by the IκBε antibody. Some variability between individual capillaries was observed, a result of the capillaries’ independence of one another, which can lead to unequal protein load or slight variations in antibody binding. This can be resolved by the addition of system control antibodies provided by the manufacturer, which detect a protein with an electrophoretic mobility of 26 kDa contained in the sample buffer. This can be used as

a loading control to normalise signal between capillaries. This experiment also allowed for a preliminary assessment of signal linearity and intra-assay variability when using NFκB-DOSCAT and a complex sample. NFκB-DOSCAT calibration curves exhibited excellent linearity, and signal variability between technical repeats of endogenous proteins was low (Figure 3.14).

These experiments demonstrate that by inserting multiple different epitope sequences into a DOSCAT protein, different antibodies can be used to detect a DOSCAT in classic and automated western blot formats. This is the first time multiple epitopes from different targets have been detected in a single protein.

The NFκB-DOSCAT was not detected by the IκBβ antibody; however, as it was not possible to test whether the antibody detected endogenous protein in cell lysate, it is difficult to say whether this is the fault of the NFκB-DOSCAT design or a poor antibody. The antibody for IκBε, however, detected endogenous protein but did not detect NFκB-DOSCAT in any experiment. As both endogenous protein and NFκB-DOSCAT were denatured and linearised as part of the sample preparation process, it is unlikely that the epitope in NFκB-DOSCAT was sterically shielded from the antibody, as can be the case in other immunoassays like ELISA. It is most likely that the epitope included in NFκB-DOSCAT did not reflect the actual epitope to which the antibody binds. Epitope details supplied by the manufacturer consisted of an amino acid around which the epitope was centered, and 25 amino acids around this central residue were included in NFκB-DOSCAT. Either the central residue data supplied was inaccurate, or the actual epitope was more than 25 residues away. Being a polyclonal antibody, however, it would be expected that at least some of the antibody population would bind to an area close to the central residue, unless there was a specific region further away that was particularly immunogenic. This is only speculation, though, and without details of the actual immunogen used or epitope mapping experiments it is difficult to say with any certainty why NFκB-DOSCAT was not detected. It does, however, highlight the importance of sourcing antibodies for which the epitope is well characterised and published before committing to make any DOSCAT protein.

Results using automated capillary western blotting indicate that DOSCATs have a propensity to readily adsorb to surfaces. This is not an unusual characteristic for proteins and peptides, especially at low concentrations (Suelter et al., 1983). Proteins are complex molecules and their tendency to bind to a surface depends on many factors such as amino acid composition, properties of the surface and properties of the local environmental (temperature, pH, buffer composition) (Rabe et al., 2011). By combining a DOSCAT with a complex matrix, other proteins are introduced that compete for and block surface binding sites, massively reducing DOSCAT adsorption. Whether this property is specific to the NFκB-DOSCAT or will be common to other DOSCATs produced in the future or remains to be seen.

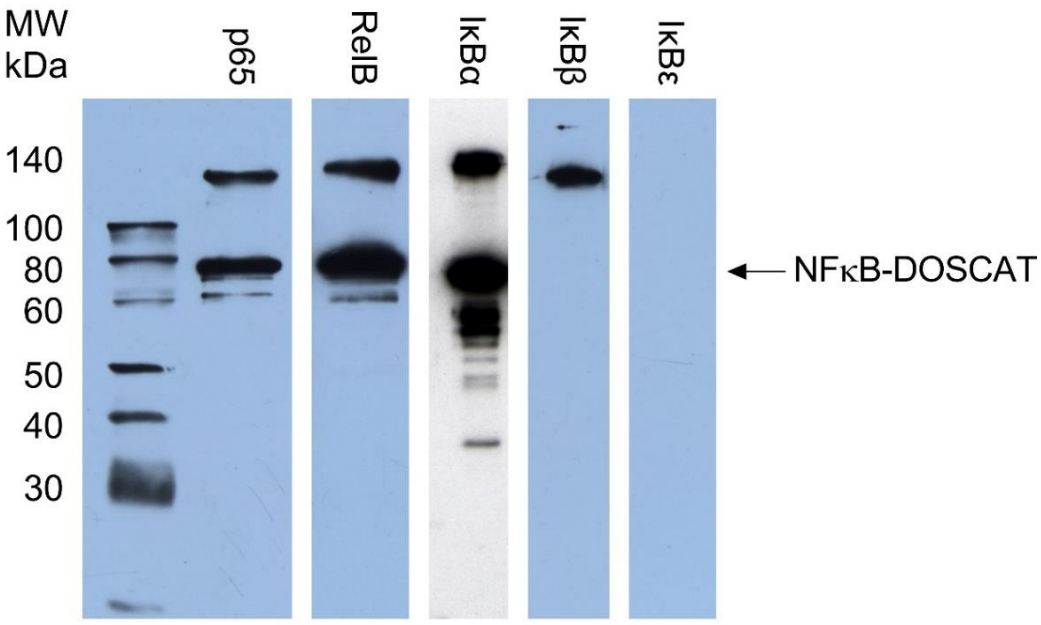


Figure 3.11 Detection of DOSCAT by NF-κB antibodies. Compositie image formed from multiple western blots of the NFκB-DOSCAT using each of the NF-κB antibodies for detection. The NFκB-DOSCAT was run on SDS-PAGE, transferred to nitrcellulose membrane, incubated with each antibody and ECL signal detected by film.

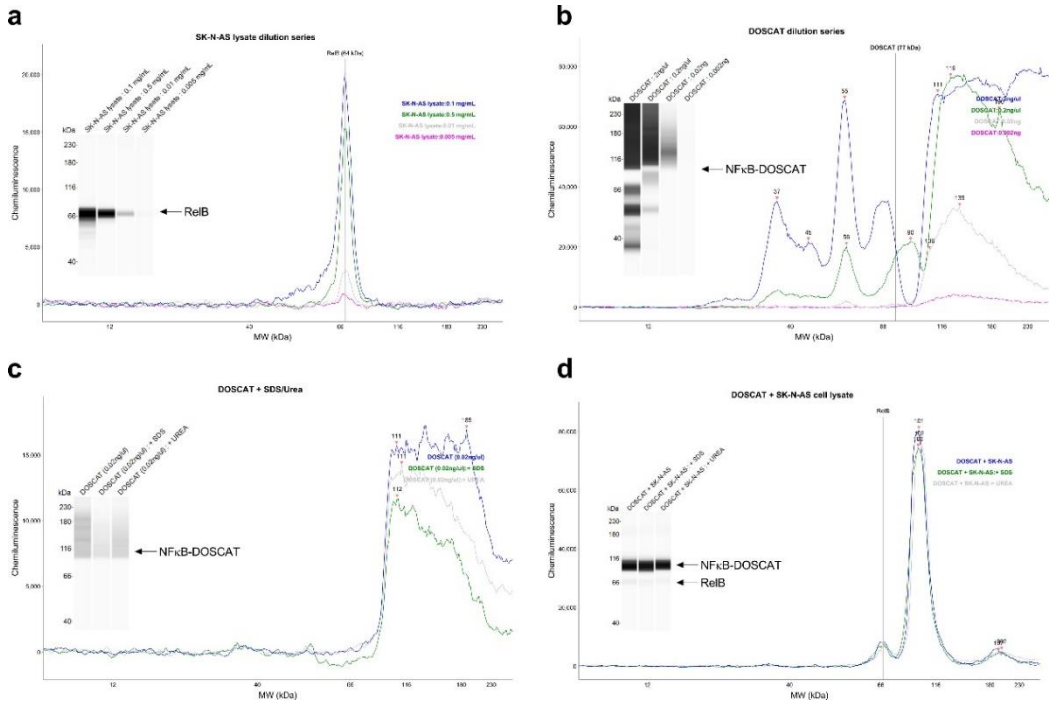


Figure 3.12 Optimisation of NFκB-DOSCAT detection on Wes. Electropherograms and gel views (inset) resultant from the analysis of the following by automated capillary western blotting, using an anti-RelB antibody for detection: a) dilution series of SK-N-AS cell lysate; b) dilution series of NFκB-DOSCAT; c) NFκB-DOSCAT (0.02 ng/µL) with no additives (blue trace), with 2% SDS (green), or with 1 M urea (grey); d) NFκB-DOSCAT (0.02 ng/µL) spiked into 0.05 µg/µL SK-NA-S cell lysate with no additives, with 2% SDS (green), or with 1 M urea (grey).

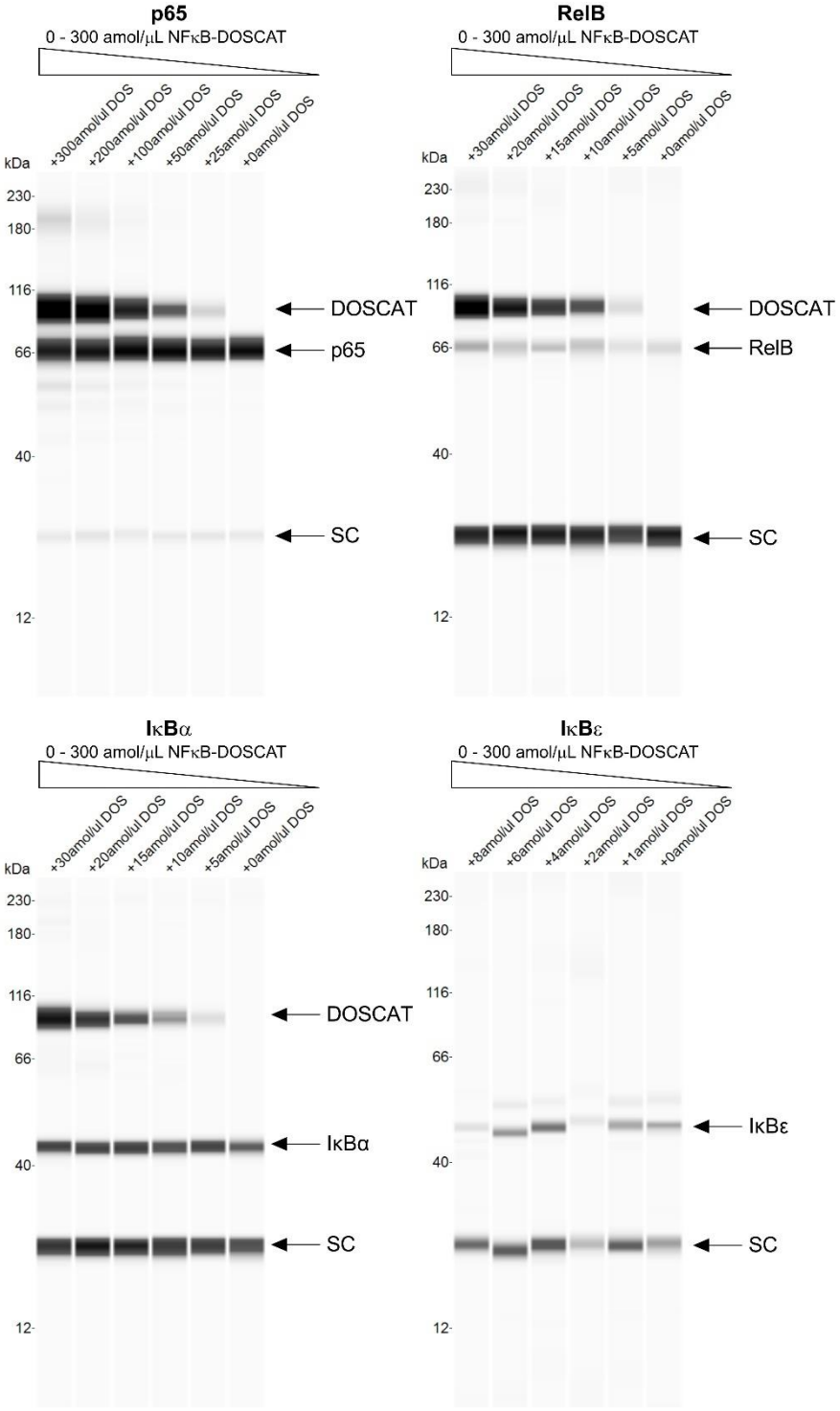


Figure 3.13 Detection of NFκB-DOSCAT by capillary WB using NF-κB antibodies. NFκB-DOSCAT was spiked into 0.2 mg/mL SK-NA-S cell lysate at various concentrations, serially diluted with the cell lysate as a diluent and analysed by automated capillary western blotting using each of the available NF-κB antibodies. System control (SC) protein contained within the sample buffer was also detected in tandem by a specific antibody to aid with signal normalisation across capillaries.

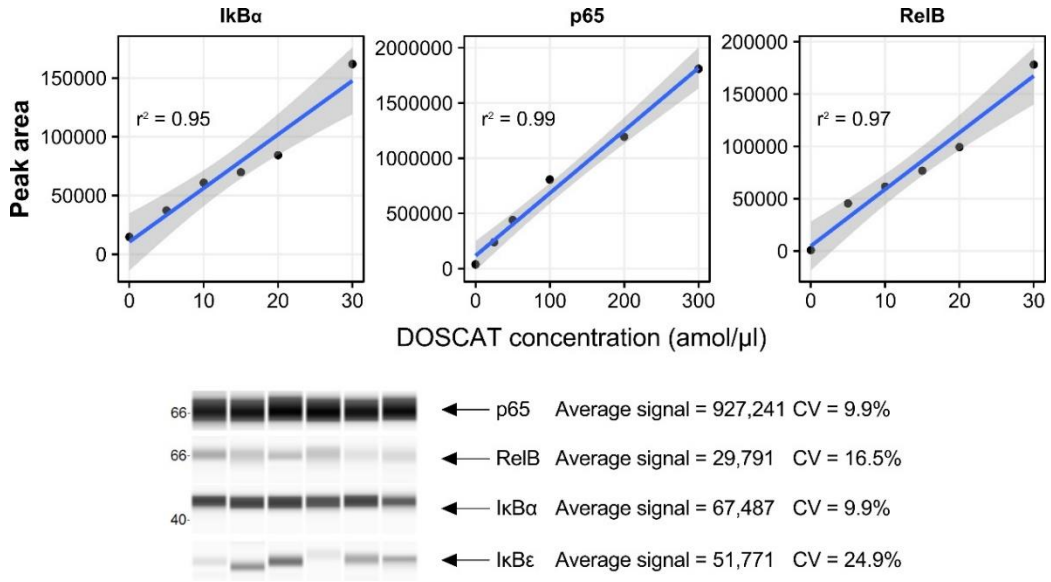


Figure 3.14 Linearity and intra-assay variability of DOSCAT and endogenous proteins detected by capillary WB. NFκB-DOSCAT spiked into 0.2 mg/mL SK-NA-S was analysed by capillary WB using each of the NF-κB Abs; the chemiluminescent signal of the NFκB-DOSCAT was plotted as a function of its concentration and the correlation coefficient calculated (top panels). The mean value and variation of the signals for technical repeats (n = 6) of detected endogenous proteins was calculated (bottom panels).

3.2.4. Restricted proteolysis of NFκB-DOSCAT to effect mobility shift

The NFκB-DOSCAT contains cleavage sites for four proteases of restricted specificity. To assess protease cleavage, NFκB-DOSCAT was incubated with each of the specific proteases at concentrations and temperatures recommended by manufactures and the literature. Samples were taken at various time points during the incubations and run on SDS-PAGE to assess cleavage. Both TEV protease and RV3C protease cleaved NFκB-DOSCAT to produce fragments at the expected molecular weights, but the cleavage specificity for the other two proteases was less clear (Figure 3.15). Proteolysis by thrombin generates many products, which amongst them may include the expected NFκB-DOSCAT cleavage products at 70 kDa and 7.5 kDa (highlighted by the boxes). However, these fragments get less intense over the course of the reaction and are not visible overnight incubation. Digestion with Factor Xa also produces many fragments that degrade further over time, meaning that it is difficult to interpret the results from the gel. It is possible that the fragment at 33 kDa is visible after overnight incubation, however whether there is a fragment a 44 kDa is less clear.

Only two of the proteases used have shown results as expected. TEV and RV3C proteases have previously demonstrated their ability to cleave protein completely with a high level of specificity (Advance, 2005; Eliseev et al., 2004; Rigaut et al., 1999; Senger et al., 1998), an outcome that is mirrored in the results obtained here. Both belong to the picornavirus 3C-like protease family, with long consensus sequence that if altered result in a major reduction in cleavage (Cordingleys et al., 1990; Dougherty et al., 1989).

Although commonly use to cleave fusion tags from recombinantly expressed proteins (Abdullah et al., 2005; Hefti et al., 2001), both thrombin and factor Xa have demonstrated poor specificity when incubated with NFκB-DOSCAT. There is evidence that both these proteases recognise sequences other than their commonly published consensus sequence (Jenny et al., 2003). It has been suggested that Factor Xa has a very loose selectivity, and as well as recognising the sequence IEGR, it may cleave at sequences as diverse as *XX*GR, *X*(G/A/S)R and *XX*(L/W/F/Y)R (Bianchini et al., 2002; He et al., 1993; Hsu et al., 2008). This would mean there are as many as 10 additional sites within DOSCAT that Factor Xa could cleave at, which would provide an explanation for the observed results. The consensus sequence of thrombin is longer than that of Factor Xa at 6 amino acids (LVPR/GS; cleavage at C-terminal of R), which would suggest a higher degree of specificity. Although there is a strong preference for Pro in P2 and Arg in P1 (Gallwitz et al., 2012; Petrassi et al., 2005), the amino acids in other positions can be less restricted (Gallwitz et al., 2012). Within the NFκB-DOSCAT sequence there are two other regions at which thrombin may also cleave. The first, KVPR/GS is in a region that forms two IκBα Q-peptides and their flanking regions. Other than substituting Lys for Leu at P4 this is identical to the consensus sequence. There is a preference for aliphatic acids at this position (Backes et al., 2000; Petrassi et al., 2005), and only negatively charged residues are known to diminish activity (Gallwitz et al., 2012). Therefore, it is highly likely that thrombin would cleave this sequence. The other potential cleavage sequence is PAPR/AG.

Although P4 is not an aliphatic residue, this is not an absolute requirement and studies have shown that Pro occupy this position whilst activity is maintained (Gallwitz et al., 2012). Ala being at the P3 and P1' position and Gly at P2' would also still permit thrombin cleavage (Gallwitz et al., 2012; Petrassi et al., 2005; Stephens et al., 1988). It can therefore be concluded that this is another viable cleavage site for thrombin. Taken together, these three sites of cleavage could result in up to 9 distinct fragments being produced, consistent with the observations by SDS-PAGE.

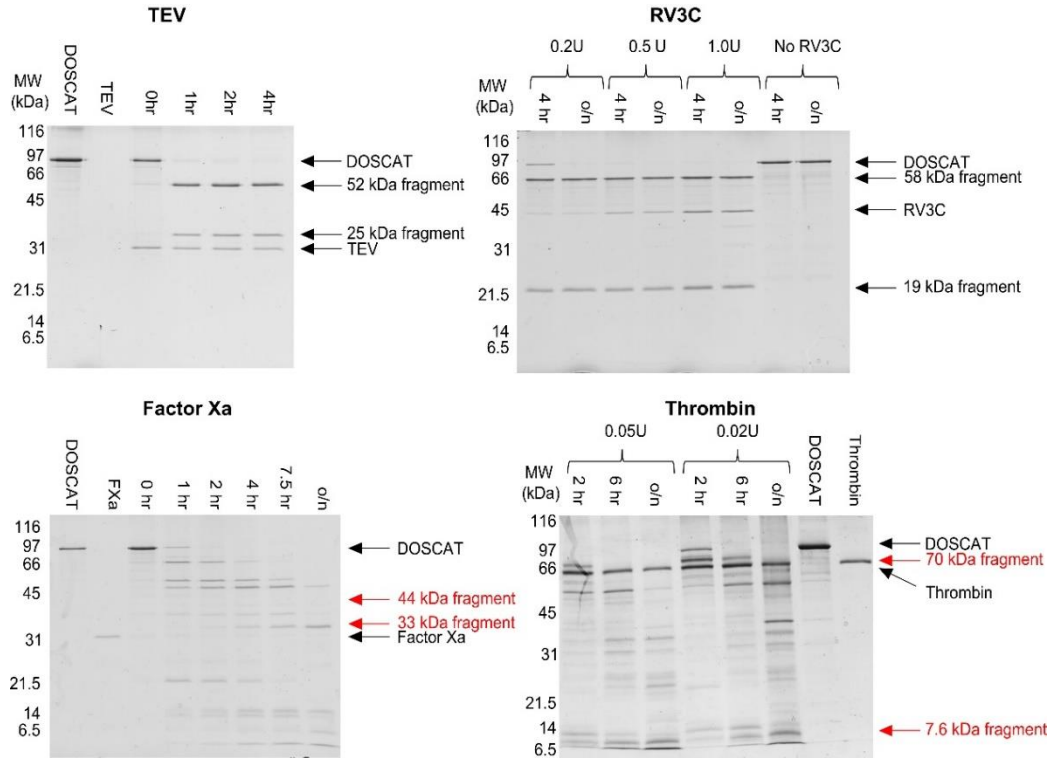


Figure 3.15 Digestion of NFκB-DOSCAT by restricted specificty proteases. NFκB-DOSCAT (10 µg) was incubated with a) 0.5 units (U) TEV protease; b) 0.2 U, 0.5 U or 1.0 U RV3C protease; c) 0.05 U, 0.02 U Thrombin; d) 0.1 U Factor Xa. Samples were taken at indicated time points and analysed on a 12% SDS-PAGE gel. Expected band sizes of proteolytic fragments are displayed in black when the fragments were observed or in red where they were not observed.

3.3. Conclusions

This work defines the first time a DOSCAT standard has been designed, expressed and purified in a recombinant system. Furthermore, the utility of DOSCAT as a dual standard has been demonstrated in that the Q-peptides and epitopes contained in the NFκB-DOSCAT can be observed by MALDI-MS and western blotting respectively. Additionally, it has been shown that compared to classic western blotting using ECL or fluorescent detection, automated capillary western blotting offers much improved characteristics for absolute protein quantification.

Despite the encouraging results, some results highlight shortcomings in the NFκB-DOSCAT design that need to addressed. Out of five epitopes inserted into the NFκB-DOSCAT sequence, only three were recognised by target antibodies. Additionally, 3 of the antibodies used have now been discontinued by the manufacturer: p65, IκBβ and IκBε. This highlights the importance of antibody selection prior to finalising a DOSCAT sequence and committing to express the protein. Ideally, antibodies that are used will have been raised using a peptide immunogen, the sequence of which has been published by the manufacturer or, the antibody binding site would be known through epitope mapping experiments. This would take away any guesswork associated with selecting epitope sequences to use in a DOSCAT. Another aspect of DOSCAT design that did not perform to expectations was the inclusion of restricted protease sites; only two out of four proteases cleaved NFκB-DOSCAT specifically at one site. Factor Xa and thrombin have a cleavage sequence that is less rigid than that of TEV and RV3C protease. Therefore, there were sites in NFκB-DOSCAT that by chance that Factor Xa and thrombin recognised, leading to many fragments being produced. These two proteases will not be used in future, and when selecting other proteases greater care will be taken to fully understand all possible cleavage sequences and not accidently include additional protease binding sites in the DOSCAT sequence that may be contained in epitopes or Q-peptides.

Due to these limitations, this NFκB-DOSCAT will not be used to in experiments to quantify the target protein by western blotting and SRM-MS. Instead, a second iteration of DOSCAT will be designed to incorporate new antibody epitopes and restricted protease sites to replace Factor Xa and thrombin. The protocols for expression and purification validated in this chapter will be used again, and, when it comes to performing quantitative experiments, automated western blotting will be used alongside SRM-MS with confidence. This will test whether QWB and SRM will generate agreeable quantitative data, the next big test for DOSCAT technology.

Chapter 4: Quantification of NF-κB proteins using DOSCAT

4.1. Introduction

The results from the initial proof of principle work (Chapter 3) were encouraging in establishing the potential of using DOSCATs in both MS and WB assays. The original objective was to use a DOSCAT to quantify target proteins in the NF-κB pathway, however, limitations in NFκB-DOSCAT design prevented this aim from being achieved. Therefore, a re-design of NFκB-DOSCAT was required to incorporate new epitopes and protease cleavage sites. Section 4.2. sets out the design, expression, purification and use of a second iteration of NFκB-DOSCAT, termed NFκB-DOSCAT-2, to quantify the five target proteins. This is described through a paper published on the matter in a peer reviewed journal (Bennett et al., 2017). Additional data that was not contained in this paper, detailing optimisation work and demonstrations of the wider utility and application of NFκB-DOSCAT-2, is presented in section 4.3.

When designing a QWB experiment, several parameters must be optimised to ensure accurate quantification. Antibody conditions must be optimised so that the antibody is at saturating conditions, meaning that any observed signal changes can only be due to changes in analyte amount. Ideally, the amount of protein loaded would be in the linear dynamic range of the assay and consideration must also be given to data analysis and normalisation (Gassmann et al., 2009; Taylor SC, 2014). Furthermore, calibrant stability is key in quantitative assays; degradation, proteolysis or loss of material through adsorption to surfaces can lead to inaccurate quantification. Degradation of proteins can occur during long-term storage, sub-optimal pH or temperature or freeze-thaw cycles (Cao et al., 2003; R. J. Simpson, 2010). Proteolysis is an issue when the calibrant is spiked into a cell lysate that has not been treated with protease inhibitors. Endogenous proteases may be active within the lysate and act upon the calibrant protein. Protein adsorption to surfaces such as plastic tubes is a well-known phenomenon, especially when proteins are at low concentrations (Rabe et al., 2011), and adsorption has been observed in the previous iteration of NFκB-DOSCAT.

Protein adsorption occurs in two stages: 1) reversible adsorption to a surface which is driven primarily by electrostatic and hydrophobic interactions, 2) conformational change leading to irreversible adsorption (Norde et al., 1995). The protein undergoes conformation change and denaturation after initial adsorption, so that more regions of the protein previously hidden by folding are exposed and able to interact with the surface (Moulin et al., 1999). Protein adsorption can be affected by the local environmental conditions (pH, temperature etc.), properties of the protein itself or by those of the surface material (e.g. plastic or glass) (Rabe et al., 2011). To prevent or reduce adsorption, the properties of the surface or local environmental conditions can be altered. Addition of detergents or high salt concentrations (J. A. Smith et al., 1978), addition of proteins such as bovine serum albumin (Felgner et al., 1976; K. J. Kramer et al., 1976), or coating tubes with polyethylene glycol (PEG) or siliconizing reagents have all been described to prevent protein interactions with surfaces. Different

surface materials and additives will be trialled to ascertain which if any can prevent DOSCAT adsorption. The intrinsic stability of DOSCAT and its stability when spiked into cell lysate will also be assessed. In addition, the optimal values for QWB parameters will be experimentally determined.

As well as this assay optimisation, the validity and flexibility of DOSCAT technology outside of internally calibrated automated capillary western blotting must be demonstrated. Although automated capillary western blotting can offer superior reproducibility and quantitative accuracy, many researchers do not have access to this technology and instead use classic WB with ECL or fluorescent detection. It is therefore important to demonstrate that DOSCAT can be used in this format, and that quantitative values generated can be favourable compared against those derived by automated QWB and SRM-MS. The calibration strategy used in quantitative assays is important; internal calibration (when standard is spiked into the sample) is favoured as it accounts for matrix effect, losses that may occur in sample preparation and drift in instrument performance (Oliveira et al., 2010; Sargent, 2013). However, in some scenarios internal calibration will not be possible and an external calibration method will be used. Therefore, NFκB-DOSCAT-2 will be used in a classic western blot using fluorescent detection (due to improved signal linearity over ECL) and its performance as a calibrator when used external to the analyte assessed.

4.2. DOSCATs: Double standards for protein quantification (Bennett et al. 2017, Scientific Reports)

The text and figures in section 4.2 are a direct replication of the paper “DOSCATs: Double standards for protein quantification” published in Scientific Reports, 2017 (Bennett et al., 2017). The manuscript and figures for the paper were prepared by myself and I am the first author of the paper. A reprint of the paper can be found in the Appendix. Note that the DOSCAT referenced in the manuscript refers to the NFκB-DOSCAT-2 iteration.

4.2.1. Abstract

The two most common techniques for absolute protein quantification are based on either mass spectrometry (MS) or on immunochemical techniques, such as western blotting (WB). Western blotting is most often used for protein identification or relative quantification, but can also be deployed for absolute quantification if appropriate calibration standards are used. MS based techniques offer superior data quality and reproducibility, but WB offers greater sensitivity and accessibility to most researchers. It would be advantageous to apply both techniques for orthogonal quantification, but workflows rarely overlap. We describe DOSCATs (**DO**uble **S**tandard con**CAT**amers), novel calibration standards based on QconCAT technology, to unite these platforms. DOSCATs combine a series of epitope sequences concatenated with tryptic peptides in a single artificial protein to create internal tryptic peptide standards for MS as well as an intact protein bearing multiple linear epitopes. A DOSCAT protein was designed and constructed to quantify five proteins of the NF-κB pathway. For three target proteins, protein fold change and absolute copy per cell values measured by MS and WB were in excellent agreement. This demonstrates that DOSCATs can be used as multiplexed, dual purpose standards, readily deployed in a single workflow, supporting seamless quantitative transition from MS to WB.

4.2.2. Introduction

Accurate quantification of proteins is of critical importance in cell biology, proteomics, clinical biomarker discovery and systems biology. Two very different approaches to quantification are routinely adopted; those based on mass spectrometry (MS) and those based on (semi)quantitative western blotting (sqWB). The two methods differ, both in the technical demands and in the complexity of the associated equipment, as well as the confidence in quantitative data generated.

Mass spectrometric methods are considered to be the gold standard for targeted protein quantification (Aebersold et al., 2013; Lawless et al., 2016; Ong et al., 2005). However, capital investment and the expertise required in setting up and executing an MS assay means that it is less widely used than sqWB. For relative MS quantification, there is increasing application of label-free quantification based on the intrinsic signal intensity of individual peptides (derived

from a digested protein) or of label-mediated quantification in which stable isotope labels are used to discriminate between two or more conditions, discriminated by the mass shift either at the level of the peptide ion or at the level of fragment ions generated within the mass spectrometer. Label based quantification methods are commonly used in conjunction with a targeted MS approach known as selected reaction monitoring (SRM). SRM utilises triple quadrupole mass spectrometers to perform two levels of mass selection, at the level of both precursor and product ion, giving much improved selectivity and sensitivity over global, 'discovery' proteomic approaches. Semi-quantitative western blotting is, by contrast, readily delivered with a small investment in equipment, and in most laboratories, requires extended sequences of manual processing steps (although there are instrumentation developments that automate the method). Although considered a semi-quantitative technique for relative quantification of signal intensity, sqWB is commonly used to draw quantitative conclusions despite the lack of calibration standards, rigorous (and standardised) methodology, and consistent data analysis (Gassmann et al., 2009; Taylor et al., 2013). However, direct comparison of sqWB results between groups is problematic as the data (effectively, the intensity of an antibody–reactive band that is generated by different chemistries and measured using different imaging devices) are dimensionless and highly variable (inter-assay) despite high levels of care and skill by the researcher. This limitation is likely to have contributed to the lack of reproducibility in pre-clinical data, which has a high cost in terms of wasted effort and delayed progress (Freedman et al., 2015; Mobley et al., 2013). Many papers that report sqWB data do not include exhaustive data that defines the specificity of the antibody-antigen interaction, linearity of response or evidence that the immunoreactive band is the target antigen. Indeed, it is common practice in publication of sqWB results to crop western blot images to the region of interest, thus obscuring other regions of cross-reactivity. In sqWB, quantification is usually relative, where one condition is compared with a second, ideally run on the same gel and developed as a single blot.

For absolute quantification, calibration standards based on stable isotope labelled proteins or peptides (for MS) or epitope bearing proteins (for WB) are required. Isotope standards for MS, based on relatively short tryptic peptides, are not suitable for western blot quantification, such that MS-based and WB workflows rarely overlap. Ideally, there would be readily deployable techniques to converge technologies, raising standards in quantitative output. There is a continuing need for appropriate calibration standards in the western blot workflow, thereby creating genuinely quantitative western blotting (QWB). Further, it would be ideal if calibration standards were capable of deployment across both MS and QWB workflows. This crossover would allow for validation of the orthogonal techniques and comparison of data between the two most common quantitative techniques. Moreover, QWB could be used to improve characterisation of research antibodies (sensitivity, dynamic range, limit of detection, consistency of batches, specificity), which is currently problematic (Bordeaux et al., 2010; Michel et al., 2009) and a major factor in the irreproducibility of pre-clinical science (Taylor et

al., 2013). Indeed, the use of orthogonal techniques such as MS to validate antibody specificity has been proposed by an international working group of scientists (Uhlen et al., 2016).

It is easy to understand why western blotting is the preferred method (Pubmed searches reveal approximately 5,000 citations using the terms 'SRM or MRM', compared to over 250,000 using the terms 'western blot' or 'western blotting'), as it is readily deployable in most laboratories and does not require access to either the specialist instrumentation or the cognate expertise that is needed to develop an MS-based method. However, there is considerable scope for studies that compare and explore the comparative performance of the two approaches. Indeed, it is a common experience with MS-based proteomics studies that reviewers request 'validation', implicitly meaning western blotting, without any evidence for the ability of both methods to deliver comparable data sets (Aebersold et al., 2013). In pursuit of reliable approaches to quantitative western blotting, there is merit in comparison of such approaches with an orthogonal method, such as those based on mass spectrometry. Further, QWB methods give added information about electrophoretic mobility (crudely indicative of protein mass), information that is absent from 'bottom-up' proteomics.

We have previously designed and employed artificial proteins to solve problems in MS-based analyses, including QconCATs (Beynon et al., 2005) for absolute quantification and QCAL (Eyers et al., 2008), QCAL-IM (Chawner et al., 2012) and RePLiCal (Stephen W Holman et al., 2016) as universal standards for MS-based techniques. In particular, we developed QconCAT proteins to facilitate multiplexed protein quantification by MS (Beynon et al., 2005; Julie M Pratt et al., 2006). QconCATs are artificial genes encoding proteins that are concatenations of (usually tryptic) peptides from multiple proteins (typically up to 25) that act as 'quantotypic' standards when the calibrator protein is expressed in bacteria and labelled with stable isotope amino acids (P. Brownridge, Holman, et al., 2011). After co-digestion of the analyte (in a biological sample) and the QconCAT, the resultant pairs of unlabelled (analyte) and labelled (standard) peptides can be analysed by mass spectrometry, permitting accurate quantification of the analyte protein abundance.

Here, we extend QconCAT principles to design DOSCATs ('**DO**uble **S**tandard con**CAT**amers', dual-purpose calibration standards for MS-based quantification and/or QWB), also designed *de novo* and expressed heterologously in *Escherichia coli*. DOSCATs concatenate epitope sequences from one or more proteins recognised by multiple antibodies. They also embed quantotypic peptides (Q-peptides) for MS quantification of the same proteins (Figure 4.1) and thus act as a single multiplexed standard that can be used for MS-based or QWB quantification. For added flexibility, the epitopes can be interspersed with restricted specificity endoproteolytic sequences, which permit generation of quantification standards of optimal mobility in sized-based or charge-based separation platforms. The DOSCAT design includes a His-tag for purification and [Glu1]-Fibrinopeptide B (Glu-Fib) sequence for quantification of the standard (Simpson and Beynon 2012) and contain at least two Q-peptides for each target

protein, chosen according to well-defined criteria (Brownridge et al. 2011; Holman, Sims, and Eyers 2012). It is axiomatic that the epitopes can only be used if they are relatively short peptide sequences, either used as an immunogen or identified as the linear sequence that is recognised by a monoclonal antibody. The aim of this work is to demonstrate that DOSCATs can be used as a calibration standard across both SRM and QWB workflows to deliver equivalent quantitative results.

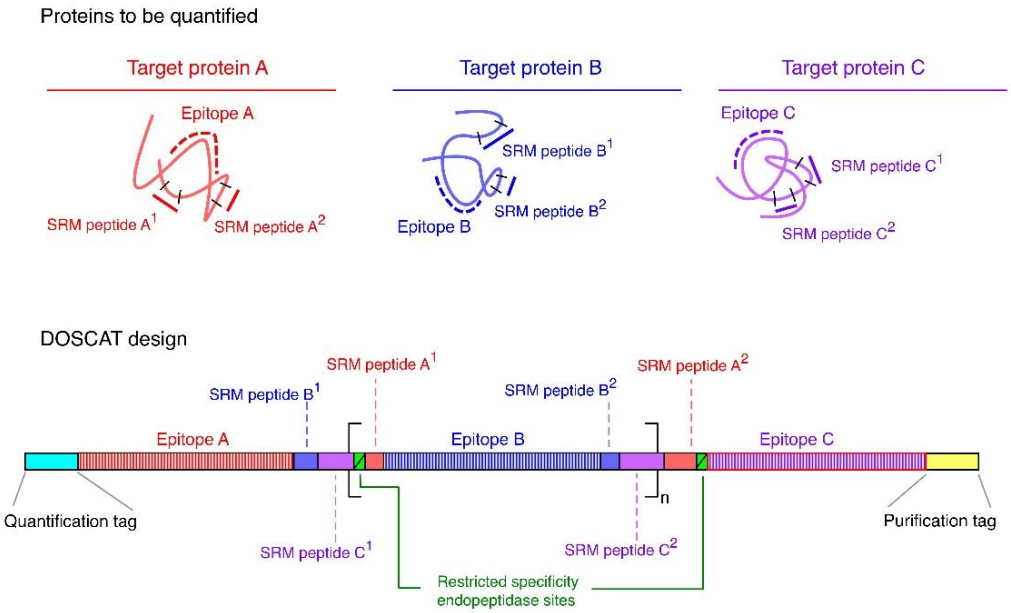


Figure 4.1 Principle of DOSCAT design. Peptides for MS-based quantification and antibody epitopes are selected for each target protein and sequences concatenated into a single artificial protein (DOSCAT). Tag sequences for quantification and purification of DOSCAT are inserted at each terminus. Restricted specificity endopeptidase sites are interspersed throughout the sequence so as to permit generation of quantification standards of optimal mobility during sized-based separation analysis.

4.2.3. Results and Discussion

DOSCAT design, expression and purification

For DOSCATs, epitopes for a small panel of antibodies are inserted into the protein in addition to Q-peptides so that the standard can be deployed across both MS and QWB workflows. A minimum of two Q-peptides were chosen for each target protein based on a well-defined process (P. Brownridge, Holman, et al., 2011; Steven W Holman et al., 2012). The data repositories PeptideAtlas (http://www.peptideatlas.org/) and Global Proteome Machine (http://www.thegpm.org/) were initially consulted and peptides identified. Where peptides could not be identified using this approach (for IκBα and RelB), peptides were selected on the basis of a computational prediction of their quantotypic propensity based on physio-chemical properties and predicted observability (PeptideSieve score (http://tools.proteomecenter.org/wiki/index.php?title=Software:PeptideSieve)). For all selected Q-peptides uniqueness in the proteome was confirmed by BLAST searches. Each peptide was supplemented with natural flanking sequences of 3 amino acids in length as a strategy to improve quantitative accuracy by equalising digestion efficiency between standard and analyte (Kito et al., 2007) (Table 4.1). For p65 and IKBβ, selected Q-peptides were adjacent in the protein sequence so flanking regions were not possible or required. Epitopes included in the DOSCAT sequence were chosen based on the immunogen sequence supplied by the manufacturer (Table 4.2). Where the precise immunogen sequence was disclosed by the manufacturer, this was used in the DOSCAT - this was the case for IκBβ and IκBε. For some antibodies, only the specific residue around which the epitope was centred was disclosed and in these instances, up to 25 amino acids flanking each side of the central residue in the protein sequence were used to ensure inclusion of the epitope in the DOSCAT sequence. This was required for the epitopes for p65, RelB and IκBα.

To introduce the possibility of manipulation of mobility in QWB, several restricted specificity endopeptidase sites were included in the DOSCAT design. These sites were selected based on the primary sequence specificity and the commercial availability of the protease. Four endopeptidases were used (Table 4.3) and for each, the target sequence was unique to a single site in the DOSCAT and was absent from any of the target proteins. Cleavage sites were inserted into the sequence between epitopes so that proteolysis would generate two fragments that could be detected by different antibodies and which would have a different electrophoretic mobility to the parent DOSCAT, permitting a mobility shift to prevent overlap between standard and analyte. Additional amino acids were added at the N-terminus to provide an initiator methionine and a sacrificial N-terminal region as well as [Glu1]-Fibrinopeptide B peptide (EGVNDNEEGFFSAR) for MS-based quantification of the standard[3]. A hexahistidine tag was added at the C-terminus to allow affinity purification. Once the protein sequence was designed (Figure 4.2a) and the codon optimised gene was synthesised, the DOSCAT was expressed successfully in *E.coli* in minimal medium (Figure 4.2b), accumulating

in inclusion bodies, and after solubilisation in a chaotrope was readily purified by Ni-NTA affinity chromatography (Figure 4.2c).

Table 4.1 SRM transitions, collision energies and dwell times for the analysis of each target peptide. Natural flanking sequences of peptides are shown in red.

Protein	Peptide sequence	Precursor ion (m/z)	Measured product ions (m/z)	Charge	Collision energy (eV)	Dwell time (msec)
p65	EGRSAGSIPGERSTD	437.2	y7 (715.4), y6 (658.4), y4 (458.2)	2	13	38
	SQRIQTNNNPFQVPIEEQR	964.0	y10 (1242.6), y7 (870.5), y6 (771.4)	2	37	30
	GDYDLNAVRLCF	511.7	y7 (850.4), y6 (687.4), y5 (572.4)	2	16	38
	DCRDGFYEAELCPDRCIH	736.3	y9 (1152.5), y8 (989.4), y7 (860.4)	2	27	30
RelB	EGRSAGSILGESSTEASKTLP	712.3	y10 (1008.5), y9 (895.34), y7 (709.3)	2	26	30
	MLRSGPASGPSVPTGRAMV	585.3	y9 (857.4), y8 (770.4), y7 (713.4)	2	20	30
	LQRLTDGVCSEPLPFTYLPRDHD	983.0	y7 (893.5), y9 (1103.6), y12 (1479.7)	2	38	163
IκBα	EEKALTMEVIRQVK	466.8	y6 (748.4), y5 (647.4), y4 (516.3)	2	25	30
	VPRGSEPWKQQL		**PEPTIDE NOT DETECTED**			
IκBβ	DAGADLDKPEPTCGR		**PEPTIDE NOT DETECTED**			
	SPLHLAVEAQAADVLELLLRAGA	720.1	**PEPTIDE NOT DETECTED**			
IκBε	QDRHGDTALHVACQRQHL	455.6	y6 (770.4), y5 (633.3), y4 (534.2)	3	9	38
	SGKTALHLAVETQERQLV	684.4	y8 (945.5), y7 (832.4), y6 (761.4)	2	24	163

Table 4.2 Antibodies used in western blotting and the epitopes built into DOSCAT sequence.

Protein	Manufacturer	ID	Lot	Clonality	Species	Antibody dilution	Immunogen	Epitope in DOSCAT
p65	Cell Signalling Technology	#8242S	4	Monoclonal	Rabbit	1:1000	Centred around Glu 498 in human p65 (Q04206)	**L478** NQGIPVAPHTTEPMLMEYP EITRLVTGAQRPDPAPAP **L518**
RelB	Cell Signalling Technology	#4954	5	Polyclonal	Rabbit	1:500	Centred around Ser424 in human RelB (Q01201)	**D401** HDSYGVDKKRKRGMPDVL GELNSSDPHGIESKRRKKK PAILD **H444**
IκBα	Cell Signalling Technology	#9242	10	Polyclonal	Rabbit	1:100	Centred around Arg29 in human IκBα (P25963)	**M1** FQAAERPQEWAMEGPRDG LKKERLLDDRHDSGLDSMK DEEYEQMVKELQEI **R53**
IκBβ	Bethyl Laboratories Inc.	A301-828A	1	Polyclonal	Rabbit	1:100	Between residue 306 and 356 in human IκBβ (Q15653)	**E306** KSGPCSSSSDSDSGDEGD EYDDIVVHSSRSQTRLPPT PASKPLPDDPRP **V356**
IκBε	Abcam	ab51147	GR974 07-5	Polyclonal	Rabbit	1:50	IESLR synthetic peptide – corresponds to residues 158 to 163 in human IκBε (O00221)	**Q154** YDSGIESLRSLRS **L168**

Table 4.3 Restricted specificity endoproteases used and their cleavage sequence inserted into DOSCAT.

Protease	Manufacturer	Cleavage sequence
Tobacco etch virus (TEV)	Invitrogen, USA	ENLYFQG
Human rhinovirus 3C (RV3C)	GE Healthcare, UK	LEVLFQGP
Tobacco vein mottling virus (TVMV)	Biomol, Germany	TVRFQS
Enteropeptidase	Novagen, UK	DDDDK

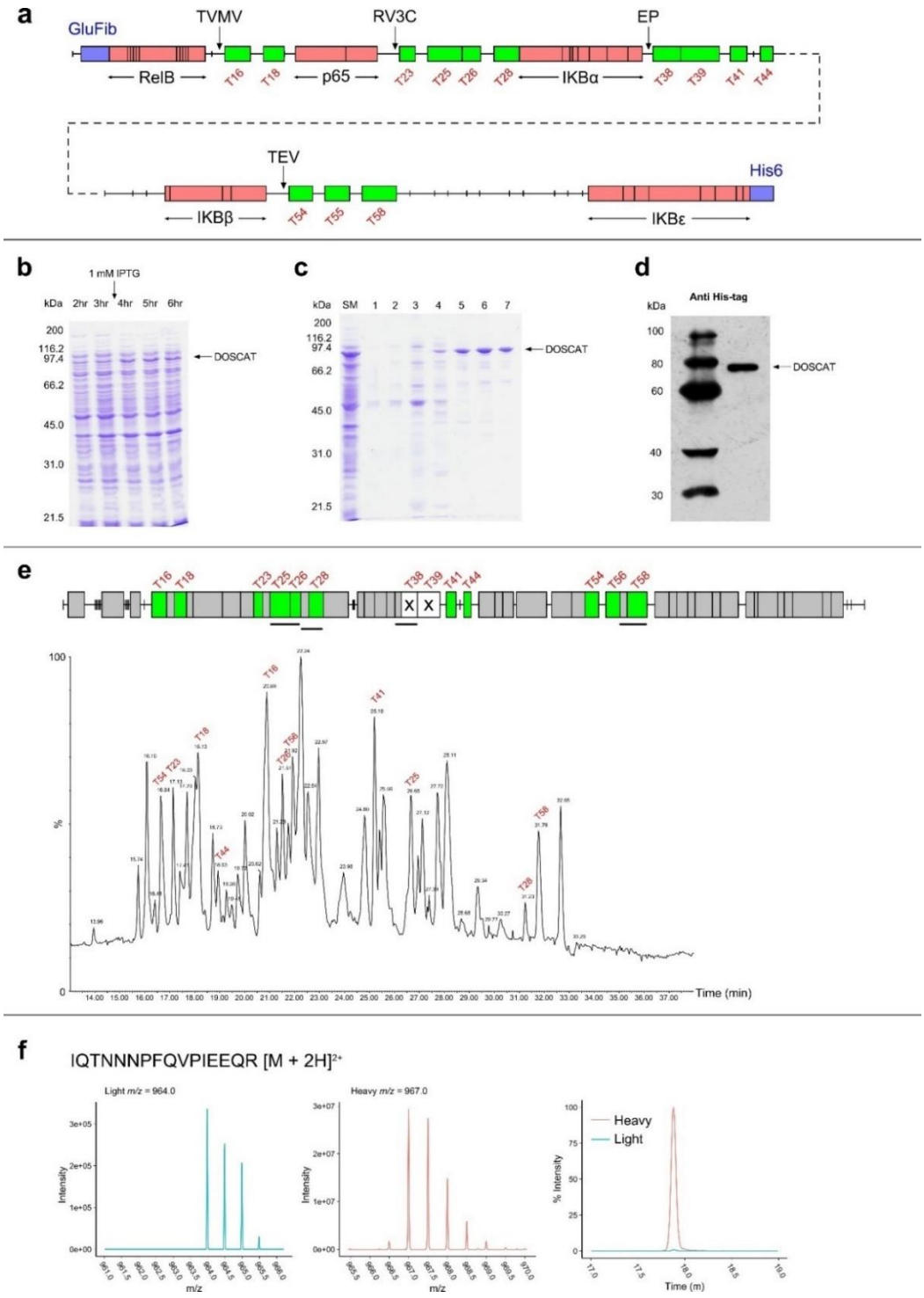


Figure 4.2 Design of a DOSCAT for quantification of members of the NFκB pathway. (a) Protein map of the NF-κB DOSCAT. Green boxes represent quantotypic peptides, red boxes define the extent of antibody binding epitopes and purple boxes the Glu1]-Fibrinopeptide B calibration peptide at the N-terminus and the His6 purification tag at the C-terminus. Arrows indicate the location of cleavage sites for each of the specific proteases (TVMV, tobacco vein mottling virus protease; RV3C, human rhinovirus 3 C protease; EP, enteropeptidase; TEV, tobacco etch virus protease. SDS-PAGE analysis of (b) E. coli culture time points 2–6 h after inoculation, with expression induced by IPTG after 3 h and (c) pre-purification starting material (SM) alongside elution fractions 1–7 from His-Trap column using an elution gradient 0–100% elution buffer over 20 min. (d) Western blot analysis of 50 ng purified DOSCAT using an anti His-tag antibody. (e) DOSCAT peptide map highlighting Q-peptides identified (green) and not identified (white with cross) by MS/MS, alongside the MS1 total ion chromatogram signifying the elution profile of each Q-peptide. (f) Mass spectra and SRM chromatogram for a representative Q-peptide demonstrating high stable isotope labelling efficiency.

Validation and performance in QWB and SRM

After purification, the DOSCAT was homogenous on SDS-PAGE and western blotting using an anti His-tag antibody (Figure 4.2d) confirmed that there was no proteolysis of the protein during expression or purification. Moreover, western blotting using p65 and RelB antibodies (the epitopes to which are at the N-terminal end), demonstrated there was no degradation or fraying at the N-terminus of the protein (Figure 4.3). After expression and purification, the DOSCAT was subject to tryptic digestion and the peptides were analysed by LC-MS/MS on a Synapt G2. This confirmed that the protein sequence was correct with 83% sequence coverage and 11 of the 13 nominated Q-peptides being identified by LC-MS/MS (Figure 4.2e). The peptide DAGADLDKPEPTCGR was not identified due to a miscleavage event at the N terminus. Only the miscleaved peptide, containing the preceding residues DDDDK, the cleavage site for enteropeptidase, was detected. It is known that acidic residues around the scissile bond can inhibit proteolysis (Lawless et al., 2012; Siepen et al., 2007); therefore, it is likely that the sequence context around the peptide was the reason for the miscleavage. Care should be taken in placing Q-peptides next to such residue sequences when designing future iterations of DOSCATs and this observation casts doubt on the utility of enteropeptidase as a restricted specificity proteinase in this application. Additionally, the peptide SPLHLAVEAQAADVLELLLR was not detected either by database searching or in the raw data. This was further investigated by running the same DOSCAT digest on a QExactive Orbitrap instrument. The peptide was detected by database searching and in the raw data, but with a signal intensity ~0.1 % of the base peak intensity. This lack of detectability on two separate instrument platforms confirm poor ionisation efficiency or fragmentation of the peptide after ionisation. When the DOSCAT was expressed in minimal media containing [$^{13}C_6$]Lys and [$^{13}C_6$]Arg as the sole source of these amino acids, complete labelling (> 99%) of the protein was confirmed by the examination of MS1 data and extracted ion chromatograms (Figure 4.2f).

For a DOSCAT to be used for the accurate quantification of target proteins, complete and equivalent digestion of standard and analyte is crucial. Although it is possible that peptides may be released more quickly from either standard or analyte, the rate of digestion in both should reach a plateau before analysis (P. Brownridge & Beynon, 2011). To ascertain whether there was complete release of Q-peptides from the DOSCAT, the standard was spiked into SK-NA-S cell lysate and the digestion mixture was sampled at regular intervals. Rates of excision of standard and analyte peptides varied, but in all instances the proteolysis attained a stable plateau before the end of the overnight incubation period (Figure 4.4) Differences in the rate of digestion between standard and analyte are thus rendered moot by the simple expedient of establishing conditions that allow the reactions to proceed to completion.

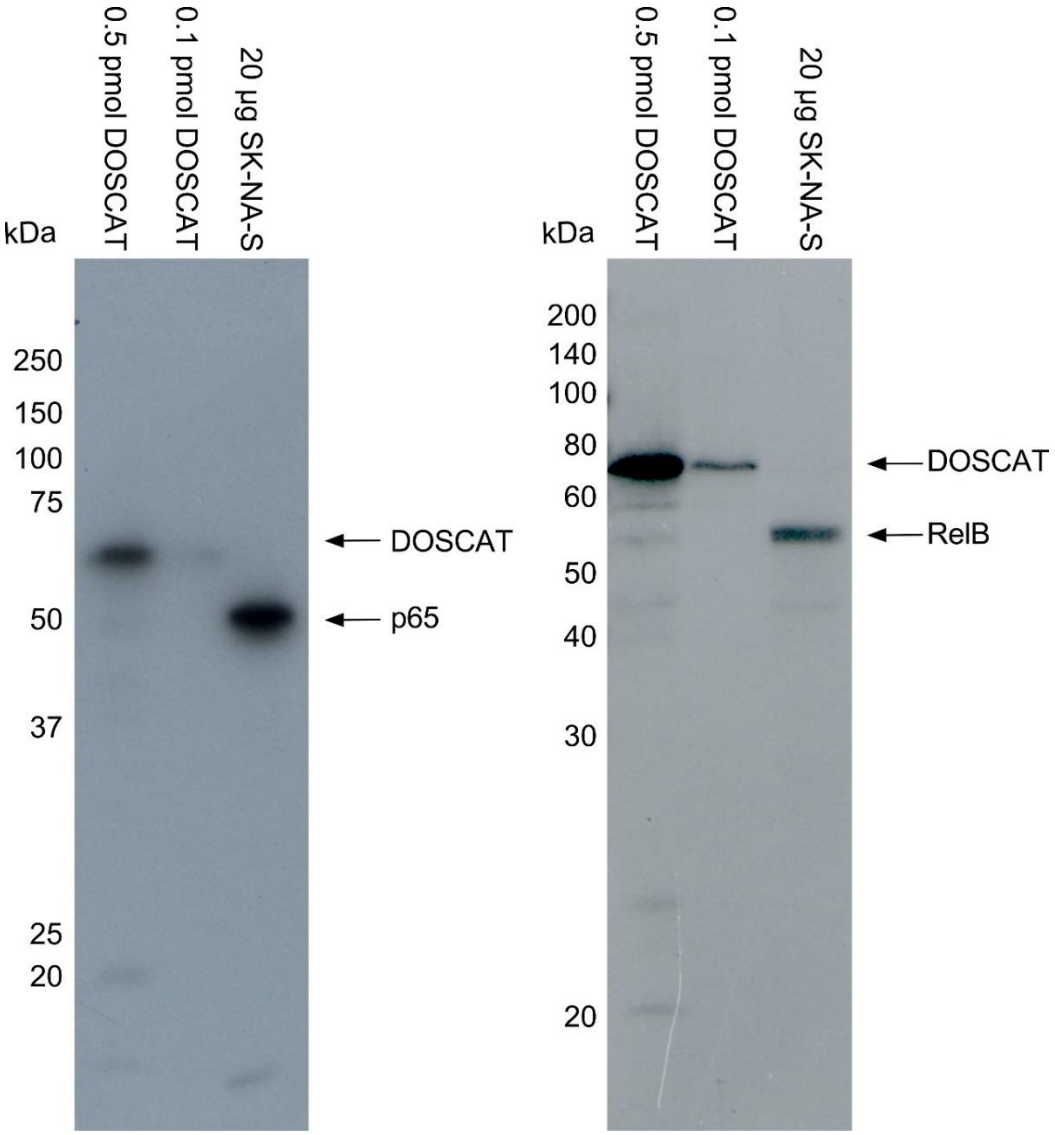


Figure 4.3 Classic gel-based western blot of DOSCAT and SK-NA-S cell lysate. Samples were separated on a 12% SDS-PAGE gel, transferred onto nitrocellulose membrane and incubated with anti p65 and RelB antibodies.

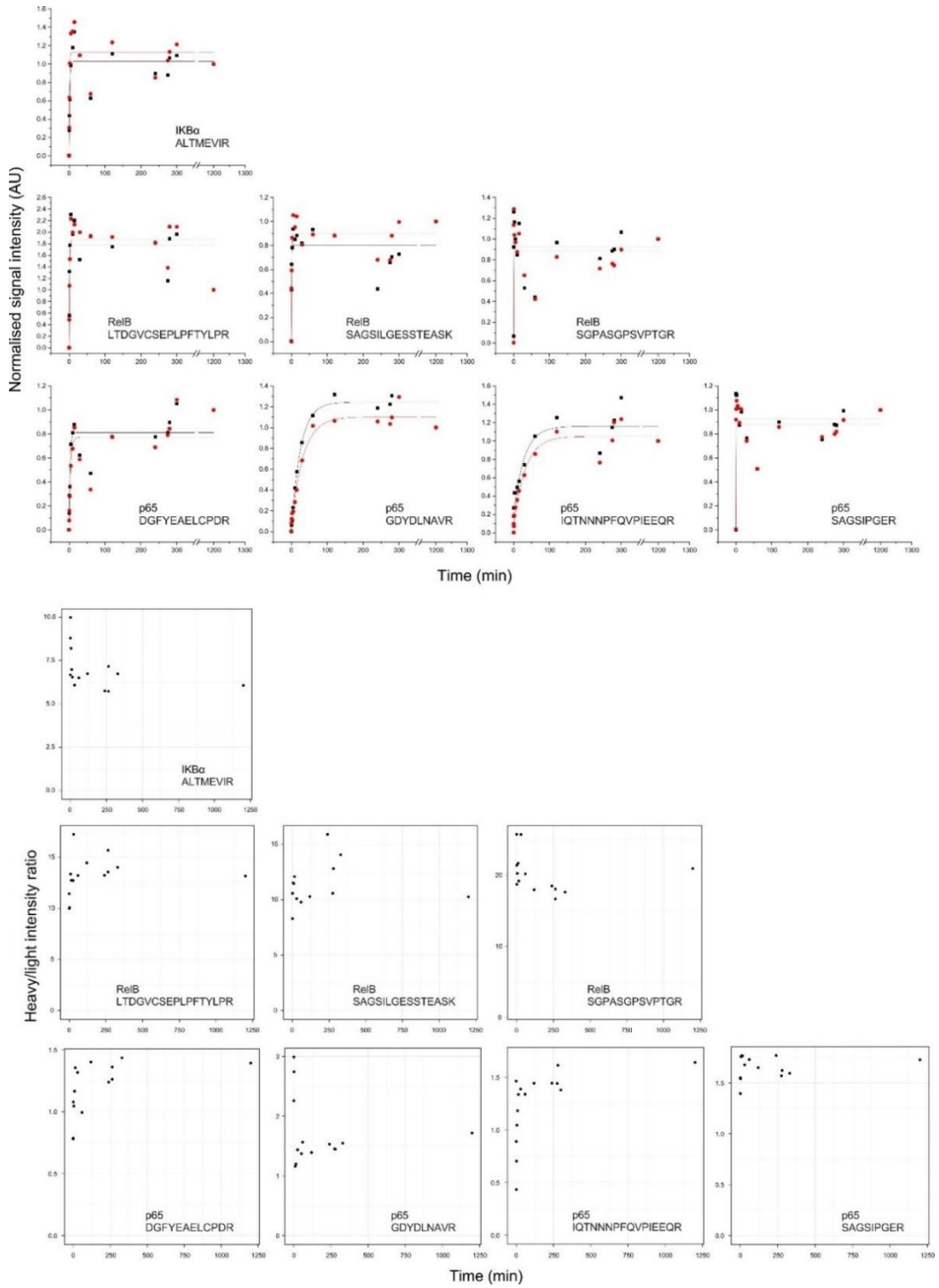


Figure 4.4 DOSCAT digestion. Time course analysis showing the release of each peptide used for quantification during tryptic digestion. Stable isotope labelled DOSCAT and SK-N-AS lysate was co-digested, with samples taken at a series of time points and analysed by SRM-MS. Displayed are signal intensities normalised to the final time point for heavy (DOSCAT) and light (endogenous) peptides (top), and the ratio of heavy to light peptide at each time point (bottom).

