UNIVERSITY OF BIRMINGHAM

Surface analysis for proteomics via liquid extraction surface analysis mass spectrometry and liquid chromatography mass spectrometry

By

Nicholas Joseph Martin

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

The School Of Biosciences
College of life and Environmental Sciences
University of Birmingham
January 2016

Abstract

Liquid extraction surface analysis (LESA) is an ambient ionisation technique which allows direct analysis of surfaces coupled with mass spectrometry. LESA mass spectrometry has been used successfully to analyse small molecules, but there are a limited number of examples where the approach has been applied to protein analysis. The work presented here aims to develop novel applications of LESA mass spectrometry of proteins. LESA mass spectrometry was used to analyse intact proteins from polymeric membranes. The rationale for these experiments was the potential application to analyse proteins electroblotted following polyacrylamide gel electrophoresis, i.e., top-down proteomics, and in air monitoring. The subsequent focus was dried blood spot (DBS) analysis. An automated LESA based trypsin digestion protocol was developed and coupled with liquid chromatography tandem mass spectrometry to enable DBS proteomics, i.e., untargeted global protein identification via a bottom-up approach. Approaches for DBS proteomics (in the absence of LESA) were explored further using conventional digestion procedures coupled with protein depletion. LESA was also applied for targeted analysis of proteins from DBS, to determine variants of alpha-1-antitrypsin. Finally, native LESA mass spectrometry was developed to analyse non-covalent complexes from dried surfaces. Native LESA mass spectrometry successfully identified the haemoglobin tetramer directly from DBS.

The work presented in this thesis (Chapters 4 and 7) resulted in the publication of two articles in peer-reviewed journals on which I am first author and texts may be similar. The work in these papers was carried out by me and the articles were written by me in consultation with my co-authors. During this research, I have also published a review article which was written by me in consultation with my co-author.

Original research articles on which I am first author:

- Martin, N.J.; Griffiths, R.L.; Edwards, R.L.; Cooper, H.J. 'Native liquid extraction surface analysis mass spectrometry: Analysis of non-covalent protein complexes directly from dried substrates' ; Journal of the American society for mass spectrometry 2015, 26, 1320-1327
- Martin, N.J.; Bunch, J.; Cooper, H.J. 'Dried blood spot proteomics: Surface extraction of endogenous proteins coupled with automated sample preparation and mass spectrometry analysis' Journal of the American society for mass spectrometry, 2013, 24, 1242-1249.

Review articles on which I am first author*:*

- Martin, N.J.; Cooper, H.J. 'Challenges and opportunities in mass spectrometric analysis of proteins from dried blood spots' Expert Reviews in proteomics, 2014, 11, 685-695.

Nicholas J. Martin
First author

Helen J. Cooper
Senior Corresponding Author

Acknowledgments

There have been a number of people who I would to thank for their help and support over the 4 years in which I have been completing my PhD.

Firstly I would like to thank my supervisor Helen Cooper, who has given me fantastic guidance and keeping me motivated. In particular I would like to thank her patience and all the opportunities I received through being part of her group.

Secondly I'd like to thank my secondary supervisor Josephine Bunch for all her support and all the ideas she's given me. Mark Allen and Mark Baumert from Advion biosciences for their technical input and funding. I would also like to thank the EPSRC for their funding and Professor Rob Stockley for supplying me with dried blood spots. Many thanks too to the whole of Helen's research group past and present. In particular to Cleidiane Zampronio for all her day to day assistance with instrumentation as well as Andrew Creese and Rebecca Edwards for their technical discussions. I would also like to thank the whole of biosciences fifth floor for their friendship and creating such a brilliant working environment.

I would like to finish off by thanking my parents and my sister Kate for all their support and helping me establish an interest in science to begin with. I'd like to thank some special friends who've helped me through the more difficult periods of this PhD, Jack and Magda Charlton, Allan West and Nicola White. Finally I'd like to thank James Haycocks for being such a brilliant flatmate and to Lorna Thorne for being such a great girlfriend and for all the strength she's given me.

Contents

Abstract i

Chapter 1 : Introduction 1

1.1 Overview 1

1.2 Mass spectrometry 2

1.2.1 Introduction to mass spectrometry 2

1.2.2 Ionisation 3

1.2.2.1. Electrospray ionisation 3

1.2.2.2 Ambient ionisation 9

1.2.3.1The Orbitrap mass spectrometer 14

1.2.3.2 Ion trap mass spectrometry 19

1.2.3.3 Quadrupole mass spectrometry 21

1.2.3.4 Time of flight mass spectrometry 24

1.2.4 Tandem mass spectrometry 28

1.2.4.1 Collision induced dissociation 28

1.2.4.2 Selected reaction monitoring assays 30

1.2.5 Native mass spectrometry 31

1.3 Proteomics 33

1.3.1 Overview of proteomics 33

1.3.2 Gel electrophoresis 34

1.3.3 Liquid chromatography 36

1.3.4 Bottom up proteomics 37

1.3.5 Intact protein analysis 41

1.4 Dried blood spot analysis 44

1.4.1 DBS as a sampling format 45

1.4.2 Analysis of DBS by mass spectrometry 46

1.4.3 Protein analysis of DBS by mass spectrometry 47

1.4.3.1 Targeted analysis of proteins from DBS 47
1.4.3.2 Untargeted analysis of proteins 50
1.4.3.3 Future prospects for DBS proteomics 50
1.5 Alpha-1-antitrypsin deficiency 51
1.6 Aims and objectives 54
Chapter 2 Materials and Methods.....55
2.1 Materials 55
Reagents 55
2.2 Methods 56
2.2.1 LESA.....56
2.2.2 LC MS/MS.....56
2.2.3 Mass spectrometry 58
2.2.3.1. Direct infusion and LESA mass spectrometry 59
2.2.3.2 Data dependent CID 60
2.2.3.3 Selective reaction monitoring (SRM).....60
2.2.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE).....61
2.2.5 Zip tip purification.....62
2.2.6 Native PAGE 63
2.2.7 Solution phase digestions 63
2.2.8 Database searching.....64
2.2.9 Protein depletion 65
Chapter 3 : Liquid extraction surface analysis of intact proteins from polyvinylidene fluoride membranes 68
3.1 Introduction.....68
3.1 Optimisation of extraction solvent composition 69
3.2 Recovery of intact proteins from PVDF membranes 72

3.3 Recovery of intact proteins blotted from SDS PAGE....76

3.4 Negative ion mode LESA analysis of PVDF membranes....78

3.5 Elimination of SDS signal suppression....80

3.5.1 Reversed phase C_4 Zip tips....80

3.5.2 Cetyltrimethyl ammonium bromide....83

3.6 Recovery of intact proteins from PVDF membranes after native PAGE.....85

3.7 Optimisation of extraction parameters for intact protein analysis from PVDF membranes....87

3.7.1 Calculation of surface areas and surface protein concentrations....87

3.7.2 Sampling time....88

3.7.3 Liquid microjunction volume....95

3.7.4 Liquid microjunction dispension height....98

3.7.5 Determination of limit of detection limit for LESA analysis of βlg from PVDF membranes....102

3.8 Conclusion....106

Chapter 4 Bottom-up proteomics analysis of dried blood spots by liquid extraction surface analysis mass spectrometry....107

4.1 Introduction....107

4.2 LESA of dried blood and plasma spots.....108

4.2.1 Dried blood spots....108

4.2.2 Dried plasma spots....110

4.3 Bottom-up proteomics....111

4.3.1 DBS proteolysis....114

4.3.2 On card proteolysis....115

4.3.3 Surface sampling followed by proteolysis....116

4.4 Adaption for Automation 120

4.4.2 Effect of proteolysis times 129

4.5 Chapter 4 conclusion 136

Chapter 5 Bottom up proteomics of dried blood spots: solution phase chemistry and immunodepletion 139

5.0 Introduction 139

5.1 Proteolysis of blood proteins following punch and elution" from DBS samples 140

5.2 Immunodepletion 146

5.2.1. Treatment of samples with Plasma Proteome Purify 12 kit 147

5.2.2 Treatment of samples with Hemovoid kit 152

5.2.3 Combined immunodepletion of abundant plasma proteins and haemoglobin depletion 156

5.3 Conclusion 165

Chapter 6 Determination of alpha-1-antitrypsin variants through bottom up proteomic analysis of DBS 167

6.0 Introduction 167

6.1 A1AT coverage 169

6.2 Untargeted data dependent LC MS/MS of A1AT deficient DBS digests 172

6.3 Targeted data dependent LC MS/MS through inclusion lists 175

6.4 Multiple reaction monitoring analysis for determining A1AT variants 178

6.4.1. Analysis of synthetic peptides through direct infusion MRM analysis 179

6.4.2 LC MRM analysis of DBS digests 185

6.5 Chapter 6 conclusion 188

Chapter 7 Analysis of non-covalent protein complexes by liquid extraction surface analysis mass spectrometry 190

7.1 Protein standards 191

7.1.1 Myoglobin 193

7.1.2 Haemoglobin 200

7.2 Dried blood spots 211

7.2.3 Orbitrap analysis 211

7.2.4 Q-TOF analysis 215

7.3 Chapter 7 conclusions 220

Chapter 8 Conclusions 222

Future work 225

Chapter 9 References 229

Appendices 243

Appendix 5 255

Appendix 6 297

Appendix 7 Published Journal Articles 300

List of figures

Figure 1.1 Illustration of electrospray ionisation

Figure 1.2 Illustration of LESA sampling routine

Figure 1.3 Illustration of ESI chip

Figure 1.4. Schematic diagram of the Orbitrap cell

Figure 1.5 Schematic diagram of Orbitrap Velos™ ETD mass spectrometer

Figure 1.6 Schematic diagram of a linear ion trap

Figure 1.7 Schematic diagram of a quadrupole

Figure 1.8 Schematic diagram of Waters Synapt™ mass spectrometer

Figure 1.9 Illustration of nomenclature for peptide fragmentation

Figure 1.10 Illustration of SRM scanning method

Figure 1.11 Image of dried blood spot

Figure 3.1 LESA of intact proteins from PVDF membranes

Figure 3.2 LESA of intact myoglobin from PVDF membranes

Figure 3.3 LESA of PVDF membrane following electroblotting

Figure 3.4 Negative ion mode LESA of PVDF membrane following electroblotting

Figure 3.5 Zip tip purification for clean-up

Figure 3.6 Treatment of PVDF membranes with CTAB

Figure 3.7 LESA of PVDF membrane blotted from native PAGE

Figure 3.8 Effect of sampling time on the full scan signal to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.9 Effect of sampling time on the targeted SIM window to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.10 Effect of the number of mix cycles on the full scan signal to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.11 Effect of the number of mix cycles on the targeted SIM window to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.12 Effect of the microjunction volume on the full scan signal to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.13 Effect of the microjunction volume on the targeted SIM window to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.14 Effect of the microjunction dispense height on the full scan signal to noise ratio of beta lactoglobulin extracted from PVDF

Figure 3.15 Effect of the microjunction dispense height on the targeted SIM window to noise of beta lactoglobulin extracted from PVDF

Figure 3.16 LESA of β lactoglobulin from PVDF membranes

Figure 4.1 LESA of intact proteins from DBS

Figure 4.2 LESA of dried plasma spot

Figure 4.3 Direct infusion ESI analysis of dried plasma spot digests

Figure 4.4 Direct infusion mass spectrum of DBS digestions

Figure 4.5. TIC following LC MS/MS of DBS digests

Figure 4.6 Illustration of LESA based digestion procedure

Figure 4.7 Flow chart showing aspiration/dispense sequence for the digestion of a DBS

Figure 4.8 TIC following LC MS/MS of LESA based DBS digest

Figure 4.9 Categorisation of proteins identified from DBS

Figure 4.10 CID spectrum of peptide DLQNFLK

Figure 4.11 Reproducibility of protein identifications from DBS

Figure 4.12 TICs of LESA based digests across 5 time points

Figure 4.13 Average number of proteins identified from DBS across 5 digestion time points

Figure 4.14 Pairwise t-test of proteins identified from LC MS/MS of DBS digestion across different time points

Figure 4.15 Average number of peptides identified from DBS across 5 digestion time points

Figure 4.16 Pairwise t-test of peptides identified from LC MS/MS of DBS digestion across different time points

Figure 5.1 TICs of solution phase digestions across 5 time points

Figure 5.2 Average number of proteins and peptides identified from DBS across 5 digestion time points

Figure 5.3 Average number of proteins identified from LESA and solution phase digestions across 5 time points

Figure 5.4 Percentage of PSMs corresponding to haemoglobin peptides in DBS digests

Figure 5.5 TIC of DBS digests depleted of the top 12 plasma proteins

Figure 5.6 Average number of proteins identified from DBS following depletion of the top 12 plasma proteins

Figure 5.7 Average percentage of PSMs originating from haemoglobin and the top 12 plasma proteins

Figure 5.8 TIC of DBS depleted of haemoglobin

Figure 5.9 Average number of proteins identified from DBS following depletion of haemoglobin

Figure 5.10 Average percentage of PSMs originating from haemoglobin and the top 12 plasma proteins

Figure 5.11 TIC of DBS depleted of the top 12 plasma proteins and haemoglobin

Figure 5.12 Average number of proteins identified from DBS following depletion of the top 12 plasma proteins and haemoglobin

Figure 5.13 Average percentage of PSMs originating from haemoglobin and the top 12 plasma proteins

Figure 5.14 TIC of DBS depleted of haemoglobin and the top 12 plasma proteins

Figure 5.15 Average number of proteins identified from DBS following depletion of haemoglobin and the top 12 plasma proteins

Figure 5.16 Average percentage of PSMs originating from haemoglobin and the top 12 plasma proteins

Figure 6.1 TIC of LC MS/MS A1AT deficient DBS digestions

Figure 6.2 TIC of targeted LC MS/MS of A1AT deficient DBS digestions

Figure 6.3 Ion counts of 6 diagnostic A1AT peptides measured by MRMs

Figure 6.4 SRM spectra showing detection of synthetic peptides

Figure 6.5 Interfering peaks observed by SRM

Figure 6.6 Relative intensity of diagnostic peptides compared to blank samples

Figure 6.7 Ion counts of 6 diagnostic A1AT peptides measured by LC MRM

Figure 6.8 Relative intensity of diagnostic peptides compared to blank samples measured by LC MRM

Figure 7.1 Crystal structures of equine myoglobin and human haemoglobin

Figure 7.2 Native direct infusion ESI mass spectrum of equine skeletal muscle myoglobin

Figure 7.3 Native LESA mass spectrum of equine myoglobin from glass microscope slide acquired on Orbitrap Velos™

Figure 7.4 Native LESA mass spectrum of equine myoglobin from PVDF membrane acquired on Orbitrap Velos™

Figure 7.5 Native LESA mass spectrum of equine myoglobin from glass microscope slide acquired on Synapt™

Figure 7.6 Native LESA mass spectrum of equine myoglobin from PVDF membrane acquired on Synapt™

Figure 7.7 Native direct infusion ESI mass spectrum of human haemoglobins

Figure 7.8 Native LESA mass spectrum of human haemoglobins from glass microscope slide acquired on Orbitrap Velos™

Figure 7.9 Native LESA mass spectrum of human haemoglobins from PVDF membrane acquired on Orbitrap Velos™

Figure 7.10 Native LESA mass spectrum of human haemoglobins from glass microscope slide acquired on Synapt™

Figure 7.11 Native LESA mass spectrum of human haemoglobins from PVDF membrane acquired on Synapt™

Figure 7.12 Direct infusion ESI mass spectrum of whole blood diluted in ammonium acetate

Figure 7.13 Isolation of m/z 3795 from direct infusion ESI mass spectrum of whole blood

Figure 7.14 Native LESA mass spectrometry of DBS obtained on Orbitrap Velos™

Figure 7.15 Native LESA mass spectrometery of DBS obtained on Synapt™

List of tables

Table 2.1 SRM transitions used for identifying 6 diagnostic peptides of the alpha-1-antitrypsin protein

Table 3.1. Formation of liquid microjunctions at different methanol concentrations

Table 3.2 Formation of liquid microjunctions at different Acetonitrile concentrations

Table 3.3. Diameters and areas of droplets deposited onto PVDF membranes

Table 3.4. Optimum LESA parameters for the recovery of β-lactoglobulin

Table 3.5 Signal to noise ratios of β-lactoglobulin

Table 5.1. Proteins targeted by Proteome Purify 12 immunodepletion kit

Table 5.2. Two tailed unpaired t-test on protein identifications before and after depletion

Table 5.3. Two tailed unpaired t-test on the percentage of PSMs corresponding to haemoglobin

Table 5.4. Two tailed unpaired t-test on protein identifications before and after haemoglob depletion

Table 5.5. Two tailed unpaired t-test on protein identifications before and after depletion of the top 12 plasma proteins and haemoglobin

Table 5.6. Two tailed unpaired t-test on protein identifications before and after depletion of haemoglobin and the top 12 plasma proteins

Table 5.7. Two tailed unpaired t-test on protein identifications before and after depletion of haemoglobin only and using two kits together

Table 6.1 Serum A1AT concentrations expected from different A1AT genotypes

Table 6.2 Sequences and theoretical m/z ratios of the 6 diagnostic peptides required for determining A1AT variants

Table 6.3 Diagnostic peptides yielded from tryptic digestions of different A1AT genotypes

Table 6.4 Coverage of A1AT obtained from LC MS/MS analysis of DBS digest of a healthy individual

Table 6.5 Identifications of diagnostic peptides from LC MS/MS analysis of DBS digest of A1AT deficiency sufferers

Table 6.6 m/z ratios added to an inclusion list to identify diagnostic peptides required for determining A1AT variants

Table 6.7 Identification of diagnostic peptides from A1AT deficient DBS digests by targeted LC MS/MS

Table 6.8 Precursor ion m/z ratios and transitions monitored in MRM assay for determining A1AT variants

Table 7.1 Haemoglobin species identified from native ESI from haemoglobin standards obtained on the Orbitrap Velos™

Table 7.2 Haemoglobin species identified from native ESI from haemoglobin standards obtained on the Synapt™

Table 7.3 Haemoglobin species identified from native LESA of haemoglobin standards from glass microscope slide obtained on the Orbitrap Velos™

Table 7.4 Haemoglobin species identified from native LESA of haemoglobin standards from PVDF membrane obtained on the Orbitrap Velos™

Table 7.5 Haemoglobin species identified from native LESA of haemoglobin standards from glass microscope slide obtained on the Synapt™

Table 7.6 Haemoglobin species identified from native LESA of haemoglobin standards from PVDF membrane obtained on the Synapt™

Table 7.7 Haemoglobin species identified from native LESA of DBS in 10 mM ammonium acetate obtained on the Synapt™

Table 7.8 Haemoglobin species identified from native LESA of DBS in 10 mM ammonium acetate and 5% methanol obtained on the Synapt™

List of equations

Equation 1.1 The Rayleigh limit

Equation 1.2 The electrostatic field generated by the Orbitrap cell

Equation 1.3 Axial ion oscillations of ions in the Orbitrap cell

Equation 1.4 Quadrupolar field

Equation 1.5. Measurement of m/z by a time of flight mass analyser

List of abbreviations

A1AT	Alpha-1-antitrypsin
ACN	Acetonitrile
AGC	Automatic gain control
AUI	Advanced user interface
BSA	Bovine serum albumin
βlg	Beta-lactoglobulin
CI	Chemical ionisation
CID	Collision induced dissociation
CTAB	Cetyltrimethylammonium bromide
DBS	Dried blood spots
DPS	Dried plasma spot
DESI	Desorption electrospray ionisation
EI	Electron ionisation
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width half maximum

HPLC	High performance liquid chromatography
IGF	Insulin like growth factor
ITRAQ	Isobaric tagging for relative and absolute quantification
LC	Liquid chromatography
LC MS/MS	Liquid chromatography tandem mass spectrometry
LESA	Liquid extraction surface analysis
LESA MS	LESA mass spectrometry
LMJ SSP	Liquid microjunction surface sampling probe
MALDI	Matrix associated laser desorption ionisation
MAla	M Alanine
MVal	M Valine
MeOH	Methanol
MRM	Multiple reaction monitoring
MSUD	Maple syrup urine disease
MS/MS	Tandem mass spectrometry
PAGE	Poly acrylamide gel electrophoresis
PKU	Phenylketonuria
PMF	Peptide mass fingerprint

PPM	Parts per million
PSMs	Peptide spectral match
PVDF	Polyvinylideneflouride
QTOF	Quadrupole time of flight
SDS	Sodium dodecyl sulphate
SDS PAGE	Sodium dodecyl sulphate polyacrylamide el electrophoresis
SIM	Selective ion monitoring
SISCAPA	Stable isotope standards and capture by anti-peptide antibodies
SILAC	Stable isotope labelling of amino acids in cell culture
SRM	Selective reaction monitoring
TFA	Trifluoroacetic acid
Th	Thomsons
TIC	Total ion chromatogram
TLC	Thin layer chromatography
TMT	Tandem mass tags
TrisHCL	Tris(hydroxylmethyl)aminomethane hydrochloride
TSQ	Triple stage quadrupole
TOF	Time of flight

UHPLC	Ultra high performance liquid chromatography

Chapter 1 : Introduction

1.1 Overview

In the first decade of the 21st century a range of ambient ionisation techniques were introduced to mass spectrometry as a means of investigating surface associated analytes. Surface analysis made a range of analytically relevant surfaces accessible to mass spectrometry. Several ambient ionisation techniques, including liquid extraction surface analysis (LESA) [1], are based on electrospray ionisation. Electrospray ionisation and LESA are both soft ionisation techniques which makes them ideally suited for ionising non-volatile macromolecules such as proteins and peptides. The aim of the work contained within this thesis is to develop new applications of LESA mass spectrometry in the field of protein analysis.

LESA has previously been successfully used for the analysis of intact haemoglobin proteins from DBS [2]. This thesis aims to extend this work and develop novel methodology and applications of LESA. The work presented in Chapter 3 focuses on the analysis of intact proteins from polyvinylideneflouride (PVDF) membranes, and investigation of the possibilities of analysing proteins electroblotted from sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDSPAGE). The work presented in Chapter 4 investigates the application of LESA for the bottom up proteomics of dried blood spots. Bottom-up proteomics of dried blood spots is explored further in Chapter 5, with the

emphasis on solution phase chemistry, specifically the incorporation of immunodepletion strategies to aid analysis of low abundance proteins in DBS. The work presented in Chapter 6 is aimed at targeted analysis of proteins found in DBS, i.e., variants of the alpha-1-antitrypsin (A1AT), using LESA. The aim of the work presented in Chapter 7 is to develop LESA as a means for investigating non-covalent protein complexes directly from surfaces, focussing on tetrameric haemoglobin from DBS.

1.2 Mass spectrometry

1.2.1 Introduction to mass spectrometry

Mass spectrometry is a powerful tool for the analysis of molecular structure based on measuring the mass to charge ratio of ions. There are 3 basic components in mass spectrometry. The first stage is the formation of ions from a neutral substance. The next stage is mass analysis which separates ions on the basis of their mass-to-charge (m/z) ratio before the final stage which is detection of the separated ions. Mass analysis and detection always takes place under a very high vacuum and so too do several ionisation techniques [3].

The first mass spectrometers were built around a hundred years ago by particle physicists studying the electrical conductivity of gases. The very first mass spectrometer was built in 1912 by J.J. Thomson and his assistant F W Aston [4]. They developed an instrument they called the parabola spectrograph [5]. Thomson successfully obtained mass spectra of

several different gases and demonstrated that neon is composed of 2 different isotopes. The instrument was further improved by Aston who proved that 53 of the known elements had multiple isotopes and was later awarded the 1922 Nobel prize in chemistry for his efforts [4]. The field of mass spectrometry has changed immensely since then. It changed from largely a physical to chemical analysis technique in the 1940s and became an important tool for analysing biological samples in the 1980s [6].

1.2.2 Ionisation

Analysis of proteins by mass spectrometry was not possible until the late 1980s when two new soft ionisation techniques were developed: electrospray ionisation (ESI) and matrix associated laser desorption ionisation (MALDI) [3]. Forming ions from large non-volatile macromolecules is perhaps the biggest challenge for an ionisation method [7]. Before the introduction of ESI and MALDI, the most established ionisation mechanisms were electron ionisation (EI) and chemical ionisation (CI). These mechanisms vaporise a substance then bombard it with electrons in EI or reagent gases in CI. Both are able to produce ions of volatile species with high vibrational energy states. Ions formed by either technique are liable to undergo extensive decomposition and so these techniques are unsuitable for most biological samples [3]. As a result of their efforts the inventors of ESI and MALDI, John Fenn and Koichi Tanaka were awarded the 2002 Nobel prize in chemistry [8].

1.2.2.1. Electrospray ionisation

Electrospray ionisation (ESI) is an electrochemical process which produces gas phase ions from solution phase solubilised analytes. It has become the most widely used ionisation

method for the analysis of proteins and peptides and is largely responsible for advancements in proteomics research during the past few decades. The first report of using ESI for protein and peptide analysis was published in 1989,[9] but the ideas have been under development since the 1960s by Dole and co workers [10]. The group proposed a method for creating ions by using a pressurised gas to spray solutions of analytes through a highly charged hypodermic needle tip at atmospheric pressure. The effect of this was to disperse the solution into a mist composed of fine droplets. Solvent molecules within the droplets would evaporate and leave behind charged gas phase ions. Dole used this technique to ionise polystyrene molecules in solutions of acetone and benzene. It successfully produced gas phase ions of large polymers and, because the mechanism did not rely on bombardment with electrons or harsh chemicals, the ions remained intact and did not fragment [10]. Nevertheless the method needed refining and it was another 20 years before the idea was taken further. The biggest problem was that the process formed highly solvated clusters rather than individual gas phase ions, and the technology available lacked the ability to detect them. Mass analysers available in the 1960s were not sensitive enough to detect the ion currents provided by a low energy ionisation mechanism and lacked the resolution for measuring m/z ratios of multiply charged macromolecules [11]. ESI was further refined in the 1980s by John Fenn and co workers [9, 12]. Fenn's method differed from Dole's in a number of features: samples were sprayed in an aqueous organic mixture, a drying gas was used to desolvate the ions and samples were sprayed through a glass capillary with metal electrodes at either end to create a potential [12]. The drying gas successfully eliminated the problem of forming solvated ion clusters. The gas flowed against the direction of spray which helped desorb

the ions from the solvent molecules. The drying gas helped form individual gas phase ions rather than solvated ion clusters that were produced in Dole's method [12, 13]. The process is illustrated in Figure 1.1. By the 1980s mass analysers had become more sensitive, and the focus of mass spectrometry research had begun to change from organic compounds to biological samples and the technique was demonstrated for use in protein and peptide analysis [9].

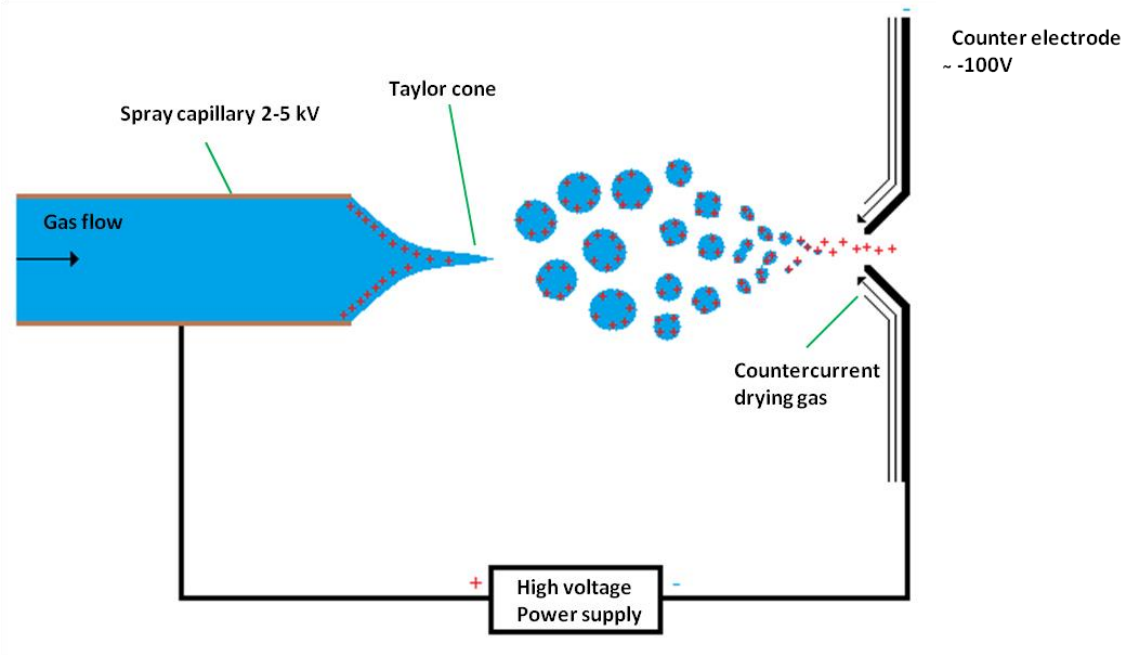


Figure 1.1. Illustration of the electrospray ionisation process

There are two mechanisms proposed to explain ion formation during ESI. Their influence appears to depend on the analyte being ionised [14]. Both explain ion formation in a similar way, and can be understood by considering the effect that electrical charging has upon liquids. A positive electric potential is applied to the capillary which generates

electrical repulsion forces within the liquid [15]. The repulsive forces push the liquid away from the capillary and the surface tension at the liquid air boundary of the droplet pulls back. When the two forces are equal, a conical shape known as the Taylor cone forms at the liquid air boundary [15]. This phenomenon was described by Taylor in 1964 [16]**.** If the electrical charging increases and the repulsion becomes greater than the surface tension, small droplets are emitted in a jet from the cone apex. The counter electrode of the mass spectrometer is held at a negative potential, which attracts the droplet emitted from the capillary and a small current is generated [17]. This process disperses analytes from the solution towards the inlet of a mass spectrometer within small highly charged droplets. Charge transfer can be explained by either the charged residue or ion evaporation model. After the charged droplets have left the capillary solvent molecules begin to evaporate. As the volume of the droplets decreases the charge density increases. Eventually the repulsive forces of the charged droplets rises above the surface tension of the liquid and the droplets explode [18]. The point where the surface tension is equal to the electrostatic repulsion is known as the Rayleigh limit and is shown in equation 1.1. After each droplet explodes several new droplets form and the process repeats producing smaller droplets after each fission cycle [17].

$$Q_R = 8\pi(\varepsilon_0 \gamma R^3)^{1/2}$$

Equation 1.1 The Rayleigh limit, where Q_R = the number of surface charges of the droplet at the Rayleigh limit, ε_0 = the permittivity of the solvent, γ = the surface tension of the solvent and R = the radius of the droplet [18].

Fission and evaporation of the droplets takes place in both models but the final stages are slightly different. The ion evaporation model assumes that the charged droplets become so small and the charge density becomes so great that the repulsive forces are strong enough to push a charged solute ion out from the droplet [19]. Evaporation of a charged ion introduces a desolvated gas phase ion into the vacuum of the instrument [20]. Ion evaporation is different to the continual evaporation of neutral solvent molecules that forms the small droplets. Ion evaporation takes place when the droplets are smaller than 20 nm in diameter [21]. At this point the droplet must contain more than one solvent molecule, the electrostatic repulsion must be large enough to push the solute molecule out of the droplet and it must be larger than the critical radius of the Rayleigh limit to prevent further droplet fission [20]. The charged residue model was proposed in Dole's 1968 paper on ESI and assumes that droplet fission becomes a more important factor in ion formation. In this case droplets repeatedly burst and reform, becoming so small that they reach a point when the final droplets contain only one solvent molecule. The final solvent molecule then evaporates and the charge is retained on the non-volatile solute molecule [19]. High molecular weight multiply charged ions are ionised by the charged residue model and smaller singly charged species are ionised by ion evaporation [14].

ESI is a very versatile ionisation technique. It is capable of ionising a wide range of different compounds of almost any molecular weight [22]. Its most significant application has been for the analysis of proteins and peptides which presents a challenge for most

other ionisation techniques. ESI suits protein and peptide analysis for several reasons. It is a soft ionisation mechanism that produces intact quasimolecular ions with low internal energy which causes little or no fragmentation [23, 24]. ESI is also capable of producing ions from non-volatile and thermally labile species [24]. ESI is unique amongst ionisation mechanisms in that it produces multiply charged species of high molecular weight [25]. Multiply charged peaks are particularly useful for analysing proteins and peptides as the operating m/z ranges of many commercial available mass spectrometers such as ion traps, orbitraps and quadrupoles have upper m/z limits of around 2000 Th[26-28]. Proteins and peptides are large molecules so the presence of multiple charges helps bring these large ions into the operating ranges of these instruments. Introducing ions directly from the solution phase into the gas phase means that ESI is easily coupled to a liquid chromatography system[13]. Proteomics experiments often require analysis of complex mixtures of proteins and peptides which require separation before analysis. Liquid chromatography can be performed with online ESI mass spectrometry analysis. The liquid chromatograph separates the complex mixture into its constituents which then directly elute from the column and are introduced into the mass spectrometer [29].

Since ESI was introduced it has become widely-used for the introduction of samples at lower flow rates. Fenn's original technique used flowrates of around 1-20µl per minute [9]. ESI flow rates have now dropped to a few tens of nanolitres per minute which is referred to as nanoelectrospray [30]. Electrospraying samples at lower flow rates has a number of advantages: the approach is more sensitive, couples to HPLC more efficiently and consumes less sample [30]. Most of the advantages are due to the fact that the process forms a smaller Taylor cone which produces smaller droplets in the ESI. Smaller

droplets have a greater charge per surface area and require fewer fission cycles to produce gas phase ions. Smaller droplets also contain less solvent which are easier to desolvate [31]. The smaller Taylor cone means that the ESI can be placed closer to the orifice of the mass spectrometer so more of the ions are introduced into the instrument [26]. The smaller Taylor cone also reduces the amount solvent entering the instrument which is less demanding on the vacuum system [26]. Nanoelectrospray is less sensitive to the effects of ion suppression from contaminants and ensures greater charging of the analyte molecules than conventional flowrates [32].

1.2.2.2 Ambient ionisation

ESI became established as the ionisation method of choice for the analysis of proteins and peptides in the 1990s and proteomics became an important field of biomedical research. The 21st century saw the development of a range of techniques for sampling analytes from the surface of a solid object at ambient conditions, then directly ionising and introducing them into a mass spectrometer. These became known as ambient ionisation techniques [33]. The term is slightly misleading: Many previously existing ionisation techniques such as conventional ESI, atmospheric pressure chemical ionisation and atmospheric pressure photo ionisation take place at ambient pressures but are not included in this category [34]. What they all have in common is that they directly sample a surface at ambient conditions with little or no prior sample preparation. Most ambient ionisation techniques incorporated a sampling procedure into a previously existing ionisation technique such as ESI or laser desorption [35]. There are many

different variations available and over 40 different ambient ionisation techniques have been described [35]. The development of ambient ionisation techniques became very popular after the invention of desorption electrospray ionisation (DESI) in 2004 [36]. DESI is an ESI based technique. A solvent is electrosprayed towards a surface by use of an electric field and nebulising gas. When the charged droplets strike the surface they desorb analytes from the surface which are directed into the inlet of a mass spectrometer. DESI can be used to analyse a surface without any sample preparation [36]. The technique has been used in analysis of dried blood spots, plant tissues, drugs and many other surfaces [37-39]. It was significant in demonstrating the utility of directly sampling surfaces of interest outside the instrument and several other techniques were developed to investigate this idea further, including direct analysis in real time in 2005, [40] paper spray ionisation in 2009 [41] and LESA.

LESA is a sampling method that is often considered to be an ambient ionisation technique [42]. It was launched commercially by Advion Biosciences as one of the modes of function on their Triversa Nanomate™ nanoelectrospray platform [1]. The principle behind the extraction technique was developed in 2002 by Gary Van Berkel [43]. It was an ESI based process that was originally designed to extract small molecules from thin layer chromatography (TLC) plates for analysis by mass spectrometry. The researchers constructed a homemade surface sampling probe composed of two capillary tubes, one inside the other. A solvent would be applied to the surface through the outer capillary which would be aspirated through the central capillary and interfaced with an ESI source.

Depositing the solvent onto the surface of the TLC plate created a wall-less liquid microjunction between the surface, the analytes and the probe. The liquid microjunction allows passive diffusion of analyte molecules from the surface which could then be aspirated into the sampling probe and introduced into the mass spectrometer [43]. The process was originally known as the liquid microjunction surface sampling probe (LJM SSP). In 2010 the technique was adapted to be merged onto the Triversa Nanomate™ [1]. Van Berkel's original probe used a continuous flow source that would continually apply solvent to the surface from one capillary which would then be continually aspirated through another. The Triversa Nanomate™ could not accommodate continuous flow as its probe consists of a robotic pipettor which could only perform discrete aspiration or dispense cycles. As an alternative a single droplet was aspirated from a solvent reservoir, and applied to the surface. The single droplet formed the liquid micro junction between surface sampling probe and the surface, and would be held in contact for a few seconds. Analytes would diffuse into the solvent via passive diffusion and after the delay the droplet would be re aspirated and directly injected into the mass spectrometer via chip based ESI [1]. An outline of the extraction process can be shown in Figure 1.2.

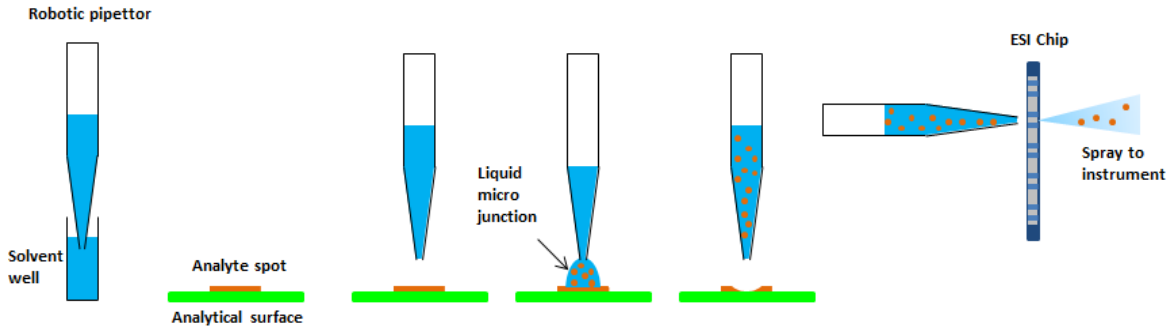


Figure 1.2. Illustration of LESA sampling routine

The Triversa Nanomate™ infuses samples via a variant of nanoelectrospray known as chip based ESI. In this process, the glass capillaries used in conventional ESI are replaced with a conductive pipette tip and a small chip composed of monolithic silicon. The chip has a series of 10 µm nozzles etched into it and act as the ESI emitters[44]. The robotic pipette tip filled with the ESI solution is sealed against the chip, a gas and a voltage are applied which spray the solution through the micro fabricated nozzles [45]. The chip based ESI is similar to conventional ESI, but produces highly stabile ESI currents and easy control of ionisation parameters [45]**.** An image of the chip is shown in fig 1.3.

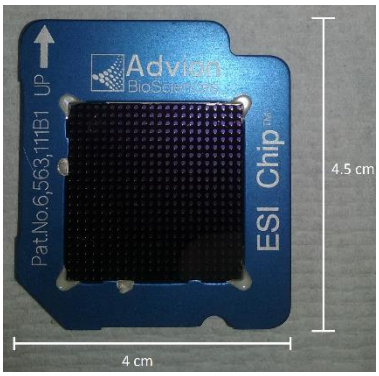


Figure 1.3 Triversa Nanomate™ ESI chip

LESA uses a soft ionisation technique that makes it applicable for a wide range of analytes including non-volatile macromolecules. Data have been published on the analysis of small organic molecules, lipids, peptides and intact proteins [2, 46-48]. LESA makes it possible to analyse a range of substances directly from surfaces in a manner that requires no sample preparation. It is non-destructive with the regards to the surface; it does consume the analyte but only a tiny amount is extracted and this often allows repeated sampling [2, 49]. LESA has been used successfully in the analysis of tissue

sections to monitor drug metabolism in pharmacokinetic studies[46]. After a drug has been administered to an animal, different areas of the tissue section can be sampled to identify the presence of the drug in different organs. LESA can be used as an alternative to whole body autoradiography, with the added advantage that it requires no sample preparation other than sectioning the tissue, it can distinguish the parent drug from its metabolites and can provide relative quantification [46, 49]. Applications of LESA have also been demonstrated in assessing the safety of foods by Eikel *et al*. [50]. The group used it to identify the presence of pesticides on fruits and vegetables and the method could distinguish whether the pesticides were on the surface or within the flesh [50]. LESA has been used for environmental applications, such as identifying the presence of brominated flame retardants in waste plastics [51]. Other applications have been demonstrated for use in biomedicine such as identifying biomarkers and for diagnostic purposes. Stegemann *et al.* [47] used LESA as a method for identifying lipids directly from the surface of human endarterectomies to identify biomarkers of atherosclerosis. The group studied arterial tissue from healthy controls and patients suffering from atherosclerosis. They identified 150 different lipids, 25 of which were unique to endarterectomies and found that LESA performed as well as direct infusion ESI of tissue lipid extracts [47]. Edwards *et al.* [2]demonstrated an application for the analysis of intact haemoglobin proteins from DBS by use of liquid extraction surface analysis. The work was the first time intact proteins had been analysed using LESA, and was able to distinguish clinically significant haemoglobin variants such as sickle cell disease. The method could analyse the proteins directly from the surface of the DBS with no sample preparation in a time period as little as 3 minutes per sample [2].

1.2.3 Mass analysers

Three different mass spectrometers were used in the work carried out in this thesis. These three were an Orbitrap Velos™ ETD mass spectrometer, a triple stage quadrupole (TSQ) Vantage™ mass spectrometer and a Waters Synapt™ Quadrupole-time of flight mass spectrometer.

1.2.3.1The Orbitrap mass spectrometer

In 2005 a new mass spectrometer with a new method of mass analysis was released by Thermo Electron. The instrument was known as the orbitrap mass spectrometer and had been under development by Alexander Makarov since the mid 1990s and patented in 1999 [52].

The orbitrap mass spectrometer uses an electrostatic trapping principle for storing ions. It uses image current based detection and Fourier transformations for signal processing [53]. The orbitrap cell is a small unit composed of two different electrodes, an inner spindle electrode encased within a larger barrel like electrode [53]. An illustration of the orbitrap is shown in Figure 1.4.

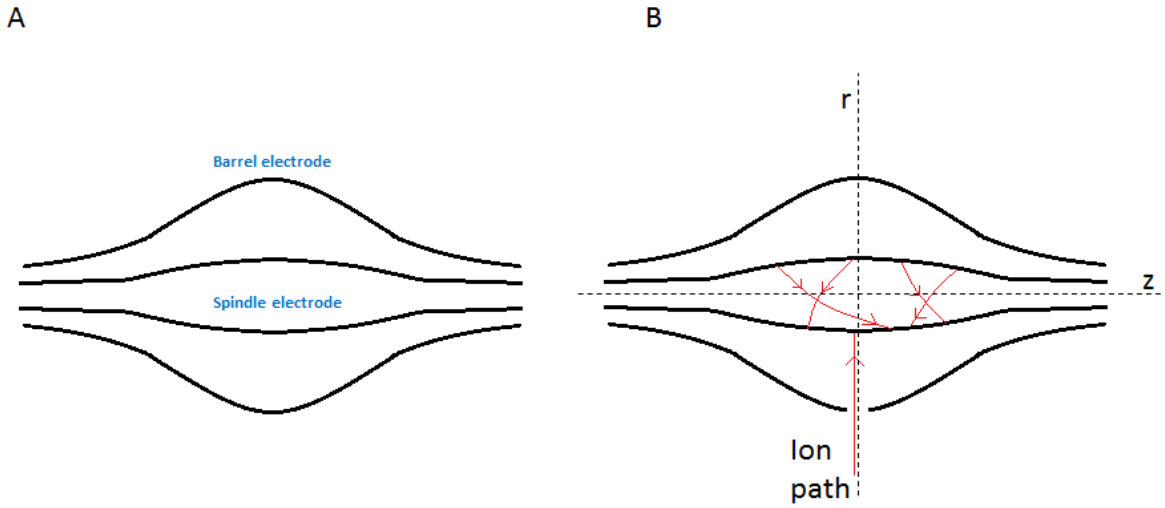


Figure 1.4 Schematic diagram of the orbitrap cell (A) Geometry of trap, (B) Geometry of trap showing axis of movemant and charachterstic ion motion

There is a small space between the two electrodes for ion storage and the trap measures a diameter of a few centimetres [54]. A potential is applied between the two electrodes which generates a static electric field. The central electrode is held at a negative potential of around -5 kV and the outer electrodes are held at a ground potential. Ions are injected into the trap at a tangent to the spindle electrode and form an orbit around it. The motion traps the ions within an electrostatic field which can be described by equation 1.2.

$$U_{(r,z)} = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 \ln\left[\frac{r}{R_m}\right] + C$$

Equation 1.2. The electrostatic field of within the orbitrap cell. U = electrostatic potential, r and z are the cylindrical co-ordinates z is the distance along axis of the central spindle and r is the radial distance out from the centre of the electrode, C is a constant, k is the field curvature and R_m is the characteristic radius [55]. The characteristic radius is the radius of an orbit when the field stops attracting ions to the centre of the spindle electrode and begins to repel them [56].

The trapping approach was developed in 1923 by K H Kingdon, where it was used as a means of storing ions and not as a mass analyser [57]. In the late 1990s when Alexander Makarov investigated its potential as a mass analyser and developed a new means of mass analysis based on axial harmonic oscillations of ions along the spindle electrode during orbit. As the ions orbit the spindle they are free to move up and down its z axis. When the oscillation is combined with the rotation of the orbit the ion paths becomes a series of complicated spirals. The two types of motion, oscillation up and down the z axis of the cell and rotation of ions around the spindle are completely independent of each other [55]. The total kinetic energy of an ion at orbit is affected by both types of motion but the frequency of ion oscillations down the z axis is dependent on the ions mass to charge ratio. The motion along the z axis can be described as a simple harmonic oscillator, a type of motion also present in a swinging pendulum or a load on a spring [53]. Equation 1.3 a explains this motion and 1.3 b shows how the mass to charge ratio of an ion be calculated from the frequency of axial oscillations.

a $$z(t) = z_0 cos(\omega t) + \sqrt{(2E_z/k)}\, sin(\omega t)$$

b $$\omega = \sqrt{(q/m)k}$$

Equation 1.3 Axial ion oscillations along the z axis of the orbitrap spindle electrode, (a) equation of motion (b) isolation of frequency of oscillations to give the charge to mass ratio of an ion. Where ω= the frequency of axial oscillations in radians/ per second E_z = energy characteristic of ion motion along the z axis k is the field curvature of the orbitrap cell and q/m = charge per mass ratio of an ion [55].

Mass analysis is possible because the frequency of axial oscillations along the z axis of the trap is dependent upon m/z and not the kinetic energy of the ions or their conditions upon injection into the cell [54]. Once a packet of ions is injected into an orbitrap the ions will separate into narrow rings of the same m/z. The rings oscillate in both halves of the cell and induce a current in the outer electrodes of the cell known as an image current. The image current forms the basis of signal detection in the orbitrap cell [58]. The current is then amplified and converted from an analogue to a digital signal which displays the intensity of the current against time. Displaying signal intensity in a temporal fashion is known as a transient. The transient is the processed with a fast Fourier transform. The fast Fourier transform generates a signal in the frequency domain which shows the frequency of the image current. The frequency of the image current is proportional to the frequency of axial oscillations of the ions and as shown by equation 1.3b this frequency is directly proportional to the ions charge per mass ratio which can generate a mass spectrum[53].

Using image currents for detection and Fourier transformations for signal processing gives the orbitrap has a high mass resolution and accuracy. Orbitrap mass spectrometers have a mass accuracy up to 2 parts per million (ppm) with external calibration and a mass resolution of 240,000 for a 768 ms transient at m/z 400 [59]. Previously Fourier Transform Ion Cyclotron Resonance (FT ICR) had been the only instruments available for high resolution mass spectrometry.

Orbitrap mass spectrometers have been used highly effectively in conjunction with other mass analysers and a range of hybrid instruments are available. Combining other mass analysers in tandem with the orbitrap provides additional options for fragmentation and scanning. Linear ion trap-orbitrap hybrids are available with the Orbitrap Velos™ and Orbitrap Elite™ [59, 60] The former was the instrument used for the majority of the work in this thesis and a schematic diagram is shown in Figure 1.4.

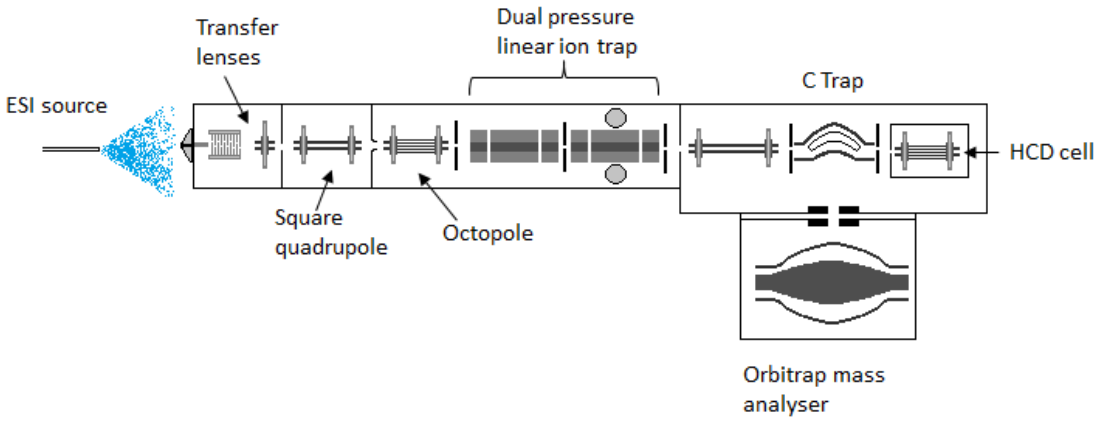


Figure 1.5 Schematic diagram of an Orbitrap Velos™ mass spectrometer

The most crucial development in orbitrap mass spectrometry, other than the mass analyser itself, was the C trap as a means of introducing ions into the orbitrap. The orbitrap mass analyser functions at best when coupled to a pulsed ion source. Makarov's prototype orbitrap mass spectrometer was coupled to a source that generated ions with

a pulsed laser so the two coupled well [55]. ESI was the most effective source in generating ions of large macromolecules like proteins and peptides which is where the resolution of the orbitrap would be most useful. ESI forms ions in a continuous rather than a pulsed manner so combining the two was a major challenge. The problem was eventually solved by using a radio frequency only quadrupole which was used to store ions produced in the spray prior to injecting them into the orbitrap. The ions were injected into the trap in a short pulse orthogonally from the spray source and the quadrupole. The quadrupole effectively decoupled the orbitrap mass analyser from the ESI source. Ions were injected by use of a high voltage electric pulse which lasted only a few hundred nanoseconds. The quadrupole had a curved shape similar to that of the letter "C" and was coined the C trap [61]. It is a highly successfully means of coupling the trap to an ESI source and has been used in every orbitrap mass spectrometer released commercially.

1.2.3.2 Ion trap mass spectrometry

The Orbitrap Velos™ mass spectrometer used in the bulk of the work shown here in this thesis was a hybrid instrument composed of an orbitrap and a linear ion trap mass analyser. The ion trap is composed of a conglomeration of four quadrupolar rod electrodes with an end cap at either end. A diagram of the linear ion trap is shown in Figure 1.6. Application of a radio frequency waveform to the quadrupolar rods generates an electric field that traps ions within the radial (xy) dimension and the electrodes at either end trap the ions within the axial (z) dimension by applying a DC voltage [62].

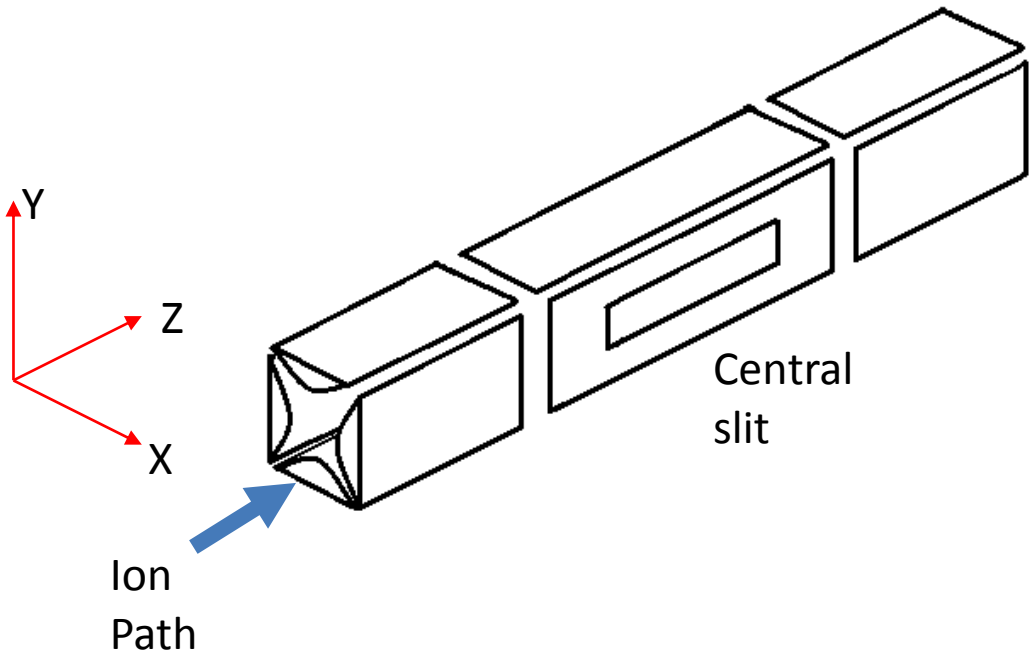


Figure 1.6 Schematic diagram and ion of a linear ion trap and axis of motion

The pressure of linear ion traps is often higher than most other mass analysers; a bath gas is often added to the trap to help improve ion transmission. Low energy collisions help the ions lose kinetic energy so they spend more time in the middle of the trap which traps the ions more effectively [63]. The trapping method means that ions have to be either injected in or ejected out of the trap. Two of the planar quadruple rods have slits cut into them which means that ions can be ejected through the slits in a radial or out the ends in an axial fashion [63]. Ejection through either means is mass selective and is achieved by applying an AC current to resonate ions of a specific m/z [64]. During axial ejection the current is generated between the two planar rods with the slits and when performing radial ejection the AC is generated between the planar rods and the end cap. Mass selective ejection is the principle of mass analysis. A particular ejection voltage will

only eject ions of a specific m/z from the trap, so to scan a chosen m/z range these voltages are ramped to eject a chosen range of m/z values [65]. Once ejected from the trap they strike an electron multiplier detector [62] which generates a small current. The current is then amplified and the whole mass range can be scanned to generate a mass spectrum.

By using mass selective ejection ions can be manipulated in a number of different ways. It is possible to eject one specific m/z, eject a range of m/z values or isolate a specific ion within the trap and ejecting all other m/z ranges [63]. Fragmentation of the trapped ions can be induced by collisions with the bath gas. Without excitation the collisions are low energy and insufficient to yield successful fragmentation. Excitation can be induced in the ions by either increasing the strength of the quadrupolar field generated by the planar rods or by inducing a small alternating auxiliary voltage between the two end caps. Excitation increases ion motion which increases the energy of collisions with the bath gas which become great enough to produce unimolecular fragmentation [63, 66]. The ability of an ion trap to perform mass selective ejection enables it to perform multiple rounds of fragmentation [67].

1.2.3.3 Quadrupole mass spectrometry

Some of the work carried out in chapter 6 was performed on a triple quadrupole mass spectrometer. Quadrupoles are common components of many different mass spectrometers, where they are used as mass filters, ion guides or mass analysers. They were invented in the 1950s by Wolfgang Paul. A quadrupole is composed of a ring of four

hyperbolic rods, each of which has an alternating radiofrequency potential applied to it. At any given time two of the rods will be a positive and two will be negative which produces a potential difference between the two pairs [68]. An ion path is aimed through the centre of the device along the z axis and the radio frequency fields stabilise the trajectories of the ions along the x and y axes. The ions passing through the device will be momentarily attracted to the rods of the opposite potential, but will be counteracted by a repulsive force as soon as the polarities switch. A diagram of a quadrupole is shown in Figure 1.7. Both the attractive and repulsive forces are proportional to the distance from the rods which focus the ions in the centre of the device [68].

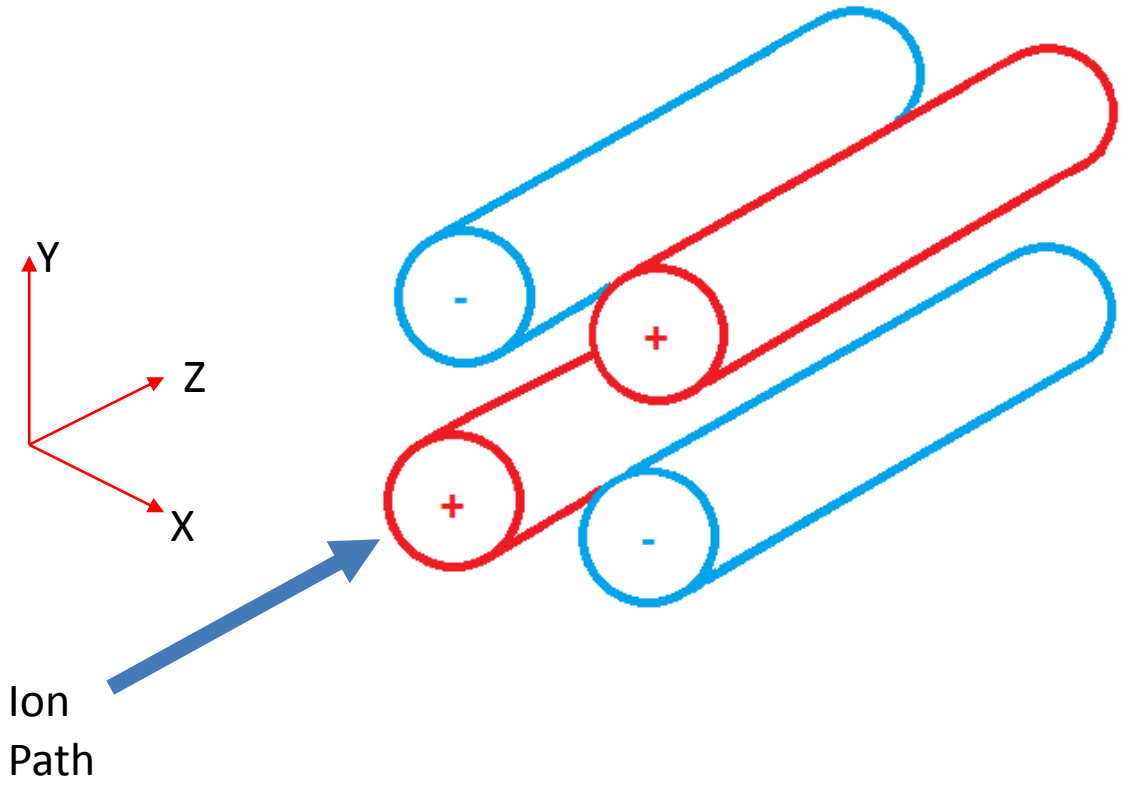


Figure 1.7 Schematic diagram and axis of motion of ions in a quadrupole

The field generated by a quadrupole maintains the ions trajectories in the X and Y dimension. Unlike the ion trap, motion in the Z dimension is unaffected so quadrupoles cannot store ions in any way. The field generated between the rods is shown in equation 4.

$$\Phi_0 = +(U - V\cos\omega t) \ \ and - \Phi_0 = -(U - Vcos\omega t)$$

Equation 4. Quadrupolar field. Where ϕ_0 is the potential of the rods, ω is the angular frequency of the field in radians per second, t is any time point, U is the direct potential between the two rods and V is the zero to peak amplitude of the RF voltage [68].

Equation 4 explains how quadrupoles can act as ion guides, but quadrupoles are also commonly used as mass filters. Mass to charge selection can be achieved by selectively stabilising the x and y trajectories to ions of a certain m/z range. For a given potential, only ions of a specific m/z will have a stable trajectory and be transmitted through the device whereas others will be deflected away [6]. When the ions reach the end of the device they strike an electron multiplier detector which generates a current. At any one voltage only one m/z reaches the detector so in each scan the voltages applied to the rods are steadily increased so the full m/z range reaches the detector [6].

A quadrupole can be used as a standalone instrument but is often more powerful when used in an array of 3 quadrupoles in series known as a triple quadrupole. The triple

quadrupole mass spectrometer was first demonstrated in the late 1970s and opened up several new options for fragmentation [69]. In this set-up the second quadrupole acts as a collision cell and both the first and the third quadrupole can be used to select a specific m/z or scan the whole range which are examples of tandem mass spectrometry experiments. The triple quadrupole was the first instrument available to put this idea into practice. Tandem mass spectrometry experiments available on triple quadrupole mass spectrometers include product ion scanning, precursor ion scanning, neutral loss scanning, and selected reaction monitoring experiments [68].

1.2.3.4 Time of flight mass spectrometry

The final type of mass spectrometer used for work in this thesis was a quadrupole time of flight (QTOF) mass spectrometer. This is a hybrid instrument, composed of a quadrupole and a time of flight mass analyser in tandem.

A time of flight (TOF) mass spectrometer contains a drift tube. An ion pulse is injected into the drift tube The ions receive an acceleration pulse on injection in the tube. The tube has a detector such as a microchannel plate at the far end [70]. The applied field is constant so all ions no matter what m/z have the same kinetic energy. According to Newtonian mechanics the kinetic energy of a moving object is proportional to its mass and velocity, so ions of a different m/z will have a different velocity when the same kinetic energy applied. The acceleration voltage separates out the ions in a spatial manner as lower m/z and faster ions will reach the detector first. The time it takes each

ion to reach the detector is measured and as the distance of the flight tube is known the velocity of the ions can be calculated [68]. The mass to charge ratio of the ion can then be calculated by the formula shown in equation 5.

$$(m/z)^{1/2} = \left(\frac{\sqrt{2eV_s}}{L}\right)t$$

Equation 5. Measurement of m/z in a time of flight mass spectrometer, where V_s =accelerating potential, L = distance of ion path t = flight time [68].

The first TOF mass spectrometer was described in 1948 [71]. Early instruments had linear flight tubes that could be several meters long and had problems with poor resolution. They did possess good sensitivity, a fast scanning rate and could analyse ions of almost any m/z range. Theoretically there is no upper m/z limit of a TOF mass analyser. TOF mass spectrometers have undergone a number of refinements to improve their resolution. Early instruments had problems applying a completely uniform acceleration voltage across all ions. The ions would receive a range of different kinetic energies from the source so when the acceleration pulse was applied, the total kinetic energies of the ions would not be identical. Ions of the same m/z could have a different kinetic energy and flight time so the peaks in a spectrum would be very broad [70]. A solution is to accelerate the ions after a short delay, which is known as delayed extraction. In the delay period ions with different kinetic energies separate slightly, so receive different energies from the acceleration pulse. The energy applied depends on the ions' proximity from the acceleration electrode and corrects for initial kinetic energies applied by the source. It

ensured that the energy distribution of ions is more uniform and peaks became less broad [70]. The resolution of a TOF can also be increased by adding a reflectron in the drift tube. A reflectron changes the ion path of the drift tube from a linear to a V shape. Both the accelerating field and the detector are at the same end of the drift tube and the reflectron is placed at the other [72]. Ions are injected into the drift tube are then reflected back to the detector by the reflectron which acts as an ion mirror. Reflectrons increase the path length without increasing the size of the drift tube. It helps increase mass resolution as the ions separate more over a greater path and the reflectron further corrects the kinetic energy balance of ions with the same m/z [72]. More energetic and faster moving ions spend more time in the reflectron so less energetic ions of the same m/z catch up. Instruments can contain multiple reflectrons for this purpose [73].

Time of flight mass spectrometers are the instruments of choice for pulsed ionisation techniques such as laser desorption ionisation and MALDI. Combining a pulsed method of mass analysis with a continuous ion source such as ESI is a challenge which was solved by use of orthogonal acceleration [74]. In this case the flight tube is set orthogonally to the ESI source and the ion beam. A “pusher” is used which provides an accelerating voltage pulse off axis to the ion beam and injects ions into the drift tube. The velocity component is independent of that produced to the spray so it is not effected by initial kinetic energies formed upon ionisation [74]. The ion pulse last only 10-100 nanoseconds, but a new population of ions is not injected until the ions of the highest m/z have reached the detector. In this time period the ion beam will begin to fill the orthogonal pusher before

they are injected into the flight tube. Waiting for the previous scan to finish does mean the instrument has a slower scanning rate compared to conventional TOF mass spectrometer [74].

Orthogonal acceleration time of flight mass spectrometers have been used highly successfully as hybrid QTOF mass spectrometers. The TOF mass analyser is used in tandem with two quadrupoles. The first is used to select a specific m/z which is fragmented in the second and then analysed by the TOF [75]. The Waters Synapt™ mass spectrometer used in this thesis is an example of such an instrument; however it only contains a single quadrupole used in tandem with a travelling wave ion mobility cell as shown in Figure 1.8 [76]. Quadrupole time of flight mass spectrometers have proved to be highly effective in analysing high m/z ions of native protein complexes [77].

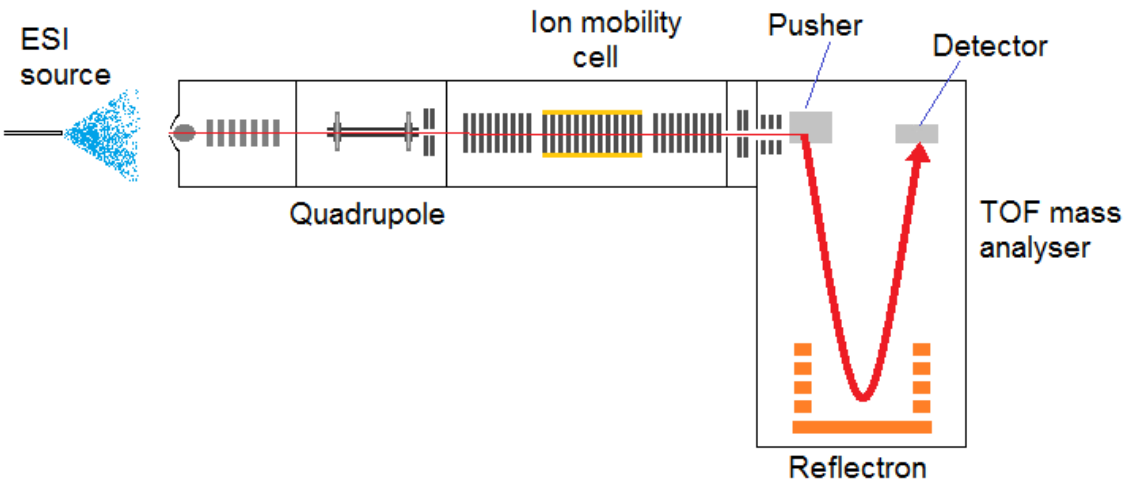


Figure 1.8 Schematic diagram of a Waters Synapt™ mass spectrometer

1.2.4 Tandem mass spectrometry

ESI mass spectrometry is very useful for obtaining molecular weight information of intact proteins and proteolytic peptides. Structural information however can be obtained from fragment ions. The process by which a precursor ion is characterised according to its fragments is known as tandem mass spectrometry (MS/MS) [6]. There are multiple fragmentation methods available including collision induced dissociation and electron transfer dissociation. One of the most commonly used MS/MS protocols is product ion scanning. In these experiments, a mass spectrum of the precursor ions is acquired. Ions of a particular m/z are then fragmented and then a second mass spectrum is acquired from which the m/z of the fragments can be determined [78].

1.2.4.1 Collision induced dissociation

There are a number of techniques available for fragmenting ions; the most commonly used technique is collision induced dissociation (CID). CID is achieved by colliding ions with an inert gas such as helium or argon in the mass spectrometer [79]. Increasing the kinetic energy of the ions will lead to higher energy collisions and this kinetic energy is converted to internal energy in the collision. In proteomics, research is focused on peptide fragmentation. The mechanism behind CID peptide fragmentation has been explained by the mobile proton model [80]. The model assumes that the protons on a charged peptide are retained on the side chains of basic amino acid residues or the N

terminus of the peptide and that fragmentation takes place when the proton is transferred to the peptide backbone. Transfer of the proton requires energy, [80] so CID increases the internal energy of ions through collisions as they pass through the mass spectrometer [81]. One of the downsides of CID is that the fragmentation pattern is non-specific: the molecule breaks at the weakest or most labile bond [82]. This can make CID difficult for identifying labile post translational modifications, and certain amino acids such as proline fragments very differently to others. The presence of proline in a peptide tends to favour fragmentation at the N-terminal bond upstream of the proline residue[83]. There are a number of positions where a peptide can fragment and a system of nomenclature has been described by Roepstorff *et al.* [84] as shown in figure 1.9.

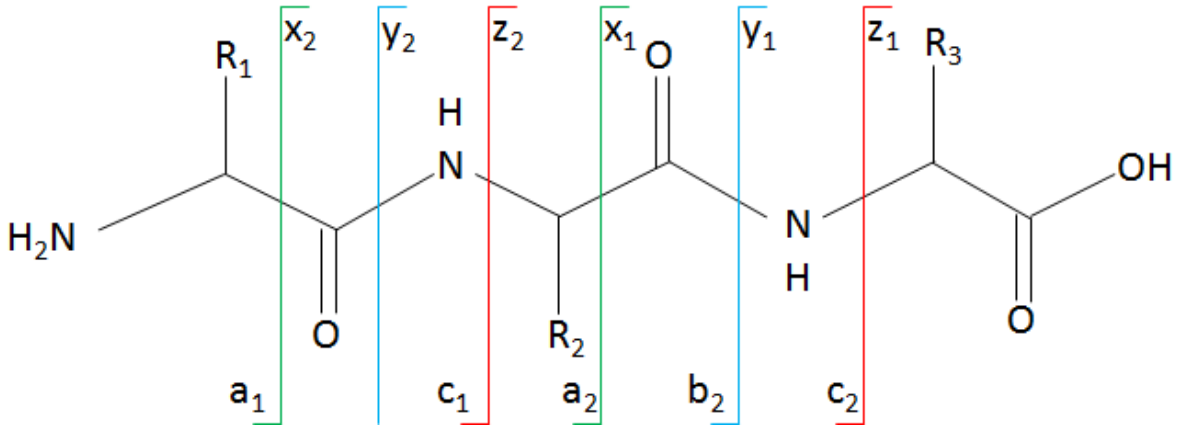


Figure 1.9 Nomenclature of peptide fragmentation and possible cleavage points [83]

CID cleaves peptides at the weakest bonds, which are usually the peptide bond between the carboxyl group of one amino acid residue, and the amino group of another [85]. These fragment ions are termed either b or y ions. In a y ion the charge remains on the

fragment which contains the C terminus of the peptide. The opposite is known as a b ion, in which the charge is retained within the N terminal fragment [85]. A population of ions will undergo fragmentation at various different points along the peptide, leaving a series of b and y fragments containing different number of residues. By finding the difference in mass between the peaks in a fragmentation spectrum, it is possible to deduce which residues were lost from each fragment and determine the sequence of the peptide. Each ion is given a number which counts how many amino acid residues from either the C or N Terminus the fragment contains. For example a C terminal fragment retaining 3 amino acid residues will be referred to as a y_3 ion and a fragment with the N terminus and 3 residues will be referred to as a b_3 ion [85].

1.2.4.2 Selected reaction monitoring assays

Selective reaction monitoring assays (SRM) are a specialised tandem mass spectrometry experiment available on triple quadrupole mass spectrometers. They are highly sensitive scanning methods used for targeted quantification [86]. As they are targeted approaches, the precursor mass of the analyte and its subsequent fragmentation pattern must be understood prior to analysis. To ensure maximum specificity the first quadrupole is tuned to only transmit the precursor m/z of the analyte, which is then fragmented in the second quadrupole. The third quadrupole is then tuned to only transmit a specific fragment or a few fragments of the analyte. The principle of double m/z scanning eliminates unwanted background or interfering ions from the spectrum which means the instrument is highly sensitive to the analyte of interest [86]. SRMs are now commonplace

techniques which are used successfully for a range of analytes from small molecules to peptides and proteins [87]. A diagram illustrating the process can be seen in figure 1.10.

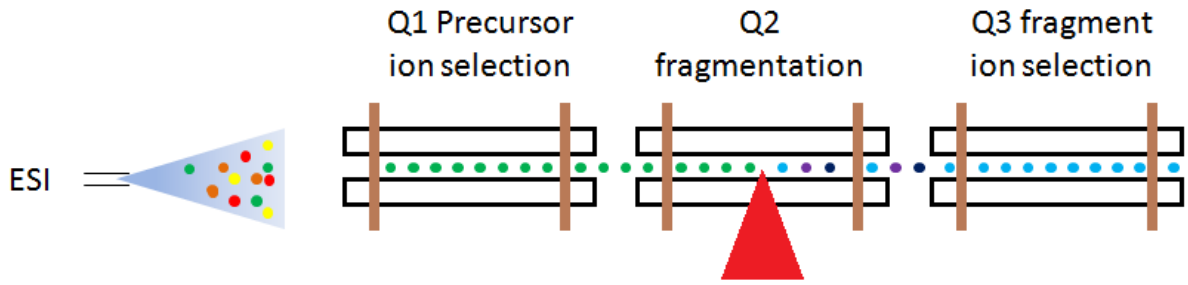


Figure 1.10 illustration of an SRM scanning method

1.2.5 Native mass spectrometry

Since the early 1990s mass spectrometry has been used as a tool for the analysis of non-covalent interactions. Analysing non-covalent interactions is possible because ESI is soft enough to retain non-covalent interactions after ionisation and this area of research has been dubbed native mass spectrometry. The concept was first demonstrated in 1991 with observation of receptor-ligand, enzyme-substrate and protein-haem interactions [88-90]. Since then the field has expanded and many different non-covalent complexes have been analysed with mass spectrometry, such as protein cofactor, protein nucleic acid and large multi-protein complexes such as intact virus capsids [91, 92]. Native mass spectrometry is becoming its own distinct field of research and is beginning to become a useful tool in structural biology [93]. Native mass spectrometry can be used to provide data on the stoichiometry of a protein and interacting partners in a complex. Mass

spectrometry has the added advantage that it can be done with much smaller amounts of material and in a much shorter period of time [93]. Native mass spectrometry is often combined with ion mobility analysis, which is very powerful tool that can be used to calculate the collision cross sectional area of an ion [94]. The collisional cross section of an ion can be used with molecular modelling techniques to calculate the proteins shape, and measure the rate of unfolding in a native protein in the gas phase [94, 95].

Retaining non-covalent interactions after ionisation can be very challenging. The complex must be electrosprayed from native conditions, as close to the physiological environment as possible [96]. In a typical intact protein mass spectrometry experiment, the protein is infused in an aqueous organic mixture at a low pH. These conditions strongly denature the protein and expose all residues to the solvent. Basic residues are protonated, which completely disrupts any native interactions [96]. In a native mass spectrometry experiment the protein is electrosprayed from an aqueous solution composed only of volatile pH buffers such as ammonium acetate. In these conditions the pH is near neutral, and the protein will not be denatured. Electrospraying a sample at neutral pH retains non-covalent interactions, but only mildly protonates the protein [97]. Proton exchange can only take place between highly basic residues such as lysine at the surface of the protein. In native conditions the complex is ionised very gently and will have a very low number of charges [97]. The m/z ratio of a particular protein in the native state may be double that of a typical denaturing experiment which can be challenging given the operating ranges of some mass analysers. Many mass spectrometers such as orbitraps have an upper m/z limit of 4000 [98]. These instruments offer high resolution but the m/z ratios of native complexes are usually above this. TOF mass analysers have become the

instruments of choice for analysing native complexes. TOF mass analysers have very high upper mass limits and are now available with high resolution. The resolution of a TOF mass analyser can now be as high as 50,000 full width half maximum (FWHM) [99] which makes them ideal for analysing large ions with high m/z ratios [100].

1.3 Proteomics

1.3.1 Overview of proteomics

Over the last 20 years improvements in mass analysers and the development of soft ionisation techniques have changed the focus of mass spectrometry. Early applications of mass spectrometry were more focused towards small organic compounds and petrochemicals, nowadays it is also a biological tool [5, 101]. The focus on biological analytes has given birth to a number of new fields including metabolomics, lipidomics and proteomics [102]. Proteomics is the biggest field in mass spectrometry and has become particularly significant since the human genome was sequenced in 2001 [103]. The aims are to characterise all the proteins expressed by a particular cell at specific times and conditions. The term proteome was originally defined to describe all the proteins coded for by the genome and is a challenging area of research. The genome can only act as a library to tell us which genes can be expressed into proteins and cannot provide any information on how cellular function changes at any specific point [104]. The number of proteins that can be expressed by an organism is far greater than the number of genes it contains. Messenger RNA transcripts may be alternatively spliced, or modified to give many distinct products from a single gene [105]. A cell will not produce all these products

at once; the proteome is dynamic [106]. An organism has multiple proteomes, different cells and tissues types have their own distinct proteomes and these change with disease, environmental conditions, available nutrition and the cycle of cell growth [107]. The field requires a completely different method of analysis to genomics. Proteins cannot be amplified in the same way the DNA can by using the polymerase chain reaction [108]. Proteomics techniques must be highly sensitive, highly specific and be able to cope with a large dynamic range of protein concentrations. They must be amenable to multiplexing, quantification and be able to identify sites of modification. Mass spectrometry fulfils all of these roles and has become the technique of choice in most proteomics research [101]. Without mass spectrometry analysis the field would not exist to the extent it does today.

Most proteomics experiments require the analysis of complex mixtures of proteins, perhaps from a cell lysate or biofluid. These samples can contain thousands of different proteins [109]. Even with high mass accuracy and automated database searching, it becomes important to separate these mixtures of proteins and peptides prior to mass spectrometry analysis [109]. For bottom up proteomics, there are two options for protein separation, gel based approaches which separate out intact proteins prior to digestion and solution based liquid chromatography techniques which can separate mixtures of proteolytic peptides [110].

