



University
of Glasgow

Lang, Alastair Michael (2013) *Developing tissue proteomics: Differential in gel electrophoresis in biomarker discovery and proteomic degradation*. PhD thesis.

<http://theses.gla.ac.uk/4642/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given



College of Medical, Veterinary and
Life Sciences

Thesis submitted for the degree of

Doctor of Philosophy

Year of submission 2013

Presented by

Alastair Michael Lang B.Sc Hons. M.Res. PGCE. M.Ed (Cantab)

The School of Life Sciences: Institute of
Molecular Cell and Systems Biology

Developing tissue proteomics: Differential in gel
electrophoresis in biomarker discovery and
proteomic degradation

Student: Alastair Michael Lang

Matriculation No: 9906621

Supervisor: Dr Andrew Pitt

Main body word count: ~85216

“Science is built up of facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house.”

Henri Poincare, *Science and Hypothesis*, 1905

i. Acknowledgements

Writing a thesis and doing a Ph.D is a long journey and like any long journey you meet a lot of people along the way who help guide you in the right direction. Also sufficing to say over such a long time the memory suffers and so if you feel you should be in these acknowledgements you are probably right and my memory has lapsed. Please don't take it personally, the omission is unintentional.

I would like to start by thanking my supervisor Dr Andrew Pitt for the opportunity to work in such a well-stocked, friendly lab and for all the guidance over the past years. Equally, a very large thank you needs to go to Dr Richard Goodwin for his constant good humour, advice and time. It was so frequently beyond the call and to this I will always be grateful.

I would like to give an all-inclusive thank you to; All the staff in the Functional Genomics facility in Glasgow University and the North lab. In particular I would like to give my deepest thanks to Dr Jayawardena and Dr Richard Burchmore for all the training using the DiGE equipment and a constant source of knowledge regarding DiGE gels and DeCyder analysis. I must also thank Dr Karl Burgess for his help with and training on numerous mass spectrometers, LC systems and any technical issues I encountered. I must also give a big thank you to Dr Sarah Cummings and Dr Susan Horne for all of their assistance with performing western blots. I would also like to thank Dr Heather Allingham for her collaboration during the three years of lab work.

A large thanks must go to Dr Martin McBride at the cardiovascular research unit at Glasgow University for the supplying of kidney tissue and meetings regarding the Hypertension part of this thesis and to Mr Bruno Bacher of GE Healthcare, Germany for his assistance in data extraction and DiGE statistics.

Last but certainly not least I need to thank my family and friends. My parents for their support both financial and emotional over the years, in addition to the gentle push I often have needed. The most important is last. To my darling wife, who for most of this write up process has been a "thesis widow" and has provided me with the emotional encouragement and support, and a lot of hot food. You are the reason I have finished this. Therefore I dedicate this thesis to you with love.

ii. Abstract

The field of proteomics and functional genomics has developed steadily since the completion of the human genome project. The wealth of genomic information and the pace at which it was compiled was astounding. Proteomics, despite considerable effort, on the other hand has not seen quite the same pace of development. The progress being considerably hindered by the lack of an amplification process and the relative complexity of the proteome in comparison to the genome. These intrinsic difficulties have led to the sensitivity of proteomic techniques being pushed closer to physical limits. There is therefore a further need to re-evaluated techniques such as sample preparation and integrity, analytical methods and collaborative strategies to maximise the effectiveness and quality of data collected.

The importance of tissue in scientific and clinical research is unequivocal. However, tissue is difficult to collect, store and work with due to issues with proteomic degradation and storage. Good lab practices can minimise the effect of degradation but degradation of proteins can be rapid. Strategies to minimise degradation include freezing, formalin fixing and microwave treatment which all have their relative advantages and disadvantages. The importance of sample preparation as being the top of the workflow is often acknowledged but improvements are not well described in the literature.

The main aim of this thesis is to present investigative studies into the mitigation of some of the limitations in tissue sample degradation, analytical approaches in differential in gel electrophoresis and accessing DiGE spot and tissue profile data. Presented is the evaluation of the effectiveness of rapid and controlled heating of intact tissue to inactivate native enzymatic activity and to aid in the cessation of proteomic degradation. A multifaceted analytical approach of differential in Gel electrophoresis spot data is assessed, giving proteomic profiles of mouse brain tissue. Preliminary data is presented showing that the process of heat-treatment has had a predominantly beneficial effect on mouse brain tissue, with a higher percentage of spots stabilised in heat-treated samples compared to snap-frozen samples. However, stabilisation did occur in snap-frozen samples for different protein spot so the appropriateness of using heat-treatment is as yet not fully determined and requires further analysis.

In addition, the variation in tissue profiles of WKY, SP.WKYGla.2a and SHRSP rat model for hypertension is investigated with the future prospect of providing that vital connection between genomic and proteomic data and link phenotype and genotype preliminarily investigated. A number of putative markers were identified and quantified using DiGE analysis. In order for these markers to be accepted as biomarkers, more downstream validation is required, however this study provides a good spring board as a proof of concept in using DiGE as an global putative biomarker discovery platform.

iii. Declaration

I hereby declare that the thesis that follows is of my own making, that it is a record of the work done by myself, unless otherwise indicated and referenced and that it has not been presented in any previous application for a higher degree.

Signature: 

iv. Contents

i.	Acknowledgements.....	i
ii.	Abstract.....	ii
iii.	Declaration	iv
iv.	Contents.....	v
v.	Table of figures.....	x
vi.	Abbreviations	xxix
1	Introduction.....	1
1.1	The study of proteomics.....	2
1.2	Aim and overview of this thesis.....	4
1.2.1	Protein degradation in mouse brain tissue and DiGE.....	4
1.2.2	Putative biomarker discovery in hypertension and moving towards linking genomic and proteomic data.....	4
1.2.3	Overview of thesis	4
1.3	Proteomics and biomarker discovery.....	6
1.3.1	Biomarker discovery in fluids.....	8
1.3.2	Biomarker discovery in tissue.....	13
1.3.3	Biomarker discovery using proteomics: A success storey?.....	17
1.3.4	Proteomic work flow for biomarker discovery.....	19
1.3.5	Summary.....	28
1.4	Quantitative proteomic technologies used in biomarker discovery.....	29
1.4.1	Differential in Gel Electrophoresis (DiGE).....	31
1.4.2	Quantitation in mass spectrometry	40
1.5	Other proteomics techniques	59
1.5.1	The importance of sample collection and storage.....	59
1.5.2	Sample fractionation and preparation	60
1.5.3	Separation in proteomics.....	61
1.5.4	Mass Spectrometry	72

1.6	Renal proteomics	81
1.6.1	Application of proteomics in renal research.	81
1.6.2	Challenges for proteomics in studying the kidney proteome	100
1.7	Statistical Analysis	102
1.7.1	Types of error	102
1.7.2	The Multiple testing problem and false discovery rate.	103
1.7.3	Statistical Analysis in Proteomics and DiGE	104
1.7.4	Univariate Analysis	107
1.7.5	Multivariate Analysis performed in the extended data analysis module	109
1.7.6	The problem of missing values.....	110
1.8	Summary	113
2	Material and Methods	114
2.1	Proteomic Methods	114
2.1.1	DiGE: Differential in Gel Electrophoresis.....	114
2.1.2	Saturation/scarce sample labelling.	120
2.1.3	Protein identification using Mass Spectrometry.....	129
2.1.4	Laser micro dissection	134
2.1.5	Western blotting.....	134
2.2	Chemicals, Consumables and Solutions.....	137
2.2.1	Tris/Glycine Electrophoresis Running Buffer	137
2.2.2	Agarose Sealing Solution.....	137
2.2.3	Bind Silane Solution	137
2.2.4	Displacing Solution.....	138
2.2.5	Equilibration Buffer	138
2.2.6	Fixing Solution.....	138
2.2.7	Lysis buffer for minimal labelling	139

2.2.8	Lysis buffer for saturation labelling	139
2.2.9	Polyacrylamide Gel Solution (12.5%).....	140
2.2.10	Rehydration Buffer	140
2.2.11	Spot handling workstation solutions	141
2.2.12	Sypro Orange Staining Solution.....	141
2.2.13	Other stock solutions for DiGE	141
3	Differential in gel electrophoresis analysis of wild type mouse brain tissue: A multi-faceted analytical approach to assessing the effect of heat treatment on the degradation of the proteome.....	142
3.1	Aims.....	142
3.2	Structure of the Chapter	144
3.3	Introduction.....	145
3.3.1	Sample preparation and protein degradation.....	145
3.3.2	Targeted tissue	150
3.4	Methods	151
3.4.1	Sample collection, extraction and processing.....	151
3.4.2	Identification of proteins from gel spots	153
3.4.3	Western Blot analysis	154
3.4.4	Experimental design.....	155
3.4.5	Statistical methods	157
3.5	Rationale of Analysis	159
3.5.1	Introduction.....	159
3.5.2	Methodology of DiGE analysis	159
3.6	Results and Discussion	174
3.6.1	Introduction to Results	174
3.6.2	Profile Classification.....	174
3.6.3	Pilot Investigation	176

3.6.4	Main Investigation.....	193
3.6.5	Summary of Venn analysis for the main investigation.....	235
3.6.6	Validative results.....	237
3.6.7	LC-MS with Label free quantitation.....	256
3.6.8	Mass Spectrometry Imaging	258
3.7	Summary and Conclusions.....	262
4	Biomarker discovery and the assessment of variation in the proteomic profiles of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model.....	264
4.1	Aims.....	265
4.2	Structure of the chapter.....	266
4.3	Introduction.....	267
4.3.1	Epidemiology of hypertension and control of blood pressure.....	267
4.3.2	Genetic factors: Monogenic.....	270
4.3.3	Rat models of human hypertension: The use of stroke-prone spontaneously hypertensive Rat (SHRSP).....	271
4.3.4	Proteomics and Hypertension	274
4.4	Methods	276
4.4.1	Sample collection, extraction and processing.....	277
4.4.2	Identification of proteins from gel spots	285
4.4.3	Macro-dissection of kidney tissue	286
4.4.4	Experimental Design.....	287
4.4.6	Statistical Methods	289
4.5	Rationale for Analysis.....	290
4.5.1	Pilot investigation and Main investigation.....	290
	Profile analysis for differences and similarities between strains.....	291
4.6	Results and discussion.....	295
4.6.1	Pilot investigations.....	296

4.6.3	Main investigation.....	325
4.6.4	Validation	346
4.6.5	Mass spectrometry Imaging.....	358
4.6.6	Label free quantitation using LC-MS of rat plasma.....	359
4.6.7	Biological significance of potential biomarkers discovered	362
4.7	Conclusion and Summary.....	371
5	General Conclusions.	372
5.1	Introduction to general discussion.....	372
5.2	Major findings.....	373
5.2.1	Proteomic degradation in mouse brain tissue.....	373
5.2.2	Proteomic Profiling of Kidney tissue	375
5.3	Limitations of DiGE	387
5.4	DiGE as a biomarker discovery platform.....	388
5.5	Future Investigation	390
5.5.1	Saturation labelling of kidney tissue	390
5.5.2	Determination of Protein concentration.....	392
5.5.3	Laser Micro-dissection of tissue and Mass spectrometry	393
5.6	Summary	395
6	References	396
7	Appendices.....	443
7.1	Tables of Identifications.....	443
7.1.1	Hypertension pilot study identifications.....	443
7.1.2	Hypertension main study identifications.....	461
7.2	Published work	471
7.3	Poster Presentation: HUPO 2008, Amsterdam	482

v. Table of figures

- Figure 1-1: Genome vs. Proteome. The Peacock butterfly (*Inachis Io*) emerges from the caterpillar. Both have the same genome but have markedly different proteomes, highlighting the dynamic and complex interactions of proteins and changes in phenotype.1
- Figure 1-2: Typical Proteomics Work flow. The first stage is sample preparation. Often samples must be desalted ready for protein separation and detection. After the sample has been separated it is then taken through to identification and analysis. .3
- Figure 1-3: The dynamic range of the plasmaome. This graph demonstrates the massive dynamic range of the plasmaome, from 10^{10} down to 10^{-1} which presents a problem for proteomic analysis of biomarkers. The blue area represents the most abundant plasma proteins such as albumin and various coagulation factors. The green area represents tissue leakage products such as cytokeratin and gastrin. Finally, the purple section represents interleukins and cytokines, such as calcitonin and somatostatin. Adapted from Schiess et al, 2008.10
- Figure 1-4: The biomarker discovery pathway. This displays a now often referenced approach to biomarker discovery and validation proposed by Rifai et al, 2006 called the biomarker pipeline. This involves several steps from discovery to validation. At the discovery phase of the pipeline, a more global approach is taken, in an attempt to examine 1000s of potential markers using small samples. As the pipeline approaches the validation stage a more targeted approach is required, with few markers being scanned but with a greater number of samples included. This was adapted from Rifai et al, 2006 and Thongboonkerd. 2009.21
- Figure 1-5: Schematic of the experimental stages in biomarker discovery in proteomics with the movement towards validation for clinical purposes. Samples come with a huge dynamic range and mixture of various proteins and peptides. These samples are then simplified by various possible methods of fractionation including; using liquid chromatography, depletion columns or antibody enrichment. The proteomic methods of choice are picked and separation and quantitation performed. Then the process of identification and validation can be executed. This is followed by a variety of possible steps toward clinical assay development.23
- Figure 1-6: Schematic of a Global proteomics work flow in biomarker discovery (Phase 1 of Figure 1-4).24
- Figure 1-7: Diagram showing the strategies employed in a quantitative proteomics work flow. It show where the introduction of labels and mixing occurs and illustrated the advantages of metabolic labelling reproduced from (Gouw et al., 2010).30
- Figure 1-8: Schematic Work flow of a Minimal DiGE gel. Sample from two or more different cell or tissue samples can be simultaneously compared. A combination of

all samples is mixed and labelled as a pool internal standard. These are then missed and loaded into a standard 2DE work flow that includes isoelectric focusing and molecular separation on SDS-PAGE gels. The gels are then scanned and analysed using GE Healthcare proprietary software. Figure was taken from DiGE manual produced by GE Healthcare, UK.	32
Figure 1-9: A schematic of labelling reaction for DiGE minimal labelling. The bond is created between a lysine and the NHS ester reactive group of the Cy dye forming an amide bond. Figure was taken from the DiGE manual produced by GE Healthcare, UK.	33
Figure 1-10: Schematic of Saturation DiGE Labelling. The disulphide bridges of the thiol groups on the Cysteines must be broken before Labelling can take place. A one-hour incubation in tris-(2-carboxyethyl) phosphine hydrochloride (TCEP) prevents the thiol forming the disulphide bridges.	34
Figure 1-11: Schematic of quantitation using SILAC. Cells are grown in the SILAC media which is incorporated into the cells as they grow. Cells are lysed, solubilised, digested and then separated before MS/MS is performed. The relative peak shifts in the light, middle and heavy chains allows quantitation to be performed. Adapted from (Geiger et al., 2011).....	42
Figure 1-12: The process of ICAT. A) Shows the ICAT Label with its three components. B) A flow schematic of a typical labelling process with the heavy and light ICAT reagent chains red and green indicate light and heavy chains respectively. Normal and diseased samples are labelled with light and heavy chains respectively. They are then typically digested and analysed by mass spectrometry. The mass shift is detected and samples can be relatively quantitated. Combined and adapted from (Gygi et al., 1999).	43
Figure 1-13: A schematic of the process of quantitation using iTRAQ™. After collection and solubilisation samples are typically digested and labelled with one of the eight iTRAQ™ reagents. After separation, samples are introduced into the mass spectrometer and identified. Identified proteins labelled with the various iTRAQ™ reagents can then be quantified. Adapted from Applied Bioscience Inc.....	47
Figure 1-14: Schematic of O-18 Labelling by reversible binding. Taken from (Fenselau, 2007).....	50
Figure 1-15: The method of selective reaction monitoring (SRM). The sample is introduced into the mass spectrometer. A specific m/z is selected in Q1 (quadropole 1) and fragment in Q2. The fragments are then selected as a transition ion and collided with the detector. Adapted from (Veenstra, 2006).	55

Figure 1-16: MRM experimental design logic flow chart. A particular protein is targeted. This is done by targeting specific known ions after tryptic digest and scanning for them within a given mass tolerance. Taken from (Kuzyk et al., 2009)	57
Figure 1-17: A schematic of the SELDI-TOF-MS strategy for profiling samples. The samples of interest are washed over SELDI target plates. These plates are effectively modified MALDI plates which can have various surface chemistries. This causes certain peptides to adhere and others to wash off. The sample is introduced into the mass spectrometer for identification and can be compared against each other.	67
Figure 1-18: Typical setup of a Capillary Electrophoresis system. A buffer filled silica fused capillary is immersed in an Electrolyte Buffer Reservoir at either end. Applied across the two reservoirs is a high voltage. As the electrolytes move along the capillary they pass the detector window where the trace data is sent to a PC. The detector is usually a UV detector but others are available such as fluorescence. <i>Diagram reproduced from introduction to CE volume 1 Beckman Coulter, High Wycombe.</i>	68
Figure 1-19: Laminar flow profile. Cross section of a capillary showing the profile of a fluid within a micro tube.	69
Figure 1-20: Electro-osmotic flow and its double layer. S- Stern layer and D- Double layer. This dual layer of ions causes the characteristic of electro-osmotic flow.	70
Figure 1-21: The chemistry of the silanol group in a high ph.	70
Figure 1-22: Electro-osmotic flow profile. Cross-section of a capillary showing the profile of a fluid within a fused silica capillary.	71
Figure 1-23: Schematic of the workflow of mass spectrometry. Introduction of a sample is followed by its ionisation using a technique such as MALDI, and then by the injection of the ions into a mass analyser. The ion is subsequently detected and analysed.	72
Figure 1-24: Main fragmentation paths of peptides. This figure adapted from (Hoffman) .	74
Figure 1-25: Illustration of Matrix Assisted laser desorption ionisation MALDI. A sample is placed in a vacuum and positioned to receive a laser pulse. The laser is fired and ablates the crystals, which have co-crystallised with the sample. The energy is transferred and proton exchanged. A potential difference is applied and the ions are accelerated into the mass analyser.	74
Figure 1-26: Schematic of electrospray ionisation. Sample is introduced through an electrospray needle, usually from the HPLC. The high voltage applied accelerated the ions and solvent. This caused a plume of solvent and ions. Solvent	

evaporation causes the destabilisation causing droplets. The ever decreasing size of the drops coupled with Coulombs repulsion causes an ever decreasing droplet size.....	76
Figure 1-27: Illustration showing the time it takes for ions of different molecular mass to reach the detector plate and thus the m/z can be measured. Taken from (Chaurand, 2005).....	77
Figure 1-28: Schematic illustration of a Time of Flight tube. The ion is accelerated over regions by a given potential difference and enters from the source into a field free region marked d. The mass analyser may be set in linear or reflector mode. It is then detected by the detector plate. The equation denoting how this is calculated is shown in equation 1.....	78
Figure 1-29: Schematic of a typical quadrupole mass analyser. Two opposing DC voltages are applied adjacent to each set of poles. An opposing RF voltage is applied to paired rods to cause stability or instability of ions with a particular m/z adapted from (Hoffman).....	79
Figure 1-30: This shows the 4 major areas of kidney proteomics. The sizes of the circles are relative to the size of the field to which proteomics is currently applied. The greatest area of proteomic involvement in renal research is discovery of biomarkers.	82
Figure 2-1: Position of gel reference markers. The fluorescent reference markers are used to position the plate and ensure accurate spot picking.	117
Figure 2-2: Table showing relative concentration used in saturation dye optimisation experiment. ".....	121
Figure 2-3: Reduction of a disulphide bridge.	123
Figure 2-4: Home screen of DeCyder 2D software version 7. The home screen illustrates the sequence of workflow.	125
Figure 2-5: Order of membrane and gel stacking for western blot membrane protein transfer to nitrocellulose membranes. The proteins will migrate towards the anode transversely from the SDS-PAGE Gel onto the nitrocellulose membrane.	135
Figure 3-1: A) Stabilzor T1 denaturing device (Denator ABDenator™, Sweden) used to rapidly apply heat treatment to tissue samples and B) maintainer cards, used to place the tissue into vacuum and for each storage. Taken from http://www.denator.com	151
Figure 3-2: Schematic of experimental work flow. Brain tissue hemispheres were bisected and either heat treated and snap-frozen or just snap-frozen. Samples were then	

stored at -80°C. They were then taken onto DiGE analysis. MALDI-MSI was performed separately by Dr R. J. A. Goodwin, Research Associate, University of Glasgow and LC-MS using label free quantitation was performed by Miss H. Allingham, Ph.D student, University of Glasgow. All experiments were performed in parallel. The larger proteins considered using DiGE and the smaller markers considered using MALDI-MSI and LC-MS.154

Figure 3-3: Schematic of Analysis work flow. If spots are to be considered stabilised by the treatment, then treated, snap-frozen = 0 minutes and treated = 10 minutes would be expected to have no statistically significant difference. The bulk of analysis was performed in DeCyder software. The analysis was split into three main areas; Profile, Venn and principal component analysis.160

Figure 3-4: Schematic of the possible effect of degradation on protein spots on a 2D gel map. Degradation occurs either at the terminus or in-between or both in a protein chain. This has a consequential effect on how they appear on a gel spot map, and the detected intensity for that gel spot.162

Figure 3-5: Predicted profiles of spot intensities. Treatment stabilisation would expect to be characterised by no change between T= 0, T =10 and S=0 minutes and a change in S = 10. Snap-frozen stabilisation is characterised by a change in T = 10. Red=treated and Blue=snap frozen.163

Figure 3-6: Example of three way Venn diagram. This particular Venn diagram is targeting treated stabilisation by using t-tests to focus on candidate spots.164

Figure 3-7: Schematic of Analysis work flow. If spots are to be considered for stabilised by the treatment, then treated = 0, 10, 20 and snap-frozen = 0 minutes would be expected to have no statistically significant difference. The bulk of analysis was performed in DeCyder software. The analysis was split into three main areas; Profile, Venn and principal component analysis.165

Figure 3-8: Predicted profiles on spot intensities. Showing treated and snap-frozen stabilisation. Red=treated and Blue=snap frozen.167

Figure 3-9: Predicted profiles on spot intensities. This figure displays the different possible profiles exhibited by spots on the DiGE gels. Showing delayed degradation in treated and snap-frozen and stabilisation in treated and snap-frozen with immediate degradation displayed respectively. Red=treated and Blue=snap frozen.168

Figure 3-10: Predicted profiles on spot intensities. Showing profiles of no change and treated and snap-frozen stabilised with other respectively unclassified. Red=treated and Blue=snap frozen.169

- Figure 3-11: Example of three way and four Venn diagrams. The 3-way Venn here is an example of targeting treated stabilisation by using t-tests to focus on candidate spots. The 4-way Venn allows multiple profiles to be targeted simultaneously. ...170
- Figure 3-12: Shows the possible evolution of protein spots when degradation occurs. As degradation occurs over time the original protein spot intensity falls, whilst other increase. If this occurs solely in one experimental group expected profiles are found as above.....171
- Figure 3-13: Schematic showing the effect of changes in molecular weight and pI may have on the location of protein/marker spots on a DiGE gel map. These changes may be due to proteolytic enzymes performing digestion leading to degradation as outlined in section 3.5.2.2.....172
- Figure 3-14: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels treated = 0, 10 and Snap-frozen = 0,10 respectively from pilot investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) 0 minutes and Cy5 (red) 10 minutes. Acidic to basic left to right.177
- Figure 3-15: Showing total amount of spots detected in pilot investigation and total spots included in profile analysis using 1-way ANOVA as a sorting methods.179
- Figure 3-16: Example of the graphical representation of variance. For the denator group time point 0, for the standardised abundance using 683 different spots and n=4. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers. 182
- Figure 3-17: A) Shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 3-5 placed in the relevant categories following profile analysis of pilot investigation. A total of 1072 spots were included. B) Shows the positions of identified spots of examples given in Figure 3-18.....184
- Figure 3-18: Example profiles for pilot investigation. Ai, ii and iii), Bi, ii and iii) and Ci, ii and iii) match predicted profiles given Figure 3-5. The positions on the spot map are given in Figure 3-17-B. All x-axis denoted time points (heat treated=0 min, 10 mins and snap-frozen=0 min, 10min respectively). A) Gives example profiles of heat-treated stabilised. B) Gives examples of snap-frozen stabilized and C) Gives examples of both stabilised or both changed.185
- Figure 3-19: Venn analysis for pilot investigation: A) and B) give two independent but comparative strategies for matching profiles in c) searching for predicted profiles in sector G focusing on treated stabilisation. D) and E) give two independent but comparative strategies for matching profiles in F) searching for predicted profiles in sector G focusing on snap-frozen stabilisation. Profile shown in C) and F) are

those shown in Figure 3-5. Red spots = heat treated and blue spots = snap-frozen.....	190
Figure 3-20: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels treated = 0, 20 and treated = 0,10 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) 0 minutes and Cy5 (red) 10 and 20 minutes respectively. Acidic to basic left to right.	193
Figure 3-21: Set of typical DiGE gels showing A) Fluorescent image with Cy2, 3 and 5 for gels treated = 10, 20 and snap-frozen= 0, 20 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) treated = 10 minutes and snap-frozen = 0 minutes respectively and Cy5 (red) treated and snap-frozen 20 minutes respectively. Acidic to basic left to right.	194
Figure 3-22: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels snap-frozen = 0, 10 and snap-frozen= 20,10 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) snap-frozen = 0 and 20 minutes respectively and Cy5 (red) treated and snap-frozen 10 minutes respectively. Acidic to basic left to right.	195
Figure 3-23: Showing total amount of spots detected in the main investigation and Total spots included in profile analysis using 1-way and 2-way ANOVA as a sorting method.....	196
Figure 3-24: Graphical representation of variance for the denator group time point 0, for the standardised abundance using 62 different spots and n=6. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.	199
Figure 3-25: A) Profile distribution for the main investigation using 1-way ANOVA as a sorting method of categories shown in Figure 3-8, Figure 3-9 and Figure 3-10. A) Shows the proportions of various experimental group categories considered in the predicted profiles then further broken down into proportions assigned to individual profiles. B) The proportion of all stabilised groups compiled for both treated and snap-frozen. This gives an overview of relevant category amalgamation. As can be seen, the treated stabilised proportion is the greatest at 57% while the snap-frozen stabilisation is only 7%. 623 spot were included in total.	202
Figure 3-26: A) Profile distribution for the main investigation using 2-way ANOVA as a sorting method of categories shown in Figure 3-8, Figure 3-9 and Figure 3-10. A) Shows the proportions of various experimental group categories considered in the predicted profiles then further broken down into proportions assigned to individual profiles. B) The proportion of all stabilised groups compiled for both treated and	

snap-frozen. This gives an overview of relevant category amalgamation. As can be seen, the treated stabilised proportion is the greatest at 61% while the snap-frozen stabilisation is only 4%. 573 spot were included in total.205

Figure 3-27: 4-Venn analysis for the main investigation: Shows cross-over of various ANOVA methods used to sort data. The low proportion of spots in common between all 4 methods of sorting is significant as data will be lost if only one is considered.207

Figure 3-28: Example profiles from the main investigation matching predicted profiles given in Figure 3-8 for treated stabilised Ai-vi). Example profiles of actual spots and their location on the gel map below. Blue lines denote snap-frozen and red heat-treated.....209

Figure 3-29: Example profiles from the main investigation matching predicted profiles given in Figure 3-9 and Figure 3-10 for treated stabilised Ai-iv) Example profiles of actual spots and their location on the gel map of treated stabilised with snap-frozen showing immediate degradation. Bi-ii) Example profiles of actual spots and their location on the gel map of treated stabilised with snap-frozen unclassified. Blue lines denote snap-frozen and red heat-treated.214

Figure 3-30: Example profiles from the main investigation matching predicted profiles given in Figure 3-8, Figure 3-9 and Figure 3-10 for snap-frozen stabilised Ai-ii) Example profiles of actual spots and their location on the gel map. Bi) Example profiles of actual spots and their location on the gel map of treated stabilised with treated showing immediate degradation. Ci-ii) Example profiles of actual spots and their location on the gel map of snap-frozen stabilised with treated unclassified. Blue line denotes snap-frozen and red heat treated.....219

Figure 3-31: Example profiles from the main investigation matching predicted profiles given in Figure 3-8, Figure 3-9 and Figure 3-10 for delayed degradation and no changes Ai) Example profiles of actual spots and their location on the gel map for delayed degradation in treated. Bi-ii) Example profiles of actual spots and their location on the gel map of delayed degradation in snap-frozen. Ci-ii) Example profiles of actual spots and their locations on the gel map of spots classified as not changing or have no difference. Blue lines denote snap-frozen and red heat-treated.....224

Figure 3-32: Venn analysis for the main investigation. A) Shows Venn analysis corresponding to red and green boxed profiles in B). Ci-ii-Di-iii) Show profile distribution of protein spots taken from areas G and E of the Venn diagram. The pie charts show how starting particular areas on the Venn diagram allow for the quick discovery of particular profile types. The one targeted here being treatment stabilised spots. This can significantly reduce manually sorting through profiles.230

- Figure 3-33: 4-Venn analysis for the main investigation: A) Red and Green boxes isolate profiles targeted in Sectors in B) relating to the profile distributions shown via pie chart in C) and D). Both Venn diagrams in Figure 3-32 and Figure 3-33 are successful at finding the predicted profiles indicated. The pie charts show how starting particular areas on the Venn diagram allow for the quick discovery of particular profile types. The one targeted here being treatment stabilised spots. This can significantly reduce manually sorting through profiles.....231
- Figure 3-34: Venn analysis for the main investigation. A and B) give two independent but comparative strategies for matching profiles in C) searching for predicted profiles in sector G focusing on treated stabilisation. D) Gives crossover of the two G sectors. Ei-ii) Show profile distribution of protein spots taken from areas G Venn diagram.....233
- Figure 3-35: Venn analysis for the main investigation. A) and B) give two independent but comparative strategies for matching profiles in C), searching for predicted profiles in sector G focusing on snap-frozen stabilisation. D) Gives crossover of the two G sectors234
- Figure 3-36: Venn analysis for the main investigation. Comparison of spot distribution between A), B) and C).....235
- Figure 3-37: Spot maps A) main investigation and B) pilot investigation indicating two example spots that have been defined in pilot investigation as no change and in the main investigation as treated stabilised. The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance.237
- Figure 3-38: Examples and positions of gel map of protein appearing in 1 experimental group only (pilot investigation). Pie chart shows the distribution of these profiles with total number of 39 found. All were in snap-frozen = 0 min and 10 group. The blue lines denote snap-frozen samples.238
- Figure 3-39: Examples and positions of gel map of protein appearing in 1 experimental group only in main investigation. Pie chart shows the distribution of these profiles with total number of 6 found. All were in snap-frozen = 0, 10 and 20 group. Profile number 1 has been identified as 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]. The blue lines denoted snap-frozen samples.239
- Figure 3-40: Gel map showing multiple identifications and spots shifts in pilot investigation of Dihydropyrimidinase-like 2 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.240

Figure 3-41: Gel map showing multiple identifications and spots shifts in pilot investigation of Peroxiredoxin 6 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.....	241
Figure 3-42: Gel map showing multiple identifications and spots shifts in pilot investigation of creatine kinase, brain [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.....	242
Figure 3-43: Gel map showing multiple identifications and spots shifts in pilot investigation of gamma-actin [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.....	242
Figure 3-44: Gel map showing multiple identifications and spots shifts in pilot investigation of alpha-tubulin isotype M-alpha-2 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.....	243
Figure 3-45: Gel map showing multiple identifications and spots shifts in pilot investigation of tubulin, beta [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.....	244
Figure 3-46: Gel map showing multiple identifications and spots shifts in the main investigation of PREDICTED: similar to pORF2 [Mus musculus], with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance.....	244
Figure 3-47: Gel map showing multiple identifications and spots shifts in the main investigation of ATP5b protein or Mitochondrial ATP synthase beta subunit [Mus musculus] with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance	245
Figure 3-48: Gel map showing multiple identifications and spots shifts in the main investigation of 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus] with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance.....	245

Figure 3-49: PCA score plots. A) Shows score plots including all detected spots for pilot investigation and B) Shows score plots including only 1-way ANOVA spots included	248
Figure 3-50: PCA score plots for the main investigation A) PCA score plots including all spots matched B) PCA score plots sorted by One-way ANOVA C) PCA score plots sorted by two-way ANOVA. As can be seen the different experimental groups can be clearly defined, with a greater variation in the snap-frozen samples.....	250
Figure 3-51: An example of one of the three repeats of 1-D Coomassie stain gel and western blot of heat-treated and snap-frozen samples blotted for peroxiredoxin 6. As can be seen heat-treated 0, 10 and 20 minutes and snap-frozen 0 minutes show the presence of a strong band at approximately between 24-32.5KDa and snap-frozen 10 and 20 minutes the band is absent. Note: the PM labels refer only to the Coomassie stain gel.	252
Figure 3-52: Graph displaying relative fold change of Peroxiredoxin 6 generated from Table 3-21. The error bars show +/- 1 standard deviation. n=3.	255
Figure 3-53: A repeat of Figure 3-54: Gel map showing multiple identifications and spots shifts in pilot investigation of Peroxiredoxin 6 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.	255
Figure 3-55: LC-MS using label free quantitation: performed by Miss H. Allingham, Ph.D Student's, University of Glasgow. A) Show the average number of identified features from intact sample results in T=0, 10 and 20 minutes and S=0, 10 and 20 minutes. The number of features across both groups is comparable. B) Give the percentage of features showing a significant change. Ci-ii) shows a profile consistent with the profiles observed in the two DiGE experiments in sections 3.6.3 and 3.6.4.	256
Figure 3-56: LC-MS using label free quantitation: performed by Miss H. Allingham, Ph.D student, University of Glasgow. A) Show the average number of identified features from digested sample results in T=0, 10 and 20 minutes and S=0, 10 and 20 minutes. The number of features across both groups is comparable. B) Give the percentage of features showing a significant change. Ci-ii) Shows a profile consistent with the profiles observed in the two DiGE experiments in sections 3.6.3 and 3.6.4.	257
Figure 3-57: MSI experiment, Performed by Dr R. J. A. Goodwin, Research Assistant, University of Glasgow. A) Treated (heat-treated) and snap-frozen (non-heat-treated) sections at 0 minutes for an example marker. B) Treated (heat-treated) and snap-frozen (non-heat-treated) sections at 20 and 10 minutes respectively for the same example marker in (A). C) Shows time course series of sections warmed	

consecutively from 0-5 minutes for treated and D) Snap-frozen tissue have been duplicated. Difficulties in cutting treated samples precluded duplicates.....	258
Figure 3-58: MSI experiment, Performed by Dr R. J. A. Goodwin, Research Assistant, and University of Glasgow. A-G) Shows snap frozen and treated sections typically digested and non-digested for different marker masses. F) Shows the optical images of those slices	260
Figure 4-1: Summary of the Renin Angiotensin Aldosterone System and the use of ACE inhibitors. The system allows for control of blood pressure by measuring increase and decrease glomerular blood flow. This is one of the main mechanisms for controlling blood pressure in mammals. This figure was reproduced from (Brewster and Perazella, 2004).	269
Figure 4-2: Chromosome 2 genetic map in SP.WKYGla2a, SP.WKYGla2k, and SP.WKYGla2c* strains. This shows the genetic makeup of the congenic strains in relation to Wky and SHRSP. Showing the congenic strain SP.WKYGla2k employed in this study, has a largely SHRSP genetic makeup. Taken from (Graham D, 2007a).....	273
Figure 4-3: Schematic of genetic model employed. A congenic strain has been developed with a SHRSP background with sections of Chromosome 2 from WKY. Figure provided by Dr Martin McBride, Cardiovascular research unit.....	278
Figure 4-4: Systolic blood pressure phenotyping using radiotelemetry. Systolic blood pressure was significantly increased in the salt-loaded SHRSP compared to salt-loaded SP.WKYGla2a and WKY rats. Figure provided by Dr Martin McBride, Cardiovascular research unit.	278
Figure 4-5: Sample work flow for the pilot investigation replicates. Showing that both cortex and medulla tissue was used in addition to salt and non-salt treatment.	279
Figure 4-6: Sample workflow for main investigation. Showing that only salt treatment samples were used but tissue segregation into cortex and medulla regions was still employed.....	279
Figure 4-7: Schematic of section sequence. Kidney tissue was mounted (A) after measurements of the dimensions were taken (B and C). Sections were taken at regular intervals. (D) Blue: ITO slides for MSI (collaborator), Green: Glass slides for macro dissecting for DiGE minimal labelling, Red: Glass slides for Saturation labelling, Laser Micro dissection slides and Glass slices for IHC (future validation of MSI). The first 100 slices taken into Eppendorfs, 6 for MSI then up to 200 slices into Eppendorfs (¼ depth). Then 6 sections for MSI, 74 sections for DiGE, LMD and IHC. Up to 300 th section into Eppendorfs. Following this pattern till ½ depth. Then a further 40 sections for DiGE.	280

- Figure 4-8: Optical image of kidney section defining regions of macro-dissection. The region within the red dotted line gives the “medulla” and blue dotted line “cortex.” It is recognised the undulations and convolutions mean regions are difficult to assign, while the definition visually is distinct.281
- Figure 4-9: Visible regions of the kidney with the eye: These regions of the kidney are noticeable as having a colour difference and show the dynamic and complex tissue of the kidney organ. This is only as eye level and the complexity deepen when using light microscopes. Therefore there intrinsically and intuitively is a need to try and separate tissue types in order to fully understand their functional significance.282
- Figure 4-10: Schematic of experimental work flow. Kidney tissue was processed for DiGE analysis. MALDI-MSI was performed separately by Dr R. J. A. Goodwin, Research Associate, University of Glasgow. The larger proteins considered using DiGE and the smaller markers considered using MALDI-MSI.284
- Figure 4-11: Schematic for the assessment of macro dissection. Kidney tissue was sliced and placed in eppendorfs and on glass slides. The area of the tissue was calculated using image j. An assessment of protein per unit area was made using a Bradford assay with a BSA standard curve of whole slices from eppendorfs, whole dissection from a glass slide and macro dissection of medulla and cortex from a glass slide. Results are shown in 4.6.4.1 on page 346.286
- Figure 4-12: Predicted Profiles of spot intensities. Shows the predicted relationship between WKY, congenic and SHRSP strains. A i-ii) Shows possible intermediate effect in the congenic strain compared to WKY and SHRSP (Intermediate effect i-ii). B i-ii) No change between the WKY and Congenic strains (WKY maintained i-ii). C i-ii) No change between the congenic and SHRSP strains (SHRSP maintained i-ii). D i-ii) Congenic strains presenting possible different spot intensity in comparison to WKY and SHRSP strains (no change in terms of WKY and SHRSP). Profiles can be overlaid to compare two different tissue types or salt treatment.292
- Figure 4-13: Targeted Venn analysis. Different crossover regions of the Venn diagram can be used to mine data by correlating them with predicted profiles. Additionally, statistical tests have already been run and cross checking is not required.293
- Figure 4-14: Example of targeted Venn analysis to illicit profiles showing differences and similarities in tissue types. C shows the changes which cortex and medulla have in common.293
- Figure 4-15: Example of targeted Venn analysis to illicit profiles showing differences and similarities between treatment with salt and no salt treatment. C) Shows any similarities between treatment of salt and non-salt.294

Figure 4-16: Set of typical DiGE gels showing. This particular gel shows A) Fluorescent image with Cy2,3 and 5 for gels Cortex WKY Salt treated and Medulla WKY Salt treated respectively from the pilot investigation B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) Cortex WKY Salt treated (red) Medulla WKY Salt treated. Acidic to basic left to right.297

Figure 4-17: Graphical representation of variance for the Congenic Cortex non-salt treated group, for the standardised abundance using 67 different spots and n=3. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.....301

Figure 4-18: Shows profile analysis for Cortex tissue and non-salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12, placed in the relevant categories following profile analysis of pilot investigation. A total of 355 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.....303

Figure 4-19: Gel map of example spot given in Figure 4-18 for Cortex non-salt treated. .304

Figure 4-20: Shows profile analysis for Cortex tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of pilot investigation. A total of 359 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.307

Figure 4-21: Gel map of example spot given in Figure 4-20 for Cortex salt treated.308

Figure 4-22: Shows profile analysis for medulla tissue and non-salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of pilot investigation. A total of 238 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.310

Figure 4-23: Gel map of example spot given in Figure 4-22 for medulla non-salt treated311

Figure 4-24: Shows profile analysis for medulla tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of pilot investigation. A total of 219 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.313

Figure 4-25: Gel map of example spot given in Figure 4-24 for medulla salt treated.....314

Figure 4-26: Comparison of salt and non-salt treatment in both cortex and medulla within the same strain. Pie charts show the direction of change.316

Figure 4-27: Comparison of tissue types in both salt and non-salt treatment within the same strain. Pie charts show the proportion and direction of change.318

Figure 4-28: Venn analysis showing spots in common between salt/non-salt treatment across tissue within strains. The proportion and direction of the change is shown by pie charts. It can be seen that the crossover of markers between tissue types is small.....319

Figure 4-29: Venn analysis comparing cortex and medulla when p-value ≤ 0.05 . The proportion of change in each sector is shown by pie charts. Salt and non-salt are compared separately.....321

Figure 4-30: Example markers from the pilot study where salt treatments lead to a change in profile shape. A)-C) show how profiles have been changed due to salt treatment Blue line shows non-salt treated samples and red line shows salt treated samples. 1, 2 and 3 on the x-axis are WKY, congenic and SHRSP respectively.323

Figure 4-31: Set of typical DiGE gels showing. This particular gel shows A) Fluorescent image with Cy2,3 and 5 for gels Cortex SHRSP Salt treated and Medulla SHRSP Salt treated respectively from the Main investigation B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) Cortex SHRSP Non-Salt treated (red) Medulla SHRSP Non-Salt treated. Acidic to basic left to right.326

Figure 4-32: Graphical representation of variance for the Congenic medulla salt treated group, for the standardised abundance using 522 different spots and n=3. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.....329

Figure 4-33: Shows profile analysis for cortex tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12 placed in the relevant categories following profile analysis of main investigation. A total of 100 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.331

Figure 4-34: Gel map of example spot given in Figure 4-33 for cortex salt treated. Proteins identified are; 1550. rCG25777, isoforms CRA_a / aminoacylase 1. 2058. Regucalcin. 2187. Mercaptopyruvate sulfurtransferase and 2620. Isoamyl acetate-hydrolyzing esterase 1 homolog.332

Figure 4-35: Shows profile analysis for medulla tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of main investigation. A total of 174 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.333

Figure 4-36: Gel map of example spot given in Figure 4-35 for medulla salt treated. Identification of spots are; 644. dnaK-type molecular chaperone hsp72-ps1 / Heat shock protein 8. 2709. 14-3-3 zeta isoform / typtohan 5-monooxygenase activation protein. 3169. Mitochondrial ribosomal protein L51.....334

Figure 4-37: Comparison of tissue types within the same strain. Pie charts show both the proportion of the direction of change.335

Figure 4-38: Venn analysis comparing cortex and medulla and shown strains when p-value = <0.05. The proportion of change in each sector is shown by pie charts.....336

Figure 4-39: Combination of Venn and profile analysis in mining down of data between tissue type. Example and of candidate marker profile is given for sector F that match predicted profile in Figure 4-12. Gel map shows the position and id of possible marker. Identified candidate marker is Predicted: Similar to Actin Cytoplasmic 2 (gamma Actin).338

Figure 4-40: Combination of Venn and profile analysis in mining down of data. Example and candidate marker profiles are given for each sector that matches predicted profiles in Figure 4-12. Venn analysis is with cortex tissue however medulla tissue profiles have been overlaid for comparison.340

Figure 4-41: Gel map of example candidate markers given in Figure 4-40.341

Figure 4-42: Combination of Venn and profile analysis in mining down of data. Example and candidate marker profiles are given for each sector that matches predicted profiles in Figure 4-12. Venn analysis is with medulla tissue, however cortex tissue profiles have been overlaid for comparison.342

Figure 4-43: Gel map of example candidate markers given in Figure 4-42. The two identified markers are; 2635. Uracil-DNA Glycosylase, isoforms CRA_a and 2698. Hypothetical Protein LOC619574343

Figure 4-44: Venn analysis for identified spots only. As can be seen there is little crossover of identified candidate markers between tissue types as the master spot numbers are different (with the exception of one). This helps to indicate the need for segregation of tissue types.....344

Figure 4-45: Venn diagrams comparing strains using an alternative strategy. An assumption of no significant difference is used between WKY and Congenic strains in order to target Sector (G) as being the intensity profile shown. A) Cortex and B) medulla. A possible strategy for increasing the number of candidate markers discovered	345
Figure 4-46: Validation of macro-dissection A) shows average protein content per 14um slice of tissue from a WKY rat. n=6 by comparing macro dissection cortex and medulla (green), a whole slice macro dissected (red) and a slice placed directly into lysis buffer from an eppendorf tube B) Shows the protein content per given area from a WKY rat. Error bars in both A) and B) are of two standard errors. n=6 for both A) and B).....	346
Figure 4-47: Variation of protein concentration from Kidney tissue from LMD tissue discs. As can be seen the variation is considerable but still overlapping. The error bars displayed are two standard errors. There were 6 slices in total and they are compared against each other individually and not averaged.	347
Figure 4-48: PCA score plot containing all 5777 protein spot in the pilot investigation for Cortex Tissue	348
Figure 4-49: PCA score plot of all one way anova spots 802 protein spot in the pilot investigation for Cortex Tissue.....	349
Figure 4-50: PCA score plot of all 5777 protein spot in the pilot investigation for Medulla Tissue	350
Figure 4-51: PCA score plot all one way anova spots 802 protein spot in the pilot investigation for Medulla Tissue.	351
Figure 4-52: PCA score plot of the main investigation. All 3174 spots included	352
Figure 4-53: PCA score plot of main investigation. Only spots with one-way anova $p < 0.05$ considered. No of spots = 189.	353
Figure 4-54: figure showing matches between micro array data and proteomic data. Red indicates upward expression in SHRSP compared to congenic and grey indicates downward expression. The highlight number is the fold change in expression. ...	354
Figure 4-55: Profiles of spot intensities in the main investigation, which match with RNA micro array data from cardiovascular research unit.	356
Figure 4-56: Gel map giving position for possible candidate markers from Figure 4-55....	356
Figure 4-57: Comparison of tissue types and matches with microarray expression data.	358

Figure 4-58: MALDI-mass spectrometry imaging (MSI) in non-salt and salt treated WKY (WT), Congenic (Con) and SHRSP strains. A) This is composite images showing all the mass filters giving an overview of the distribution of markers across the three strains. B) Mass filter of (2007 m/z) with heat-map display showing regions of highest intensity. Mass distribution is even across all kidneys. C) Mass filter of (2039 m/z) Mass seen to vary significantly between the SHRSP kidney and that of wild type and congenic Salt treatment affects the intensities across the kidneys. D) ms/ms identification of Histone H1 direct from tissue for 2039 m/z. Figure provided by Dr Richard Goodwin, Research Associate, Glasgow University.	359
Figure 4-59: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Show the average number of identified features from intact sample results across all groups	360
Figure 4-60: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Expression profile from quantitative western blot for Fibrinogen matching profiles from the label free quantitation LC-MS.....	360
Figure 4-61: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Expression profile from quantitative western blot for Hemopexi matching profiles from the label free quantitation LC-MS.....	361
Figure 4-62: Positions and profiles of multiple identifications of protein Aminoacylase 1, from the main investigation. . The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance...364	364
Figure 4-63: Positions and profiles of multiple identifications of protein Eno1, from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.....	365
Figure 4-64: Positions and profiles of multiple identifications of ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.	367
Figure 4-65: Positions and profiles of 3 identified heat shock proteins, from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.	368
Figure 5-1: Set of graphs showing the relative expression of Wky, Congenic and SHRSP strains. Columns are an expression of the ratio of themselves against Wky stain.	

It is proposed that these identifications might serve to aid future work into looking for a profile pattern from 2D-DiGE gels. These are for Cortex tissue.	381
Figure 5-2: Set of graphs showing the relative expression of Wky, Congenic and SHRSP strains. Columns are an expression of the ratio of themselves against Wky stain. It is proposed that these identifications might serve to aid future work into looking for a profile pattern from 2D-DiGE gels. These are for medulla tissue.	386
Figure 5-3: Figure showing Dye concentration determination for DiGE saturation labelling (labelling for scares samples). All images where scanned using GE Healthcare Typhoon Scanner at the same scan settings. Fluorescent and traditional gel displays are shown.	391
Figure 5-4: Figure showing gel representation of the effect of DiGE lysis buffer on the resolution of Amershams Bioscience, Bioanalyser Chip reader. If true miniaturisation and quantitation is to be achieved a reliable form of protein concentration determination is required for small samples. The lysis Buffer used in DiGE labelling contains a number of components and is not compatible with the Bioanalyser Chip reader, as the concentration of Chaps rises the resolution is lost making calibration curves too inaccurate for protein concentration determination. Lane 1. Protein marker, 2-7 is BSA standard in water. 8-10 wild type mouse brain samples in solution labelled above.	393
Figure 5-5: Laser Micro-Dissection. The effect of LMD slide plastic on downstream mass spectrometry. A) Discs of Tissue adhered to LMD slide Polyethylene film, and spotted in various concentrations of matrix on an ITO slide ready for MALDI-MS. 50/50 ACN Solution B) Mass spectra at various laser intensities. As can be seen resolution is very poor. C) Mass spectra of tissue after it was disturbed using Sterile Gilson tip.	394

vi. Abbreviations

1D	One Dimensional
2D	Two Dimensional
2DE/2DGE	Two Dimensional Gel Electrophoresis
2DLC	Two Dimensional Liquid Chromatography
3D	Three Dimensional
ACN	Acetonitrile
ANOVA	Analysis of Variance
BSA	Bovine Serum Albumin
CHAPS	3-[3-Cholamidopropyl]dimethylammonio]-1-propanesulfonate
CHCA	Alpha-Cyano-4-hydroxycinnamic acid
CID	Collision Induced Dissociation
CSF	Cerebral Spinal Fluid
ddH ₂ O	deionised and Distilled Water
DiGE	Differential Gel Electrophoresis
DMF	Dimethyl Formamide
DNA	Deoxyribo Nucleic Acid
ESI	Electrospray Ionisation
FA	Formic Acid
H ₂ O	Water
HPLC	High Performance Liquid Chromatography
ICAT	Isotope Coded Affinity Tag
IEF	Isoelectric Focusing
IMS/MSI	Imaging Mass Spectrometry
IPG	Immobilised pH Gradient
ITO	Indium Tin Oxide
iTRAQ TM	Isotope Tagging for Relative and Absolute Quantitation
LC	Liquid Chromatography

LMD	Laser Micro-dissection
m/z	Mass to charge ratio
MALDI	Matrix Assisted Laser Desorption Ionisation
MOWSE	Molecular Weight Search
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MuDPIT	Multidimensional Protein Identification Technique
NCBI	National Center for Biotechnology Information
PAGE	Poly-Acrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PET	Polyethylene Terephthalate
PMF	Peptide Mass Fingerprint
PTM	Post-Translational Modification
RF	Radio Frequency
RNA	Ribonucleic Acid
RP	Reversed Phase
SA	Sinapinic Acid
SDS	Sodium Dodecyl Sulphate
SHRSP	spontaneously hypertensive rat stroke prone
SILAC	Stable Isotope Labelling by Amino Acids in Cell Culture
TCA	Trichloroacetic Acid
TFA	Trifluoroacetic Acid
TOF	Time of Flight
UV	Ultraviolet
WKY	Wistar Kyoto rat

1 Introduction

“Because technology provides the tools and biology the problems, the two should enjoy a happy marriage”

(Fields, 2001): 10051

In the evolution of scientific thinking, science has come from the advent of modern thought with the Greek philosophers to the highly rigorous experimental science of today. Human kind has amassed a staggering multitude of theories which have been pieced together over generations and ordered into our understanding of the world today. One defining characteristic of human beings is their self-awareness and the extent to which they will endeavour to understand not only the environment in which they live but also their internal environment.