To define the transition coordinates of the SRM assays, a tryptic digest of the DOSCAT was analysed by LC-MS/MS using a Q-TOF mass spectrometer. For each peptide, MS/MS fragmentation data were used to identify optimal transitions for each peptide (Figure 4.5). These transitions were then used to build a scheduled SRM profile and a programme of timed transitions resulted in cleanly isolated peaks specific for all but two of the Q-peptides (Figure 4.6). To determine performance in a complex matrix, DOSCAT was spiked in SK-NA-S lysate at concentrations over a 1000-fold range and analysed by SRM. The signal linearity varied for each peptide, with some peptides exhibiting a linear relationship of the entire measured range. Moreover, the limits of detection (LOD) (based on a signal to noise ratio of 3) also differed for each peptide (Figure 4.7) with an average LOD of 100 amol. Based on the number of cells loaded onto the column, this corresponds to about 10,000 copies per cell.

To validate the DOSCAT for western blotting, dilution series of the standard protein were loaded onto the capillary western blotting system and were detected independently by each of the five commercial antibodies (Figure 4.8). Sensitivities varied for different antibodies, ranging from 5 amol to 1.5 fmol in each lane, defined as the lower limit of quantification. Based on the number of cell equivalents that were loaded into the capillaries, this equates to about 1000 copies per cell for the lowest limit of detection. IκBε could not be quantified using this methodology as the antibody did not detect the epitope in DOSCAT. However, when used against SK-NA-S cell lysate in both capillary and classic western blotting, the antibody detected a band at ~90 kDa (data not shown), very different from the expected molecular weight (53 kDa). We conclude that the antibody was unreliable and it was not used in further experiments.

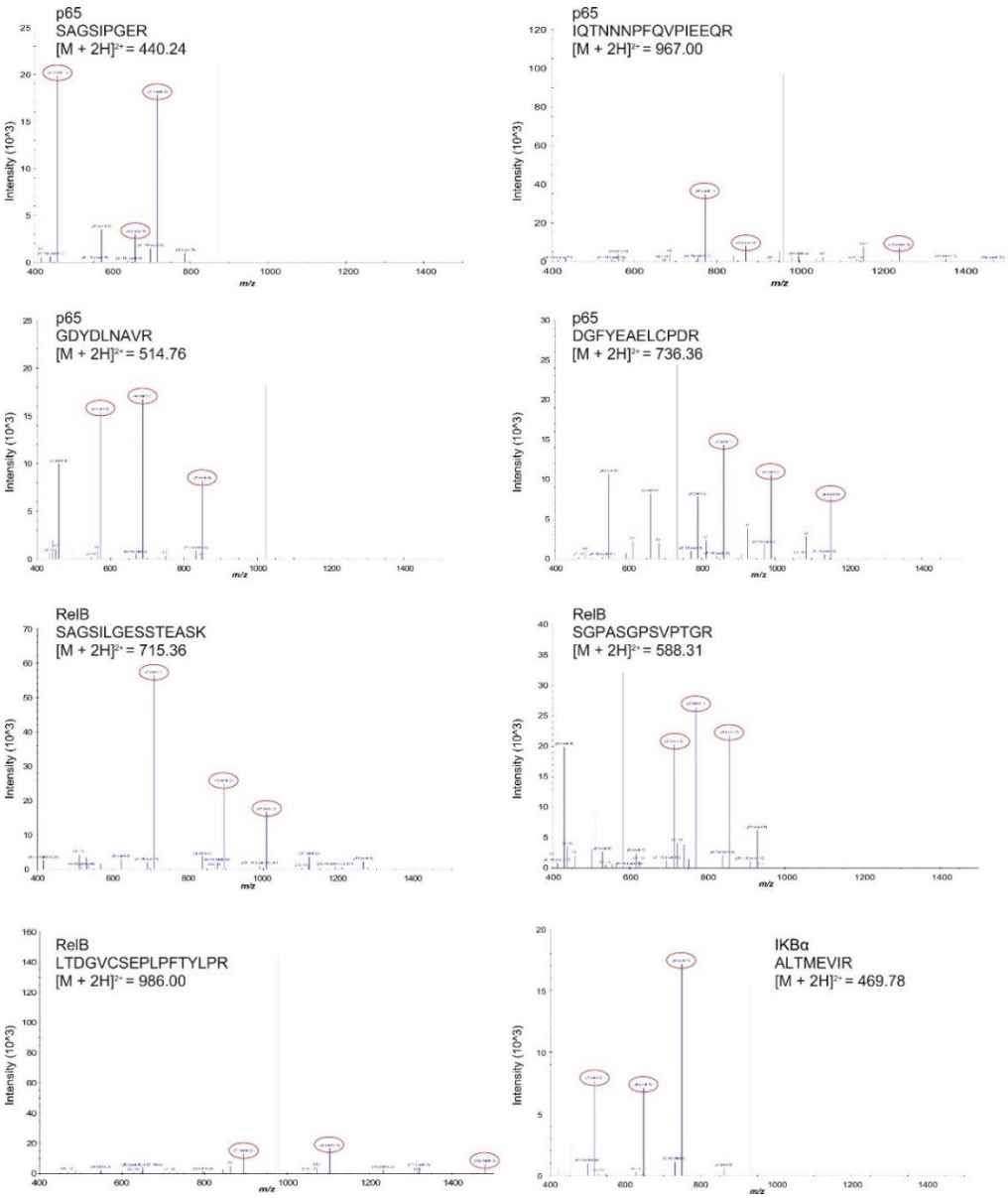
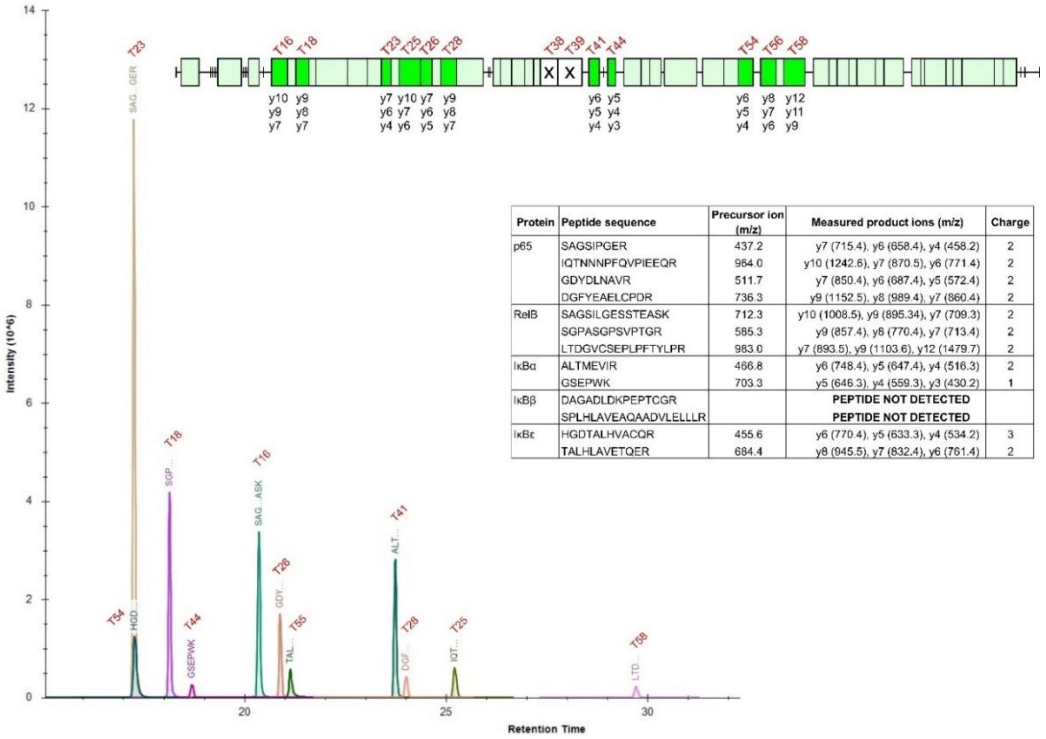


Figure 4.5 MS/MS spectra of DOSCAT Q-peptides. Tryptic digest of DOSCAT was analysed by MS^E by a Q-TOF mass spectrometer. Observed product ions chosen for analysis by SRM-MS are highlighted in red.



Protein	Peptide sequence	Precursor ion (m/z)	Measured product ions (m/z)	Charge
p65	SAGSIPGER	437.2	y7 (715.4), y6 (658.4), y4 (458.2)	2
	IQTNNNPFQVPIEEQR	964.0	y10 (1242.6), y7 (870.5), y6 (771.4)	2
	GDYDLNAVR	511.7	y7 (850.4), y6 (687.4), y5 (572.4)	2
	DGFYEAELCPDR	736.3	y9 (1152.5), y8 (989.4), y7 (860.4)	2
RelB	SAGSILGESSTEASK	712.3	y10 (1008.5), y9 (895.34), y7 (709.3)	2
	SGPASGPSVPTGR	585.3	y9 (857.4), y8 (770.4), y7 (713.4)	2
	LTDGVCSEPLPFTYLPR	983.0	y7 (893.5), y9 (1103.6), y12 (1479.7)	2
IκBα	ALTMEVIR	466.8	y6 (748.4), y5 (647.4), y4 (516.3)	2
	GSEPWK	703.3	y5 (646.3), y4 (559.3), y3 (430.2)	1
IκBβ	DAGADLDKPEPTCGR		**PEPTIDE NOT DETECTED**	
	SPLHLAVEAQAADVLELLLR		**PEPTIDE NOT DETECTED**	
IκBε	HGDTALHVACQR	455.6	y6 (770.4), y5 (633.3), y4 (534.2)	3
	TALHLAVETQER	684.4	y8 (945.5), y7 (832.4), y6 (761.4)	2

Figure 4.6 Summary of scheduled SRM-MS assays for Q-peptides contained in DOSCAT. Extracted ion chromatogram from SRM-MS analysis of 1 fmol digested DOSCAT loaded onto the column, detailing the peak intensity and retention time for all Q-peptides for which scheduled SRM-MS assays were built. Inset: DOSCAT peptide map illustrates all Q-peptides for which SRM-MS assays were designed (bright green boxes) out of all observed peptides (pale green boxes) alongside product ions monitored for each peptide. Two Q-peptides were not detectable by SRM-MS (white box with cross). Inset table contains a detailed list of transitions for each peptide.

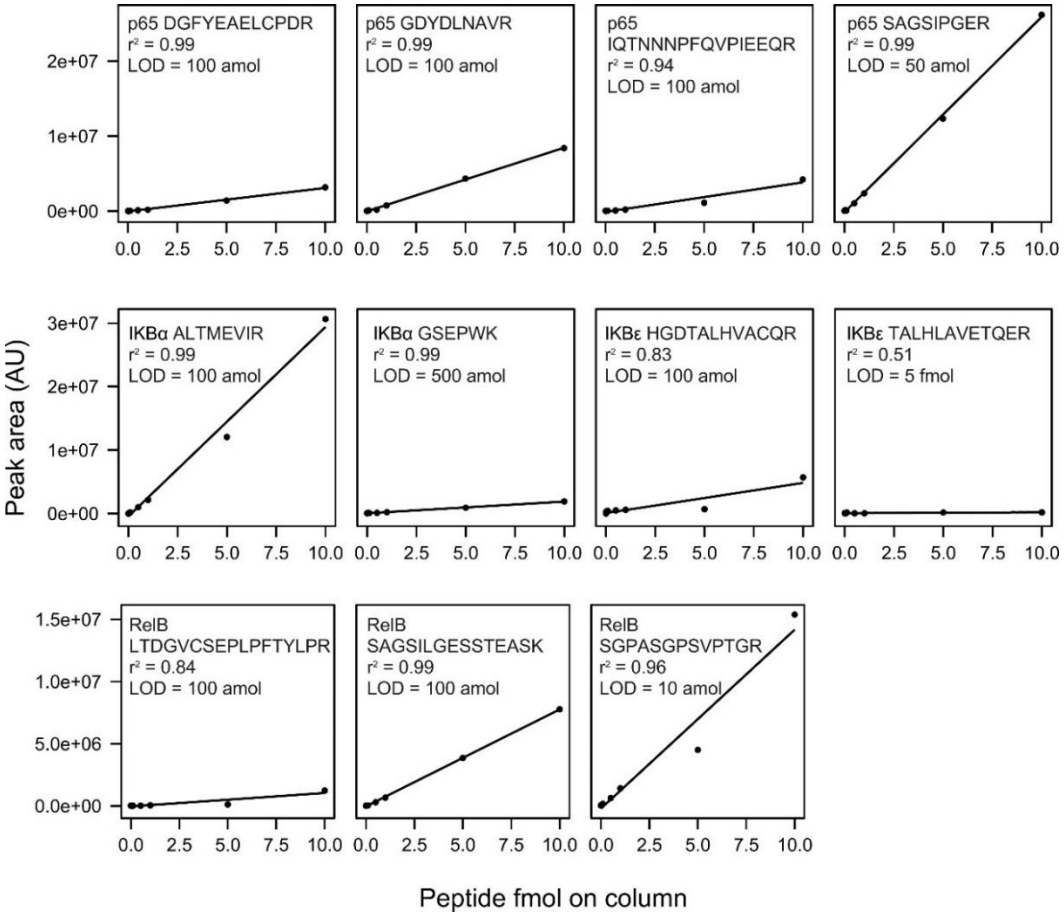


Figure 4.7 Standard curves for Q-peptides in SRM-MS assays. DOSCAT and SK-NA-S cell lysate was co-digested, serially diluted in SK-NA-S cell lysate digest and analysed by scheduled SRM-MS. Also displayed are measures of limit of detection (where signal:noise = 2) and linearity of response from 10 amol to 10 fmol material on column.

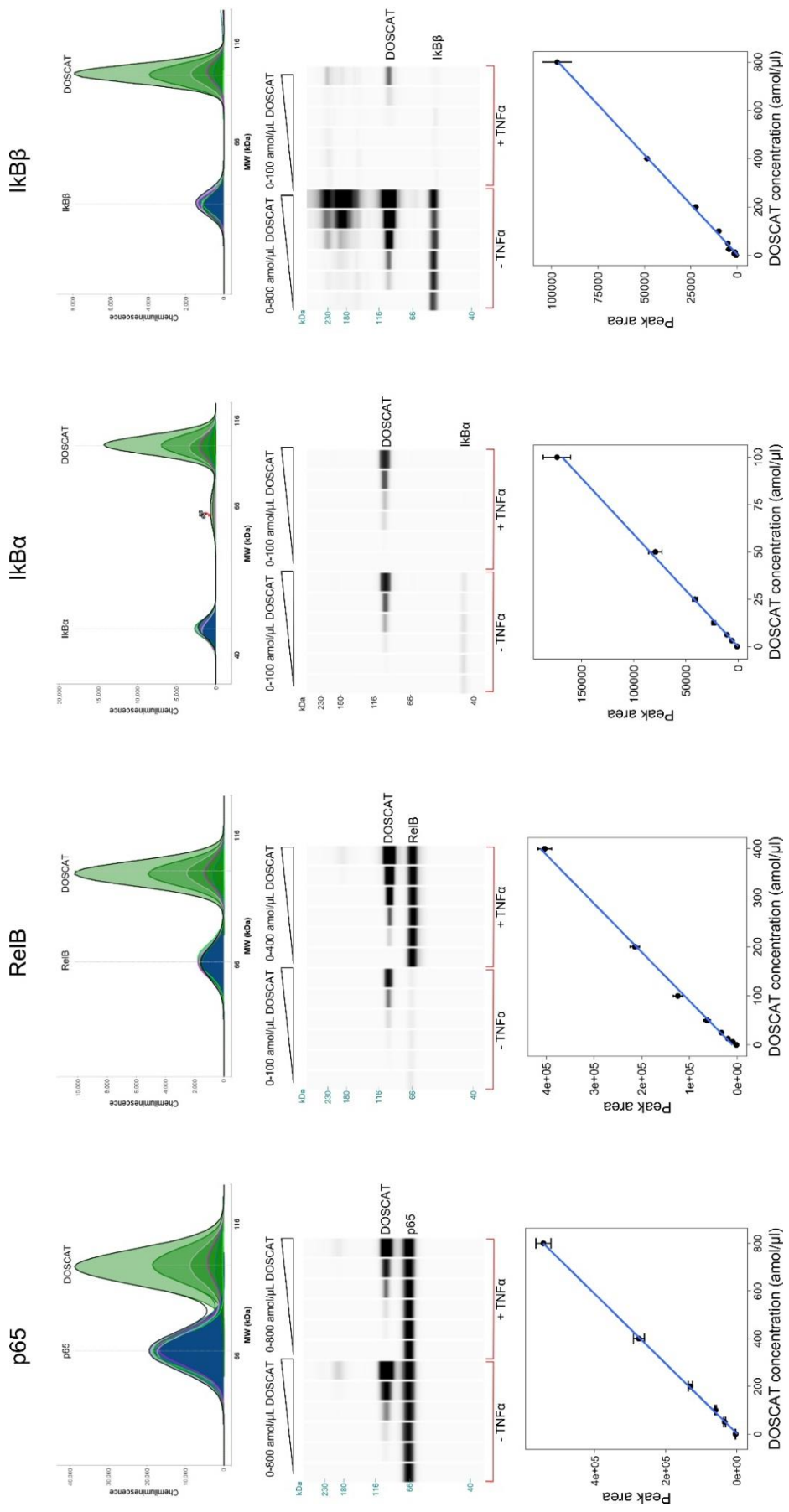


Figure 4.8 Quantitative western blots. Representative Simple Western data for 0.4 µg/µl SK-NA-S cell lysate ± TNFα stimulation spiked with a dilution series of DOSCAT and probed with antibodies for each target protein, shown in electropherogram view (top) and gel view (middle). Bottom: Calibration curves generated by DOSCAT for each protein. Data presented as mean ± standard error (n = 6).

Although this DOSCAT met many of the criteria of a dual-purpose standard, the critical test of such a standard is in the quantification of endogenous proteins. Specifically, would both QWB and MS assays yield acceptable quantification, and secondarily, would the presence of a complex background such as a cell homogenate impair the ability to use the standard? Further, the migration of the DOSCAT in the capillary QWB system had the potential to lead to interference with the signal for the endogenous analyte. To explore a realistic scenario for deployment of the artificial dual standard, DOSCAT was spiked into protein extracts of unstimulated and TNFα-stimulated SK-N-AS cells. The same cell lysates, spiked with the DOSCAT standard, were used for MS or QWB quantification, the difference being that for QWB, the spiked lysates with analysed without further treatment, whereas for MS quantification, the mixture was reduced, alkylated and proteolysed with trypsin. The DOSCAT was accurately quantified against an unlabelled Glu-Fib standard (Simpson and Beynon 2012; Brownridge et al. 2012) before subsequent SRM-MS and QWB analysis.

QWB was performed on the same samples in parallel with SRM-MS. The DOSCAT was serially diluted in SK-N-AS cell lysates to create an internal calibration series for each lysate. DOSCAT/lysate mixtures were analysed by automated western blotting system with antibodies that were specific and had been validated for each of the target proteins. The DOSCAT was designed to migrate at an apparent molecular weight (Mr) that differed from endogenous target proteins, allowing separation of signal between the exogenous DOSCAT standard and endogenous analytes. In practice, we observed some baseline interference at very high loadings of standards when the DOSCAT and analyte were similar in mobility, leading to less accurate quantification of peak area (Figure 4.9). This was readily resolved by restricting the standard data points to exclude the highest internal DOSCAT concentrations (Figure 4.10). This adjustment was applied for all target proteins, regardless of similarity in mobility between analyte and standard. Moreover, analyte levels were always within the linear calibration region of the standard curve. The DOSCAT signals were used to construct calibration curves that were linear in all cases (r^2>0.98, exemplar in Figure 4.8, all data in Figure 4.11).

For SRM-MS, the extracted ion chromatograms for unlabelled analyte and the labelled Q-peptide (released from DOSCAT) were used to calculate the amount of analyte present, as copy number per cell (cpc) values (using the calculation described in Materials and Methods). Protein level quantification values were calculated by averaging peptide values across technical replicates and in turn, taking the mean of the values obtained for different peptides. More than one peptide per protein was detected for p65 and RelB; for p65 the peptide cpc values were in good agreement but for RelB, SGPASGPSVPTGR gave lower cpc values than the other peptides (Figure 4.12). This may be due to the proximity of the endogenous peptide to the N-terminus or the presence of a previously unknown post translational modification on the peptide, and such the peptide was not included in calculating the final cpc value. Despite nominating peptides based on experimental evidence, neither IκBβ, IκBε nor the IκBα peptide

GSEPWK could be quantified by SRM-MS due to a combination of miscleavage potential in the standard, poorly performing peptides and the intrinsic low abundance of the target proteins (P. Brownridge, Holman, et al., 2011).

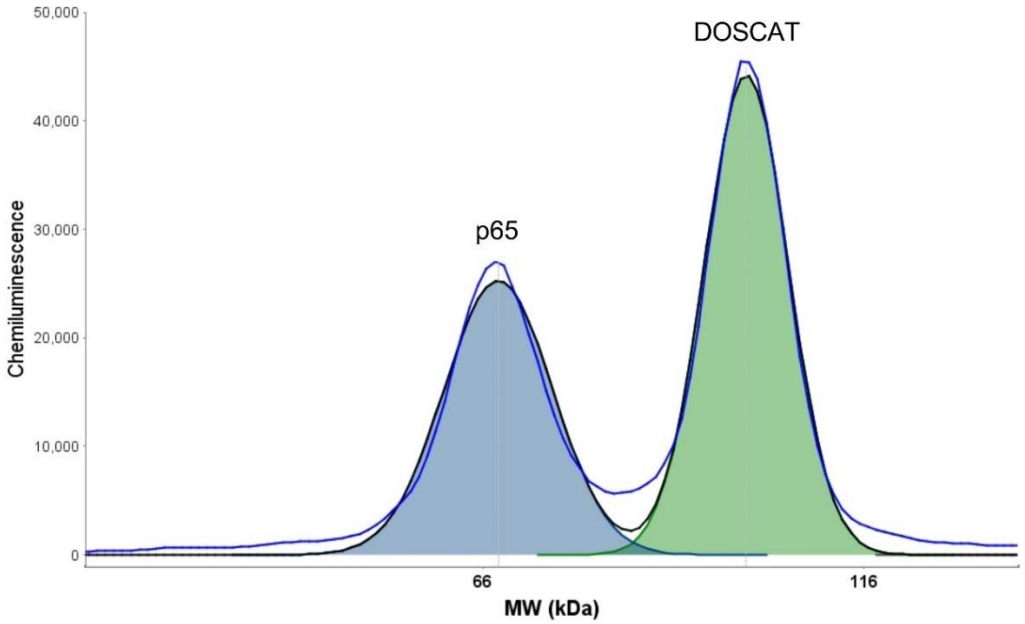


Figure 4.9 Baseline interference in Wes analysis. Electropherogram trace generated by automated western blotting for endogenous p65 and DOSCAT standard highlighting the overlap in peak fitting between the analyte and standard.

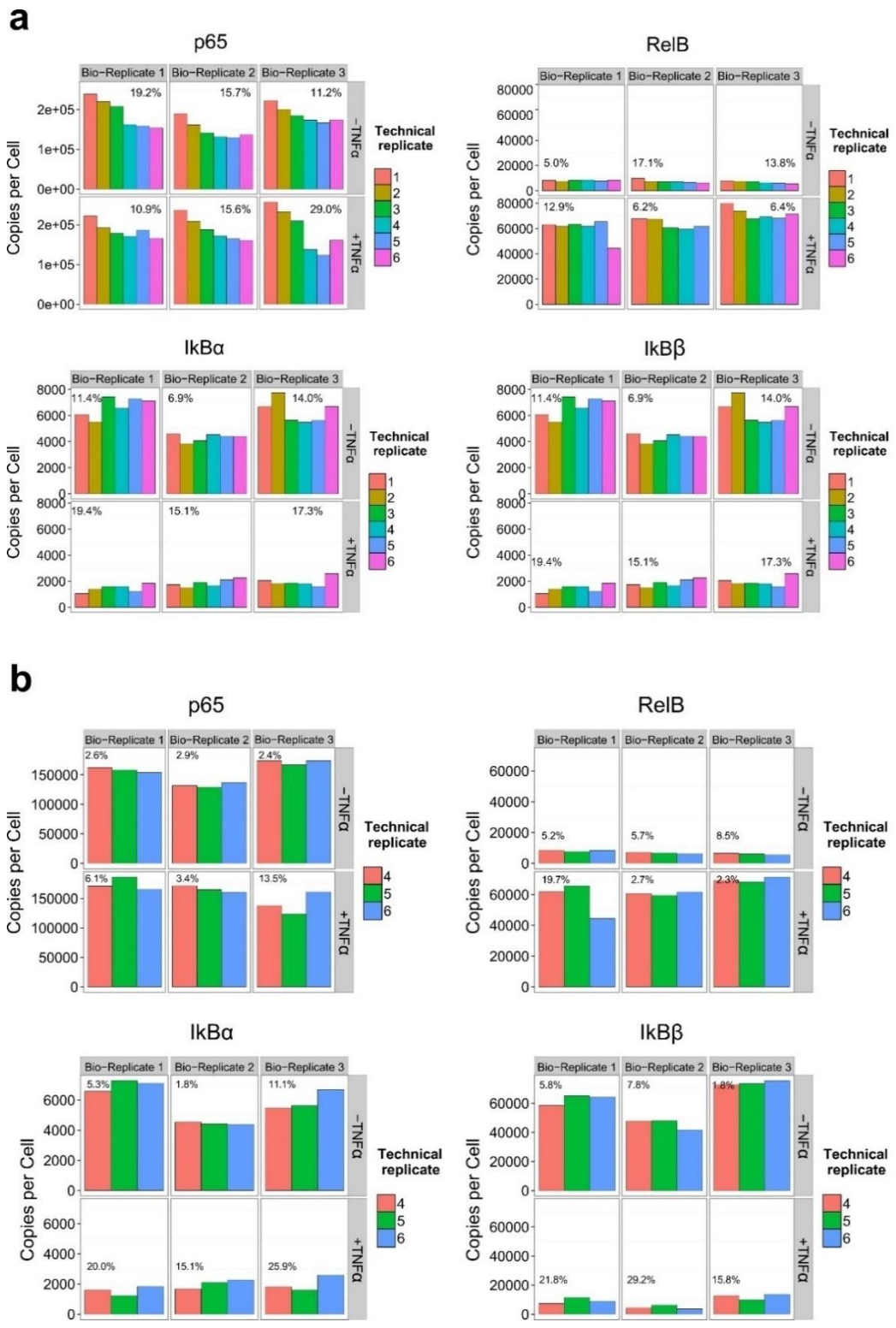


Figure 4.10 Data point exclusion of technical replicates. Quantification values determined by quantitative western blotting factored by biological replicate and treatment type (±TNFα) for a) all six technical replicates; b) the three technical replicates paired with the lowest spike-in DOSCAT concentrations. Percentage value in each panel represents % CV across the technical replicates.

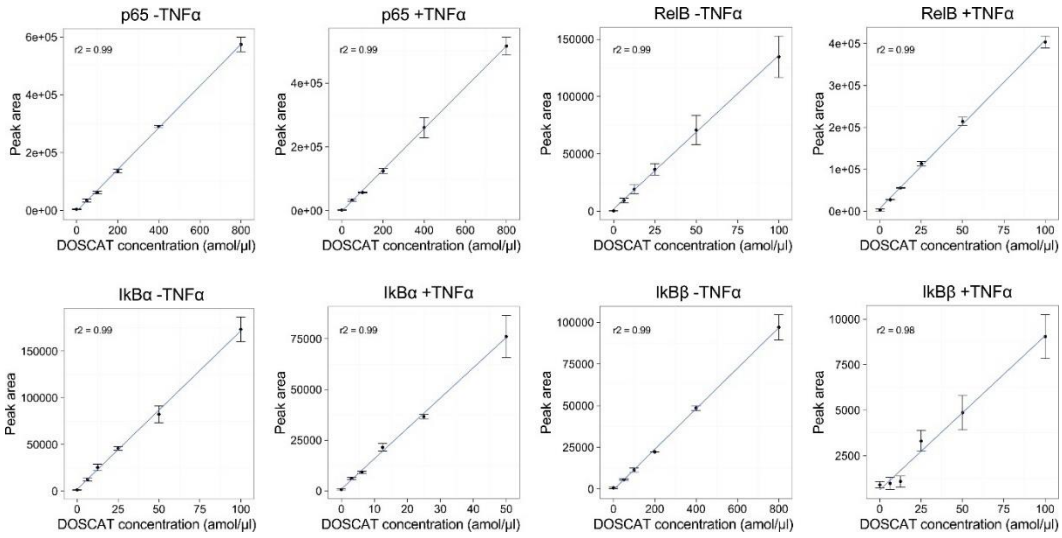


Figure 4.11 DOSCAT calibration curves in QWB. Dilution series of DOSCAT standard spiked into SK-NAS lysate +/- TNFα was analysed by automated capillary western blotting using four target antibodies against DOSCAT. Data presented as mean ± standard error for each treatment type (n = 3).

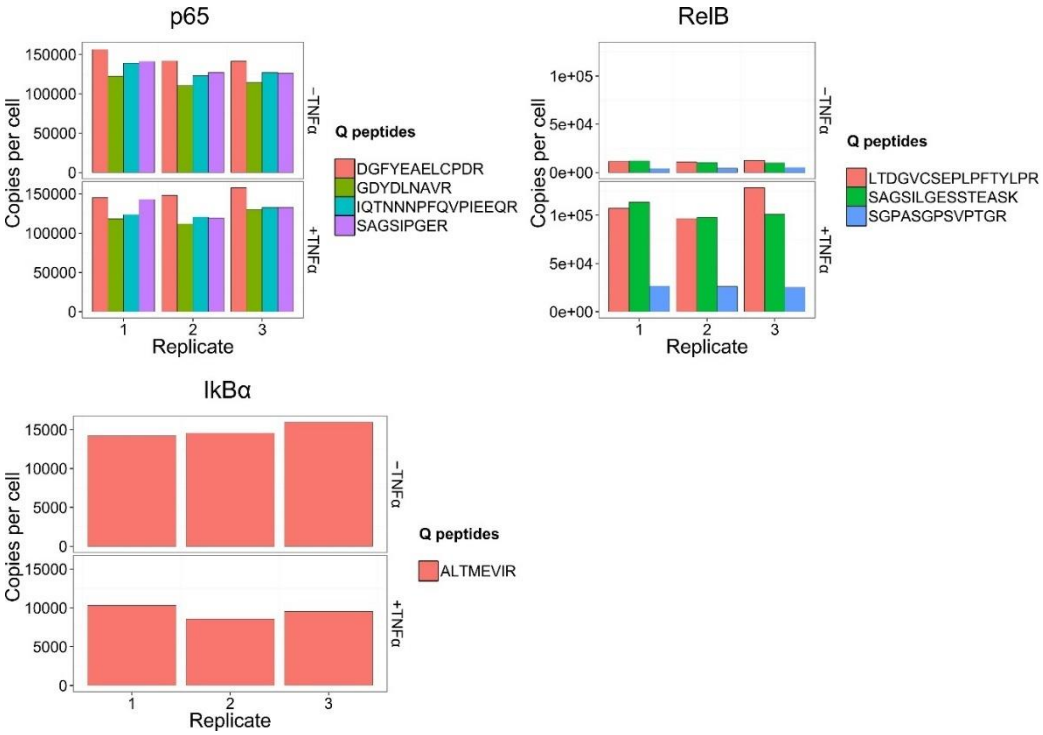


Figure 4.12 Peptide level quantification of target protein. DOSCAT spiked into SK-N-AS lysate was digested and analysed by scheduled SRM-MS. Quantification values in copies per cell for each detected Q-peptide were determined by analysing heavy:light ratios.

Two measures were used to compare the performance of the orthogonal analytical approaches. First, each method yielded absolute quantification values in copies per cell. Secondly, the changes in protein abundance after stimulation by TNFα could be compared. The utility of DOSCAT was demonstrated by quantification of five target proteins in the NF-κB pathway using both quantitative platforms. Exposure of cells to the cytokine TNFα significantly increases the levels of endogenous RelB and elicits degradation of the inhibitor proteins of NF-κB, IκBs (Lawrence, 2009). DOSCAT standardised quantification by SRM-MS or QWB were consonant with this expectation, giving confidence to the method. Although only a relatively small number of proteins were quantified in this study, agreement in copy numbers between the two techniques were improved compared to other studies that have compared MS and WB quantification (Kiel et al., 2014; Lawless et al., 2016; Picotti et al., 2009). In terms of both copies per cell and relative fold change, quantitative values generated by QWB and SRM-MS were in agreement for the 3 out of 5 proteins where comparison was possible (Figure 4.13a, b). Moreover, both techniques demonstrated high precision with a mean % CV across both techniques of 14 % (biological replicates, n=3). This is a clear demonstration that careful design of standards, coupled with appropriate technology and experimental design can converge these orthogonal methodologies.

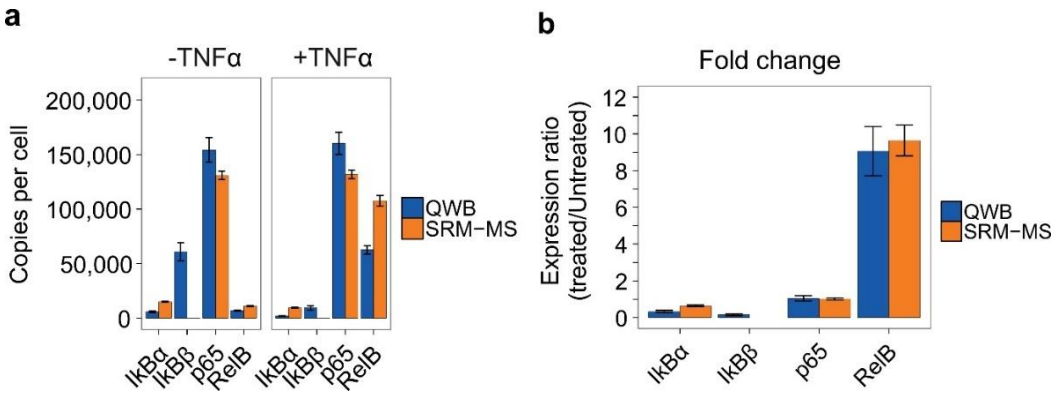


Figure 4.13 Target protein quantification. (a) Comparison of absolute quantification values for each target protein as obtained by SRM-MS and QWB. Quantification of IκBβ using SRM was not possible. Data presented as mean ± standard error (n = 3). (b) Relative fold change of proteins in TNFα treated and untreated SK-NA-S cells as measured by quantitative western blotting (QWB) and selected reaction monitoring (SRM-MS). Data are presented as mean ± standard error (n = 3).

Accurate protein quantification by orthogonal techniques is desirable to increase accuracy and robustness of quantitative data, especially important in systems analysis of specific pathways or biomarker validation. Quantification by two methodologies in parallel in a single experiment is difficult, though, due to differences in calibration standards. To address this problem we have proposed DOSCATs as a single calibration standard to support quantification by both QWB and SRM-MS assays across a single workflow. In some instances, QWB allowed quantification whereas SRM-MS did not, due to lack of peptide detection at low protein abundance and poor peptide ionisation. This further emphasises the value of using orthogonal techniques to quantify proteins across a wide dynamic range, or small proteins where options for Q-peptides are restricted. In our hands both quantitative platforms demonstrated an equivalently high level of precision using DOSCAT, which was within the range typically reported for SRM and QWB (using Simple Western technology) assays (J. Q. Chen et al., 2013; Lawless et al., 2016; Loughney et al., 2014). Using 'classic' western blotting, it would be anticipated that reproducibility would be lower due to the greater number of manual handling steps. With rigorous experimental design and well validated antibodies (Ghosh et al., 2014; Taylor et al., 2013), there is no reason why similarly equivalent quantitative data could be obtained using DOSCAT as a calibration standard.

Based on well-established QconCAT technology for which much of the route to deployment is well characterised (P. J. Brownridge et al., 2012; Scott et al., 2016) (~900 citations citing QconCATs are recorded in Google Scholar as of December 2016), the DOSCAT workflow is simple to implement (a generic workflow is presented in Figure 4.14). Since the initial publication in 2005 (Beynon et al., 2005), over 200 QconCATs have been utilised in a range of quantitative studies. QconCATs differ to natural proteins in amino acid composition (D. M. Simpson et al., 2012) and almost always accumulate in inclusion bodies. Expression in inclusion bodies can provide a useful pre-purification step but can also create problems of insolubility in long-term storage post purification – the use of MS-compatible detergents such as RapiGest *SF* and a low concentration of a reducing agent can minimise such problems (D. M. Simpson et al., 2012). The NF-κB DOSCAT expression is resonant with these previous observations, and we would expect future DOSCATs to follow this trend. In terms of design, selection criteria for Q-peptides are well documented (Brownridge et al. 2011; Holman, Sims, and Eyers 2012) and can equally be applied to DOSCATs. Of course, the availability of antibodies that are cross-reactive to known linear epitopes is a limiting factor when constructing a DOSCAT. However, a search for the term 'synthetic peptide' in the CiteAb (https://www.citeab.com/) database revealed over half a million antibodies raised to linear synthetic peptides. Thus, in an environment that was permissive to collaboration with antibody manufacturers, or where custom antibodies are made in-house, these peptide sequences could be included in DOSCATs. We encourage antibody manufacturers and suppliers to release the peptide sequences used for antibody generation as a contribution to reproducible research but also, to ease the construction of standards such as DOSCATs. If restricted proteolytic sites are built into the DOSCAT sequence, there is also an element of adaptability

in the workflow so that fragments of a predictable electrophoretic mobility are produced upon incubation with specific proteases. This not only allows for additional flexibility if the DOSCAT standard migrates at a similar mobility to endogenous proteins but the ability to create standards containing epitopes separable by limited proteolysis can assist in multiplexed protein detection in a single lane or capillary of a western blot.

DOSCATs offer a new calibration tool for protein quantification by both SRM-MS and QWB and unite two disparate workflows by a single calibration standard yielding equivalent quantification. Western blotting is one of the most widely used research techniques practised by the majority of cell biologists, despite previous limitations in delivering quantitative data. The DOSCAT approach has the potential to enhance the rigour of QWB that is more readily applied after MS validation, to generate reliable quantitative information particularly relevant for systems biology studies and contribute to the desired increase in reproducibility of biological research.

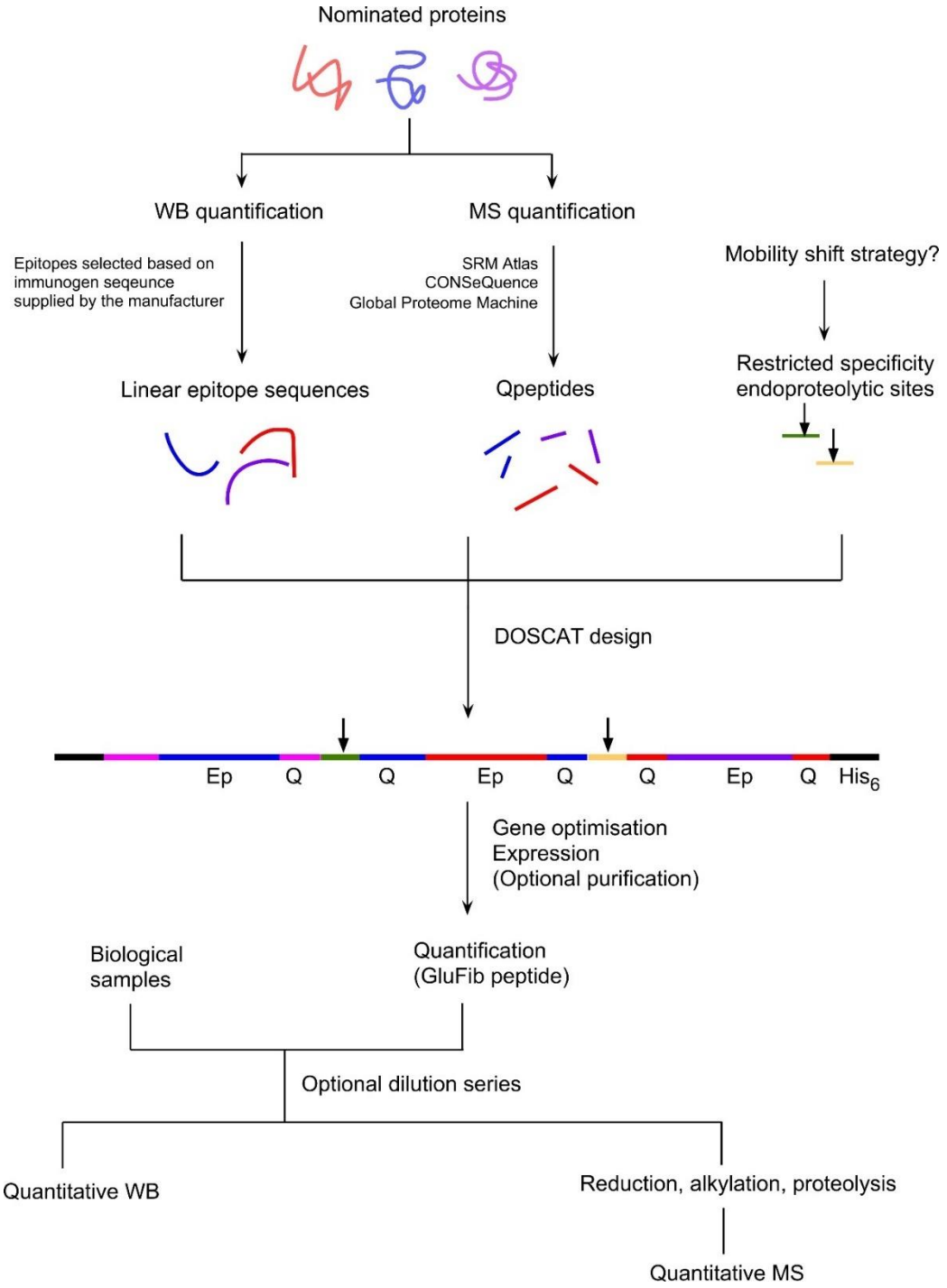


Figure 4.14 Overall DOSCAT workflow. The deployment of DOSCAT quantification can be resolved into three phases; design, expression and assay development. This workflow summarises the major steps.

4.3. Additional results and discussion

The paper making up section 4.2 describes the deployment of NFκB-DOSCAT-2 to successfully quantify target NF-κB proteins. However, additional work was also performed that was not included in the manuscript. Section 4.3 contains the details of this work, made up of a study of NFκB-DOSCAT-2 stability, the performance of NFκB-DOSCAT-2 when used in both classic western blotting with fluorescent detection, or in capillary WB as an external standard, and experiments assessing the new restricted proteolysis sites incorporated into NFκB-DOSCAT-2.

4.3.1. NFκB-DOSCAT-2 stability

The stability of a DOSCAT is critical for its use, so a series of experiments were carried out to determine NFκB-DOSCAT-2 stability when incubated by itself and in cell lysate, and how protein adsorption affects NFκB-DOSCAT-2 concentration.

To test whether NFκB-DOSCAT-2 was being proteolysed by lysate, NFκB-DOSCAT-2 was incubated by itself, with SK-N-AS cell lysate or with heat-inactivated cell lysate over 2 hours at room temperature. Samples were taken over a period of 120 mins and analysed by automated western blot using a p65 antibody (Figure 4.15). The signal for endogenous p65 in lysate stayed constant over the time course, as did the signal for NFκB-DOSCAT-2 spiked into cell lysate and heat-inactivated cell lysate. This ruled out proteolysis by cell lysate as an explanation for NFκB-DOSCAT-2 loss. When NFκB-DOSCAT-2 was incubated by itself, however, the signal decreased significantly over the time course. This suggested that NFκB-DOSCAT-2 was degrading or adsorbing to the tube surface when not contained in a cell extract. To assay potential protein adsorption to plastic tubes, an experiment was designed in which NFκB-DOSCAT-2 was passaged through five low bind tubes, with a sample removed at each tube (Figure 4.16). To investigate whether it was degrading over time, NFκB-DOSCAT-2 was incubated at RT and samples removed at time points over 2 hours. All samples were analysed by SDS-PAGE and densitometry performed on NFκB-DOSCAT-2 bands to quantify levels of protein in each sample (Figure 4.17). The tube passage experiment demonstrated that a significant amount of NFκB-DOSCAT-2 was being lost as it moved from tube to tube. By tube 4 over 80% of NFκB-DOSCAT-2 had been lost relative to the starting material (tube 0). The incubation experiment resulted in 40% of NFκB-DOSCAT-2 being lost after 2 h. However, no bands relating to degradation products became visible on the SDS-PAGE gel, so these losses, when taken in conjunction with the tube passage results, were attributed to adsorption onto the tube and pipette tip.

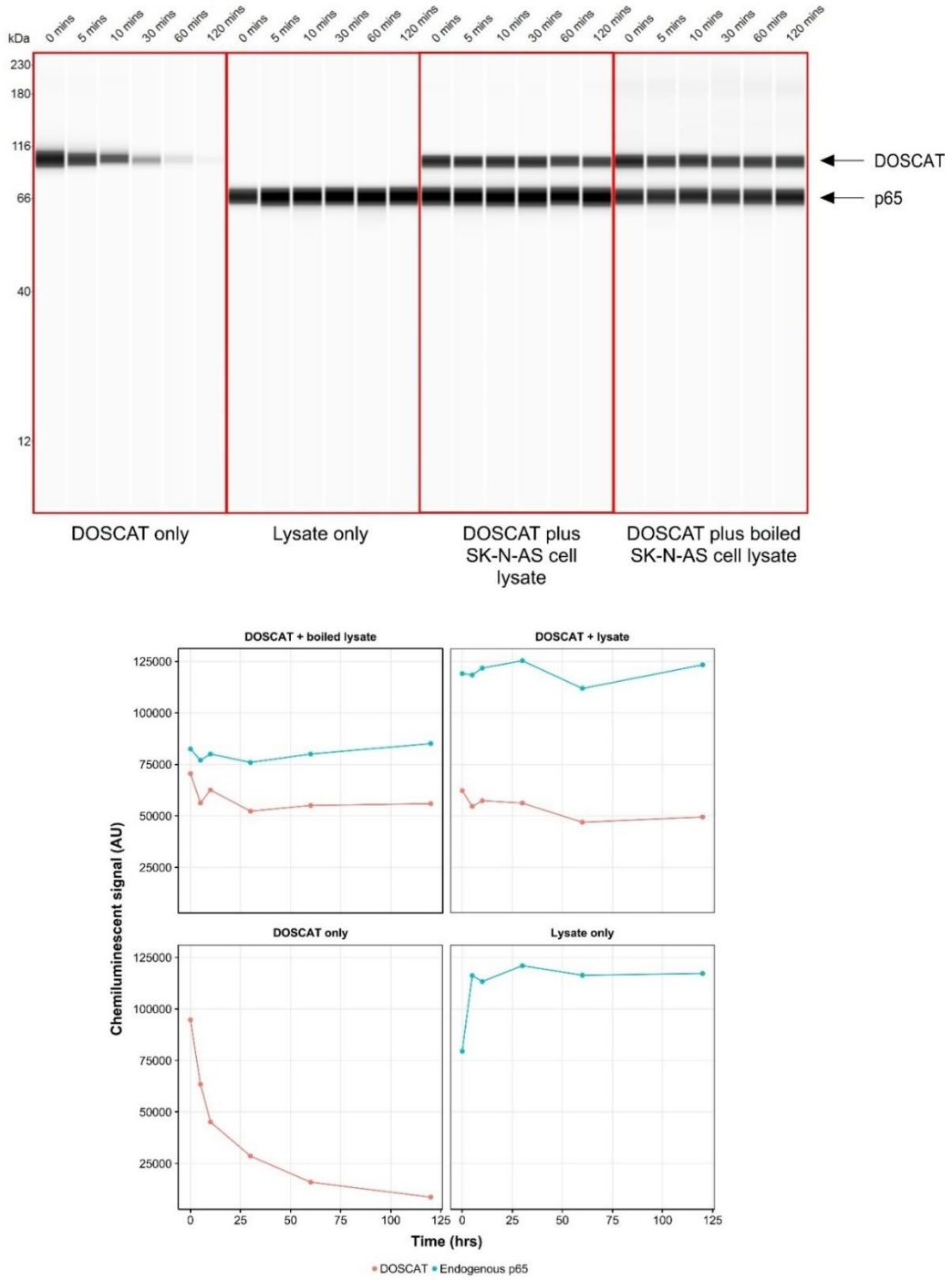


Figure 4.15 DOSCAT stability during incubation. SK-N-AS cell lysate ± 95°C heating for 10 min and DOSCAT were incubated either individually or mixed together for 2 h at RT. Samples were taken at time points throughout the time course and analysed by automated capillary western blotting using a p65 antibody (top panel). Chemiluminescent signal calculated by Compass software was plotted over time (bottom panel).

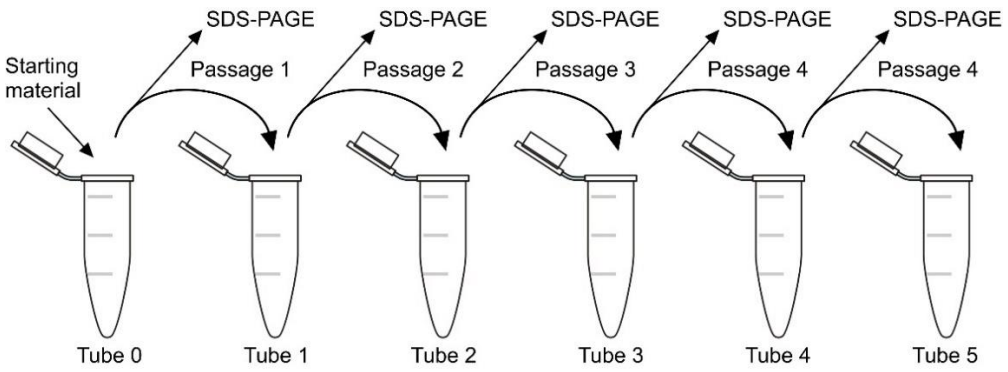


Figure 4.16 Tube passage experiment. Starting material (tube 0) is pipetted through a series of tubes or glass vials, with a sample removed at each tube for analysis by e.g. SDS-PAGE.

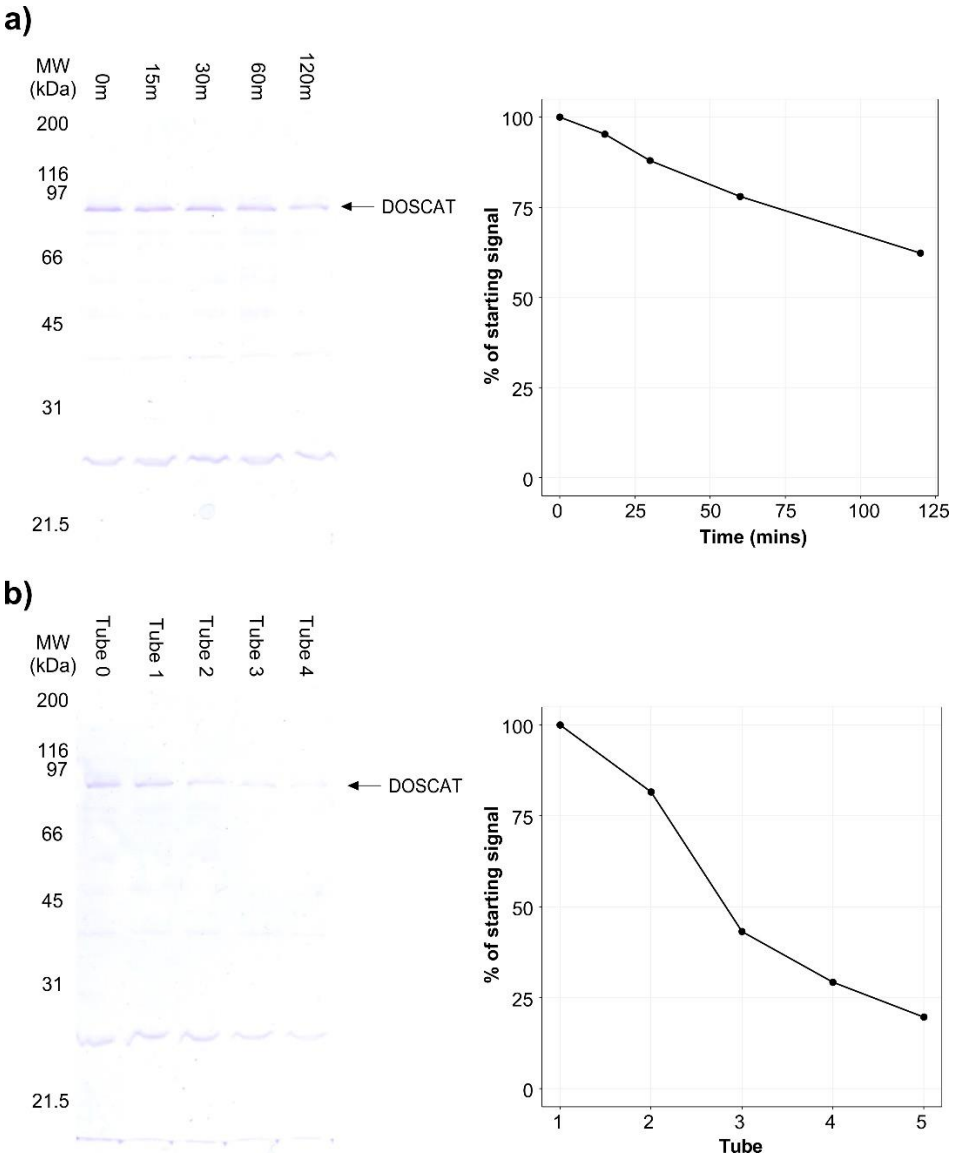


Figure 4.17 DOSCAT tube passage and incubation. NFκB-DOSCAT-2 was either a) incubated for 2 h or b) passaged through four tubes, with samples taken at each tube or time-point. Samples were analysed by SDS-PAGE and densitometry performed to determine signal relative to starting material.

In an initial attempt to stop these losses, the surface of the tube was modified. Regular non-low bind Eppendorf tubes (made from polypropylene) were siliconised using Sigmacote and the tube passage experiment was repeated alongside low bind tubes (Figure 4.18). The siliconisation technique did prevent losses slightly, however they were still very high (70% relative to tube 0). Glass vials, both siliconised and non-siliconised, were then trialled. These both performed better than plastic tubes, however over 50% of NFκB-DOSCAT-2 was still being lost by tube 5 relative to the starting material (Figure 4.19). The surface material of tips was also investigated. Tips were siliconised by pipetting Sigmacote up and down before being washed in MilliQ water. A volume of DOSCAT was placed in a siliconized glass vial and the entire volume taken up in a siliconized pipette tip before being dispensed back into the vial. A sample of DOSCAT was taken and the procedure repeated four more times. Siliconised tips performed much better than regular tips (Figure 4.20) with little NFκB-DOSCAT-2 lost over the course of the experiment. This indicated that tip material did have an effect, although less NFκB-DOSCAT-2 was lost on a tip surface compared to that of a tube.

As modifying the tube surface had only a limited effect, an alternative approach was tested in which Rapigest *SF*, an acid labile surfactant regularly used in MS workflows, was added to the NFκB-DOSCAT-2 solution. Rapigest *SF* acts like a detergent to solubilise and stabilise proteins, but unlike detergents such as SDS or Triton it is compatible with downstream MS analysis. Rapigest *SF* at 0.1% (w/v) was added to NFκB-DOSCAT-2 and the tube passage experiment repeated. Losses were limited to <20% of starting material by tube 5, even when non-siliconised tubes and tips were used, demonstrating the effectiveness of Rapigest *SF* to prevent DOSCAT adsorption (Figure 4.21). The tube passage experiment was repeated on the Wes system to show that Rapigest *SF* was a compatible reagent with the system (Figure 4.22). Some detergents such as SDS can affect the running of the capillary WB system through inaccurate molecular weight calibration or diminished chemiluminescent signal; however, no adverse effects in data quality were observed when Rapigest *SF* was contained in the sample. As observed on the SDS-PAGE gels there were minimal losses throughout the tube passage experiment, and the overall signal was higher relative to the non-Rapigest *SF* control, presumably due to fewer DOSCAT losses during dilution from stock material.

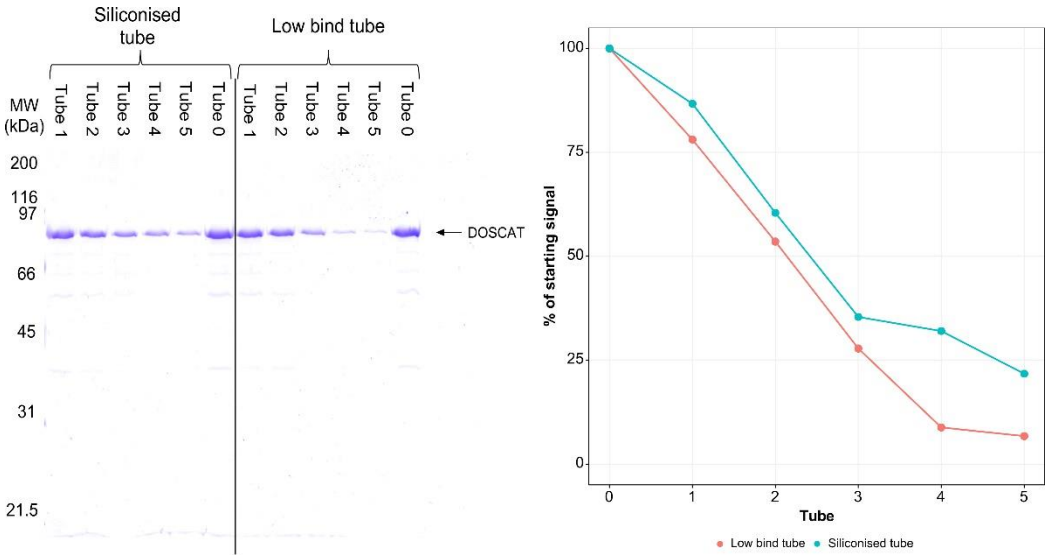


Figure 4.18 Effect of tube modification on NFκB-DOSCAT-2 adsorption. NFκB-DOSCAT-2 was passaged through five Eppendorf low-bind tubes or Eppendorf tubes siliconized using Sigmacote and samples at each tube were analysed by SDS-PAGE (left). Densitometry was performed and signal relative to starting material was plotted against tube number (right).

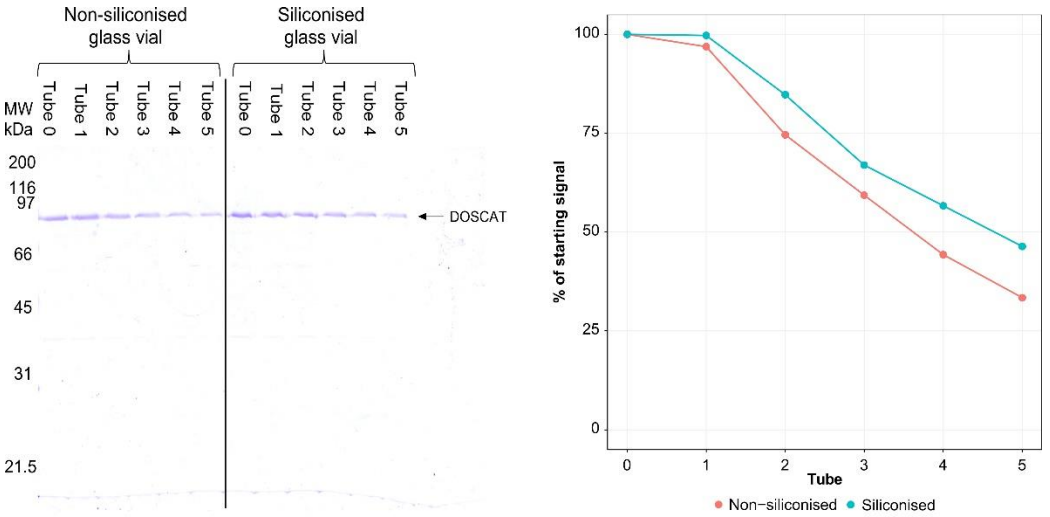


Figure 4.19 NFκB-DOSCAT-2 adsorption to glassware. NFκB-DOSCAT-2 was passaged through five glass vials with and without siliconized surfaces and samples at each tube were analysed by SDS-PAGE (left). Densitometry was performed and signal relative to starting material was plotted against vial number (right).