1.3.2 Gel electrophoresis

Polyacrylamide gel electrophoresis (PAGE) is a powerful technique for resolving proteins and can be carried out in either one or two dimensions. One dimensional PAGE uses the

surfactant sodium dodecyl sulphate to separate proteins on the basis of their molecular weight [109]. It is known as SDS PAGE, and is well-established having been separating proteins since the late 1960s [111, 112]. If PAGE is carried out in 2 dimensions the proteins are separated out on the basis of their isoelectric point in the first dimension, and then by their molecular weights after the addition of sodium dodecyl sulphate (SDS) in the second dimension [113]. Both techniques separate proteins in a spatial manner on a gel. Two dimensional PAGE separates proteins more successfully than one dimensional gel PAGE as it separates proteins out on the basis of two different criteria. PAGE separates a complex mixture of intact proteins into individual spots or bands. The proteins can be visualised on a gel by staining with Coomassie blue [113]. Once resolved into individual proteins, bands or spots of interest can be cut out for mass spectrometry analysis. The proteins are destained to remove the Coomassie blue and then a proteolytic digestion is carried out in the gel. Proteolytic peptides can then be extracted, desalted and analysed via either ESI or MALDI mass spectrometry [114]. The approach has been commonly used when combined with MALDI peptide mass finger printing (PMF) [115].

One or 2D gels coupled with mass spectrometry is a useful and low cost means of identifying a particular protein of interest from a complex mixture. The downsides are that the technology is limited to individual proteins and impractical for the characterisations of complex mixtures [116]. Gel electrophoresis is a low throughput technique and in the past decade or so shotgun proteomics techniques have become

more popular [116]. Shotgun proteomics relies on performing an online liquid chromatography tandem mass spectrometry experiment (LC MS/MS).

1.3.3 Liquid chromatography

ESI ionises analyte directly from the solution phase so as a result it couples to liquid chromatography (LC) very effectively [117]. LC or high performance liquid chromatography (HPLC) is a widely used laboratory tool for separating mixtures into their individual components [118]. There are many different methods available but proteomics research typically use reversed phase C_{18} or reversed phase C_4 columns. A sample is pumped through a column packed with beads of silica with octadecyl hydrocarbons in C_{18}, or butyl groups in C_4 chromatography, under a high pressure. The beads are immobilised within a capillary and are referred to as the stationary phase [119]. A solution known as the mobile phase is pumped through the column. In reversed phase chromatography the mobile phase is typically composed of an aqueous organic mixture such as water and acetonitrile [120]. Peptides and proteins bind to the beads of the stationary phase. The mobile phase changes composition from a highly aqueous solution to become more organic [121]. The composition gradient separates out peptides and proteins on the basis of their hydrophobicity. More hydrophilic species will elute from the column earlier in more aqueous solutions than more hydrophobic species, which separates the analytes in a temporal fashion [121]. In an online LC MS/MS experiment the eluted species are directly ionised and introduced into a mass spectrometer [122]. LC

MS/MS with shotgun proteomics are commonly used techniques. In these experiments a complex mixture of intact proteins is digested into peptides, yielding a mixture of many different peptides before LC MS/MS analysis [122]. An individual peptide may elute from the column in a time window of around 20-30 seconds and the mass spectrometer will be constantly scanning to detect peptides as they elute from the column [123]. Data dependent acquisition is a commonly used mass spectrometry method. The instrument performs a survey scan and individual peaks detected in the survey scan are selected for subsequent fragmentation [123]. Shotgun proteomics experiments are useful for characterising complex mixtures of proteins. As a result they have become powerful tools for the analysis of biofluids and discovering biomarkers of disease [124]. The downsides are that LC MS/MS usually requires long run times of over an hour or more which places great demands on an instrument's time. Liquid chromatography alone may not be sufficient to fully separate highly complex mixtures such as plasma, salvia and cerebrospinal fluid. Multidimensional chromatography experiments are available for further separation but this can still be insufficient and separation remains as a limiting factor for complex samples [125].

1.3.4 Bottom up proteomics

Mass spectrometry based proteomics experiments are usually based on analysing small peptides formed from a proteolytic digests of larger proteins rather than analysing the proteins themselves. The analysis of proteolytic peptides is known as bottom up proteomics[126]. Analysing intact proteins is far more challenging as most mass analysers have insufficient resolving power. Intact protein analysis is a distinct field in its own right known as top down proteomics. Proteolytic peptides are smaller, which enables them to

be investigated with low resolution instruments and they are often as unique as the proteins they are digested from [127]. A range of proteases are available that cleave the peptide backbone of intact proteins at specific amino acid residues [128]. Trypsin is the most commonly used protease in proteomics research and cleaves proteins at lysine and arginine residues. Others are available however such as Lys-C which cleaves C terminal to lysine residues [128, 129]. These proteases are ubiquitous, cleaving almost any protein at the same amino acid residues to leave a mixture of peptides that can be used for identification. Trypsin is popular because it has a high rate of activity and most peptides it produces are in a size range suitable for mass spectrometric analysis. Trypsin also has the advantage in that it cleaves lysines or argines, so all tryptic peptides contain one of these two basic amino acid residues which ionise very strongly under ESI [101]. Tryptic peptides are on average 8.4 amino acid residues long and up to 97% of all are between 7 and 35 residues [130].

The simplest type of bottom-up proteomics experiment is protein identification. Both ESI and MALDI mass spectrometry are highly effective techniques for peptide analysis. Mass spectrometry can be used to measure the precursor m/z ratio of the peptides, or if required fragmentation can be induced within the peptide to generate sequence data [126]. Either is sufficient to identify the peptide. Proteases cleave proteins at specific residues and yield a number of unique peptides. The unique mass of the peptide can be measured by mass spectrometry and used to identify the protein from a protein database. The procedure is known as PMF [78] PMF is sufficient to identify proteins but

in many cases it is more useful to generate sequence data which can be achieved by fragmentation techniques described in section 1.2.4.

After a PMF or a fragmentation spectrum has been obtained the protein must be identified. Identification can be very challenging and for this reason a number of search algorithms are available for database searching of mass spectrometry data [131]. Bioinformatics is now becoming incredibly important in proteomics research. It is crucial to have a detailed protein database to efficiently identify proteins. Two of the most common algorithms, both of which were used in this thesis, are Mascot and Sequest. Sequest was first described in 1994 [132] and Mascot was released a few years later in 1999 [133]. The two work in slightly different ways. Sequest searches a protein database for peptides which match the precursor mass of the observed peptide. Sequest then generates a virtual fragmentation spectrum for all peptides that match the precursor mass and generates a score for how closely the virtual and observed spectra match. The peptide with the closest match to the spectrum is then provided as the identity [134]. Mascot is a probability based search tool. The software matches the observed spectra to theoretical peptides, and then calculates the probability that these matches arose by chance. The data is then filtered out to a significance threshold, for example $P<0.05$ [133]. Since both algorithms were published it has become commonplace to search experimental data against a decoy database to estimate the false discovery rate of the peptide identifications. In many cases the decoy database is a copy of the original protein database with the same number of sequences, each with the same amino acids but with

randomised sequences. Data are searched against both the real and decoy databases. The experimental data will contain some peptides that are matched by chance, the number of which is approximately equal to the number of decoy peptides that match the experimental. Decoy databases cannot eliminate false positives; they can only give an estimation of how many peptide matches are false but the false discovery rate can be used to filter the data to a desired confidence level [135].

Mass spectrometry can be used for the quantification of proteins as well as identification. The most effective methods involve the use of isotopically labelled internal standards which behave the same chemically but have a different mass so can be distinguished by mass spectrometry. Cell culture based techniques are available in the form of stable isotope labelling of amino acids in cell culture (SILAC) experiments, and isobaric labelling methods such as isobaric tagging for relative and absolute quantification (ITRAQ) or tandem mass tags (TMT) tags are available if cell culture is not an option [136]. Both methods allow untargeted quantification of proteins but were not used in this thesis so shall not be discussed in any great detail. The most accurate and most sensitive quantitative proteomics experiment are multiple reaction monitoring assays (MRM) [136]. These are targeted approaches analogous to the selective reaction monitoring assays mentioned in section 1.2.4.2. MRMs and SRMs differ in that an SRM is used to detect a single ion whereas an MRM is the use of multiple SRMs to detect several different ions [87]. MRMs can be used to quantify a known protein from a mixture. If the sequence of the protein is known, then so too are the proteolytic peptides it yields after

digestion. A range of isotopically labelled standards are synthesised with the same sequence as the endogenous peptides, but are labelled with stable isotopes to give them a different mass. The two react the same chemically and ionise in the same fashion but can be distinguished as separate by the mass spectrometer. The internal standard is spiked into the sample at a known concentration. One SRM is used to measure the intensity of endogenous analyte and the other is used to measure the intensity of the internal standard [86]. A calibration curve is generated from a range of standard samples with known concentrations of the labelled peptides. The amount of endogenous analyte in an unknown sample can be measured from its position on the calibration curve [87]. These are very effective tools for the targeted quantification of known proteins. They can multiplexed, are highly sensitive and have become popular for quantifying biomarkers in complex biofluids such as plasma [137].

1.3.5 Intact protein analysis

The field of mass spectrometry proteomics is not limited to the analysis of proteolytic peptides. The analysis of intact proteins via mass spectrometry is a less commonplace technique known as top down proteomics [138]. Top down and bottom up proteomics generate different types of data and are used for different purposes. Rather than being a tool for the identification or quantification of proteins top down proteomics is more commonly used as a tool to characterise particular proteins of interest [127]. Bottom up proteomics experiment only identify proteins on the basis of a few short peptide sequences. The majority of information about the protein is lost, including information on

the molecular weight of the protein and the presence of any modifications [139]. Many proteins are covalently modified in various different biological processes and this cannot be predicted from the genome. Phosphorylation of proteins is important in many signalling events, glycosylation is often important in cell-cell recognition and methylation of proteins can be important in regulating gene expression [140]. Post translational modifications can be identified by bottom up approaches, but because the approach only involves identifying the protein from a few peptides many areas of the protein are not covered. The post translational modifications identified by a bottom up approach are never complete [141]. In a top down approach the quasimolecular ion can be used to provide the molecular weight of the intact protein including any modifications [126]. This can then be fragmented to provide sequence data in the same way as peptide fragmentation. Fragmentation is unlikely to take place at every residue along the protein backbone, but the approach does not miss large regions in the same way bottom up proteomics approaches can [142]. Top down protein analysis provides more extensive identifications of post translational modifications than data provided by bottom up proteomics techniques. The confidence in protein identifications are also higher than by bottom up approaches, due to the fact that molecular weight information is obtained and the sequence coverage of the protein is greater [141].

There are a number of factors preventing top down proteomics from becoming adopted as a widespread technique. These include sample preparation and fractionation, mass spectrometry resolving power and tools for data analysis [139]. Fractionating intact proteins in a format that makes them amenable for mass spectrometry analysis can be very difficult. The gel electrophoresis available for bottom up proteomics do not work for

top down proteomics. Proteins can be highly resolved but it is almost impossible to extract large proteins from a polyacrylamide gel [138]. Reversed phase chromatography can be performed on intact proteins, but does not work as well as it does for proteolytic peptides. Intact protein chromatography has a number of drawbacks, intact proteins can become irreversibly adsorbed to the beads on a column; which can lead to carry over between different samples and the chromatographic resolution is often poor [139, 143]. Top down proteomics is also very demanding in terms of instrumentation, particularly in mass resolving power. Intact proteins are much larger than tryptic peptides, have a wider distribution of charge states and a more complicated isotope spacing pattern. The field has been dominated by highly expensive Fourier transform ion cyclotron resonance instruments. The development of the orbitrap mass spectrometer has provided an alternative, but it is most applicable for analysing small to medium size proteins [144]. Even high resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometers struggle with very large proteins and have an upper limit to size of proteins they can resolve, which are currently around 170 kDa [127]. The bioinformatics tools are more limited than those available for bottom up proteomics [145]. There is only one commercially available algorithm available for identifying top down spectra, Prosight. The algorithm considers the experimentally measured molecular mass of the protein to proteins that match it within a specified range. It then considers the fragmentation spectrum expected from each of the shortlisted proteins and chooses the closest match to the experimental spectrum as the identity [146]. Although fewer algorithms are available for identifying intact proteins, the matches provided by top down techniques tend to be more confident than identifying the protein from bottom up data [141].

1.4 Dried blood spot analysis

One of the most popular applications investigated by groups working with LESA and ambient ionisation mass spectrometry has been dried blood spot analysis. DBS are a routine blood sampling format used in the clinic. They were initially developed as a means of collecting blood samples from newborn babies, but have since been adopted for other applications [147]. The paediatrician Robert Guthrie was the first to use blood soaked dried paper discs as a sampling tool in 1963 whilst he was working on the disease phenylketonuria (PKU) [148]. PKU treatment required a means of diagnosing the disease based on testing blood in the first few days of life and required a new sampling method that was safe for use on newborn babies. Guthrie's solution was to acquire the blood from a heel prick and apply it directly to small filter paper disc. An image of a DBS is shown in figure 1.11 The blood could then be dried for storage and later analysed by applying it to bacterial culture plate that identified the disease through a growth inhibition assay [148]. The assay identified the disease before symptoms emerge and soon became adopted as the basis for many different newborn screening programmes [149]. Every child born in a particular population had a DBS taken a few days after birth and was tested for PKU.

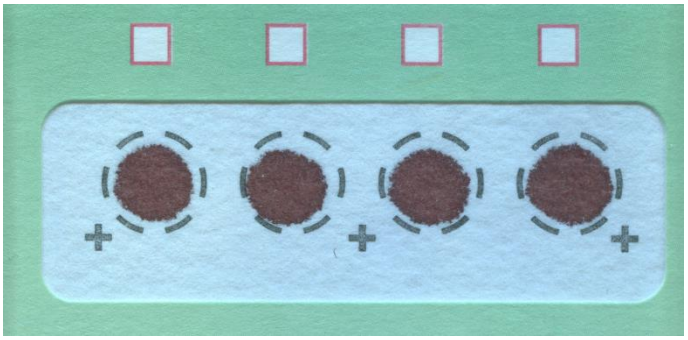


Figure 1.11 A filter paper DBS

1.4.1 DBS as a sampling format

DBS have become popular in the clinic because they are far easier to work with than whole blood. They are easier to acquire, easier to store and ship, and are safer for the laboratory analyst [147]. Whole blood for plasma assays needs to be acquired by a specially trained phlebotomist and large volumes need to be acquired by a fairly invasive procedure [150]. Many young children and adults alike find phlebotomy fairly traumatic and painful. It must also be done in a clinical setting with anticoagulant and centrifugation being required to separate the blood into plasma [151]. Acquiring a DBS, however, is much simpler. The blood can be acquired via a finger prick for adults or a heel prick for newborns [152]. The person taking the blood requires very little training. It can even be taken by the patient themselves and does not need to be done in a clinical setting [152]. DBS do not require any special storage conditions. Once the blood is applied to the filter paper it is dried in air and they can then be stored at ambient conditions or at -20˚C [153]. In comparison whole blood must be separated into plasma and transported on dry ice [151]. DBS are stable at ambient conditions and can even be transported through standard postal services [153]. Using dried blood also helps reduce

the dangers of biohazards for the user. Many different blood borne viruses including HIV and hepatitis C are inactivated upon drying [149].

1.4.2 Analysis of DBS by mass spectrometry

Newborn DBS screening expanded after the 1960s to include a range of other inborn errors of metabolism such as medium chain acyl coenzyme dehydrogenase deficiency (MCAD), and maple syrup urine disease (MSUD) [154, 155]. In the 1990s mass spectrometry was introduced as a means of analysis. Diagnosis of inborn errors of metabolism required the accurate quantification of a few small amino acids and organic acids [156]. The early 1990s had seen the introduction of ESI and the field of metabolomics which was a field aimed to characterise small molecules produced by biological organisms [157]. ESI combined with triple quadrupole MRMs replaced bacterial inhibition assays as a means of analysing small molecule metabolites and diagnosing inborn errors of metabolism. The assay is similar to that of targeted MRMs for quantitative proteomics mentioned in section 1.3.4. A small disc (a few millimetres in diameter) was cut out of a DBS and the sample eluted in a solvent [158]. SRMs were used to measure the intensity of the endogenous metabolite relative to an isotopically labelled derivative of the metabolite used as an internal standard. The ratio between the two was then used to calculate the absolute concentration of the endogenous metabolite [158]. These tests were higher throughput than existing technologies, could be easily automated and most importantly could be multiplexed. A single LC MRM assay could be

used for the quantification of several different metabolites and diagnose several different diseases in a single test [159].

1.4.3 Protein analysis of DBS by mass spectrometry

The potential for using mass spectrometry for analysing metabolites from DBS was recognised immediately and has been adopted worldwide the method of choice for screening inborn errors of metabolism across the globe [160]. The idea has also been used by the pharmaceutical industry. DBS are now often used as an alternative to plasma based screening assays in drug metabolism and pharmacokinetic studies [161]. DBS have largely remained as a means of small molecule analysis only and very few attempts have been made to utilise mass spectrometry to analyse the protein content of DBS [162]. It is a particular surprise that the idea has not been explored further, considering the growth of proteomics in this time and the focus on plasma as a source for disease biomarkers [163].

1.4.3.1 Targeted analysis of proteins from DBS

In the few examples where groups have investigated DBS for protein analysis the majority of the work has focused on haemoglobin. Haemoglobin is of interest because it is high abundance protein and is useful for clinical diagnosis. Newborn DBS programmes screen haemoglobin for pathologic variants such as sickle cell disease and thalassemias which are a range of disorders caused by a reduced synthesis of haemoglobin chains [164]. These are currently screened for by techniques such as isoelectric focusing and HPLC

[165]. A triple quadrupole mass spectrometry method to screen different haemoglobin variants was developed by Daniel *et al.* in 2005, [166] and later the method was applied to DBS for providing relative quantification of haemoglobin A2 [167]. In this method a small punch was cut out of a DBS and eluted in water under shaking at 37˚C. A small aliquot was taken, treated to a short trypsin digestion and analysed with a LC MRM approach which was composed of SRMs to detect for the presence of common healthy haemoglobin peptides and pathogenic variants [166]. The method was very quick and more specific than conventional haemoglobin screening methods, but as only the most common variants were added to the MRMs, uncommon variants would be missed. The method was later refined by Boemer *et al.* [168] who successfully developed this further and used for the mainstream haemoglobin variant screening programme in Belgium[169]. Boemer's method worked well on this screening programme, proving to be reliable, robust and cost effective [169].

The opportunities in a proteomic analysis of DBS go far beyond that of just haemoglobin analysis. Blood is a very rich biofluid composed of thousands of different proteins [170]. Prior to the work in this thesis a few groups had investigated targeted MRM approaches for the quantification of specific proteins of interest. The first example was in 2008 by De Wilde *et al.* who used MRMs for the quantification of ceruloplasmin [171]. Ceruloplasmin is a protein responsible for transporting copper in the blood which can be affected in Wilson’s disease. Wilson’s disease is caused by insufficient synthesis of mature

ceruloplasmin and a reduction in the amount in the plasma [172]. De Wilde's method aimed to quantify the amount of ceruloplasmin from a DBS. Similar in principle to that of Daniels *et al*. [167], a DBS punch was eluted in solvent and treated to a trypsin digest before internal standard peptides were added. The digests were then analysed by a ultra high performance liquid chromatography (UHPLC) MRM assay to detect both endogenous and internal standard peptides[171]. Similar approaches have been carried out for investigations in pharmaceutical drug development studies. Kehler *et al*. in 2011 sampled DBS for quantification of the exendin 1, a 4 kDa peptide, from monkey plasma [173]. Sleczka *et al.* [174] investigated the use of a multiple reaction monitoring assay to monitor the degradation of mice dosed with therapeutic proteins [174]. In 2013, Cox *et al.* demonstrated a different application: detection of sports enhancing drugs in sporting events. Insulin like growth factor (IGF) is a common doping agent [175]. IGF needs to be tested for regularly, even during periods off training and the need to give regular whole blood samples is very demanding for the athletes. The authors developed a DBS MRM approach as an alternative [175].

One of the great advantages of metabolic DBS analysis is that MRMs can be multiplexed to detect for the presence of multiple different analytes [159]. Multiplexed MRMs were demonstrated on proteins in 2012 in proof-of-principle experiments by Chambers *et al.* [176]. As seen previously, proteins were eluted from a DBS and digested into peptides. In this case the MRM method was used to measure the concentration of 65 peptides from 60 different proteins. The instrument response showed a linear relationship with

endogenous concentration for 40 of the 65 targeted peptides, but many of the concentrations measured by the assay did not agree with the established reference concentrations [176].

1.4.3.2 Untargeted analysis of proteins

In 2011 LESA was used for sampling of DBS to analyse intact haemoglobin proteins. The method could successfully distinguish different haemoglobin variants [2]. The advantages of using an untargeted approach for the analysis of haemoglobin variants was a major advantage as it could be used to characterise less well understood and even unknown variants [177]. LESA sampling meant that the DBS could be analysed directly without any sample pre-treatment or preparation. The only downside was that because the technique was based on analysing intact proteins, it was demanding in terms of resolution and required a high performance orbitrap mass spectrometer for analysis [2].

1.4.3.3 Future prospects for DBS proteomics

Research into DBS proteomics is still in its early stages, but there has been a lot of effort to characterise the human plasma proteome to identify new biomarkers of disease [163]. Plasma is a highly dynamic and complicated biofluid. It is highly affected by disease and many clinical assays are based on measuring the amount of a particular protein in the plasma[178]. DBS could be used as an alternative sampling format to plasma, but would be more convenient in the clinic. DBS analyses face the same challenges as analysis of plasma. The number of proteins in plasma and blood is enormous. Over 10,000 different

distinct proteins are known to have been identified in plasma so far [179] and some speculate that it could contain the entire human proteome at some level [180]. Characterising plasma is made particularly difficult by the fact that the concentrations of these proteins have a high dynamic range [180]. There are thousands of low abundance proteins and a few high abundance proteins that make up most of the total protein in plasma. Of the total protein in plasma, 99% is made up of 22 high abundance proteins [179]. The presence of high abundance proteins makes it more difficult to analyse the proteins that make up the remaining 1% of total protein such as interleukins, cytokines and receptor ligands. Low abundance proteins can be more useful diagnostically and can be expressed at the pg/ml level. The difference in concentration between the most and least abundant proteins in plasma can be more than 10 orders of magnitude [170]. Generating an assay that performs accurately across this range is an incredible challenge [181] and analysing the full proteome of a DBS is likely to be even more difficult than analysing the plasma proteome.

1.5 Alpha-1-antitrypsin deficiency

The protease inhibitor A1AT was studied in chapter 6 in the investigation of LESA based methods for the diagnosis of the disease alpha-1- anti trypsin deficiency. A1AT is a 52 kDa glycoprotein which is produced in the liver where it is secreted into the blood plasma. In a healthy human 34 mg for every kg of body mass is made per day. A1AT is one of the more abundant plasma proteins at around 20-48 μM in a healthy adult [182] and has a circulating half-life of around six days [183]. It belongs to a family of protease inhibitors

known as the serpins, an abbreviation of serine protease inhibitors. The primary biological function of the protein is to act as an inhibitor of defensive proteases secreted by phagocytes during an immune response. The activity of A1AT is highest against neutrophil elastase and acts mostly in the lungs. The lungs receive a constant immune challenge due to continual inhalation of airborne particulates and pathogens [184]. The immune system responds by recruiting phagocytic cells such as neutrophils into lung tissues [185]. Proteases such as elastase are released during phagocytosis and can degrade the lung tissues. If left unchecked they can damage alveolar structures [184]. A1AT prevents prolonged proteolytic activity by denaturing elastase; elastase binds to A1AT at its active site and attempts to cleave the protein. The A1AT protein however remains irreversibly bound to the protease and denatures its tertiary structure. The inhibition process is suicidal for both proteins but prevents unwanted and prolonged proteolysis of self-tissues [186]. The A1AT gene is polymorphic. Several natural variants exist to this protein which result in a reduction of synthesis of the protein that can cause a pathogenic condition known as alpha-1-anti trypsin deficiency [187]. There are 2 recognised healthy variants of the protein known as the M variant. One of these variants has a valine at the 213 position and the other codes for an alanine [188]. These genotypes will result in a normal production and excretion of A1AT so plasma levels will be normal between 20 and 50 μM. There are two well understood clinically significant mutants which each possess a single nucleotide polymorphism that cause in a decrease in the synthesis of the protein. The first is known as S, which is a mutation of the M valine (M val) 213 allele with a glutamic acid to valine transition at amino acid 264. The second is known as the Z variant and has the most serious implications. It is modification of the

M alanine 213 (M ala) allele with has a glutamic acid to lysine transition at the 342 position [188]. The possession of either of the two alleles results in a decrease in serum concentration of A1AT, but the amount can vary. A homozygous MS individual has serum A1AT concentration between 20 and 48 μM, which is close to the amount expected in a healthy individual. A typical MZ genotype has a concentration between 12 and 35 μM. The concentration of A1AT in an SS genotype is normally between 15 and 33 μM whereas SZ and ZZ variants are the most serious. An SZ genotype normally results in a concentration of A1AT around 8-19 μM and the concentration of a ZZ individual can be as low as 2.5-7 μM [189]. Lower serum concentrations result in an increased likelihood of clinical problems, but is possible to have a diseased genotype and not suffer any symptoms [184]. If left untreated A1AT deficiency can lead to a prolonged attack of lung tissues by neutrophil elastase resulting in emphysema and chronic obstructive pulmonary disease [190]. Damage can also be done in the liver; A1AT deficiency can be followed by cirrhosis or jaundice in children. These problems are caused by the synthesis of the protein [189]. Certain variants such as the Pi ZZ variant result in a change to the 3D structure of the protein which can aggregate to form a series of A1AT polymers. These cannot be secreted by a liver cell and begin to accumulate in the endoplasmic reticulum. Accumulation of A1AT polymers can be toxic to the cell which induces apoptosis [189]. A1AT deficiency is not a rare disease and frequently goes undiagnosed [191]. It is not currently screened for and is usually diagnosed following the presentation of symptoms. The means of diagnosis differ from laboratory to laboratory, but rely on multi-step processes. Many laboratories use DBS for sampling [192].

1.6 Aims and objectives

The overall aim of the work presented in this thesis was to develop new methods using LESA and mass spectrometry for protein analysis and to combine LESA with previously existing proteomics techniques.

- The aim of the work presented in Chapter 3 was to analyse intact proteins from PVDF membranes using LESA. Analysis of PVDF had applications in investigating proteins electroblotted from SDS PAGE as a technique for intact protein fractionation and as a means for detecting intact proteins from air filters.
- The work presented in Chapter 4 was aimed at developing the use of LESA for the analysis of proteolytic peptides from DBS as a means of untargeted protein identification (proteomics). The procedure behind LESA was adapted to incorporate a trypsin digestion into the sampling process.
- The work presented in Chapter 5 further investigated the potentialsof DBS proteomics using solution phase chemistry rather than LESA and the use of protein depletion to increase the number of proteins that could be identified from DBS.
- The work presented in Chapter 6 focused on targeted analysis of proteins from DBS by LESA mass spectrometry to determine alpha-1-antitrypsin variants.
- The aim of the work presented in Chapter 7 was to investigate the potential of using LESA to retain native protein interactions from dried surfaces for mass spectrometry analysis and analyse the native haemoglobin tetramer directly from DBS.

Chapter 2 Materials and Methods

2.1 Materials

Reagents

All solvents used for mass spectrometry, i.e., water, acetonitrile and methanol were HPLC grade obtained from J. T. Baker, (Deventer, The Netherlands). Formic acid (HPLC grade), ammonium bicarbonate, dithiothretiol and iodoacetamide were obtained from Sigma Aldrich (Gillingham, UK). Purified lypholised proteins, bovine cytochrome C, bovine alpha casein, equine myoglobin, yeast alcohol dehydrogenase and human hemoglobins were also obtained from Sigma Aldrich (Gillingham, UK). Pre cast 4-12% gradient SDS PAGE and precast 2-8% native PAGE gels, 4X LDS sample buffer and 20X SDS running buffer were obtained from Expedeon (Cambridge, UK). Tris(hydroxylmethyl)aminomethane hydrochloride (tris HCl) , tricine, glycine, SDS, reagent grade methanol, Coomassie Plus Bradford assay reagent and plain glass microscope slides were obtained from Fisher Scientific (Loughborough, UK). Immobilon P PVDF membranes, reversed phase C_{18} and reversed phase C_4 zip tips were obtained from Millipore (Watford, UK). Mass spectrometry grade trypsin gold was obtained from Promega (Southampton, UK) Blood spot cards were standard NHS blood spot (Guthrie) cards, made from Ahlstrom grade 226 filter paper (ID Biological Systems Greenville, SC, USA). Customised synthetic peptides were made to a purity >70% and purchased from GenicBio (Shanghai, China). Healthy adult DBS were acquired from anonymous donors via finger prick. The work was

approved by the University of Birmingham STEM ethical review committee ERN 12-0782A. The blood was applied to standard NHS DBS cards made of Ahlstrom grade 226 filter paper (ID Biological Systems, Greenville,SC, USA). After blood was spotted onto the card, DBS were dried overnight prior to analysis. DBS from previously screened SZ and ZZ genotype alpha-1-antitrypsin deficiency patients were provided by Prof R Stockley, Queen Elizabeth Hospital, Lung Function and Sleep department, which was approved by STEM ethical review committee LREC 3359.

2.2 Methods

2.2.1 LESA

LESA was carried out using the Triversa Nanomate™ nanoelectrospray platform (Advion Biosciences Ithaca NY, USA). The analytical surface was loaded onto a universal LESA adapter plate (Advion Biosciences Ithaca NY, USA) and scanned on an Epson perfection V300 photo scanner. The LESA points software was used to plot the exact position of the dried sample spot and perform sampling at this precise location. Surface sampling was carried out as described in each individual experiment.

2.2.2 LC MS/MS

Three liquid chromatography methods were used in the work presented in this thesis. Two were used for online analysis with the Orbitrap Velos™ and the third method was used online with analysis on the TSQ. In all cases, the HPLC unit was coupled to the mass spectrometer via a Triversa Nanomate™, infusing the samples in positive mode with an ionisation voltage of 1.7 kV.

Method one was used online with the Orbitrap Velos™ and was composed of a 30 minute 3.2-44% acetonitrile gradient. A Dionex Ultimate 3000 was used with an analytical column and a trapping column on a 6 port valve. Samples were pumped through a Pepmap 100 reversed phase C18 analytical column at a flow rate of 350 nanolitres per minute. The trapping column used was a nanoviper pepmap 100, C18 75 µm, samples were loaded onto it at a flow rate of 4 µl/minute in 0.1% formic acid. After loading, the valves switched and the analytical gradient was applied starting at 3.2% acetonitrile, which increased to 44% over a 30 minute period. The column was then washed by increasing the acetonitrile gradient to 90% for 10 minutes before being lowered to 3.2%. After a further 10 minutes the valves switched back to load position. The total run time of the method was 56 minutes.

Method two was similar to method one but included an online desalting stage to eliminate the need for clean-up with C_{18} zip tips before analysis. Samples were loaded onto a Pepmap 100, C18 100 µm nanoviper trapping column at a flow rate of 20 µl/minute in 0.1% formic acid. Before the valves switched the trap was washed for 5 minutes in 0.1% formic acid to eliminate salts and contaminants before being switched online with the analytical column. The analytical column and the analytical gradient was the same as the one used in the first method. The gradient started at 3.2% acetonitrile and was increased to 44% over a 30 minute period. The column was washed in 90% acetonitrile for 10 minutes before being reduced back to 3.2%. After a further 10 minutes

the valves switched back to load position and remained this way until the end of the run. The total run time of the method was 61 minutes, 5 minutes longer than method one. Eluents from both methods one and two were infused into the mass spectrometer and analysed by a Top 7 CID data dependent acquisition described in section 2.2.3.2.

Method three was used in conjunction with a triple quadrupole MRM method and used a Dionex Ultimate 3000 UHPLC unit coupled to a 1:300 splitter. The analytical column was a Pepmap 100 reversed phase C18 column, used with a Pepmap 100, C18 100 µm nanoviper trapping column. Samples were loaded onto the trapping column at a flow rate of 10 µl/minute in 0.1% formic acid. The valve switched and the analytical gradient was applied starting at 3.2% acetonitrile, and increased to 44% over a 17 minute period at a flow rate of 0.35 µl/min. The column was washed with 90% acetonitrile (ACN) for 10 minutes, before being reduced to 3.2% ACN for 13 minutes. Eight minutes later the valves switched back to load position. The total time of the method was 40 minutes. The eluent was infused into the mass spectrometer and analysed by the SRM method described in section 2.2.3.3.

2.2.3 Mass spectrometry

Three mass spectrometers were used in this work. The majority of the work was performed on an Orbitrap Velos™ ETD mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The work presented in Chapter 6 required the use of a TSQ Vantage™ triple quadrupole mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) and

the work presented in Chapter 7 made use of a Synapt™ G2S (Waters Manchester, UK) quadrupole-time of flight mass spectrometer.

2.2.3.1. Direct infusion and LESA mass spectrometry

In direct infusion electrospray, and LESA mode the Orbtirap Velos was operated in positive ion mode unless otherwise stated. Samples were infused by the Triversa Nanomate™ with an ionisation voltage of 1.7 kV and a gas pressure of 0.3 PSI. Data were acquired in the orbitrap with a mass resolution of 100,000 at m/z 400. The maximum injection time was 1000 ms with an automatic gain control (AGC) target of $1x10^6$ charges. When data was acquired in full scan mode the number of microscans used was either 5, 30 or 50 as defined in each particular experiment.

A customised method was used for the detection of β-lactoglobulin in Chapter 3. The customised method was 7 minutes long and composed of 5 minutes full scan in the m/z region of 600-2000, followed by 1 minute isolation of the +10 charge state of the protein (m/z 1837.23) with an isolation window of 10 thomsons (Th), and one minute scanning in selected ion monitoring (SIM) mode of m/z 1825-1845 to monitor the same +10 charge state. Each scan in this method was made up of 5 microscans.

When using the Waters Synapt™ G2-S, samples were infused via the Triversa Nanomate™ with an ionisation voltage of 1.7 kV and a gas pressure of 1.0 PSI. The instrument was used for both direct infusion ESI and LESA analysis. The capillary temperature was set to 30°C and the cone voltage was 45 V. Samples were acquired in time of flight mode with scan time of 2 seconds.

2.2.3.2 Data dependent CID

A top 7 data dependent CID method was used in Chapters 4, 5 and 6 for analysis by direct infusion ESI and online LC MS. The method comprised 8 scan events beginning with a survey scan in the orbitrap cell followed by collision induced dissociation in the ion trap of the 7 most abundant peaks identified from the survey scan. The survey scan was acquired at a resolution of 60,000 at m/z 400 with an AGC of $1x10^6$ charges. Ion trap fragmentation was carried out with a normalised collision energy of 35% and an AGC target of 30,000 charges and a precursor isolation width of 2 Th. Dynamic exclusion was applied to prevent any peak with the same m/z from being selected for fragmentation within a period of 60 seconds. Only multiply charged peaks with an intensity greater than 5000 arbitrary units were selected for fragmentation.

2.2.3.3 Selective reaction monitoring (SRM)

SRMs were used to detect targeted peptides from samples analysed through direct infusion electrospray ionisation and LC MS method 3 described in section 2.2.2. The mass spectrometry method comprised 7 scan events: the first was a full scan in Q3 in the m/z range of 50-1500 with a scan time of 1 second. The following 6 scan events were SRMs to detect the tryptic peptides covering the 213, 264 and 342 residues of the M Valine 213, M Alanine 213, Z and S variants of the alpha-1-antitrypsin protein. Three transitions were used for monitoring each peptide, each with a scan time of 1 second. Table 2.1 shows

which peptides were targeted, the selected precursors, the transitions monitored and the collision energies of the fragmentation stage.

Peptide	Sequence	Selected Precursor m/z	Transitions			Collision energy (v)
MAla 213	**DTEEEDFHVDQATTVK**	[M+3H] 621.94	y_9 998.52	y_{10} 1145.59	y_{11} 1260.62	35
MVal 213	**DTEEEDFHVDQVTTVK**	[M+3H] 631.29	y_7 790.43	y_8 889.49	y_9 1026.55	35
M264	**LQHLENELTHDIITK**	[M+2H] 902.48	y_8 940.54	y_9 1069.58	y_{10} 1183.63	33
S264	**LQHLVNELTHDIITK**	[M+2H] 887.49	y_9 1069.58	y_{10} 1183.63	y_{11} 1282.70	33
M342	**AVLTIDEK**	[M+2H] 444.75	y_4 504.26	y_5 605.31	y_6 718.39	30
Z342	**AVLTIDK**	[M+2H]380.23	y_3 375.22	y_4 476.41	y_5 589.35	30

Table 2.1 SRM transitions of the 6 peptides required for determining variants of the alpha-1-antitrypsin protein

2.2.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE)

SDS PAGE was performed in chapter 3. Proteins solutions for SDS page were dissolved in a 50 mM pH 7.5 TrisHCl solution. Twenty microlitre aliquots were loaded onto the gel along with 5 µl of 4X LDS sample buffer. Prior to loading the sample was heated to 95˚C for 5 minutes and centrifuged at 13,000 rpm for 1 minute. Protein solutions were added with a gel loading pipette tip. Precast 4-12% gradient polyacrylamide gels were immersed in TrisHCl/tricine SDS running buffer with 60 mM TrisHCl, 80 mM tricine, 2.5mM sodium bisulphate and 0.1% SDS. Electrophoresis was run at 180 volts, 50 mA for 38 minutes. After electrophoresis, proteins were blotted onto an Immobilon P PVDF

membrane. The electroblot apparatus was immersed in a transfer buffer composed of 192 mM glycine, 25 mM TrisHCl and 10% methanol. Blotting was carried out at 100 volts, 40 mA and 25 W for 75 minutes.

2.2.5 Zip tip purification

Purification by using Zip tips was performed in chapters 3, 4 and 6. The procedure for sample clean-up with reversed phase Zip tips was identical for both C_4 and C_{18} tips. Tips were loaded onto a 10 µl pipette which was set at 10 µl. Pure acetonitrile was aspirated and dispensed to waste to prepare the tip. This was repeated once. The tips were rinsed through with a wash solution composed of 0.1% Trifluoroacetic acid (TFA) which was dispensed to waste. This process was repeated 3 times. Next the proteins or peptides were bound to the tips via 7 repeated aspiration and dispense cycles of sample solution. The tips were then rinsed by aspirating aliquots of 0.1% TFA and dispensing them to waste. This process was also repeated 7 times. After rinsing, the proteins or peptides were eluted. This was achieved by aspirating and dispensing 4µl of a 50/0.1% acetonitrile/TFA solution in a clean tube. This process was repeated 7 times. After elution, samples were dried down in a speedvac running at 60˚C for 30 minutes and resuspended in 10 µl 49.4/49.5/1% water/methanol/formic acid for analysis by direct infusion electrospray ionisation mass spectrometry or 0.1% formic acid for LC-MS/MS analysis.

2.2.6 Native PAGE

Native page was carried out in chapter 3. Proteins solutions for native page were dissolved in a 50 mM pH 7.5 Tris-HCl solution. Twenty microlitre aliquots were loaded onto the gel along with 5 µl of native PAGE sample buffer. The native PAGE sample buffer was composed of 40% glycerol, 4% ficoll 400, 2 mM EDTA, 0.8 M triethanolamine chloride pH 7.7 0.025% Commassie blue G250 and 0.025% phenol red.

phenol red. Protein solutions were loaded onto precast 2-8% acrylamide gradient gels with a gel loading pipette tip. Gels were immersed in trisHCl/tricine running buffer with 60 mM TrisHCl and 80 mM tricine. Electrophoresis was run at 150 V, 50 mA for 90 minutes. After electrophoresis, proteins were blotted onto an immobilon P PVDF membrane. The blotting apparatus was immersed in a transfer buffer composed of 192 mM glycine and 25 mM TrisHCl. Blotting was carried out at 100 V, 40 mA and 25 W for 75 minutes.

2.2.7 Solution phase digestions

Where stated that proteins from DBS were digested in solution (see Chapter 5), DBS were prepared by a "punch, elute and digest" procedure that is based on the method described by Sleczka *et al.* [174]. A 6 mm disc was cut out of a DBS and transferred to an eppendorf tube where the proteins were eluted in 150 µl 50 mM NH_4HCO_3 on a thermomixer shaking at 800 RPM and 40˚C. Protein concentration was determined via a Bradford assay. The Bradford assay was carried out by measuring the UV absorbance of Coomassie Plus reagent at 595 nm. A standard curve was generated from a series of

standards with concentrations of 1.0, 0.8, 0.6, 0.4, 0.2 and 0 mg/ml of bovine serum albumin (BSA). Ten microliters of each standard solution was mixed with 200 µl coomassie plus reagent and the UV absorbance at 595 nm was measured by spectrophotometry. Ten microliters of DBS eluent was also mixed with 200 µl Coomassie Plus reagent and the absorbance was measured on the standard curve to calculate the protein concentration. An aliquot corresponding to 20 µg of protein was taken from the DBS eluent for digestion. This aliquot was diluted tenfold in 50 mM NH_4HCO_3 and 1 µg of trypsin was added to give a 1:20 ratio of enzyme to protein. The sample was incubated on a thermo mixer shaking at 800 RPM and 40˚C. The incubation time varied from 30 minutes to 8 hours and is defined in each experiment. Samples were stored at -20˚C prior to analysis. Prior to injection onto the HPLC, samples were diluted by a factor of 200 in 0.1% formic acid.

2.2.8 Database searching

Database searching of data dependent CID was carried out in chapters 4 and 5. Bottom up proteomics data were searched against protein databases using the Mascot and Sequest algorithms in Proteome Discoverer (versions 1.3 and 1.4). The search parameters were as follows. Trypsin was selected as the protease and up to 2 missed cleavages were allowed in the digestion. The precursor ion mass accuracy tolerance was 10 ppm, the fragment ion mass accuracy tolerance was of 0.8 Da. Methionine oxidation was included as a variable modification and where disulphide bond reduction had been used carbamidomethylation of cysteines was included as a static modification. Data were searched against the SwissProt human database composed of 20,233 sequences (downloaded in December 2012). The false discovery rates of the peptides were

calculated with the percolator function of Proteome Discoverer and filtered to a false discovery rate of 1%.

2.2.9 Protein depletion

Depletion of the top 12 plasma proteins and haemoglobin was carried out on samples eluted from DBS in chapter 5. Samples were eluted as described previously in section 2.2.7**.** The sample was subsequently depleted of abundant proteins via four methods.

1. The first method depleted the 12 most abundant plasma proteins with the Proteome Purify 12 kit. The entire 150 µl DBS eluent was mixed with 1 ml Immunodepletion resin for 1 hour. The mixture was transferred to a spin column and centrifuged at 4600 rpm for 2 minutes. This eluted approximately 500 µl of depleted sample from the resin.
2. The second method depleted DBS of haemoglobin using Hemovoid: Fifty mg Hemovoid matrix was weighed out into a spin column, mixed with 250 µl Hemovoid binding buffer and vortexed for 5 minutes. The mixture was centrifuged at 3000 rpm for 2 minutes, the flow through was discarded and this procedure was then repeated. The 150 µl DBS eluent was added to the resin along with 300 µl binding buffer, this mixture was vortexed for 10 minutes and spun for 4 minutes at 10,000 rpm. The flow through was discarded. Five hundred microlitres of Hemovoid wash buffer was vortexed with the sample for 5 minutes and then centrifuged at 10,000 rpm for 4 minutes. The flowthrough was discarded and this was repeated twice before elution. Three hundred microlitres of elution

buffer was added to the sample, which was vortexed for 10 minutes and then centrifuged for 4 minutes at 10,000 rpm to elute the depleted proteins.

3. In the third option Hemovoid was used to deplete haemoglobin followed by the top 12 plasma proteins with proteome purify 12. Fifty milligrams of Hemovoid matrix was weighed out into a spin column, mixed with 250 µl Hemovoid binding buffer and vortexed for 5 minutes. The mixture was centrifuged at 3000 rpm for 2 minutes, the flowthrough was discarded and this procedure was then repeated. The 150 µl DBS eluent was added to the resin along with 300 µl binding buffer. This mixture was vortexed for 10 minutes and centrifuged for 4 minutes at 10,000 rpm. The flowthrough was discarded. Five hundred microlitres of Hemovoid wash buffer was vortexed with the sample for 5 minutes and then centrifuged at 10,000 rpm for 4 minutes. The flowthrough was discarded and this was repeated twice before elution. Three hundred microlitres of elution buffer was added to the sample, which was vortexed for 10 minutes and then centrifuged for 4 minutes at 10,000 rpm to elute the depleted proteins. After the proteins were depleted of haemoglobin, the eluted sample was then depleted of the top 12 plasma proteins. The sample was mixed with 1 ml proteome purify 12 immunodepletion resin for 1 hour. The mixture was transferred to a spin column and centrifuged at 4600 rpm for 2 minutes.
4. In the final method the Proteome purify 12 kit was used first, before being followed by depletion with Hemovoid. The DBS eluent was mixed with 1 ml Immunodepletion resin for 1 hour. The mixture was transferred to a spin column and centrifuged at 4600 rpm for 2 minutes. This eluted approximately 500 µl of

depleted sample from the resin. Following depletion of the top 12 plasma proteins the sample was then depleted of Haemoglobin. Fifty milligrams of Hemovoid matrix was weighed out into a spin column, mixed with 250 µl Hemovoid binding buffer and vortexed for 5 minutes. The mixture was centrifuged at 3000 rpm for 2 minutes, the flowthrough was discarded and this procedure was then repeated. The 150 µl DBS eluent was added to the resin along with 300 µl binding buffer, this mixture was vortexed for 10 minutes and centrifuged for 4 minutes at 10,000 rpm. The flowthrough was discarded. Five hundred microlitres of Hemovoid wash buffer was vortexed with the sample for 5 minutes and then centrifuged at 10,000 rpm for 4 minutes. The flowthrough was discarded and this was repeated twice before elution. Three hundred microlitres of elution buffer was added to the sample, which was vortexed for 10 minutes and then centrifuged for 4 minutes at 10,000 rpm to elute the depleted proteins.

After processing, samples were dried down entirely in a speed vac, before a trypsin digestion and disulphide bond reduction was carried out as instructed by the manufacturer: The dried samples were resuspended in 100 µl 200 mM NH_4HCO_3. Twenty microlitres of 50 mM DTT solution was added to samples which were then incubated at 60˚C for 45 minutes. One hundred microlitres 22 mM iodoacetamide was added and samples were incubated in the dark for 25 minutes. Residual iodoacetamide was then quenched with 2.8 µl of DTT and then 2µg of trypsin. Samples were incubated at 40˚C and 800 RPM. Samples were stored at -20˚C prior to analysis.

Chapter 3 : Liquid extraction surface analysis of intact proteins from polyvinylidene fluoride membranes

3.1 Introduction

The aim of the work presented in this chapter was to investigate whether LESA can be used to extract intact proteins from PVDF membranes. The work was performed for two reasons: firstly, to try and develop a new technique for the fractionation of intact proteins based on sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) to be used in top down proteomics research. Proteomics experiments ordinarily require the analysis of complex mixtures of proteins that require separation prior to mass spectrometry analysis. Although SDS PAGE is a useful method for separating proteins, there are no reliable techniques for extracting intact proteins out of a gel in a format that is compatible with mass spectrometry analysis [139].

Electroblotting is a commonly used molecular biology technique for extracting proteins from SDS PAGE gels. PAGE separates out a mixture of proteins into narrow bands which are then blotted onto an absorbent PVDF membrane by use of an electric current. Blotting immobilises the protein onto the membrane which can then be probed with antibodies to identify a particular protein. This process is commonly carried out during Western blotting [193]. The sample format, i.e., protein on PVDF, has previously been incompatible with electrospray ionisation analysis because the proteins are immobilised on a solid support. These proteins could be accessible for analysis through a surface

analysis technique such as LESA and this could become a means for incorporating the power of PAGE separation into top down proteomics research.

In order to achieve this goal, four principles must be demonstrated: 1. Intact proteins must be successfully extracted from PVDF membranes. 2. The extraction process must be demonstrated on proteins blotted from an SDS PAGE gel. 3. The technique must be able to identify unknown proteins. 4. It must be demonstrated to be effective on a range of different proteins of different molecular weights.

A second potential application of the technique is analysis of proteins from membrane filters. The key points of interest here are to determine the most suitable LESA and mass spectrometry parameters for optimum sensitivity and consequently to determine the sensitivity of the approach. The requirements of this latter application clearly have some overlap with those of the former.

3.1 Optimisation of extraction solvent composition

The solvent composition used in the LESA sampling and electrospray process has an important effect on the recovery of proteins from the membrane and the subsequent ionisation. PVDF membranes are hydrophobic but absorbent to organic solvents. ESI solvents are ideally composed of an aqueous organic mixture, but the membrane will absorb high concentrations of organic solvents, rendering formation of a liquid microjunction in the extraction process impossible. To deduce appropriate solvent

compositions, 5 µl aliquots of solutions with varying methanol and acetonitrile compositions were spotted onto a PVDF membrane by manual pipetting to investigate whether a liquid microjunction could be formed.

Solvent composition	microjunction formed Y/N		
	Replicate 1	Replicate 2	Replicate 3
100% H_2O	Y	Y	Y
10 % MeOH	Y	Y	Y
20 % MeOH	Y	Y	Y
30 % MeOH	Y	Y	Y
40 % MeOH	N	N	N
50 % MeOH	N	N	N
60 % MeOH	N	N	N
70 % MeOH	N	N	N
80 % MeOH	N	N	N
90 % MeOH	N	N	N
100 % MeOH	N	N	N

Table 3.3.1 Formation of liquid microjunctions at different methanol compositions

Solvent	microjunction formed Y/N		
	Replicate 1	Replicate 2	Replicate 3
100% H_2O	Y	Y	Y
10 % ACN	Y	Y	Y
20 % ACN	Y	Y	Y
30 % ACN	N	N	N
40 % ACN	N	N	N
50 % ACN	N	N	N
60 % ACN	N	N	N
70 % ACN	N	N	N
80 % ACN	N	N	N
90 % ACN	N	N	N
100 % ACN	N	N	N

Table 3.3.2 formation of liquid microjunctions at different acetonitrile compositions

Tables 3.1 and 3.2 show that liquid microjunctions with a high organic content could not be formed on PVDF membranes. The maximum organic content of a solution that could form a microjunction was 30% methanol, and 20% acetonitrile. Above these limits, the applied solutions were absorbed by the membrane. ESI is more efficient with organic compositions approaching 50% [13]. Consequently, solutions comprising 30% methanol were used for the majority of this work because those solutions were capable of maintaining a liquid microjunction on PVDF with a higher organic content than solutions containing acetonitrile.

3.2 Recovery of intact proteins from PVDF membranes

To investigate whether intact proteins could be recovered from PVDF membranes, a range of solutions of lyophilised proteins were spotted onto PVDF and dried before being sampled by LESA. Experiments were performed on 4 purified proteins, Bovine cytochrome C, Equine myoglobin, Bovine alpha casein and Yeast alcohol dehydrogenase. These 4 proteins were of molecular weights in the range of 12-33 kDa. Five microlitre aliquots of a 10 μM 70/30% water/methanol solution of each protein were dried onto the membrane at marked positions, to give a total of 50 pmol of protein in each spot. Protein spots were sampled with LESA by use of the Triversa Nanomate™ coupled to the Orbitrap Velos™ mass spectrometer, see Section 2.2.3.1. The samples were extracted in a solvent composed of 69.7/29.3/1% water/methanol/formic acid. In each extraction process, 4.5μl of solution was aspirated from the solvent reservoir and 2.5μl was dispensed onto the surface of the PVDF with the pipette tip held 1.6 mm above the surface to form a liquid microjunction. The liquid microjunction was held in contact for 2 seconds before being aspirated back into the pipette tip. This cycle was repeated twice before the samples were infused into the mass spectrometer. Mass spectra shown below were acquired for 3 minutes in full scan mode, with each scan being composed of 5 co-added microscans.

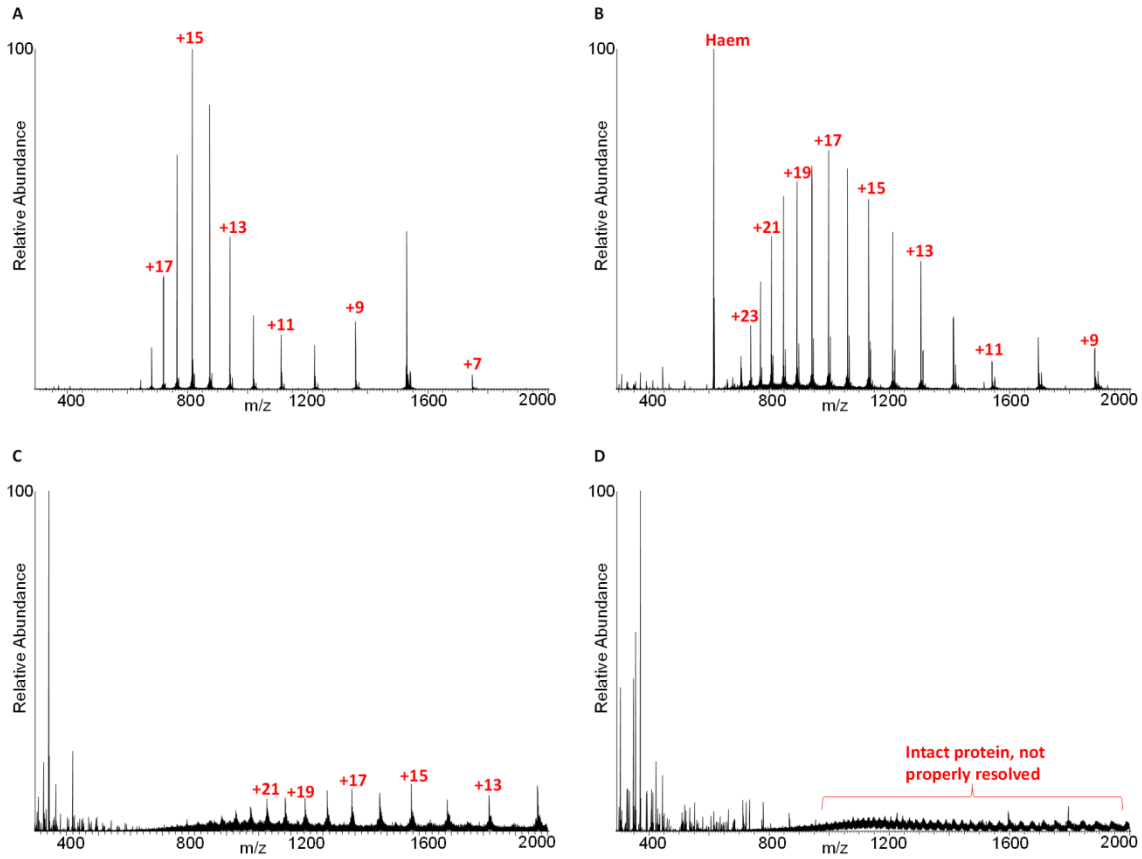


Figure 3.1 LESA mass spectra of intact proteins extracted from PVDF membranes (a) bovine cytochrome C 12.3 kDa, (b) equine myoglobin 16.9 kDa, (c) bovine alpha casein 22 kDa (d) yeast alcohol dehydrogenase 33 kDa.

Fig 3.1 shows the mass spectra from four different proteins spots on PVDF. These spectra show that proteins up to 22 kDa can be successfully analysed via LESA from PVDF. The mass spectrum obtained following LESA of alcohol dehydrogenase (33 kDa) showed peaks corresponding to the protein although they were not well-resolved. Peaks corresponding to the 22 kDa alpha casein protein were also less well-resolved than myoglobin or cytochrome C. These results demonstrate that intact proteins can be successfully extracted from PVDF membranes and that it can be used to analyse proteins of different

sizes, but this particular instrument could not successfully analyse proteins from a PVDF membrane larger than 30 kDa.

Myoglobin was selected as the model protein to investigate compatibility of the LESA technique with electroblotting. Prior to blotting the protein from a gel, the sensitivity of LESA extraction from PVDF was investigated. A serial dilution of myoglobin solutions were prepared with concentrations of 100 µM 10 µM 1 µM, 100 nM 10 nM, 1 nM , 100 pM 10 pM and 1 pM. Samples were made up in 70/30% water/methanol and 5 µl aliquots of each were deposited onto a PVDF membrane at marked positions. The samples spread over an area of approximatley 0.07 cm^2 and were left to dry. This left a series of spots with 500 pmol, 50 pmol 5 pmol, 500 fmol, 50 fmol, 5 fmol, 500 amol 50 amol and 5 amol of protein on the surface. Protein spots were analysed by use of the LESA mass spectrometry method described above.

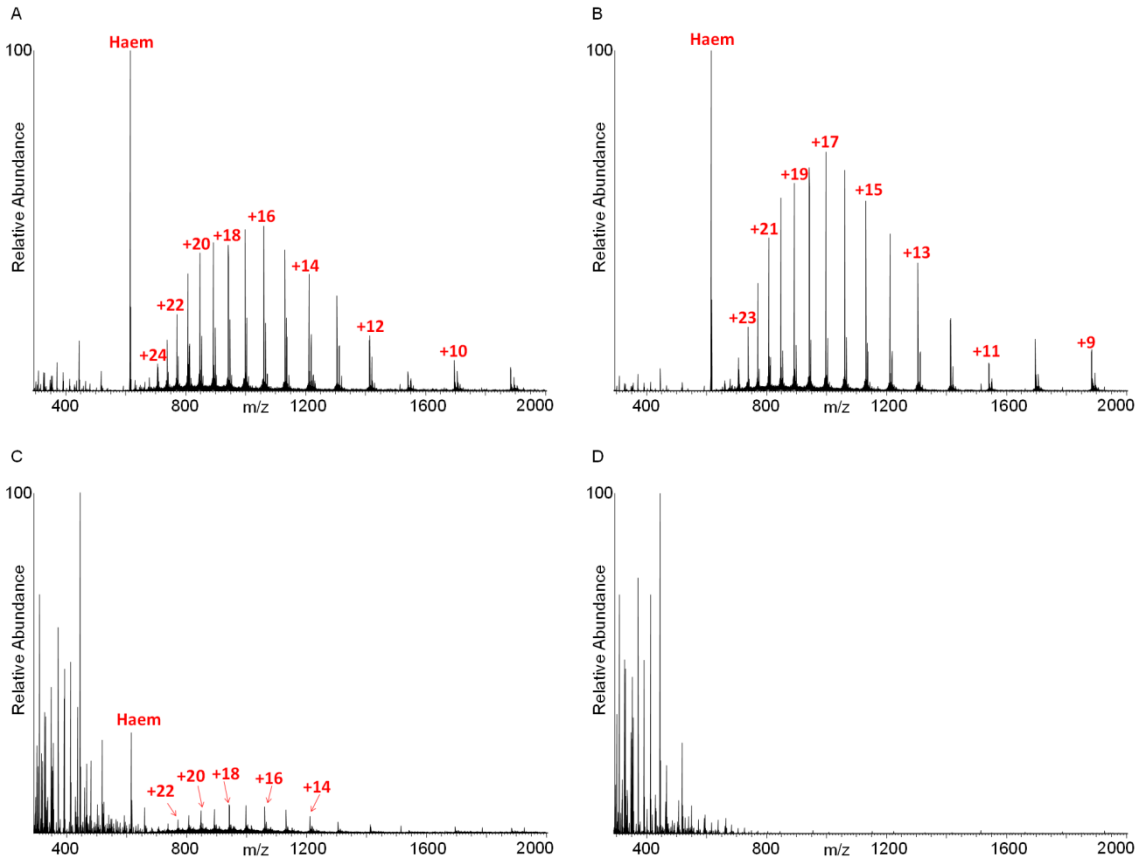


Figure 3.2 LESA of intact myoglobin from PVDF with different amounts of protein (a) 500 pmol (b) 50 pmol (c) 5 pmol (d) 500 fmol

Fig 3.2 shows LESA spectra of myoglobin from PVDF membranes with different amounts of protein. Signals were observed from spots with as little as 5 pmol of protein. A typical PVDF membrane can recover around 33% of the protein loaded into a gel [194]. At least 15 pmol of protein must be loaded into the gel for 5 pmol of protein to be recovered on the membrane. Fifteen pmol of myoglobin is approximately 0.256 µg of protein. A typical Western blot may have around 20-40 µg of a mixture of proteins loaded onto the gel [195]. This could be a very complex mixture such as a cell lysate that could possibly

contain thousands of proteins [196, 197]. In a very complex sample many low abundance proteins will be below this detection limit,[197] but the 5 pmol limit should be sensitive enough to detect some of the high abundance proteins, or be useful for less complex samples.

3.3 Recovery of intact proteins blotted from SDS PAGE

In section 3.2 it was shown that the sensitivity of LESA extraction of intact proteins from PVDF membranes was sufficient for the detection of proteins blotted from a gel. Consequently, LESA was used to analyse an electroblot of myoglobin. An SDS PAGE and electroblot was carried out as described in section 2.2.4 with 10 μg of protein loaded per lane onto a gel. Electroblotted protein samples were analysed by LESA mass spectrometry as described in section 3.2.

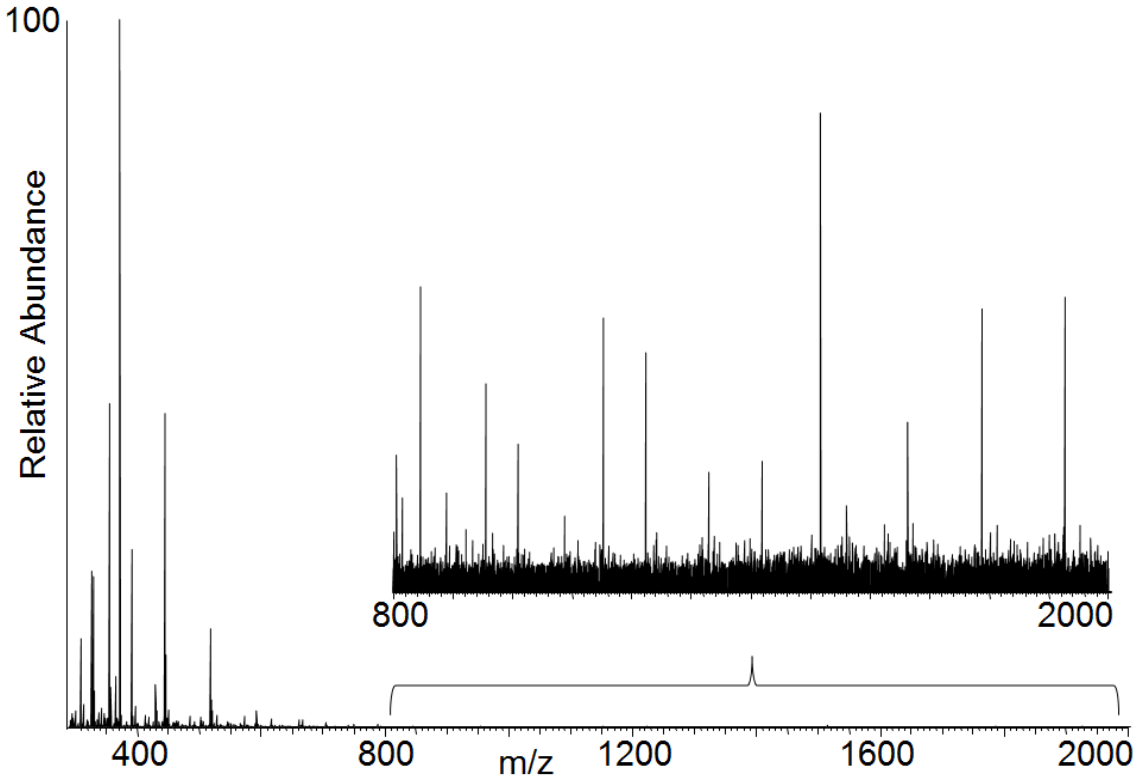


Figure 3.3 LESA mass spectrum obtained from PVDF membrane after electroblotting of myoglobin from an SDS PAGE gel. Ten micrograms of myoglobin was loaded onto the gel.

The mass spectrum in fig 3.3 shows that no signal corresponding to myoglobin was observed following LESA of a PVDF membrane blotted from an SDS PAGE gel loaded with 10 µg of protein. This amount of protein is around 40 fold higher than that needed to generate a signal according to levels determined in the experiments of section 3.2. To investigate the lack of observed protein signal, further experiments were conducted as described below.

3.4 Negative ion mode LESA analysis of PVDF membranes

One reason for the absence of signal observed in the spectrum shown in fig 3.3 could be the presence of dodecyl sulphate adducts extracted from the gel along with the protein. Sodium dodecyl sulphate is a notorious signal suppressor in ESI mass spectrometry that inhibits the ionisation of proteins and peptides. The presence of SDS on ESI has two separate negative effects: Firstly, it is a detergent which reduces the surface tension of a solution. When a solution is charged during the ESI process any surfactants present will destabilise the formation of the Taylor cone [198]. Secondly, the surfactant aggregates at the liquid air boundary and fewer ESI droplets will be emitted in the desorption process of ESI. In the droplets that are emitted, the surfactant accumulates at the liquid air boundary preventing the analytes from reaching the surface. In ESI, charge transfer from solvent to solute takes place at the surface of these droplets and the presence of SDS prevents this, thereby inhibiting ionisation [198]. SDS is responsible for providing a uniform charge to the proteins that is responsible for the electrophoretic mobility when a current is applied. This mobility is required for resolving proteins during PAGE and drawing extracting the proteins from the gel during electroblotting, however high concentrations of SDS can reduce the binding affinity between the protein and the PVDF [194]. The transfer buffers contain methanol to attempt to strip it from the protein. The removal may not be entirely complete and some SDS may remain bound to the protein [194]. The presence of SDS could explain the absence of signal in section 3.3. If SDS is retained on the membrane and is extracted along with the protein during LESA, it may be inhibiting the ionisation of the proteins. To investigate this hypothesis, an electroblot of

myoglobin was analysed via negative ion mode LESA mass spectrometry (LESA MS) to investigate the presence of dodecyl sulphate ions. SDS PAGE and electroblotting was carried out as described in section 2.2.4 with 10 µg of protein loaded per lane. Samples were analysed with LESA parameters as described in section 3.2, with the exception that samples were ionised and analysed in negative ion mode. The ionisation voltage was -1.7 kV and the gas pressure was 0.3 PSI. LESA extraction parameters remained the same as in section 3.2 with the exception that the extraction solvent composition was 30/70% water/methanol. The solvent did not contain formic acid as samples were being ionised in negative ion mode.

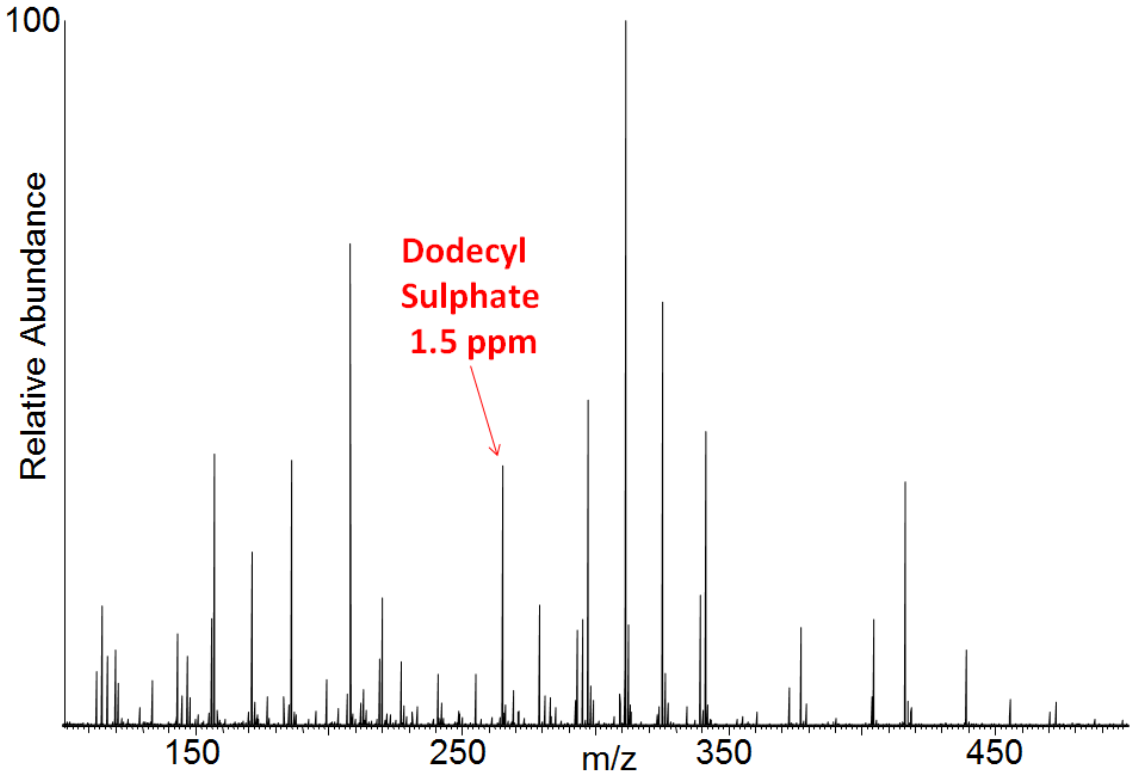


Figure 3.4 Negative ion mode LESA mass spectrum obtained from PVDF membrane after electroblotting myoglobin from SDS PAGE gel. Ten micrograms of myoglobin was loaded onto the gel.

Dodecyl sulphate anions were detected from myoglobin bands on the PVDF membrane as shown on the mass spectrum in fig 3.4. The measured m/z of the dodecyl sulphate peak was 265.1482, 1.5 ppm greater than the theoretical m/z of 265.1478. There is no indication whether these ions were protein bound adducts or free ions present on the membrane surface. This observation supports the hypothesis that SDS could be responsible for suppressing the myoglobin signal in section 3.2. Future efforts were aimed at developing means for eliminating SDS signal suppression. Eliminating signal suppression could require the incorporation of clean-up stages into the sampling routine. Although LESA is an automated process, one of the advantages of the procedure is that it is possible to decouple sampling and ionisation.

3.5 Elimination of SDS signal suppression

3.5.1 Reversed phase C_4 Zip tips

Reversed phase C_4 zip tips are a commercially available approach for desalting intact proteins. They are specialised pipette tips with beads of C_4 silica immobilised in the tip. Peptides and proteins will bind to the tips in aqueous solutions. Salts and detergents will not bind in these conditions. After the proteins are bound to the beads, the tips can be rinsed with an aqueous solution to eliminate these contaminants before the proteins are eluted. Elution is achieved by passing an organic solution through the beads. The

advanced user interface section (AUI) of the Chipsoft software used for controlling the Triversa Nanomate™ can be used to customise the robotic aspiration and dispense cycles of the sampling process. This feature makes it viable to extract blotted proteins from a PVDF membrane using a LESA sampling cycle and place them back into a well of a microtitre plate. The samples could then be purified with C_4 zip tips manually before final analysis by direct infusion ESI. SDS PAGE and electroblots were carried out as described in section 2.2.4 with 10 µg of protein loaded per lane. An AUI sampling routine was written to allow extraction of proteins from the surface of the PVDF before dispensing into a well. In this sampling method, a PVDF membrane was placed on the surface of a microtitre plate with the myoglobin bands located on the centre of the surface of the wells. One well was filled with an extraction solvent and the remaining wells were left empty. Eight microlitres of a 69.7/29.3/1% water/methanol/formic acid solution was extracted from the solvent reservoir before 7 µl was dispensed onto the surface of a PVDF membrane. Larger microjunction volumes were used to account for sample loss from dispensing the solution into a well in a microtitre plate after LESA. After a delay of approximately 3 seconds, 6.5 µl of sample was reaspirated into the pipette tip and the entire solution was deposited back into one of the empty wells of the microtitre plate. Samples were then purified by reversed phase C_4 Zip tips as described in section 2.2.5. After samples were purified by C_4 zip tips, they were analysed by direct infusion electrospray ionisation, at a gas pressure of 0.3 PSI and an ionisation voltage of 1.7 kV.

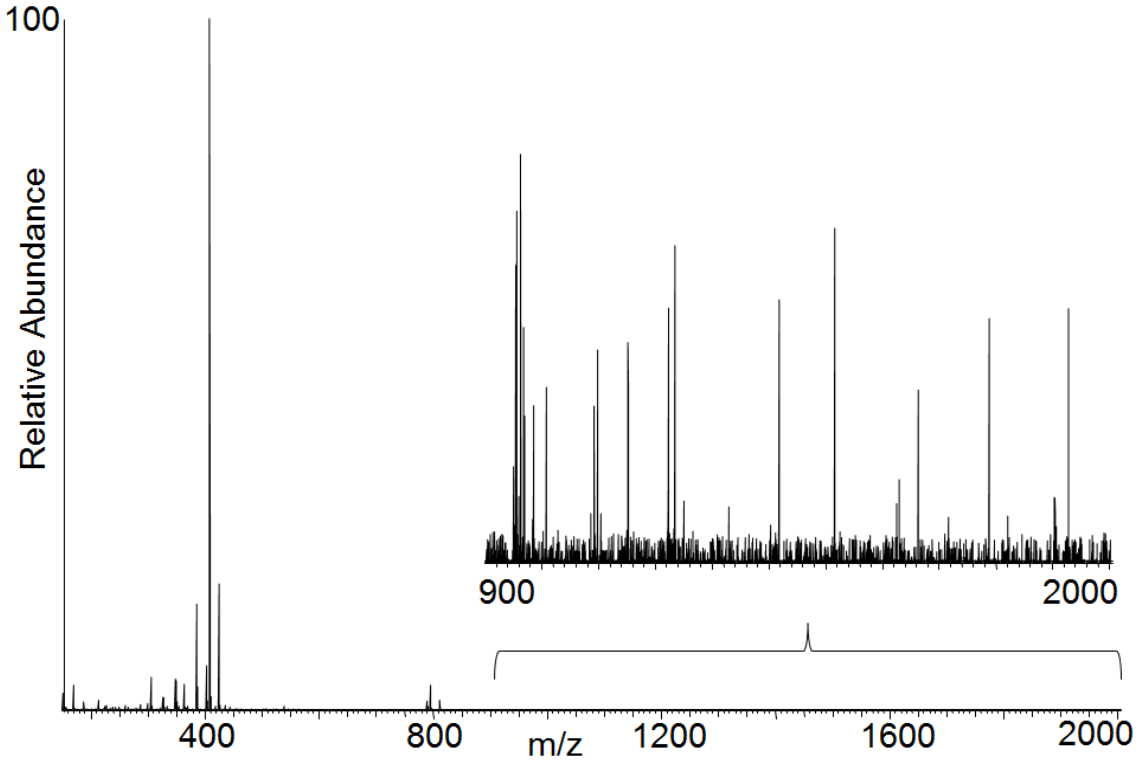


Figure 3.5 Direct infusion ESI mass spectrum of myoglobin afer surface extraction and zip tipping of a PVDF membrane electroblotted from SDS PAGE gel. Gel was loaded with 10 µg of myoglobin

Fig 3.5 shows the mass spectrum obtained from using c_4 zip tips to purify a sample extracted from an electroblot with LESA. The results show that zip tipping did not eliminate signal suppression. The composition of SDS in the PAGE running buffers is approximately 0.1% w/v which is much higher than the concentration of the proteins. During page one SDS molecule binds along the protein at every two amino acids [199]. To remove all the SDS will be a major challenge. The most likely reason for the failure of this treatment appears to be that the zip tipping was could not to remove all the SDS from the protein. Other treatments were investigated to remove the SDS before the surface extraction process, as described below.

3.5.2 Cetyltrimethyl ammonium bromide

A method to eliminate the effects of SDS from ESI solutions has recently been presented by Shieh *et a. l*[200]. In this method the cationic surfactant cetyltrimethyl ammonium bromide (CTAB) is added into the ESI solution at a concentration equivalent to the concentration of SDS. When in solution, SDS and CTAB form ion pairs which are stronger than the interactions between the protein and the SDS. These ion pairs are insoluble in aqueous solutions and precipitate out [200]. If CTAB is used for this purpose it must be used carefully: it acts as a ESI signal suppressor in its own right and to effectively eliminate the SDS it must be used in an equivalent concentration [200]. The concentration of SDS within the ESI after LESA is unknown and to reduce the possibility of causing signal suppression, CTAB was used to eliminate the SDS from the PVDF membrane after the proteins had been electroblotted onto it. SDS PAGE and electroblotting was carried out as described in section 2.2.4. With 10 μg of myoglobin loaded per lane. After blotting, the entire PVDF membrane was rinsed in a 1% solution of CTAB for 10 minutes which was then dried for 2 hours. The membrane was rinsed in methanol for 30 seconds to remove residual CTAB and air dried. Membranes were analysed by the LESA mass spectrometry method described in section 3.2.