In terms of achievement in the biological sciences, the human genome project could be considered one of the pinnacles of scientific discovery and a testament to collaboration. Since the penultimate completion in 2001 (Venter, 2001., Mundy, 2001) and finally in 2003, around 50 years after the discovery of the double helix (Watson, 1953.), we have seen the explosion of the post genomic era and Pandora’s box is deeper and broader than anybody could have imagined. The information obtained has allowed us to see the coding sequences and the blue print for the functional entities; proteins. The focus, therefore, has now shifted from the relatively simple genome, and characterising genotype, to the more complex proteome, and examining phenotype (displayed in Figure 1-1). As the later stages of gene sequencing drew to a close, it became clear the once thought mantra of; one gene-one protein-one function was incorrect. Kolch et al has predicted that there could be anything up to 1,000,000+ functional entities at the protein level, compared to about 40,000 genes (Kolch, 2005).

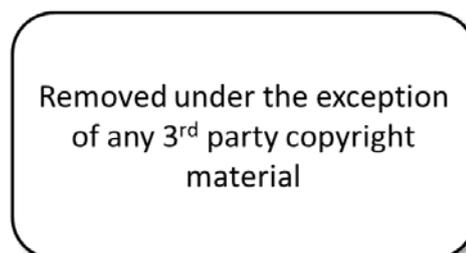


Figure 1-1: Genome vs. Proteome. The Peacock butterfly (*Inachis lo*) emerges from the caterpillar. Both have the same genome but have markedly different proteomes, highlighting the dynamic and complex interactions of proteins and changes in phenotype.

1.1 The study of proteomics

We have now emerged from the post-genomic era which has spawned a collection of -omic fields. One such -omic being proteomics and rather than genomics capturing all the limelight, the focus has expanded into decoding proteins and how they interact (Hochstrasser, 1998). The term proteome was defined some 16 years ago by Marc Wilkins as “all proteins expressed by a genome or tissue.”(Wilkins et al., 1997) Thus proteomics is the study of proteins and their interactions within a cell, tissue, organ, system or organism.

As there has been progression through history, man has developed more and more elaborate technologies, which have allowed us to delve deeper and reveal what makes us work. The technological breakthrough enjoyed by genomics research after the development of polymerase chain reaction (PCR) by Kary Mullis (who was awarded the Nobel prize in 1993) (Saiki, 1985.) has not been repeated in proteomics. This leaves the reality of having to work with the material that can be extracted from cells or tissue directly.

This comparatively slow progression in proteomics compared to genomics is largely due to the relative complexity of the proteome. Proteins, by their very nature, have a large dynamic range and engage in complex interactions, form, function and properties (Wu L, 2006, Garbis et al., 2005). They also exhibit a number of post translational modifications, such as phosphorylation, glycosylation and ubiquitination (Spickett, 2006). In addition, there are limitations regarding the bioinformatics analysis (Cristoni, 2004) required and the ever increasing need to combat and improve quantitation (Hamdan and Righetti, 2003) is palpable. This makes proteins intrinsically more difficult to investigate and the issue of balancing high throughput with quality of data is imperative(Wilkins, 2009).

However, with the increasing need to study protein interactions and phenotypical representation of the genome, there has not been an accompanied progression in the advances in technology and methods at the rate required and therefore, in many respects, proteomics techniques all have their limitations (Beranova-Giorgianni, 2003, Garbis et al., 2005). The problems and limitations are present at every point in a typical proteomics work flow (Figure 1-2). In sample preparation, often samples are of low quality or abundance with a high dynamic range (Garbis et al., 2005). In the analysis of blood samples albumin is of high abundance and often masks any subtle changes in

other protein levels, often making biomarker discovery difficult where one or two proteins dominate. Some proteins have very short life spans, complicating analysis even further (Bachmair, 1986). This heterogeneity in samples makes separation essential. The variety in solubility, pH and polarity of different proteins make standardization of methodology impractical. The issues and challenges surrounding proteomics are reviewed in a number of articles (Kavallaris, 2005, Hong, 2006, Zhou, 2005).

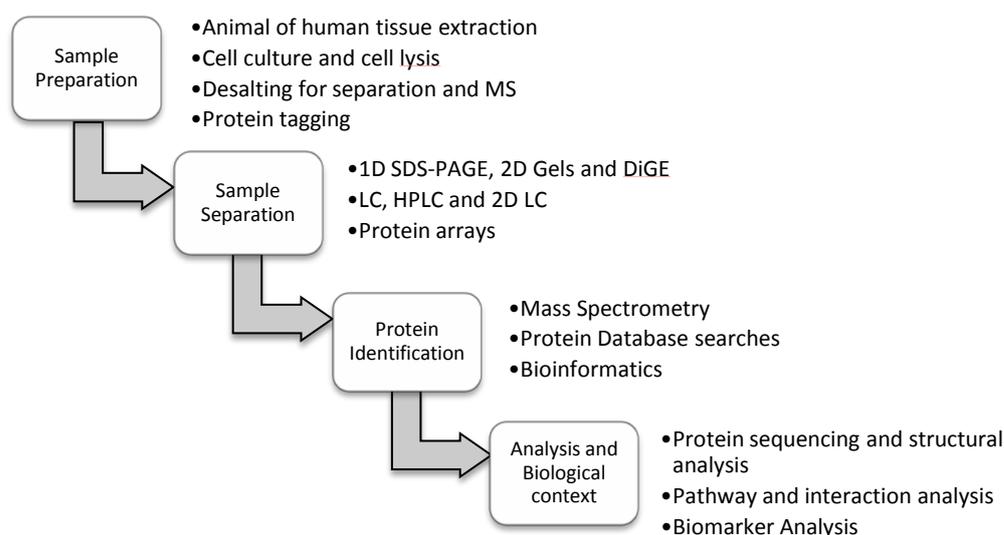


Figure 1-2: Typical Proteomics Work flow. The first stage is sample preparation. Often samples must be desalted ready for protein separation and detection. After the sample has been separated it is then taken through to identification and analysis.

As outlined above, proteomics is a field of importance and limitations. Interestingly and importantly, the development of proteomics potentially would allow for the marrying of proteomic and genomic data, genotype to phenotype, and subsequently to pathology. Importantly, proteomic analysis can give information about the level of protein expression as opposed to gauging the level of gene expression, which does not necessarily gain information about protein expression. As proteins are the direct functional entities of a cell, distinct from DNA being the blue print, they give us a greater insight into what cellular processes are actually occurring and allow the study of an organism's phenotype. However, this huge explosion of complexity and heterogeneity presents major problems with regards to deconvolution of information.

It is from the heterogeneity of proteins and their limitation that allows for the aims of this thesis to be explored.

1.2 Aim and overview of this thesis

It is from the limitations outlined above that bare the need to reduce the impact of those impediments. In this thesis two important areas in tissue proteomics are to be considered; degradation and biomarker discovery.

1.2.1 Protein degradation in mouse brain tissue and DiGE.

Perhaps the most important stage in the proteomics work flow is that of sample preparation. The need for sample integrity from the beginning is vital to give a representative view of data downstream. This is often an area that is overlooked. A method for the reduction of protein degradation and stability of the proteome is considered in Chapter 3 and reviewed in section 3.3.

1.2.2 Putative biomarker discovery in hypertension and moving towards linking genomic and proteomic data.

1.2.3 Two areas of increasing interest are that of biomarker discover and also linking genomic and proteomic data. The impact of markers for diagnosis, tracking and treating disease is substantial and the ability to link phenotype and genotype is required for development, collaboration and pursuit. Biomarker discovery is reviewed in section 1.3, renal proteomics in 1.6 and Hypertension will be reviewed in Chapter 4.

Over the coming introduction chapter I will introduce the various areas and techniques that will be applicable to attaining the aims outlined in section 1.8 and in my experimental chapters.

Chapter 2 outlines the general methods and materials employed in carrying out the investigations in the experimental portion of this thesis. Any methods which are particular to a chapter are further elaborated in the applicable section of the experimental chapter.

There are two experimental chapters in this thesis

Chapter 3: Differential in gel electrophoresis analysis of wild type mouse brain tissue: A multi-faceted analytical approach to assessing the effect of heat treatment on the degradation of the proteome.

This chapter investigates the use of heat treatment as a mode for the reduction or cessation of proteomic degradation in comparison to the widely used method of snap-freezing tissue samples. The tissue used was wild type mouse brain.

In addition, a multi-faceted approach has been used to evaluate degradation and to try and develop a method for mining Differential in Gel Electrophoresis data.

Chapter 4: Accessing the proteomic profiles of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model in the search for candidate markers.

This chapter uses the multifaceted analytical approaches developed in the previous chapter to mine DiGE data for candidate markers for hypertension. It is an investigation that has the potential to link genomic and proteomic data together at a later date.

1.3 Proteomics and biomarker discovery

Biotechnology and proteomics has emerged from the post-genomic era and puristically has the goal of furthering the understanding of the complexities of the cellular world and the interactions and processes of life. Practically, however, this understanding is largely required to inform clinicians about the presence or state of a disease. A primary role of the clinician is to diagnose a patient's symptoms. Ideally this would be done quickly and as non-invasively as possible (Tyther et al., 2009, Petricoin et al., 2002, Calvo et al., 2005). Early diagnosis is important, so that a suitable treatment can be selected and the onset of a disease state can be halted. Depending on the type of disease, early detection can mean the differences between a successful or fatal outcome. According to CRUK (Cancer Research UK) the percentage survival rate for stage 1A non-small cell carcinoma is between 58-73% compared to 2-13% with stage 4. This highlights the need for early and specific diagnoses.

Since certain diseases are more treatable if diagnosed early, it is of great importance that new diagnostic methods are developed which are quick and easy to administer and time-effective to process, such efficiency and accuracy can be seen in a home pregnancy test. Therefore the clinicians would regard an ideal diagnostic test as one that could be administered simply with as little invasiveness as possible and that gives quick results without the need for further lab work, all performed at the bedside, with a high degree of specificity and accuracy (i.e. how good a test is at differentiating false positives) (Vblokeswar and Soloway, 2001, Wu et al., 2002). In order for this to become a reality, it is necessary to identify a molecule or collection of molecules which would be unique to a disease state or disease progression or otherwise so called biomarkers.

A biomarker can be defined as cellular, biochemical, molecular, or genetic alterations by which a normal, abnormal, or biologic process can be recognized or monitored. This could be a protein, peptide, protein fragment, post translational modification or combination of peptides, which indicate the onset or state of a disease (Wu, 2002) (Jones MB, 2002, Craven et al., 2013, Zhang et al., 2010, Matt et al., 2008). In a state of disease a biomarker could give an indication of not only the presence of a certain disease but the stage of its progression or the therapeutic response to a treatment Thus allowing doctors to give better treatments faster with a greater ability to predict prognosis. Ideally,

a biomarker would be present in a disease state and absent in a healthy patient. However, as cellular interactions are complicated, at best, this is rarely the case. Due to the complex nature of cellular interaction coupled with degradative process, the biomarkers of a disease are likely to be equally complex in nature. This could be in the form of multiple abundance changes both up and down regulation or a pattern of biomolecules rather than simply an absent or presence (Issaq et al., 2011, Veenstra, 2006). Therefore, quantitation becomes essential with as high as specificity (how good a test is at differentiating false positives) and resolution as possible.

The obvious place to start when looking for biomarkers is in fluids from around the body which includes; urine, saliva, blood plasma, blood, nipple aspirant fluid (NAF) and cerebral spinal fluid (CSF) (Huck, 2006, Plebani, 2005). Proteomics offers analytical solutions to these fluids due to the lack of a genomic element. Body fluids are of course advantageous in nature due to their ease of extraction from the body (animal or human) and their plentiful supply. Some body fluids are more complex in nature, although each has its own challenges; therefore different proteomics work flows are required. The alternative to biofluids is the use of cell lysate or tissue samples, which have the advantage of having a closer proximity to the site of disease.

Traditionally, biomarker studies have focused on those which are present in body fluids. However the use of tissue is gaining significant ground due to the relatively slow progress of biofluid proteomics. Using tissue would allow for the tracking of biomarker evolution from source to fluid and help the understanding of how a disease evolves over time.

In proceeding sections of this literature review the discovery of biomarkers in fluids and tissues will be examined, comparing the relative advantages, limitations and gaps in the literature. Additionally, the biomarker discovery “pipeline” will be considered and whether or not biomarker discovery has progressed substantially since its inception. Finally, the importance of biomarker discovery using differential in gel electrophoresis is assessed alongside other quantitation methods.

1.3.1 Biomarker discovery in fluids

The concept of searching for biomolecules in fluids is far from new; glucose has been considered an indicator of diabetes for hundreds of years since diagnoses were attempted by “water tasters” in the efforts to detect sweetness (King and Rubin, 2003). Thankfully, these days, diabetes can relatively easily be checked for by using a glucose tolerance test (Stern et al., 2002). Due to their ease of collection and abundance, body fluids have long been an attractive prospect for gaining a “point of time” view of the state of the body’s health (Thomas et al., 2010, Afkarian et al., 2010, Kim et al., 2009, Veenstra et al., 2005, Ploussard and De La Taille, 2010). In terms of proteomic research, urine is appealing as it is less invasive to collect than blood, requires less skilled labour to collect and extract and is less complicated due to the much lower protein content than blood or CSF (Humpel, 2011, Huzarewich et al., 2010, Mischak and Schanstra, 2011). Also, due to the direct relationship physiologically to the kidneys, urine is of prime interest to the research of kidney proteomics. Of course the benefits of using urine are also paradoxically its limitation. The simplicity and low number of proteins in the voided urine reduce the chances of seeing disease specific markers at all. The kidneys act as a barrier by filtering out proteins and peptides reducing the likelihood of successfully “fishing” for markers. Equally, a great deal of the proteolytic activity has already taken place in the bladder. While this leads to a more stable sample, it also means that a lot of small peptide fragments are present which may be harder to disseminate in downstream analysis. Additionally, due to the fluctuations of fluid intake, the concentrations of the proteins present varies meaning quantitation becomes less straight-forward. This could be controlled for in laboratory animals, but not as easily in human studies and clinical samples. Therefore, the robustness necessary for a “real world” environment may not be attainable. Positively, from a procedural perspective, urine is a relatively stable fluid at room temperature making lab protocols easier for researchers and at less risk of changing the sample due to bad practice. Some of the most successful biomarker discoveries are those of single proteins that identify a disease such as the prostate specific antigen (PSA) as an indicator for prostate cancer (Petricoin III et al., 2002, Partin et al., 1996, Benson et al., 1992, Adam et al., 2002). However, it should be noted that even this is not a definitive predictor of prostate cancer. PSA is not a

highly specific biomarker of prostate cancer and often leads to unnecessary exploratory biopsies (Srinivas et al., 2002, Goo and Goodlett, 2010)

Urine is commonly used as a sample substance in diseases associated with the renal system. The use of urine as a biofluid for biomarker discovery in kidney disease is discussed in section 1.6.1.2.4, however urine is also used in the search of biomarkers in numerous other diseases, which are not directly associated with the kidney.

A number of proteomic technologies are being applied to the field of ovarian cancer and using urine as the biofluid to look for specific markers. 2D-DiGE was employed and discovered afamin as a putative marker (Jackson et al., 2007) and the use of SELDI-TOF-MS of urine uncovered the up regulation of three markers; fibrinogen alpha fragment, collagen alpha 1 (III) fragment and fibrinogen beta NT fragment in human samples from ovarian cancer sufferers (Petri et al., 2009).

Although urine is the most easily accessible fluid, the most common body fluid used in proteomic experiments is plasma or serum. It is often used in studies relating to cancer (Taguchi et al., 2011, Pan et al., 2011, Piersma et al., 2010, Anderson, 2010, Honda et al., 2005), the cardiovascular system (Whiteaker et al., 2011, Májek et al., 2011, Gerszten et al., 2011, Addona et al., 2011, Vaisar et al., 2010, Loo et al., 2010, Dardé et al., 2010) and along with CSF for neurological studies (Polman et al., 2011, McKhann et al., 2011, de Souza et al., 2011, Rinne et al., 2010, Cedazo-Minguez and Winblad, 2010, Blennow et al., 2010). Plasma is more complex than serum due to its additional components, so a balance between simplicity and losing markers must be met. The majority of journal articles in this field cite the use of plasma, particularly for unguided proteomics, where gaining a wealth of data is considered of importance and analytical techniques can be used to mine down into the data. It should be noted that statistical techniques can be used to mine this data, but data sets of such high magnitude tend to present challenge and limitation, these are discussed in section 1.7.

One major issue associated with using plasma in biomarker discovery is the presence of high abundance proteins, such as albumin. This constitutes over half of the total blood proteins and consequently masks possible biomarkers which can often constitute small peptide fragments in low abundance (Diamandis and van der Merwe, 2005). This is indeed a universal problem with

regards to biomarker discover but is particularly the case for plasma. In efforts to minimise this “masking” effect of these abundant proteins, it is possible to use a technique where these proteins are depleted using a depletion column connected to an LC (Plavina et al., 2006). This has spurred the necessity for various separation, depletion, enrichment, and quantitative efforts, which has seen some significant improvements but yet still remain one of the largest challenges in biomarker discovery, perhaps only second to dealing with massive data sets. In addition to the high abundance and quantity of some plasma proteins, it also contains a great degree of dynamic range (estimated in the range of 10^{10}) (Jacobs et al., 2005).

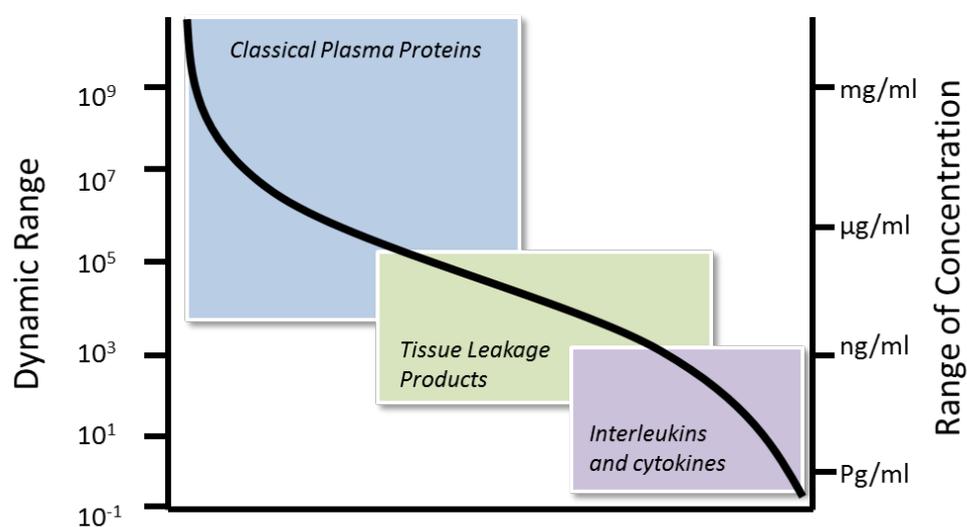


Figure 1-3: The dynamic range of the plasmaome. This graph demonstrates the massive dynamic range of the plasmaome, from 10^{10} down to 10^{-1} which presents a problem for proteomic analysis of biomarkers. The blue area represents the most abundant plasma proteins such as albumin and various coagulation factors. The green area represents tissue leakage products such as cytokeratin and gastrin. Finally, the purple section represents interleukins and cytokines, such as calcitonin and somatostatin. Adapted from Schiess et al, 2008.

This dynamic range means the use of one technique to picture all is unlikely to be achievable. The dynamic range of plasma is illustrated in Figure 1-3

The use of plasma over serum is a continuous discussion. Serum is essentially plasma with clotting factors removed. It is common practice to halt clotting of plasma at collection by using heparin which prevents the cleavage of fibrinogen. It has been shown that degradation can be problematic when using heparin as intrinsic coagulation can occur in its presence, up stream of thrombin in the coagulation process, causing the activation of proteases. (Koomen et al., 2005). The problems with degradation has great potential to

hinder proteomics from successfully discovering biomarkers. This is reviewed in section 3.3.

Serum and plasma have also been employed in the search of ovarian cancer biomarker. Scholler et al, 2008 used a combination of 2DE and LC-MS/MS to analyse serum and discovered the up regulation of catabolic fragments of complement factors, Von Willebrand factor, PEBP1 (RKIP) and EMILIN2 (Scholler et al., 2008). Serum biomarkers have been described in a number of studies in bladder cancer, breast cancer and pancreatic cancer to mention a few (Li et al., 2002b, Rosty et al., 2002, Zhang et al., 2004b).

With regards to CSF as a viable biomarker biofluid. The major drawback is concerned with attaining this fluid. It requires a more invasive approach than the former urine and plasma, however, it is closely associated with the CNS and spinal cord which make the discovery of protein markers more associable with their origins, particularly due to the fact that CSF does not have protein infiltration from a variety of organs, unlike the blood and urine. These advantages have yielded some success, all-be-it limited. This was shown by the 14-3-3 brain protein found in CSF fluid being the first biomarkers which can aid in the diagnosis of a transmissible spongiform encephalopathies group of diseases such as Creutzfeldt–Jakob disease (Hsich et al., 1996, Lescuyer et al., 2007) by aiding the development of a premortem immunoassay. Other successes have been reported in Guillain-Barré syndrome (Petzold et al., 2009) and Alzheimer’s disease (Mattsson et al., 2009).

The study, which is undertaken in this thesis, was done in conjunction with a number of individuals in the laboratory, in order to try and use a multiple methods approach. This investigation has focused on the use of tissue in biomarker discovery while others concentrated on biofluids and MS imaging of the tissue. This was in an attempt to counter any limitations and analyse a broader set of the proteome. The results of the biofluids analysis are described by Dr Heather Allingham in her thesis Development of proteomic techniques for biomarker discovery accepted published in 2012. The MS imaging work was performed by Dr Richard Goodwin, Research Associate and has been published with part of this work in The Journal of Proteomics, titled “Stopping the clock on proteomic degradation by heat treatment at the point of tissue excision.”

It would appear that the closer the proximity of the fluid to the sight of disease the greater the probability of success with regards to discovery will be. In recent years, therefore, there has been a switch from using biofluids to the use of cell and tissue in proteomics, perhaps in the hope of moving the field of biomarker discovery using proteomics on from a what has been suggested as a slow state of progression (Lescuyer et al., 2007) .

1.3.2 Biomarker discovery in tissue

It is with the difficulties and relatively low successes of proteomics of biofluids that has led to researchers turning toward the use of tissue and cell lines (Lescuyer et al., 2007). Although body fluids are the ideal candidate from a diagnosis point of view as they are relatively easy to obtain, there is an argument that tissue has the greatest chance of harvesting results, at least in a non-targeted or unguided approach (Phillips and Wellner, 2007, Chen et al., 2005), due to the proximity of tissue with the diseased area. Although fluids are generally less complicated, they are often a long way from the site of disease. Tissue taken from the point of origin has the potential to yield more biomarkers and in higher concentration that are more likely to be specific to that condition. When sampling body fluids it is not always clear where those markers originate or whether another disease or process is occurring. Additionally, biomarkers may have gone through a number of degradative steps, with the exception of urine, which has been shown to be relatively stable, but does show variability depending on time of day and if the sample is taken mid-stream or not (Kentsis et al., 2009, Thongboonkerd, 2007, Schaub et al., 2004, Papale et al., 2007). By taking a biopsy from the site of pathology, it is possible with greater certainty to attribute markers to that condition (de Roos, 2008). This, of course, is not without difficulties, tissue is made up of many cell types and as such there will always be a lack of certainty in linking disease to specific cell types.

Tissue proteomics in the search for biomarkers has largely concentrated on diseases which have a specific disease site such as cancer (Kondo, 2008b, Hwang et al., 2006, Zheng et al., 2003, Jones et al., 2002, Paweletz et al., 2001, Uhlén et al., 2005), stroke (Foerch et al., 2009, Sironi et al., 2004, Guerrini et al., 2002, Sironi et al., 2001, Rabek et al., 2009), or ischemic conditions (White et al., 2006, Fentz et al., 2004, Sawicki and Jugdutt, 2004), although, admittedly, these diseases are not exclusively researched using tissue. In other conditions such as those that the area affected is not specific such as Alzheimer's disease (Zellner et al., 2009, Puchades et al., 2003, Castaño et al., 2006) and CVD (Abonnenc and Mayr, 2012, Blumenstein et al., 2009, Koenig et al., 2005, Berhane et al., 2005, Ridker et al., 2000, Rohrer et al., 2004) there is more evidence of the use of biofluid. However, the use of tissue in biomarker discovery is on the rise, even for diseases such as hypertension which is systemic in nature due to the limitations of biofluids previously

discussed. These specific sites of tissue growth, such as in cancer, or infarct, as in stroke, can be biopsied in human patients and dissected out in animals. Obviously this is a more costly, time dependent and an invasive procedure in the experiment workflow and thus is one of the major disadvantages of using tissue, whereas biofluid can be extracted relatively quickly and placed into storage. The idiosyncrasies in tissue extraction via either biopsy or dissection between hospitals or laboratory, along with the time taken can have an effect on sample quality, potentially affecting downstream results. It has been shown that the degradation of tissue can occur within 60 seconds of dissection. Therefore, there is a need for strict sample management and techniques to halt degradation. (Elliott and Peakman, 2008). This is explored in Chapter 3: “Differential in gel electrophoresis analysis of wild type mouse brain tissue: A multi-faceted analytical approach to assessing the effect of heat treatment on the degradation of the proteome.”

Similarly, as for biofluids, the use of tissue has had numerous applications within biomarker discovery. Much of the work in renal cell carcinoma (RCC) has been done using tissue samples. In an interesting and large 2D-DiGE experiment, principal component analysis was used to determine between control and RCC patient groups. 520 spots were identified as being differentially expressed, while 121 of these were present in the majority of cases. The most significant of these were; annexin A2, peptidylprolyl cis–trans isomerase A (cyclophilin A), brain FABP and galectin-1 and these were further validated by western blotting (Raimondo et al., 2012a). 2D-DiGE on tissue samples has also been employed in Oesophageal cancer. A comparison was made between primary oesophageal cancer tissue and tissue near the site of the cancer which was shown to be non-cancerous. A protein called perioplakin was identified as being significantly down-regulated in cancer samples and therefore is a putative marker for oesophageal cancer (Nishimori et al., 2006). A number of other cancers have been researched using tissue. This is an advantage over biofluids, as cancers tend to have defined histological stages for direct comparison.

Another advantage of using tissue is that you can preserve the spatial resolution of markers. This may be achieved by dividing the sample into anatomically recognised areas, such as the medulla and cortex in the kidney or by using a technique which gives spatial resolution of tissues such as mass spectrometry

imaging (MSI). In theory, this type of technique could be very powerful in detecting markers and then giving information about their location and evolution over time (Sanders et al., 2008, Goodwin et al., 2008a, Chaurand et al., 2008, Meistermann et al., 2006, Chaurand et al., 2006, Chaurand et al., 2002, Reyzer and Caprioli, 2005). It is accepted however, that MSI has its limitations; this is reviewed in section 1.5.4.5. Additionally, the dynamic range of protein concentrations is thought to be lower in comparison to biofluids which is both an advantage and limitation (Jan Eriksson, 2007).

One strategy for separating anatomically distinct tissues is laser micro dissection (LMD). This has the potential to be used in the conjunction with proteomics technologies to get specific extraction from tissue slices. A method has been described in which LMD has been used to excise small amounts of tissue (and therefore protein) coupled with saturation labelling in DiGE (Sitek et al., 2005b). This technique is gaining interest and allowing the greater sensitivity (the ability to identify a positive result) of protein detection; so called small sample proteomics. This has been used in the discovery of putative biomarkers in cancer (Coca, 2008) and Atherosclerotic Coronary Intima (de la Cuesta et al., 2011). However, saturation labelling uses no more than 500µg of protein on each gel, causing problems with gaining identifications using mass spectrometry. Meaning if this style of small sample proteomics is to be employed successfully, then more sensitive MS methods and efficient protein extraction methods will be required. The use of LMD is gaining pace in the search for biomarkers due to the targeted nature of this technique but does require more development (Kawamura et al., 2010, Romanuik et al., 2009, Sanders et al., 2008, Cheng et al., 2008, Kwapiszewska et al., 2004, Kondo et al., 2003, Li et al., 2004).

One reason that the use of cell lines or tissues is so successful is because experiments are getting consistent richness of data within the 1000s of proteins, these increasing datasets obtained by such experiments is encouraging, however, it has left a large problem; how do we deal with large data sets? Currently there are statistical issues alongside problems with throughput and validation, holding back the tide of a biomarker discovery explosion (Nesvizhskii et al., 2007, Nesvizhskii and Aebersold, 2004, Cargile et al., 2004, Hancock et al., 2002).

A strong reason for the increase in the amount of tissue proteomics for biomarkers is the development of methods in order to unlock a large bank of information stored in formalin fixed/paraffin embedded (FFPE) tissue. There are many examples across the world of pathology which has been preserved using formalin. If tissue proteomics could develop, then this bank of material could serve as a great repository of information to enhance the field of biomarker discovery. Unfortunately, not only is there limitations in fresh tissue proteomics, there are additional challenges in regards to FFPE tissue. FFPE causes cross-linkages with proteins, which prevents degradation but make a lot of proteomic techniques difficult at best (Azimzadeh et al., 2010, Sprung et al., 2009, Lemaire et al., 2007, Palmer-Toy et al., 2005, Guo et al., 2007).

With a massive amount of data generated via both tissue and biofluids, how successful has proteomics been at producing biomarkers that can go from “bench to bedside?”

1.3.3 Biomarker discovery using proteomics: A success storey?

Proteomics has now been employed in the detection of biomarkers for many years and across a multitude fields, but has it been successful? There is no doubt that compared to 10 years ago there is a plethora of information regarding the association of biomolecules with a vast array of disease states. However, if the success is to be judged on the basis of using the discoveries of the research scientist to aid the clinician, a conclusion would have to be made that the reality has fallen short of the ideal. The number of clinically applicable markers discovered by proteomic methodologies is virtually non-existent. One of the main issues is that the markers being discovered are not disease specific enough for use within a clinical setting (Rifai and Gerszten, 2006).

It could be argued that the problem lies not with proteomics, but with the cross disciplinary nature of research on top of an undefined route from bench to bedside. With a more clearly defined “pipelines” there may be greater success in bringing these possible biomarkers to the clinicians and used in diagnostics (Rifai et al., 2006a, Addona et al., 2011, Gerszten et al., 2011, Makawita and Diamandis, 2010, Paulovich et al., 2008, Whiteaker et al., 2011). The pipeline needs to be a clear distinct pathway from discovery to clinical assay. Rifai et al, 2006 describe "six essential process components—discovery, qualification, verification, research assay optimization, clinical validation and commercialization” In the last 10 years the proteomic community has produced countless putative biomarkers with only a few making it to clinical assay. This connection of different fields with a dedicated pipeline might serve to bridge the gap between discovery and clinical application.

However, paradoxically, one of the unforeseen products of searching for biomarkers in fluids and tissues is the necessity to push our current technologies, methods and analytical techniques. In this respect it could be claimed to be successful in furthering these technologies. It might be also argued that the technological developments just simply have not been enough. The dynamic range of proteins is simply vast. Multiple techniques can be employed to investigate low, medium and high mass ranges but ultimately there is no complete cross over. Also, biomarkers are literally the preverbal “needle in the haystack.” In plasma 22 proteins make up the majority of the protein mass (Lescuyer et al., 2007). If markers from damaged tissue are leaked into the blood, they are masked and current proteomic technology lack

the sensitivity or resolution to detect them from the large abundant proteins (Ilyin et al., 2004)

Therefore, if biomarker discovery is going to be more successful, the clinicians and proteomists need to consider focusing on clinical outcomes which can be validated after discovery. Additionally, if biofluids and tissues are to be used, different methods need to be developed to either increase the sensitivity of detection or reduce the complexity of samples without removing potential markers. Therefore, a focus in the coming years might be in reviewing each step in the proteomic workflow and work to optimisation.

There are two main proteomic tactics in the search for biomarkers

- A global, unbiased approach using high through-put methods (such as 2DE, DiGE or SELDI)
- Targeted approach, where targets are probed using antibody related methods such as ELISA.

Although the quantitative methods used in proteomics and biomarker discovery are reviewed in section 1.4, these two different approaches are considered in section 1.3.4.

1.3.4 Proteomic work flow for biomarker discovery

As is described in section 1.3, proteomics offers the potential to uncover numerous biomarkers in a number of different diseases. The advantage of a proteomics approach as opposed to a genomic method is that the gene transcription level does not necessarily bare any resemblance to the protein levels and therefore phenotypical representation in the organism. Additionally, proteomics can gain information regarding PTMs and protein-protein interactions (Fliser et al., 2007, Issaq et al., 2002).

The proposed proteomic approach to biomarker discovery is summarised in Figure 1-4 and was described by Rifai et al, 2006. It was described as a pipeline and proposes the necessity to link proteomic discovery and validation to clinical assay development. In many ways, the field has only started to embrace this paradigm, with the vast majority of biomarker investigation forcing on phase 1; the discovery. The minority take this forward to validation and practically none have led to clinical assay development. Figure 1-4 has also been constructed with additions shown in Thongboonkerd, 2009. This has been further discussed in a number of papers and the idea is gaining credence (Addona et al., 2011, Gerszten et al., 2011, Makawita and Diamandis, 2010, Paulovich et al., 2008, Whiteaker et al., 2011).

The first step is the discovery phase to identify candidate markers, usually this is an untargeted or unguided approach using technologies that can gain a snapshot of 1000s of proteins, and peptides at a single time point. This might include technologies such as 2DE coupled to MALDI-TOF-MS or the use of SELDI-TOF-MS, or a number of quantitative technologies such as DiGE or isotopic labelling for use with MS (discussed in sections 1.4.1). This step could be achieved using cell lines, animal models or human fluid or tissue samples. This could also include steps to cause a simplification of the sample. This is often done by comparison of normal vs. disease states. This process can yield lists of many differentially expressed proteins. These lists can be 100s of proteins or peptides long and will often contain many false positives. This is especially the case for lower-abundant sets of proteins, as these often are in the lower section of the MS dynamic range (Keshishian et al., 2007)

After the initial discovery of putative markers, then there is a qualification phase. This phase is used to confirm or deny any of the generated candidate markers from the initial list of discoveries. It is suggested that this might be

achieved by using alternative technologies as a form of validation. Also, at this point, it is necessary to try and gauge the sensitivity and specificity of the biomarkers to the described diseases. That is usually expressed as a percentage measure that a diseased sample will test positive or a normal person will test negative. This stage is still performed with relatively low sample numbers.

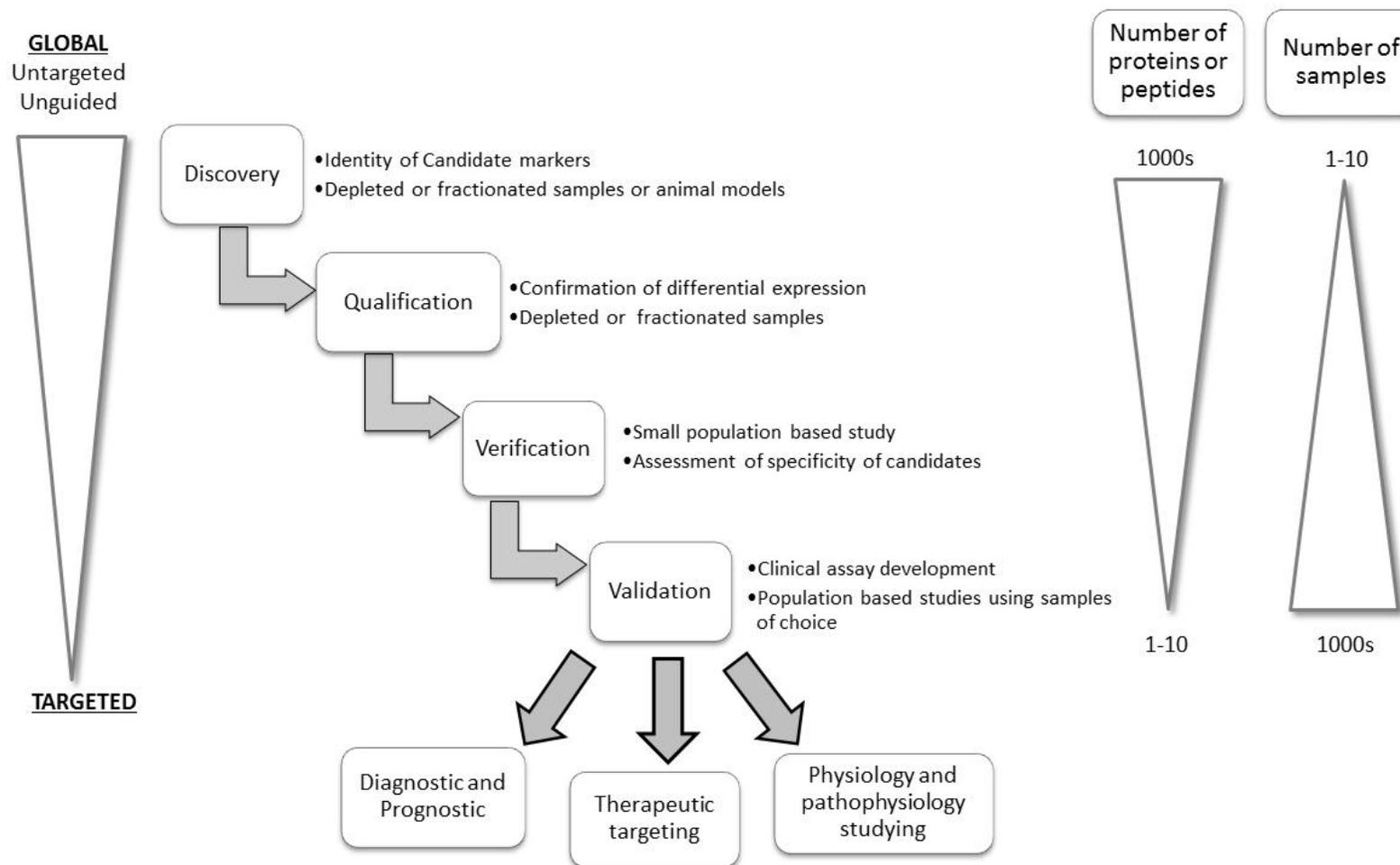


Figure 1-4: The biomarker discovery pathway. This displays a now often referenced approach to biomarker discovery and validation proposed by Rifai et al, 2006 called the biomarker pipeline. This involves several steps from discovery to validation. At the discovery phase of the pipeline, a more global approach is taken, in an attempt to examine 1000s of potential markers using small samples. As the pipeline approaches the validation stage a more targeted approach is required, with few markers being scanned but with a greater number of samples included. This was adapted from Rifai et al, 2006 and Thongboonkerd. 2009.

It is also suggested that the verification stage uses larger, usually human, samples to confirm the sensitivity and specificity from the qualification set.

The final phase involved the development of an immunoassay to finally validate markers. This assay then has the potential to be developed to meet any clinical requirements (Rifai et al., 2006a, Rifai and Gerszten, 2006).

With this in mind, the typical theoretical work flow from proteomics to clinical application is summarised in Figure 1-5, where proteins are collected, fractionated, quantified and validated.

From the literature it is clear that there is a top heavy approach to biomarker discovery with much research focused on the discovery as opposed to the validation of markers. Rifai et al, 2006 makes a sensible suggested bridge between the clinical necessity and the proteomic discovery. This is also highlighted by Mishak et al, 2007 in the need of a clearer definition of the field of clinical proteomics to bridge the gap between the scientists and the clinicians (Gerszten et al., 2008, Rifai et al., 2006a, Mischak et al., 2007).

This work flow relies on a set of discovery methodologies. In the application of these quantitative methods there are two broad approaches to biomarker discovery in proteomics; unguided global proteomics, which usually involves high through put methodologies to try and uncover a large number of putative markers or a more targeted approach, which drills down into a smaller number. This process is summed up in Figure 1-4.

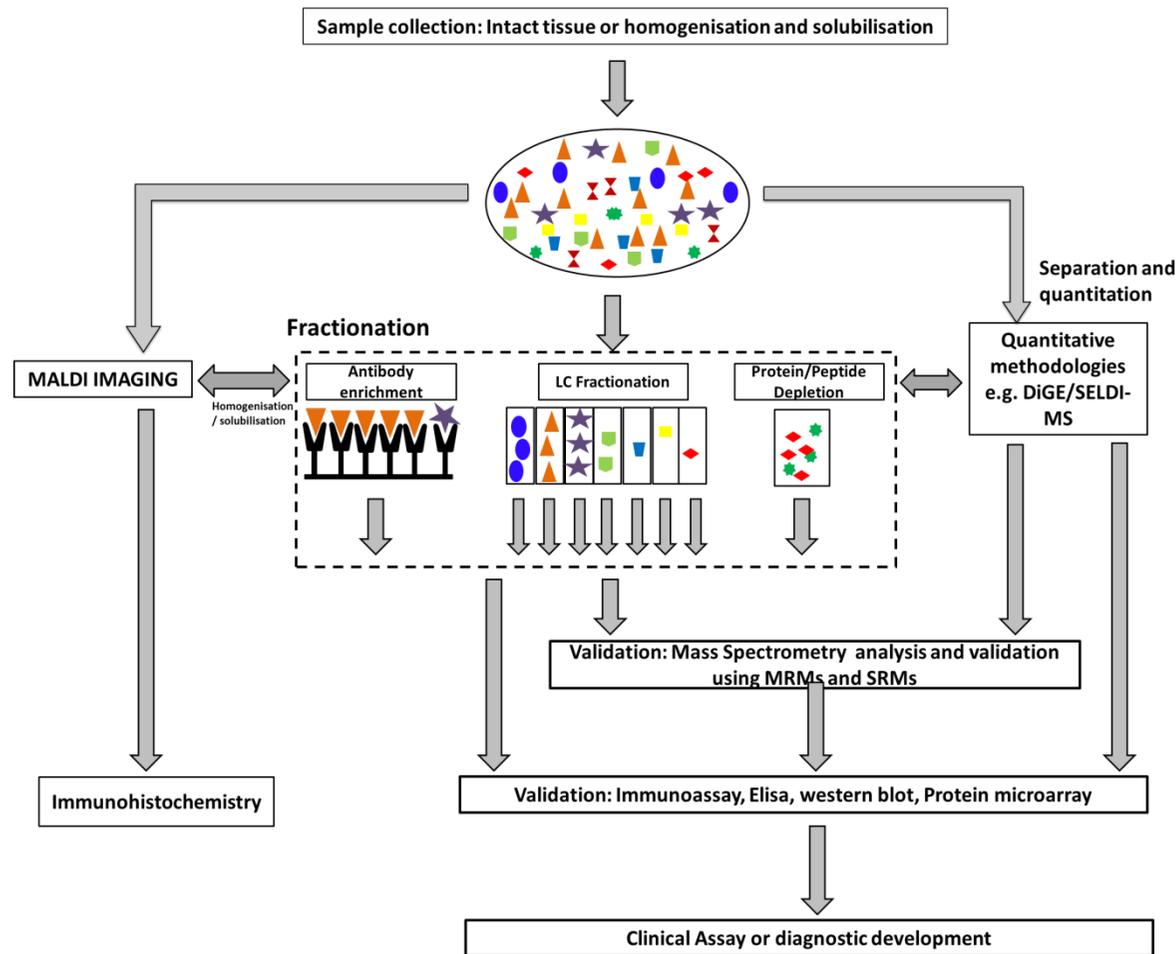


Figure 1-5: Schematic of the experimental stages in biomarker discovery in proteomics with the movement towards validation for clinical purposes. Samples come with a huge dynamic range and mixture of various proteins and peptides. These samples are then simplified by various possible methods of fractionation including; using liquid chromatography, depletion columns or antibody enrichment. The proteomic methods of choice are picked and separation and quantitation performed. Then the process of identification and validation can be executed. This is followed by a variety of possible steps toward clinical assay development.

1.3.4.1 Global unguided proteomics and biomarker discovery

Global assessment of a subset of proteins has become possible since the inception of proteomic separation and detection technologies such as; 2DE, HPLC coupled to various forms of chromatographic separation columns and MALDI-TOF-MS, SELDI-TOF-MS and ESI-QUAD-TOF-MS to mention but a view (Veenstra, 2006). One of the problems that used to be associated with biomarker discovery is the need of previous knowledge to guide the production of immunoassays. This has slowed progression in the discovery of biomarkers. However, proteomics provides solutions to this issue with the ability to screen populations due to high throughput approaches (Pan et al., 2005, Whiteaker et al., 2010). These technologies, coupled with software and bioinformatics, are able to visualise 1000s of proteins at once, thus allowing for the analysis of protein maps and patterns to compare normal and disease samples. The technologies employed in viewing whole proteomes all have their relative advantages and disadvantages, which will be described in 1.4.1 along with examples of how they have been utilised for biomarker discovery. A schematic of a typical approach to biomarker discovery without any prior knowledge can be seen in Figure 1-6. Most methods employed by proteomics are global methods for discovery and quantification and these are the primary focus of section 1.4.

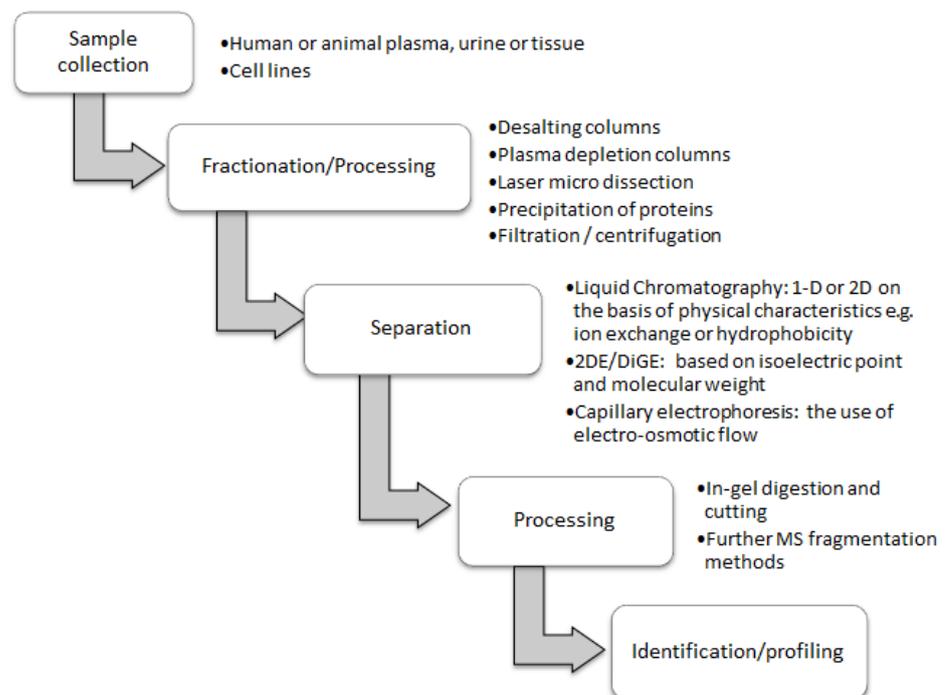


Figure 1-6: Schematic of a Global proteomics work flow in biomarker discovery (Phase 1 of Figure 1-4).

1.3.4.2 Targeted proteomics for biomarker validation

The traditional targeted approach to biomarker validation is the use of antibodies. Antibodies take a great deal of time and expense to produce and it is not inevitable that the antibody produced will have a high affinity and specificity. In addition, antibody detection can have an issue with cross-reactivity and batch production variation. Currently, antibodies still prove to be a popular way to validate biomarkers; however there has been some considerable promise in using MS to validate novel makers. In this section the more traditional validation methods are considered. In section 1.4.2.10 the targeted MS method of multiple reaction monitoring is explored (Veenstra, 2006).

1.3.4.2.1 Protein microarrays

A protein microarray is a silicon chip containing a collection of immobilised protein spots. The spots contain a set of capture molecules; this could be protein or peptide fragments or antibodies. It allows for precise protein/protein signalling. The chip is washed with either a labelled antibody or biofluid (plasma, cell lysate or tissue homogenate). Normal and disease states can be compared. The query molecules are tagged with fluorescent molecules and can be excited. A picture is taken using the appropriate excitation wavelengths which indicates a positive or negative interaction. They are useful in revealing protein-protein interactions. The main advantage is that large amounts of proteins can be validated simultaneously (Joos and Bachmann, 2009). However, protein chips have proved to be difficult to develop compared to their DNA counterparts. With a more diverse range of chemistries involving protein interactions on binding surfaces compared to DNA, researchers and industries have faced considerable problems in binding proteins and peptides to surfaces while maintaining their structures required for interactions. Once again there is an issue with antibody production, cross-reactivity and non-specific binding.

1.3.4.2.2 Immunoassays

Immunoassays measure the concentration of a solute of interest within a solution by using antigen-specific antibodies. The concentration is determined via the reaction between the antigen and antibody. The antibody can be labelled with a radioisotope, fluorescent marker or magnetic label. Although they are easy to use, they are limited in the number of samples that can be scrutinised and therefore are not efficient for high throughput purposes.

Immunoassays were first described in the 1950s by Yalow, R and Berson, SA. Yalow later accepted the Nobel Prize in 1977 (Berson et al., 1956). These are a powerful and inexpensive clinical test for known biomarkers associated with a specific disease; however, if an antibody does not already exist for a potential biomarker, it can be a lengthy process to raise them with no guarantee of a specific antibody being produced. This is also because biomarkers in a diseased state maybe little different from a normal state, perhaps only a PTM difference, although antibodies can be specific, it is not always the case. This means that immunoassays need to be specifically developed for each potential new marker. This is possibly one of the reasons a gap exists between the mass of discovery style biomarker research and the development of clinical assays, however the field has had some successes (Ray et al., 2005, Aston et al., 2002).

Another issue with antibody based techniques is the chance of cross-reactivity. The smaller the biomolecule, the greater a problem this can be, resulting in either false positives or poor quantitation. Also with regards to antibody production, consistent sources are not guaranteed as antibodies are specific to the serum of the animal they are raised in. This is minimised by having serum pools that are well characterised but this does not eliminate the problem totally. Additionally, the specificity of immunoassay is limited by the non-specific binding of the reporter molecules (Fredriksson et al., 2007).

1.3.4.2.2.1 *ELISA: Enzyme-Linked Immunosorbent Assay*

ELISA is an extension of the technology described regarding immunoassay. There are many different kinds of ELISA, however, the general principle is the detection of a solubilised analyte. The detection is done via a colour change which is quantifiable. A molecule, such as an antigen or antibody, is usually bound to a solid phase (this could be a well in a 96-well plate). In the case of biomarker discovery, a biofluid is washed over an antibody which binds to a specific target. The remainder of the fluid is washed away. A primary antibody is then washed over it and binds to the analyte as a detection method. A secondary antibody raised against the primary antibody has the conjugated enzyme to it. The enzyme reacts proportionally to an added dye. This is known as a sandwich ELISA. In an indirect ELISA an antigen or protein is adhered to the solid phase. This allows for a sensitive technique compared to a straight immunoassay as the enzymes act to amplify an initial signal as even one bound conjugated enzyme antibody will cause many dye reactions. The sandwich ELISA also has the benefit of not causing serum proteins to stick to

the well plate, as with indirect ELISAs. Quantitation can be done using a standard curve. (Diamandis, 2004a, Rifai et al., 2006b)

1.3.4.2.3 Immunohistochemistry

First employed by Coons et al, 1941 immunohistochemistry is the detection of antigen by the use of antibodies which gives spatial resolution. It is a useful method in the validation of biomarker discovery using MSI. The antibodies are tagged with a fluorescent dye which can be excited to reveal the presence of detection. In addition to fluorescent dyes, antibodies may be labelled with radioactive elements or enzymes. This technique is not only employed in the validation of markers, but in a clinical setting to diagnose and monitor the progression of disease (Fitzgibbons and Cooper, 2009, Zhang et al., 2004c, Diamandis, 2004b).