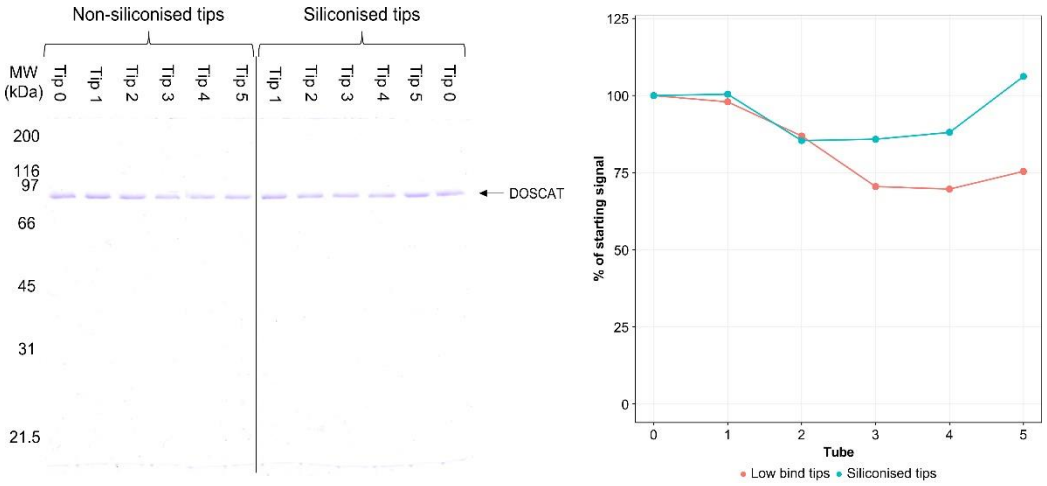


Figure 4.20 Effect of tip modification on NFκB-DOSCAT-2 adsorption. Siliconised and non-siliconised tips were used to take up and dispense a volume of NFκB-DOSCAT-2 from a glass vial. A sample of NFκB-DOSCAT-2 was taken and the procedure repeated four times using fresh tips. Samples were analysed by SDS-PAGE (left), densitometry performed and signal relative to starting material plotted against tube number (right).

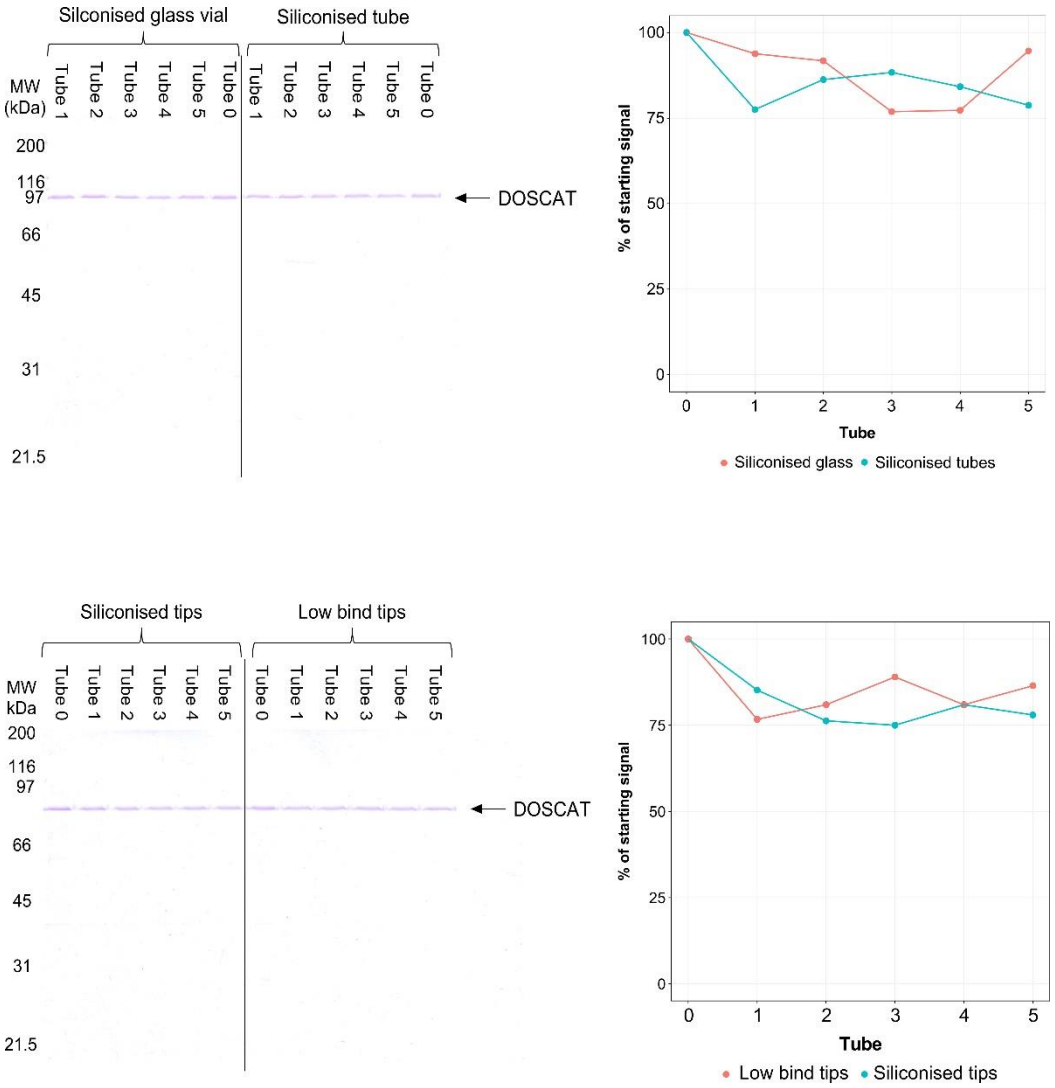


Figure 4.21 Use of Rapigest *SF* to prevent NFκB-DOSCAT-2 adsorption. Rapigest *SF* (0.1% *w/v*) was added to NFκB-DOSCAT-2, which was passaged through five siliconised glass vials or glass tubes using regular tips, and five low-bind tubes using siliconised or low bind tips. Samples at each tube were analysed by SDS-PAGE (left panels), densitometry performed and signal relative to starting material plotted against tube number (right panels).

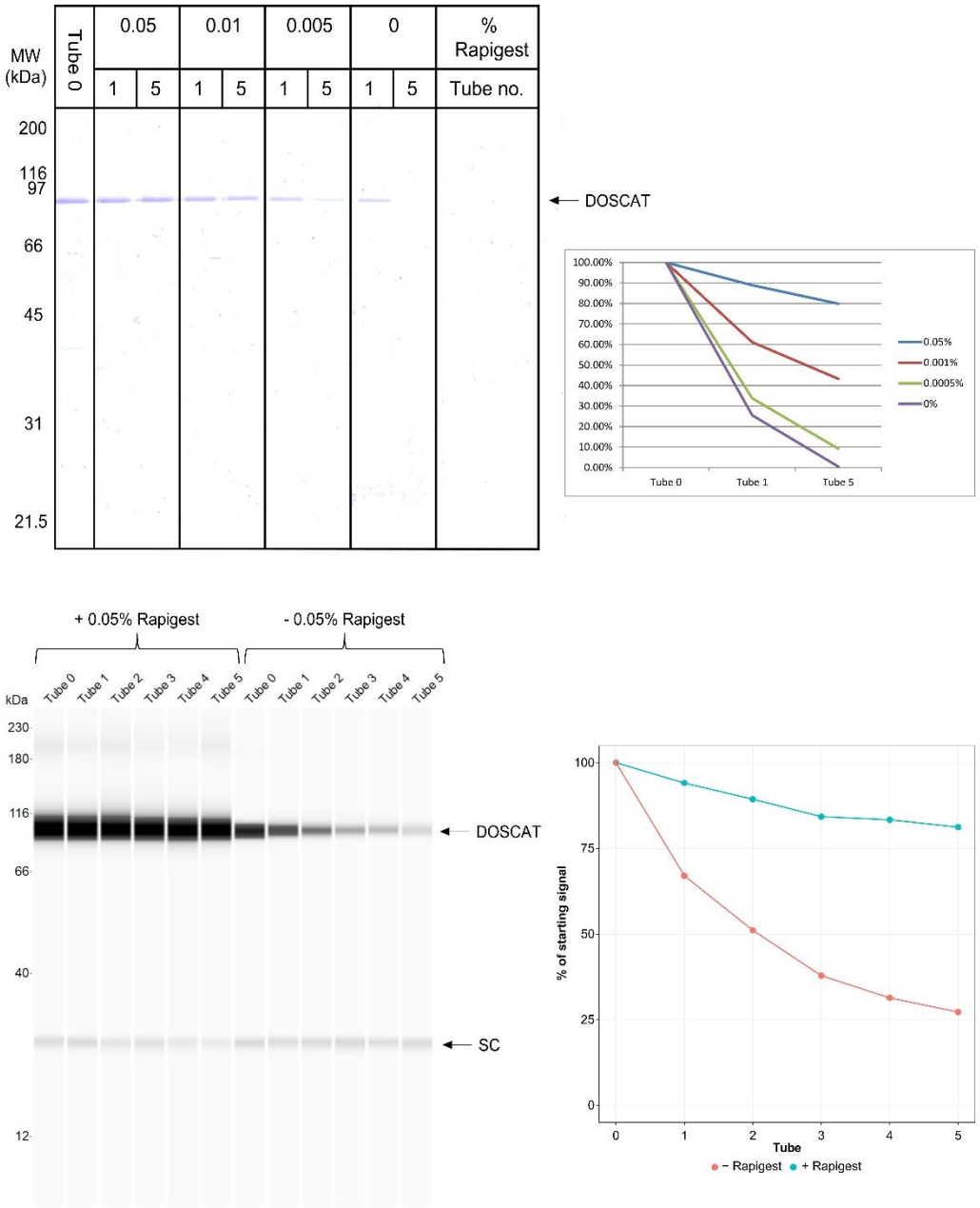


Figure 4.22 Rapigest *SF* compatibility with Wes system. a) Rapigest *SF* was added to NFκB-DOSCAT-2 at varying concentration from 0 – 0.05% *w/v* and this solution was passaged through five low bind tubes using regular tips. Samples at each tube were analysed by SDS-PAGE, densitometry performed and signal plotted against tube number. b) NFκB-DOSCAT-2 ± 0.05% *w/v* Rapigest SF was passaged through five low-bind tubes and analysed by automated capillary western blotting using a p65 antibody. Chemiluminescent signal relative to starting material was calculated by Compass software and plotted against tube number.

The tendency for NFκB-DOSCAT-2 to adsorb to surfaces has been demonstrated, as has an effective solution in the addition of Rapigest *SF* to the buffer solution. Understanding this property of DOSCATs and how to control it is hugely important, as it gives greater confidence in quantitative assays calibrated by a DOSCAT. If the concentration of a DOSCAT (or any calibrant) is measured inaccurately, resultant quantification values will also be inaccurate. Unlike an SRM experiment, there is not a way to quantify DOSCATs internally in a western blot, so it is vital that there is confidence that DOSCAT concentration will not change during the QWB workflow. This work has demonstrated that the NFκB-DOSCAT-2 is stable in a buffer solution containing Rapigest *SF*, does not degrade at room temperature and is not proteolysed when contained in biological sample; therefore, quantitative values generated by QWB are reliable.

The results presented here are unsurprising as it has been long known that proteins will readably adsorb to surfaces (Rabe et al., 2011). Indeed, protein adsorption is a requisite in many biological processes such as trans-membrane signalling and control of cell adhesion and proliferation (Allan L. T. et al., 2006; Hinderliter et al., 2006). Moreover, within the field of medical devices protein adsorption can affect the biocompatibility of an implant leading to its premature degradation (Brash et al., 2012). Similar observations were also made with the first iteration of NFκB-DOSCAT, which seemed to adsorb to the capillary wall in the Wes system (section 3.2.3.). Changing the surface properties of the tubes or tips did not make much difference to NFκB-DOSCAT-2 adsorption. Although glass performed slightly better than polypropylene plastic here, there is evidence to suggest that the optimal plastic or glassware is specific to the protein or peptide and that there is no clear correlation between amino acid composition and optimal surface (Goebel-Stengel et al., 2011). Siliconisation of surfaces led to a small improvement in protein recovery, but overall losses were still considerable. This goes against the view that siliconisation reduces protein adsorption (Seed 2001), although other studies have also found siliconisation does not prevent adsorption and in some cases actually increases it (Goebel-Stengel et al., 2011). Changing the solvent composition through the addition of Rapigest *SF* all but prevented protein adsorption. This concurs with other results that suggest that modifying the solvent rather than the surface is a much more effective way to prevent protein-surface interactions (Suelter et al., 1983). This is because efficiency of surface treatment may not always be complete, and different proteins will have unique affinities for different surfaces.

Surfactants such as Rapigest *SF* can improve protein stability (Bummer et al., 2000), and Rapigest *SF* has previously been shown to improve protein solubility during enzymatic digestion (Meng et al., 2002; Yu et al., 2003). Protein stability can affect rate of adsorption, as the conformation of a folded protein is restricted and the conformational entropy is low (Willem, 1986). Protein unfolding after initial contact with a surface leads to a conformational entropy gain, driving adsorption even at hydrophilic surfaces. It has been demonstrated that a more stable protein will be less likely to denature on a surface, so protein adsorption will be lower

(Billsten et al., 1997; Karlsson et al., 2005; Malmsten, 1998). As an artificial protein, a DOSCAT does not have a defined folded conformation and is inherently unstable, which may explain why it will so readily adsorb, even to modified surfaces. By stabilising the conformational state of a DOSCAT using Rapigest *SF*, the rate of irreversible protein adsorption has been reduced. A similar phenomenon has been observed using sugar excipients to stabilise protein states in solution, thus decreasing adsorption (Wendorf et al., 2004).

These experiments give assurances about DOSCAT stability over the short term, but they do not address the long-term stability of DOSCATs, nor the effect of pH and temperature on DOSCAT stability and adsorption. Further work is required to establish these properties, and strategies to mitigate protein loss over time must be conceived.

4.3.2. Optimisation of QWB assays

For each protein to be analysed by the Wes system, three parameters were optimised: cell lysate concentration, DOSCAT concentration range for the calibration curve and antibody concentration.

To find the optimal antibody concentration, a dilution series of each antibody was run against a fixed concentration of NFκB-DOSCAT-2 spiked into SK-NA-S cell lysate. The peak area and average baseline was derived using Compass software and the peak area/baseline value calculated for each antibody concentration (Figure 4.23). As the antibody concentration increases, the signal increases until the point of saturation. A concentration near this saturation point common to both endogenous analyte and NFκB-DOSCAT-2 standard was chosen and used as the optimal antibody concentration for all future assays. NFκB-DOSCAT-2 was not detected by the IκBε Ab, so optimal concentration was based on data from the endogenous analyte alone.

The optimal lysate concentration falls within the linear dynamic range of the assay, that being between the limit of detection and saturation level. This ensures that changes in signal are directly proportional to changes in protein level. To determine the optimal loading concentrations a dilution series of TNFα treated and untreated SK-N-AS cell lysate was analysed by each antibody (Figure 4.24). Optimal lysate concentration was defined as 0.4 mg/mL as it fell within the linear dynamic range for each protein. The range of NFκB-DOSCAT-2 concentration for the calibration curve was then determined for each target protein by running a dilution series of NFκB-DOSCAT-2 and probing with each antibody. The signal of the endogenous protein at optimal lysate concentration (in both untreated and treated lysate) was used to select the range of NFκB-DOSCAT-2, as the endogenous signal should be within the range of the calibration curve. This calibration curve range was different for each protein, for example p65 ranged from 0-800 amol/µL and IκBα 0-100 amol/µL.

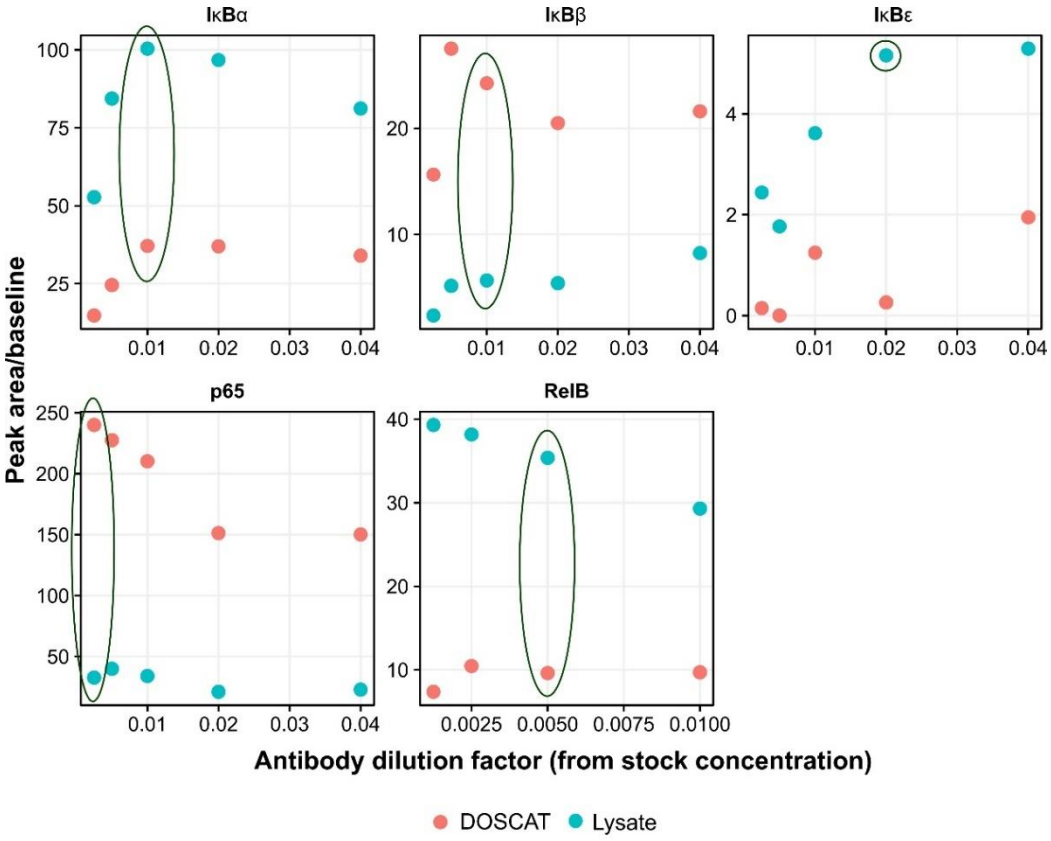


Figure 4.23 Antibody concentration optimisation. A dilution series of each antibody was prepared and used to analyse 100 fmol/µL DOSCAT spiked into 0.5 mg/mL SK-N-AS by automated capillary western blotting. Chemiluminescent signal was calculated by Compass software and plotted against each antibody concentration.

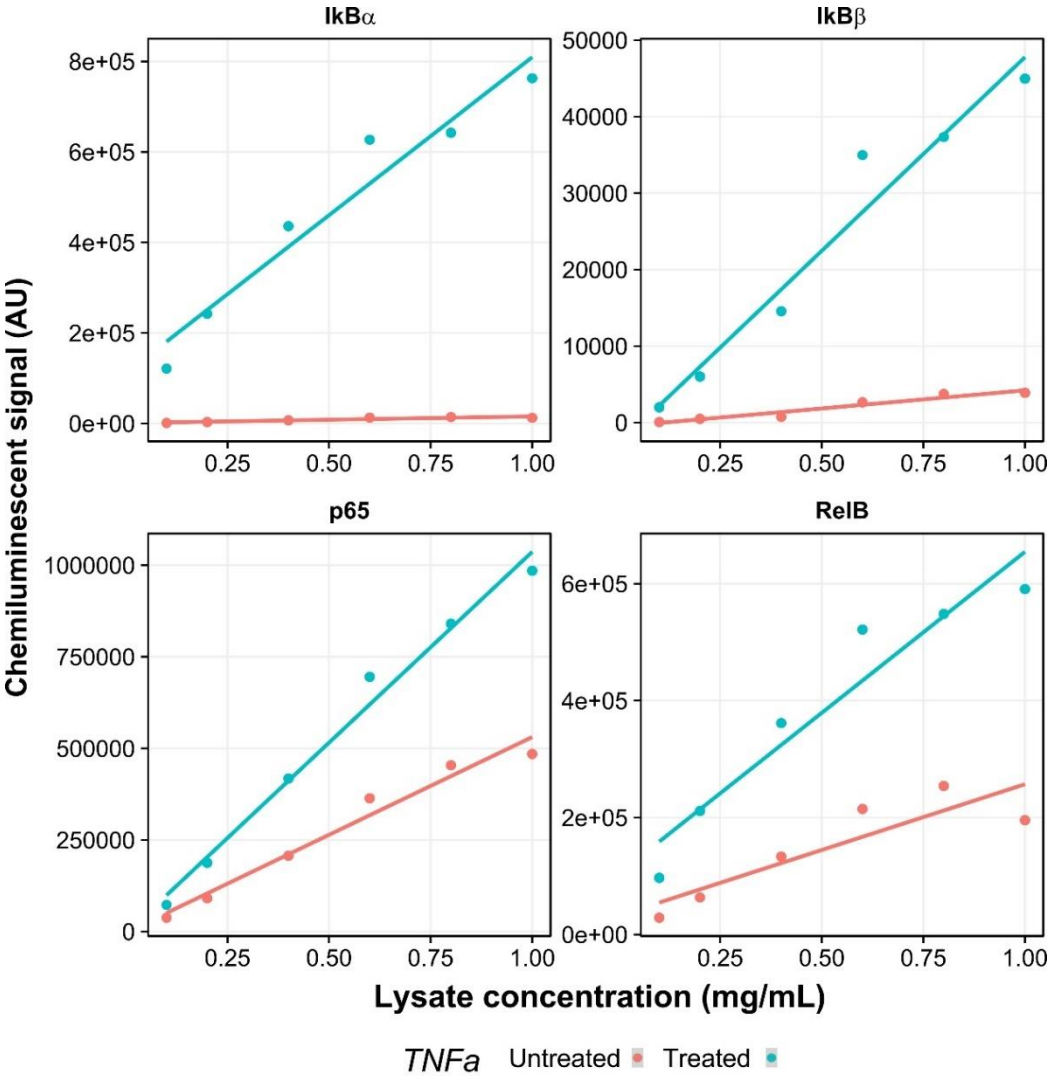


Figure 4.24 Optimisation of lysate concentration for QWB analysis. Dilutions series of TNFα treated and untreated SK-N-AS cell lysate from 0.1 - 1.0 mg/mL were prepared and analysed by automated capillary western blotting using the target NF-κB antibodies.

4.3.3. DOSCAT as an external standard

To further demonstrate the flexibility of DOSCAT as a standard, experiments were performed to show that NFκB-DOSCAT-2 could be used to calibrate protein quantification when not spiked into the sample analyte. DOSCAT surface adsorption is a known problem that is abated when it is contained within a complex matrix, therefore, both *E.coli* and yeast lysate were trialled as suitable diluents. Both lysates were prepared and heated at 80°C to inactivate any protease activity. After cooling, automated western blotting was performed on each lysate with and without spiked-in NFκB-DOSCAT-2, using each of the NF-κB antibodies (Figure 4.25). Non-specific binding was observed in *E.coli* lysate when probed with IκBβ and IκBε antibodies, but no signal was detected when yeast lysate was used as a background matrix. NFκB-DOSCAT-2 was detected when spiked in to both lysates, however, the signal was much stronger when yeast lysate was used as a matrix. Due the lack of background signal and increased standard signal, yeast lysate was selected as the background matrix for quantitative experiments.

For protein quantification, a NFκB-DOSCAT-2 stock was accurately quantified by digesting with Glu-Fib standard and analysis by SRM. The stock was spiked into a heat-inactivated yeast lysate matrix containing Rapigest *SF* and serially diluted in a yeast lysate/Rapigest *SF* diluent. Three biological replicates of TNFα treated and untreated SK-N-AS lysate alongside NFκB-DOSCAT-2 standards were prepared for Wes analysis and run using a p65 antibody for detection. Within a single QWB experiment, a six-point NFκB-DOSCAT-2 dilution series and three technical repeats of each biological replicate were run (Figure 4.26a). The experiment was repeated three times. Chemiluminescent signal was calculated using Compass software and NFκB-DOSCAT-2 calibration curves were constructed for each replicate. The three calibrations curves for each replicate agreed very well with one another and linearity was excellent (Figure 4.26b). Using the NFκB-DOSCAT-2 calibration curve within the same replicate, endogenous protein was quantified in terms of cpc (using the same calculation described in section 4.2.3.). Figure 4.26c displays the cpc values for each sample (denoted by shape) across each of the three QWB replicates (denoted by colour). Intra-assay variance between the technical repeats in each run was very low; across the technical repeats the average CV for untreated lysate was 4.4% and for TNFα treated lysate 8.0%. Inter-assay variance for each biological replicate was much higher, ranging from 24.0% (for C2) to 63.8% (T3), with an average of 46.4%. It is worth noting that values obtained in the third run (blue points in Figure 4.26c) were higher than the other two, which would have an adverse effect on overall inter-assay variance. Copy per cell values were quantified across all technical and biological replicates to give final cpc values. Compared to values obtained by QWB and SRM-MS where DOSCAT was employed as an internal standard, measured cpc were higher across both TNFα treated and untreated samples (Figure 4.26d). Moreover, overall variance was higher, with CVs of 22.7% and 22.5% for untreated and TNFα treated samples respectively.

Differences in copy per cell values compared with those generated with internally calibrated QWB and SRM-MS could be attributed to inaccurate concentrations of NFκB-DOSCAT-2 resultant from changes to the workflow. In this experiment, stock NFκB-DOSCAT-2 was quantified (by digestion with Glu-Fib and MS analysis) before being added to yeast lysate. This gave rise for an opportunity for unaccounted NFκB-DOSCAT-2 losses by adsorption. A lower actual concentration of NFκB-DOSCAT-2 than measured would lead to inflated cpc values for unknown analyte. Furthermore, for each gel replicate NFκB-DOSCAT-2 was transferred into fresh yeast lysate, so differential losses each time would account for the increased inter-assay variation. For more accurate quantification and improved reproducibility, NFκB-DOSCAT-2 should have been quantified when in the yeast lysate diluent. This highlights the careful consideration that must be given to experimental design when dealing with a calibrant that is very prone to surface adsorption.

Owing to the resources that would be required to quantify the entire set of proteins (i.e. 15 Wes runs compare to 6 for internal calibration), only p65 was quantified. This of course limits the ability to draw broad conclusions about accuracy and reproducibility but does offer a proof of principle of the specific approach. Despite this, the importance of the results stand in the demonstration that a DOSCAT can be used to calibrate a quantitative western blot when it is external to the analyte. This is a very useful attribute if a DOSCAT has the same electrophoretic mobility as an analyte (and it cannot be suitably adjusted by proteolysis) or if non-specific binding interferes with the DOSCAT signal.

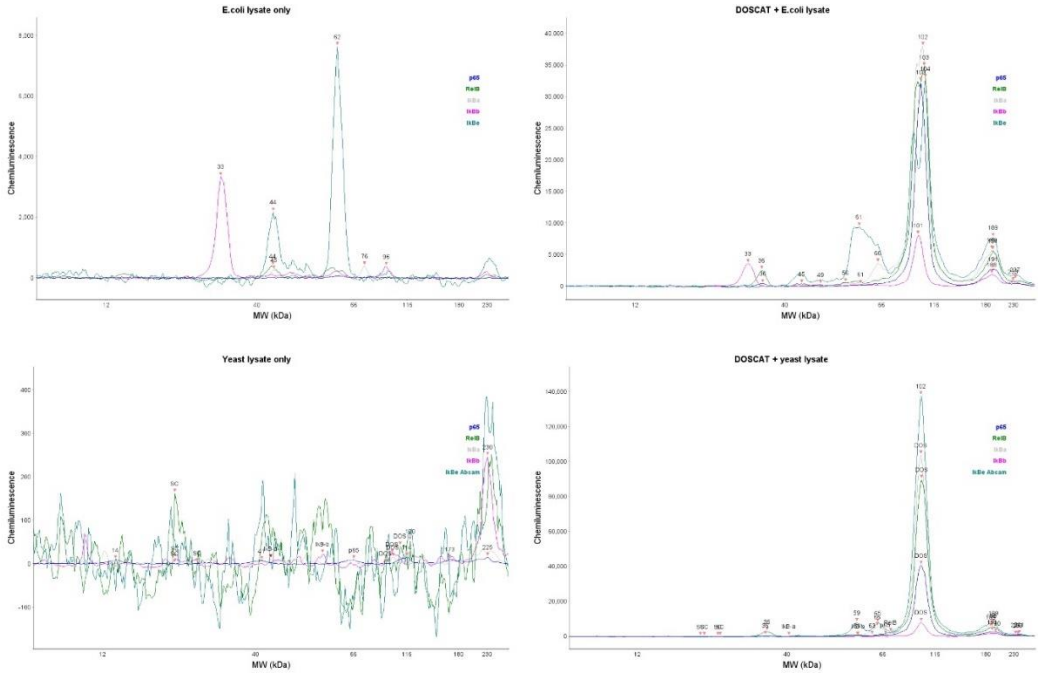


Figure 4.25 Optimisation of surrogate matrix for DOSCAT. Electropherograms resultant from heat-inactivated *E.coli* and yeast lysate ± 100 fmol/µL DOSCAT analysed by automated capillary western blotting using NF-κB antibodies.

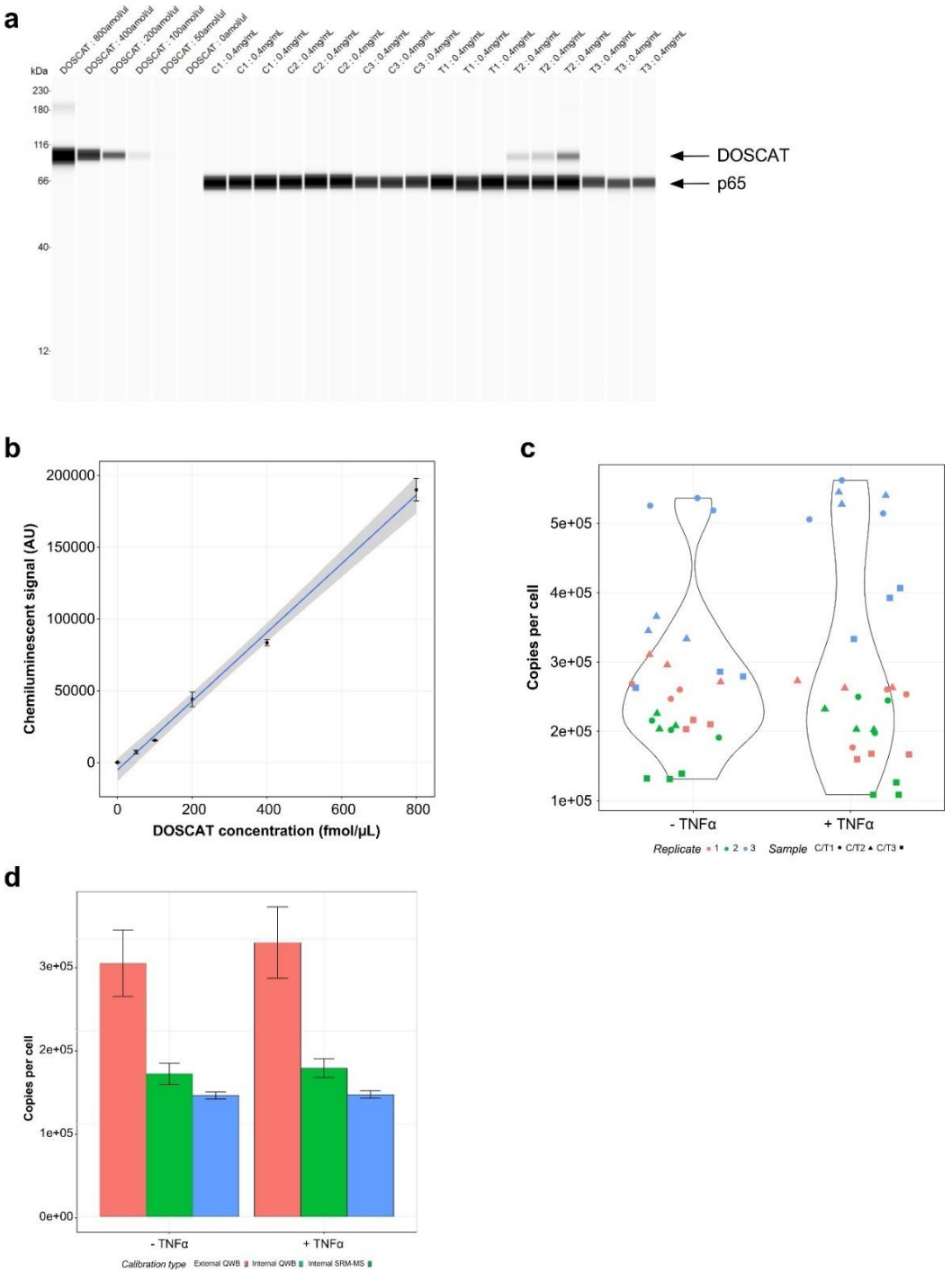


Figure 4.26 External calibration by DOSCAT in QWB. a) Representative pseudo-gels resultant from 0-800 fmol/µL DOSCAT and 0.4 mg/mL ± TNFα treated SK-N-AS cell lysate (biological replicates, n = 3; technical repeats, n = 3) analysed by automated capillary western blot using a p65 antibody. b) DOSCAT calibration curve, presented as mean ± standard error (n = 3). c) Copies per cell of p65 calculated using QWB and DOSCAT calibration, factored by gel replicates (colours) and sample replicates (shapes). d) Comparison of cpc values of p65 as determined by externally calibrated QWB and internally calibrated QWB and SRM-MS. Data presented as mean ± standard error (n = 3).

4.3.4. QWB using fluorescent detection

All QWBs using DOSCATs thus far have been performed using automated capillary western blotting due to its superior quantitative performance. However, many researchers do not have access to this technology, so it is important to demonstrate that a DOSCAT can be used in classic WB formats. To do this, NFκB-DOSCAT-2 was used for quantification using the same antibodies and cell line, but westerns blots were performed by electro-transfer onto PVDF membrane and subsequent fluorescent detection.

NFκB-DOSCAT-2 was spiked into TNFα treated or untreated SK-N-AS cell lysate and DOSCAT concentration accurately obtained by digesting a sample of this mixture with a Glu-Fib peptide standard and analysing the heavy:light ratio by SRM-MS. The mixture was serially diluted 7-fold in a cell lysate diluent and, along with a cell lysate only sample, was run on an SDS-PAGE gel, giving an 8-point standard curve and 8 technical repeats. Within a single gel, two untreated and two TNFα treated samples (C4, C5, T4, T5) were run and subsequently electro-transferred onto a single PVDF membrane together. SK-N-AS cell lysate from different biological replicates to those used previously had to be used due to lack of availability of the original samples. This would mean that a direct comparison of cpc values would not be possible, but it would be a fair assumption that these values should be broadly similar within the same cell line. The membranes were blocked together before being cut in half, each half containing one untreated and one TNFα treated samples. These were then processed in parallel through primary and secondary antibody incubation and wash steps before being imaged. This process was repeated for all NF-KB antibodies apart from IκBε, which has been shown not to detect NFκB-DOSCAT-2.

Using this experimental set-up NFκB-DOSCAT-2 and endogenous protein was successfully detected by all four of the antibodies that were used. An exemplar blot using a p65 antibody to detect NFκB-DOSCAT-2 and endogenous protein TNFα treatment is displayed in Figure 4.27. Using Image Studio software (LI-COR), densitometry was performed on bands and calibration curves were constructed based on the NFκB-DOSCAT-2 signal (Figure 4.28). For IκBα and IκBβ, signal saturation meant that for certain replicates no signal could be recorded and so these data points in the standard curve were omitted. Signal linearity was mostly excellent with r^2 values of 0.97 and above, although in two cases r^2 values were 0.95 and 0.80 for IκBβ and IκBα respectively. Using these calibration curves the signal for endogenous protein was converted to copy per cell values. Across all samples, average intra-assay variability in terms of CV was 19.2%, however this ranged from 6.6% to 45.6%. Technical repeats and biological replicates were averaged to give final quantification values for each protein (Figure 4.29a). Variability between biological replicates of nearly all samples was very high (CV > 40%); only IκBβ and IκBα post TNFα treatment had CVs < 20%. This increased variability was attributed to biological variation, as values in sample 5 were consistently inflated compared to sample 4. Values obtained by fluorescent QWB were compared to those

obtained by automated capillary QWB (Wes) and SRM-MS (Figure 4.29b). For p65, RelB and IκBβ, quantification values were higher than those previously recorded, however the large variance for p65 and RelB made it difficult to accurately assess how well the values agreed. Copy per cell values for IκBα agreed well with Wes QWB, but again the error associated with the measurement made it difficult to draw firm conclusions.

This study establishes that DOSCATs can be used in a classic western blot to quantify proteins with a good degree of accuracy. Although classic western blotting is not generally seen as quantitative, these results demonstrate that with proper calibration, quantification is achievable with values that broadly match those generated by gold standard platforms. Not having access to the same sample set used in previous quantitative experiments is a clear limitation in assessing quantitative accuracy, and this is compounded by the large variability in results. Repeating this study with an increased number of samples and performing repeats of the same gel would allow for a better evaluation of accuracy and variability.

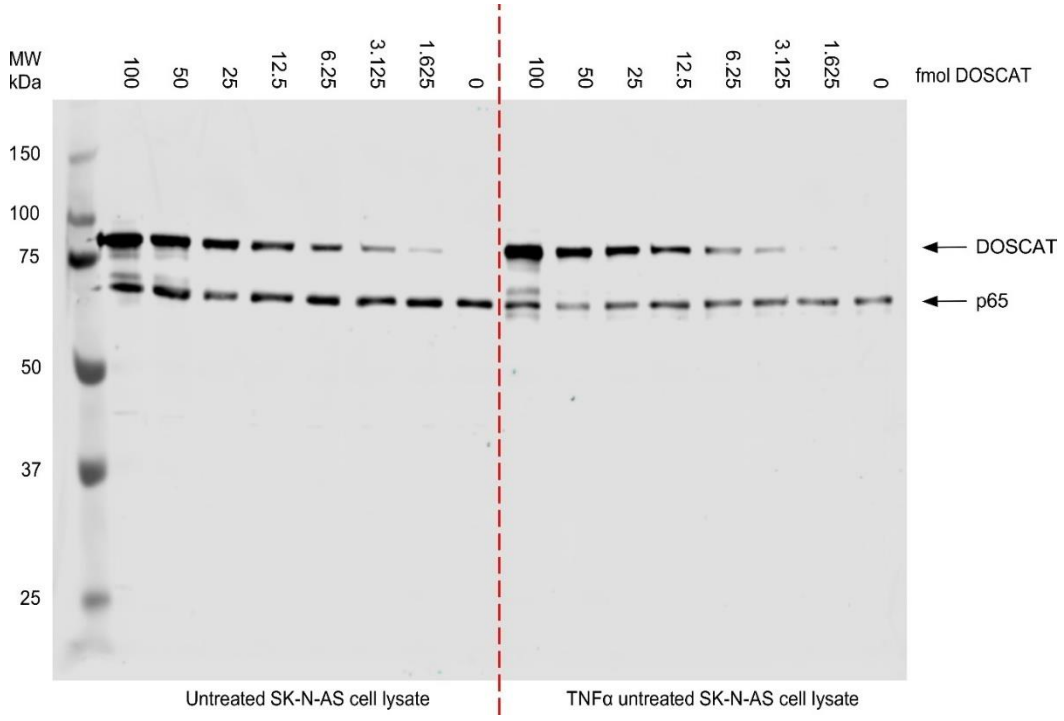


Figure 4.27 Quantitative western blotting using NFκB-DOSCAT-2 and fluorescent detection. Dilution series of NFκB-DOSCAT-2 spiked into 5 µg SK-N-AS cell lysate analysed by classic western blotting, using an anti p65 primary antibody and a fluorescently labelled secondary antibody that was detected by Odyssey Fc scanner (LI-COR).

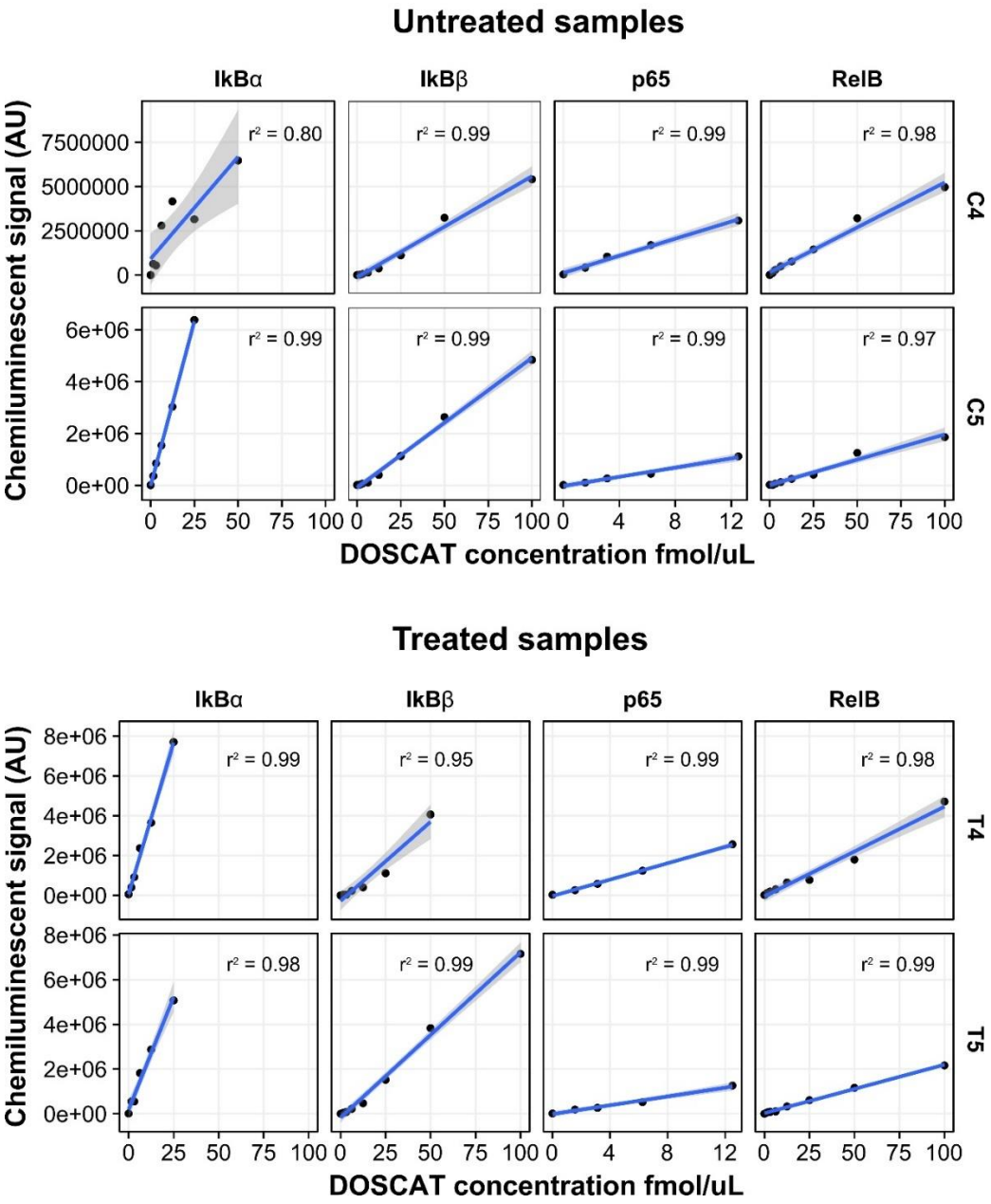


Figure 4.28 NFκB-DOSCAT-2 calibration curves in fluorescent QWB. Dilution series of NFκB-DOSCAT-2 standard spiked into yeast lysate was analysed as an external standard to SK-N-AS cell lysate samples by automated capillary western blotting, using four target antibodies against NFκB-DOSCAT-2. Data presented as mean ± standard error for each sample (n = 3).

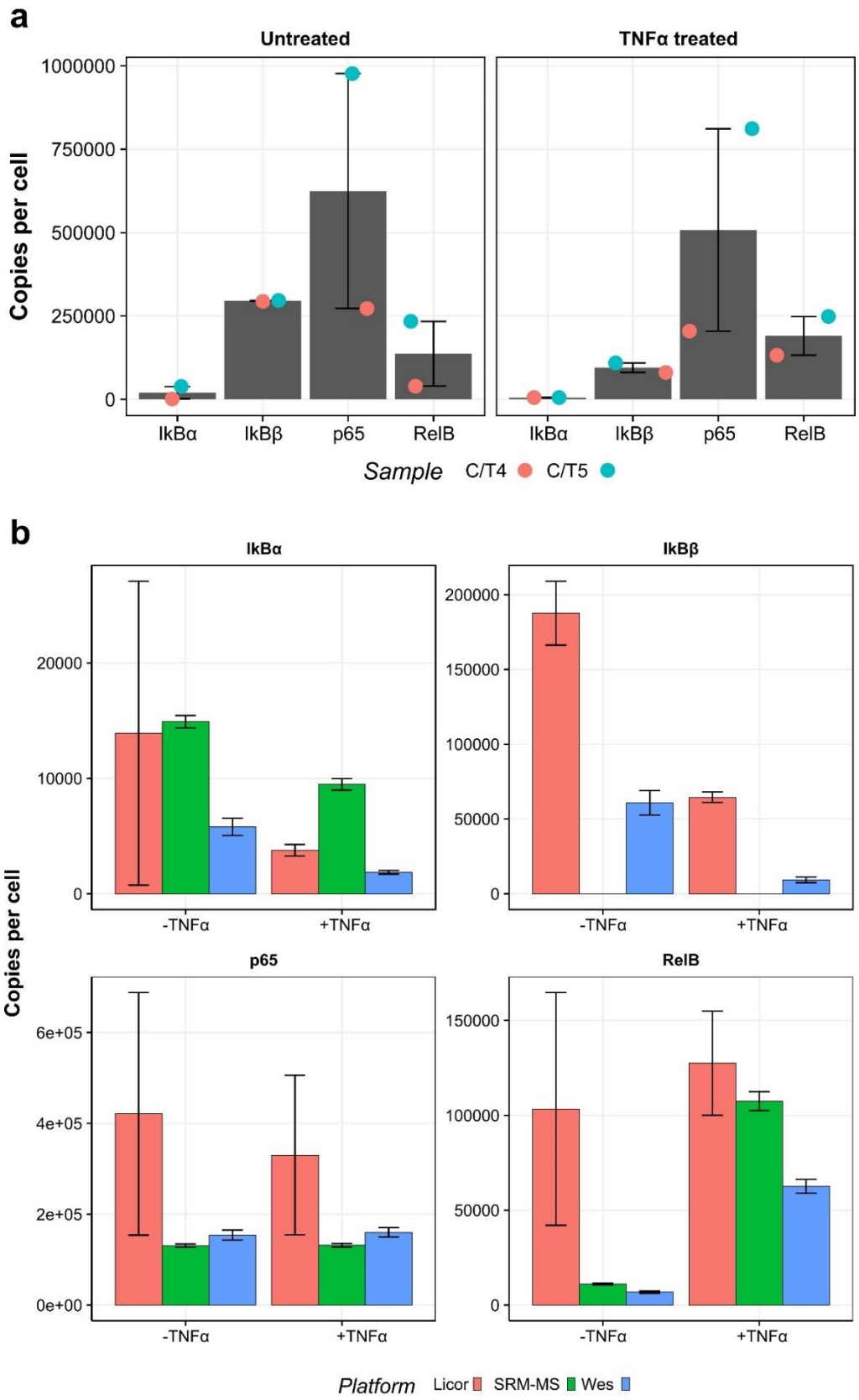


Figure 4.29 Quantification of target protein by fluorescent QWB. a) Quantification values for target proteins derived from fluorescent QWB. Data in bars are mean of biological replicates ± standard error (n = 2). Data points are mean values from technical repeats of individual biological replicates (n = 8). b) Copy per cell values from fluorescent QWB compared to values resultant from automated capillary WB and SRM-MS analysis.

4.3.5. **Restricted proteolysis of NFκB-DOSCAT-2**

This iteration of NFκB-DOSCAT-2 contained four protease cleavage sites (Table 4.3), each of which was assessed for specificity and efficiency of cleavage. NFκB-DOSCAT-2 was incubated with each of the proteases and samples taken at defined time points, which were subsequently analysed by SDS-PAGE (Figure 4.30). As with NFκB-DOSCAT, both RV3C and TEV proteases cleaved NFκB-DOSCAT-2 specifically after 4 hours, yielding two fragments of a distinct and predictable molecular weight. TVMV protease also specifically cleaved DOSCAT, yielding fragments as 67.5 kDa and 8.0 kDa. However, NFκB-DOSCAT-2 was not totally digested after overnight incubation, even when increasing amounts of protease well above the recommended protease:protein ratio were used. Incubation with enteropeptidase resulted in total digestion of NFκB-DOSCAT-2 within 1 hour with no proteolytic fragments visible on the gel at any time point. Only the contaminant that is present in the starting material is visible post-digestion. The heating of an enzyme during SDS-PAGE sample preparation may accelerate its activity, even at the high temperatures used. To investigate whether this was happening with enteropeptidase, the enzyme was incubated with NFκB-DOSCAT-2 and samples subjected to a TCA precipitation. Precipitate was then resuspended in SDS sample buffer and analysed by SDS-PAGE (Figure 4.31). As before, no bands for NFκB-DOSCAT-2 or expected proteolytic fragments were visible at the time points sampled. This discounted the hypothesis that enteropeptidase was being rendered hyper-active by heating. As specific cleavage could not be achieved, enteropeptidase was discontinued from further experiments.

Three out of the four proteases had been shown to cleave NFκB-DOSCAT-2 specifically, producing two fragments of differing electrophoretic mobility that could be resolved on an SDS-PAGE gel. The next step was to test whether these fragments could be detected by antibodies in western blots. NFκB-DOSCAT-2 was digested with each protease, and at each time point a sample was removed, diluted and prepared for Wes analysis. Six time points were taken for each protease and analysed by the five NF-KB antibodies (Figure 4.32). For TEV and RV3C proteases, antibodies successfully detected fragments that contained their cognate epitope. Fragments of a lower molecular weight had a lower signal intensity than larger fragments. For TVMV, the 67.5 kDa was detected by the correct antibodies but the 8 kDa fragment was not detected by RelB, most likely as it lies outside of the analytical range of the instrument (12 – 240 kDa). The RelB antibody also detected intact NFκB-DOSCAT-2, suggesting that digestion was incomplete in this case.

Compared with the earlier NFκB-DOSCAT, there is a gain in the introduction of protease sites. TEV and RV3C proteases worked well, consistent with results observed with the NFκB-DOSCAT, as did the newly introduced TVMV protease. TVMV is closely related to TEV protease and previous studies have demonstrated its high specificity (Nallamsetty et al., 2004; Sun et al., 2010), consistent with the results presented here. The totality of proteolysis by enteropeptidase would initially suggest the presence of a protease contaminant, due to the

high specificity of the canonical cleavage sequence DDDDK. However, using an approach in which cellular libraries of peptides substrates were screened, enteropeptidase has been shown to have a broader than thought substrate specificity with a strong preference for arginine at P1 and one or more Asp or Glu residues at P2 or P3 (Boulware et al., 2006). This would introduce many more potential sites of proteolysis in the DOSCAT sequence and may explain these results.

It has been demonstrated that epitopes contained in different proteolytic fragments can be detected by antibodies in a western blot, with the expected shift in electrophoretic mobility. It would be expected that the chemiluminescent signal for an intact DOSCAT and a proteolytic fragment would be the same as there is an equivalent molar amount of epitope loaded on to the system; however, this was not the observed result. The reasons for this are not clear, antibodies may bind differently to proteolytic fragments than the full length DOSCAT sequence, or the results may be an artefact of using the Wes system for analysis. Further work that replicated the experiment on different DOSCATs and using classic western blotting on proteolytic fragments might help to explain the data presented here.

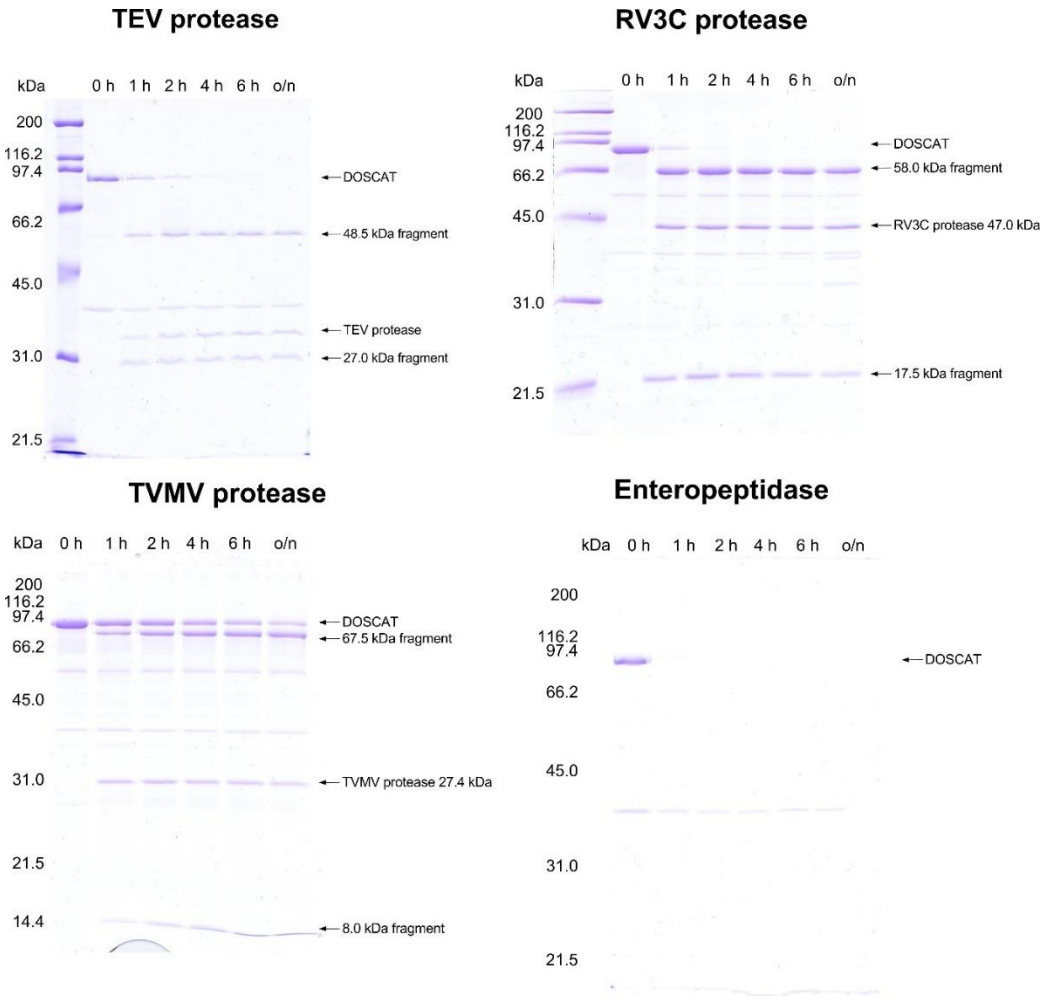


Figure 4.30 Digestion of DOSCAT by restricted specificity proteases. DOSCAT (9 µg) was incubated with each protease, samples taken at indicated time points and analysed on a 12% (TEV protease, enteropeptidase) or 15% (RV3C, TVMV proteases) SDS-PAGE gels.

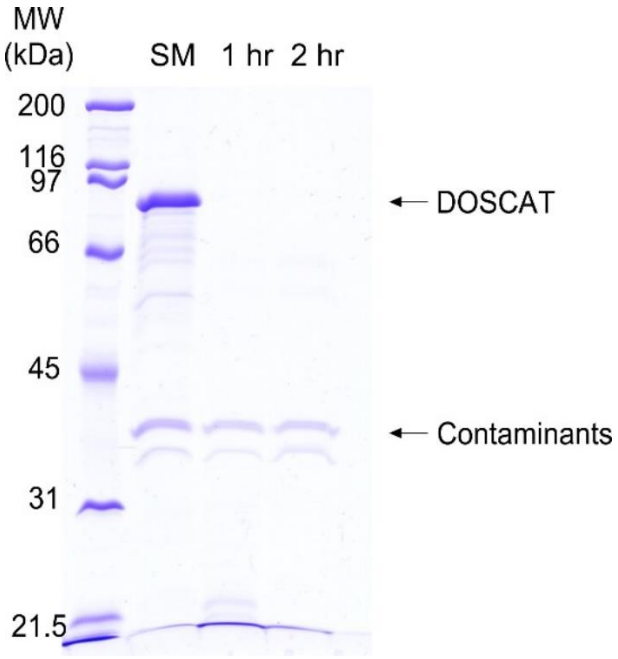


Figure 4.31 TCA precipitation of enteropeptidase proteolysis. DOSCAT was incubated with TVMV protease in the appropriate buffer. Starting material (SM) and samples at indicated time points were added to an equal volume of 10% TCA. Precipitate was pelleted by centrifugation and washed with acetone twice, acetone evaporated and pellet resuspended in SDS sample buffer for SDS-PAGE analysis.