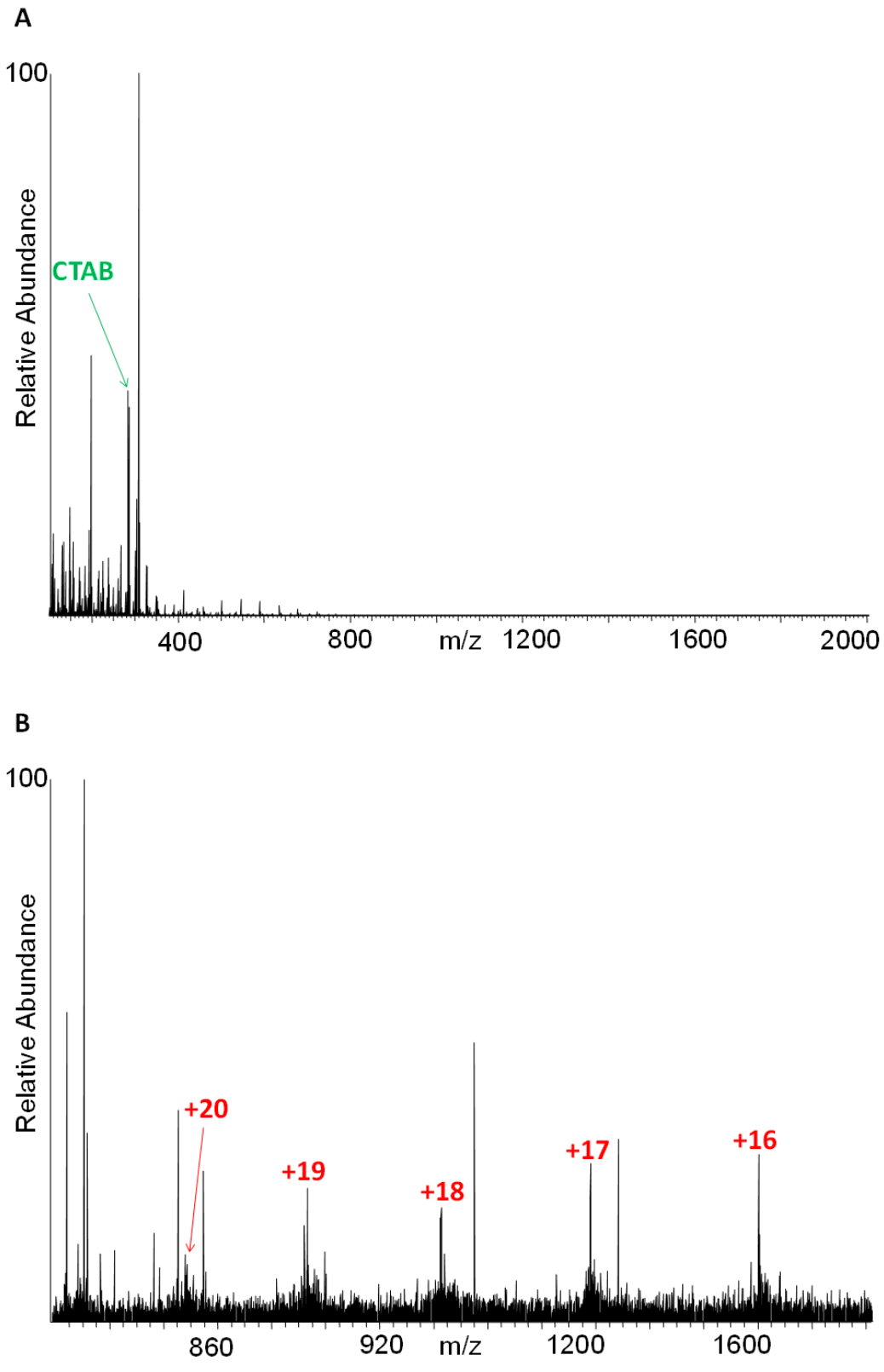


Figure 3.6 LESA mass spectrum of PVDF membrane plotted from an SDS PAGE gel after treatment with CTAB (A) full mass spectrum in the mass range 190- 2000 (B) magnified region of m/z 750-1100

Fig 3.6 shows the LESA mass spectrum of a PVDF membrane electroblot rinsed with CTAB prior to extraction. Low abundance peaks corresponding to myoglobin were detected. These peaks could only be seen under magnification by approximately 770 fold of the m/z 800-1100 region, see fig 3.6 B. Residual CTAB was observed in the full scan mass spectrum, see fig 3.6 A. The signal intensity is too low to carry out fragmentation and if the protein was unknown it could not be identified from a mass spectrum of this quality. The method was repeated, but this signal could not be reproduced or improved.

3.6 Recovery of intact proteins from PVDF membranes after native PAGE

Considering attempts to eliminate SDS had proved to be unsuccessful, native PAGE was investigated as an alternative means of separating proteins prior to electroblotting. Native PAGE is an alternative to SDS page that, rather than using a surfactant to denature proteins, separates intact proteins on the basis of their native charge or following the addition of mild surfactants such as Coomassie blue dyes [201]. Native PAGE lacks the molecular weight based separation that is provided by SDS but can still be used to separate mixtures of proteins into their individual complexes [201]. Interestingly it may also retain non-covalently bound subunits [202]. Native PAGE could eliminate the problem of signal suppression and importantly confirm whether the presence of SDS was responsible for the non-detection of proteins described above. A native PAGE and electroblot was run as described in section 2.2.6, with 100 μg of protein loaded on the

gel. Samples were analysed with the LESA MS method described in section 3.2, with the exception that samples were analysed by a mass spectrometry method composed of 30 microscans in the mass range of 450-2000 for 8 minutes.

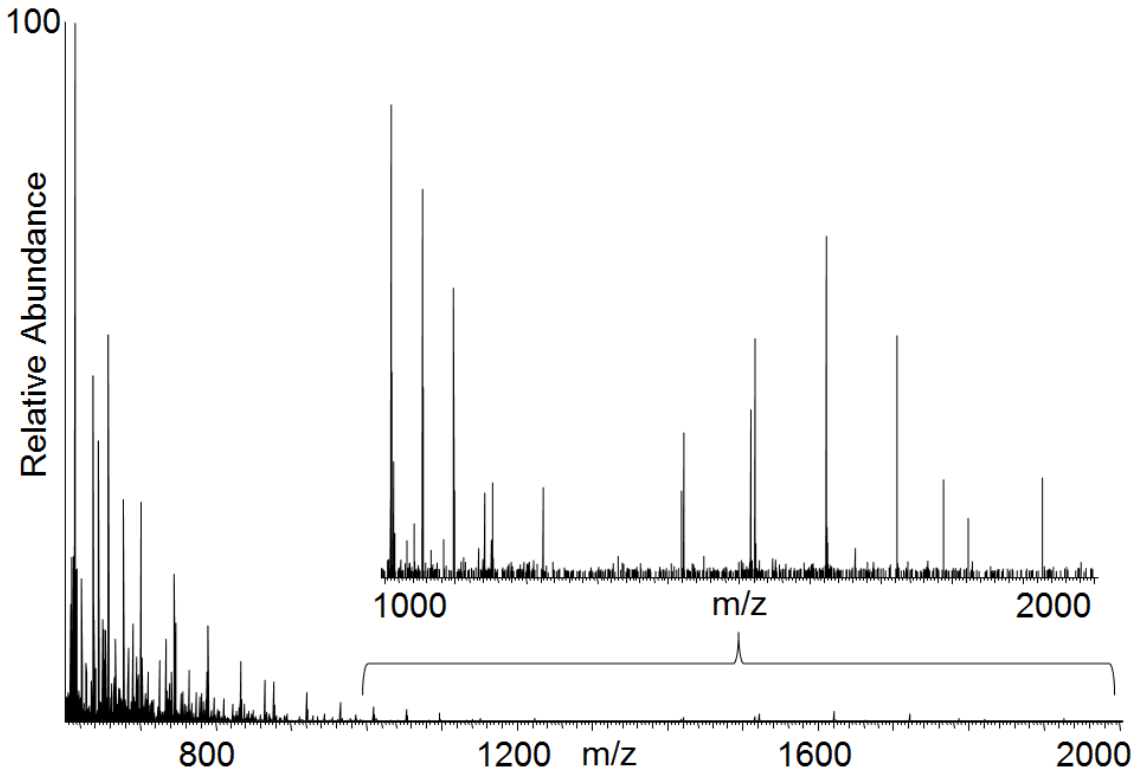


Figure 3.7 LESA mass spectrum of myoglobin extracted from a PVDF membrane blotted from a native gel loaded with 100 µg of protein.

Fig 3.7 shows a LESA mass spectrum of a PVDF membrane blotted from native PAGE of myoglobin. No peaks were observed that corresponded with myoglobin. This demonstrates the lack of signal observed in previous experiments was not entirely due to signal suppression from SDS. There are other problems associated with electroblotting that may make the technique incompatible for analysis by a surface sampling technique such as LESA. Firstly, proteins do not always bind to the surface of the PVDF and may become transmitted through the polymeric matrix. A second problem may be that when

the proteins do bind to the PVDF, they bind irreversibly [203]. LESA is a surface sampling technique, only capable of extracting soluble analytes weakly associated with the surface of a substrate. It cannot be used to solubilise strongly bound analytes buried deep within the matrix of a solid polymer.

3.7 Optimisation of extraction parameters for intact protein analysis from PVDF membranes

The results described previously show that electroblotting is not a suitable application for LESA, so the work presented in the second half of this chapter was aimed at analysing intact proteins from PVDF membrane based air filters used in air quality monitoring programmes. The remainder of the work presented in this chapter was performed in collaboration with Unilever plc. Initial experiments focussed on optimising the extraction conditions for the model protein β lactoglobulin (βlg) from PVDF membranes. In section 3.1, signals were observed for the protein myoglobin from membranes spotted with as little as 5 pmol of protein over an area of approximately 0.07 cm^2. In order to improve sensitivity, each of the LESA extraction parameters was optimised.

3.7.1 Calculation of surface areas and surface protein concentrations

To calculate the surface concentration of protein in a spot, a range of droplets with different volumes were measured to show how far they spread on the PVDF. The

droplets were composed of 70/30% water/methanol and were applied to a membrane and the diameters of these spots were measured. The areas of the spots were calculated and are shown in table 3.3, this was performed in triplicate.

Volume deposited / µl	Diameter / mm			Area / cm^2		
	Replicate 1	Replicate 2	Replicate 3	Replicate 1	Replicate 2	Replicate 3
5	3	3	3	0.07	0.07	0.07
4.5	3	3	3.5	0.07	0.07	0.096
4	3	3	3	0.07	0.07	0.07
3.5	3	3	3	0.07	0.07	0.07
3	3	3	2.5	0.07	0.07	0.05
2.5	2.5	2.5	2.5	0.05	0.05	0.05

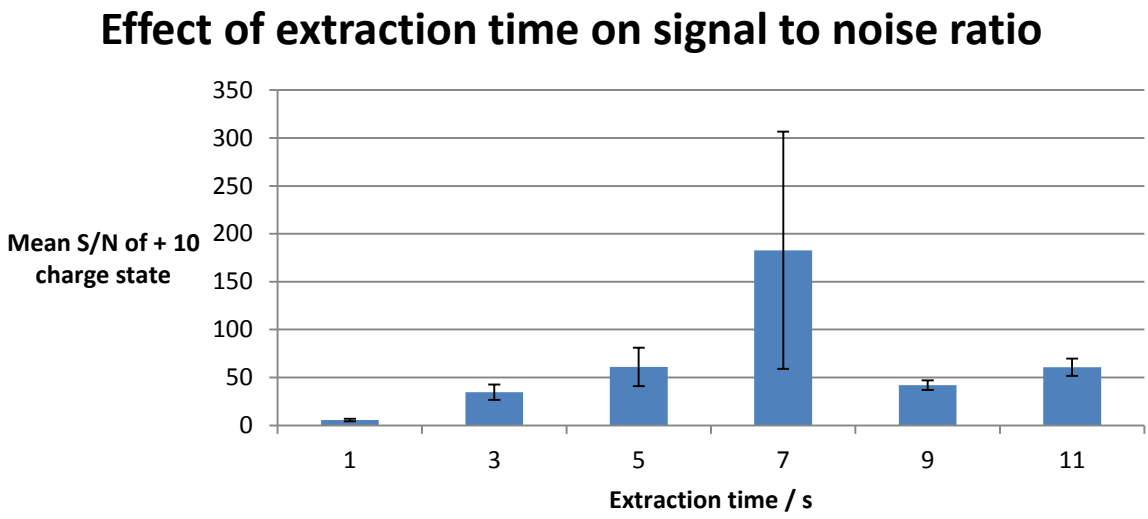
Table 3.2 Diameters and areas of sample spots following deposition of droplets of varying volumes onto PVDF membranes.

Table 3.3 shows the areas covered by sample spots of different volumes. A five microlitre droplet produces a sample spot of area 0.07 cm^2. For protein concentrations of 1 pg/µL, the surface concentration of the sample spot will be 70 pmol/ cm^2.

3.7.2 Sampling time

The sampling time of a LESA sampling routine can be controlled either by increasing the post dispense delay or increasing the number of repeated mix cycles. The post dispense delay is the period of time between dispensing the solvent onto the surface of the PVDF and re-aspirating it into the pipette tip. A series of 70 pmol/cm^2 spots were dried onto a

PVDF membrane and sampled by LESA whilst varying the post dispense delay. Experiments were performed in triplicate. Seven microlitres of a 69.7/29.3/1% water/methanol/formic acid was aspirated from the solvent reservoir and 4.5 µl were deposited onto the surface of the membrane at a height of 1.6 mm above the surface. (Note that a 2.5 µl microjunction was used in the experiments described in sections 3.2-3.6; however the results shown in section 3.7.1 suggest that a larger microjunction is required to fully cover the protein spot). The delay before aspiration of 5 µl was varied between 1 and 11 seconds at 2 second intervals. The aspiration and dispense cycle was repeated once before being infused into the mass spectrometer. Samples were analysed by a 7 minute mass spectrometry method composed of 5 minutes full scan in the m/z region of 600-2000, followed by 1 minute isolation of the +10 charge state of the protein, followed by a one minute SIM to monitor the same charge state. The method is described in detail in section 2.2.3.1 and was used throughout all future optimisation experiments. In all cases, each scan was composed of 5 co-added microscans. The optimum conditions were chosen on the basis of the signal to noise ratio of the + 10 charge state of the protein.



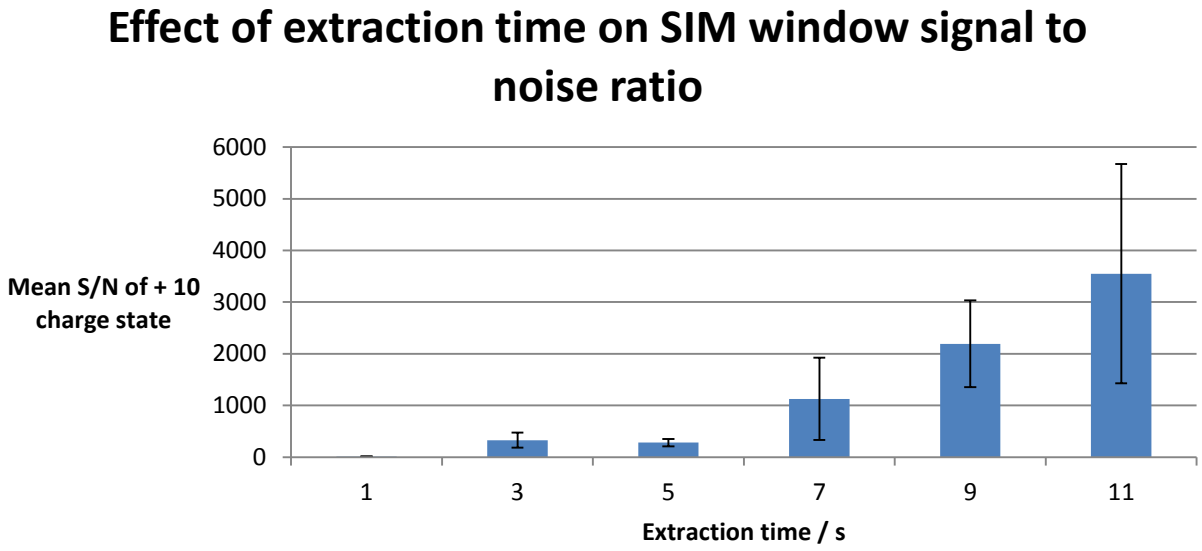
SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1 Second	3	16.62	5.54	8.4196
3 Seconds	3	104.05	34.683333	261.48243
5 Seconds	3	182.9	60.966667	1596.7582
7 Seconds	3	548.562	182.854	61323.779
9 Seconds	3	126.1	42.033333	100.49923
11 Seconds	3	181.98	60.66	323.2557

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	56714.863	5	11342.973	1.069853	0.42376744	3.105875
Within Groups	127228.39	12	10602.366			
Total	183943.25	17				

Figure 3.8. Chart showing effect of extraction time on mean signal to noise ratio of the +10 charge state of βlg obtained over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.



SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1 Second	3	33.1	11.033333	56.389233
3 Seconds	3	984.88	328.29333	84522.947
5 Seconds	3	844.29	281.43	20418.662
7 Seconds	3	3379.56	1126.52	2535023.8
9 Seconds	3	6581.41	2193.8033	2800978.3
11 Seconds	3	10645.16	3548.3867	18004956

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	28533853	5	5706770.5	1.4604064	0.27296147	3.105875
Within Groups	46891912	12	3907659.3			
Total	75425764	17				

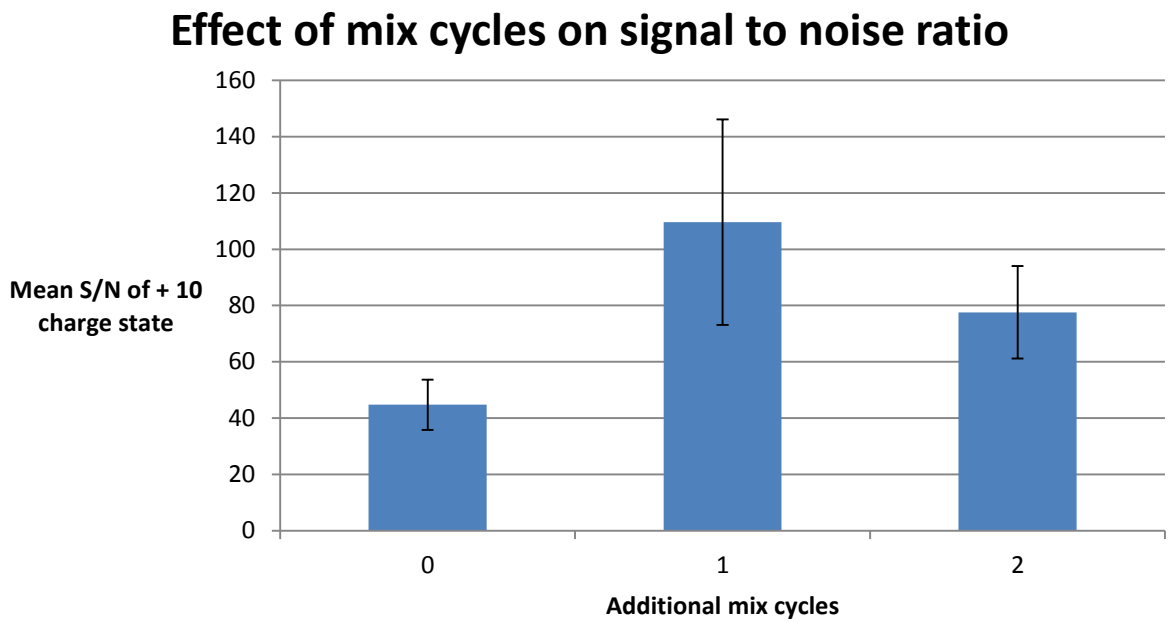
Figure 3.9 Chart showing effect of extraction time on mean signal to noise ratio of the +10 charge state of βlg obtained from targeted SIM window over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.

The effect of sampling time on the signal to noise ratio of the + 10 charge state of βlg is shown in fig 3.8 and fig 3.9. The results of the one way ANOVAs however show that the

effect of sampling time did not have a statistically significant effect on the signal to noise ratio of either the full scan (p=0.42) or the targeted SIM window (p=0.27) Both fig 3.8 and 3.9 show that as the signal to noise ratio increases they error bars also get larger and the variance amongst the data also increases. Despite the results of the ANOVA, figs 3.8 and 3.9 show that the average signal to noise ratio of the +10 charge state was greatest in the full scan when the post dispense delay was 7 s whereas in the SIM mode mass spectrum the average signal to noise ratio was greatest when the delay was 11 seconds. This difference between the results is likely to be due to the increased influence of polymeric contaminants extracted in longer the sampling times. Longer extraction times will also allow the diffusion of more contaminants from the membrane into the extraction solvent along with the protein thus reducing the signal to noise ratio of the protein. The targeted SIM-mode analysis eliminates some of this noise and ensures that the ions contributing to the AGC target will primarily be protein ions, leading to spectra with a higher signal to noise. For known proteins, such as those likely to be analysed in air monitoring programmes, targeted mass spectrometry approaches such as the SIM-mode analysis are ideal. Despite that the results were not statistically significant the greatest average signal to noise ratio acquired with the SIM-mode analysis was found at 11 seconds so this was used in further experiments.

The sampling time of LESA can also be increased by increasing the number of mix (dispense/reaspirate) cycles which also mixes the sample in the pipette tip. An experiment was carried out to determine the optimum number of mix cycles between 0,

1 and 2 on a series of 70 pmol/cm^2 spots. Experiments were performed in triplicate, the remainder of the sampling routine was the same as described previously with the exception that a post-dispense delay of 11 seconds was applied.



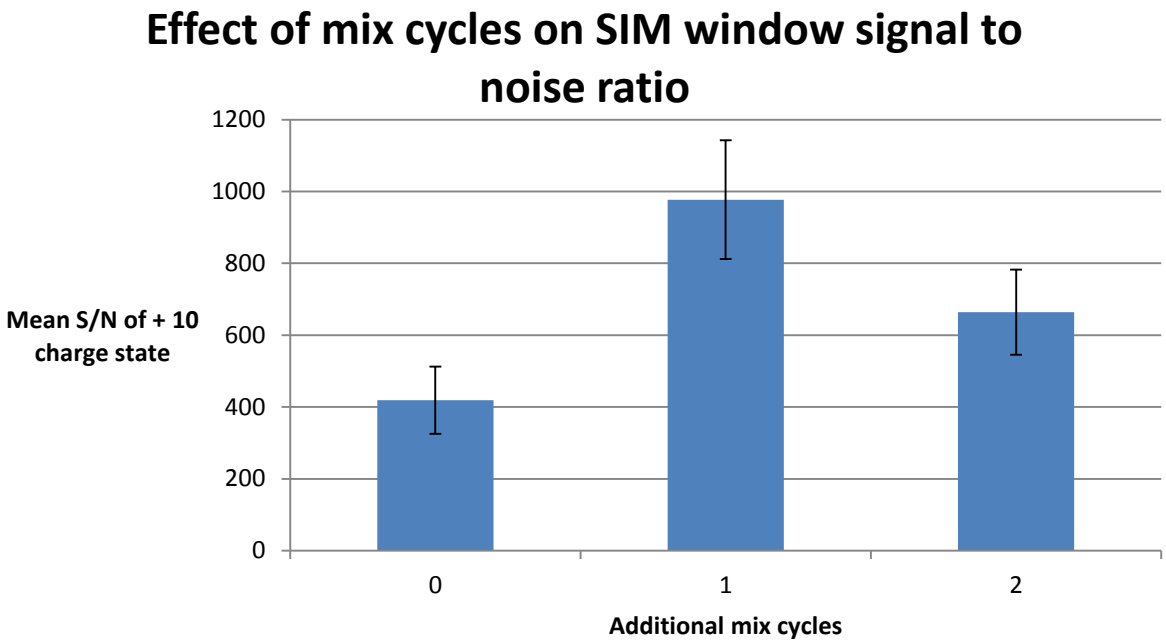
SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0 mixings	3	134.27	44.75667	321.7572
1 mixings	3	328.84	109.6133	5323.698
2 mixings	3	232.74	77.58	1079.59

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	6309.8929	2	3154.946	1.407402	0.315367	5.143253
Within Groups	13450.091	6	2241.682			
Total	19759.984	8				

Figure 3.10 Chart showing effect of additional mix cycles on mean signal to noise ratio of the +10 charge state of βlg obtained over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.



SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0 mixings	3	1256.94	418.98	35094.16
1 mixings	3	2931.28	977.0933	109200.6
2 mixings	3	1992.75	664.25	56378.76

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	469518.82	2	234759.4	3.509572	0.097883	5.143253
Within Groups	401347.08	6	66891.18			
Total	870865.9	8				

Figure 3.11 Chart showing effect of additional mix cycles on mean signal to noise ratio of the +10 charge state of βlg obtained from targeted SIM window over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.

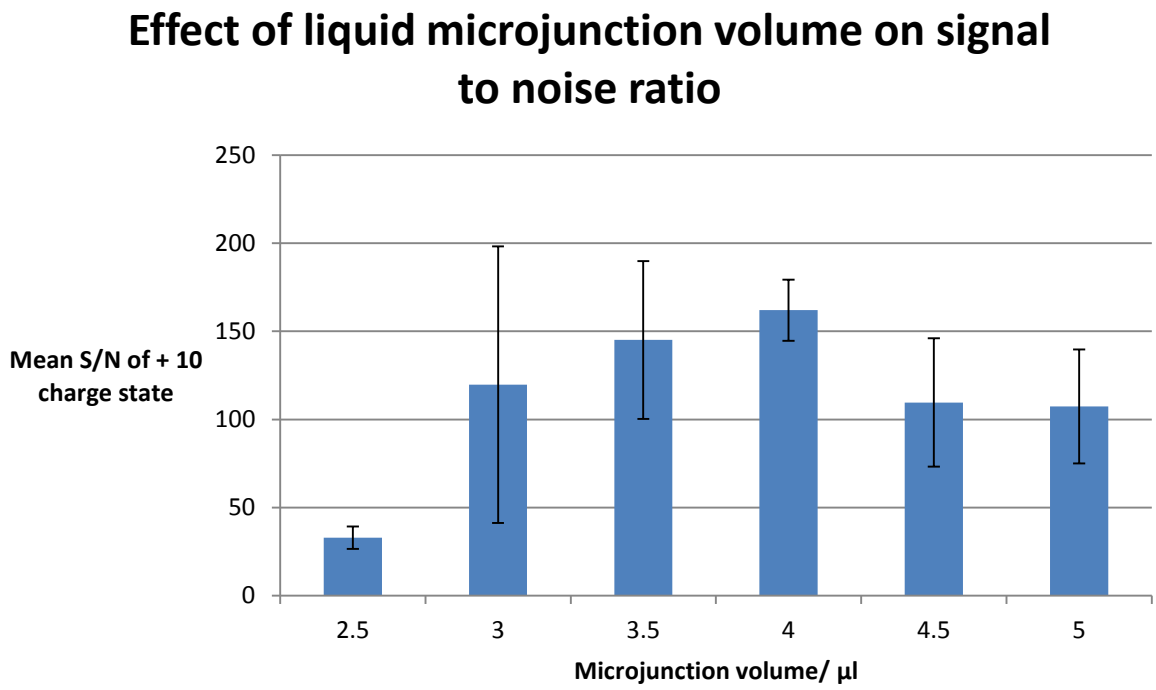
The effect of adding additional mix cycles into the LESA sampling routine on the signal to noise ratio of the + 10 charge state of βlg is shown in fig 3.10 and 3.11, along with corresponding one way ANOVAs. The average signal to noise ratio's for both the full scan

mass spectra and the SIM-mode mass spectra was at its greatest following one repeat mix cycle. The one way ANOVAs however show that this effect was not statistically significant on either the full scan ($p=0.31$) or the SIM mode mass spectrum ($p=0.09$). As mentioned previously, adding mix cycles increases the length of time that the microjunction is in contact with the surface of the PVDF and also mixes the sample within the pipette tip and mixes the sample in the pipette tip. This resulted in an improved signal to noise ratio following the first mix cycles, but adding a second did not provide any improvement. The ANOVA showed this effect was not statistically significant but the greatest average signal to noise ratio from this experiment was obtained following one mix cycle, which was retained for future experiments.

3.7.3 Liquid microjunction volume

The volume of solvent applied to surface was varied to determine the optimum liquid microjunction volume. Changing the microjunction volume affects the area of PVDF being sampled. βlg spots were analysed with a solution composed of 69.3/29.7/1% water/methanol/formic acid. Seven microlitres of this solution was aspirated from the solvent well and a range of different volumes between 2.5 and 5µl at intervals of 0.5µl were applied to the surface of the PVDF. Experiments were performed in triplicate. The solution was dispensed at a height of 1.6 mm above the surface. A volume of solvent 0.5µl greater than that of the amount deposited was then reaspirated into the pipette tip. One repeated mix cycle was performed before the solution was infused into the mass

spectrometer at a gas pressure of 0.3 PSI and an ionisation voltage of 1.7 kV. All experiments were performed in triplicate



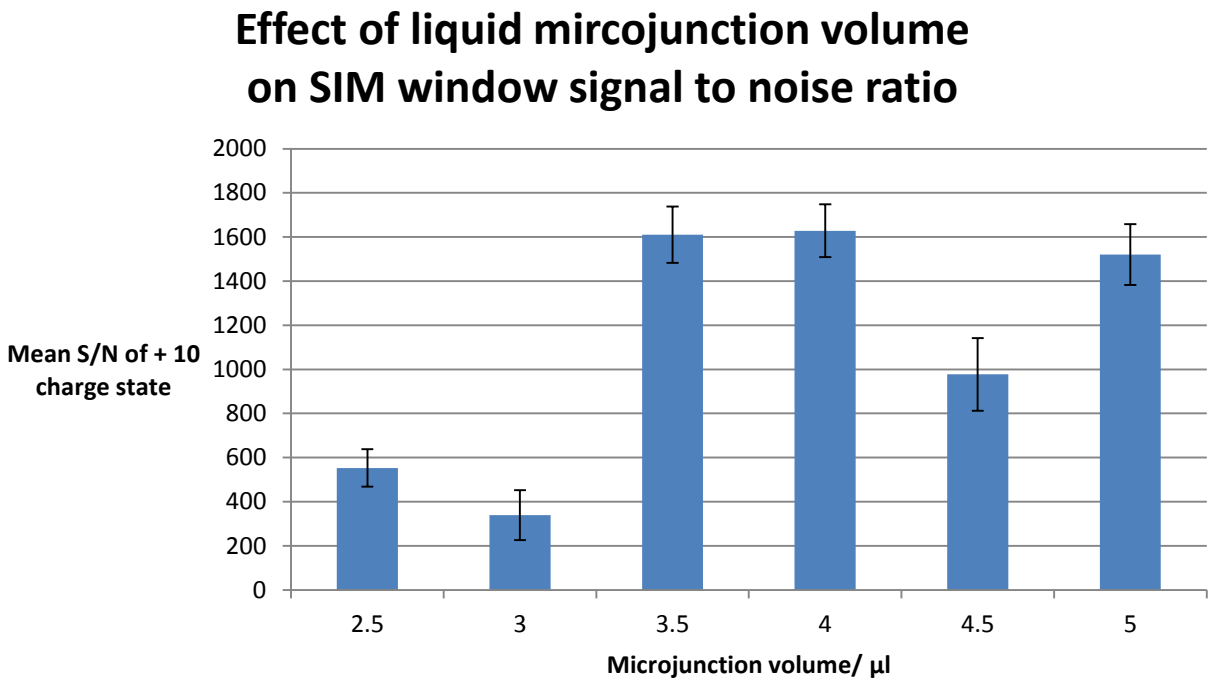
SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
2.5µl	3	65.77	21.92333	442.1056
3µl	3	359.18	119.7267	24621.11
3.5 µl	3	435.21	145.07	8006.199
4 µl	3	485.91	161.97	1209.363
4.5µl	3	328.84	109.6133	5323.698
5 µl	3	322.16	107.3867	4183.901

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	35353.43	5	7070.686	0.968888	0.47433	3.105875
Within Groups	87572.76	12	7297.73			
Total	122926.2	17				

Figure 3.12 Chart showing effect of liquid microjunction extraction volume on mean signal to noise ratio of the +10 charge state of βlg obtained over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.



SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
2.5µl	3	1106.19	368.73	116228.9
3µl	3	1017.19	339.0633	51243.77
3.5 µl	3	4830.24	1610.08	65267.37
4 µl	3	4884.73	1628.243	57694.5
4.5µl	3	2931.28	977.0933	109200.6
5 µl	3	4563.26	1521.087	76027.65

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5524147	5	1104829	13.93629	0.00012	3.105875
Within Groups	951325.6	12	79277.14			
Total	6475473	17				

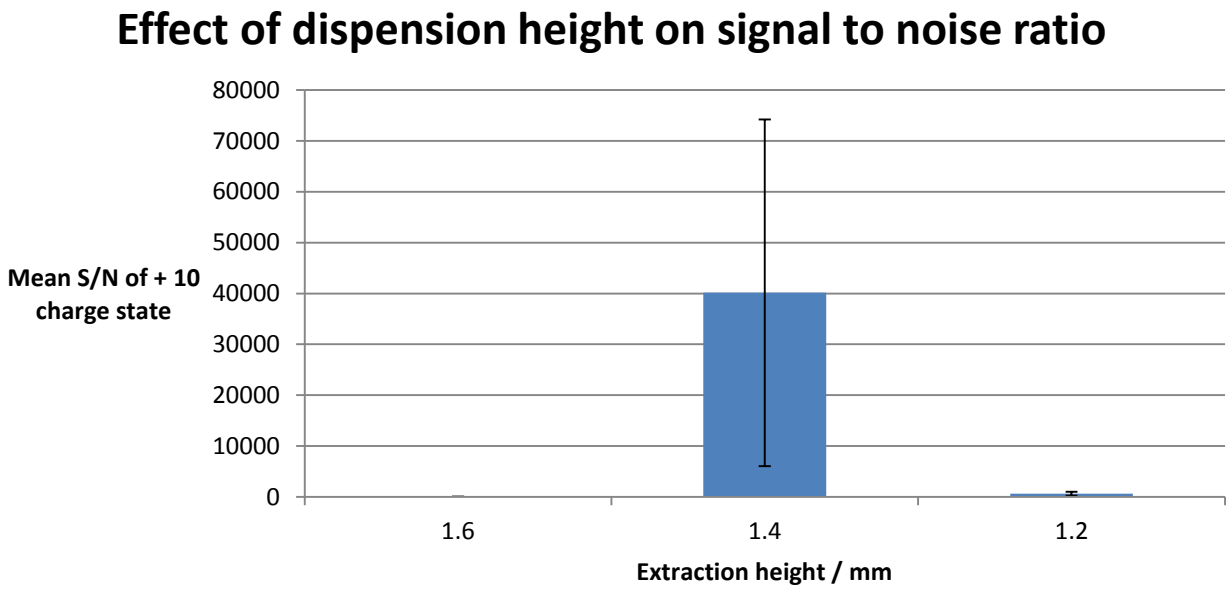
Figure 3.13 Chart showing effect of liquid microjunction volume on mean signal to noise ratio of the +10 charge state of βlg obtained from targeted SIM window over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.

Figs 3.12 and figs 3.13 show the effect of sampling volume on the signal to noise ratio of βlg on both the full scan and the targeted SIM window. The results of the one way ANOVA show that the effect of sampling volume did not have a statistically significant effect on the signal obtained from the full scan (p=0.474) but did have a statistically significant effect on the targeted SIM approach. The significance of the p value was further supported by the fact the F value (13.93) was greater than the F critical (3.10). Fig 3.13 shows that the greatest average signal to noise ratio obtained from the SIM mode mass spectrum was from liquid microjunctions with a volume of 4 μl. Larger micro junctions may be more successful than smaller microjunctions because there is a greater area of the spot can be sampled so that more protein can diffuse into the microjunction during extraction. Smaller microjunctions will also leave some of the solution remaining in the tip which never comes into contact with the surface and the protein. The reasons as to why the signal to noise ratio decreased with micro junctions with volumes of 4.5 and 5 μl are less obvious. The reason may be that a 4 μl microjunction was enough to sample the whole spot, and that larger microjunctions do not provide any additional benefit in sampling.

3.7.4 Liquid microjunction dispension height

The sampling area of a liquid microjunction can be affected by the height from which the microjunction is applied. A droplet can spread further across the surface when applied closer to the membrane o spread less when applied further away. This effect of

dispensing a microjunction at a height of 1.2, 1.4 and 1.6 mm was investigated on a series of 70 pmol/cm^2 spots. The extraction parameters were used as previously, with a dispension volume of 4 µl, an aspiration volume of 4.5 µl, a post dispense delay of 11 seconds and the inclusion of one repeat mix cycle. In this experiment the pipette dispension and aspiration heights were adjusted between 1.2 and 1.6 mm, in 0.2 mm intervals. This range was chosen because at 1.8 mm the liquid microjunction was too far above the surface for sampling and at 1.0 mm the pipette tip was pushed directly onto the surface of the PVDF. The PVDF membrane is a floppy surface that warps slightly when applied to the sampling plate, so the microjunction needed to be applied more than 1 mm above the surface. Experiments were recorded in triplicate.



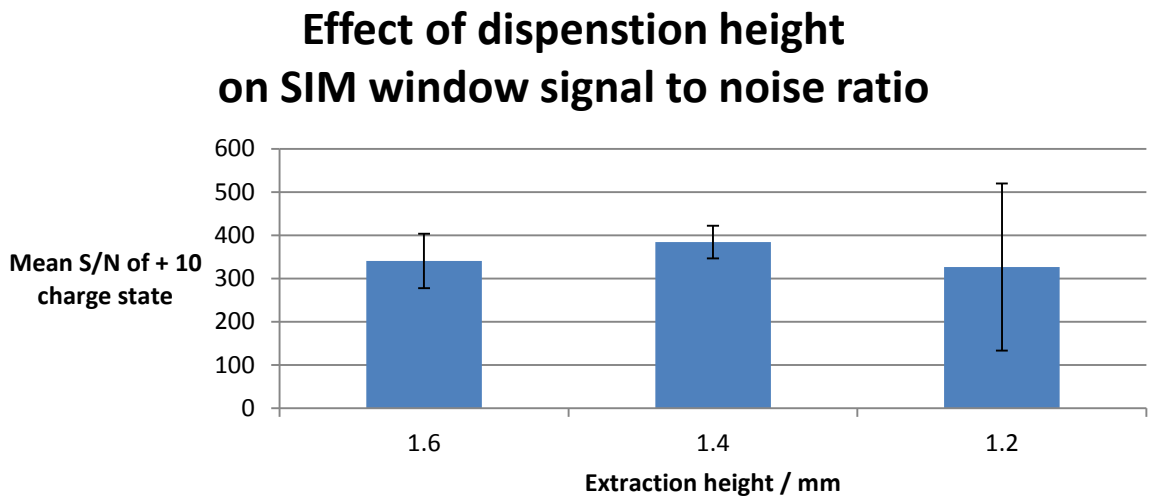
SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1.2 mm	3	160.24	53.41333	166.3297
1.4 mm	3	120411.5	40137.18	4.65E+09
1.6 mm	3	1916.86	638.9533	610944.2

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	3.17E+09	2	1.58E+09	1.020918	0.415325	5.143253
Within Groups	9.31E+09	6	1.55E+09			
Total	1.25E+10	8				

Figure 3.14 Chart showing effect of liquid microjunction dispension height on mean signal to noise ratio of the +10 charge state of βlg obtained over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.



SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1.2 mm	3	980.27	326.7567	149091.4
1.4 mm 1	3	1153.09	384.3633	5642.792
1.6 mm	3	1022	340.6667	15674.3

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5421.415	2	2710.707	0.047721	0.953758	5.143253
Within Groups	340816.9	6	56802.82			
Total	346238.3	8				

Figure 3.15 Chart showing effect of liquid microjunction dispension height on mean signal to noise ratio of the +10 charge state of βlg obtained from targeted SIM window over 3 replicates and corresponding one way ANOVA. Error bars represent 1 standard deviation.

Figs 3.14 and 3.15 show the effect of extraction/dispension height on the signal to ratio of βlg acquired with the full scan and the Targeted SIM approach. The results of the one way ANOVA show that the sampling height does not have a statistically significant effect on the signal to noise ratio of the protein on either the full scan (p=0.415) or the targeted

SIM window (p=0.95). The error bars at an extraction height of 1.4 mm on fig 3.14 are very large, which is also reflected by the variance at 1.4 mm as shown in the ANOVA. This variance is due to the fact the noise obtained from one of the three replicates was unusually low, resulting in an usually high signal to noise ratio. The exact reasons behind this are not known but this error was not reflected in the targeted SIM approach. The targeted SIM approach is more sensitive and therefore more appropriate for detecting small amounts of known proteins. Although the effect of dispension height did not have a statistically significant effect the greatest average signal to noise ratio was obtained when the microjunction was dispensed at a height of 1.4 mm which was carried forward in future experiments. Applying the liquid microjunction closer to the surface of the membrane ensures that it spreads across a larger area, ensuring more of the spot is sampled. This may explain why the signal to noise ratio was higher at 1.4 mm than it was at 1.6. There was a small fall in signal to noise ratio when sampling was decreased from 1.4 to 1.2 mm above the surface of the membrane. This decrease is less easy to explain, but may be simply that dispensing the microjunction at 1.4 mm sufficiently covered the whole spot and that dispensing the sample below this provided no additional benefit.

3.7.5 Determination of limit of detection limit for LESA analysis of βlg from PVDF membranes

Having determined the optimum extraction parameters for LESA surface sampling parameters, a PVDF membrane with spots containing a serial dilution of different βlg

concentrations was analysed. Concentrations ranged from 1-70 pmol/cm^2 at 10 pmol/cm^2 intervals plus additional samples at 45 and 35 pmol/cm^2. Samples were analysed by LESA using the optimum extraction parameters determined by the experiments above, see table 3.3. Experiments were performed in triplicate.

Sample volume	7 µl
Dispense volume	4 µl
Aspiration volume	4.5 µl
Post dispense delay	11 seconds
Repeat mix cycles	1
Aspiration/dispense height	1.4 mm

Table 3.3 Optimum LESA parameters determined by previous experiments

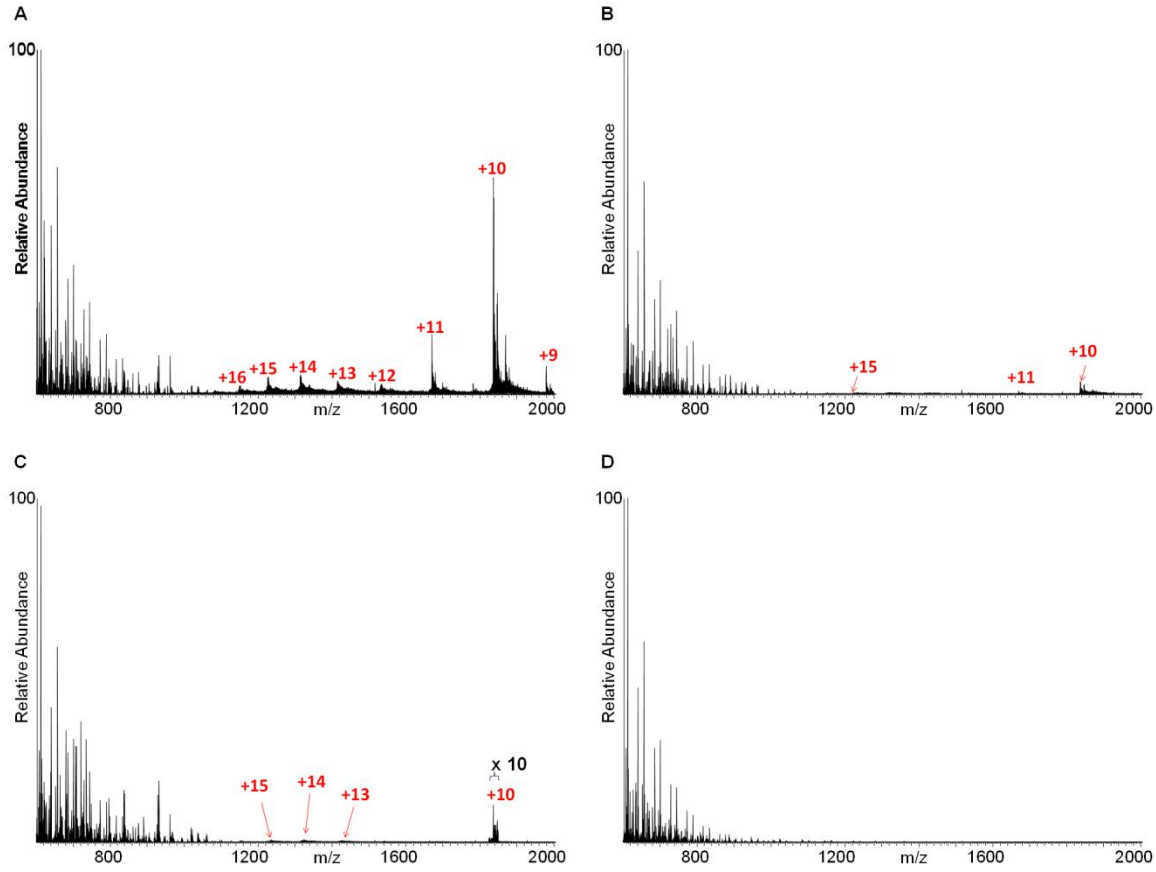


Figure 3.16 LESA mass spectra of βlg spots at different concentrations: (A) 70 pmol/cm² (B) 45 pmol/cm² (C) 40 pmol/cm² (D) 35 pmol/ cm²

	Signal to noise ratio of peak							
Sample	+15	+14	+13	+12	+11	+10	+10 Isolation	+10 SIM
70 pmol/cm^2 1	7.35	7.35	4.78		18.75	66.9	85.74	581.74
70 pmol/cm^2 2	6.33	6.78	5.03		53.29	216.27	4.16	167.66
70 pmol/cm^2 3	7.64	7.52	5.47		17.57	56.27	72.5	3085.05
60 pmol/cm^2 1	6.14	6.61	4.24		10.55	38.28	28.17	2650.8
60 pmol/cm^2 2	2.91	4.14	3.87		331.92	10106.67	-	73.87
60 pmol/cm^2 3	3.65	3.74	2.03		8.6	44.69	-	178.06
50 pmol/cm^2 1	2.78	3.07	2.71		32.63	102.06	-	81.5
50 pmol/cm^2 2	-	-	-	-	5.19	25.6	-	66.34
50 pmol/cm^2 3	2.96	3.43	3.66		422.1	2617.57	-	78.27
45 pmol/cm^2 1	-	-	-	-	4.24	16.44	-	122.93
45 pmol/cm^2 2	-	5.37	6.86	4.49	3.56	18.05	-	54.1
45 pmol/cm^2 3	-	2.77	-	-	-	10.52	-	47.91
40 pmol/cm^2 1	-	-	-	-	-	-	-	-
40 pmol/cm^2 2	3.79	3.81	3.48	-	-	20	-	6.68
40 pmol/cm^2 3	-	-	1.9	-	-	4.48	-	-
35 pmol/cm^2 1	-	-	-	-	-	-	-	-
35 pmol/cm^2 2	-	-	-	-	-	-	-	-
35 pmol/cm^2 3	-	-	-	-	-	-	-	-
30 pmol/cm^2 1	-	-	-	-	-	-	-	-
30 pmol/cm^2 2	-	-	-	-	-	-	-	-
30 pmol/cm^2 3	-	-	-	-	-	-	-	-
20 pmol/cm^2 1	-	-	-	-	-	-	-	-
20 pmol/cm^2 2	-	-	-	-	-	-	-	-
20 pmol/cm^2 3	-	-	-	-	-	-	-	-
10 pmol/cm^2 1	-	-	-	-	-	-	-	-
10 pmol/cm^2 2	-	-	-	-	-	-	-	-
10 pmol/cm^2 3	-	-	-	-	-	-	-	-
1 pmol/cm^2 1	-	-	-	-	-	-	-	-
1 pmol/cm^2 2	-	-	-	-	-	-	-	-
1 pmol/cm^2 3	-	-	-	-	-	-	-	-

Table 3.4 Signal to noise ratios of peaks corresponding to β lactoglobulin from protein spots of various concentrations Results shown for individual replicates.

Fig 3.16 and table 3.4 show that with the optimised LESA extraction parameters, βlg was detected down to a concentration level of 40 pmol/cm^2. At this concentration, the proteins was only observed in one of three replicates and fig 3.16 C shows that the signal

is very weak. To ensure a consistent and reproducible signal, at least 45 pmol/cm^2 is needed on the surface.

3.8 Conclusion

The work presented in this chapter demonstrates that the analysis of intact proteins from PVDF membranes is possible by use of LESA MS. The initial aim was to develop a new approach for the purification of intact proteins from PVDF membranes following electroblotting from SDS PAGE. The approach was not able to recover proteins blotted from a gel but it was possible to extract intact proteins directly applied to the surface of the membrane. Both SDS PAGE and native PAGE, were attempted which suggests that the problem is due to an inherent problem caused during the blotting process rather than as a result of signal suppression by SDS anions. The second half of the chapter focused on determining the optimum LESA parameters for extracting proteins from membranes and investigating the limits of detection, with a view to application in air monitoring. The limit of detection for b lactoglobulin was shown to be 40 pmol/cm ^{2}.

Chapter 4 Bottom-up proteomics analysis of dried blood spots by liquid extraction surface analysis mass spectrometry

4.1 Introduction

Prior to the work contained in this thesis, LESA had been used effectively for the analysis of intact globin proteins from DBS. The technique was demonstrated for use in the clinic for determining the presence of newborn haemoglobin variants[2]. The approach focused on the analysis of haemoglobins and no other proteins had been identified from DBS. Blood is a complex biofluid composed of thousands of different proteins and a top-down approach may be unsuitable for identifying other proteins. As mentioned in section 1.3.5 the analysis of intact proteins is usually used as a tool for the characterisation of particular proteins of interest and is less successful in untargeted protein identification. The aim of the work presented in this chapter was to incorporate a bottom-up proteomics strategy for the identification of proteins from DBS by LESA in order to identify proteins other than the highly abundant haemoglobins. In doing so, the methodology must maintain the automated nature of LESA and reduce any required manual intervention as much as possible. Part of this work presented here was published as an article in the Journal of the American Society for Mass Spectrometry [204].

4.2 LESA of dried blood and plasma spots.

4.2.1 Dried blood spots

To confirm that no proteins other than haemoglobin could be detected from a DBS by using LESA, the methods described by Edwards *et al.* [2] were repeated. Seven microlitres of a solution composted of 48.5/48.5/3% methanol/water/ formic acid was aspirated from the LESA solvent well, and 6 microlitres were dispensed onto the spot at a height of 1.6 mm above the surface. Five microlitres were reaspirated after a delay of 5 seconds. The aspiration and dispense cycle was repeated once before the solution was infused by electrospray into the mass spectrometer. Samples were analysed for a period of 8 minutes in full scan mode. Each scan comprised 30 microscans.

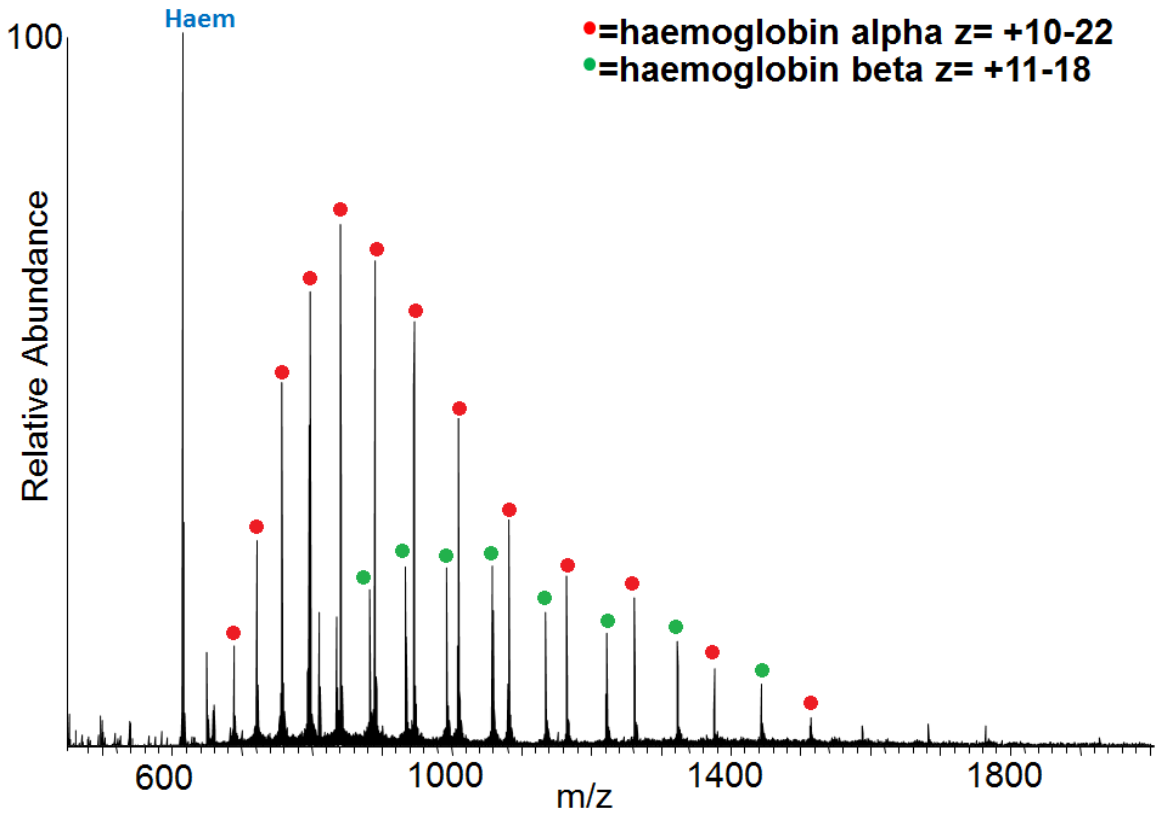


Figure 4.1 LESA mass spectrum obtained from DBS on filter paper.

The mass spectrum shown in fig 4.1 reveals peaks corresponding to alpha globin, beta globin and the free haem peak. No other proteins were identified. Haemoglobin is by far the most abundant protein in blood, with a concentration of approximately 130 mg/ml for a healthy adult male and 120 mg/ml for a healthy adult female [205]. The concentration of haemoglobin is significantly higher than any other protein: the next most abundant protein is albumin which has a concentration of 50 mg/ml [206]. The lack of signal for any other proteins could be down to two factors: Either the top-down method could not cope with the dynamic range between haemoglobin and the next most

abundant proteins, or that the remaining proteins are too large to be detected by top-down methods with the resolution available on the orbitrap.

4.2.2 Dried plasma spots

To investigate if any other blood proteins could be detected by a top-down technique, a dried plasma spot (DPS) was analysed with LESA. Plasma is the liquid portion of the blood which is depleted of red blood cells and haemoglobin so its analysis does not face the problem of an overwhelming haemoglobin concentration. Ordinary DBS cards could not be used for LESA analysis of plasma. They are suitable for DBS analysis because the blood clots to form a matrix on top of the filter paper. Without this matrix, the liquid microjunction will be immediately absorbed into the filter paper rendering surface sampling impossible. PVDF was used as an alternative to filter paper for DPS as sampling of intact proteins had already been demonstrated on this surface, as described in Chapter 3. Twenty microlitre aliquots of plasma were spotted onto PVDF at marked positions, left to dry and analysed with LESA using the same method applied for analysis of DBS with the exception that the extraction solvent was composed of 69.3/29.7/1% water/methanol /formic acid and that each scan was composed of 50 microscans.

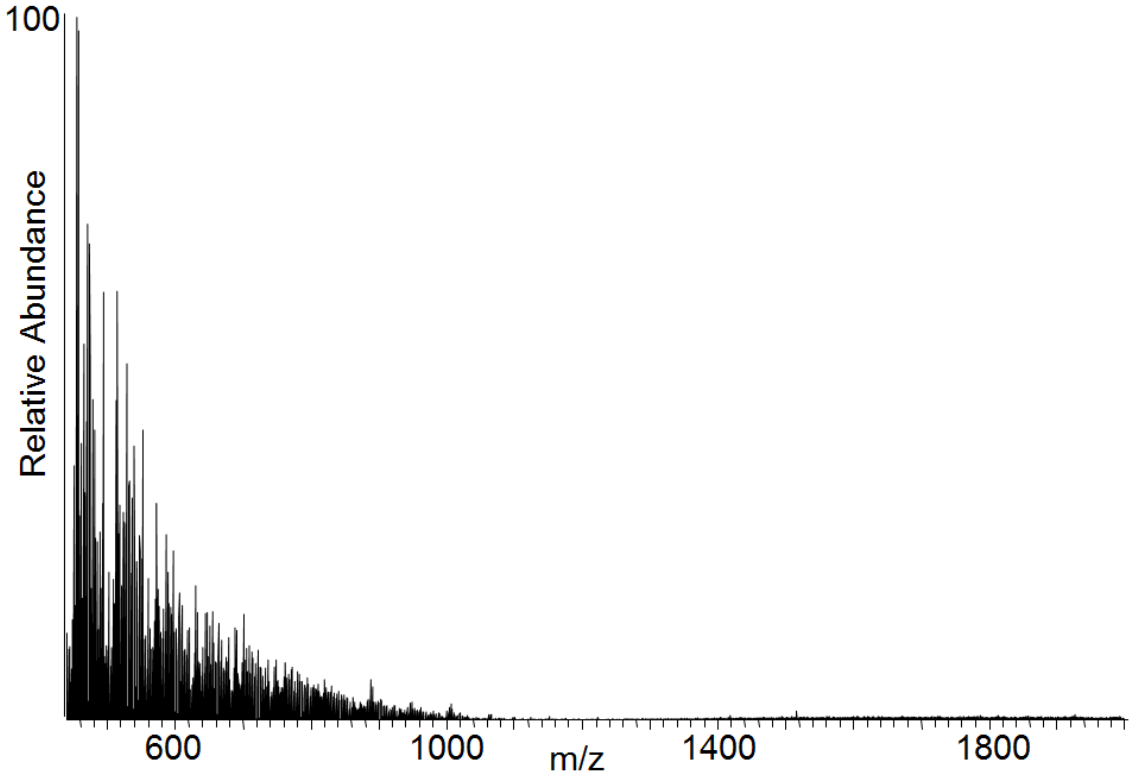


Figure 4.2 LESA mass spectrum obtained from a dried plasma spot on PVDF.

The mass spectrum in fig 4.2 shows that no intact proteins could be identified after LESA analysis of a dried plasma spot. This result suggests that even if the issue of haemoglobin abundance could be overcome, the top down method is not a viable option for identifying other proteins from a DBS.

4.3 Bottom-up proteomics

The results above suggest that top down proteomics is not viable for identifying other proteins from DBS, therefore, bottom up proteomics was investigated as an alternative. An initial experiment on a DPS was carried out to investigate the potential for protein

identification. Mass spectrometry had been used for the targeted identification of particular proteins from DBS previously [171, 174]. In both of these analyses, a small punch was cut out of a DBS, the proteins were eluted into solution which was then digested with trypsin. A modification of the procedure described in [174] was carried out on a DPS. Twenty microlitres of human plasma was spotted onto a filter paper dried blood spot card. After drying, 6 mm punches were cut out and the proteins were eluted from the card. Proteins were eluted into 75 µl 50 mM NH_4 HCO_3 10% methanol for 2 hours on a thermomixer heated to 40˚C shaking at 800 rpm. After elution, a Bradford assay was carried out to measure the protein concentration. A 40 µg aliquot of protein was taken and 5µg of trypsin was added to create a 1:20 ratio of enzyme to protein. Samples were digested for periods of 1 hr and 12 hr. Proteolysis was quenched by the addition of 10% formic acid and acetonitrile. The digests were diluted tenfold in 49.5/49.5/1% water/methanol/formic acid and directly infused into the mass spectrometer. Samples were analysed by a 5 minute direct infusion top 7 CID method described in section 2.2.3.2 and database searching was carried out as described in section 2.2.8.

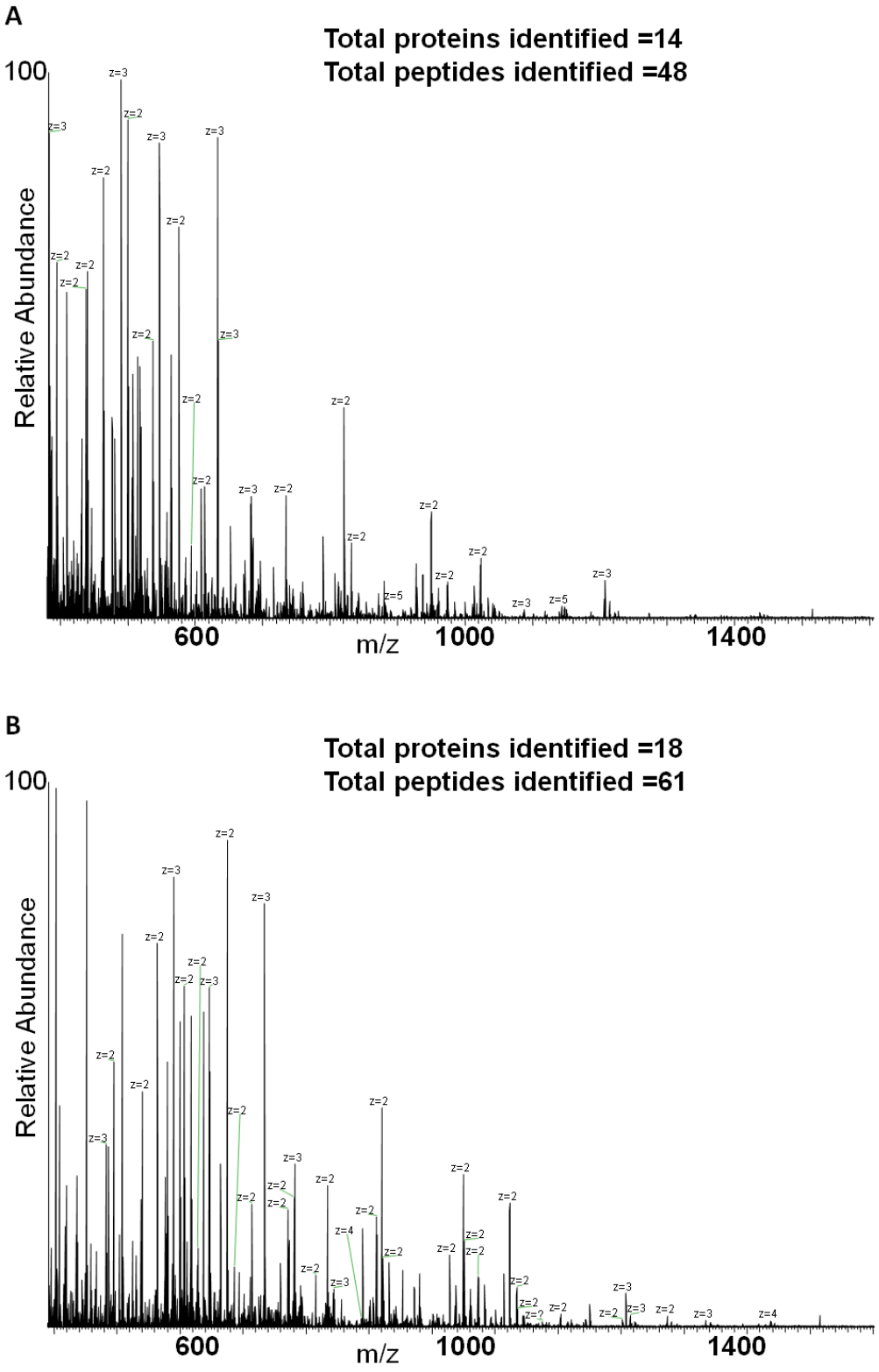


Figure 4.3 Direct infusion ESI analysis of dried plasma spot digest: (A) 1 h digest (B) 12 h digest.

Multiple proteins were successfully identified by direct infusion analysis of a tryptic digestion of a dried plasma spot as shown on fig 4.3. The full list of proteins identified from the 2 digests are shown in appendix 4.1 and 4.2. The proteins identified were mostly high abundance plasma proteins such as albumin and immunoglobulins. That result is perhaps to be expected as the method relied of direct infusion ESI analysis without HPLC separation. This result demonstrates that bottom up protein identification could be suitable for identifying other proteins from DBS. Future efforts were aimed at incorporating a DBS digestion procedure that is suitable for analysis with LESA.

4.3.1 DBS proteolysis

There are three possible means of incorporating a digestion stage into the DBS-LESA sampling routine. The blood could be digested before it is applied to the card, whilst the blood is on the card, or after the proteins have been extracted and before they are introduced into the mass spectrometer. The first option of digesting the blood before it is applied to the card is not a viable option as it would negate the advantages of DBS sampling such as the ability to sample blood straight from a finger prick. To digest the blood in this fashion would require the blood to be collected via phlebotomy and digested in solution before spotting. In which case there would be no added advantage of applying it to the DBS card. An on card digestion is a viable option as it would maintain compatibility with finger prick sampling, but would be difficult to maintain compatibility with LESA sampling. Filter paper DBS cards are absorbent so the digestion process must proteolytically cleave proteins on the DBS surface but not degrade the clot matrix

required for LESA sampling. The final option would be to digest the proteins after the sampling process. This option is feasible but would include extra time demands in the sampling procedure. The second and third approaches were attempted on a series of DBS with the aim of identifying multiple proteins.

The two different digestion protocols were carried out manually with the emphasis on methods that could subsequently be performed by a robotic sequence carried out on the Triversa Nanomate™.

4.3.2 On card proteolysis

In this procedure, trypsin solutions were applied to the surface of the DBS which must be kept wet to permit proteolysis. The procedure was carried out on DBS placed on a heating block at 37˚C. A 20 µl aliquot of methanol was applied to denature the proteins prior to digestion. The continual heating meant that applied solvents slowly evaporated. After applying the methanol, twenty microliters of 0.5 µg/µl trypsin in 25 mM acetic acid and 25 mM NH_4HCO_3 were applied to the surface of the spot. After 10 minutes, this solution had begun to evaporate and an additional 20 µl was added. A further ten minutes followed before this solution had dried and from there on 20 µl aliquots of 50 mM NH_4HCO_3 were added at 10 minute intervals for a period of 4 hours. After digestion the peptides were eluted from the spot. LESA sampling was attempted but the surface matrix had been degraded too much to permit the formation of a liquid microjunction. The elution process was carried out by cutting out a 5 mm punch and eluting it in 75 µl 10% methanol for 2 hours on a thermomixer operating at 800 rpm at 40˚C. After elution, digestion was inhibited by the addition of 15µl 10% formic acid and 20µl acetonitrile.

4.3.3 Surface sampling followed by proteolysis

In this approach, intact proteins were extracted from the surface of the DBS and subsequently digested in solution. The surface extraction was performed by a hand-held pipette. Seven microliters of a 50 mm NH_4HCO_3 solution was aspirated into the pipette tip and approximately half was applied to the surface of a DBS. The liquid microjunction was held in contact with the surface for a period of 5 seconds before reaspiration. The sample was placed in an Eppendorf tube and 1 µg of trypsin was added. The mixture was incubated in a thermomixer for 1 hour at 37˚C whilst being shaken at 800 RPM. After 1 hour, the digestion was inhibited by addition of 1.5 µl 10% formic acid and 2 µl acetonitrile prior to direct infusion analysis.

Both samples were then treated to a tenfold dilution with 49.5/49.5/1% water/methanol/formic acid and analysed by a 5 minute direct infusion top 7 CID method described in section 2.2.3.2 and searched as described in section 2.2.8.

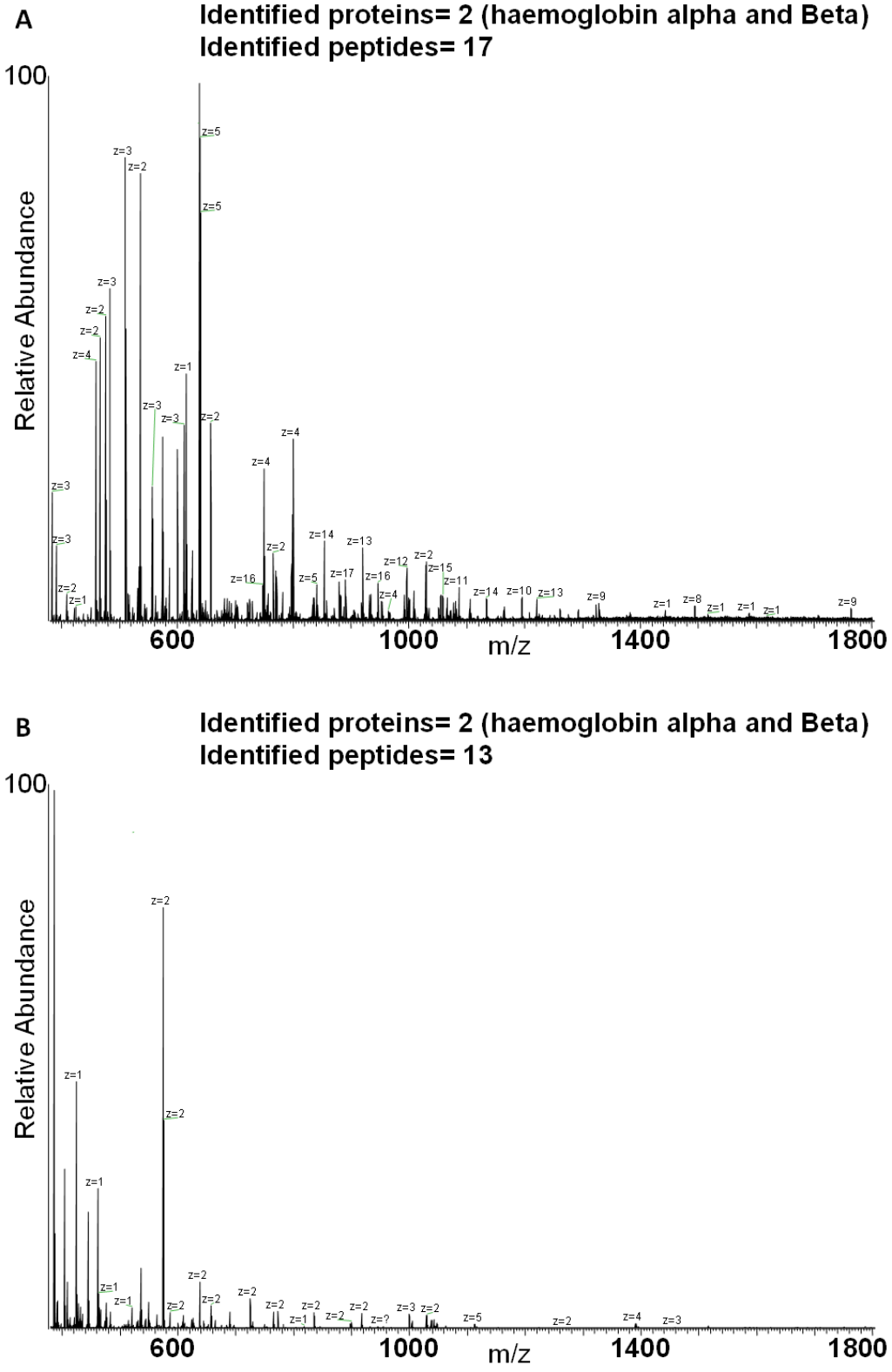


Figure 4.4 Direct infusion mass spectra obtained following (A) on card digestion of DBS (B) LESA sampling and subsequent digestion

Fig 4.4 shows that both on-card proteolysis and proteolysis following LESA extraction yielded tryptic peptides; however, only alpha globin and beta globin could be identified. The mass spectrum obtained following on card digestion shows peaks corresponding to undigested intact proteins, with charges of +16, +15 and +14, which can be identified as haemoglobin alpha and beta. The limiting factor is likely the high abundance of haemoglobin. LC MS/MS was therefore employed as a means of peptide fractionation with the aim of identifying more proteins. Digests were prepared as described previously, however prior to analysis samples were purified with C_{18} zip tips as described in section 2.2.5 and analysed with the LC MS/MS method described in section 2.2.2.

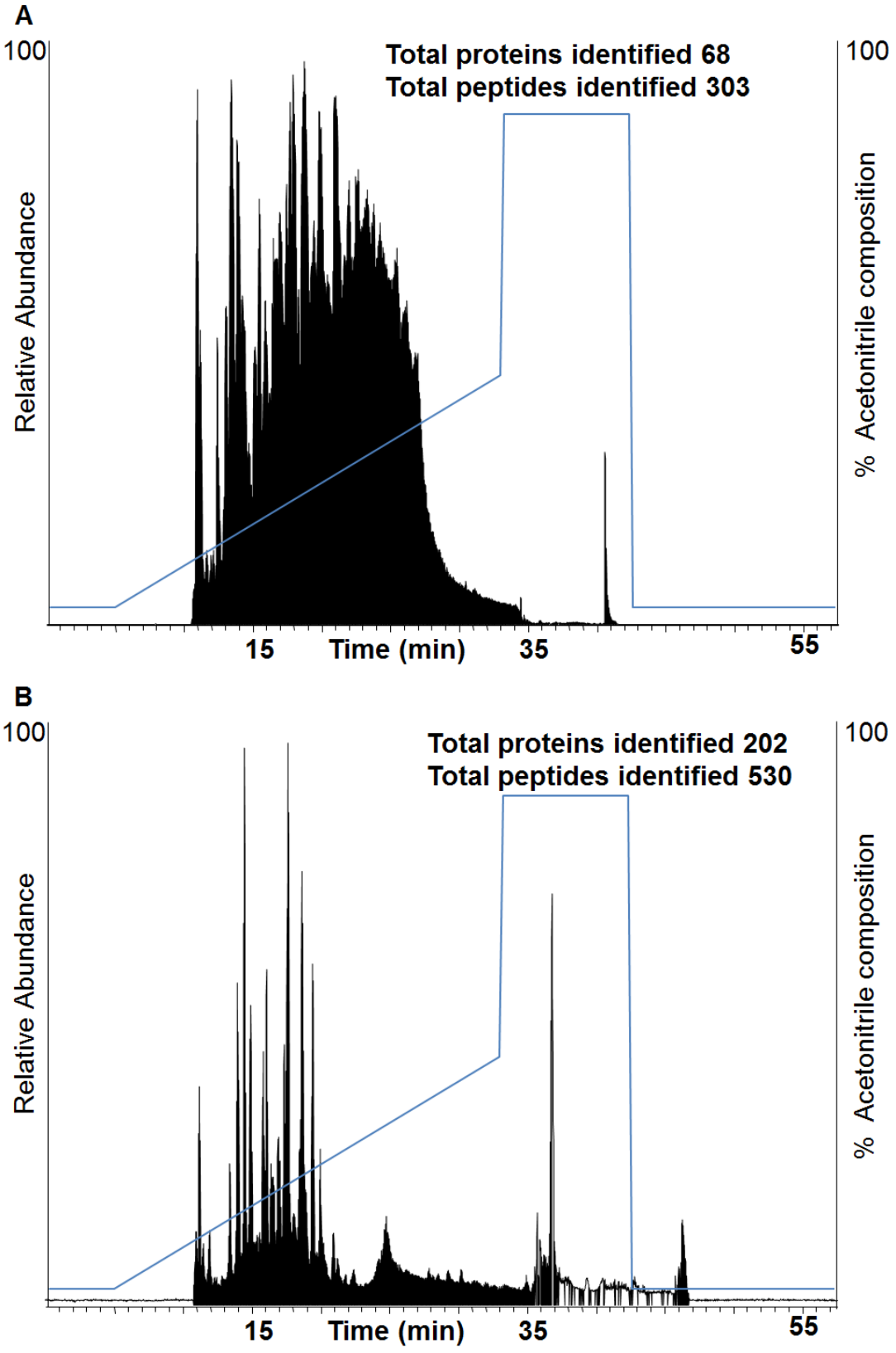


Figure 4.5 Total ion current chromatograms showing acetonitrile composition obtained following LC MS/MS analysis of DBS digest: (A) on card proteolysis (B) LESA extraction followed by proteolysis

Fig 4.5 shows the total ion chromatograms (TICs) of both the on card and the surface sampling based digestions. Sixty eight different proteins were successfully identified from the on card digest and 202 proteins were identified from the surface sampling based digest. The full list of all the proteins identified can be seen in appendix 4.3 and 4.4. The on card digestion identified fewer proteins than the surface sampling and solution phase digestion. This digest protocol was also more compatible with LESA sampling so was selected for further development.

4.4 Adaption for Automation

The manual surface sampling and solution phase digestion procedure was adapted for execution on the Triversa Nanomate™. There were four features of the manual procedure described in section 4.3 that needed to be adapted: surface sampling, incubation of the proteolysis mixture, clean-up of the digestion products and finally transfer of the samples to the HPLC unit. The sampling and incubation were performed by using a sequence written with the AUI feature of the Triversa Nanomate™. The whole process was performed on a 96 well microtitre plate: DBS were loaded onto the surface of a wellplate, with a separate well acting as a solvent reservoir. The digestion/extraction solvent was aspirated from the solvent well, a LESA extraction process was performed on the DBS surface and the solution containing the intact proteins was then dispensed into an empty well which acted as a reaction vessel. The extraction and digestion solvent was composed of 50 mM NH_4HCO_3. In each extraction process 7 µl of solvent was aspirated

from the solvent well and 6 µl were applied to the surface of the DBS. Five microlitres was reaspirated and the sample was dispensed into the reaction well. After extraction, the previous pipette tip was ejected to prevent carryover, before 4.5 µl of 0.1 µg/µl trypsin was aspirated from the trypsin well and added to the reaction well for the incubation to begin. The sample was mixed once by an aspiration and dispense cycle of 4 µl. The Triversa Nanomate™ contains a temperature control unit that can heat or cool the plate to any desired temperature between 4 and 60°C. In this experiment, the temperature control unit was used to heat the plate to 40˚C to incubate the enzyme. In the manual procedure described in section 4.3, the sample was digested on a thermomixer to mix the sample. This approach would not be a viable option in the automated procedure as the Triversa Nanomate™ does not feature a shaker. As an alternative, the sample was mixed at regular intervals by performing an aspiration and dispense cycle with the pipettor. Performing an aspiration and dispense cycle mixes the sample at intermittent periods but does not create the constant agitation that was provided in the manual digestion. The aspiration and dispense cycle also proved useful because it was noticed that as the plate was being heated the small amounts of solvent in the well were prone to evaporation. These mixing stages also allowed additional solvent to be added to prevent the sample from drying out. Seven and a half microlitres of 50 mM NH_4HCO_3 was added from the solvent well to the reaction wells at 30 minute intervals and mixed with an aspiration and dispense cycle of 4 µl.

Automation of steps 3 and 4 was less simple. The manual procedure used C_{18} zip tips to remove salts and contaminants produced during the digestion. These contaminants need to be removed prior to LC MS analysis, but zip tipping requires manual intervention. To

prevent this, an online desalting procedure was incorporated into the LC MS/MS analysis (described in detail section 2.2.2). The procedure involved loading the sample onto a trapping column that was rinsed through with 0.1% formic acid prior to being switched in line with an analytical column. The wash cycle helped remove salts and contaminants before the valve switched and the analytical acetonitrile gradient was applied to elute the peptides. It is analogous to the effect of zip tipping, but is performed online during the LC process to eliminate user intervention. As a result it does however increase LC MS/MS analysis time by 5 minutes. Transferring the plate in step 4 was the only issue that could not be overcome. Combining injection into the LC unit via the Triversa Nanomate™ was not a viable option. Transferring the plate remained the only manual intervention remaining in the procedure. The samples were digested in an automated fashion on the Triversa Nanomate™ and then transferred to the HPLC autosampler for analysis. It can be carried out in a batch process where multiple samples are digested in parallel by using the nanomate™ and then the plate containing multiple samples can be transferred to the HPLC unit for analysis. To demonstrate multiplexing, the process was performed on a series of 3 DBS that were digested in parallel for a period of 1 hour as described above. Data was searched as described in section 2.2.8. A diagram illustrating the sampling routine that was developed is shown in fig 4.6 and is described by a flow chart in fig 4.7.