1.3.4.2.4 Western blots

The process of western blotting is summarised in section 2.1.5. Western blotting is a common technique for identifying (both qualitatively and quantitatively) the presence of a protein in a sample that has been separated on a gel. Therefore this can be done in both one dimension and two dimensions. A membrane transfer step is used to allow the proteins to migrate from the gel to the membrane. The membrane undergoes a series of washes with primary and secondary antibodies. The antibodies are tagged with a fluorescent marker and after excesses have been washed off, an image is taken with the corresponding excitation wavelength and camera. Quantitation can be achieved either relatively, against the total protein loaded across all samples, or absolutely, by using internal standard samples of known quantity.

Similar to immunoassays, western blots can only be performed if a suitable antibody is available and the specificity is dependent on the quality of the antibody (Gravett et al., 2004, Nirmalan et al., 2009b)

1.3.5 Summary

Despite proteomics generating countless data containing potential biomarkers, progress has been slower than expected. The vast majority of the biomarkers discovered have not been employed in a clinical assay. The reasons are varied. There is a lack of a definitive path for the development of discoveries to the lack of validation occurring.

The movement away from biofluids to tissue proteomics is clear, any secreted biomarker near, in or surrounding the tissue will be in considerably higher abundance compared to after travelling around the bodies in 1000s of miles of arteries, veins and capillaries. Using tissue, all be it invasive, presents a more realistic prospect for biomarker discovery as a whole. Meanwhile, sensitive technologies are required to be developed to then detect them in body fluids, if they actually exist in that form in biofluids.

However, the hard truth is that biomarker discovery is complicated and involves many processes with different organisations and institution, which mean collaboration if anything meaningful is to be achieved at a faster rate. A biomarker needs to be quantified with sufficient accuracy, in order to be deemed suitable for validation and clinical assay. Therefore, in this next section the methodologies of quantitative proteomics are considered and their relative advantages and limitations.

1.4 Quantitative proteomic technologies used in biomarker discovery.

In order to make best use of the identifications and make meaningful conclusions about the biological effect of the identified proteins it is essential to be able to quantitate changes between proteins or peptides. There is a number of proteomic technologies employed in the discovery of biomarkers. In this section these technologies are reviewed. Quantitation comes in two forms, absolute or relative. Absolute involving the knowledge of quantities by comparing a signal to a known value (this could be a set of standards used to generate a standard curve), and relative meaning the comparison of the sample of interest to another sample (this could be an internal standard).

There are a number of quantitative techniques in proteomic research which perform absolute or relative quantitation, some of which are mentioned in section 1.3.4.2. With regards to other methods of quantitation, given below is a synopsis of Differential in Gel electrophoresis or DiGE, which is the main technique presented in this thesis. In addition, there is a resume of quantitative methods used in conjunction with mass spectrometry in biomarker discovery with their relative advantages and disadvantages explored and applications.

In order to employ quantitation in proteomics it is important to understand the strategies available when labelling a sample for quantitation. Broadly speaking there are three methods of labelling for quantitation; metabolic labelling, protein labelling and peptide labelling. This is summed up in Figure 1-7. Metabolic labelling affords the advantage of reducing variation to a minimum by introducing the label further up the proteomic workflow. Metabolic labelling therefore is preferential to reduce variation and is performed in vivo with isotopically enriched elements. These will be incorporated into proteins. This necessity to perform in vivo labelling places a limitation on type of experiments possible with metabolic labelling. It is also relatively expensive in terms of money and time, having to raise cell lines or organisms on media or food containing enriched elements. In vitro labelling in comparison is cheaper and experimentally more diverse in its applications as cells, tissue or organisms can be obtained from numerous sources without the need for prior planning. The labelling of proteins and peptides, therefore, is independent of planning and preparation but has a greater source of variation due to the induction of labels further down the proteomic workflow (Gouw et al., 2010, Beynon and

Pratt, 2005, Ong and Mann, 2005). Additionally, there is a vast array of tissue already stored in tissue banks which would benefit from analysis. Metabolic labelling of this tissue is not possible so despite the disadvantages of using chemical labels, their greater flexibility in different applications is of great advantage over metabolic labelling.

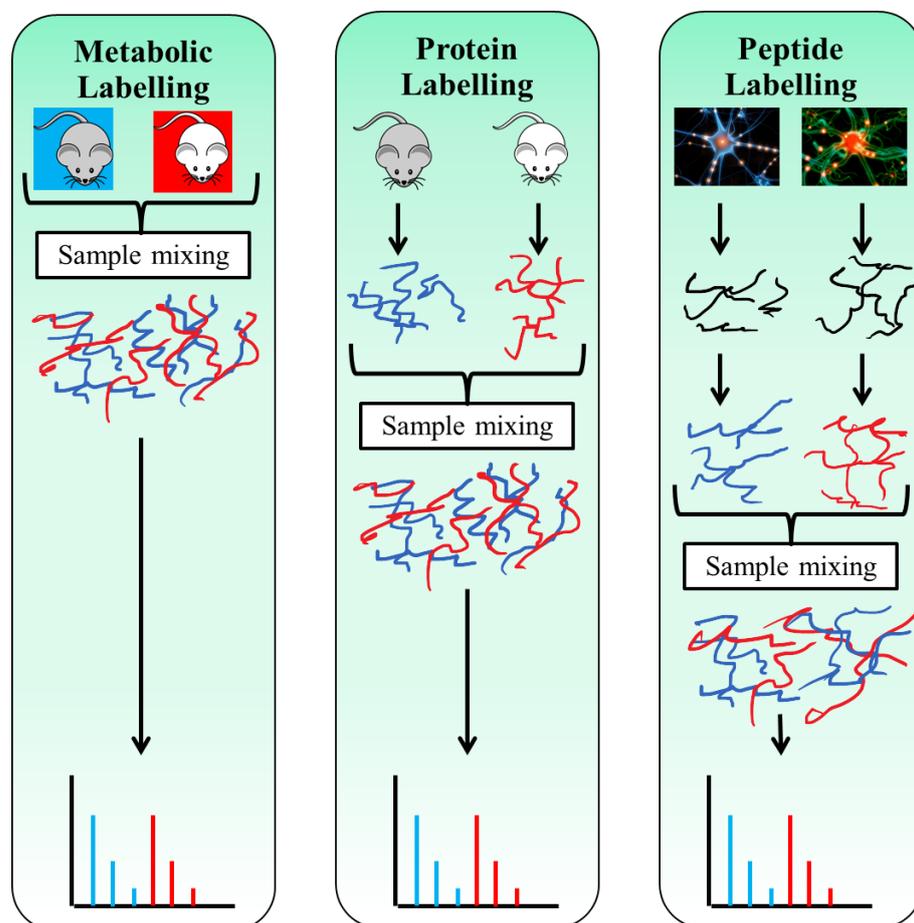


Figure 1-7: Diagram showing the strategies employed in a quantitative proteomics work flow. It show where the introduction of labels and mixing occurs and illustrated the advantages of metabolic labelling reproduced from (Gouw et al., 2010).

1.4.1 Differential in Gel Electrophoresis (DiGE)

Differential in Gel Electrophoresis is an adaptation of 2DE (discussed in 1.5.3.1.1) and born out of the limitations associated with it. It is a relative quantitative technique and harnesses the usefulness of 2DE, while combating the limitations and allows for multiplexed analysis. It achieves this by using fluorescent markers and an internal standard. DiGE was first described by Ünlü in 1997 (Ünlü et al., 1997) and then later developed by GE Healthcare (Marouga et al., 2005). It allows for the detection of 1000s of protein or peptide spots while harnessing the quantitative and sensitive capabilities of fluorescence and thus gaining the ability to compare multiple gels with greater ease and less time in comparison to 2DE. The workflow is identical to that of 2DE (outlined in 1.5.3.1.1), except samples are labelled with different fluorescent markers before separation. An internal standard, usually being a mixture of every sample, is labelled and loaded on each gel with the differentially labelled samples. This internal standard allows for the comparison across gels and quantitation. A schematic of the work flow can be seen in Figure 1-8.

Removed under the exception
of any 3rd party copyright
material

Figure 1-8: Schematic Work flow of a Minimal DiGE gel. Sample from two or more different cell or tissue samples can be simultaneously compared. A combination of all samples is mixed and labelled as a pool internal standard. These are then mixed and loaded into a standard 2DE work flow that includes isoelectric focusing and molecular separation on SDS-PAGE gels. The gels are then scanned and analysed using GE Healthcare proprietary software. Figure was taken from DiGE manual produced by GE Healthcare, UK.

The dyes employed in DiGE are cyanine dyes. There are two forms of DiGE experiments. DiGE using minimal labelling and DiGE using scarce sample labelling (saturation labelling). The most common and practically easier is minimal labelling and GE Healthcare have made three commercially available minimal Cy dyes; Cy2, Cy3 and Cy5, although there are two dyes that have not been developed for proteomic applications. They have distinct emission and absorption spectra. The dyes are size and charges matched and are spectrally resolvable. Therefore, the same protein or peptide labelled with either of the Cy dyes migrates practically to the same spot. The dyes are linear in excitation

and have a wide dynamic range. Minimal labelling occurs on 1-2% of available lysines in the sample and is bonded to the fluorescent marker at an NHS ester reactive group (see Figure 1-9). The +1 charge of the lysine is replaced by the +1 charge of the Cy dye. It has, however, been shown that there will possibly be a small proportion of unlabelled protein or peptide that will migrate and show up as a different spot (Tonge et al., 2001) and lead to extra analysis, time and expense in spot picking and identification (Wheelock AM, 2006).

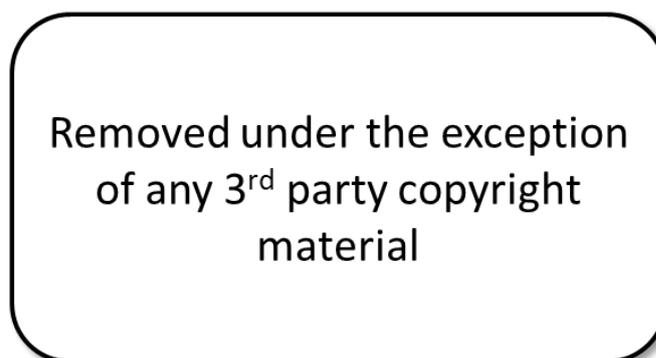


Figure 1-9: A schematic of labelling reaction for DiGE minimal labelling. The bond is created between a lysine and the NHS ester reactive group of the Cy dye forming an amide bond. Figure was taken from the DiGE manual produced by GE Healthcare, UK.

Saturation DiGE currently only has two available Cy dyes (Cy3 and Cy5) but allows for a far more sensitive approach. It is even more sensitive than silver staining, needing as little as 0.025ng of protein. Saturation labelling is 10x more sensitive than minimal labelling (needing only 5µg per a gel compared to 50µg in minimal labelling) making it an exciting prospect for small sample tissue proteomics, using laser micro dissection (Sitek et al., 2005a) as shown by Helmet Meyer Lab in an impressive study into pancreatic ductal adenocarcinoma. It was shown that saturation DiGE managed to resolve 2500 proteins from the protein equivalent from 1000 cells. However, there is still some way to go as in this study it was still required to laser micro dissect a significant amount more than 1000 cells in order to get the required level of protein. This means that the investigation still suffered from limitation of an average effect which small sample proteomics is trying to get away from (see chapter 5, section 5.5).

In comparison to minimal labelling, saturation labelling is technically more challenging. The dye concentration must first be determined for the sample in question, meaning that more optimisation is required. Saturation dyes are designed to bond to cysteine residues and all available cysteines must be

labelled. The thiol group on the cysteines covalently bond with the maleimide group on the Cy dyes. A schematic of the reaction can be seen in Figure 1-10.



Figure 1-10: Schematic of Saturation DiGE Labelling. The disulphide bridges of the thiol groups on the Cysteines must be broken before Labelling can take place. A one-hour incubation in tris-(2-carboxyethyl) phosphine hydrochloride (TCEP) prevents the thiol forming the disulphide bridges.

In a review on quantitation techniques it was stated, that DiGE is expensive in both time and money, however, depending on your sample size is no greater in expense than other labelling techniques such as iTRAQ™. Equally it stated that:

“This technique is probably more sensitive to protein degradation in the sample than are the techniques based on peptide detection, because truncated proteins would show up as separate spots on the 2D gel. Conversely, this technique is probably better at detecting isoforms than iTRAQ™ or the other peptide-based techniques, because these would probably be reported as a single protein”(Elliott et al., 2009):1637

This helps to confirm its use in chapter 3. DiGE does not circumvent all the limitation of 2DE. Gels still are not able to display very low abundance protein particularly well, low molecular weight proteins or peptides would simply run off the end of the gel and hydrophobic proteins may not enter the gel.

A rival to DiGE as a quantitation method is called ALIS (alexa molecule internal standard) and bypasses the need for labelling all the proteins in a sample by adding an internal standard, made up of a set of proteins labelled with the Alexa molecule. It is added into the sample before the first dimension separation (Wheelock et al., 2006a). Therefore, meaning that there will be no migration shift of the sample proteins due to a labelling process. This of course is assuming that all the proteins are absorbed into the IPG strip during the first dimension with equal efficiency. Once the gels have undergone the second dimension of separation, the gels are further stained with a spectroropically different fluorophore which labels proteins, such as Sypro Ruby. The gels can then be normalised against the internal standard. ALIS is cheaper in comparison to the propriety Cy dyes, however, it has not been conclusively shown whether AlexFluors are better at quantifying proteins in comparison to cyanine dyes. It is known that AlexFluor labelled proteins migrate differently depending on the proteins and this is why a form of DiGE using these dyes must be done using it as an internal standard, not using a cyanine experimental design. Therefore, doubling the amount of gels needed to be run for an equivalent cyanine dye labelled DiGE experiment (Ballard et al., 2007, Waggoner, 2006, Wheelock et al., 2006a, Wheelock et al., 2006b, Berlier et al., 2003).

AlexFluors come in a variety of excitation wavelengths. With regards to the comparison of Alexa and cyanine dyes, it is clear that there is controversy within the scientific community. Some studies comparing Cy5 and the AlexaFluor647, state that Cy5 is brighter but the AlexaFluor647 is more photostable. This comparison was done in DNA microarray use. A study was undertaken due to the problem of dye bias in cyanine dyes. Cyanine dyes 3 and 5 presented a significantly better signal than the AlexFluor counterparts (Ballard et al., 2007). As a whole DiGE based research continues to use cyanine dyes as results with AlexFluor appear to be more variable (Attard et al., 2004).

There are several advantages of using DiGE as a quantitative method. Firstly DiGE can be used to simultaneously visualise and quantitate 1000s of proteins with a reasonable dynamic range and high sensitivity. Also, gels display proteins with PTM, seen as a shift on the gel. Furthermore, the proteins are being visualised intact and not as fragments in the MS, therefore they are in

there intact state. MS can detect PTMs, however, when sample mixtures are complex (as with biomarker discovery in plasma), it becomes a challenge to gain enough spectral information to distinguish them. Therefore the main advantage and differences to MS based quantitation is the visualisation of intact proteins which can reflect their biological significance. DiGE does have its limitations. None of the proteomic technologies are truly global due to inherent limitation. DiGE can only visualise proteins in a narrow range of pIs and in the middle molecular weight range (250KDa – 10KDa), as small proteins and peptide fragments run off the gel and large proteins do not enter. Whereas, many MS method offer greater sensitivity over a larger dynamic range. Also, DiGE gels do not readily resolve hydrophobic proteins and membrane proteins very readily. However, DiGE and MS are not in total competition - they can be used in synergy as comparative and complementary methods. DiGE is suited to biomarker discovery due to the advantage of visualising large amounts of proteins in their intact state (Fenselau, 2007, Lilley and Friedman, 2004, Martyniuk et al., 2011, Wu et al., 2006b) and MS can view the smaller mass range.

1.4.1.1 DiGE and Biomarker discovery.

The process of DiGE has been explained in section 1.4.1; however, it seems prudent to review how DiGE has been employed in biomarker discovery. DiGE has been employed in a large number of biomarker studies, in a number of different fields such as: primary gastrointestinal stromal tumours using tissue, prostate cancer using tissue, bladder cancer using urine, lung cancer using saliva, breast cancer using serum, hepatic fibrosis in hepatitis C using serum and the use of material blood for biomarkers indicating down syndrome as outlined below.

DiGE has been used in the investigation of prostate cancer. Currently PSA is a relatively low specificity marker for prostate cancer. To avoid unnecessary biopsy operations, a more specific biomarker for prostate cancer would be preferable. An approach with DiGE using protein patterning has been shown to distinguish malignant cancer from benign cancer tissue using a PCA analysis. They showed 79 differentially expressed spots in the malignant cancer sample, compared to the benign cancer tissue. The majority being heat shock proteins, signal transmitting proteins, metabolic enzymes, tumour associated proteins, cytoskeletal and oxidative stress controlling proteins which had previously been linked to cancer processes. A specific example of PPAP

was found to be differentially expressed, which is a known prostate cancer marker. Different classes of peroxiredoxins were also found. Preliminary validation was done using western blots, but this study suffers from the lack of validation and therefore the gap between bench-side to bed-side is still present (Ummanni et al., 2011).

DiGE has also shown to be a useful technique in the monitoring of successful treatments. In a study into primary gastrointestinal stromal tumours using GI tissue, pftin was identified as a prognostic biomarker. Those patients with pftin showed positive results after treatment having a 5 year survival rate of 93.6% compared to a negative reading showing only 36.3% (Kikuta et al., 2010). This was further validated using immunohistochemistry techniques and in subsequent studies at other hospitals (Kubota et al., 2011).

DiGE has also been employed in investigating bladder cancer. Current diagnosis of bladder cancer is costly in terms of clinical time and often uncomfortable for the patients. DiGE has been used to identify a novel cancer biomarker of Gc-globulin from urine, for the detection of infiltrating urothelial carcinoma of bladder cancer. This is of interest due to the ease of sample collection and speed of extraction. In addition, other markers were also found to differentially express in bladder cancer patients and correlated with other studies (Li et al., 2012a).

In breast cancer, DiGE has been used with success to identify proapolipoprotein A-I, transferrin, and haemoglobin being up-regulated and apolipoprotein A-I, apolipoprotein C-III, and haptoglobin $\alpha 2$ as down-regulated in patients suffering from breast cancer. Interestingly, from an immunochemical reaction used in validation, only one protein correlated well to the original DiGE expression results. Two markers, apolipoprotein A-I and haptoglobin, did not get detected using the immunoassay, due to the lack of isoform specificity in the antibody. A considerable advantage exhibited by DiGE over the use of antibodies (Huang et al., 2006a).

Lung cancer is often fatal due to its asymptomatic nature in early stages. Late stage detection represents a high mortality rate. The need for an early detection mechanism would be beneficial. DiGE has been employed in the search of biomarkers in human saliva. In this study, DiGE analysis coupled with immunoassay methods revealed 16 candidate markers. Two of these proteins were further validated using lung cancer cell lines. The proteins were; AZGP1

and human calprotectin. Both had reasonably high specificity scores between cancer and non-cancer patients, however, these markers have also been described in association with smoking and smoking related illnesses. Something that a high percentage of lung cancer patients do (Xiao et al., 2012).

A dual technique employing both DiGE and ITRAQTM was used to investigate hepatic fibrosis in hepatitis C patients using serum as a biofluid. 305 spots were identified using ITRAQTM, of which 66 were seen as being differentially expressed compared to 704 protein spots, of which only 66 were excised, identifying 135 proteins. Two overlapping proteins were identified; complement C4-A and inter-alpha-trypsin inhibitor heavy chain H4, demonstrating the dual validation of two proteomic technologies (Yang et al., 2011a).

Heywood et al, 2011 employed DiGE to determine differential expression of maternal blood plasma proteins in fetal Down Syndrome. This was to try and avoid the risky prenatal tests for Down Syndrome diagnosis. Plasma was depleted prior to isoelectric focusing. A number of changes manifested in Down Syndrome samples taken in the second trimester: increased levels of ceruloplasmin, inter-alpha-trypsin inhibitor heavy chain H4, complement proteins C1s subcomponent, C4-A, C5, and C9 and kininogen 1. These have been described in previous studies. However, sample sizes were relatively low (Heywood et al., 2011).

The use of DiGE to analyse the proteome of cell lysates and tissue proteomics in research and clinical samples is also well documented, with much of the biomarker discovery using DiGE being geared towards studies of biomarkers for cancer and cardiovascular markers (Nordon et al., 2010). The combination of proteomic and genomic profiling, using DiGE, is also increasing (Hariharan et al., 2010). As stated above, the use of DiGE in hypertension and proteomic degradation is addressed in chapter 3 and 4.

The use of DiGE in the detection of candidate markers for renal diseases and hypertension in the kidney is not usually performed using kidney tissue but mostly urine; however there is a growing trend toward tissue as a sample source (Wang L, 2010, Varghese et al., 2010, Bañón-Maneus et al., 2010). The use of biofluids has its advantages regarding ease of collection and cost effectiveness, however analysing the kidney tissue itself gives an upstream view of what is happening and has the greater prospect of finding biomarkers

in higher abundances due to the proximity of the site of disease. It also has the possibility of isolating proteomic activity to a certain area of tissue. This approach in using DiGE in conjunction with tissue, further separating the areas of tissue and qualitative trait loci, is therefore a novel approach in this study. Renal proteomics is reviewed in 1.6 including further examples of the application in DiGE.

1.4.2 Quantitation in mass spectrometry

For many years the gold standard for separation has been in 2DE, however, disadvantages of high labour cost and time in using this technique, coupled with limitations in quantitation and software analysis, has led to the development of a range of different quantitative methods using mass spectrometry. These are largely broken into labelled and non-labelled quantitation. The majority of labelled methodologies for quantitation using mass spectrometry involved the use of stable isotope labelling (Carr and Anderson, 2008, Schulze and Usadel, 2010, Veenstra, 2006). The concept being that a protein or peptide in normal and disease samples are labelled using two differently distinguishable isotope tags. Therefore this known mass shift can be detected and used to quantitate relatively. It is largely true that the majority of mass spectrometry quantitation is relative in nature, with the exception of selective reaction monitoring and iTRAQ™, which does both. It is possible to combine a relative quantitation method such as Metal-coded tags (MeCAT) and use it in combination with inductively coupled plasma mass spectrometry (ICP-MS) to gain absolute quantitation (Ahrends et al., 2007, Bergmann et al., 2012, Schwarz et al., 2011).. Below is a summary of the most common methods of quantitation using mass spectrometry.

1.4.2.1 SILAC : Stable isotope labelling by amino acids in cell culture

SILAC is a form of metabolic labelling and was first developed by Oda et al, 1999 and first performed by Ong et al, 2002 (Oda et al., 1999, Ong et al., 2002). It is a process which introduces amino acids which carry a heavy stable isotope. Cells are cultured in a media which usually has lysine and arginine with a heavy stable isotope labelled with heavy or light versions of carbon (heavy form being ^{13}C) or nitrogen (heavy form being ^{15}N). This ensures that every peptide, except for the carboxyl-terminal peptide is labelled. They are, of course, cultured separately, lysed and processed for enrichment, depletion or simplified in line with experimental design. Then proteins are digested and all mixed together. A modified version of SILAC is summarised in Figure 1-14. Labelled cell lines can thus be compared with non-labelled cells to compare a heavy and light version of the amino acids. As the majority of amino acids are now labelled with a heavy version of the amino acid sequence coverage in the mass spectrometer is greatly increase. As it has been developed in the following years there are light, medium and heavy versions of arginine and lysine. The relative quantities can then be determined by mass spectrometry as

with ICAT. Alternatively, there is also a method known as spike-in (Geiger et al., 2011), where samples are cultured without the use of reagents and the SILAC tags are spiked into the sample later. This reduces the cost of experiments using SILAC reagents.

The major advantage is that all amino acids and therefore peptides are labelled so unlike ICAT there should be total sequence coverage. SILAC has proven a popular method, however, it does have its limitation mainly with regards to cost, as heavy media is expensive. This is to some extent combated by only labelling one sample whilst comparing a normal unlabelled version. Also some cell lines do not grow well in SILAC media and therefore it is of limited uses in these cases. It has been shown that in some cases in vitro conversion of labelled arginine to proline has occurred (Van Hoof et al., 2007). One of the major barriers to its use is the need to plan its use. This is not a method that can be used if sample have been given or used in retrospective analysis. Outside of cells the use of SILAC has seen limited success in higher animals and plants with too little incorporated heavy amino acids or taking too much time and expense. The inability to use SILAC in a large number of cases has been primarily the reason for the large number of chemical labelling alternative available .

SILAC has been successfully employed in numerous biomarker discovery studies including; various forms of cancer using human cells (Zhao et al., 2009, Grønberg et al., 2006, Kulasingam and Diamandis, 2008),

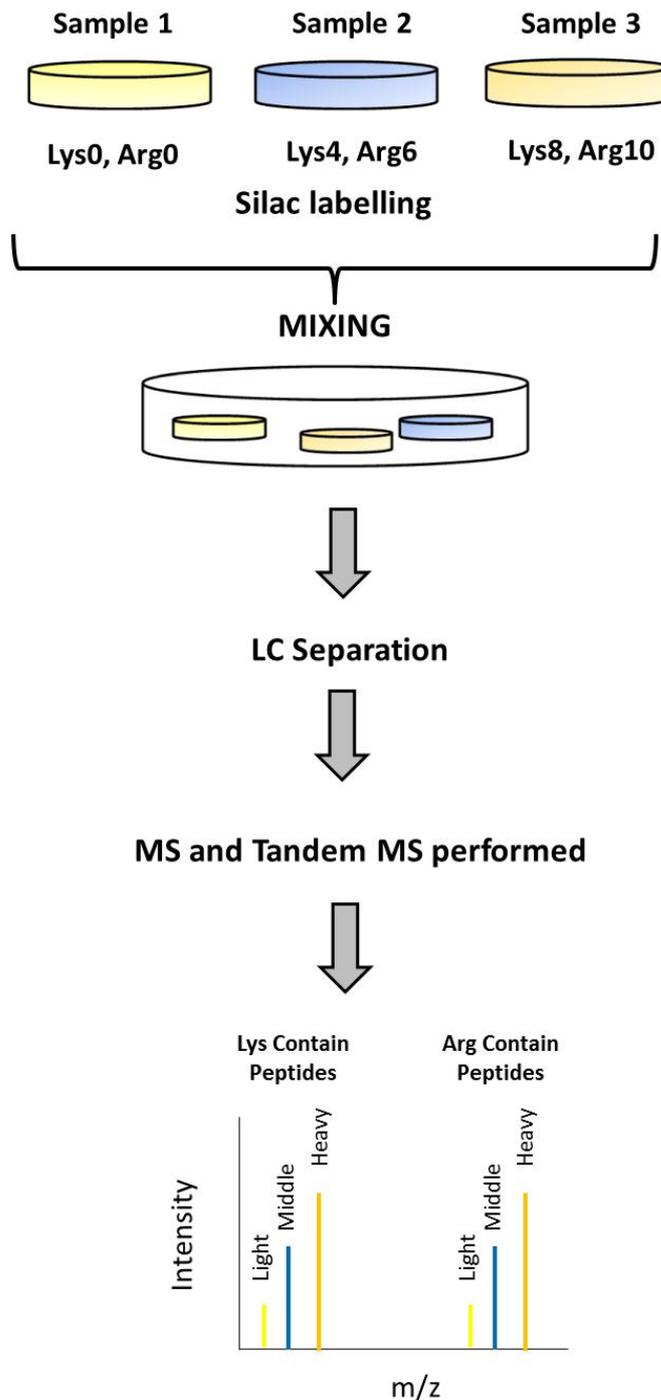


Figure 1-11: Schematic of quantitation using SILAC. Cells are grown in the SILAC media which is incorporated into the cells as they grow. Cells are lysed, solubilised, digested and then separated before MS/MS is performed. The relative peak shifts in the light, middle and heavy chains allows quantitation to be performed. Adapted from (Geiger et al., 2011)

1.4.2.2 ICAT: Isotope Coded Affinity Tags

ICAT was first utilised for use in proteomics by Aebersold et al, 1999. It uses the relative difference between labelled cysteine containing peptides labelled with either light or heavy ICAT reagents. These reagents are chemically identical. Typically, a normal sample is labelled chemically with either the light or heavy chain or the disease with the other. Once labelled, the samples

are combined and digested (usually tryptically). The excess ICAT reagent is removed using a cation exchange column. An avidin column is subsequently used to isolate the ICAT labelled peptides. Then they are introduced via LC to a mass spectrometer (Gygi et al., 1999, Kang et al., 2010a, Turtoi et al., 2010). This process is summed up in Figure 1-12.

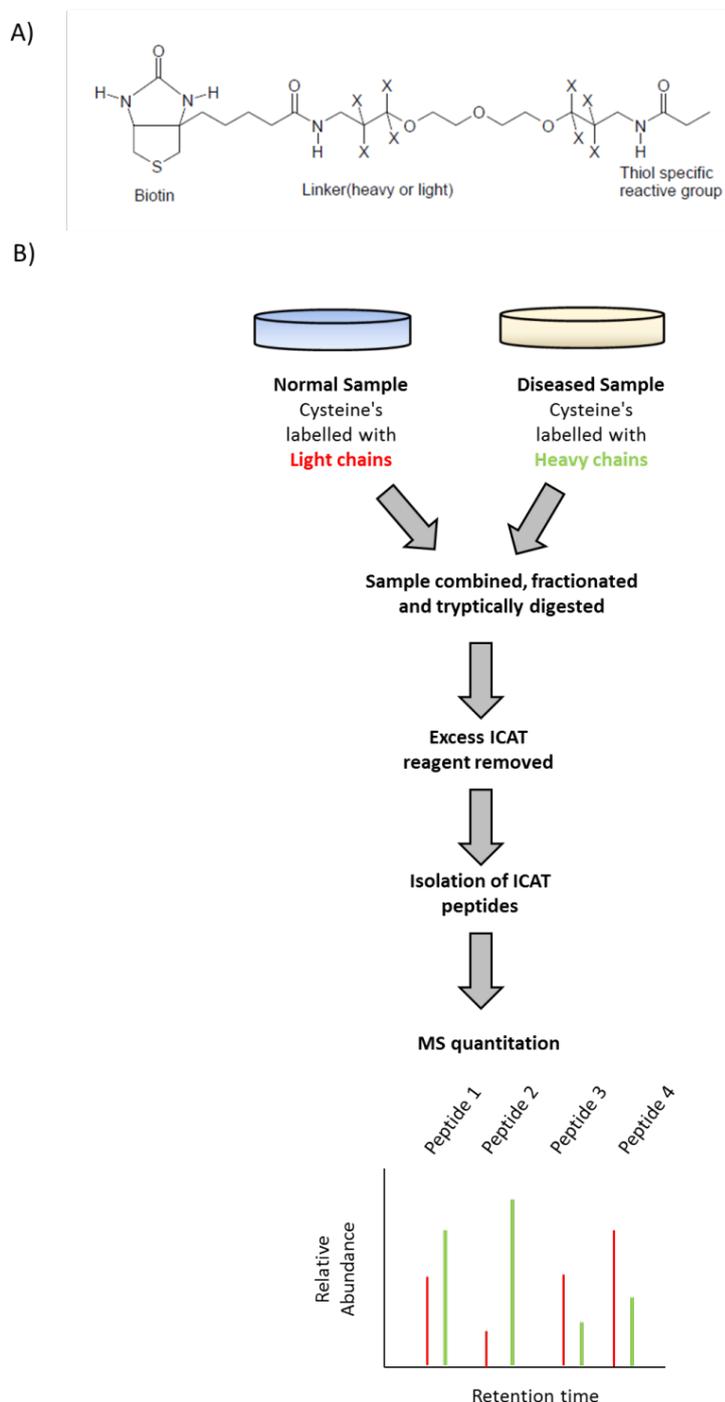


Figure 1-12: The process of ICAT. A) Shows the ICAT Label with its three components. B) A flow schematic of a typical labelling process with the heavy and light ICAT reagent chains red and green indicate light and heavy chains respectively. Normal and diseased samples are labelled with light and heavy chains respectively. They are then tryptically digested and analysed by mass spectrometry. The mass shift is detected and samples can be relatively quantitated. Combined and adapted from (Gygi et al., 1999).

The ICAT reagents consist of three distinct parts (see Figure 1-12 part A). The ICAT reagents are made up of a reactive group, a linker and affinity tag. The most common linker contains either a light chain with a (^{12}C) section or a heavy chain containing (^{13}C) (Zhang et al., 2005). These are commonly referred to as cICAT or just ICAT. This causes a mass shift of 9 Da. Originally, hydrogen and deuterium was used to create the light and heavy chains respectively, but this led to issues with co-migration of paired peptides during separation. The reactive group is a thiol which reacts and attaches to a cysteine. The peptides can be relatively quantitated against each other by either peak height or more accurately comparing the peak areas. Sequence information can be determined by further fragmentation and interrogated against databases. The shift in m/z ratio is due to the difference in mass between the light and the heavy chains.

The most recent generation of ICAT reagents are modified versions called visible ICAT reagents (VICAT). These replace the linker with ^{13}C or ^{15}N and allow for the photocleavage of the reagents to reduce the problems associated with fragmentation and co-elution. The visible tag can be monitored during separation (Bottari et al., 2003). A further version using ^{14}C and NBD fluorophore can be used to absolutely quantitate.

ICAT labelling has been employed in a number of studies including those in the search for biomarker. In one such study, biotinidase ICAT were utilised as a discovery tool and used to find a serological biomarker for breast cancer. This was performed using blood plasma from 6 patients suffering from breast cancer and 6 normal subjects. After discovery of 155 proteins, 5 were chosen for follow up with immunoblot assay and a blind study using 21 breast cancer patients and controls. Biotinidase was shown to be markedly down regulated in cancer sufferers (Kang et al., 2010a). Another investigation using ICAT combined with tandem MS was in the screening of hepatocellular carcinoma (HCC). In this study, they used cICAT to search for Alpha-1-acid glycoprotein combining AFP in patients with both liver cirrhosis and HCC patients, in order to diagnose HCC (Kang et al., 2010b).

There are certain issues that using ICAT presents. The largest restriction for this technique is that only proteins containing cysteines are quantifiable and only two samples can be compared at once. Cysteine is a relatively rare amino acid, so this presents a large problem. This can, however, be circumvented by

incorporating a sulfhydryl group onto an N-terminus of the cleaved amino acid. Additionally, there are problems with peptide co-eluting affecting quantitation. Also, the ionization process can cause different charge states for paired ions which need correction in the software. The advantages of quantifying proteins using ICAT are multiple. Firstly, any variability in ionization is eliminated as the peptides labelled with both the heavy and light ICAT reagents are ionized simultaneously (at least in theory, although it has been shown that due to the mass difference the peptides can elute from the LC at slightly different retention times, thus causing issues with correct quantitation. This can be minimized but at the expense of sequencing). Another benefit is that secondary variation in sample processing is reduced in the majority of ICAT protocols, as the mixture of the standards is early in the sample processing. Consequently, due to these limitations, an alternative method using isobaric tags for relative and absolute quantitation was developed, to minimize these limitations.

1.4.2.3 Tandem mass tags (TMT)

Tandem mass tags (TMT) are a form of isobaric labelling used in conjunction with tandem mass spectrometry, which give an alternative to traditional proteomic gel based strategies. They help to minimise the disadvantages which are coupled with performing ICAT. Independent of the enzyme used to digest the protein sample of choice, the TMT allow the amine-reactive, NHS-ester-activated compounds to covalently attach the amino terminus of the peptide or the free amino termini of lysine residues. This allows the labelling of all peptides in a sample. Due to the matched mass and structural properties of the TMT's co-elution occurs from the LC meaning that there will be fewer missed peptides when performing MS/MS. The advantages of the TMT is that they produce unique reporter ions in the MS when fragmented to allow sample quantitation. This is realised by comparing the relative intensities of the reporter ions in the tandem MS spectra. The tag itself is constructed of different groups which allow co-migration and co-elution. A reactive group and mass reporter group. In order to balance the overall masses a third group is added as a mass normalisation group. This identical mass provides a great advantage over the ICAT reagents and overcomes many of the issues associated with ICAT. Additionally, a greater overall proportion of peptides become labelled meaning quantitation is more accurate and higher sensitivity can be achieved due to the fact MS signals are not separated in two peaks.

Finally the charge states are the same with TMT, unlike ICAT, meaning there is no need for any adjustments for differently charged peptides. Currently, the most developed tags contain a set of six isobaric tags allowing multiplexing in the mass spectrometer (Bantscheff et al., 2007, Han et al., 2001, Thompson et al., 2003).

1.4.2.4 iTRAQ™: Isobaric tags for relative and absolute quantitation and tandem mass tagging

One of the limitations in ICAT is the exclusion of peptides or proteins without a cysteine amino acid. To counteract this, a method using isobaric reagents, which are reactive to amine groups and alter the peptides at the N-terminus therefore tagging all peptides, was developed (Ross et al., 2004). These isobaric tags have identical masses and therefore do not separate during LC elution. A popular method employed using isobaric tagging is the use of tandem mass tags TMTs (described in sections 1.4.2.3) or ITRAQ™. As with ICAT, this employs the concept of heavy and light chains. Once run through an LC, the heavy and light versions co-elute at the same time will undergo tandem MS in order to be quantitated. In addition, iTRAQ™ can also be multiplexed for up to 8 samples (named 113-119 and 121 due to their m/z ratio). The iTRAQ™ reagents are made from 3 elements; a reporter group, balance group and a peptide reactive group. The process is summarised in Figure 1-13. Quantitation is performed by comparing the peak areas in the MS-MS reporter ion spectra.

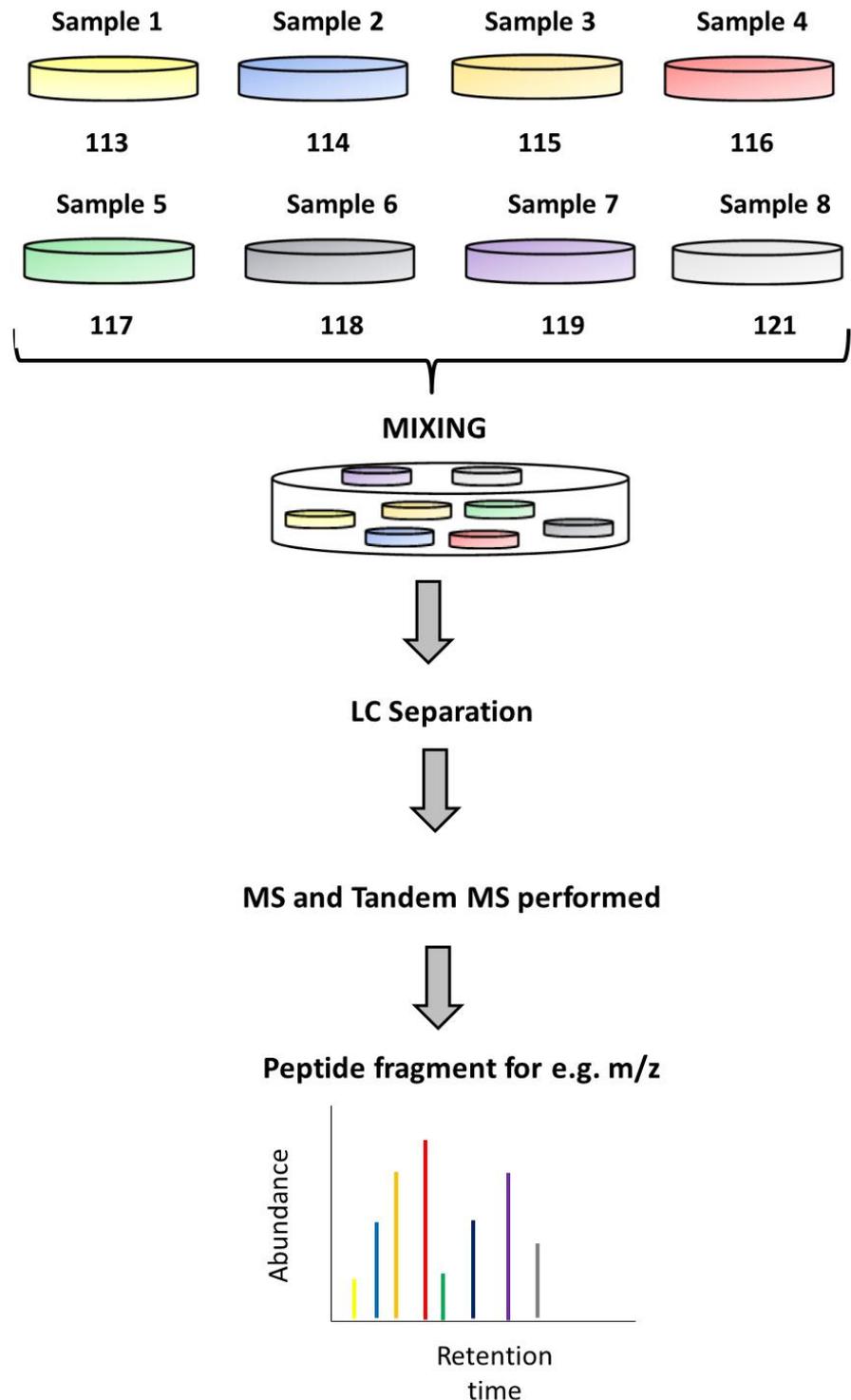


Figure 1-13: A schematic of the process of quantitation using iTRAQ™. After collection and solubilisation samples are typically digested and labelled with one of the eight iTRAQ™ reagents. After separation, samples are introduced into the mass spectrometer and identified. Identified proteins labelled with the various iTRAQ™ reagents can then be quantified. Adapted from Applied Bioscience Inc.

The reporter portion has a unique mass and structure which makes it discernable. The balance group is a mass normaliser and varies, depending on the reporter group, these two groups are collectively known as the isobaric tag, which has an identical mass across all the tags. When fragmentation occurs

during MS/MS, the reporter ion fragments off the rest of the molecule and its intensity allows for quantitation. The iTRAQ™ protocol is relatively straight forward to follow. The control and test samples are denatured with SDS, reduced with tris-(2-carboxyethyl) phosphine (TCEP), the cysteines are blocked with thiols and digested with an enzyme of choice separately. Following digestion, each sample is labelled with a different iTRAQ™ reagent. Currently up to 8 different samples. These are then combined into one sample mixture ready for introduction to the mass spectrometer. Fractionation is performed using LC. The fractionated mixture has tandem MS performed on it. This allows for the reporter ion to be fragmented off the isobaric tag and quantified.

There are considerable advantages to iTRAQ™ including the labelling of all tryptic peptides, the multiplexing of up to 8 samples, the ability to analyse post translational modifications and improved MS/MS fragmentation, which allows for a greater confidence in the identification of peptides (Shen et al., 2008). However, iTRAQ™ still has a number of limitations. There are an increased number of peptides being generated, so this is expensive in terms of operational time on the mass spectrometer. Also the process of labelling is a time consuming and intricate one, which requires strict adherence to protocols. A non-isobaric alternative to iTRAQ™ is mTRAQ (DeSouza et al., 2010). As with iTRAQ™, mTRAQ reagents are chemically identical and give labelled peptides identical retention time on the LC and the same ionisation characteristics on the MS, but contrary to iTRAQ™ have different masses. The produced sequence specific ions can then be used with MRM (see section 1.4.2.10) quantitation (DeSouza et al., 2010) to give a higher specificity than with iTRAQ™. However, it has been shown that the iTRAQ™ protocol has quantified significantly more phosphopeptides (3X) and proteins (2X) than mTRAQ labelling (Mertins et al., 2012). Both iTRAQ™ and mTRAQ have been employed in biomarker discovery. iTRAQ™ has been used to analyse urine for biomarkers in bladder cancer (Chen et al., 2010b), in plasma for diabetic nephropathy (Overgaard et al., 2010) and in prostate cancer cells (Glen et al., 2010) to highlight a few examples. Finally, a major consideration for any laboratory is expense, and iTRAQ™ reagents are proprietary and therefore expensive, a major downside.

ICenS: Isotope Coded N-terminal Sulphonation

This method of quantitation uses ^{13}C -labelled 4-sulphophenyl isothiocyanate (^{13}C -SPITC) and unlabelled 4-sulphophenyl isothiocyanate (^{12}C -SPITC) as the labelling tags. This method is said to yield a promising combination of simultaneous de novo sequencing and quantitation (Lee et al., 2004). In this technique two samples of choice are labelled independently and mixed back together in a similar fashion as described in ICAT and iTRAQTM. After running LC-MS/MS, isotopically labelled peptides should exhibit a mass shift of 6 Da. This technique has the advantage of quantitating on the basis of comparing the N-terminal sulfonated peptides and unlabelled peptides in addition to the N-terminal fragments and C-terminal fragments by comparing the fragmentation ions of the isotopic pairs. The disadvantage is of course that you cannot multiplex in the same fashion as iTRAQTM labelling. Therefore there is a trade-off between the accuracy in identification, multiplexing and quantitation. Additionally, ICenS also co-elutes at almost identical retention times which surpass the ICAT labelling using hydrogen and deuterium labelling.

1.4.2.5 O-18 labelling

Labelling with ^{18}O is a relative quantitation method which allows paired protein/peptide comparison from a variety of sample sources (Yao et al., 2003, Heller et al., 2003, Staes et al., 2004, Bantscheff et al., 2004). Samples are prepared beforehand and are typically digested prior to labelling. The advantage of this being (sample dependant) that identification runs can be done prior to labelling to try and pre-search for prospective precursor ions. Therefore the sample can be split and labelled after another separation method such as 2DE. Two atoms of ^{18}O are introduced causing a mass shift of 4 Da. These atoms label the carboxylic acid of every cleaved peptide. These can be subsequently mixed with ^{16}O for peak comparison. This is a great advantage, as serine proteases catalyse the reaction therefore trypsin aids the labelling process, as do other common enzymes used in mass spectrometry. The reaction which is caused by disruption of a water molecule with ^{18}O is shown in Figure 1-14. This continues until equilibrium is reached.

Removed under the exception
of any 3rd party copyright
material

Figure 1-14: Schematic of O-18 Labelling by reversible binding. Taken from (Fenselau, 2007)

There are considerable advantages to using this quantitative method. Every peptide is labelled meaning global inclusion. The labelling of the carboxyl terminal increases the presence of y ions in tandem MS spectra. The sensitivity is increased due to the secondary production of water rather than a chemical component which causes adulteration. Also, this method can be used in conjunction with affinity tags for peptide fractionation. This method is limited by the need for high resolution MS to detect the 4 Da shift in mass.

The use of this form of quantitation has been demonstrated to be effected by testing healthy plasma samples with lipopolysaccharide (LPS) administration (LPS-treated) and no plasma depletion. LPS is known to be a bacterial endotoxin to induce inflammatory responses. In this study the ¹⁸O labelling technique was combined with accurate mass and time tags (AMT). Using this technique, it was shown that it was possible to precisely quantitate across the global proteome in a high throughput manner, which showed great promise for biomarker discovery (Qian et al., 2005). This study also claims to eliminate the problem of back exchange by destroying the trypsin by boiling it. This does, however, raise some questions regarding sample disruption caused by heat.

1.4.2.6 Dimethyl labelling

A popular alternative to O-18 labelling is dimethyl labelling. Samples are first digested as normal and then labelled with a isotopomeric dimethyl labels. This labelling occurs due to a reaction with the primary amines of the peptide. formaldehyde is used to make a Schiff base. The labelling reaction occurs at the N-terminus of peptides and the epsilon amino group of lysine's residues. A reduction of the Schiff base occurs when cyanoborohydride is added to form a secondary amine. The greater reactivity of the secondary amine causes a reaction with an additional formaldehyde which is reduced to form a dimethylamino group. The reaction takes a sort period of time and is complete in five minutes. It is suitable for global proteomic quantitation and the ionic

state of the peptides remains unchanged which affords a considerable advantage. The molecular increases are 28 Da or 32 Da depending for the H-labelling and D-labelling respectively, creating a mass difference of 4 Da for each labelled lysine. This is a competitive choice of labelling compared to other expensive methods. The materials utilised are relatively cheap, it is a fast process, the ionisation efficiency is conserved due to the conservation of charge state and globally labels peptides at the N-terminus (Hsu et al., 2003). Initially, only two samples could be labelled but over the last 10 years this labelling procedure has been developed to triplex and then multiplex version of labelling (Boersema et al., 2008, Boersema et al., 2009, Huang et al., 2006b). This has been achieved by using different isotopomers of formaldehyde to increase the number of available labels.

1.4.2.7 MeCAT: Metal coded affinity tags

In this method of quantitation, the concept is similar to those already described, in how the method is employed. However, instead of chemically labelling an isotope to the peptides of interest, different lanthanide ions are utilised. This is coupled with inductively coupled plasma (ICP) MS to gain absolutely quantitative data. Protein identification can be then obtained via more proteomic traditional tandem MS (via any number of MS instruments). MeCAT can be used to quantitate both relatively and absolutely, and although the idea was present for a number of years, it was first described in 2007 in relation to proteomics (Ahrends et al., 2007). ICP-MS has been used for a number of years to identify metals; however it is not until recently that a strategy of labelling proteins with metal affinity tags has been employed. Peptides can be quantified down to the attomol range with the potential to go to the zeptomol range. This method has the advantage of multiplex analysis, by using different lanthanoids within the chelate complex or by using DOTA (1,4,7,10-tetraazacyclododecane- N,N^I,N^{II},N^{III} -tetraacetic acid) metal complexes. The original MeCATs contained a maleinimide reactivity for labelling thiol groups of cysteines, however, there are now MeCAT-IA which use iodoacetamide, allowing for better labelling efficiency and the elimination of a diastereomer forming (Schwarz et al., 2011, Bergmann et al., 2012). The MeCAT reagents consist of a metal portion (DOTA macrocycle), a spacer and a maleimido group for thiol reactivity. Additionally, they also can contain a biotin group to allow for purification.

1.4.2.8 AQUA peptides and QconCAT.

If researchers are to compare results across laboratories, then an absolute quantitative method is required. A strategy for absolute quantitation in the validation of biomarkers using MS is the use of an internal standard to generate a standard curve. These internal standards come in the form of synthetic peptides. There are now a variety available, however, two common versions are AQUA (absolute quantitation) peptides (Gerber et al., 2003), chemically synthesised peptides and QCAT/QconCAT (quantitation using concatenation of tryptic peptides) (Beynon et al., 2005), concatamers of tryptic peptides in an artificial protein which are labelled with stable isotope-precursors. The chemically synthesised AQUA peptides are labelled with a heavy isotope. AQUA peptides are more expensive than QconCAT, particularly if required for large studies, in larger experiments as QconCAT can be produced in batches more cheaply. Also, AQUA peptides need to be quantitated individually, taking considerable time (Pratt et al., 2006). However, the quantitation between AQUA and QconCAT has been shown to be comparable (Mirzaei et al., 2008).

1.4.2.9 Label free quantitation

Most quantitation in proteomics is performed using a label, often a fluorescent or stable isotope labelling, as discussed. As an alternative to labelling samples with a stable isotope, which is expensive and requires meticulous planning, a technique of label free quantitation is simple, safe and cost effective (at least during the initial stages). This form of quantitation uses mass spectrometry coupled to a LC or CE system, to quantitate either peptide intensities (matching LC and MS retention times) or a total spectral count is used to quantitate against (Lundgren et al., 2010, Schulze WX, 2010, Wienkoop S, 2006). Normalisation is applied to minimise errors but spectral counting has been shown to have good intra-experimental reliability and reproducibility. Using peak area measurements, however, has fewer errors when measuring protein peak ratios. Using this method is advantageous when labelling is not possible but quantitation using other methods is more sensitive. Just as with stable isotope labelling, it is important to have reproducible retention times with LC elution if quantitation is to be accurate (Turtoi et al., 2010). The label free method is likely to be popular, due to the reduction in sample processing and the unlimited number of samples that can be analysed, therefore it is likely to interest biomarker discovery researchers looking at global proteome expression

differences. This is shown by the number of publications in different fields using label free quantitation (Yang et al., 2011b, Beer et al., 2011b, Beer et al., 2011a, Wang et al., 2010, Mouton-Barbosa et al., 2010, Washburn et al., 2009, Ishihara et al., 2011).