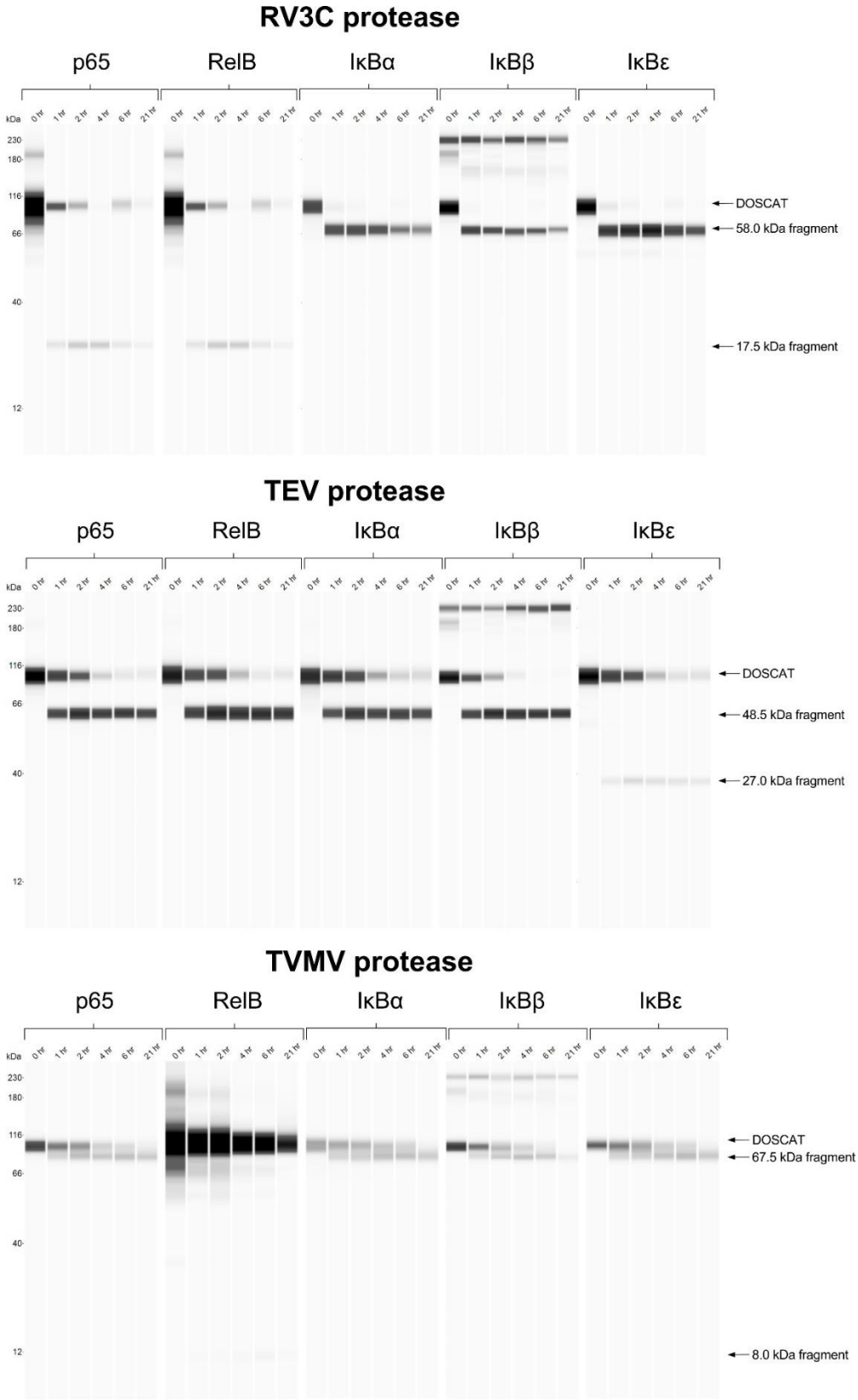


Figure 4.32 Western blots of DOSCAT proteolytic fragments. DOSCAT was incubated with each protease, samples taken at indicated time points and analysed by automated capillary western blotting using each of the NF-κB antibodies. Presented are pseudo-gel views of data generated by Compass software.

4.4. Conclusions

The work in this chapter builds on the initial proof of principle work by demonstrating that DOSCATs can be utilised to accurately and reproducibly quantify multiple target proteins by both QWB and SRM-MS, with quantification values measured by each technique being in good agreement. This shows not only that the DOSCAT approach is viable, but both automated and classic western blotting, generally only considered semi-quantitative, are as analytically rigorous as the current gold standard for targeted quantification. There is also flexibility in how DOSCATs can be used, whether spiked-in with a sample or external to the sample. Additionally, more has been understood about the properties of DOSCAT, such as its propensity to adsorb to surfaces, which will inform how experiments and workflows should be designed to minimise inaccuracies arising from erroneous DOSCAT concentration. This all represents a successful deployment of DOSCAT technology to fulfil the aims of absolute targeted protein quantification by MS and immunoblotting.

Following on from the NF-κB DOSCAT, the next stage in development will be to design a new DOSCAT that will quantify a different set of target proteins in an expanded sample size. Using all that has been learnt from this deployment of DOSCAT, it should be anticipated that the next deployment will have minimal issues and perform equally in terms of experimental accuracy and precision. This will give further assurance of the technique; reinforcing and refining design principles, increasing understanding of DOSCAT behaviour at a protein level, and having greater confidence in the accuracy and reproducibility of the orthogonal techniques.

Chapter 5: DOSCAT technology to quantify of putative pneumococcal meningitis biomarkers

5.1. Introduction

The previous chapter detailed how a DOSCAT could be designed, expressed and utilised to quantify target proteins using SRM-MS and automated capillary QWB, yielding results that were comparable in accuracy and reproducibility. Whereas this work quantified well studied proteins derived from a cell line, there is potentially a lot of value in using DOSCAT technology to aid in the rapid quantification of disease biomarkers in complex clinical samples. The advent of label-free quantitative proteomics has led to a sharp rise in the identification of putative biomarkers for many disease areas, however, few of these biomarkers are translated into the clinic for routine testing due to lack of robust verification and validation studies. The next stage in the advancement of DOSCATs is to use the technology to quantify potential biomarker proteins in a complex sample, further demonstrating how the methodology can be employed to improve data quality and existing practices.

This chapter will focus on the targeted quantification of proteins implicated in the pathogenesis of acute bacterial meningitis (ABM) caused by *Streptococcus pneumoniae* infection. Pneumococcal meningitis has a poor prognosis, and is particularly prevalent in sub-Saharan Africa where co-infection with HIV further increases the fatality rate (Nyasulu et al., 2011). Indeed, *Streptococcus pneumoniae* infection leading to pneumonia, septicaemia and meningitis is estimated to cause approximately 1 million deaths in children under the age of 5 years every year (O'Brien et al., 2009). Current diagnostic testing is based on Gram stain and culture of CSF obtained by lumbar puncture from patients suspected of meningitis, with diagnosis also based on several biochemical indicators including an increased white blood cell count and increased protein concentration in CSF (Scarborough et al., 2008). However, such tests are time-consuming, and as rapid diagnosis and treatment is associated with increased survival, there is a clear need for a point of care diagnostic test based on a panel of protein biomarkers. Previous studies have employed 2D-electrophoresis based proteomics of CSF to identify several proteins associated with pneumococcal meningitis (Cordeiro et al., 2015; Goonetilleke et al., 2010; Jesse et al., 2010). However, 2D-electrophoresis can lack sensitivity compared to more contemporary LC-MS techniques. More recently, MS based label-free quantitative proteomics was used to identify a number of proteins that were differentially expressed in the CSF of patients diagnosed with pneumococcal meningitis. A subset of these proteins was further analysed by automated capillary western blotting using recombinant protein standards, which confirmed the proteomics data and offered a panel of putative protein biomarkers; cathelicidin, cystatin C, ceruloplasmin, myeloperoxidase, S100A8 and S100A9. Other than cystatin C, abundances of these proteins increased in CSF from patients with pneumococcal meningitis infection. However, taken in isolation, such changes in

protein abundance are not clinically valuable in determining *Streptococcus pneumoniae* infection as most are involved in the innate immune response and reflect a non-specific antimicrobial response. They would, therefore, be upregulated in any bacterial infection and so not be specific to infection by *Streptococcus pneumoniae*. Myeloperoxidase is essential for the rapid activation of antimicrobial activity in neutrophils upon detection of pathogens through catalysing the generation of hypochlorous acid from hydrogen peroxidase and chloride anion (Nauseef, 2014). Cathelicidin is an inactive precursor protein that is cleaved by neutrophil elastase into antimicrobial peptides that are released from neutrophils when required (Kościuczuk et al., 2012; Treffers et al., 2005). S100A8 and S100A9 form a heterodimer known as calprotectin, which by sequestering the essential nutrients manganese and zinc at sites of infection restricts the ability of pathogens to establish themselves in a host (Brophy et al., 2015; Corbin et al., 2008). Ceruloplasmin is a ferroxidase that has been implicated in the acute phase response to inflammation, although its exact physiological role is not well understood (Gabay et al., 1999; Jain et al., 2011). When analysed together in a panel, it is hoped that these proteins along with cystatin C (which is down regulated upon infection) will aid the diagnosis of *Streptococcus pneumoniae* infection with the specificity and sensitivity required for a clinical assay.

The aim of this chapter is to use these six proteins as the targets to be quantified using a new DOSCAT, again utilising SRM-MS and automated capillary QWB as platforms for quantification. To achieve this, over 40 CSF samples from patients who tested positive and negative for pneumococcal meningitis will be used; such an expanded sample set will permit a more robust comparison of quantitative performance between the two assays as well as demonstrating that DOSCATs can be used in complex samples such as CSF. Moreover, quantitative data generated using a DOSCAT can be directly compared to label-free proteomics and QWB data generated in the previous study to further validate the approach.

A weakness with the DOSCAT approach is successfully selecting well validated antibodies with defined epitopes that will recognise DOSCATs with the epitope built in to the sequence. Although there are now resources available to aid the selection of antibodies that are validated and likely to specifically detect endogenous proteins (Björling et al., 2008; Helsby et al., 2014), manufacturers do not always disclose epitopes, and when they do, they may not always be accurate. As discussed in the previous chapter, selected antibodies did not detect the NFκB-DOSCAT even though they detected endogenous protein. There is no way to know if the antibody will detect a DOSCAT until after it is expressed, and so such failure represents a waste of resources and a reduction in experimental output. This limitation could be overcome if the epitope sequences built into a DOSCAT were used as an immunogen to create an anti-DOSCAT antibody. Previously, recombinantly expressed epitopes fragments (PrESTs) have been used as immunogens to produce monospecific antibodies for affinity proteomics in multiplex by immunising rabbits with up to 10 PrESTs at once (Larsson et al., 2006). The work

in this chapter aims to build on this by creating designer proteins that concatenate epitope sequences into a single protein that can be expressed recombinantly and used to immunise a single rabbit. This would yield a polyclonal serum of antibodies that would recognise the epitopes built into a cognate DOSCAT standard, with the added benefit of reducing the number of animals required, as a single animal would produce the antibody to detect all epitopes. This chapter will attempt to demonstrate the viability of this strategy through the design and expression of an immunogen that concatenates the epitopes used in the DOSCAT standard against the six protein biomarkers. The immunogen will be used to produce an anti-DOSCAT antisera of which the specificity and sensitivity to DOSCAT standard and endogenous proteins in western blotting will be evaluated.

5.2. Results and discussion

5.2.1. Preparation of a meningitis-DOSCAT standard and meningitis-DOSCAT immunogen

Two new constructs were built for this work: a DOSCAT targeting six proteins implicated in the pathophysiology of paediatric *Streptococcus pneumoniae* meningitis infection and a protein incorporating the chosen epitopes to act as an immunogen to create an anti-DOSCAT antibody. The standard and immunogen will subsequently be referred to as M-DOSCAT-S and M-DOSCAT-i respectively. Design principles refined in the previous iterations of NFκB-DOSCATs were employed in the design of M-DOSCAT-S. A shortlist of candidate Q-peptides for each protein were selected using the previously described criteria, and tools such as CONSEQUENCE and SRM-Atlas were also used to refine the shortlist so that two or three Q-peptides for each protein were selected (Table 5.1). Each selected peptide was present in SRM-Atlas (Kusebauch et al., 2016) as a peptide that had been observed previously in SRM experiments. Each peptide was assessed for potential sites of post-translational modifications using NetPhos (Blom et al., 1999) and PhosphoSite plus (Hornbeck et al., 2017) and subsequently cross-checked against selected epitopes to ensure that antigenic regions were not replicated throughout the sequence.

Table 5.1 M-DOSCAT-S peptides and transitions. SRM transitions (for light peptides), collision energies and dwell times for the analysis of each target peptide contained in M-DOSCAT-S. Natural flanking sequences of peptides are shown in red, where a flanking sequence is not displayed the Q-peptide is adjacent to another Q-peptide.

Protein	Peptide sequence	Precursor ion (m/z)	Measured product ions (m/z)	Charge	Collision energy (eV)	Dwell time (msec)
Cathelicdin	DKDNKR**FALLGDFFR**KSKEKI	543.3	y7 (867.5), y6 (754.4), y5 (641.3)	2	19	101
	KEAVLR**AIDGINQR**SSDANL	443.7	y5 (569.3), y4 (432.3), y3 (345.2)	2	15	66
	GMKTQR**DGHSLGR**WSLVLL	371.2	y6 (702.4), y5 (587.3), y3 (417.2)	2	13	205
Ceruloplasmin	IDIFTK**ENLTAPGSDSAVFFEQGTTR**IGGSYG	709.7	y9 (1084.5), y8 (985.5), y7 (838.4)	3	24	48
	GRLYKK**ALYLQYTDETFR**TTIEKP	760.4	y8 (1059.4), y7 (931.4), y6 (768.4)	2	27	48
	GITYYK**EHEGAIYPDNTTDFQR**ADDKVY	631.6	y9 (1093.5), y7 (881.4), y5 (666.3)	3	21	66
Cystatin C	EEGVRR**ALDFAVGEYNK**ASNDMY	613.8	y7 (780.4), y6 (709.4), y5 (610.3)	2	22	48
	VVRARK**QIVAGVNYFLDVELGR**TTCTK	897.0	y13 (1452.7), y12 (1381.7), y10 (1225.6)	2	32	101
	TQPNLDNCPFHDQPHLKRKAFCS	687.7	y11 (1392.6), y9 (1118.6), y6 (737.4)	3	23	66
Myeloperoxidase	DSVDPR**IANVFTNAFR**YGHTLI	576.8	y8 (968.5), y6 (755.4), y5 (608.3)	2	20	48
	QISLPR**IICDNTGITTVSK**NNIFMS	711.4	y11 (1195.6), y10 1035.5), y9 (920.5)	2	25	66
	RVPSLR**VFFASWR**VVLEGG	456.7	y6 (813.4), y5 (666.3), y4 (519.3)	2	46	48
S100A8	LTELEK**ALNSIIDVYHK**YSLIK	424.9	y6 (774.4), y5 (661.3), y3 (447.2)	3	14	48
	GNFHAVYRDDLKKL	482.2	y5 (645.3), y4 (508.3), y3 (437.3)	2	17	101
S100A9	KELVRK**DLQNFLK**KENKNE	439.2	y5 (649.4), y5-18 (637.4), y4 (521.3)	2	15	48
	MSQLER**NIETIINTFHQYSVK**LGHPDT	603.0	y10 (1236.6), y9 (1123.6), y6 (761.4)	3	20	48

Antibodies were selected primarily on whether data relating to their epitopes were available (Table 5.2). Each selected antibody was raised using a known synthetic peptide as an immunogen and the sequences of these peptides were inserted into M-DOSCAT-S as the epitopes. By using this more specific approach rather than inserting epitopes centred around an amino acid, it was hoped that antibodies would be more likely to recognise epitopes inserted into the DOSCAT sequence. Compared to NFκB-DOSCAT-2, a higher proportion of selected antibodies were monoclonal (3/6), which is preferred as they will bind to a specific epitope only, and so non-specific binding to other regions within the DOSCAT will be less likely. Before committing to synthesising the gene for M-DOSCAT-S, the selected antibodies were tested against recombinantly expressed target proteins using automated capillary WB (Figure 5.1). All target proteins except for cystatin C were detected, although only a low signal for the light chain of myeloperoxidase was observed. Increasing amounts of protein loaded onto the Wes system with higher concentrations of antibody did improve the signal for myeloperoxidase, but cystatin C was still not detected. Cystatin C is an important target protein as it is the only protein that increases in abundance in healthy control samples, so it would be advantageous for the study if it was kept in the analysis. At the time, no other antibodies with a known epitope could be found, but we were in possession of an antibody that was known to us to work against the recombinant standard and in CSF samples. As cystatin C is a small protein (15.8 kDa), it was decided to include the full protein sequence within M-DOSCAT-S. This would ensure that M-DOSCAT-S could be detected by WB, and would also facilitate the inclusion of Q-peptides for SRM-MS analysis. The same restricted specificity protease sites included in NFκB-DOSCAT-2 were selected except for enteropeptidase, which when previously tested resulted in the complete proteolysis of the DOSCAT protein. When assembling the sequence for M-DOSCAT-S, the locations of the restricted proteolytic sites were chosen to maximise multiplexing capability and to ensure epitopes were not at termini of fragments, which may affect antibody binding. As before, His-tag for purification and Glu-Fib peptide for quantification were included in the final sequence (Figure 5.2a).

M-DOSCAT-i was designed as an immunogen to create an anti-DOSCAT antibody that would detect epitopes contained in M-DOSCAT-S. Epitopes for four proteins, myeloperoxidase, ceruloplasmin, cathelicidin and S100A8, were taken from M-DOSCAT-S and included in the M-DOSCAT-i sequence along with a Met initiator at the N-terminus and His-tag at the C-terminus (Figure 5.2b). No spacer regions were inserted between epitopes as this may have produced non-specific antibodies.

Table 5.2 Antibodies used with M-DOSCAT-S.

Protein	Supplier	Product no.	Host species	Clonality
Cathelicidin	Abcam	ab58387	mouse	monoclonal
Ceruloplasmin	Abgent	AP7340C	rabbit	polyclonal
Cystatin C	Novus	NBP1-26402	goat	polyclonal
Cystatin C	Abcam	ab109508	rabbit	monoclonal
Myeloperoxidase	Abcam	ab45977	rabbit	polyclonal
S100 A8	Abcam	ab92331	rabbit	monoclonal
S100 A9	Abcam	ab92507	rabbit	monoclonal

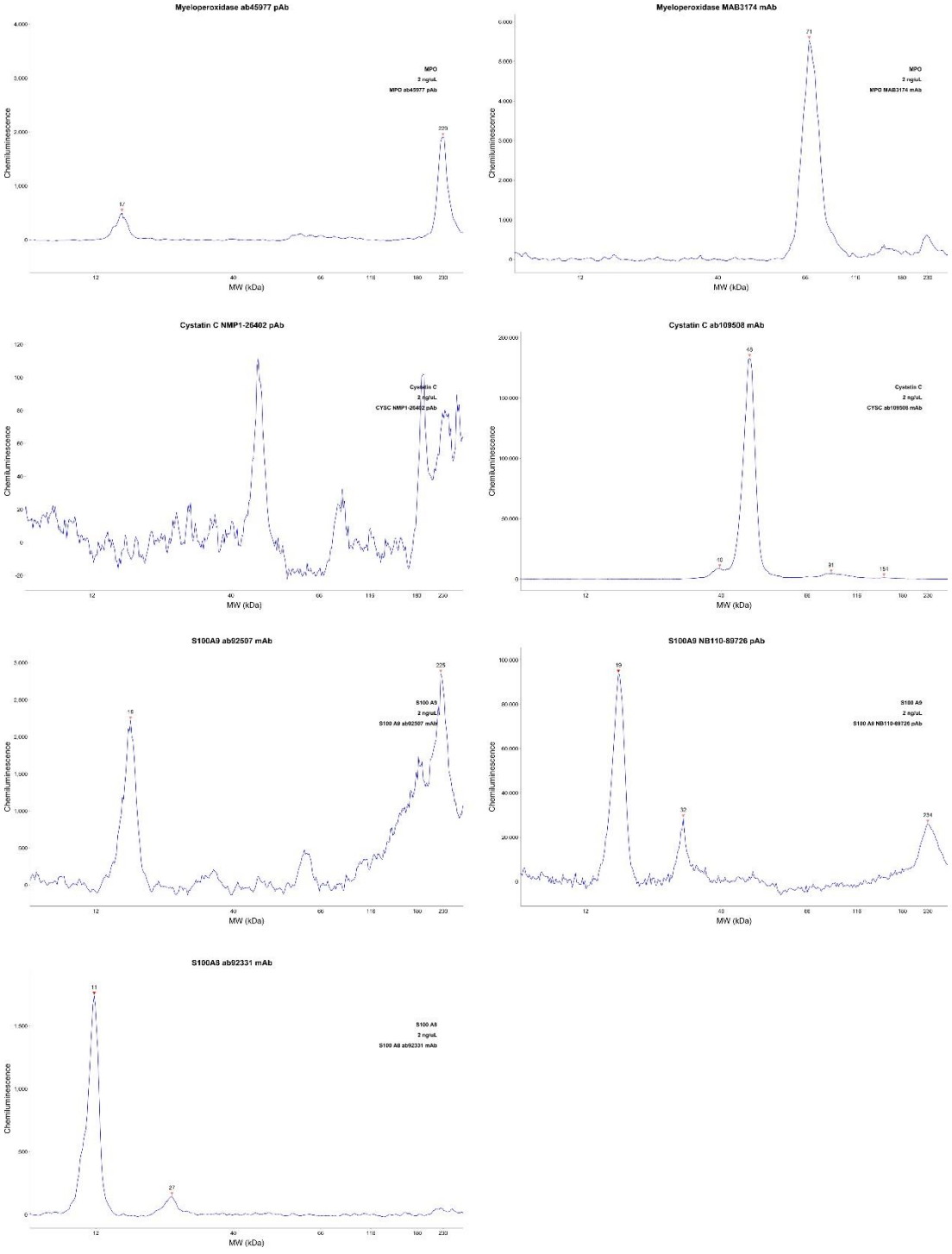


Figure 5.1 Validation of antibodies against recombinant standards. Electropherograms resultant from the capillary WB analysis of recombinant protein standards (loaded onto the Wes system at 0.2 µg/µL) with their cognate antibody used at a 1:50 dilution.

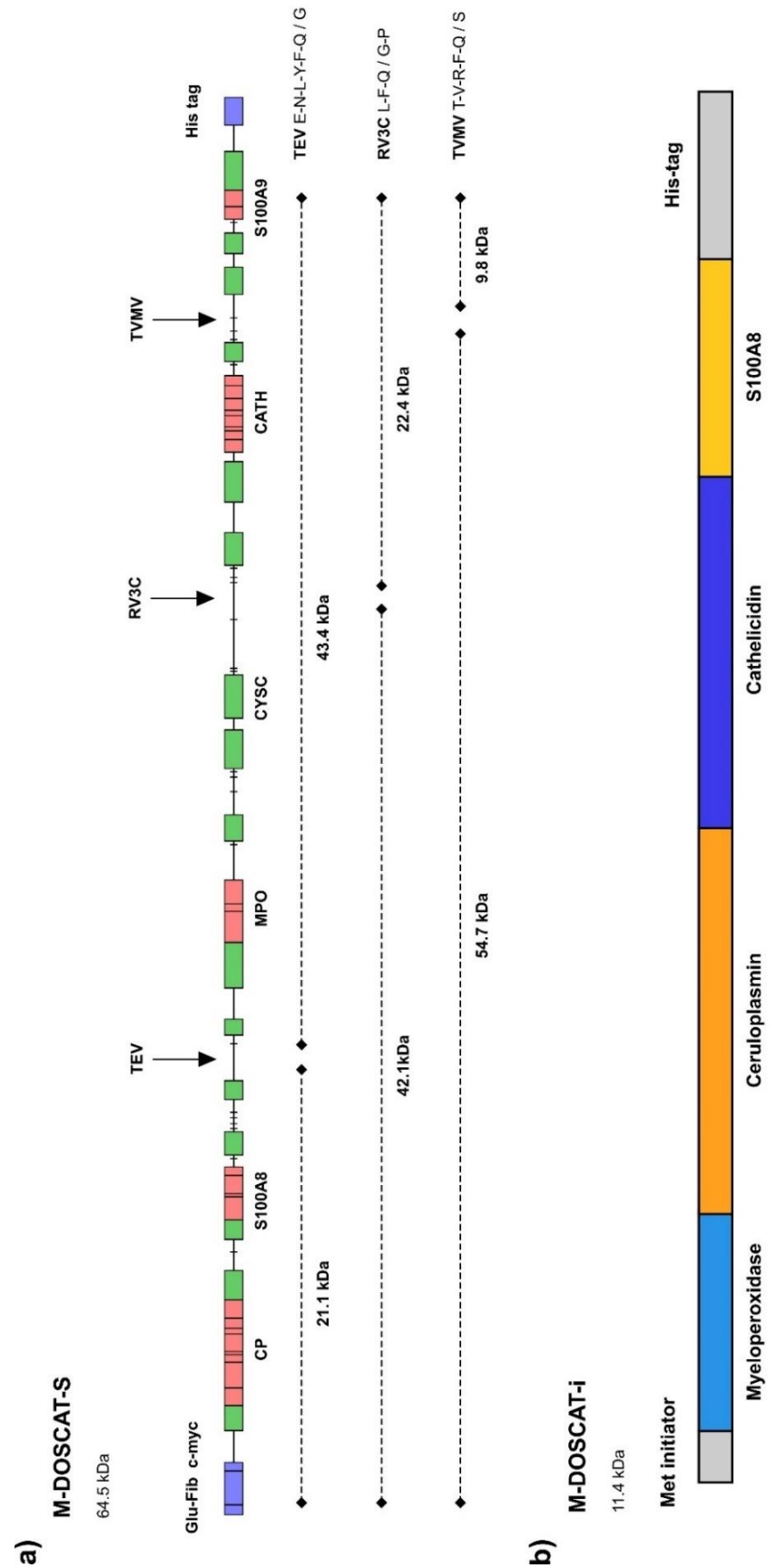


Figure 5.2 Design of meningitis DOSCAT standard and immunogen. (a) Protein map of M-DOSCAT-S. Green boxes represent quantotypic peptides, red boxes define the extent of antibody binding epitopes and purple boxes the [Glu1]-Fibrinopeptide B calibration peptide at the N-terminus and the His6 purification tag at the C-terminus. Arrows indicate the location of cleavage sites for each of the specific proteases and dotted lines the molecular weight of the fragments they generate upon cleavage (TVMV, tobacco vein mottling virus protease; RV3C, human rhinovirus 3 C protease; EP, enteropeptidase; TEV, tobacco etch virus protease. b) Protein map of M-DOSCAT-i. Epitopes for each protein (coloured boxes) are concatenated into a single sequence flanked by a Met initiator and His tag (grey boxes) at N and C termini respectively.

Codon optimised genes for both M-DOSCATs were synthesised, cloned into pET21a vectors (Eurofins Genomics) and transformed into BL21 DE3 *E.coli* cells. As the immunogen did not need to be stable isotope labelled, cells transformed with DNA for M-DOSCAT-i were grown in LB media at 37°C and expression induced by the addition of 1 mM IPTG when OD at 600 nm reached 0.6. After 3 h no expression of M-DOSCAT-i was evident (Figure 5.3a), so expression was attempted again using minimal media at 37°C, but again, no expression could be detected by SDS-PAGE (Figure 5.3a). It was possible that M-DOSCAT-i was degrading within the cell after being expressed, so in an attempt to prevent this *E.coli* cells were grown in LB media at 18°C overnight with different concentrations of IPTG to induce expression, however there was no indication of protein expression under any of the conditions (Figure 5.3b). Five different *E.coli* cell lines were then used: SL BL21, C43 (DE3), BL21(DE3)-pLysS (pLysS), Rosetta-pLysS (Rosetta) and Rosetta-gami-pLysS (Rosetta-gami). The pLysS strain expresses T7 lysozyme so that basal expression of genes controlled by the lac promoter is reduced and the C43 (DE3) strain contains mutations on the lac promoter that reduce levels of expression (Miroux et al., 1996; Wagner et al., 2008), thus making both strains more tolerant to protein toxicity. The Rosseta and Rosseta gami strains, developed by Novagen, supply tRNAs for the rare codons AUA, AGG, AGA, CUA, CCC, GGA, improving expression levels for proteins with these codons present. The M-DOSCAT-i plasmid was successfully transformed into each cell line, which were grown in LB media containing 1% glucose at 37°C with expression induced using 0.5 mM IPTG when OD_{600nm} reached 0.6 and culture was continued for 5 h. All cell lines grew as expected except for Rosetta Gami, which did not progress past lag phase (Figure 5.4a). In all cases, no expression of M-DOSCAT-i was evident (Figure 5.4b). Western blot analysis using a His-tag antibody detected a single protein expressed even before the addition of IPTG, although due to failure of the MW markers it is unclear if the visible band is at the correct MW for M-DOSCAT-i (Figure 5.4c). However, as the bands are present in sample pre-IPTG induction it is likely that they are indicative of a non-specific contaminant signal.

Cells transformed with the plasmid for M-DOSCAT-S were grown in minimal media containing [$^{13}C_6$]Arg/[$^{13}C_6$]Lys at 37°C and expression induced during mid-log phase by the addition of 1 mM IPTG. The M-DOSCAT-S protein was expressed after 3 h (Figure 5.5a) and subsequently determined to be contained in the insoluble fraction of the *E.coli* cell lysate. M-DOSCAT-S was purified by Ni-NTA affinity chromatography (Figure 5.5b) and dialysed into a 50 mM ammonium bicarbonate, 1 mM DTT, 0.1% Rapigest *SF* (*w/v)* storage buffer. Concentration was estimated by analysing the DOSCAT protein alongside a dilution series of BSA standards on an SDS-PAGE gel (Figure 5.5c). As well as further demonstrating the purity of the protein, the concentration was estimated to be 0.04 µg/µL, or 0.6 pmol/µL. Purified M-DOSCAT-S was digested with trypsin and analysed by MS/MS on an Orbitrap Elite. Sequence coverage was high (83%) with all Q-peptides observed (Figure 5.6a). Moreover, labelling efficiency of isotope

labelled Arg/Lys was shown to be > 95% (Figure 5.6b). When analysed by automated capillary WB, M-DOSCAT-S was detected by all target antibodies apart from S100A8 (Figure 5.7).

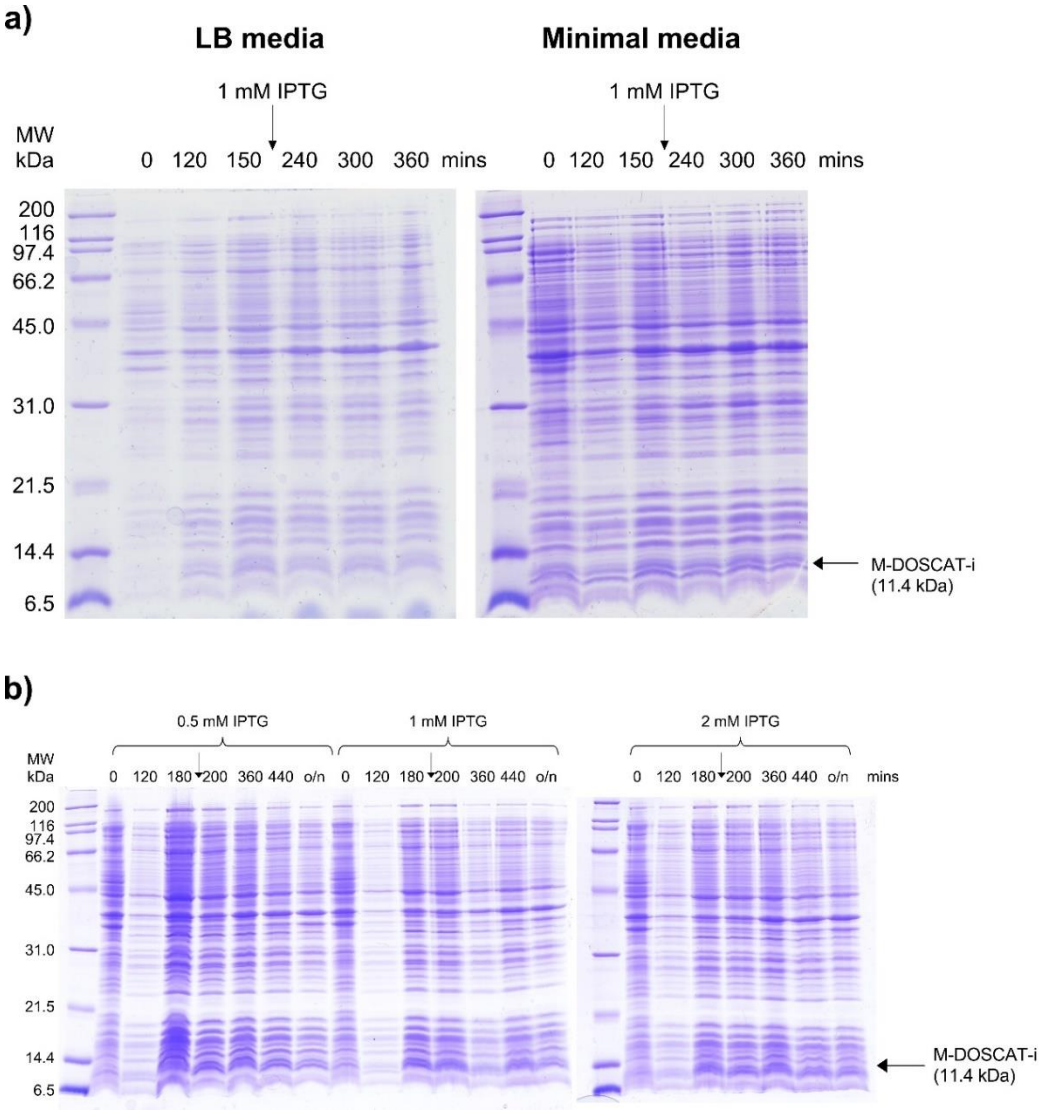


Figure 5.3 Expression of M-DOSCAT-i in BL21 *E.coli* cells. SDS-PAGE analysis of time points from *E. coli* culture containing M-DOSCAT-i plasmid, grown in a) LB and minimal media for 6 h after inoculation, with expression induced by 1 mM IPTG at indicated time points; b) minimal media at 18°C overnight, with expression induced by indicated concentration of IPTG 3 h after inoculation, as indicated by arrows.

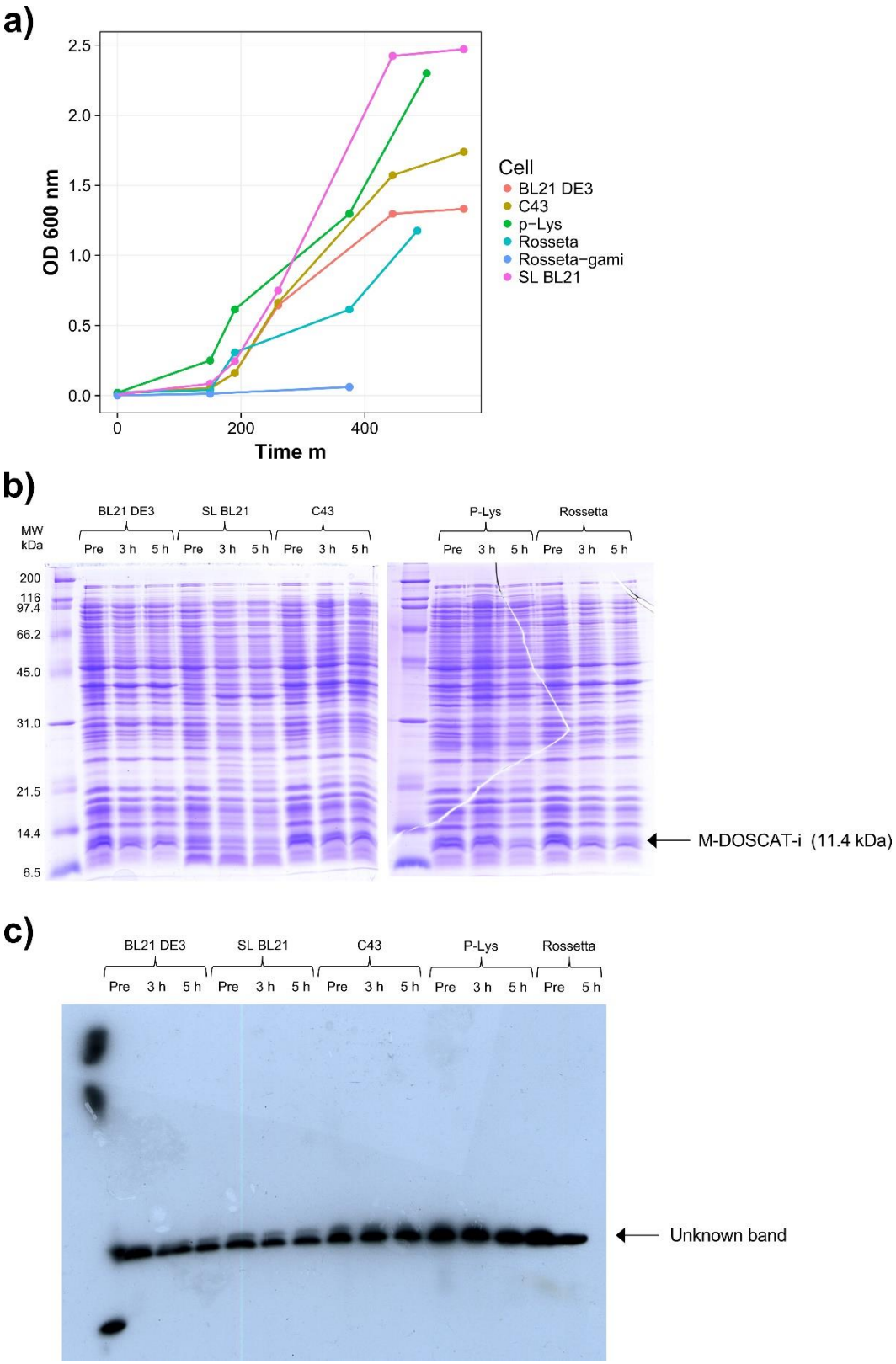


Figure 5.4 Expression of M-DOSCAT-i in multiple cell lines. a) Growth curves of each culture measured by optical density at 600 nm. Samples from each culture were analysed pre-induction by 0.5 mM IPTG (Pre) and up to 5 h after induction by b) SDS-PAGE and c) western blot using an anti His-tag antibody.

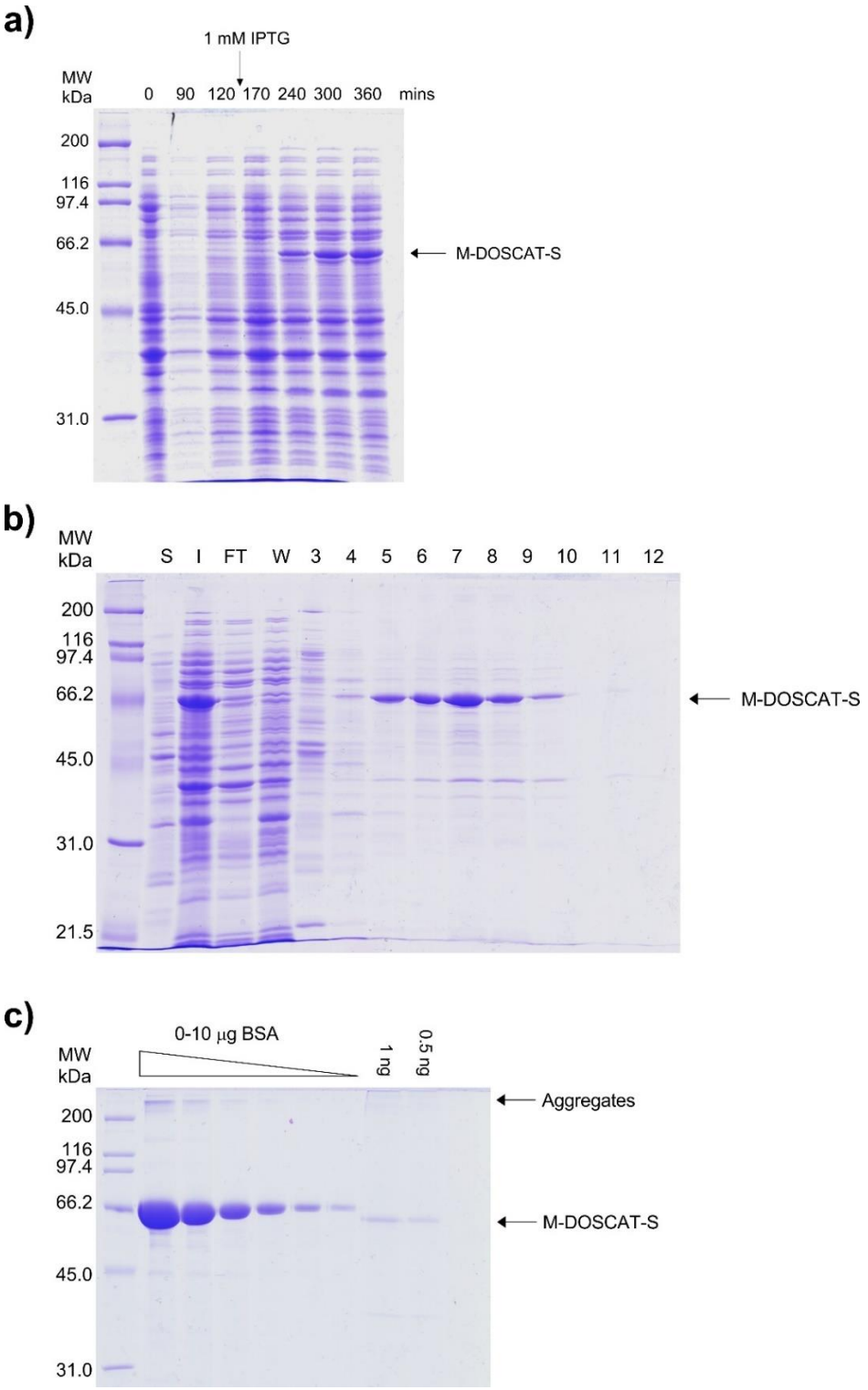


Figure 5.5 Expression and purification of M-DOSCAT-S. SDS-PAGE analysis of a) time points from *E. coli* culture containing M-DOSCAT-S plasmid, grown in minimal media containing [$^{13}C_6$]Arg/[$^{13}C_6$]Lys for 6 h after inoculation, with expression induced by 1 mM IPTG at indicated time point; b) soluble (S) and insoluble (I) fractions of *E.coli* cell extract, flow through (FT), wash fraction (W) and elution fractions 3 – 12 from His-Trap column using an elution gradient 0–100% elution buffer over 20 min; c) purified M-DOSCAT-S alongside albumin standards.

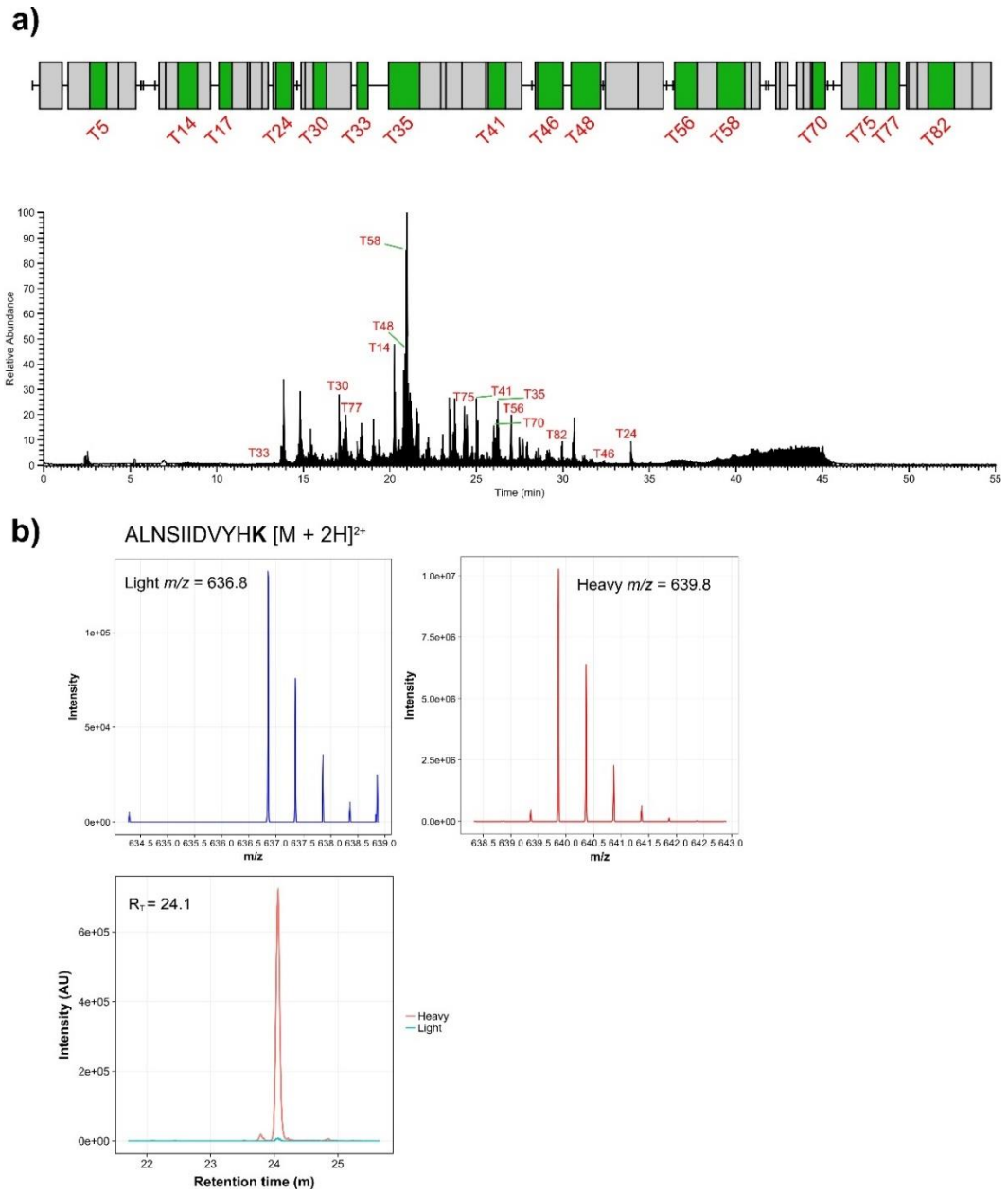


Figure 5.6 MS verification of M-DOSCAT-S expression. a) M-DOSCAT-S peptide map highlighting identified Q-peptides (green) and other peptides (grey) by MS/MS, alongside the MS1 total ion chromatogram signifying the elution profile of each Q-peptide. b) Precursor ion mass spectra from MS/MS analysis and SRM chromatogram for a representative Q-peptide, ALNSIIDVYHK, demonstrating high efficiency of stable isotope labelling.

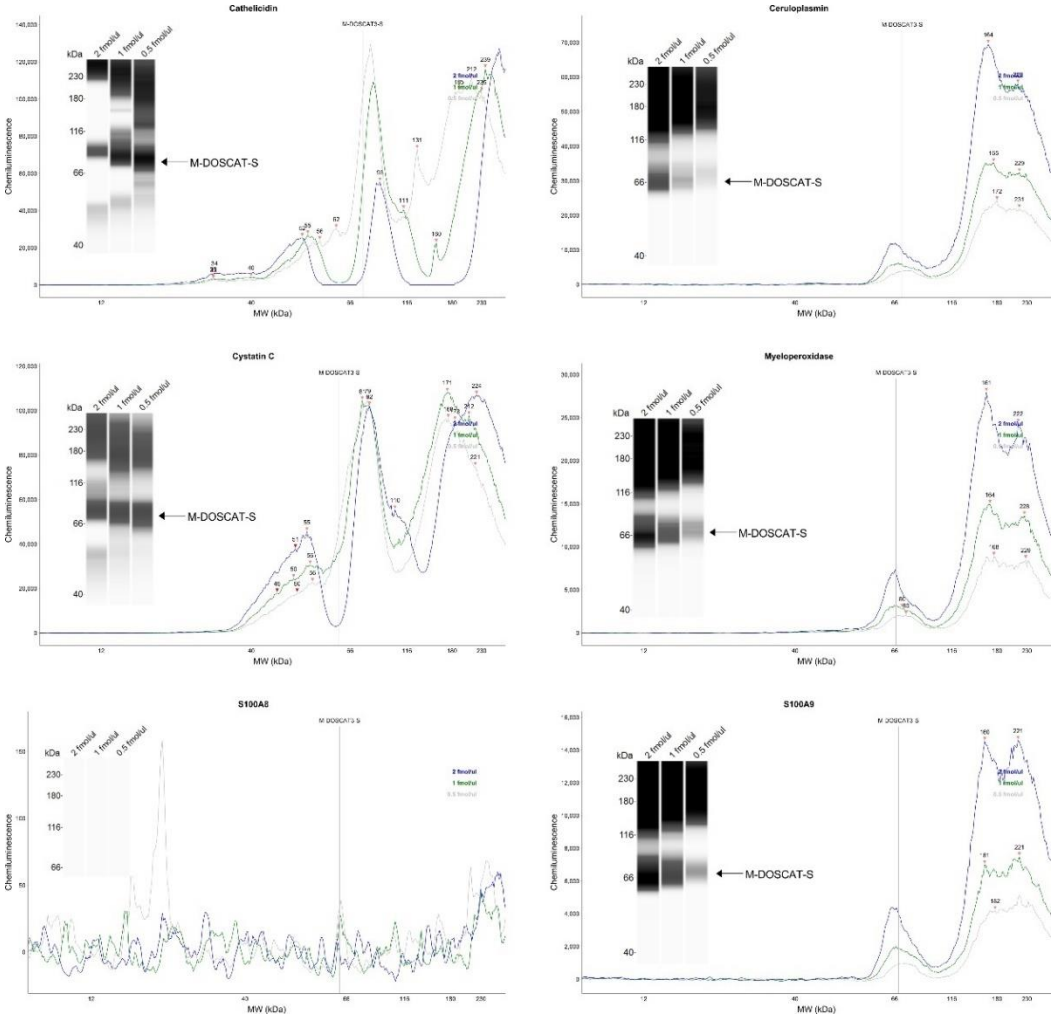


Figure 5.7 Detection of M-DOSCAT-S by target antibodies. Electropherogram and pseudo-gel images resultant from the capillary western blot analysis of M-DOSCAT-S loaded onto the Wes system at 2, 1 and 0.5 fmol/µL and analysed by target antibodies at a 1:50 concentration.

In all blots containing M-DOSCAT-S, non-specific bands were observed at 150 kDa that weren't present when CSF alone was run. When analysed by capillary WB without the presence of primary antibodies, no signal was observed, confirming that the bands were not caused by non-specific detection of *E.coli* proteins by secondary antibodies (Figure 5.8). Instead, their presence was hypothesised to be due to the formation of aggregates in the M-DOSCAT-S sample. Bands visible at high MW on an SDS-PAGE provided further evidence of this (Figure 5.5c). The presence of such aggregates could adversely affect QWB as bands for endogenous target proteins would be obscured and there would also be a discrepancy between MS quantification of M-DOSCAT-S stock and the signal measured by QWB. This is because aggregates would be digested during trypsin digestion and so be included in quantification using Glu-Fib peptide. Without knowledge of the proportion of M-DOSCAT-S to aggregates in a sample, it would be difficult to correct for this. Therefore, strategies were needed to overcome or prevent aggregation.

Aggregation can be the result of disulphide bond formation between cysteine residues, and so reduction of the disulphide bond may disrupt existing aggregates. Sample buffers for SDS-PAGE and western blotting include DTT, so if aggregates could be removed by modifying levels of DTT within the sample buffer this would provide a simple solution as it would just be an amendment to the existing WB workflow. Sample buffers containing normal (100 mM) and increased (400 mM) concentration of DTT were incubated with M-DOSCAT-S at different times and temperatures (Figure 5.9). There was no effect on signal intensity in either M-DOSCAT-S or aggregates between each sample preparation procedure, even with four times as much reducing agent used. This ruled out the approach of modifying sample buffer to reverse aggregation. The next approach taken was to attempt to remove aggregates using a molecular weight cut off (MWCO) filter. A spin filter with MWCO of 100 kDa was selected, so that M-DOSCAT-S would pass through the column but aggregates would not. This was not the case however, with both M-DOSCAT-S and aggregates being retained by the filter and no protein detected in the flow through (Figure 5.10). Efforts then moved on to prevent M-DOSCAT-S aggregating in the first place. It was observed that when solubilised in guanidinium hydrochloride after purification, no aggregates were present (Figure 5.5b). If disulphide bond formation could be prevented at this stage then aggregation might not occur. M-DOSCAT-S was purified from cell pellets and whilst still in elution buffer containing Gu-HCl, cysteine residues were reduced and alkylated by incubation with DTT and IAM respectively. After dialysis into storage buffer, no evidence of aggregation was observed in DTT/IAM treated M-DOSCAT-S when analysed by SDS-PAGE (Figure 5.11a) or automated capillary western blotting (Figure 5.11b). The completeness of cysteine modification was demonstrated by analysing M-DOSCAT-S by MS/MS (Figure 5.11c), and modified epitopes within recombinant standards were still recognised by all antibodies although the signal was reduced compared to unmodified protein (Figure 5.11d).

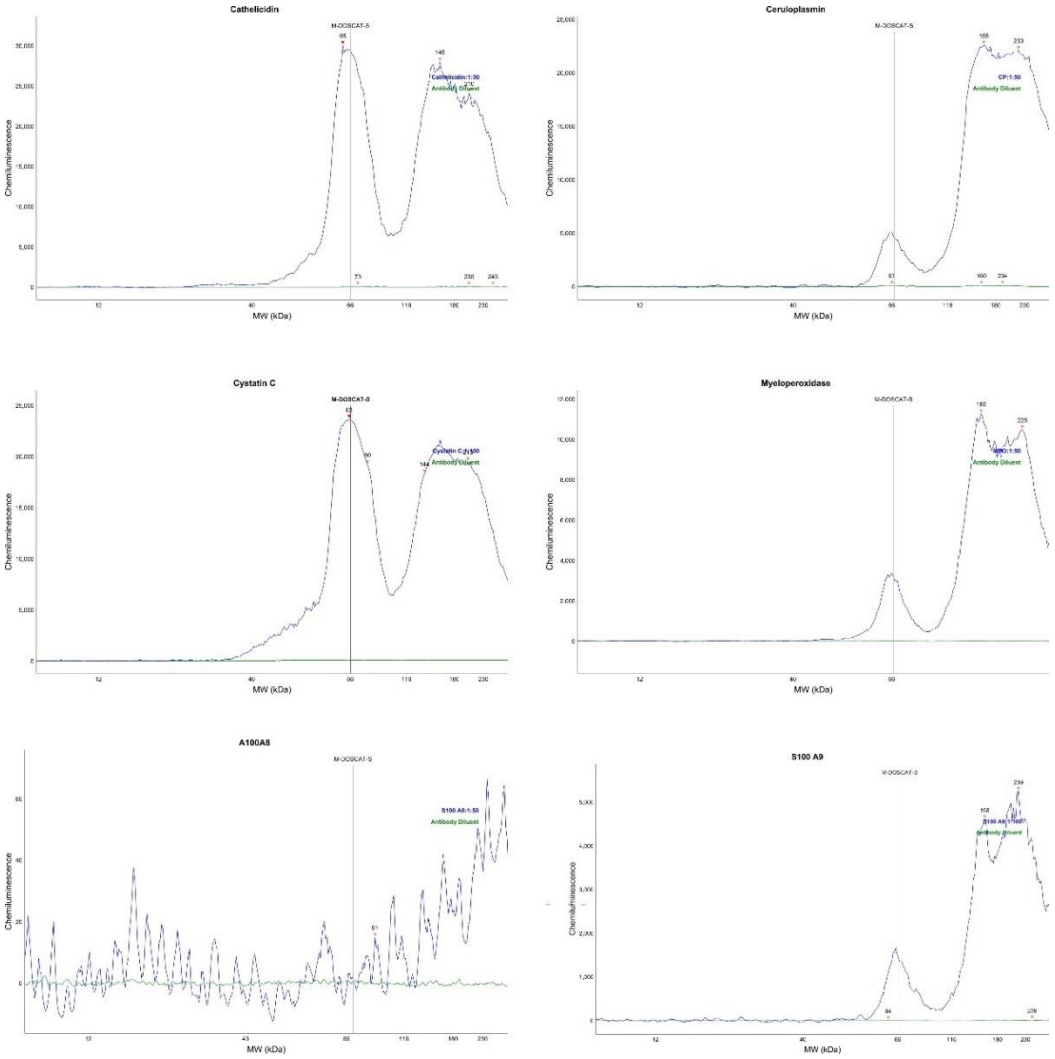


Figure 5.8 Analysis of secondary antibody detection of M-DOSCAT-S. Electropherograms resultant from capillary western blot analysis of M-DOSCAT-S (1 fmol/µL) using each of the target antibodies (blue traces) or blocking buffer (also referred to as antibody diluent, green traces) as the primary antibody. The secondary antibody was either HRP conjugated anti-rabbit or anti-mouse dependent on the source of the primary antibody.

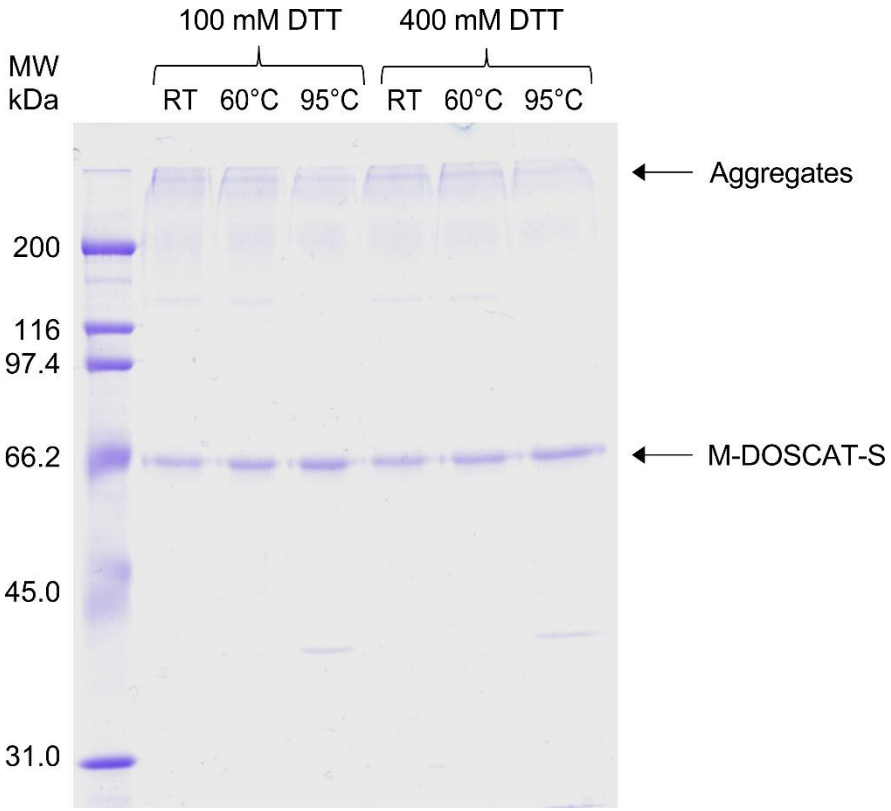


Figure 5.9 The effect of DTT concentration on M-DOSCAT-S aggregation. M-DOSCAT-S was incubated in SDS sample buffer containing 100 or 400 mM DTT and incubated at RT 15 mins, 60°C 10 mins or 95°C 5 mins before analysis by SDS-PAGE.

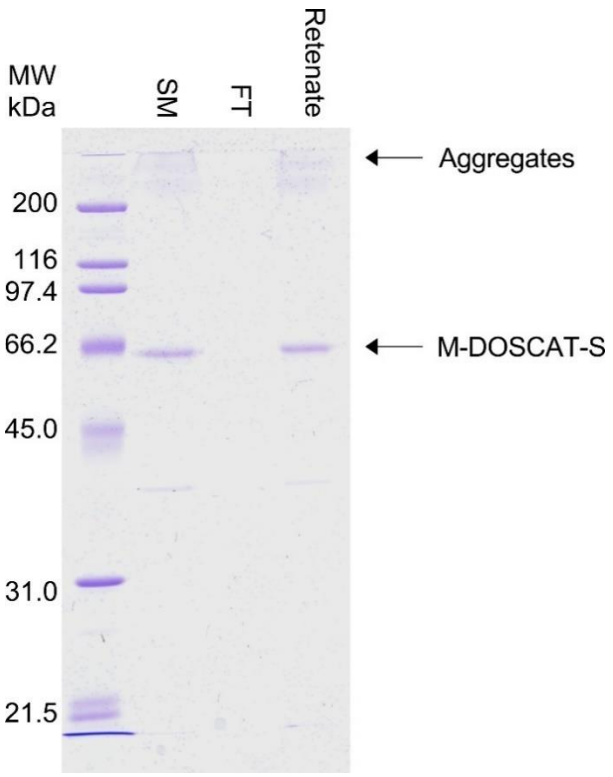


Figure 5.10 MWCO to remove M-DOSCAT-S aggregates. M-DOSCAT-S was transferred into a spin filter with a MWCO of 100 kDa and centrifuged at 15,000 x *g*, 15 m. Starting material (SM), flow through (FT) and retentate was analysed by SDS-PAGE.