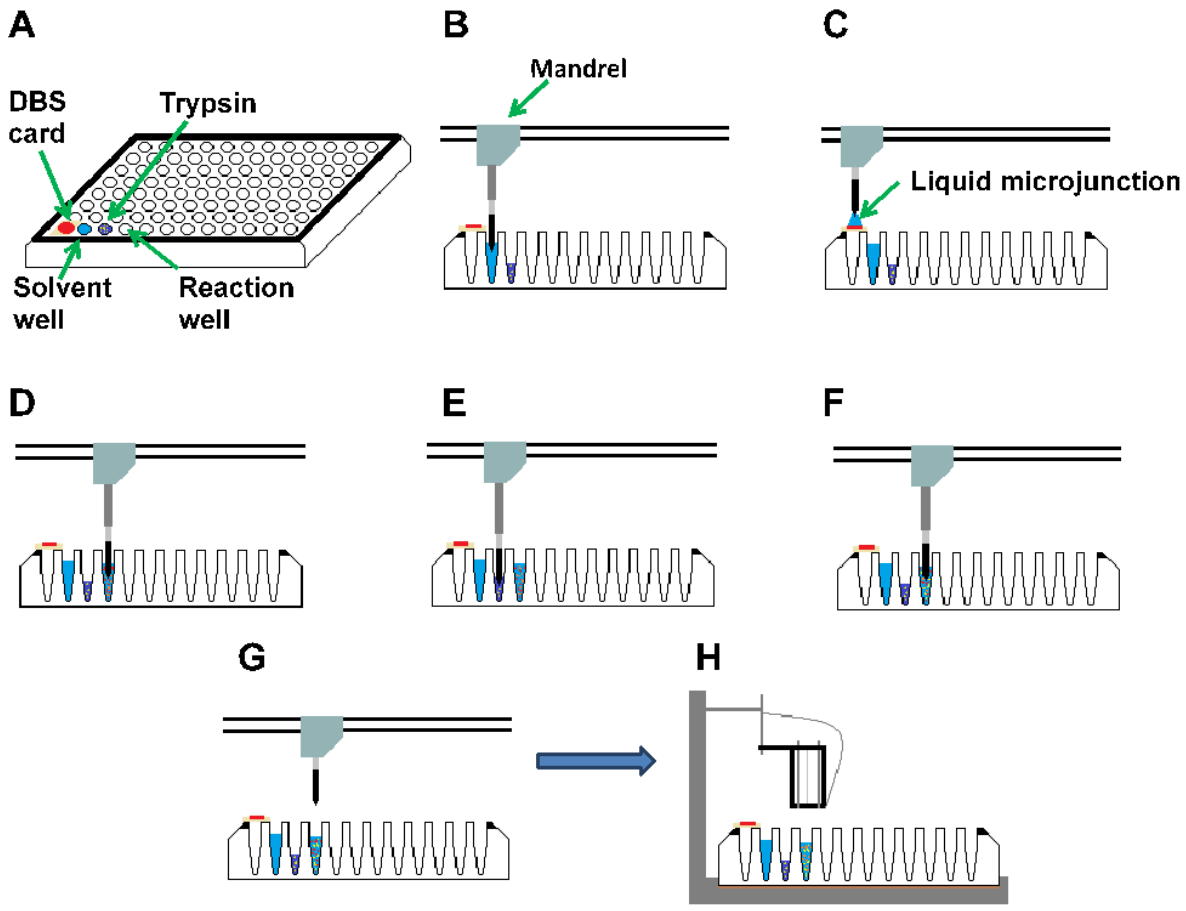
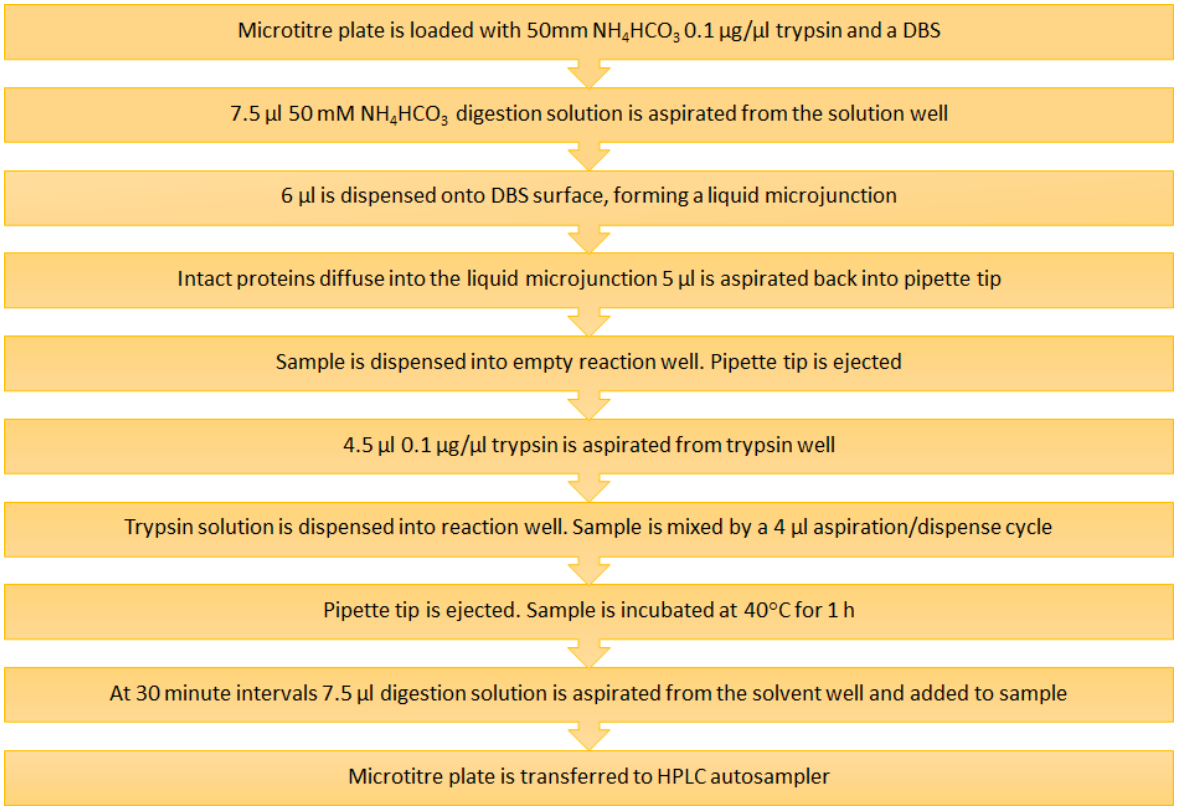


Figure 4.6 Illustration of the digest procedure adapted for LESA sampling (A) the 96 well plate is placed in the autosampler of the triversa nanomate™. The temperature control unit heats the plate to 40˚C. A DBS card is loaded onto the surface of the well, a solvent well is filled with 50 mM NH_4HCO_3 and a trypsin well is filled with 0.1 µg/µl trypsin. Remaining wells are left blank to act as reaction vessels. (B) Seven microlitres of solvent is aspirated from the solvent well. (C) Six microlitres of solvent is applied to the surface of the DBS forming a liquid microjunction and allowing the diffusion of intact proteins from the surface of the spot into the solvent. (D) Five microlitres of solvent was reaspirated from the spot and the whole sample was then deposited into one of the reaction wells. (E) After the tip has been ejected and replaced 4.5 µl of 50 mM NH_4HCO_3 was aspirated from the trypsin well (F) trypsin was deposited into sample reaction well and digestion begins. (G). Sample is left to digest for 1 hour, additional 7.5 µl of 50 mM NH_4HCO_3 is added to reaction well to account for solvent evaporation at 30 minute intervals (H) sample is transferred to HPLC autosampler for LC MS analysis.



4.7 Flow chart showing the aspiration and dispense cycles for the digestion of a single DBS

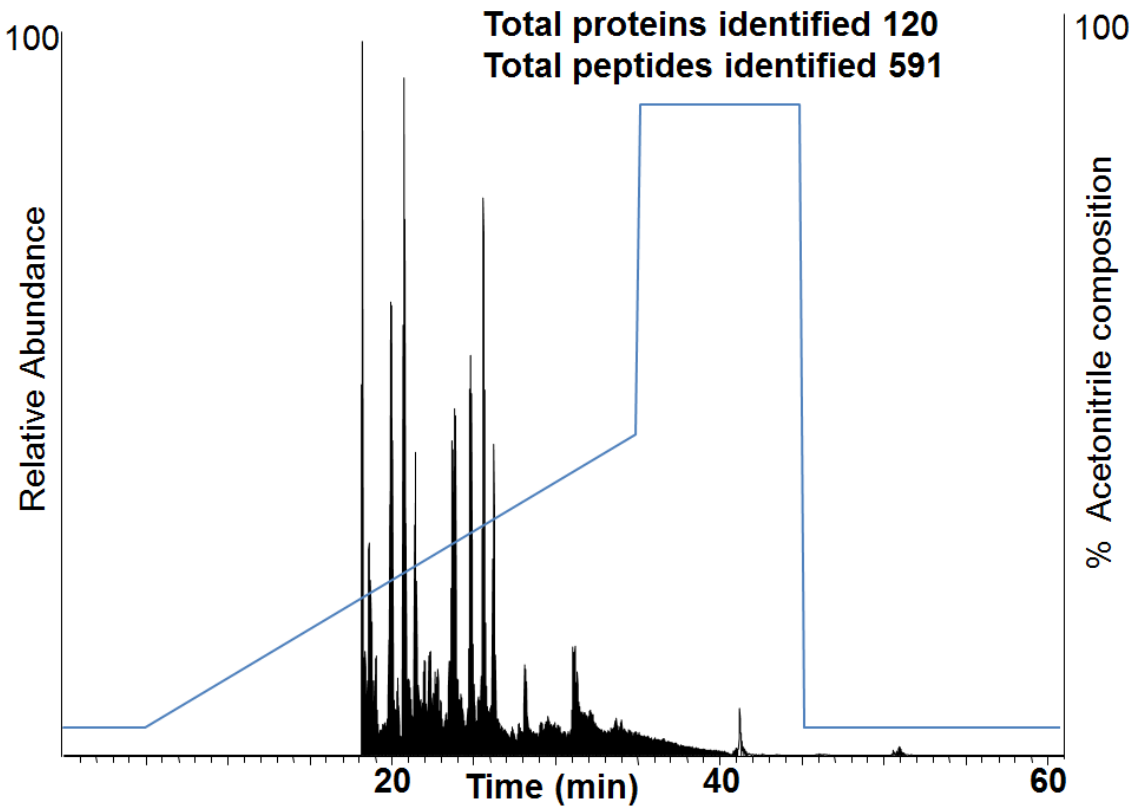


Figure 4.8 Total ion current chromatogram showing acetonitrile gradient obtained from LC MS/MS of DBS digest prepared by automated LESA sampling and trypsin digest.

Fig 4.8 shows the TIC of an automated digestion of a DBS performed by surface sampling and subsequent LESA analysis. The full list of non-redundant proteins identified across the 3 replicates is shown in appendix 4.5. These results demonstrate that the LESA coupled with proteolysis can be successfully adapted onto an automated platform. In the 3 replicates used demonstrate the success of the automated digest, 120, 115 and 107 proteins were identified after LC MS/MS analysis. Previously reported examples of protein analysis of DBS have all been targeted to particular proteins of interest [167, 171, 174, 176]. The great advantage of this untargeted approach is that the analysis is not limited to a few proteins of interest and as demonstrated here, it becomes possible to

identify over 100 proteins per LC MS/MS analysis, compared to a previous record of 37 [176].

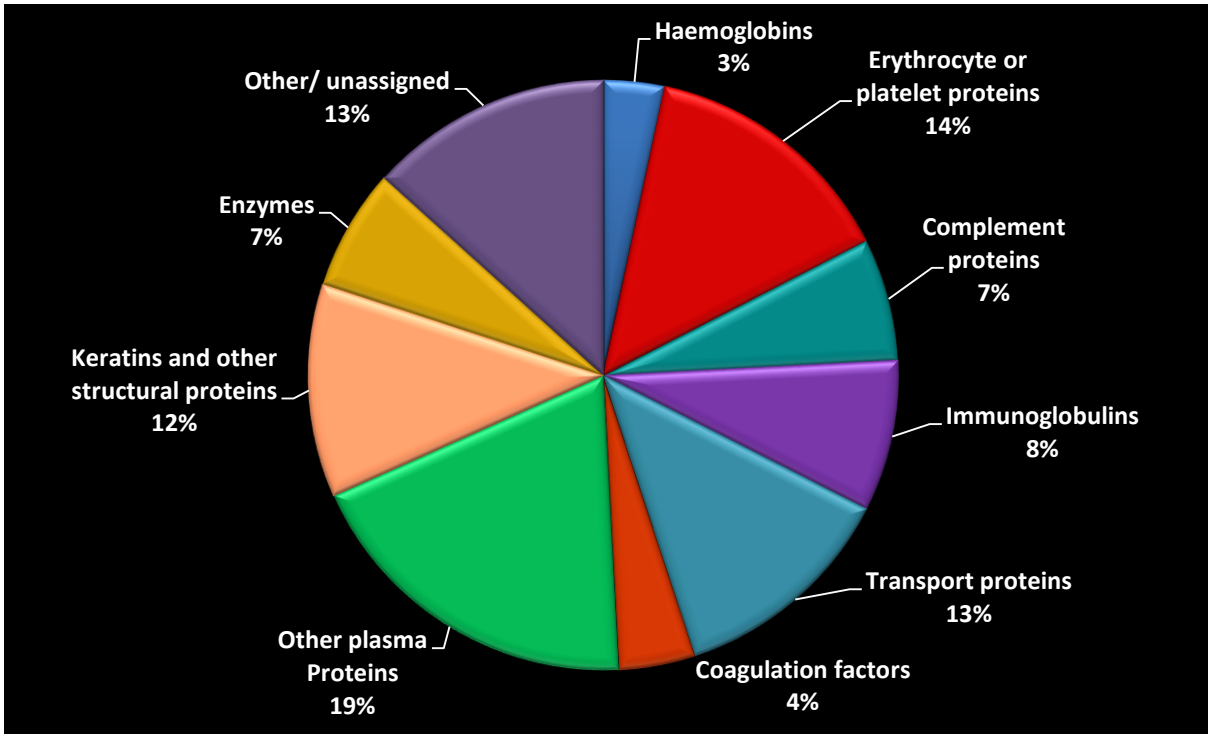


Figure 4.9 Categorisation of different proteins identified from the DBS.

Fig 4.9 shows the different categories of proteins that were identified from the DBS. The analysis identified a number of different types of plasma proteins. Approximately 50% of all the proteins identified were plasma proteins. A number proteins originated from erythrocytes or platelets. The analysis also identified various different structural proteins and a number of proteins whose cellular origin was unconfirmed. Many proteins are known to leak into the plasma from the surrounding tissues. Tissue leakage takes place primarily in the capillary region and these proteins account for some of the less abundant

proteins in human plasma [170, 207]. DBS proteomics also faces the challenge that the blood is acquired with a finger prick, which releases skin cells onto the card. These epithelial cells could be the source for some of the proteins identified in this analysis. The least abundant protein identified in this analysis was protein S100 A9. In healthy human plasma this protein is expressed at a concentration level of a 0.05 µM [208] which has a concentration 4 orders of magnitude less than serum albumin which is the most abundant plasma protein. The CID MS/MS spectrum of a peptide from this protein is shown in fig 4.10.

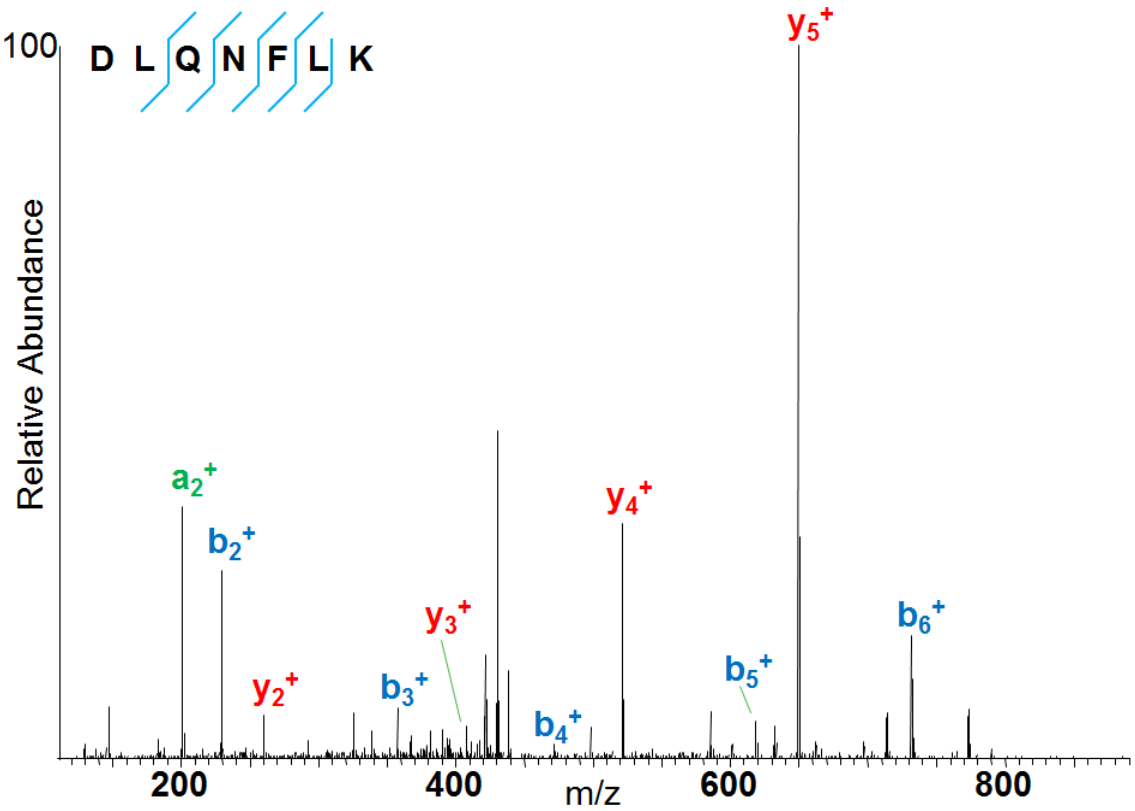


Figure 4.10 CID mass spectrum obtained from 2+ ions of peptide DLQNFLK. The peptide is assigned to protein S100 A9, the least abundant protein identified in this analysis.

The reproducibility of this LESA-based DBS proteomics approach is shown on the Venn diagram in fig 4.11. Three replicate analyses resulted in the identification of 120, 115 and 107 proteins. Sixty four proteins identified were common to all 3 replicates and approximately 80 proteins were common to any 2 replicates. A reproducibility of 64 proteins out of 120 proteins over 3 replicates may not seem particularly high, however poor reproducibility is a common criticism of data dependent acquisition techniques in proteomics[123].

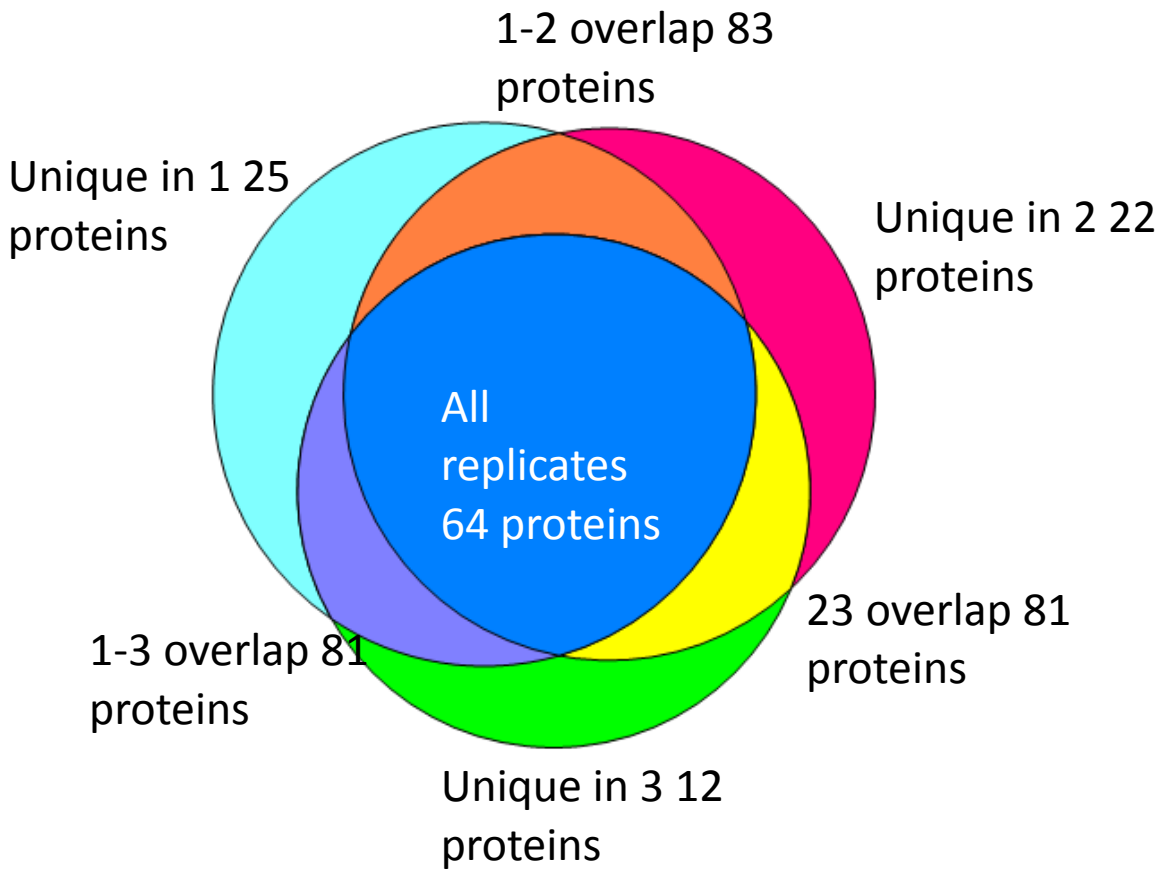


Figure 4.11 Reproducibility of protein identifications over 3 replicates

4.4.2 Effect of proteolysis times

The work presented above demonstrated that LESA based bottom-up proteomics of DBS can be successfully adapted onto an automated platform, and that the procedure is capable of identifying approximately 100 proteins per LC MS/MS run.

An advantage of the method presented here over previously published methods analysing proteins from DBS is that the process was highly economical with the amount of trypsin used. Previous examples used between 25 and 200 µg to digest each sample [167, 171, 174, 176] whereas this method required only 0.45 µg. This means it becomes possible to use high purity proteomics grade trypsin such as trypsin gold at a very low cost per sample.

The 1 hour time period used in the digestion is short compared to the digestion times used for typical solution phase proteomics experiments. A 1 hour digest fits in well with the run time of HPLC. Longer digestion times would be more demanding in instrument time but could possibly result in a greater number of protein identifications. The influence of digestion time was investigated in a time course experiment where DBS were treated the a digestion routine described in section 4.4, but with incubation periods of 30 minutes, 1 hour, 2 hours, 4 hours and 8 hours. This was performed in triplicate and

samples were analysed by the LC MS/MS method described in section 2.2.2 and the database searching as described in section 2.2.8.

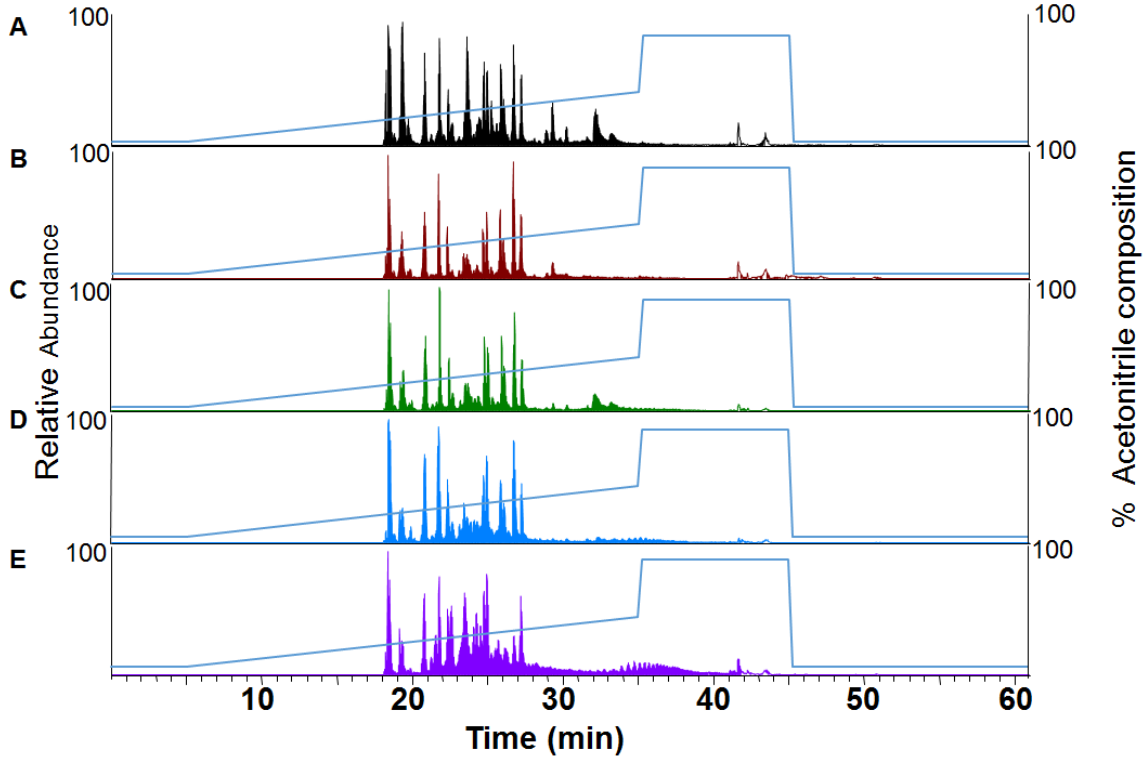
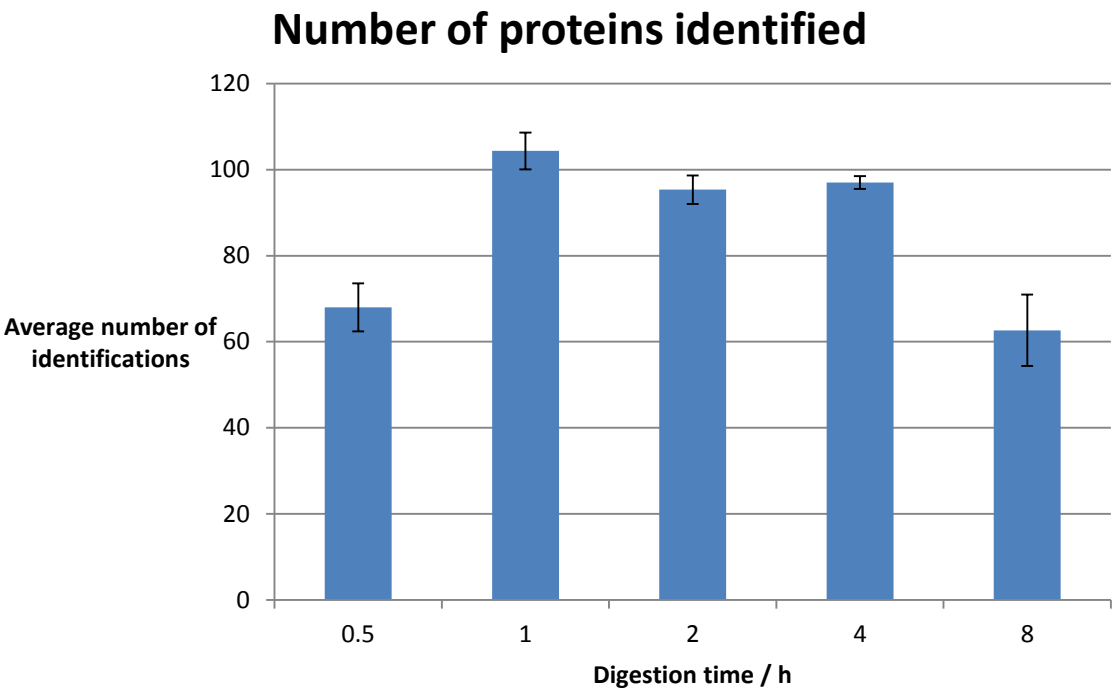


Figure 4.12 total ion current chromatograms showing acetonitrile compositions of LESA based DBS digests across different time points (A) 30 minute digest (B) 1 hour digest (C) 2 hour digest (D) 4 hour digest (E) 8 Hour digest

Fig 4.12 shows representative TICs obtained following LC MS/MS analyses of samples with different proteolysis times. The results obtained from the 30 minute, 1 hour and 2 hour digests shown in fig 4.12 A B and C show peaks corresponding to small amounts of intact protein eluting shortly after 30 minutes. As digestion time is increased (Fig 4.12 D and E) these peaks disappear as the digestion comes closer to completion.



SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
30 mins	3	204	68	124
1 hour	3	313	104.3333	72.33333
2 hours	3	286	95.33333	44.33333
4 hours	3	291	97	9
8 hours	3	188	62.66667	274.3333

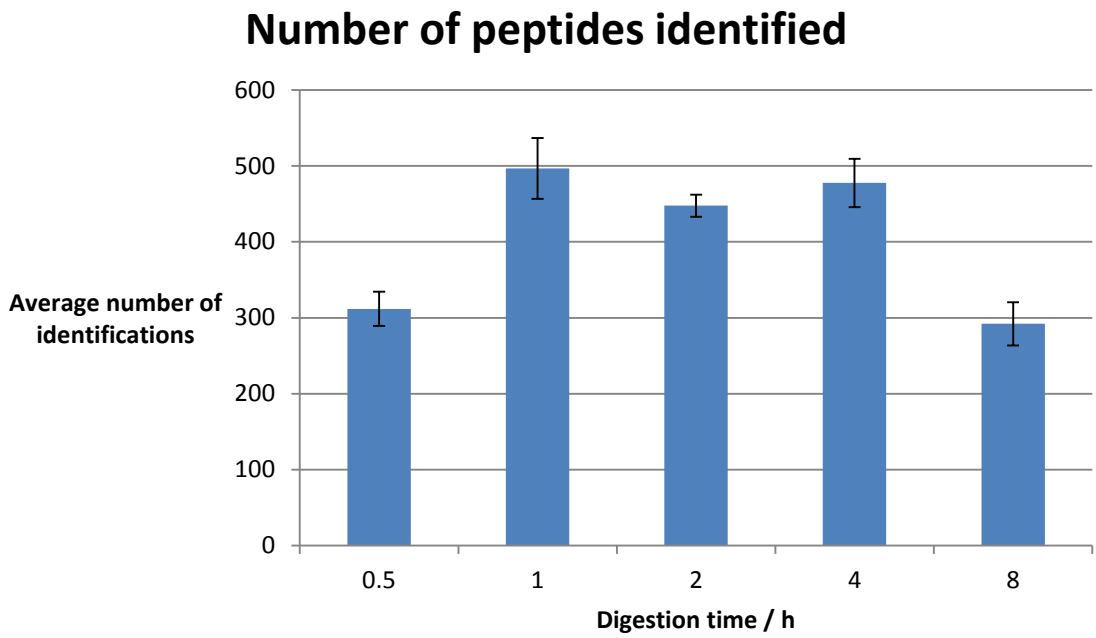
ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4233.733	4	1058.433	10.09955	0.00154	3.47805
Within Groups	1048	10	104.8			
Total	5281.733	14				

Figure 4.13 Average number of proteins identified across 3 replicates for each digestion time point and corresponding one way ANOVA. Error bars represent one standard deviation

Pairwise t-test Pair	p
30 mins-1 hour	0.00145
30 mins-2 hours	0.00843
30 mins-4 hours	0.00603
30 mins-8 hours	0.5377
1 hour- 2 hours	0.30689
1 hour- 4 hours	0.40089
1 hour- 8 hours	0.00055
2 hours- 4 hours	0.84595
2 hours- 8 hours	0.00292
4 hours- 8 hours	0.00212

Figure 4.14 Pairwise t-test of proteins identified from digestion each time point



Groups	*Count*	*Sum*	*Average*	*Variance*
30 mins	3	935	311.6667	2032.333
1 hour	3	1490	496.6667	6434.333
2 hours	3	1343	447.6667	849.3333
4 hours	3	1433	477.6667	4034.333
8 hours	3	876	292	3199

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	110951.1	4	27737.77	8.380328	0.003107	3.47805
Within Groups	33098.67	10	3309.867			
Total	144049.7	14				

Figure 4.15 Average number of peptides identified across 3 replicates for each digestion time point and corresponding one way ANOVA. Error bars represent one standard deviation

Pairwise t-test Pair	p
30 mins-1 hour	0.0028
30 mins-2 hours	0.016
30 mins-4 hours	0.0054
30 mins-8 hours	0.6843
1 hour- 2 hours	0.3215
1 hour- 4 hours	0.6944
1 hour- 8 hours	0.0014
2 hours- 4 hours	0.5374
2 hours- 8 hours	0.0078
4 hours- 8 hours	0.0027

Figure 4.16 Pairwise t-test of peptides identified from each digestion time point

Fig 4.13 shows that digestions of one hour resulted in the greatest number of proteins identified across the whole time range. The figure shows that the number of proteins identified increases when digestion time increases from 30 minutes to one hour, followed by a slow decrease between one and four hours and then decreases further between four and eight hours. The results of the ANOVA prove that the effect of digestion time does have a statistically significant effect (p=0.00154) on the number of proteins identified following LC MS/MS of the digests. The significance of the p Value is also reinforced by the fact that the F value (F=10.099) is greater than the F Critical of 3.47. The pairwise t-test shown in fig 4.14 revealed that there were significant differences between some time points and not others. There was no significant difference between 30 minutes and 8 hours, 1 and 2 hours, 1 and 4 hours and 2 and 4 hours. The differences between all other time points were statistically significant. There was no significant difference between the one hour time point and the two or four hour time points despite the fact

the one hour digest resulted in the greatest average number of protein identifications. Fig 4.15 shows the how the number of peptides identified changes with digestion time. Unsurprisingly this follows the trend of the number of proteins identified. The results of the ANOVA and the pairwise t-test of the peptide identifications are also similar to that of the protein identifications. The ANOVA showed that the digestion time does have a statistically significant effect on the number of peptides identified by LC MS/MS. The pairwise t-test found the same significant difference between the different time points. I.e. there was no significant difference between time points 30 minutes and 8 hours, 1 and 2 hours, 1 and 4 hours and 2 and 4 hours whereas all other pairs were significantly different.

It is perhaps surprising that the number of proteins identified did not increase greatly with proteolysis time. The expectation is that a more complete digestion would result in a greater number of protein identifications. The fact that it was not is likely due to the complexity of the sample. Blood and plasma are incredibly complex samples, composed of tens of thousands of different proteins. Without any fractionation beyond liquid chromatography, an untargeted shotgun proteomics approach will only be capable of identifying the high abundance proteins. The technique is not sensitive enough to cope with the dynamic range present in blood or plasma. If a digestion of a complex sample is carried out to completion the number of rare peptides from low abundance proteins will become highly saturated by peptides yielded from high abundance proteins. Previous groups have reported that for a complex sample, longer digestion times do not

necessarily lead to an increase in protein identifications [209]. The decrease in the number of proteins identified in the later time point may also be due to the effect of proteases present in blood. Blood contains a number of proteases, such as those found in the complement system [210] and a large number of clotting factors such as thrombin and Factor Xa [211]. Blood proteases do not cleave proteins ubiquitously at lysine or arginine residues in the way that trypsin does [212]. If proteolytic activity is regained after drying and after the proteins have been extracted from the DBS, a longer digestion time may result in continued proteolysis of endogenous blood proteases may cleave proteins at many different residues [212]. This effect may prevent peptides being recognisable in a database search.

4.5 Chapter 4 conclusion

The work described in this chapter shows that LESA can be used for the successful identification of over 100 different proteins from a DBS after the incorporation of an automated trypsin digest. It is also the first time in which DBS had been probed with untargeted shotgun proteomics by any means at all. Since publishing this work this idea has since been explored by others [213]. Untargeted analysis is commonly performed on plasma samples to identify new disease biomarkers [163] with the majority of groups using multi-dimensional fractionation or immunodepletion techniques to increase protein identifications. In comparison to the number of proteins identified by analysing plasma with multi-dimensional fractionation the number of proteins identified from this analysis of DBS is very small. By using multi-dimensional fractionation of plasma, over a thousand proteins can be identified from a plasma sample. Shen *et al.* [214] identified 1348 proteins from a plasma sample by using strong cation exchange alongside reverse

phased LC MS/MS. Without multi-dimensional fractionation the numbers are less impressive. Using reversed phase LC MS/MS analysis only Shen *et al*. [214] identified 110 proteins from a plasma sample and haudek *et al.* [215] identified 204 proteins from plasma. This is more comparable to the number of proteins identified from DBS in this chapter. There are far fewer examples where mass spectrometry has been used to analyse whole blood or DBS. The majority of groups that have analysed whole blood by mass spectrometry have done so to analyse haemoglobin for variant determination [166, 216, 217] and have not attempted deep level proteome analysis. Hemoglobin analysis can even involve analysing intact proteins rather than tryptic digests [216, 217]. Following the work carried out in this thesis however chambers *et al*. [213] have carried out an analysis of whole blood by an untargeted bottom up proteomics approach. The group analysed whole blood and DBS with an LC MS/MS shotgun proteomics approach. The authors identified an average of 223 proteins from a whole sample and 253 proteins from a DBS both over 3 replicates [213]. They did not carry out multi-dimensional fractionation, merely reversed phase liquid chromatograph. Their analysis identified a greater number of proteins than the numbers of proteins identified from the DBS in this chapter however the LC gradient used in the method was longer than that carried out here so did they did benefit from improved separation.

The work carried out in this chapter demonstrates that DBS could be used as an alternative platform for biomarker discovery with the added advantages such as ease of use that are associated with DBS. The method used here successfully identified a number

of proteins that are already used as or are under investigation to be used as biomarkers for clinical assays. These include haemoglobin, cerruloplasmin, alpha-1-anti-trypsin and proteins S100 A8 and S100 A9 [166, 171, 184, 218]. Alpha-1-antitrypsin was investigated for this purpose in Chapter 6. The surface sampling based digest is highly economical in regards to the amount of trypsin required and is applicable for many other surfaces. Following on from this work, the approach has been successfully applied for the analysis of thin tissue sections [219]. The aim of this thesis is to investigate new applications for using LESA in proteomics. Previously applications of LESA in proteomics were limited entirely to top down protein analysis. It is a major challenge when success is limited to the chance that the sample contains a few proteins that are small enough and abundant enough to be identified without proteolytic digestion. A bottom up approach does not face the same limitation. Performing a trypsin digestion of a DBS enabled identification of proteins with molecular weights in the range 9-515 kDa and endogenous concentrations between 600 and 0.05 µM in human plasma. The method retained the major advantages of LESA sampling. The digestion method was able to analyse a surface with no sample preparation in an automated fashion that kept manual user intervention to a minimum.

Chapter 5 Bottom up proteomics of dried blood spots: solution phase chemistry and immunodepletion

5.0 Introduction

The aim of the work presented in this chapter was to further develop DBS proteomics by using solution phase techniques as opposed to the surface sampling based method described in Chapter 4. As described in the introduction, other groups have employed a "punch and elute" strategy for the analysis of targeted proteins from DBS and even untargeted shotgun proteomics since the work in Chapter 4 was published [213]. Untargeted shotgun proteomics of DBS was carried out by Chambers *et al*. The researchers compared proteins identified from whole blood, plasma, serum and 20 µl spots of each of the three biofluids dried onto filter paper. After LC MS/MS analysis, 253 proteins were successfully identified from DBS. This number is greater than were identified by the LESA based digest carried out in Chapter 4, but was obtained following use of a longer LC gradient (90 minutes) [213]. In the punch and elute methods, a small disk (typically 2-6 mm) is cut out of a DBS card and the blood sample is eluted into solution and digested before LC MS/MS analysis. The "punch and elute" method was investigated here for two reasons: Firstly, to act as a comparison to the LESA-based approach described in Chapter 4, and secondly, to investigate the potential of immunodepletion as a means for identifying low abundance proteins. The results of Chapter 4 demonstrated the successful identification of approximately 120 proteins per

LC MS/MS run, however this represents a fraction of the proteins that are present in blood. Blood is an incredibly complex sample that is dominated by a few high abundance proteins, as mentioned in section 1.4.3.3 99% of the total protein content comprises just 22 different proteins and the remaining one percent is composed of several thousand [179]. Consequently, many low abundance proteins are undetectable by shotgun proteomics techniques in the absence of enrichment techniques. One approach to address this challenge is to remove some of the high abundant proteins before digestion, thus simplifying the sample and improving detection of lower abundance proteins. A range of immunodepletion kits are available commercially for removing high abundance proteins from human plasma. The R&D Systems Proteome Purify 12 kit which targets the 12 most abundant human plasma proteins, is investigated in this chapter. A separate depletion kit, known as Hemovoid that is designed for removing haemoglobin from erythrocytes, is also available commercially and was attempted for use on DBS.

5.1 Proteolysis of blood proteins following punch and elution" from DBS samples

A series of blood samples eluted from DBS were digested in solution as described in section 2.2.7. After elution of the proteins from the DBS, the proteins were digested for a range of different time points (30 minutes, 1 hour, 2 hours, 4 hours and 8 hours). These timepoints were the same as those described for optimisation of the LESA based digest described in section 4.4. The resulting digests were analysed by use of the same LC MS/MS method with online desalting as described in section 2.2.2. This was performed in triplicate and representative total ion current chromatograms obtained following

analysis of the resulting digests are shown in fig 5.1. The full list of proteins identified for the 3 replicates across the various time points is shown in appendix 5.1-5.5.

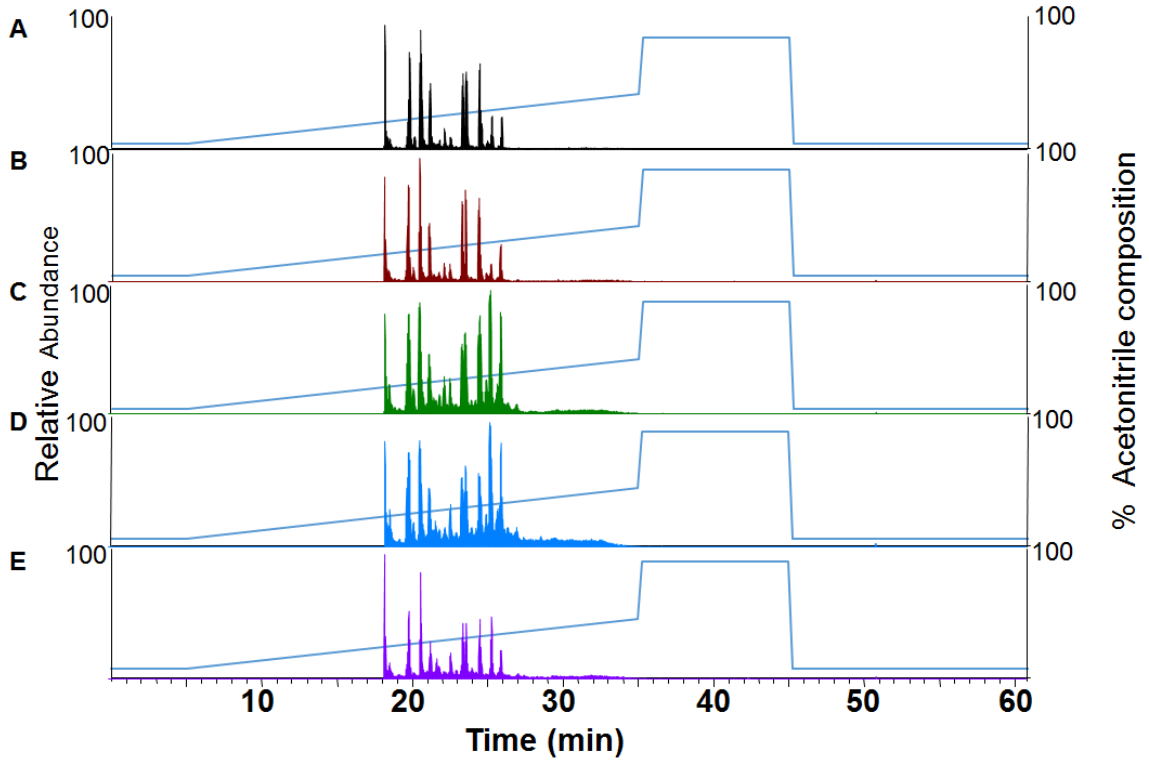


Figure 5.1 Total ion chromatograms showing acetonitrile composition obtained following LC MS/MS analysis of samples extracted from DBS ("punch and elute") and digested for (A) 30 mins (B) 1 hour (C) 2 hours (D) 4 hours (E) 8 hours.

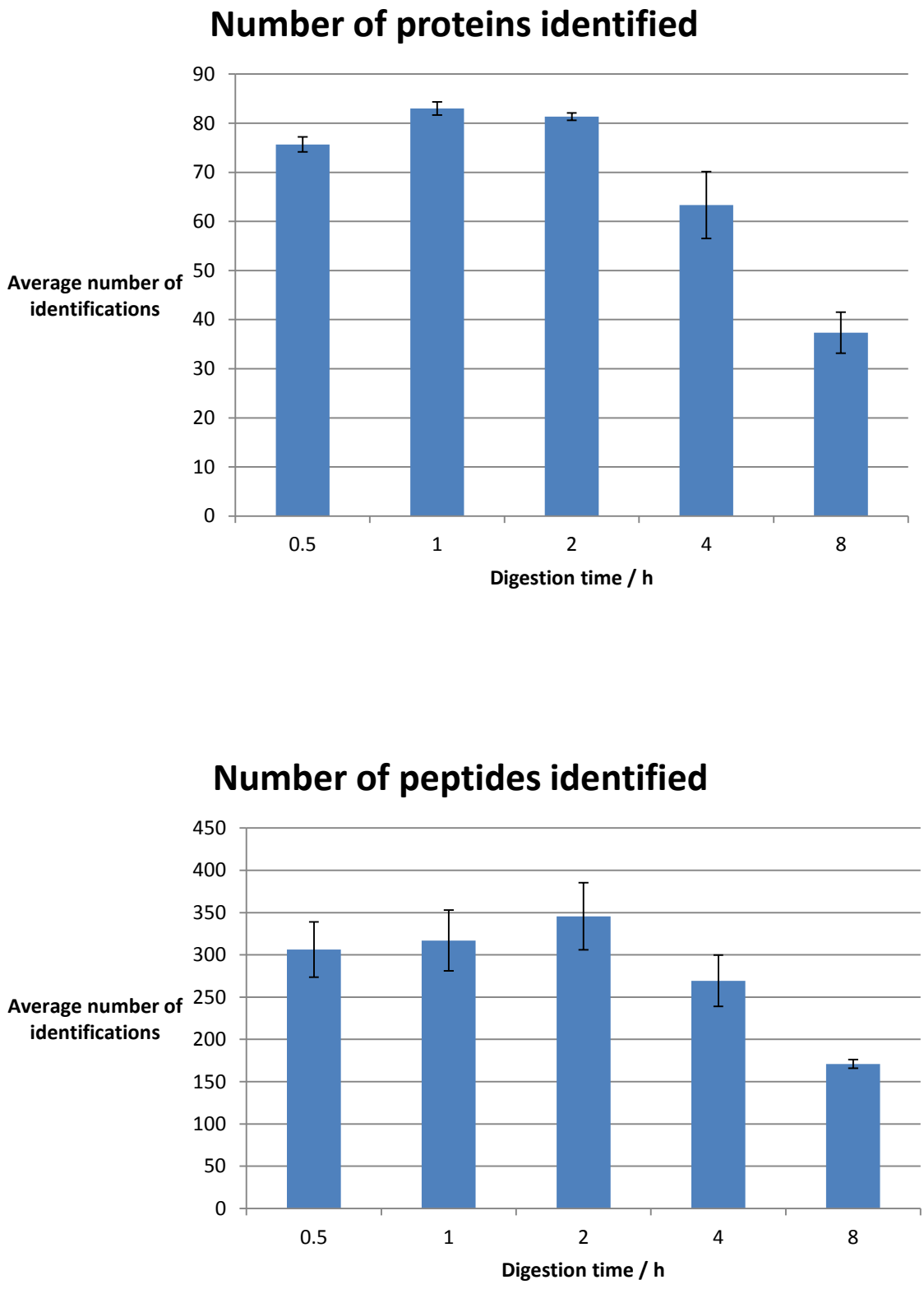


Figure 5.2 Average number of identifications across 3 replicates for each time point (A) number of proteins identified (B) number of peptides identified. Error bars represent 1 standard deviation

As shown in fig 5.2, the number of proteins and peptides identified after LC MS/MS decreased at digestion times of greater than two hours. The same phenomenon was reported in the LESA based digests described in Chapter 4. This observation, i.e., that the number of proteins identified decreases as the digestion comes to completion, may seem counter-intuitive. It is worth bearing in mind that, in a complex sample such as blood, the number of different peptides in the digest greatly exceeds the scanning capability of data dependent MS/MS analysis, which will target the most abundant peptides in a sample [122]. This effect where longer digestion times result in fewer protein identifications has been reported previously in the analysis of complex mixtures [209]. That paper also reported that the greatest number of proteins is identified with a short digestion of just 1 hour. The exact reasons behind this effect are unknown but plausible explanations include increased deamidation of proteins in digestion buffers during prolonged proteolysis, and the fact that shorter digests can increase the number of peptides with one or two missed cleavages which can improve protein identifications [220, 221].

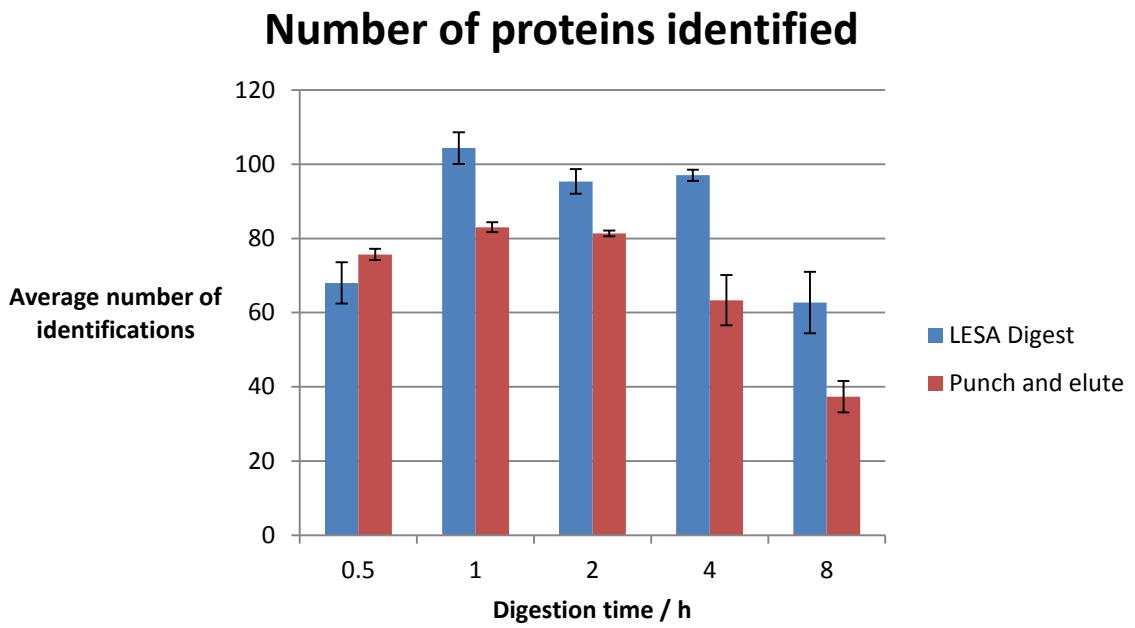


Figure 5.3 Average number of proteins identified across 3 replicates from different digestion times using the LESA based digest described in Chapter 4 and a 'punch and elute' approach. Error bars represent 1 standard deviation.

Fig 5.3 shows that the number of proteins identified from the "punch and elute" digest is significantly less than that from the LESA digest for all time points other than 30 minutes. Again, this result is counterintuitive; however, it is worth bearing in mind that in the LESA based digest, only the surface proteins are extracted from the DBS for digestion whereas in the "punch and elute" procedure the whole sample, including all the dried erythrocytes and platelets are eluted from the card. The second approach will therefore result in samples that are more enriched with cellular proteins such as haemoglobin. Fig 5.4 shows that the percentage of peptide spectral matches corresponding to haemoglobin is higher in the "punch and elute" approach than it is in the LESA based process. This result suggests that the number of identified proteins from the LESA approach is greater than

with the “punch and elute” method because the former is less saturated with haemoglobin.

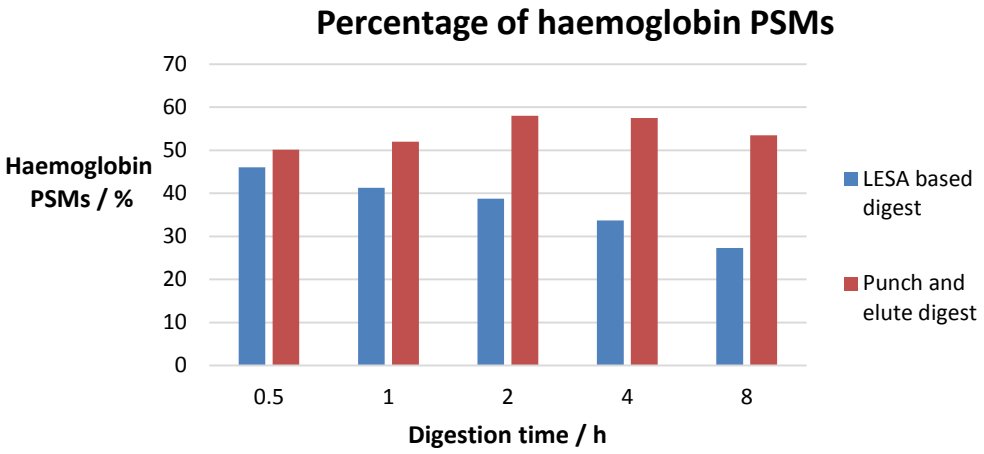


Figure 5.4 Percentage of peptide spectral matches corresponding to haemoglobin peptides in DBS digests. Error bars represent 1 standard deviation.

The results presented above demonstrate the successful identification of multiple proteins from DBS via a “punch and elute” method, although fewer proteins were identified than by digesting with LESA. The results demonstrate that DBS proteomics can be achieved through more traditional solution phase digestions and could be carried out in laboratories that do not have the facilities for surface sampling.

5.2 Immunodepletion

There are many plasma immuno-depletion kits available commercially, which remove the most abundant proteins from a sample. The kits can target a range of proteins from albumin alone to the top 20 most abundant plasma proteins [222, 223].

The R&D Systems Proteome Purify 12 kit was used here for the depletion of the high abundance plasma proteins. It targets the top 12 most abundant plasma proteins, see Table 5.1,which together make up 95% of the total protein in human plasma [224].

Alpha-1-acid glycoprotein	Alpha-1-antitrypsin
Alpha-2-Macroglobulin	Albumin
Apolipoprotein A-I	Apolipoprotein A-II
Fibrinogen	Haptoglobin
IgA	IgG
IgM	Transferrin

Table 5.1 Proteins targeted by the Proteome Purify 12 kit

The Proteome Purify 12 kit does not target haemoglobin, plasma is depleted of red blood cells so it contains very little haemoglobin. A separate depletion kit known as Hemovoid is available for haemoglobin depletion. This kit uses a spin column composed of a silica enrichment matrix and was originally designed for erythrocyte proteomics [225]. Unlike the Proteome Purify 12 kit, Hemovoid does not use antibodies. The silica beads contain a range of ionic, hydrophobic, aromatic and polymer ligands. Almost all proteins have an

affinity to one of these types of ligand and will bind to the silica beads. Of all the different ligands only a small percentage will have suitable chemistry for an affinity to haemoglobin, and these ligands quickly become saturated with haemoglobin when the concentration is high. Other proteins are able to bind to the free ligands which enriches the sample in low abundance proteins and eliminates up to 98% of haemoglobin present [225].

Neither of the two depletion kits were investigated for use with the LESA based digest. They both require use of centrifugation during sample preparation so are not practicable for use in conjunction with LESA. The two kits were each used in isolation for treatment of DBS samples and were subsequently used in tandem. When used in tandem, the plasma proteins could be depleted first followed by the haemoglobin, or vice versa. Both approaches were investigated.

5.2.1. Treatment of samples with Plasma Proteome Purify 12 kit

Proteins were eluted from the DBS as described in section 2.2.7 After elution the proteins recovered from the DBS were depleted of the top 12 most abundant plasma proteins as described in section 2.2.19. The depleted sample was reduced, alkylated and digested with trypsin. Two proteolysis time points were used: 1 hour and 8 hours. As described in

section 4.1, the 1 hour time proteolysis resulted in the greatest number of protein identifications (in the absence of immunodepletion); however, since a more complete digest would be less complex after depletion, an 8 hour digest was investigated as well. All data were acquired in triplicate.

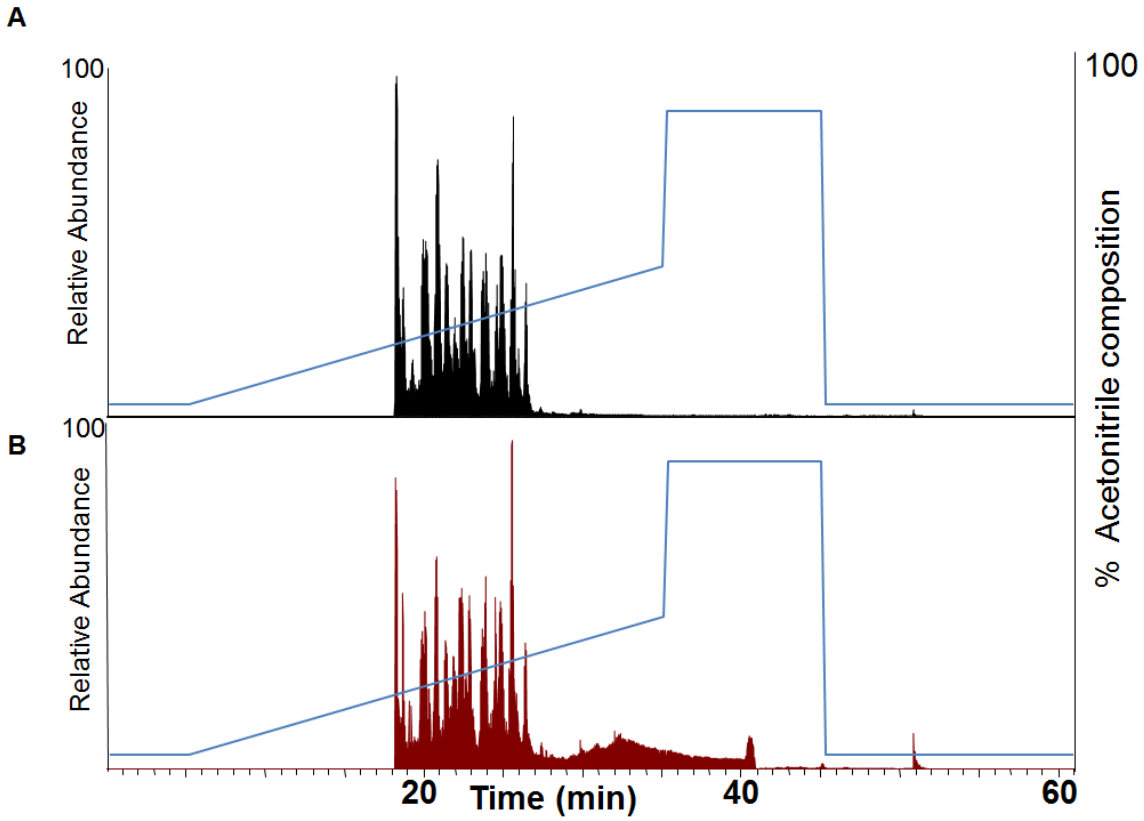


Figure 5.5 Total ion chromatograms showing acetonitrile composition obtained following LC MS/MS analysis of samples extracted from DBS and immunodepleted of the top 12 plasma proteins (A) 1 hour digest (B) 8 hour digest

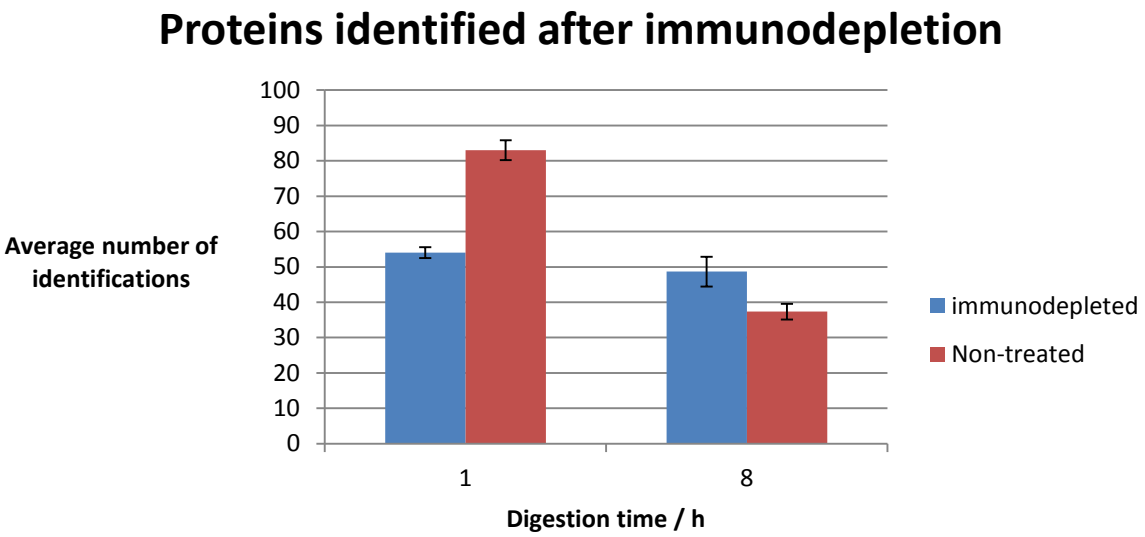


Figure 5.6 Average number of proteins identified across 3 replicates after immunodepletion of the top 12 plasma proteins and non-treated DBS digest. Error bars represent 1 standard deviation.

Digestion time	P Value
1 hour	0.001
8 hour	0.1

Table 5.2 Two tailed unpaired t-test on number of protein identifications before and after depletion of the top 12 plasma proteins

Representative total ion chromatograms are shown in fig 5.5. Fig 5.6 shows that there was no major improvement in the number of proteins identified after immunodepletion. The full lists of proteins identified after immunodepletion for the two time points are shown in appendix 5.6 and 5.7. The number of proteins identified from immunodepleted

samples actually decreased in the 1 hour digest and there was only a small and non-significant (p=0.1) increase in the number of proteins identified from the 8 hour digest. There are two factors that could be responsible for this observation. Firstly, the depletion kit pulls out the abundant plasma proteins. Across the 3 replicates for the 1 hour digests, a total of 8,9 and 9 of the top 12 plasma proteins were not identified from the samples and a total of 9 of the top 12 proteins were not identified from each replicate of the 8 hour digests. Secondly, problems with non-specific protein removal have been reported with plasma depletion kits [224], i.e., removal of proteins other than those targeted for by the antibodies in a non-specific manner during the immunodepletion step.

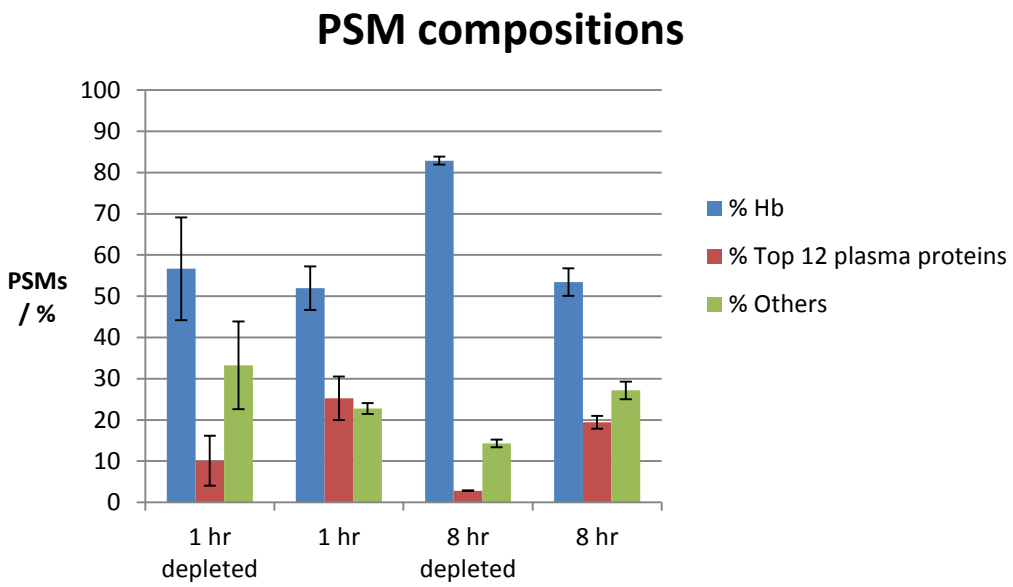


Figure 5.7 Average percentage of PSMs over 3 replicates corresponding to peptides originating from haemoglobin and the 12 most abundant plasma proteins. Error bars represent 1 standard deviation.

The success of the immunodepletion can be assessed by considering the percentage of peptide spectral matches (PSMs)s made up of the top 12 plasma proteins as shown in fig 5.7. In the untreated samples 25 % of all PSMs from the 1 hour digest and 19% of those from the 8 hour digest originated from the top 12 plasma proteins. After immunodepletion, these values are 10% and 3 % respectively. This result shows that the immunodepletion was successful in reducing the amount of high abundant plasma proteins; however, there is a concomitant increase in the number of haemoglobin PSMs. For the 1 hour digest, the increase was small (from 52 to 57%) (p=0.7) but the increase for the 8 hour digest was much larger - from 53 to 83% and was significant (p=0.001), as shown in Table 5.3. This observation may explain why the number of identified proteins

did not increase considerably. After the high abundance plasma proteins have been removed, haemoglobin dominates the digest more so than previously, thus preventing detection of the low abundance peptides and further protein identification.

Digestion time	P Value
1 hour	0.7
8 hour	0.001

Table 5.3 Two tailed unpaired t-test on the percentage of PSMs corresponding to haemoglobin before and after depletion of the top 12 plasma proteins.

5.2.2 Treatment of samples with Hemovoid kit

In these experiments, samples were depleted by Hemovoid only. A punch was cut out of a DBS and proteins were eluted as described in section 2.2.7. After elution the proteins recovered from the DBS were depleted of haemoglobin by using Hemovoid as described in section 2.2.9. The depleted sample was reduced, alkylated and digested with trypsin as described in section 2.2.9. Samples were digested for 1 hour and 8 hours as above. Data was acquired in triplicate and representative total ion chromatograms obtained following LC MS/MS of the peptide digests are shown in fig 5.8.

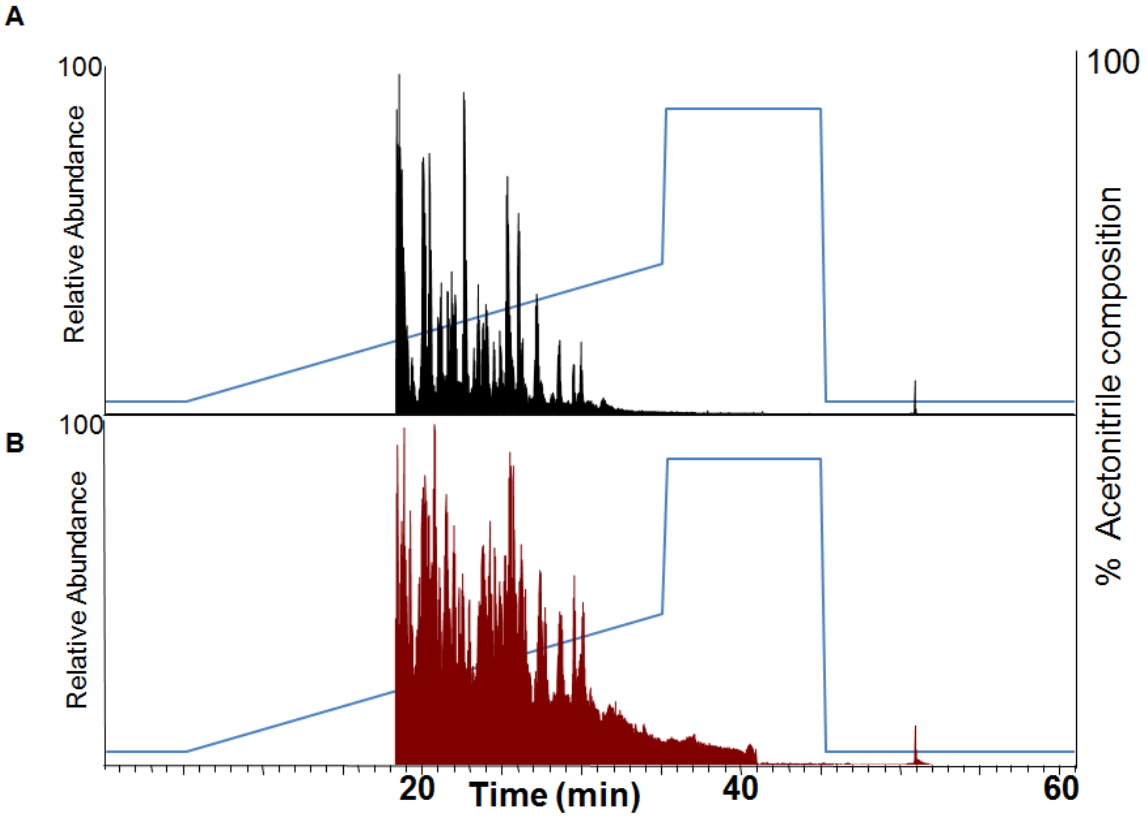


Figure 5.8 Total ion chromatograms showing acetonitrile composition obtained following LC MS/MS of DBS samples depleted of haemoglobin (A) 1 hour digest (B) 8 hour digest.

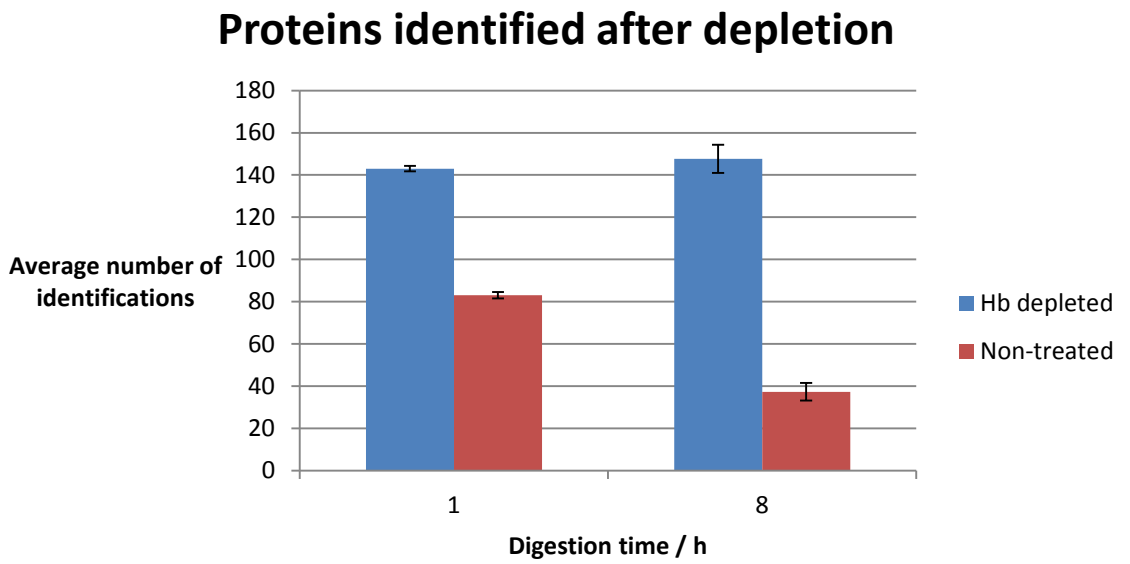


Figure 5.9 Average number of proteins identified from DBS samples across 3 replicates after haemoglobin depletion and non-treated DBS digest. Error bars represent 1 standard deviation

Digestion time	P Value
1 hour	0.00001
8 hour	0.0002

Table 5.4 Two tailed unpaired t-test on protein identifications before and after depletion of haemoglobin

Treatment of the samples with Hemovoid successfully resulted in an increase in the number of proteins identified from the DBS samples as shown in fig 5.9. The full list of proteins identified following haemoglobin depletion are shown in appendix 5.8 and 5.9. Treatment with Hemovoid was more successful than with the plasma proteome depletion kit (see fig 5.7). After depletion of the top 12 plasma proteins an average of 54 and 48 proteins were identified from the samples digested for 1 hour and 8 hours respectively, whereas haemoglobin depletion resulted in the identification of an average of 143 proteins for the 1 hour and 147 proteins for the 8 hour digestion time points.

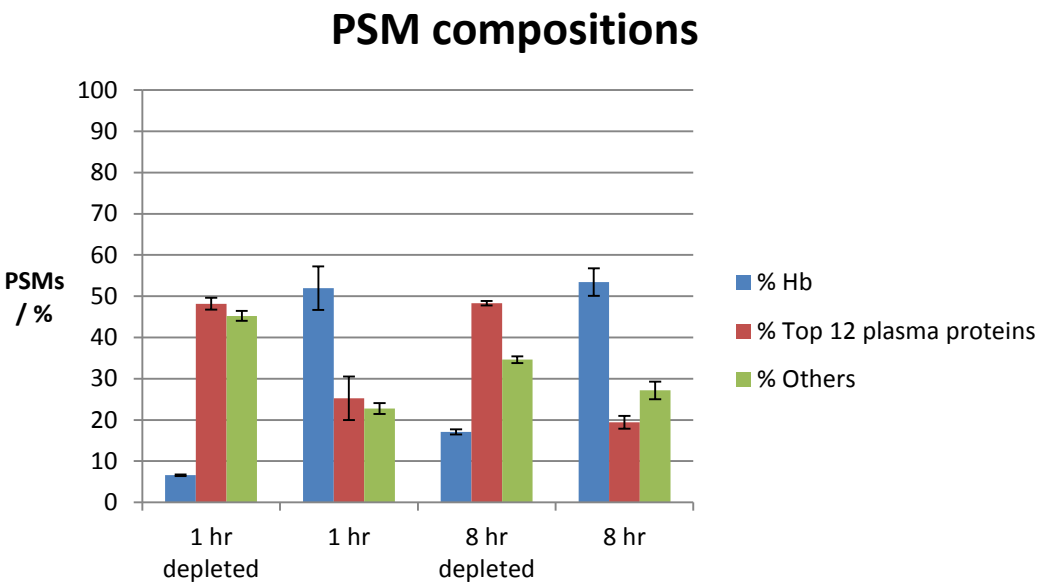


Figure 5.10 Average percentage of PSMs over 3 replicates corresponding to peptides originating from haemoglobin and the top 12 plasma proteins Error bars represent 1 standard deviation

Hemovoid successfully reduced the concentration of haemoglobin in the sample as illustrated in fig 5.10. The percentage of PSMs corresponding to haemoglobin peptides in the 1 hour digest dropped from 52 to 7 % for the 1 hour digest and from 53 to 17% for the 8 hour digest. The remaining amount of haemoglobin after depletion may seem high but it should be noted that Hemovoid removes only 98% of the total haemoglobin. Considering haemoglobin is such an abundant protein, for a healthy adult male with approximately 130 mg/ml of haemoglobin in their blood [205], the remaining 2% would leave a concentration of 2.6 mg/ml which would make it the next most abundant protein after albumin [208]. Following haemoglobin depletion, fig 5.10 shows that the percentage of PSMs corresponding to high abundance plasma proteins increased, as did the percentage of PSMs corresponding to the other proteins. The relative increase in the

PSMs from the top 12 plasma proteins is similar to the increase in PSMs from haemoglobin following immunodepletion of the abundant plasma proteins.

5.2.3 Combined immunodepletion of abundant plasma proteins and haemoglobin depletion

Workflows combining the two depletion strategies were considered. Proteins were eluted from a DBS punch as described in section 2.2.7, and treated by immunodepletion followed by haemoglobin depletion, or by haemoglobin depletion followed by immunodepletion.

5.2.3.1 Treatment of samples with Proteome purify 12 and subsequent treatment with hemovoid

After elution the samples recovered from the DBS were immunodepleted of the twelve most abundant proteins by use of Proteome Purify 12. The eluent was then depleted of haemoglobin by use of hemovid as described in section 2.2.9. The twice-treated sample was then treated to a reduction, alkylation and trypsin digest as described in section 2.2.9. Samples were digested with trypsin for 1 hour and 8 hour and data were acquired in triplicate as described previously.