This approach is starting to gain favour as it has application for shotgun proteomics. It does, however, have its disadvantages. Although the initial stages are relatively simple, due to lack of labelling protocols, the post experimental analytical work is time consuming. Additionally, the software is still at the relatively early stages of development. This is being tackled and currently one of the best packages available mimics the DeCyder Software and is called DeCyder MS, GE Healthcare. This has a very familiar interface in comparison to the DiGE software.

Despite being in its relatively infancy, it is still being applied to the issue of searching for biomarkers in clinical samples for cancer (Washburn AL, 2009, Washburn et al., 2010) amongst others.

The use of label free quantitation has been investigated in our laboratory in parallel to this project in an attempt to integrate data down-stream. This method has been applied to the proteomic degradation and hypertension studies (Allingham, 2012).

1.4.2.9.1 Methods in label free quantitation

There are three main label free quantitation methods available. Absolute protein expression profiling (APEX), Exponentially modified protein abundance index (emPAI), intensity based absolute quantitation (iBAQ) and T3PQ. Although at present the label free quantitation methods are less accurate than their isotopically labelled counterparts, it does afford a more straightforward, quicker and cheaper alternative. Label free strategies are split into two different ways. Either the precursor ions current area is measured such as in iBAQ or those based on tandem MS data like APEX or emPAI. Standards can be employed but if omitted the assumption is made that each protein identified contributes to the total protein pool. Traditional methods such as 2D-gel electrophoresis have been compared to various label free methods and showed a reasonable correlation but suffered when proteins were in large abundance (Grossmann et al., 2010, Kuntumalla et al., 2009).

APEX, a spectral counting method, was described by Lu et al, 2007 to estimate the relative contributions of transcriptional factors (Lu et al., 2006). APEX

works by proportionating the fraction of peptides expected and observed from a given identified protein. Proteins are analysed by introducing tryptically digested proteins from the sample of interest. Once digested they are fractionated and introduced to the MS and further fragmented by performing MS/MS. The steps to introduce and perform MS/MS itself introduces a bias which affects the overall probability of observing each peptide, this is corrected using and training a classifier to estimate the expected number of peptides. This can be used to correct the final estimate. This has been shown to provide a high throughput, cost effective and quick methods for obtaining absolute quantitation (Vogel and Marcotte, 2008). An advantage of this method, as well as emPAI is that it can be applied to existing data. However using APEX has been simplified by the APEX Quantitative Proteomics Tool which is used to perform the calculation (Arike et al., 2012). emPAI is a spectral counting method and was described by Ishihama et al, 2005 for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein (Ishihama et al., 2005). This is a modified version of a protein abundance index PAI. It is a relatively accurate quantitation tool as it takes into account that larger proteins will proportionally produce a greater number of observed peptides. PAI was used as a relative quantitative strategies but the modified version allows absolute quantitation. The emPAI uses a fraction of the molecular weight of the interested protein to allow absolute quantitation.

An alternative to APEX and emPAI is iBAQ which employs a method based on using intensity data from MS. The sum of peak intensities of all peptides matching to a specific protein are taken and divided by the number of theoretically observable peptides, this value provides an accurate estimate for the amount of protein. This has been shown to give accurate protein quantitation over four orders of magnitude (Schwanhäusser et al., 2011). A study showing parallel processing of protein and mRNA expression should high levels of correlation between mRNA and protein expression using this method (Schwanhäusser et al., 2011). The iBAQ algorithm is integrated within the MaxQuant software, used for SILAC analysis and therefore direct comparison can be made between absolute and relative quantitation data. In another studies assessing the three methods iBAQ turned out to have the greatest level of correlation between biological replicate compared to APEX and emPAI, as well as normalised distribution among protein abundances and lowest variation in ribosomal protein abundance (Kuntumalla et al., 2009).

1.4.2.10 Selective and multiple reaction monitoring

The use of MS in validation of biomarkers is performed using either selective reaction monitoring (SRM) or multiple reaction monitoring (MRM). This can be used when no antibody is available for immuno-validated techniques. These methods are often used in conjunction with stable isotope labelling to determine quantitation. A sample is traditionally separated using LC, however, other methods may be employed. If a LC is used, the molecules elution time should be known from the biomarker discovery or verification phase, therefore at the time of interest the instrument isolates the m/z ratio of concern in the first quadrupole and it is fragmented by the second quadrupole, usually by collision induced dissociation or CID. These fragments are then scanned to calculate their specificity and sensitivity. The technique of SRM is summed up in Figure 1-15. SRMs are performed using a triple-quad-MS. In Q1 only a specific peptide is selected and allowed to enter the collision cell at Q2, when CID is performed the peptide fragments, which is dependent on the amino acid sequence of the precursor. One of these fragments, at a time, is selected to pass through into the Q3 and hit the detector. This process has been extended by allowing a repeated cycling through a list of ion pairs, which are associated with a certain retention time. This allows the targeting of multiple peptides in one experiment (Monica et al., 2009).

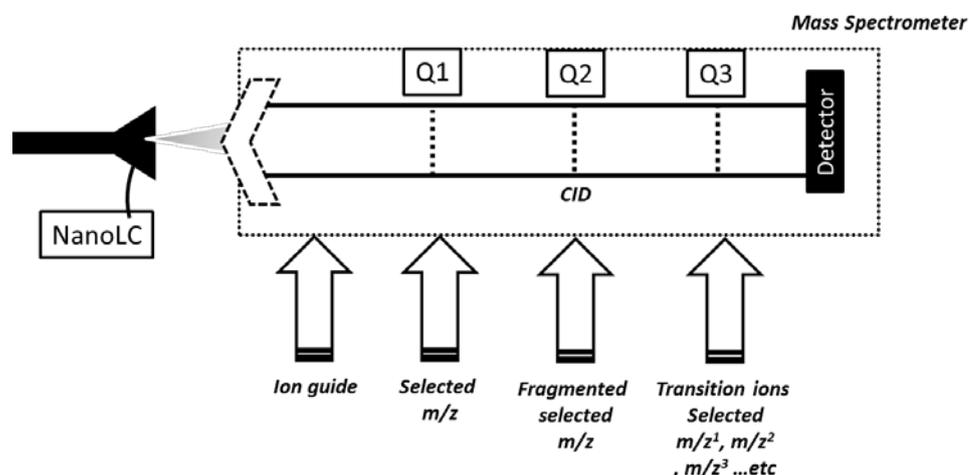


Figure 1-15: The method of selective reaction monitoring (SRM). The sample is introduced into the mass spectrometer. A specific m/z is selected in Q1 (quadrupole 1) and fragmented in Q2. The fragments are then selected as a transition ion and collided with the detector. Adapted from (Veenstra, 2006).

This is called multiple reaction monitoring or MRM. This has the advantage of targeting certain peptides in a complex mixture, which is ideal for validation or possibly even diagnostic purposes. In order to quantitate, this process is used

in conjunction with a stable isotope labelled protocol as described above. This can give relative quantitation if used in conjunction with ICAT or it can return absolute quantitation if used with AQUA peptides. This is a particularly sensitive way of monitoring ions, as much of the noise is reduced. MRMs and SRMs have been used for a number of years, but are a more recent focus in proteomics. The MRM methodology has been employed to create an assay. The methodology of this is summarised in Figure 1-16. The use of SRMs and MRMs in biomarker discovery is often as a verification step in the biomarker work flow as opposed to the discovery phase. Additionally, as biomarker verification demands high throughput methodology's MRM has been developed to make it, which has been dubbed, ultrahigh throughput (Yao et al., 2010).

Removed under the exception
of any 3rd party copyright
material

Figure 1-16: MRM experimental design logic flow chart. A particular protein is targeted. This is done by targeting specific known ions after tryptic digest and scanning for them within a given mass tolerance. Taken from (Kuzyk et al., 2009)

Due to this kind of high throughput and verification use, there have been the use of the phrase MRM assay as they are seen to be rapid, specific and sensitive as demonstrated with plasma samples (Kuzyk et al., 2009, Addona et al., 2009) and protein phosphorylation-mediated signalling networks (Wolf-Yadlin et al., 2007) . MRM and SRM are by their nature a targeted approach

and therefore used for verification, so a limitation is that prior knowledge is required and a shotgun approach cannot be used for them. Due to the selection of transition pairs, the MRM method produces results that are sensitive and selective detection of the peptides, however, there is a limit to the number of transition that can be run as the greater the number of transition chosen, the less sensitive and accurate the quantitation becomes due to the trade-off between dwell time and cycle time (Lange et al., 2008).

1.4.2.11 PSAQ: Protein standard absolute quantification

Protein standard absolute quantification was first described by Brun et al, 2007. It is a strategy to quantify absolutely trace amounts of protein in complex mixtures (Brun et al., 2007). PSAQ works by using a full length isotopically labelled proteins which can be analogues of proteins that intend to be assayed. These can be used to quantitate against after tandem MS. They come in two forms; proteins uniformly labelled with ^{15}N or a set of standards labelled on an amino acids (Picard et al., 2012). Variations exist such as absolute silac and FLEXQuant method. One advantage of using an analogues protein as a standard is that it can be introduced early in the proteomic work flow. Therefore should any of the samples protein me lost the PSAQ used will also be long, therefore a greater accuracy of quantitation occurs, this circumvents having to assess the efficiency of tryptic digestion that is necessary when using AQUA or QconCAT (Van Oudenhove and Devreese, 2013). This was shown in the assessment of staphylococcal toxins when comparing three method of quantitation of PSAQ, AQUA and QconCAT. PSAQ strategy was shown to be considerably better at quantifying these toxins. This is partly down to PSAQ compatibility with pre-fractionation techniques. This is down to the fact that the natural proteins behaviour at each point in the work flow is mimicked by the recombinant standard. It appears to work with a number of varied downstream techniques including SDS-PAGE and immunocapture (Brun et al., 2009). Additionally, it also offers the largest sequence coverage of the three methods and thus isoforms and variants can be seen. However limitations prevail regarding the cost of production and in order to use the correct PSAQ standard you need to have a protein of interest to target. Making this less suitable for exploratory investigations.

1.5 Other proteomics techniques

In order to overcome the limitations and issues associated with the complexity and heterogeneity of proteomic samples, a number of techniques have been developed. Within the proteomics work flow there is a necessity to maintain sample integrity, reduce sample complexity using separation technologies and to identify proteins. The means of separation and identification are reviewed below. The issue of sample integrity will be discussed in section 3.3. However, it is prudent to briefly review sample collection and fractionation here.

1.5.1 The importance of sample collection and storage

Sample collection and storage is of vast importance to a proteomics workflow. As highlighted above, biomarker discovery within proteomics has many issues to overcome if a greater and more promising throughput of biomarkers is to be discovered. The sample integrity is of central priority if valid results are to be obtained. It has been demonstrated that biofluids such as urine (Zhou et al., 2006), plasma (Fischerstrand et al., 1993) and CSF (Andreasen et al., 1999) have a greater degree of stability when it comes to proteolytic degradation than tissues. It has been shown using mass spectrometry imaging (MSI) that markers in mouse brain that had been allowed to warm between 30 seconds and 3 hours varied considerable across the time period, with some of the biomarkers varying from as little as 30 seconds (Goodwin et al., 2008b) and supported in (Ahmed and Gardiner, 2011, Grassl et al., 2009, Kultima et al., 2011, Sköld et al., 2007a, Svensson et al., 2007) .

The greater synchrony in how a sample is handled the greater the comparability of results is going to be; both inter and intra-comparability with laboratories. This is why HUPO are arguing for standardisations to be made in dealing with plasma samples (Rai et al., 2005). This lack of standardisation makes comparing and validating studies problematic and therefore another “bump in the road” in bringing biomarkers to a clinical setting. Although it has been shown in plasma and serum samples that the storage has little effect on proteomic characteristics even during long term storage. However, the handling, such as freeze thawing, container and laboratory practices, has had an effect in both tissue and biofluids (Arakawa et al., 2001) (Flower et al., 2000, Arakawa et al., 2001, Holten-Andersen et al., 2003, Schwartz et al., 2003, Rai et al., 2005, West-Nielsen et al., 2005).

This discussion for the need of standardisation contradicts the findings of Hsieh et al, who have shown that the effect of sample handling might actually be reasonably negligible, or at least of less importance than other technical processes (Hsieh et al., 2006).

With regards to proteomic degradation, the research into sample integrity, degradation and preservation is largely under-researched and therefore relatively small amounts of literature are available. With regards to gel based research, a strategy of 2DE and DiGE has been employed for neuropeptides analysis (Sköld et al., 2007b) but the use of DiGE to gauge the global proteomic profiles is novel.

Further the need for the development of analytical technique and methodology with regards to tissue is apparent, due to the importance of tissue as a mode of pathological information. Subsequently there is now a drive to regain information from fixed tissue (Nirmalan et al., 2009a).

Therefore, if specific clinically relevant biomarkers are to be discovered, then an important problem to overcome is sample stability and degradation. This is considered and reviewed in Chapter 3.

1.5.2 Sample fractionation and preparation

The dynamic range and complexity of biological samples postures a great deal of issues in the proteomic work flow in biomarker discovery. In order to simplify biological samples and analyse a particular subset of the proteome, there are a variety of techniques to fractionate samples prior to running analysis. This can be beneficial, particularly in phase I of biomarker discovery as it allows the simplification of sample analysis. There are a number of ways to simplify samples such as; centrifugation, precipitation, protein enrichments, protein depletion and filtration/isolation. The technique required depends on the biological question of the study. The start of fractionation usually requires the solubilisation of the sample of choice. This is achieved by disrupting the cell membranes using both physical (e.g. pestle and mortar for tissue or freeze thawing and chemical means (detergents) to solubilise proteins, nucleic acids and cellular components. This creates a homogenate which can later be applied to the fractionation method of choice.

For instance, centrifugation is used to separate blood samples into plasma and other subcellular fractions, to separate organelles (Brunet et al., 2003, Huber et al., 2003) or to concentrate cellular matter from the liquid fractions, such as

exosomes from urine (Hoorn et al., 2005, Moon et al., 2011).. However, total purity is not guaranteed and cross contamination is likely. Protein precipitations are a common technique for maximising protein content in a solution. There are a number of chemicals and protocols, but using organic solvents such as TCA or acetone is common place (Klose, 1999, Pasquali et al., 1999, Janini and Veenstra, 2002, Jiang et al., 2004a, Jiang et al., 2004b, Stasyk and Huber, 2004, Cox and Emili, 2006, Cañas et al., 2007, Rappsilber et al., 2007). This causes the increased attraction between positive and negative charges leading to the precipitation in solution. This is not a discriminatory process so cannot be used to pull down selective proteins and reproducibility is often questionable.

A more advanced method of fractionation is the use of depletion columns (Echan et al., 2005, Tu et al., 2010). These are commonly used when working with plasma samples. These deplete the higher abundance proteins in an attempt to enrich low, abundant proteins which are more likely to be the fraction biomarkers are present in. These depletion technologies can be in the form of columns that work under gravity, or more commonly those that use centrifugation to drive the low abundant smaller sized proteins through a membrane that traps the higher molecular weight proteins. This will trap higher molecular weight proteins such as albumin, reducing its masking effects in any down-stream analysis. However, important high molecular weight proteins may also be lost as well as protein-protein interactions. In addition, there is also antibody affinity removal and chromatographic systems (Zhang et al., 2007b, Romig et al., 1999) and the use of magnetic beads (Li et al., 2002a, Zhang et al., 2004a) All of these systems allow the simplification of the sample in order to improve identification and detection of less abundant proteins. However, there are always trade-offs, with the potential of losing important information and interactions.

1.5.3 Separation in proteomics

Traditionally separation techniques have been limited to strategies involving filtrations, evaporation of liquids and precipitation and distillation. However, for the molecular scientist there are a number of chemical and physical properties which allow for the separation of proteins or molecules in a sample. Many of the proteomic technologies utilise either one or more of; Molecular size, charge, polarity and shape or varying solubility or volatility.

The need for simplifying complex mixtures by separation is crucial if we are to analyse with any degree of accuracy a particular analyte molecule. One of the main modes for separation in biology has been the use of gels. Gels have been used to separate analytes for many years in DNA work and are also an essential technique employed in proteomics work. The use of gel technology is essentially a form of filtration using molecular size, filtered through a colloid and electrophoretic mobility to move sample through the gel. Ever since the advent of sodium dodecyl sulphate (SDS) in gels in 1969 (Weber and Osborn, 1969) the use of SDS-PAGE gels has been cited in countless articles. The most successful incarnation being that of the 2-Dimensional-gel electrophoresis (O'Farrell, 1975) and more recently DiGE (see section 1.4.1) (Unlü et al., 1997). Other commonly used forms of separation include Liquid Chromatography (see section 1.5.3.2) and Capillary Electrophoresis (Kuhr and Monnig, 1992) which also have their advantages and limitations.

1.5.3.1 Gel Electrophoresis theory

One of the major used strategies for separation in the biosciences is the use of Electrophoresis and gels. A gel is a colloid like a micro sponge with regular sized holes. These holes act as a filtration device, allowing for the discrimination of analytes on the basis of size. By using a detergent such as SDS, this unravels the proteins and gives an overall negative charge. Those gel pores act as a molecular size filter. If a charge is applied, the proteins or peptides will then migrate towards the negative electrode (cathode). Equation 1 shows the magnitude of the force (F) on the charged ion (q) in an applied electric field (E). This force drives the sample through the gel and the proteins will migrate faster (and therefore further) the smaller they are.

$$F = qE \quad \text{(Equation 1)}$$

1.5.3.1.1 2 Dimensional Gel Electrophoresis (2DE)

First described by O'Farrell in 1975, the use of 2-dimensional gel electrophoresis (O'Farrell, 1975) has proved an important part in the proteomists arsenal for the separation of large numbers of proteins. Described in literally 1000s of journal articles (Gorg, 2004, Oh-Ishi and Maeda, 2007, Oh-Ishi and Maeda, 2002), alongside other methods such as liquid chromatography (discussed in section 1.5.3.2).

2-dimensional gel electrophoresis is commonly used to try and map the proteomes of whole cells or tissue. As the name eludes, it has two dimensions

of separation. The first stage after protein extraction is isoelectric focusing, separating proteins on the basis of the isoelectric point. An immobilized pH gradient (IPG) strip is used. A protein will migrate in an electric field until the charge of the protein is neutral. The pI is the pH at which the protein has an overall charge of 0. Where this occurs is different for a given protein. A protein's pI generally falls in the range of 3-12 with most between 4-7. The pH gradient is created by ampholytes, these are molecules with a set pK. The strips immobilize the ampholytes using acrylamide molecules, which are cast into gels. The voltage is increased in steps to improve the focusing of the proteins. The next stage is to run the 1D separation. The staining may be Coomassie, Sypro ruby or even radio/UV labelling and placed in an image reader. The image can be placed into a software package to pick out the spots ready for automated cutting, digestion and then Mass Spectrometry (O'Farrell, 1975).

2DE was once the most commonly used form of proteomic separation and was seen as a good approach for separating proteins in the thousands range, now with increased applications and methodologies developed in MS 2DE has lost favour. If a protein is to be viewed using a 2D gel, it should be within an average pH range and not have long stretches of hydrophobic polypeptides. 2DE can often yield in excess of 3000 protein spots (O'Farrell, 1975), however it often falls short of this. It also maintains information regarding PTM (Rodriguez-Pineiro, 2006) and is "unsurpassed if one wants a global view of cellular activity" (Fey, 2001).

However, ever since the inception of 2DE there have been a number of limitations repeatedly referred to in the literature. Although it is a good way of gaining a global snap shot, 2DE does not have the ability to resolve all proteins in a particular sample. Highly basic or acidic proteins are lost at the fringes of the gel and membrane proteins are notoriously difficult to solubilise and therefore not readily resolvable on a 2DE gel, however, advances are being made here (Molloy, 2000, Wenge et al., 2008). Also, proteins that have a low molecular mass could be lost off the end of the gel. Equally, gels often get over-crowded with spots merging. There are attempts to resolve issues with narrow range gels where software can overlap the gel images (Wildgruber R, 2000). This can resolve up to 10000 spots, however, it is obviously a time costly activity on an already time consuming lab procedure.

There is also an issue of reproducibility both inter and intra lab. The reproducibility gets worse for the less abundant species of proteins (Garbis et al., 2005). There is significant variation of success in the isoelectric focusing stage, as not all protein enter the IPG strips. Whereas procedures can more easily be standardised within a lab, more variation occurs between different labs. In addition, comparison between labs has limited usefulness as 2DE gels show the proteome at a particular point in time, showing data only as good as the upstream preparation. A 2DE-gel map cannot always be compared directly as patient heterogeneity would need to be taken into account.

Quantitation has been a major issue regarding 2DGE, with artefacts and interlab difference casting doubt in the ability to quantitate effectively. Software and careful operator spot detection is necessary, but with the advent of DiGE the issues surrounding comparability are less potent (discussed in 1.4.1).

In 2DE only the abundant proteins are well-resolved and visualised, in addition downstream detection of low-abundance proteins using MS is not likely and therefore would not be able to be looked at (Fey, 2001). Additionally, hydrophobic proteins are not kept easily in solutions, strong ionic detergent can keep them in solution but cannot be used in IEF, and they tend to be lost in the exchange of detergents.

Throughput is also a big issue in 2DE-gels, as a run can take up to 3 days in large format. The gels are fragile and easily broken when moving onto staining and subsequent steps.

With all these limitations present, there has been a need for replacement or complementary separation technologies. One such technique is the other standard of proteomic separation of liquid chromatography.

1.5.3.2 Liquid chromatography

Most school child in the UK will have heard the term “Chromatography” with regards to paper chromatography, separation of black ink into its respective colours or “chroma”. Chromatography has advanced far beyond this realm. Separation in proteomics using a chromatographically method is common place but, as with 2DE, has its relative advantages and disadvantages for which solutions are being continually developed (Fröhlich and Arnold, 2006). The field itself is wide and deep and is beyond the scope of this introduction; however a brief synopsis is given below.

(LC) as a more familiar incarnation to the bimolecular laboratory was first developed by Михайл Семёнович Цвет (Mikhail Semyonovich Tsvet) and published in 1905 on this work in the separation of plant pigments (Tswett, 1905). Now many methods and strategies are used in proteomics for the implementation of different LC methods. It works by separating the analyte molecules on the basis of their physical and chemical properties. A liquid, or mobile phase (soluble sample) is passed through a column or stationary phase. Different stationary phases are available, depending on the properties the operator wishes to select for. Reverse phase separates samples on the basis of hydrophobicity and ion exchange columns on the analytes' charge (Liu, 2002, Scott, 1992, Zhang X, 2010, Jungbauer and Hahn, 2009).

The most commonly employed use of LC in the proteomics work flow is arguably shot gun reverse phase high performance (or pressure) liquid chromatography HPLC (2D and 1D). Protein samples are digested by a proteolytic enzyme such as trypsin and are separated on the basis of hydrophobicity (1D) and on charge if a cation exchange column is used as well (2D). In 1D reverse phase LC, a digested sample is solubilised and run via a mobile phase usually starting with an aqueous solution and run through a gradient to a non-polar solvent (usually acetonitrile as this compatible with mass spectrometry). The sample is first loaded onto the column and subsequently analytes with greater hydrophobicity are eluted off as the non-polar solvent concentration increases. The rate of passage depends on the hydrophobicity of the analyte, which will increase with the introduction of organic solvent (Vailaya, 2005). Often reverse phase LC is used after some other form of separation. In this investigation RPLC was employed to gain further identification using an ESI-QUAD-TOF-MS in addition to MALDI-TOF-TOF-MS.

The principles of how chromatography works lie principally with two theories; plate theory (Martin, 1941) and rate theory (van Deemter, 1959).

LC is often coupled with the mass spectrometer particular in use with RPLC as a compatible organic solvent can be used in tandem with removing contaminants using a trapping column prior to separation.

1.5.3.3 SELDI-MS: Surface enhanced laser desorption ionisation mass spectrometry

The use of protein profiling has increased in popularity recently, particularly with biomarker discovery (Issaq et al., 2002, Caffrey, 2010, Fung, 2010, Goo and Goodlett, 2010). These profiles are obtained by comparing mass spectra across a variety of normal and diseased replicates. The idea simply being that diseased fluids or tissue will have a different mass spectra signature and thus can be differentiated. One such technology used in these profiling experiments is SELDI-TOF-MS. This is a particularly popular method for examining biofluids in biomarker discovery; however, tissue homogenates can also be employed. SELDI is an extension of the popular MALDI but instead of just a using a standard plate, SELDI target plate have bound chemical and biochemical substances which causes proteins or peptides from the samples to bind to them depending on different chemical properties. A schematic of the workflow is given in Figure 1-17. Attached probes that can be used can be for example; hydrophobic, ion exchange, metal binding, antibody based, DNA fragments, enzymes or receptor molecules. Once the samples are washed over the probes, any unbound sample is washed off and the plate is spotted with matrix and ran in the same manner as in MALDI-MS. This, coupled with bioinformatics pattern recognition, can be used to differentiate healthy and diseased samples, meaning this can be used for high throughput clinical diagnosis or biomarker discovery. It is recognised that bioinformatics patterning seems to work well within labs but lab-lab reproducibility has not been demonstrated (Veenstra, 2006, Zhou, 2005). SELDI-TOF-MS has been utilised in a number of studies with regards to biomarkers. In one such study it played a central role in the development of an assay called the OVA1 Test for ovarian cancer, by profiling hundreds of samples and profiling potential marker patterns (Fung, 2010). SELDI does however have issues with dynamic range, which is particularly important when using biofluids like plasma.

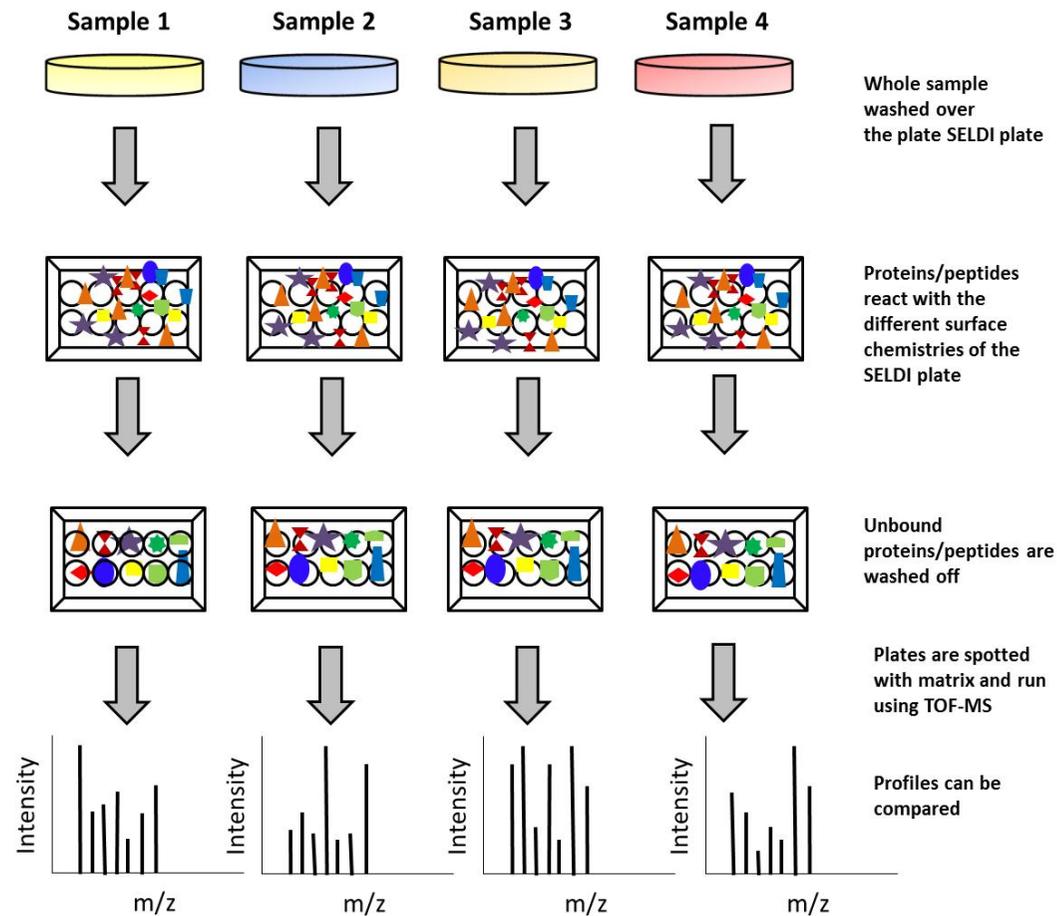


Figure 1-17: A schematic of the SELDI-TOF-MS strategy for profiling samples. The samples of interest are washed over SELDI target plates. These plates are effectively modified MALDI plates which can have various surface chemistries. This causes certain peptides to adhere and others to wash off. The sample is introduced into the mass spectrometer for identification and can be compared against each other.

There are methods to help minimise this limitation. The main method is to reduce the complexity of the sample. This does however produce run-to-run variation. There are many different kits available for protein enrichment and depletion. One such kit has seen some good results regarding reliability when used with SELDI-TOF-MS. The ProteoMiner™ Protein Enrichment Kit which proved to provide reproducible results on whole serum (Fröbel et al., 2010).

1.5.3.4 CE-MS: Capillary electrophoresis coupled to mass spectrometry

Capillary electrophoresis is another technique which is currently being utilised in biomarker discovery. Capillary electrophoresis is a high-resolution separation technique used to separate molecules on the basis of their electrophoretic mobility in a capillary. At its simplest a CE system consists of a narrow bore (in the order of 20-200µm ID) fused silica capillary, typically 50-90 cm long, filled with a buffer solution. The choice and pH of the buffer depends on the molecules being separated and as CE has developed the range of buffers and techniques used has become more elaborate (Altria, 1999). The filled capillary is placed in reservoirs at either end or coupled at one end to a Mass Spectrometer for online analysis. A schematic of a typical set up is shown in Figure 1-18.

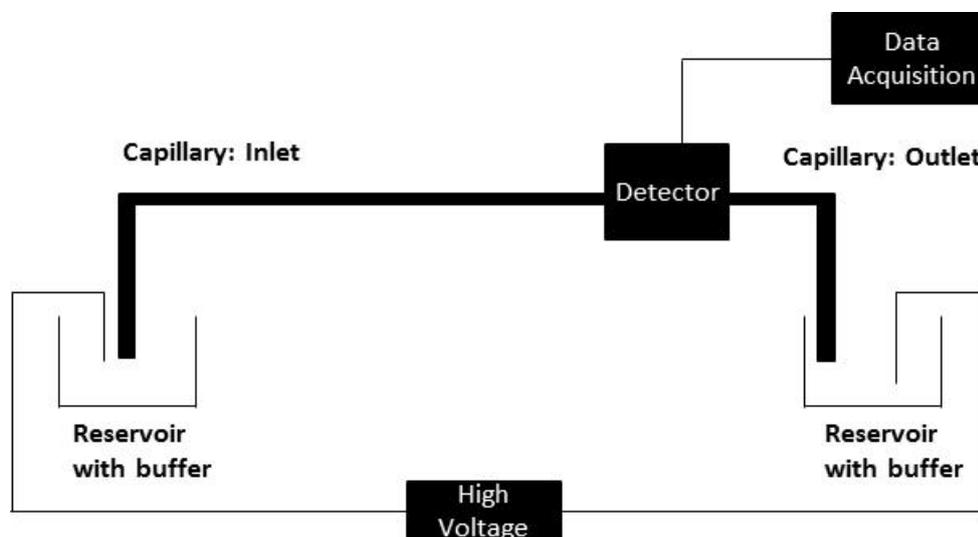


Figure 1-18: Typical setup of a Capillary Electrophoresis system. A buffer filled silica fused capillary is immersed in an Electrolyte Buffer Reservoir at either end. Applied across the two reservoirs is a high voltage. As the electrolytes move along the capillary they pass the detector window where the trace data is sent to a PC. The detector is usually a UV detector but others are available such as fluorescence. *Diagram reproduced from introduction to CE volume 1 Beckman Coulter, High Wycombe.*

The sample of interest is injected at one end usually by a pressure injection. The capillary immersed in the buffer reservoirs effectively completes a circuit

and when a high voltage is applied to the reservoirs, and current is conducted along the length of the buffer solution. This drives the electrophoretic separation. Cations migrate towards the cathode, anions to the anode and the neutral ions are not attracted to either. This separation occurs on the basis of the mass / charge of the ion. The surface chemistry of the capillary plays a major role in CE. When liquid flows through a pressure driven micro-tube, as in LC, the flow of the liquid is governed by laminar flow or parabolic flow. This is due to the frictional forces at the interface of the tube. This results in a substantial pressure drop, resulting in the flow velocity in the middle of the tube being the highest and near zero at the walls. This causes band broadening and loss of resolution. Laminar flow is shown below in Figure 1-19.

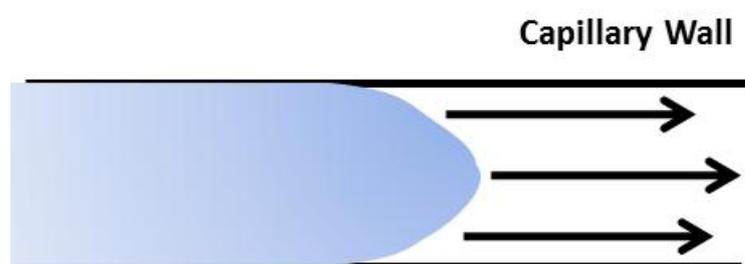


Figure 1-19: Laminar flow profile. Cross section of a capillary showing the profile of a fluid within a micro tube.

Capillary electrophoresis, however, is governed primarily by electro-osmotic flow (EOF). EOF is the consequence of the ionisable silanol groups of the fused silica capillary. The degree of this ionization is dependent on the pH of the buffer. The negative charge silanol groups attract positively charged ions to form an electrical double layer; the bottom layer, or Stern layer, is strongly adhered to the capillary wall but the top, or diffuse layer, is not. This is shown in Figure 1-20 and Figure 1-21 shows the silanol chemistry:

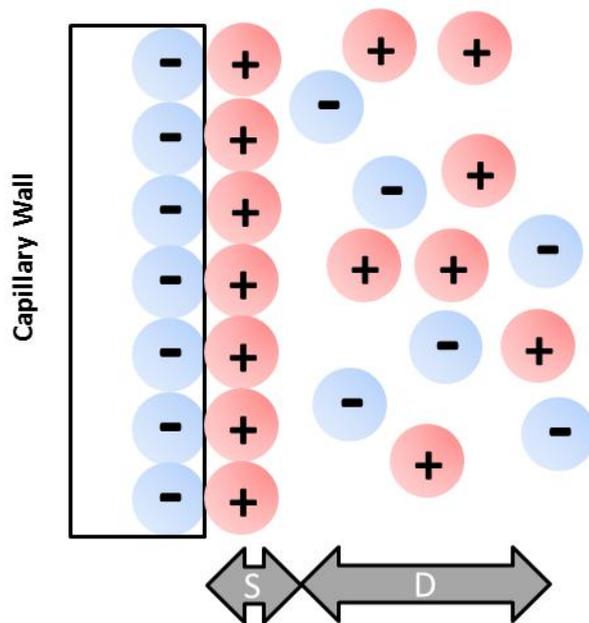


Figure 1-20: Electro-osmotic flow and its double layer. S- Stern layer and D- Double layer. This dual layer of ions causes the characteristic of electro-osmotic flow.

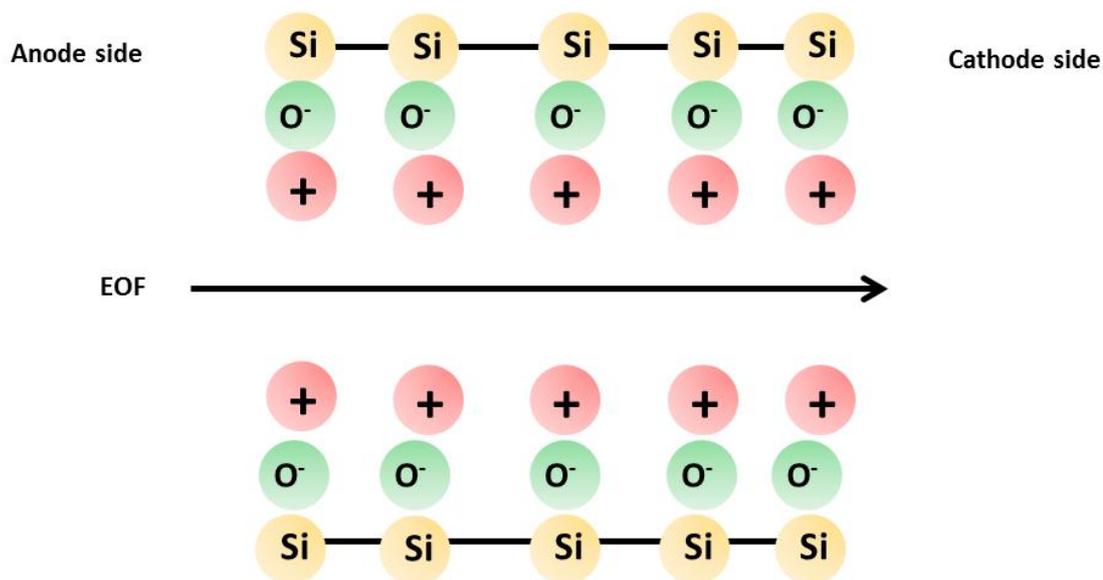


Figure 1-21: The chemistry of the silanol group in a high pH.

When a potential difference is applied to the capillary, the cations in the diffuse layer migrate toward the cathode. At neutral to high pH this causes a net flow of buffer towards the cathode carrying along with it the neutral followed by the positive ions (only if the flow is greater than the migration to cathode).

A further benefit of EOF is that flow is no longer governed by laminar flow, due to the reduced frictional forces as the forces from the voltage are uniformly

distributed across the capillary and mean that there is velocity is uniform (except very close to the walls) and the resolution is increased due to the flow profile created. Figure 1-22 shows the flow profile of an electro-osmotically driven system:

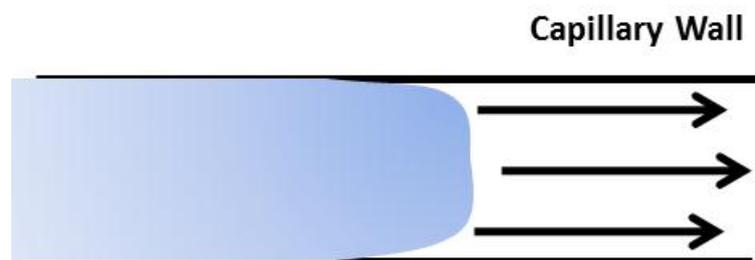


Figure 1-22: Electro-osmotic flow profile. Cross-section of a capillary showing the profile of a fluid within a fused silica capillary.

Run at high-voltage, CE is a technique that is quick and has high resolution with small samples sizes and low cost makes it a desirable separation technology.

Detection of the migrated analytes is commonly done with a UV-visible absorbance, however, recently CE has employed the use of laser induced fluorescence which use native fluorescence from aromatic components of amino acid which can also be a mode for quantitation (Szökő and Tábi, 2010, Albrecht et al., 2010).

Due to the high resolution of CE as a separation method, it has become popular in biomarker experiments. Including in diabetic nephropathy (Mischak and Rossing, 2010), Paediatric renal disease (Decramer et al., 2006), polycystic kidney disease (Kistler et al., 2009) and acute kidney disease (Metzger et al., 2010)

1.5.4 Mass Spectrometry

1.5.4.1 Introduction to Mass spectrometry

Arguably, the most important technological impact in proteomics is that of the mass spectrometer. Since the inception of MS with Thomson's work on cathode rays (Thomson, 1897) and Aston's work on the mass spectra of elements (Aston, 1919) it has evolved into a powerful tool in the field of Proteomics with the addition of Tanaka and Fenn's work (Tanaka, 1988a). The basic principle of a mass spectrometer is that ion movement is dependent on the mass to charge ratio of a particular ion and can be measured as such, the work flow for proteomics and the mass spectrometer is summarised in Figure 1-23 below.



Figure 1-23: Schematic of the workflow of mass spectrometry. Introduction of a sample is followed by its ionisation using a technique such as MALDI, and then by the injection of the ions into a mass analyser. The ion is subsequently detected and analysed.

This, coupled with the development of informatics, has proven a powerful tool to identify proteins. Soft ionisation has had an important impact in proteomics, allowing for the ionisation of organic molecules without their obliteration. The development of soft ionisation techniques; which was pioneered by Fenn and Tanaka in 1988 (Tanaka, 1988a) with the development of electrospray ionisation (ESI) and matrix assisted desorption ionisation (MALDI) (John Fenn Receiving the Nobel Prize in 2002 for ESI and Tanaka for work his work on MALDI) (Dole, 1968) has been essential (see 1.5.4.3) for the investigation of biological macromolecules using MS. Ionization plays an essential role in mass spectrometry as an analyte's mass is measured as a mass to charge ratio (m/z), in essence no ionization, no mass detection. There are several different kinds of mass spectrometers, which work in different ways, however, all of them need a source of ionization.

1.5.4.2 Protein identification

Before the theory of mass spectrometry is discussed, it is prudent to look at protein identification. After all, without the identification of proteins, the information provided by the MS would be limited. It is one thing measuring

the m/z ratio in a mass spectrometer but how is a particular protein, ion or biomolecule identified? The principal method for doing this is to generate a peptide mass fingerprint (PMF) (Pappin, 1993). This is achieved by using a known proteolytic enzyme such as trypsin, which predictably cleaves at amino acids arginines and lysines. The resultant spectra obtained from each digested peptide is compared against a theoretical digest (in-silico digest), a statistical relationship or closeness is determined from a bioinformatics database for example Mascot (Perkins, 1999). This method does have issues; the presence of any form of contaminants in the sample, PTMs and the completeness of the database being searched against. Further the scoring of the particular algorithm and its ability to distinguish hits and false positive/true positives also limit the accuracy and interpretation of PMF. Mascot uses a MOWSE score to give a statistical identification by comparing the protein databases (such as NCBI or Swissprot) with the identified peptide fragments.

Another form of identification, which provides statistically more robust data, is tandem mass spectrometry (MS/MS). This uses the principle of fragmentation of the parent peptide ion. There are a number of fragmentation techniques - the most common of which is collision induced dissociation (CID) (Johnson, 1987). Different methods of fragmentation have predictable outcomes, but fragmentation is not completely predictable. The fragments cause a series of ions. Named respectively depend on whether the fragmentation occurs at the C-terminus (x, y and z ions) or the N-terminus (a, b and c ions).

When proteins or peptide undergo fragmentation of one or several bonds we can retrieve structural information from this. The fragments can be labelled depending on the cleavage site. There are two possible scenarios: A cleavage of one or more bonds in the peptide chain or amino acid lateral chain. In a peptide chain the cleavage can occur at C α -C, C-N or N-C α giving an, bn and cn fragments if the positive charge is on the N-terminus and xn, yn and zn when the positive is on the C-terminus (the n giving the number of amino acids in the chain). Figure 1-24 below gives the possible fragment products.

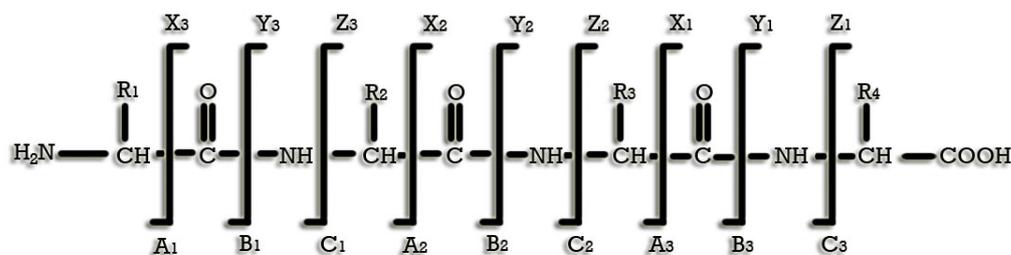


Figure 1-24: Main fragmentation paths of peptides. This figure adapted from (Hoffman)

The mass difference between successive fragments allows the deduction of the peptide sequence, except for leucine and isoleucine which are isomers and Glutamine and Lysine which are isobars. Other situations to keep in mind are the presence of post translation modifications, internal fragments, ammonium ions and the loss of water molecules, thus making interpretation of mass spectra difficult.

1.5.4.3 Ionisation techniques

1.5.4.3.1 Matrix Assisted laser desorption ionisation.

Matrix Assisted Laser Desorption ionization (MALDI) is a form of soft ionisation which has proved invaluable in the field of biomolecules and along with ESI has allowed the use of the mass spectrometer into the field of biomolecules and proteomics(Tanaka, 1988a).

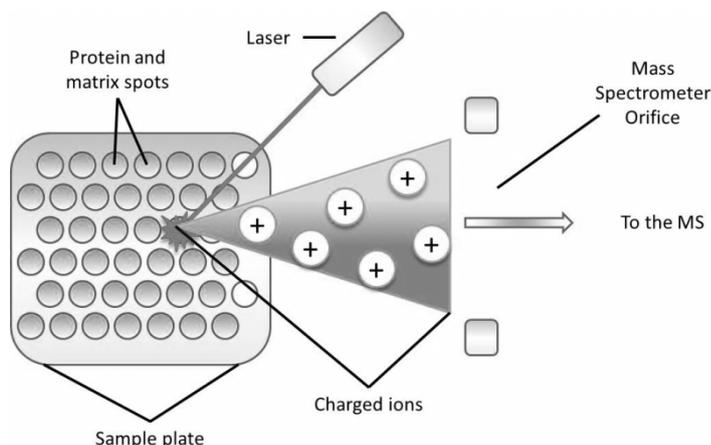


Figure 1-25: Illustration of Matrix Assisted laser desorption ionisation MALDI. A sample is place in a vacuum and positioned to receive a laser pulse. The laser is fired and ablates the crystals, which have co-crystallised with the sample. The energy is transferred and proton exchanged. A potential difference is applied and the ions are accelerated into the mass analyser.

MALDI utilizes co-crystallisation of a matrix with an analyte. A schematic of the process is shown in Figure 1-25. The matrix must have the properties of being able to co-crystallise with the sample and have an absorption wavelength

corresponding to that of the laser in order to transfer protons to the sample. The sample and matrix are spotted onto a metal plate, or in IMS a glass slide coated with indium tin oxide (ITO). The matrix and sample are spotted together in a soluble form; the matrix, commonly sinapinic acid for molecules greater than 4Kda and α -Cyano-4-hydroxycinnamic acid for less than 4KDa, is placed in a mixture of deionised water and organic solute such as acetonitrile (ACN). The ACN allows the hydrophobic molecules to dissolve into solution and the water solubilises the hydrophilic portion. Also in the matrix mix will be acid (about 0.1%) commonly trifluoroacetic acid, this gives an excess of free H^+ for the protonation of the analyte. The crystals are then ablated with the laser, the matrix absorbing the energy and transferring a proton to the analytes. There is a subsequent vaporised collection of the analyte which can then be accelerated by a potential difference in to mass analyser (Hoffman, Chang, 2007). MALDI has its advantages and disadvantage over electrospray. Electrospray is a real time event with a continuous plume into the mass analyser, whereas MALDI allows the operator to optimise setting for that particular sample and then select the precursor of interest. In relation to tissue proteomics it is possible to maintain morphological information by placing a slice of tissue directly on the MALDI target. This gives the operator the ability to profile or image and not lose morphological information. This has its own issues discussed in section 1.5.4.5.

1.5.4.3.2 Electrospray ionisation

This form of ionisation is caused by the introduction of the sample to the mass analyser via a micro fluidic needle with a high voltage applied. The theory of vaporisation dates back to Lord Rayleigh (Rayleigh, 1882). The ion evaporation model suggests that the evaporation of the charged solvent and analytes (helped by the heat created by the high voltage) would cause repulsion between the droplets. This repulsion referred to as coulombs repulsion causes the droplets to disperse into ever decreasing size as it overcomes the surface tension of the drop, until it reaches a point where the surface tension and coulombs repulsion are balanced forces. On splitting, the droplets lose a small percentage of their size (around 2%) (Fenn, 2002, Kebarle, 2009, Iribarne and Thomson, 1976).

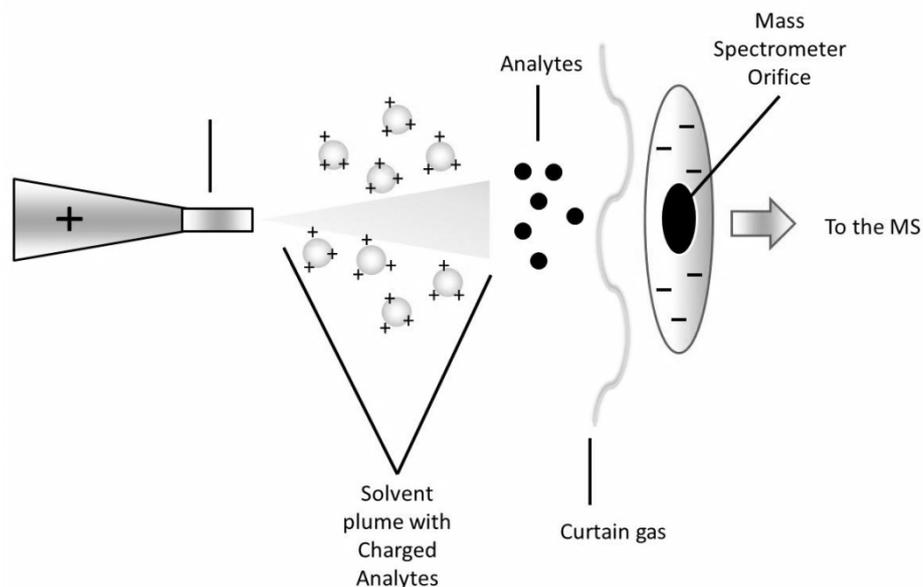


Figure 1-26: Schematic of electrospray ionisation. Sample is introduced through an electro spray needle, usually from the HPLC. The high voltage applied accelerated the ions and solvent. This caused a plume of solvent and ions. Solvent evaporation causes the destabilisation causing droplets. The ever decreasing size of the drops coupled with Coulombs repulsion causes an ever decreasing droplet size.

1.5.4.4 Mass analysers

Once ionisation has taken place, the ions are introduced to the mass analyser. This is part of the mass spectrometer that allows for the filtering and selection of precursor ions based on the ions mass. The mass spectrometers used in this thesis contained time of flight and quadropole mass analysers. These are described below.

1.5.4.4.1 Time of flight

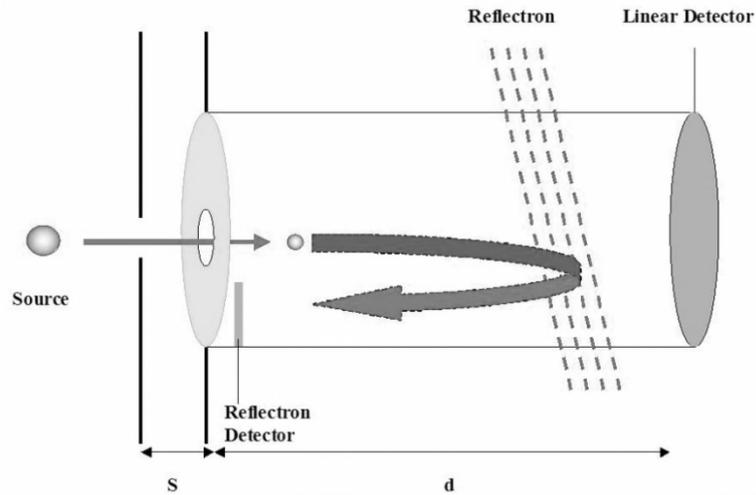
Time of flight was first described by Stephens, in 1946, but under the title “Pulsed Mass Spectrometer with time dispersion” (Stephens, 1946).