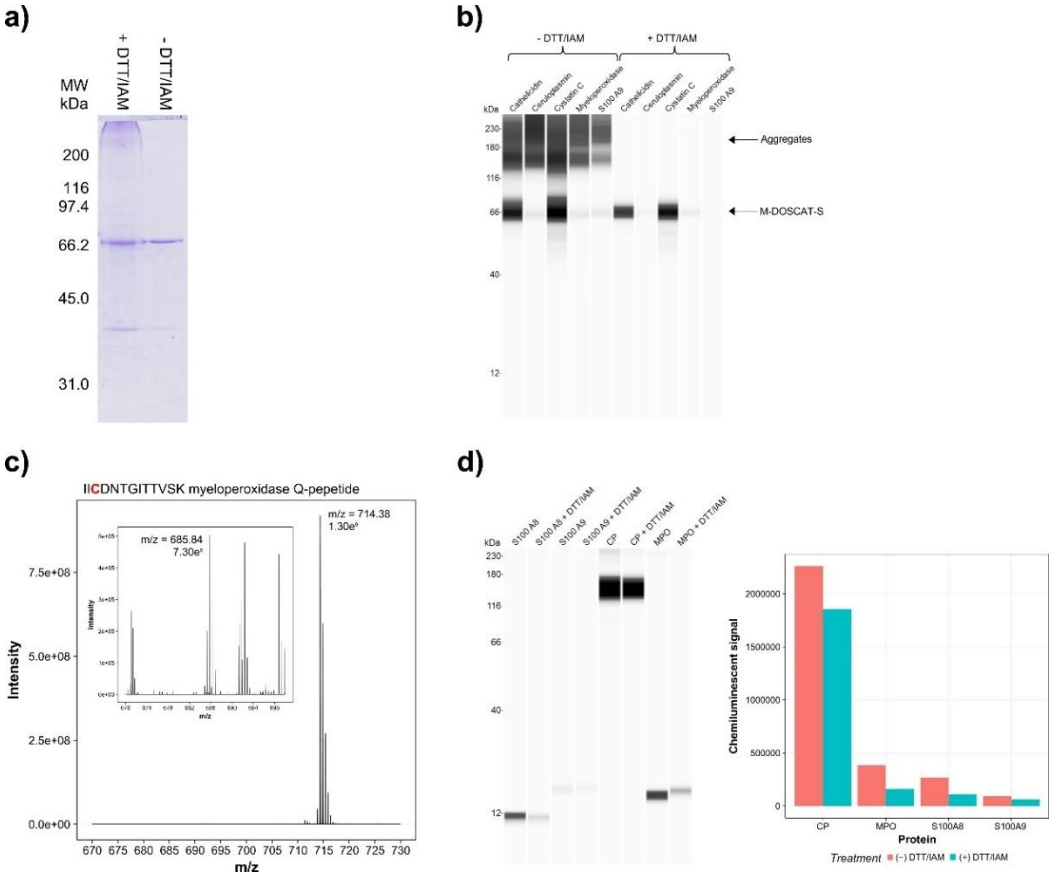


Figure 5.11 Prevention of aggregation in M-DOSCAT-S by DTT/IAM incubation. M-DOSCAT-S contained in elution buffer was incubated with 10 mM DTT overnight followed by 20 mM IAM for 30 m and dialysed into storage buffer. Treated M-DOSCAT-S was analysed alongside untreated protein by a) SDS-PAGE and b) automated capillary western blotting using target antibodies. c) XIC of a Q-peptide from the MS1 trace of DTT/IAM treated M-DOSCAT-S digested with trypsin and analysed by MS/MS, highlighting the completeness of cysteine alkylation. d) Recombinant protein standards ± DTT/IAM treatment were analysed by automated capillary WB with their cognate antibody; Compass software was used to extract chemiluminescent signal for comparison.

The aims of the initial work in this chapter were to design and express a DOSCAT standard targeting six proteins implicated in pneumococcal meningitis and an immunogen containing a concatenation of epitopes that would be used for the generation of an anti-DOSCAT antibody. Whereas the DOSCAT standard expressed well and could be used for further studies, the immunogen did not express in multiple cell lines and conditions. Lack of expression of recombinant proteins is usually attributed to leaky expression, protein toxicity to the cell, or codon bias (Rosano et al., 2014). Leaky expression occurs when promotors are not tightly regulated and so proteins are expressed before induction, which can lead to plasmid instability and the culture becoming saturated with cells that no longer contain the plasmid. M-DOSCAT-i, though, was not expressed in cell lines with highly regulated promotors (such as C43 and PLysS) that are known to supress leaky expression and express toxic proteins when BL21 DE3 cells will not (Dumon-Seignovert et al., 2004). This contradicts the view that toxicity of M-DOSCAT-i prevented expression, compounded by the fact that growth curves for all cell lines used were normal when arrested growth or cell death would be expected (Doherty et al., 1993; H. Dong et al., 1995). Codon bias describes the differences in the frequency of occurrence of synonymous codons in coding DNA, with the tRNA population of a cell closely reflecting this bias (Hengjiang Dong et al., 1996). During overexpression, if codons in heterologous DNA are significantly different from that of the host, there may be a deficiency in particular tRNAs, leading to amino acid misincorporation or translation termination (Gustafsson et al., 2011). However, M-DOSCAT-i DNA had been optimised to include high abundance codons as well as to minimise mRNA secondary structure features that may prevent translation. Additionally, M-DOSCAT-i did not express in several cell lines that have been developed for the expression of toxic proteins or mRNA containing rare codons.

It is therefore unclear exactly why M-DOSCAT-i did not express. Protein degradation is known to increase under cell stress, which may explain why cells continued to grow despite protein toxicity. Leaky expression or codon bias cannot be completely ruled out despite the protein not expressing in cell lines designed to prevent these problems, as they are not completely infallible. Some QconCATs fail to express (Beynon and Harman, personal communication, 2017), although it is difficult to estimate an overall failure rate as not every failure will be reported in the literature. As reported here, there is evidence that changing expression conditions or cell strains does not overcome lack of QconCAT expression (Russell et al., 2013), but shuffling the order of peptides in the sequence and resynthesizing the gene can (Mirzaei et al., 2008; Russell et al., 2013). This presumably alters the properties of the QconCAT, making it less toxic to the cell. For financial and time considerations it was decided not to attempt this with M-DOSCAT-i, especially as success was not guaranteed (Mackenzie et al., 2016), but it is certainly a viable strategy for future work. Another approach that could be considered is the use of cell-free expression systems, which have been employed to express proteins recombinantly (Carlson et al., 2012), including QconCATs (Takemori et al., 2016). In such systems the crucial cellular components for protein synthesis are isolated and

expression is carried out *in vitro*. This approach overcomes the problem of protein toxicity, making it an attractive proposition for future studies.

In contrast to the immunogen, the DOSCAT standard behaved mostly as expected based on experiences with previous DOSCATs and QconCATs. This DOSCAT, though, formed higher order aggregates after purification, a behaviour which is commonly observed in QconCATs but is usually overcome by solubilisation in a chaotrope during the purification workflow (D. M. Simpson et al., 2012). Although protein aggregation can originate from several mechanisms, intrinsically disordered proteins lacking a unique tertiary structure are much more likely to aggregate (Peng et al., 2014; Sigalov, 2010). As DOSCATs and QconCATs are unique artificial proteins, they are likely to lack a defined tertiary structure and so aggregate. However, M-DOSCAT-S formed aggregates after solubilisation in a chaotrope whereas the NFκB-DOSCATs did not. Aggregation was overcome by reduction and alkylation of thiol groups, suggesting that it was caused by disulphide bond formation. There is the same proportion (~1.35%) of cysteine residues in M-DOSCAT-S (8 Cys residues) and NFκB-DOSCAT-2 (9 Cys residues), but cross-linking only occurred in M-DOSCAT-S. It may be that the unique structural properties of M-DOSCAT-S allowed for cysteine residues to come into the required proximity to form disulphide bonds whereas this could not occur the NFκB-DOSCATs, for example if cysteines were sterically hidden or otherwise unable to contact one another. When designing future DOSCATs it would be prudent to take the number of cysteines into account and even attempt to model the structure of the protein to see if disulphide crosslinking is likely to be an issue. Reduction and alkylation of cysteine overcomes the problem of aggregation, but at the cost of chemically modifying the protein and potentially altering antibody affinity for the epitope. This is accounted for if the sample (e.g. CSF) is also treated in the same way, as the epitopes in standard and analyte will be the same. This will, however, lower the sensitivity of QWB assays, potentially limiting the number of proteins that can be reliably quantified by the technique. Despite this drawback, in the light of the problem presented and the ease of which it can be overcome by DTT/IAM treatment, reduction in sensitivity can be deemed an acceptable compromise and so DTT/IAM treatment will be included in all future workflows using M-DOSCAT-S.

5.2.2. Restricted proteolysis of M-DOSCAT-S

M-DOSCAT-S contained three protease cleavage sites, the specificity of which had been confirmed in NFκB-DOSCAT-2. An initial experiment was performed that aimed to demonstrate that the proteases cleaved M-DOSCAT-S with the required specificity when using the same conditions that had been successful previously. However, after overnight digestion with each protease M-DOSCAT-S was not proteolysed and no proteolytic fragments were detected upon SDS-PAGE analysis (Figure 5.12a). The M-DOSCAT-S signal was very weak, though, and there was evidence of aggregation in the sample that may have influenced proteolytic activity. The experiment was repeated using a fresh preparation of M-DOSCAT-S that had not been DTT/IAM treated but did not yet exhibit aggregation (determined by SDS-PAGE), freshly made protease buffers and increased concentration of proteases. Again, though, M-DOSCAT-S remained intact even after overnight incubation (Figure 5.12b). Due to time constraints and the need to achieve the primary aim of using M-DOSCAT-S to quantify target proteins, further experimentation attempting to accomplish complete digestion was not performed. Clearly, there is scope for future work in both optimising protease digests and subsequently analysing protease fragments by western blotting.

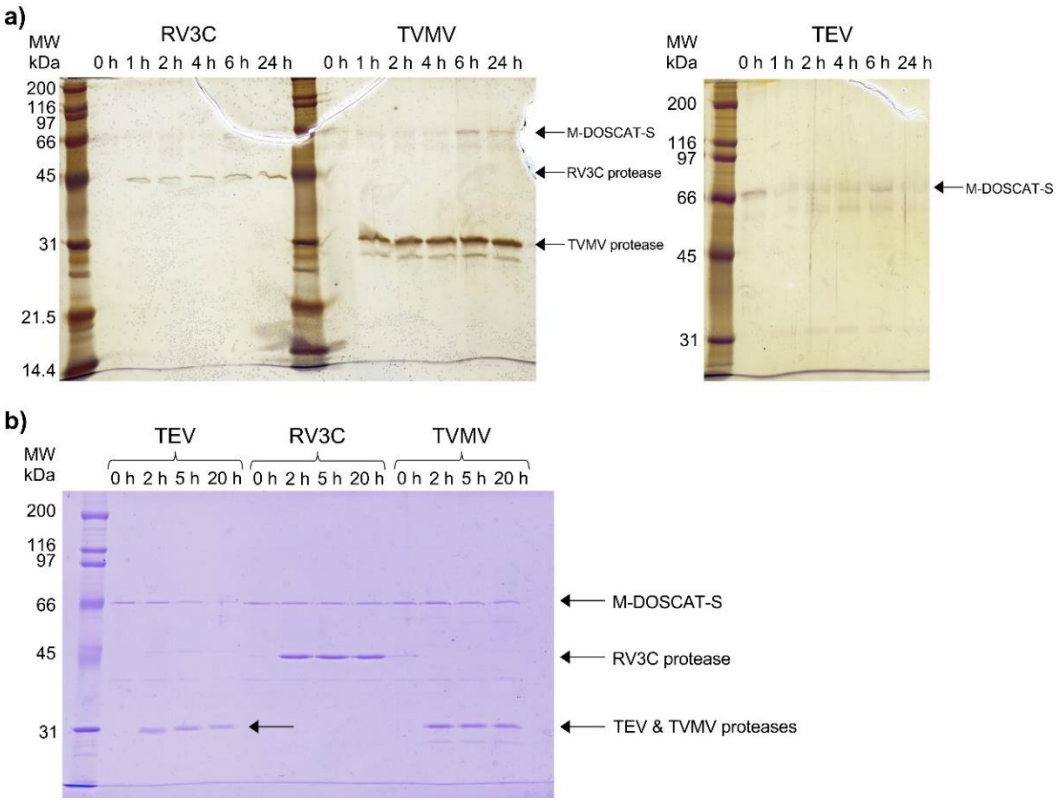


Figure 5.12 Restricted proteolysis of M-DOSCAT-S. a) M-DOSCAT-S (5 µg) was incubated with 1 U of each restricted protease and analysed on 15% (left) 12% (right) SDS-PAGE gels, using Silver stain for protein visualisation. b) M-DOSCAT-S (10 µg) was incubated with 10 U of each restricted protease and analysed on a 12% SDS-PAGE gel.

5.2.3. Optimisation of quantitative assays

As previously explained (Chapter 4.3.2), antibody, sample and DOSCAT concentration must be optimised for automated capillary QWB. To find optimal antibody concentrations, a dilution series of each antibody was run against a fixed concentration of recombinant protein and the peak area/baseline value for each antibody concentration was calculated (Figure 5.13). The antibody concentration at the saturation point was selected as the optimal concentration and used in all future experiments. To assess the range of M-DOSCAT-S to use for calibration curves for each target protein, M-DOSCAT-S was spiked into CSF (to provide a suitable matrix background) and serially diluted. The signal was linear across the range of concentrations measured for all antibodies, and endogenous proteins did not interfere with M-DOSCAT-S signal (Figure 5.14).

The next optimising step involved finding the optimal amount of CSF to load so that resultant signal was measurable, fell within the range of the standard curve and was not saturated. There was a wide range of total protein concentrations across all CSF samples, evidenced when the CSF samples were analysed by SDS-PAGE (supplementary data). Rather than normalising each sample for protein concentration, a fixed volume of CSF was to be used as any future diagnostic test would be applied to a fixed volume and so the analysis should take into account the natural range of protein content between samples. This was consistent with the approach taken in the discovery proteomics and WB verification work (Gómez-Baena et al., 2017). The abundance of target proteins would differ greatly between samples and sample types (case vs control), so samples required different dilutions on a protein specific basis.

To help guide this process, two CSF samples with the highest and lowest total protein concentrations (based on the SDS-PAGE data) were diluted to 1:5 and 1:50, treated by DTT/IAM and analysed by automated capillary WB by each antibody (results summarised in Table 5.3). S100A9 and MPO were not detected in either CSF samples, indicating a very low protein abundance in these samples or a relatively poor antibody that could not detect the proteins in a complex background. For other proteins, resultant signals were compared with the M-DOSCAT-S titration data and dilutions for CSF samples were selected based on the requirement for signal to be in the linear range of the standard curve (Table 5.4). For ceruloplasmin, myeloperoxidase and S100A9 it was clear that higher concentrations of M-DOSCAT-S were required for standard curves to cover all analytes, therefore a titration from 0 – 100 fmol/µL was performed to demonstrate linearity across this range. For control proteins, it was anticipated that target proteins other than cystatin C would be at a very low abundance, so no sample dilution was required.

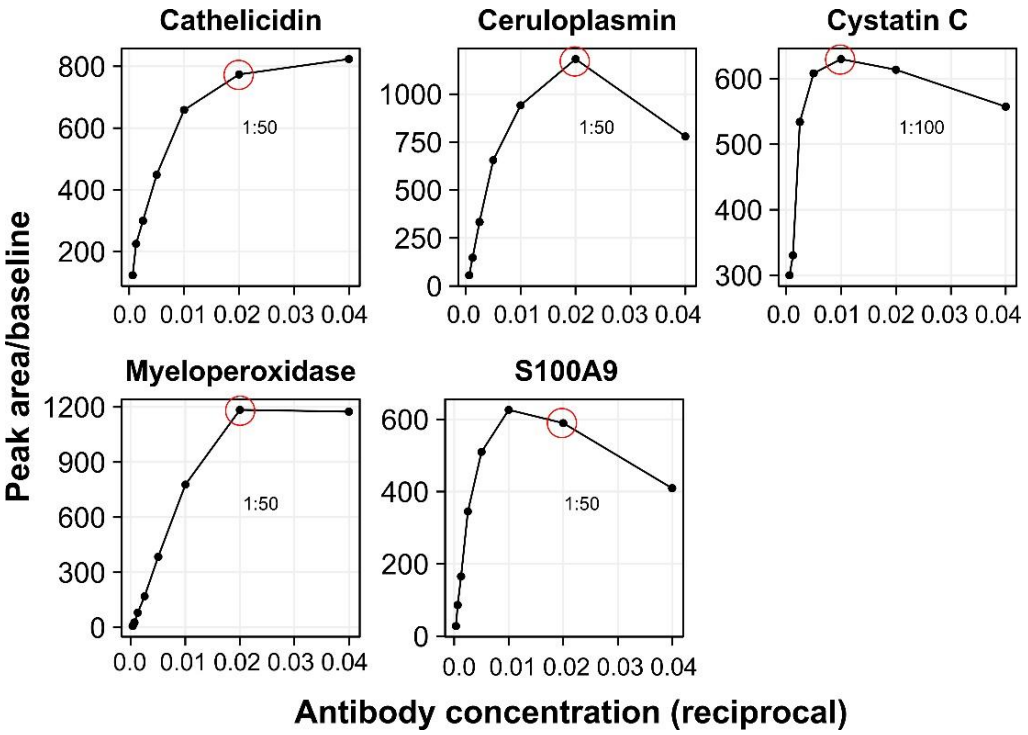


Figure 5.13 Antibody concentration optimisation. A dilution series of each antibody was prepared and used to analyse 2 ng/µL recombinant protein standard by automated capillary western blotting. Chemiluminescent signal was calculated by Compass software and plotted against each antibody concentration. The optimal concentrations selected for each antibody for use in the Wes system are circled on each plot.

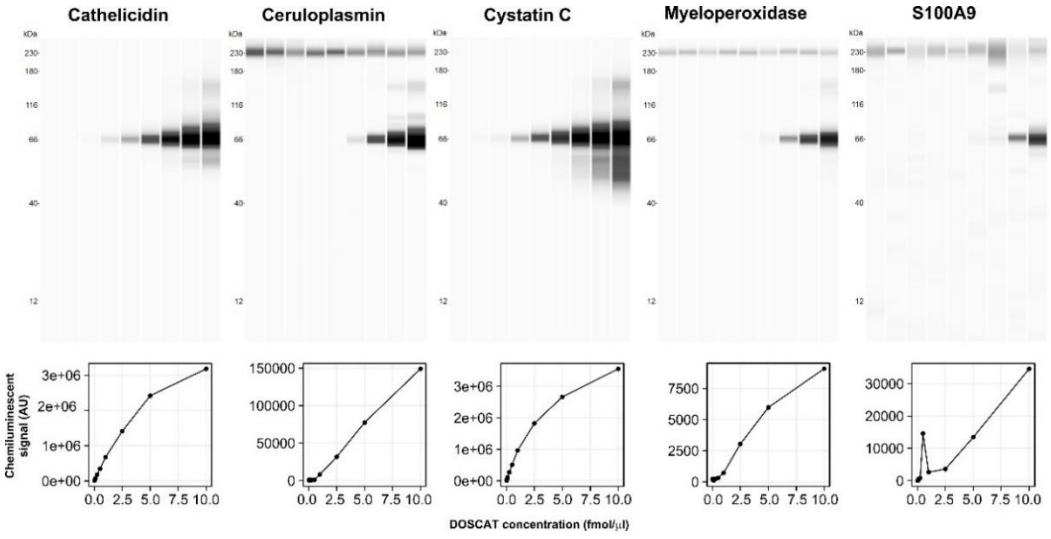


Figure 5.14 Assessment of M-DOSCAT-S linearity in capillary western blot analysis. Pseudo-gels and standard curves resultant from the automated western blot analysis of 0 – 10 fmol/µL M-DOSCAT-S using each target antibody.

Table 5.3 Optimisation of CSF dilutions for Wes analysis. CSF samples with low and high total protein concentration were diluted 1:5 and 1:50 and analysed by capillary WB using each target Ab. Integrated peak areas determined using Compass are displayed for each sample and target protein. An X denotes that no signal was detected and B/O denotes signal burn out due to excess protein loaded onto the capillary.

Protein	CSF sample 107 1:5 dilution	CSF sample 107 1:50 dilution	CSF sample 140 1:5 dilution	CSF sample 140 1:50 dilution
Cathelicidin	1,255,222	88,990	B/O	168,635
Ceruloplasmin	1,918,261	86,978	B/O	3,067,633
Cystatin C	B/O	1,224,451	B/O	1,611,013
Myeloperoxidase	X	X	X	X
S100A9	X	X	X	X

Table 5.4 Optimal dilutions of CSF samples for capillary WB analysis.

Protein	Case samples dilution	Hospital control samples dilution	Healthy control samples dilutions
Cathelicidin	1 in 50	no dilution	no dilution
Ceruloplasmin	1 in 100	no dilution	no dilution
Cystatin C	1 in 20	1 in 40	1 in 40
Myeloperoxidase	1 in 10	no dilution	no dilution
S100A9	1 in 5	no dilution	no dilution

Using the spectral library generated from the analysis of a tryptic digest of M-DOSCAT-S by LC-MS/MS on an Orbitrap Elite, three transitions were selected for each peptide based on signal intensity with a preference for product ions with a higher *m/z* than the precursor ions (all MS/MS spectra in supplementary data). A scheduled SRM method was built based on the retention time for each peptide, which resulted in all peptides being detected other than DGHSLGR. Peaks were mostly of high intensity and cleanly isolated, although QIVAGVNYFLDVELGR and TQPNLDNCPFHDQPHLK (both cystatin C peptides) exhibited poor chromatography and weak signal compared to other peptides (Figure 5.15a). To attempt to explain why these peptides performed poorly, digested M-DOSCAT-S was analysed on a QExactive HF Orbitrap instrument and miscleaves and chemical modifications to the peptides searched for. DGHSLGR was not detected by database searching or in the raw data, and signal for the peptide was very low within the raw data from the Orbitrap Elite run. This lack of

detectability on two separate instrument platforms was highly suggestive of poor ionisation efficiency or fragmentation. For this reason, DGHSLGR was discarded from future analysis. The presence of asparagine deamidation on the remaining two peptides was reported by database searching, and this was confirmed by inspecting the raw data (Figure 5.15b).

The digestion profile of each peptide was assessed in standard and analyte by spiking M-DOSCAT-S into *Streptococcus pneumoniae* positive (SPP) CSF and sampling the digestion mixture regularly. Two ceruloplasmin peptides, EHEGAIYPDNTTDFQR and ENLTAPGSDSAVFFEQGTTR were not observed in the analyte, so a standard:analyte could not be calculated, explaining the very low heavy:light ratio for these peptides compared to the other ceruloplasmin peptide. Of the 13 remaining peptides, all reached a stable plateau after overnight incubation with trypsin (Figure 5.16). This was the expected result, given that spacer regions of six amino acids were used around each peptide. To determine signal linearity and assay sensitivity in the sample matrix, M-DOSCAT-S was spiked into CSF over a 100-fold range and analysed by SRM. Most peptides exhibited a linear relationship across the entire range with limits of detection measured between 100 and 250 amol (Figure 5.17). The peptides QIVAGVNYFLDVELGR and TQPNLDNCPFHDQPHLK exhibited non-linear behaviour, this was due to very low signal intensity resultant from asparagine deamidation, as explained above.

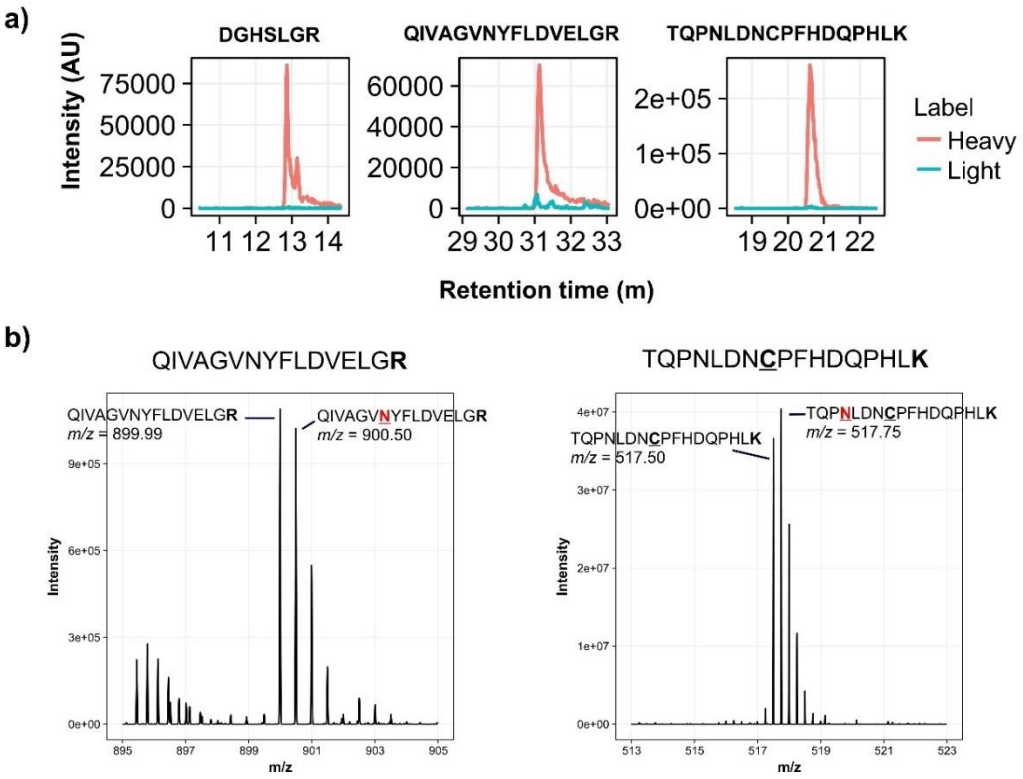


Figure 5.15 Analysis of underperforming Q-peptides. a) SRM chromatograms of three Q-peptides that displayed poor chromatography and/or low signal intensity in the analysis of digested M-DOSCAT-S by SRM-MS. b) Mass spectra of two Q-peptides resultant from MS/MS analysis of digested M-DOSCAT-S, highlighting the presence of asparagine deamidation in both peptides.

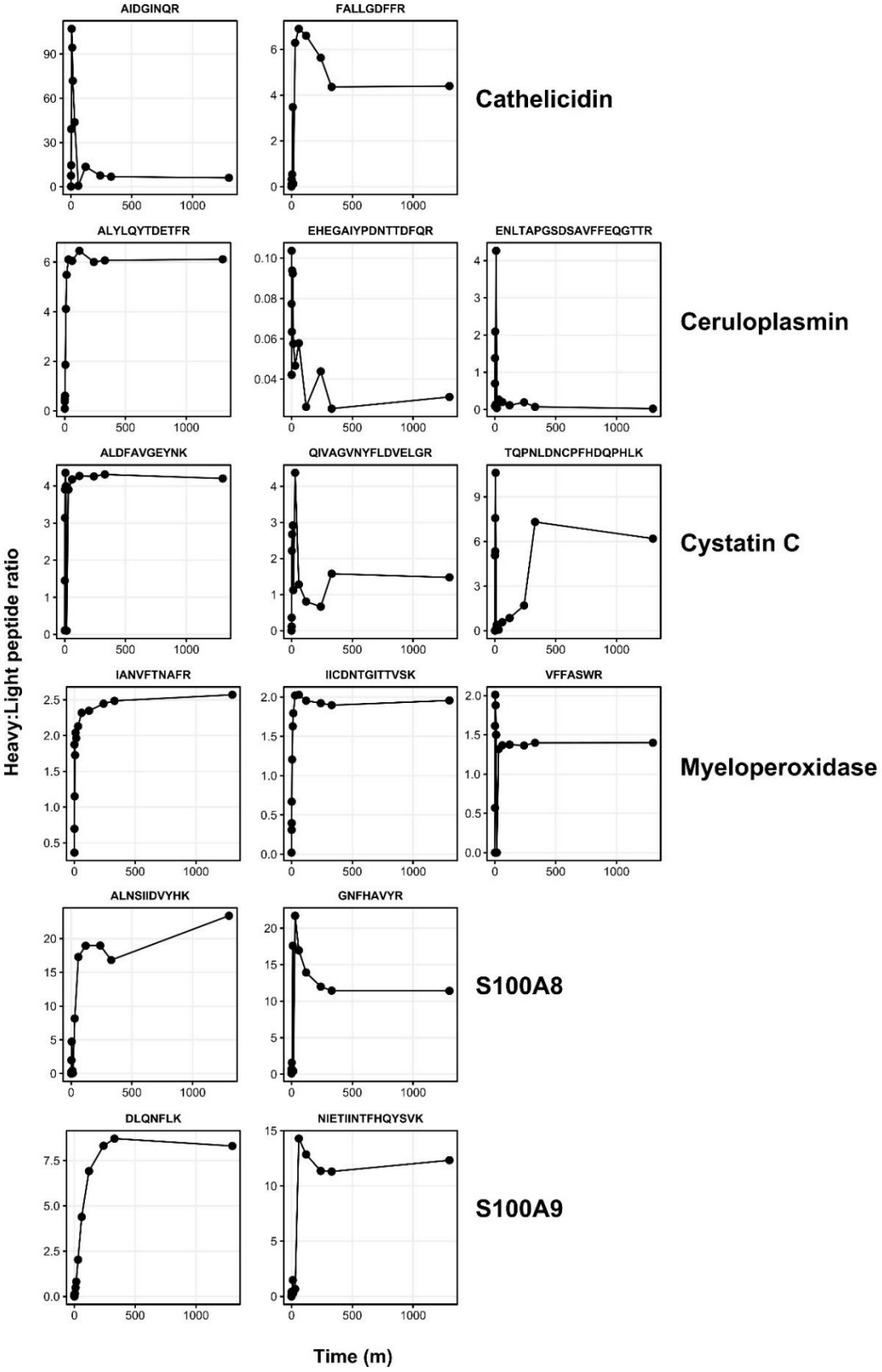


Figure 5.16 M-DOSCAT-S digestion time course analysis. Time course analysis showing the release of each peptide used for quantification during tryptic digestion. Stable isotope labelled M-DOSCAT-S and a SPP CSF sample were co-digested, with samples taken at a series of time points and digestion halted by acidification. Levels of target peptides in samples from each time point were analysed by scheduled SRM-MS. Displayed are ratio of heavy (M-DOSCAT-S) to light (endogenous) peptides at each time point.

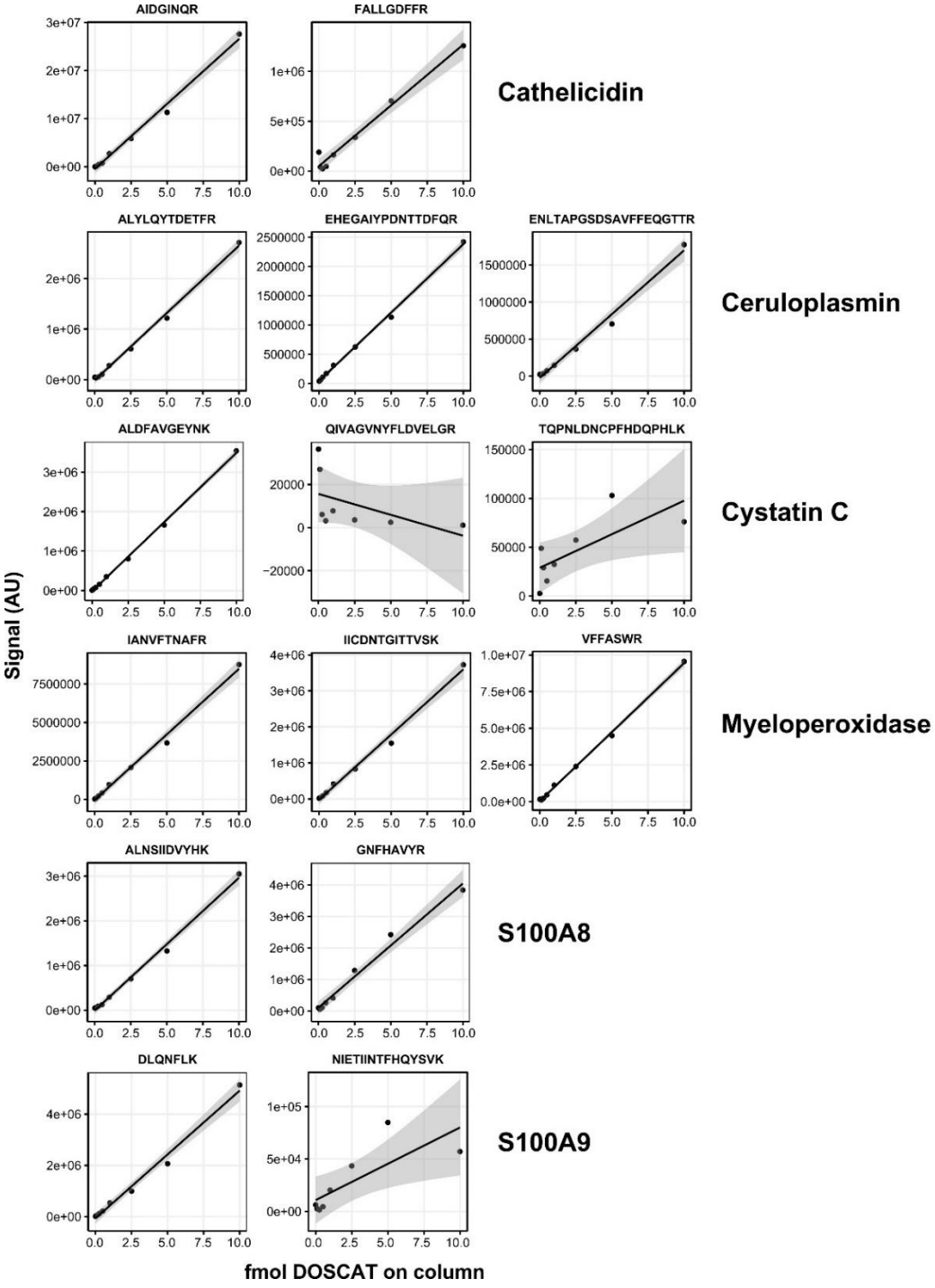


Figure 5.17 Q-peptides limit of detection in SRM-MS. M-DOSCAT-S (10 fmol/µL final) spiked in to a SPP CSF background was digested by trypsin, serially diluted in digested CSF and levels of each peptide were analysed by scheduled SRM-MS. The response of each Q-peptide from 0 – 10 fmol M-DOSCAT-S on the column is displayed.

5.2.4. Quantification of target proteins in CSF

Due to the large number and the variable protein concentration of samples, it was decided to perform quantitative experiments with the standard external to the analyte. Performing QWB using internal standard curves for each sample would require 44 runs compared with 10 for external calibration, making it unfeasible in terms of time and money. This is based on a six-point internal calibration curve being constructed for each of the 35 samples, requiring 210 data points for each antibody, of which there are five. This totals 1,050 Wes lanes required for analysis; a single run can accommodate 24 lanes, thus 44 runs would be required not including any repeats that might be needed. For external calibration, all the samples could be analysed by a single antibody across two 24-lane runs with six lanes on each run dedicated to the DOSCAT standard. For five analytes, this would total 10 runs. SRM-MS analysis could feasibly be performed using internal calibration, but external calibration was selected so as to be analogous with the QWB experiments. A total of 35 CSF samples were used for analysis, made up of 15 hospital controls, 5 healthy controls and 15 SPP patients.

For QWB analysis using Wes, accurately quantified M-DOSCAT-S was spiked into a CSF background (CSF diluted 1 in 40) and serially diluted to levels reflective of endogenous protein in CSF samples based on assay optimisation results. CSF samples were also diluted so the resultant signal was within the linear range of the assay and the standard curves, such that extrapolation was not required. For each Wes run, a six-point standard curve was analysed alongside 18 unknown CSF samples, which were run in a random order. This was repeated with antibodies for each target protein other than S100A8, which had been shown to not detect M-DOSCAT-S (Figure 5.7). M-DOSCAT-S was detected by all antibodies and resultant standard curves were linear and highly reproducible across runs (Figure 5.18). However, endogenous proteins were not detected in every sample (Figure 5.19). Myeloperoxidase was not detected in any SPP or control sample and S100A9 was detected in only 7 SPP samples. Cathelicidin was detected in all but one SPP sample, but in none of the control samples. Cystatin C was detected in all but one of the SPP and control samples and ceruloplasmin was detected in every SPP and control sample. There was a large amount of non-specific binding using the ceruloplasmin antibody; to determine if this was irrelevant protein or ceruloplasmin degradation fragments, CSF was run on an SDS-PAGE gel and segments around the expected ceruloplasmin band cut out for in-gel digestion. Analysis of the digests by MS/MS revealed no evidence of ceruloplasmin fragments at lower molecular weights, confirming that the observed non-specific were other proteins and could be ignored for the purposes of quantification (Figure 5.20). From each Wes run endogenous protein concentration was calculated from the calibration curve generated within that run only.

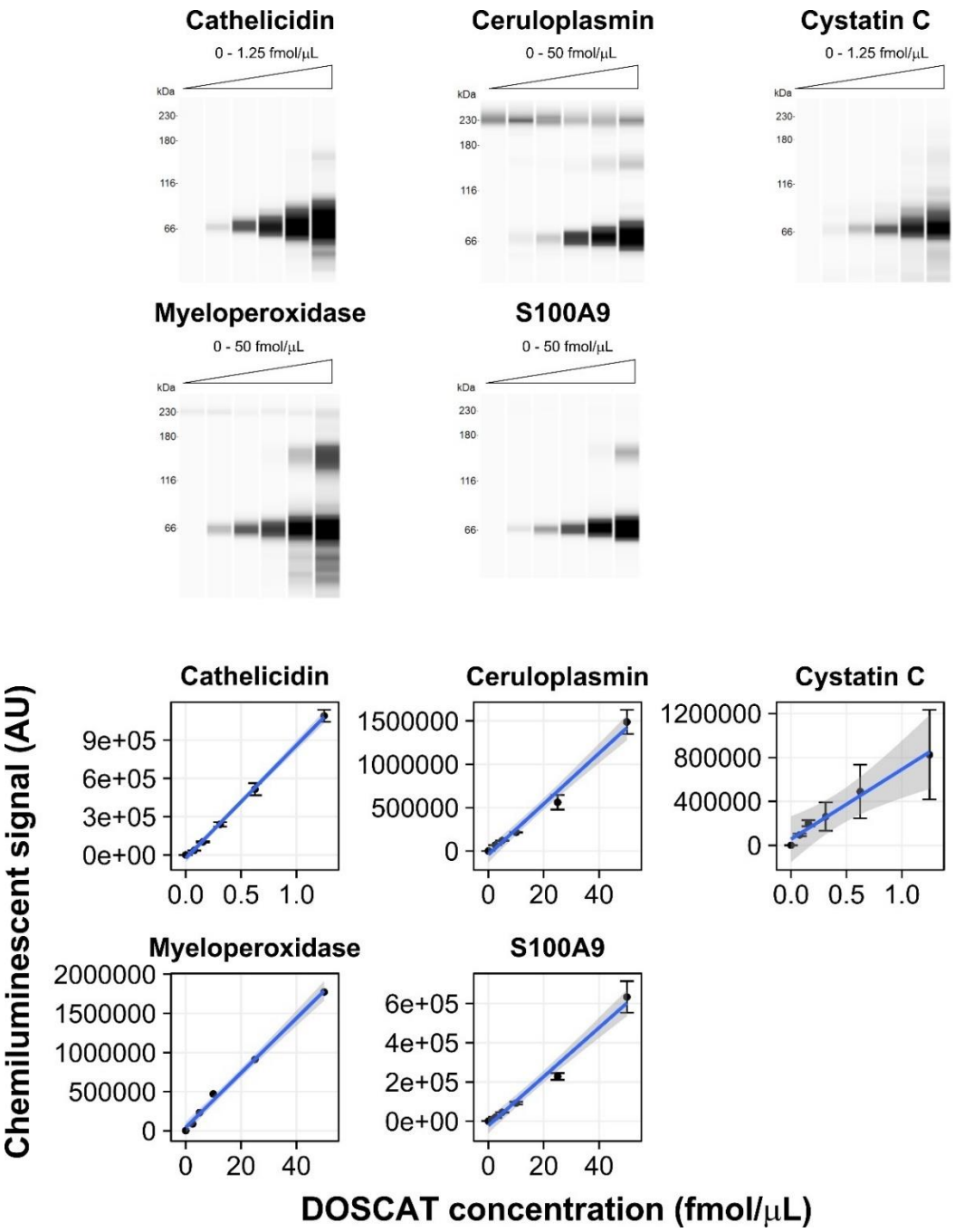


Figure 5.18 M-DOSCAT-S standard curves for QWB assays. Pseudo-gels and associated standard curves resultant from multiple dilution series of M-DOSCAT-S in a CSF background, run alongside CSF analytes in automated capillary western blotting. Raw data were analysed by Compass software and standard curves are presented as mean ± standard error (n = 3).

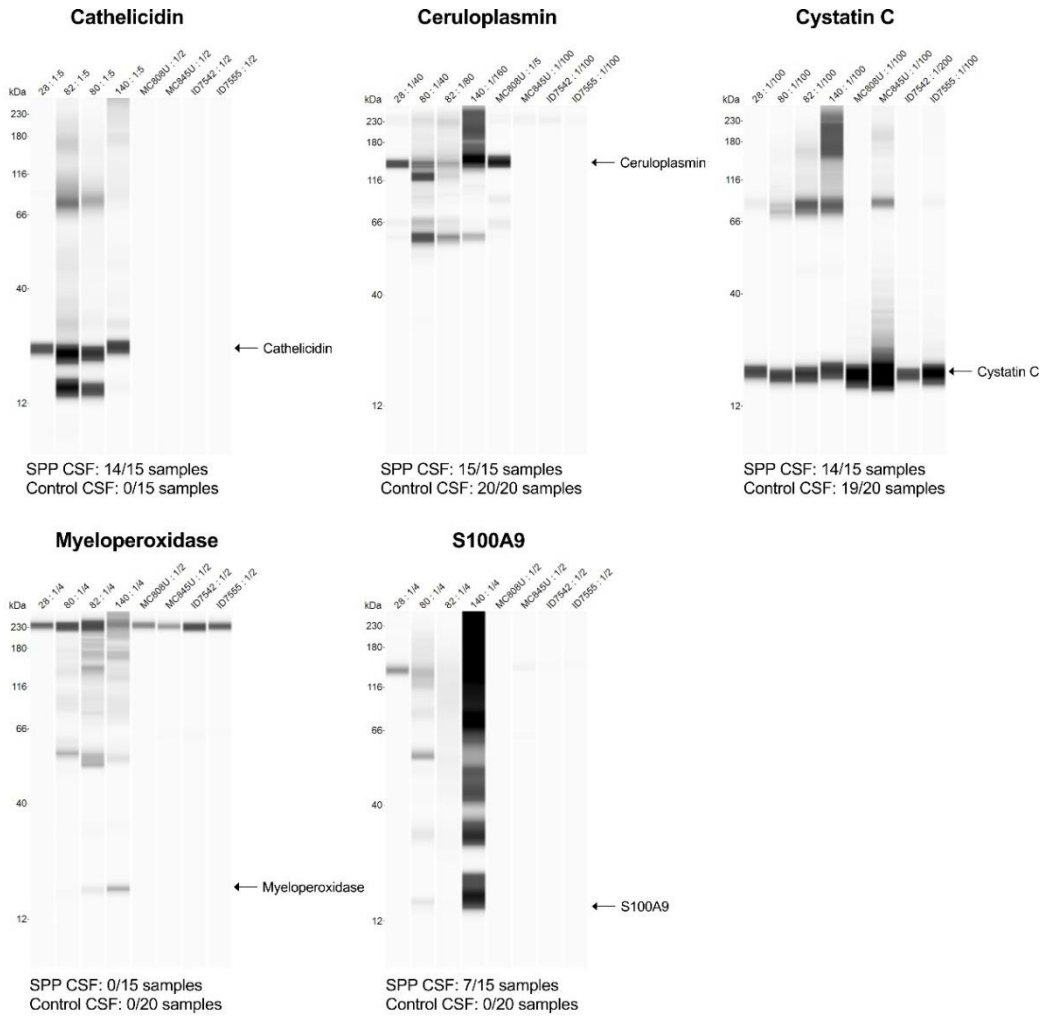
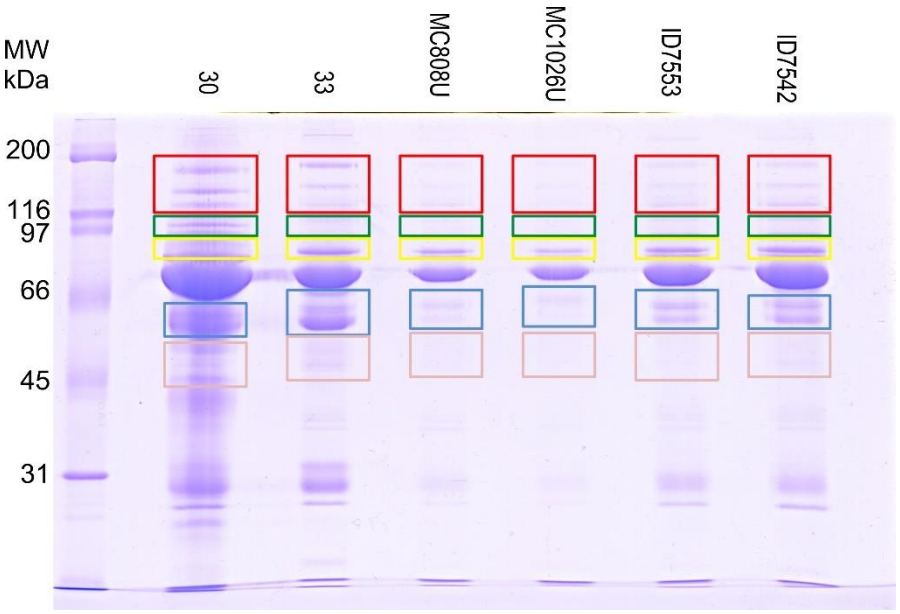


Figure 5.19 Western blots of CSF samples. Representative pseudo-gels of eight selected CSF samples at various dilutions resultant from analysis by automated capillary western blotting using five different antibodies to target proteins. Text under each blot refers to the total number of samples in which endogenous protein was detected by each antibody.



	30	33	MC808U	MC1026U	ID7553	ID7542
Red	✓	✓	✓	✓	✓	✓
Green	✗	✗	✗	✗	✗	✗
Yellow	✗	✗	✗	✗	✗	✗
Blue	✗	✗	✗	✗	✗	✗
Orange	✗	✗	✗	✗	✗	✗

Figure 5.20 In-gel digestion and MS/MS analysis of CSF samples. CSF samples were run on an SDS-PAGE gel (top) and the highlighted sections subjected to in-gel trypsin digestion and analysis by MS/MS. The bottom table summarises whether the protein ceruloplasmin was detected in each gel section defined in the above image.

For analysis by SRM-MS, M-DOSCAT-S was spiked into the same CSF matrix background as used for QWB and serially diluted in CSF to form an eight-point standard curve. Each CSF sample was digested with a final dilution of 1 in 10 and analysed by SRM-MS in a random order with blanks between samples. The calibration curve samples were analysed in triplicate alongside the unknown samples at the beginning, middle and end of the sample list. Calibration curves using M-DOSCAT-S signal were constructed for each peptide revealing poor reproducibility between the replicates (Figure 5.21). Each replicate was drawn from the same sample, so such poor reproducibility is surprising, indicative of a decline in instrument performance or sample degradation over time. As the standards were external to the sample, there was no way to normalise for the drop in assay performance between samples run at the beginning and end of the sample list. Therefore, it was decided to repeat the experiment using internal standardisation; this would be more expensive in terms of instrument time but would

yield more reliable data. CSF with a final dilution of 1 in 5 was co-digested with either 1 fmol and 0.1 fmol M-DOSCAT-S; this amount of standard was calculated to encompass the full dynamic range of target proteins. M-DOSCAT-S and CSF mixtures were digested with trypsin and to ensure complete digestion samples of each digest and the pellet generated post-acidification were analysed by SDS-PAGE (Figure 5.22). Undigested protein would precipitate upon acidification by TFA, so bands in the pellet samples would indicate incomplete digestion. This was evident in eight samples, all of which were SPP samples and so likely to have an increased total protein concentration. Digestion of these samples was repeated using an increased 1 in 20 dilution of CSF, and repeat analysis demonstrated that this permitted complete digestion.

Digests were analysed by SRM-MS and for each peptide quantification in terms of µg/mL was calculated from the ratio of peak areas in heavy and light extracted ion chromatograms (Figure 5.23). Peptide quantification values across the same protein agreed well in most instances, however there were some notable exceptions (Figure 5.24). The cathelicidin peptide DGHSLGR was known to ionise poorly and was not detected in heavy or light samples (Figure 5.15a); it was therefore removed from analysis when calculating protein level quantification. As previously observed in the digest time course data, the ceruloplasmin peptides EHEGAIYPDNTTDFQR and ENLTAPGSDSAVFFEQGTTR were not detected at all in the analyte, explaining the lack of quantification values (Figure 5.15b). These peptides were removed from analysis so that ALYLQYTDETFR alone was used for protein level quantification. Higher quantification values were measured by the cystatin C peptide TQPNLDNCPFHDQPHLK compared to the other two peptides for the protein. However, chromatography and signal intensity were good so the peptide was included in further analysis. Protein level quantification was calculated by taking the mean of the values obtained for different peptides, other than the ones removed for the reasons described.

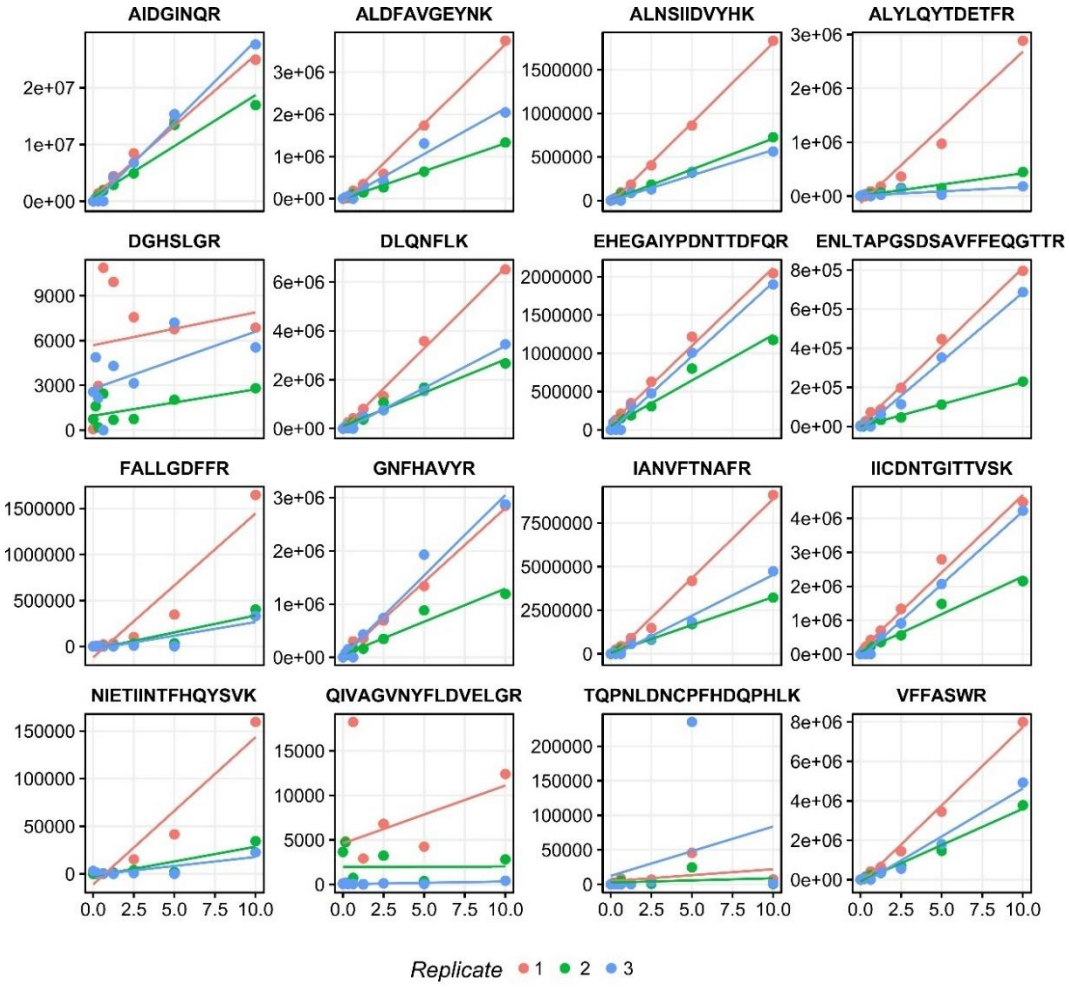


Figure 5.21 External standard curve reproducibility in SRM-MS. Technical replicates of standard curves resultant from the SRM-MS analysis of a dilution series of digested M-DOSCAT-S in a CSF background, analysed in between CSF analytes over the course of a 5-day acquisition.

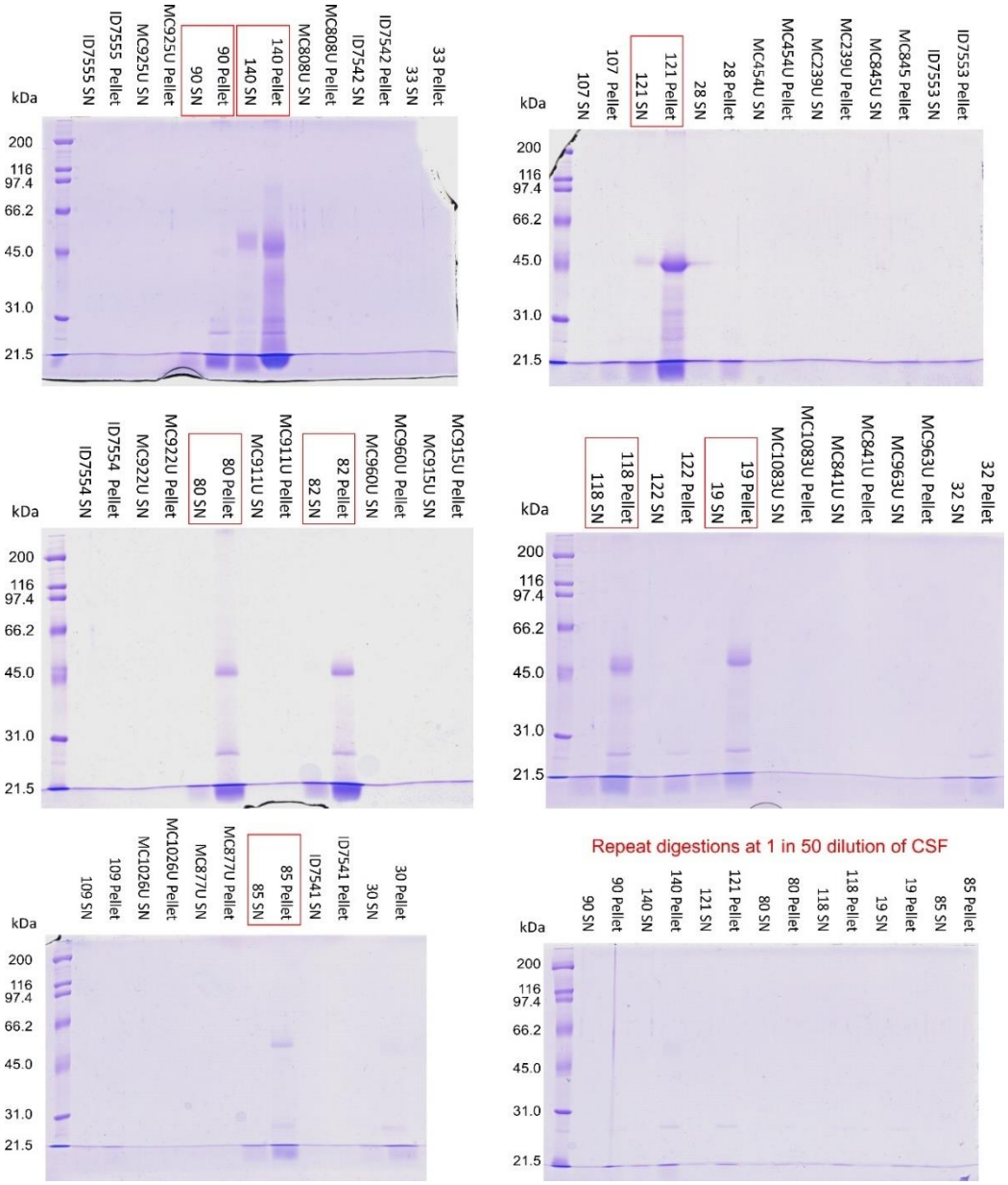


Figure 5.22 Assessment of CSF digestion by trypsin. CSF samples diluted 1 in 5 and M-DOSCAT-S (10 fmol/µL final) were co-digested with trypsin, acidified with TFA and centrifuged to separate the soluble and insoluble fractions. For all samples the supernatant (SN) was removed and the pellet resuspended in 10µL 2 x SDS sample buffer, and both SN and pellet were analysed on 12% SDS-PAGE gels. Samples highlighted in red boxes demonstrate incomplete digestion. These samples were digested again using a 1 in 50 dilution of CSF and analysed by SDS-PAGE (bottom right panel).

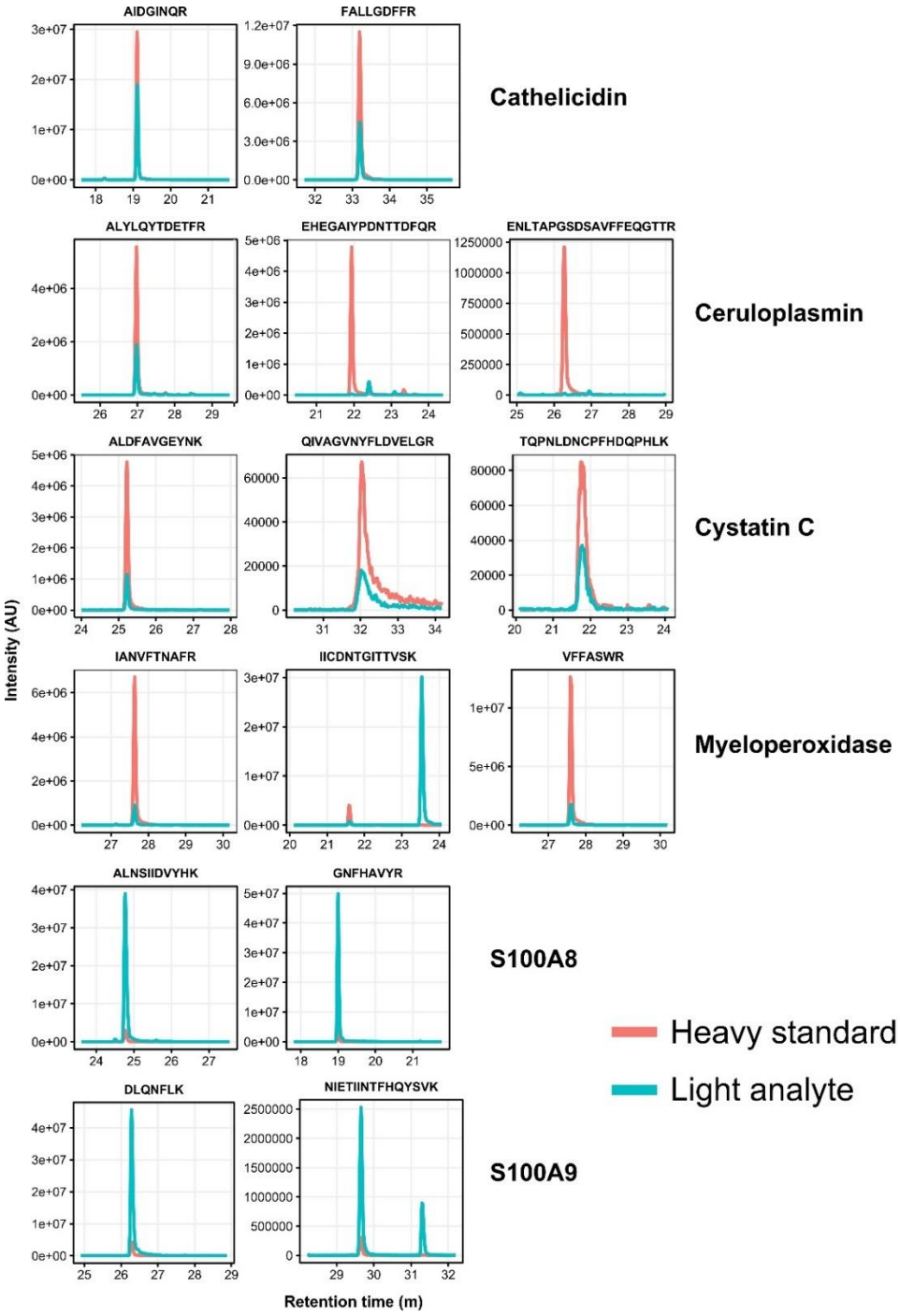


Figure 5.23 SRM-MS analysis using internal standardisation. Extracted SRM chromatograms for each Q-peptide from M-DOSCAT-S (red) and endogenous analyte (blue) resultant from the scheduled SRM-MS analysis 10 fmol M-DOSCAT-S spiked into SPP CSF sample #30.