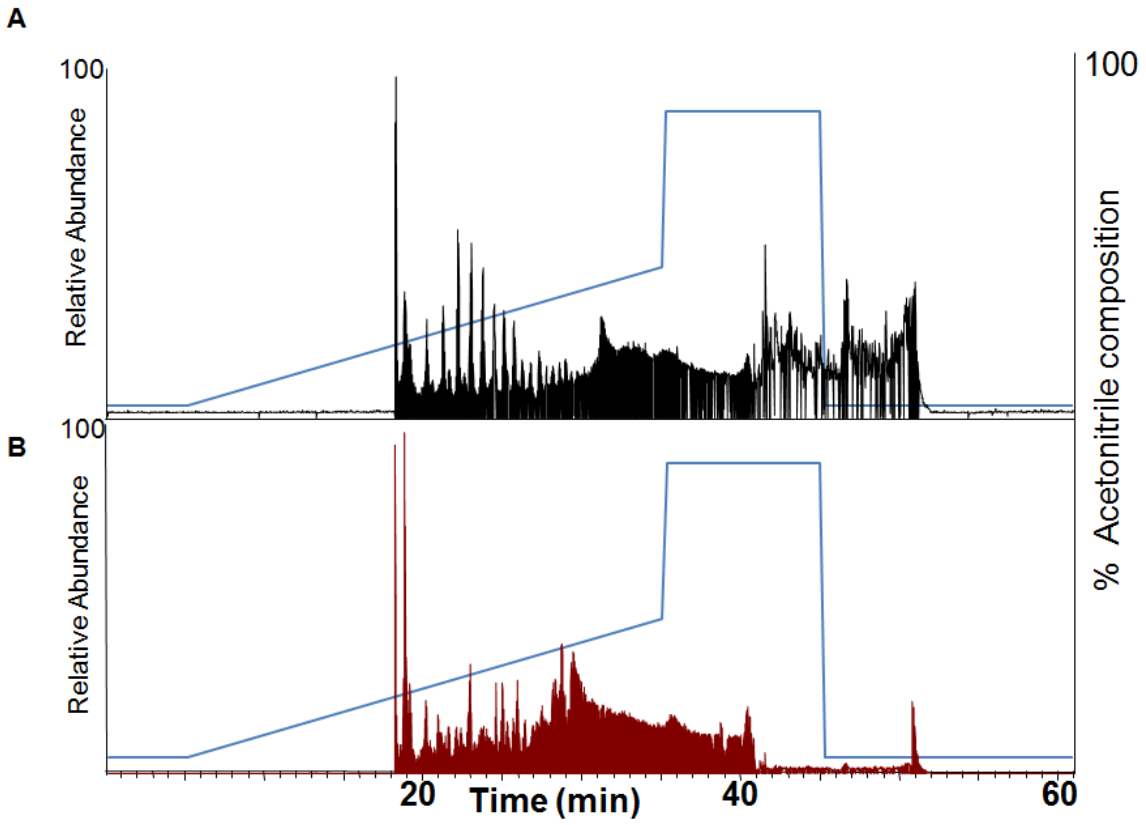


Figure 5.11 Total ion chromatograms showing acetonitrile compositions obtained following DBS digestions depleted of the top 12 plasma proteins and haemoglobin (A) 1 hour digest (B) 8 hour digest

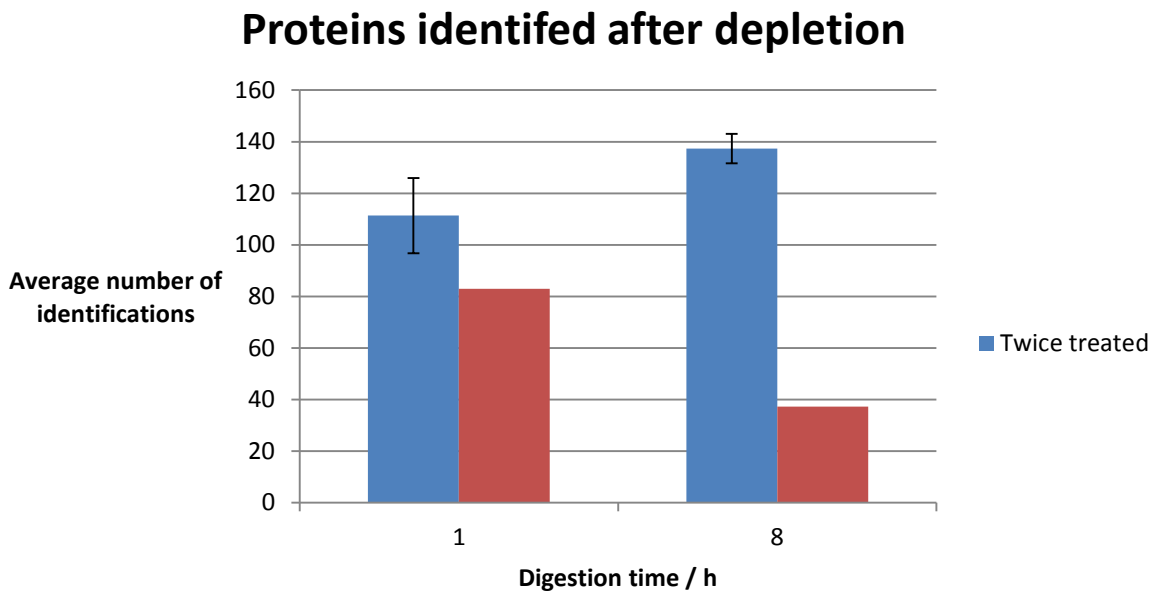


Figure 5.12 Average number of proteins identified across 3 replicates after twice treated and non-treated DBS digest. Error bars represent 1 standard deviation

Digestion time	P Value
1 hour	0.17
8 hour	0.0001

Table 5.5 Two tailed unpaired t-test on protein identifications before and after depletion of the top 12 plasma proteins and haemoglobin.

Representative total ion chromatograms obtained following LC MS/MS of the resulting peptide digests are shown in fig 5.11. The chart on fig 5.12 showed that performing depletion by both kits actually increased the number of proteins identified compared to the untreated sample. The full lists of proteins identified after depleting both the top 12 plasma proteins and haemoglobin are shown in appendix 5.10 and 5.11. This result shows that even after the high abundance plasma proteins and haemoglobin have removed there was still enough sample for the remaining low abundance proteins to be detected. The increase in the number of protein identifications was however smaller than performing haemoglobin depletion only. The average number of identifications from using Hemovoid only was 143 for the 1 hour time point and 147 for the 8 hour digest, whereas when using the Proteome Purify 12 kit and Hemovoid the average number of identifications were 111 and 137 for the 1 hour and 8 hour digests respectively.

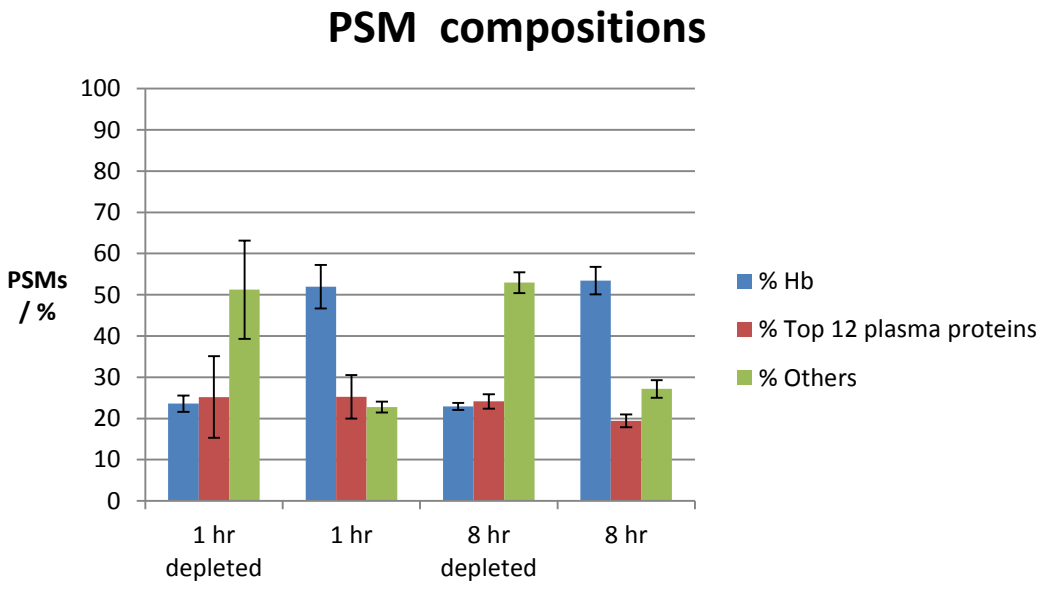


Figure 5.13 Average percentage of PSMs over 3 replicates corresponding to peptides originating from the top 12 plasma proteins and haemoglobin. Error bars represent 1 standard deviation.

Fig 5.13 shows how the composition of the PSMs changed according to treatment. The percentage of PSMs corresponding to haemoglobin dropped from 52 to 24 for the 1 hour digest and from 53 to 22 for the 8 hour digest. This decrease was less than that observed when treating with Hemovoid alone. The percentage of PSMs originating from the top 12 plasma proteins did not vary. For the samples digested for 1 hour, 25% of all PSMs were composed of the top 12 plasma in the untreated samples compared with 25% in the depleted samples. The percentage of PSMs from other proteins increased for both 1 hour and 8 hour digestions.

5.2.3.2 Treatment of samples with Hemovoid and subsequent treatment with Proteome Purify 12

After elution, the samples recovered from the DBS were depleted of haemoglobin by use of Hemovoid as described in section 2.2.9**.** The eluent was then depleted of the high abundance plasma proteins by using Proteome Purify 12. The sample was then reduced, alkylated and digested as described in section 2.2.9. Samples were digested for 1 hour and 8 hours. All data were acquired in triplicate as performed previously.

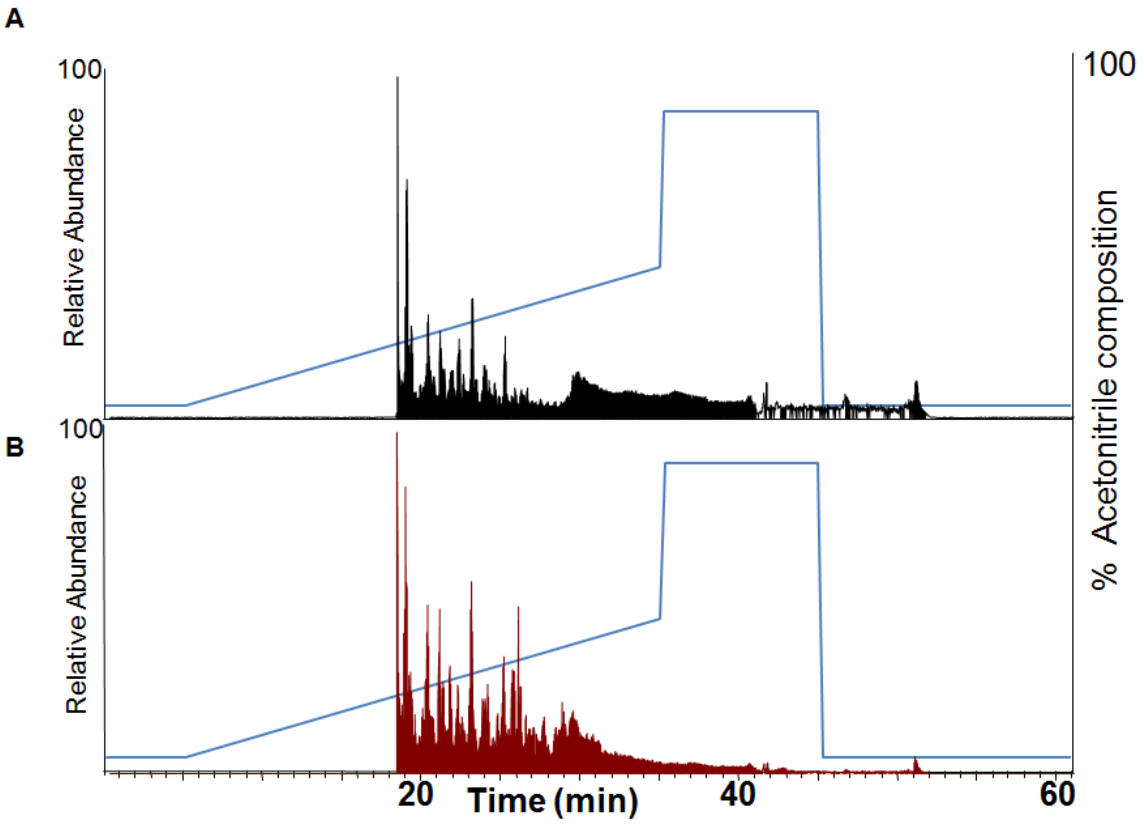


Figure 5.14 Total ion chromatograms showing acetonitrile composition obtained following DBS digestions depleted of haemoglobin and the high abundance plasma proteins (A) 1 hour digest (B) 8 hour digest

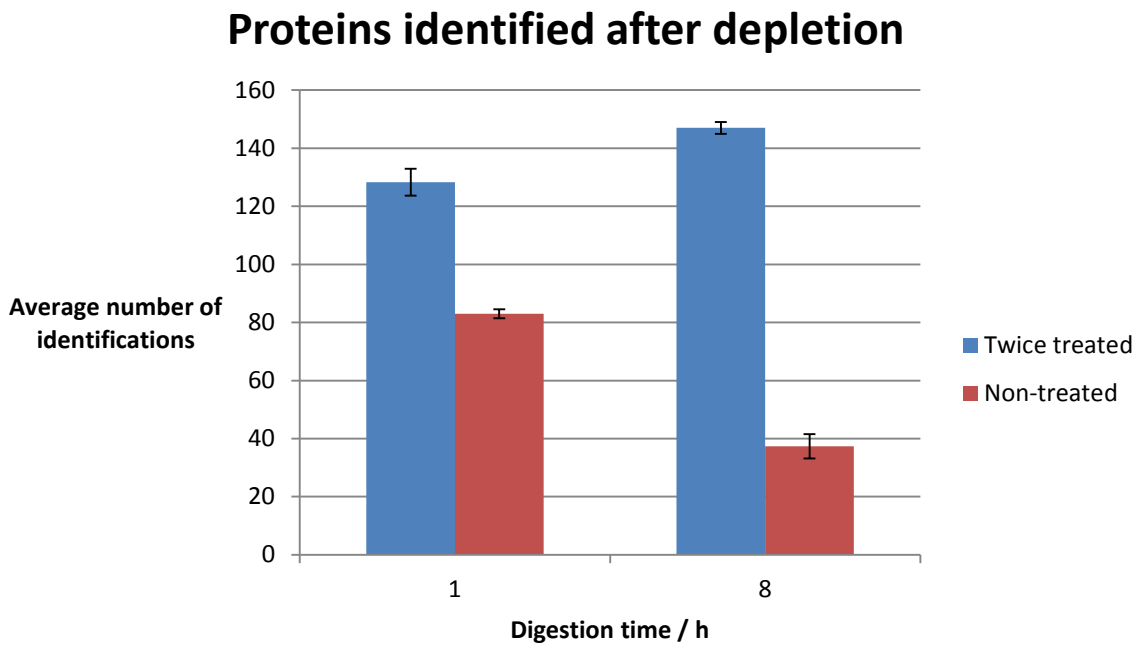


Figure 5.15 average number of proteins identified across 3 replicates after from twice treated and non-treated DBS digests. Error bars represent 1 standard deviation

Digestion time	P Value
1 hour	0.0012
8 hour	0.00005

Table 5.6 Two tailed unpaired t-test on protein identifications before and after depletion of haemoglobin and the top 12 plasma proteins.

Representative total ion chromatograms obtained following LC MS/MS of the resulting peptide digests are shown in fig 5.14. The chart on fig 5.15 showed that the number of proteins identified from samples depleted of haemoglobin followed by high abundance

plasma proteins increased compared to the untreated sample. The number of proteins identified by depleting haemoglobin followed by the 12 most abundant plasma proteins identified more proteins than performing the two rounds of depletion in the opposite order as carried out in section 5.2.3.1. The full lists of proteins identified after depleting haemoglobin and the top 12 plasma proteins can be seen in appendix 5.12 and 5.13. For the sample obtained following an 8 hour digest, the number of proteins identified from the combined depletion approach was equal to the number of proteins identified following depletion by Hemovoid alone. For the sample obtained following a 1 hour digest, fewer proteins were identified following the combined treatment than the Hemovoid alone. There was no statistically significant difference in the number of protein identifications between using Hemovoid and the two treatments together as shown in table 5.7.

Digestion time	P Value
1 hour	0.06
8 hour	0.95

Table 5.7 Two tailed unpaired t-test on protein identifications after depletion of haemoglobin prior to the top 12 plasma proteins and haemoglobin only.

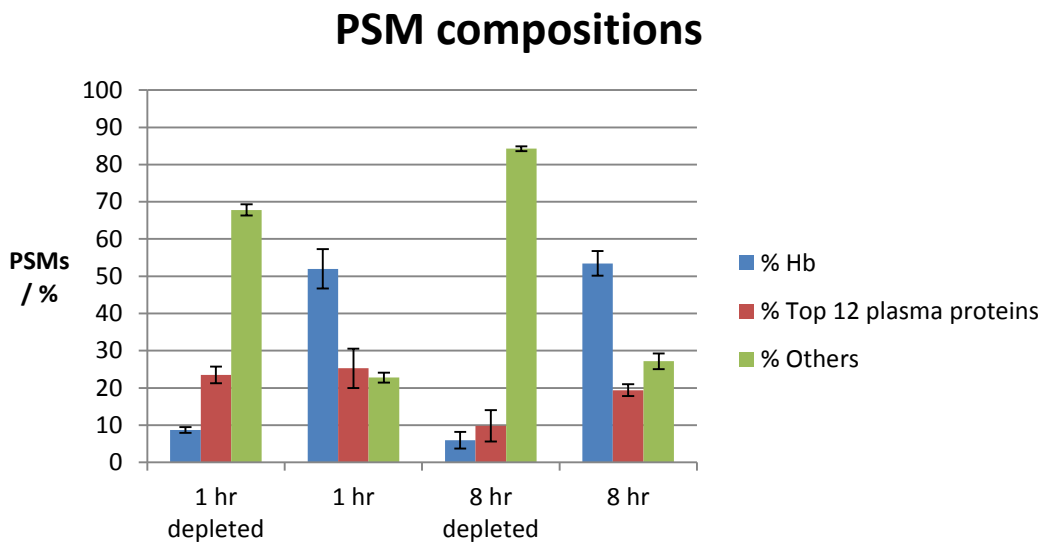


Figure 5.16 Average percentage of PSMs over 3 replicates corresponding to peptides originating from haemoglobin and the top 12 plasma proteins. Error bars represent 1 standard deviations

Fig 5.16 shows the composition of PSMS according to treatment. For the depleted samples, only 9% of PSMs from the 1 hour digest sample and 6% of PSMs from the 8 hour digest sample originated from haemoglobin. Of the PSMs that corresponded to the high abundance plasma proteins, these made up 23% of the total for the 1 hour digest sample and 10% of the total in the 8 hour digest sample. In both cases the remaining peptides made up more than 50% of all PSMs. Whilst this option has provided the most successful depletion, the increase in protein identifications was less impressive. The average number of proteins identified was 147 for the 8 hour digest and 128 for the 1 hour digest. The reason for this is likely to be the same as described when depletion was carried out in the opposite order. There is a finite limit to the initial amount of sample loaded for a DBS, and once both the haemoglobin and high abundance plasma proteins have been depleted, very little protein is left. Only 20µl of blood is applied to a DBS and each 6 mm punch corresponds to around 7µl of blood. The same challenge is not faced when

depleting plasma: it is possible to begin with larger volumes (up to 100 µl) to account for the amount of protein lost [226].

5.3 Conclusion

The work presented in this chapter has demonstrated that a proteomics analysis of DBS can be achieved not just by LESA sampling methods but also a “punch and elute” approach. The LESA sampling method consistently identified more proteins per LC MS/MS run than by using the “punch and elute” procedure without depleting the sample of its most abundant proteins. One advantage of analysing DBS via the “punch and elute” approach is that it does offer additional options for sample manipulation that would be impractical for a surface sampling approach. A range of different options for the depletion of high abundance proteins have been attempted on DBS eluents. Depleting the high abundance plasma proteins alone does not increase the number of protein identifications. Haemoglobin depletion did increase the number of protein identifications, and successfully identified more proteins than using the LESA based digest described in Chapter 4. Depleting both haemoglobin and the high abundance plasma proteins does not dramatically improve the number of identifications compared to just depleting haemoglobin. In each elution from a DBS only a small amount of blood is recovered, corresponding to approximately 7 µl. If both haemoglobin and the high abundance plasma proteins are removed from these eluents, then the remaining amount protein is

very small and relatively few extra proteins can be identified. By incorporating immunodepletion the number of proteins identified from the analysis can be increased when haemoglobin is removed. Further depletion of the plasma proteins is viable but does not necessarily increase the number of protein identifications [226].

Chapter 6 Determination of alpha-1-antitrypsin variants through bottom up proteomic analysis of DBS

6.0 Introduction

The work presented in this chapter aimed to investigate the use of LESA coupled with bottom-up proteomics of dried blood spots for the determination of variants of the alpha-1-antitrypsin protein. As mentioned in section 1.5, A1AT is the most abundant protease inhibitor secreted into human plasma. Several variants of this protein exist which result in a reduction in synthesis of the protein that leads to a pathological condition known as alpha-1- antitrypsin deficiency. Clinical manifestations of the disease can result in emphysema, chronic obstructive pulmonary disease and liver failure [227]. The A1AT protein is highly polymorphic. There are two common healthy variants of the A1AT protein which are known as the M variant. The two M variants differ in that one has a valine at the 213 position and the other has an alanine residue [188]. There are two common clinically significant variants that increase the likelihood of suffering from A1AT deficiency, known as S and Z. The S variant is a mutation of the healthy M valine 213 (MVal 213) form with a glutamic acid to valine transition at the 264 residue and the Z variant is a mutant of the M alanine 213 (MAla 213) form with a glutamic acid to lysine transition at residue 342 [188]. Inheriting either an S or Z variant of the protein decreases the serum concentration of A1AT of an individual. An individual with an SZ or ZZ genotype has a very low serum concentration of A1AT and are likely to suffer from A1AT deficiency [189]. The typical concentrations associated with each of the different genotypes can be seen in table 6.1.

A1AT genotype	Serum A1At concentration / μ M
MM (MVal 213, MAla213 or heterozygous)	20-50
SM	20-48
MZ	12-35
SZ	8-19
ZZ	2.5-7

Table 6.1 typical Serum concentrations associated with different A1AT genotypes [189].

Variants of A1AT are currently diagnosed by using isoelectric focusing. Isoelectric focusing measures the electrophoretic mobility of A1AT against a range of laboratory standard proteins for each of the common variants [192]. In addition to determining the variant of A1AT, quantification of the protein is also required for diagnosis , which is typically carried out by nephelometry or radial immune diffusion [192]. In some cases the measured concentration of A1AT disagrees with that which would be expected from the variant determined by the isoelectric focusing. In such an event genotyping of the A1AT is required for confirmation which can be achieved by PCR or DNA sequencing [228]. Mass spectrometry could be used as an alternative approach. Ultimately it could provide both quantification of the endogenous A1AT protein and be used to determine the different phenotypic variants, similar to methods involving the determination of haemoglobin variants previously reported [167, 168]. A report of a mass spectrometry MRM method has been published to quantify clinically significant variants of the A1AT protein, but the researchers used whole plasma rather than DBS for their analysis [229]. A method for determining clinically significant variants of A1AT from DBS could be a

useful application of the LESA based digest described in Chapter 4. The work in this chapter aimed to investigate the potential for of LESA mass spectrometry of DBS for determining the different genotypic variants.

6.1 A1AT coverage

The most common healthy and clinically significant variants of A1AT MAla 213, MVal 213, S and Z differ in 3 different amino acid residues: Ala or Val at position 213, Glu or Val at position 264 and Glu or Lys at position 342. To determine the variant of the A1AT protein these 3 regions of interest on the protein must be identified. Following trypsin digestion of the proteins there are 6 different diagnostic peptides across the 4 variants (MVal 213, MAla 213, S and Z) that cover these regions. The sequences and m/z ratios of the +2 and +3 charge states of these 6 peptides are shown in table 6.2. Table 6.3 shows the diagnostic peptides which would be yielded from a tryptic digestion of each of the 4 variants. If other A1AT peptides are identified, this will not provide any information on which variants are present and detection of them is therefore irrelevant in this context. The challenge for a mass spectrometry method for determination of A1AT variants is to reliably and consistently identify the 6 diagnostic peptides shown in table 6.2.

		m/z monoisotopic	
Peptide	Sequence	$[M+2H]^{2+}$	$[M+3H]^{3+}$
MAla 213	**DTEEEDFHVDQATTVK**	**946.4314**	**631.2900**
MVal 213	**DTEEEDFHVDQVTTVK**	**932.4158**	**621.9462**
M264	**LQHLENELTHDIITK**	**902.4836**	**601.9914**
S264	**LQHLVNELTHDIITK**	**887.4965**	**592.0001**
M 342	**AVLTIDEK**	**444.7555**	**296.8394**
Z342	**AVLTIDK**	**380.2342**	**253.8252**

Table 6.2 The sequences and theoretical m/z values of the 6 diagnostic peptides required to determine variants of A1AT.

Variant	Peptides yielded					
MAla 213	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	AVLTIDEK			
MVal 213	DTEEEDFHVDQVTTVK	LQHLENELTHDIITK	AVLTIDEK			
Heterozygous MAla 213, MVal 213	DTEEEDFHVDQVTTVK	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	AVLTIDEK		
S	DTEEEDFHVDQVTTVK	LQHLVNELTHDIITK	AVLTIDEK			
Heterozygous S, MVal 213	DTEEEDFHVDQVTTVK	LQHLVNELTHDIITK	LQHLENELTHDIITK	AVLTIDEK		
Heterozygous S, MAla 213	DTEEEDFHVDQVTTVK	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	LQHLENELTHDIITK	AVLTIDEK	
Heterozygous S, Z	DTEEEDFHVDQVTTVK	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	LQHLENELTHDIITK	AVLTIDEK	AVLTIDK
Z	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	AVLTIDK			
Heterozygous Z, MAla 213,	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	AVLTIDK	AVLTIDEK		
Heterozygous Z, MVal 213,	DTEEEDFHVDQVTTVK	DTEEEDFHVDQATTVK	LQHLENELTHDIITK	AVLTIDK	AVLTIDEK	

Table 6.3 Diagnostic peptides yielded from a tryptic digestion of different A1AT variants.

The A1AT protein was detected in the LESA-based DBS proteomics analyses described in Chapter 4. See Appendix 4.5. The likelihood of this method to be able to cover the 3

regions was considered by observing the coverage of A1AT detected from a healthy individual by the LESA analysis based digest in Chapter 4. A full sequence diagram can be shown in appendix 6.2.

	Replicate		
	1	**2**	**3**
MAla 213	-	-	-
MVal 213	+	+	+
M264	-	-	+
S264	-	-	-
M342	+	+	+
Z342	-	-	-

Table 6.4 coverage of A1AT of a healthy individual and identifications of diagnostic peptides from DBS digest of healthy donor. + = identification of the peptide - = peptide not identified

Table 6.4 shows the coverage of the A1AT protein obtained from a healthy individual over 3 technical replicates, by means of the LESA based digest described in chapter 4. Not all the expected peptides were identified across all 3 replicates despite the fact that this blood was provided by a presumably healthy patient with a presumably healthy A1AT concentration at around 20-50 µM [227]. Replicates 1 and 2 were missing the M264 peptide, but replicate 3 did detect all the 3 peptides that were expected assuming the A1AT genotype of the donor was homozygous MVal 213.

6.2 Untargeted data dependent LC MS/MS of A1AT deficient DBS digests

The data acquired in chapter 4 had not successfully detected all the peptides that would be expected from a healthy variant of A1AT in all 3 replicates, but it had successfully identified 2 of the 3 peptides that were expected from 2 replicates and all 3 of the expected peptides from the third. Therefore the method seemed worth attempting to identify variants from DBS of A1AT deficient individuals. The genotypes of these individuals were known to be either SZ or ZZ. Three biological replicate DBS were prepared by use of the 1 hour LESA based digestion procedure described in section 4.4 and analysed via the same LC MS/MS method with online desalting and top 7 CID. Data were searched manually for the presence of all the 6 diagnostic peptides and those covering the same regions with 1 or 2 missed cleavages. This term refers to larger peptides that covered the same regions of the protein (213, 342, or 432) but where the digestion had not gone to completion, resulting in a longer peptide with up to two un-cleaved lysine or arginine residues.

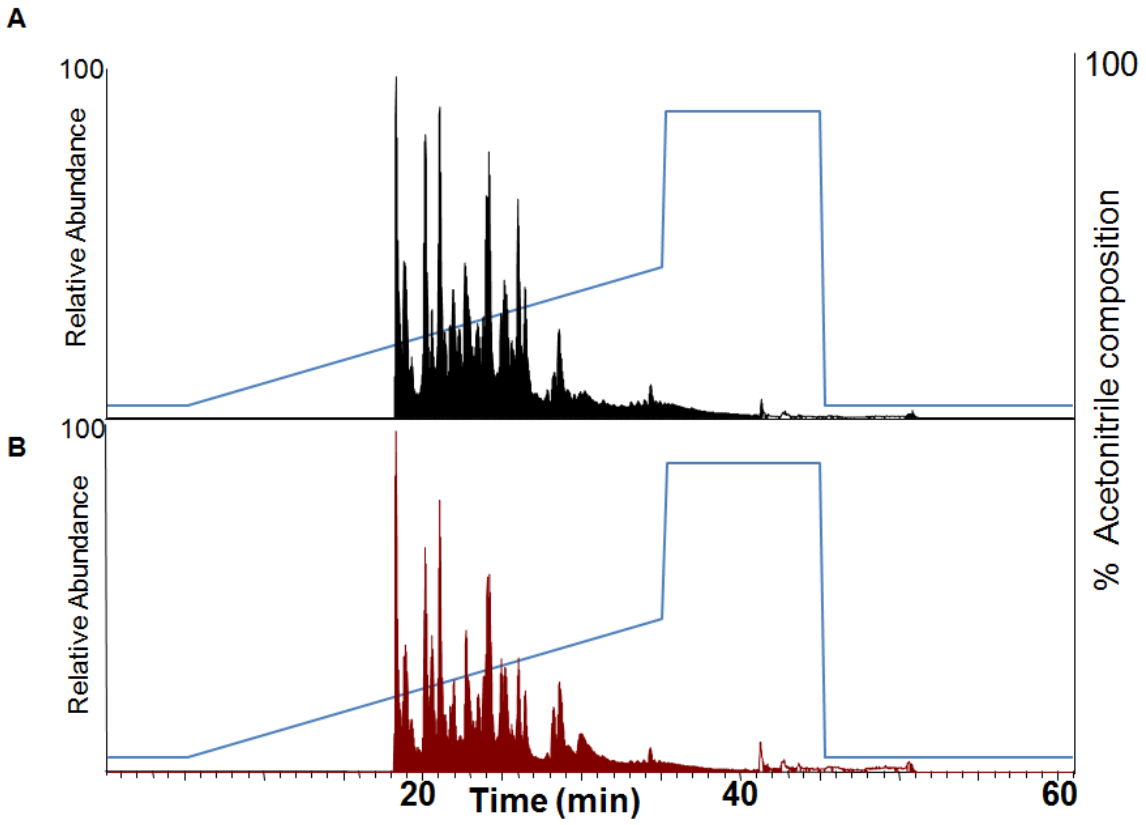


Figure 6.1 Representative total ion chromatograms showing acetonitrile composition obtained following LC MS/MS analysis of DBS digests from A1AT deficient DBS with (A) SZ and (B) ZZ genotypes. Samples were prepared by a 1 hour LESA based trypsin digest

	Sample						Expected	
Peptide	SZ1	SZ2	SZ3	ZZ1	ZZ2	ZZ3	SZ	ZZ
MAla 213	-	+	+	-	-	+	+	+
MVal 213	+	+	+	-	-	-	+	-
M264	+	+	+	+	-	-	+	-
S264	-	-	-	-	-	-	+	+
M342	+	+	+	-	-	-	+	+
Z342	-	-	-	-	-	-	+	-

Table 6.5 Identification of diagnostic peptides from A1AT deficient DBS digests and theoretical sample showing the peptides that would be expected to be identified from the two genotypes. + = identification of the peptide - = no identification of the peptide

Table 6.5 shows which peptides were identified from the samples. Not all the peptides that were expected to be produced by digestion of these variants were identified. More of the expected peptides were identified from the SZ variants than the ZZ variants. This is to be expected as the serum concentration of A1AT is higher in a SZ individual than it is in a ZZ. A trypsin digestion of A1AT from an SZ genotype should be expected to yield all 6 of the diagnostic peptides. From the DBS of the SZ individuals, 4 of the 6 expected peptides were identified in 2 replicates and 3 of the 6 peptides were detected in the other. Digestion of a ZZ genotype should yield the MAla 213 peptide, the M264 peptide and the Z342 peptide. The data from this experiment only resulted in the identification of the M264 peptide from one replicate, the MAla 213 peptide from the other and none of the 6 diagnostic peptides from the final. Other A1AT peptides were detected so the protein can be identified from the sample, but as mentioned previously this cannot provide any information on which variants are present. Appendix 6.2 shows the full sequence coverage of A1AT obtained for each replicate. Determinination of protein variants by data dependent LC MS/MS of a tryptic digest faces a number of challenges. The method fragments the most abundant peptides observed in the survey scan and in a complex sample like blood the number of peptides present in the sample exceeds the scanning capabilities of the instrument [122]. Lower abundant peptides will not be selected for fragmentation. When comparing the data from this experiment to the data in section 6.1, it becomes clear that the identification of the diagnostic peptides is becoming less successful as the concentration of the protein drops from the healthy donors, to the SZ genotypes and then on to the ZZ individuals. Untargeted data dependent acquisition

favours identification of abundant proteins, and because it could not successfully determine the genotypes of A1AT deficient individuals, targeted approaches were deemed more suitable and selected for further development.

6.3 Targeted data dependent LC MS/MS through inclusion lists

In a data dependent LC MS/MS experiment, an inclusion list can be used to select ions of specific m/z values for fragmentation prior to other ions regardless of their relative abundance peaks. An inclusion list was used here to attempt to identify A1AT variants from DBS digests of A1AT deficient individuals. Samples were digested as described in section 6.1, and then analysed with the same LC parameters but a different MS/MS method. The mass spectrometry method comprised a full mass spectrum survey scan in the orbitrap, which was followed by up to 7 fragmentation events in the ion trap in the event that any precursor m/z from a series of peptides added to an inclusion list were detected. The inclusion list was composed of the m/z values shown in table 6.2 and additional peptides covering the same regions with 1 or 2 missed cleavages. The full list of m/z values added to the inclusion list can be found in table 6.6. Data were acquired in duplicate.

	Sequence	$[M+2H]^{2+}$	$[M+3H]^{3+}$
MAla 213	**GKWERPFEVKDTEEEDFHVDQATTVK**	**1560.749**	**1040.835145**
	WERPFEVKDTEEEDFHVDQATTVK	**1468.191**	**979.12967**
	WERPFEVKDTEEEDFHVDQATTVKVPMMK	**1761.339**	**1174.56204**
	DTEEEDFHVDQATTVK	**932.4158**	**621.946282**
	DTEEEDFHVDQATTVKVPMMK	**1225.564**	**817.378652**
	DTEEEDFHVDQATTVKVPMMKR	**1303.615**	**869.412355**
MVal 213	**GKWERPFEVKDTEEEDFHVDQVTTVK**	**1574.765**	**1050.178912**
	WERPFEVKDTEEEDFHVDQVTTVK	**1482.207**	**988.473437**
	WERPFEVKDTEEEDFHVDQVTTVKVPMMK	**1775.355**	**1183.905807**
	DTEEEDFHVDQVTTVK	**946.4314**	**631.290048**
	DTEEEDFHVDQVTTVKVPMMK	**1239.58**	**826.722418**
	DTEEEDFHVDQVTTVKVPMMKR	**1317.631**	**878.756122**
M264	**YLGNATAIFFLPDEGKLQHLENELTHDIITK**	**1770.922**	**1180.950621**
	LQHLENELTHDIITK	**902.4836**	**601.991497**
	LQHLENELTHDIITKFLENEDR	**1354.188**	**903.127712**
	LQHLENELTHDIITKFLENEDRR	**1432.238**	**955.161416**
	YLGNATAIFFLPDEGKLQHLENELTHDIITKFLENEDR		**1482.086836**
	LSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITK		**1533.483762**
S264	**YLGNATAIFFLPDEGKLQHLVNELTHDIITK**	**1755.935**	**1170.959228**
	LSSWVLLMKYLGNATAIFFLPDEGKLQHLVNELTHDIITK		**1523.492369**
	LQHLVNELTHDIITK	**887.4965**	**592.000104**
	LQHLVNELTHDIITKFLENEDR	**1339.201**	**893.136319**
	LQHLVNELTHDIITKFLENEDRR	**1417.251**	**945.170023**
	YLGNATAIFFLPDEGKLQHLVNELTHDIITKFLENEDR		**1472.095443**
M342	**LSKAVHKAVLTIDEK**	**826.4907**	**551.329562**
	AVHKAVLTIDEK	**662.3852**	**441.925877**
	AVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVK	**1782.446**	**1188.633247**
	AVLTIDEK	**444.7555**	**296.839409**
	AVLTIDEKGTEAAGAMFLEAIPMSIPPEVK	**1564.817**	**1043.54678**
	AVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTK		**1655.866667**
Z342	**AVLTIDKKGTEAAGAMFLEAIPMSIPPEVK**	**1564.343**	**1043.230903**
	AVLTIDKK	**444.2817**	**296.523532**
	AVHKAVLTIDKK	**661.9114**	**441.61**
	AVHKAVLTIDK	**597.8639**	**398.911679**
	LSKAVHKAVLTIDIK	**818.5114**	**546.010052**
	AVLTIDK	**380.2342**	**253.825211**

Table 6.6 m/z values added to an inclusion list to identify the 6 diagnostic peptides required for determining A1AT variants.

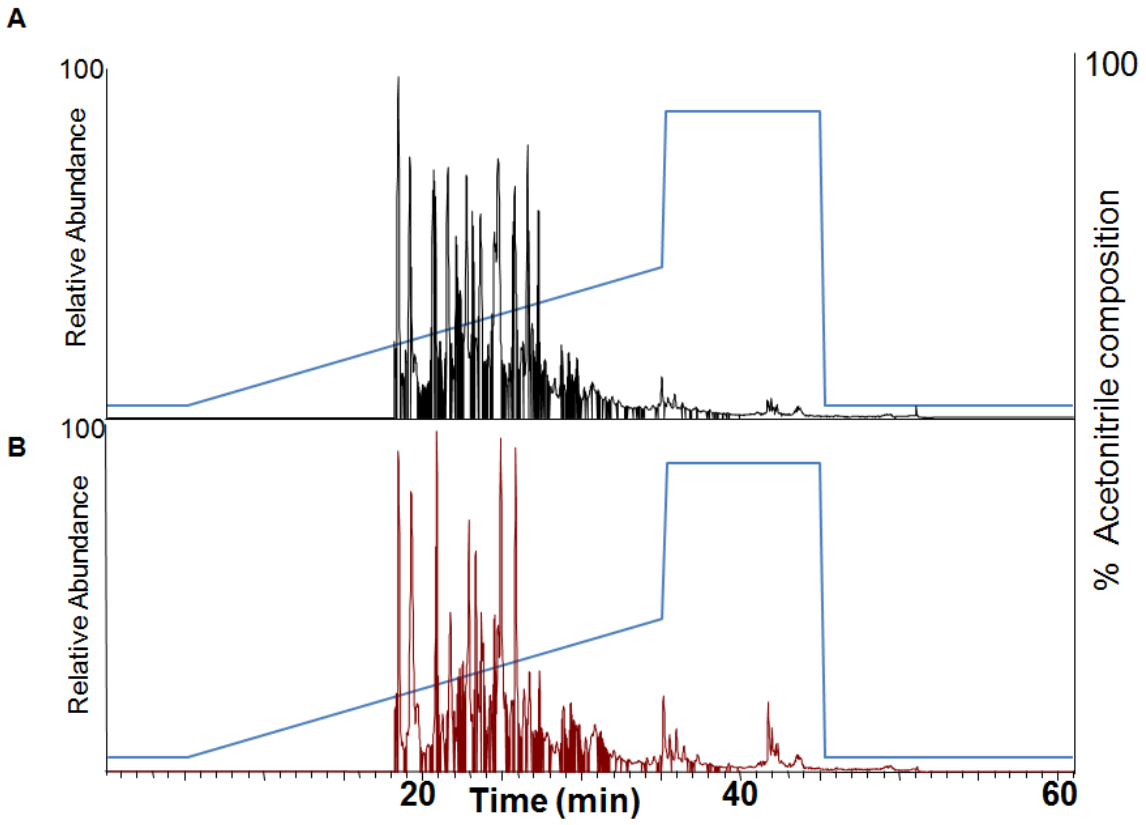


Figure 6.2 Representative total ion chromatograms showing acetonitrile composition obtained following LC MS/MS analysis, incorporating the inclusion list described in table 6.6, of DBS digests from A1AT deficient individuals from (A) SZ and (B) ZZ A1AT genotypes.

	Sample				Expected	
Peptide	SZ1	SZ2	ZZ1	ZZ2	SZ	ZZ
MAla 213	-	+	-	+	+	+
MVal 213	-	+	-	-	+	-
S264	-	-	-	-	+	-
M264	-	-	-	-	+	+
M342	-	+	-	-	+	-
Z342	-	-	-	-	+	+

Table 6.7 Identification of diagnostic peptides from A1AT deficient DBS digests by LC MS/MS analysis with an inclusion list. + = identification of the peptide - = no identification of the peptide

Table 6.7 shows that the peptides necessary for diagnosis of the variant were not identified by LC MS/MS analysis incorporating an inclusion list. More peptides were detected from the SZ variant than the ZZ, which is to be expected due to the higher serum concentrations of A1AT associated with the SZ genotype. The mass spectrometer was only fragmenting the targeted diagnostic peptides and not others, so when none of the targeted peptides were detected, the protein could not be identified. As discussed above, one issue preventing the detection of the other peptides may be the lack of sensitivity. An individual with a ZZ phenotype may have as little as 2.5 µM of A1AT in their serum. Only 20 µl of blood is applied to a DBS, which would amount to 50 pmol of protein in the whole spot in a ZZ individual. LESA only samples a small area of the DBS surface and as the extraction efficiency is likely to be low anyway, the total amount of A1AT protein in the digest is likely to be very low indeed. To improve detection of the diagnostic peptides, a more sensitive and selective means of analysis was needed.

6.4 Multiple reaction monitoring analysis for determining A1AT variants

SRM assays and the analogous multiple reaction monitoring assays are considered to be the most sensitive mass spectrometry techniques and can be targeted to individual analytes. SRMs are more commonly used for quantification than identification, but it is possible to directly measure the abundance of specific peptides to determine their presence. An LC MRM method is a viable option for determining A1AT variants by using

with 6 different SRMs to detect each of the six diagnostic peptides shown in table 6.2. Presence of the diagnostic peptide in the digest would be indicated by a signal in the SRM targeted to the peptide and absence of the peptide would be indicated by absence of signal. The approach is similar to the method described by Chen *et al.* [229] with the exception that SRMs were used to measure the intensity of endogenous peptides only and not the amount relative to isotopically labelled internal standard peptides spiked into the sample.

6.4.1. Analysis of synthetic peptides through direct infusion MRM analysis

Prior to analysing DBS samples from A1AT deficient individuals, the SRM methods were validated by direct infusion electrospray ionisation of the targeted peptides. Synthetic analogues of each of the 6 diagnostic peptides shown in table 6.2 were analysed individually, along with one sample composed of a mixture of all 6 peptides and another sample in which the six peptides were spiked into a proteolytic digest standard. The digest standard comprised a tryptic digest of a mixture of six proteins (cytochrome C, lysozyme, alcohol dehydrogenase, serum albumin, apotransferrin and β galactosidase) at a concentration of 50 fmol/µl. The individual peptides were made up in a 0.1% formic acid solution to a concentration of 1 mg/ml and in the mixtures, each peptide was at a concentration of 0.1 mg/ml. Both these concentrations were significantly higher than those expected in a DBS digest of an A1AT deficient individual. The concentration of A1AT in the serum of a ZZ individual is 527 fold lower than the concentration of the peptide used in this experiment, but this experiment was designed to validate the SRMs only. Ten

µl aliquots of each sample were purified with C_{18} zip tips as described in section 2.2.5 prior to analysis. After drying down, samples were resuspended in 49.95/49.95/0.1% water/ acetonitrile/ formic acid and samples were analysed by direct infusion with an MRM method composed of 3 transitions for each of the 6 diagnostic peptides. The precursor masses and the three transitions used to detect each peptide are shown in table 6.8, and the method is fully described in section 2.2.3.3.

Peptide	Sequence	Selected Precursor m/z	Transitions			Collision energy (v)
MAla 213	**DTEEEDFHVDQATTVK**	[M+3H] 621.94	y_9 998.52	y_{10} 1145.59	y_{11} 1260.62	35
MVal 213	**DTEEEDFHVDQVTTVK**	[M+3H] 631.29	y_7 790.43	y_8 889.49	y_9 1026.55	35
M264	**LQHLENELTHDIITK**	[M+2H] 902.48	y_8 940.54	y_9 1069.58	y_{10} 1183.63	33
S264	**LQHLVNELTHDIITK**	[M+2H] 887.49	y_9 1069.58	y_{10} 1183.63	y_{11} 1282.70	33
M342	**AVLTIDEK**	[M+2H] 444.75	y_4 504.26	y_5 605.31	y_6 718.39	30
Z342	**AVLTIDK**	[M+2H]380.23	y_3 375.22	y_4 476.41	y_5 589.35	30

Table 6.8 Precursor ion masses and the 3 transitions monitored in the MRM assay for detection of the 6 diagnostic peptides.

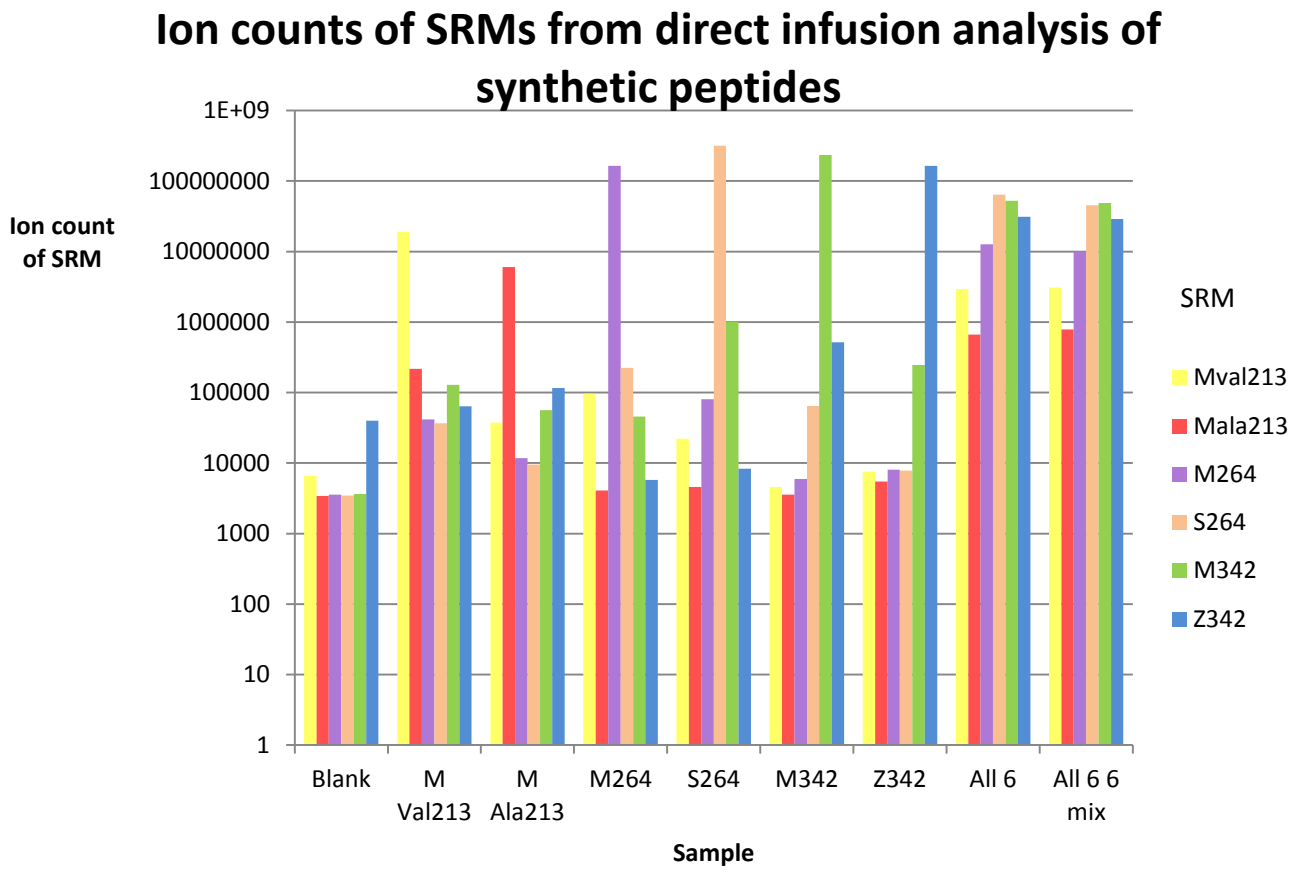


Figure 6.3 Ion counts for each diagnostic peptide measured by direct infusion electrospray MRM.

Fig 6.3 shows the ion counts of each of the SRMs for each different peptide. When infusing the individual peptides, signal was detected by the correct SRM to show the peptide was present, which is demonstrated by the spectra in fig 6.4. Interference was encountered in some samples. In certain cases signals were observed for peptides that were not present in the sample. For example infusing the Z342 peptide shows a signal for the M342 peptide, and infusing the S264 peptide also showed a signal for the M342. Fig 6.5 demonstrates this interference.

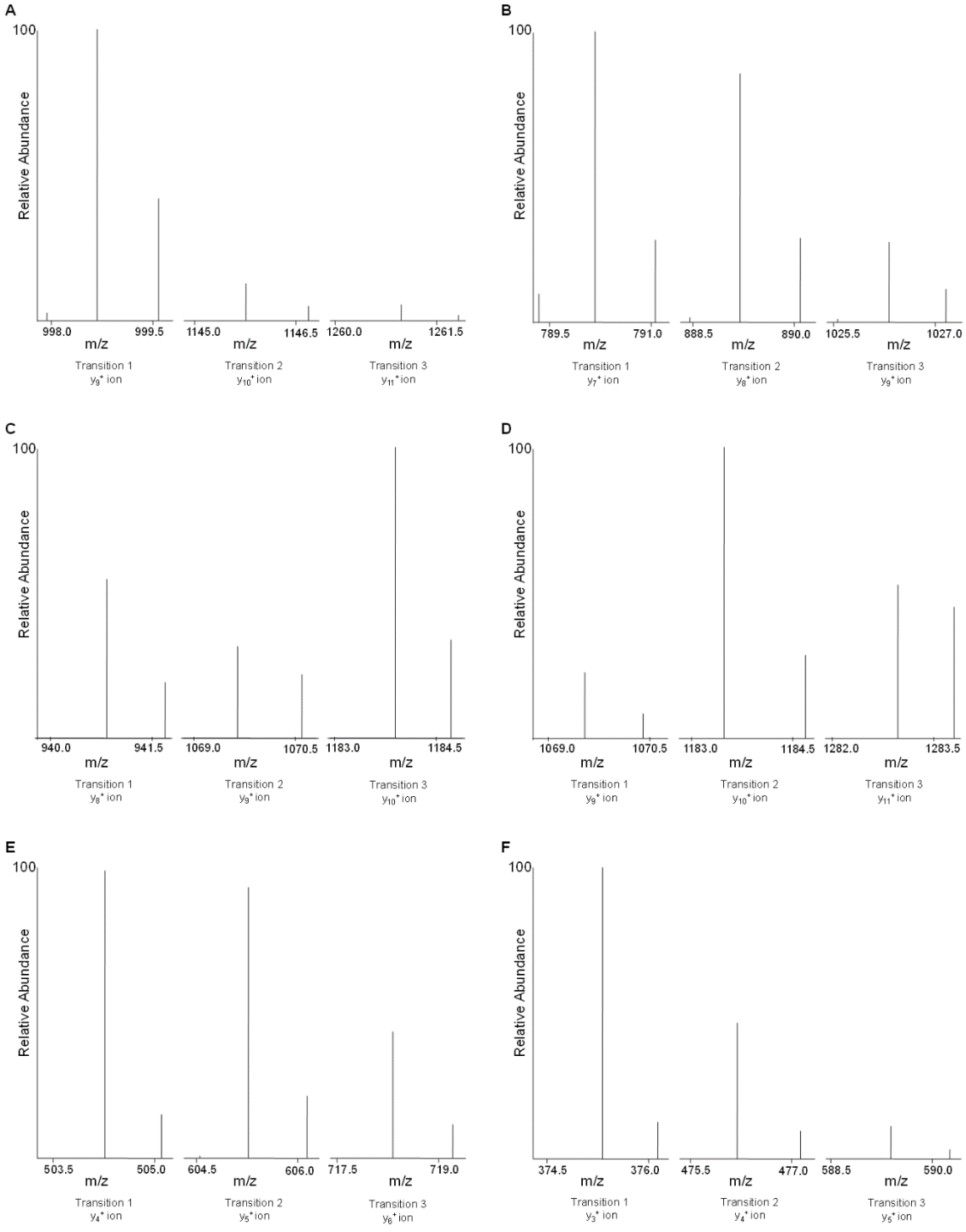


Figure 6.4 SRM spectra showing detection of synthetic peptides. (A) MAla 213 SRM showing signal for MAla 213 peptide. (B) MVal 213 SRM showing signal for MVal 213 peptide (C) M264 SRM showing signal for M264 peptide (D) S264 SRM showing signal for S264 peptide (E) M342 SRM showing signal for M342 peptides (F) Z342 SRM showing signal for Z342 peptide.

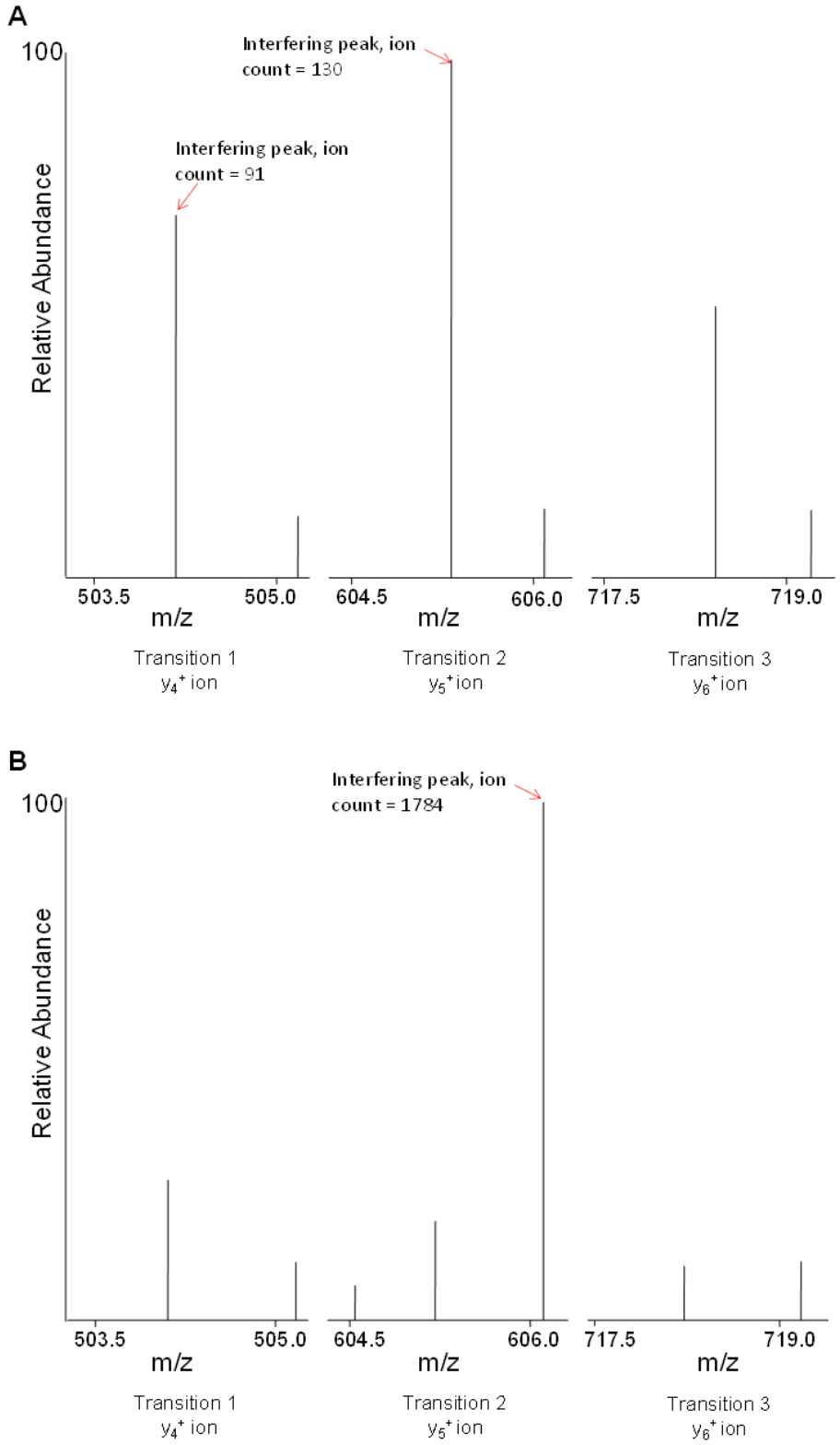


Figure 6.5 interfering peaks observed by direct infusion SRM assay, (A) transition monitoring presence of the M342 peptide whilst infusing the Z342 peptide. Interfering peaks are shown with ion counts of 130 and 91. (B) Transition monitoring the M342 peptide whilst infusing the S264 peptide. Spectrum shows an interfering peak with an ion count of 1784

The data provided by direct infusion electrospray MRM analysis may not be robust enough to indicate the presence of variants in an unknown sample. A better means of analysis may be to compare the intensity of each SRM relative to the intensity of the SRM measured in a blank sample, see Fig 6.6.

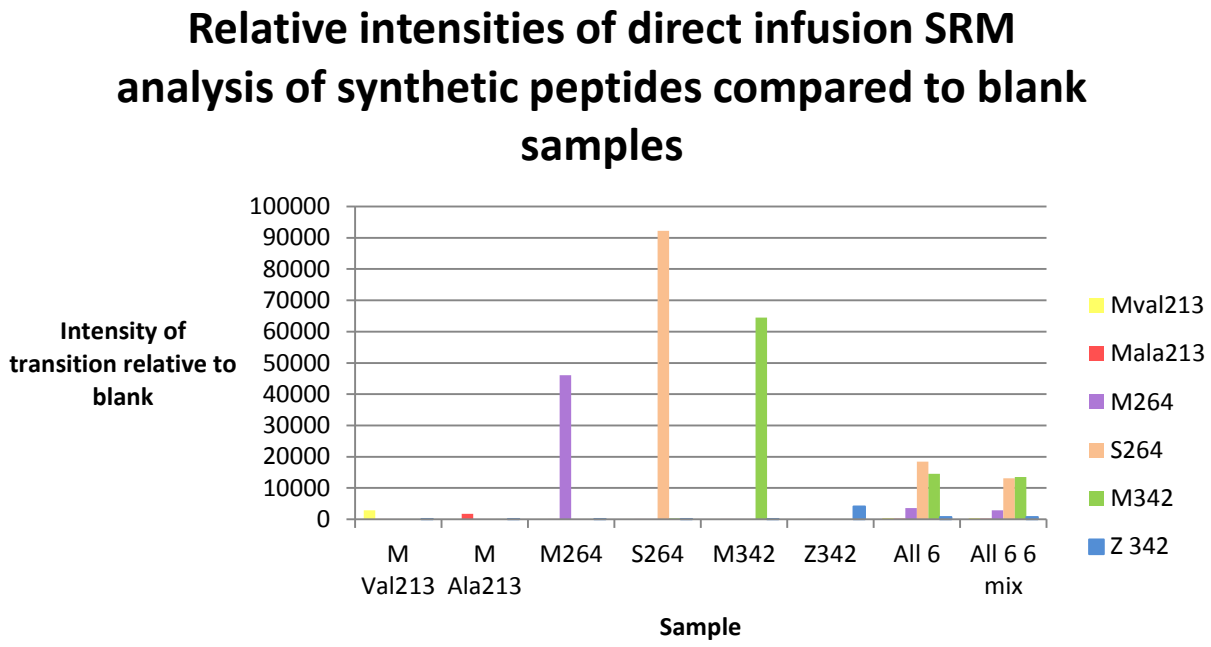


Figure 6.6 Relative intensities of diagnostic peptides compared to blank samples, measured detected by direct infusion MRM.

Fig 6.6 shows the ion counts of each of the SRMs from each sample compared to the ion counts measured by the SRMs when infusing blank solvent. The ion counts of SRMs of the M264, S264 and M342 were over 40,000 fold higher than the signal measured for the blank sample, indicating that these SRMs are very sensitive to their targeted peptides. The ion counts of the MVal 213, MAla 213 and Z342 SRMs were over 1000 fold higher than the ion counts measured from the blank sample, but were less sensitive to their

targeted peptides than the SRMs for the M264, S264 and M342 peptide. Lack of sensitivity of the SRM could be a problem for detecting the low abundance peptides from the blood of A1AT deficient individuals. Comparing the signal measured by an SRM to a blank sample did help avoid problems of interference.

6.4.2 LC MRM analysis of DBS digests

Liquid chromatography concentrates low abundance species from a mixture into a short chromatographic time window, which could help improve the sensitivity of the MRM assay used in section 6.4.1. An LC MRM method was used to attempt to determine the variants of the A1AT protein from a series of DBS digests from A1AT deficient individuals and healthy donors to act as controls. Three different biological replicates were used for each genotype of the A1AT deficient variants, and the control samples were technical replicates from the same donor. The exact genotype of the healthy control was not known. Samples were digested as described in section 6.1 and analysed by the LC method described in section 2.2.2 and the mass spectrometry MRM method described in section 2.2.3.3. The LC MS/MS method did not include an online desalting stage so samples were purified with C_{18} zip tips as described in section 2.2.5 prior to analysis.

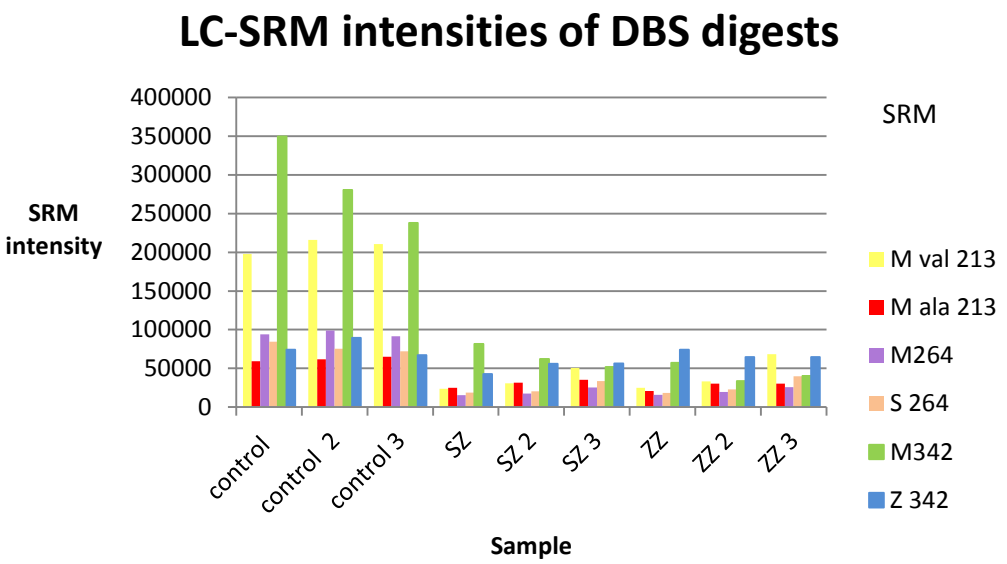


Figure 6.7 Ion counts of each of the 6 diagnostic peptides measured by LC MRM analysis from DBS digests of healthy controls and A1AT deficient individuals.

Fig 6.7 shows that the LC MRM method did not detect the peptides that would be expected to be present in -either the variant or the healthy control samples. The control blood samples show strong signals for the MVal 213 and M342 peptides but no others and signals for all the A1AT deficiency samples are very weak. The ion counts of the SRMs were then compared to those measured from blank samples, shown in fig 6.8.

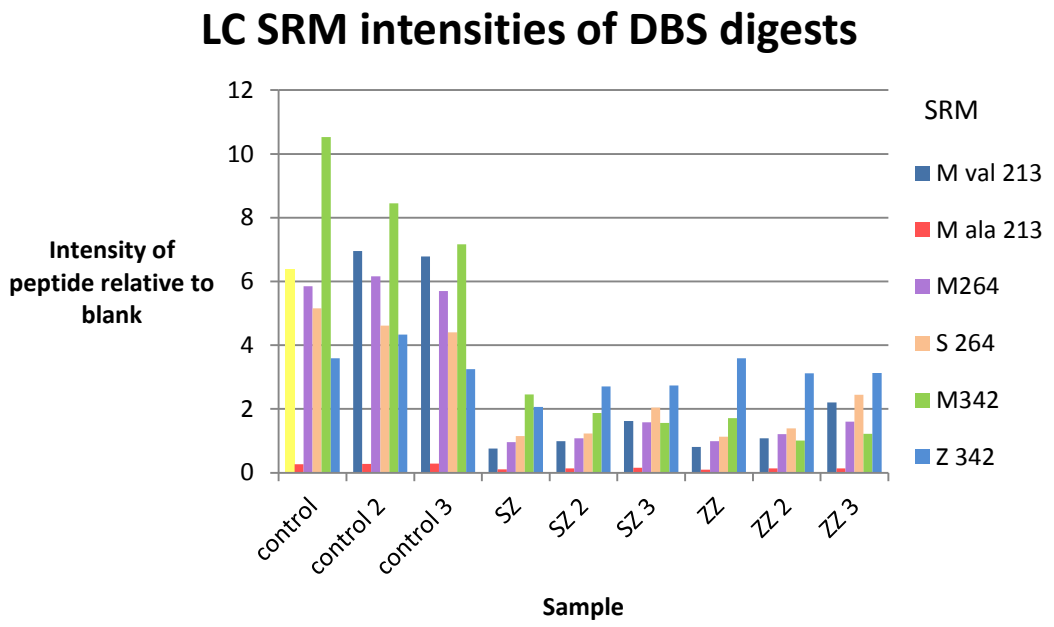


Figure 6.8 Relative intensities of diagnostic peptides compared to blank samples measured from DBS digest of healthy controls and A1AT deficient individuals by LC MRM assay.

Comparing the intensities of the diagnostic peptides from DBS digests relative to the blank samples cannot determine which variants of A1AT are present in the blood. The most intense signal was an M342 peptide detected from a healthy control, which was only 10.53 fold higher than the ion count measured from a blank sample. This signal is not strong enough signal to reliably confirm the presence of the peptide in the digest. In comparison, the ion counts measured from the synthetic peptides were all over 1000 fold more intense than those acquired from the blank samples. Additionally the results suggest that the Z342 and S264 peptides were also present in the control patients, which is unlikely to be the case. Signals for the healthy controls had a greater intensity relative to blank than any of the diseased variants. The increase in signal intensity in healthy controls does suggest that the instrument response increases in concordance with the concentration of the protein. The results suggests that the SRMs were not sensitive

enough to detect the correct diagnostic peptides from DBS digests of healthy controls, so it would be unable to determine the variant of A1AT from a diseased individual. SRM analysis of the variant samples did not identify the correct peptides that would be expected from their genotypes. The ZZ variants did show a small signal for the Z342 peptide, but also showed signals corresponding to the S264 peptides and the M342 peptides as well. Neither the S264 nor M342 peptides would be expected to be present in a ZZ genotype individual. The SZ type individuals would be expected to show that all 6 diagnostic peptides were present. Fig 7.6 suggests that all peptides other than the MAla 213 peptides were present. The ion counts for the observed peptides were approximately two fold greater than the ion counts measured from a blank sample, so the presence of these peptides cannot be reliably confirmed from this data.

6.5 Chapter 6 conclusion

Several different methods have been attempted for to determine genotypic variant of the A1AT protein from DBS. The methods attempted included untargeted data dependent LC MS/MS, targeted data dependent LC MS/MS and targeted LC MRMs. None of these methods were able to determine which variant of A1AT was present from a DBS. Determining the variant of a proteins is a challenge for a bottom up proteomics technique, it requires 3 different regions of the protein to be consistently and reproducibly identified. The amount of A1AT expressed by a deficient individual can be as low as 2.5 µM, and identifying 3 regions of a protein at such a low concentration in a

mixture as complex as blood was difficult. The coverage of A1AT could be improved by incorporating a depletion strategy as demonstrated in section 5.3. By using hemovoid to deplete the haemoglobin proteins, the analysis was more sensitive to the high abundant plasma proteins which could be investigated to increase the coverage of A1AT and determine the variants.

Chapter 7 Analysis of non-covalent protein complexes by liquid extraction surface analysis mass spectrometry

The work presented in the final chapter of this thesis was aimed at demonstrating the use of LESA for analysing non-covalent protein complexes from dried surfaces. The field of native mass spectrometry is becoming a commonplace technique for the analysis of non-covalent complexes. Native mass spectrometry has been carried out using solution phase ESI which was discussed in detail in section 1.2.5. Prior to this work, no ambient ionisation techniques have been applied to the analysis of protein complexes directly from surfaces. Zhang *et al.* investigated paper spray ionisation, but the group have not been able to analyse surface associated analytes. The method was only capable of analysing complexes prepared in solution [230]. LESA is a possible tool for native mass spectrometry because, unlike some other ambient ionisation techniques it forms ions by conventional electrospray ionisation which is soft enough to retain fragile interactions. Solution phase native mass spectrometry is not normally used to analyse surface associated proteins so the additional challenge for a LESA approach is that the technique must be able to retain interactions from dried proteins. Drying a protein denatures its tertiary structure and disrupts non-covalent interactions [231]. In order for these non-covalent complexes to be retained into the gas phase they must either be strong enough to survive in the drying process or reform upon sampling with LESA. Much of the work shown in this chapter has been accepted for publication as an article for the Journal of the American Society for Mass Spectrometry [232].

7.1 Protein standards

A range of lypholised protein standards dried onto surfaces were analysed. The standards comprised two small well-defined proteins: equine skeletal muscle myoglobin and human haemoglobin. Myoglobin forms a non-covalent complex with a prosthetic haem cofactor and haemoglobin is comprised of a non-covalent protein heterotetramer composed of 2 alpha and 2 beta subunits each of which is non-covalently bound to a haem group. The crystal structures of these proteins can be seen in fig 7.1.

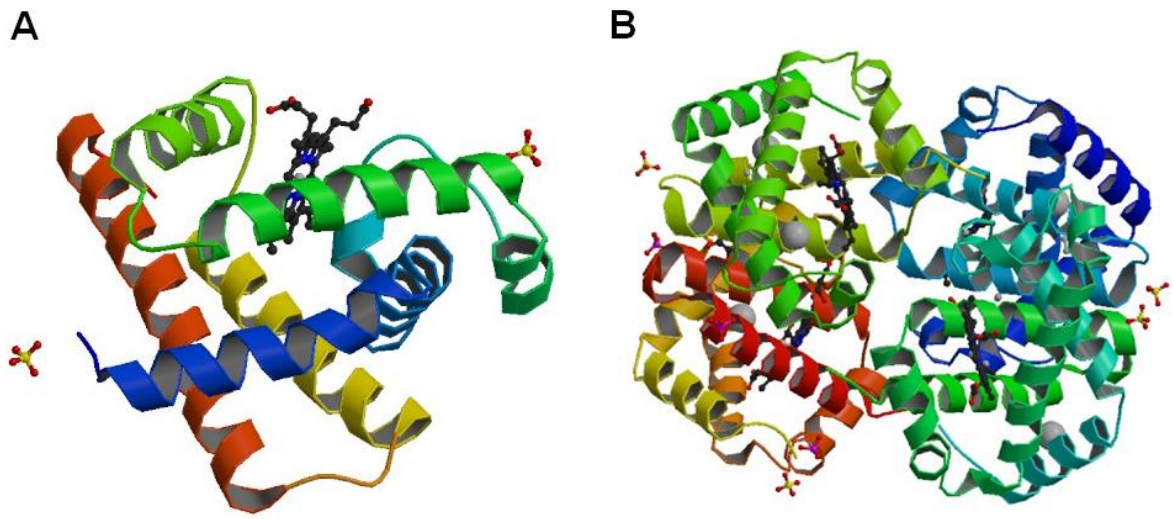


Figure 7.1 Crystal structures of (A) equine myoglobin obtained at 1.45 Å resolution, generated using PQS software [233] (B) human haemoglobin tetramer obtained at 1.8 Å resolution, generated using PISA software [234] Both Images Courtesy of Protein Data Bank

Proteins were dried onto a surface and then analysed with LESA mass spectrometry. LESA analyses were validated against direct infusion ESI. Glass microscope slides and

PVDF membranes were used as model surfaces as LESA MS of intact proteins had previously been demonstrated from these surfaces as shown in chapter 3. Five microlitre aliquots of a 100 µM aqueous solution of each protein was spotted onto the surface, leaving 500 pmol of protein. The spots were left to dry for 4 hours before analysis. Unless otherwise stated the extraction solvent was composed of 10 mM NH_4AC, pH balanced to 6.8. In direct infusion ESI, each protein was electrosprayed at a concentration of 10 µM in 10 mM NH_4AC (pH 6.8). Mass spectrometry analysis was performed on two different instruments: the Thermo Fisher Orbitrap Velos™ and the Waters Synapt™ G2-S. In the LESA extraction process, 7 µl of solvent was extracted from the solvent well, 4 µl were then deposited from a height of 1.6 mm onto the surface (Orbitrap) or from 0.6 mm above the surface (Synapt™). After a delay of 3 seconds, 4.5 µl was reaspirated. The cycle was repeated once before being infused into the mass spectrometer. Samples were analysed in full scan mode on both instruments as described in section 2.2.3.1. Data were acquired for a period of 5 minutes on the Synapt™ and 10 minutes (each scan was comprised of 50 microscans) on the Orbitrap Velos™.

7.1.1 Myoglobin

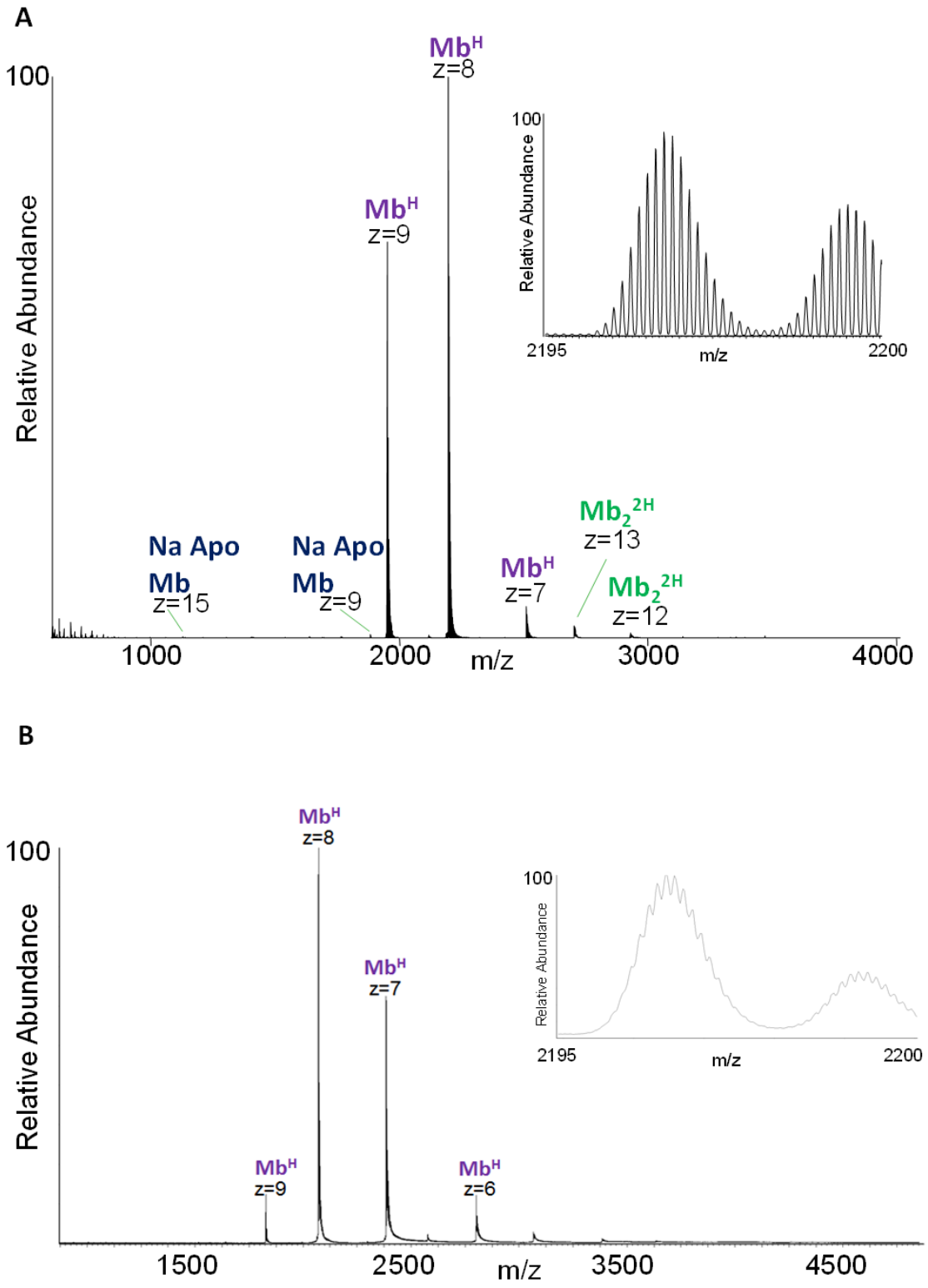


Figure 7.2 Direct infusion native ESI mass spectrum of equine skeletal muscle myoglobin magnification of the +8 charge state of the myoglobin-heme complex obtained on (a) Orbitrap Velos™ (B) Synapt™ G2 S.