Once the analyte is ionised, it is then accelerated into the mass analyser by a potential difference, giving it a given kinetic energy. In time of flight (TOF) the analyte is allowed to fly in a field free tube under vacuum, which removes any interactions with any gas molecules. In a TOF analyser the mass to charge ratio is obtained by measuring the time that the ion takes to move through this field-free region between the source and a detector plate. The lower the mass of the ion the faster it will travel and hit the detector (Figure 1-27).

Removed under the exception
of any 3rd party copyright
material

Figure 1-27: Illustration showing the time it takes for ions of different molecular mass to reach the detector plate and thus the m/z can be measured. Taken from (Chaurand, 2005)

TOF analysers have a high sensitivity but low mass resolution. Lengthening the TOF tube can increase mass resolution, which is done in reflector mode, as opposed to linear mode. If the mass spectrometer is used in reflector mode, the ions are refocused and this minimises the aberrations that occur due to any kinetic energy variations of the ions but these also act to increase the length of the flight tube and increase mass resolution (see Figure 1-28 for a schematic of the time of flight tube). As the ions travel further, a greater resolution is achieved, but as the distance increases the ion undergoes drifting. To correct this drifting the use of a reflector was developed in 1973 (Mamyrin, 1973) allowing the refocusing of the ions. The m/z can then be worked out using the equation 2 shown in Figure 1-28 (Hoffman, Goudsmit, 1948).



$$M/z = e \ 2eEs(t/d)^2 \quad \text{(Equation 2)}$$

- m/z*** is mass-to-charge ratio of the ion
- E** is the extraction pulse potential
- s** is the length of flight tube over which E is applied
- d** is the length of field free drift zone
- t** is the measured time-of-flight of the ion

Figure 1-28: Schematic illustration of a Time of Flight tube. The ion is accelerated over region s by a given potential difference and enters from the source into a field free region marked d. The mass analyser may be set in linear or reflector mode. It is then detected by the detector plate. The equation denoting how this is calculated is shown in equation 1.

1.5.4.4.2 Quadrupole

A quadrupole consists of 4 poles opposite and adjacent in position to each other. The ions are introduced through the aperture and held in the mass analyser by the use of an opposing RF frequency and DC voltages. As the ions are introduced, only ions of certain *m/z* will remain stable to reach the detector, this allows the operator to select the given precursor masses. Any other ions will be destabilised and collide with the rods before reaching the detector. The mass filter window can be widened or narrowed by reducing the DC voltage (widen) or by increasing the voltage (narrow). A schematic can be seen in Figure 1-29 below (Slono and Volmer, 2004).

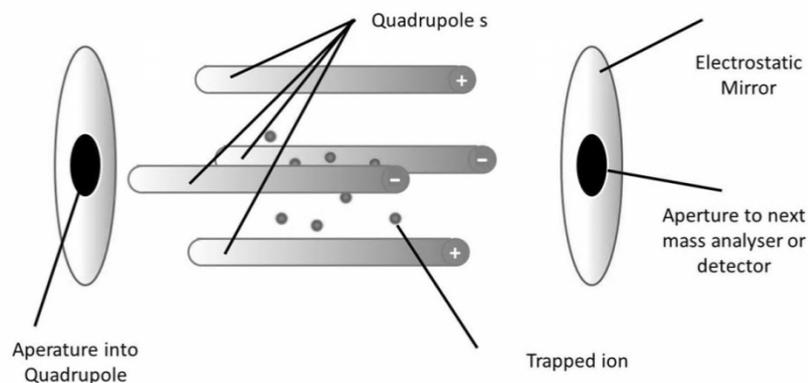


Figure 1-29: Schematic of a typical quadrupole mass analyser. Two opposing DC voltages are applied adjacent to each set of poles. An opposing RF voltage is applied to paired rods to cause stability or instability of ions with a particular m/z adapted from (Hoffman).

Now that the general proteomic work flow has been considered, it is time to look at how quantitation is performed in proteomics and how DiGE compares to other methods of quantitation.

1.5.4.5 MALDI-Imaging

As this project was collaborative in nature, it is prudent to briefly overview the two main techniques employed by the proteomic collaborators involved.

1.5.4.5.1 MALDI-Imaging and Biomarker discovery

One of the most recent applications using MALDI and is gaining momentum is the use of MALDI-TOF-MS to image tissue directly for biomolecules and profiling of tissue. MALDI imaging was first described in 1997 by Caprioli (Caprioli, 1997, Stoeckli, 2001.) and has been demonstrated as a powerful potential tool in the detection of biomarkers by imaging or profiling tissue sections, defining particular regions. It is proving to have great potential in mapping change in tissue and showing morphological relationships. Imaging mass spectrometry is a relatively new technology and improvements in areas such as sample preparation; matrix application, data collection and analysis are needed. Improvement of certain instrument parameters would also benefit including spot-to-spot sample repositioning and data processing, however even though a run takes up 7 hours, the amount of data retrieved is large. A widely studied tissue in imaging is brain tissue (Pierson, 2004, Stoeckli et al., 2002, Rohner et al., 2005, Ceuppens et al., 2007, Binz, 1999, Altelaar, 2006) as it is ideally suited to the imaging methodology due to its well-defined structures and bilateral symmetry, which can provide an automatic control, however, other tissues have been imaged including kidney (Zoriy et al., 2007) and soft tissue and cancer biopsies (Cornett, 2006).

The ability to view the spatial distribution of biomolecules in tissue sections is exciting, as it has great applications in the field of biomarker discovery due to the fact of not needing any prior information about the biomarkers involved, before being able to start investigating a tissue, and thus it can be used to characterise normal and disease tissue in tissue specific disease such as various forms of cancers or Parkinson's (Cantuti-Castelvetri and al, 2002). This ability to go "fishing" for biomolecules is a major advantage over other imaging technologies such as immuno-fluorescence, which requires antibodies (Coons, 1961). IMS has great potential as a biomarker discovery platform, but it is not a replacement for other technologies, it must be used in synergy with other technologies, such as 2 dimensional gel electrophoresis (2DGE), in order to move discovery to identification (Lescuyer et al., 2004).

1.6 Renal proteomics

With the inception of the Human Urine and Kidney Proteome Project within HUPO (Yamamoto et al., 2008) and the establishment of EUROKUP: European kidney and urine proteomics (Vlahou et al., 2009), it is clear kidney proteomics has a high profile within the Proteomics community. This has been spurred in part by the increasing number of people suffering from some form of renal disease and the limitation of current diagnostic and therapeutic treatments (Coresh et al., 2003). Therefore, renal proteomics is one field filling the gap and a fast growing and emerging subset of the proteomics field (Schaub et al., 2005, Knepper, 2002). In this section of the literature review the applications of proteomics within renal research is explored, analysing the various aspects of biomarker discovery and the use in various renal biology such as; nephrology, physiology and biomarker discovery in various forms of renal disease.

1.6.1 Application of proteomics in renal research.

Currently, there are two fundamental areas in renal research with regards to the involvement of proteomics. Scientists are either engaged in studying the physiology or pathophysiology of the kidney or they are involved in the search for biomarkers. There are also some more juvenile areas of growing interest; the uses of proteomics in the development of therapeutic targets and personalised medicine. This is summarised in Figure 1-30: This shows the 4 major areas of kidney proteomics.

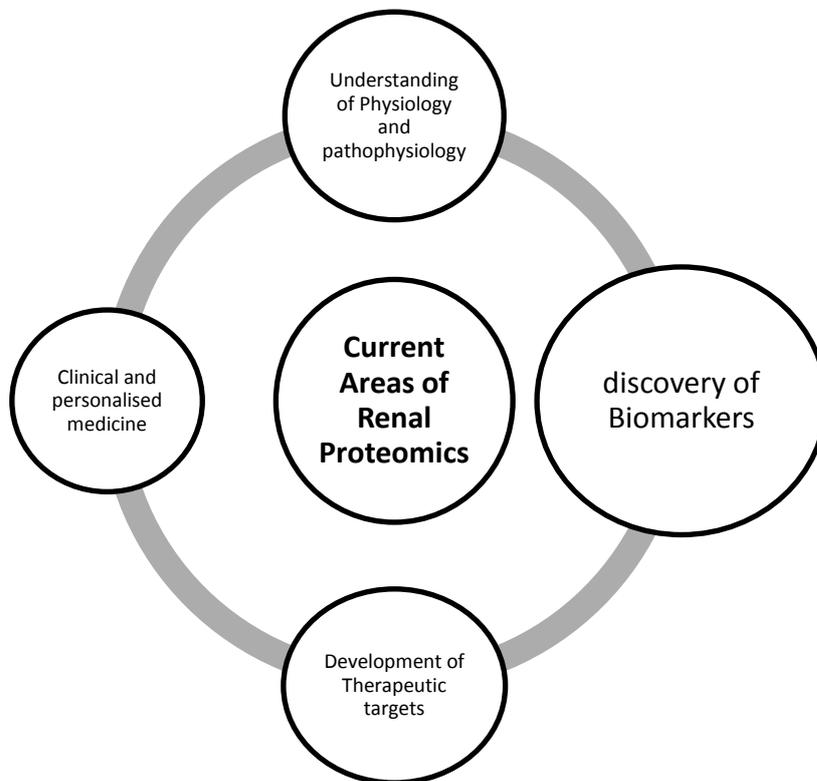


Figure 1-30: This shows the 4 major areas of kidney proteomics. The sizes of the circles are relative to the size of the field to which proteomics is currently applied. The greatest area of proteomic involvement in renal research is discovery of biomarkers.

The use of proteomics is a relatively new field, so renal proteomics is also relatively in its infancy; however it has had some success in mapping and understanding the physiology and pathophysiology of the kidney (Craven et al., 2013, Thongboonkerd and Malasit, 2005, Janech et al., 2007, Thongboonkerd, 2010). Proteomics has also discovered numerous candidate markers, principally in the analysis of urine, but has had no real success in defining novel biomarkers that could be used or taken into a clinical setting. There is some work in the development of new therapeutic targets but has not yet been fruitful (Han et al., 2008, Schaub et al., 2008, Thongboonkerd and Malasit, 2005). Also, the literature makes mention of using proteomics for personalized medicine in kidney disease, however, there is little evidence to suggest this will be occurring any time soon (Susztak and Böttinger, 2006, Jain, 2004).

1.6.1.1 Nephrology: Physiology and pathophysiology.

One of the first investigations using proteomics in nephrology is thought to originate by Witzmann et al, 1995 (Witzmann et al., 1995). Witzmann et al compared heat shock proteins (stress factors) by using normal and stress tissue in the kidney and liver, with the possibilities of being used in toxicological screening. Since then a PubMed search of kidney and proteomics returns over 1000 papers, in approximately 15 years.

A great deal of work in proteomics and molecular techniques has been conducted in understanding the role of aquaporin and vasopressin (Knepper, 2002, Knepper and Masilamani, 2001, Nielsen et al., 2002a, Pisitkun et al., 2004, Pisitkun et al., 2006, Terris et al., 1996, Terris et al., 1995, Knepper, 1994, Knepper, 1997, Terris et al., 1997, Marples et al., 1995, Schnermann et al., 1998, Nielsen et al., 2002b, Masilamani et al., 1999, Kim et al., 1999, Nielsen et al., 1993a, Nielsen et al., 1995). This work into aquaporins and vasopressin solved a long standing mystery into how water crosses the membranes and how the kidneys maintain an osmoregulatory balance in the body. Aquaporins are situated at intervals along the kidney nephron. Different aquaporin isoforms are situated at various abundances along different parts of the nephron. The interaction of AQP2 (aquaporin 2) and vasopressin (otherwise known as ADH) is crucial in the reabsorption of water in the collecting duct. This key discovery in the physiological role of the aquaporins in water control has allowed the description and characterisation of a number of water balance disorders such as nephrogenic diabetes insipidus and the effects of non-kidney related disease that with associated water balance defects like congestive heart failure. These proteins are of extreme importance and are relatively abundant. AQP1 is present in 1% of the total membrane proteins in the cortex of the kidney (Nielsen et al., 1993b). Much of these discoveries surrounding the function of aquaporins were made using a technique called immunoblotting; a targeted proteomics approach, with purification of samples this method is powerful and can be specific. With relatively low protein populations immunoblotting can be used to quantitate or semi quantitate multiple samples and has been used to characterise much of the transporter in the regulation of aldosterone and salt transporters. It is known that aldosterone causes the stimulation of sodium transport in the collecting duct via the interaction with epithelial sodium channel (ENaC). Using an immunoblot technique it was found that aldosterone caused a change in the molecular weight of gENaC from 85 kDa to 70 kDa. This is a subunit of ENaC. Additionally, the alpha subunit shows a marked increase in the presence of aldosterone with the consistencies of the beta and gamma subunits (Masilamani et al., 1999).

In order to better understand the physiology of the kidney, proteomics has been performed on whole kidney homogenate. Additionally, regions of kidney have also been described and we now have a number of 2DE reference gels for future comparison in both whole kidney, kidney cortex and medulla regions

(Arthur et al., 2002, Xu et al., 2005a, Yoshida et al., 2005). Arthur et al described the differential expression of these regions (Arthur et al., 2002). These studies centre on using rat or murine kidney. These maps coupled to the identification of proteins will provide a useful reference for future studies. Proteomic mapping of normal sample is not limited to tissue but also to urine. Adachi et al used LTQ-FT MS to evaluate over 1500 proteins of the urinary proteome. Examples of how using this global proteomic approach to mapping kidney physiology includes; (Knepper and Brooks, 2001, Weissinger et al., 2004, Hoorn et al., 2006)

Proteomics has been successfully used to study the physiological effects that toxins have on the kidney. The change in the proteome can be used to determine the underlying mechanism of toxins on the protein expression of the kidney. One such study looked at the gentamicin treatment on renal cortical protein expression using a 2DE approach (Charlwood et al., 2002a). This study gave a useful insight into the role of mitochondrial proteins and those involved in production of ATP. In addition, these changes in protein expression could be used as toxicity markers to identify causes of poisoning. Clinically identifying a toxin is difficult and often a crucial, time dependent step in the application of suitable treatments. Toxicity markers could be extremely important in this regard. Other proteomic studies into toxicity include lead (Witzmann FA and LS, 1999) , fluoride (Xu et al., 2005b) uremic toxins (Kaiser et al., 2003) among others.

This information on the whole proteome has limited uses due to averaging effects and the high amount of cell types in the renal structures. Due to the issues associated with gaining enough sample for proteomics experiments, there are more limited numbers of studies addressing specific areas of the kidney such as parts of the nephron. Proteome mapping has taken place on the human glomeruli (Musante et al., 2002, Potthoff et al., 2008), brush border membrane vesicles (Cutillas et al., 2004b), the ascending and descending loop of Henle (Dihazi et al., 2005) and collection ducts (Hoffert et al., 2007) amongst others. By isolating certain proteins thought to be important for various processes, a proteomic analysis could be performed on any associate proteins with their extraction. Further to this, MS or western blot analysis can be used to identify these proteins. For instance, denzins association with

nephron was determined using this approach and MALDI-TOF-MS (Ahola et al., 2003).

Much of medicine has its basis in discovering the pathological nature of tissue and proteomics has been used to investigate this in an attempt to characterise the pathophysiological nature of certain diseases. Information has been gathered regarding; diabetic nephropathy (Thongboonkerd et al., 2004, Susztak and Böttinger, 2006, Alkhalaf et al., 2010, Randa et al., 2010, Overgaard et al., 2010, Isabel Padrão et al., 2012), glomerular disease (Ahola et al., 2003, Nabity et al., 2011), urological and hepatocellular cancers (Adam et al., 2001, Poon et al., 2003, Kommu et al., 2004, Ren et al., 2010),

With regards to studying glomerular disease, the glomeruli are isolated using a sieving technique. Currently, there is a glomerular data base linking identification to 2D gel images. This “normal” reference map can be used to compare against when studying the pathology of the glomerulus. Proteomics has been largely responsible for the uncovering of the pathophysiological mechanism in proteinuria in glomerular disease. Focal segmental glomerulosclerosis (FSGS) factors suspected in causing the damage to the glomerular barrier were confirmed using multidimensional LC and electrophoretic approach identifying 6 key proteins which maintain strong permeability activity to albumin and albumin fragments (Musante et al., 2002).

There has also been an increase in understanding regarding the pathophysiology and mechanism underlying diabetic nephropathy. Thongboonkerd et al, 2004 described over 30 proteins which were differentially expressed in the diseased kidney of type 1 diabetic patients, a range of roles were proposed with novel discovery of elastase inhibitor expression was shown to increase with the corresponding decrease in elastase. Elastin is thus controlled within a pathway involving an elastase enzyme and inhibitor. Elastin maintain integrity of the glomerulus. Therefore this discovery gives important information about the mechanism of this disease state. This takes treatments a step closer.

There is also currently discussion regarding proteomics in the role of diagnostic pathology, and using it to monitor the progression of diseases, particularly using processes such as MALDI-IMS; (Stoeckli, 2001, Goodwin et al., 2008a), There have been some interesting investigations with this regard into the diagnostic and prognostic assessments of cancers. MALDI-TOF-MSI

was used to sub classify tumours and also place a prediction of prognoses to different classifications of lung cancer (Yanagisawa et al., 2003). It is also possible to map the progression and treatment characteristics of cancer as shown in mice (Kasper et al., 1998).

1.6.1.2 Biomarker discovery in renal research

The field of renal proteomics is heavily biased towards the discovery of biomarkers, with the vast majority of citations regarding kidney and proteomics involving the search for biomarkers. The largest source for this is the use of urine as a biofluid to explore for markers in kidney disease. This is due to the obvious association with the kidney. Biomarkers have been described in renal disease, cancer (see section 1.3.1), renal injury and hypertension and cardiovascular disease, along with others.

1.6.1.2.1 Acute kidney Injury Biomarkers

Acute kidney injury (AKI) or acute renal failure is a collective name given to processes surrounding the failure of the renal system leading to death. This occurs within 48 hours (Bellomo et al., 2004). AKI has a number of causes, including but not limited to low blood volume, possibly due to dehydration, trauma, restructured blood flow, obstruction of the urinary tract and other intrinsic causes (Han and Bonventre, 2004, Chertow et al., 2005a, Varghese et al., 2010, Siew et al., 2011). It is therefore a broad class of disease. It is known that AKI is associated with a change in the serum creatinine level in the blood stream (Addis et al., 1947, Perrone et al., 1992, Han and Bonventre, 2004, Siew et al., 2011, Mårtensson et al., 2012). This is all despite a multifaceted nature to this disease. However, the current use of serum creatinine as a diagnostic measure is hindered by limitations in diagnostic techniques. Proteomics has driven an insurgence to look into new biomarkers associated with AKI. Small rises in the serum creatinine have been associated with patients increased chances of death (Chertow et al., 2005b, Xue et al., 2006). There is a need for more specific biomarkers as serum creatinine is not a reliable marker due to variations occurring, due to factors like body weight or gender (Coca et al., 2007). Due to these limitations, it has been deemed a high priority in nephrology research (Mehta et al., 2007).

Seeing as AKI is of increasing occurrences in the population, a biomarker specific to acute events is necessary. Currently there is not a specific biomarker indicating an “acute vs. chronic” kidney condition. As well as a

marker that can differentiate between other kidney based conditions such as urinary tract infections and the effects of toxins. A number of markers have been discovered using a transcriptomics approach with associated proteomics downstream such as; NGAL neutrophil gelatinase-associated lipocalin (Supavekin et al., 2003, Mishra et al., 2003, Mishra et al., 2004), Although NGAL seems to have great prospects as a predictive biomarker, studies thus far have had small samples and uncomplicated populations. With regards to biofluid analysis SELDI-TOF-MS is the most prominent proteomic technique in the search for biomarkers, with the advantages of a high throughput and the ability to detect those low molecular weight biomarkers which gel-based technologies lack. Most studies concerning NGAL are performed using cardiac patients post-surgery due to the increased risk of AKI. Nguyen et al, 2005 employed this strategy and showed pattern analysis differences in AKI patients between normal and affected a patients (Nguyen et al., 2005). In addition to these markers, Interleukin (IL)-18 has also been identified as a putative biomarkers. After renal injury this marker is secreted into the urine prior to the occurrence of any renal dysfunction (Parikh et al., 2006). Kidney injury molecule 1 (KIM-1) is expressed in the proximal renal tubular cells in response to injury discovered in the up regulation in the rat model (Ichimura et al., 2008).

The causes of AKI are vast in nature. One such cause is nephrotoxicity. Zhou et al, 2006 investigated the urinary exosomes for this case. Using 2D-DiGE and subsequent MS identification they discovered 74 differentially expressed proteins. The two most promising of which, confirmed by western blotting, were Fetuin-A, which increased with AKI and annexin V, which decreased. An extremely promising biomarker for AKI is α 1-microglobulin and cystatin C which both were found to be a predictive indicator of AKI in urine. On top of which cystaitin C is stable in the blood and not affected by age, sex or weight. It therefore has many characteristic, that are seen as ideal in a biomarker (Koyner et al., 2008). However, NGAL has been shown to be detected significantly earlier post-surgery than cystatin C.

Currently, proteomics has produced 20+ “quality” candidature biomarkers for AKI (Siew et al., 2011) using a variety of techniques including ELISA, 2DE and LC-MS. The challenge still remains in the validation of these to a sensitive and selective standard for clinical diagnosis and prognostic purposes.

1.6.1.2.2 Chronic kidney disease Biomarkers

Chronic kidney disease (CKD) is the collective name given to a variety of diseases that amounts to the progressive decline in kidney function due to the rise in glomerular filtration rate of greater than or equal to 90mL/min, persistent for greater than 90 days (Vassalotti et al., 2007). Proteinuria is also a consistent symptom of CKD. The vast majority of forms of CKD are fatal seeing patients progress to renal failure, fortunately this is in the minority of cases. One major concern is that those who do progress to end stage renal failure, the diagnosis is late and often the renal disease is advanced by the time patients experience symptoms. This makes the need for specific biomarkers palpable. Biomarker discovery of a class of disease such as CKD is difficult at best due to the multifactorial nature of the underlying pathophysiology (Coresh et al., 2003, Vassalotti et al., 2007, Jantos-Siwy et al., 2008, Iwanaga and Miyazaki, 2010). CKD patients also have greater incidents of other cardiovascular incidents. (Stenvinkel et al., 2008). Biomarkers of CKD have been identified, some have more specific properties than others. serum creatinine is used but it is non-specific and presents itself across a number of kidney problems. One of the most promising set of biomarkers discoveries using a proteomics methodology is that of multiple profiles of collagen files to distinguish between two classes of CKD and diabetes. Diabetes, diabetic nephropathy, and non-diabetic proteinuric renal diseases. A strategy of multiple markers was employed for the distinction between the different types. It total 40 markers differentiated diabetic patients and healthy controls, 65 markers differentiated that diabetic patients have nephropathy with 97% sensitivity and specificity. The collagen type 1 fragments showed a reduced presence in urine, giving the prospect of a proteomic diagnostic method (Rossing et al., 2008, Jantos-Siwy et al., 2008).

Renal diseases are heavily associated with CVD and hypertension, due to the kidney's osmoregulatory function and interaction with vasopressin. In the next section biomarkers in hypertension and CVD are considered.

1.6.1.2.3 Hypertension and cardiovascular disease

Hypertension and cardiovascular diseases are a leading cause of death in western and developed countries. These classification of diseases are multifactorial in nature, with associations with both environmental factors such as; socio-economic status, diet and obesity and geographical locations as well as genetic factors (Burt et al., 1995, Carretero and Oparil, 2000). Hypertension has been clinically defined by the World Health Organisation, amongst others, as sustained raised blood pressure of $\geq 140/90$ mm Hg. Hypertension is one of the main symptoms and causes of cardiovascular disease which has been defined as the largest cause of death in the world. Correlations between hypertension and other diseases such as coronary heart disease, stroke, cerebrovascular disease, peripheral vascular diseases, type 2 diabetes and renal disease. Seeing as the prevention and treatment of hypertension may have wide implications in reducing the impact of these diseases, the requirement for diagnostic and therapeutic markers along with understanding the pathophysiology of these processes would be extremely beneficial. Proteomics and integration of the -omics has a significant part to play in the future of CVD and hypertension research (Thongboonkerd, 2005, Abdul-Salam et al., 2006, Elliott, 2007, Kuklinska et al., 2009, Delles et al., 2012). The topic of hypertension will be considered in greater detail in section 4.3. However, in the majority of patients suffering from hypertension there is no single cause. Therefore this multifactorial problem of “essential hypertension” lends itself towards a global unbiased approach which is offered by a great deal of proteomic techniques, as opposed to the traditional hypothesis driven techniques used in kidney and hypertension research. Whereas the genome of an individual will give an indication of “risk” associated with different conditions. The proteome and transcriptome will change dynamically with the stages of CVD and hypertension and any levied treatments. This gives the potential for a systems biology approach combining data for transcriptomics and proteomics to map changes and progression of diseases and treatments. However, this is posing a challenge in terms of marrying the data and mining the difference between “real” and false positive changes. This is compounded by the inconsistencies of similar studies within proteomics and transcriptomics before even trying to compare the datasets.

Increasingly, there has been greater attempt to profile multiple markers to identify risk potentials, however, these have had varying success. This varying

success is indicative of the lack of reproducibility and relatively small scale of studies undertaken. There is also a limited benefit to most studies as validation of markers is done so infrequently.

The pathophysiology of rare forms of hypertension is relatively well described in the literature due to the single gene mutation cause. However, essential hypertension poses a multitude of problems for both genomic and proteomic studies. It is a relatively simple process of diagnosing hypertension and comparing hypertensive patients with those not suffering is a straight forward prospect. However, due to the fact hypertension is considered a risk factor of an underlying CVD, gaining samples of known etiologic is difficult. The best approach using proteomics is to examine the global proteome by using an animal models of known genetic background. This will be discussed in greater detail in 4.3, however it is expedient to consider some of the discoveries of biomarkers in hypertension here.

There have been a number of studies that have identified markers in various forms of hypertension. One of the most accepted biomarkers of essential hypertension is the inflammatory marker C-reactive protein (Sesso et al., 2003, Wang et al., 2005). This is present in serum aldosterone, however, is not always present in all patients and therefore is not necessarily specific enough. This is especially the case when it comes to incident hypertension where the patient did not previously know. To overcome this limitation, Wang et al, 2007 used a multiple biomarker approach using “C-reactive protein (inflammation); fibrinogen (inflammation and thrombosis); plasminogen activator inhibitor-1 (fibrinolytic potential); aldosterone, renin, B-type natriuretic peptide, and N-terminal proatrial natriuretic peptide (neurohormonal activity); homocysteine (renal function and oxidant stress); and urinary albumin/creatinine ratio (glomerular endothelial function)” to investigate incident hypertension and if a pattern of biomarkers can be used to determine its presence (Wang et al., 2007).

Another set of prospective markers was investigated by Kuklinska et al, 2009 on the basis that hypertension causes endothelial damage. This damage is associated with the decrease in nitric oxide, prostacyclin, oxidised-LDL and peroxide. A relatively small scale study of 62 patients was undertaken. Although all prospective markers showed a significant degree in the

hypertensive patients compared to the healthy subjects, prostacyclin and oxidised-LDL had the best prognostic value (Kuklinska et al., 2009).

With regards to pulmonary hypertension, which is a rarer form of progressive hypertension, sufferers are subjected to an increasing vascular resistance in the pulmonary circulation. The diagnosis is often relatively late, thus reducing the effectiveness of prospective treatments. Diagnosis is often delayed and it would therefore benefit by biomarkers detection in biofluids. There has been a number of possible markers found including; Natriuretic peptides, Endothelin-1, Troponin T, Nitric oxide, Uric acid, Asymmetric dimethylarginine, and cGMP among others (Warwick et al., 2008). However, many of these markers are seen in numerous forms of hypertension and CVD. Thus the clinician would not readily be able to differentially diagnose a patient using these biomarkers alone.

Many of the current clinically accepted biomarkers for hypertension have not been generated using proteomic methodologies at all and although there has been some considerable success in obtaining biomarkers for hypertension, it is clear that a greater wealth of information and prospective candidates are needed to be taken onto validation. Therefore in this investigation in Chapter 4 biomarkers for hypertension are investigated via an unguided global proteomic analysis. This is in conjunction with other studies using complementary proteomic approaches, in order to overcome some of the limitations described in the literature. It is hoped, that by using an unguided approach, new novel biomarkers may be discovered and also that the methods developed may be employed in numerous different disease states.

One of the main current strategies in proteomics when investigating cardiovascular diseases is to compare patients with confirmed disease with normal tissue. Coronary artery disease or CAD is relatively easy to confirm with angiography. These patients can then be compared against patients without CAD. This is considered a good model due to the link between cardiovascular health and CAD (Zimmerli et al., 2008a). This study involving 88 patients revealed a set of biomarkers indicating the possibility of CAD. Healthy controls showed significantly lower high-density lipoprotein levels than CAD sufferers. Additionally, C-reactive protein was also lower. However these are indicators of a number CVDs. This study was performed using CE-LC-MS to generate urine profiles and identifications. 47 patients carried onto

the blind assessment study. A strength of this study is that centre specific bias was ruled out by using two control groups from geographically different locations. CE time based traces were used to generate a polypeptide signature which was successfully used to distinguish between healthy and control patients. Interestingly, those sufferers who exercised more showed a healthier profile of peptides. It was noted that, when using drugs, patients exhibited a more variable profile and therefore this could present a significant barrier to more general clinical use. The polypeptide patterns found were predominantly fragments of collagen (Zimmerli et al., 2008b, Danesh et al., 2004).

Cardiovascular disease or CVD is a huge area of study, but they are closely related to diseases of the kidney and hypertension. Patients who have CKD have a predisposition to various forms of CVD such as stroke and peripheral vascular disease. Biomarkers are necessary as studies have shown that current treatment strategies do not increase the mortality of patients suffering from a combination of CKD and CVD (Stenvinkel et al., 2008). Other similar studies are described in numerous journal articles (Anderson, 2005, Ridker et al., 2000, Mayr et al., 2006, Blumenstein et al., 2009).

Despite this wealth of new information regarding candidate markers, there is a notable lack of markers that have been validated as novel and moved onto assay development. This is a severe barrier and a bridge that needs to be built between candidate biomarker discovery and clinical assay development.

1.6.1.2.4 Urinary proteome and biomarkers in renal proteomics

Urinary proteomics has been occurring for more time than the term proteomics has existed, however, recently there has been a massive interest in searching for biomarkers in urine. Therefore urinary proteomics within renal proteomics is wide and deep in nature, covering the physiology and pathophysiology of disease and the search for markers in such areas as; Renal cancer, acute kidney injury, chronic kidney injury, diabetic nephrology, urological cancer, glomerular disease, IgA nephropathy and membranous glomerulonephritis to mention but a few. In addition, there is urinary analysis in a number of different types of cardiovascular disease. Therefore, it would be superfluous to mention every possible area where urinary proteomics has been employed, as it is vast; however in this section some of the more recent applications will be considered. Within proteomics, urine analysis has been performed using 2DE and DiGE (Sharma et al., 2005), CE-MS (Coon et al., 2008) and numerous quantitative methods including SELDI-TOF-MS (Cadieux et al., 2004) , iCAT, iTRAQ™ and label free (Yang et al., 2011b).

The Holy Grail for clinical diagnostic is to examine a non-invasive sample such as urine, to determine a specific diagnosis. This has driven the requirement to examine urine as the principle biofluid in renal proteomic and where the greatest number of biomarker studies concentrates on. Urine can be prepared in a number of ways for a proteomic analysis. This might be as a liquid form, desalted, or in a pellet form containing the urinary exosomes. Urine, however, does present difficulties regarding the stability of protein content which can depend on voiding time of day, gender, diet and exercising. This presents a serious problem in gaining reproducibility (Kentsis, 2011).

The first proteomic based experiment on human urine was performed in 1979 (before the coining of the phrase proteomics). It was the first in a number of 2DE mapping experiments (Anderson et al., 1979). Since then there has been 1000s of articles on the urinary proteome.

One area proteomics has been employed in is the biomarker discovery of renal cancer markers in urine. This area is very important, as renal cancer is relatively unresponsive to more traditional cancer treatments such as chemotherapy and radiotherapy. However, currently there are no renal biomarkers for cancer of the kidney in clinical use. This is due to the lack of specificity and complexity of the disease. The most common form of renal

cancer is renal cell carcinoma (RCC) (Craven et al., 2013). There has been some success, however, compared to other aspects of oncoproteomics, RCC remain a challenging area. A combination of fragments of uromodulin were found to be present in human urine and distinguished between control and diseased patients, both in urine and serum. It was also shown that malignant and benign masses have the prospect of being differentiated (Bosso et al., 2008). Additionally, it was also demonstrated that RCC sufferers showed the expression of two forms of Mn-superoxide dismutase (SOD), compared to only one variant in healthy controls (Raimondo et al., 2012b, Sarto et al., 2001). This is in contrast to other forms of urological cancers which have seen some greater success. In bladder squamous cell carcinoma (SCC), a protein called psoriasin was seen to be expressed in only patients suffering from bladder SCC (Rasmussen et al., 1996). Another investigation in bladder cancer using iTRAQ™ has been performed by Chen et al, 2010. Three bladder cancer subgroups were compared to control samples. A total of 638 proteins were identified. Out of the proteins identified, apolipoprotein A-I (APOA1), apolipoprotein A-II, heparin cofactor 2 precursor and peroxiredoxin-2 were identified as being significantly elevated in the bladder cancer group. Further, APOA1 was confirmed using ELISA (Chen et al., 2010b).

In addition to iTRAQ™, capillary electrophoresis coupled to mass spectrometry (CE-MS) is a method that has seen considerable success in the area of urinary proteomics. A study comparing 230 patients with chronic renal disease and 379 controls, resulted in over 634 peptide being identified as differentially expressed between the two groups. Approximately a third of these markers were sequenced and 144 validated. A large proportion being down-regulated collagen fragments(Good et al., 2010).

Another area of considerable interest is that of diabetic nephropathy. Diabetic nephropathy is the biggest contributor to end stage renal disease, which is usually fatal. As with other types of kidney disease, diabetic nephropathy (DN) is characterised by an increased presence of protein in the urine (Alkhalaf et al., 2010). The current clinically accepted marker of DN is the occurrence of microalbuminuria (MA). It is however still an unspecific marker and can only be seen in about 40% of type two diabetic patients. Although a number of studies have been undertaken, a review from Ameer et al, 2010 highlights the need for validation steps, as less than half the studies observed did any kind of

validation (Randa et al., 2010). Diseases like DN pose another issue. As DN is a collection of diseases, finding a specific biomarker is not necessarily likely. This is prevalent throughout biomarker research. As more information is discovered about disease, the classification get more and more differential, thus meaning more specific markers are needed, however, some attainment is occurring in this area. A study from Papale et al, 2010 described differential markers for a form of DN called glomerulosclerosis. They managed to identify and validate ubiquitin and 2-microglobulin as specific markers. It should be noted however, that proteins like ubiquitin, are by their nature likely to occur in many other disease states (Papale et al., 2010). There has already been considerable interest in DN (Alkhalaf et al., 2010, Isabel Padrão et al., 2012, Mischak and Rossing, 2010, Rossing et al., 2008, Zürbig et al., 2009), as this is expect to rise due to the continuing rise in diabetes (particularly type 2) in westernised countries. In addition to biomarker discovery in DN, it is also looking more likely that proteomic methodologies may be used directly in a clinical setting. This is due to the fact there is evidence to suggest processes such as CE-MS of urine can differentiate patients using peptide profiles (Alkhalaf et al., 2010).

Urinary proteomics is also engaging in examining the exosomes within urine. An exosomes is a membrane bound vesicle which originates from the epithelial lining of the collecting ducts and urinary tract. They are obtained by differential centrifugation of urine. A number of applications are described in the literature (Zhou et al., 2006, Hoorn et al., 2005, Moon et al., 2011, Pisitkun et al., 2004). The exosomes can contain nucleic acids which can be used as biomarkers (Miranda et al., 2010) or protein and peptide fragments. This isolation of these exosomes allows for the reduction in complexity and increased stability of the protein content in the excretion when compared to urine, making it an excellent candidate for superseding urine as a sample of choice.

Proteomic techniques have also been employed in renal disease, which are associated with other ailments. Systemic lupus erythematosus (SLE), for example, is known to cause nephritis. In particular Glomerulonephritis, which is the most common manifestation in SLE, has a diminished 5 and 10 year survival rate. Current diagnostic techniques and therapies are unsatisfactory, therefore, finding diagnostic and therapeutic biomarkers would be

advantageous. Monocyte chemoattractant protein-1 (MCP-1) has been shown to indicate glomerular inflammation in mice (Hasegawa et al., 2003), which is higher in those with active renal damage, however, as it is also a marker for intestinal inflammation, it does not provide an absolutely definitive diagnosis of renal dysfunction. Elevated levels have also been shown in humans (Kiani et al., 2009). Another disease that causes serious renal problems is Fabry disease. Elevated levels of globotriaosylceramide (Gb3) have been identified as being present in patients with Fabry disease as glomerular permeability is thought to increase. However, the source of this biomarker is in question and has not as yet been linked to a specific renal problem (Schiffmann et al., 2010).

As can be seen the range of areas that urinary proteomics is involved in is vast. However, using tissue in renal research is a growing area. In this next section tissue proteomics is examined within the context of renal research

1.6.1.3 The use of tissue in renal proteomics

Within kidney proteomics, the majority of citations focus on urine as the primary source of material for investigation. There is, however, a slow but significant change in this trend. Over several decades of clinical investigation tissue has proven useful in viewing pathological changes. It gives spatial awareness of markers and many discoveries have been made using immunohistochemistry and histological staining. Proteomics, until recently, did not have a technique to keep this spatial resolution, but with imaging mass spectrometry coupled to the traditional techniques, the use of tissue as a material is gaining ground in proteomic nephrology.

Naturally, the topics of interest for tissue and cellular proteomics are those similar to proteomic analysis of urine such as; urological cancers (Craven et al., 2006), problems associated with; diabetes (Bugger et al., 2009), AKI (Reeves et al., 2008), nephrotoxicity (Lei et al., 2008), cancer (Castronovo et al., 2006) and CKD (Perco et al., 2006).

With regards to whole kidney expression proteomics, there are considerable advantages to using a proteomic approach as opposed to a hypothesis driven approach. The main being that it allows for an unknown hypothesis to be discovered. This is particularly the case when a certain disease is not yet linked to a causative region of the body. Thongboonkerd et al, 2002 described how alterations in the Renal Kallikrein Pathway during Hypoxia caused by sleep apnoea induced hypertension (Thongboonkerd et al., 2002). This was

investigated using a rat model. Using a similar approach using a murine model Thongboonkerd et al, 2004 also demonstrated the increased deposition of elastin in diabetic nephropathy (Thongboonkerd et al., 2004).

One severe limitation in using tissue is the multitude of cell types. In the kidney this is broadly defined into two regions; the medulla and cortex. Homogenisation of whole kidney may not lead to the most successful biomarker discovery investigation, as up or down regulation could be considerably masked by homogenisation of multiple cell types. As yet, the most common solution is to separate the kidney into the two stated gross anatomical regions, however as proteomic technologies get more sensitive the future is almost certainly leading to the isolation of individual cell types from tissue. This of course has advantages but currently the limits of technology make this approach too challenging. Therefore in Chapter 4 macro dissection was performed to separate the kidney in the medulla and cortex to allow comparison with other studies in the literature.

Xu et al, 2005 reported a study utilising a 2DE approach to profiling normal kidney proteome expression, that highlights the need for tissue type segregation. The kidney was separated into three distinct regions; the cortex, medulla and glomerulus. They showed substantial differences in the three cell types with around 50% of the spots detected in each structure being differentially expressed. Separating the cortex and medulla is relatively straightforward, due to their anatomical distinct regions, however, separating the glomerulus increased the complicity of this study, increasing wash and filtration steps, thus risking the loss of any solubilised proteins. Therefore, for this study it was decided to only separate into the medulla and cortex regions (Xu et al., 2005a). A preceding study supported Xu et al, 2005 findings but found fewer differentially expressed proteins between the cortex and medulla regions (Arthur et al., 2002). In future, techniques which employ separation of cell types with more precise and less destructive methodology than washing protocols are needed. One successful technique that is often underutilised, therefore underdeveloped, is laser micro-dissection. Although it has been used for a number of years now, there is limitation regarding loss of sample due to eppendorf static, preventing tissue collection. Dealing with such small pieces of tissue is exceptionally challenging. However, it does show great potential to be used in conjunction with sensitive techniques such as DiGE saturation

labelling (Banks et al., 1999, Sitek et al., 2005b). There is a gradual movement toward the macro dissection and LMD of tissue in proteomics research to prevent the averaging effects of homogenisation. This is no less true in renal proteomics.

Another benefit emerging from the development of methods using tissue as the main source of investigative material is the unlocking of a fixed tissue banks. There has been some small success with respect to this but it is an emerging field with huge potential (Nirmalan et al., 2008) both in understanding the kidney proteome and in biomarker discovery (Nirmalan et al., 2010)

There is a lack of exploration of the kidney proteome using renal tissue connected to transcriptomic and genomic studies. This spawn the necessity in this study to look into biomarker discovery using kidney tissue in Chapter 4 “Biomarker discovery and the assessment of variation in the proteomic profiles of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model.”

1.6.1.4 Therapeutic targets in renal proteomics

In addition to biomarker discovery, an area under development is the search for therapeutic targets. These are cellular elements, such as extracellular membrane proteins, which can lead to a therapeutics effect. However, currently within proteomics this is an emerging field and as such has reported some limited success (Thongboonkerd, 2004). Such as using elastin in the renal elastin–elastase system and the use of collagen fibre fragments in diabetic caused CKD (Thongboonkerd et al., 2004, Rossing et al., 2008). However currently there is a lack of specificity.

At present the majority of the stronger potential therapeutic candidates come from genomic studies. Wang et al, 2004 describes the overexpression of spermidine/spermine N-1-acetyl-transferase (SSAT) in cultured kidney cells which has the detrimental effect of decreasing cell growth and increases the occurrence of kidney ischemia-reperfusion injury. The knowledge of this gene target could potentially lead to drug development of the SSAT gene as a therapeutic target (Wang et al., 2004). The integration of this genomic data with proteomic data remain sizeable (Stojnev et al., 2009) .

It is clear though, that this will be of greater importance when the proteomic technologies and techniques are more refined, along with the integration of the –omics data-sets.

There has been promise shown in this field by using quantitative multiplexed strategies, by combining gel-assisted digestion with iTRAQ™ in the characterisation of membrane proteins. Han et al, 2008 reported considerable success in using this relatively high throughput quantitative methodology to improve the coverage of the proteome viewed. A number of membrane proteins were quantified, which could be used as therapeutic targets for autosomal dominant polycystic kidney disease. Although it should be noted this was using HeLa cells, not kidney tissue.

Another promising study illustrated the connection between biomarker discovery and the detection of therapeutic targets. Holy et al, 2006 used a DiGE methodology to quantitate and discover biomarkers in sepsis induced ARF (now known as AKI). Sepsis is a big problem in hospitals, particularly post-surgery, with renal failure being a common complication of sepsis. Holy et al, 2004 used a rat model to discover urinary biomarkers and the possible therapeutic drug targets of albumin, enzymes in the brush border in particular meprin-1-alpha and serine protease inhibitors. The meprin-1-alpha inhibitor actinonin prevented ARF in the older mice. Also it showed that serum creatinine levels varied across animals and only 24% of those rats that suffered ARF had an elevated serum creatinine level (Holly et al., 2006).

1.6.2 Challenges for proteomics in studying the kidney proteome

The challenges that face renal proteomics are similar to those that confront any proteomic investigation. They can be broken down into technological, biological, statistical and practical issues.

Many of the technological considerations have been covered in section 1.4. Currently, there are a large number of renal studies in proteomics that rely on a gel based methodology. Therefore, many of the practical and technological considerations are similar to other studies. With gel based approaches being relatively labour intensive, there has been a move towards using more automated online approaches with mass spectrometry such as SELDI-TOF-MS and other quantitative methods (Janech et al., 2007). In addition, it is likely that in the near future there will be more functionally based proteomic experiments to characterise protein-protein interactions, complexes and PTM. However, as yet, the expression proteomic approach is currently one that predominates. Gel-based analysis has a wealth of information available for comparison and is still likely to be an integral way of imaging the proteome, particularly with the powerful technique of DiGE, but “real” high throughput proteomics using gels is in need of greater software automation and more powerful statistical applications to minimise the limitations of the statistical analyses (see section 1.7). There is currently a real change of “over fitting” the data to the experiments.

There is also an argument for the need for a more defined availability of tissue. Often experiments are performed using tissue or biofluids of patients, which have a late stage of disease. This is useful; however, a proposal of sample banking has been suggested in order to track the progression of disease from high risk patient through to when some of these patients develop the disease. This could provide powerful data on how the proteome changes over time. There are ethical barriers to this kind of approach, not to mention financial implications for the collection and storage of massive tissue and biofluid banks from possibly healthy patients.

Biologically, the kidney is made up of several cell types and has nephron structures which bridge the cortex and the medulla. Therefore, there is debate on how useful whole tissue experiments are. The challenge of extracting specific cell types is going to be paramount in gaining higher quality proteomic information.

On top of sampling challenges, there are also technological limitations. In the use of mass spectrometry to gain identifications, there are fundamental limits to the cycle time of 10Hz, this is combated by running multiple runs of the same sample. However, this is expensive in time and is biased towards abundant proteins. The use of exclusion list is therefore often done, but this requires prior knowledge and is not really appropriate for an unguided approach. Other challenges include; selection bias in the sample selection, dynamic range of the proteins in the sample and statistical issue in multidimensional data (Kentsis, 2011).

Renal proteomics has come a long way since its inception, however there is a need to explore and develop proteomic methodologies and to elicit biologically significant marker in hypertension. This process of biomarker discovery is explored in Chapter 4 “Biomarker discovery and the assessment of variation in the proteomic profiles of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model.”

1.7 Statistical Analysis

The field of statistics is, of course, of central importance to the conclusions which can be drawn from scientific experiments. However, due to the enormity of the field it is beyond the scope of this introduction to cover every aspect, as such an overview is given of the main techniques employed in this investigation, all of which are well described and accepted in the field. Despite the central importance to the sciences, statistics are often misused, misinterpreted or simply the limitations ignored. In this section some of the limitations of the techniques employed in this study are explored. Some of these limitations are inherent in the statistical model, some are due to experimental design.

1.7.1 Types of error

Statistical significance tests whether something has or has not occurred by chance alone. Errors by chance manifest in one of two ways; type I errors or type II errors. Type I errors are known as false positives. That is to say a result which is given as statically significant within the statistical test of choice, when it is by chance alone. Type II errors are known as false negatives. These are not given as a significant test result but are in fact biologically significant. Chance errors like these can be minimised by increasing sample sizes to reduce the chance of error. It is also known that running statistical tests multiple times results in a greater chance of false positives. This can be demonstrated with a simple example. If you have 10 balls in a lottery (number 1-10) what is the chance of getting a number 1? It is of course 1/10 or 10%. However, what is the chance of getting a number one if you do the same draw 20 times (replacing the ball each time)? It works out at 88% ($= (1-0.092)^{20} * 100$). This is why running tests multiple times, such as Student's t-test, increases the number of false positive results. The multiple testing problem is discussed in section 1.7.2.

Another danger to the validity of an experiment is external validity. This needs to be considered when trying to transfer results into wider applications. Biomarker discovery is susceptible to this as discovery of putative biomarkers in say animal models or certain conditions are not necessarily applicable to wider applications in human models. Equally, certain diseases have a high association with other factors. For example lung cancer patients often are smokers and therefore biomarkers discovered may be associated with smokers

as opposed to lung cancer sufferers. Another particular problem in biomarker discovery platforms (such as DeCyder) is the over fitting of data due to multivariate techniques like PCA or many forms of the hierarchical analysis.

Another form of error which can occur is bias. Bias can be experimentally based or analytically based. Experimental bias is difficult to eliminate and minimise, often because experimenters are unaware of doing it. To minimise it, researchers need to be rigorous in following protocols and treating samples to same and randomisation when running samples is important (Chich et al., 2007). In DiGE experiments, bias can occur due to technical reasons such as differential labelling. Therefore bias can be minimised in this case by performing a dye switch, therefore it is important to be aware of inherent bias within techniques as well. Bias can occur in all sorts of ways including; analytical bias by round up or down or choosing to use the mean or the median values, observer bias, positive or negative results bias from literature and broadly experimenter bias.

1.7.2 The Multiple testing problem and false discovery rate.

Although univariate testing is relatively easy to use, a problem exists in practice when 1000s of hypotheses are being tested simultaneously. This problem is known as the multiple testing problems. When performing hypothesis testing 1000s of times, the chances of discovering false positives increases. This problem arises due to the interdependence of proteins in the in vivo state, whereas the statistical test assumes independence. A correction can be made for this, but is often ignored in the proteomics fields. In genomics, due to the high use of microarray experiments, this problem has been given a great deal of consideration. If the most is to be gathered from proteomic data in the future, along with increasing the validity of conclusion, more consideration is necessary in order to reduce the number of false positive discovered. Thankfully this idea is gaining significant ground with proteomic researchers performing DiGE experiments. It is possible with unadjusted P-values and $\alpha=0.05$ significance levels that 1 in 20 tests are expected to give false positive results and as the number of tests increases, the number of false positives increases. One method of dealing with this is to use the Bonferroni correction (a form of FWER – Familywise error rate), this has been considered to be somewhat conservative in its approach and has lost favour with many researchers. To consider why this correction method is conservative, an

example is prudent. The reduced significance level using Bonferroni's adjusted formula of $(1-\alpha)^{1/n}$ is 1.71×10^{-5} (where you have $n=3000$ protein spots and α is given as the chance of type I errors or 0.05). This is a reduced significant level which could reduce data sets to a situation with no significant results. Also with decreasing type I errors, it is likely that type II errors will increase.