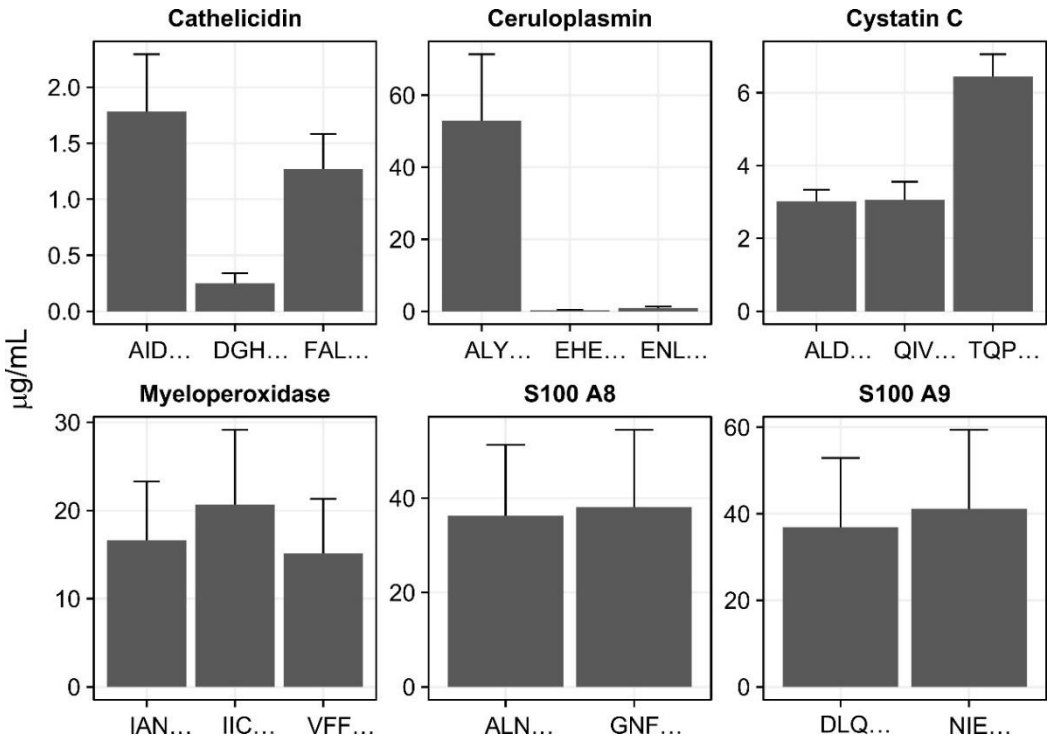


Figure 5.24 Peptide level quantification of target proteins. Quantification values for each detected Q-peptide were determined by analysing heavy:light ratios and converting to µg/mL for each sample. Data are from SPP samples only and presented as mean ± standard error (n = 15).

Quantification using QWB was possible for four proteins, whereas SRM-MS quantified all proteins. In some control samples, specific proteins were below the limit of detection in both assays and so quantification was not possible. Quantification by both QWB and SRM-MS resulted in similar relative quantification between SPP and control samples; moreover, protein fold changes were in the direction expected, in all proteins except cystatin C increased in SPP samples (Figure 5.25). Absolute quantification values, though, were remarkably different between the two platforms (Figure 5.26). Other than for ceruloplasmin, values obtained by QWB were substantially lower for cystatin C (~ 8-fold lower) and cathelicidin (~ 20-fold). As the differences were not consistent across all proteins, it was unlikely that there was a systematic error in M-DOSCAT-S concentration causing the discrepancy. It was hypothesised that proteases in CSF were degrading the DOSCAT at either the peptide or protein level, leading to inaccuracies in SRM or QWB assays respectively. To assess potential proteolysis at the peptide level, digested M-DOSCAT-S was spiked into four different undigested CSF samples at RT, and samples removed at 0, 1 and 4 h into 5% TFA for instant protein denaturation. SRM-MS analysis of the samples revealed that whilst there was a drop in signal intensity between the digest starting material and the first time point in the incubation, there was no drop in signal intensity throughout the incubation for most peptides (Figure 5.27). For some peptides, the signal intensity fluctuated or appeared to increase throughout the incubation, however this can be ascribed to inaccuracies in the measurement of peak area

due to a low signal to noise ratio. The initial signal loss can be attributed to adsorption of peptide to plastic tubes and tips during the spike-in process, so it was concluded that CSF was not proteolysing M-DOSCAT-S peptides. To investigate proteolysis at a protein level, intact M-DOSCAT-S was incubated with CSF at room temperature and at 0, 1 and 4 h samples were removed, incubated with Wes sample buffer at 95°C to inhibit proteolysis and subsequently analysed by automated capillary WB (Figure 5.28). When contained within a CSF matrix the signal intensity for M-DOSCAT-S was decreased and the electrophoretic mobility shifted compared to M-DOSCAT-S on its own. There was no decrease in signal intensity throughout the incubation, however. As M-DOSCAT-S had been spiked into a relatively high concentration of CSF (diluted 5-fold dilution against 40-fold dilution used for the standard curve samples), it was likely that the mobility shift of M-DOSCAT-S was due to increased total protein concentration and matrix effect rather than proteolysis of the standard. Proteolysis of M-DOSCAT-S at both a peptide and protein level was therefore ruled out as an explanation for the QWB results.

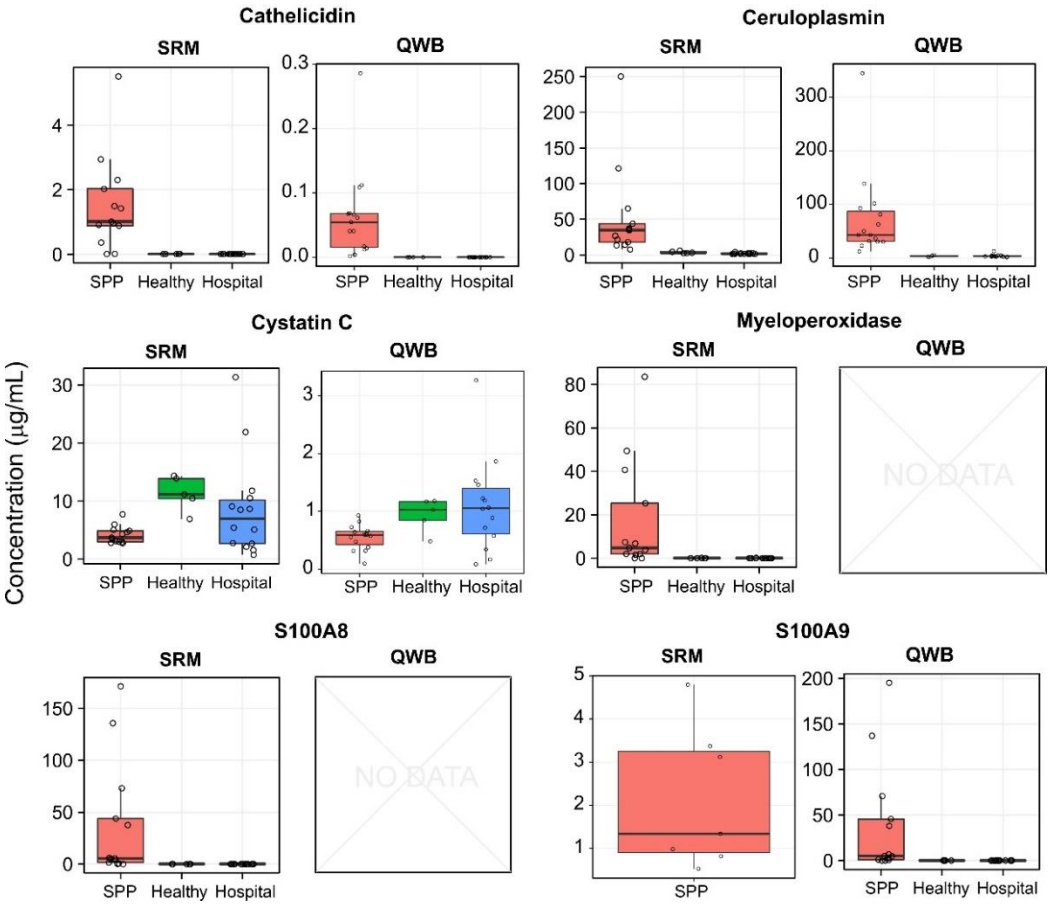


Figure 5.25 Target protein quantification summary. Quantification data for each target protein resultant from the analysis of SPP and control CSF by SRM-MS and automated capillary WB using M-DOSCAT-S as a calibrant. Boxplots display median, 25% and 75% quantiles, and spread of data. Individual data points represent specific CSF samples. Blank plots with 'no data' represent occasions in WB analysis where the antibody did not detect endogenous protein in any CSF sample.

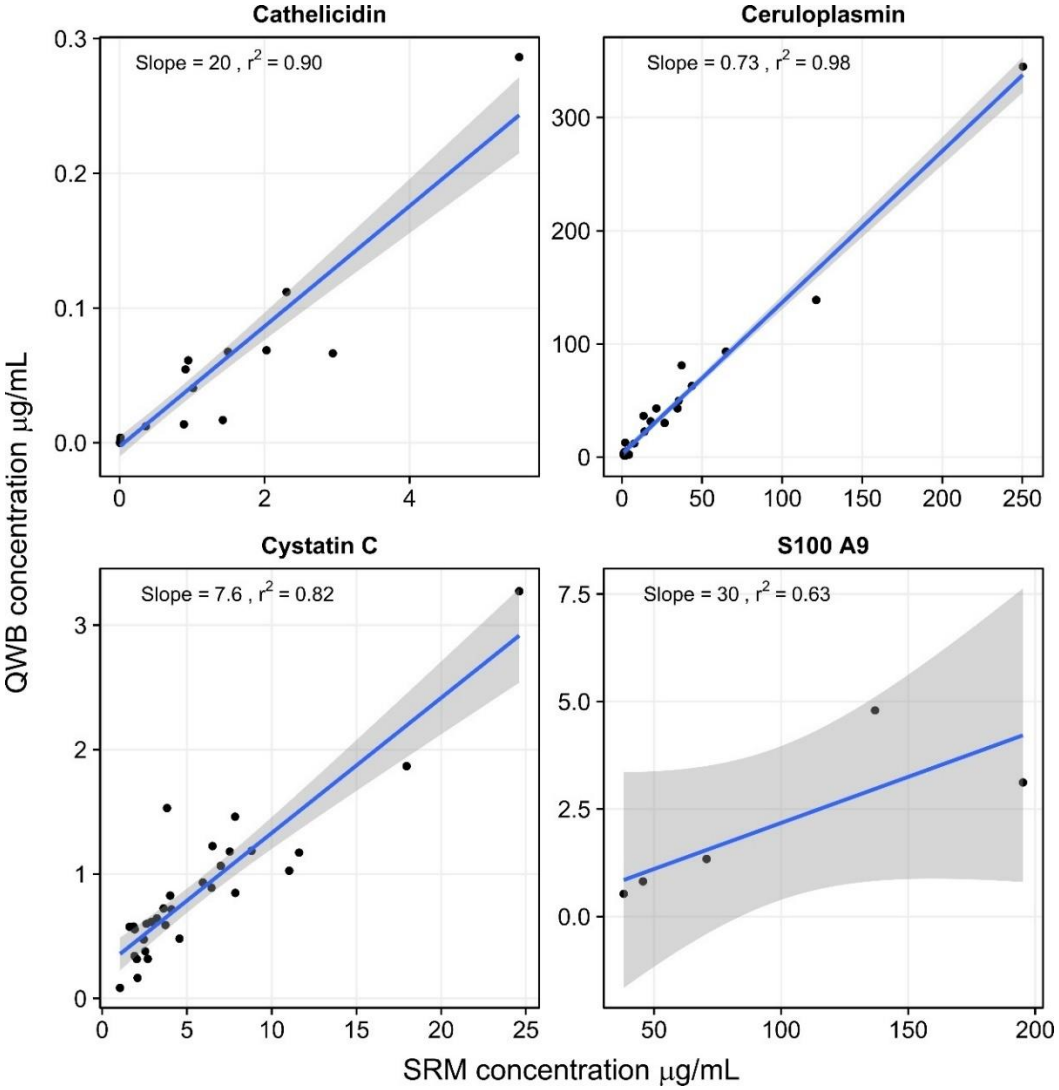


Figure 5.26 Platform comparison of absolute quantification. Quantification values for each sample as derived by SRM-MS and QWB were plotted against one another and modelled using linear regression. Gradient of the slope and r^2 values are displayed for each plot.

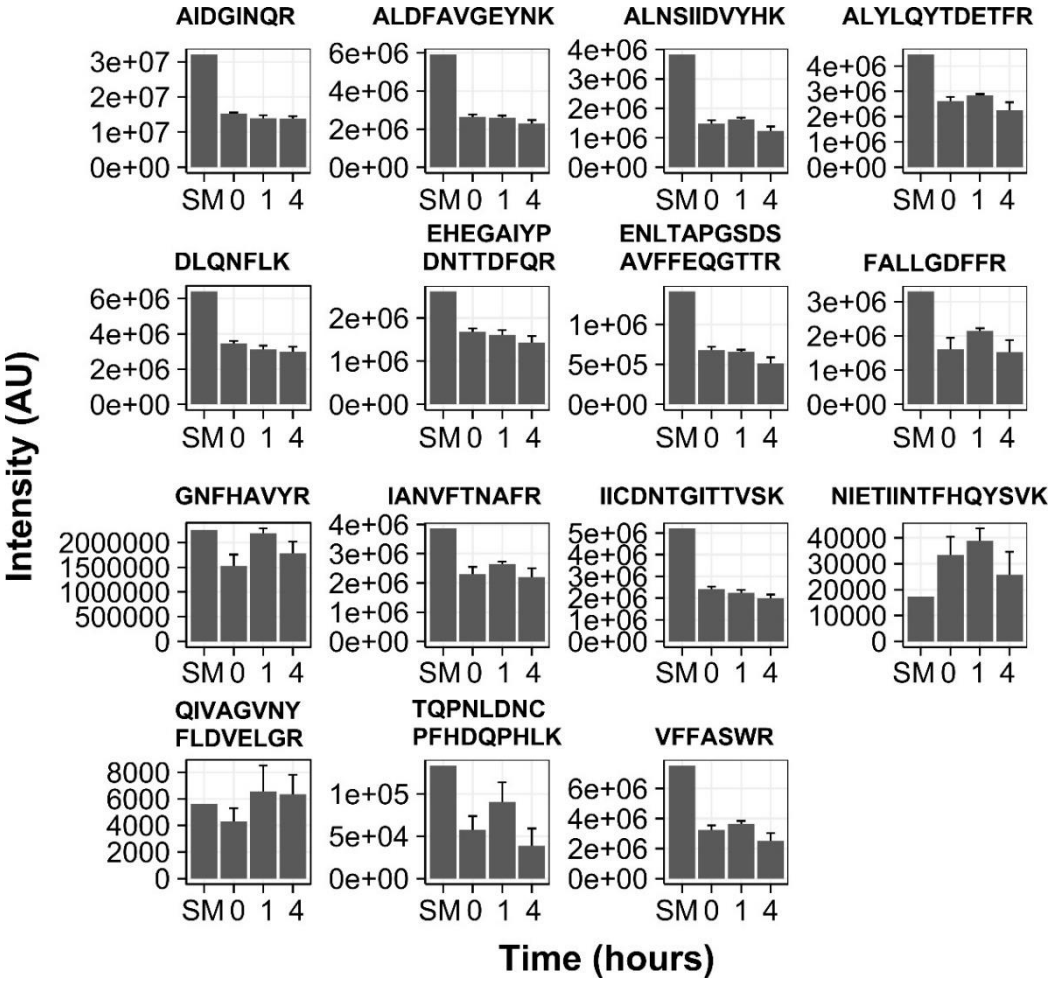


Figure 5.27 Incubation of Q-peptides with active CSF. Digested M-DOSCAT-S incubated with CSF for 0, 1 and 4 h alongside a digest independent of CSF incubation (SM) was analysed by SRM-MS. Displayed are mean values of signal intensity for each Q-peptide ± standard error (n = 4).

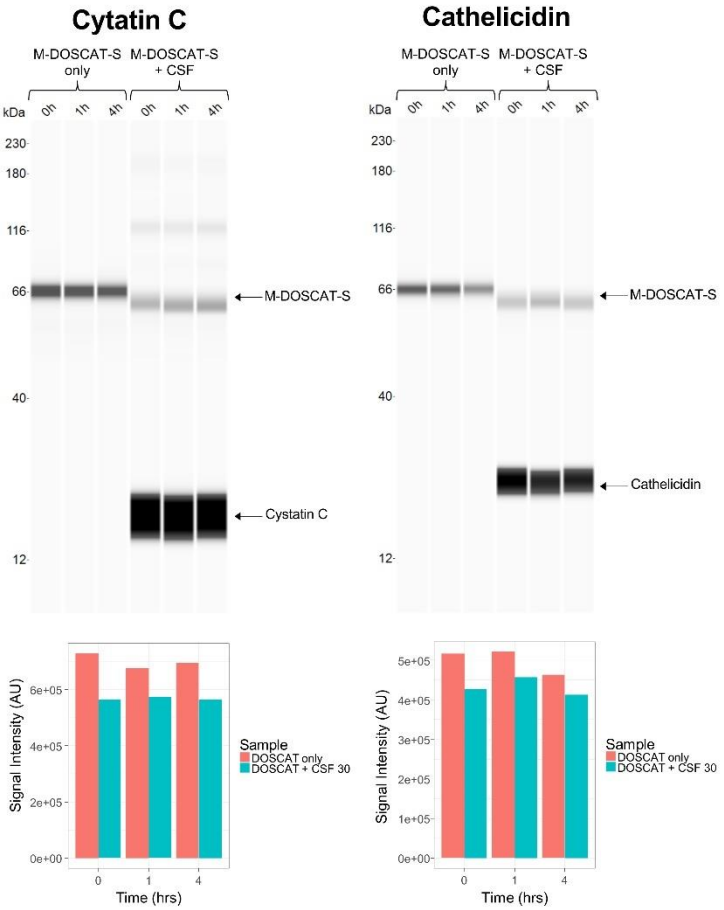


Figure 5.28 Incubation of M-DOSCAT-S with active CSF. M-DOSCAT-S alone and M-DOSCAT-S spiked into SPP CSF (#30) was incubated for 4 h at RT, with samples removed at 3 time points. Top panels: pseudo gels resultant from the capillary western blot analysis of WB of each time point sample using antibodies against two target proteins. Bottom panels: Chemiluminescent signal intensity as calculated by Compass software for each time point sample.

Two other datasets resultant from the related proteomics study (Gómez-Baena et al., 2017) were available to compare quantitative values generated by M-DOSCAT-S. Hi3 label-free quantification was possible using the MS data as samples had been mixed 1:1 with 50 fmol/µL of yeast alcohol dehydrogenase as a reference protein and quantification performed as previously described (J. C. Silva et al., 2006). The study also contained data from QWB experiments, using full length recombinant protein standards for calibration. These QWB experiments used the same set of CSF samples as described earlier (15 SPP and 20 control), whereas the label-free experiments used a different set of 16 CSF samples (8 SPP and 8 control). Quantification values for SPP samples only were averaged across all samples and compared between platforms (Figure 5.29). For all six proteins, concentrations obtained by Hi3 quantification and SRM-MS agreed very well. Quantitative values derived by QWB, however, were much more variable in a protein dependent manner. Results for QWB

calibrated by recombinant standards (QWB-std) were significantly higher compared to MS values for cathelicidin, cystatin C and S100A9, but were within the same range for ceruloplasmin and myeloperoxidase. Conversely, values for QWB calibrated by M-DOSCAT-S (QWB-DOSCAT) were significantly lower for the same three proteins, but values for ceruloplasmin were similar to those obtained by MS.

The elevated concentrations measured by QWB-std could be due to inaccurate concentrations of recombinant standards. If the standard concentration was actually lower than thought, erroneously high values of analyte would be measured. To test for this, three of the recombinant proteins that were used in the QWB experiments were analysed on an SDS-PAGE gel alongside BSA and affirmer standards, with samples prepared by adding sample buffer straight to aliquots to minimise losses (Figure 5.30). Affimers are based on cystatin C structure and so were deemed more suitable to act as a standard for cystatin C. BSA was used as a standard for the other two proteins. The affimer aggregated despite prior incubation with DTT and heating with sample buffer; to account for this the fraction of monomer intensity out of the sum of the five main aggregation species was used to calculate monomer concentration from the total protein concentration. Concentration of protein standards were calculated to be much lower than expected, compared to the amounts that were thought to be loaded onto the gel. Using these data, correction factors were calculated for each recombinant protein and applied to the QWB-std results (Figure 5.31). The corrected values agree much better with the SRM-MS and Hi3 results, suggesting that the abnormal QWB-std results were due to incorrect standard concentrations.

These results do not account for QWB-DOSCAT results; incorrect standard concentration is not a likely explanation as differences in measured concentration are not uniform across all proteins. Standard curves generated in QWB-DOSCAT and QWB-std (using corrected concentrations) were compared, revealing a discrepancy between the two (Figure 5.32). The differences in slopes (i.e. much higher signal for DOSCAT for the same amount of protein in cathelicidin and cystatin C) match the relative difference in measured analyte concentration between QWB-DOSCAT and QWB-std. It should be noted that data presented here are from recombinant proteins not treated with DTT/IAM, although the conclusions as still valid as such treatment would only reduce rather than increase signal.

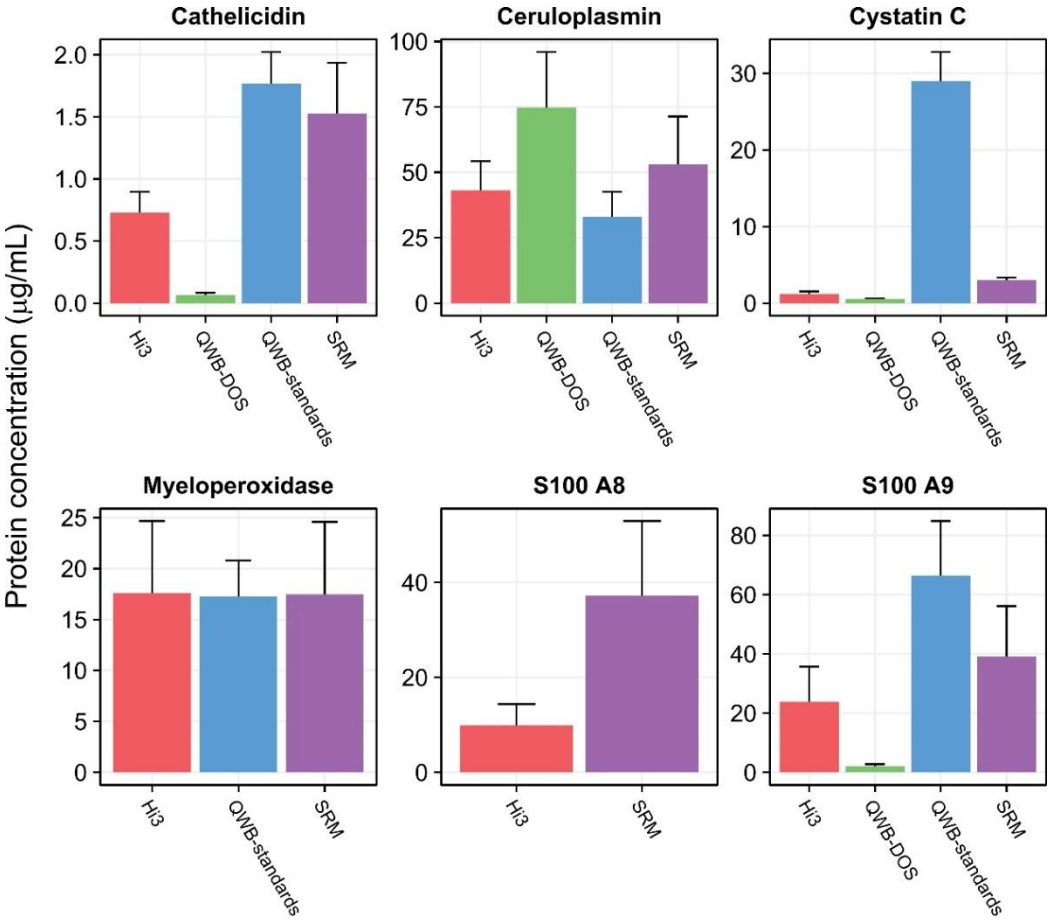


Figure 5.29 Comparison of quantification by different datasets. Protein level quantification values as calculated by Hi3 methodology using a label-free MS^E dataset, QWB calibrated by both recombinant protein standards (QWB-standards) and M-DOSCAT-S (QWB-DOS), and SRM-MS calibrated by M-DOSCAT-S. Data presented are mean values from SPP samples only ± standard error (n = 15).

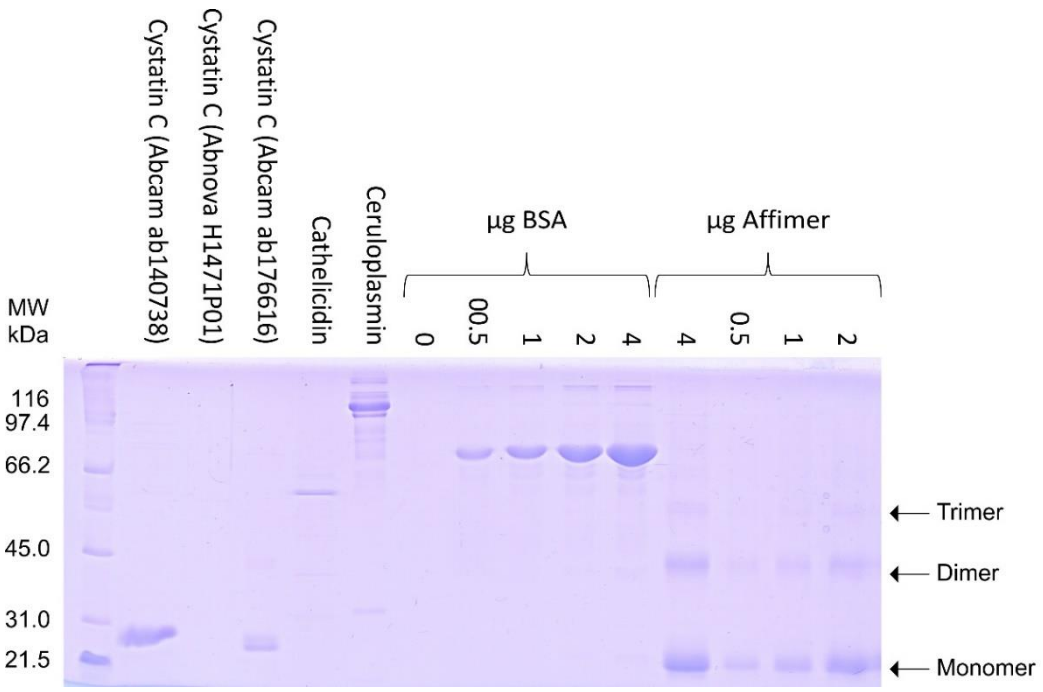


Figure 5.30 Quantification of recombinant standards. Recombinant Cystatin C, cathelicidin and ceruloplasmin alongside albumin and affimer standards were analysed on a 12% SDS-PAGE gel.

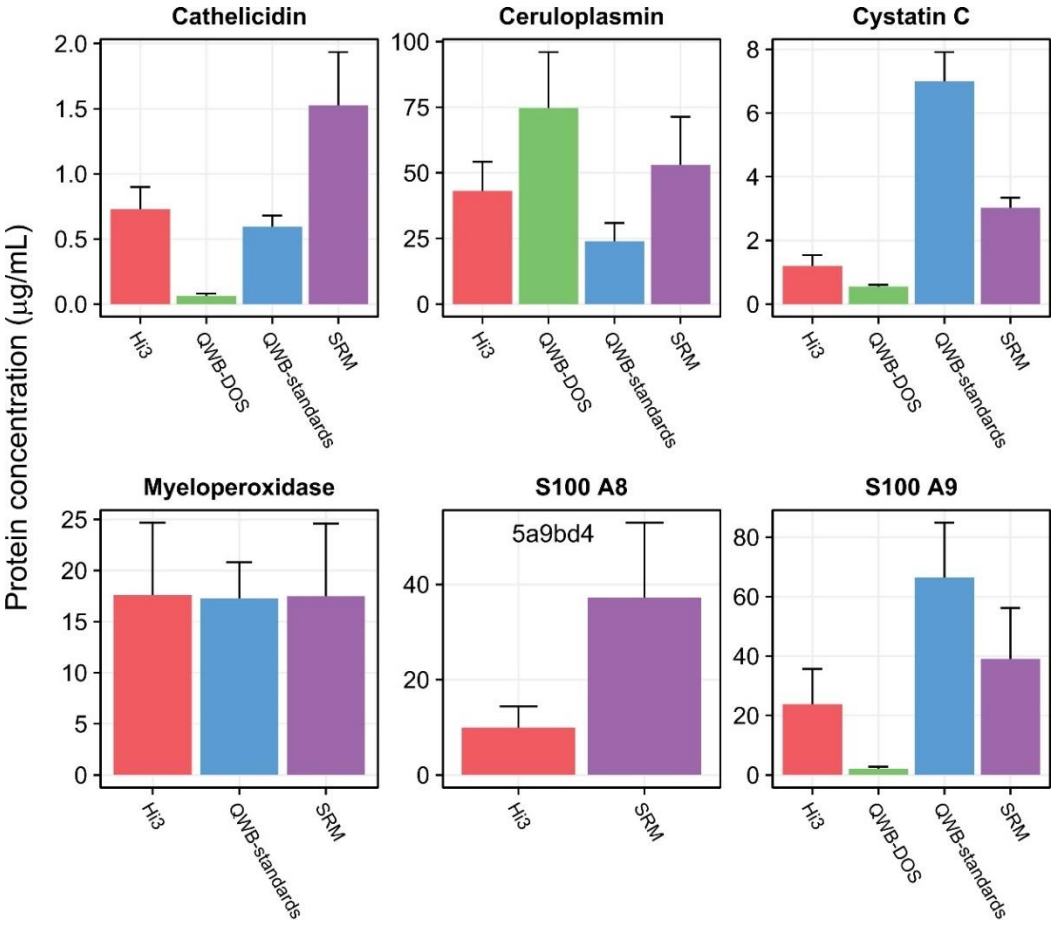


Figure 5.31 Comparison of quantification from different datasets with correct protein standard abundances. Protein level quantification values as calculated by Hi3 methodology using a label-free MS^E dataset, QWB calibrated by both M-DOSCAT-S and recombinant protein standards with corrected concentrations, and SRM-MS calibrated by M-DOSCAT-S. Data presented are mean values from SPP samples only ± standard error (n = 15).

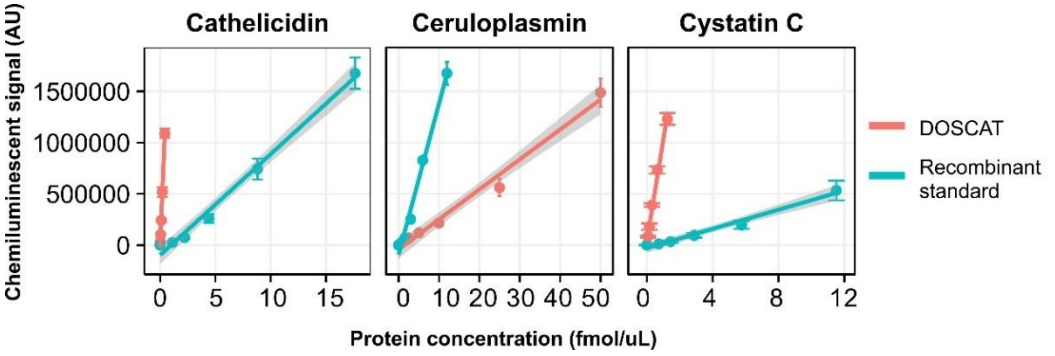


Figure 5.32 Standard curves in QWB from M-DOSCAT-S and recombinant standards. Standard curves resultant from dilution series of recombinant protein standards and M-DOSCAT-S analysed by cathelicidin, ceruloplasmin and cystatin C antibodies in automated capillary WB, processed using Compass software.

Until this point DOSCAT technology had been used to quantify a set of well characterised target proteins in a relatively small sample size. This work both demonstrates the effectiveness and highlights the issues of DOSCAT technology to calibrate a different set of six potentially clinically relevant proteins in human samples over an expanded sample size. In contrast to the NF-κB work, SRM-MS proved more effective in detecting and quantifying target proteins. This was due to good peptide selection and under-performing antibodies that either did not detect proteins with required sensitivity or had poorly characterised epitopes so that M-DOSCAT-S was not detected. Protein fold changes measured in both assays were consistent between one another and with previous measurement by quantitative proteomics and western blotting (Gómez-Baena et al., 2017), adding confidence to the DOSCAT methodology and further validity to the biomarker panel as a diagnostic test. There is still further work required to validate the biomarker panel for use as a diagnostic test, however. Future validation studies will have to be carefully designed and used with an increased sample size so that the study is of sufficient statistical power for results to be deemed clinically relevant (Alonzo et al., 2002; Skates et al., 2013).

In addition to further demonstrating the utility of DOSCAT technology in a clinically relevant area, this work also compares absolute values derived by DOSCAT calibration to values generated by Hi3 quantification and QWB calibrated by recombinant standards from datasets produced in a previous complementary study (Gómez-Baena et al., 2017). This aids in improving accuracy and evaluating the true concentration of the target analytes in CSF. The comparison highlights a substantial difference in absolute quantification values derived from QWB calibrated by M-DOSCAT-S and the three other independent measurements. Why this is the case is difficult to understand. If an inaccurate amount of M-DOSCAT-S was loaded in QWB experiments so in actuality there was more standard loaded than thought, quantification of the analyte would be artificially low, as was observed in the results. This is unlikely, though, as inaccurate standard concentration would lead to a systematic error in analyte quantification that would be consistent across each protein, which was not reflected in the data (Figure 5.26). Moreover, M-DOSCAT-S concentration was carefully determined against a Glu-Fib peptide standard and workflows were designed to minimise losses of standard.

The same amount of M-DOSCAT-S or recombinant protein effects a very different signal in capillary WB when the same antibodies are used (Figure 5.32). The presence of extra unexpected antibody binding sites in the M-DOSCAT-S sequence or antibodies binding to the epitope in M-DOSCAT-S with greater affinity than the recombinant standard (perhaps due to interference from nearby residues) would explain the different signals. However, both cathelicidin and cystatin C antibodies are monoclonal, therefore would not be expected to contain populations of non-specific antibodies that may bind to other regions in the DOSCAT like a polyclonal antibody might. Incidentally, the ceruloplasmin antibody was polyclonal, yet the same differences in signal were not observed. Additionally, the full sequence for Cystatin

C was included in the M-DOSCAT-S sequence so the epitope sequence context was the same between standard and analyte.

Another explanation, based on previous observation based on automated capillary WB of DOSCAT proteolytic fragments, is that epitopes in smaller proteins could be compromised by UV cross linking of residues to the capillary surface. If a greater proportion of the epitope in endogenous analyte has been compromised by this chemical modification, then the resultant signal would be lower, accounting for the observed results. Both cystatin C (15.8 kDa) and cathelicidin (19.3 kDa) are much smaller than M-DOSCAT-S (65.0 kDa), whereas ceruloplasmin (122.2 kDa), for which M-DOSCAT-S signal was lower than recombinant protein, is larger, adding further credence to this hypothesis. Whilst this is far from conclusive evidence, it does warrant further investigation. Classic western blot analysis of the same molar amounts of standard and M-DOSCAT-S will allow for the comparison of signal when there is no possibility of epitopes being affected by modifications.

By comparisons with other datasets, it might be concluded that QWB-DOS data are not accurate, however, values obtained by the other three methodologies agree well. Although the remaining three approaches all offer slightly different quantitative values, they are in general agreement. SRM-MS calibrated by M-DOSCAT-S is arguably the most robust data set and so can be the considered the most accurate. This is because quantification is based on two or three peptides, each measured by multiple transitions and against a well characterised stable isotope labelled standard.

The overall agreement of Hi3 quantification with the other datasets is consistent with previous findings comparing Hi3 with quantification using QconCATs and immunoassays (Smith et al. 2016; Kramer et al. 2015; Carroll et al. 2011). Hi3 quantification is consistently lower, though, and the difference is more pronounced in smaller protein such as S100A8 and S100A9. Underestimation of protein levels using Hi3, especially in smaller proteins, has previously been reported (Carroll et al. 2011; Kramer et al. 2015; Ahrné et al. 2013), and can result from miscleaves or poor ionisation efficiencies in the most intense three peptides selected. Smaller proteins would have fewer potential peptides to choose from, so these effects would be amplified leading to less accurate quantification. It has also been shown that different bioinformatic pipelines can affect accuracy of the data (Ahrné et al., 2013; Rami et al., 2017), so it is possible that reprocessing of the data may yield a different result.

Although performing poorly when calibrated by M-DOSCAT-S, QWB values calibrated by recombinant standards agreed fairly well with both MS methodologies, validating the use of automated capillary WB for quantitative assays. This was only true, though, after protein standard concentrations were corrected from the inaccurate information given by the manufacturers, demonstrating the need to perform in-house validation of protein concentration. It would be anticipated that the values for S100A9 would decrease if calibrant

concentration was also adjusted; this was not possible here as stocks of the standard had been exhausted.

5.3. Conclusions

The work presented in this chapter is another example of how DOSCATs can be used as multiplexed standards to facilitate the quantification of target proteins. Building on previous work that has demonstrated the proof of principle and initial success of the approach, a DOSCAT has now been used to quantify a putative biomarker panel in complex human CSF samples. This is significant given the trend towards using panels of biomarkers rather than single proteins; it is easy to see the benefits a multiplexed standard would have over individual protein or peptide standards in the routine quantification of such panels. Having said this, the difference in quantification values obtained by SRM-MS and automated QWB is a result that requires resolution to ensure future confidence in using DOSCATs for calibration. Forthcoming studies should seek to determine whether this difference in results between analytical platforms is due to inadequacies in the DOSCAT principle and/or design, or if it is an artefact of using the automated capillary WB system, Wes. If it is the former, a re-think on epitope selection, experimental design and workflows will be required and if it is the latter the use of classic WB over automated capillary WB must be considered, unless a workaround can be found.

Despite the WB results requiring further clarification, SRM-MS data agreed very well with label-free MS and QWB datasets from a previous study. This gives general confidence in automated capillary WB as a platform for quantification, even if data presented in this chapter exposes some unresolved issues. For some proteins, SRM-MS and Hi3 data agree so well that one could question the purpose of performing time-consuming targeted assays using expensive stable isotope labelled reagents when fast, seemingly accurate quantification is possible by label-free proteomic approaches. Indeed, studies evaluating the merits of label-free quantification have found it performs favourably in terms of accuracy and reproducibility, although not to the same level as SRM-MS (Distler et al., 2016; L et al., 2013; Lawless et al., 2016). However, clinical validation of biomarkers, to date almost exclusively performed by ELISAs (Drabovich et al., 2015), requires a level of accuracy and reproducibility that label-free workflows cannot provide. Therefore, despite how well label-free workflows can perform, targeted MS and immunoassay approaches alongside highly quality and well characterised reference standards such as DOSCATs are still required and will continue to be in the future.

Throughout the optimisation work described in this chapter more unforeseen challenges in the deployment of DOSCATs have been recorded, namely the formation of higher order aggregates post-purification. A viable strategy has been devised to overcome this issue, using DTT and IAM to reduce and alkylate thiol groups in cysteine residues so that they cannot form disulphide bonds, which is postulated to cause the aggregation. The knowledge that

aggregation is a potential problem, though, can aid the design process in future DOSCAT builds in an attempt to minimise the possibility of aggregation occurring in the first place.

Some of the aims of the chapter were not realised, namely the generation of an anti-DOSCAT antibody and a satisfactory conclusion to the restricted proteolysis work. The failure of the immunogen protein to express prevented the development of the antibody from ever being attempted. As some of the commercial antibodies failed to detect DOSCAT, it would have been interesting to see if an anti-DOSCAT antibody could have fared better. It would be certainly worth attempting expression of the immunogen again in future work, using the sequence reshuffle or cell-free expression strategies described earlier. Optimising conditions for successful restricted proteolysis should also be a goal for work in the near future. As well as realising the aims of giving improved flexibility and multiplexing capabilities, WB of proteolytic fragment, with the same epitope but smaller size of intact DOSCAT protein, could resolve the issue of whether signal is reduced in smaller protein when analysed by capillary WB. This could have big implication for further use of automated capillary western blotting in the DOSCAT workflow.

Chapter 6: Commercialisation of DOSCAT technology

6.1. Introduction

6.1.1. Trends in technology transfer

It has long been acknowledged that the commercialisation of publicly funded, university based research delivers considerable economic and social benefits in terms of industrial growth, job creation, increased funding opportunities for universities the creation of products and services that improve public health (Caulfield et al., 2015; Salter et al., 1999). Commercialisation of academic research is most often achieved through licensing of intellectual property (IP) to existing companies or setting up new spin-out companies. There are now numerous examples of products arising from basic scientific research including recombinant protein expression technology (Ratzkin, 1977; Vapnek et al., 1977), the 'super material' graphene (Novoselov et al., 2004) and CRISPR-CAS9 gene editing (Jinek et al., 2012).

The positive effects of commercialising research have not gone unnoticed by world governments; indeed there has been a concerted effort at a statutory and institutional level to encourage the transfer and commercialisation of public research in order to drive economic growth (Mowery et al., 2006). A fundamental instrument for the commercialisation of a technology by universities and researchers is the patenting of IP linked to an invention. Prior to the 1980s, patents derived from public research were owned by the government, who licensed out the technology on a non-exclusive basis. This resulted in a low take up of product commercialisation (due to non-exclusivity), with only 5% of patents owned by the USA government being used by industry (Schacht, 2012). The legislative groundwork to overcome this was contained in the Bayh-Dole Act, passed in the USA in 1980 (Public Law 1980). The Act permitted universities or small companies to have ownership of an invention derived from public funding rather than the government. This enabled institutions to form exclusive licensing agreements with industry, incentivising them to commercialise technology to generate revenue streams. Moreover, the private sector had a greater incentive to develop inventions as they have the exclusive rights to market them, and governments benefit from increased tax revenue generated from a stronger private sector.

Since the passing of the Bayh-Dole Act the number of patents of issued to universities increased exponentially (Berkeley, 2004) and the legislation has now been replicated in most countries across the world (Meyer et al., 2007; OECD, 2013). Since then governments have introduced additional regulation to create a clear legal framework for IP ownership and extra policies to incentivise universities to file patents. For example, some government have introduced a legal 'grace period' where disclosure of an invention in a journal or conference is permitted 6-12 months prior to applying for a patent (Edmondson, 2013). Although this is a growing trend it is not yet embraced by all countries, so great care should still be taken prior to disclosure depending on location.

As well as IP legislation, governments have also been proactive in forming bridging organisations to facilitate the transfer of technology and knowledge between academia and industrial partners. The formation of technology transfer offices (TTO) has taken place in nearly all research institutions and they have been credited with the increase in IP ownership and entrepreneurial activity in institutions (Grimaldi et al., 2011; Siegel et al., 2007). At a basic level, TTOs act to identify commercialisation opportunities arising from academic research and assist in the transfer of this knowledge to the private sector, where it can be translated into products and services. This is commonly achieved by performing administrative tasks such as patent filing and IP portfolio management, however some TTOs have expanded their activities to providing seed funding for start-up companies and proof of concept studies (OECD, 2013).

As well as government policy encouraging commercialisation, there has also been a push at the institutional level to change the culture within the academic community to encourage researchers to be more entrepreneurial, disclose inventions and begin the commercialisation process. Disclosure refers to a confidential document describing the invention in detail that can be used by the university or a patent attorney to determine whether patent protection can and should be sought. Researchers may not disclose their inventions due to lack of commercial-mindedness (i.e. not realising the commercial value of their research) or though concerns about delays to publications or personal perception by academic peers (Wright et al., 2012). One obvious way to encourage researchers to disclose is to offer a financial incentive. Most often a rate of revenue generated from IP will be offered, although other incentives such as a lump-sum payment, promise of additional funding or career progression have also been used (Lach et al., 2008; Zuniga, 2011). Another tack institutions have taken is to foster an entrepreneurial culture among staff and students. Workshops, mentoring, seminars and business plan competitions, such as Biotech YES in the UK (BBSRC, 2017), have all been used to educate and create skills for entrepreneurship (A. Nelson et al., 2010).

6.1.2. Intellectual property rights

Intellectual property refers to an idea or creation of the mind, and is assigned to an individual by law. In the process of scientific research it is inevitable that IP will be created, which for the purpose of commercialisation needs to be properly protected and managed from an early stage, much before any products come to market. Therefore, IP strategy, which simply refers to a plan to acquire and extract maximum value from IP, is a vital consideration during the commercialisation process. There are several types of intellectual property rights (IPRs) (summarised in Table 6.1) that can be granted to the creators of IP, affording protection and exclusivity when bringing the idea to market. Usually more than one type of IPR is employed in an IP strategy.

Table 6.1 Most common types of intellectual property rights (based on UK law).

IPR type	Protection offered	Cost	Length of protection
Patent	Exclusive right to inventions and products	Filing, legal and ongoing renewal fees can total £000s	20 years from filing date
Copyright	Original works e.g. literacy creations, video/sound recordings, art, software	Free	Life of the creator + 70 years
Trademark	Product names, brands or company logos	UK registration is approximately £200, but with international registration and ongoing legal fees can be £1000s	Must be renewed every 10 years
Registered designs	How a product appears e.g. specific packaging including colours, patterns etc.	Up to £150	Must be renewed every 5 years up to 25 years
Unregistered design rights	Shape or orientation of an original design	Free	10 years after product was first sold or 15 years after it was created
Trade secrets	A formula, process, design etc. that is not readably identifiable without confidential 'know-how'	Free	As long as a company can guard its secret

Patent protection offers a monopoly in the market for a fixed period of time, and so gaining one is usually (although not exclusively) a key strategic step in the successful commercialisation of an idea. A patent acts as a contract between state and inventor, granting exclusive rights to manufacture and sell an invention. In return for this, a full and detailed disclosure of the invention must be made available in the public domain. There are, however, criteria that must be met for a patent to be granted. Firstly, the invention must be novel, which is ascertained by assessing it against the "State of the Art". This means everything that it is the public domain in any way (e.g. written, oral, in use) before the date the patent is filed. This can include the invention itself, so if it is disclosed to the public domain (e.g. via publication, conference talk or poster etc.) a patent cannot be filed. Secondly, the invention must contain

an inventive step. This means that it must not be obvious to a skilled person trained in the art who has knowledge of the current state of the art. Inventiveness can be hard to assess and is to an extent subjective, so the counsel of a Patent Attorney is often sought. Finally, an invention must have an industrial application and be in a tangible, physical form. It cannot be, for example, a mathematical theorem, a computer program or an artistic piece of work.

If an invention is deemed suitable a patent application can be filed and the process can begin. The application process is long and costly, taking several years between filing and granting of the patent (Figure 6.1). During the application process a detailed description of the invention with claims and technical drawings is submitted to the Patent Office, usually with the assistance of a Patent Attorney. Initial searches against the prior art are carried out to establish novelty and inventiveness, and if no infringements are found the application is taken forwards. If the invention is to be patented in international territories, this must be done within 12 months of the original application. Further, more detailed, examination of the application then takes place before publication and eventual granting of the patent can occur. Once the patent is granted ongoing costs are incurred through renewal fees for each territory in which the patent is filed and any legal challenges that may have to mounted because of third party infringement. Owing to the time and financial implications businesses (particularly small ones) must consider very carefully whether it is worth pursuing patent protection. A lot of money can be invested without the outcome of the patent being granted and, if granted, enforcement may not be possible due to the substantial legal costs involved. It is therefore worth examining whether other forms of IPR can be used to protect the idea.

Copyright and unregistered design rights are automatically assigned to a creation; however, they only offer protection for written/artistic works and original product shapes respectively. Trademarks and registered design rights can, for a relatively small fee, act to protect a strong brand identity for a product in terms of product names and logos and how the product packaging appears. These IPRs can be very valuable in certain industries (e.g. music, as songs cannot be patented), or if a product has an extremely strong brand that elicits a high level of consumer trust and confidence (which may confer an advantage over any ‘copycat’ competitors). However, these rights will not prevent a competitor using the fundamental idea of an invention in a different form. A patent may not be necessary if an invention relies on a formulation or process that is not obvious or relies on a large amount of research and development. In these cases, keeping this a trade secret will prevent competitors from copying the idea. However, this may be difficult in a publicly funded academic setting as there is pressure to publish all research outcomes in journals.

6.1.3. Aims

DOSCATs have the potential to be commercialised into a valuable product within the life science and biotechnology industry. The technology has been developed in collaboration with an industrial partner, Badrilla, a life sciences company that specialises in producing high quality affinity reagents.

This chapter aims to look at some of the key considerations to make when devising a strategy to exploit DOSCATs for commercial gain. The process for protecting DOSCATs though intellectual property rights will be explored, as will the procedure and challenges in bringing DOSCAT technology to market.

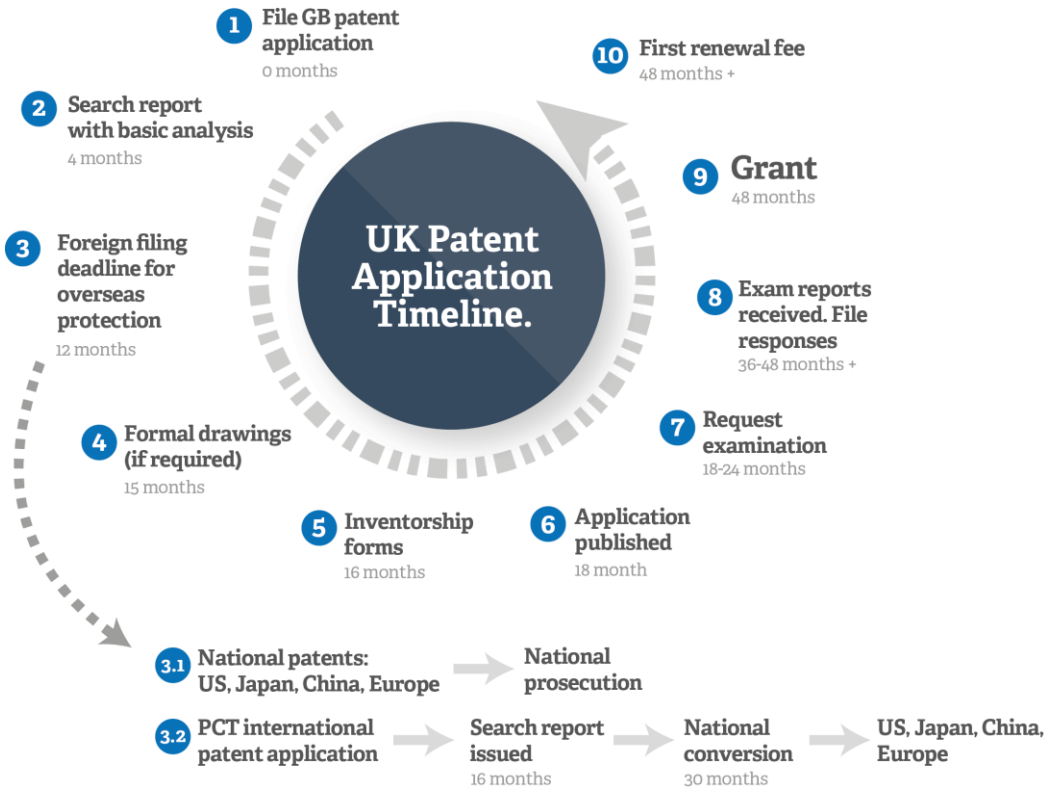


Figure 6.1 Timeline from patent filing to granting in the UK. Figure replicated from Albright IP (2017).

6.2. DOSCAT IP strategy

For DOSCAT technology there are multiple types of IPRs with which it can be associated. Patent protection would be the most valuable in that it would prevent competitors from copying the idea, thus maximising the commercial opportunity. This section describes a series of work carried out to investigate whether DOSCAT was a patentable invention, and what form the IP strategy for the product should take.

As previously described, an invention is patentable only if it is novel and there is no prior art relating to it. Before embarking on a costly patent application process, it is vital to perform searches to ascertain the state of the IP landscape and evaluate if there are any existing disclosures in the scientific literature or patent library that would prevent a patent being granted, or restrict the freedom to operate; that is the right to market a technology without infringing on another's IP.

The first step taken in embarking on the prior art search was to define the core technology of the invention and list all possible uses. From this, relevant search terms can be easily identified. For the case of DOSCATs, the core technology was defined as a calibration standard material that comprises calibrators for assays conducted on different analytical platforms. This does not necessarily limit the invention to SRM and western blotting, so every possible use for DOSCATs were thought of. Multiple variants of immunoassays and targeted mass spectrometry were identified, which were included in searches and could be encompassed in the wording of a future patent (Table 6.2).

As described extensively in this thesis DOSCAT can be used in western blotting with a variety of detection methods and platforms, as well as with targeted SRM-MS assays. Additionally, although not yet tested, it is anticipated DOSCAT could also be used in other targeted MS assays such as PRM. ELISAs are an often-used quantitative immunoassay technique and DOSCATs in their current configuration would be compatible with a direct ELISA experiment. Many ELISA assays use the sandwich format, in which two antibodies are used; one to capture the antigen and another to detect it. This will require two separate epitopes or domains to be incorporated into the DOSCAT standard. The epitopes in the standard must be available to bind the antibodies simultaneously and in an equivalent manner compared to the endogenous protein. For this, strategies to fold DOSCAT so that it is equivalent to the endogenous protein must be developed and evaluated.

Additionally, DOSCATs could be used in planar and suspension multiplex ELISA assays described in detail in Chapter 1. The same principle applies that DOSCAT would need to contain epitopes for capture and detection antibody recognition. Moreover, it is likely that multiple DOSCAT proteins will be required if many analytes are to be measured in multiplex.

Guided by the definition of the core technology and the list of uses for DOSCAT, search terms were derived and the literature was searched. This returned numerous results for each term

(Table 6.3); for some there were too many results to analyse so only top 50 results were analysed when results were ordered by 'relevance'. Each paper was analysed to assess whether it was pertinent to DOSCAT technology and if so whether it was likely to constitute prior art that could negatively affect patent granting. Papers were considered pertinent if they combined immunoassays and mass spectrometry in a single workflow to improve data quality or if they described a single calibration material in a multiplexed assay. Contained in the results were multiple descriptions of immuno-MS workflows (described in Chapter 1), in which protein or peptide antibodies are used to enrich for specific proteins prior to targeted SRM analysis using SIL peptide or protein standard. Although these bring the concept of MS and immunoassay together there is no quantitative measurement by immunoassay and no novel calibration material is used and so do not anticipate the DOSCAT concept. There were numerous examples of both immunoassay and MS data being compared to assess the accuracy of the data, but these data were always acquired independently and there was no mention of combining the two workflows in a single experiment. No novel calibration material to allow for quantification by MS and immunoassay could be identified in any of the results.

Table 6.2 Technologies to which DOSCAT could be applied.

Application	**Comments**
Western blot	Describes all variations of western blots e.g. different transfer (wet, semi-dry), detection (ECL, fluorescent) and imaging (film, CCD camera) methodologies. Also includes newer automated capillary based Simple Western technology.
ELISA	DOSCAT protein will require additional engineering
Planar multiplex assay	
Suspension multiplex assay	
Targeted mass spectrometry	Stable isotope labelled DOSCAT digested and analysed at a peptide level; applicable in a range of MS methods such a SRM, PRM or SWATH

Table 6.3 DOSCAT prior art search. Search terms and number of results for literature search using Web of Science, using the Core Collection database with all Citation Indexes from 1900 to 2017, accurate as of September 2017.

Web of Science search term	Results
TOPIC: (absolute quant*) AND TOPIC: (mass spec*)	4,936
TOPIC: (multiplex* immunoassay)	2,533
TOPIC: (selected reaction monitoring) AND TOPIC:(mass spec*) AND TOPIC: (quant*)	2,505
TOPIC: (multiplex* immunoassay) AND TOPIC: (quant*)	824
TOPIC: (targeted mass spec*) AND TOPIC: (selected reaction monitoring)	800
TOPIC: (orthogonal calibration standard)	349
TOPIC: (absolute quant* western blot*)	315
TOPIC: (multiplex* immunoassay luminex)	216
TOPIC: (protein standard concat*)	130
TITLE: (quant* western blot*)	130
TOPIC: (quant* multiplex standard calibration)	78
TOPIC: (QconCat)	76
TOPIC: (multiplex* immunoassay luminex) AND TOPIC: (quant*)	70
TOPIC: (western blot calibration)	63
TOPIC: (absolute quant*) AND TOPIC: (immunoblot)	40
TOPIC: (multiplex* protein calibration standard)	36
TOPIC: (multiplex* immunoassay) AND TOPIC: (absolute quant*)	28
TOPIC: (western blot calibration curve) OR TOPIC: (immunoblot calibration curve)	27
TOPIC: (multiplex* immunoassay calibration standard)	22
TOPIC: (epitope concat*)	21
TOPIC: (dual calibration standard) AND TOPIC: (protein quant*)	17
TOPIC: (orthogonal calibrat* standard quant* assay)	16
TOPIC: (orthogonal multiplex calibration standard)	14
TOPIC: (protein quant* calibration standard) AND TOPIC: (orthogonal)	12
TOPIC: (parallel multiplex calibration standard)	11
TOPIC: (multiplex* immunoassay luminex) AND TOPIC: (absolute quant*)	3

Patent searches were carried out using the same search terms as the literature search. The results of the patent search unveiled 8 key patents that could be judged as prior art and so necessitated further evaluation (Table 6.4).

Two of the patents were anticipated to appear in the search. The first (EP 1904517) relates to QconCAT technology, which has been described and discussed at length in previous chapters. The patent protects the idea of an artificial protein comprising a sequence of quantotypic peptides separated by an enzymatic or chemical cleavage sequence. The other patent, held by Badrilla (EP 1711835), describes a scaffold material containing multiple sites or domains to which a target moiety is covalently attached. The moiety could be a protein, peptide, antibody epitope, or a sequence of nucleic acids, and is designed to mirror a target moiety present in the sample. Therefore, the entire structure can be used as a multiplexed calibrator for numerous different assays.

Patent US 20140322732 describes an assay or kit for the diagnosis, therapy, prognosis or patient stratification of prostate cancer based a biomarker panel of four proteins contained in human serum, blood or plasma samples. It was deemed relevant as it describes a biomarker discovery workflow using label-free discovery proteomics followed by a validation step in which targeted MS or an immunochemical assay could be used. In the patent text SRM, western blotting and ELISAs are all specifically mentioned. This patent comes close to anticipating the DOSCAT concept as it proposes linking MS and immunoblotting as orthogonal techniques for biomarker validation, despite each technique being employed independently from another. Calibration standards are mentioned for each technique but there is no reference of unifying them to combine the assays.