Fig 7.2 shows the native mass spectra of myoglobin when analysed by direct infusion ESI. The holomyoglobin complex was retained in the gas phase in both instruments, and the + 8 charge state was the most abundant peak in both cases. The mass spectrum recorded on the Orbitrap Velos™ shows peaks corresponding to holomyoglobin in the +9, +8 and +7 charge states at m/z 1952.80, 2196.77, and 2510.45 respectively. The theoretical m/z for the the +9, +8 and +7 charge states are 1952.87, 2196.85 and 2510.54 (each m/z both theoretical and measured corresponds to the most abundant isotope). The mass spectrum recorded on the Orbitrap Velos™ also showed evidence of dimerization of myoglobin, with peaks at m/z 2703.15 and 2928.82. Mass spectra showing evidence of the formation of myoglobin dimers have been reported elsewhere [221]. The dimers are likely to be an artefact of the ESI process as myoglobin only exists as a monomer in its native environment [220]. The mass spectrum recorded on the Waters Synapt™ showed peaks corresponding to the + 9, +8 +7 and +6 charge states of the holomyoglobin complex at m/z 1952.70, 2196.48, 2509.68 and 2927.00. The theoretical m/z for 9+, +8, +7 and and +6 charge states are 1952.87, 2196.85, 2510.54 and 2928.80. Once again each m/z corresponds to the most abundant isotope. The different isotopologues could not be distinguished from the mass spectra acquired using the Synapt™, unlike those acquired using the Orbitrap Velos™. Sodium adducts were observed in addition to the protonated species on both instruments. The presence of sodiated peaks are a common problem in native mass spectrometry. Non-specific cation adducts form alongside the protonated species during ESI [235].

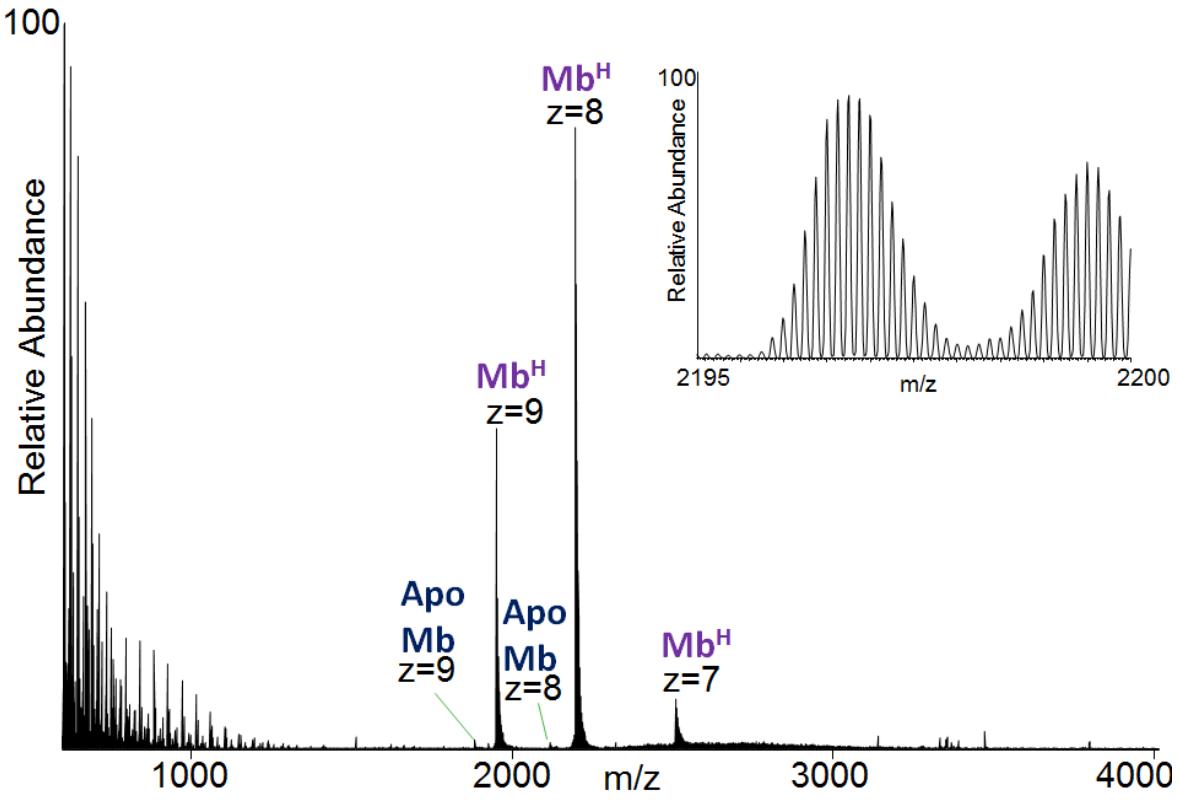


Figure 7.3 Native LESA mass spectrum of equine myoglobin and magnification of the +8 Charge state of the myoglobin-heme complex obtained from glass microscope slide acquired on the Orbitrap Velos™

The holomyoglobin complex was observed following native LESA sampling of a glass microscope slide as seen in the mass spectrum shown in fig 7.3. The holomyoglobin complex was observed in the + 9, +8 and + 7 charge states at m/zs 1952.78, 2196.75 and 2510.42. The theoretical m/z of these charge states were 1952.87, 2196.85, 2510.54 and as previously all of which refer to the most abundant isotope. The mass spectrum shows polymeric contaminants in the low m/z range and the protein peaks show evidence of sodiated adducts, but otherwise appear very similar to when analysed by direct infusion. Sodium adducts and other cationated species of the complex appear often when analysing non-covalent complexes. This is a result of the electrospray conditions [235]. This takes place during ion evaporation, where the concentration of salts can accumulate at the liquid air boundary of the charged droplets in an ESI source. The salts are the liable

to undergo ion paring with the recently formed protein ions [236]. The dimer artefact was absent from the native LESA MS analysis of myoglobin on glass.

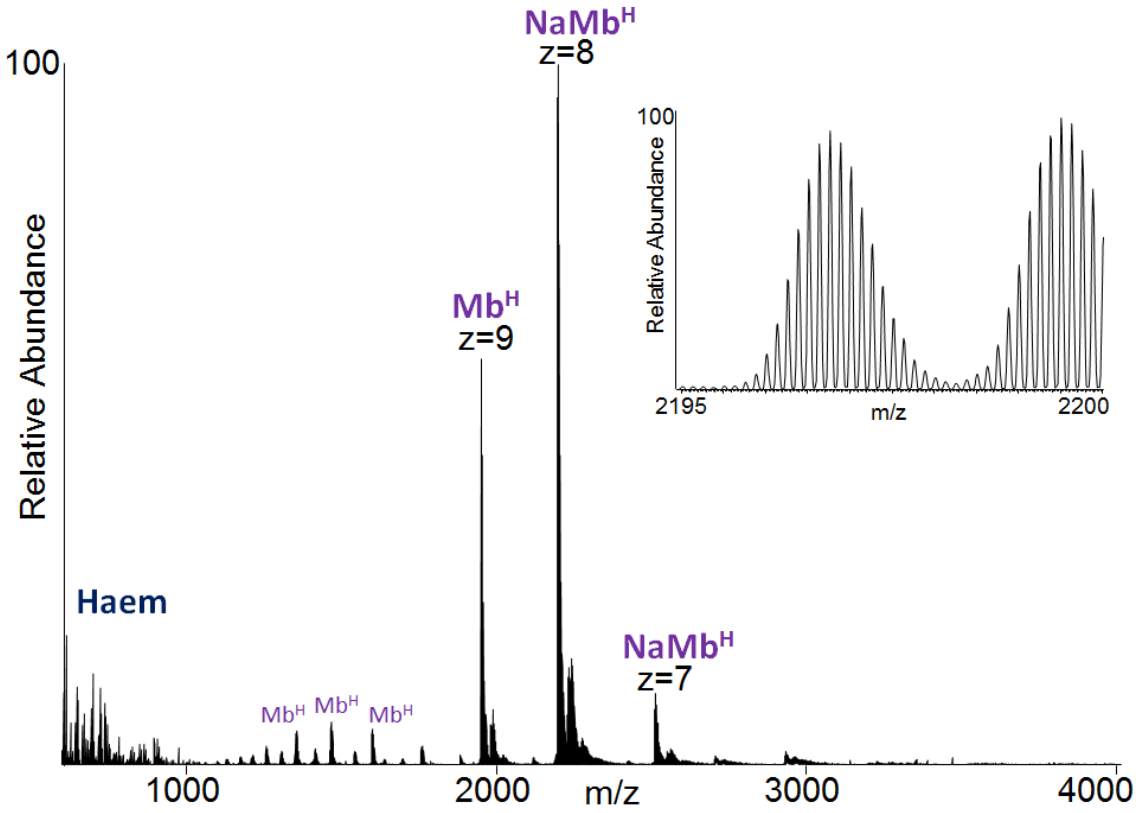


Figure 7.4 Native LESA mass spectrum of equine myoglobin and magnification of the +8 charge state of the myoglobin-heme complex obtained from PVDF membrane acquired on the Orbitrap Velos™

LESA mass spectrometry of the intact myoglobin complex was less successful from PVDF than from glass, see Fig 7.4. The LESA mass spectrum shows peaks corresponding to apomyoglobin and a peak corresponding to the haem group at m/z 616.17. The most intense peaks were the sodiated forms of the holomyoglobin complex, which was observed in the +8 and + 7 charge states at m/zs 2199.51 and 2513.58. The protonated

species was also observed in the + 9 charge state at m/z 1952.79, but was less intense than the + 8 charge state of the sodiated adduct. The results suggest that PVDF is less suitable for native LESA mass spectrometry than the glass microscope slide. PVDF is a very hydrophobic surface which can denature proteins as a result of interactions between the hydrophobic amino acid residues and the surface [237]. There were also more sodium adducts detected. The monosodiated species was more abundant than the protonated adducts of the + 8 and + 7 charge state. Whilst not as successful as the analysis from glass microscopes slides, this mass spectrum demonstrates that non-covalent complexes can be extracted from different types of surface.

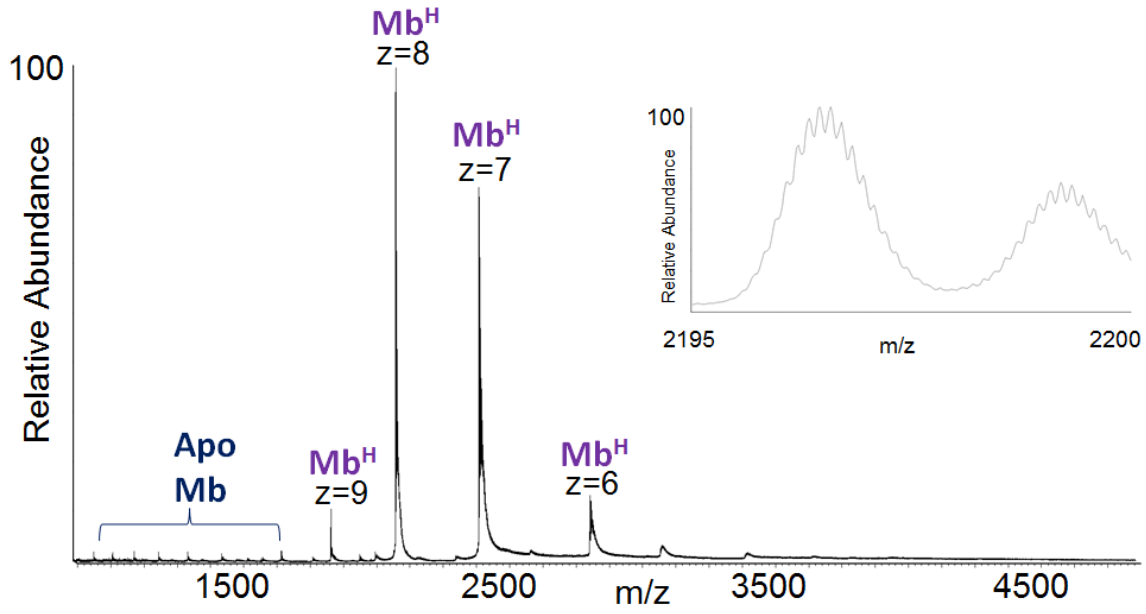


Figure 7.5 Native LESA mass spectrum of equine myoglobin and magnification of the +8 charge state of the myoglobin-heme complex obtained from glass microscope slide acquired on the Synapt™ G2S

Fig 7.5 shows the LESA mass spectrum of myoglobin acquired from the glass microscope slide using the Synapt™. This mass spectrum is very similar to that acquired from the

Orbitrap Velos™, see fig 7.3. The holomyoglobin complex was observed in the +9, +8 +7, +6 charge states at m/z 2927.06, 2509.66, 2196.88 and 1952.80. The theoretical m/z of these peaks are 2928.80, 2510.54, 2196.85 and 1952.87. The holomyogoblin species were by far the most abundant peaks, but there is evidence of denaturation as well. The charge state distribution of the holomyoglobin has shifted slightly towards lower charge states when compared to the mass spectrum obtained on the Orbitrap Velos™. The most abundant peak in the mass spectrum obtained on the Orbitrap Velos™ (fig 7.3) was the +8 peak, followed closely by the + 9 and with only a small peak corresponding to the +7 charge state. On the Synapt™, the +8 charge state was the most abundant followed by the +7 and then the + 6. The variation in the charge state distribution is likely to be due to the difference in the temperature of the ion transfer tube of the two instruments. The Orbitrap's ion transfer tube is heated to 250°C whereas the Synapt's™ is 30°C. Higher capillary temperatures lead to more efficient desolvation of the protein complex during ESI which can lead to more efficient protonation and the production of higher charge states [97, 238, 239].

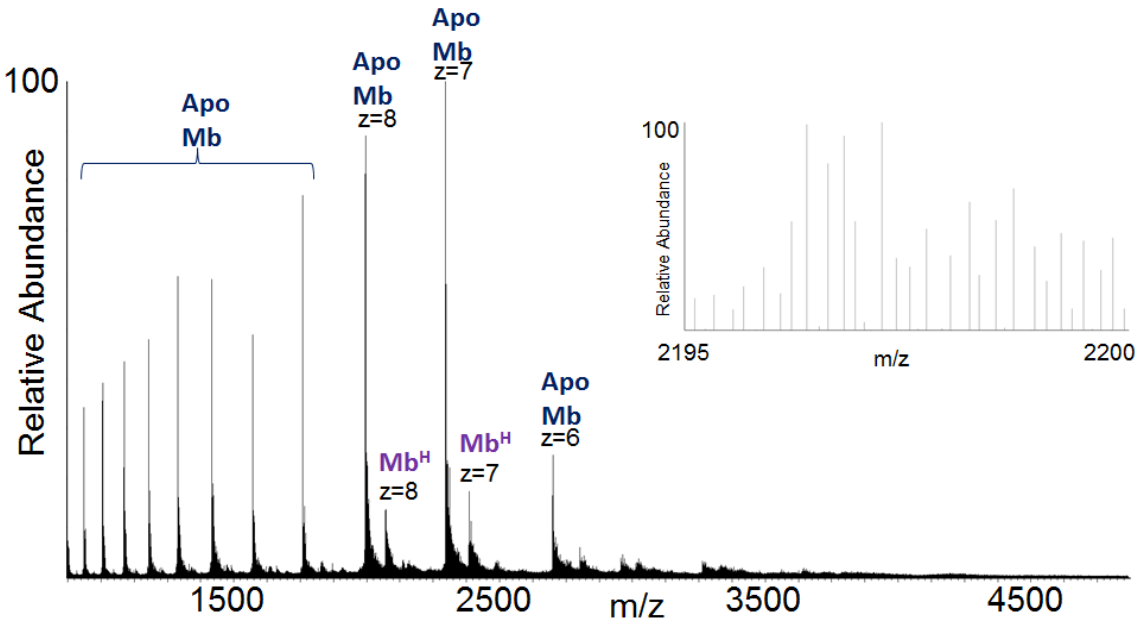


Figure 7.6 Native LESA mass spectrum of equine myoglobin and magnification of the +8 charge state of the myoglobin-heme complex obtained from PVDF membrane slide acquired on the Synapt™ G2S

Fig 7.6 shows the native LESA mass spectrum obtained from myoglobin spots on PVDF using the Synapt™. It is clear there has been significant denaturation of the complex. The extent of denaturation was greater than that seen with the Orbitrap Velos™, see fig 7.4. The holomyoglobin complex was observed in the + 8 and +7 charge states m/z 2197.14 and 2510.42. Theoretical m/z values were 2196.85 and 2150.54. The holomyoglobin species were less intense than peaks corresponding to apomyoglobin. There is a Gaussian-like distribution of charge states between m/z 1000- 1700 which is a similar appearance to that expected with a fully denatured protein The most abundant peaks were the + 7 and +8 charge states of apomyoglobin at m/z of 2422.32 and 2120.43 (theoretical m/z 2422.62 and 2119.91). This denaturation is likely to be induced by the

surface, as discussed above, as the instrument was capable of detecting the holomyoglobin complex from glass and by direct infusion electrospray.

7.1.2 Haemoglobin

In its native state, haemoglobin forms a heterotetramer composed of two alpha-globin subunits and two beta-globin sub units each non-covalently bound to a haem group, see fig 7.1 (B). Haemoglobin samples were analysed in the same way as myoglobin, beginning with analysis by direct infusion electrospray.

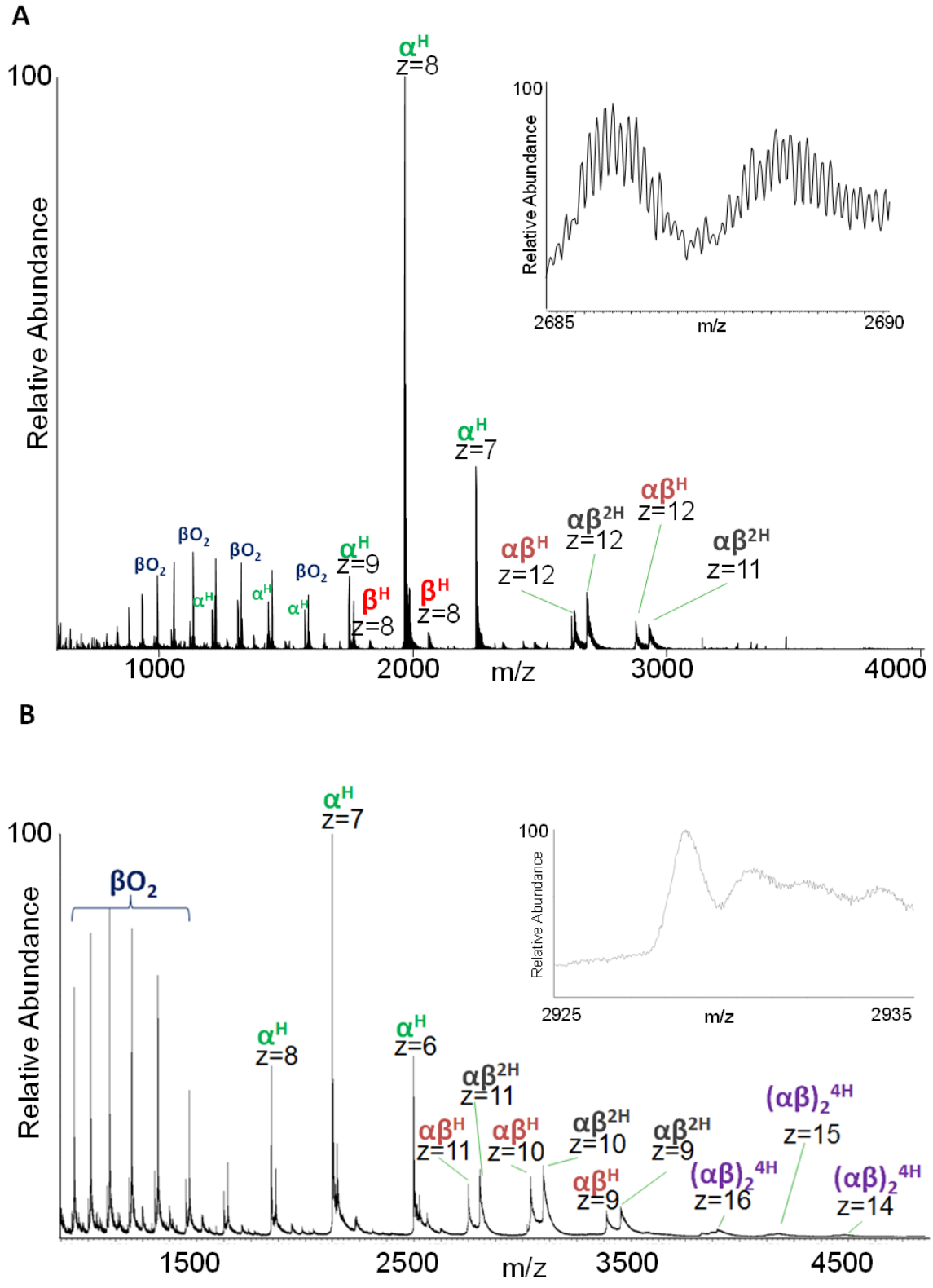


Figure 7.7 Direct infusion ESI native mass spectra of human haemoglobin obtained on (a) Orbitrap Velos™ and magnification of the +12 charge state of the $\alpha\beta^{2H}$ dimer (B) Synapt™ G2 S and magnification of the +11 charge state of the $\alpha\beta^{2H}$ dimer

Haemoglobin species	Observed charge states	Observed m/z (most abundant isotope)	Calculated m/z (most abundant isotope)
$\alpha\beta^{2H}$	+12, +11	2686, 2930	2686, 2930
$\alpha\beta^{H}$	+12, +11	2637, 2877	2635, 2874
α^{H}	+9, +8, +7,	1750, 1968, 2249	1750, 1968, 2249
β^{H}	+9, +8	1832, 2061	1832, 2061
β	+18 -+10	884, 936, 994, 1060 1136, 1223 1325, 1446, 1590	882, 934, 1058, 1134, 1221, 1323, 1443, 1587

Table 7.1 Haemoglobin species detected by the Orbitrap Velos™ mass spectrometer following direct infusion electrospray of haemoglobin standards in ammonium acetate.

Species	Observed Charge states	Observed m/z's (most abundant isotope)	Calculated m/z (most abundant isotope)
$(\alpha\beta)_2{}^{4H}$	+16, +15, +14	4031, 4298, 4605,	4029, 4298, 4604
$\alpha\beta^{2H}$	+11 ,+10, +9	3576, 3220, 2928	3581, 3223, 2930
$\alpha\beta^{H}$	+11 ,+10, +9	2875, 3511, 3162	2874, 3513, 3161
α^{H}	+8, +7, +6	1968, 2249, 2623	1968, 2249, 2624
β	+16 -10	1061, 1136, 1224, 1325, 1446,-1590	1058, 1134, 1221, 1323, 1443, 1587

Table 7.2 species acquired on the Waters Synapt™ from direct infusion analysis of haemoglobin standards

Haemoglobin complexes were successfully detected following from direct infusion electrospray of the protein standard as shown in fig 7.7. Peak assignments are shown in tables 7.1 and 7.2. The mass spectrum obtained on the orbitrap reveals the presence of $(\alpha\beta)^{2H}$ and haem-deficient $(\alpha\beta)^{H}$ dimers. Three different monomers were detected: α^{H}, β^{H} and additionally a β only subunit. The spectrum showed the beta subunit had undergone an oxidation. Haemoglobin tetramers were not detected. The mass spectrum acquired using the Synapt™ revealed the same peaks as observed using the orbitrap but without the presence of the intact β^{H} monomer and low abundance peaks that could be due to the presence of the intact tetramers. These peaks are at the baseline level and cannot be assigned with great confidence. Having successfully obtained multi-protein haemoglobin complexes through direct infusion electrospray, standards were analysed with native LESA mass spectrometry.

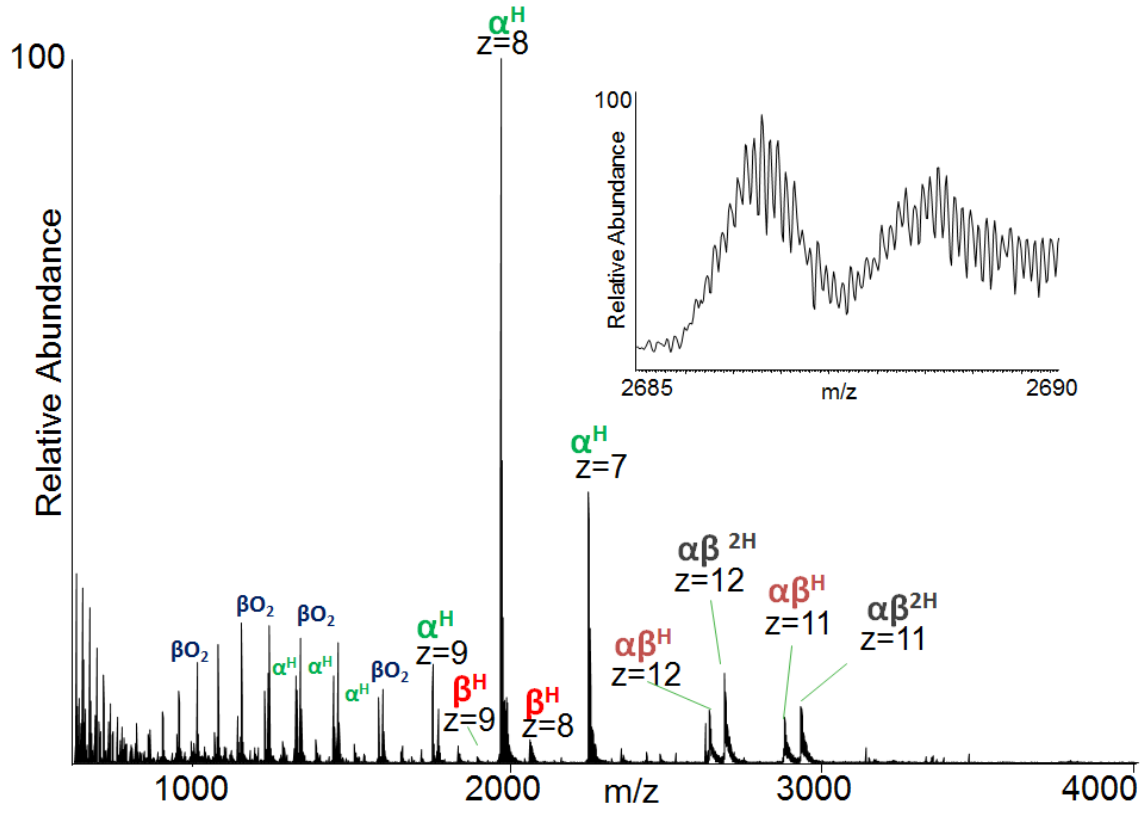


Figure 7.8 Native LESA mass spectrum of human haemoglobin and magnification of the +12 charge state of the $\alpha\beta^{2H}$ dimer obtained from glass microscope slide acquired on the Orbitrap Velos™

Species	Observed Charge states	Observed m/z (most abundant isotope)	Calculated m/z (most abundant isotope)
$\alpha\beta^{2H}$	+12, +11	2686, 2930	2686, 2930
$\alpha\beta^{H}$	+12, +11	2637, 2877	2635, 2874
α^{H}	+9, +8, +7,	1750, 1968, 2249	1750, 1968, 2249
β^{H}	+9, +8	1832, 2061	1832, 2061
β	+18 -10	884, 936, 994, 1060, 1136, 1224, 1325, 1446, 1590	882, 934, 992, 1058, 1134, 1221, 1323, 1443, 1587

Table 7.3 Species acquired on the Orbitrap Velos™ from LESA of haemoglobin standards from glass microscope slides

As described above, glass microscope slides were the most suitable substrate for native LESA mass spectrometry of myoglobin. The native LESA mass spectrum of haemoglobin corroborates that finding, see fig 7.8 and table 7.3. The same complexes observed by direct infusion ESI were also observed following native LESA from glass: $(\alpha\beta)^{2H}$, haem deficient $(\alpha\beta)^{H}$ dimers, and three types of monomer. The same 3 types of monomer, α^{H} β^{H} and the oxidised apo beta monomer were seen as well. The presence of dimers is particularly significant as it demonstrates that multi-protein complexes can be reanalysed

from a surface after drying. Salt adducts were present, as seen in the case of myoglobin. No tetramers were observed in this mass spectrum, but this is unsurprising considering they were absent when electrosprayed by direct infusion (see fig 7.7). Having been demonstrated on glass, native LESA mass spectrometry of haemoglobin was repeated on PDVF.

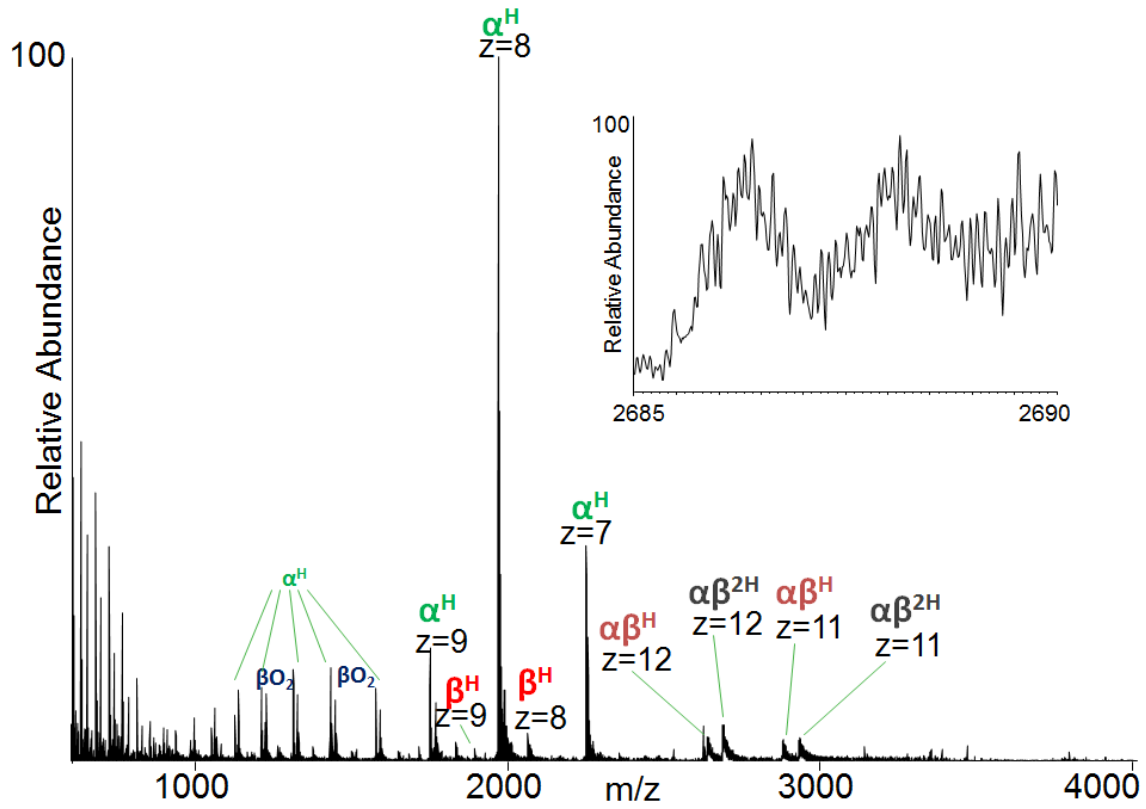


Figure 7.9 Native LESA mass spectrum of human haemoglobin and magnification of the +12 charge state of the $\alpha\beta^{2H}$ dimer obtained from PVDF membrane acquired on the Orbitrap Velos™

Species	Observed Charge states	Observed m/z's (most abundant isotope)	Calculated m/z (most abundant isotope)
$\alpha\beta^{2H}$	+12, +11	2686, 2930	2686, 2930
$\alpha\beta^{H}$	+12, +11	2637, 2877	2635, 2874
α^{H}	+9, +8, +7,	1750, 1968, 2249	1750, 1968, 2249
β^{H}	+9, +8	1832, 2061	1832, 2061
β	+18 -10	884, 936, 994, 1060, 1136, 1223, 1325, 1446, 1590	882, 934, 992, 1058, 1134, 1221, 1323, 1443, 1587,

Table 7.4 Species acquired on the Orbitrap Velos™ from LESA of haemoglobin standards from PVDF membrane

The native LESA mass spectrum of haemoglobin standard from PVDF (fig 7.9 and table 7.4) shows the presence of the same complexes as that obtained from glass (fig 7.8). Myoglobin was denatured when applied to PVDF, but the effect on haemoglobin has not been as great. Dimers were successfully detected although they are less intense than those obtained from glass. Dimers were present in +12 and +11 charge states as both $(\alpha\beta)^{2H}$ and haem deficient $(\alpha\beta)^{H}$ dimers. Monomers observed in this mass spectrum included α^{H}, β^{H} and the apo beta monomer.

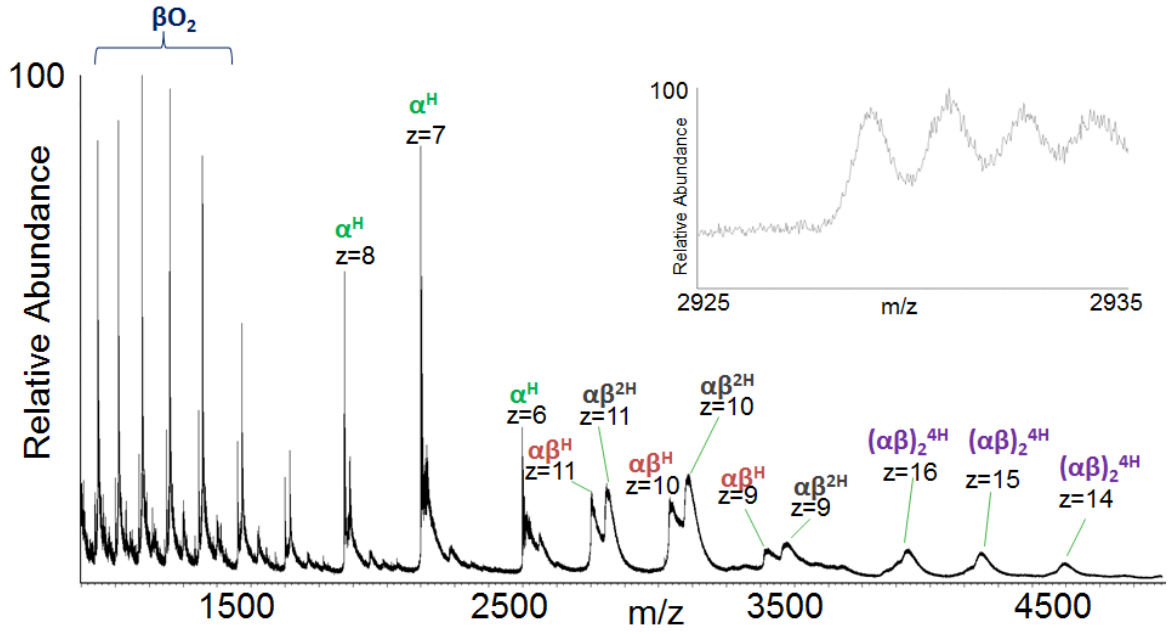


Figure 7.10 Native LESA mass spectrum of human haemoglobin and magnification of the +11 charge state of the $\alpha\beta^{2H}$ dimer obtained from glass microscope slide acquired on the Synapt™

Species	Observed Charge states	Observed m/z (most abundant isotope)	Calculated m/z (most abundant isotope)
$(\alpha\beta)_2{}^{4H}$	+16, +15, +14	4046, 4309, 4614	4029, 4297, 4604
$\alpha\beta^{2H}$	+11 ,+10, +9	2934, 3234, 3596,	2930, 3223, 3581
$\alpha\beta^{H}$	+11 ,+10, +9	2875, 3162, 3511,	2874, 3161, 3513
α^{H}	+8, +7, +6	1968, 2249, 2623	1968, 2249, 2624
β	+16 -10	1061, 1136,-1224, 1325, 1446, 1590	1058, 1134, 1221, 1323, 1443, 1587,

Table 7.5 species acquired on the Waters Synapt™ from LESA of haemoglobin standards from glass microscope slide

Fig 7.10 shows a mass spectrum obtained following native LESA of a haemoglobin spot on a glass microscope slide acquired on the Synapt™. The mass spectrum shows peaks corresponding to the intact tetramer in addition to the dimers and monomers. Although of low abundance, these peaks are more intense than observed in the mass spectrum acquired from solution (fig 7.7). The dominant peaks in this mass spectrum corresponded to the β monomers. Intact α^{H} monomers were observed but β^{H} monomers were not whereas both were obtained with the Orbitrap Velos™. Heterodimers were present as both $(\alpha\beta)^{2H}$ and haem deficient $(\alpha\beta)^{H}$ dimers, which is similar to the spectrum acquired with the orbitrap. The Synapt™ was not able to distinguish the isotopic spacing of the peaks, unlike the orbitrap.

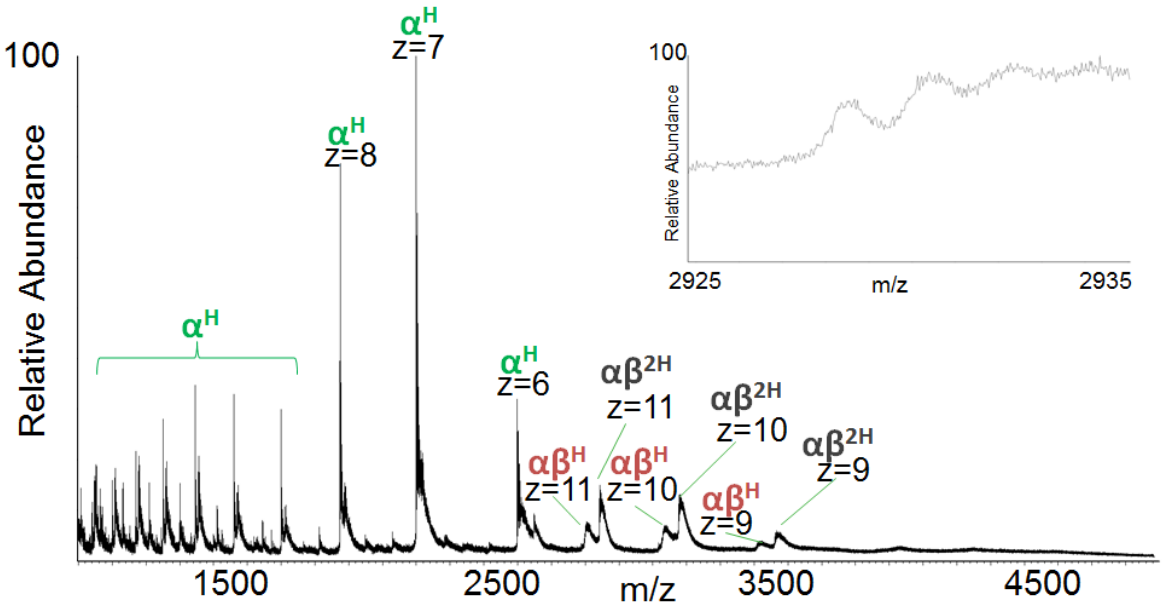


Figure 7.11 Native LESA mass spectrum of human haemoglobin and magnification of the + 11 charge state of the $\alpha\beta^{2H}$ dimer obtained from PVDF membrane acquired on the Synapt™

Species	Observed Charge states	Observed m/z (most abundant isotope)	Calculated m/z (most abundant isotope)
$\alpha\beta^{2H}$	+11 ,+10, +9	2930, 3220, 3576,	2930, 3223, 3581
$\alpha\beta^{H}$	+11 ,+10, +9	2872, 3161, 3520,	2874, 3161, 3513,
α^{H}	+8, +7, +6 , +12-9	1312, 1432, 1575, 1750, 1968, 2249, 2623,	1312, 1432, 1575, 1750, 1968, 2249, 2624

Table 7.6 species acquired on the Waters Synapt™ from LESA of haemoglobin standards from glass microscope slide

The haemoglobin standard was analysed by native LESA from a PVDF membrane. The resulting mass spectrum is shown on fig 7.11. No peaks corresponding to the intact tetramer were observed. The dominant peak was an α^{H} monomer. Peaks corresponding to dimers were observed as both the $(\alpha\beta)^{2H}$ and haem deficient $(\alpha\beta)^{H}$ forms but were less intense than those observed in the mass spectrum obtained from glass. These results are consistent with reports that a hydrophobic surface can act as a denaturant.

7.2 Dried blood spots

After non-covalent complexes had been successfully analysed from dried purified proteins by native LESA mass spectrometry, the next step was to apply the approach to a sample of biological interest. DBS were chosen due to the successful results obtained after analysing haemoglobin standards, and because analysing haemoglobin tetramers could be of clinical interest for investigating disorders affecting haemoglobin synthesis. Work presented in previous PhD thesis from this group had attempted using LESA orbitrap mass spectrometry for this purpose detecting haemoglobin dimers from DBS, but not tetramers [240]. The remainder of this chapter focuses on native LESA mass spectrometry from DBS.

7.2.3 Orbitrap analysis

In order to determine whether detection of haemoglobin tetramers was possible by use of the orbitrap, diluted whole blood samples were analysed. By using whole blood, the challenges of drying and coagulation are avoided. The upper m/z limit is 4000 Th and previously published examples where the native haemoglobin tetramer have been detected have peaks close to or above this limit [230, 241]. Human whole blood has been used to analyse haemoglobin tetramers previously [242]. In that work, human whole blood was acquired via finger prick, diluted by a factor of X 500 in ammonium acetate, and analysed via direct infusion ESI.

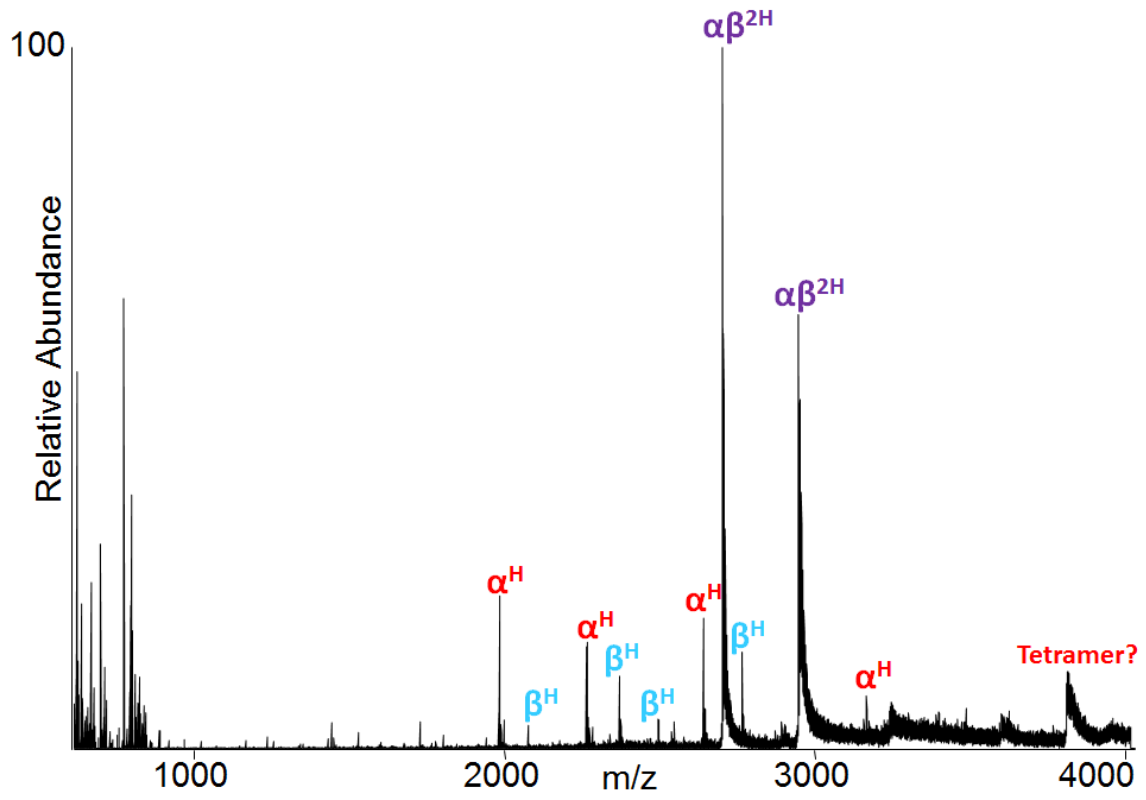


Figure 7.12 Direct infusion native ESI mass spectrum of whole blood diluted x500 in 10 mM ammonium acetate acquired on the Orbitrap Velos™

Fig 7.12 shows the mass spectrum obtained following direct infusion ESI of whole blood diluted by 500 in 10 mM ammonium acetate. The dominant peaks correspond to the $(\alpha\beta)^{2H}$ dimer. In addition, there are peaks corresponding to α^H and β^H monomers and a low abundance poorly-resolved peak at m/z 3795 corresponding to a tetramer in the + 17 charge state. The ions at m/z 3795 were isolated (isolation width 50 Th) in an attempt to improve resolution.

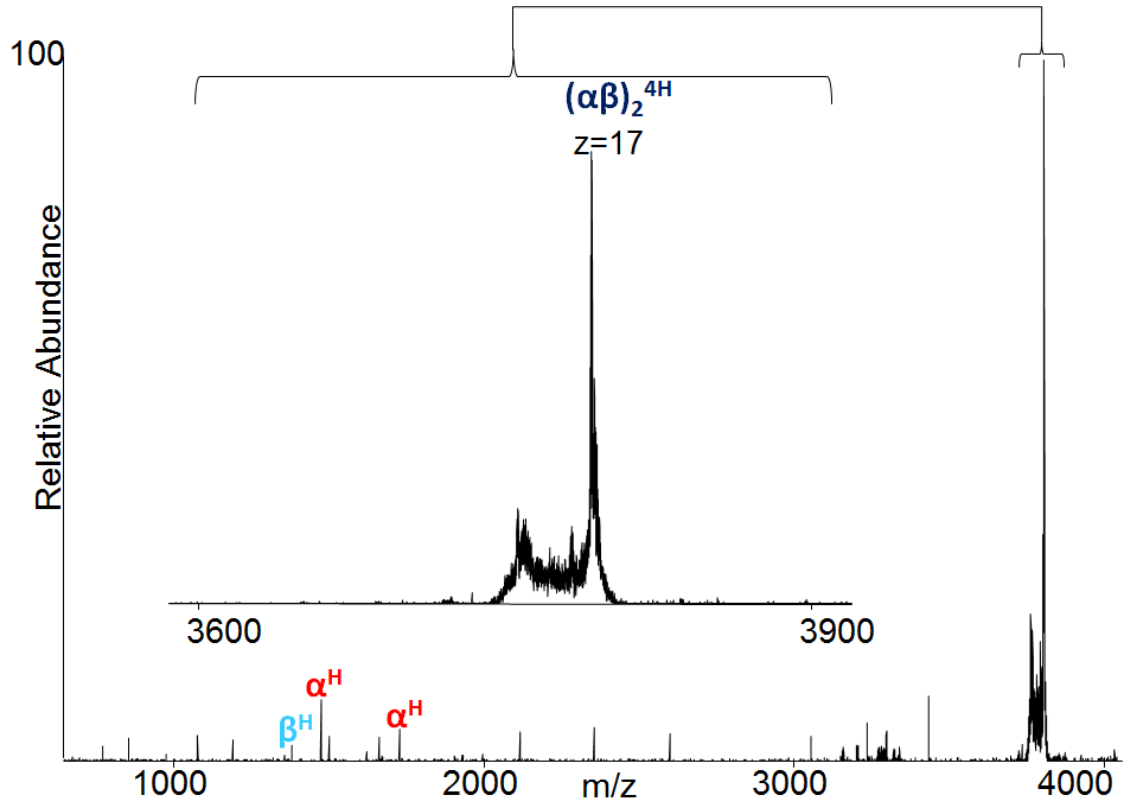


Figure 7.13 Isolation of 3795 peak following direct infusion native mass ESI analysis of whole blood diluted X500 in 10 mM ammonium acetate acquired on the Orbitrap Velos™

The mass spectrum obtained following peak isolation is shown in fig 7.13. The measured m/z was 3792.22, (theoretical m/z is 3792.33 , Δ 22.8 ppm) just and the charge state was determined to be +17 so can be confirmed to be a haemoglobin tetramer. Isolation could be used as a means for targeted identification of the haemoglobin tetramer from DBS.

A series of DBS were analysed by this approach using LESA mass spectrometry. Surface sampling was carried out as described in section 7.1. Samples were analysed first via a full scan mass spectrum, followed by isolation centred at m/z 3792.22 with a 50 Th window and a normalised collision energy of 0%.

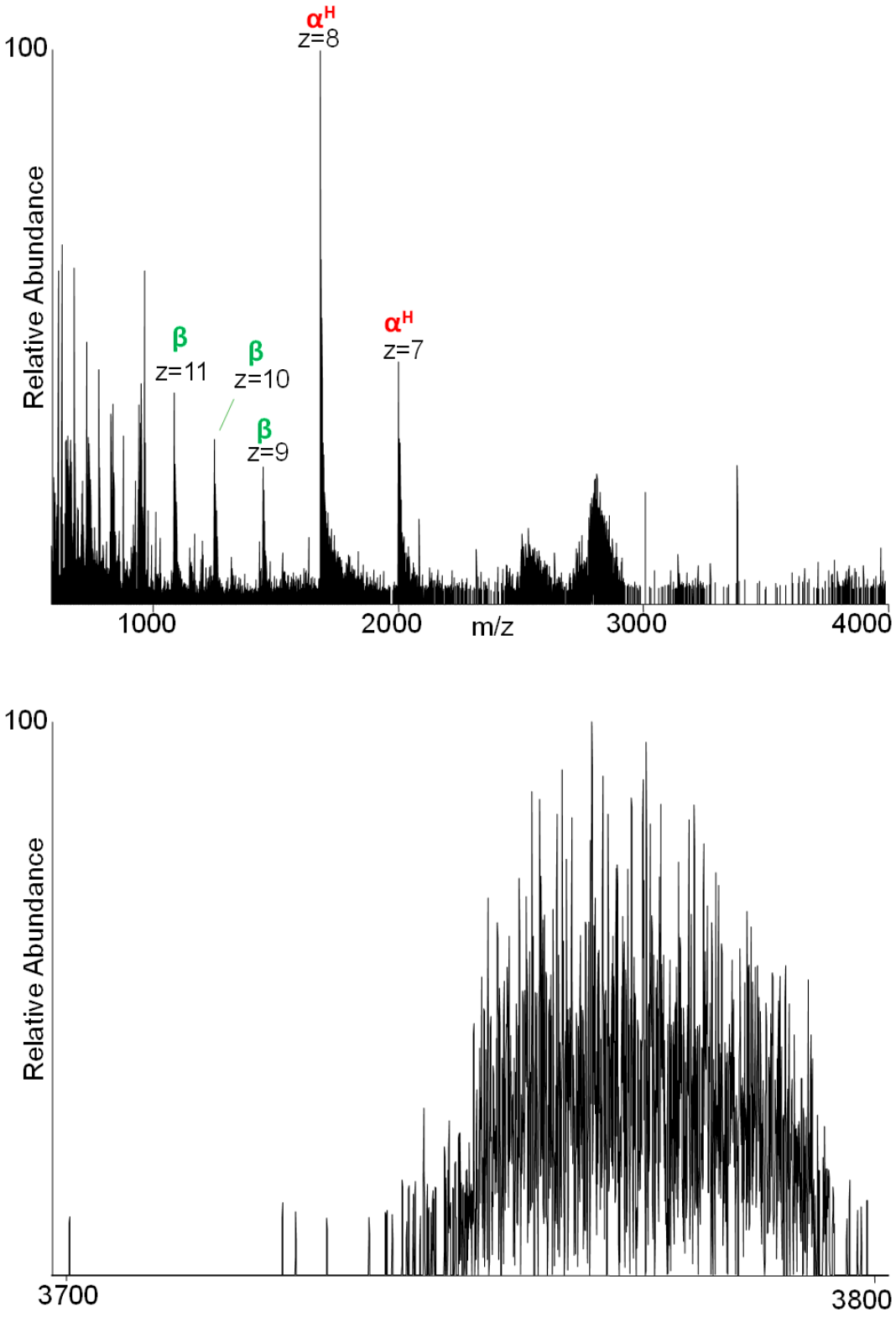


Figure 7.14 Native LESA mass spectrometry of DBS: (A) Full scan mass spectrum (B) Isolation m/z 3792.22

LESA analysis of DBS on the orbitrap did not result in the successful detection of the native human haemoglobin tetramer. The full scan mass spectrum shown in fig 7.14 A is of very low intensity: the dimers could not successfully be resolved. There were peaks corresponding to the α^H monomer, but this was the only non-covalent complex that could confidently be identified. Apo beta monomers were also present. The isolation of the m/z 3792.22 peak was also unsuccessful because there was no peak present in the full scan.

7.2.4 Q-TOF analysis

Haemoglobin tetramers had not successfully been obtained from DBS via LESA mass spectrometry on the orbitrap, so the Synapt™ mass spectrometer was used as an alternative. As mentioned in section 1.2.5, TOF mass spectrometers are the most common mass analysers used for the analysis of non-covalent protein complexes. The Synapt™ is more likely to be able to detect fragile multi-protein complexes than the orbitrap: Its inlet capillary is capable of operating at lower temperatures which are less likely to disrupt fragile interactions and the upper mass limit is higher, allowing the analysis of higher m/z ions.

DBS were acquired in full scan TOF mode as described in section 2.2.3.1 and acquired for a period of 5 minutes. Problems occurred with formation of liquid microjunctions. When the solvent was applied to the surface, it was followed by the collapse of the microjunction and absorption of the solvent into the DBS. When this took place, spots were left to dry for 5 minutes before the same area was resampled using the same method described in section 7.1 but with only one dispension and aspiration cycle. Two

different solvent systems were used in the LESA analyses. The first was composed of a 10 mM NH_4AC solution, pH balanced to 6.8, and the second was also composed of a 10 mM NH_4AC solution, balanced to pH 6.8, but with the addition of 5% methanol.

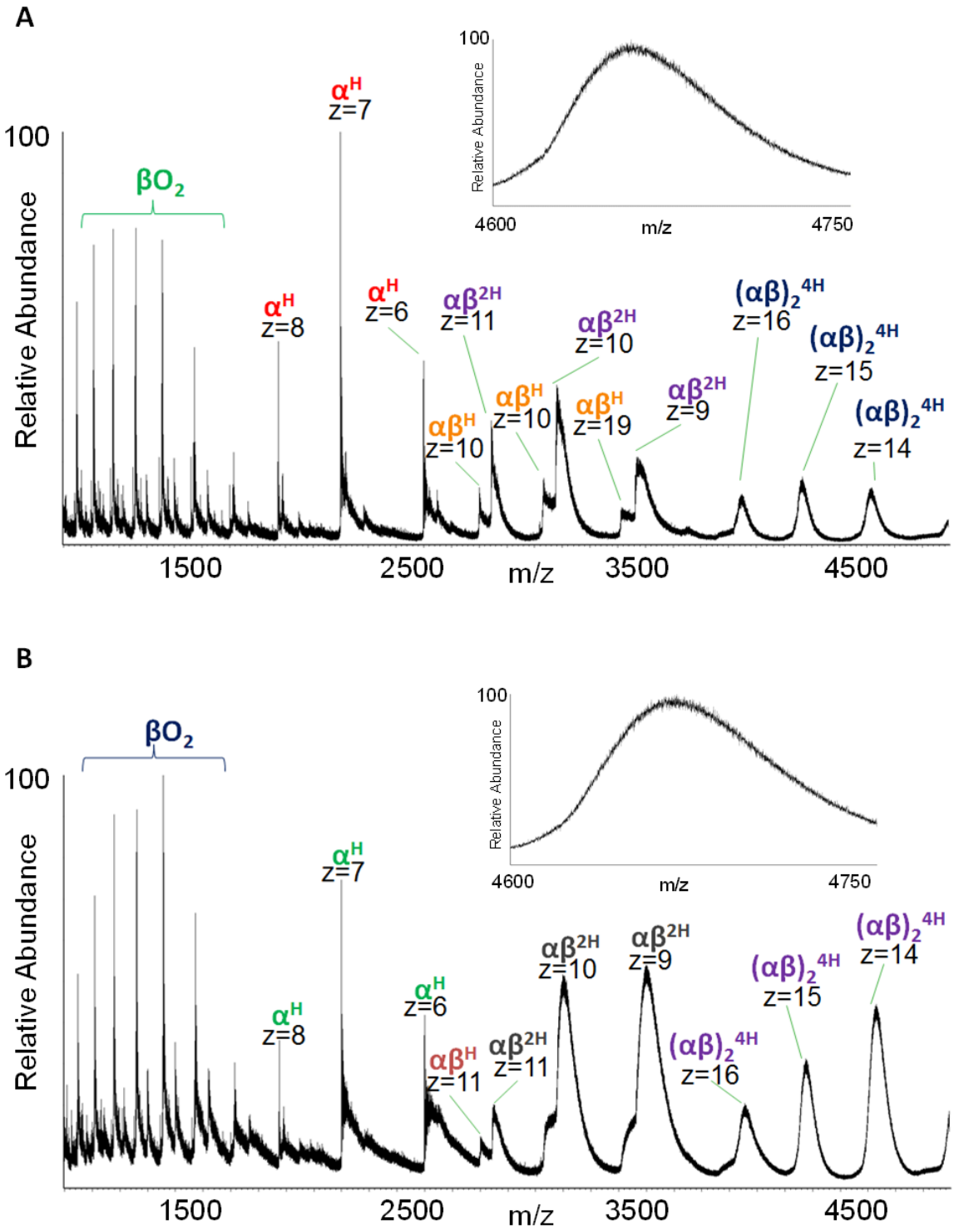


Figure 7.15 Native LESA mass spectrum obtained from DBS and magnification of the +14 charge state of the $\alpha\beta_2^{4H}$ tetramer acquired using the Synapt™ G2S. (A) 10 mM NH_4AC extraction solvent (B) 10 mM NH_4AC 5% methanol extraction solvent

Species	Observed Charge states	Observed m/z (most abundant isotope)	Calculate m/z (most abundant isotope)
$(\alpha\beta)_2{}^{4H}$	16, 15, 14	4065, 4338, 4652	4029, 4297, 4604
$\alpha\beta)^{2H}$	+11 ,+10, +9	2929, 3222, 3612	2930, 3223, 3581
$\alpha\beta^{H}$	+11 ,+10, +9	2876, 3163, 3514	2874, 3161, 3514
α^{H}	+8, +7, +6	1968, 2249, 2623, 1312-1750	1968, 2249, 2624
β	+16-+10	1060, 1136, 1224, 1325, 1446, 1590	992, 1058, 1134, 1221, 1323, 1443, 1587

Table 7.7 Species acquired on the Waters Synapt™ after LESA of DBS with 10 mM NH_4AC

Species	Observed Charge states	Observed M/z's	Calculate m/z (most abundant)
$(\alpha\beta)_2{}^{4H}$	16, 15, 14	4069, 4345, 4662	4029, 4297, 4604
$\alpha\beta^{2H}$	+11 ,+10, +9	2929 3247, 3623	2930, 3223, 3581
$\alpha\beta^{H}$	+11 ,	2876	2874, 3161, 3514
α^{H}	+8, +7, +6	1968, 2249, 2623	1968, 2249, 2624
B	+16-+10	1060, 1136, 1223, 1325, 1446, 1590	992, 1058, 1134, 1221, 1323, 1443, 1587

Table 7.8 Species acquired on the Waters Synapt™ after LESA of DBS with 10 mM NH_4Ac and 5% MeOH

Haemoglobin tetramers were detected from DBS following native LESA mass spectrometry on the Synapt™ mass spectrometer using both solvent conditions. Tetramers were observed in the +16, +15 and +14 charge states. The tetramer peaks were less intense than those corresponding to the dimers and monomer but significantly more intense than those observed from the purified protein on glass microscope slides shown in fig 7.10. The tetramer peaks were very broad (~ 100 Th) which is similar to previously reported examples[230, 241]. Broadening of the peaks is likely to be due to insufficient desolvation of the complex, and aggregations of salts. The temperature of the Synapt's™ ion transfer tube is set at 30°C, compared with 250°C for the orbitrap. Lower temperatures prevent dissociation of unstable interactions which may be why it is possible to obtain tetramers on the Synapt™. Lower temperatures can lead to the retention of interactions between the solvent and the protein into the gas phase resulting in peak broadening. Higher temperatures would result in narrower peaks but may make it impossible to retain the tetramer into the gas phase. Salts can be another problem: Salts form adducts along with the protonated form of the protein, as seen in previous spectra. Here the different adducts could not be distinguished from the tetramer due to the limited resolving power. The resolving power of TOF mass analysers are not great enough to distinguish different adducts for a native protein complex with a mass higher than 65 kDa [243]. Salts are an unavoidable issue when analysing DBS. Blood contains a physiologically high concentration of salts. The most concentrated endogenous salt is sodium at between 135-145 mM in a healthy individual [244]. Peaks corresponding to tetramers were more intense when analysed in a solution composed of 5% methanol. Adding small amounts of organic solvents in native mass spectrometry can improve

ionisation efficiency,[245] but in this case it could also help improve protein extraction and solubility.

7.3 Chapter 7 conclusions

This work presented in this chapter demonstrates the successful use of LESA mass spectrometry for the analysis of non-covalent complexes from proteins dried onto surfaces. Multi-protein and protein-cofactor complexes were observed in the gas phase following native LESA sampling with a spectral quality similar to that obtained following direct infusion ESI. Two different mass spectrometers were used to analyse these non-covalent complexes, a high resolution orbitrap mass spectrometer and a high resolution QTOF. Both w--ere capable of analysing non-covalent interactions, but the QTOF mass spectrometer was better able to transmit larger complexes of higher m/z ratios. This was work was some of the first demonstrating the analysis of non-covalent interactions from a surface by an mass spectrometry surface sampling tool, but this concept has since been explored by others. LESA has recently been used to sample native protein complexes from thin tissue sections [246].

In this chapter the haemoglobin tetramer was observed from a DBS, which is of particular interest as the approach could be of use for investigating disorders affecting haemoglobin. Certain disorders such as thalassemias result in abnormal haemoglobin

subunit stochiometries [247]. Thalassemias can result in larger amounts of haemoglobin A2, haemoglobin H or haemoglobin Barts [248]. Haemoglobin A2 is a tetramer composed of two alpha chains and two delta-globin chains, haemoglobin H is made up of four delta globin chains and haemoglobin Barts is a tetramer composed of four gamma haemoglobin chains[248, 249]. If the healthy haemoglobin tetramer can be identified then it may be possible to identify others too and native LESA mass spectrometry could be used to these tetramers directly from a DBS without any additional sample preparation.

Chapter 8 Conclusions

The work carried out in this thesis has shown a number of new applications for LESA mass spectrometry in proteomics research. Prior to beginning this work, there had only been a few published examples of using LESA for proteomics analysis and the majority of applications were aimed at the analysis of small molecules, such as lipids, pesticides and drugs. This work in this thesis demonstrates the development of new methods and substrates showing that LESA is a useful tool for analysing proteins from surfaces.

The work presented in Chapter 3 was carried out to attempt to improve fractionation techniques available for top down proteomics, by using LESA to extract intact proteins blotted onto PVDF membranes from SDS PAGE gels. LESA could not be used for the analysis of proteins blotted from gels but it was possible to reanalyse proteins after they have been directly deposited onto PVDF. The ability to analyse intact proteins from PVDF led to formation of an external collaboration and proved to be useful in for analysing proteins in air filters.

The most successful part of this PhD was presented in Chapter 4 in which untargeted bottom up proteomics analysis of DBS by use of LESA coupled with automated proteolysis was demonstrated. Until this work was undertaken, applications of LESA in proteomics were restricted to the analysis of intact proteins. Top down protein analysis can be very limiting: many proteins are too large to be detected without any proteolytic digestion and

it is most effective at detecting the most abundant proteins in a sample. In a complex sample, the majority of proteins will not be accessible to top down analysis. Incorporating a trypsin digestion into the LESA sampling routine made it possible to greatly increase the number of proteins that can be identified from a DBS via LESA. It was possible to identify over 100 proteins through an LC MS/MS analysis of a DBS digest whereas using a LESA top down proteomics approach had only ever resulted in the identification of haemoglobin proteins. The research also unexpectedly demonstrated a new approach for analysing DBS that had never been investigated in any detail: proteomics. This approach could be of major advantage in the clinic as DBS could be used as a convenient source for many different protein biomarkers. The work demonstrated that these protein biomarkers are accessible to mass spectrometric analysis. DBS could become as a sampling format to rival plasma in identifying new biomarkers which is a major focus of research in proteomics. It is surprising that until now the DBS had not even been considered as an alternative.

The idea of pursuing a proteomics analysis of DBS were further explored in Chapter 5 by considering more established means of sample preparation. Proteins were eluted from punches cut out of DBS and digested in solution. If DBS were to be used as an alternative to plasma the majority of laboratories would be unlikely to use LESA sampling and would most likely use the “punch and elute” approach. Data from Chapter 5 showed that the “punch and elute” approach was a viable option for identifying proteins from a DBS, but less successful than the LESA based digest. Solution phase digests were amenable to

enrichment of low abundance proteins by immunodepletion. Sample complexity is a major challenge when analysing blood and plasma and the number of proteins identified was improved by removing the higher abundance proteins. Several different options were attempted but the most successful appeared to be removing haemoglobin only. Care is needed when depleting DBS. DBS contain only a small amount of protein, and even though there are many different types, if too much protein is pulled out by immunodepletion, there may very little of the low abundance proteins available for detection.

The work presented in Chapter 6 aimed to demonstrate the clinical use of the LESA based trypsin digest developed in Chapter 4 by determining variants of the alpha-1-antityrypsin protein. This was unsuccessful. The different variants linked to disease result in a significant reduction in synthesis of the protein which becomes difficult to characterise at low concentrations. The results did demonstrate, however, that the LESA based trypsin digest can be analysed by a variety of different mass spectrometry methods. It can be analysed by targeted approaches which could be particularly important for clinical applications.

The work presented in the final chapter demonstrated the application of LESA to analyse non-covalent complexes from proteins dried onto surfaces, overcoming the challenge of denaturation upon drying. The work presented in this chapter saw protein cofactor and multi-protein interactions retained into the gas phase from a dried surface. Most exciting

of all the spectra obtained were those where a haemoglobin tetramer had been recovered directly from a dried blood spot, exactly how it exists in its native state. This was all achieved without performing any sample preparation or pre-treatment.

Future work

Much of the work carried out in this thesis could be investigated further. The LESA based digest described in Chapter 4 could be applied to a range of different samples and not just DBS. It has already been used successfully for the analysis of thin tissue sections [219]. Similar approaches have been published, where LESA has been used to identifying proteins from tryptic peptides originating from tissues but were not able to couple sampling and digestion so had to perform them in a multi-step process. Applications have been demonstrated for analysing meat surfaces in the authentication of foods and discovering biomarkers from cancerous tissue [250, 251]. Montowska *et al.* [250] focused on analysing meat surface and developed a method by which a meat sample were digested in solution and then printed onto a surface and dried prior to re-analysis with LESA. Wisztorski [251] used a similar approach for identifying of cancer biomarkers from ovarian tissue, their method involved performing an on tissue digest with a chemical inkjet printer, which was then followed up by LESA sampling to extract the peptides, solid phase extraction for clean-up and then finally LC MS/MS. By using the method described in Chapter 4 the sample preparation performed in these two multi-step methods could be carried out in a single stage. Chapter 4 demonstrated that LESA sampling can be

incorporated alongside sample processing by using the Triversa Nanomate™ as a liquid handling robot. Only a trypsin digest has been experimented with in detail so far, but it would be interesting to see if other processing stages could be incorporated as well. Other processing stages could include a disulphide bond reduction, immunodepletion such as that carried out in Chapter 5, or even sample clean-up such as zip tipping.

In terms of advancing the field of DBS analysis, there are a number of options for taking this work further. A new range of specialised DBS cards incorporating membrane filtration have been released. These contain specialised filters that filter out red blood cells from the plasma during the blood spotting and drying process, leaving a sample that is more similar to a dried plasma spot [252, 253]. These spots could be an interesting sample to analyse from a proteomics perspective as it would be a less complicated sample than a standard DBS but still retain the advantages associated with fingerprick sampling. These samples would also be interesting to analyse with immunodepletion methods such as those described in Chapter 5.

The immunodepletion methods used in Chapter 5 were interesting in showing that identification of low abundance proteins could be improved by sample enrichment. Removing the high abundance proteins was a viable option, but an alternative could be to pull out specific proteins of interest for analysis. This approach has been applied to plasma samples, such as the stable isotope standards and capture by anti-peptide antibodies (SISCAPA) assays [254]. This would be a targeted means of analysis, but it can be multiplexed and crucially can be used for the detection of very low abundance

proteins such as the interleukins and cytokines [254] and have been experimented for use on DBS [255].

A1AT variants could not be characterised by LESA based DBS digestions, however there are a number of additional options that could be attempted. The failure was hypothesised to be due to low concentrations of the A1AT protein rendering it impractical for detailed characterisation by bottom up proteomics techniques. Chapter 5 did demonstrate that depletion of haemoglobin proteins does increase the detection of some of the less abundant plasma proteins. If haemoglobin depletion were carried out on an A1AT deficient DBS then it may enrich the A1AT protein enough for bottom up characterisation and variant determination.

Chapter 7 also raises a number of new possibilities. The native haemoglobin tetramer was detected from LESA sampling of DBS. This method could be of interest in sampling DBS from patients suffering from thallasemias and could be used for detecting abnormal haemoglobin stoichiometries. The results in this chapter also showed LESA analysis of non-covalent complexes from purified proteins. Protein nucleic acid complexes were not investigated in this work, but are commonly used in research on native mass spectrometry. It would be interesting to investigate if these interactions could be analysed using LESA. A final avenue of research would be to combine LESA sampling with the travelling wave ion mobility spectrometry that is available on the Synapt™. If used

effectively it can be used as a structural tool to calculate the collision cross section of an ion, and could reveal the extent of denaturation proteins undergo when drying.

Chapter 9 References

1. Kertesz, V. & Van Berkel, G.J. Fully automated liquid extraction-based surface sampling and ionization using achip-based robotic nanoelectrospray platform. *Journal of mass spectrometry.* **45**, 252-260 (2009).
2. Edwards, R.L., Creese, A.J., Baumert, M., Griffiths, P., Bunch, J. & Cooper, H.J. Hemoglobin variant analysis via direct surface sampling of dried blood spots coupled with high resolution mass spectrometry. *Analytical Chemistry.* **83**, 2265-2270 (2011).
3. Glish, G.L. & Vachet, R.W. The basics of mass spectrometry in the twenty first century. *Nature Reviews in Drug Discovery.* **2**, 140-150 (2003).
4. Grayson, M.A., *Measuring mass: from positive rays to proteins.* (2002), Philadelphia. Chemical Heritage press.
5. Griffiths, J. A brief history of mass spectrometry. *Analytical Chemistry.* **80**, 5678-5683 (2008).
6. El-Aneed, A., Cohen, A. & Banoub, J. Mass spectrometry, review of the basics: electrospray, MALDI and commonly used mass analysers. *Applied Sprectroscopy Reviews.* **44**, 210-230 (2009).
7. Siuzdak, G. The emergence of mass spectrometry in biochemical research. *Proceedings of the National Academy of Sciences.* **91**, (1994).
8. Zenobi, R. Chemistry nobel prize 2002 goes to analytical chemistry. *Chimia International Journal of Chemistry.* **57**, 73 (2003).
9. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. & Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science.* **246**, 64-71 (1989).
10. Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D. & Alice, M.B. Molecular beams of macroions *The Journal of Chemical Physics.* **49**, 2240-2249 (1968).
11. Fenn, J.B. Electrospray wings for molecular elephants (Nobel lecture). *Angewandte Chemie.* **42**, 3871-3894 (2003).
12. Wong, S.F., Meng, C.K. & Fenn, J.B. Multlple charging in electrospray ionization of poly(ethy1ene glycols) *Journal of Physical Chemistry.* **92**, 546-550 (1988).
13. Cech, N.B. & Enke, C.G. Practical implications of some recent study in electrospray ionization fundamentals. *Mass Spectrometry Reviews.* **20**, 362-387 (2001).
14. Hogan, C.J., Carroll, J.A., Rohrs, H.W., Biswas, P. & Gross, M.L. Combined charged residue-field emission model of macromolecular electrospray ionization. *Analytical chemistry.* **81**, 369-377 (2009).
15. Kebarle, P. & Verkerk, U.H. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews.* **28**, 898-917 (2009).
16. Taylor, G. Disintegration of water drops in an electric field. *Proceedings of the Royal Society of London series A, Mathematical and Physical Sciences.* **280**, 383-397 (1964).
17. Rohner, T.C., Lion, N. & Girault, H.H. Electrochemical and theoretical aspects of electrospray ionisation. *Physical Chemistry Chemical Physics.* **6**, 3056-3068 (2004).
18. Smith, J.N., Flagan, R.C. & Beauchamp, J.L. Droplet evaporation and discharge dynamics in electrospray ionization. *Journal of Physical Chemistry A.* **106**, 9957-9967 (2002).
19. Nguyen, S. & Fenn, J.B. Gas-phase ions of solute species from charged droplets of solutions. *Proceedings of the National Academy of Sciences.* **104**, 1111-1117 (2007).
20. Iribarne, J.V. & Thomson, B.A. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics.* **64**, 2287-2294 (1975).
21. Wang, G. & Cole, R.B. Charged residue versus ion evaporation for formation of alkali metal halide cluster ions in ESI. *Analytica Chimica Acta.* **406**, 53-65 (2000).

22. Mallet, C.R., Lu, Z. & Mazzeo, J.R. A study of ion suppression effects in electrospray ionization from mobile phase additives and solid-phase extracts. *Rapid Communications in Mass Spectrometry.* **18**, 49-85 (2003).
23. Gaskell, S.J. Electrospray: principles and practice. *Journal of Mass Spectrometry.* **32**, 677-688 (1997).
24. Gabelica, V. & Pauw, E.W. Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrometry Reviews.* **24** 566-587 (2004).
25. Iavarone, A.T., Jurchen, J.C. & Williams, E.R. Effects of solvent on the maximum charge state and charge state distribution of protein ions produced by electrospray ionization. *Journal of the American Society for Mass Spectrometry.* **11**, 976-985 (2000).
26. Hendrickson, C.L. & Emmett, M.R. Electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Annual reviews in Physical Chemistry.* **50**, 517-536 (1999).
27. Hardman, M. & Makarov, A.A. Interfacing the orbitrap mass analyzer to an electrospray ion source. *Analytical chemistry.* **75**, 1699-1705 (2003).
28. Coon, J.J., Ueberheide, B., Syka, J.E.P., Dryhurst, D.D., Ausio, J., Shabanowitz, J., Hunt, D.F. & McLafferty, F.W. Protein ientification usng sequential ion/ion reactions and tandem mass spectrometry. *Proceedings of the National Academy of Sciences.* **102**, 9463-9468 (2005).
29. Mitulovic G & Mechtler K HPLC techniques for proteomics analysis a short overview of latest developments. *Briefings in functional genomics and proteomics.* **5**, 249-260 (2006).
30. El-Faramawy, A., Michael Siu, K.W. & Thomson, B.A. Efficiency of nano-electrospray ionization. *Journal of the American Society for Mass Spectrometry.* **16**, 1702-1707 (2005).
31. Gibson, G.T.T., Mugo, S.M. & Oleschuk, R.D. Nanoelectrospray emitters: trends and perspective. *Mass Spectrometry Reviews.* **28**, 918-936 (2009).
32. Schmidt, A., Karas, M. & Dulcks, T. Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI? *Journal of the American Society for Mass Spectrometry.* **14**, 492-500 (2003).
33. Harris, G.A., Nyadong, L. & Fernandez, L.M. Recent developments in ambient ionization techniques for analytical mass spectrometry. *The Analyst.* **133**, 1297-1301 (2008).
34. Alberici, R.M., Simas, R.C., Sanvido, G.B., Romão, W., Lalli, P.M., Benassi, M., Cunha, I.B.S. & Eberlin, M.N. Ambient mass spectrometry: bringing MS into the “real world". *Analytical Bioanalytical Chemistry.* **398**, 265-294 (2010).
35. Monge, M.E., Harris, G.A., Dwivedi, P. & Fernandez, F.M. Mass spectrometry: recent advances in direct open air surface sampling/ionization. *Chemical Reviews.* **113**, 2269-2308 (2013).
36. Takats, Z., Wiseman, J.M., Gologan, B. & Cooks, G.R. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science.* **306**, 471-473 (2004).
37. Takats, Z., Wiseman, J.M. & Cooks, R.G. Ambient mass spectrometry using desorptionelectrospray ionization (DESI): instrumentation,mechanisms and applications in forensics, chemistry,and biology. *Journal of Mass Spectrometry.* **40**, 1261-1275 (2005).
38. Williams, J.P. & Scrivens, J.H. Rapid accurate mass desorption electrospray ionisation tandem mass spectrometry of pharmaceutical samples. *Rapid Communications in Mass Spectrometry.* **19**, 3643-3650 (2005).
39. Wiseman, J.M., Evans, C.A., Bowen, C.L. & Kennedy, J.H. Direct analysis of dried blood spots utilizing desorption electrospray ionization (DESI) mass spectrometry. *The Analyst.* **135**, 720-725 (2009).
40. Cody, R.B., Larame , J.A. & Dupont Durst, H. Versatile new ion source for the analysis of materials in open air under ambient conditions. *Analytical Chemistry.* **77**, 2297-2302 (2005).