A method which is generally considered more favourable and practical, is a false discovery rate (FDR). The FDR approach allows for a trade-off between power and errors. The FDR allows for a declaration of the false positives expected as part of the p-value, thus when performed it adjusts the p-value for every test run. This has been performed for a number of years on microarray data and is making head way into proteomics, which is now dealing with increased size of data sets, once only seen in genomics. DeCyder 2D, GE Healthcare now has an option built in to run an FDR which adjusts the p-values. The algorithm used was described by Benjamini and Hochberg, 2000. Although it was first described by Benjamini and Hochberg, 1995 (Benjamini and Hochberg, 1995, Benjamini and Hochberg, 2000, Karp et al., 2007). It has been paid relatively little attention as a whole in the proteomics community but is going to become of increasing importance as the size of data sets increase. This has been an issue for genomics and the use of microarrays for some time and has been discussed in microarray literature. This proves to be a rich source of information for the proteomic community to take advantage (Karp et al., 2007, Liu and Hwang, 2007, Allison et al., 2006, Pawitan et al., 2005a, Nadon and Shoemaker, 2002, Pawitan et al., 2005b)

1.7.3 Statistical Analysis in Proteomics and DiGE

With regards to the statistical analysis of 2D gels or DiGE the most commonly used analytical software is produced by GE Healthcare and is called DeCyder 2D (version 7.0). However, there are a number of alternatives. This application has different modules for different analytical tasks, depending on experimental design; this is discussed in section 2.1.2.7 (Fodor et al., 2005) This software allows the assessment by univariate or multivariate analysis of DiGE data. However, the statistical tests available tend to be of the standard parametric tests such as t-tests and ANOVA. These have been commonly applied to many scientific fields and proteomics is no exception. In order for parametric tests to be run an assumption of normal distribution and

homogeneity of variance is necessary. The majority of 2DE and DiGE studies make little attempt to assess their particular data for these assumptions (Dautel et al., 2010). Additionally, many studies look for fold (typically 1.5 or 2x) change as an indicator of significance, however, this is not necessarily a robust measure of change. There have, however, been some researchers who have investigated the use of statistics in DiGE. This is especially since proteomics is encountering some of the same issues as DNA microarray data analysts have been dealing with for some time. Namely issues such as multiple testing and normalisation (as discussed in 1.7.2) (Fodor and Nelson, 2005, Kreil et al., 2004, Chich et al., 2007, Taylor et al., 2003)

One such study that assessed some of these important assumptions for which most studies simply assume, was performed by Karp et al, 2005. In this study it was found that the assumption of normality held as approximately 5% or less spots fell under the normality threshold. The assumption of homogeneity was also found to be a robust one. Equally, this study also suggests technical replicates are independent of sample and variance, due to the technique is minimal. Karp et al, 2005 suggests many DiGE experiments are therefore a cost vs. significantly balance. This is backed up in general across biomarker discovery, as the discovery phase is generally the cheapest compared to validation, giving credence to the top heavy nature in favour of discovery research compared to clinical assay development. The paper provides a powerful framework for DiGE users to show greater validity in their results by extending data analysis and prove the use of correct statistics. This is important, as currently there is no availability in the DeCyder 2D 7.0 platform to perform non-parametric tests (Karp and Lilley, 2005) in the BVA module at least. It would be suggested to do this for every new sample type to make sure that the correct statistical test is employed and therefore minimising false positive results. This approach, or those similar, is starting to make its way into the workflow of the standard DiGE gels, however, the vast majority still do not engage with the data much more than superficially and still rely on using an arbitrary fold change result as a cut for DiGE or 2DE data.

Certainly the use of internal standards in DiGE has increased the reproducibility. The use of standards and therefore the need for normalisation raises issues about how to normalise data. Keeping et al, 2010 investigated various normalisation techniques and their ability to reduce background noise.

They discovered that the use of an internal standard was the best way to influence noise in a DiGE experiment, equally variance was maximally reduced by using an internal standard and performing any type of normalisation method (as opposed to none). The recommendation of the authors were to use various methods before deciding on the chosen method, however this is likely to lead to a far larger analysis time. This is particularly true in the proprietary analysis software such as DeCyder 2D, GE Healthcare, as analysts would be required to engage a complication process of normalisation and match using different programs. Thus it could be argued, that the benefits would be outweighed by time and costs, coupled with the risk of user error (Keeping and Collins, 2010). Despite these questions that arise, the use of an internal standard has improved quantitation using 2DE and has been shown to reduce technical variation immensely, particularly in comparison to biological variation (Zech et al., 2011) supported by Karp et al, 2005. It was further argued that by dropping technical replicates and replacing them with biological replication, the statistical power of the experiment can be increased (Engelen et al., 2010). However, Karp et al, 2009 further investigated the idea of biological variation reduction by using a strategy of sub pooling and thus show that this technique does not lead to systemic bias, but the balance of technical and biological variance is dependent on tissue type and the specific experiment (Karp and Lilley, 2009) . Therefore running technical replicate may still be an important step in the DiGE workflow.

An alternative to assessing normalisation and homogeneity or heterogeneity of variance has been proposed by Anderson et al, 2006 in the search for biomarkers. They use a strategy called relative expression reversal (RER). This is a set of machine learning algorithms which can assess relative protein expression levels. The algorithm has a classifier that accounts for monotonic expression variability which eliminates the need for normalisation and dye swapping, a considerable advantage in the reduction of experimental bias. A process of learning occurs via a set of validative standards and normal and disease samples, which returned a generalisation score of zero. They managed to discover that tropomyosin isoforms 3 and 4 and a-enolase as the best method for discriminating between normal and cancer samples. This method has been applied to microarray data with similar success (Fodor et al., 2005). Some of these issues highlighted are not just of concern to DiGE experiments. Any proteomics method that attempts large numbers of protein quantitation is

vulnerable to multiple testing. Also the assumption of normality and homogeneity of variance are often also taken for granted by many experimenters.

So using the recommendation of Biron et al, 2006, this study has assessed the normality of the DiGE spot data, the homogeneity of variance and applied a correction of multiple testing using a false discovery rate in order to validate the use of t-test and ANOVA tests (Biron et al., 2006).

1.7.4 Univariate Analysis

1.7.4.1 Student's t-test

The Student's t-test is one of the most commonly applied tests in science. It was first described in *Biometrika* in 1908 anonymously (Student, 1908) with the author A Student hence the acquired name. The principal of the Student's t-test is to show any significance between sample means. The Student's t-tests employed in this study were calculated within the DeCyder 2D software developed by GE Healthcare. The form of t-test employed depends on what is being tested. In this investigation, an independent t-test was used to compare two separate sample population means from each other. The assumption for using the t-test is that the data is normally distributed with homogeneous variance. This was calculated and results shown in chapter 3 and 4.

The Student's t-test calculated in DeCyder 2D, GE healthcare was a two-tailed test in which the program calculated the p-value. The formula for the Student's t-test and degrees of freedom is given in Equation 3 and Equation 4. In Equation 3 the term $\mu_a - \mu_b$ is the two sample means and the δ_{a-b} is the deviation term. The degrees of freedom are given by the sum to the two samples - 2.

$$t = \frac{\mu_a - \mu_b}{\delta_{a-b}} \quad (\text{Equation 1})$$

$$df = N_a + N_b - 2 \quad (\text{Equation 2})$$

1.7.4.2 Analysis of variance (ANOVA)

The use of a t-test is useful but limited in the number of populations that can be compared. This was overcome by using a technique first described in 1918 by Fisher, "The Correlation Between Relatives on the Supposition of Mendelian Inheritance" (Fisher, 1918). This has evolved and is collectively termed analysis of variance, although there are different forms. It employs much of

the theory of Student's t-tests but allows the inclusion of, at least in theory, any number of populations compared to each other. Once again the analysis of variance (one-way and two-way) was performed using the DeCyder Software developed by GE Healthcare. The DeCyder ANOVA algorithms are implemented using multiple linear regression analysis to handle unbalanced data sets. The null hypothesis for the ANOVA is that there is no difference between the group means of proteins spot data. It should be noted that it indicates a difference in means between groups but does not give information of which groups are different. The ANOVA is performed by calculating the F-ratio. This F-ratio is the difference ratio of between group variability and within group variability. This F-ratio is given by the formula:

$$F = \frac{\sum_i n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1)}{\sum_{ij} (Y_{ij} - \bar{Y}_i)^2 / (N - K)}, \quad (\text{Equation 5})$$

The \bar{Y}_i denoted the sample mean within the i th group and n_i gives the number of observations in the i th group. K signifies the total number of groups and \bar{Y} describes the overall mean. If this F-ratio is above a critical value then the null hypothesis is rejected. The F-value can be converted into the p-value by integrating the area of the F-distribution curve to the right of the f-ratio value. This conversion to a p-value allows comparison to a given alpha value (e.g. 0.05). Once again this p-value can be adjusted using a FDR.

1.7.4.3 Normality Testing

One of the assumptions that the Student's t-test and ANOVA make is that the data is normally distributed. Therefore, in order to use these tests it is important to assess the data for normality. There are a number of ways of doing this, however, in this thesis the Shapiro-Wilk goodness to fit test was employed. This method uses the calculation of a W statistic that tests whether a sample come from a normally distributed set. Therefore, it tests the null hypothesis that the samples, $x_1, x_2 \dots x_n$, come from a normally distributed population. The W statistic is calculated with the formula;

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (\text{Equation 6})$$

The x_i denotes the ordered sample value, a_i is a constant calculated from the means, variances and covariance's of the order samples. The term \bar{x} , is calculated by the formula $\frac{(x_1+x_2\dots xn)}{n}$, which gives the sample means, where n is the sample size. From this, SPSS, IBM (Version 17.0.1) converts the W statistic into the more commonly understood p -value (Shapiro and Wilk, 1965). If the returned p -value is lower than the assigned alpha level (often 0.05), then we reject the null hypothesis of normality.

1.7.4.4 Homogeneity of variance testing

Another assumption that needs to be tested before running parametric tests is homogeneity of variance. This is important, as if means are to be compared there needs to be a similar degree of variance between samples. The method employed in this thesis to assess homogeneity of variance is the Levene's test. The Levene's test tests the null hypothesis that the population variances are equal and uses the sample variances to do this. If the returned p -value falls below the set alpha level, then the null hypothesis of normality is rejected. Once again a W statistic is calculated with the formula;

$$W = \frac{(N - k) \sum_{i=1}^k N_i (Z_i - Z_{..})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - Z_i)^2}, \quad (\text{Equation 7})$$

The N_i denotes the total number of samples in the i th group, N is the total number of samples and K is the total number of groups that the samples belong to. Y_{ij} is the value of the j th sample within the i th group. The Z_{ij} term is the difference between Y_{ij} and the mean of the i th group. $Z_{..}$ is given by the mean of all Z_{ij} and Z_i is the mean of Z_{ij} for the i th group. From this, SPSS, IBM (Version 17.0.1) converts the W statistic into the more commonly understood p -value by running a quartile of an F -test distribution using the degrees of freedom and alpha level.

1.7.5 Multivariate Analysis performed in the extended data analysis module

1.7.5.1 Principal Component Analysis (PCA)

Principal components analysis (PCA) is a technique that has been used in genomics research, when searching for candidate genes and expression of genes (Kawano et al., 2010, Tanaka, 2009, Shimada Y, 2009) . It has also been employed in proteomic analysis with regards to LC (Cserháti, 2010, Stasiak J, 2010) It has now been developed into the DeCyder Software,

developed by GE Healthcare for the use with spot detection and analysis for DiGE experiments. It is part of the extended data analysis module (EDA), with the prospect of clustering groups of proteins or peptides for biomarker discovery and group validation. Therefore it is with the use of this module that PCA will be employed in this investigation.

PCA was developed by Karl Pearson in 1901 (Pearson, 1901), and is an exploratory statistical method for mining data to display relationships between groups. It is known as an orthogonal linear transformation. This means that it splits the relationship of data into principal components. The first component has the greatest degree of variance between groups. Therefore groups that lie closer together have less variance from each other. Each component thereafter shows lesser variance between groups (Jolliffe, 1986, Werner Dubitzky, 2007). The PCA generated in this thesis was performed using the EDA module software from GE Healthcare. It is known as an orthogonal linear transformation which causes a transformation in coordinates to simplify the visualisation of large data sets. This can be defined mathematically. The idea is to reduce a higher order dimensional space, \mathbb{R}^p to \mathbb{R}^q , for a given set of data points, $x_1, x_2 \dots x_n$, we reconstruct the data using;

$$f(\lambda) = \mu + v_q \lambda \quad \text{(Equation 8)}$$

This is where $f(\lambda)$ is the new dimensional space of \mathbb{R}^q , μ is the mean of the dimensional space \mathbb{R}^p , v_q is a $p \times q$ matrix with q orthogonal unit vectors and λ is low dimensional data point that is being projected. Choosing μ , v_q and λ can be done in a number of ways depending if you want to limit error or maximise variance.

1.7.6 The problem of missing values

Two-dimensional gel electrophoresis and DiGE can experience a high insistence of missing spot data, which is a problem across other proteomic methodologies too. There are multiple strategies for dealing with this situation, but in much of proteomic analyses it is simply ignored. One strategy to deal with this could be to concentrate on a limited number of instances where all the data is available across all the gels replicates and condition. This has the advantage of maintaining a complete set of replicates and make statistical analysis easier. However, it also means that you reduce significantly the number of spots that you include in an analysis, and is seen as a conservative

method for dealing with missing values, which is likely to mean missing spots with significant expression changes. Equally, there is a high likelihood that even complete data is due to the presence of mismatches in the software or variants in abundances or even an absence of protein, so false positives cannot be eliminated this way. This conservative approach also may not be appropriate for “fishing style” experiments, where missing markers may be considered more detrimental than a gain in false positives. Dealing with mismatches globally is fraught with difficulties and requires spot by spot inspection. This possesses practical difficulty in gel data, which could contain 1000s of spots on each gel. This makes it very labour intensive. Other alternatives exist. Fortunately, a complete set of data is not necessary to perform differential analysis. Some scientists propose replacing missing values with zero, but this assumes a lack of protein. Additionally, others decide to use the nearest neighbour approach. A statistical approach is to use a simulation; for example a combination of the lowest median values and the variance. However, caution is necessary depending on the extent of the missing data as this may produce bias in any results (Herbert, 2001). A number of methods for dealing with missing values in DiGE have been assessed before multivariate analysis including; nonlinear estimation by iterative partial least squares (NIPALS) algorithm or imputing them by using either k-nearest neighbour or Bayesian principal component analysis (BPCA) (Pedreschi et al., 2008b). BPCA turned out to be the most successful method in replacing missing values, however, the scientific and statistical community is yet to decide on the best strategy for this. It is, however, recognised that this question will become of more importance seeing as there is greater and greater sized data sets being generated (Chich et al., 2007). A common mode for replacing missing values is to use the algorithm K-Nearest Neighbour (KNN). This can be used in pattern recognition and therefore employed in replacing missing values. It is used to classify cases based on their similarities to other cases. A K value is set which is the number of neighbours you wish the algorithm to use. In this study the K number is set to automatic as the number of nearest neighbours is likely to change for each spot. KNN uses a weighted average of all neighbours assigned to estimate the missing values. One of the advantages of using KNN to replace values is that it can estimate both quantitative and qualitative attributes. KNN is a popular method of missing value estimation in DNA micro-array experiments (Aittokallio, 2010, Kim et al., 2005, Pan and Li, 2010, Valafar,

2002). The method selects missing spot values for a given proteins and find K other spots which have a value most similar to the missing spot value from other gels. A weighted average depending on the relative distances calculated is then applied, which is done by the similarity in the spots expression values (Troyanskaya et al., 2001).

1.8 Summary

The field of quantitation and biomarker discovery in proteomics is gathering pace. With new technologies such as imaging mass spectrometry and label free quantitation to back up the more traditional and well established techniques such as 2DE. DiGE, on the other hand, still remains a powerful technique for harnessing the advantages of 2DE and whilst gaining reliable quantitation and multiplexing across gels. The level of data gained from DiGE is large and has potential for biomarker discovery. Developing analytical approaches would therefore benefit many different applications of DiGE and could be used in prospective data mining when searching for biomarkers.

In this thesis the DiGE technique was applied to investigating how the application of heat treatment applied to tissue compares to the traditional technique of snap-freezing. This was used to generate a list of potential markers that could occur during tissue degradation. Additionally, the DiGE technique was also applied to investigating potential biomarkers of hypertension by studying the difference between a WKY, congenic strains and SHRSP rat model. The specific aims are given in the relevant chapters; however the general aims of this thesis are summarized:

- To employ and develop the quantitative method of DiGE to look at tissue profiles.
- To develop analytical strategies to mine quantitative proteomic data from tissue samples gained using DiGE.
- To access the proportion of change in proteomic degradation in heat-treated and snap-frozen wild type mouse brain tissue.
- To find candidate examples of proteins or markers of degradation of wild type mouse brain tissue.
- To find putative examples of biomarkers in hypertension.

2 Material and Methods

2.1 Proteomic Methods

2.1.1 DiGE: Differential in Gel Electrophoresis

2.1.1.1 Tissue collection and extraction

All tissue harvested for experimentation in this thesis was collected in line with the U.K. Animals (Scientific Procedures) Act, 1986 and local ethical guidelines of Glasgow University. Specific details of tissue collection can be seen in the methods sections of the experimental chapters. All animals (mice and rats) were euthanatized by cervical dislocation and dissection was performed immediately. Any material that was taken for experimental purposes was snap-frozen in liquid nitrogen immediately after dissection (unless where specified in chapter 3) and then stored at -80oC.

Sectioning was performed using a cryostat microtome (Leica Microsystems CM 1900UV, UK). Typically, tissue was mounted using an embedding medium supplied by Leica (Leica Microsystems CM 1900UV, UK cat#14020108926). Any tissue in direct contact with this embedding medium was not used for experimental purposes, in order to limit contamination of upstream processes. Sectioning was typically at 14µl thickness unless otherwise stated.

After sectioning of tissue samples was performed, either a small pre-chilled brush was placed in 1.5ml eppendorf tubes, or sections were thaw mounted onto slides. The eppendorf tubes used were pre-chilled and an antistatic gun was used to prevented loss of sample due to repulsion.

2.1.1.2 Sample clean-up, desalting and precipitation

After tissue samples were solubilised in sample lysis buffer, they were taken into a clean-up process to minimise salts, which may affect first dimension separation (isoelectric focusing). A 2D clean up and preparation kit was used to achieve desalting and precipitation of proteins (GE Healthcare, UK cat#80-6484-51). The 50µl of solubilised sample and lysis buffer was taken and placed in a 1.5ml eppendorf tube in preparation for micro centrifugation. Then 300µl of precipitant solution was added and mixed by vortexing. The mixture was then rested on ice for 15 minutes. After resting a 300µl of co-precipitant was added and mixed by vortexing. The samples were then centrifuged for 5 minutes at 12000 x g or until a white pellet was clearly visible. The tubes were

carefully placed with the hinges on the outside of the circumference. The supernatant was removed by carefully pipetting using a Gilson pipette. If the pellet was disturbed centrifugation was repeated. The tubes were placed back in the same orientation and spun for 1 minute at 12000 x g to bring out any more supernatant. Any remaining supernatant was removed, using a pipette as before. Co-precipitant of 40µl was added to the top of the pellet and the tubes were left on ice for a minimum of 5 minutes and not longer than 8 minutes. The tubes were centrifuged for a further 5 minutes at 12000 x g in the same orientation. The supernatant was removed as before. Once the supernatant was discarded 25µl of d_4H_2O was placed on top. The tubes were then vortexed until the pellet was dispersed. If the vortexing took longer than 30 seconds, another wash step with co-precipitant was repeated. Once the pellet was dispersed 1ml of pre-chilled wash-buffer and 5µl of additive was placed on top of the pellet and placed at -20°C for at least one hour, but usually overnight. The tubes were vortexed every 10 minutes for 1 hour. After that, the tubes were taken and centrifuged for 5 minutes at 12000 x g. The supernatant was removed and discarded. The pellet was air dried for approximately two minutes. The pellet could then be stored at -80°C or re-suspended in DiGE lysis buffer (section 2.2.7).

2.1.1.3 Protein concentration Determination

2.1.1.3.1 Bradford Assay

The main mode for determining the protein concentration of samples in this thesis was the use of a Bradford Assay. The Bradford Assay has a linear range over 0.125–1.5 mg/ml or 0.125-1.5µg/µl. If necessary, the sample was diluted into this range for accurate analysis. The bovine serum albumin (BSA) standards used were obtained from Biorad, UK (catalogue #500-0207) with 7 pre-diluted standards of 0.125, 0.25, 0.5, 0.75, 1.0, 1.5, 2.0 mg/ml. A standard curve was prepared in triplicate, and needed to have an R^2 value of 0.95 or more in order to be used for quantitation of protein concentration. Samples were diluted to the same volume of 100µl using water. Then 100µl of Bradford Assay Reagent (Biorad, UK) was used to cause the colour change. A Tecan Genios (Tecan, UK) plate reader was used and wave length set to 595nm. Standard and sample were all performed in triplicate.

2.1.1.3.2 Bioanalyser Chip

In the movement towards small sample proteomics, an initial attempt was made to access the validity of using Agilent's, UK 2100 Bioanalyser chip technology for use with DiGE to move into using saturation labelling with small protein tissue samples. The findings are discussed in the discussion in section 5.5.2. Wild type OCR 21 weeks mouse brain tissue was sectioned and solubilised in Phosphate Buffered Saline (PBS), DiGE saturation labelling lysis buffer (see Figure 5-4) full and half concentration, 6M urea in PBS and 30mM tris in PBS. The solutions and lab-on-a-chip, called Agilent Protein 230 Kit came from Agilent, UK cat# 5067-1517. This kit can be used to quantify a number of proteins between 14-230KDa in molecular weight.

Firstly, a gel dye mix and destaining solution must be made. The dye concentrate and gel matrix from the kit were allowed to rise to room temperature for 30 minutes. The dye concentrate was vortexed and centrifuged briefly then 25µl was pipetted into the gel matrix solution. This was stored on ice in the dark. This mixture was vortexed and centrifuged for 15 seconds. This has created the working stock of gel/dye mix, this was placed into a spin filter. The destaining solution was prepared by transferring 650µl of gel matrix into a spin filter. When required, these spin filters were centrifuged at 2500 x g for 15 minutes. They can then be stored at -20°C until required.

When the chip was run, all solutions were allowed to thaw and warm to room temperature for 10 minutes. The denaturing solution was prepared by placing 200µl of sample buffer into a vial (supplied) and adding 7 µl of 1 M Dithiothreitol (DTT) solution and Vortex for 5 seconds. The samples are prepared by placing 4µl of sample and 2µl of denaturing solution into a vial and centrifuging this for 15 seconds. For the protein ladder, 6µl should be added in place of the sample. All samples and ladder were heated for 5 minutes at 100°C in a heating block. Samples and ladder were allowed to cool and then briefly centrifuged. Finally add 84µl of ddH₂O was added and vortexed.

Once all solutions were complete, the samples were loaded onto a chip. The chip was placed ready in the chip priming station. To load the chip, first 12µl of gel/dye solution was loaded into the section marked G and closed the priming station at the 1ml mark. 60 seconds were allowed and then the plunger was released. After waiting a further 5 seconds, the plunger was pulled back and any excess solution was removed. This process was repeated for the 4

wells marked with G. Then 12µl of destaining solution was pipetted into well DS. Now the samples were loaded into wells labelled 1 to 10 using 6 µl as prepared above. Wells 1-6 were BSA standards used in section 2.1.1.3.1, 7-10 were mouse brain tissue. Then the ladder was placed in the well marker with a ladder. The chip was inserted into the bioanalyser chip reader and the programme was started by File>electrophoresis>start. A standard curve can be produced by the separation and unknown samples quantified.

2.1.1.4 Glass plate preparation

Prior to gel casting, the glass plates were cleaned using 10% v/v De-con solution and thoroughly rinsed in very hot water and allowed to dry by evaporation. Gel plates used for preparative gels were taken and the long/backing plate wiped with approximately 4 ml/plate of bind saline solution (section 2.2.3). This made sure the gel adhered to the plate for spot picking when the front plate was removed. There was no need to prepare analytical gel plates with bind silane, as the removal of the front plate was unnecessary. In addition to using bind silane, two fluorescent reference markers (GE Healthcare, UK) were placed onto the gel side of the long/backing plate as shown in Figure 2-1. Positioning of the markers was not made too centrally to avoid any interference with protein spots. The markers were used to provide coordinates for the spot handling work station in order to pick any selected protein spots. It was crucial to place them correctly and in line with each other in order for camera recognition to take place.

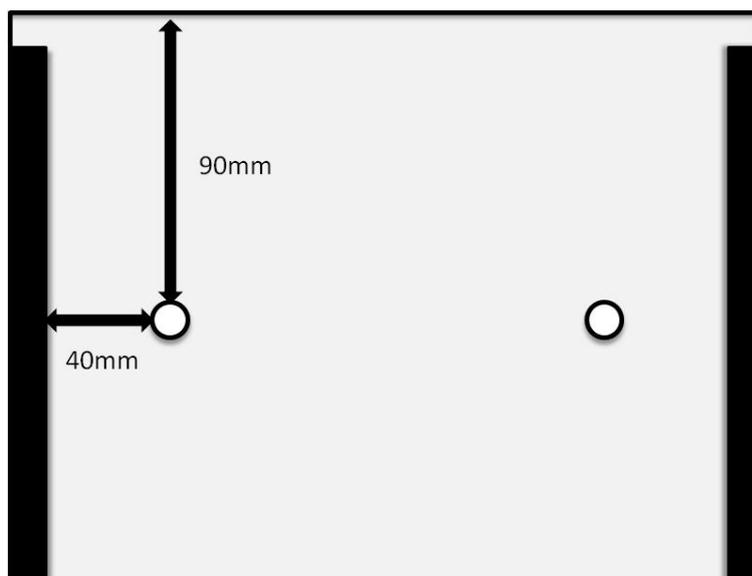


Figure 2-1: Position of gel reference markers. The fluorescent reference markers are used to position the plate and ensure accurate spot picking.

2.1.1.5 Polyacrylamide Gel Solution preparation and gel casting assembly

A 12.5% Sodium Dodecyl Sulphate -Polyacrylamide Gel Solution was prepared using the proportions of the relevant solutions described in section 2.2.9. The solution was filtered using a steripur and vacuum pump (Millipore, UK). The filter pore size was 0.22 μm and filtered out any contaminants that may impede the path of proteins.

An Ettan Dalt II gel caster was used for gel plate assembly and pouring of SDS-PAGE gels. The gel caster is a large plastic casing that allows for the casting of up to 12 gels at once. This required 900ml of solution. When more than 12 gels were needed, a second caster was employed but enough solution was made to fill both casters to maintain consistency in the gels. The caster was assembled before the addition of the TEMED and Ammonium persulphate (APS). The TEMED APS was added just before pouring to avoid the gel setting in the bottle.

The caster was assembled by the alternation of a thick acetate sheet, a back plate (with rubber spacers of $\sim 1\text{-}2\text{mm}$), short front plate and then another acetate sheet. The first acetate prevents the glass plate adhering to the caster, while the other acetate sheets stop adhering of glass gel plate to gel plate. Any gap was filled using further acetates or a blank plastic gel plate. To allow for a small amount of expansion, a gap must be left at the front of about 1mm. The front of the caster is screwed into place maintaining the gasket into the pre-machined groove, which maintains the seal and prevents the gel solution leaking.

Once assembly was complete, the TEMED and APS was added and mixed thoroughly by vortexing. A funnel was inserted into a grommet at the top of the gel caster and the gel solution was poured at a steady rate. As the caster fills from the bottom, the gel rises. At this point it was important to minimise the air that was allowed to enter the gel caster. The caster was tapped occasionally, to encourage the solution to rise and to displace the odd air bubble. When the rising gel line reached approximately 15mm from the top (but not at the top) about 50-100mls of displacing solution (section 2.2.4) was added to the funnel which brought the gel line up with about 5-7mm to spare ready to accept the IPG strip from the first dimension. Once the gel line had stopped, a water saturated iso-propanol solution was gently poured or pipetted over the top of the gel plates. This allowed for a straight gel line, which is

important to enable transfer of all the protein from the first to second dimension. The top was covered to prevent evaporation of the water saturated iso-propanol and to prevent dehydration. The gels were then allowed to polymerise for a minimum of 4 hours, but usually overnight. At this point the gels were removed from the caster and stored after cleaning with d_4H_2O . For storage, d_4H_2O was pipetted into the 7mm gap and gels were then wrapped and stored at 4°C. The gels were always used within 1 week of manufacture. All polyacrylamide waste from the gel caster was stored for appropriate disposal.

2.1.1.6 Labelling and preparation of Dyes

All supplies for DiGE were obtained by GE Healthcare, UK. For DiGE using minimal labelling a kit was purchased containing all three Cy dyes required (cat# 25-8010-65). The DiGE process uses fluorescent Cyanine dyes. The process of minimal labelling is covered in the introduction of this thesis (section 1.5.3.1.1).

There are 3 spectrally resolvable Cy dyes for minimal labelling. In the experiments outlined in this thesis Cy 2 was employed to label the internal standard and Cy 3 and 5 were used to label experiment samples. The employment of Cy 2 for the internal standard is usual practice in the literature reviewed. The internal standards comprise of a combination of all samples used in the experiments described in chapters 3 and 4. For example in chapter 3 in the pilot investigation (see section 3.4.4.1) for 9 analytical gels a total of 450 μ g is required (50 μ g per a gel). In order to achieve this, 150 μ g must be taken from the each of the 3 sample groups and pooled for the internal standard. In addition a further 500 μ g is required for preparation gels to pick spots for MS identification.

2.1.1.7 Reconstitution of Minimal Cy dyes

The labelling procedure follows closely the protocol outlined in the product booklet entitled “Amersham Cy dye DIGE Fluors (minimal dyes) for Ettan DIGE” produced by GE Healthcare, UK. The first process in minimal labelling is the reconstitution of the Cy dyes using anhydrous dimethylformamide (DMF). This must be of the highest quality and no older than 3 month due to the risk of contamination with water from the atmosphere. The Cy dyes were removed from storage at -20°C and thawed for 5 minutes at room temperature. After thawing, 25 μ l of DMF is added to each vial of Cy 2, Cy3 and Cy5 dyes. This gives a working dye concentration of 0.4nM. The

tube was then vortexed for 30 seconds. Subsequently, to avoid continuous thawing the Cy dyes were divided into 1µl aliquots using 0.5ml eppendorfs. 1µl of working dye solution was sufficient to label 50µg of protein sample. When ready to label, the 0.5ml eppendorfs were centrifuged briefly to ensure all working dye solution was in the bottom of the tube, ready for pipetting.

2.1.1.8 Reconstitution of Saturation Cy dyes

The procedure for the reconstitution of saturation Cy dyes is similar to that for minimal Cy dyes. The reconstitution of saturation dyes also used anhydrous DMF. The Cy dyes were removed from storage at -20°C and thawed for 5 minutes at room temperature. After thawing, 50µl of DMF was added to each vial of Cy3 and Cy5 dyes giving a working concentration of 100nmol. The tube was then vortexed for 30 seconds. Subsequently, to avoid continuous thawing, the Cy dyes were divided into 1µl aliquots using 0.5ml eppendorfs. 1µl of working dye solution was sufficient to label 5µg of protein sample

2.1.1.9 Minimal labelling

The minimal labelling protocol requires a concentration of protein to be between 1µg/µl and 20µg/µl. All experimental samples used in this investigation had a protein concentration in this range. Minimal labelling was performed for samples in chapters 3 and 4. Experimental groups were prepared by placing a total protein content of 50µg of each sample into a 1.5ml eppendorf for each of the specified Cy dyes (see experimental design in chapters 3 and 4 for corresponding sample groups). The total volume needs to be made up to 10µl using the addition of an appropriate volume of DiGE lysis buffer (section 2.2.7). One 1µl of the aliquoted Cy dye was added to the respective samples. The dye and sample were vortexed and centrifuged briefly and incubated on ice for 30 minutes in the dark. After incubation 1µl of 10mM lysine was added to stop the reaction by binding to any excess dye. It was mixed by pipetting and briefly centrifuging. It is left for a further 10 minutes on ice in the dark. Samples were always taken straight into first dimension in this investigation. However, it is possible to store samples in the dark at -70°C for up to 3 months.

2.1.2 Saturation/scarce sample labelling.

Saturation was performed in this investigation only at the end in order to start the optimisation for future work (see chapter 5). Labelling using saturation

dyes is slightly more time consuming and complicated in comparison to minimal labelling.

A 5µg sample was added to the required 1µl of working dye solution from the reconstituted dyes aliquots (as described in section 2.1.1.8). The volume was made up to 9µl using DiGE lysis buffer (section 2.2.8). At this point a certain volume of TCEP was added. The amount of TCEP added depended on the gel in the Dye optimisation experiment (see section 2.1.2.1). A 2mM TCEP solution (Sigma Aldrich, UK) was prepared by dissolving 2.8mg of TCEP in 5ml of ddH₂O. The respective volumes were added to create different concentrations of dye for each gel of 2nmol, 4nmol and 8nmol.

2.1.2.1 Saturation Dye concentration determination

For the dye optimisation experiment OCR male 21 week old mouse brain tissue was used. TCEP is used to reduce the formation of disulphide bridges, therefore samples with a high Cysteine content require more TCEP. The ratio of dye to TCEP should always be at 2:1, therefore dye concentration needs to be optimised. Figure 2-2 shows the relative amounts of TCEP and dye used in the dye optimisation experiment. The concentrations of 2nmol, 4nmol and 8nmol concentration were tested.

Gel	2mM TCEP (µl)	TCEP (nmol)	2nM Dye (µl)	Dye (nmol)
1	0.5	1	1	2
2	1	2	2	4
3	2	4	4	8

Figure 2-2: Table showing relative concentration used in saturation dye optimisation experiment. "

2.1.2.2 1st Dimension and rehydration loading

After labelling the three respective Cy dyes 2, 3 and 5 for 1 gel were mixed by vortexing and centrifugation. At this point 440µl of rehydration buffer (section 2.2.10), Dithiothreitol (DTT) and IPG buffer were added fresh to stocks) was added to the mixed Cy dye and protein sample giving a total volume of 452µl. The first dimension separates the sample on the basis of their charge using isoelectric focusing. The process was performed using GE Healthcare, UK;

Amersham's Ettan IPGphor system. 450µl of the sample, Cy dye and rehydration buffer for 1 gel was loaded onto an immobilised pH gradient dry strip gel. The gradient used in the minimal labelling experiments of chapter 3 and 4 was pH 4-7. The strip used in the saturation dye concentration determination experiment had gradient of pH3-10 (see chapter 5). The IPG strips were 24cm in length. In order to perform isoelectric focusing the 450µl sample mixture was placed in a ceramic strip holder by slowly pipetting an even amount along the length of the holder, the IPG strips were removed from the cover exposing the gel face. The gel face was laid down onto the sample, with the positive end of the strip to the pointed end of the ceramic strip holder. Care was taken not to introduce air bubbles. Any air bubbles were then expelled using a syringe needle. After all bubbles were expelled, 1.5ml of mineral oil (Amersham Bioscience, UK) was laid on top of the strips and sample to prevent evaporation. The positive end was placed on the anode of the IPGphor machine and the negative tail onto the cathode. There were up to 12 ceramic strip holders run at any one time. The programme used for IEF was done in steps. First, the strip was stepped to 30 volts and held for 14 hours, then stepped and held at 500 volts for 1 hour, and then to 1000 volts for 1 hour. After, the voltage was ramped up in a gradient to 8000 volts over 1 hour and then held for 10 hours. Strips were allowed to accumulate a minimum of 75000 volt hours and never taken off with less than this. Strips were typically only stored for 1-2 days at -20°C if not immediately taken into the 2nd dimension.

2.1.2.3 IPG strip equilibration

Before the second dimension IPG strips must be equilibrated. IPG strips were placed in plastic tubes following isoelectric focusing. Strips were then either frozen and stored, or taken straight into IPG strip equilibration buffer (SEB) in preparation for the 2nd dimension. SEB was poured into the plastic storage tubes (frozen or non-frozen) and placed on a slow speed shaker at approximately 30 Cycles/min. Equilibration was split into two steps. First, strips were equilibrated in 10mls of SEB with 10mg/ml of dithiothreitol (DTT) for 15 minutes. Secondly, strips were equilibrated in 10mls of SEB with 25mg/ml of iodoacetamide for 15minutes. Intermittently between equilibration, strips were washed using 1x running buffer. The first stage of equilibrium with the addition of the reduction agent DTT causes the reduction of the thiol groups causing the disulphide bridges to break. This occurs by two

sequential thiol-disulfide exchange reactions resulting in DTT becoming a six-member ring structure as shown in

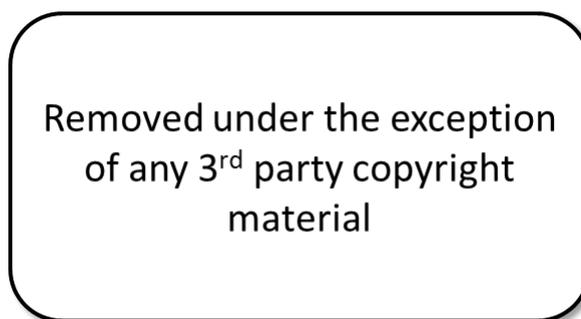


Figure 2-3: Reduction of a disulphide bridge.

The second wash with iodoacetamide (which is an alkylating sulfhydryl reagent) prevents the reformation of disulphide bridges by binding covalently to free sulphides during the second wash during equilibrium. Absence of this step leads to gel distortion and an aberrant gel image (Herbert, 2001). The bromophenol blue contained in the SEB allowed for the dye front to be tracked when the gel was run in the 2nd dimension.

2.1.2.4 2nd Dimension

The second dimension separates proteins on the basis of their molecular weight. After equilibration strips were rinsed and placed horizontal across the top of 24 cm (large format) gels, which had to cast prior to running strip equilibration. The top of the gels were filled with 1x running buffer (section 2.2.1) to aid the insertion of the IPG strips. IPG strips were rinsed in 1 x running buffer to take off any of the equilibration buffer from the previous step. Using a pair of tweezers, the strips were carefully placed onto the top of the SDS-PAGE gel, being careful to allow any bubbles to rise above the strip. The acidic end was placed to the left hand side with the gel side facing out. Finally, the strips were held in place with 2ml of Agarose sealing solution. The gel tank was supplied by GE Healthcare, UK; Ettan Dalt II system. The slots of the gel tank were lubricated with 1x running buffer to allow smooth entry of the glass plates and to prevent damage to the seals and electrodes. Before the gels were placed into the gel tank, 1x running buffer was poured into the bottom half of the tank (approximately 8 litres) and then the gels were slotted into position. The tank was then topped up with 2x running buffer in the top half. When the lid was secured the gels were ran at 1-2 Watts per gel overnight or until the dye front reached the bottom of the gel. The gel tank was covered to avoid dye degradation for the period of running.

2.1.2.5 DiGE gel scanning

Gel scanning was performed using Typhoon 9400 variable mode imager (GE Healthcare, UK). Glass plates were cleaned and rinsed, in order to discard any residue of running buffer before placing the glass plates onto the scanner. A metal frame is placed on the scanner's glass platen to hold the glass plates in place. Plates were placed long/backing plate down on the glass platen with acidic end of the strip in the top left hand corner. Scanning for analytical gels was performed immediately after 2nd dimension in order to minimise protein spot diffusion and bleaching of fluorescent signal. Despite this, subsequent scans up to 4 days later showed very little diffusion and strong signal providing they had been stored out of the light. Scanning of preparative gels took place directly after Sypro Orange staining. Initially, when scanning either analytical or preparative gels, a low resolution "pre-scan" of 1000 microns was used for each DiGE channel or Sypro orange. This was done to assess the optimal PMT voltage to use. Gels were subsequently saved in a DiGE format as .gel file. For analytical gels the three channels for Cy2, 3 and 5 (or just Cy3 and 5 for saturation labelling) were also saved as a data set file or .ds. For the analytical gels as part of the pre-scan PMT voltage was also balanced between the three Cy channels. In order to perform accurate quantitation analysis, gel channels must be on average no more than 15% different from the corresponding spots within that gel, this allows accurate normalisation downstream in the analytical software DeCyder. This task was performed in Image Quant software (version 5.2, Amersham Bioscience, UK). Spots with a greater maximum intensity value than 100,000 were classed as saturated and are not quantifiable. The number of spots that had an intensity value of 100,000 was minimised. Once the PMT voltages were determined for all channels, all the gels were then scanned at those specific voltages at 100 micron resolution. The different CyDyes need scanning using different wave lengths. Cy2, 3 and 5 are scanned at 488,532 and 633nm using a blue, green and red laser respectively. Analytical gels were then stored in the dark in anticipation of any need to rescan. Preparative gels were stored in 7.5% acetic acid solution, before proceeding to picking on the Spot handling work station.

2.1.2.6 Preparation gel staining with Sypro orange

On completion of the second dimension, preparative gels were placed into approximately 500ml of fixative solution (10% methanol and 7.5% acetic acid) for a minimum of 2hrs. If stored overnight a 1hr wash in 500ml of 0.05% SDS

solution was preformed, as Sypro orange binds to SDS. Staining was performed using Sypro Orange (Molecular Probes, UK) at a dilution of 1/10000 in 7.5% acetic acid for 2hrs before being washed in ddH₂O removing excess stain from the surface. Afterwards, washing gels were scanned, in order to visualise protein spots to allow for matching and picking downstream. Staining and storage of gels was performed in a clean polyethylene try to avoid bleaching from light. Scan of preparative gels was performed at 580nm using the green laser, once again scanning at a 1000 microns to determine a PMT voltage and then at the higher resolution of 100 microns.

2.1.2.7 Software Analysis: DeCyderDeCyder

The software used for matching, image analysis and statistical analysis was Decyder2D version 7 upgraded after using Decyder5 (GE Healthcare, UK). The software is made up of several modules; Image loader, Batch processor, differential in gel analysis (DIA), biological variance analysis (BVA) and extended data analysis (EDA). A summary of the work flow is illustrated by the home screen of the DeCyder Software shown in Figure 2-4.

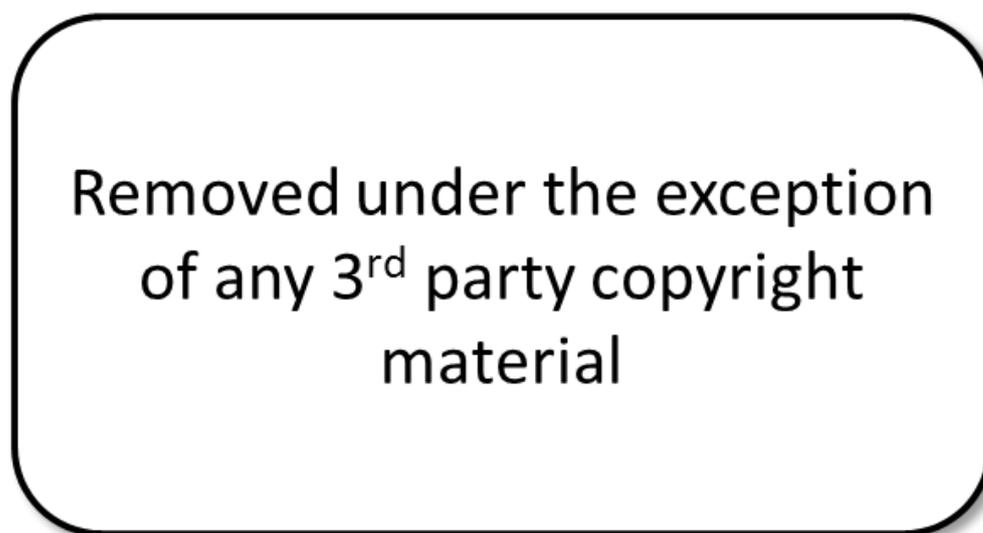


Figure 2-4: Home screen of DeCyder 2D software version 7. The home screen illustrates the sequence of workflow.

Before analysis using DeCyder, the gels' images were cropped after image acquisition using ImageQuant tools (molecular Dynamics, UK), this removed any unusable regions of the gels such as the extremities, which may affect spot detection. An analysis in the DeCyder software is started by using the image loader to input the .gel / .ds files from image acquisition. The batch processor was employed to open and save DIA module files. This process involved invoking the algorithm (algorithm detection version 6.0) for spot detection by

search for the maximum 10,000 spots, which improves the algorithms capacity for detection. Additional spots less than 29000 volume were excluded as dust particles. This batch processing saved a great deal of time, compared to manually inputting up to 18 gel images into the DIA. Additionally, as the DIA is used solely for the comparison of two gels/channels, only limited amount of time was spent using this module, however, files must be saved in .DIA format prior to exporting into the BVA module. DeCyder version 7 is more powerful than version 5, with an updated algorithm and better automated warping and spot detection. Although a small amount of analysis on the pilot study of chapter 3 was analysis using DeCyder 5, the rest of the experiment work for chapter 3 and 4 was done using DeCyder 7. The batch processor also allowed for the creation of .BVA files that can be entered into the next part of the DeCyder workflow.

2.1.2.8 Biological Variance Analysis (BVA) module

After batch processing and the creation of .BVA files, they were then opened using the BVA module. The DIA modules algorithm allowed for the identification of spots on a particular gel, however, the BVA module is used for analysis and matching. Matching is a semi-automated process in reality. The BVA module in the most part is successful in matching spots across gels, by using the internal standard. However, in practice, a great deal of time is spent disassociating poorly matched spots and re-associating them with the correct spots across all channels. With up to 18 gels and a preparation gel this could be 55 different images. This took about 1-2 weeks of constant matching and checking. Once again DeCyder 7 performed matching better than DeCyder 5. The matching process starts by assigning some reference spots manually, the more chosen the greater the matching efficiency, about 50 spots from each channel was initially matched manually. This was performed by defining a master Cy2 image and matching all the other internal standards to this. The master image was picked on the basis of visually being the best, with well defined, numerous spots and the least streaking. The automated process could then be initiated and then manually checked for correct associations.

Once matching was complete, the analysis could start. The type of analysis performed in this thesis is discussed in sections 3.4.5 and 4.4.4.2 of chapter 3 and 4 respectively. A pick list was then generated from spots of interest (see section 2.1.2.10).

2.1.2.9 Extended Data Analysis (EDA) module

The EDA accepts BVA files and works on the basis of defining a working base set. A number of filter options are available and groups can be redefined from the original BVA files. The only analysis performed in the EDA module (which was newly acquired towards the end of this investigation) was the performance of principal component analysis. The processes using the EDA are relatively automated. Once the base set is defined, a number of statistical operations and comparisons can be made, from differential analysis (which could be performed in the BVA) to looking at k-means and hierarchical relationships. The PCA analysis was performed by selecting the required display and clicking calculate.

2.1.2.10 Pick list generation and spot picking

Pick list generation for spots of interest was done using DeCyder Software (GE Healthcare, UK). Pick list provided coordinates in relation to the place fluorescence reference markers, which were either automatically or manually selected in the software. The pick list can then be exported to the spot handling work station.

The Ettan Spot Handling Workstation is a fully automated system for handling, picking, digesting and ultimately spotting digested protein samples on MALDI target plates or into a 96 well plate for LC-ESI-MS from 2D or DiGE gels. The gels are placed in definable positions in the unit, and all identification is done by barcodes. The preparative gels are placed into a gel holder which is barcoded with the short plate removed. The gel was covered with water to prevent dehydration. The reference markers need to be placed between two parallel lines for camera recognition. All setup needed to take place before starting the work flow. All preparative gels needed for picking were placed in the designated places and barcodes scanned. The appropriate number of barcoded 96 well plates are placed in the designated position and scanned. The 96 well plates are needed for placing picked gel pieces, performing digestion drying and subsequent spotting. Therefore, for every 96 spots picked, two 96 well plates are required, one for the picked gel piece and one for the final protein solution. In addition a 192 MALDI plate was inserted into the designed barcode holder, placed in the relevant slot in the unit and scanned. All solutions were topped up (section 2.2.11). When the sequence started, the work station confirmed all the scan elements by rescanning all the barcode (thus linking that object to that position). The gel was then moved ready for

spot picking. The camera confirms the reference markers and spot picking begins using the 1.4mm picking needle, going through the assign pick list. The plugs were incubated in 300µl of 50mM Ammonium bicarbonate in 50% methanol for 30 minutes, followed by a drying phase for 30 minutes. The digestion phase followed by adding 20mM ammonium bicarbonate followed by trypsin (Promega, UK) 1:100, by weight. in 25 mM ammonium bicarbonate. The gel plug and trypsin were then incubated at 37°C for 4 hours to cause the digestion of protein to peptide allow their extraction from the gel. To wash the peptides out 125µl of 50% ACN/0.1 trifluoroacetic acid (TFA) was placed with the gel plug and incubated for 30 minutes. This solution was then removed and placed into the second 96 well plate and dried. The spotting took place (when required) by the addition of 3µl of 50% ACN/0.2%TFA. A small fraction of this (0.5µl) was drawn up and spot with 0.5µl of matrix α -Cyano-4-hydroxycinnamic acid and spotted onto 192 4700 target plates. The target plates were not left for longer than 1 hour before placing into the MS and into vacuum.

2.1.3 Protein identification using Mass Spectrometry

Two types of mass spectrometer were employed in the identification of protein samples during this investigation. Firstly, gel spots excised and digested by the spot handling workstation were spotted onto 192 target plate for MS analysis using 4700 Proteomics Analyser (Applied Biosystems, UK). This is a MALDI-TOF-TOF-MS with automated plate cassette. This was an excellent start in protein identification of a large amount of spots. 1000 spots could be analysed in about 4 days, compared to 100 using an LC-ESI-MS. Any spot of importance undetected were further analysed using QStar Pulsar i (Applied Biosystems, UK). This was a ESI-QUAD-TOF-MS. The mass spectrometer was coupled to a LC Packings, NL LC system with FAMOS auto-sampler and injector. The column used was a 15cm, 75 μm C18 3 μm particle size, 100Å pore size PepMap (Dionex, UK). The column was stored at 37 °C in an oven. There were two solvents used to form a gradient for elusion. Solvent A was a 2 % acetonitrile, 0.1 % formic acid solution and solvent B was an 80 % acetonitrile, 0.1 % formic acid solution. The samples were first loaded onto a trapping column, for 5 minutes and washed to protect the more expensive C18 column and to hold the sample for valve switching. The loading solution for the trapping column was 2 % acetonitrile, 0.5 % tri-fluoroacetic acid solution. The flow rate for loading onto the trap was set at 200 $\mu\text{l}/\text{min}$ and the flow through the column was 0.3 $\mu\text{l}/\text{min}$. Solvent A and B were introduced, using the following gradient. The ratio of solvents was 95 % solvent A and 5% solvent B for the first 10 minutes, this ratio ramped up to 80% solvent B and 20% solvent A over 30 minutes.

2.1.3.1 Matrix-assisted laser desorption ionisation time of flight mass spectrometric analysis

Analysis was carried out using the 4700 Proteomics Analyser (Applied Biosystems, UK). The MS was operated in positive reflector mode using a relative laser intensity of around 2500-4500 depending on the signal observed. The MS and MS/MS data was generated using the following methods:

2.1.3.2 MS Acquisition method:

For the acquisition of spectra the following settings were used;

- Mass range (Da): 800 to 4000
- Focus mass (Da): 1700
- Matrix: alpha-Cyano-4-hydroxycinnamic acid
- Total Spectrum: 2000 shots (100 shots/sub spectrum) accepting every sub spectra discard the first 50
- Automatic control: Random with centre bias at fixed laser intensity of 5250
- Final detector voltage: 1916

2.1.3.3 Processing Method:

For the calibration and processing of the collected spectra the following settings were used;

- Calibration performed internally using common trypsin peaks e.g.. 842.510, 1045.564, 2211.105
- Mass tolerance: +/- 1 m/z
- Max outlier error: 20ppm

2.1.3.4 MS/MS Interpretation and Acquisition method:

For the collection and interpretation criteria for MS/MS to be performed on precursor ions the following settings were used;

- Exclusion list: All common/known trypsin peaks.
- First 10 precursor masses selected on the basis of highest intensity first.
- Fragmentation: Collision Induced Dissociation
- Mass range (Da): 10-1361.535
- Focus mass (Da): 1296
- Matrix: alpha-Cyano-4-hydroxycinnamic acid

- Total Spectrum: 5000 shots (250 shots/sub spectrum) accepting every sub spectra discard the first 10
- Automatic control: Random with centre bias at automatic laser intensity adjustment to maximum of 6000
- Final detector voltage: 1916

2.1.3.5 LC-ESI-QUAD-TOF Mass Spectrometry

If identification was not achieved using 4700 Proteomics Analyser, then an Electrospray ionisation MS was performed using the QStar Pulsar i (Applied Biosystems, UK) which is an ESI-QUAD-TOF-MS. The 96 well digested sample plates created by the spot handling workstation were used and the peptides solubilised and injected into the MS via an LC system (described in section 2.1.3). The methods used to generate the MS and MS/MS data is shown below.

The method is made up of 1 TOF MS+ scan and 4 Product Ion scans.