Two patents in the list are focussed on immuno-MS; that is the selection of proteins or peptides by antibodies (or other affinity reagents) prior to MS analysis. US 20120171782 describes the advantages of immuno-MS in that it offers an extra dimension of selectivity compared to one assay alone. It also depicts how the assay could be quantitative if an internal reference species is combined with the sample. The patent offers protection for the immuno-MS workflow, where proteins are captured, eluted, and analysed by MS at the protein or peptide level. US 20100267069 builds on this, but specifically protects MSIA (mass spectrometric immunoassay) tip technology. Affinity reagents are covalently bound to porous solid supports that are contained within pipette tips through which sample can be passed and subsequently eluted for MS analysis. This increases the throughput and allows for automation of immuno-MS assays. In both these inventions the concept of immunoassays and mass spectrometry have been combined, however, analytical measurements are only ever made by MS. Although the addition of reference standards for quantification is discussed, there are no claims in the patents for a reference standard that can act across the two techniques. It is therefore a tandem use of the two techniques rather than an orthogonal use, and so these patents were not considered to anticipate the DOSCAT concept.

The other inventions are a step beyond immuno-MS as they use both immunoassay and MS as an analytical technique to detect and quantify proteins. US 20080003599 describes a biological microchip for the multiplexed analysis of compounds. The microchip comprises an array of three-dimensional hydrogel elements, each containing an immobilised affinity ligand for immunoassay analysis. Compounds can be captured on the microchip in parallel and subsequently analysed *in situ* by sequential techniques such as immunoassay, MALDI-TOF MS and PCR.

The invention contained in US 2007009292 constitutes a method for the detection and quantification of a specific protein isoform in a complex sample. A proteolytic peptide unique to the isoform is selected, the sample is them digested and the peptide detected by MS and/or immunoassay (ELISA, western blot, dot-blot are all specifically mentioned in the text) using anti-peptide affinity reagents. For quantification, synthetic isoform-specific peptides would be used to build calibration curves for whichever assay was employed.

US20110287446 discloses a novel device that contains an immunoanalytical section and a mass spectrometry section and is designed to eliminate the problem of cross-reactivity in immunoassay. The immunoanalytical section uses antibodies to capture target molecules and quantify them in the form of an ELISA. Quantification is achieved through recombinant standards that are contained within the device. Captured molecules are then released and subsequently analysed by mass spectrometry. This comprises identification of the target molecule plus any non-specific molecule that was captured by the antibody. If non-specific species are found by MS, the quantitative value derived by ELISA can be adjusted based on the ratio of specific analyte to non-specific contaminant.

These three inventions are particularly pertinent with regards to DOSCAT technology as they explicitly advocate using both MS and immunoassay as analytical techniques to improve the identification and quantification of target analytes. In particular, the invention by Kanda et al illustrates an invention designed to improve the accuracy of an immunoassay by using an orthogonal technique. Calibration standards referred to in the patents comprise different material for each assay type e.g. recombinant protein for immunoassay and stable isotope labelled peptide or protein for MS analysis.

In conclusion, the prior art search has revealed that combining immunoassay and MS analysis to improve performance is a known concept and practice. It must be assumed that a person skilled in the art would have knowledge of this. When this is taken in tandem with the description of multiplexed calibration standards for MS (Pratt et al) and immunoassay (Colyer), it could be reasonably argued that a skilled person motivated to find a more accurate method of assay calibration would combine the described teachings and arrive at a product akin to DOSCATs. Therefore, although novel, there cannot said to be a significant inventive step in the creation of DOSCAT technology, thus rendering the invention not patentable.

This decision gives clear guidance to the IP strategy for DOSCAT. Rather than pursuing patent protection for DOSCAT, protection could come from the Badrilla and Polyquant patents. Of course, collaboration and discussion between the two companies must take place so that Badrilla has the freedom to operate, and not risk litigation from infringement of another company's patent. Additional IPR could come in the form of trademarking the DOSCAT names and any design rights that associated with the logo, brand and product design.

Table 6.4 Patents relevant to DOSCAT technology.

Patent code/reference	Title	Inventors	Reference
US 20140322732 A1	Method for biomarker and drug-target discovery for prostate cancer diagnosis and treatment as well as biomarker assays determined therewith	Wilhelm Krek, Igor Cima, Rudolf Aebersold, Ralph Schiess, Thomas Cerny, Silke Gillessen	(Thomas Cerny, 2009)
US 20120171782 A1	Mass spectrometric immunoassay	Randall W. Nelson, Peter Williams, Jennifer Reeve Krone	(R. W. Nelson et al., 2011)
US 20100267069 A1	Analysis of proteins from biological fluids using mass spectrometric immunoassay	Urban A. Kiernan, Eric E. Niederkofler, Kemmons A. Tubbs, Dobrin Nedelkov, Randall W. Nelson	(Kiernan et al., 2003)
US 20080003599 A1	Biological Microchip for Multiple Parallel Immunoassay of Compounds and Immunoassay Methods Using Said Microchip	Ekaterina Dary, Ekaterina Dementieva, Veronika Butvilovskaya, Alexandr Zasedatelev, Alla Rubina, Andrei Stomakhin, Elena Savvateeva	(Angenedt et al., 2002)
EP1904517	Artificial protein for absolute quantification of protein and uses thereof	Julie Pratt, Robert Beynon, Simon Gaskell	(J M Pratt et al., 2009)
US 20110287446 A1	Immunoanalytical method and system using mass spectrometry technology	Katsuhiro Kanda, Makoto Nogami, Izumi Waki	(Kanda et al., 2011)
US 20070092926 A1	Analysis of protein isoforms using unique tryptic peptides by mass spectrometry and immunochemistry	Michail Alterman, Boris Kornilayev	(Alterman et al., 2006)
EP 1711835 B1	Agents for and method of quantifying binding	John Colyer	(Colyer, 2011)

6.3. Bringing DOSCAT technology to market

Bringing a new technology to any market is a difficult task as potential customers must be convinced to change their existing practices and start using a new product. To maximise chances for success, target customers must be identified and their requirements understood. This purpose of this section is to critically interrogate the invention to define where the commercial value within it lies, and to discuss the work required to develop the technology into a marketable product that fulfils the target consumers' needs.

To fully exploit an ideas potential, careful though must be given to an invention's applications, the fundamental uses or unique selling points (USPs) and all the conceivable ways the invention can be applied. Overlap between these areas delivers the benefit to the consumer that can be marketed and sold (Figure 6.2). As part of the prior art search strategy the applications for DOSCAT have already been listed (Table 6.2) and are summarised again in Figure 6.2. The USPs of DOSCAT are, that compared to currently available techniques and products, it confers improved data quality in terms of quantitative accuracy and precision, and that it increases the success of obtaining quantitative data (i.e. one platform can provide data if the other one fails). Also, using one standard to quantify multiple proteins is more financially economical for the end user. DOSCATs could be used in numerous different applications and industries. DOSCATs could be at the basic research level to, for example, quantify entire signalling pathways or protein subunits for complex stoichiometry analysis. They could also be used in quality controls processes to quantify protein contaminates in industrially produced samples or allergens in food products, for instance. Within the diagnostics field DOSCATs could be designed to quantify several proteins in a panel of biomarkers, thus creating a dual diagnostic assay with one standard. It may even be possible to create a point of care diagnostic kit with DOSCAT as the calibration standard for multiple analytes.

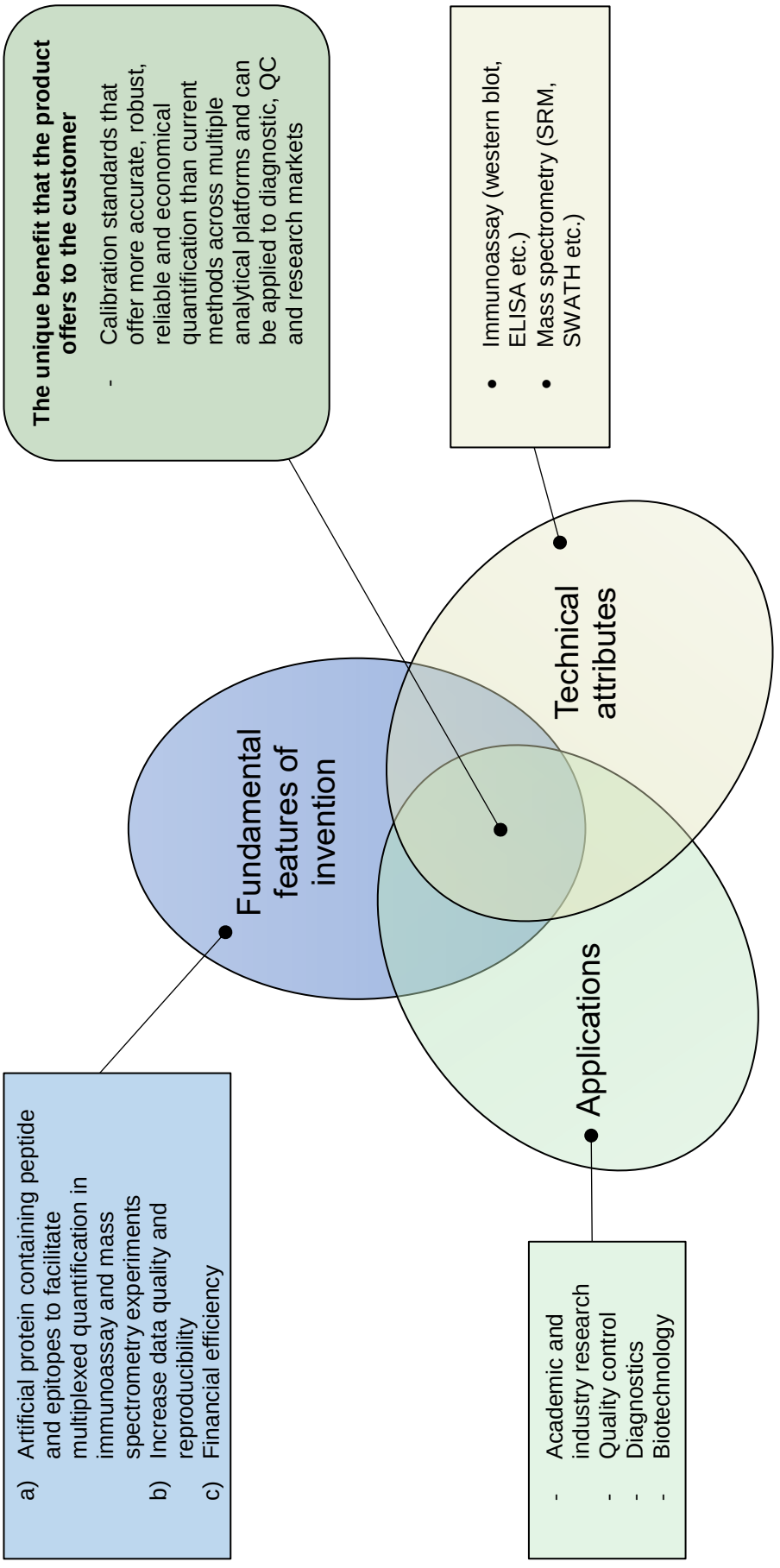


Figure 6.2 Process to define a new invention. Figure replicated from Colyer, personal communication (2017).

This exercise clearly defines a market, a consumer base and the motivation for a consumer to use the product. This must then be placed in the context of the market and the competition. The laboratory diagnostics market alone was worth $72 billion in 2012, with 5-10% annual growth (Wild et al., 2013). Although the market for calibration standards will make up a small percentage of this market, it is still a sizable area to tap in to if the consumer can be persuaded to use DOSCAT technology rather than competing techniques. Within the market the competition consists of reference standards that are identical or very close to the analyte (i.e. recombinant proteins), and of a known identity, concentration and purity. These standards can be grouped into one of three categories (Validation, 1998):

1) Certified reference standards that adhere to the highest regulatory standard (primary standards)
2) Commercially available standards from reputable companies (secondary standards)
3) Custom made standards from non-commercial analytical laboratories

Currently, DOSCATs would fall into group (3). Despite the advantages that DOSCATs can deliver, it is unlikely a consumer would switch from using current practices without clear evidence assays calibrated by DOSCATs can perform as well as what is currently on the market. Therefore, work would be required to validate the DOSCAT standard itself and the quantitative assays that DOSCAT is built for.

A calibration standard must be well-characterised with respect to its identity, molecular weight and purity (Lee et al., 2009). This can be achieved with analytical methods such as HPLC and mass spectrometry. The standard must also be producible in sufficient amounts, with high consistency of production (i.e. low lot-to-lot variability). Stability of the standard should be well understood in terms of freeze-thaw cycles, short term exposure to room temperature and long term stability at storage conditions.

Once characterised, the standard should be validated in the assay in which it will be used (in the case of DOSCAT, automated western blotting and SRM-MS). For diagnostic grade assays, there are key parameters that need to be assessed and determined: selectivity, sensitivity, accuracy and precision (Plant et al., 2014; Valentin et al., 2011; Wild et al., 2013). Each parameter must be individually tested and the results documented.

Sensitivity

Sensitivity is the level at which quantification is not possible due to noise in the system. It is defined by the lower limit of quantification (LLOQ), which is the lowest analyte concentration that can be measured above the system noise. This is determined by measuring serial dilutions of standard spiked into a suitable background matrix until the signal to noise ratio reaches a specific level.

Selectivity

This is the ability to differentiate and quantify an analyte within a sample in the presence of other compounds. Selectivity can be evaluated by measuring at least six biological matrices both blank and with spiked in reference material at different concentrations. The blank matrix sample should not contain interferences and the reference material should be present at the LLOQ.

Accuracy

Accuracy describes the closeness of the measured value of a target analyte to its true value within a sample. It is determined by the analysis of samples containing known amounts of the analyte at a range of levels, with at least five repeat measurements at each concentration. At each concentration, the mean value should be within 15% of the actual value except for the LLOQ where it must not exceed 20%.

Precision

Precision encompasses both repeatability, that is variation when measurements are made one after the other with the same experimental conditions, and reproducibility, which describes variance when experiments are repeated by different analysts at another time and place. Precision should be determined by the repeated measurement of analytes in a biological matrix across the working range of the assay, using at least five measurements per concentrations. Acceptable precision at each concentration is 15% except for the LLOQ where it must not exceed 20%.

For DOSCATs to be used in diagnostic grade assays, there would have to be stringent validation per the above criteria for each analyte and analytical platform. This would create a considerable amount of development work, but result in a product that is superior to its competitors. Alternatively, DOSCAT could be used at research grade, which requires less stringent validation of assays. This would limit the market, but perhaps allow for DOSCATs to a wider range of targets to be created.

6.4. Summary and conclusions

The transfer of technology from bench to a commercial product is important for social and economic reason, and has been encouraged by governments and universities by numerous laws and policies. DOSCAT is a prime example of a technology emerging from a university that can be commercialised and brought to market. Whilst the DOSCAT idea itself is not patentable due to insufficient inventiveness, it is protected by previous patents owned by Badrilla and Polyquant as well as various other types of IPR. From the identification of the target market and consumer base that DOSCAT will be marketed to, it is clear that there is further work required at a commercial level to demonstrate the stability and robustness of the assays that DOSCAT is built for.

Chapter 7: General conclusions

7.1. General conclusions

The primary aim of the work presented in this thesis was to demonstrate that DOSCATs could be designed, expressed and used to calibrate both quantitative western blotting and mass spectrometry assays, uniting the two orthogonal techniques into a single workflow. This has largely been achieved through two deployments of DOSCATs, first targeting the quantification of NF-κB proteins in SK-N-AS cell lysate and secondly, host proteins associated with pneumococcal meningitis in human CSF samples. This has required the development of three DOSCAT proteins, with each iteration further refining design principles in terms of epitope and Q-peptide selection and inclusion of restricted specificity proteolytic sites, as well as solidifying procedures for the successful expression and purification of the DOSCAT protein. As such, it should now be possible for a person skilled in the art to use the design principles laid out in this thesis to design, express and utilise their own DOSCAT to any set of target proteins.

As with any new technology, the development of DOSCATs has not been problem-free, with several challenges emerging that required resolution. Some issues related to the behaviour of the DOSCAT protein itself, such as adsorption to surfaces and irreversible aggregation. DOSCATs are artificial proteins and so will not necessarily behave like proteins that have evolved within the constraints of their local environment. Similar issues have been extensively observed and documented with QconCAT technology (Brownridge et al. 2011; Simpson and Beynon 2012). Both adsorption and aggregation result in a discrepancy between the measured concentration of the standard and the amount of standard that is present in an assay so that there is less standard present than thought, resulting in an overestimation of analyte abundance. These issues were overcome by relatively trivial modifications to the DOSCAT production workflow; addition of Rapigest *SF* to buffer solutions to prevent excessive protein adsorption and the reduction and alkylation of DOSCAT immediately after purification to prevent the formation of disulphide bonds that enhance aggregation.

Reliable selection of antibodies is another challenge when it comes to deploying DOSCAT technology. Across the three DOSCATs made in this thesis, 14 different antibodies were selected, of which 12 recognised endogenous protein and 10 recognised the DOSCAT protein in western blots. This is perhaps reflective of an industry-wide issue of antibody reliability, with many antibodies not meeting requirements in terms of selectivity. The main criteria for selecting antibodies to use with a DOSCAT are that they must a) specifically bind the target analyte with high affinity, and b) bind the linear peptide epitope contained in the DOSCAT sequence. Fulfilling criterion a) is perhaps the easier of the two, as there have now emerged several tools and initiatives to improve validation of commercial antibodies and highlight highly cited or well-validated antibodies. Finding such an antibody that is also accompanied by published data relating to its epitope further diminishes the number of possible antibodies, and even then, as demonstrated in this thesis, it is not always the case that the given epitope data

is accurate. The design and expression of a DOSCAT is a substantial commitment in terms of time and money, and so to discover that it is not recognised by one or more target antibodies represents a considerable waste of resources. Therefore, more work is needed to mitigate against the risk of this occurring. It is perhaps understandable that antibody manufacturers do not want to share epitope data as substantial work can go into finding a suitable and well performing immunogen, making such information commercially sensitive, although this practice does not meet the principles of open and reproducible science (Munafò et al., 2017). Closer collaboration with antibody manufacturers could overcome this hurdle, perhaps leading to DOSCATs being built that compliment well characterised, highly validated and widely used antibodies. Another approach would be to produce antibodies (or alternative protein binders) in-house, although this is costly, requires a high level of expertise and does not automatically resonate with 3R's (reduction, replacement and refinement) initiatives (Törnqvist et al., 2014).

The inclusion of restricted proteolytic sites to shift electrophoretic mobility has had mixed results, with numerous proteases not cleaving DOSCAT specifically and proteolytic fragments not having the same chemiluminescent signal as intact DOSCAT in capillary WBs. The inclusion of proteolytic sites is a supplementary benefit rather than a core component of DOSCAT function, and whether they should be included in the design of future DOSCATs is for debate. Whilst certainly having the potential to add to the multiplexing capability, further work is required to understand the performance of the proteolytic fragments in western blotting. This could be performed using the existing NFκB-DOSCAT-2 rather than creating new DOSCATs. Once this is known, a more informed judgement can be made as to whether they add enough value to warrant inclusion in further DOSCAT designs.

As well as overcoming the technical challenges associated with creating and implementing DOSCATs, another stated aim of this thesis was to compare data generated by QWB and SRM, which has been facilitated by the success of the DOSCAT approach. Using the NFκB-DOSCAT-2, it has been demonstrated that QWB using automated capillary WB instrumentation can produce quantitative data with comparable reproducibility and accuracy comparable to the gold standard MS-based technique. Moreover, DOSCATs have been used to compare SRM to 'classic' western blotting, which although did not quite meet the same analytical standards of automated WB, produced quantitative data that was highly agreeable with SRM approaches. However, the second deployment of DOSCATs to quantify meningitis proteins resulted in a wide discrepancy between QWB and SRM quantification values that remains unexplained. Further experimentation is required to explain this result and determine whether it is a technical issue that can be resolved or a fundamental problem with the DOSCAT approach. This is clearly a vital task and will be discussed further in the future work section.

Now that experience has been accumulated with the use of DOSCATs, an assessment can be made with regards to whether the technology is mature enough to be regularly used by the research community. Each deployment of DOSCAT as described in this thesis has presented

technical challenges that, whilst being ultimately resolved, have taken time and resources that would not be tolerated if DOSCATs were a commercial product. One would assume that as more DOSCATs are made, the frequency of these issues would decrease, but currently it cannot be assured that all possible issues can be anticipated. There also needs to be a resolution as to why WB and MS data generated using M-DOSCAT-S differ and do not give complete confidence in the accuracy of the DOSCAT approach. For these reasons, it can be concluded that whilst DOSCATs offer great potential more work is required to fully understand all possible issues surrounding their manufacture so that they can be produced and deployed with regularity and consistency.

Assuming these technical challenges can be met, there remains the question of how DOSCATs might be applied in research. Some of the results presented in this thesis challenge the previously postulated notion that western blotting is a vastly inferior technique to SRM, and thus cannot be a truly orthogonal technique for data comparison (Aebersold et al., 2013). With careful experimental design and, importantly, a well characterised calibration standard, western blotting can be used with confidence to verify MS data or even replace MS entirely, if the assay can be well validated. The addition of DOSCATs into workflows would be beneficial for studies that require highly accurate and robust quantitative data, for example molecular systems biology studies or experiments elucidating protein complex subunit stoichiometry. It is easy to anticipate the motivation to include DOSCATs in the workflows of such studies to facilitate the orthogonal quantification of targeted MS data by immunoblotting, a common request from reviewers upon publication.

It is also worth considering how DOSCATs might be used in biomarker development workflows and the routine analysis of biomarkers in the clinic. Such workflows follow a well-documented path of discovery, verification and validation stages with a decreasing set of target proteins and an expanding sample size as the pipeline progresses (Parker et al., 2014). Discovery methods, based on shotgun (DDA) MS or more recently DIA methods such as SWATH, are tensioned against targeted assays for the process of verification and validation. These assays can take the form of SRMs for the verification for up to 10-20 markers, but for clinical validation where only 1-5 biomarkers will be studied, immunoassays are still preferred by the clinical community and regulatory bodies. DOSCATs could be used to ease this transition from MS to antibody based assays as the use of a consistent, highly characterised standard between two assays would lessen the immunoassay development time as a well improve confidence in resultant data as generated by the two platforms. There are some issues that would preclude the use of DOSCATs in biomarker discovery pipelines, however. Between the verification and validation stages numerous proteins are dropped from the analysis, so there would be many epitopes built into DOSCATs that would be redundant for the immunoassay stage. Moreover, ELISAs rather than western blots are commonly used for clinical validation studies due to their robustness and high throughput capacity. Therefore, DOSCATs would have to be re-engineered to work in the sandwich immunoassay format that is commonly used in clinical

ELISAs. Having said this, automated high-throughput western blot technologies such as the Wes platform used throughout the work presented in this thesis offer an attractive alternative to ELISA and may be well-placed to eventually replace ELISAs for validation and routine testing of clinical biomarkers. It is clear from this work that the Wes system can be relied upon for quantification and a single system can realistically be used to analyse up to 72 samples daily by a single operator in an unsupervised manner. Although not as high-throughput as ELISAs, protein separation by electrophoresis reduces issues of cross-reactivity that can dramatically affect ELISA accuracy. Some Wes-based assays for clinical applications have already been developed (J.-Q. Chen et al., 2015) and it remains to be seen if they are incorporated into clinical laboratories for routine use.

For much of the research community, MS based techniques for targeted quantification are not considered due to inaccessibility; the high-cost and expertise required to design and execute an SRM assay means that it is only practised by a limited number of research laboratories and core facilities. For many research groups, western blotting remains the primary method for generating quantitative data and MS approaches are not used, so the introduction of DOSCATs would not be necessary or beneficial. DOSCATs could, however, be used to improve the quality of western blot assays so that confidence in quantitative values is on par with those generated by gold-standard MS-based techniques. A DOSCAT could be used to demonstrate that a western blot carried out using specific reagents and a defined SOP could perform equally, in terms of quantitative accuracy and precision, to an SRM assay. Such assay validation and reporting of reagents and operating protocols in a detailed manner is likely to contribute to an increase in reproducibility (Helsby et al., 2013). Validation could be performed commercially and details of these 'DOSCAT-validated' WB made available alongside the reagents (e.g. specific mAbs) and SOPs they were validated alongside. To improve such assays further, a version of DOSCAT containing only epitopes could be supplied to enable absolute quantification and so cross-comparison of outputs with other laboratories.

In a similar vein, DOSCATs could also be employed in the process of antibody validation. Antibody-independent validation has recently been proposed by The International Working Group for Antibody Validation as a method to validate antibody specificity (Uhlen et al., 2016). They propose using targeted proteomics with internal standards to quantify target protein expression across a set of samples with variable expression of the target protein. Western blots can then be performed using the same samples and if relative protein abundances correlate with the targeted proteomics data, specificity of the antibody is confirmed. DOSCATs could be used to improve the robustness of this process by facilitating the absolute quantification of target proteins in both assays and so allowing cross-comparison of data generated by different users in different laboratories.

7.2. Future work

The immediate next stages of work must concentrate on the resolution of different quantitative values for QWB and SRM using M-DOSCAT-S. As explained in Chapter 5, the working hypothesis is that for Wes analysis, epitopes in smaller proteins are compromised by UV cross linking to the capillary surface in a way that does not occur in the larger DOSCAT standard. This can easily be assessed by classic western blotting of an equimolar mixture of recombinant protein and DOSCAT. If this experiment results in equal signal between DOSCAT and analyte rather than a larger signal for DOSCAT as observed in Wes results, this would demonstrate that the DOSCAT principle still stands and that there are technical issues with using capillary WB that need to be addressed. An understanding of the crosslinking chemistry, which is currently proprietary, would be required to see if there was a way to circumvent the problem, perhaps by engineering the DOSCAT so it would bind to the capillary wall in the same manner as the analyte. If this is not possible other methods of classic western blotting or technologies such as µWesterns would have to be explored for use with DOSCATs. Alternately, experiments might show that capillary WB technology is not an issue and the antibodies have a different affinity for M-DOSCAT-S compared to analyte proteins. This would present a more serious issue as it would undermine the principle that DOSCATs are based on. Further experiments to pinpoint the flaw in the design and see if it applies to other DOSCATs would then have to be carried out.

If these technical difficulties can be overcome the next stage of work would focus on ensuring DOSCATs can be designed and manufactured consistently and maintain stability during an extended storage period. As discussed, a better understanding of antibody selection is required to increase the reliability that DOSCATs that will be detected by immunoassay. Further work will also focus on improving the solubility and stability of DOSCAT, which is especially important if DOSCATs are to be made into a commercial product. Different buffer solutions, solubility tags or vaccine preservation approaches could all be employed towards this goal.

All the work presented in this thesis has been towards developing DOSCATs for use with western blotting, but there are many different immunoassays for which multiplexed calibration standards could be used with, as discussed in Chapter 6.2. ELISAs are a well-known immunoassay and are extensively used in clinical laboratories, so there is clear motivation for developing ELISA-compatible DOSCATs. Significant re-engineering of DOSCATs would be required to include epitopes for capture and detection antibodies and, perhaps most challengingly, present these epitopes in their native state as ELISAs are not performed under denaturing conditions. A possible approach might be to use an inflexible scaffold region, with capture and detection epitopes presented at either side, with Q-peptides contained within the scaffold and released upon proteolysis.

Chapter 8: References

Abdullah, N., & Chase, H. A. (2005). Removal of poly-histidine fusion tags from recombinant proteins purified by expanded bed adsorption. *Biotechnology and Bioengineering*, *92*(4), 501–513. https://doi.org/10.1002/bit.20633

Acharya, P., Quinlan, A., & Neumeister, V. (2017). The ABCs of finding a good antibody: How to find a good antibody, validate it, and publish meaningful data. *F1000Research*, *6*, 851. https://doi.org/10.12688/f1000research.11774.1

Advance, T. (2005). An alternative tandem af nity puri cation strategy applied to Arabidopsis protein complex isolation. *The Plant Journal*, *41*(5), 767–778. https://doi.org/10.1111/j.1365-313X.2004.02328.x

Aebersold, R., Burlingame, A. L., & Bradshaw, R. A. (2013). Western Blots *versus* Selected Reaction Monitoring Assays: Time to Turn the Tables? *Molecular & Cellular Proteomics*, *12*(9), 2381–2382. https://doi.org/10.1074/mcp.E113.031658

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, *422*(6928), 198–207. https://doi.org/10.1038/nature01511

Aebersold, R., & Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature*, *537*(7620), 347–355. https://doi.org/10.1038/nature19949

Ahrné, E., Molzahn, L., Glatter, T., & Schmidt, A. (2013). Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics*, *13*(17), 2567–2578. https://doi.org/10.1002/pmic.201300135

Albright IP. (2017). UK patent application procedure.

Allan L. T., Tosetto, M., Miller, I. S., O'Connor, D. P., Penney, S. C., Lynch, I., Keenan, A. K., Pennington, S. R., Dawson, K. A., & Gallagher, W. M. (2006). Surface induced changes in protein adsorption and implication for cellular phenotypic responses to surface interaction. *Biomaterials*, *27*(16), 3096–3108. Retrieved from http://dx.doi.org/10.1016/j.biomaterials.2006.01.019

Alonzo, T. A., Pepe, M. S., & Moskowitz, C. S. (2002). Sample size calculations for comparative studies of medical tests for detecting presence of disease. *Stat Med*, *21*(6), 835–852. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11870820

Alterman, M., & Kornilayev, B. (2006, October 13). Analysis of protein isoforms using unique tryptic peptides by mass spectrometry and immunochemistry. Google Patents.

Anderson, N. G., & Anderson, N. L. (1996). Twenty years of two-dimensional electrophoresis: past, present and future. *Electrophoresis*, *17*(3), 443–453.

Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., & Pearson, T. W. (2004). Mass Spectrometric Quantitation of Peptides and Proteins Using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *Journal of Proteome Research*, *3*(2), 235–244. https://doi.org/10.1021/pr034086h

Anderson, N. L., Jackson, A., Smith, D., Hardie, D., Borchers, C., & Pearson, T. W. (2009). SISCAPA peptide enrichment on magnetic beads using an in-line bead trap device. *Molecular & Cellular Proteomics : MCP*, *8*(5), 995–1005. https://doi.org/10.1074/mcp.M800446-MCP200

Angenedt, P., Glokler, J., Murphy, D., Lehrach, H., & Cahill, D. J. (2002). Biological microchip for multiple parallel immunoassay of compounds and immunoassay methods using said microchip. *Anal. Biochem*, *309*, 253–260.

Anjo, S. I., Santa, C., & Manadas, B. (2017). SWATH-MS as a tool for biomarker discovery:

From basic research to clinical applications. *Proteomics*, *17*(3–4). https://doi.org/10.1002/pmic.201600278

Antharavally, B. S., Carter, B., Bell, P. A., & Mallia, A. K. (2004). A high-affinity reversible protein stain for Western blots. *Analytical Biochemistry*, *329*(2), 276–280. https://doi.org/10.1016/j.ab.2004.02.049

Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., & Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on Escherichia coli. *Journal of Proteomics*, *75*(17), 5437–5448. https://doi.org/10.1016/j.jprot.2012.06.020

Backes, B. J., Harris, J. L., Leonetti, F., Craik, C. S., & Ellman, J. a. (2000). Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. *Nature Biotechnology*, *18*(2), 187–193. https://doi.org/10.1038/72642

Banerjee, S., & Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry*, *2012*, 1–40. https://doi.org/10.1155/2012/282574

Bantscheff, M., Lemeer, S., Savitski, M. M., & Kuster, B. (2012). Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Analytical and Bioanalytical Chemistry*, *404*(4), 939–965. https://doi.org/10.1007/s00216-012-6203-4

BBSRC. (2017). Biotechnology Young Entrepreneurs Scheme (YES) competition.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533. https://doi.org/10.1038/483531a

Bennett, R. J., Simpson, D. M., Holman, S. W., Ryan, S., Brownridge, P., Eyers, C. E., Colyer, J., & Beynon, R. J. (2017). DOSCATs: Double standards for protein quantification. *Scientific Reports*, *7*, 45570. https://doi.org/10.1038/srep45570

Berglund, L., Björling, E., Oksvold, P., Fagerberg, L., Asplund, A., Al-Khalili Szigyarto, C., Persson, A., Ottosson, J., Wernérus, H., Nilsson, P., Lundberg, E., Sivertsson, Å., Navani, S., Wester, K., Kampf, C., Hober, S., Pontén, F., & Uhlén, M. (2008). A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. *Molecular & Cellular Proteomics*, *7*(10), 2019–2027. https://doi.org/10.1074/mcp.R800013-MCP200

Berkeley, U. C. (2004). The Bayh-Dole Act of 1980 and University-Industry Technology Transfer : A Policy Model for Other Governments? *Essays in Honor of Edwin Mansfield*, 1–35. https://doi.org/10.1007/0-387-25022-0_18

Beynon, R. J., Doherty, M. K., Pratt, J. M., & Gaskell, S. J. (2005). Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nature Methods*, *2*(8), 587–589. https://doi.org/10.1038/NMETH774

Bianchini, E. P., Louvain, V. B., Marque, P. E., Juliano, M. A., Juliano, L., & Le Bonniec, B. F. (2002). Mapping of the catalytic groove preferences of factor Xa reveals an inadequate selectivity for its macromolecule substrates. *Journal of Biological Chemistry*, *277*(23), 20527–20534. https://doi.org/10.1074/jbc.M201139200

Billsten, P., Freskgård, P. O., Carlsson, U., Jonsson, B. H., & Elwing, H. (1997). Adsorption to silica nanoparticles of human carbonic anhydrase II and truncated forms induce a molten-globule-like structure. *FEBS Letters*, *402*(1), 67–72. https://doi.org/10.1016/S0014-5793(96)01431-7

Björling, E., & Uhlén, M. (2008). Antibodypedia, a portal for sharing antibody and antigen validation data. *Molecular & Cellular Proteomics : MCP*, *7*(10), 2028–37.

https://doi.org/10.1074/mcp.M800264-MCP200

Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology*, *294*(5), 1351–1362. https://doi.org/10.1006/jmbi.1999.3310

Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., & Heck, A. J. R. (2009). Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols*, *4*(4), 484–494. https://doi.org/10.1038/nprot.2009.21

Boesl, U. (2017). Time-of-flight mass spectrometry: Introduction to the basics. *Mass Spectrometry Reviews*, *36*(1), 86–109. https://doi.org/10.1002/mas.21520

Bondarenko, P. V., Chelius, D., & Shaler, T. A. (2002). Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography - Tandem mass spectrometry. *Analytical Chemistry*, *74*(18), 4741–4749. https://doi.org/10.1021/ac0256991

Bordeaux, J., Welsh, A. W., Agarwal, S., Killiam, E., Baquero, M. T., Hanna, J. A., Anagnostou, V. K., & Rimm, D. L. (2010). Antibody validation. *BioTechniques*, *48*(3), 197–209. https://doi.org/10.2144/000113382

Boulware, K. T., & Daugherty, P. S. (2006). Protease specificity determination by using cellular libraries of peptide substrates (CLiPS). *Proceedings of the National Academy of Sciences of the United States of America*, *103*(20), 7583–8. https://doi.org/10.1073/pnas.0511108103

Bourmaud, A., Gallien, S., & Domon, B. (2016). Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *Proteomics*, *16*(15–16), 2146–2159. https://doi.org/10.1002/pmic.201500543

Braisted, J. C., Kuntumalla, S., Vogel, C., Marcotte, E. M., Rodrigues, A. R., Wang, R., Huang, S.-T., Ferlanti, E. S., Saeed, A. I., Fleischmann, R. D., Peterson, S. N., & Pieper, R. (2008). The APEX Quantitative Proteomics Tool: Generating protein quantitation estimates from LC-MS/MS proteomics results. *BMC Bioinformatics*, *9*(1), 529. https://doi.org/10.1186/1471-2105-9-529

Brash, J. L., & Horbett, T. a. (2012). Protein Surface Interactions and Biocompatibility: A Forty Year Perspective. In *ACS Symposium Series* (Vol. 1120, pp. 277–300). ACS Publications. https://doi.org/10.1080/01932699708943757

Bromage, E., Carpenter, L., Kaattari, S., & Patterson, M. (2009). Quantification of coral heat shock proteins from individual coral polyps. *Marine Ecology Progress Series*, *376*, 123–132. https://doi.org/10.3354/meps07812

Brophy, M. B., & Nolan, E. M. (2015). Manganese and microbial pathogenesis: Sequestration by the mammalian immune system and utilization by microorganisms. *ACS Chemical Biology*, *10*(3), 641–651. https://doi.org/10.1021/cb500792b

Brownridge, P., & Beynon, R. J. (2011). The importance of the digest: Proteolysis and absolute quantification in proteomics. *Methods*, *54*(4), 351–360. https://doi.org/10.1016/j.ymeth.2011.05.005

Brownridge, P., Holman, S. W., Gaskell, S. J., Grant, C. M., Harman, V. M., Hubbard, S. J., Lanthaler, K., Lawless, C., O'cualain, R., Sims, P., Watkins, R., & Beynon, R. J. (2011). Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy. *Proteomics*, *11*(15), 2957–2970. https://doi.org/10.1002/pmic.201100039

Brownridge, P. J., Harman, V. M., Simpson, D. M., & Beynon, R. J. (2012). Absolute multiplexed protein quantification using qconcat technology. In K. Marcus (Ed.), *Methods in Molecular Biology* (Vol. 893, pp. 267–293). Totowa, NJ: Humana Press.

https://doi.org/10.1007/978-1-61779-885-6_18

Brownridge, P., Lawless, C., Payapilly, A. B., Lanthaler, K., Holman, S. W., Harman, V. M., Grant, C. M., Beynon, R. J., & Hubbard, S. J. (2013). Quantitative analysis of chaperone network throughput in budding yeast. *Proteomics*, *13*(8), 1276–1291. https://doi.org/10.1002/pmic.201200412

Bruce, C., Stone, K., Gulcicek, E., & Williams, K. (2013). Proteomics and the analysis of proteomic data: 2013 overview of current protein-profiling technologies. *Current Protocols in Bioinformatics*, (41). https://doi.org/10.1002/0471250953.bi1321s41

Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., & Garin, J. (2007). Isotope-labeled Protein Standards. *Molecular & Cellular Proteomics*, *6*(12), 2139–2149. https://doi.org/10.1074/mcp.M700163-MCP200

Brun, V., Masselon, C., Garin, J., & Dupuis, A. (2009). Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, *72*(5), 740–749. https://doi.org/10.1016/j.jprot.2009.03.007

Bummer, P. M., & Koppenol, S. (2000). Chemical and physical considerations in protein and peptide stability. In *Drugs Pharm. Sci.* (Vol. 99, pp. 5–69). CRC Press. https://doi.org/doi:10.3109/9780849379529-3

Cao, E., Chen, Y., Cui, Z., & Foster, P. R. (2003). Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions. *Biotechnology and Bioengineering*, *82*(6), 684–690. https://doi.org/10.1002/bit.10612

Carlson, E. D., Gan, R., Hodgman, C. E., & Jewett, M. C. (2012). Cell-free protein synthesis: Applications come of age. *Biotechnology Advances*, *30*(5), 1185–1194. https://doi.org/10.1016/j.biotechadv.2011.09.016

Carr, S. A., Abbatiello, S. E., Ackermann, B. L., Borchers, C., Domon, B., Deutsch, E. W., Grant, R. P., Hoofnagle, A. N., Hüttenhain, R., Koomen, J. M., Liebler, D. C., Liu, T., MacLean, B., Mani, D., Mansfield, E., Neubert, H., Paulovich, A. G., … Weintraub, S. (2014). Targeted Peptide Measurements in Biology and Medicine: Best Practices for Mass Spectrometry-based Assay Development Using a Fit-for-Purpose Approach. *Molecular & Cellular Proteomics*, *13*(3), 907–917. https://doi.org/10.1074/mcp.M113.036095

Carroll, K. M., Simpson, D. M., Eyers, C. E., Knight, C. G., Brownridge, P., Dunn, W. B., Winder, C. L., Lanthaler, K., Pir, P., Malys, N., Kell, D. B., Oliver, S. G., Gaskell, S. J., & Beynon, R. J. (2011). Absolute quantification of the glycolytic pathway in yeast: deployment of a complete QconCAT approach. *Molecular & Cellular Proteomics : MCP*, *10*(12), M111.007633. https://doi.org/10.1074/mcp.M111.007633

Caulfield, T., & Ogbogu, U. (2015). The commercialization of university-based research: Balancing risks and benefits. *BMC Medical Ethics*, *16*(1), 70. https://doi.org/10.1186/s12910-015-0064-2

Chang, W. C., Huang, L. C. L., Wang, Y.-S., Peng, W.-P., Chang, H. C., Hsu, N. Y., Yang, W. Bin, & Chen, C. H. (2007). Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. *Analytica Chimica Acta*, *582*(1), 1–9. https://doi.org/10.1016/j.aca.2006.08.062

Chawner, R., McCullough, B., Giles, K., Barran, P. E., Gaskell, S. J., & Eyers, C. E. (2012). QconCAT standard for calibration of ion mobility-mass spectrometry systems. *Journal of Proteome Research*, *11*(11), 5564–5572. https://doi.org/10.1021/pr3005327

Chen, J.-Q., Wakefield, L. M., & Goldstein, D. J. (2015). Capillary nano-immunoassays: advancing quantitative proteomics analysis, biomarker assessment, and molecular diagnostics. *Journal of Translational Medicine*, *13*(1), 182. https://doi.org/10.1186/s12967-015-0537-6

Chen, J. Q., Heldman, M. R., Herrmann, M. A., Kedei, N., Woo, W., Blumberg, P. M., & Goldsmith, P. K. (2013). Absolute quantitation of endogenous proteins with precision and accuracy using a capillary Western system. *Analytical Biochemistry*, *442*(1), 97–103. https://doi.org/10.1016/j.ab.2013.07.022

Cheung, C. S. F., Anderson, K. W., Wang, M., & Turko, I. V. (2015). Natural flanking sequences for peptides included in a quantification concatamer internal standard. *Analytical Chemistry*, *87*(2), 1097–1102. https://doi.org/10.1021/ac503697j

Clackson, T., Hoogenboom, H., Griffiths, A., & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature*, *352*(6336), 624–628. https://doi.org/10.1038/352624a0

Colella, A. D., Chegenii, N., Tea, M. N., Gibbins, I. L., Williams, K. A., & Chataway, T. K. (2012). Comparison of Stain-Free gels with traditional immunoblot loading control methodology. *Analytical Biochemistry*, *430*(2), 108–110. https://doi.org/10.1016/j.ab.2012.08.015

Colyer, J. (2011, August 14). Agents for and method of quantifying binding. Google Patents. Retrieved from https://www.google.com/patents/US7939265

Coon, J. J., Shabanowitz, J., Hunt, D. F., & Syka, J. E. P. (2005). Electron transfer dissociation of peptide anions. *Journal of the American Society for Mass Spectrometry*, *16*(6), 880–882. https://doi.org/10.1016/j.jasms.2005.01.015

Corbin, B. D., Seeley, E. H., Raab, A., Feldmann, J., Miller, M. R., Torres, V. J., Anderson, K. L., Dattilo, B. M., Dunman, P. M., Gerads, R., Caprioli, R. M., Nacken, W., Chazin, W. J., & Skaar, E. P. (2008). Metal Chelation and Inhibition of Bacterial Growth in Tissue Abscesses. *Science*, *319*(5865), 962–965. https://doi.org/10.1126/science.1152449

Cordeiro, A., Pereira, R., Chapeaurouge, A., Coimbra, C., Perales, J., Oliveira, G., Candiani, T., & Coimbra, R. (2015). Comparative proteomics of cerebrospinal fluid reveals a predictive model for differential diagnosis of pneumococcal, meningococcal, and enteroviral meningitis, and novel putative therapeutic targets. *BMC Genomics*, *16*(Suppl 5), S11. https://doi.org/10.1186/1471-2164-16-S5-S11

Cordingleys, G., Colonno, J., & Callahan, P. L. (1990). Substrate Requirements of Human Rhinovirus 3C Protease for PeptideCleavage in Vitro *. *Journal of Biological Chemistry*, *265*(16), 9062–9065.

Deutsch, E. W., Lam, H., & Aebersold, R. (2008). PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Reports*, *9*(5), 429–434. https://doi.org/10.1038/embor.2008.56

Dhiman, A., Kalra, P., Bansal, V., Bruno, J. G., & Sharma, T. K. (2017). Aptamer-based point-of-care diagnostic platforms. *Sensors and Actuators, B: Chemical*. https://doi.org/10.1016/j.snb.2017.02.060

Dias, A. M. G. C., & Roque, A. C. A. (2017). The future of protein scaffolds as affinity reagents for purification. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.26090

Distler, U., Kuharev, J., Navarro, P., & Tenzer, S. (2016). Label-free quantification in ion mobility–enhanced data-independent acquisition proteomics. *Nature Protocols*, *11*(4), 795–812. https://doi.org/10.1038/nprot.2016.042

Doherty, A. J., Connolly, B. A., & Worrall, A. F. (1993). Overproduction of the toxic protein, bovine pancreatic DNasel, in Escherichia coli using a tightly controlled T7-promoter-based vector. *Gene*, *136*(1–2), 337–340. https://doi.org/10.1016/0378-1119(93)90491-K

Dong, H., Nilsson, L., Kurland, C., Dong, H., Nilsson, L., & Kurland, C. (1995). Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome

destruction . These include : Gratuitous Overexpression of Genes in Escherichia coli Leads to Growth Inhibition and Ribosome Destruction. *Journal of Bacteriology*, *177*(6), 1497–1504.

Dong, H., Nilsson, L., & Kurland, C. G. (1996). Co-variation of tRNA Abundance and Codon Usage inEscherichia coliat Different Growth Rates. *Journal of Molecular Biology*, *260*(5), 649–663. https://doi.org/10.1006/jmbi.1996.0428

Dougherty, W. G., Cary, S. M., & Parks, T. D. (1989). Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology*, *171*(2), 356–364. https://doi.org/Doi: 10.1016/0042-6822(89)90603-x

Drabovich, A. P., Mart??nez-Morillo, E., & Diamandis, E. P. (2015). Toward an integrated pipeline for protein biomarker development. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1854*(6), 677–686. https://doi.org/10.1016/j.bbapap.2014.09.006

Du, Y., Parks, B. A., Sohn, S., Kwast, K. E., & Kelleher, N. L. (2006). Top-down approaches for measuring expression ratios of intact yeast proteins using Fourier transform mass spectrometry. *Analytical Chemistry*, *78*(3), 686–694. https://doi.org/10.1021/ac050993p

Dumon-Seignovert, L., Cariot, G., & Vuillard, L. (2004). The toxicity of recombinant proteins in Escherichia coli: a comparison of over expression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification*, *37*(1), 203–206. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15294299

Edmondson, G. (2013). A Grace Period For Patents Could It Help European Universities Innovate ? *The Science|Business*, 40.

Edwards, B. M., & He, M. (2012). Evolution of antibodies in vitro by ribosome display. *Methods in Molecular Biology*, *907*, 281–292. https://doi.org/10.1007/978-1-61779-974-7-16

El-Aneed, A., Cohen, A., & Banoub, J. (2009). Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews*, *44*(3), 210–230. https://doi.org/10.1080/05704920902717872

Elbaggari, A., McDonald, K., Alburo, A., & Martin, C. (2008). Imaging of Chemiluminiscent Western Blots: Comparison of Digital Imaging and X-ray Film. *Imaging and Blotting*, (Figure 1), 1–4.

Eliseev, R., Alexandrov, A., & Gunter, T. (2004). High-yield expression and purification of p18 form of Bax as an MBP-fusion protein. *Protein Expression and Purification*, *35*(2), 206–209. https://doi.org/10.1016/j.pep.2004.01.015

Engholm-Keller, K., & Larsen, M. R. (2013). Technologies and challenges in large-scale phosphoproteomics. *Proteomics*. https://doi.org/10.1002/pmic.201200484

Engvall, E., & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry*, *8*(9), 871–874. https://doi.org/10.1016/0019-2791(71)90454-X

Eyers, C. E., Lawless, C., Wedge, D. C., Lau, K. W., Gaskell, S. J., & Hubbard, S. J. (2011). CONSeQuence: Prediction of Reference Peptides for Absolute Quantitative Proteomics Using Consensus Machine Learning Approaches. *Molecular & Cellular Proteomics*, *10*(11), M110.003384. https://doi.org/10.1074/mcp.M110.003384

Eyers, C. E., Simpson, D. M., Wong, S. C. C., Beynon, R. J., & Gaskell, S. J. (2008). QCAL-a Novel Standard for Assessing Instrument Conditions for Proteome Analysis. *Journal of the American Society for Mass Spectrometry*, *19*(9), 1275–1280. https://doi.org/10.1016/j.jasms.2008.05.019

Fahnert, B., Lilie, H., & Neubauer, P. (2004). Inclusion Bodies: Formation and Utilisation. *Physiological Stress Responses in Bioprocesses*, 93–142.

https://doi.org/10.1007/b93995

Felgner, P. L., & Wilson, J. E. (1976). Hexokinase binding to polypropylene test tubes. Artifactural activity losses from protein binding to disposable plastics. *Analytical Biochemistry*, *74*(2), 631–635. https://doi.org/10.1016/0003-2697(76)90251-7

Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. (2015). The economics of reproducibility in preclinical research. *PLoS Biology*, *13*(6), 1–9. https://doi.org/10.1371/journal.pbio.1002165

Frenzel, A., Hust, M., & Schirrmann, T. (2013). Expression of recombinant antibodies. *Frontiers in Immunology*. https://doi.org/10.3389/fimmu.2013.00217

Gabay, C., & Kushner, I. (1999). Acute-Phase Proteins and Other Systemic Responses to Inflammation. *New England Journal of Medicine*, *340*(6), 448–454. https://doi.org/10.1056/NEJM199902113400607

Gallien, S., Bourmaud, A., Kim, S. Y., & Domon, B. (2014). Technical considerations for large-scale parallel reaction monitoring analysis. *Journal of Proteomics*, *100*, 147–159. https://doi.org/10.1016/j.jprot.2013.10.029

Gallien, S., Duriez, E., Crone, C., Kellmann, M., Moehring, T., & Domon, B. (2012). Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics*, *11*(12), 1709–1723. https://doi.org/10.1074/mcp.O112.019802

Gallien, S., Kim, S. Y., & Domon, B. (2015). Large-Scale Targeted Proteomics Using Internal Standard Triggered-Parallel Reaction Monitoring (IS-PRM). *Molecular & Cellular Proteomics*, *14*(6), 1630–1644. https://doi.org/10.1074/mcp.O114.043968

Gallwitz, M., Enoksson, M., Thorpe, M., & Hellman, L. (2012). The extended cleavage specificity of human thrombin. *PLoS ONE*, *7*(2), e31756. https://doi.org/10.1371/journal.pone.0031756

Gassmann, M., Grenacher, B., Rohde, B., & Vogel, J. (2009). Quantifying Western blots: Pitfalls of densitometry. *Electrophoresis*, *30*(11), 1845–1855. https://doi.org/10.1002/elps.200800720

Gauthier, M. S., Pérusse, J. R., Awan, Z., Bouchard, A., Tessier, S., Champagne, J., Krastins, B., Byram, G., Chabot, K., Garneau, P., Rabasa-Lhoret, R., Faubert, D., Lopez, M. F., Seidah, N. G., & Coulombe, B. (2015). A semi-automated mass spectrometric immunoassay coupled to selected reaction monitoring (MSIA-SRM) reveals novel relationships between circulating PCSK9 and metabolic phenotypes in patient cohorts. *Methods*, *81*, 66–73. https://doi.org/10.1016/j.ymeth.2015.03.003

Gebauer, M., & Skerra, A. (2009). Engineered protein scaffolds as next-generation antibody therapeutics. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/j.cbpa.2009.04.627

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, a, Dephoure, N., O'Shea, E. K., & Weissman, J. S. (2003). Global analysis of protein expression in yeast. *Nature*, *425*(6959), 737–741. https://doi.org/10.1038/nature02046\nnature02046 [pii]

Ghosh, R., Gilda, J. E., & Gomes, A. V. (2014). The necessity of and strategies for improving confidence in the accuracy of western blots. *Expert Review of Proteomics*, *11*(5), 549–560. https://doi.org/10.1586/14789450.2014.939635

Giansanti, P., Tsiatsiani, L., Low, T. Y., & Heck, A. J. R. (2016). Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols*, *11*(5), 993–1006. https://doi.org/10.1038/nprot.2016.057

Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., & Aebersold, R. (2012). Targeted Data Extraction of the MS/MS Spectra Generated by Data-

independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics*, *11*(6), O111.016717. https://doi.org/10.1074/mcp.O111.016717

Gilroy, K. L., Cumming, S. A., & Pitt, A. R. (2010). A simple, sensitive and selective quantum-dot-based western blot method for the simultaneous detection of multiple targets from cell lysates. *Analytical and Bioanalytical Chemistry*, *398*(1), 547–554. https://doi.org/10.1007/s00216-010-3908-0

Glish, G. L., & Vachet, R. W. (2003). The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery*, *2*(2), 140–150. https://doi.org/10.1038/nrd1011

Goebel-Stengel, M., Stengel, A., Taché, Y., & Reeve, J. R. (2011). The importance of using the optimal plasticware and glassware in studies involving peptides. *Analytical Biochemistry*, *414*(1), 38–46. https://doi.org/10.1016/j.ab.2011.02.009

Gómez-Baena, G., Bennett, R. J., Martínez-Rodríguez, C., Wnęk, M., Laing, G., Hickey, G., McLean, L., Beynon, R. J., & Carrol, E. D. (2017). Quantitative Proteomics of Cerebrospinal Fluid in Paediatric Pneumococcal Meningitis. *Scientific Reports*, *7*(1), 7042. https://doi.org/10.1038/s41598-017-07127-6

Goonetilleke, U. R., Scarborough, M., Ward, S. a, & Gordon, S. B. (2010). Proteomic analysis of cerebrospinal fluid in pneumococcal meningitis reveals potential biomarkers associated with survival. *The Journal of Infectious Diseases*, *202*(4), 542–50. https://doi.org/10.1086/654819

Gotrik, M. R., Feagin, T. A., Csordas, A. T., Nakamoto, M. A., & Soh, H. T. (2016). Advancements in Aptamer Discovery Technologies. *Accounts of Chemical Research*, *49*(9), 1903–1910. https://doi.org/10.1021/acs.accounts.6b00283

Grimaldi, R., Kenney, M., Siegel, D. S., & Wright, M. (2011). 30 years after Bayh – Dole : Reassessing academic entrepreneurship. *Research Policy*, *40*(8), 1045–1057. https://doi.org/10.1016/j.respol.2011.04.005

Grimm, S., Yu, F., & Nygren, P. Å. (2011). Ribosome display selection of a murine IgG1 fab binding affibody molecule allowing species selective recovery of monoclonal antibodies. *Molecular Biotechnology*, *48*(3), 263–276. https://doi.org/10.1007/s12033-010-9367-1

Grossmann, J., Roschitzki, B., Panse, C., Fortes, C., Barkow-Oesterreicher, S., Rutishauser, D., & Schlapbach, R. (2010). Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *Journal of Proteomics*, *73*(9), 1740–1746. https://doi.org/10.1016/j.jprot.2010.05.011

Gupta, S., Manubhai, K. P., Kulkarni, V., & Srivastava, S. (2016). An overview of innovations and industrial solutions in Protein Microarray Technology. *Proteomics*. https://doi.org/10.1002/pmic.201500429

Gustafsson, C., Govindarajan, S., & Minshull, J. (2011). Codon bias and heterologous protein expression. *Trends in Biotechnology*, *22*(7), 346–353.

Gygi, S. P. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.*, *17*(10), 994–999. Retrieved from http://dx.doi.org/10.1038/13690

Hamm, M., Ha, S., & Rustandi, R. R. (2015). Automated capillary Western dot blot method for the identity of a 15-valent pneumococcal conjugate vaccine. *Analytical Biochemistry*, *478*, 33–39. https://doi.org/10.1016/j.ab.2015.03.021

Hancock, D. C., & O'Reilly, N. J. (2005). Synthetic peptides as antigens for antibody production. In R. Burns (Ed.), *Methods in molecular biology (Clifton, N.J.)* (Vol. 295, pp. 13–26). Totowa, NJ: Humana Press. https://doi.org/10.1385/1592598730

Hanes, J., & Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences*, *94*(10), 4937–4942. https://doi.org/10.1073/pnas.94.10.4937

Hansen, K. C., Schmitt-ulms, G., Chalkley, R. J., Hirschţ, J., Baldwin, M. A., & Burlingame, A. L. (2003). Mass Spectrometric Analysis of Protein Mixtures at Low Levels Using Cleavable C-Isotope-coded Affinity Tag and Multidimensional Chromatography. *Molecular & Cellular Proteomics*, *2*(5), 299–314. https://doi.org/10.1074/mcp.M

Hayden, M. S., & Ghosh, S. (2012). NF-κB, the first quarter-century: Remarkable progress and outstanding questions. *Genes and Development*, *26*(3), 203–234. https://doi.org/10.1101/gad.183434.111

He, M., Jin, L., & Austen, B. (1993). Specificity of Factor Xa in the cleavage of fusion proteins. *Journal of Protein Chemistry*, *12*(1), 1–5. https://doi.org/10.1007/BF01024906

Hefti, M. H., Van Vugt-Van der Toorn, C. J. G., Dixon, R., & Vervoort, J. (2001). A Novel Purification Method for Histidine-Tagged Proteins Containing a Thrombin Cleavage Site. *Analytical Biochemistry*, *295*(2), 180–185. https://doi.org/doi:10.1006/abio.2001.5214

Helsby, M. A., Fenn, J. R., & Chalmers, A. D. (2013). Reporting research antibody use: how to increase experimental reproducibility. *F1000Research*. https://doi.org/10.12688/f1000research.2-153.v2

Helsby, M. A., Leader, P. M., Fenn, J. R., Gulsen, T., Bryant, C., Doughton, G., Sharpe, B., Whitley, P., Caunt, C. J., James, K., Pope, A. D., Kelly, D. H., & Chalmers, A. D. (2014). CiteAb: a searchable antibody database that ranks antibodies by the number of times they have been cited. *BMC Cell Biology*, *15*(1), 6. https://doi.org/10.1186/1471-2121-15-6

Hinderliter, A., & May, S. (2006). Cooperative Adsorption of Proteins Onto Lipid Membranes. *J. Phys.: Condens. Matter*, *18*(28), S1257--S1270.