41. Liu, J., Wang, H., Manicke, N.E., Lin, J.L., Cooks, R.G. & Ouyang, Z. Development, characterization, and application of paper spray ionization. *Analaytical Chemistry.* **82**, 2463-2471 (2010).
42. Van Berkel, G.J., Pasilis, S.P. & Ovchinnikova, O. Established and emerging atmospheric pressure surface sampling/ionization techniques for mass spectrometry. *Journal of mass spectrometry.* **43**, 1161-1180 (2008).
43. Van Berkel, G.J. Thin layer chromatography and electrospray mass spectrometry coupled using a surface sampling probe. *Analytical Chemistry.* **74**, 6216-6223 (2002).
44. Schultz, G.A., Corso, T.N., Prosser, S.J. & Zhang, S. A fully integrated monolithic microchip electrospray device for mass spectrometry. *Analytical Chemistry* **72**, 4058-4063 (2000).
45. Van Pelt, C.K., Zhang, S. & Henion, J.D. Characterization of a fully Automated nanoelectrospray System with mass spectrometric detection for proteomic analyses. *Journal of Biomolecular Techniques.* **13**, 72-84 (2002).
46. Schadt, S., Kallbach, S., Almeida, R. & Sandel, J. Investigation of figopitant and Its metabolites in rat tissue by combining whole body autoradiography with liquid extraction surface analysis mass spectrometry. *Drug metabolism and disposition.* **40**, 419-425 (2012).
47. Stegemann, C., Drozdov, I., Shalhoub, J., Humphries, J., Ladroue, C., Didangelos, A., Baumert, M., Allen, M., Davies, A.H., Monaco, C., Smith, A., Xu, Q. & Mayr, M. Comparative lipidomics profiling of human atherosclerotic plaques. *Circulation: cardiovascular genetics.* **4**, 232-242 (2011).
48. Walworth, M.J., Stankovich, J.J., Van Berkel, G.J., Schulz, M. & Minarik, S. High-performance thin-layer chromatography plate blotting for liquid microjunction surface sampling probe mass spectrometric analysis of analytes separated on a wettable phase plate. *Rapid Communications in Mass Spectrometry.* **26**, 37-42 (2011).
49. Eikel, D., Vavrek, M., Smith, S., Bason, C., Yeh, S., Korfmacher, W.A. & Henion, J.D. Liquid extraction surface analysis mass spectrometry (LESA-MS) as a novel profiling tool for drug distribution and metabolism analysis: the terfenadine example. *Rapid Communications in Mass Spectrometry.* **25**, 3587-3596 (2011).
50. Eikel, D. & Henion, J. Liquid extraction surface analysis (LESA) of food surfaces employing chip based nano electrospray mass spectrometry. *Rapid Communications in Mass Spectrometry.* **25**, 2345-2354 (2011).
51. Paine, R.L.M., Rae, I.D. & Blanksby, S.J. Direct detection of brominated flame retardants from plastic e-waste using liquid extraction surface analysis mass spectrometry. *Rapid Communications in Mass Spectrometry.* **28**, 1203-1208 (2014).
52. Makarov, A. *Orbitrap mass spectrometer*. United States patent. United states (1999).
53. Perry, R.H., Cooks, G.R. & Noll, R.J. Orbitrap mass spectrometry: instrumentatnion, ion motion and applications. *Mass Spectrometry Reviews.* **27**, 661-699 (2008).
54. Hu, Q., Noll, R.J., Li, H., Makarov, A., Hardmanc, M. & Cooks, G. The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry.* **40**, 430-443 (2005).
55. Makarov, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical Chemistry.* **72**, 1156-1172 (2000).
56. Makarov, A., Denisov, M. & Lange, O. Performance evaluation of a high field orbitrap mass analyzer. *Journal of the American Society for Mass Spectrometry.* **20**, 1391-1396 (2009).
57. Kingdon, K.H. A method for the neutralization of electron space charge by posotive ionization and very low gas pressures. *Physical review.* **21**, 408-418 (1923).
58. Zubarev, R.A. & Makarov, A. Orbitrap Mass Spectrometry. *Analytical Chemistry.* **85**, 5288-5296 (2013).

59. Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Muller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J.J., Cox, J., Horning, S., Mann, M. & Makarov, A. Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Molecular and Cellullar Proteomics.* **11**, (2012).
60. Olsen, J.V., Schwartz, J.C., Griep-Raming, J., Nielsen, M.L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E.R., Senko, M., Makaov, A., Mann, M. & Horning, S. A dual pressure linear ion trap orbitrap instrument with very high sequencing speed. *Molecular and Cellullar Proteomics.* **8**, 2759-2769 (2009).
61. Olsen, J.V., De Godoy, L.M.F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S. & Mann, M. Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. *Molecular and Cellullar Proteomics.* **4**, 2010-2021 (2005).
62. Schwartz, J.C., Senko, M.W. & Syka, J.E.P. A two-dimensional quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry.* **13**, 659-669 (2002).
63. Douglas, D.J., Frank, A.J. & Mao, D. Linear ion traps in mass spectrometry. *Mass Spectrometry Reviews.* **24**, 1-29 (2003).
64. Hashimoto, Y., Hasegawa, H., Baba, T. & Waki, I. Mass selective ejection by axial resonant excitation from a linear ion trap. *Journal of the American Society for Mass Spectrometry.* **22**, 685-690 (2006).
65. Hager, J. A new linear ion trap mass spectrometer. *Rapid Communications in Mass Spectrometry.* **16**, 512-526 (2002).
66. Laskay, U.A. & Jackson, G.P. Resonance excitation and dynamic collision-induced dissociation in quadrupole ion traps using higher-order excitation frequencies. *Rapid Communications in Mass Spectrometry.* **22**, 2342-2348 (2008).
67. Olsen, J.V. & Mann, M. Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proceedings of the National Academy of Sciences.* **101**, 13417-13422 (2004).
68. De Hoffmann, E. & Stroobant, V., *Mass spectrometry principles and applications.* Third eddition. (2007), Chichester. Wiley.
69. Yost, R.A. & Enke, C.G. Tripple Quadrupole mass spectrometry for direct mixture analysis and structure eludication. *Analytical Chemistry.* **51**, 1251-1264 (1979).
70. Burlingame, A.L., *Biological mass spectrometry.* (2005), Amsterdam. Elsevier.
71. Cameron, A.E. & Eggers, D.F. An ion "velocitron". *Reviews of Scientific Instrumentation.* **19**, 605-607 (1948).
72. Weickhardt, C., Moritz, F. & Grotemeyer, J. Time of flight mass spectrometery: state of the art chemical analysis and molecular science. *Mass spectrometry Reviews.* **15**, 139-162 (1996).
73. Vlasak, P.R., Beussman, D.J., Ji, Q., . & Enke, C.G. Method for the design of broad energy range focusing reflectrons *Journal of the American Society for Mass Spectrometry.* **7**, 1002-1008 (1996).
74. Guilhaus, M., Selby, D. & Mlynski, V. Orthogonal acceleration time of flight mass spectrometry. *Mass Spectrometry Reviews.* **19**, 65-107 (2000).
75. Chernushevich, I.V., Loboda, A.V. & Thomson, B.A. An introduction to quadrupole–time-of-flight mass spectrometry. *Journal of Mass Spectrometry.* **36**, 849-865 (2001).
76. Pringle, S.D., Giles, K., Wildgoose, J.A., Williams, J.P., Slade, S.E., Thalassinos, K., Bateman, R.H., Bowersc, R.H. & Scrivens, J.H. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry.* **261**, 1-12 (2007).

77. Heck A J R & Van den Heuvel R H H Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368-389 (2003).
78. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature.* **422**, 198-207 (2003).
79. Steen, H. & Mann, M. The abc's (and xyz's) of peptide sequencing *Nature Reviews.* **5**, 699-711 (2004).
80. Wysocki, V.H., Tsaprailis, G., Smith, L.L. & Breci, L.A. Mobile and localized protons: a framework for understanding peptide dissociation. *Journal of Mass Spectrometry.* **35**, 1399-1406 (2000).
81. Vachet, R.W., Winders, A.D. & Glish, G.L. Correlation of kinetic energy losses in high energy collision induced dissociation with observed peptide product ions. *Analytical Chemistry.* **68**, 522-526 (1996).
82. Sleno, L. & Volmer, D.A. Ion activation methods for tandem mass spectrometry. *Journal of Mass Spectrometry.* **39**, 1091-1112 (2004).
83. Paizs, B. & Suhai, S. Fragmantation pathways of protonated peptides. *Mass Spectrometry Reviews.* **24**, 508-548 (2005).
84. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *BIomedical Mass Spectrometry.* **11**, 601 (1984).
85. Standing, K.G. Peptide and protein de novo sequencing by mass spectrometry. *Current Opinions in Structural Biology.* **13**, 595-601 (2003).
86. Lange, V., Picotti, P., Domon, B. & Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular Systems Biology.* **4**, 1-14 (2008).
87. Gallien, S., Duriez, E. & Domon, B. Selected reaction monitoring applied to proteomics. *Journal of Mass Spectrometry.* **46**, 298-312 (2011).
88. Ganem, B., Li, Y.T. & Henion, J.D. Observation of noncovalent enzyme substrate and enzyme product complexes by ion spray mass spectrometry. *Journal of the American Chemical Society.* **113**, 7818-7819 (1991).
89. Ganem, B., Li, Y.T. & Henion, J.D. Detection of noncovalent receptor ligand complexes by mass spectrometry. *Journal of the American Chemical Society.* **113**, 6294-6296 (1991).
90. Katta, V. & Chait, B.T. Observation of the heme-globin complex in native myoglobin by electrospray ionization mass spectrometry. *Journal of the American Chemical Society.* **113**, 8534-8535 (1991).
91. Ashcroft, A.E. Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. *Natural Product Reports.* **22**, 452-464 (2004).
92. Uetrecht, C., Versluis, C., Watts, N.R., Roos, W.H., Wuite, G.J.L., Wingfield, P.T., Steven, A.C. & Heck, A.J.R. High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. *Proceedings of the National Academy of Sciences.* **105**, 9216-9220 (2008).
93. Heck, A.J.R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods.* **5**, 927-933 (2008).
94. Uetrecht, C., Rose, R.J., Van Duijn, E., Lorenzen, K. & Heck, A.J.R. Ion mobility mass spectrometry of proteins and protein assemblies. *Chemical Society Reviews.* **39**, 1633-1655 (2010).
95. Hopper, J.T.S. & Oldham, N.J. Collision induced unfolding of protein ions in the gas phase studied by ion mobility mass spectrometry: the effect of ligand binding on conformational stability. *Journal of the American Chemical Society.* **20**, 1851-1858 (2009).
96. Pramanik, B.M., Bartner, P.L., Mirza, U.A., Liu, Y.H. & Ganguly, A.K. Electrospray ionization mass spectrometry for the study of non-covalent complexes: an emerging technology. *Journal of Mass Spectrometry.* **33**, 911-920 (1998).

97. Loo, J.A. Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *International Journal of Mass Spectrometry.* **200**, 175-186 (2000).
98. Heck, A.J.R. & Van den Heuvel, R.H.H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrometry Reviews.* **23**, 368-389 (2004).
99. Pelander, A., Decker, P., Baessmann, C. & Ojanperä, I. Evaluation of a high resolving power time-of-flight mass spectrometer for drug analysis in terms of resolving power and acquisition rate. *Journal of the American Society for Mass Spectrometry.* **22**, 379-385 (2011).
100. Hilton, G.R. & Benesch, J.L.P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of the Royal society Interface.* **9**, 801-816 (2012).
101. Mann, M., Hendrickson, R.C. & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annual Reviews of Biochemistry.* **70**, 437-73 (2001).
102. Griffiths, W.J. & Wang, Y. Mass spectrometry: from proteomics to metabolomics and lipidomics. *Chemical Society Reviews.* **38**, 1882-1896 (2008).
103. International human genome sequencing consortium. Initial sequencing and analysis of the human genome. *Nature.* **409**, 860-921 (2001).
104. Huber, L.A. Is proteomics heading in the wrong direction. *Nature Reviews in Molecular Cell Biology.* **4**, 75-80 (2003).
105. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science.* **213**, 212-217 (2006).
106. Gromov, P.S. & Celis, J.E. From genomics to proteomics *Molecular Biology.* **34**, 508-20 (2000).
107. Haynes, P.A., Gygi, S.P., Figeys, D. & Aebersold, R. Proteome analysis: biological assay or data archive. *Electrophoresis.* **19**, 1862-1871 (1998).
108. Martin, D.B. & Nelson, P.S. From genomics to proteomics: techniques and applications in cancer research *Trends in Cell Biology.* **11**, S60-S65 (2001).
109. Stasyk, T. & Huber, L.A. Zooming in: fractionation strategies in proteomics. *Proteomics.* **4**, 3704-3716 (2004).
110. Issaq, H.J. The role of separation science in proteomics research. *Electrophoresis.* **22**, 3629-3638 (2001).
111. Raymond, S. Acrylamide gel electrophoresis. *Annals of the New York Academy of Sciences.* **121**, 350-365 (1964).
112. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-685 (1970).
113. Görg, A., Weiss, W. & Dunn, M.J. Current two dimensional electrophoresis technology for proteomics. *Proteomics.* **4**, 3665-3685 (2004).
114. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols.* **1**, 2856-2860 (2006).
115. Stults, J.T. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) *Current Opinions in Structural Biology.* **5**, 691-698 (1995).
116. Wu, C. & MacCoss, M.J. Shotgun proteomics: tools for the analysis of complex biological systems. *Current Opinions in Molecular Theraputics.* **4**, 242-250 (2002).
117. Yates, J.R., Ruse, C.I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annual Review of Biomedical Engineering.* **11**, 49-79 (2009).
118. Mitulovic, G. & Mechtler, K. HPLC techniques for proteomics analysis a short overview of latest developments. *Briefings in Functional Genomics and Proteomics.* **5**, 249-260 (2006).

119. Kirkland, J.J. Development of some stationary phases for reversed-phase high-performance liquid chromatography. *Journal of Chromatography A.* **1060**, 9-211 (2004).
120. Shi, Y., Xiang, R., Horvath, C. & Wilkins, J.A. The role of liquid chromatography in proteomics. *Journal of Chromatography A.* **1053**, 27-36 (2004).
121. Chen, H. & Horwith, C.S. High-speed high-performance liquid chromatography of peptides and proteins. *Journal of Chromatography A.* **705**, 3-20 (1995).
122. Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data dependent LC-MS/MS. *Journal of Proteome Research.* **10**, 1785-1793 (2011).
123. Geromanos, S.J., Vissers, J.P.C., Silva, J.C., Dorschel, C.A., Li, G.Z., Gorenstein, M.V., Bateman, R.H. & Langridge, J.I. The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics.* **9**, 1683-1695 (2009).
124. McDonald, W.H. & Yates, J.R. Shotgun proteomics and biomarker discovery. *Disease Markers.* **18**, 99-105 (2002).
125. Ly, L. & Wasinger, V.C. Protein and peptide fractionation, enrichment and depletion: Tools for the complex proteome. *Proteomics.* **11**, 513-534 (2011).
126. Wysockia, V.H., Resingb, K.A., Zhanga, Q. & Cheng, G. Mass spectrometry of peptides and proteins. *Methods.* **35**, 211-222 (2005).
127. Bogdanov, B. & Smith, R.D. Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrometry Reviews.* **24**, 168-200 (2004).
128. Olsen, J.V., Ong, S.E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular and Cellullar Proteomics.* **3**, 608-614 (2004).
129. Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.C. & Yates, J.R. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews.* **113**, 2343-2394 (2013).
130. Swaney, D.L., Wenger, C.D. & Coon, J.J. Value of using multiple proteases for large scale mass spectrometry based proteomics. *Journal of Protrome Research.* **9**, 1323-1329 (2009).
131. Sadygov, R.G., Cociorva, D. & Yates, J.R. Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nature Methods.* **1**, 195-202 (2004).
132. Eng, J.K., McCormack, A.L. & Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database *Journal of the American Society for Mass Spectrometry.* **5**, 976-989 (1994).
133. Prerkins, D.N., Pappin, D.J.C., Creasy, D.M. & Cottrell, J.S. Probability based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* **20**, 3551-3567 (1999).
134. James, P., *Proteome Research: Mass Spectrometry.* (2001), Berlin. Springer.
135. Choi, H. & Nesvizhskii, A.I. False discovery rates and related statistical concepts in mass spectrometry-based proteomics. *Journal of Proteome Research.* **7**, 47-50 (2008).
136. Bantscheff, M., Lemeer, S., Savitski, M.M. & Kuster, B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Analytical and Bioanalytical Chemistry.* **404**, 939-965 (2012).
137. Kitteringhama, N.R., Jenkins, R.E., Lane, C.S., Elliott, V.L. & Park, K.B. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *Journal of Chromatography A.* **877**, 1229-1239 (2009).
138. Siuti, N. & Kelleher, N.L. Decoding protein modifications using top-down mass spectrometry. *Nature Methods.* **4**, 817-821 (2007).
139. Armirotti, A. & Damonte, G. Achievements and perspectives of top-down proteomics. *Proteomics.* **10**, 3566–3576 (2010).

140. Mann, M. & Jensen, O.N. Proteomic analysis of post-translational modifications. *Nature Biotcechnology.* **21**, 255-261 (2003).
141. McLafferty, F.W., Breuker, K., Jin, M., Han, X., Infusini, G., Jiang, H., Kong, X. & Begley, T.P. Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *The FEBS Journal.* **274**, 6256-6268 (2007).
142. Cui, W., Rohrs, H.W. & Gross, M.L. Top down mass spectrometry: recent developments, applications and perspectives. *The Analyst.* **136**, 3854-3864 (2011).
143. Capriotti, A.L., Cavaliere, C., Foglia, P., Samperi, R. & Laganà, A. Intact protein separation by chromatographic and/or electrophoretic techniques for top-down proteomics. *Journal of Chromatography A.* **1218**, 8760-8776 (2011).
144. Ahlf, D.R., Compton, P.D., Tran, J.C., Early, B.P., Thomas, P.M. & Kelleher, N.L. Evaluation of the compact high field orbitrap for top down proteomics of human cells. *Journal of Proteome Research.* **11**, 4308-4314 (2012).
145. Zhou, H., Ning, Z., Starr, A.E., Abu-Farha, M. & Figeys, D. Advancements in top-down proteomics. *Analytical chemistry.* **84**, 720-734 (2011).
146. LeDuc, R.D., Taylor, G.K., Kim, Y.B., Januszyk, T.E., Bynum, L.H., Sola, J.V., Garavelli, J.S. & Kelleher, N.L. ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Research.* **32**, W340-W345 (2004).
147. Li, W. & Tse, F.L.S. Dried blood spot sampling in combination with LC-MS/MS for quantitative analysis of small molecules. *Biomedical Chromatography.* **24**, 49-65 (2009).
148. Guthrie, R. & Susi, A. A simple pheynlalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics.* **32**, 338-343 (1963).
149. Parker, S.P. & Cubitt, W.D. The use of the dried blood spot sample in epidemiological studies. *Journal of Clinical Pathology.* **52**, 633-639 (1999).
150. Bhatti, P., Kampa, D., Alexander, B.H., McClure, C., Ringer, D., Doody, M.M. & Sigurdson, A.J. Blood spots as an alternative to whole blood collection and the effect of a small monetary incentive to increase participation in genetic association studies. *BMC Medical Research Methods.* **9**, 76-81 (2009).
151. Mcdade, T.W., Williams, S. & Snodgrass, J. What a drop can do: dried blood spots as a minimally invasive mathod for integrating biomarkers into population based research. *Demography.* **44**, 899-925 (2007).
152. Déglon, J., Thomas, A., Mangin, P. & Staub, C. Direct analysis of dried blood spots coupled with mass spectrometry: concepts and biomedical applications. *Analytical Bioanalytical Chemistry.* **402**, 2485-2498 (2012).
153. Mei, J.V., Alexander, R., Adam, B.W. & Hannon, H. Use of filter paper for the collection and analysis of human whole blood specimens. *Journal of Nutrition.* **131**, 1631-1637 (2001).
154. Naylor E W & Guthrie R Newborn screning for maple syrup urine disease (branched-chain ketoaciduria). *Pediatrics.* **61**, 262-266 (1978).
155. Mortona, H. & Kelley, R.I. Diagnosis of medium-chain acyl-coenzyme A dehydrogenase deficiency in the neonatal period by measurement of medium-chain fatty acids in plasma and filter paper blood samples. *Journal of Pediatrics.* **117**, 439-442 (1990).
156. Wilcken, B., Wiley, V., Hammond, J. & Carpenter, K. Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. *The New England Journal of Medicine.* **348**, 2304-2312 (2003).
157. Werner, E., Francois -Heilier, J., Ducruixa, C., Ezana, E., Junot, C. & Tabet, J.C. Mass spectrometry for the identification of the discriminating signals from metabolomics: Current status and future trends. *Journal of Chromatography B.* **871**, 143-163 (2008).

158. Zytkovicz, T.H., Fitzgerald, E.F., Marsden, D., Larson, C.A., Shih, V.E., Johnson, D.M., Strauss, A.W., Comeau, A.M., Eaton, R.B. & Grady, G.F. Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two year summary from the new england newborn screening program. *Clinical Chemistry.* **47**, 1945-1955 (2001).
159. Chace, D.H., Kalas, T.A. & Naylor, E.W. Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. *Clinical Chemistry.* **49**, 1797-1817 (2003).
160. Schulze, A., Lindner, M., Kohlmuller, D., Olgemoller, K., Mayatepek, E. & Hoffmann, G.F. Expanded newborn screening for inborn errors of metabolism by electrospray ionization tandem mass spectrometry: results, outcome, and implications. *Pediatrics.* **111**, 1399-1406 (2003).
161. Spooner, N., Lad, R. & Barfield, M. Dried blood spots as a sample collection technique for the determination of pharmacokinetics in clinical studies: considerations for the validation of a quantitative bioanalytical method. *Analytical Chemistry.* **81**, 1557-1563 (2009).
162. Martin, N.J. & Cooper, H.J. Challenges and opportunities in mass spectrometric analysis of proteins from dried blood spots. *Expert Reviews in Proteomics.* **11**, 685-695 (2014).
163. Jacobs, J.M., Adkins, J.N., Qian, W.J., Liu, T., Shen, Y., Camp, D.G. & Smith, R.D. Utilizing Human Blood Plasma for Proteomic Biomarker Discovery. *Journal of Proteome Research.* **4**, 1073-1085 (2005).
164. Edwards, R.L., Martin, N.J. & Cooper, H.J. Hemoglobin variant analysis of whole blood and dried blood spots by MS. *Bioanalysis.* **5**, 2043-2052 (2013).
165. Ryan, K., Bain, B.J., Worthington, D., James, J., Plews, D., Mason, A., Roper, D. & Rees, D.C. Significant haemoglobinopathies: guidelines for screening and diagnosis. *British Journal of Haematology.* **149**, 35-49 (2010).
166. Daniel, Y.A., Turner, C., Haynes, R.M., Hunt, B.J. & Dalton, N.R. Rapid and specific detection of clinically significant haemoglobinopathies using electrospray mass spectrometry–mass spectrometry. *British Journal of Haematology.* **130**, 635-643 (2005).
167. Daniel, Y.A., Turner, C., Haynes, R.M., Hunt, B.J. & Dalton, N.R. Quantification of hemoglobin A2 by tandem mass spectrometry. *Clinical Chemistry.* **53**, 1448-1454 (2007).
168. Boemer, F., Ketelslegers, O., Minon, J.M., Bours, V. & Schoos, R. Newborn screening for sickle cell disease using tandem mass spectrometry. *Clinical Chemistry.* **54**, 2036-2041 (2008).
169. Boemer, F., Cornet, Y., Libioulle, C., Segers, K., Bours, V. & Schoos, R. 3-years experience review of neonatal screening for hemoglobin disorders using tandem mass spectrometry. *Clinica Chimicha Acta.* **412**, 1476-1479 (2011).
170. Anderson, N.L. & Anderson, N.G. The human plasma proteome. *Molecular and Cellullar Proteomics.* **1**, 845-867 (2002).
171. Dewilde, A., Sadilkova, K., Sadilek, M., Vasta, V. & Houn Hahn, S. Tryptic peptide analysis of ceruloplasmin in dried blood spots using liquid chromatography tandem mass spectrometry: application to newborn screening. *Clinical Chemistry.* **54**, 1961-1968 (2008).
172. Ala, A., Walker, A., Keyoumars, A., Dooley, J.S. & Schilsky, M.L. Wilson's disease. *The Lancet.* **369**, 397-408 (2007).
173. Kehler, J.R., Bowen, C.L., Boram, S.L. & Evans, C.A. Application of DBS for quantitative assessment of the peptide exendin-4; comparison of plasma and DBS method by UHPLC-MS/MS. *Bioanalysis.* **2**, 1461-1468 (2010).

174. Sleczka, B.G., D'Arienzo, C.J., Tymiak, A.A. & Olah, T.V. Quantitation of therapeutic proteins following direct trypsin digestion of dried blood spot samples and detection by LC–MS-based bioanalytical methods in drug discovery. *Bioanalysis.* **4**, 29-40 (2012).
175. Cox, H.D., Rampton, J. & Eichner, D. Quantification of insulin-like growth factor-1 in dried blood spots for detection of growth hormone abuse in sport. *Analytical Bioanalytical Chemistry.* **405**, 1949-1958 (2013).
176. Chambers, A.G., Percy, A.J., Yang, J., Camenzind, A.G. & Borchers, C.H. Multiplexed quantitation of endogenous proteins in dried blood spots by multiple reaction monitoring mass spectrometry. *Molecular and Cellullar Proteomics.* **12**, 781-791 (2013).
177. Edwards, R.L., Griffiths, P., Bunch, J. & Cooper, H.J. Top down proteomics and direct surface sampling of neonatal dried blood spots: diagnosis of unknown hemoglobin variants. *Journal of the American Society for Mass Spectrometry.* **23**, 1921-1930 (2012).
178. Anderson, N.L. The Clinical Plasma Proteome: A Survey of Clinical Assays for Proteins in Plasma and Serum. *Clinical Chemistry.* **56**, 177-185 (2010).
179. Nanjappa, V., Thomas, J.K., Marimuthu, A., Muthusamy, B., Radhakrishnan, A., Sharma, R., Khan, A.A., Balakrishnan, L., Sahasrabuddhe, N.A., Kumar, S., Nitinbhai Jhaveri, B., Sheth, K.V., Khatana, R.K., Shaw, P.G., Manda Srikanth, S., Mathur, P.P., Sunkar, S., Nagaraja, D., Christopher, R., Mathivanan, S., Raju, R., Sirdeshmukh, R., Chatterjee, A., Simpson, R.J., Harsha, H.C., Pandey, A. & Keshava Prasad, T.S. Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic Acids Research.* **42**, D959-D965 (2014).
180. Anderson, N.L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R.S., Conrads, T.P., Veenstra, T.D., Adkins, J.N., Pounds, J.G., Fagan, R. & Lobley, A. The human plasma proteome, a non redundant list composed of 4 separate sources. *Molecular and Cellullar Proteomics.* **3**, 311-326 (2004).
181. Anderson, L. & Hunter, C.L. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular and Cellullar Proteomics.* **5**, 573-588 (2006).
182. Brantly. M, Nukiwa, T. & Crystal, R.G. Molecular basis of alpha-l -antitrypsin deficiency *American Journal of Medicine.* **84**, 13-31 (1988).
183. Carrell, R., W., Jeppsson, J.O., Laurell, C.B., Brennan, S.O., Owen, M.C., Vaughan, L. & Boswell, R.D. Structure and variation of human α1-antitrypsin. *Nature.* **298**, 329-334 (1982).
184. Köhnlein, T. & Welte, T. Alpha-1 antitrypsin deficiency: pathogenesis, clinical presentation, diagnosis, and treatment. *American Journal of Medicine.* **121**, 3-9 (2008).
185. Mizgerd, J.P. Molecular mechanisms of neutrophil recruitment elicited by bacteria in the lungs. *Semminars in Immunology.* **14**, 123-132 (2002).
186. Carrell, R.W. & Lomas, D.A. Alpha-1-antitrypsin deficiency—a model for conformational diseases. *The New England Journal of Medicine.* **346**, 45-53 (2002).
187. Kelly, E., Greene, C.M., Carroll, T.P., McElvaney, N.G. & O'Neill, S.J. Alpha-1 antitrypsin deficiency. *Respiratory Medicine Case Reports.* **4**, 1-8 (2011).
188. Ferrarotti, I., Zorzetto, M., Scabini, R., Mazzola, P., Campo, I. & Luisetti, M. A novel method for rapid genotypic identification of alpha1-antitrypsin variants. *Daignostic Molecular Pathology.* **13**, (2004).
189. Kelly, E., Greene, C.M., Carroll, T.P., McElvaney, N.G. & O'Neill, S.J. Alpha-1 antitrypsin deficiency. *Respiratory Medicine.* **104**, 763-772 (2010).
190. Abboud, R.T. Alpha1-antitrypsin deficiency. *Respiratory Medicine.* **1**, 80-87 (2006).
191. De Serres, F.J. Alpha-1 antitrypsin deficiency is not a rare disease but a disease that is rarely diagnosed. *Environmental Health Perspectives.* **111**, 1852-1854 (2003).

192. Ferrarotti, I., Scabini, A., Campo, I., Ottaviani, S., Zorzetto, M., Gorrini, M. & Luisetti, M. Laboratory diagnosis of alpha1-antitrypsin deficiency. *Translational Research.* **150**, 267-274 (2007).
193. Kuriena, B.T. & Scofield, R.H. Protein blotting: a review. *Journal of Immunological Methods.* **274**, 1-15 (2003).
194. Mozdzanowski, J. & Speicher, D.W. Microsequence analysis of electroblotted proteins, comparison of electroblotting recoveries using different types of PVDF membranes. *Analytical Biochemistry.* **207**, 11-18 (1992).
195. Schägger, H. Tricine–SDS-PAGE. *Nature Protocols.* **1**, 16-22 (2006).
196. Jensen, E.C. The basics of western blotting. *The Anatomical Record.* **295**, 369-371 (2012).
197. Pandey, A. & Mann, M. Proteomics to study genes and genomes. *Nature.* **405**, 837-846 (2000).
198. Rundlett, K.L. & Armstrong, D.W. Mechanism of signal suppression by anionic surfactants in capillary electrophoresis electrospray ionization mass spectrometry. *Analytical Chemistry.* **68**, 3493-3497 (1996).
199. Gudiksen, K.L., Gitlin, I. & Whitesides, G.M. Differentiation of proteins based on characteristic patterns of association and denaturation in solutions of SDS. *Proceedings of the National Academy of Sciences* **103**, 7968-7972 (2006).
200. Shieh, I.F., Lee, C.Y. & Shiea, J. Eliminating the interferences from TRIS buffer and SDS in protein analysis by fused-droplet electrospray ionization mass spectrometry. *Journal of Proteome Research.* **4**, 606-612 (2005).
201. Ilka, W., Braun, H. & Schagger, H. Blue native PAGE. *Nature Protocols.* **1**, 418-428 (2006).
202. Wittig, I. & Schagger, H. Native electrophoretic techniques to identify protein–protein interactions. *Proteomics.* **9**, 5214-5223 (2009).
203. Mozdzanowski, J., Hembach, P. & Speicher, D.W. High yield electroblotting onto polyvinylidene diflouride membranes from polyacrylamide gels. *Electrophoresis.* **13**, 59-64 (1992).
204. Martin, N.J., Bunch, J. & Cooper, H.J. Dried blood spot proteomics: surface extraction of endogenous proteins coupled with automated sample preparation and mass spectrometry analysis. *Journal of the American Society for Mass Spectrometry.* **24**, 1242-1249 (2013).
205. WHO, *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity*, WHO, Editor. 2011: Geneva, Switzerland. p. 6.
206. Greenough, C., Jenkins, R.E., Kitteringham, N.R., Pirmohamed, M., Park, K. & Pennington, S.R. A method for the rapid depletion of albumin and immunoglobulin from human plasma. *Proteomics.* **4**, 3107-3111 (2004).
207. Nagy, J.A., Benjamin, L., Zeng, H., Dvorak, A.M. & Dvorak, H.V. Vascular permeability, vascular hyperpermeability and angiogenesis. *Angiogenesis.* **11**, 109-119 (2008).
208. Hortin, G.L., Sviridov, D. & Anderson, N.L. High abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clinical Chemistry.* **54**, 1608-1616 (2008).
209. Klammer, A.A. & MacCoss, M.J. Effects of modified digestion schemes on the identification of proteins from complex mixtures. *Journal of Proteome Research.* **5**, 695-700 (2006).
210. Sim, R.B. & Laich, A. Serine Proteases of the complement system. *Biochemical Society Transactions.* **28**, 545-550 (2000).
211. Chambers R C, L.G.J. Coagulation cascade proteases and tissue fibrosis. *Biochemical society transactions.* **30**, 194-200 (2002).
212. Wildes, D. & Wells, J.A. Sampling the N-terminal proteome of human blood. *Proceedings of the National Academy of Sciences.* **107**, 4561-4566 (2010).

213. Chambers, A.G., Percy, A.J., Hardie, D.B. & Borchers, C.H. Comparison of proteins in whole blood and dried blood spot samples by LC/MS/MS. *Journal of the American Society for Mass Spectrometry.* **24**, 1338-1345 (2013).
214. Shen, Y., Jacobs, J.M., Camp, D.G., Fang, R., Moore, J.M. & Smith, R.D. Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Analytical Chemistry.* **76**, 1134-1144 (2004).
215. Haudek, V.J., Slany, A., Gundacker, N.C., Wimmer, H., Drach, J. & Gerner, C. Proteome maps of the main human peripheral blood constituents. *Journal of Proteome Research.* **8**, 3834-3843 (2008).
216. Wild, B.J., B., G.N., Cooper, E.K., Lalloz, M.R.A., Erten, S., Stephens, A.D. & Layton, D.M. Rapid identification of hemoglobin variants by electrospray ionization mass spectrometry. *Blood Cells Molecules and Diseases.* **27**, 691-704 (2001).
217. Rai, D.K., Griffiths, W.J., Landin, B., Wild, B.J., Alvelius, G. & Green, B.N. Accurate mass measurement by electrospray ionization quadrupole mass spectrometry: detection of variants differing by <6 Da from normal in human hemoglobin heterozygotes. *Analytical Chemistry.* **75**, 1978-1982 (2003).
218. Foell, D., Froschb, M., Sorga, C. & Roth, J. Phagocyte-specific calcium-binding S100 proteins as clinical laboratory markers of inflammation. *Clinica Chimicha Acta.* **344**, 37-51 (2004).
219. Sarsby, J., Martin, N.J., Lalor, P.F., Bunch, J. & Cooper, H.J. Top-down and bottom-up identification of proteins by liquid extraction surface analysis mass spectrometry of healthy and diseased human liver tissue. *Journal of the American Society for Mass Spectrometry.* **25**, 1953-1961 (2014).
220. Brownridge, P. & Beynon, R.J. The importance of the digest: proteolysis and absolute quantification in proteomics. *Methods.* **54**, 351-360 (2011).
221. Siepen, J.A., Keevil, E.J., Knight, D. & Hubbard, S.J. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *Journal of Proteome Research.* **6**, 399-408 (2006).
222. Tonacka, S., Aspinall-O`Deaa, M., Jenkinsb, R.E., Elliot, V., Murraya, S., Lanec, C.S., Kitteringhamb, N.R., Neoptolemosa, J.P. & Costello, E. A technically detailed and pragmatic protocol for quantitative serum proteomics using iTRAQ. *Journal of Proteome Research.* **73**, 352-356 (2009).
223. Polaskova, V., Kapur, A., Khan, A., Molloy, M.P. & Bake, M.S. High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. *Electrophoresis.* **31**, 471-482 (2010).
224. Liu, T., Qian, W.J., Mottaz, H.M., Gritsenko, M.A., Norbeck, A.D., Moore, R.J., Purvine, S.O., Camp, D.G. & Smith, R.D. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Molecular and Cellullar Proteomics.* **5**, 2167-2174 (2006).
225. Walpurgis, K., Kohler, M., Thomas, A., Wenzel, F., Geyer, H., Schanzer, W. & Thevis, M. Validated hemoglobin-depletion approach for red blood cell lysate proteome analysis by means of 2D PAGE and Orbitrap MS. *Electrophoresis.* **33**, 2537-2545 (2012).
226. Jones, K.A., Kim, P.D., Patel, B.B., Kelsen, S.G., Braverman, A., Swinton, D.J., Gafken, P.R., Jones, L.A., Lane, W.S., Neveu, J.M., Leung, H.C.E., Shaffer, S.A., Leszyk, J.D., Stanley, B.A., Fox, T.E., Stanley, A., Hall, M.J., Hampel, H., South, C.D., De la Chapelle, A., Burt, R.W., Jones, D.A., Kopelovich, L. & Yeung, A.T. Immunodepletion plasma proteomics by tripleTOF 5600 and orbitrap elite/LTQ-orbitrap velos/Q exactive mass spectrometers. *Journal of Proteome Research.* **12**, 4351-4365 (2013).

227. Stoller, J.K. & Aboussouan, L.S. 1-antitrypsin deficiency. *The Lancet.* **365**, 2225-2236 (2005).
228. Costa, X., Jardi, R., Rodriguez, F., Miravitlles, M., Cotrina, M., Gonzalez, C., Pascual, C. & Vidal, R. Simple method for a1-antitrypsin deficiency screening by use of dried blood spot specimens. *European Respiratory Journal.* **15**, 1111-1115 (2000).
229. Chen, Y., Snyder, M.R., Zhu, Y., Tostrud, L.J., Benson, L.M., Katzmann, J.A. & Bergen, R.H. Simultaneous Phenotyping and Quantification of-1-Antitrypsin by Liquid Chromatography–Tandem Mass Spectrometry. *Clinical Chemistry.* **57**, 1161-1168 (2011).
230. Zhang Y, Ju Y, Huang C & Wysocki V H Paper spray ionization of noncovalent protein complexes. *Anal. Chem.* **86**, 1342-1346 (2014).
231. Arakawa, T., Prestrelski, S.J., Kenney, W.C. & Carpenter, J.F. Factors affecting short-term and long-term stabilities of proteins. *Advanced Drug Delivery Reviews.* **46**, 307-326 (2001).
232. Martin N J, Griffiths R L, Edwards R L & Cooper H J Native liquid extraction surface analysis mass spectrometry: analysis of noncovalent protein complexes directly from dried substrates. *J. Am. Soc. Mass. Spectrom.* (2015).
233. Chu K, Vojtechovsky J, Mcmahon B H, Sweet R M, Berendzen J & Schlichting I Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature.* **403**, 921-923 (2000).
234. Savino C, Miele A E, Draghi F, Johnson K A, Sciara G, Brunori M & Vallone B Pattern of cavities in globins: the case of human hemoglobin. *Biopolymers.* **91**, 1197-1107 (2009).
235. Moini, M. Metal displacement and stoichiometry of protein-metal complexes under native conditions using capillary electrophoresis/mass spectrometry. *Rapid Communications in Mass Spectrometry.* **24**, 2730-2735 (2010).
236. Pan J, Kun Xu, Yang X, Choy W Y & Konermann L Solution phase chelators for suppressing nonspecific protein-metal interactions in electrospray mass spectrometry. *Anal. Chem.* **81**, 5008-5015 (2009).
237. Moulin, A.M., O'Shea, S.J., Badley, R.A., Doyle, P. & Welland, M.E. Measuring surface-induced conformational changes in proteins. *Langmuir.* **15**, 8776-8779 (1999).
238. Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *International Journal of Mass Spectrometry.* **219**, 39-62 (2002).
239. Benesch, J.L.P., Sobott, F. & Robinson, C.V. Thermal dissociation of multimeric protein complexes by using nanoelectrospray mass spectrometry. *Analytical Chemistry.* **75**, 2208-2214 (2003).
240. Edwards, R.L., *Direct surface sampling of dried blood spots coupled with mass spectrometry for haemoglobin analysis*, in *School of Biosciences*. 2013, The University of Birmingham: Birmingham.
241. Simmons, D.A., Wilson, D.J., Lajoie, G.A., Doherty-Kirby, A. & Konermann, L. Subunit disassembly and unfolding kinetics of hemoglobin studied by time-resolved electrospray mass spectrometry. *Biochemistry.* **43**, 14792-14801 (2004).
242. Scarff, C.A., Patel, V.J., Thalassinos, K. & Scrivens, J.H. Probing hemoglobin structure by means of traveling-wave ion mobility mass spectrometry. *Journal of the American Society for Mass Spectrometry.* **20**, 625-631 (2009).
243. Lössl, P., Snijder, J. & Heck, A.J.R. Boundaries of mass resolution in native mass spectrometry. *Journal of the American Society for Mass Spectrometry.* **25**, 906-917 (2014).
244. Marshall, W.J. & Bangert, S.K., *Clinical Chemistry.* Sixth Edition. (2008), Edinburgh. Elsevier.

245. Veenstra, T.D. Electrospray ionization mass spectrometry in the study of biomolecular non-covalent interactions. *Biophysical Chemistry.* **79**, 63-79 (1999).
246. Griffiths, R.L. & Cooper, H.J. Direct tissue profiling of protein complexes: toward native mass spectrometry imaging. *Analytical Chemistry.* **88**, 606-609 (2016).
247. Cao, A. & Galanello, R. Beta-thalassemia. *Genetics in Medicine.* **12**, 61-76 (2010).
248. Munice, H.L. & Campbell, J.S. Alpha and beta thalassemia. *American Acadamey Family Physicians.* **80**, 339-344 (2009).
249. Galanello, R. & Origa, R. Beta thalassemia. *Orphanet Journal of Rare Diseases.* **5**, 1-15 (2010).
250. Montowska, M., Rao, W., Alexander, M.R., Tucker, G.A. & Barrett, D.A. Tryptic digestion coupled with ambient desorption electrospray ionization and liquid extraction surface analysis mass spectrometry enabling identification of skeletal muscle proteins in mixtures and distinguishing between beef, pork, horse, chicken, and turkey Meat. *Analytical Chemistry.* **86**, 4479-4487 (2014).
251. Wisztorski, M., Fatou, B., Franck, J., Desmons, A., Farre, I., Leblanc, E., Fournier, I. & Salzet, M. Microproteomics by liquid extraction surface analysis: application to FFPE tissue to study the fimbria region of tubo-ovarian cancer. *Proteomics Clinical Applications.* **7**, 1-7 (2013).
252. Kim, J.H., Woenker, T., Adamec, J. & Regnier, F.E. Simple, miniaturized blood plasma extraction method. *Analytical Chemistry.* **85**, 11501-11508 (2013).
253. Li, Y., Henion, J., Abbott, R. & Wang, P. The use of a membrane filtration device to form dried plasma spots for the quantitative determination of guanfacine in whole blood. *Rapid Communications in Mass Spectrometry.* **26**, 1208-1212 (2012).
254. Kuhn, E., Addona, T., Keshishian, H., Burgess, M., Mani, D.R., Lee, R.T., Sabatine, M.S., Gerszten, R.E. & Carr, S.A. Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry. *Clinical Chemistry.* **55**, 1108-1117 (2009).
255. Anderson, L. Six decades searching for meaning in the proteome. *Journal of Proteome Research.* **107**, 24-30 (2014).

Appendices

Appendix 4

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Ig kappa chain C region	49.06 %	2	98.30	48.65
Serum albumin	32.18 %	25	1969.35	323.34
Apolipoprotein A-II	21.00 %	2		4.01
Uncharacterized protein C7orf13	20.37 %	1		0.00
Apolipoprotein A-I	16.48 %	4	409.85	33.26
Ig gamma-1 chain C region	10.91 %	3	58.63	25.90
Ig lambda-2 chain C regions	9.43 %	1	65.62	16.62
Alpha-1-antitrypsin	7.89 %	3		32.17
Alpha-1-acid glycoprotein 1	7.46 %	1		0.00
Haptoglobin	3.94 %	2	34.43	7.95
Hemicentin-2	2.92 %	1		0.00
Metal transporter CNNM1	2.80 %	1		4.28
Fibrinogen alpha chain	1.96 %	1	90.95	28.55
Serotransferrin	1.72 %	1		3.16

Appendix 4.1. Proteins identified from direct infusion ESI mass spectrometry of dried plasma spot digested for 1 hour.

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Ig kappa chain C region	49.06 %	2	269.85	40.06
Serum albumin	32.68 %	24	1886.53	249.15
Ig lambda-2 chain C regions	27.36 %	2	36.53	6.72
Apolipoprotein A-I	26.97 %	6	584.63	51.49
Ig gamma-1 chain C region	26.67 %	4	267.51	37.74
Ig gamma-4 chain C region	17.13 %	1	124.58	16.39
Alpha-1-antitrypsin	16.51 %	5	54.50	34.19
Apolipoprotein A-II	11.00 %	1		2.44
Fibrinogen beta chain	7.72 %	1		3.32
Heterogeneous nuclear ribonucleoprotein L	7.64 %	1	23.53	
ATP-binding cassette sub-family G member 5	6.78 %	1		0.00
Haptoglobin	5.91 %	3	65.29	9.83
Isoform Gamma-A of Fibrinogen gamma chain	5.26 %	1		0.00
28S ribosomal protein S5, mitochondrial	3.49 %	1		1.74
Coiled-coil domain-containing protein 54	3.35 %	1		1.85
Roundabout homolog 2	2.98 %	1		1.99
Serotransferrin	1.72 %	1		2.87
Fibrinogen alpha chain	1.50 %	1	29.06	2.78

Appendix 4.2 Proteins identified from direct infusion ESI mass spectrometry of dried plasma spot digested for 12 hours.

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Hemoglobin subunit delta	99.32 %	20	2645.90	176.47
Hemoglobin subunit beta	99.32 %	29	6942.99	649.81
Hemoglobin subunit alpha	99.30 %	27	4062.92	505.25
Apolipoprotein A-I	67.42 %	20	714.65	68.89
Ig kappa chain C region	66.98 %	6	316.99	25.96
Carbonic anhydrase 1	62.45 %	13	813.29	79.76
Carbonic anhydrase 2	44.23 %	8	271.74	24.38
Flavin reductase (NADPH)	41.75 %	5	183.80	15.63
Alpha-1-antitrypsin	37.80 %	11	265.69	29.31
Peroxiredoxin-	36.36 %	8	343.73	29.35
Serum albumin	31.86 %	22	829.90	68.46
Protein S100-A6	31.11 %	1	45.61	2.93
Alpha-synuclein	27.14 %	2	101.94	6.53
Glyceraldehyde-3-phosphate dehydrogenase	21.49 %	5	99.94	3.06
Catalase	21.06 %	8	208.81	19.52
Apolipoprotein C-III	19.19 %	2	132.87	12.13
Ig lambda-7 chain C region	18.87 %	2	85.57	7.66
Actin, cytoplasmic 1	17.07 %	5	74.59	8.26
Fibrinogen gamma chain	16.56 %	4	131.40	17.45
Transthyretin	15.65 %	1	41.02	3.11
Granulysin	15.12 %	1		0.00
Hemoglobin subunit gamma-2	14.97 %	2	748.88	40.83
N-glycosylase/DNA lyase	14.21 %	1		0.00
Leucine-rich repeat-containing protein 55	14.15 %	1		0.00
Fibrinogen beta chain	13.44 %	5	106.56	14.77
Apolipoprotein C-I	13.25 %	1	25.22	3.10
Serotransferrin	12.75 %	8	272.09	33.29
Ig gamma-1 chain C region	12.73 %	3	110.13	9.70
Ig alpha-1 chain C region	12.18 %	4	139.07	8.75
Apolipoprotein F	12.01 %	1		0.00
Ig mu chain C region	11.95 %	5	140.44	15.99
Band 3 anion transport protein	11.31 %	6	224.24	22.59
Acyl-coenzyme A thioesterase 1	10.45 %	1		0.00
Olfactory receptor 2F1 [	9.15 %	1		0.00
Fibrinogen alpha chain	9.01 %	6	254.17	22.99
Protein Smaug homolog 1	8.12 %	1		0.00
Isoform 6 of Osteoclast-associated immunoglobulin-like receptor	7.94 %	1		2.65
Peptidyl-prolyl cis-trans isomerase	7.88 %	1	25.76	2.83
Complement C3	7.58 %	9	261.27	24.08
Isoform 2 of Tetratricopeptide repeat protein 8	7.20 %	1		0.00
Alpha-2-macroglobulin	6.65 %	6	178.99	11.90
Insulinoma-associated protein 1	6.27 %	1		0.00

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Ig gamma-2 chain C region	5.83 %	2	57.11	2.33
Alpha-1B-glycoprotein	5.66 %	2	33.02	2.34
Apolipoprotein A-IV	5.56 %	2	89.53	4.98
Alpha-2-HS-glycoprotein	5.45 %	1	71.91	5.17
Serum amyloid P-component	4.48 %	1	42.89	2.31
Isoform 2 of Cingulin	3.89 %	1		0.00
Uncharacterized protein C7orf63	3.40 %	1	27.02	
Complement C4-B alpha chain	3.36 %	1		1.66
Haptoglobin	3.20 %	1	47.06	3.64
Serine/threonine-protein kinase MRCK alpha (Fragment)	3.11 %	1		2.02
Spectrin beta chain, erythrocytic	2.90 %	5	96.77	9.27
L-lactate dehydrogenase B chain	2.69 %	1	67.51	2.11
Spectrin alpha chain, erythrocytic 1	2.44 %	5	88.18	12.27
Kininogen-1	2.30 %	1		1.69
T-box transcription factor TBX18	2.30 %	1		0.00
Ankyrin-1	2.18 %	2	34.54	3.70
Solute carrier family 2, facilitated glucose transporter member 1	2.03 %	1	58.50	2.79
Alpha-1-antichymotrypsin	1.65 %	1	22.67	1.61
35 kDa inter-alpha-trypsin inhibitor heavy chain H4 (Fragment)	1.53 %	1		0.00
Putative uncharacterized protein C12orf63	1.36 %	1		0.00
Inter-alpha-trypsin inhibitor heavy chain H1	1.32 %	1	55.99	2.48
DNA-binding protein RFX5	1.22 %	1		0.00
Isoform 2 of Protein TANC1	0.63 %	1		0.00
Isoform 2 of Histone-lysine N-methyltransferase, H3 lysine-36 and H4 lysine-20 specific	0.58 %	1		0.00
Apolipoprotein B-100	0.55 %	2	61.07	2.55

Appendix 4.3 Proteins identified from LC MS/MS analysis of on card proteolysis

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Hemoglobin subunit alpha	74.65 %	16	737.74	124.29
Hemoglobin subunit beta	69.39 %	7	1059.32	96.86
Transthyretin	59.86 %	7	123.41	31.63
Hemoglobin subunit delta	57.82 %	3	724.67	55.26
Carbonic anhydrase 1	54.02 %	9	233.00	24.65
Ig kappa chain C region	49.06 %	3	68.52	13.82
Apolipoprotein A-I	42.32 %	13	416.75	40.70
Hemoglobin subunit gamma-1	39.46 %	3	317.79	23.11
Glyceraldehyde-3-phosphate dehydrogenase	37.31 %	5	147.81	25.76
Alpha-1-antitrypsin	34.45 %	13	337.62	37.11
ADP-ribosylation factor-like protein 6-interacting protein 1	31.82 %	1		2.53
Flavin reductase (NADPH)	30.58 %	3	20.69	0.00
Actin, cytoplasmic 1	29.87 %	8	95.47	26.41
Apolipoprotein C-III	27.27 %	2	78.81	4.71
Apolipoprotein C-I	26.51 %	2		2.37
Complement C3	25.56 %	27	668.82	86.91
Serum albumin	25.29 %	18	543.74	59.86
Dermcidin	24.55 %	3	65.30	8.40
14-3-3 protein zeta/delta	23.67 %	3	77.82	14.37
dCTP pyrophosphatase 1	23.53 %	2	35.17	5.86
Peroxiredoxin-2	23.23 %	5	122.02	12.76
Profilin	22.73 %	1		0.00
Ig gamma-1 chain C region	22.42 %	3	98.02	13.93
Fibrinogen beta chain	22.20 %	8	89.79	13.44
Fibrinogen gamma chain	22.08 %	7	199.59	26.20
Heat shock 70 kDa protein 1A/1B	20.75 %	6	106.68	18.04
RAC-alpha serine/threonine-protein kinase (Fragment)	20.00 %	1		0.00
Complement component 1 Q subcomponent-binding protein, mitochondrial	19.86 %	2	79.42	3.98
14-3-3 protein epsilon	19.22 %	2	95.47	18.75
Alpha-2-macroglobulin	19.20 %	18	489.04	59.80
Mitochondrial fission 1 protein	19.08 %	1	39.32	2.89
Isoform 2 of COP9 signalosome complex subunit]	18.75 %	1		0.00
Apolipoprotein A-IV	18.69 %	5	120.28	15.36
Ceruloplasmin	18.59 %	11	220.05	37.29
Carbonic anhydrase 2	18.46 %	3	124.89	14.09
Hemopexin	18.40 %	7	157.52	19.30
Alpha-1B-glycoprotein	18.18 %	4	117.28	7.77
Protein FAM98B	17.82 %	1		2.31
Phosphoglycerate mutase 1	17.72 %	2	44.94	11.89
60 kDa heat shock protein, mitochondrial	17.63 %	5	195.46	16.92
Abhydrolase domain-containing protein 10, mitochondrial	17.56 %	1		0.00
Peroxiredoxin-6	16.96 %	2	46.12	2.56
Heat shock cognate 71 kDa protein	16.87 %	4	109.98	23.70
Inter-alpha-trypsin inhibitor heavy chain H2	16.70 %	9	159.11	29.79
L-lactate dehydrogenase B chain	16.47 %	3	105.69	14.11
Nuclear autoantigenic sperm protein (Fragment)	16.00 %	1		3.48
Glutathione reductase, mitochondrial (Fragment)	15.86 %	1		2.24
Alpha-synuclein	15.71 %	1	37.24	3.27
Heat shock protein HSP 90-beta	15.61 %	2	158.68	16.11
V-type proton ATPase subunit D (Fragment)	15.03 %	1		0.00
Eukaryotic translation initiation factor 5A-1	14.94 %	1	28.22	2.76
Isoform 2 of Acidic leucine-rich nuclear phosphoprotein 32 family member B	14.87 %	1		2.46
Ig kappa chain V-III region SIE	14.68 %	1	74.61	3.20
Fibrinogen alpha chain	14.67 %	10	269.47	36.33
Angiotensinogen	14.64 %	4	163.04	11.73
Keratin, type I cytoskeletal 10	14.55 %	5	59.90	13.84
Alpha-trypsin chain 2 (Fragment)	14.08 %	1		2.02
Clusterin	14.03 %	4	73.11	9.43
26S proteasome non-ATPase regulatory subunit 13 (Fragment)	13.57 %	1		0.00
Stress-induced-phosphoprotein 1	13.56 %	1		0.00

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Elongation factor 1-gamma	12.81 %	3	23.10	5.13
Spliceosome RNA helicase DDX39B (Fragment)	12.77 %	1		0.00
39S ribosomal protein L12, mitochondrial	12.12 %	1	20.86	2.52
Protein SET	12.08 %	1		0.00
L-lactate dehydrogenase A chain	12.05 %	2	72.78	12.99
Translin (Fragment)	11.89 %	1		0.00
Vesicle-fusing ATPase (Fragment)	11.80 %	1		2.62
Platelet basic protein	11.72 %	1	35.76	3.14
Acidic leucine-rich nuclear phosphoprotein 32 family member A	11.65 %	1	23.85	2.44
Poly(rC)-binding protein 2 (Fragment)	11.41 %	1		2.40
Protein S100-A9	11.40 %	1		2.73
Heat shock protein HSP 90-alpha	10.66 %	1	276.57	24.98
Eukaryotic translation initiation factor 3 subunit H (Fragment)	10.64 %	1		2.06
T-complex protein 1 subunit beta	10.47 %	3	96.42	13.68
DNA damage-binding protein 1 (Fragment)	10.44 %	1		4.99
Keratin, type II cytoskeletal 1	10.40 %	5	137.78	14.85
Inosine-5'-monophosphate dehydrogenase 2 (Fragment)	10.39 %	1		0.00
Ig alpha-1 chain C region	10.20 %	3	68.27	4.47
60S ribosomal protein L10a	10.14 %	1	23.43	0.00
Ig gamma-3 chain C region	10.08 %	1	70.11	6.69
Stress-70 protein, mitochondrial	10.01 %	2	85.76	14.59
Apolipoprotein A-II	10.00 %	1	55.27	2.49
Heterogeneous nuclear ribonucleoprotein U (Fragment)	9.92 %	1		5.69
T-complex protein 1 subunit alpha	9.89 %	2	116.36	6.30
HBS1-like protein (Fragment)	9.88 %	1		0.00
ADP/ATP translocase 2	9.73 %	1		2.09
Inter-alpha-trypsin inhibitor heavy chain H4	9.68 %	7	159.84	22.13
Inter-alpha-trypsin inhibitor heavy chain H1	9.44 %	5	54.35	12.42
Gelsolin	9.34 %	4	109.25	15.69
Coatomer subunit epsilon	9.09 %	1	22.82	2.07
Methylosome protein 50 (Fragment)	8.93 %	1		0.00
Fructose-2,6-bisphosphatase TIGAR	8.52 %	1	0.00	
Histidine-rich glycoprotein	8.38 %	3	29.93	2.03
Fructose-bisphosphate aldolase A	8.24 %	1	27.43	2.55
Complement C4-A	8.03 %	8	238.00	
Serotransferrin	8.02 %	5	96.14	13.80
Tubulin alpha-4A chain	7.81 %	2	61.83	2.83
Alcohol dehydrogenase class-3	7.49 %	1		0.00
Transcription factor SOX-17	7.49 %	1		0.00
Complement factor B	7.46 %	4	67.34	9.80
Alpha-2-HS-glycoprotein	7.36 %	2	60.27	9.43
Proteasome subunit alpha type-6	7.32 %	1	20.17	2.10
Alpha- and gamma-adaptin-binding protein p34	7.30 %	1	26.83	3.08
Exportin-2	7.21 %	3	26.14	2.10
Nucleosome assembly protein 1-like 4	7.20 %	1	39.02	3.56
T-complex protein 1 subunit zeta	7.16 %	2	41.57	5.23
2-hydroxyacyl-CoA lyase 1	7.12 %	1		2.03
Isoform 4 of Apoptosis-inducing factor 1, mitochondrial	7.10 %	1		0.00
Ig gamma-2 chain C region	7.06 %	1	53.32	4.00
Serpin H1	6.94 %	1	23.58	2.33
Endoplasmin	6.85 %	2	23.50	3.20
T-complex protein 1 subunit eta	6.81 %	2	27.86	2.65
Adenylyl cyclase-associated protein 1	6.74 %	2	65.10	6.86
X-ray repair cross-complementing protein 5	6.69 %	2	109.18	20.55
Tubulin-specific chaperone D	6.64 %	1		0.00
Acetyl-CoA acetyltransferase, cytosolic	6.55 %	1	28.67	2.28
Phosphoribosylformylglycinamidine synthase	6.29 %	2		0.00
Tropomodulin-3	6.25 %	1	20.44	2.12
Importin subunit alpha-1	6.24 %	1	36.69	5.35
Ig mu chain C region	6.19 %	3	46.65	1.84
Haptoglobin	6.16 %	2	63.28	4.89
Importin subunit beta-1	6.16 %	2	33.14	2.04
cAMP-dependent protein kinase catalytic subunit alpha	5.98 %	1	24.04	0.00

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Tetranectin	5.94 %	1	94.71	3.19
Complement component C8 beta chain	5.75 %	2	35.32	6.22
Apoptosis inhibitor 5	5.53 %	1	24.85	2.06
Vitamin D-binding protein	5.49 %	2	38.55	7.62
Trifunctional enzyme subunit beta, mitochondrial	5.49 %	1	35.91	5.78
Copine-3	5.40 %	1	34.09	2.97
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	5.39 %	1	41.11	2.62
Elongation factor 2	5.36 %	2	116.88	16.59
Mitogen-activated protein kinase 3	5.01 %	1	29.38	2.90
General transcription factor 3C polypeptide 5	5.01 %	1	0.00	2.13
Serum amyloid P-component	4.93 %	1		2.16
Leucine-rich alpha-2-glycoprotein	4.85 %	1		0.00
26S proteasome non-ATPase regulatory subunit 2	4.83 %	1		0.00
Eukaryotic peptide chain release factor subunit 1	4.81 %	1	121.07	13.63
Eukaryotic translation initiation factor 3 subunit M	4.81 %	1	27.27	2.25
Monofunctional C1-tetrahydrofolate synthase, mitochondrial	4.70 %	2	53.82	4.12
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	4.61 %	1	22.93	5.46
Hypoxanthine-guanine phosphoribosyltransferase	4.59 %	1		1.99
ATP synthase subunit beta, mitochondrial	4.54 %	1	72.26	2.47
Splicing factor 3A subunit 3	4.46 %	1		0.00
Alanine--tRNA ligase, cytoplasmic	4.24 %	2	33.92	7.98
Eukaryotic initiation factor 4A-I	4.19 %	1	288.73	39.25
Transketolase	4.17 %	1	27.09	0.00
Fatty acid synthase	4.10 %	5	53.23	8.84
T-complex protein 1 subunit epsilon	4.02 %	1		0.00
Glucose-6-phosphate isomerase	3.94 %	1		0.00
SUMO-activating enzyme subunit 2	3.91 %	1		2.78
Phosphoglycerate kinase 1	3.84 %	1	32.25	2.87
Squalene monooxygenase	3.83 %	1	23.56	5.09
Cystathionine beta-synthase	3.81 %	1	34.85	4.57
Dihydrolipoyl dehydrogenase, mitochondrial	3.73 %	1	63.18	11.55
Heat shock 70 kDa protein 6	3.73 %	1	49.56	7.50
Puromycin-sensitive aminopeptidase	3.67 %	1		2.10
ATP-citrate synthase	3.63 %	2	118.81	16.62
Heterogeneous nuclear ribonucleoprotein K	3.46 %	1	47.22	2.35
Plasma protease C1 inhibitor	3.40 %	1	32.42	0.00
DNA replication licensing factor MCM6	3.29 %	1	26.07	2.09
Elongation factor 1-alpha 1	3.25 %	1	23.23	2.66
Alpha-enolase	3.23 %	1	82.39	3.74
Citrate synthase, mitochondrial	3.22 %	1	44.25	3.88
Epidermal growth factor receptor kinase substrate 8	3.20 %	1		0.00
Complement component C8 beta chain	5.75 %	1	42.75	2.88
Catalase	3.04%	1	26.84	0.00
Pyruvate kinase PKM	3.01 %	1	130.33	8.78
tRNA-splicing ligase RtcB	2.97 %	1		2.73
Probable phospholipid-transporting ATPase VD	2.88 %	1		0.00
Peroxisomal multifunctional enzyme type 2	2.85 %	1	70.46	5.07
Ubiquitin carboxyl-terminal hydrolase 5	2.80 %	1	24.93	4.98
C-1-tetrahydrofolate synthase, cytoplasmic	2.78 %	1	44.08	10.64
78 kDa glucose-regulated protein	2.75 %	1	89.83	4.31
Importin-9	2.69 %	1		0.00
Alpha-1-antichymotrypsin	2.60 %	1	55.35	3.50
Exportin-T	2.60 %	1		2.55
Insulin-like growth factor-binding protein complex acid labile subunit	2.48 %	1	55.87	4.09
Neutral alpha-glucosidase AB	2.44 %	1	57.91	3.53
Microtubule-associated protein 4	2.34 %	1	20.13	5.50
Complement C1s subcomponent	2.33 %	1	47.04	9.58
Afamin	2.17 %	1		2.22
Coatomer subunit beta	2.10 %	1	97.14	6.95
Filamin-B	1.96 %	2	33.23	7.62

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
Protein KIAA0368	1.90 %	1		0.00
N-alpha-acetyltransferase 15, NatA auxiliary subunit	1.85 %	1	25.05	5.97
Importin-5	1.73 %	1	79.90	9.58
Apolipoprotein B-100	1.71 %	5	176.50	18.15
Eukaryotic translation initiation factor 5B	1.64 %	1	78.59	5.50
Band 3 anion transport protein	1.43 %	1	49.60	3.05
Kininogen-1	1.40 %	1	33.98	0.00
DNA-dependent protein kinase catalytic subunit	1.17 %	2		4.51
Filamin-A	1.02 %	1	23.70	2.59
E3 ubiquitin-protein ligase HUWE1 (Fragment)	0.56 %	1		4.10
Kinesin-like protein KIF1C	0.54 %	1		0.00
Cytoplasmic dynein 1 heavy chain 1	0.50 %	1		0.00
Spectrin beta chain, erythrocytic	0.47 %	1	28.21	0.00
Heterogeneous nuclear ribonucleoprotein U (Fragment)	9.92 %	1		5.69
T-complex protein 1 subunit alpha	9.89 %	2	116.36	6.30
HBS1-like protein (Fragment)	9.88 %	1		0.00
ADP/ATP translocase 2	9.73 %	1		2.09
Inter-alpha-trypsin inhibitor heavy chain H4	9.68 %	7	159.84	22.13
Inter-alpha-trypsin inhibitor heavy chain H1	9.44 %	5	54.35	12.42
Gelsolin	9.34 %	4	109.25	15.69
Coatomer subunit epsilon	9.09 %	1	22.82	2.07
Methylosome protein 50 (Fragment)	8.93 %	1		0.00
Fructose-2,6-bisphosphatase TIGAR	8.52 %	1	0.00	
Histidine-rich glycoprotein	8.38 %	3	29.93	2.03
Fructose-bisphosphate aldolase A	8.24 %	1	27.43	2.55
Complement C4-A	8.03 %	8	238.00	
Serotransferrin	8.02 %	5	96.14	13.80
Tubulin alpha-4A chain	7.81 %	2	61.83	2.83
Alcohol dehydrogenase class-3	7.49 %	1		0.00
Transcription factor SOX-17	7.49 %	1		0.00
Complement factor B	7.46 %	4	67.34	9.80
Alpha-2-HS-glycoprotein	7.36 %	2	60.27	9.43
Proteasome subunit alpha type-6	7.32 %	1	20.17	2.10
Alpha- and gamma-adaptin-binding protein p34	7.30 %	1	26.83	3.08
Exportin-2	7.21 %	3	26.14	2.10
Nucleosome assembly protein 1-like 4	7.20 %	1	39.02	3.56
T-complex protein 1 subunit zeta	7.16 %	2	41.57	5.23
2-hydroxyacyl-CoA lyase 1	7.12 %	1		2.03
Isoform 4 of Apoptosis-inducing factor 1, mitochondrial	7.10 %	1		0.00
Ig gamma-2 chain C region	7.06 %	1	53.32	4.00
Serpin H1	6.94 %	1	23.58	2.33
Endoplasmin	6.85 %	2	23.50	3.20
T-complex protein 1 subunit eta	6.81 %	2	27.86	2.65
Adenylyl cyclase-associated protein 1	6.74 %	2	65.10	6.86
X-ray repair cross-complementing protein 5	6.69 %	2	109.18	20.55
Complement component C8 beta chain	5.75 %	2	35.32	6.22
Apoptosis inhibitor 5	5.53 %	1	24.85	2.06
Vitamin D-binding protein	5.49 %	2	38.55	7.62
Trifunctional enzyme subunit beta, mitochondrial	5.49 %	1	35.91	5.78
Copine-3	5.40 %	1	34.09	2.97
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	5.39 %	1	41.11	2.62
Elongation factor 2	5.36 %	2	116.88	16.59
Mitogen-activated protein kinase 3	5.01 %	1	29.38	2.90
General transcription factor 3C polypeptide 5	5.01 %	1	0.00	2.13
Serum amyloid P-component	4.93 %	1		2.16
Leucine-rich alpha-2-glycoprotein	4.85 %	1		0.00
26S proteasome non-ATPase regulatory subunit 2	4.83 %	1		0.00
Eukaryotic peptide chain release factor subunit 1	4.81 %	1	121.07	13.63
Eukaryotic translation initiation factor 3 subunit M	4.81 %	1	27.27	2.25
Monofunctional C1-tetrahydrofolate synthase, mitochondrial	4.70 %	2	53.82	4.12
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	4.61 %	1	22.93	5.46
Hypoxanthine-guanine phosphoribosyltransferase	4.59 %	1		1.99

Protein Description	Total Coverage	Unique peptides	Mascot Score	Sequest Score
ATP synthase subunit beta, mitochondrial	4.54 %	1	72.26	2.47
Splicing factor 3A subunit 3	4.46 %	1		0.00
Alanine--tRNA ligase, cytoplasmic	4.24 %	2	33.92	7.98
Eukaryotic initiation factor 4A-I	4.19 %	1	288.73	39.25
Transketolase	4.17 %	1	27.09	0.00
Fatty acid synthase	4.10 %	5	53.23	8.84
T-complex protein 1 subunit epsilon	4.02 %	1		0.00
Glucose-6-phosphate isomerase	3.94 %	1		0.00
SUMO-activating enzyme subunit 2	3.91 %	1		2.78
Phosphoglycerate kinase 1	3.84 %	1	32.25	2.87
Squalene monooxygenase	3.83 %	1	23.56	5.09
Cystathionine beta-synthase	3.81 %	1	34.85	4.57
Dihydrolipoyl dehydrogenase, mitochondrial	3.73 %	1	63.18	11.55
Heat shock 70 kDa protein 6	3.73 %	1	49.56	7.50
Puromycin-sensitive aminopeptidase	3.67 %	1		2.10
ATP-citrate synthase	3.63 %	2	118.81	16.62
Heterogeneous nuclear ribonucleoprotein K	3.46 %	1	47.22	2.35
Plasma protease C1 inhibitor	3.40 %	1	32.42	0.00
DNA replication licensing factor MCM6	3.29 %	1	26.07	2.09
Elongation factor 1-alpha 1	3.25 %	1	23.23	2.66
Alpha-enolase	3.23 %	1	82.39	3.74
Citrate synthase, mitochondrial	3.22 %	1	44.25	3.88
Epidermal growth factor receptor kinase substrate 8	3.20 %	1		0.00
Complement component C8 beta chain	5.75 %	1	42.75	2.88
Catalase	3.04%	1	26.84	0.00
Pyruvate kinase PKM	3.01 %	1	130.33	8.78
tRNA-splicing ligase RtcB	2.97 %	1		2.73
Probable phospholipid-transporting ATPase VD	2.88 %	1		0.00
Peroxisomal multifunctional enzyme type 2	2.85 %	1	70.46	5.07
Ubiquitin carboxyl-terminal hydrolase 5	2.80 %	1	24.93	4.98
C-1-tetrahydrofolate synthase, cytoplasmic	2.78 %	1	44.08	10.64
78 kDa glucose-regulated protein	2.75 %	1	89.83	4.31
Importin-9	2.69 %	1		0.00
Alpha-1-antichymotrypsin	2.60 %	1	55.35	3.50
Exportin-T	2.60 %	1		2.55
Insulin-like growth factor-binding protein complex acid labile subunit	2.48 %	1	55.87	4.09
Neutral alpha-glucosidase AB	2.44 %	1	57.91	3.53
Microtubule-associated protein 4	2.34 %	1	20.13	5.50
Complement C1s subcomponent	2.33 %	1	47.04	9.58
Afamin	2.17 %	1		2.22
Coatomer subunit beta	2.10 %	1	97.14	6.95
General vesicular transport factor p115	2.08 %	1	22.12	4.63
Filamin-B	1.96 %	2	33.23	7.62
Protein KIAA0368	1.90 %	1		0.00
N-alpha-acetyltransferase 15, NatA auxiliary subunit	1.85 %	1	25.05	5.97
Importin-5	1.73 %	1	79.90	9.58
Apolipoprotein B-100	1.71 %	5	176.50	18.15
Eukaryotic translation initiation factor 5B	1.64 %	1	78.59	5.50
Band 3 anion transport protein	1.43 %	1	49.60	3.05
Kininogen-1	1.40 %	1	33.98	0.00
DNA-dependent protein kinase catalytic subunit	1.17 %	2		4.51
Filamin-A	1.02 %	1	23.70	2.59
E3 ubiquitin-protein ligase HUWE1 (Fragment)	0.56 %	1		4.10
Kinesin-like protein KIF1C	0.54 %	1		0.00
Cytoplasmic dynein 1 heavy chain 1	0.50 %	1		0.00
Spectrin beta chain, erythrocytic	0.47 %	1	28.21	0.00

Appendix 4.4 Proteins identified from LC MS/MS analysis of surface sampling followed by proteolysis

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	Y	N	N
Actin, cytoplasmic 1	Y	Y	Y
Actin, gamma-enteric smooth muscle	N	Y	N
Adenylate kinase isoenzyme 1	Y	N	N
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	N	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-enolase	Y	Y	N
Alpha-synuclein	Y	Y	N
Angiotensinogen	Y	N	N
Annexin A2	N	Y	Y
Antithrombin-III	N	Y	Y
AP-2 complex subunit beta	N	Y	N
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	N	Y
Apolipoprotein B-100	Y	Y	N
Apolipoprotein C-I	Y	N	Y
Apolipoprotein C-II	Y	N	Y
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	N	Y	Y
Apolipoprotein E	Y	N	N
Apolipoprotein-L1	Y	N	N
Arginase-1	Y	Y	N
Astrotactin-1	N	N	Y
Bisphosphoglycerate mutase	Y	N	Y
Calmodulin-like protein 5	N	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Cathepsin B	N	Y	N
Ceruloplasmin	Y	Y	Y
Clusterin	Y	Y	Y
Cofilin-1	Y	N	N
Coiled-coil domain-containing protein 27	N	N	Y
Complement C1s subcomponent	Y	N	N
Complement C3	Y	Y	Y
Complement C4-A	N	Y	N
Complement C4-B	Y	N	Y
Complement component C9	Y	N	N
Complement factor B	Y	Y	Y
Complement factor H	Y	N	N
Cystatin-A	N	Y	Y
Delta-aminolevulinic acid dehydratase	Y	N	N
Dermcidin	Y	Y	Y
Desmoglein-1	N	Y	Y
Desmoplakin	N	Y	N
Extracellular glycoprotein lacritin	N	N	Y
Fatty acid-binding protein, epidermal	N	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Fatty acid-binding protein, liver	N	Y	N
Fibrinogen alpha chain	Y	Y	Y
Fibrinogen beta chain	Y	Y	Y
Fibrinogen gamma chain	Y	Y	Y
Fibroblast growth factor 2	Y	Y	Y
Filaggrin-2	N	Y	N
Flavin reductase	Y	Y	Y
FRAS1-related extracellular matrix protein 2	N	Y	N
Fructose-bisphosphate aldolase A	Y	N	Y
Gelsolin	Y	N	Y
Glutathione S-transferase A1	Y	Y	N
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
GTP-binding nuclear protein Ran	Y	N	Y
Haptoglobin	Y	Y	Y
Heat shock 70 kDa protein 1A/1B	N	Y	N
Heat shock protein beta-1	Y	Y	N
Heat shock-related 70 kDa protein 2	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Heparin cofactor 2	Y	N	Y
Histidine-rich glycoprotein	Y	Y	Y
Histone H2A type 1-H	Y	Y	Y
Hornerin	N	N	Y
Ig alpha-1 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	N	Y
Ig gamma-3 chain C region	Y	Y	N
Ig heavy chain V-III region TEI	N	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region Lay	Y	Y	Y
Ig kappa chain V-III region SIE	Y	Y	Y
Ig lambda chain V-I region WAH	N	N	Y
Ig lambda chain V-III region LOI	Y	N	N
Ig lambda-7 chain C region	Y	Y	Y
Ig mu heavy chain disease protein	Y	N	N
Inter-alpha-trypsin inhibitor heavy chain H1	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H2	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H4	Y	N	Y
Junction plakoglobin	N	Y	N
Kallikrein-7	N	Y	N
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 13	Y	N	N
Keratin, type I cytoskeletal 14	Y	Y	Y
Keratin, type I cytoskeletal 16	N	Y	N
Keratin, type I cytoskeletal 17	N	Y	Y
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 2 oral	Y	Y	N
Keratin, type II cytoskeletal 5	Y	Y	Y
Keratin, type II cytoskeletal 6A	Y	Y	Y
Keratin, type II cytoskeletal 6C	N	Y	Y
Kininogen-1	Y	Y	Y
Lactotransferrin	Y	Y	Y
Lethal(3)malignant brain tumor-like protein 2	N	Y	N
Lipocalin-1	N	Y	Y
Liver carboxylesterase 1	Y	Y	Y
Lumican	Y	N	N
Lysozyme C	Y	Y	Y
Mammaglobin-B	N	N	Y
Mucin-like protein 1	Y	Y	Y
N-acetylmuramoyl-L-alanine amidase	N	Y	N
Nesprin-1	N	Y	N
Nicotinate phosphoribosyltransferase	N	Y	Y
Nucleoside diphosphate kinase A	Y	N	N
Peptidyl-prolyl cis-trans isomerase A	Y	N	N
Peroxiredoxin-1	Y	N	N
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	N
Phosphoglycerate kinase 1	Y	Y	N
Plasma kallikrein	Y	N	N
Plasma protease C1 inhibitor	Y	Y	Y
Plasminogen	Y	Y	Y
Plastin-2	N	Y	N
Polymeric immunoglobulin receptor	N	N	Y
Polyubiquitin-C	Y	Y	Y
Prolactin-inducible protein	N	Y	Y
Proline-rich protein 4	N	N	Y
Protein disulfide-isomerase A2	Y	N	N
Protein DJ-1	Y	Y	N
Protein S100-A4	N	N	Y
Protein S100-A6 OS=Homo sapiens	Y	N	N
Protein S100-A7	N	Y	Y
Protein S100-A8	Y	Y	Y
Protein S100-A9	Y	Y	Y
Prothrombin	Y	Y	Y
Purine nucleoside phosphorylase	Y	N	N
Purine nucleoside phosphorylase	N	N	Y
Putative lipocalin 1-like protein 1	N	Y	N
Rab GDP dissociation inhibitor beta OS=Homo sapiens	Y	N	N
Rab-related protein Rab-14	Y	N	N
Retinal dehydrogenase 1	Y	N	N
Secretoglobin family 1D member 2	N	Y	N
Semenogelin-1	Y	Y	Y
Semenogelin-2	Y	Y	N
Serotransferrin	Y	Y	Y
Serpin B3	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Serum albumin	Y	Y	Y
Serum amyloid A-4 protein	Y	N	N
Serum amyloid P-component	N	N	Y
Serum paraoxonase/arylesterase 1	Y	N	Y
Sialic acid-binding Ig-like lectin 16	N	Y	N
Testis-specific serine/threonine-protein kinase 4	N	Y	Y
Thioredoxin	Y	N	N
Transketolase	N	Y	N
Transthyretin	N	Y	Y
Trypsin-1	Y	Y	N
Uncharacterized protein UNQ773/PRO1567	N	Y	N
Vitamin D-binding protein	Y	Y	Y
Vitronectin	Y	Y	Y
Zinc finger protein 611	Y	N	N
Zinc-alpha-2-glycoprotein	Y	Y	Y
Zymogen granule protein 16 homolog B	Y	N	N

Appendix 4.5 Proteins identified from LC MS/MS run of automated surface sampling based digestion of dried blood spot.