- **MS+ scan**
- accumulation time: 3.000069seconds
- polarity: positive
- Cycle time: 14.9959
- Cycles: 136
- Delay time: 0
- tof masses: 400amu to 1500
- Transmission window: 380amu at 100%

2.1.3.6 Switch Criteria

The switch criteria were as follows:

- with charge state: 2 to 4
- which exceeds: 15 counts

- exclude former target ion: 180 seconds
- mass tolerance of: 300mmu
- **4x Product ion scans**
- Product of: 1
- accumulation time 3.000069 seconds
- polarity: positive
- Cycle time: 14.9959
- Cycles: 136
- Delay time: 0
- tof masses: 50amu to 2000

2.1.3.7 Advanced MS

The advanced MS criteria and transmission window were as follows;

- 30 amu 16.7%
- 75amu 16.7%
- 150amu 16.7%
- 300 amu 16.7%
- 600amu 16.7%
- 1200amu 16.7%

2.1.3.8 Identification using Mascot database searching.

The data for the MALDI-TOF-TOF was submitted to mascot database searching version 2.1 using Applied Biosystems GPS explorer version 3.5(build321) and for ESI-QUAD-TOF instrument Mascot Daemon version 2.2.0 using the following settings:

- Taxonomy: Mus./ Rattus

- Database: NCBIInr
- Enzyme: Trypsin
- Missed cleavages: 1
- Fixed Modifications: Carbamidomethyl(C)
- Variable Modifications: Oxidation(M)
- Precursor tolerance: 100ppm and MS/MS Fragment tol. 0.8 Da

2.1.4 Laser micro dissection

Once again, in order to move towards small sample proteomics, a reliable method of small sample tissue collection was trialled and discussed in chapter 5 section 5.5. A Leica LMD6000 (Leica, UK) was used to perform tissue collection of wild type OCR 21 Weeks mouse brain tissue. In order to assess the compatibility of the Polyethylene terephthalate (PET) membrane LMD slides with MALDI-MS. Tissue was sectioned in 14µm slices using a cryostat microtome (Leica Microsystems CM 1900UV, UK). The sections were then thaw mounted onto the LMD slides, often with considerable difficulty. The slides were then placed under the LMD objective. Circular discs with 600µm diameter were cut and allowed to drop on ITO MALDI glass target slides and immediately spotted with sinapinic acid matrix at 5mg/ml and 20mg/ml in 50/50 and 70/30 ACN and water. These slides were taken on for MS analysis. A sub-section of the disc was mechanically disturbed with a pipette tip.

2.1.5 Western blotting

The steps involved in western blotting are described below. The samples used were taken from the prepared sample used in DiGE before the labelling protocol was performed. The western blot performed was repeated 3x.

2.1.5.1 1D gel electrophoresis

1D electrophoresis was performed using Invitrogens NuPAGE Novex 4-12% Bis-Tris Gel 1.0 mm, 12 well pre-cast gels (Invitrogen, UK, cat#NP0322BOX) in conjunction with diluted NuPAGE MOPS SDS Running Buffer (20X) cat#NP0001 in water and Invitrogen gel tanks. The sample was prepared by using an appropriate volume of 6x Laemmli loading dye (50 glycerol, 10% SDS, 0.375 M Tris pH6.8, 10% β-mercaptoethanol and a trace of bromphenol blue dye). A ColorPlus™ Prestained Protein Marker (BioRad, UK) with a range of protein molecular weights was included into the first lane of every gel used with a range of 6.5-175KDa. Samples were heated on a block at 95°C for 5 minutes to denature the proteins and allow SDS access along the length of the proteins. A total volume of 10µl was loaded into each well and the gels were ran at 200 volts for 30 minutes or until the dye front had reached the bottom.

2.1.5.2 Protein-membrane transfer

After electrophoresis the gels were removed and trimmed appropriately. 3mm chromatography paper was cut to size (Whatman, UK) and soaked in transfer buffer Nitrocellulose membrane 0.45µm (Milipore, UK. Product

#HATF00010) was also cut to size and soaked in methanol and then water until it did not float on the surface. Then the membrane was moved to soak in transfer buffer (0.025M Tris at pH10.5, 0.2M Glycine, 0.1% SDS, 20% methanol). The membrane and gel stack was prepared in the order shown in Figure 2-5 and performed using Pierce Fast Western System (Thermo, UK), which is a dry membrane transfer system. The Semiphor set up requires the use of 27 mA of current, per gel, for 1 hour.

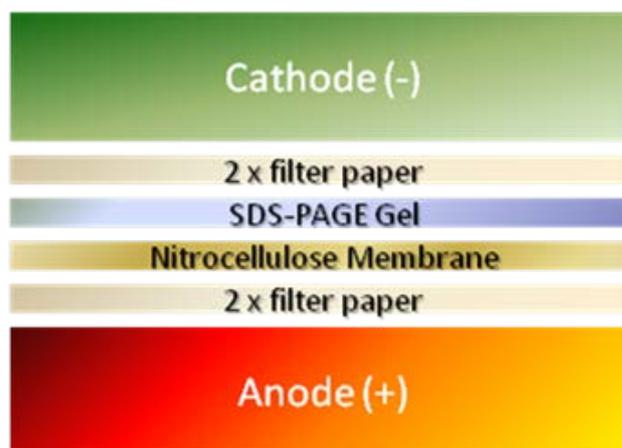


Figure 2-5: Order of membrane and gel stacking for western blot membrane protein transfer to nitrocellulose membranes. The proteins will migrate towards the anode transversely from the SDS-PAGE Gel onto the nitrocellulose membrane.

2.1.5.3 Antibody probing and Chemiluminescence

Following protein-membrane transfer, the membranes were placed into Tris-Buffered Saline Tween (TBST)(0.025M Tris at pH7.5, 0.0028 M KCl, 0.14 M NaCl, 0.1% Tween, 5% (w/v) Marvel dried milk) for blocking (approximately 10ml) on a shaker at 60 oscillation/minute for at least 1 hour.

The membrane was then moved for primary antibody incubation. The primary antibody used was an anti-peroxiredoxin 6 raised in rabbit (monoclonal) (Abcam, UK. Product # EPR3755) to a dilution of 1:1000. Primary antibody exposure was performed using approximately 10ml of TBST and 5% Marvel. This was placed with the membranes into a falcon tube and placed back on the shaker for a minimum of 3 hour, usually overnight

The membranes were then washed with 20ml TBST and 5% marvel in 10 x 2.5 minute washes. A horse-radish peroxidase conjugated with a secondary antibody targeted to rabbit raised in goat (Abcam, UK. Product # ab97200) was used to attach to the primary antibody. The membrane was washed for 1 hour minimum with 10ml of TBST with a 1:2000 dilution secondary antibody.

The secondary antibody solution was then removed and the membranes were then washed with 20ml TBST and 5% marvel in 10 x 2.5 minute washes again.

2.1.5.4 Gel dock image acquisition

Image acquisition was performed using G:Box (Syngene, UK) gel documentation system. Each membrane was then incubated in Pierce enhanced chemiluminescence (ECL) reagent for 2 minutes. This is used to detect the binding of the HRP-conjugated secondary antibody. Membranes were then carefully dried and placed into the G:Box. The membranes were exposed for a 2 minute exposure to give good exposure.

2.2 Chemicals, Consumables and Solutions

Below is a set of tables showing the main solution used throughout experimental work.

2.2.1 Tris/Glycine Electrophoresis Running Buffer

	Concentration	Quantity	Supplier
Tris-Cl (1.5M pH8.8)	25mM	30.25g	Sigma, UK
Glycine	192mM	144g	Sigma, UK
SDS	0.1% (w/v)	10g	GE Healthcare UK

* Made up to 5L with ddH₂O at 10X

This solution was diluted for 1x and 2x Running buffer as required

2.2.2 Agarose Sealing Solution

	Concentration	Quantity	Supplier
SDS Buffer		100ml	GE Healthcare UK
Agarose	0.5%	0.5g	Sigma, UK
Bromophenol blue	trace	trace	Fisher Scientific, UK

2.2.3 Bind Silane Solution

	Concentration	Quantity	Supplier
Ethanol	80%	8ml	Fisher Scientific, UK
Acetic Acid	1.9%	200µl	Sigma, UK
ddH ₂ O	17.6%	10µl	LAB
Bind Saline	0.50%	5µl to 1ml	GE Healthcare UK

2.2.4 Displacing Solution

	Concentration	Quantity	Supplier
SDS Buffer		100ml	GE Healthcare UK
Agarose	0.5%	0.5g	Sigma, UK
Bromophenol blue	trace	trace	Fisher Scientific, UK

2.2.5 Equilibration Buffer

	Concentration	Quantity	Supplier
Urea	6M	72.07g	GE Healthcare UK
Tris-Cl (1.5M pH8.8)	50mM%	6.7ml	GE Healthcare UK
Glycerol	30%	69ml	Sigma, UK
SDS	2%	4g	GE Healthcare UK
Bromophenol Blue	trace	trace	Fisher Scientific, UK
DTT*	10mg/ml	100mg in 10ml	Sigma, UK
Iodoacetamide*	25mg/ml%	250mg in 10ml	GE Healthcare UK

2.2.6 Fixing Solution

	Concentration	Quantity	Supplier
Methanol	10%	100ml	Fisher Scientific, UK
Acetic Acid	7.50%	75ml	Sigma, UK

2.2.7 Lysis buffer for minimal labelling

	Concentration	Quantity/10ml	Supplier
Urea	7M	4.20g	GE Healthcare, UK
Thiourea	2M	1.52g	GE Healthcare, UK
CHAPS	4% (w/v)	0.4g	GE Healthcare, UK
Tris	30mM	0.06g	GE Healthcare, UK
Acetate - check	5mM	1.2mg	GE Healthcare, UK
Adjust to pH 8.5 with dilute HCl			

2.2.8 Lysis buffer for saturation labelling

	Concentration	Quantity/10mls	Supplier
Urea	7M	4.20g	GE Healthcare, UK
Thiourea	2M	1.52g	GE Healthcare, UK
CHAPS	4% (w/v)	0.4g	GE Healthcare, UK
Tris	30mM	0.06g	GE Healthcare, UK
Acetate - check	5mM	1.2mg	GE Healthcare, UK
Adjust to pH 8.0 with dilute HCl			

2.2.9 Polyacrylamide Gel Solution (12.5%)

	Quantity	Supplier
Acrylamide-Bis(37.5:1)	375ml	Bio-Rad
Tris-Cl (1.5M pH8.8)	225ml	Sigma, UK
ddH ₂ O	281ml	Lab
10% SDS	9ml	GE Healthcare UK
10% APS	9ml	GE Healthcare UK
10% TEMED	1.25ml	GE Healthcare UK

*Made up to 1L or amount required

2.2.10 Rehydration Buffer

	Concentration	Quantity/10mls	Supplier
Urea	8M	4.85g	GE Healthcare UK
CHAPS	2%	0.2g	GE Healthcare UK
Bromophenol Blue	trace	trace	Sigma, UK
DTT	2mg/ml	2mg/ml	Sigma, UK
IPG buffer	0.50%	5µl to 1ml	GE Healthcare UK

2.2.11 Spot handling workstation solutions

	Concentration	Quantity	Supplier
AmBic/Methanol	50mM in 50%	As required	Sigma, UK and Fisher Scientific, UK
AmBic	20mM	As required	GE Healthcare UK
ACN/TFA	50% in 0.1%	As required	Sigma, UK
ACN/TFA	50% in 0.5%	As required	Sigma, UK
Trypsin	1ml AmBic in 1 vial	As required	Promega, UK
CHCA	20mg/ml	As required	Sigma, UK
Spotter	59%ACN with 0.5%TFA	As required	Sigma, UK

2.2.12 Sypro Orange Staining Solution

	Concentration	Quantity	Supplier
Sypro Orange Staining	1/10000	100µl	Molecular Probes, UK
Acetic Acid	7.50%	75ml	Sigma, UK

2.2.13 Other stock solutions for DiGE

- 10% Ammonium persulphate.
- 10% Sodium DodeCyl Sulphate
- 10% Tetramethylethylenediamine

3 Differential in gel electrophoresis analysis of wild type mouse brain tissue: A multi-faceted analytical approach to assessing the effect of heat treatment on the degradation of the proteome.

3.1 Aims

The aims of this investigation were to assess the use and effectiveness of heat treatment as a mode of preventing proteomic degradation of wild type mouse brain tissue. A comparison of the traditional method of snap-freezing in liquid nitrogen and the concept of heat-treating tissue are considered.

In addition, a review and assessment of the effectiveness of different analytical approaches using Differential in Gel Electrophoresis (DiGE) data and the effectiveness and limitation in deconvolution of spot map and intensity data in respect to proteomic degradation of wild type mouse brain tissue were investigated.

The use of DiGE to assess the effectiveness of heat treatment is novel. The approach employed was split into two. A pilot study with 4 technical replicates for each condition was used to validate the use of DiGE for this application and to guide the main investigation using 6 biological replicates of each condition. DiGE is a well-documented and effective strategy to look at the wide dynamic range of the proteome at a given point. In order to assess the effectiveness of a technique in reducing degradation it seems pertinent to attempt to evaluate across a high number of proteins with a technique, which covers a large proportion of the proteome.

Specifically the aims of the experiments detailed in this chapter are:

- To evaluate heat treatment as a method in the reduction or cessation of protein degradation of brain tissue in comparison with snap freezing.
- To determine the proportion of change in proteomic degradation in heat-treated and snap frozen samples.

- To determine the effect on the proteome of heat-treatment of wild type mouse brain tissue samples.
- To find candidate examples of proteins or markers of degradation of wild type mouse brain tissue.
- To provide a strategy for future studies in proteomic degradation.
- To determine the validity of different methods for analysing and representing DiGE data. So this can be employed in a wide range of DiGE applications in the future.

3.2 Structure of the Chapter

Due to the complexity and size of the work undertaken, it is prudent to lay out a structure to help with the clarity of the chapter. This chapter has been broken up from one large chapter to 5 sub chapters. The following allows for a point of reference to aid the reading of this investigation.

3.5 Rationale of Analysis

This outlines the methods taken to analyse the data giving an overview of what has guided the approach to analysing the DiGE gels, how and why profile analysis was undertaken and the logic used to construct the targeted profiles.

3.6 Results and Discussion

The results section is split into 3 subsections:

3.6.3 Pilot Study

The results of the pilot study help to inform and guide the strategy for the main investigation.

3.6.4 Main Investigation

This section extends the investigation from the pilot study. It looks more deeply into the use of profile analysis, Venn analysis and the benefit of the additional time point.

3.6.6 Validative results

The validative section corroborates the results discussed in the main investigation and pilot study.

3.7 Summary and Conclusions

3.3 Introduction

In the last 100 years there has been an exponential explosion in technologies in everyday life and in scientific research. Proteomics is not an exception. Ever since the development of techniques such as the soft ionisation technique of MALDI (Tanaka, 1988b), the dye had been cast for technological and methodological developments. This rapid growth includes the increased sensitivity and resolution of the mass spectrometer, reverse phase HPLC (Anderson and Murphy, 1976, Eschelbach and Jorgenson, 2006), 2-dimensional gel electrophoresis (O'Farrell, 1975) and other quantitative separation techniques such as DiGE (Ünlü et al., 1997) among others. Sadly, this rush in the development of technology has not been matched in downstream laboratory processes such as sample preparation which can be a severe limiting factor for obtaining good quality proteomics data (Kikuchi, 2007, Fountoulakis, 2001, Palmer, 1988). In addition there is an ever increasing interest in tissue as being a biochemical vault for proteomic, genomic, biomarker and clinical diagnostic studies. As a consequence, in order to capitalise on these technological advances, there is now a higher demand than ever for maintaining stringent sample integrity to match. This is particularly true in areas such as biomarker discovery which has a requirement for an ever increasing level of specificity, sensitivity and reproducibility (Kikuchi, 2007, Lemaire, 2007).

3.3.1 Sample preparation and protein degradation

An exigent area of particular challenge is in maintaining sample integrity of fresh tissue samples where proteomic degradation can be rapid and variable (Sköld et al., 2007a, Richard, 2008). The deviation and complexity of the average tissue sample or cell lysate makes assessing degradation a very demanding prospect. Whereas technology and methodologies can be assessed relatively easily with calibration standards and known samples, the complexity and inconsistency of endogenous proteolytic enzymes coupled with sample handling, short and long term storage and post processing can cause the fluctuation of sample constitution making it too erratic for any meaningful quantitative proteomic analysis. Understanding how tissue degrades and markers of it occurring, could lead to strategies for limiting such degradation of the newly named degradome (Scholz et al., 2010b). The importance of sample preparation is unequivocal but up until recent years has been neglected.

However, there is a new trend for research in this field due to the potential impact on most areas of biochemistry and molecular biology. Neuro-tissue remains the most characterised area but studies regarding other tissues are starting to emerge (Che, 2005, Sköld et al., 2007a, Scholz et al., 2010b, Svensson, 2003).

Tissue integrity compromises can occur at the point of collection to the point of analysis (Fountoulakis, 2001, Palmer, 1988). The development of a treatment applied to samples at the early stages of handling after collection would provide the best solution to limit proteome variability. Freeze-thawing and tissue sectioning causes an untargeted and non-reproducible cell fracture. Processing can lead to micro and macro heat fluctuations. In addition, processing methods such as freeze thawing multiple times has been shown to enhance certain degradative products in some studies, as does long term storage (Gao et al., 2007), although for completeness some markers are robust for a limited number of cycles (Bao and Zuo, 2009). A number of strategies have been used to process tissue in various ways including but not exclusively; fixing and embedding, freezing, microwave treatment, snap-freezing, freeze drying and protease inhibition (Troiano et al., 2009, Ruijter ET, 1997, Mizuhira V, 1996, Login and Dvorak, 1988). However proteolytic enzyme activity can be rapid (Sköld et al., 2007a) and cause a detrimental effect on further analysis. Another strategy which has received limited attention, is the use of heat treatment. For heat treatment to be successful, it would need to be rapid and evenly applied. Previously described approaches including microwave treatment have had problems regarding selective heat dispersion leading to “hot and cold spots”.

The progression of tissue degradation and the proteolytic pathways post-mortem would be useful. Recently, post-mortem changes have been shown to display a high variability, some occurring up to the 6 hour time frame. Whereas degradation within 6 hours is generally expected, the period surrounding the onset of death has also been shown to have major influences on the preservation and degradation of the sample (Sköld et al., 2007a, Richard, 2008). The presence of a marker to quantify the condition of post-mortem tissue would be most useful

An alternative technology to prevent proteomic degradation and allow stable storage has been developed and warrants further investigation. A proteomic

stabilisation system call Stabilzor T1 denaturing device (Denator AB, Gothenburg, Sweden) rapidly and homogeneously raises tissue to a temperature of 95°C under a high pressure vacuum in order to try and curb endogenous proteolytic and other enzyme activity, such as phosphatases. The inactivation of phosphatases would inhibit the cleavage of phosphate groups, thus leaving a greater degree of proteins with this PTM. This could lead to large amount of additional spots on the gels that might otherwise not be seen. A phosphate has a molecular mass of 94.97 Da. This would cause a shift up the gel compared to the protein without the PTM. It may also affect the pI and therefore the horizontal position on the gel.

On the inception of this study the concept of heat stabilization of tissue was a relatively new idea. Since the publication of the paper by our group in 2010: “*Stopping the clock on proteomic degradation by heat treatment at the point of tissue excision*” (Goodwin et al, 2010), there has been an increase in the number of publications, both accessing and using the Denator Stabilzor T1, from approximately 16 to 29, showing that this idea is gaining ground (Spellman et al., 2013, Zhang et al., 2012, Ye et al., 2012b, Ye et al., 2012a, Lundby et al., 2012b, Lundby et al., 2012a, Goodwin et al., 2012, Ahmed et al., 2012, Smejkal et al., 2011, Scholz et al., 2011, Kultima et al., 2011, Kokkat et al., 2011, Kennedy et al., 2011, Finoulst et al., 2011, Colgrave et al., 2011, Ahmed and Gardiner, 2011, Scholz et al., 2010a, Rountree et al., 2010, Lull et al., 2010, Goodwin et al., 2010, Chughtai and Heeren, 2010, Svensson et al., 2009b, Robinson et al., 2009, Kultima et al., 2009, Kaletaş et al., 2009, Grassl et al., 2009, Goodwin et al., 2008c, Sköld et al., 2007a, Svensson et al., 2007).

The stabilizing effect of the denator system on proteomic degradation was assessed using murine brain and heart tissue. Samples were either snap-frozen and stored at -80°C, heat-treated and stored at -80°C or both heat-treated and snap-frozen stored at -80°C. They were then compared using a 2DE approach by Robinson et al, 2009. Interestingly, in the heart tissue the snap-freezing and heat treatment were found to be both beneficial but with no real advantage at using the heat treatment over the traditional method of snap freezing. However, in murine brain, there were a significantly greater number of differences between the three methods. They found that brain responded far better to the heat-treatment by reducing the protein fragmentation and preserving the amount of higher weight proteins. This validates the use of

murine brain as a tissue type in the following study (Robinson et al., 2009). Also DiGE suffers from fewer technical limitations compare to 2DE and therefore is a good technique to help validate the finding of this study.

This positive effect in brain tissue has implication to neurological clinical studies and neuroproteomic studies. The stabilising device has been employed in neuroproteomic studies in the search for neuropeptides which can be elusive. Zhang et al, 2012 managed to identify 500 endogenous peptides from mouse hypothalamus and whole brain samples which is 3 times higher than other reported studies. However, the stabilisation of tissue cannot necessarily be attributed to this increase in success, as an array of new extraction protocols was also being employed. As a positive note for the system, it was explained that some of the peptides discovered are those that are often lost to post-mortem changes (Zhang et al., 2012). It has also been shown to preserve the phosphorylation's of neuropeptides (Ahmed and Gardiner, 2011, Lundby et al., 2012b). Much of the work conducted with the stabilisation device has been done in brain tissue using a variety of methods including MS imaging, 2DE, DiGE and LC-MS (Lull et al., 2010, Smejkal et al., 2011, Goodwin et al., 2010, Sköld et al., 2007a, Goodwin et al., 2008c, Ye et al., 2012b, Svensson et al., 2007, Scholz et al., 2010a).

Despite the heavy emphasis on neuroproteomics, the system has also been employed in liver and pancreas tissue in examining the so called degradome. Scholz et al, 2011 compared the snap-freezing of liver and pancreas tissue to heat-treatment using the denator stabilisation device using DiGE and label free MS. They found that snap-freezing produces a degradome with a greater amount of degradative products when compared to heat treatment. They also showed that subsequently heat treating a previously snap frozen sample has an effect of lowering degradative products in both the peptidome and proteome. They used a similar approach as described in the DiGE portion of our published work and the work further described in this thesis chapter (Scholz et al., 2011).

The system does also have some notable limitations. By causing the inactivation of enzymatic activity any studies which use enzyme activity are obviously not going to work. Also, anything that involved needing structural information of the protein or investigation protein-protein interaction is not possible. However, the unfolding of protein structures is used in a lot of

proteomic techniques already. Additionally, the extraction of organelles from treated tissue has not been achieved as yet. Also the use of the stabilisation device also has had detrimental effects of RNA extraction, so any studies which hope to link genomic, transcriptomic and proteomic data is not possible. Finally, as a relatively young idea, it would be necessary to evaluate the technical and experimental limitations in various tissue types in order to state the potential benefits and limitations of this technique in various areas of clinical and proteomic research (Kultima et al., 2011).

An ideal proteomic strategy to assess degradation would allow for the assessment of the global proteome, while at the same time looking for a specific marker to characterise degradation. This is due to the fact that degradation is a systemic, as well as localised problem. There are many possible proteomic strategies for evaluation of this system, but a quantitative approach using 2D-PAGE like Differential in gel electrophoresis allows scope for a broad investigation, while allowing and giving the ability to investigate specific proteins or markers. In order to assess proteomic degradation a strategy using DiGE was initiated, comparing heat-treatment with a more traditional alternative of snap-freezing in liquid nitrogen. DiGE is a well-established technique for gauging proteomic expression differences in different samples (see section 1.4.1). Traditionally, DiGE is used in multiplexed time course or disease versus control experiment, however, it equally lends itself towards accessing protein profile changes. However, with a process such as degradation, having a plethora of contributing factors a single approach will also struggle. DiGE and traditional 2DE are essentially limited to the mid-high mass region and therefore a more dynamic approach is desirable. In addition, with any homogenisation approach to proteomics, there is a loss of the low abundant signals, localised signals and an average effect occurs. Therefore this investigation has been collaborative which has allowed for a multi-faceted approach involving DiGE, Mass Spectrometry Imaging (MSI) and label free quantitation using LC-ESI-MS. MSI is useful as it maintains spatial resolution meaning you can track markers across the tissue and time states. Whereas LC and gel based approaches lose this information, MSI also retains localised variation that would otherwise be lost. It is a relatively new technique that has been pioneered by Caprioli et al. Once again, as a standalone technique it has its own limitations regarding a low mass range of up to 30KDa and gaining

identifications. The low mass range and spatial resolution is complementary to a DiGE approach.

Although 2DE and DiGE has its limitation, it is still recognised as an effective technique and one of the few proteomic methodologies for visualising 1000's of proteins and/or peptides in a single experiment. Ironically, the very limitation of DiGE may afford an advantage when studying degradation. Degradation has been shown to rapidly produce many protein and peptide fragments which masked an attempt to detect neuropeptides in a recent study (Scholz et al., 2010a), in DiGE this complexity will be reduced as they would simply not appear on the gel. Although there is always a concern when losing information, it can be also beneficial to start viewing a problem from a more simplified perspective. LC-MS allows for the inclusion of these smaller molecules.

Traditionally, DiGE analysis consists of looking for intensity differences between groups but degradation is more complex, with protein or peptide spots appearing, disappearing and relative changes being difficult to visualise, all make degradation a greater challenge, therefore the development of a specific set of analysis would further target the way degradation can be visualised and broaden DiGEs remit in terms of its application.

3.3.2 Targeted tissue.

Brain tissue is intrinsically of interest with a multitude of studies involving diseases and disorders with complex aetiology which are poorly understood and preserving tissue integrity is of increasing importance (Richard, 2008, Che, 2005, Svensson, 2003, Fountoulakis, 2001). There is further pressure with the regarding post-mortem time of collection of tissue, where regulation governing autopsy is required. Prediction of such changes would be useful in determining protein stability within different tissue types, along with enzymes patterns within different structures. Predicting such patterns is nearly impossible with careful sample preparation, let alone in tissue collection. Any changes that could be mapped would be of interest to anybody involved in sample preparation of tissue in addition to more specially the Neuropathologist.

3.4 Methods

A detailed account of general methodology is outlined in Chapter 2 of this thesis. Methodology specific to this chapter is given below.

3.4.1 Sample collection, extraction and processing

In line with the U.K. Animals (Scientific Procedures) Act, 1986 and local ethical guidelines male ICR mice (6-8 weeks) were euthanized by cervical dislocation and dissection of the brain was rapidly performed. The brain was removed and bisected into the two hemispheres. One whole hemisphere was taken to be “treated” with focused, rapid and homogeneous heating under pressure using a Stabilzor T1 denaturing device (Denator ABDenator™, Sweden), whilst the other was immersed in liquid nitrogen and snap-frozen. The device is preheated at 95°C in preparation for the tissue. Upon dissection of the tissue, tissue for heat treatment was placed onto the maintainer tissue cards. These cards have a plastic film for maintaining vacuum. The cards are then placed into the slot of the Stabilzor T1 denaturing device. The device rapidly engages a pump to cause a vacuum seal around the tissue. The device has a reactive heating algorithm which heats the tissue rapidly, but in a controlled manner as not to overheat and destroy the tissue. This occurs within 30 seconds from the time the card is placed in the device. This process can be performed on fresh tissue (as was the case in this study) or frozen tissue. The device is a bench top device and reasonably portable. This can be seen in Figure 3-1

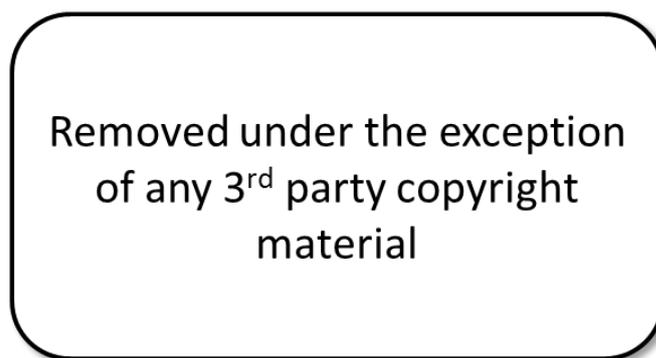


Figure 3-1: A) Stabilzor T1 denaturing device (Denator ABDenator™, Sweden) used to rapidly apply heat treatment to tissue samples and B) maintainer cards, used to place the tissue into vacuum and for each storage. Taken from <http://www.denator.com>.

Subsequent to treatment, the treated hemisphere was also immersed in liquid nitrogen. Sacrifice, dissection and treatment was accomplished in less than 60 seconds. Samples were then stored at -80°C in anticipation of use

downstream. Samples were always handled with care to prevent external factors affecting the possibility of degradation. Consequently protease and phosphatase inhibitor (Phosphatase Inhibitor Cocktail 1, Sigma Aldrich, UK and Protease Inhibitor Cocktail for use with mammalian cell and tissue extracts, DMSO solution, Sigma Aldrich, UK. Both are in tablet form and half a tablet dissolved in 10ml of lysis buffer) were added at the earliest possible time point and all work was carried out on dry ice to minimize any effect external heat may have had.

Sectioning was performed using a cryostat microtome (Leica Microsystems CM 1900UV, UK). Multiple sections were taken and divided up between eppendorfs and MALDI ITO coated glass slides. Tissue sections were cut at a thickness of 14µm in a pre-chilled chamber of -20°C and a sample stage temperature of -19°C. Eppendorfs used for storage were kept and allowed to pre-chill in the chamber as well, before being transferred to dry ice. Sections used in MALDI-Mass Spectrometry Imaging (MSI), were thaw mounted and stored on dry ice further to long term storage at -80°C. Sections were taken from the mid-brain section between ¼ and ½ the coronal depth of the brain. It is at this point half were raised to room temperature for 10 minutes, in pilot investigation (Table 3-1) and for 10 and 20 minutes, in the main investigation (Table 3-2). Processing at this point was separated and all samples were treated individually to allow replicates to be performed. Four experimental replicates for pilot investigation (n=4) and 6 experimental replicates for the main investigation (n=6). Protein lysis and extraction was performed immediately after treatment in a DiGE compatible lysis buffer with the addition of protease inhibitor and DNase 7 M Urea, 2 M thioruea, 4% CHAPS [w/v], 30 mM Tris base. Following addition of lysis buffer, the samples were subject to 3 Cycles of snap freezing, thawing and 4 x 5 minute cycles in an iced sonication bath with 1 minute cooling on ice between sonication.

Extractions were followed by protein precipitation and clean-up using Ettan™ 2D Clean-up Kit (GE Healthcare, Bucks, UK cat #80-8484-51). Minimal labelling reactions were performed with Cyanine dyes 2, 3 and 5 using protocol specified in the product booklet (GE Healthcare, Bucks, UK #25-8009-83/84.) and in the general methods. In short 50µg of protein was used for analytical gels and 500µg for preparative gels which were stained using Sypro Orange (Sigma Aldrich, Dorset, UK, cat #S5692-500UL). The reaction was performed

at pH 8.5, using 400 pmol of CyDye incubated on ice for 30 minutes in the dark. The reaction was stopped using the addition of 1 mM lysine. They were loaded (as shown in Table 3-1 and Table 3-2) for 1st dimension separation which was carried out using IEF on 24 cm IPG strips of pH 4-7 [GE Healthcare, Bucks, UK cat # 17-6002-46]) with a minimum of 75000 Volt hours at 20°C with a program of 30 V step and hold for 12 hours, 300 volts step and hold for 2 hours, 1000 volts gradient for 2 hours, 8000 volt gradient for 5 hours, 8000 volts step and hold for 8 hours and 1000 volts step and hold for up to 24 hours to avoid strip diffusion prior to next step. Prior to 2nd dimension separation by SDS-PAGE, an equilibration was performed using SDS equilibration buffer (50 mM Tris-HCL, pH 8.8, 6 M urea, 30% glycerol [w/v] 2% SDS [w/v] and 0.002% BPB [w/v]) with 10 mL/gel followed by reaction with DTT (10 mg/mL) for 15 minutes then with iodoacetamide (25 mg/mL) for 15 minutes to prevent reduction/alkylation. 2nd dimension separation was performed at 1-2 watts per gel for approximately 12 – 15 hours or until the dye front reached the bottom of the gel. Gels were imaged using GE Healthcare Typhoon 9400 Series Variable Imager at 100µm resolution after optimization of photomultiplier voltages using a pre-scan at a resolution of 1000µm. Gels were then loaded, (see chapter 2) matched and analysed, (DeCyder Version 5.01.01, GE Healthcare, Bucks, UK), and spots selected for picking using Ettan™ Gel Handling Work Station and MS identification. The methodology of this analysis is given in 3.5.2. A schematic of the workflow can be seen in Figure 3-2.

3.4.2 Identification of proteins from gel spots

Spots were picked, tryptically digested and spotted with *o*-Cyano-4-hydroxycinnamic acid in 50% acetonitrile/0.5% trifluoroacetic acid using a GE Healthcare Ettan Spot Handling Work Station and then analysed on 4700 Proteomics Analyser (Applied Biosystems, Cheshire, UK) MALDI-ToF-ToF-MS using standard settings. MS/MS was performed on the top 10 precursor ions in each spot. Any unidentified spots of particular interest were further analysed by LC-MS/MS on a Dionex Ultimate+ LC system coupled to a QStar Pulsar I (Applied Biosystems, Cheshire, UK). GPS Explorer and MACOT Daemon Software was used to automated submission of collected data to MASCOT database searching software for searching with fixed modification of carbamidomethyl (C) and variable modification of oxidation (M), peptide and MS/MS +/- tolerances of 0.8 Da searching NCBI database on Mus

musculus taxonomy with 1 missed cleavage allowed. Spots identified as keratins were excluded.

3.4.3 Western Blot analysis

To aid validation western blot analysis was performed on treated and snap-frozen samples 0, 10 and 20 minutes. A detailed protocol for western blotting can be seen in section 2.1.5 of the general methods in chapter 2.

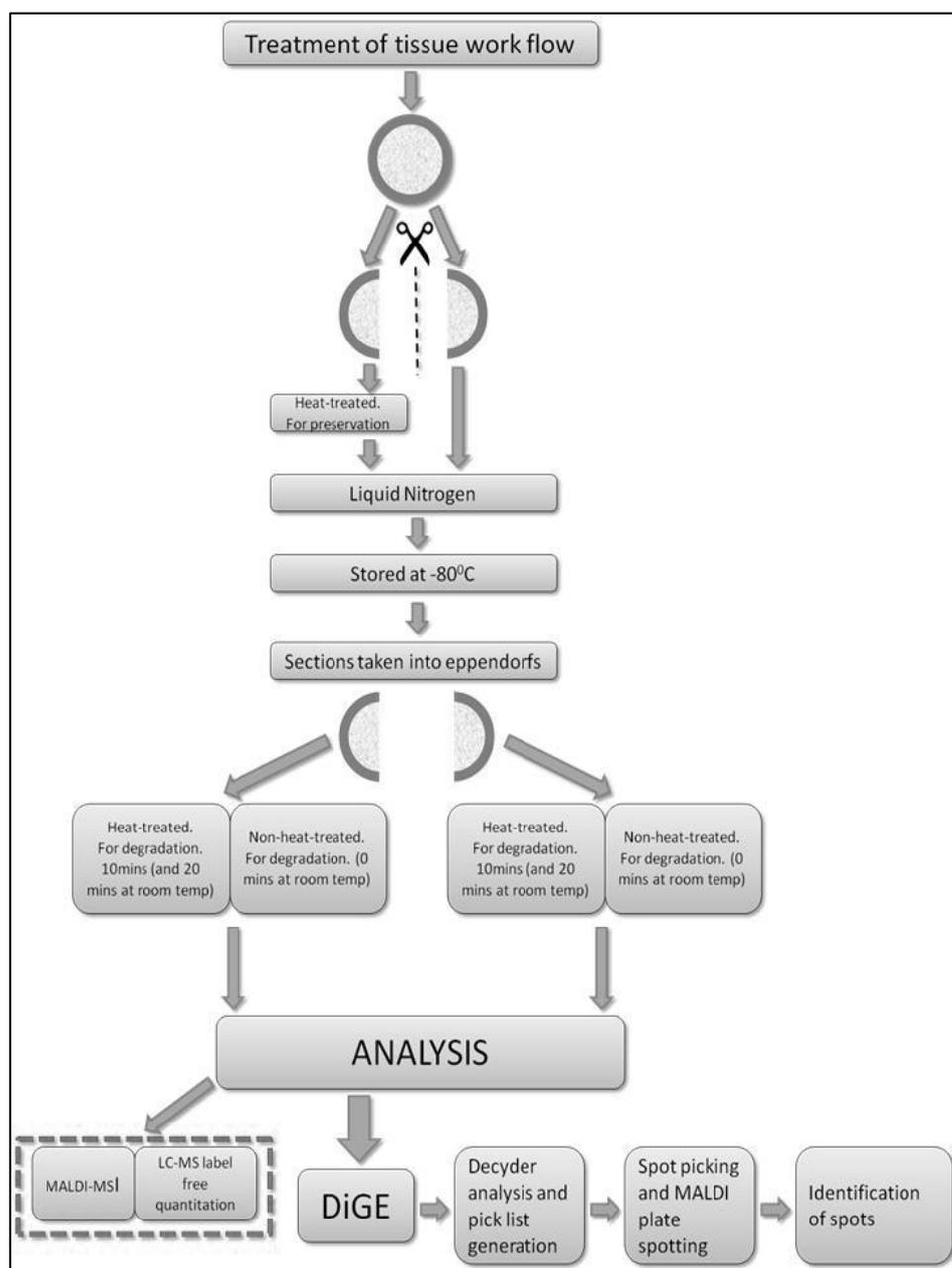


Figure 3-2: Schematic of experimental work flow. Brain tissue hemispheres were bisected and either heat treated and snap-frozen or just snap-frozen. Samples were then stored at -80°C. They were then taken onto DiGE analysis. MALDI-MSI was performed separately by Dr R. J. A. Goodwin, Research Associate, University of Glasgow and LC-MS using label free quantitation was performed by Miss H. Allingham, Ph.D student, University of Glasgow. All experiments were performed in parallel. The larger proteins considered using DiGE and the smaller markers considered using MALDI-MSI and LC-MS.

3.4.4 Experimental design

The investigation was effectively broken up into two sections. A pilot investigation to help develop the laboratory and analytical techniques and then the main investigation which eliminates and minimises any of the difficulties experienced in the pilot investigation. It is noted that the experimental design in this chapter does not have a dye switch applied and therefore the design may have introduced a greater level of experimental bias that would be ideal. This was the first set of DiGE experiments undertaken and therefore the experimenter was not aware of the issues with differential labelling until the experiment. This will of course be reflected in the conclusions that can be drawn.

The experimental design used for the DiGE gels in the pilot investigations were as follows:

3.4.4.1 Pilot investigation: Time points Treated/Snap-frozen = 0, 10 minutes.

	Cy5	Cy3	Cy2
gel 1	Treated at 10 minutes	Treated at 0 minutes	internal standard
gel 2	Treated at 10 minutes	Treated at 0 minutes	internal standard
gel 3	Treated at 10 minutes	Treated at 0 minutes	internal standard
gel 4	Treated at 10 minutes	Treated at 0 minutes	internal standard
gel 5	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 6	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 7	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 8	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 9	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
			n=4

Table 3-1: Experimental design of gels ran for pilot investigation: Time points Treated/Snap-frozen = 0, 10 minutes.

Additionally, the experimental design used for the DiGE gels in the main investigation was as follows:

3.4.4.2 Main investigation: Time points Treated/Snap-frozen = 0, 10, 20 minutes.

	Cy5	Cy3	Cy2
gel 1	Treated at 20 minutes	Treated at 0 minutes	internal standard
gel 2	Treated at 20 minutes	Treated at 0 minutes	internal standard
gel 3	Treated at 20 minutes	Treated at 0 minutes	internal standard
gel 4	Treated at 0 minutes	Treated at 10 minutes	internal standard
gel 5	Treated at 0 minutes	Treated at 10 minutes	internal standard
gel 6	Treated at 0 minutes	Treated at 10 minutes	internal standard
gel 7	Treated at 10 minutes	Treated at 20 minutes	internal standard
gel 8	Treated at 10 minutes	Treated at 20 minutes	internal standard
gel 9	Treated at 10 minutes	Treated at 20 minutes	internal standard
gel 10	Snap-frozen at 20 minutes	Snap-frozen at 0 minutes	internal standard
gel 11	Snap-frozen at 20 minutes	Snap-frozen at 0 minutes	internal standard
gel 12	Snap-frozen at 20 minutes	Snap-frozen at 0 minutes	internal standard
gel 13	Snap-frozen at 0 minutes	Snap-frozen at 10 minutes	internal standard
gel 14	Snap-frozen at 0 minutes	Snap-frozen at 10 minutes	internal standard
gel 15	Snap-frozen at 0 minutes	Snap-frozen at 10 minutes	internal standard
gel 16	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 17	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
gel 18	Snap-frozen at 10 minutes	Snap-frozen at 20 minutes	internal standard
			n=6

Table 3-2: Experimental design of gels ran for the main investigation: Time points Treated/Snap-frozen = 0, 10 minutes.

3.4.5 Statistical methods

In order to assess the validity of using the parametric tests of Student's t-test and ANOVA, it is important to evaluate the assumptions of these tests. As discussed in section 1.7 it would be erroneous to assume that data follows the rules of assumed normality and homogeneity of variance. If the data does not follow these assumptions, the validity of conclusions made is compromised and then the numbers of false positives are likely to increase. All statistical calculations have been performed in either the DeCyder 2D 7.0, GE Healthcare software directly with further statistical analysis done using SPSS (Statistical Package for the Social Sciences) version 17.0.1 (2008).

In order to perform statistical operation in SPSS, the data was exported via xml exporter module from the DeCyder work space. The results are exported in the format of normalised spot volume ratios between the cy3/cy2 and cy5/cy2 channels which is equivalent to log standardised abundances ($\log_{10}SA$). This was confirmed by GE Healthcare. All data was exported for matched gel spots. This data was then converted, using an Excel macro from GE Healthcare, to convert this into standardised abundances (SA). Normality testing was then performed in SPSS 17.0.1 on $\log_{10}SA$ using the Shapiro-Wilk statistical test. This is an accepted robust test and has been employed in other studies concerning DiGE (Karp and Lilley, 2005). A spot was considered of non-normal distribution if a p-value was returned of 0.05 or less. This was performed on spots that had a full set of repeats across all gels, those that had a full set of repeats across all gels and an ANOVA score of 0.05 or less and those that had an ANOVA score of 0.05 or less which has missing values replaced using k-Nearest Neighbour (KNN) algorithm in SPSS 17.0.1 to replace values. The attribute k was set to automatic. This is a method of machine learning also used in other DiGE studies and is considered a good way of replacing missing values (Pedreschi et al., 2008a).

To assess the homogeneity of variance (an assumption of equal sample variance must hold if the parametric t-test and ANOVA can be used) the SA and $\log_{10}SA$ data across all groups was analysed using the Levene's test for homogeneity. This was performed on spots that had a full set of repeats in every group for each master spot. A spot was considered not to have homogeneity of variance when it has a p-value of 0.01 or less. Additionally graphs were produced to assess visually the spread of data.

The ANOVA calculation and student t-test was performed within DeCyder 2D 7.0, GE Healthcare in the BVA module. In addition to testing these assumptions, the issue of multiple testing of data has been addressed by applying a p-value correction. This has been applied directly in DeCyder via the BVA module option for the application of a false discovery rate. The FDR in DeCyder is an adaptive approach to recalculate all p-values to reflect a FDR but maintains a selectable threshold of 0.05. Thus meaning the FDR is set to the equivalence of 5% for all results shown. The FDR algorithm applied is described in detail in Benjamini and Hochberg, 2000. In the bulk of the analysis no adjustments have been made for missing values as the DeCyder software does not cater for this.

3.5 Rationale of Analysis

3.5.1 Introduction

Due to the complexity of this investigation and to aid the deconvolution of the aims outlined in section 3.1 on page 142 (the biological study of degradation and the review of analytical options using DiGE); it is prudent to delineate the approach to analysis.

3.5.2 Methodology of DiGE analysis

3.5.2.1 DiGE Pilot investigation: Time points Treated/Snap-frozen = 0, 10 minutes

Principally, the basis for analysis for the pilot investigation is outlined in Figure 3-3. In order to consider if treatment has been successful, any protein or peptide spots that are present and used in analysis should show no difference between treated, snap-frozen = 0 minutes and treated = 10 minutes would be expected to have no statistically significant difference. This can be visualised and mined for using different methods. Profile analysis allows for a visual validation by seeing the relationship graphically. The Venn analysis allows for statistical differences or no differences to be seen and compared with the profile matches. Principal Component Analysis (PCA) can then be used as a second statistical validation tool to look for differences between groups.

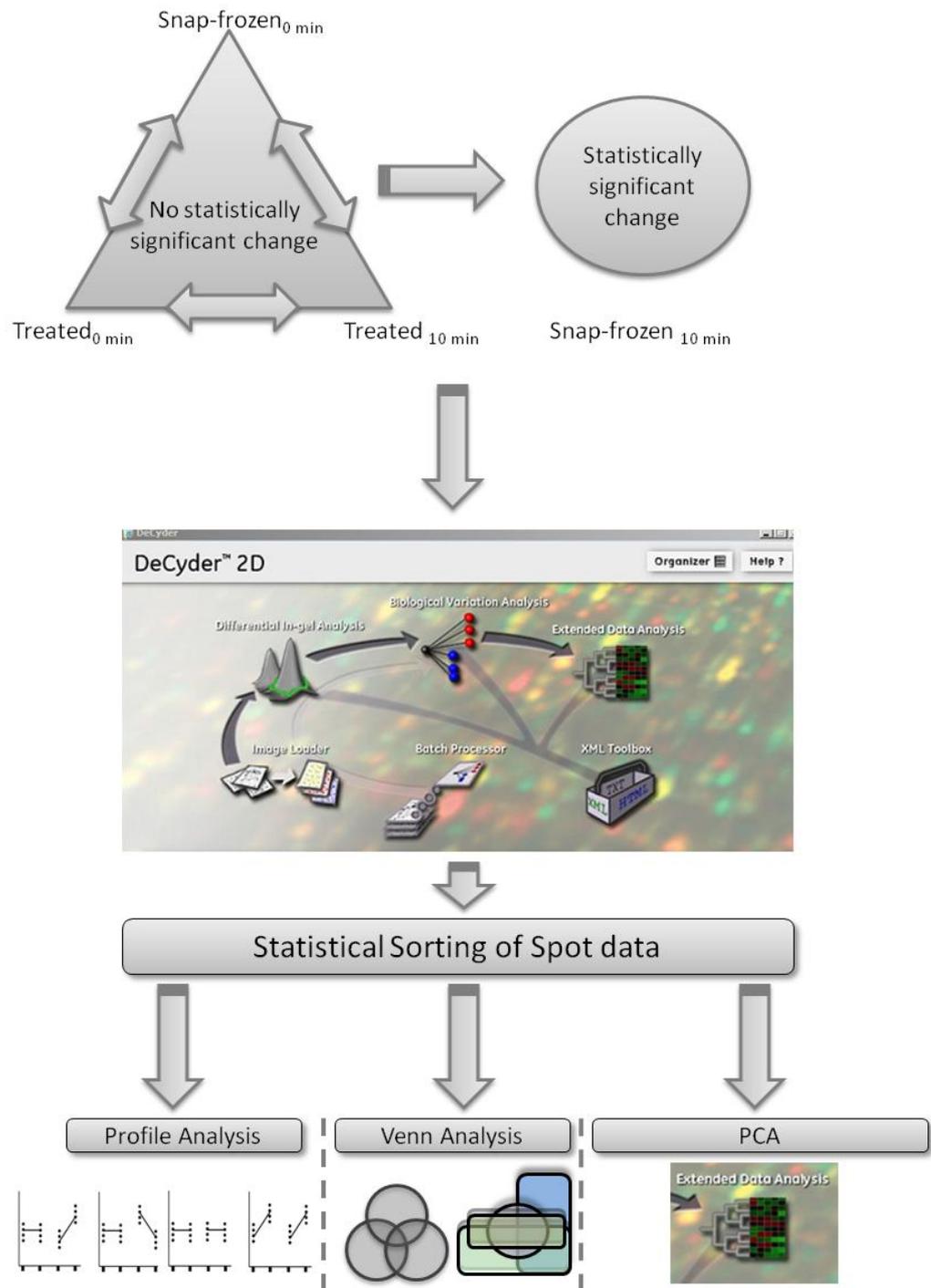


Figure 3-3: Schematic of Analysis work flow. If spots are to be considered stabilised by the treatment, then treated, snap-frozen = 0 minutes and treated = 10 minutes would be expected to have no statistically significant difference. The bulk of analysis was performed in DeCyder software. The analysis was split into three main areas; Profile, Venn and principal component analysis.

Once experimental work has been completed, the gel files are uploaded onto the DeCyder software database. Initially, profile analysis is performed. The various Cy channels are sorted into corresponding experimental groups; Treated = 0 minutes, Treated = 10 minutes, Snap-frozen = 0 minutes and Snap-frozen = 10 minutes. One-way ANOVA is performed across groups, giving a p-score/value for each spot across all groups. This allows the protein spots to

be sorted in the spot table on the basis of their ANOVA score. A cut-off score of $p \leq 0.05$ was used. At this stage, only protein spots with a significant ANOVA cut-off p-value are considered. Experimentally obtained spot profiles are manually compared and checked against predicted spot profiles, the key profile is shown in Figure 3-5. Stabilisation was surmised to have occurred when a marker's intensity distribution profile of a particular spot across groups matched these predicted profiles. It should be noted profiles are deemed inclusive on the basis of relative intensity not absolute intensity. Profiles identified as being stabilised (in either treated or snap-frozen) were selected for picking and MS identification. The proportions of spots found for each profile was noted.

Additionally Principal Component Analysis (PCA) was performed using DeCyder Extended Data Analysis module (EDA). PCA was simply used for discerning whether the different groups were distinguishable from each other globally. Following profile analysis, an approach using Venn diagrams was performed. This required performing numerous t-tests and exporting and sorting data in excel. This is discussed in section 3.5.2.4.

3.5.2.2 Predicted mechanisms of degradation.

In order to consider the expected DiGE analysis profiles it is necessary to appraise how degradation may affect the distribution of spots on a 2D gel map.

Degradation occurs when endogenous proteolytic enzymes cleaves a protein or peptide chain at a particular site. Broadly, there are three possible modes for this to occur, exoprotease and endoprotease activity, or both. Exoproteases cleave at one of the end termini and endoprotease cleaving in-between. The effects on 2D spot maps are summarised in Figure 3-4. Exoprotease activity on the spot map would be characterised by the appearance of the large fragment spot close to the site of the original protein and the small fragment spot further down or even off the bottom of the gel. The spots may have shifted horizontally due to a change in their isoelectric focusing point. With endoproteases the production of two fragments may be expected, smaller than the parent protein but more even in size giving rise to two spots further down the gel, possibly in close proximity to each other. However, with various different kinds of proteases it is far more feasible that a combination of events is happening, giving rise to multiple spots from a single parent protein, this is one of the reasons accessing degradation globally is difficult.