Hoesel, B., & Schmid, J. A. (2013). The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer*, *12*(1), 86. https://doi.org/10.1186/1476-4598-12-86

Holman, S. W., Mclean, L., & Eyers, C. E. (2016). RePLiCal : A QconCAT protein for retention time standardisation in proteomics studies RePLiCal : A QconCAT protein for retention time standardisation in proteomics studies. *Journal of Proteome Research*, *15*(3), 1090–1102. https://doi.org/10.1021/acs.jproteome.5b00988

Holman, S. W., Sims, P. F., & Eyers, C. E. (2012). The use of selected reaction monitoring in quantitative proteomics. *Bioanalysis*, *4*(14), 1763–1786. https://doi.org/10.4155/bio.12.126

Hornbeck, P. V, Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Psp, P. R. (2017). PhosphoSitePlus , 2014 : mutations , PTMs and recalibrations. *Nucleic Acids Research*, *43*(December 2014), 512–520. https://doi.org/10.1093/nar/gku1267

Hsu, H.-J., Tsai, K.-C., Sun, Y.-K., Chang, H.-J., Huang, Y.-J., Yu, H.-M., Lin, C.-H., Mao, S.-S., & Yang, A.-S. (2008). Factor Xa active site substrate specificity with substrate phage display and computational molecular modeling. *The Journal of Biological Chemistry*, *283*(18), 12343–53. https://doi.org/10.1074/jbc.M708843200

Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., & Cooks, R. G. (2005). The Orbitrap: A new mass spectrometer. *Journal of Mass Spectrometry*. https://doi.org/10.1002/jms.856

Hughes, A. J., & Herr, A. E. (2012). Microfluidic Western blotting. *Proceedings of the National Academy of Sciences*, *109*(52), 21450–21455. https://doi.org/10.1073/pnas.1207754110

Hung, C. W., & Tholey, A. (2012). Tandem mass tag protein labeling for top-down identification

and quantification. *Analytical Chemistry*, *84*(1), 161–170. https://doi.org/10.1021/ac202243r

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., & Mann, M. (2005). Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein. *Molecular & Cellular Proteomics*, *4*(9), 1265–1272. https://doi.org/10.1074/mcp.M500061-MCP200

Jain, S., Gautam, V., & Naseem, S. (2011). Acute-phase proteins: As diagnostic tool. *Journal of Pharmacy and Bioallied Sciences*, *3*(1), 118. https://doi.org/10.4103/0975-7406.76489

Jenny, R. J., Mann, K. G., & Lundblad, R. L. (2003). A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expression and Purification*, *31*(1), 1–11. https://doi.org/10.1016/S1046-5928(03)00168-2

Jeong, S., Heu, W., Kim, J., & Kim, H.-S. (2016). Protein Binders Specific for Immunoglobulin G from Different Species for Immunoassays and Multiplex Imaging. *Analytical Chemistry*, *88*(23), 11938–11945. https://doi.org/10.1021/acs.analchem.6b03851

Jesse, S., Steinacker, P., Lehnert, S., Sdzuj, M., Cepek, L., & Tumani, H. (2010). A proteomic approach for the diagnosis of bacterial meningitis. *PLoS One*, *5*(4), e10079. https://doi.org/10.1371/journal.pone.0010079

Jia, W., Ren, C., Wang, L., Zhu, B., Jia, W., Gao, M., Zeng, F., Zeng, L., Xia, X., Zhang, X., Fu, T., Li, S., Du, C., Jiang, X., Chen, Y., Tan, W., Zhao, Z., & Liu, W. (2016). CD109 is identified as a potential nasopharyngeal carcinoma biomarker using aptamer selected by cell-SELEX. *Oncotarget*, *7*(34), 55328–55342. https://doi.org/10.18632/oncotarget.10530

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Johnson, A. R., & Carlson, E. E. (2015). Collision-Induced Dissociation Mass Spectrometry: A Powerful Tool for Natural Product Structure Elucidation. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.5b01543

Juncker, D., Bergeron, S., Laforte, V., & Li, H. (2014). Cross-reactivity in antibody microarrays and multiplexed sandwich assays: Shedding light on the dark side of multiplexing. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/j.cbpa.2013.11.012

Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*, *96*(4), 317–323. https://doi.org/10.1016/S1389-1723(03)90130-7

Kanda, K., Nogami, M., & Waki, I. (2011, November 24). Immunoanalytical method and system using mass spectrometry technology. Google Patents. Retrieved from https://www.google.com/patents/US20110287446

Kane, J. F., & Hartley, D. L. (1988). Formation of recombinant protein inclusion bodies in\textit{ Escherichia coli}. *Trends in Biotechnology*, *6*(5), 95–101.

Karas, M., Bachmann, D., Bahr, U., & Hillenkamp, F. (1987). Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes*, *78*(C), 53–68. https://doi.org/10.1016/0168-1176(87)87041-6

Karin, M. (2009). NF- B as a Critical Link Between Inflammation and Cancer. *Cold Spring Harbor Perspectives in Biology*, *1*(5), a000141–a000141. https://doi.org/10.1101/cshperspect.a000141

Karin, M., & Delhase, M. (2000). The I kappa B kinase (IKK) and NF-kappa B: key elements of proinflammatory signalling. In *Seminars in immunology* (Vol. 12, pp. 85–98). Elsevier. https://doi.org/10.1006/smim.2000.0210

Karlsson, M., Ekeroth, J., Elwing, H., & Carlsson, U. (2005). Reduction of irreversible protein adsorption on solid surfaces by protein engineering for increased stability. *Journal of Biological Chemistry*, *280*(27), 25558–25564. https://doi.org/10.1074/jbc.M503665200

Katafuchi, T., Esterházy, D., Lemoff, A., Ding, X., Sondhi, V., Kliewer, S. A., Mirzaei, H., & Mangelsdorf, D. J. (2015). Detection of FGF15 in plasma by stable isotope standards and capture by anti-peptide antibodies and targeted mass spectrometry. *Cell Metabolism*, *21*(6), 898–904. https://doi.org/10.1016/j.cmet.2015.05.004

Kebarle, P., & Verkcerk, U. H. (2009). Electrospray: From Ions in solution to Ions in the gas phase, what we know now. *Mass Spectrometry Reviews*, *28*(6), 898–917. https://doi.org/10.1002/mas.20247

Kettenbach, A. N., Rush, J., & Gerber, S. A. (2011). Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nature Protocols*, *6*(2), 175–186. https://doi.org/10.1038/nprot.2010.196

Kiel, C., Ebhardt, H. A., Burnier, J., Sabido, E., Zimmermann, T., Aebersold, R., & Serrano, L. (2014). Quanti fi cation of ErbB Network Proteins in Three Cell Types Using Complementary Approaches Identi fi es Cell-General and Cell-Type-Speci fi c Signaling Proteins. *Journal of Proteome Research*, *13*(1), 300–313. Retrieved from http://pubs.acs.org/doi/pdfplus/10.1021/pr400878x

Kiernan, U. A., Niederkofler, E., Tubbs, K. A., Nedelkov, D., & Nelson, R. W. (2003, October 21). Analysis of proteins from biological fluids using mass spectrometric immunoassay. Google Patents. Retrieved from https://www.google.com/patents/US20100267069

Kito, K., Ota, K., Fujita, T., & Ito, T. (2007). A synthetic approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. *Journal of Proteome Research*, *6*(2), 792–800. https://doi.org/10.1021/pr060447s

Kockmann, T., Trachsel, C., Panse, C., Wahlander, A., Selevsek, N., Grossmann, J., Wolski, W. E., & Schlapbach, R. (2016). Targeted proteomics coming of age – SRM, PRM and DIA performance evaluated from a core facility perspective. *Proteomics*, *16*(15–16), 2183–2192. https://doi.org/10.1002/pmic.201500502

Köhler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, *256*(5517), 495–497. https://doi.org/10.1038/256495a0

Koller, A., & Wätzig, H. (2005). Precision and variance components in quantitative gel electrophoresis. *Electrophoresis*, *26*(12), 2470–2475. https://doi.org/10.1002/elps.200500024

Kondrat, R. W., McClusky, G. A., & Cooks, R. G. (1978). Multiple Reaction Monitoring in Mass Spectrometry/Mass Spectrometry for Direct Analysis of Complex Mixtures. *Analytical Chemistry*, *50*(14), 2017–2021. https://doi.org/10.1021/ac50036a020

Kościuczuk, E. M., Lisowski, P., Jarczak, J., Strzałkowska, N., Jóźwik, A., Horbańczuk, J., Krzyzewski, J., Zwierzchowski, L., & Bagnicka, E. (2012). Cathelicidins: family of antimicrobial peptides. A review. *Molecular Biology Reports*, *39*(12), 1–14. https://doi.org/10.1007/s11033-012-1997-x

Kramer, G., Woolerton, Y., Van Straalen, J. P., Vissers, J. P. C., Dekker, N., Langridge, J. I., Beynon, R. J., Speijer, D., Sturk, A., & Aerts, J. M. F. G. (2015). Accuracy and reproducibility in quantification of plasma protein concentrations by mass spectrometry without the use of isotopic standards. *PLoS ONE*, *10*(10), e0140097. https://doi.org/10.1371/journal.pone.0140097

Kramer, K. J., Dunn, P. E., Peterson, R. C., Seballos, H. L., Sanburg, L. L., & Law, J. H. (1976). Purification and characterization of the carrier protein for juvenile hormone from the hemolymph of the tobacco hornworm Manduca sexta Johannson (Lepidoptera: Sphingidae). *Journal of Biological Chemistry*, *251*(16), 4979–4985.

Krastins, B., Prakash, A., Sarracino, D. A., Nedelkov, D., Niederkofler, E. E., Kiernan, U. A., Nelson, R., Vogelsang, M. S., Vadali, G., Garces, A., Sutton, J. N., Peterman, S., Byram, G., Darbouret, B., P??russe, J. R., Seidah, N. G., Coulombe, B., … Lopez, M. F. (2013). Rapid development of sensitive, high-throughput, quantitative and highly selective mass spectrometric targeted immunoassays for clinically important proteins in human plasma and serum. *Clinical Biochemistry*, *46*(6), 399–410. https://doi.org/10.1016/j.clinbiochem.2012.12.019

Krause, E., Wenschuh, H., & Jungblut, P. R. (1999). The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Analytical Chemistry*, *71*(19), 4160–4165. https://doi.org/10.1021/ac990298f

Krey, J. F., Wilmarth, P. A., Shin, J. B., Klimek, J., Sherman, N. E., Jeffery, E. D., Choi, D., David, L. L., & Barr-Gillespie, P. G. (2014). Accurate label-free protein quantitation with high- and low-resolution mass spectrometers. *Journal of Proteome Research*, *13*(2), 1034–1044. https://doi.org/10.1021/pr401017h

Kurien, B. T., & Towbin, H. (2015). From little helpers to automation. In *Western Blotting: Methods and Protocols* (pp. 31–40). https://doi.org/10.1007/978-1-4939-2694-7_6

Kusebauch, U., Campbell, D. S., Deutsch, E. W., Chu, C. S., Spicer, D. A., Brusniak, M. Y., Slagel, J., Sun, Z., Stevens, J., Grimes, B., Shteynberg, D., Hoopmann, M. R., Blattmann, P., Ratushny, A. V., Rinner, O., Picotti, P., Carapito, C., … Moritz, R. L. (2016). Human SRMAtlas: A Resource of Targeted Assays to Quantify the Complete Human Proteome. *Cell*, *166*(3), 766–778. https://doi.org/10.1016/j.cell.2016.06.041

L, I. J., Stoop, M. P., Stingl, C., Sillevis Smitt, P. A., Luider, T. M., & Dekker, L. J. (2013). Comparative study of targeted and label-free mass spectrometry methods for protein quantification. *J Proteome Res*, *12*(4), 2005–2011. https://doi.org/10.1021/pr301221f

Lach, S., & Schankerman, M. (2008). Incentives and invention in universities. *RAND Journal of Economics*, *39*(2), 403–433. https://doi.org/10.1111/j.0741-6261.2008.00020.x

Laemmli, U. K. (1970). Cleavage of structural proteins during the sssembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680–685. https://doi.org/10.1038/227680a0

Larsson, K., Wester, K., Nilsson, P., Uhlén, M., Hober, S., & Wernérus, H. (2006). Multiplexed PrEST immunization for high-throughput affinity proteomics. *Journal of Immunological Methods*, *315*(1–2), 110–120. https://doi.org/10.1016/j.jim.2006.07.014

Law, P. The University and Small Business Patent Procedures (Bayh-Dole) Act of 1980, 96 Public Law § (1980).

Lawless, C., Holman, S. W., Brownridge, P., Lanthaler, K., Harman, V. M., Watkins, R., Hammond, D. E., Miller, R. L., Sims, P. F. G., Grant, C. M., Eyers, C. E., Beynon, R. J., & Hubbard, S. J. (2016). Direct and Absolute Quantification of over 1800 Yeast Proteins via Selected Reaction Monitoring. *Molecular & Cellular Proteomics*, *15*(4), 1309–1322. https://doi.org/10.1074/mcp.M115.054288

Lawless, C., & Hubbard, S. J. (2012). Prediction of missed proteolytic cleavages for the selection of surrogate peptides for quantitative proteomics. *Omics : A Journal of Integrative Biology*, *16*(9), 449–56. https://doi.org/10.1089/omi.2011.0156

Lawrence, T. (2009). The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harbor Perspectives in Biology*, *1*(6), a001651. https://doi.org/10.1101/cshperspect.a001651

Lee, J. W., & Hall, M. (2009). Method validation of protein biomarkers in support of drug development or clinical diagnosis / prognosis. *J Chromatogr B Analyt Technol Biomed Life Sci.*, *877*(13), 1259–1271.

Li, H., Han, J., Pan, J., Liu, T., Parker, C. E., & Borchers, C. H. (2017). Current trends in quantitative proteomics–an update. *Journal of Mass Spectrometry*, *52*(5), 319–341.

Li, Q., Huo, Y., Guo, Y., Zheng, X., Sun, W., & Hao, Z. (2017). Generation and applications of a DNA aptamer against Gremlin-1. *Molecules*, *22*(5), 706. https://doi.org/10.3390/molecules22050706

Liu, H., Sadygov, R. G., & Yates, J. R. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry*, *76*(14), 4193–4201. https://doi.org/10.1021/ac0498563

Liu, Y., Buil, A., Collins, B. C., Gillet, L. C. J., Blum, L. C., Cheng, L.-Y., Vitek, O., Mouritsen, J., Lachance, G., Spector, T. D., Dermitzakis, E. T., & Aebersold, R. (2015). Quantitative variability of 342 plasma proteins in a human twin population. *Molecular Systems Biology*, *11*(1), 786. https://doi.org/10.15252/msb.20145728

Liu, Y., Hüttenhain, R., Surinova, S., Gillet, L. C. J., Mouritsen, J., Brunner, R., Navarro, P., & Aebersold, R. (2013). Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics*, *13*(8), 1247–1256. https://doi.org/10.1002/pmic.201200417

Loughney, J. W., Lancaster, C., Ha, S., & Rustandi, R. R. (2014). Residual bovine serum albumin (BSA) quantitation in vaccines using automated Capillary Western technology. *Analytical Biochemistry*, *461*, 49–56. https://doi.org/10.1016/j.ab.2014.05.004

Loureiro, L. R., Carrascal, M. A., Barbas, A., Ramalho, J. S., Novo, C., Delannoy, P., & Videira, P. A. (2015). Challenges in antibody development against Tn and sialyl-Tn antigens. *Biomolecules*. https://doi.org/10.3390/biom5031783

Lu, P., Vogel, C., Wang, R., Yao, X., & Marcotte, E. M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology*, *25*(1), 117–124. https://doi.org/10.1038/nbt1270

Lundberg, M., Eriksson, A., Tran, B., Assarsson, E., & Fredriksson, S. (2011). Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Research*, *39*(15). https://doi.org/10.1093/nar/gkr424

Lundberg, M., Thorsen, S. B., Assarsson, E., Villablanca, A., Tran, B., Gee, N., Knowles, M., Nielsen, B. S., González Couto, E., Martin, R., Nilsson, O., Fermer, C., Schlingemann, J., Christensen, I. J., Nielsen, H.-J., Ekström, B., Andersson, C., … Fredriksson, S. (2011). Multiplexed Homogeneous Proximity Ligation Assays for High-throughput Protein Biomarker Research in Serological Material. *Molecular & Cellular Proteomics*, *10*(4), M110.004978. https://doi.org/10.1074/mcp.M110.004978

Mackenzie, R. J., Lawless, C., Holman, S. W., Lanthaler, K., Beynon, R. J., Grant, C. M., Hubbard, S. J., & Eyers, C. E. (2016). Absolute protein quantification of the yeast chaperome under conditions of heat shock. *Proteomics*, *16*(15–16), 2128–2140. https://doi.org/10.1002/pmic.201500503

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., & MacCoss, M. J. (2010). Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, *26*(7), 966–968. https://doi.org/10.1093/bioinformatics/btq054

Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Analytical Chemistry*, *72*(6), 1156–1162. https://doi.org/10.1021/ac991131p

Malmsten, M. (1998). Formation of Adsorbed Protein Layers. *Journal of Colloid and Interface Science*, *207*(2), 186–199. https://doi.org/10.1006/jcis.1998.5763

Marx, V. (2012). Targeted proteomics. *Nature Methods*, *10*(1), 19–22. https://doi.org/10.1038/nmeth.2285

Mcqueen, P., Spicer, V., Schellenberg, J., Krokhin, O., Sparling, R., Levin, D., & Wilkins, J. A. (2015). Whole cell, label free protein quantitation with data independent acquisition: Quantitation at the MS2 level. *Proteomics*, *15*(1), 16–24. https://doi.org/10.1002/pmic.201400188

Megger, D. A., Bracht, T., Meyer, H. E., & Sitek, B. (2013). Label-free quantification in clinical proteomics. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1834*(8), 1581–1590. https://doi.org/10.1016/j.bbapap.2013.04.001

Meng, F., Cargile, B. J., Patrie, S. M., Johnson, J. R., McLoughlin, S. M., & Kelleher, N. L. (2002). Processing complex mixtures of intact proteins for direct analysis by mass spectrometry. *Anal Chem*, *74*(13), 2923–9. https://doi.org/Doi 10.1021/Ac020049i

Messiha, H. L., Kent, E., Malys, N., Carroll, K. M., Swainston, N., Mendes, P., & Smallbone, K. (2014). Enzyme characterisation and kinetic modelling of the pentose phosphate pathway in yeast. *PeerJ PrePrints*, *2*, e146v4. https://doi.org/10.7287/peerj.preprints.146v4

Method of the Year 2012. (2013). *Nat Meth*, *10*(1), 1.

Meyer, M. S., & Tang, P. (2007). Exploring the “value” of academic patents: IP management practices in UK universities and their implications for Third-Stream indicators. *Scientometrics*, *70*(2), 415–440. https://doi.org/10.1007/s11192-007-0210-9

Michel, M. C., Wieland, T., & Tsujimoto, G. (2009). How reliable are G-protein-coupled receptor antibodies? *Naunyn-Schmiedeberg’s Archives of Pharmacology*, *379*(4), 385–388. https://doi.org/10.1007/s00210-009-0395-y

Miroux, B., & Walker, J. E. (1996). Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *Journal of Molecular Biology*, *260*(3), 289–298. https://doi.org/10.1006/jmbi.1996.0399

Mirzaei, H., McBee, J. K., Watts, J., & Aebersold, R. (2008). Comparative Evaluation of Current Peptide Production Platforms Used in Absolute Quantification in Proteomics. *Molecular & Cellular Proteomics*, *7*(4), 813–823. https://doi.org/10.1074/mcp.M700495-MCP200

Mitchell Wells, J., & McLuckey, S. A. (2005). Collision-induced dissociation (CID) of peptides and proteins. *Methods in Enzymology*, *402*, 148–185. https://doi.org/10.1016/S0076-6879(05)02005-7

Mobley, A., Linder, S. K., Braeuer, R., Ellis, L. M., & Zwelling, L. (2013). A Survey on Data Reproducibility in Cancer Research Provides Insights into Our Limited Ability to Translate Findings from the Laboratory to the Clinic. *PLoS ONE*, *8*(5), e63221. https://doi.org/10.1371/journal.pone.0063221

Mollica, J. P., Oakhill, J. S., Lamb, G. D., & Murphy, R. M. (2009). Are genuine changes in protein expression being overlooked? Reassessing Western blotting. *Analytical Biochemistry*, *386*(2), 270–275. https://doi.org/10.1016/j.ab.2008.12.029

Moulin, A. M., O’shea, S. J., Badley, R. A., Doyle, P., & Welland, M. E. (1999). Measuring surface-induced conformational changes in proteins. *Langmuir*, *15*(26), 8776–8779.

Mouratou, B., B??har, G., & Pecorari, F. (2015). Artificial affinity proteins as ligands of immunoglobulins. *Biomolecules*. https://doi.org/10.3390/biom5010060

Mowery, D. C., & Sampat, B. N. (2006). Universities in National Innovation Systems. In *Business* (pp. 1–38). Georgia Institute of Technology. https://doi.org/10.1093/oxfordhb/9780199286805.003.0008

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 21. https://doi.org/10.1038/s41562-016-0021

Murphy, R. M., Larkins, N. T., Mollica, J. P., Beard, N. A., & Lamb, G. D. (2009). Calsequestrin content and SERCA determine normal and maximal Ca $^{2+}$ storage levels in sarcoplasmic reticulum of fast- and slow-twitch fibres of rat. *The Journal of Physiology*, *587*(2), 443–460. https://doi.org/10.1113/jphysiol.2008.163162

Murphy, R. M., Mollica, J. P., Beard, N. a, Knollmann, B. C., & Lamb, G. D. (2011). Quantification of calsequestrin 2 (CSQ2) in sheep cardiac muscle and Ca2+-binding protein changes in CSQ2 knockout mice. *American Journal of Physiology. Heart and Circulatory Physiology*, *300*(2), H595-604. https://doi.org/10.1152/ajpheart.00902.2010

Musheev, M. U., & Krylov, S. N. (2006). Selection of aptamers by systematic evolution of ligands by exponential enrichment: Addressing the polymerase chain reaction issue. *Analytica Chimica Acta*, *564*(1), 91–96. https://doi.org/10.1016/j.aca.2005.09.069

Nakamura, K., Hirayama-Kurogi, M., Ito, S., Kuno, T., Yoneyama, T., Obuchi, W., Terasaki, T., & Ohtsuki, S. (2016). Large-scale multiplex absolute protein quantification of drug-metabolizing enzymes and transporters in human intestine, liver, and kidney microsomes by SWATH-MS: Comparison with MRM/SRM and HR-MRM/PRM. *Proteomics*, *16*(15–16), 2106–2117. https://doi.org/10.1002/pmic.201500433

Nallamsetty, S., Kapust, R. B., Tözsér, J., Cherry, S., Tropea, J. E., Copeland, T. D., & Waugh, D. S. (2004). Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expression and Purification*, *38*(1), 108–115. https://doi.org/10.1016/j.pep.2004.08.016

Nauseef, W. M. (2014). Microreview Myeloperoxidase in human neutrophil host defence. *Cellular Microbiology*, *16*(June), 1146–1155. https://doi.org/10.1111/cmi.12312

Nelson, A., & Byers, T. (2010). Challenges in University Technology Transfer and the Promising Role of Entrepreneurship Education. *University of Oregon*, 33. https://doi.org/10.2139/ssrn.1651224

Nelson, R. W., & Borges, C. R. (2011, July 5). Mass spectrometric immunoassay revisited. *Journal of the American Society for Mass Spectrometry*. Google Patents. https://doi.org/10.1007/s13361-011-0094-z

Neubert, H., Muirhead, D., Kabir, M., Grace, C., Cleton, A., & Arends, R. (2013). Sequential protein and peptide immunoaffinity capture for mass spectrometry-based quantification of total human ??-nerve growth factor. *Analytical Chemistry*, *85*(3), 1719–1726. https://doi.org/10.1021/ac303031q

Ng, E. W. M., & Adamis, A. P. (2006). Anti-VEGF aptamer (pegaptanib) therapy for ocular vascular diseases. In *Annals of the New York Academy of Sciences* (Vol. 1082, pp. 151–171). https://doi.org/10.1196/annals.1348.062

Nguyen, U., Squaglia, N., Boge, A., & Fung, P. A. (2011). The Simple Western ™ : a gel-free , blot-free , hands-free Western blotting reinvention. *Nature Methods*, *8*(11), v–vi. https://doi.org/10.1038/nmeth.f.353

Nie, S., Shi, T., Fillmore, T. L., Schepmoes, A. A., Brewer, H., Gao, Y., Song, E., Wang, H., Rodland, K. D., Qian, W. J., Smith, R. D., & Liu, T. (2017). Deep-Dive Targeted Quantification for Ultrasensitive Analysis of Proteins in Nondepleted Human Blood Plasma/Serum and Tissues. *Analytical Chemistry*, *89*(17), 9139–9146.

https://doi.org/10.1021/acs.analchem.7b01878

Nohmi, T., & Fenn, J. B. (1992). Electrospray Mass Spectrometry of Poly(ethylene glycols) with Molecular Weights up to Five Million. *Journal of the American Chemical Society*, *114*(9), 3241–3246. https://doi.org/10.1021/ja00035a012

Norde, W., & Haynes, C. A. (1995). Reversibility and the Mechanism of Protein Adsorption (pp. 26–40). ACS Publications. https://doi.org/10.1021/bk-1995-0602.ch002

Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V., & Firsov, A. A. (2004). Electric Field Effect in Atomically Thin Carbon Films. *Science*, *306*(5696), 666–669. https://doi.org/10.1126/science.1102896

Ntai, I., Kim, K., Fellers, R. T., Skinner, O. S., Smith, A. D., Early, B. P., Savaryn, J. P., Leduc, R. D., Thomas, P. M., & Kelleher, N. L. (2014). Applying label-free quantitation to top down proteomics. *Analytical Chemistry*, *86*(10), 4961–4968. https://doi.org/10.1021/ac500395k

Nyasulu, P., Cohen, C., De Gouveia, L., Feldman, C., Klugman, K. P., & von Gottberg, A. (2011). Increased Risk of Death in Human Immunodeficiency Virus-infected Children With Pneumococcal Meningitis in South Africa, 2003–2005. *The Pediatric Infectious Disease Journal*, *30*(12), 1075–1080. https://doi.org/10.1097/INF.0b013e31822cca05

O'Brien, K. ., Wolfson, L. ., Watt, J. ., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Mulholland, K., Levine, O. ., & Cherian, T. (2009). Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *The Lancet*, *374*(9693), 893–902. Retrieved from http://ac.els-cdn.com/S0140673609612046/1-s2.0-S0140673609612046-main.pdf?_tid=21405e68-0488-11e5-9354-00000aacb35f&acdnat=1432742022_89147a09aeed3c6464d0d0d05caf1199

O'Neill, R. a, Bhamidipati, A., Bi, X., Deb-Basu, D., Cahill, L., Ferrante, J., Gentalen, E., Glazer, M., Gossett, J., Hacker, K., Kirby, C., Knittle, J., Loder, R., Mastroieni, C., Maclaren, M., Mills, T., Nguyen, U., … Vander Horn, P. B. (2006). Isoelectric focusing technology quantifies protein signaling in 25 cells. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(44), 16153–8. https://doi.org/10.1073/pnas.0607973103

OECD. (2013). Commercialising Public Research. OECD Publishing. https://doi.org/10.1787/9789264193321-en

Oliveira, E. C. de, Muller, E. I., Abad, F., Dallarosa, J., & Adriano, C. (2010). Internal standard versus external standard calibration: an uncertainty case study of a liquid chromatography analysis. *Química Nova*, *33*(4), 984–987. https://doi.org/10.1590/S0100-40422010000400041

Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., & Mann, M. (2002). Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics*, *1*(5), 376–386. https://doi.org/10.1074/mcp.M200025-MCP200

Ong, S.-E., & Mann, M. (2005). Mass spectrometry–based proteomics turns quantitative. *Nature Chemical Biology*, *1*(5), 252–262. https://doi.org/10.1038/nchembio736

Oss, M., Kruve, A., Herodes, K., & Leito, I. (2010). Electrospray ionization efficiency scale of organic compound. *Analytical Chemistry*, *82*(7), 2865–2872. https://doi.org/10.1021/ac902856t

Palombella, V. J., Rando, O. J., Goldberg, A. L., & Maniatis, T. (1994). The ubiquitin-proteasome pathway is required for processing the NF-κB1 precursor protein and the activation of NF-κB. *Cell*, *78*(5), 773–785. https://doi.org/10.1016/S0092-8674(94)90482-0

Parker, C. E., & Borchers, C. H. (2014). Mass spectrometry based biomarker discovery, verification, and validation - Quality assurance and control of protein biomarker assays. *Molecular Oncology*. https://doi.org/10.1016/j.molonc.2014.03.006

Peng, Z., Oldfield, C. J., Xue, B., Mizianty, M. J., Dunker, A. K., Kurgan, L., & Uversky, V. N. (2014). A creature with a hundred waggly tails: Intrinsically disordered proteins in the ribosome. *Cellular and Molecular Life Sciences*, *71*(8), 1477–1504. https://doi.org/10.1007/s00018-013-1446-6

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012). Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Molecular & Cellular Proteomics*, *11*(11), 1475–1488. https://doi.org/10.1074/mcp.O112.020131

Petrassi, H. M., Williams, J. A., Li, J., Tumanut, C., Ek, J., Nakai, T., Masick, B., Backes, B. J., & Harris, J. L. (2005). A strategy to profile prime and non-prime proteolytic substrate specificity. *Bioorganic and Medicinal Chemistry Letters*, *15*(12), 3162–3166. https://doi.org/10.1016/j.bmcl.2005.04.019

Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., & Aebersold, R. (2009). Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell*, *138*(4), 795–806. https://doi.org/10.1016/j.cell.2009.05.051

Plant, A. L., Locascio, L. E., May, W. E., & Gallagher, P. D. (2014). Improved reproducibility by assuring confidence in measurements in biomedical research. *Nature Methods*, *11*(9), 895–898. https://doi.org/10.1038/nmeth.3076

Pratt, J. M., Beynon, R., & Gaskell, S. (2009, January 21). Artificial protein for absolute quantification of proteins and uses thereof. Google Patents. Retrieved from https://www.google.com/patents/EP1904517B1?cl=3Den

Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., & Beynon, R. J. (2006). Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature Protocols*, *1*(2), 1029–1043. https://doi.org/10.1038/nprot.2006.129

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, *10*(9), 712–712. https://doi.org/10.1038/nrd3439-c1

Qi, Y., & Volmer, D. A. (2017). Electron-based fragmentation methods in mass spectrometry: An overview. *Mass Spectrometry Reviews*. https://doi.org/10.1002/mas.21482

Rabe, M., Verdes, D., & Seeger, S. (2011). Understanding protein adsorption phenomena at solid surfaces. *Advances in Colloid and Interface Science*, *162*(1), 87–106. https://doi.org/10.1016/j.cis.2010.12.007

Raijmakers, R., Berkers, C. R., de Jong, A., Ovaa, H., Heck, A. J. R., & Mohammed, S. (2008). Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. *Molecular & Cellular Proteomics : MCP*, *7*(9), 1755–1762. https://doi.org/10.1074/mcp.M800093-MCP200

Rami, M. H. D., Shweiki, A., Mo, S., Majovsky, P., Thieme, D., Trutschel, D., & Hoehenwarter, W. (2017). Assessment of Label-Free Quanti fi cation in Discovery Proteomics and Impact of Technological Factors and Natural Variability of Protein Abundance. *Journal of Proteome Research*, *16*(4), 1410–1424. https://doi.org/10.1021/acs.jproteome.6b00645

Rappsilber Ryder, U., Lamond, A. I., Mann, M., J. (2002). Large-scale proteomic analysis of the human splicesome. *Genome Research*, *12*(8), 1231–1245. https://doi.org/Doi 10.1101/Gr.473902

Ratzkin, B. (1977). Functional Expression of Cloned Yeast DNA in Escherichia coli. *Proceedings of the National Academy of Sciences*, *74*(2), 487–491. https://doi.org/10.1073/pnas.74.2.487

Rifai, N., Watson, I. D., & Miller, G. (2013). Commercial immunoassays in biomarkers studies: Researchers beware! *Clinical Chemistry and Laboratory Medicine*. Clinical Chemistry. https://doi.org/10.1515/cclm-2013-0015

Rigaut, G., Shevchenko, A., Rutz, B., Matthias, W., Mann, M., & Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnology*, *17*(10), 1030–1032. https://doi.org/10.1038/13732

Ronsein, G. E., Pamir, N., von Haller, P. D., Kim, D. S., Oda, M. N., Jarvik, G. P., Vaisar, T., & Heinecke, J. W. (2015). Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *Journal of Proteomics*, *113*, 388–399. https://doi.org/10.1016/j.jprot.2014.10.017

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00172

Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., & Pappin, D. J. (2004). Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics*, *3*(12), 1154–1169. https://doi.org/10.1074/mcp.M400129-MCP200

Rozenblum, G. T., Lopez, V. G., Vitullo, A. D., & Radrizzani, M. (2016). Aptamers: current challenges and future prospects. *Expert Opinion on Drug Discovery*, *11*(2), 127–135. https://doi.org/10.1517/17460441.2016.1126244

Russell, M. R., Achour, B., McKenzie, E. A., Lopez, R., Harwood, M. D., Rostami-Hodjegan, A., & Barber, J. (2013). Alternative fusion protein strategies to express recalcitrant QconCAT proteins for quantitative proteomics of human drug metabolizing enzymes and transporters. *Journal of Proteome Research*, *12*(12), 5934–5942. https://doi.org/10.1021/pr400279u

Russell, M. R., Walker, M. J., Williamson, A. J. K., Gentry-Maharaj, A., Ryan, A., Kalsi, J., Skates, S., D'Amato, A., Dive, C., Pernemalm, M., Humphryes, P. C., Fourkala, E. O., Whetton, A. D., Menon, U., Jacobs, I., & Graham, R. L. J. (2016). Protein Z: A putative novel biomarker for early detection of ovarian cancer. *International Journal of Cancer*, *138*(12), 2984–2992. https://doi.org/10.1002/ijc.30020

Rustandi, R. R., Hamm, M., Loughney, J. W., & Ha, S. (2015). Detection of ADP ribosylation in PARP-1 and bacterial toxins using a capillary-based western system. *Electrophoresis*, *36*(21–22), 2798–2804. https://doi.org/10.1002/elps.201500173

Rustandi, R. R., Loughney, J. W., Hamm, M., Hamm, C., Lancaster, C., Mach, A., & Ha, S. (2012). Qualitative and quantitative evaluation of Simon, a new CE-based automated Western blot system as applied to vaccine development. *Electrophoresis*, *33*(17), 2790–2797. https://doi.org/10.1002/elps.201200095

Salter, A., & Martin, B. (1999). The Economic Benefits of Publicly Funded Basic Research: A Critical Review. *Research Policy*, *30*(3), 509–532.

Sargent, M. (2013). Guide to achieving reliable quantitative LC-MS measurements. *Analytical Methods Committee*, 1–61. https://doi.org/10.1017/CBO9781107415324.004

Sarkar, P. K., Prajapati, P. K., Shukla, V. J., Ravishankar, B., & Choudhary, A. K. (2009). Toxicity and recovery studies of two ayurvedic preparations of iron. *Indian Journal of*

Experimental Biology, *47*(12), 987–992. https://doi.org/10.1002/mas

Scarborough, M., & Thwaites, G. E. (2008). The diagnosis and management of acute bacterial meningitis in resource-poor settings. *The Lancet Neurology*, *7*(7), 637–648. https://doi.org/10.1016/S1474-4422(08)70139-X

Schacht, W. H. (2012). The Bayh-Dole Act: Selected issues in patent policy and the commercialization of technology. In *CRS Report* (pp. 1–25). DTIC Document. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84932093310&partnerID=40&md5=7c166324b0ef9f67381a62edfe9dc3bb

Scherer, D. C., Brockman, J. A., Chen, Z., Maniatis, T., & Ballard, D. W. (1995). Signal-induced degradation of I kappa B alpha requires site-specific ubiquitination. *Proceedings of the National Academy of Sciences*, *92*(24), 11259–11263. https://doi.org/10.1073/pnas.92.24.11259

Schiffmann, C., Hansen, R., Baumann, S., Kublik, A., Nielsen, P. H., Adrian, L., Von Bergen, M., Jehmlich, N., & Seifert, J. (2014). Comparison of targeted peptide quantification assays for reductive dehalogenases by selective reaction monitoring (SRM) and precursor reaction monitoring (PRM). *Analytical and Bioanalytical Chemistry*, *406*(1), 283–291. https://doi.org/10.1007/s00216-013-7451-7

Schilling, B., MacLean, B., Held, J. M., Sahu, A. K., Rardin, M. J., Sorensen, D. J., Peters, T., Wolfe, A. J., Hunter, C. L., MacCoss, M. J., & Gibson, B. W. (2015). Multiplexed, Scheduled, High-Resolution Parallel Reaction Monitoring on a Full Scan QqTOF Instrument with Integrated Data-Dependent and Targeted Mass Spectrometric Workflows. *Analytical Chemistry*, *87*(20), 10222–10229. https://doi.org/10.1021/acs.analchem.5b02983

Schirrmann, T., Meyer, T., Schütte, M., Frenzel, A., & Hust, M. (2011). Phage display for the generation of antibodies for proteome research, diagnostics and therapy. *Molecules*, *16*(1), 412–426. https://doi.org/10.3390/molecules16010412

Scholl, B., Liu, H. Y., Long, B. R., McCarty, O. J. T., O'Hare, T., Druker, B. J., & Vu, T. Q. (2009). Single particle quantum dot imaging achieves ultrasensitive detection capabilities for Western immunoblot analysis. *ACS Nano*, *3*(6), 1318–1328. https://doi.org/10.1021/nn9000353

Schroeder, H. W. J., & Cavacini, L. (2010). Structure and Function of Immunoglobulins (author manuscript). *Journal of Allergy and Clinical Immunology*, *125*(2), S41–S52. https://doi.org/10.1016/j.jaci.2009.09.046.Structure

Schubert, O. T., Gillet, L. C., Collins, B. C., Navarro, P., Rosenberger, G., Wolski, W. E., Lam, H., Amodei, D., Mallick, P., MacLean, B., & Aebersold, R. (2015). Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nature Protocols*, *10*(3), 426–441. https://doi.org/10.1038/nprot.2015.015

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342. https://doi.org/10.1038/nature10098

Scott, K. B., Turko, I. V., & Phinney, K. W. (2016). QconCAT: Internal Standard for Protein Quantification. *Methods in Enzymology*, *566*, 289–303. https://doi.org/10.1016/bs.mie.2015.09.022

Seed, B. (2001). Silanizing glassware. *Current Protocols in Cell Biology*, A-3E.

Senger, B., Simos, G., Bischoff, F. R., Podtelejnikov, A., Mann, M., & Hurt, E. (1998). Mtr10p functions as a nuclear import receptor for the mRNA-binding protein Np13p. *EMBO Journal*, *17*(8), 2196–2207. https://doi.org/10.1093/emboj/17.8.2196

Shi, T., Fillmore, T. L., Sun, X., Zhao, R., Schepmoes, A. A., Hossain, M., Xie, F., Wu, S., Kim,

J.-S., Jones, N., Moore, R. J., Pasa-Tolic, L., Kagan, J., Rodland, K. D., Liu, T., Tang, K., Camp, D. G., … Qian, W.-J. (2012). Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. *Proceedings of the National Academy of Sciences*, *109*(38), 15395–15400. https://doi.org/10.1073/pnas.1204366109

Shi, T., Song, E., Nie, S., Rodland, K. D., Liu, T., Qian, W. J., & Smith, R. D. (2016). Advances in targeted proteomics and applications to biomedical research. *Proteomics*, *16*(15–16), 2160–2182. https://doi.org/10.1002/pmic.201500449

Shin, S., Kim, I.-H., Kang, W., Yang, J. K., & Hah, S. S. (2010). An alternative to Western blot analysis using RNA aptamer-functionalized quantum dots. *Bioorganic & Medicinal Chemistry Letters*, *20*(11), 3322–3325. https://doi.org/10.1016/j.bmcl.2010.04.040

Siegel, D. S., Veugelers, R., & Wright, M. (2007). University commercialization of intellectual property: Policy implications. *Oxford Review of Economic Policy*, *23*(4), 640–660.

Siepen, J. A., Keevil, E. J., Knight, D., & Hubbard, S. J. (2007). Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *Journal of Proteome Research*, *6*(1), 399–408. https://doi.org/10.1021/pr060507u

Sigalov, A. B. (2010). Unusual biophysics of immune signaling-related intrinsically disordered proteins. *Self Nonself*, *1*(4), 271–281. https://doi.org/10.4161/self.1.4.13641

Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. C., & Geromanos, S. J. (2006). Absolute Quantification of Proteins by LCMS ^{E}. *Molecular & Cellular Proteomics*, *5*(1), 144–156. https://doi.org/10.1074/mcp.M500230-MCP200

Silva, J. M., & McMahon, M. (2014). The Fastest Western in Town: A Contemporary Twist on the Classic Western Blot Analysis. *Journal of Visualized Experiments*, (84). https://doi.org/10.3791/51149

Simpson, D. M., & Beynon, R. J. (2012). QconCATs: Design and expression of concatenated protein standards for multiplexed protein quantification. *Analytical and Bioanalytical Chemistry*, *404*(4), 977–989. https://doi.org/10.1007/s00216-012-6230-1

Simpson, R. J. (2010). Stabilization of Protein for Storage. *Cold Spring Harb Protocols*, *2010*(5), 20439424. https://doi.org/2010/5/pdb.top79 [pii]10.1101/pdb.top79

Singh, S., Springer, M., Steen, J., Kirschner, M. W., & Steen, H. (2009). FLEXIQuant: A novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. *Journal of Proteome Research*, *8*(5), 2201–2210. https://doi.org/10.1021/pr800654s

Sinkala, E., Sollier-Christen, E., Renier, C., Rosàs-Canyelles, E., Che, J., Heirich, K., Duncombe, T. A., Vlassakis, J., Yamauchi, K. A., Huang, H., Jeffrey, S. S., & Herr, A. E. (2017). Profiling protein expression in circulating tumour cells using microfluidic western blotting. *Nature Communications*, *8*, 14622. https://doi.org/10.1038/ncomms14622

Skates, S. J., Gillette, M. A., LaBaer, J., Carr, S. A., Anderson, L., Liebler, D. C., Ransohoff, D., Rifai, N., Kondratovich, M., Težak, Ž., Mansfield, E., Oberg, A. L., Wright, I., Barnes, G., Gail, M., Mesri, M., Kinsinger, C. R., … Boja, E. S. (2013). Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies. *Journal of Proteome Research*, *12*(12), 5383–5394. https://doi.org/10.1021/pr400132j

Smith, D. G. S., Gingras, G., Aubin, Y., & Cyr, T. D. (2016). Design and expression of a QconCAT protein to validate Hi3 protein quantification of influenza vaccine antigens. *Journal of Proteomics*, *146*, 133–140. https://doi.org/10.1016/j.jprot.2016.06.024

Smith, G. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, *228*(4705), 1315–1317.

https://doi.org/10.1126/science.4001944

Smith, J. A., Hurrell, J. G. R., & Leach, S. J. (1978). Elimination of nonspecific adsorption of serum proteins by Sepharose-bound antigens. *Analytical Biochemistry*, *87*(2), 299–305. https://doi.org/10.1016/0003-2697(78)90679-6

Stephens, A. W., Siddiqui, A., & Hirs, C. H. W. (1988). Site-directed mutagenesis of the reactive center (serine 394) of antithrombin III. *Journal of Biological Chemistry*, *263*(31), 15849–15852.

Stoltenburg, R., Reinemann, C., & Strehlitz, B. (2007). SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomolecular Engineering*. https://doi.org/10.1016/j.bioeng.2007.06.001

Straw, S., Ferrigno, P. K., Song, Q., Tomlinson, D., & Galdo, F. Del. (2013). Proof of concept study to identify candidate biomarkers of fibrosis using high throughput peptide aptamer microarray and validate by enzyme linked immunosorbant assay. *Journal of Biomedical Science and Engineering*, *6*(8), 32–42. https://doi.org/10.4236/jbise.2013.68A2005

Suelter, C. H., & DeLuca, M. (1983). How to prevent losses of protein by adsorption to glass and plastic. *Analytical Biochemistry*, *135*(1), 112–119. https://doi.org/10.1016/0003-2697(83)90738-8

Sun, P., Austin, B. P., Tözsér, J., & Waugh, D. S. (2010). Structural determinants of tobacco vein mottling virus protease substrate specificity. *Protein Science*, *19*(11), 2240–2251. https://doi.org/10.1002/pro.506

Swaney, D. L., Wenger, C. D., & Coon, J. J. (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of Proteome Research*, *9*(3), 1323–1329. https://doi.org/10.1021/pr900863u

Szájli, E., Fehér, T., & Medzihradszky, K. F. (2008). Investigating the Quantitative Nature of MALDI-TOF MS. *Molecular & Cellular Proteomics*, *7*(12), 2410–2418. https://doi.org/10.1074/mcp.M800108-MCP200

Takemori, N., Takemori, A., Tanaka, Y., Ishizaki, J., Hasegawa, H., Shiraishi, A., & Ohashi, Y. (2016). High-throughput production of a stable isotope-labeled peptide library for targeted proteomics using a wheat germ cell-free synthesis system. *Mol. BioSyst.*, *12*(8), 2389–2393. https://doi.org/10.1039/C6MB00209A

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., & Matsuo, T. (1988). Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, *2*(8), 151–153. https://doi.org/10.1002/rcm.1290020802

Taylor, S. C., Berkelman, T., Yadav, G., & Hammond, M. (2013). A defined methodology for reliable quantification of western blot data. *Molecular Biotechnology*, *55*(3), 217–226. https://doi.org/10.1007/s12033-013-9672-6

Taylor SC, P. A. (2014). The Design of a Quantitative Western Blot Experiment. *BioMed Research International*, *2014*. https://doi.org/10.1155/2014/361590

Thomas Cerny, S. G. W. K. I. C. R. A. R. S. (2009, October 30). Method for biomarker and drug-target discovery for prostate cancer diagnosis and treatment as well as biomarker assays determined therewith. Google Patents. https://doi.org/US 9377463 B2

Thompson, A., Schäfer, J. J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G. G., Neumann, T., & Hamon, C. (2003). Tandem mass tags: a novel quantificaiton strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry*, *75*(8), 1895–1904. https://doi.org/10.1021/ac0262560

Tighe, P. J., Ryder, R. R., Todd, I., & Fairclough, L. C. (2015). ELISA in the multiplex era:

Potentials and pitfalls. *Proteomics - Clinical Applications*, *9*(3–4), 406–422. https://doi.org/10.1002/prca.201400130

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, *302*(5909), 575–581. https://doi.org/10.1038/302575a0

Törnqvist, E., Annas, A., Granath, B., Jalkesten, E., Cotgreave, I., & Öberg, M. (2014). Strategic focus on 3R principles reveals major reductions in the use of animals in pharmaceutical toxicity testing. *PLoS ONE*, *9*(7), e101638. https://doi.org/10.1371/journal.pone.0101638

Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(9), 4350–4. https://doi.org/10.1002/bies.950190612

Treffers, C., Chen, L., Anderson, R. C., & Yu, P. L. (2005). Isolation and characterisation of antimicrobial peptides from deer neutrophils. *International Journal of Antimicrobial Agents*, *26*(2), 165–169. https://doi.org/10.1016/j.ijantimicag.2005.05.001

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment:RNA ligands to bacteriophage T4 DNA polymerase. *Science*, *249*(4968), 505–510. https://doi.org/10.1126/science.2200121

Uhlen, M., Bandrowski, A., Carr, S., Edwards, A., Ellenberg, J., Lundberg, E., Rimm, D. L., Rodriguez, H., Hiltke, T., Snyder, M., & Yamamoto, T. (2016). A proposal for validation of antibodies. *Nature Methods*. https://doi.org/10.1038/nmeth.3995

Uhlén, M., & Hober, S. (2009). Generation and validation of affinity reagents on a proteome-wide level. *Journal of Molecular Recognition*, *22*(2), 57–64. https://doi.org/10.1002/jmr.891

Valentin, M. A., Ma, S., Zhao, A., Legay, F., & Avrameas, A. (2011). Validation of immunoassay for protein biomarkers: Bioanalytical study plan implementation to support pre-clinical and clinical studies. *Journal of Pharmaceutical and Biomedical Analysis*, *55*(5), 869–877. https://doi.org/10.1016/j.jpba.2011.03.033

Validation, B. M. (1998). Guidance for Industry Bioanalytical Methods Validation for Human Studies Guidance for Industry Bioanalytical Methods Validation.

Van Weemen, B. K., & Schuurs, A. H. W. M. (1971). Immunoassay using antigen-enzyme conjugates. *FEBS Letters*, *15*(3), 232–236. https://doi.org/10.1016/0014-5793(71)80319-8

Vapnek, D., Hautala, J., Jacobson, J., Giles, N., & Kushner, S. (1977). Expression in Escherichia coli K-12 of the structural gene for catabolic dehydroquinase of Neurospora crassa. *Pnas 74:3508-3512.*, *74*(8), 3508–3512.

Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., & Yates, J. R. (2004). Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods*, *1*(1), 39–45. https://doi.org/nmeth705 [pii]\r10.1038/nmeth705

Vidova, V., & Spacil, Z. (2017). A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. *Analytica Chimica Acta*, *964*, 7–23. https://doi.org/10.1016/j.aca.2017.01.059

Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., & de Gier, J.-W. (2008). Tuning Escherichia coli for membrane protein overexpression. *Proceedings of the National Academy of Sciences*, *105*(38), 14371–14376. https://doi.org/10.1073/pnas.0804090105

Wang, S., Cyronak, M., & Yang, E. (2007). Does a stable isotopically labeled internal standard

always correct analyte response?. A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *Journal of Pharmaceutical and Biomedical Analysis*, *43*(2), 701–707. https://doi.org/10.1016/j.jpba.2006.08.010

Wang, W., Becker, C. H., Zhou, H., Lin, H., Roy, S., Shaler, T. A., Hill, L. R., Norton, S., Kumar, P., & Anderle, M. (2003). Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Analytical Chemistry*, *75*(18), 4818–4826. https://doi.org/10.1021/ac026468x

Wang, Y. V, Wade, M., Wong, E., Li, Y.-C., Rodewald, L. W., & Wahl, G. M. (2007). Quantitative analyses reveal the importance of regulated Hdmx degradation for p53 activation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12365–12370. https://doi.org/10.1073/pnas.0701497104

Wendorf, J. R., Radke, C. J., & Blanch, H. W. (2004). Reduced protein adsorption at solid interfaces by sugar excipients. *Biotechnology and Bioengineering*, *87*(5), 565–573. https://doi.org/10.1002/bit.20132

Whiteaker, J. R., Zhao, L., Frisch, C., Ylera, F., Harth, S., Knappik, A., & Paulovich, A. G. (2014). High-affinity recombinant antibody fragments (Fabs) can be applied in peptide enrichment immuno-MRM assays. *Journal of Proteome Research*, *13*(4), 2187–2196. https://doi.org/10.1021/pr4009404

Whitehouse, C. M., Dreyer, R. N., Yamashita, M., & Fenn, J. B. (1985). Electrospray Interface for Liquid Chromatographs and Mass Spectrometers. *Analytical Chemistry*, *57*(3), 675–679. https://doi.org/10.1021/ac00280a023

Wild, D., John, R., & Sheehan, C. (2013). *The Immunoassay Handbook Theory and applications of ligand binding*. *The Immunoassay Handbook*. Newnes. https://doi.org/10.1016/B978-0-08-097037-0.01001-0

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., & Williams, K. L. (1996). Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnology and Genetic Engineering Reviews*, *13*(1), 19–50. https://doi.org/10.1080/02648725.1996.10647923

Willem, N. (1986). Adsorption of proteins from solution at the solid-liquid interface. *Advances in Colloid and Interface Science*, *25*, 267–340. https://doi.org/10.1016/0001-8686(86)80012-4

Wold, E. D., Smider, V. V, Felding, B. H., Jolla, L., Jolla, L., & Jolla, L. (2016). Antibody Therapeutics in Oncology. *Immunotherapy (Los Angeles, Calif.)*, *2*(1), 1–18. https://doi.org/108 [pii]

Wright, M., Clarysse, B., & Mosey, S. (2012). Strategic entrepreneurship, resource orchestration and growing spin-offs from universities. *Technology Analysis and Strategic Management*, *24*(9), 911–927. https://doi.org/10.1080/09537325.2012.718665

Xu, D., Marchionni, K., Hu, Y., Zhang, W., & Sosic, Z. (2017). Journal of Pharmaceutical and Biomedical Analysis Quantitative analysis of a biopharmaceutical protein in cell culture samples using automated capillary electrophoresis (CE) western blot. *Journal of Pharmaceutical and Biomedical Analysis*, *145*, 10–15. https://doi.org/10.1016/j.jpba.2017.06.025

Yamashita, M., & Fenn, B. J. (1984). Negative ion production with the electrospray ion source. *J Phys.Chem.*, *88*(20), 4671–4675. https://doi.org/10.1021/j150664a046

Yamashita, M., & Fenn, J. B. (1984). Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry*, *88*(20), 4451–4459. https://doi.org/10.1021/j150664a002

Yeong, H. A., Ji, Y. L., Ju, Y. L., Kim, Y. S., Jeong, H. K., & Jong, S. Y. (2009). Quantitative

analysis of an aberrant glycoform of TIMP1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. *Journal of Proteome Research*, *8*(9), 4216–4224. https://doi.org/10.1021/pr900269s

Yi, E. C., Li, X. J., Cooke, K., Lee, H., Raught, B., Page, A., Aneliunas, V., Hieter, P., Goodlett, D. R., & Aebersold, R. (2005). Increased quantitative proteome coverage with13C/ 12C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics*, *5*(2), 380–387. https://doi.org/10.1002/pmic.200400970

Yu, Y. Q., Gilar, M., Lee, P. J., Bouvier, E. S. P., & Gebler, J. C. (2003). Enzyme-Friendly, Mass Spectrometry-Compatible Surfactant for In-Solution Enzymatic Digestion of Proteins. *Analytical Chemistry*, *75*(21), 6023–6028. https://doi.org/10.1021/ac0346196

Zellner, M., Babeluk, R., Diestinger, M., Pirchegger, P., Skeledzic, S., & Oehler, R. (2008). Fluorescence-based Western blotting for quantitation of protein biomarkers in clinical samples. *Electrophoresis*, *29*(17), 3621–3627. https://doi.org/10.1002/elps.200700935

Zhang, Z., Wu, S., Stenoien, D. L., & Paša-Tolić, L. (2014). High-Throughput Proteomics. *Annual Review of Analytical Chemistry*, *7*(1), 427–454. https://doi.org/10.1146/annurev-anchem-071213-020216

Zhao, Y., Widen, S. G., Jamaluddin, M., Tian, B., Wood, T. G., Edeh, C. B., & Brasier, A. R. (2011). Quantification of Activated NF-κB/RelA Complexes Using ssDNA Aptamer Affinity – Stable Isotope Dilution—Selected Reaction Monitoring—Mass Spectrometry. *Molecular & Cellular Proteomics*, *10*(6), M111.008771. https://doi.org/10.1074/mcp.M111.008771

Zuniga, P. (2011). The State of Patenting at Research Institutions in developing countries: Policy approaches and practices. *… Report to the Publication: Intellectual Property and the …*, (4), 96. Retrieved from http://www.wipo.int/econ_stat/en/economics/pdf/WP4_Zuniga_to_be_published.pdf

APPENDIX

Nucleotide sequences of DOSCATs

Nucleotide sequences are stored electronically and can be obtained by contacting the Centre for Proteome Research at the University of Liverpool.

Protein sequences of DOSCATs

NFκB-DOSCAT

```
> NFκB-DOSCAT
MGTREGVNDNEEGFFSARDHDSYGVDKKRKRGMPDVLGELNSSDPHGIESKRRKKKPAILDHRLVPRG
SEGRSAGSILGESSTEASKTLPMLRSGPASGPSVPTGRAMPQVAIVFRTPPYADPSLQAPVRVSMQLR
RPSDRELSEPMEFQYLPDTDDRHRIEEKLEVLFQGPEGRSAGSIPGERSTDSQRIQTNNNPFQVPIEE
QRGDYDLNAVRLCFDCRDGFYEAELCPDRCIHMFQAAERPQEWAMEGPRDGLKKERLLDDRHDSGLDS
MKDEEYEQMVKELQEIRIEGRTLLRDAGADLDKPEPTCGRSPLHLAVEAQAADVLELLLRAGAEEKAL
TMEVIRQVKVPRGSEPWKQQLEGRSAGSIPGEHSTDNNRTYPLLRVGADPALLDRHGDGHTALHLACR
VGAHACARALLQPRPRRPREAPDTYLAQGPDRTPDTNHTPVALYPDSDLEKENLYFQGQDRHGDTALH
VACQRQHLSGKTALHLAVETQERGLVLQRLTDGVCSEPLPFTYLPRDHDTPRSGNTNPLSSFSTRTLP
LYKLLEIPDPDNWAPVKKTTSQAHSLPLSPASTRELVNMRNDLYQTPLHLAVITKQEDNARLFGLAQR
SARSLPESTSAPASGPSDGSPQPCTHPPGPVKEPQEKEDADGERADSTYGSSLTYTLSLLGGPEAEDP
APRAGAGAAGHHHHHH
```

NFκB-DOSCAT-2

```
> NFκB-DOSCAT-2
MGTREGVNDNEEGFFSARDHDSYGVDKKRKRGMPDVLGELNSSDPHGIESKRRKKKPAILDHRTVRFQ
SEGRSAGSILGESSTEASKTLPMLRSGPASGPSVPTGRAMVFRLNQGIPVAPHTTEPMLMEYPEAITR
LVTGAQRPPDPAPAPLRLEVLFQGPEGRSAGSIPGERSTDSQRIQTNNNPFQVPIEEQRGDYDLNAVR
LCFDCRDGFYEAELCPDRCIHMFQAAERPQEWAMEGPRDGLKKERLLDDRHDSGLDSMKDEEYEQMVK
ELQEIRDDDDKDAGADLDKPEPTCGRSPLHLAVEAQAADVLELLLRAGAEEKALTMEVIRQVKVPRGS
EPWKQQLEGRSAGSIPGEHSTDNNRTYPLLRVGADPALLDREKSGPCSSSSDSDSGDEGDEYDDIVVH
SSRSQTRLPPTPASKPLPDDPRPVKENLYFQGQDRHGDTALHVACQRQHLSGKTALHLAVETQERGLV
LQRLTDGVCSEPLPFTYLPRDHDTPRSGNTNPLSSFSTRTLPLYKLLEIPDPDKNWAPVKTTSQAHSL
PLSPASTRELVNMRNDLYQTPLHLAVITKQEDNARLFGLAQRSARASGPSDGSPQPCTHPPGPVKEPQ
EKEDADGERADSTYGSSSLTYTLSLLGGPEAEDPAPRLPLPHSRQYDSGIESLRSLRSLRAGAGAAGH
HHHHH
```

M-DOSCAT-S

```
>M-DOSCAT-S
MGTREGVNDNEEGFFSAREQKLISEEDLDSVDPRIANVFTNAFRYGHTLIKDIFTGLIGPMKICKKGS
LHANGRQKDVDKEQISLPRIICDNTGITTVSKNNIFMSRVPLSRVFFASWRVVLEGGVAAHKKSHEES
HKEDKDNKRFALLGDFFRKSKEKIKEAVLRAIDGINQRSSDANLENLYFQGMKTQRDGHSLGRWSLVL
LIDIFTKENLTAPGSDSAVFFEQGTTRIGGSYGVTCPEQDKYRTITGSPGKPPRLVGGPMDASVEEEG
VRRALDFAVGEYNKASNDMYHSRALQVVRARKQIVAGVNYFLDVELGRTTCTKTQPNLDNCPFHDQPH
LKRKAFCSFQIYAVPWQGTMTLSKSTCQDALEVLFQGPGRGRLYKKALYLQYTDETFRTTIEKPGITY
YKEHEGAIYPDNTTDFQRADDKVYIGKEFKRIVQRIKDFLRNLVPRTESKELVRKDLQNFLKKENKNE
TVRFQSLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLMTCKMSQLERNIETIINTFHQYSVKL
GHPDTAGAGAGAKAGAGAAGHHHHHH
```

M-DOSCAT-i

```
>M-DOSCAT-i
MAGRGVTCPEQDKYRTITGKDIFTGLIGPMKICKKGSLHANGRQKDVDKEIGKEFKRIVQ
RIKDFLRNLVPRTESVAAHKKSHEESHKEAAGLEHHHHHH
```

Publications resultant from this thesis and personal contribution to each publication

Bennett, Richard J., Deborah M. Simpson, Stephen W. Holman, Sheila Ryan, Philip Brownridge, Claire E. Eyers, John Colyer, and Robert J. Beynon. 2017. "DOSCATs: Double Standards for Protein Quantification." *Scientific Reports* 7 (April). Nature Publishing Group: 45570. doi:10.1038/srep45570.

- I was responsible for all experimental work and data analysis and I prepared the draft manuscript and all figures.

Gómez-Baena, Guadalupe, **Richard J. Bennett**, Carmen Martínez-Rodríguez, Małgorzata Wnęk, Gavin Laing, Graeme Hickey, Lynn McLean, Robert J. Beynon, and Enitan D. Carrol. 2017. "Quantitative Proteomics of Cerebrospinal Fluid in Paediatric Pneumococcal Meningitis." *Scientific Reports* 7 (1): 7042. doi:10.1038/s41598-017-07127-6.

- I carried out the quantitative western blots and associated data analysis, and I prepared the draft figure summarising the western blotting results (Figure 6).