Appendix 5

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	Y	N	N
60 kDa heat shock protein, mitochondrial	Y	N	N
Actin, cytoplasmic 1	Y	Y	Y
ADP-ribosylation factor 5	N	N	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	N	Y	N
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	N	Y	N
Alpha-2-macroglobulin	Y	Y	Y
Alpha-enolase	N	Y	N
Alpha-synuclein	Y	Y	Y
Angiotensinogen	N	Y	N
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein B-100	Y	N	Y
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	Y	Y	Y
Band 3 anion transport protein	Y	Y	Y
Beta-actin-like protein 2	N	Y	N
Bisphosphoglycerate mutase	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	N	Y
Clusterin	N	Y	N
Complement C3	Y	Y	Y
Complement C4-A	Y	N	N
Complement factor B	Y	N	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Flavin reductase	Y	Y	Y
Fructose-bisphosphate aldolase A	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
GTP-binding nuclear protein Ran	Y	Y	Y
Haptoglobin	Y	Y	Y
Heat shock-related 70 kDa protein 2	Y	N	N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Heparin cofactor 2	N	Y	N
Hermansky-Pudlak syndrome 1 protein	N	Y	N
Histidine-rich glycoprotein	Y	N	N
Ig alpha-1 chain C region	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Ig alpha-2 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig heavy chain V-III region GAL	Y	Y	Y
Ig heavy chain V-III region TEI	N	Y	Y
Ig kappa chain V-I region Lay	N	Y	N
Ig kappa chain V-III region B6	N	Y	N
Ig kappa chain V-III region SIE	Y	N	Y
Ig kappa chain V-III region VG (Fragment)	Y	Y	Y
Ig kappa chain V-IV region Len	N	N	Y
Ig lambda chain V-I region WAH	Y	Y	Y
Ig lambda chain V-III region LOI	N	Y	Y
Ig lambda-1 chain C regions	N	Y	Y
Ig lambda-7 chain C region	Y	Y	Y
Ig mu heavy chain disease protein	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H1	N	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H2	N	Y	N
Inter-alpha-trypsin inhibitor heavy chain H4	N	N	Y
Keratin, type II cytoskeletal 1	N	N	Y
Kininogen-1	Y	Y	Y
L-lactate dehydrogenase B chain	N	Y	Y
Mucin-16	Y	N	N
Nuclear transport factor 2	N	Y	N
Nucleoside diphosphate kinase A	Y	N	N
Peroxiredoxin-1	N	Y	Y
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	Y
Phosphoglycerate kinase 1	Y	N	N
Plasma protease C1 inhibitor	Y	Y	Y
Polyubiquitin-C	Y	Y	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protein S100-A8	N	Y	Y
Prothrombin	N	N	Y
Purine nucleoside phosphorylase	Y	Y	N
Putative uncharacterized protein C19orf81	N	Y	N
Retinal dehydrogenase 1	Y	Y	N
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Spectrin alpha chain, erythrocyte	Y	Y	Y
Spectrin alpha chain, erythrocytic 1	N	Y	N
Superoxide dismutase [Cu-Zn]	Y	Y	Y
Thioredoxin	Y	N	N
Transthyretin	Y	Y	Y
Triosephosphate isomerase	N	Y	Y
UDP-glucuronosyltransferase 1-9	Y	N	N
Vitamin D-binding protein	Y	Y	Y
Zinc-alpha-2-glycoprotein	Y	N	N

Appendix 5.1 Proteins identified from LC MS/MS analysis of a DBS digest with a digestion time of 30 minutes

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	Y	Y	Y
60 kDa heat shock protein, mitochondrial	Y	Y	N
Actin, cytoplasmic 1	Y	Y	Y
ADP-ribosylation factor 5	N	N	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	N	N
Alpha-2-macroglobulin	Y	Y	Y
Alpha-enolase	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Angiotensinogen	Y	N	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein B-100	Y	N	Y
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	Y	Y	Y
Band 3 anion transport protein	Y	Y	Y
Beta-actin-like protein 2	N	N	Y
Bisphosphoglycerate mutase	Y	Y	Y
Calcium-dependent secretion activator 2	Y	N	N
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	Y	Y
Clusterin	Y	Y	Y
Complement C3	Y	Y	Y
Complement C4-A	Y	Y	N
Complement factor B	N	N	Y
D-dopachrome decarboxylase	N	Y	N
Delta-aminolevulinic acid dehydratase	Y	Y	Y
DNA mismatch repair protein Mlh1	N	N	Y
Ecto-ADP-ribosyltransferase 5	N	N	Y
Flavin reductase	Y	Y	Y
Fructose-bisphosphate aldolase A	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
GTP-binding nuclear protein Ran	Y	Y	Y
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Heparin cofactor 2	N	Y	Y
High affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8A	N	Y	N
Histidine-rich glycoprotein	Y	Y	N
Ig alpha-1 chain C region	Y	Y	Y
Ig alpha-2 chain C region	Y	Y	Y
Ig heavy chain V-III region	Y	N	N
Ig heavy chain V-III region GAL	Y	Y	Y
Ig heavy chain V-III region VH26	Y	N	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Heparin cofactor 2	N	Y	Y
High affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8A	N	Y	N
Histidine-rich glycoprotein	Y	Y	N
Ig alpha-1 chain C region	Y	Y	Y
Ig alpha-2 chain C region	Y	Y	Y
Ig heavy chain V-III region	Y	N	N
Ig heavy chain V-III region GAL	Y	Y	Y
Ig heavy chain V-III region VH26	Y	N	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region	Y	N	N
Ig kappa chain V-I region Lay	N	N	Y
Ig kappa chain V-III region SIE	Y	Y	Y
Ig kappa chain V-III region VG (Fragment)	Y	Y	Y
Ig kappa chain V-IV region Len	N	N	Y
Ig lambda chain V-I region WAH	N	Y	Y
Ig lambda chain V-III region LOI	Y	Y	Y
Ig lambda-1 chain C regions	N	N	Y
Ig lambda-7 chain C region	Y	Y	Y
Ig mu heavy chain disease protein	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H1	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H2	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H4	Y	Y	Y
Interleukin-7 receptor subunit alpha	Y	Y	N
Keratin, type II cytoskeletal 1	N	N	Y
Kininogen-1	Y	Y	Y
L-lactate dehydrogenase B chain	Y	Y	Y
Malate dehydrogenase, cytoplasmic	N	Y	N
Nucleoside diphosphate kinase A	Y	Y	N
Peptidyl-prolyl cis-trans isomerase A	N	Y	N
Peroxiredoxin-1	N	Y	Y
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	Y
Phosphoglycerate kinase 1 O	Y	Y	N
Plasma protease C1 inhibitor	Y	Y	Y
Polycystin-2	N	N	Y
Polyubiquitin-C	Y	Y	Y
Proteasome activator complex subunit 1	N	N	Y
Protein S100-A4	Y	Y	N
Protein S100-A6	Y	Y	N
Protein S100-A8	N	Y	Y
Protein S100-A9	N	Y	N
Purine nucleoside phosphorylase	Y	Y	Y
Putative protein FAM10A4	N	N	Y
Putative uncharacterized protein C19orf81	N	N	Y
Rab GDP dissociation inhibitor beta	N	Y	N
Retinal dehydrogenase 1	Y	Y	Y
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum amyloid P-component	N	Y	Y
Spectrin alpha chain, erythrocyte	Y	N	N
Spectrin alpha chain, erythrocytic 1	N	Y	Y
Spectrin beta chain, erythrocytic	Y	Y	Y
StAR-related lipid transfer protein 3	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Superoxide dismutase	Y	Y	Y
Thioredoxin	Y	N	Y
Transthyretin	Y	Y	Y
Triosephosphate isomerase	N	N	Y
Ubiquitin conjugation factor E4 A	N	N	Y
Vitamin D-binding protein	Y	N	Y

Appendix 5.2 Proteins identified from LC MS/MS analysis of a DBS digest with a digestion time of 1 hour.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	Y	Y	Y
60 kDa heat shock protein, mitochondrial	Y	N	N
Actin, alpha cardiac muscle 1	Y	Y	N
Actin, cytoplasmic 1	Y	Y	Y
Agmatinase, mitochondrial	Y	N	N
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	N
Alpha-2-macroglobulin	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Ankyrin-1	Y	N	N
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	Y	Y	Y
Band 3 anion transport protein	Y	Y	N
Beta-actin-like protein 2	N	Y	N
Bisphosphoglycerate mutase	Y	Y	Y
Carbohydrate sulfotransferase 8	Y	N	N
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	Y	Y
Clusterin	Y	Y	N
Complement C3	Y	Y	Y
Complement C4-A	Y	N	N
Complement C4-B	N	Y	Y
Complement factor B	N	Y	N
Complement factor H	N	N	Y
D-dopachrome decarboxylase	N	Y	N
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Fibulin-2	Y	N	N
Flavin reductase	Y	Y	Y
Fructose-bisphosphate aldolase A	Y	Y	Y
Gamma-enolase	Y	N	N
General transcription factor 3C polypeptide 1	N	Y	N
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
GTP-binding nuclear protein Ran	N	Y	Y
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Heparin cofactor 2	Y	N	N
Hermansky-Pudlak syndrome 1 protein	N	N	Y
Histidine-rich glycoprotein	N	N	Y
Ig heavy chain V-III region TEI	N	N	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region CAR	N	N	Y
Ig kappa chain V-I region DEE	Y	Y	N
Ig kappa chain V-III region SIE	Y	Y	Y
Ig kappa chain V-III region VG (Fragment)	N	Y	Y
Ig lambda chain V-I region WAH	Y	N	Y
Ig lambda chain V-III region LOI	N	N	Y
Ig lambda-7 chain C region	Y	Y	Y
Ig mu chain C region	Y	Y	N
Ig mu heavy chain disease protein	N	N	Y
Inter-alpha-trypsin inhibitor heavy chain H4	N	N	Y
Keratin, type I cytoskeletal 10	Y	N	Y
Keratin, type I cytoskeletal 9	N	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Kininogen-1	Y	Y	Y
Leucine-rich repeat neuronal protein 2	Y	N	N
L-lactate dehydrogenase B chain	N	Y	Y
Malate dehydrogenase, cytoplasmic	N	Y	N
Mitogen-activated protein kinase-binding protein 1	N	Y	N
Nucleoside diphosphate kinase A	Y	Y	N
Paired box protein Pax-1	Y	N	N
Peptidyl-prolyl cis-trans isomerase A	Y	Y	N
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-1	N	Y	Y
Peroxiredoxin-6	Y	Y	Y
Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform	Y	N	N
Phosphoglycerate kinase 1	Y	Y	Y
Plasma protease C1 inhibitor	Y	Y	Y
Polyubiquitin-C	Y	Y	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protein S100-A8	Y	Y	Y
Protein S100-A9	N	N	Y
Protein TANC2	N	N	Y
Purine nucleoside phosphorylase	Y	Y	Y
Putative uncharacterized protein C19orf81	N	Y	Y
Retinal dehydrogenase 1	N	Y	Y
Rotatin	Y	N	N
Sal-like protein 3	Y	N	N
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum amyloid P-component	Y	Y	Y
Solute carrier organic anion transporter family member 1A2	N	N	Y
Spectrin alpha chain, erythrocyte	Y	Y	Y
Spectrin beta chain, erythrocyte	N	Y	Y
Superoxide dismutase	Y	Y	Y
Thiamin pyrophosphokinase 1	Y	N	N
Thioredoxin	Y	N	N
Transthyretin	Y	Y	Y
Triosephosphate isomerase	N	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Ubiquitin-conjugating enzyme E2 variant 2	N	N	Y
Uncharacterized protein C6orf163	N	Y	N
Vitamin D-binding protein	Y	Y	Y
Vitronectin	N	N	Y
Zinc-alpha-2-glycoprotein	Y	N	N

Appendix 5.3 Proteins identified from LC MS/MS analysis of a DBS digest with a digestion time of 2 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	N	Y	N
Actin, cytoplasmic 1	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-synuclein	Y	Y	N
Ankyrin-1	N	Y	N
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	N	Y	Y
Bisphosphoglycerate mutase	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	Y	N
Clusterin	N	Y	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Complement C3	Y	Y	Y
Complement C4-B	N	Y	N
Complement factor B	Y	N	N
Delta-aminolevulinic acid dehydratase	Y	Y	N
Ensconsin	N	Y	N
Exportin-1	Y	Y	N
Flavin reductase	Y	Y	Y
Fructose-bisphosphate aldolase A	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
GTP-binding nuclear protein Ran	Y	Y	Y
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Ig alpha-1 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region CAR	N	Y	N
Ig kappa chain V-I region DEE	Y	Y	Y
Ig kappa chain V-III region SIE	Y	Y	Y
Ig kappa chain V-III region VG (Fragment)	N	Y	Y
Ig lambda chain V-I region WAH	N	Y	N
Ig lambda chain V-III region LOI	N	Y	N
Ig lambda-7 chain C region	Y	Y	Y
Ig mu chain C region	Y	Y	N
Immunity-related GTPase family Q protein	Y	Y	N
Inter-alpha-trypsin inhibitor heavy chain H2	N	Y	N
Interleukin-7 receptor subunit alpha	N	Y	N
Katanin p60 ATPase-containing subunit A1	Y	Y	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Keratin, type I cytoskeletal 10	Y	N	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Kininogen-1	Y	Y	Y
L-amino-acid oxidase	Y	Y	N
Methyltransferase-like protein 7B	Y	Y	N
N-glycosylase/DNA lyase	Y	Y	N
Nucleoside diphosphate kinase A	Y	Y	N
Paired box protein Pax-1	Y	Y	N
Peptidyl-prolyl cis-trans isomerase A	N	Y	N
Peroxiredoxin-1	N	Y	N
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	N
Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform	N	Y	N
Polyubiquitin-C	Y	Y	Y
Protein S100-A4	Y	Y	N
Protein S100-A9	N	Y	N
Protein unc-80 homolog	N	Y	N
Purine nucleoside phosphorylase	N	Y	N
Putative sodium-coupled neutral amino acid transporter 10	Y	Y	N
Putative uncharacterized protein C19orf81	N	N	Y
Receptor-type tyrosine-protein phosphatase F	Y	N	N
Retinal dehydrogenase 1	N	Y	N
Selenium-binding protein 1	N	Y	N
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Spectrin alpha chain, erythrocyte	Y	Y	N
Spectrin alpha chain, erythrocytic 1	N	N	Y
Spectrin beta chain, erythrocytic	N	Y	N
Superoxide dismutase [Cu-Zn]	Y	Y	Y
Tetratricopeptide repeat protein 40	Y	N	N
TGF-beta-activated kinase 1 and MAP3K7-binding protein 3	Y	N	N
Thioredoxin	Y	Y	Y
Transient receptor potential cation channel subfamily M member 3	Y	N	N
Transthyretin	Y	Y	Y
Vitamin D-binding protein	N	Y	N
Vitronectin	N	Y	N
Zinc finger protein 470	N	Y	N
Zinc finger protein 638	N	Y	Y

Appendix 5.4 Proteins identified from LC MS/MS analysis of a DBS digest with a digestion time of 4 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
60 kDa heat shock protein, mitochondrial	Y	N	N
Actin, cytoplasmic 1	Y	Y	Y
Alpha-1-acid glycoprotein 1	N	Y	Y
Alpha-1-acid glycoprotein 2	N	Y	N
Alpha-1-antitrypsin	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein C-III	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
CD166 antigen	Y	N	N
Complement C3	Y	Y	Y
Complement factor B	N	N	Y
Disrupted in schizophrenia 1 protein	Y	N	N
EF-hand domain-containing family member A1	Y	N	N
Flavin reductase	Y	Y	Y
Fructose-bisphosphate aldolase A	Y	N	Y
Fructose-bisphosphate aldolase B	Y	N	N
Glutathione reductase, mitochondrial	Y	N	N
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Haptoglobin	Y	Y	Y
HEAT repeat-containing protein 1	Y	N	N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	N
Ig alpha-1 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region DEE	Y	N	Y
Ig lambda-7 chain C region	Y	Y	Y
Ig mu chain C region	Y	N	Y
Long-chain fatty acid transport protein 6	N	N	Y
N-glycosylase/DNA lyase	Y	Y	N
Paired box protein Pax-1	N	Y	N
Peroxiredoxin-1	Y	N	N
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	N
Polyubiquitin-C	Y	Y	Y
Protein RIC1 homolog	N	Y	Y
Purine nucleoside phosphorylase	Y	N	N
Retrotransposon-like protein 1	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Spectrin alpha chain, non-erythrocytic 1	Y	N	N
Superoxide dismutase [Cu-Zn]	Y	Y	Y
Thioredoxin	Y	N	N
Transthyretin	Y	Y	Y
Uncharacterized protein KIAA1958	Y	N	N
Voltage-dependent L-type calcium channel subunit alpha-1	N	N	Y

Appendix 5.5 Proteins identified from LC MS/MS analysis of a DBS digest with a digestion time of 8 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein theta	N	N	Y
Actin, cytoplasmic 1	N	N	Y
Alpha-1-acid glycoprotein 1	Y	N	Y
Alpha-1-antichymotrypsin	Y	Y	N
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	N	Y
Alpha-2-macroglobulin	Y	N	N
Alpha-enolase	Y	Y	N
Alpha-mannosidase 2C1 (Fragment)	N	N	Y
Alpha-synuclein	N	N	Y
Antithrombin-III	N	Y	Y
Apolipoprotein A-I	N	N	Y
Apolipoprotein A-IV	N	N	Y
Apolipoprotein C-III	N	N	Y
Bisphosphoglycerate mutase	N	N	Y
Calmodulin	Y	Y	N
Calmodulin-like protein 5	Y	Y	N
Calpastatin	N	N	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Caspase-14	N	Y	N
Catalase	Y	Y	N
Cell division cycle protein 20 homolog B	N	Y	N
Ceruloplasmin	Y	Y	Y
Cofilin-1	N	N	Y
Complement factor B	Y	N	N
Cyclic AMP-responsive element-binding protein 3-like protein	Y	N	Y
D-dopachrome decarboxylase	N	N	Y
Delta-aminolevulinic acid dehydratase	N	N	Y
Dermcidin	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Elongation factor 1-alpha 1	Y	N	N
Flavin reductase	N	Y	N
Flavin reductase (NADPH)	Y	N	Y
Fructose-bisphosphate aldolase	N	N	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Histidine-rich glycoprotein	N	N	Y
Ig gamma-2 chain C region	Y	N	N
Ig kappa chain C region	Y	N	N
Importin subunit beta-1	N	N	Y
Importin-13	N	N	Y
Isoform 2 of Rab GDP dissociation inhibitor beta	N	N	Y
Keratin, type I cytoskeletal 10	Y	Y	N
Keratin, type I cytoskeletal 13	Y	Y	N
Keratin, type I cytoskeletal 14	Y	Y	N
Keratin, type I cytoskeletal 16	Y	Y	N
Keratin, type I cytoskeletal 9	Y	Y	N
Keratin, type II cytoskeletal 4	Y	Y	N
Keratin, type II cytoskeletal 5	Y	Y	N
Keratin, type II cytoskeletal 6A	Y	Y	N
Keratin, type II cytoskeletal 6B	Y	N	N
L-lactate dehydrogenase B chain	N	N	Y
MORC family CW-type zinc finger protein 1	N	Y	N
Nucleoside diphosphate kinase A (Fragment)	N	N	Y
Olfactory receptor 10K2	N	N	Y
Peptidyl-prolyl cis-trans isomerase A	Y	Y	Y
Peroxiredoxin-2	Y	N	Y
Phosphoglycerate kinase 1	N	N	Y
Phosphoglycerate mutase 2	N	N	Y
Polyubiquitin-C	N	Y	Y
ProSAP-interacting protein 1	Y	N	N
Protein 4.1	N	Y	N
Protein disulfide-isomerase A2	Y	Y	N
Protein DJ-1	N	Y	N
Protein fantom	N	N	Y
Protein S100-A4	N	N	Y
Prothrombin	Y	Y	N
Purine nucleoside phosphorylase	N	N	Y
Putative uncharacterized protein C19orf81	Y	N	N
Retinal dehydrogenase 1	N	N	Y
Retinol-binding protein 4	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Ribose-5-phosphate isomerase	N	N	Y
RNA-binding motif, single-stranded-interacting protein 1 (Fragment)	N	N	Y
Selenium-binding protein 1	Y	Y	N
Serine/threonine-protein kinase LMTK2	Y	N	N
Serotransferrin	Y	N	N
Serum albumin	Y	Y	N
Sickle tail protein homolog	N	Y	N
Spectrin beta chain, erythrocytic	N	N	Y
Superoxide dismutase	Y	Y	Y
Thioredoxin	N	N	Y
Transthyretin	N	Y	N
Triosephosphate isomerase	Y	Y	N
Tropomyosin alpha-3 chain	N	Y	N
Trypsin-3	Y	N	N
Uncharacterized protein C6orf163	N	Y	N
Vitamin D-binding protein	Y	Y	N
Vitronectin	N	N	Y
Zinc finger protein 616	N	N	Y
Zinc-alpha-2-glycoprotein	Y	Y	N

Appendix 5.6 Proteins identified from LC MS/MS analysis of a DBS digest depleted of the top 12 plasma proteins. Sample digested for 1 hour.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
45 kDa calcium-binding protein	Y	N	N
Actin, cytoplasmic 1	Y	N	N
Alpha-1-acid glycoprotein 1	Y	Y	N
Alpha-1-antitrypsin	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	N	N	Y
Apolipoprotein A-IV	N	Y	Y
Apolipoprotein C-III	Y	Y	Y
Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 11	N	N	Y
Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase, mitochondrial	N	N	Y
Bisphosphoglycerate mutase	Y	N	Y
Calpastatin	Y	N	N
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	N	Y
Cofilin-1	N	N	Y
Complement C3	Y	Y	Y
Complement factor B	Y	Y	N
C-type mannose receptor 2	Y	N	N
Cyclic AMP-responsive element-binding protein 3-like protein 4	N	N	Y
Cytochrome P450 1B1	N	N	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
DNA polymerase subunit gamma-2, mitochondrial	N	Y	N
EIF2AK2 protein (Fragment)	N	N	Y
Endothelial protein C receptor	N	N	Y
Euchromatic histone-lysine N-methyltransferase 2	N	N	Y
Flavin reductase (NADPH)	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Golgi integral membrane protein 4	N	N	Y
GS homeobox 1	N	Y	N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemoglobin subunit gamma-2	Y	Y	Y
Hemopexin	Y	Y	Y
Importin subunit beta-1	Y	N	Y
Importin-13	Y	N	N
Inactive caspase-12	N	N	Y
Insulin receptor-related protein	N	N	Y
Inter-alpha-trypsin inhibitor heavy chain H4	N	Y	N
Isoform 2 of Alpha-1B-glycoprotein	N	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Isoform 2 of Coiled-coil domain-containing protein 34	N	N	Y
Isoform 2 of Cytochrome P450 4V2	N	N	Y
Isoform 2 of Delta-aminolevulinic acid dehydratase	N	N	Y
Isoform 2 of Glioma tumor suppressor candidate region gene 1 protein	N	Y	N
Isoform 2 of GPI inositol-deacylase	N	N	Y
Isoform 2 of Sodium bicarbonate transporter-like protein 11	N	N	Y
Isoform 2 of Vacuolar protein sorting-associated protein 13B	N	N	Y
Isoform 3 of Perilipin-3	N	Y	N
Isoform 3 of Up-regulator of cell proliferation	N	N	Y
Isoform 4 of Polypyrimidine tract-binding protein 3	N	N	Y
Isoform 7 of Protein LAP2	N	N	Y
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 9	Y	N	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	N	N	Y
Keratin, type II cytoskeletal 4	N	N	Y
Keratin, type II cytoskeletal 6B	N	N	Y
Lipopolysaccharide-responsive and beige-like anchor protein	N	Y	N
L-lactate dehydrogenase B chain	Y	Y	N
NADH-cytochrome b5 reductase-like	N	N	Y
N-alpha-acetyltransferase 50	N	N	Y
Neuroendocrine convertase 1	Y	N	N
Nucleoside diphosphate kinase A	Y	Y	Y
Peptidyl-prolyl cis-trans isomerase A	Y	Y	N
Peroxiredoxin-2	Y	Y	Y
Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform	Y	N	N
Phosphoglycerate kinase	N	N	Y
Plasminogen	N	N	Y
Polyubiquitin-C	Y	Y	Y
POTE ankyrin domain family member F	N	N	Y
Protein ARMCX6	N	N	Y
Protein CNPPD1 (Fragment)	N	N	Y
Protein FAM180A	N	N	Y
Protein fantom	N	N	Y
Protein RRP5 homolog	N	N	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protocadherin Fat 2	N	N	Y
Purine nucleoside phosphorylase	Y	N	Y
Regulator of microtubule dynamics protein 2	Y	N	N
Retinal dehydrogenase 1	Y	Y	Y
Ribose-5-phosphate isomerase	N	N	Y
Selenium-binding protein 1	Y	N	Y

Sodium-coupled monocarboxylate transporter 1	N	N	Y
Solute carrier family 22 member 8	N	N	Y
Splicing factor 45 (Fragment)	N	Y	N
Superoxide dismutase [Cu-Zn]	Y	Y	Y
Synembryn-B (Fragment)	N	N	Y
Tau-tubulin kinase 1	N	N	Y
Thioredoxin	Y	Y	Y
Titin	N	Y	N
Transcription initiation factor TFIID subunit 10	N	N	Y
Transmembrane protein PVRIG	Y	N	N
Triosephosphate isomerase	N	Y	N
Uncharacterized protein	Y	N	N
Uncharacterized protein KIAA1841	N	N	Y
Vitamin D-binding protein	Y	N	Y
Voltage-dependent L-type calcium channel subunit alpha-1F	N	N	Y
Zinc finger protein 566 (Fragment)	N	Y	N
Zinc-alpha-2-glycoprotein	Y	N	N

Appendix 5.7 Proteins identified from LC MS/MS analysis of a DBS digest depleted of the top 12 plasma proteins. Sample digested for 8 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein beta/alpha	Y	y	Y
14-3-3 protein epsilon	Y	Y	Y
14-3-3 protein gamma	Y	Y	Y
14-3-3 protein theta	Y	Y	N
14-3-3 protein zeta/delta	Y	Y	Y
A disintegrin and metalloproteinase with thrombospondin motifs 13	N	N	Y
Actin, alpha cardiac muscle 1	N	Y	N
Actin, cytoplasmic 1	Y	Y	Y
Actin, gamma-enteric smooth muscle	Y	N	Y
Adenylate cyclase type 9	N	Y	N
Adenylate kinase isoenzyme 1	Y	Y	Y
Afamin	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-antiplasmin	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-hemoglobin-stabilizing protein	Y	Y	Y
Alpha-synuclein	Y	Y	Y
AMBP protein	Y	Y	Y
Angiotensinogen	Y	Y	Y
ANKRD26-like family C member 1A	N	Y	N
Ankyrin-1	Y	Y	Y
Antithrombin-III	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
ATP synthase subunit alpha, mitochondrial	N	Y	N
ATP synthase subunit beta, mitochondrial	N	Y	N
Band 3 anion transport protein	Y	Y	Y
Beta-2-glycoprotein 1	Y	Y	Y
Bisphosphoglycerate mutase	Y	Y	Y
Calmodulin	Y	Y	Y
Calpastatin	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	Y	Y
Charged multivesicular body protein 5	N	N	Y
Clusterin	Y	Y	Y
Cofilin-1	Y	Y	Y
Collagen alpha-1(IX) chain	N	Y	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Complement C1q subcomponent subunit A	Y	Y	N
Complement C1q subcomponent subunit B	Y	N	N
Complement C3	Y	Y	Y
Complement C4-B	Y	Y	Y
Complement C5	N	N	Y
Complement component C6	N	Y	N
Complement component C9	N	N	Y
Complement factor B	Y	Y	Y
Complement factor H	Y	Y	Y
Complement factor I	Y	N	Y
CUB domain-containing protein 1	Y	N	N
Cystic fibrosis transmembrane conductance regulator	N	Y	N
Cytochrome b5	Y	Y	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Deoxyribonucleoside 5'-monophosphate N-glycosidase	N	Y	N
Eukaryotic translation initiation factor 4B	N	N	Y
Eukaryotic translation initiation factor 5A-1	Y	Y	Y
Fibrinogen alpha chain	Y	N	Y
Fibrinogen beta chain	N	N	Y
Fibronectin type-III domain-containing protein 3A	N	N	Y
Flavin reductase	Y	Y	Y
Gelsolin	Y	Y	N
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Glycophorin-A	Y	Y	Y
Guanine nucleotide-binding protein alpha-12 subunit	N	N	Y
Haptoglobin	Y	Y	Y
Haptoglobin-related protein	N	N	Y
Heat shock 70 kDa protein 1	N	Y	N
Heat shock cognate 71 kDa protein	Y	Y	Y
Heat shock protein HSP 90-alpha	N	Y	N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
High affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8A	Y	N	N
Histidine-rich glycoprotein	Y	Y	Y
Histone H2A.Z	N	Y	N
Histone H2B type 1-K	N	Y	N
Ig alpha-1 chain C region	Y	Y	Y
Ig alpha-2 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig gamma-3 chain C region	Y	Y	Y
Ig gamma-4 chain C region	Y	Y	Y
Ig heavy chain V-III region GAL	N	Y	N
Ig heavy chain V-III region TEI	Y	Y	Y
Ig heavy chain V-III region TIL	Y	N	Y
Ig heavy chain V-III region VH26	Y	N	N
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region CAR	Y	N	N
Ig kappa chain V-I region DEE	Y	Y	Y
Ig kappa chain V-I region Lay	Y	Y	N
Ig kappa chain V-II region TEW	Y	Y	Y
Ig kappa chain V-III region SIE	Y	Y	Y
Ig kappa chain V-III region VG	Y	N	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Ig kappa chain V-IV region Len	Y	Y	Y
Ig lambda chain V-III region LOI	Y	N	Y
Ig lambda chain V-III region SH	Y	N	Y
Ig lambda chain V-IV region Hil	Y	Y	N
Ig lambda-1 chain C regions	Y	Y	Y
Ig mu chain C region	Y	N	Y
Ig mu heavy chain disease protein	N	Y	N
Inactive carboxypeptidase-like protein X2	N	Y	N
Inactive caspase-12	Y	N	N
Inter-alpha-trypsin inhibitor heavy chain H1	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H2	Y	Y	Y
Inter-alpha-trypsin inhibitor heavy chain H4	Y	Y	Y
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 14	N	Y	N
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 4	N	N	Y
Keratin, type II cytoskeletal 5	N	Y	Y
Kininogen-1	Y	Y	Y
Leucine-rich alpha-2-glycoprotein	Y	N	Y
Leucine-rich repeat neuronal protein 4	Y	N	N
Leucine-rich repeat-containing protein 9	N	Y	N
L-lactate dehydrogenase B chain	Y	Y	Y
Nck-associated protein 1	N	Y	N
NEDD4-binding protein 2	N	Y	N
NSFL1 cofactor p47	Y	Y	Y
Nucleoside diphosphate kinase A	Y	Y	Y
Peroxiredoxin-1	Y	Y	Y
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	Y
Peroxisomal NADH pyrophosphatase NUDT12	N	N	Y
Plasma protease C1 inhibitor	Y	Y	Y
Plasminogen	Y	Y	Y
Platelet basic protein	Y	Y	Y
Polyubiquitin-C	N	N	Y
Pre-rRNA-processing protein TSR1 homolog	Y	N	N
Proteasome activator complex subunit 1	N	Y	N
Proteasome subunit alpha type-1	N	N	Y
Proteasome subunit alpha type-2	N	N	Y
Proteasome subunit beta type-4	N	N	Y
Proteasome subunit beta type-7	Y	Y	Y
Protein 4.1	Y	Y	Y
Protein DDI1 homolog 2	N	Y	N
Protein DJ-1	Y	N	N
Protein FAM10A4	Y	N	Y
Protein FAM20A	N	Y	N
Protein S100-A10	N	Y	N
Protein S100-A12	Y	Y	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protein S100-A9	N	N	Y
Prothrombin	Y	Y	Y
Protocadherin beta-15	N	N	Y
PTB domain-containing engulfment adapter protein	N	Y	N
Purine nucleoside phosphorylase	Y	Y	Y
Putative tropomyosin alpha-3 chain-like protein	N	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum paraoxonase/arylesterase 1	Y	N	N
SH3 domain-binding glutamic acid-rich-like protein 3	Y	N	N
Small ubiquitin-related modifier 2	Y	N	N
Spectrin alpha chain, erythrocyte	Y	Y	Y
Spectrin beta chain, erythrocyte	Y	Y	Y
S-phase kinase-associated protein 1	Y	Y	N
Superoxide dismutase	Y	Y	Y
Taste receptor type 2 member 30	Y	N	N
Thioredoxin	Y	Y	Y
Thyroxine-binding globulin	Y	N	Y
Transaldolase	N	Y	Y
Transitional endoplasmic reticulum ATPase	N	N	Y
Transthyretin	Y	Y	Y
Triosephosphate isomerase	Y	Y	Y
Tropomyosin alpha-3 chain	Y	Y	Y
Tubulin alpha-3C/D chain	N	Y	N
Tubulin beta chain	N	Y	N
Tubulin-specific chaperone A	Y	N	N
Ubiquitin-associated domain-containing protein 1	N	Y	N
Ubiquitin-conjugating enzyme E2 N	N	Y	N
Uncharacterized protein C9orf40	Y	Y	Y
UV excision repair protein RAD23 homolog A	Y	Y	Y
Vacuolar fusion protein MON1 homolog A	N	N	Y
Vitamin D-binding protein	Y	Y	Y
Vitronectin	Y	Y	Y
X-ray repair cross-complementing protein 6	Y	N	N
Zinc finger CCHC domain-containing protein 5	N	Y	N
Zinc-alpha-2-glycoprotein	Y	Y	N

Appendix 5.8 Proteins identified from LC MS/MS analysis of a DBS digest depleted of haemoglobin. Sample digested for 1 hour.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein beta/alpha	Y	Y	Y
14-3-3 protein epsilon	Y	Y	Y
14-3-3 protein gamma	N	Y	Y
14-3-3 protein zeta/delta	N	Y	Y
Actin, cytoplasmic 1	Y	Y	Y
Adenylate kinase isoenzyme 1	N	Y	N
Afamin	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-hemoglobin-stabilizing protein	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Angiotensin-2	N	N	Y
Angiotensinogen	N	Y	N
Ankyrin repeat and BTB/POZ domain-containing protein 1	N	N	Y
Ankyrin-1	Y	Y	Y
Antithrombin-III	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein C-III	Y	Y	Y
Band 3 anion transport protein	Y	Y	Y
B-cell receptor CD22	Y	Y	N
Beta-2-glycoprotein 1	Y	N	Y
Beta-actin-like protein 2	N	N	Y
Beta-adducin	Y	Y	N
Beta-galactosidase-1-like protein 3	N	Y	N
BH3-interacting domain death agonist	N	N	Y
Bisphosphoglycerate mutase	Y	Y	Y
BPI fold-containing family B member 1	N	Y	Y
Calmodulin	Y	Y	Y
Calpastatin	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	N	Y	Y
Catalase	Y	Y	Y
Ceruloplasmin	Y	Y	Y
Chromosome 1 open reading frame 123	Y	N	N
Clusterin	Y	Y	Y
Cofilin-1	Y	Y	Y
Collagen alpha-1(XVIII) chain (Fragment)	N	Y	N
Complement C3	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Complement C4-B	Y	Y	Y
Complement factor B	Y	Y	Y
Complement factor H	Y	Y	Y
Complement factor I	Y	Y	N
Copper chaperone for superoxide dismutase	Y	Y	Y
Corticosteroid-binding globulin	N	N	Y
Cyclin-Y-like protein 2 (Fragment)	Y	N	N
Cytochrome b5	Y	Y	Y
Cytochrome P450 2B6	N	Y	Y
D-dopachrome decarboxylase	N	Y	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Dematin	Y	N	N
Dipeptidyl peptidase 3	N	N	Y
Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1,3-mannosyltransferase (Fragment)	N	Y	N
Dynein heavy chain 9, axonemal (Fragment)	N	Y	N
Eukaryotic translation initiation factor 4B	N	N	Y
Eukaryotic translation initiation factor 5A-1	N	Y	Y
Eukaryotic translation initiation factor 5A-2	Y	N	N
Fibrinogen alpha chain	Y	Y	Y
Flavin reductase (NADPH)	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Glycophorin Erik I-IV (Precursor)	N	N	Y
Glycophorin-A	N	Y	N
Haptoglobin	Y	Y	Y
Heat shock 70 kDa protein 1A/1B	N	N	Y
Heat shock cognate 71 kDa protein	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit delta (Fragment)	N	N	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Histidine triad nucleotide-binding protein 1	Y	N	Y
Histidine-rich glycoprotein	Y	y	Y
Hsc70-interacting protein	N	Y	N
Hyaluronan and proteoglycan link protein 3	N	N	Y
Ig alpha-1 chain C region	Y	Y	Y
Ig alpha-2 chain C region	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig gamma-3 chain C region	Y	Y	Y
Ig gamma-4 chain C region	Y	Y	Y
Ig heavy chain V-III region TEI	Y	Y	Y
Ig heavy chain V-III region TIL	Y	N	Y
Ig kappa chain C region	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Ig kappa chain V-I region DEE	Y	Y	Y
Ig kappa chain V-I region Lay	Y	Y	Y
Ig kappa chain V-III region SIE	Y	Y	Y
Ig kappa chain V-IV region Len	Y	N	Y
Ig lambda chain V-III region LOI	Y	Y	Y
Ig lambda-1 chain C regions	Y	Y	Y
Ig lambda-2 chain C regions	N	Y	Y
Ig lambda-6 chain C region	Y	N	N
Ig mu chain C region	Y	Y	Y
Inactive caspase-12	N	N	Y
Inter-alpha-trypsin inhibitor heavy chain H4	Y	Y	Y
Isoform 1 of Rho GTPase-activating protein 7	N	N	Y
Isoform 2 of Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	N	Y	N
Isoform 2 of Leucine-rich repeat-containing protein 9	N	N	Y
Isoform 2 of PRELI domain-containing protein 2	N	N	Y
Isoform 2 of Proteasome subunit alpha type-7	Y	N	N
Isoform 2 of Three prime repair exonuclease 1	N	N	Y
Isoform 3 of Interleukin-12 receptor subunit beta-1	N	Y	N
Isoform 3 of Magnesium-dependent phosphatase 1	N	N	Y
Isoform 4 of Lipid phosphate phosphatase-related protein type 3	N	Y	N
Kelch-like protein 30	N	Y	N
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 17	Y	N	N
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 6B	Y	N	N
Kinetochore-associated protein 1	Y	N	Y
Kininogen-1	Y	Y	Y
Lactoylglutathione lyase	N	Y	Y
Laminin subunit alpha-5	N	Y	N
L-lactate dehydrogenase A chain	Y	Y	Y
L-lactate dehydrogenase B chain	Y	Y	Y
Mediator of RNA polymerase II transcription subunit	Y	N	N
Myotrophin	N	N	Y
NSFL1 cofactor p47	Y	Y	Y
Nuclear transport factor 2 (Fragment)	Y	N	N
Nucleoside diphosphate kinase A	Y	Y	Y
Nucleoside diphosphate kinase B	N	Y	N
Peptidyl-prolyl cis-trans isomerase A	Y	Y	Y
Peroxiredoxin-1	Y	Y	Y
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Phosphoglycerate mutase 1	N	Y	Y
Phospholysine phosphohistidine inorganic pyrophosphate phosphatase	N	Y	Y
Pigment epithelium-derived factor	Y	Y	Y
Plasma protease C1 inhibitor	Y	Y	Y
Plasminogen	N	Y	Y
Plasminogen (Fragment)	Y	N	N
Platelet basic protein	Y	Y	Y
Polyubiquitin-C	N	Y	N
Proteasome activator complex subunit 1	Y	Y	Y
Proteasome subunit alpha type	Y	N	N
Proteasome subunit alpha type-1	N	N	Y
Proteasome subunit alpha type-2	N	Y	N
Proteasome subunit alpha type-4	Y	N	Y
Proteasome subunit alpha type-5	N	Y	N
Proteasome subunit alpha type-6	N	Y	Y
Proteasome subunit beta type-7	N	Y	Y
Protein 4.1	Y	Y	Y
Protein AMBP	N	Y	Y
Protein DDI1 homolog 2	Y	Y	N
Protein RRP5 homolog	N	Y	N
Protein S100-A12	N	N	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protein S100-A7	N	N	Y
Protein S100-A8	N	Y	N
Protein unc-13 homolog B	Y	N	N
Prothrombin	Y	Y	Y
Purine nucleoside phosphorylase	Y	Y	Y
Putative protein FAM10A4	Y	N	Y
Radixin	Y	Y	Y
Ran-specific GTPase-activating protein	N	N	Y
Retinol-binding protein 4	Y	Y	Y
SAGA-associated factor 29 homolog	N	N	Y
Selenium-binding protein 1	Y	Y	Y
Serine palmitoyltransferase 2	N	N	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum amyloid P-component	Y	Y	Y
SH3 domain-binding glutamic acid-rich-like protein 3	Y	Y	Y
Small ubiquitin-related modifier 2	Y	Y	Y
Spectrin alpha chain, erythrocytic 1	Y	Y	Y
Spectrin beta chain, erythrocytic	Y	Y	Y
S-phase kinase-associated protein 1	N	Y	Y
Stress-induced-phosphoprotein 1	N	Y	N
Superoxide dismutase [Cu-Zn]	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Syntaxin-1B (Fragment)	N	Y	N
Thioredoxin	Y	Y	Y
Thioredoxin domain-containing protein 17	Y	N	N
Transaldolase	Y	Y	Y
Transgelin-2	N	Y	N
Transthyretin	Y	Y	Y
Triosephosphate isomerase	Y	Y	Y
Tropomyosin 3	Y	Y	Y
Tubulin-specific chaperone A	Y	N	Y
Tyrosine--tRNA ligase, mitochondrial (Fragment)	N	N	Y
Ubiquitin-associated domain-containing protein 1	N	Y	N
Ubiquitin-conjugating enzyme E2 N	Y	Y	Y
Uncharacterized protein	Y	N	N
Uncharacterized protein C9orf40	N	Y	Y
Uncharacterized protein KIAA0240	N	N	Y
UV excision repair protein RAD23 homolog A	Y	Y	Y
Vitamin D-binding protein	Y	Y	Y
Vitronectin	N	Y	Y
Voltage-dependent R-type calcium channel subunit alpha-1E	Y	N	N
WD repeat-containing protein 17 (Fragment)	N	Y	N
Xaa-Pro dipeptidase	Y	N	N
YTH domain family protein 1	Y	N	N
Zinc finger CCHC domain-containing protein 5	N	Y	Y
Zinc-alpha-2-glycoprotein	Y	Y	Y

Appendix 5.9 Proteins identified from LC MS/MS analysis of a DBS digest depleted of haemoglobin. Sample digested for 8 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein epsilon	N	Y	Y
14-3-3 protein zeta/delta	N	Y	N
26S proteasome non-ATPase regulatory subunit 9	Y	N	Y
3-hydroxyanthranilate 3,4-dioxygenase	Y	N	N
60 kDa heat shock protein, mitochondrial	Y	N	N
Acidic leucine-rich nuclear phosphoprotein 32 family member C	Y	N	N
Actin, cytoplasmic 1	N	Y	N
Actin-related protein 2/3 complex subunit 2	Y	N	N
Acylamino-acid-releasing enzyme	Y	N	Y
ADP-sugar pyrophosphatase	N	Y	N
Afamin	Y	N	Y
Aflatoxin B1 aldehyde reductase member 3	Y	N	N
Alcohol dehydrogenase 1B	Y	N	N
Aldo-keto reductase family 1 member C4	Y	N	N
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	N	N	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-amylase 1	N	Y	N
Alpha-enolase	Y	N	N
Alpha-hemoglobin-stabilizing protein	Y	N	Y
Alpha-synuclein	Y	N	Y
AMBP protein	N	Y	N
Angiotensinogen	Y	N	Y
Ankyrin-1	Y	N	Y
Antithrombin-III	Y	Y	Y
Apolipoprotein A-I	N	Y	N
Apolipoprotein C-III	Y	N	Y
Attractin	Y	N	Y
Band 3 anion transport protein	N	Y	N
Beta-adducin	Y	N	Y
Betaine--homocysteine S-methyltransferase 1	Y	N	N
Bile salt sulfotransferase	Y	N	N
Bisphosphoglycerate mutase	N	Y	N
Bleomycin hydrolase	Y	N	Y
C-1-tetrahydrofolate synthase, cytoplasmic	Y	N	N
Calmodulin	Y	N	N
Calmodulin-like protein 5	Y	N	N
Calpastatin	Y	N	Y
Carbamoyl-phosphate synthase [ammonia], mitochondrial	Y	N	N
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Catalase	Y	Y	Y
CD44 antigen	Y	N	N
Ceruloplasmin	Y	Y	Y
Chromosome 1 open reading frame 123	Y	N	N
Clusterin	N	N	Y
Coagulation factor V	N	N	Y
Cofilin-1	N	Y	N
Colorectal mutant cancer protein	N	Y	N
Complement C1q subcomponent subunit A	N	Y	N
Complement C3	Y	Y	Y
Complement C4-A	N	Y	N
Complement C4-B	N	N	Y
Complement factor B	N	N	Y
Complement factor H	N	Y	N
Complement factor I	N	Y	N
Copper chaperone for superoxide dismutase	Y	N	Y
Cystatin-S	N	N	Y
Cytochrome b5	Y	N	Y
Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	Y	N	N
Delta-aminolevulinic acid dehydratase	Y	N	Y
Dematin	N	N	Y
Dematin (Fragment)	Y	N	N
Dermcidin	Y	N	Y
Dihydropteridine reductase	N	N	Y
Dipeptidyl peptidase 3	Y	N	N
Dynactin subunit 2	N	N	Y
Elongation factor 1-alpha 1	N	Y	N
Endoplasmin	Y	N	N
Enoyl-CoA hydratase, mitochondrial	Y	N	N
Eukaryotic translation initiation factor 5A-1	N	Y	N
Fatty acid synthase	Y	N	N
Fatty acid-binding protein, liver	Y	N	N
Flavin reductase (NADPH)	N	N	Y
G2/M phase-specific E3 ubiquitin-protein ligase (Fragment)	Y	N	N
Gamma-enolase	Y	N	Y
GDP-L-fucose synthase	Y	N	N
GDP-L-fucose synthase (Fragment)	N	N	Y
Glutathione S-transferase A1	Y	N	Y
Glutathione synthetase	Y	N	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	N
Glycophorin-A	N	Y	N
Haptoglobin	Y	Y	Y
Heat shock cognate 71 kDa protein	N	N	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit delta (Fragment)	N	Y	Y
Hemoglobin subunit gamma-1	N	Y	Y
Hemopexin	Y	Y	Y
Histidine-rich glycoprotein	Y	Y	Y
Ig alpha-1 chain C region	N	Y	N
Ig gamma-1 chain C region	N	Y	N
Ig gamma-2 chain C region	N	Y	N
Ig kappa chain C region	N	Y	N
Ig kappa chain V-III region SIE	N	Y	N
Ig lambda-1 chain C regions	N	Y	N
Inter-alpha-trypsin inhibitor heavy chain H1	N	Y	N
Inter-alpha-trypsin inhibitor heavy chain H2	N	Y	N
Inward rectifier potassium channel 16	N	Y	N
Isoform 2 of 14 kDa phosphohistidine phosphatase	N	N	Y
Isoform 2 of Aldo-keto reductase family 1 member C2	Y	N	N
Isoform 2 of Hydroxymethylglutaryl-CoA synthase, mitochondrial	Y	N	N
Isoform 2 of Proteasome subunit alpha type-4	Y	N	N
Isoform 2 of Tropomyosin alpha-1 chain	Y	N	Y
Isoform 5 of EF-hand calcium-binding domain-containing protein 6	N	N	Y
Keratin, type I cuticular Ha3-II	N	Y	N
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 14	N	Y	N
Keratin, type I cytoskeletal 16	N	Y	N
Keratin, type I cytoskeletal 17	N	Y	N
Keratin, type I cytoskeletal 19	N	N	Y
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cuticular Hb4	N	Y	N
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 5	N	Y	N
Keratin, type II cytoskeletal 6A	N	Y	N
Keratin, type II cytoskeletal 6B	N	Y	N
Keratin-associated protein 3-1	N	Y	N
Ketohexokinase	Y	N	N
Kininogen-1	Y	Y	Y
Leucine-rich alpha-2-glycoprotein	Y	N	N
Leucine-rich repeat neuronal protein 4	N	Y	N
LIM and SH3 domain protein 1	N	N	Y
LisH domain and HEAT repeat-containing protein KIAA1468	N	Y	N
L-lactate dehydrogenase A chain	Y	N	Y
L-lactate dehydrogenase B chain	Y	N	Y
Lysosome-associated membrane glycoprotein 2	N	N	Y
Lysozyme C	N	Y	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Nostrin	N	Y	N
NSFL1 cofactor p47	Y	N	Y
Nucleoside diphosphate kinase	Y	N	N
Peptidyl-prolyl cis-trans isomerase A	N	N	Y
Peroxiredoxin-2	Y	Y	Y
PH domain leucine-rich repeat-containing protein	Y	N	N
Phosphatidate phosphatase LPIN2	N	N	Y
Phosphoacetylglucosamine mutase	Y	N	N
Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	Y	N	Y
Pigment epithelium-derived factor	Y	N	Y
PITH domain-containing protein 1	Y	N	Y
Plasma protease C1 inhibitor	Y	Y	Y
Plasminogen	Y	N	Y
Platelet basic protein	Y	N	Y
Platelet glycoprotein Ib alpha chain	N	N	Y
Polyubiquitin-C	Y	N	Y
Probable ATP-dependent RNA helicase DDX41 (Fragment)	N	N	Y
Prolactin-inducible protein	Y	N	N
Proteasome subunit alpha type-1	Y	N	Y
Proteasome subunit alpha type-3	Y	N	Y
Proteasome subunit alpha type-5	N	N	Y
Proteasome subunit alpha type-6	Y	N	Y
Proteasome subunit alpha type-7	Y	N	N
Proteasome subunit beta type-1	Y	N	Y
Proteasome subunit beta type-2	Y	N	Y
Proteasome subunit beta type-6	Y	N	N
Protein 4.1	Y	Y	Y
Protein AMBP	Y	N	N
Prothrombin	Y	Y	Y
Putative nucleoside diphosphate kinase	N	N	Y
Ran-specific GTPase-activating protein (Fragment)	N	N	Y
Retinal dehydrogenase 1	Y	N	N
Retinol-binding protein 4	Y	Y	Y
Selenium-binding protein 1	N	N	Y
Serine hydroxymethyltransferase, cytosolic	Y	N	N
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum amyloid P-component	Y	N	Y
Serum deprivation-response protein	N	N	Y
Small ubiquitin-related modifier 2	Y	N	N
SMT3 suppressor of mif two 3 homolog 1 (Yeast), isoform CRA_b	N	N	Y
Spectrin beta chain, erythrocyte (Fragment)	N	N	Y
S-phase kinase-associated protein 1	Y	N	Y
Superoxide dismutase	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Syntaxin-3	N	Y	N
Thioredoxin	Y	Y	Y
Thymidine phosphorylase	Y	N	N
Thyroxine-binding globulin	Y	N	Y
TPR and ankyrin repeat-containing protein 1	Y	N	N
Transaldolase	Y	N	Y
Transcription elongation factor SPT6	N	Y	N
Transmembrane 9 superfamily member 4	N	Y	N
Transthyretin	Y	Y	Y
Tropomyosin 3	Y	N	Y
Tropomyosin alpha-4 chain	Y	Y	Y
Tubulin-specific chaperone A	N	N	Y
Uncharacterized protein C6orf163	Y	N	Y
UPF0587 protein C1orf123	N	N	Y
UTP--glucose-1-phosphate uridylyltransferase	Y	N	N
UV excision repair protein RAD23 homolog A	Y	N	Y
UV excision repair protein RAD23 homolog B	Y	N	N
Vitamin D-binding protein	Y	N	Y
Vitronectin	Y	N	N
WD repeat-containing protein 1	N	N	Y
Xaa-Pro dipeptidase	N	N	Y
Zinc-alpha-2-glycoprotein	Y	N	Y

Appendix 5.10 Proteins identified from LC MS/MS analysis of a DBS digest depleted of the top 12 plasma proteins and haemoglobin. Sample digested for 1 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
14-3-3 protein beta/alpha	N	N	Y
14-3-3 protein epsilon	Y	Y	Y
14-3-3 protein zeta/delta	Y	Y	Y
26S proteasome non-ATPase regulatory subunit 4	Y	N	N
26S proteasome non-ATPase regulatory subunit 9	N	Y	Y
Acidic leucine-rich nuclear phosphoprotein 32 family member A	Y	Y	N
Acylamino-acid-releasing enzyme	Y	Y	Y
Acyl-CoA:lysophosphatidylglycerol acyltransferase 1	N	N	Y
Adapter molecule crk	Y	Y	N
Afamin	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Alpha-2-macroglobulin	Y	Y	Y
Alpha-enolase	Y	Y	Y
Alpha-hemoglobin-stabilizing protein	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Angiotensinogen	Y	Y	Y
Ankyrin-1	Y	N	N
Antithrombin-III	Y	Y	Y
Apolipoprotein B-100	N	N	Y
Apolipoprotein C-III	N	N	Y
Attractin	Y	Y	Y
Band 3 anion transport protein	Y	Y	Y
Beta-adducin	Y	Y	Y
Bleomycin hydrolase	Y	Y	Y
BolA-like protein 2	N	Y	Y
Calcineurin-like phosphoesterase domain-containing protein 1	N	Y	Y
Calmodulin	Y	Y	Y
Calmodulin-like protein 5	Y	Y	Y
Calpastatin	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	N	Y	N
Catalase	Y	Y	Y
CD44 antigen	Y	Y	Y
Ceruloplasmin	Y	Y	Y
Charged multivesicular body protein 4a (Fragment)	N	N	Y
Clusterin	Y	N	N
Clusterin alpha chain (Fragment)	N	Y	Y
Complement C3	Y	Y	Y
Complement C4-B	Y	Y	Y
Complement factor H	Y	Y	Y
Copper chaperone for superoxide dismutase	Y	Y	y
Cytochrome b5	Y	Y	Y
Dedicator of cytokinesis protein 11	N	Y	N
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Dematin	Y	Y	Y
Dipeptidyl peptidase 3	Y	Y	Y
DNA damage-binding protein	Y	Y	Y
Dynactin subunit 2	Y	N	N
Epithelial discoidin domain-containing receptor 1 (Fragment)	N	N	Y
Eukaryotic translation initiation factor 5A-1	Y	Y	Y
Flavin reductase (NADPH)	Y	Y	Y
Gamma-enolase	Y	Y	Y
GDP-L-fucose synthase	Y	Y	Y
Glutathione synthetase	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit delta (Fragment)	Y	N	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Hepatoma-derived growth factor	Y	N	N
Histidine-rich glycoprotein	Y	Y	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig heavy chain V-III region TEI	Y	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-I region Lay	N	Y	Y
Ig kappa chain V-III region NG9 (Fragment)	Y	N	N
Ig lambda-1 chain C regions	Y	N	N
Ig lambda-7 chain C region	N	Y	N
Ig mu chain C region	N	Y	Y
Ig mu heavy chain disease protein	Y	N	N
Importin subunit beta-1	Y	Y	N
Insulin-like growth factor-binding protein complex acid labile subunit	Y	Y	Y
Isoform 2 of Forkhead box protein K1	N	N	Y
Isoform 2 of Lysine-specific histone demethylase 1B	N	Y	N
Isoform 2 of NIF3-like protein 1	N	N	Y
Isoform 2 of Solute carrier family 35 member	N	N	Y
Isoform 2 of Spectrin alpha chain, erythrocyte	N	N	Y
Isoform 2 of Tropomyosin alpha-1 chain	Y	Y	Y
Isoform 3 of Magnesium-dependent phosphatase 1	N	N	Y
Isoform Er3 of Ankyrin-1	N	Y	Y
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 14	Y	Y	Y
Keratin, type I cytoskeletal 15	N	Y	Y
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 5	N	N	Y
Keratin, type II cytoskeletal 6A	N	N	Y
Keratin, type II cytoskeletal 6B	Y	Y	N
Kininogen-1	Y	Y	Y
Leucine-rich alpha-2-glycoprotein	Y	N	Y
Leucine-rich repeat-containing protein 72	Y	Y	N
LIM and SH3 domain protein 1	Y	Y	Y
L-lactate dehydrogenase B chain	Y	Y	Y
Lysosome-associated membrane glycoprotein 2	N	Y	Y
Mediator of DNA damage checkpoint protein 1	N	Y	N
Myeloma-overexpressed gene 2 protein	N	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
NIF3-like protein 1	N	Y	N
NSFL1 cofactor p47	Y	Y	Y
Nucleoredoxin	Y	N	N
Nucleoside diphosphate kinase A	Y	Y	Y
Peptidyl-prolyl cis-trans isomerase A	Y	Y	Y
Peroxiredoxin-2	Y	Y	Y
Phosphatidylinositol 4-kinase type 2 beta	N	N	Y
Pigment epithelium-derived factor	Y	Y	Y
PITH domain-containing protein 1	Y	Y	Y
Plasma protease C1 inhibitor	Y	Y	Y
Plasminogen	Y	Y	N
Platelet basic protein	Y	Y	Y
Platelet glycoprotein Ib alpha chain	Y	Y	Y
Polyubiquitin-C	Y	Y	Y
Programmed cell death protein 5	Y	Y	Y
Prolactin-inducible protein	Y	Y	Y
Proteasome subunit alpha type-1	Y	Y	Y
Proteasome subunit alpha type-2	Y	Y	Y
Proteasome subunit alpha type-3	N	Y	Y
Proteasome subunit alpha type-4	Y	Y	Y
Proteasome subunit alpha type-5	Y	N	Y
Proteasome subunit alpha type-6	Y	Y	Y
Proteasome subunit alpha type-7	Y	Y	Y
Proteasome subunit beta type-1	Y	Y	Y
Proteasome subunit beta type-2	N	Y	Y
Proteasome subunit beta type-3	Y	Y	Y
Proteasome subunit beta type-4	Y	Y	Y
Proteasome subunit beta type-6	N	N	Y
Proteasome subunit beta type-7	N	Y	Y
Protein 4.1	Y	Y	Y
Protein AMBP	Y	Y	Y
Protein LZIC	Y	Y	Y
Protein phosphatase 1 regulatory subunit 7	N	N	Y
Protein S100-A6	Y	Y	Y
Prothrombin	Y	Y	Y
Putative Polycomb group protein ASXL3	N	Y	N
Ran-specific GTPase-activating protein	Y	Y	Y
Retinol-binding protein 4	Y	Y	Y
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y
Serum amyloid P-component	Y	Y	Y
Serum deprivation-response protein	Y	Y	Y
SH3 domain-binding glutamic acid-rich-like protein 3	Y	Y	Y
Small ubiquitin-related modifier 1	N	Y	N

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Solute carrier family 25 member 53	N	N	Y
Spectrin beta chain, erythrocytic	Y	Y	Y
S-phase kinase-associated protein 1	Y	Y	Y
Stathmin	N	N	Y
Superoxide dismutase [Cu-Zn]	Y	Y	Y
Thioredoxin	Y	Y	Y
Thioredoxin reductase 1, cytoplasmic (Fragment)	N	N	Y
Thrombospondin-1	N	Y	N
Thyroxine-binding globulin	Y	Y	Y
Transaldolase	Y	Y	Y
Transthyretin	Y	Y	Y
Triosephosphate isomerase	N	N	Y
Tropomodulin-1	Y	Y	Y
Tropomyosin 3	Y	Y	Y
Tropomyosin alpha-4 chain	Y	Y	Y
Tubulin-specific chaperone A	Y	Y	Y
Tyrosine--tRNA ligase, mitochondrial (Fragment)	Y	N	N
Ubiquitin-conjugating enzyme E2 variant 2	N	Y	Y
Uncharacterized protein	N	N	Y
Uncharacterized protein C9orf40	Y	Y	Y
UPF0587 protein C1orf123	Y	Y	Y
UV excision repair protein RAD23 homolog A	Y	Y	Y
UV excision repair protein RAD23 homolog B	N	N	Y
Vitamin D-binding protein	Y	Y	Y
Vitronectin	N	N	Y
WD repeat-containing protein 1	Y	N	Y
Zinc-alpha-2-glycoprotein	Y	Y	Y

Appendix 5.11 Proteins identified from LC MS/MS analysis of a DBS digest depleted of the top 12 plasma proteins and haemoglobin. Sample digested for 8 hours.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
26S proteasome non-ATPase regulatory subunit 9	N	Y	Y
Actin, cytoplasmic 1, N-terminally processed	N	N	Y
Adenosylhomocysteinase	N	N	Y
Afamin	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	N
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-adducin	Y	Y	Y
Alpha-hemoglobin-stabilizing protein	Y	Y	Y
Alpha-L-iduronidase	N	Y	N
Alpha-synuclein	Y	Y	Y
Antithrombin-III	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein C-I	Y	Y	Y
Apolipoprotein C-II	Y	N	N
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	Y	Y	Y
Apolipoprotein E	Y	Y	Y
Attractin	Y	Y	Y
Beta-2-glycoprotein 1	Y	Y	Y
Beta-actin-like protein 2	Y	N	N
Beta-adducin	Y	Y	Y
Bleomycin hydrolase	Y	Y	Y
C4b-binding protein alpha chain	Y	Y	Y
C4b-binding protein beta chain	Y	Y	N
Calmodulin	Y	Y	Y
Calmodulin-like protein 5	Y	N	N
Calpastatin	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	N	Y	Y
Catalase	Y	Y	Y
CD44 antigen	Y	Y	Y
Centrosomal protein of 192 kDa	N	Y	N
Ceruloplasmin	Y	Y	Y
Clusterin	Y	Y	Y
Complement C3	Y	Y	Y
Complement C4-B	Y	N	N
Complement factor B	Y	Y	Y
Complement factor H	N	Y	Y
Complement factor I	N	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Cystatin-A	Y	N	Y
Cystatin-S	N	Y	N
Cytochrome b5	Y	Y	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Dematin	Y	N	Y
Dermcidin	Y	Y	Y
Eukaryotic translation initiation factor 4B	Y	N	N
FGA protein	N	Y	Y
Fibrinogen alpha chain	Y	N	N
Fibrinogen beta chain	Y	N	N
Fibrinogen gamma chain	Y	Y	Y
Flavin reductase (NADPH)	Y	Y	Y
GDP-L-fucose synthase	Y	Y	Y
Haptoglobin	Y	Y	Y
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit delta (Fragment)	Y	Y	N
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Histidine-rich glycoprotein	Y	Y	N
Hsc70-interacting protein	Y	Y	Y
Ig alpha-1 chain C region	Y	Y	Y
Ig alpha-2 chain C region	N	N	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-III region NG9 (Fragment)	Y	Y	Y
Ig lambda chain V-III region LOI	N	Y	N
Ig lambda-7 chain C region	N	Y	N
Ig mu chain C region	Y	Y	Y
Immunoglobulin J chain	Y	Y	Y
Importin subunit beta-1	N	Y	Y
Insulin-like growth factor-binding protein complex acid labile subunit	Y	Y	Y
Isoform 2 of 14 kDa phosphohistidine phosphatase	Y	N	Y
Isoform 2 of Haloacid dehalogenase-like hydrolase domain-containing protein 2	Y	N	Y
Isoform 2 of Sushi domain-containing protein 3	Y	N	N
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 14	Y	Y	Y
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 5	Y	Y	Y
Keratin, type II cytoskeletal 6A	N	Y	Y
Kininogen-1	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Lactotransferrin	Y	N	Y
Leucine-rich alpha-2-glycoprotein	Y	Y	Y
LIM and SH3 domain protein 1	Y	N	N
L-lactate dehydrogenase A chain	Y	Y	Y
L-lactate dehydrogenase B chain	Y	Y	Y
Low affinity immunoglobulin gamma Fc region receptor III-A	Y	N	N
Low affinity immunoglobulin gamma Fc region receptor III-A (Fragment)	N	Y	N
L-selectin	Y	Y	Y
Lysosome-associated membrane glycoprotein 2	Y	Y	N
Lysozyme C	Y	Y	Y
Malate dehydrogenase, cytoplasmic	Y	N	N
Mitogen-activated protein kinase kinase kinase 14	N	Y	N
Myeloperoxidase	Y	N	N
Myomegalin	Y	N	N
NSFL1 cofactor p47	Y	Y	Y
Peptidyl-prolyl cis-trans isomerase A	Y	Y	N
Peroxiredoxin-1	N	Y	Y
Peroxiredoxin-2	Y	Y	Y
Peroxiredoxin-6	N	N	Y
Phosphatidylinositol-glycan-specific phospholipase D	N	Y	N
Pigment epithelium-derived factor	Y	Y	N
Plasma protease C1 inhibitor	Y	Y	N
Plasminogen	Y	Y	Y
Platelet basic protein	Y	Y	Y
Platelet glycoprotein Ib alpha chain	Y	Y	Y
Polyubiquitin-C	Y	N	Y
Probable ATP-dependent RNA helicase DDX60-like (Fragment)	N	N	Y
Prolactin-inducible protein	Y	Y	Y
Protein 4.1	Y	Y	Y
Protein AMBP	Y	Y	N
Protein S100-A4	Y	Y	Y
Protein S100-A7	Y	Y	Y
Protein S100-A8	Y	Y	N
Protein S100-A9	Y	N	N
Prothrombin	Y	Y	Y
Putative nucleoside diphosphate kinase	N	Y	N
Quinone oxidoreductase PIG3	Y	Y	N
RAD23 homolog B (S. cerevisiae), isoform CRA_a	Y	N	N
Ran-specific GTPase-activating protein	Y	Y	Y
Retinol-binding protein 4	Y	Y	Y
Secretoglobin family 1D member 2	N	Y	N
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y
Serum albumin	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Serum deprivation-response protein	Y	Y	Y
SH3 domain-binding glutamic acid-rich-like protein 3	Y	Y	Y
Spectrin alpha chain, erythrocytic 1	Y	Y	Y
Spectrin beta chain, erythrocyte	Y	N	N
Spectrin beta chain, erythrocyte (Fragment)	N	N	Y
Stathmin	Y	Y	N
Stathmin (Fragment)	N	N	Y
Stress-induced-phosphoprotein 1	Y	N	N
Superoxide dismutase	Y	Y	Y
Tetranectin	Y	N	N
Thioredoxin	Y	Y	Y
Thrombospondin-1	Y	N	N
Thyroxine-binding globulin	Y	Y	Y
Transaldolase	Y	Y	Y
Transthyretin	Y	Y	Y
Triosephosphate isomerase	Y	Y	Y
Tropomyosin 3	Y	N	Y
Tropomyosin alpha-1 chain	N	Y	Y
Tubulin-specific chaperone A	Y	Y	Y
Uncharacterized protein C13orf33	N	Y	N
Uncharacterized protein C9orf40	Y	N	N
UPF0587 protein C1orf123	Y	Y	N
UV excision repair protein RAD23 homolog A	N	Y	Y
Vitamin D-binding protein	Y	Y	N
Vitamin K-dependent protein S	N	Y	N
Vitronectin	Y	N	N
WD repeat-containing protein 1	N	Y	N
Xaa-Pro dipeptidase	N	N	Y

Appendix 5.12 Proteins identified from LC MS/MS analysis of a DBS digest depleted of haemoglobin and the top 12 plasma proteins. Sample digested for 1 hour.

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
26S proteasome non-ATPase regulatory subunit 9	N	Y	Y
Actin, cytoplasmic 1, N-terminally processed	N	Y	N
Adenylate kinase 7, isoform CRA_e	Y	N	N
Afamin	Y	Y	Y
Alpha-1-acid glycoprotein 1	Y	Y	Y
Alpha-1-acid glycoprotein 2	Y	Y	Y
Alpha-1-antichymotrypsin	Y	Y	Y
Alpha-1-antitrypsin	Y	Y	Y
Alpha-1B-glycoprotein	Y	Y	Y
Alpha-2-HS-glycoprotein	Y	Y	Y
Alpha-2-macroglobulin	Y	Y	Y
Alpha-adducin	Y	Y	N
Alpha-enolase	Y	Y	Y
Alpha-hemoglobin-stabilizing protein	Y	Y	Y
Alpha-synuclein	Y	Y	Y
Ankyrin-1	Y	Y	Y
Antithrombin-III	Y	Y	Y
Apolipoprotein A-I	Y	Y	Y
Apolipoprotein A-II	Y	Y	Y
Apolipoprotein A-IV	Y	Y	Y
Apolipoprotein C-I	N	Y	N
Apolipoprotein C-III	Y	Y	Y
Apolipoprotein D	Y	Y	Y
Apolipoprotein E	Y	Y	Y
Apolipoprotein M	N	N	Y
Aspartate aminotransferase, cytoplasmic	N	N	Y
Attractin	Y	Y	Y
Beta-2-glycoprotein 1	Y	Y	Y
Beta-adducin	Y	Y	Y
Bleomycin hydrolase	Y	Y	Y
Brevican core protein	N	Y	N
C4b-binding protein beta chain	Y	Y	Y
Calcium-regulated heat stable protein 1	N	Y	N
Calmodulin	Y	Y	Y
Calmodulin-like protein 5	Y	Y	Y
Carbonic anhydrase 1	Y	Y	Y
Carbonic anhydrase 2	Y	N	N
Caspase-14	N	Y	N
Catalase	Y	Y	Y
CD44 antigen	Y	N	N
Ceruloplasmin	Y	Y	Y
Chromodomain-helicase-DNA-binding protein 1-like	Y	N	N
Clusterin	Y	Y	Y
Complement C3	Y	Y	Y
Complement C4-B	Y	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Complement factor B	Y	Y	Y
Complement factor I	N	N	Y
Cystatin-A	Y	Y	Y
Cytochrome b5	Y	Y	Y
Delta-aminolevulinic acid dehydratase	Y	Y	Y
Dematin	N	N	Y
Dermcidin	Y	Y	Y
E3 ubiquitin-protein ligase UBR3	N	Y	N
Eukaryotic translation initiation factor 4B	Y	N	Y
Eukaryotic translation initiation factor 5A-1	Y	Y	Y
Fibrinogen alpha chain	Y	Y	Y
Fibrinogen beta chain	N	Y	Y
Fibrinogen gamma chain	N	Y	Y
Filaggrin-2	Y	N	Y
Flavin reductase (NADPH)	Y	Y	Y
GDP-L-fucose synthase	Y	Y	Y
Glyceraldehyde-3-phosphate dehydrogenase	Y	Y	Y
Haptoglobin	Y	Y	Y
Heat shock protein beta-1	N	Y	N
Hemoglobin subunit alpha	Y	Y	Y
Hemoglobin subunit beta	Y	Y	Y
Hemoglobin subunit delta	Y	Y	Y
Hemoglobin subunit delta (Fragment)	Y	Y	Y
Hemoglobin subunit gamma-1	Y	Y	Y
Hemopexin	Y	Y	Y
Histidine-rich glycoprotein	Y	Y	Y
Ig alpha-2 chain C region	Y	N	Y
Ig gamma-1 chain C region	Y	Y	Y
Ig gamma-2 chain C region	Y	Y	Y
Ig kappa chain C region	Y	Y	Y
Ig kappa chain V-III region NG9 (Fragment)	Y	Y	Y
Ig lambda chain V-III region LOI	Y	Y	Y
Ig lambda-1 chain C regions	Y	Y	Y
Ig mu chain C region	Y	Y	Y
Ig mu heavy chain disease protein	Y	Y	Y
Insulin-like growth factor-binding protein complex acid labile subunit	Y	Y	Y
Integrin alpha-IIb light chain, form 1	N	Y	N
Inter-alpha-trypsin inhibitor heavy chain H4	N	Y	N
Isoform 2 of 14 kDa phosphohistidine phosphatase	Y	Y	Y
Isoform 2 of Cullin-associated NEDD8-dissociated protein 2	N	Y	N
Isoform 2 of Dipeptidyl peptidase 3	N	Y	N
Isoform 2 of Nuclear receptor corepressor 1	N	Y	N
Isoform 2 of Spectrin alpha chain, erythrocyte	N	N	Y
Isoform 5 of cAMP-specific 3',5'-cyclic phosphodiesterase 4A	N	N	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Isoform Short of Laminin subunit gamma-2	N	Y	N
Junction plakoglobin	Y	N	N
Keratin, type I cytoskeletal 10	Y	Y	Y
Keratin, type I cytoskeletal 14	Y	Y	Y
Keratin, type I cytoskeletal 16	N	Y	N
Keratin, type I cytoskeletal 17	Y	Y	N
Keratin, type I cytoskeletal 9	Y	Y	Y
Keratin, type II cytoskeletal 1	Y	Y	Y
Keratin, type II cytoskeletal 2 epidermal	Y	Y	Y
Keratin, type II cytoskeletal 5	Y	Y	Y
Keratin, type II cytoskeletal 6A	N	Y	Y
Keratin, type II cytoskeletal 6B	Y	N	N
Mesoderm induction early response protein 2	N	Y	N
Nebulin	N	N	Y
NIF3-like protein 1	N	Y	N
NSFL1 cofactor p47	Y	Y	Y
Peptidyl-prolyl cis-trans isomerase A	N	Y	Y
Peroxiredoxin-2	Y	Y	Y
Phospholysine phosphohistidine inorganic pyrophosphate phosphatase	Y	Y	Y
Plasma protease C1 inhibitor	Y	N	N
Plasminogen	Y	Y	Y
Platelet basic protein	Y	Y	Y
Platelet glycoprotein Ib alpha chain	Y	Y	Y
Polyubiquitin-C	Y	Y	N
Prolactin-inducible protein	Y	Y	Y
Proteasome activator complex subunit 1	Y	N	N
Proteasome activator complex subunit 1 (Fragment)	N	N	Y
Protein 4.1	Y	Y	Y
Protein AMBP	Y	Y	Y
Protein LZIC	Y	N	N
Protein RRP5 homolog	N	N	Y
Protein S100-A4	Y	Y	Y
Protein S100-A6	Y	Y	Y
Protein S100-A7	Y	Y	Y
Protein S100-A8	N	Y	N
Protein S100-A9	Y	Y	N
Prothrombin	Y	Y	Y
Putative unconventional myosin-XVB	Y	N	Y
Quinone oxidoreductase PIG3	Y	N	N
Radixin	N	Y	N
Retinol-binding protein 4	Y	Y	Y
Roundabout homolog 2	Y	N	N
Secretoglobin family 1D member 2	Y	Y	Y
Selenium-binding protein 1	Y	Y	Y
Serotransferrin	Y	Y	Y

Protein description	Present replicate 1 Y/N	Present replicate 2 Y/N	Present replicate 3 Y/N
Serum albumin	Y	Y	Y
Serum amyloid A-4 protein	N	Y	N
Serum deprivation-response protein	Y	Y	Y
SH3 domain-binding glutamic acid-rich-like protein 3	Y	Y	Y
Sialic acid-binding Ig-like lectin 16	Y	N	Y
Sodium/potassium/calcium exchanger 3	N	N	Y
Solute carrier family 2, facilitated glucose transporter member 1	N	Y	N
Spectrin alpha chain, erythrocytic 1	Y	Y	N
S-phase kinase-associated protein 1	Y	Y	Y
Stathmin	N	N	Y
Superoxide dismutase	Y	Y	Y
Thioredoxin	Y	Y	Y
Thioredoxin domain-containing protein 17	Y	N	N
Thyroxine-binding globulin	Y	Y	Y
Transaldolase	Y	Y	Y
Transthyretin	Y	Y	Y
Triosephosphate isomerase	Y	Y	Y
Tropomyosin 3	Y	N	N
Tropomyosin alpha-3 chain	N	Y	Y
Tyrosine-protein kinase FRK	Y	N	N
Ubiquitin (Fragment)	N	N	Y
Uncharacterized protein C9orf40	Y	Y	Y
UPF0587 protein C1orf123	Y	Y	Y
UV excision repair protein RAD23 homolog A	Y	Y	Y
Vitamin D-binding protein	Y	Y	Y
Vitamin K-dependent protein S	Y	N	N
Vitronectin	N	Y	Y
WD repeat-containing protein 1 (Fragment)	N	N	Y
Xaa-Pro dipeptidase	Y	N	N
Zinc-alpha-2-glycoprotein	Y	Y	Y
Zyxin	N	Y	N

Appendix 5.13 Proteins identified from LC MS/MS analysis of a DBS digest depleted of haemoglobin and the top 12 plasma proteins. Sample digested for 8 hours.

Appendix 6

Replicate 1

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

Replicate 2

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

Replicate 3

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

Appendix 6.1 Sequence coverage of A1AT protein obtained from LC MS/MS analysis of DBS digestions from healthy human controls.

SZ1

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

SZ2

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

SZ3

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

ZZ1

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

ZZ2

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

ZZ3

MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI TPNLAEFAFS

LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI LEGLNFNLTE IPEAQIHEGF

QELLRTLNQP DSQLQLTTGN GLFLSEGLKL VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ

INDYVEKGTQ GKIVDLVKEL DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV

KVPMMKRLGM FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL

ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP LKLSKAVHKA

VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE QNTKSPLFMG KVVNPTQK

Appendix 6.2. Sequence coverage of A1AT from SZ and ZZ genotype DBS digestions obtained from untargeted data dependent LC MS/MS analysis

Appendix 7 Published Journal Articles

- Martin, N.J.; Bunch, J.; Cooper, H.J. 'Dried blood spot proteomics: Surface extraction of endogenous proteins coupled with automated sample preparation and mass spectrometry analysis' Journal of the American society for mass spectrometry 2013, 24, 1242-1249.
- Edwards R, Martin N, Cooper H, Hemoglobin variant analysis of whole blood and dried blood spots by MS. Bioanalysis 2013; 5:2043-2052
- Sarsby, J; Martin, N.J.; Lalor, P. F.; Bunch, J.; Cooper, H.J. 'Top-down and bottom-up identification of proteins by liquid extraction surface analysis mass spectrometry of healthy and diseased human liver tissue' Journal of the American society for mass spectrometry, 2014; 25:1953-1961
- Martin, N.J.; Cooper, H.J. 'Challenges and opportunities in mass spectrometric analysis of proteins from dried blood spots' Expert reviews in proteomics, 2014, 11, 685-695.
- Martin, N.J.; Griffiths, R.L.; Edwards, R.L.; Cooper, H.J. 'Native liquid extraction surface analysis mass spectrometry: Analysis of non-covalent protein complexes directly from dried substrates' ; Journal of the American society for mass spectrometry, 2015, 26, 1320-1327