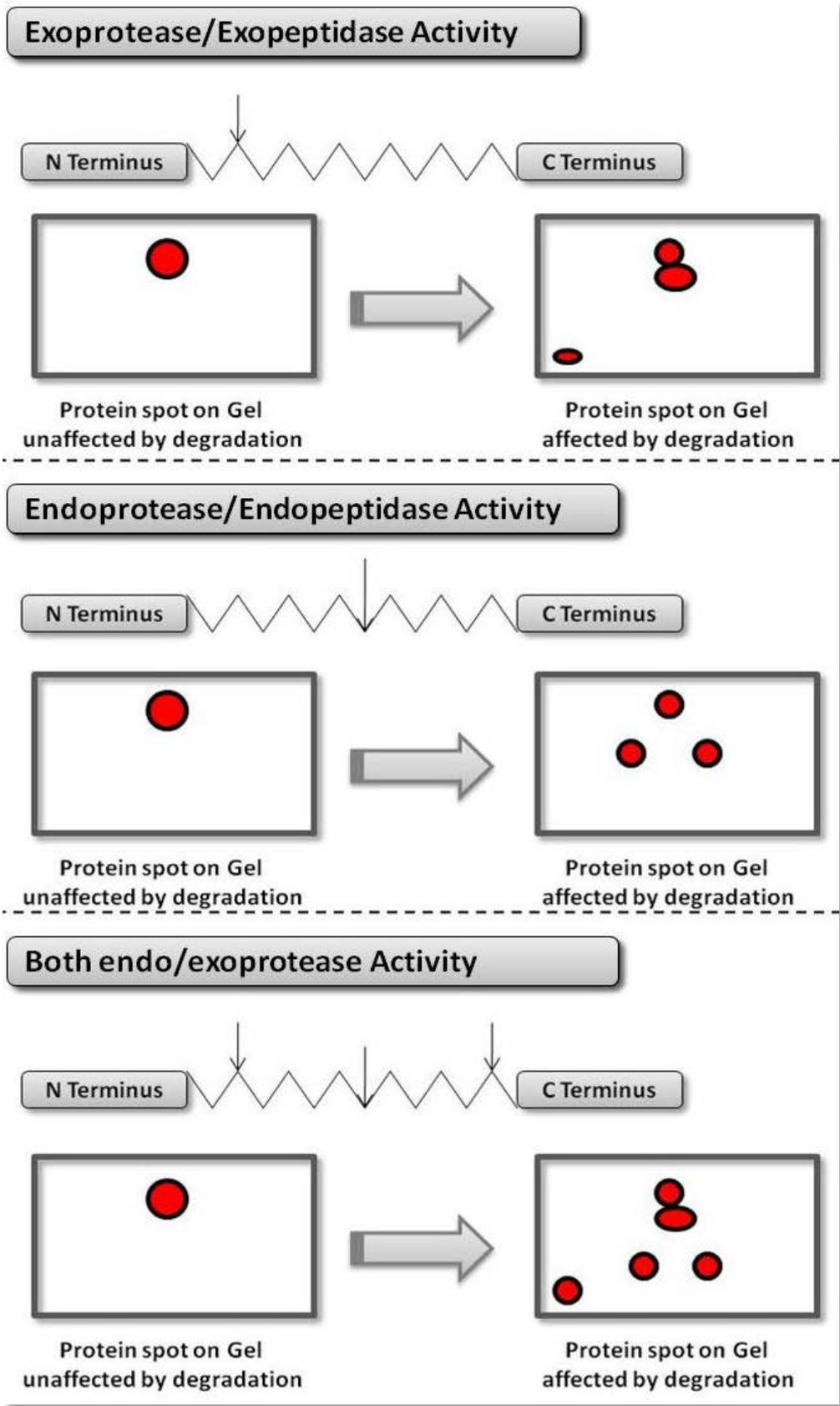


Figure 3-4: Schematic of the possible effect of degradation on protein spots on a 2D gel map. Degradation occurs either at the terminus or in-between or both in a protein chain. This has a consequential effect on how they appear on a gel spot map, and the detected intensity for that gel spot.

3.5.2.3 Predicted profiles for pilot investigation.

Experimental profiles were manually compared against predicted profiles. Broad categories have been defined however, within each of these classifications several different profiles were considered but these 4 categories were assigned; treated stabilisation, snap-frozen stabilisation, no change and unclassified. These can be seen in Figure 3-5

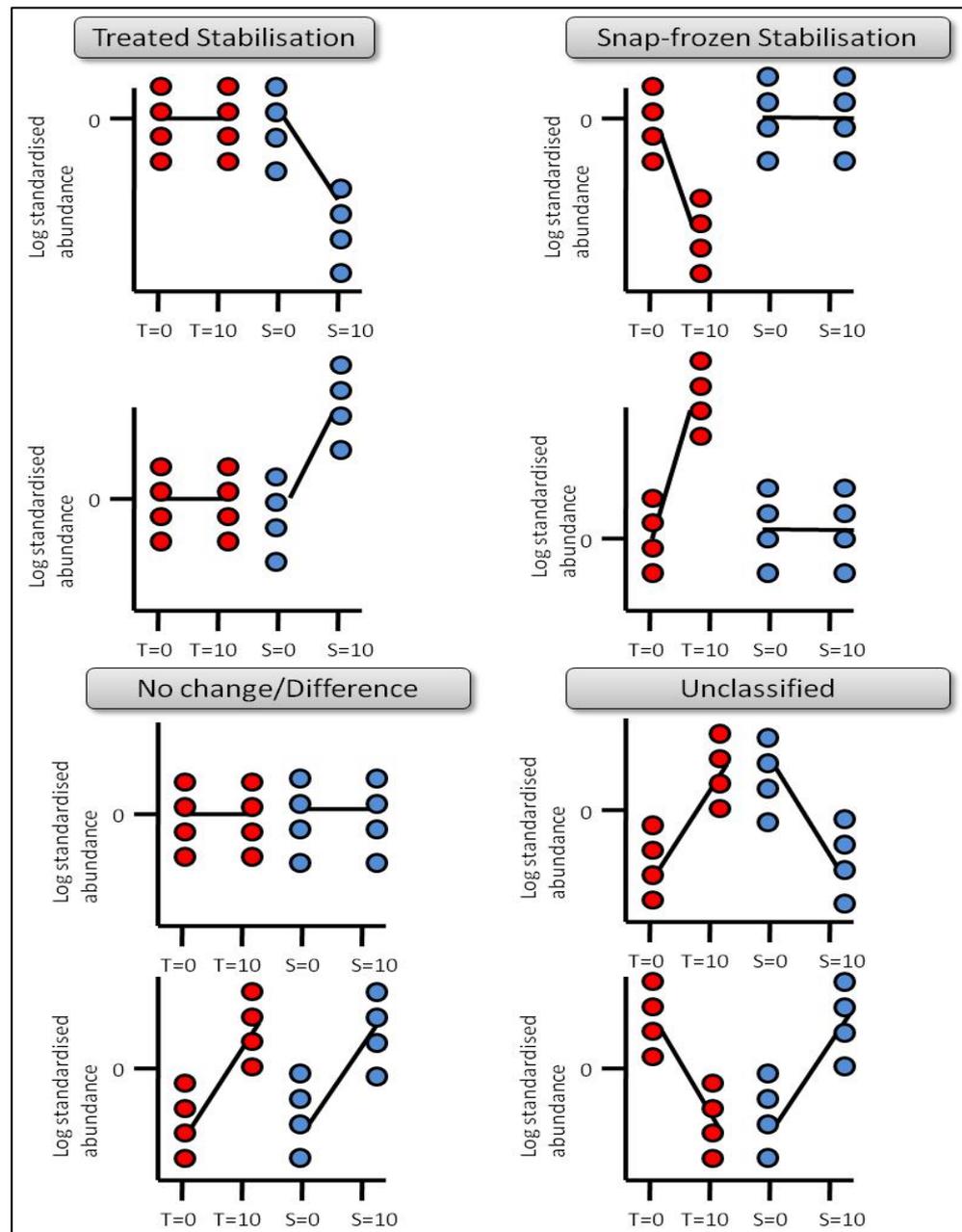


Figure 3-5: Predicted profiles of spot intensities. Treatment stabilisation would expect to be characterised by no change between T= 0, T=10 and S=0 minutes and a change in S = 10. Snap-frozen stabilisation is characterised by a change in T = 10. Red=treated and Blue=snap frozen.

3.5.2.4 Venn analysis for pilot investigation.

The Venn analysis was performed by compiling all t-test comparisons, including both significant score ($p \leq 0.05$) and non-significant score ($p > 0.05$) in order not to exclude spot data.

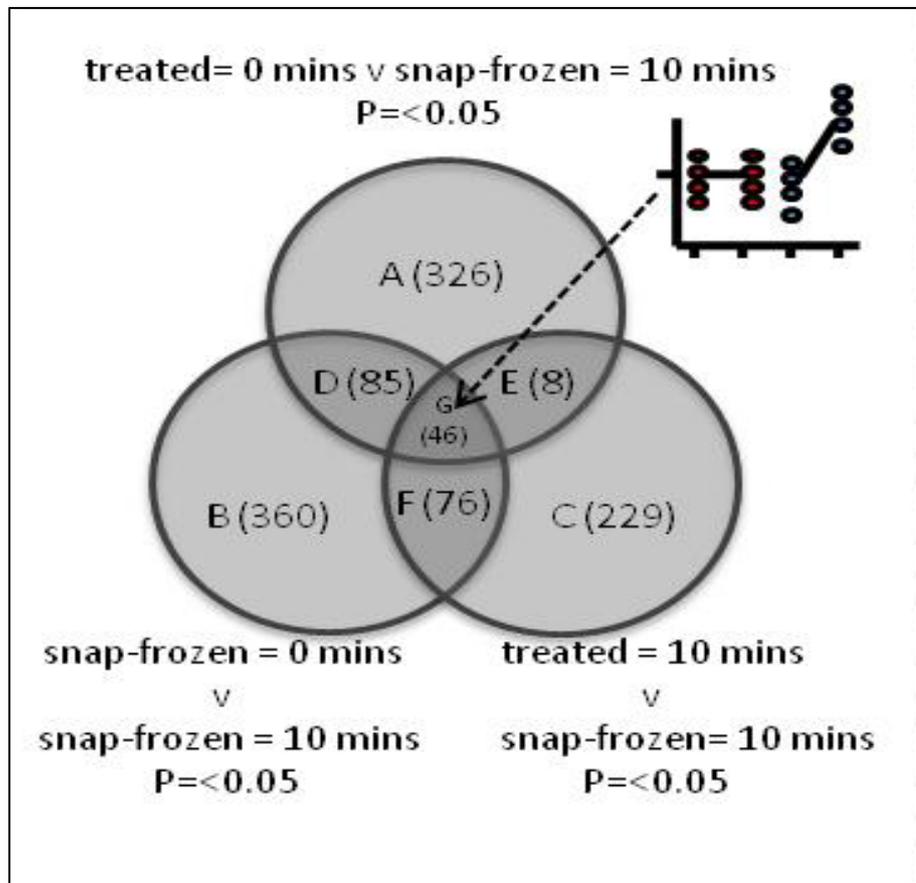


Figure 3-6: Example of three way Venn diagram. This particular Venn diagram is targeting treated stabilisation by using t-tests to focus on candidate spots.

3.5.2.5 DiGE Main investigation: Time points Treated/Snap-frozen = 0, 10, 20 minutes

Principally the basis for analysis is outlined in Figure 3-7

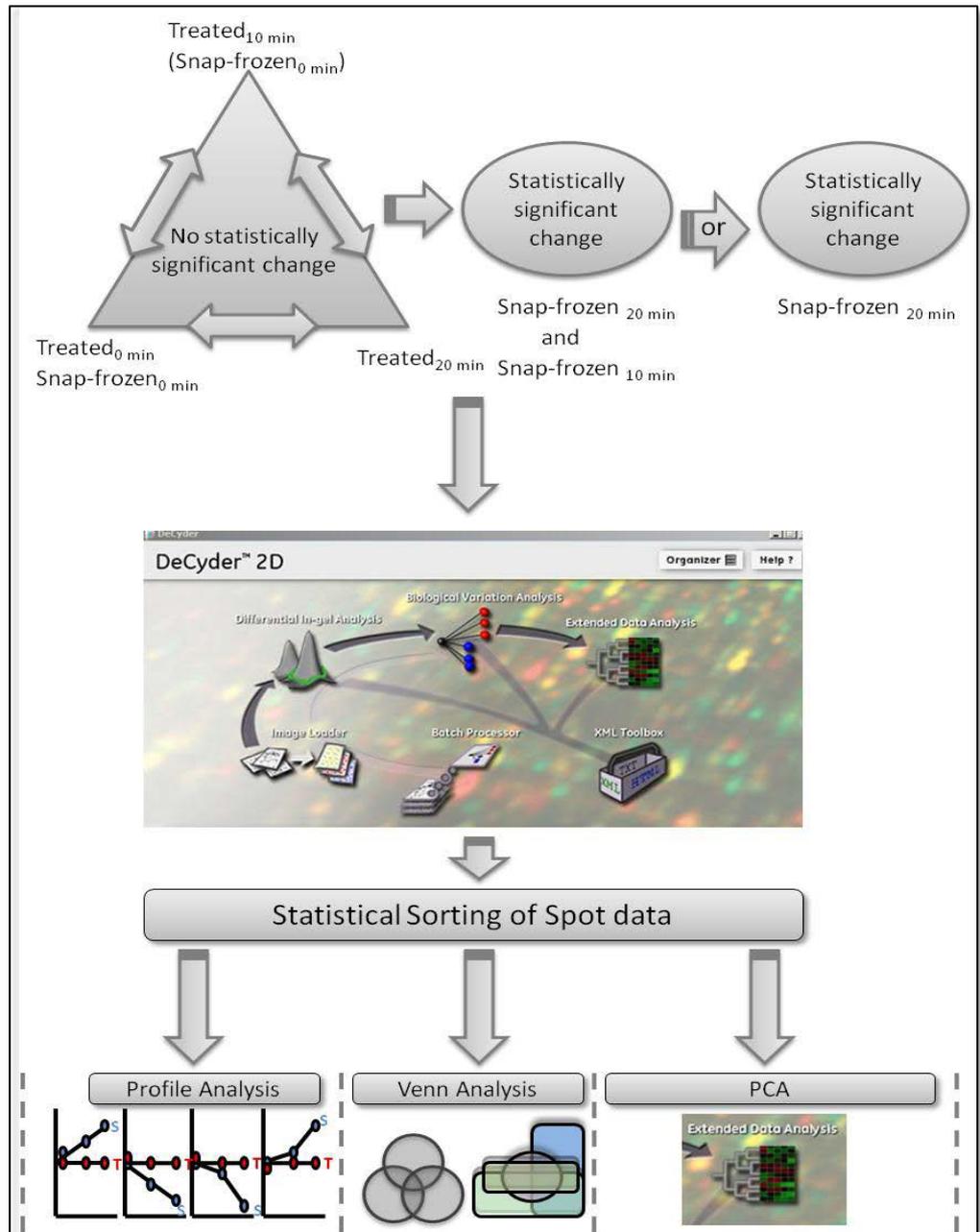


Figure 3-7: Schematic of Analysis work flow. If spots are to be considered for stabilised by the treatment, then treated = 0, 10, 20 and snap-frozen = 0 minutes would be expected to have no statistically significant difference. The bulk of analysis was performed in DeCyder software. The analysis was split into three main areas; Profile, Venn and principal component analysis.

Initially, profile analysis is performed (following profiles in section 3.5.2.6). The various Cy channels are sorted into corresponding experimental groups; Treated = 0 minutes, Treated = 10 minutes, Treated = 20 minutes Snap-frozen = 0 minutes, Snap-frozen = 10 minutes and Snap-frozen = 10 minutes. In addition, to pilot investigation, they are assigned conditions to allow them to be

sorted graphically and statistically by two-way ANOVA. The Conditions system in DeCyder allows the separation and analysis between time points and treatment. First One-way ANOVA is performed across groups, giving a p-score/value for each spot across all groups. A cut-off score of $p \leq 0.05$ was used. Further to pilot investigation a two-way ANOVA was performed and A cut-off score of $p \leq 0.05$ was used. The data was first sorted by one-way ANOVA and the experimentally obtained spot profiles are manually compared and checked against predicted spot profiles. This process was repeated using two-way ANOVA. The predicted profiles are given in Figure 3-8, Figure 3-9 and Figure 3-10 and in comparison to pilot investigation is more diverse in categories.

There are three options when using two-way ANOVA. All three options operate between all experimental groups, but this can be further applied to either; condition 1 (treated vs. snap-frozen), condition 2 (0 vs. 10 vs. 20) or interaction (condition 1 vs. condition 2).

As with pilot investigation, stabilisation was surmised to have occurred when a marker's intensity distribution profile of a particular spot across groups matched these predicted profiles. It should be noted that profiles are deemed inclusive on the basis of relative intensity, not absolute intensity. Profiles identified as being stabilised (in either treated or snap-frozen) were selected for picking and MS identification. The proportions of spots found for each profile was noted.

Principal Component Analysis (PCA) was performed using DeCyder Extended Data Analysis module (EDA). PCA was simply used for discerning whether the different groups were distinguishable from each other globally.

Following profile analysis, an approach using Venn diagrams was performed. This required performing numerous Student's t-test and exporting and sorting data in excel. This is discussed in section 3.5.2.7.

3.5.2.6 Profile analysis for the main investigation.

Experimental profiles were manually compared against predicted profiles. As with pilot investigation broad categories have been defined but in contrast with experiment a degree of the complexity has been retained in order to gain greater information. Therefore a greater number of profile categories have been defined in Figure 3-8, Figure 3-9 and Figure 3-10.

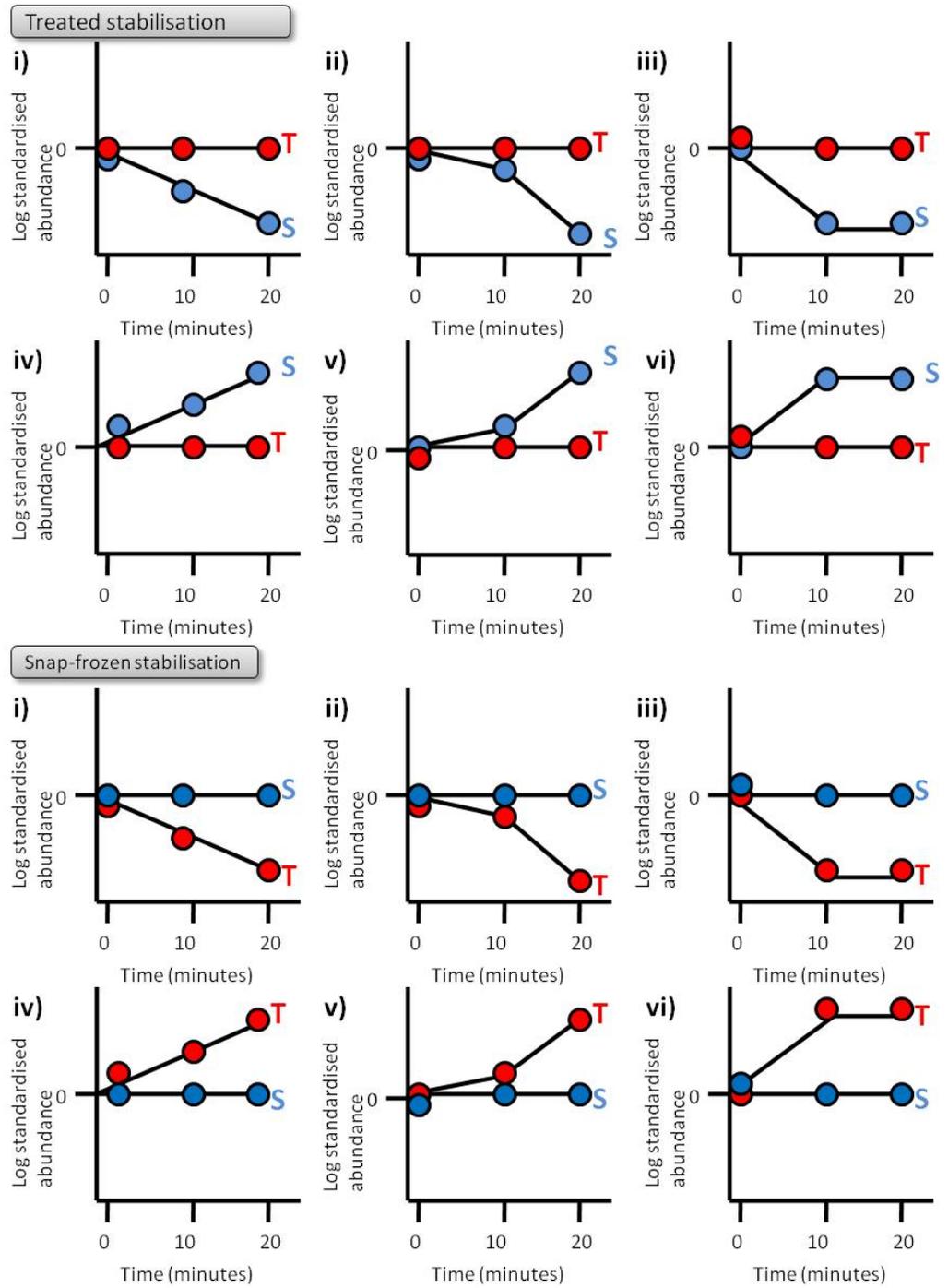


Figure 3-8: Predicted profiles on spot intensities. Showing treated and snap-frozen stabilisation. Red=treated and Blue=snap frozen.

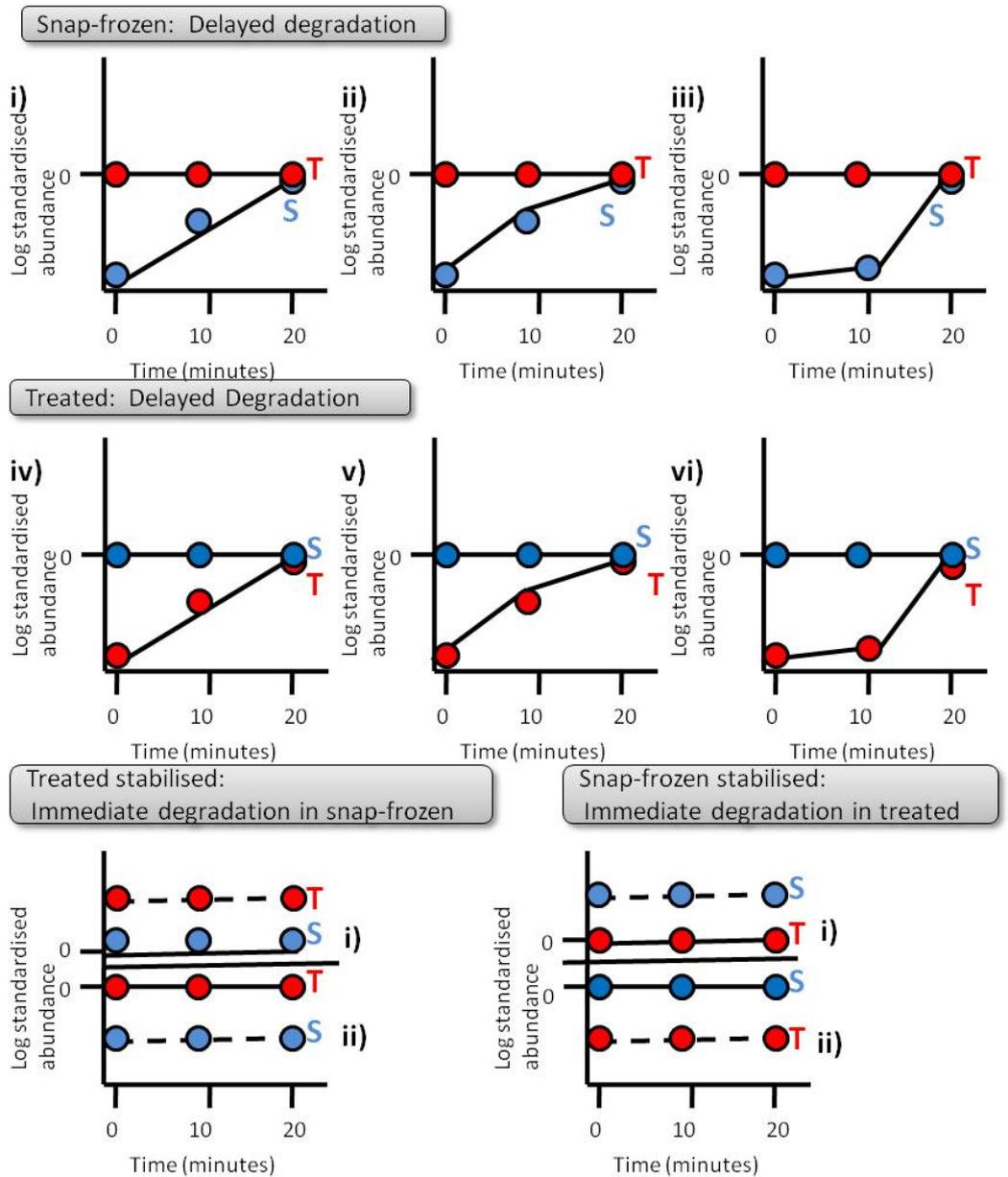


Figure 3-9: Predicted profiles on spot intensities. This figure displays the different possible profiles exhibited by spots on the DiGE gels. Showing delayed degradation in treated and snap-frozen and stabilisation in treated and snap-frozen with immediate degradation displayed respectively. Red=treated and Blue=snap frozen.

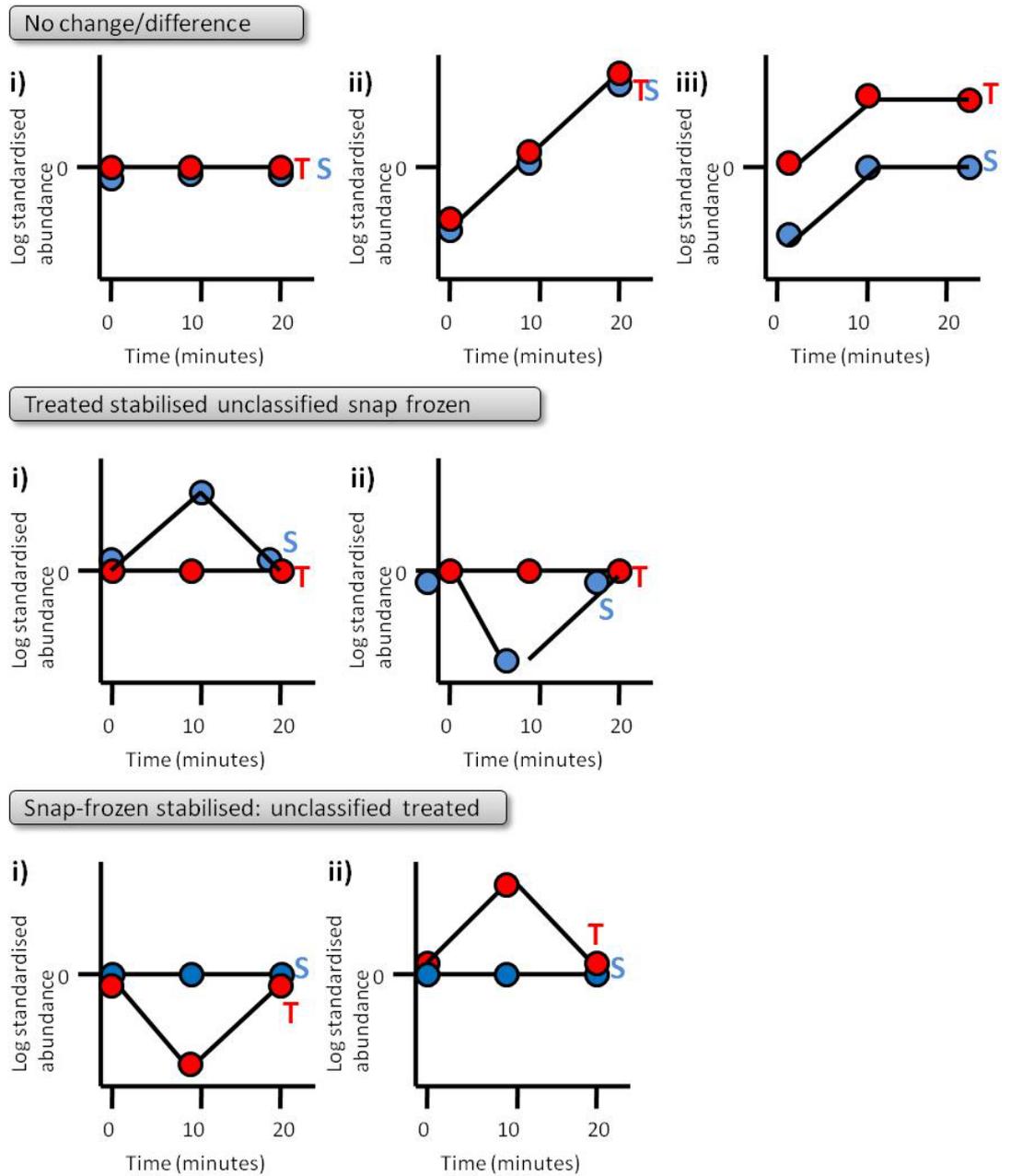


Figure 3-10: Predicted profiles on spot intensities. Showing profiles of no change and treated and snap-frozen stabilised with other respectively unclassified. Red=treated and Blue=snap frozen.

3.5.2.7 Venn analysis of the main investigation.

The Venn analysis was performed by compiling all t-test comparisons, including both significant score ($p \leq 0.05$) and non-significant score ($p > 0.05$) in order not to exclude spot data. Both 3 and 4 way Venn analysis was used in the main investigation to target specific profiles.

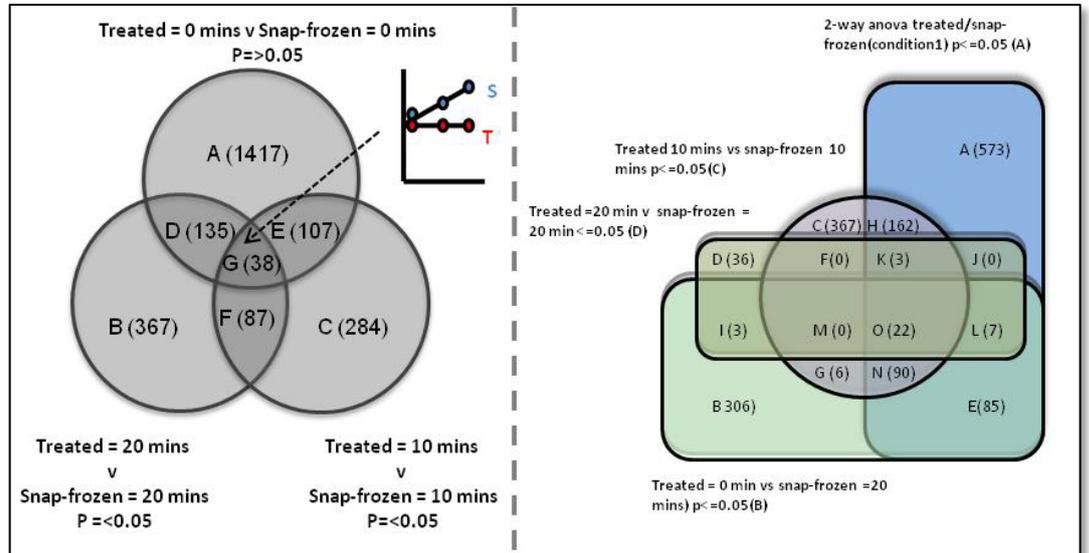


Figure 3-11: Example of three way and four Venn diagrams. The 3-way Venn here is an example of targeting treated stabilisation by using t-tests to focus on candidate spots. The 4-way Venn allows multiple profiles to be targeted simultaneously.

3.5.2.8 Predicted profiles for analysis of presence vs. absence spots in the pilot and main investigation.

As degradation occurs over time, the original protein spot intensity falls, whilst others increase. Expected profiles are given in Figure 3-12. In order to target such spots a method of combining group was used. T=0, T=10 and T=20 were combined and Student's t-test performed against S=0, S=10 and S=20. Data was then sorted using 2-way ANOVA with significance score of $p < 0.05$. In addition, this was cross referenced with those spots without an average ratio, to reveal statistically significant spots which appear in just the experimental group set (i.e. just snap frozen or treated).

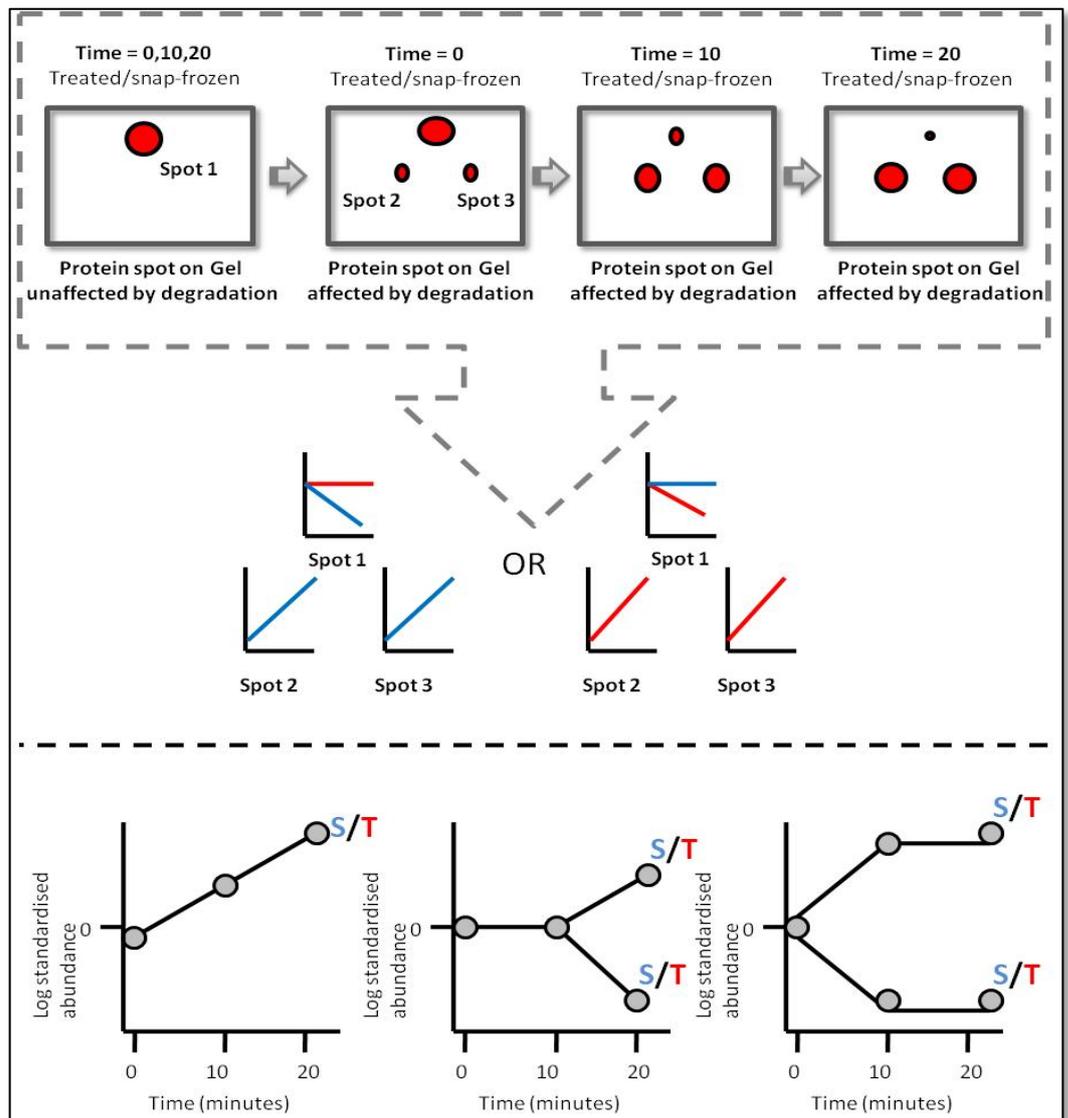


Figure 3-12: Shows the possible evolution of protein spots when degradation occurs. As degradation occurs over time the original protein spot intensity falls, whilst other increase. If this occurs solely in one experimental group expected profiles are found as above.

3.5.2.9 Gel locations of multiply identified marker spots for pilot investigation and main investigation.

Another expected characteristic of degradation, where the principal was demonstrated in section 3.5.2.2, is the movement of markers to different locations on the gel map. If the molecular weight or pI changes due to proteolytic activity this directly affects the position of the spot. Multiply identified proteins from different locations on a gel map would give an indication of possible degradation. This is demonstrated in a schematic in Figure 3-13.

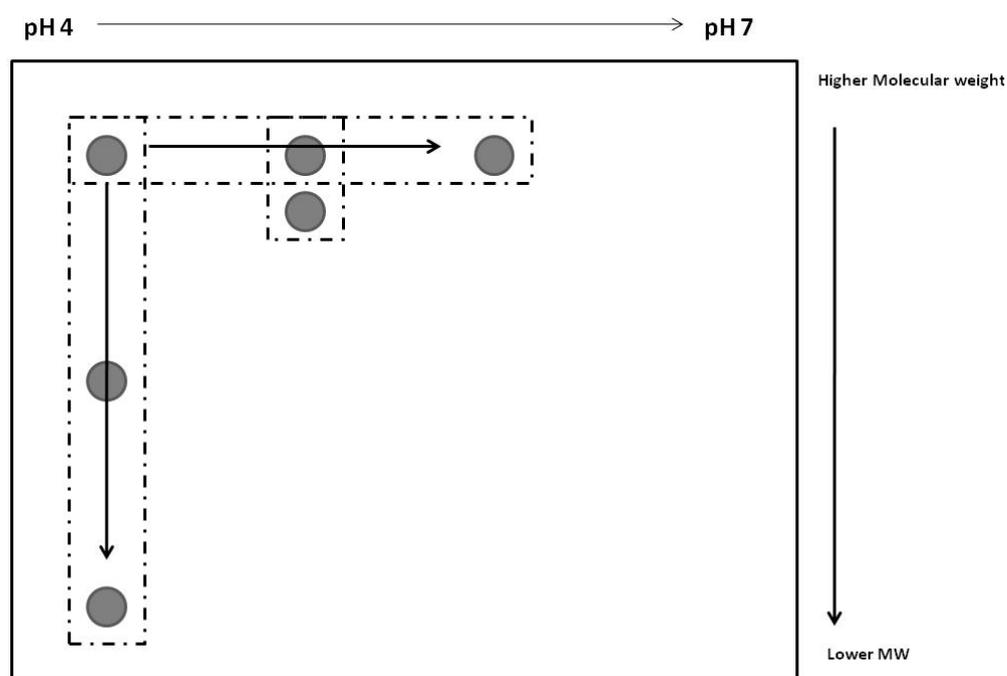


Figure 3-13: Schematic showing the effect of changes in molecular weight and pI may have on the location of protein/marker spots on a DiGE gel map. These changes may be due to proteolytic enzymes performing digestion leading to degradation as outlined in section 3.5.2.2.

Examples of multiply detected marker spots are given below in Figure 3-40, Figure 3-41, Figure 3-42, Figure 3-43, Figure 3-44, and Figure 3-45 for DiGE pilot investigation. Horizontal shifts denote a change in pI and therefore isoelectric point. This can be caused by degradative processes, particularly exoprotease but shifts in pI are not exclusive to degradation and likely to be PTMs. Vertical shifts in the gel are equally indicative of degradation and effectively denote a shift in molecular weight (directly affects size), large shifts would be symptomatic of endoprotease activity considerably reducing an amino acid chain. Circumspection is important when drawing conclusion, as many commonly identified proteins have many homologous sections of sequence and therefore identifications often are correct in broad classes but

would not always identify accurately the correct isoforms. Furthermore many of the examples shown are identified but differently expressed markers display a variety of intensity profiles and whereas in the predications in section 3.5.2.2 we have the luxury of following the parent protein, this is not necessarily the case in experimentally obtained spot data.

3.6 Results and Discussion

3.6.1 Introduction to Results

The use of DiGE to investigate non-specific and global proteomic degradation of mouse brain tissue samples treated with rapid heat-treatment is a novel approach. Although similar techniques have been described in the heat stabilisation of peptides and neuro-peptides (Svensson et al., 2009a) using the same system and the stabilisation using microwaves (Che, 2005). However, degradation research in general, perhaps due to its intrinsically non-specific and complex nature, is poorly reported. Despite being a novel approach, the techniques and technologies used are well established. DiGE has in excess of 1000 publications, with 2DE, a standard in proteomic techniques, having many more. In addition the collaborative work involving the use of other proteomic techniques (MSI and label-free quantitation) gives this study a more comprehensive approach to a multifarious problem. In a collaborative process, technological limitations are minimised allowing a larger dynamic range of protein masses to be included. The larger proteins abundances monitored using DiGE, affording information regarding quantitative changes probably resulting from proteolytic activity and PTMs. Whilst MSI has allowed the inclusion of lower protein masses (<30KDa) and maintained spatial resolution for at least the low mass range. In addition, label free-quantitation has allowed the gain of quantitation in addition to gaining MS level resolution and sensitivity without the complexity of labelling peptides. All three techniques provide an innate possibility for cross over validation that using a number of versatile approaches affords.

3.6.2 Profile Classification

The idea of scrutinising profiles in DiGE is not new. In DiGE analysis it is reasonably common to look for a pattern using the graphical display. However, the desired pattern is usually a positive or negative correlation between two points or a time course of treated sample which reviews a general trend. The use of profile classifications in this thesis is a novel approach and tries to utilise many of the possible permutations to make sense of what is happening regarding the degradation of proteins. As can be seen in section 3.4.5 on page 157 the use of profiles is complex and confusing. As such it is useful to consider the reasoning behind the selection of profiles at this junction before going on to discuss the findings of the pilot and main investigations.

The initial consideration came from how a change in a protein or peptide marker may affect its position on a 2D gel. As can be seen in Figure 3-13 there are two clear results of a change in a marker protein or peptide. Either a shift in the point of IEF, or a change in molecular weight (or a combination of the two, or indeed any fragments just run-off the end of the gel). This will lead to the numerous permutation outlined in Figure 3-8Figure 3-9Figure 3-10Figure 3-12. What makes this essentially more challenging is the possibility of spots disappearing and reappearing. Often in the DeCyder software, this will affect statistical significance as the algorithm does not notice the absence as a severe reduction in intensity, but as a spot of no importance (because it is not there) and considers it in the statistics. Therefore, looking for these present and absent spots in degradation becomes difficult once the spots have been sorted by ANOVA score alone. This problem remains somewhat unresolved.

It was decided, that for degradation studies, limiting profile classifications to positive or negative correlation was not a valid approach and would limit the usefulness of information that might be obtainable from a data rich DiGE experiment. To look at each profile class an increasing amount of information can be gained in order to make conclusions about how degradation in wild type mouse brain tissue may take place.

To consider the initial value and validity to this approach, the pilot study was engaged first.

3.6.3 Pilot Investigation

Firstly, a pilot investigation was performed in order to ascertain if heat-treating tissue lead to any noticeable change in the proteome in comparison to snap-freezing in liquid nitrogen as well as determining whether two time-points were required in the main investigation. Technical repeats were used in the pilot investigation to ascertain if DiGE was a suitable, valid and reliable method for investigation in tissue degradation.

3.6.3.1 Typical Gels: Pilot investigation.

Gels for the pilot investigation were run as described in methods using pH 4-7 IPG strips from Ettan™ (GE Healthcare, Bucks, UK) for isoelectric focusing. Cropping of gels was performed prior to importing into DeCyder software in order to minimise errors in the spot matching and allow the algorithm to function optimally. To gain maximum sensitivity, the algorithm was told to estimate 10,000 spots and filter out spot volume $\leq 29,999$, as dust falls into this range.

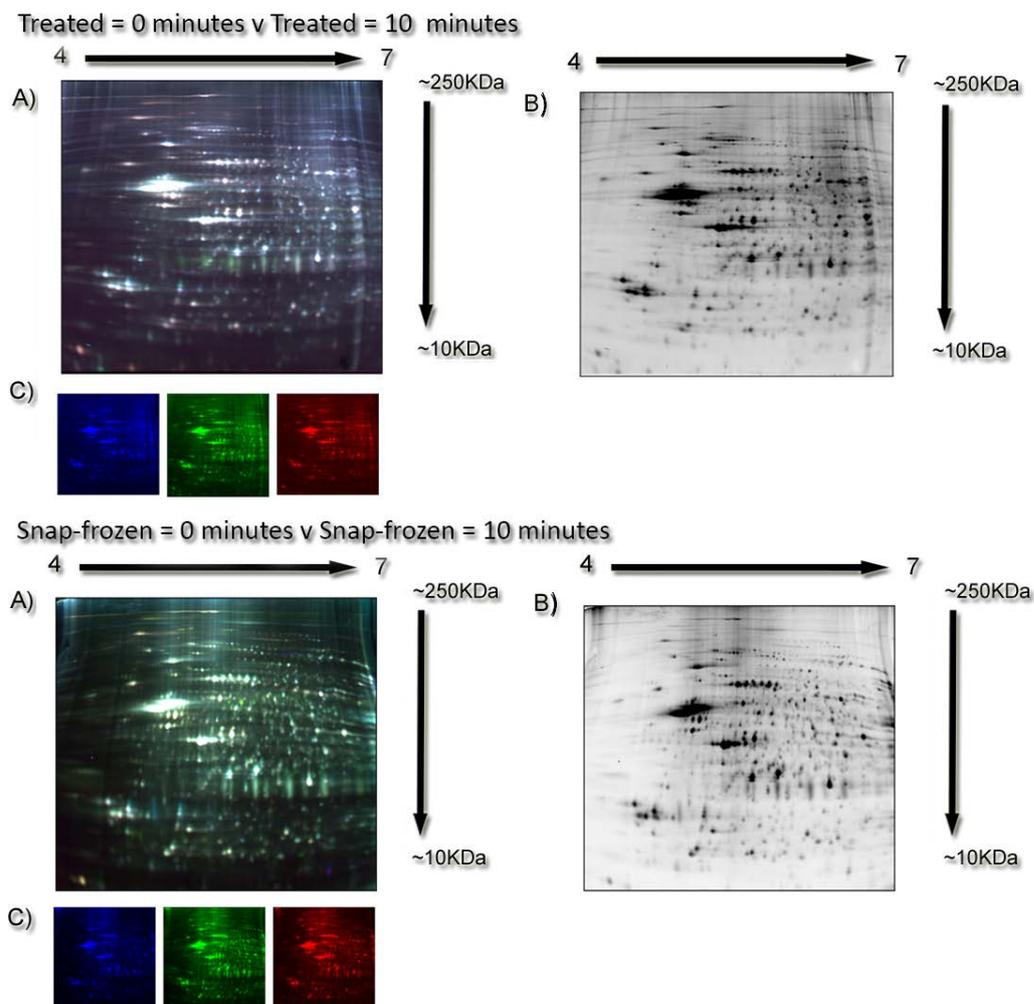


Figure 3-14: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels treated = 0, 10 and Snap-frozen = 0,10 respectively from pilot investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) 0 minutes and Cy5 (red) 10 minutes. Acidic to basic left to right.

Typical gels for the pilot investigation are presented in Figure 3-14. As can be seen, the gel maps show clear and well resolved protein spots with little to no smearing, particularly in the middle portion of the gel, indicative of using a pH4-7 strip. As expected, a small amount of precipitation and smear has occurred at the extreme fringes. Internal standard channels compare well to each other, and average protein intensity spots (relative value of 80,000) fall within 15% of each other between all channels and across gels allowing for accurate quantitation. Some small areas of the gel, corresponding to known areas of structural proteins, have saturated portions so quantitation will not be accurate for those regions. This is normal for DiGE gels as a compromise must be made between being able to visualise low intensity stop without losing more abundant spots.

The overlay of colours for the different channels allows for an initial view of the spread of change over the proteome. It can be seen in Figure 3-14 that

certain spots appear and disappear showing an initial difference between the different channels and shows the need for deeper analysis using DeCyder software to look for the intensity differences. The gels were then taken forward for quantitative analysis. The data obtainable from well-run DiGE gels, such as those above, is considerable, regarding the ability to multiplex and quantitate. About 3000 spots across all the gels are quantifiable, a significant number in a proteomic workflow.

3.6.3.2 Data set to be analysed

As with traditional analysed DiGE experiments, the data was sorted for profile analysis. For pilot investigation 5920 spots were detected and remained after filtering and profile analysis was performed on 1072 (18.11%) after 1-way ANOVA sorting on the basis of $p < 0.05$. A schematic is shown below in Figure 3-15.

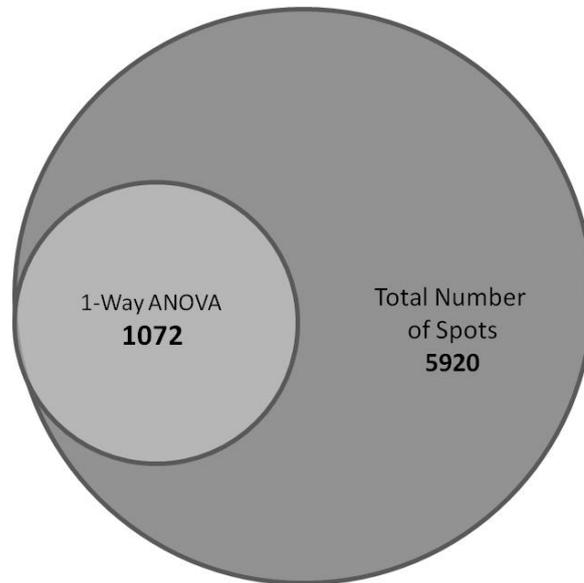


Figure 3-15: Showing total amount of spots detected in pilot investigation and total spots included in profile analysis using 1-way ANOVA as a sorting methods.

3.6.3.3 Statistical Results

3.6.3.3.1 Normality testing

Within DeCyder 2D version 7.0, GE Healthcare BVA module, there is no alternative to running parametric tests in the form of an ANOVA or Student's t-tests. Therefore, it is essential to show normality and homogeneity of variance, if false positives are to be minimised and useful conclusion drawn. Table 3-3 shows the percentage of protein spots that were out with the Shapiro-Wilk p-value of 0.05 or less. As can be seen, scores range from 7.10 - 2.52%, the average being 5.58%. Therefore it was found that the assumption of normality is true. This is because with a p value of 0.05 it would be expected that 5% of spots would fail due to random sampling alone, therefore a result of 5.58% is within an acceptable tolerance.

Dataset	Number of Proteins spots tested	Data type	Percentage spot significance score <0.05
Snap-frozen = 0	291	log ₁₀ SA *	6.19
Snap-frozen = 0	1637	log ₁₀ SA **	6.17
Snap-frozen = 0	1072	log ₁₀ SA ***	3.17
Snap-frozen = 10	291	log ₁₀ SA *	6.87
Snap-frozen = 10	1637	log ₁₀ SA **	7.09
Snap-frozen = 10	1072	log ₁₀ SA ***	4.20
Denator = 0	225	log ₁₀ SA *	6.67
Denator = 0	1296	log ₁₀ SA **	7.10
Denator = 0	1072	log ₁₀ SA ***	2.52
Denator = 10	225	log ₁₀ SA *	7.11
Denator = 11	1296	log ₁₀ SA **	5.86
Denator = 10	1072	log ₁₀ SA ***	4.01

*Spots included had log₁₀SA for all 4 repeats and an 1-way ANOVA score of <0.05 or better

**Spots included had log₁₀SA for all 4 repeats

*** Spots included had log₁₀SA for all 4 repeats and an 1-way ANOVA score of <0.05 or better with missing values replaced

Assessing normality of the log₁₀SA using the Shapiro-Wilk goodness-of-fit test

Table 3-3: Assessing normality of log₁₀SA using the statistical test Shapiro-Wilk, results for denator pilot study. These statistical test was performed with n=4

3.6.3.3.2 Homogeneity of Variance

The second assumption of parametric testing is that samples have an equal (or approximately equal) variance. DeCyder transforms standardised abundance values into $\log_{10}SA$ in order to stabilise variance and statistical analysis in DeCyder is performed only on the $\log_{10}SA$. A number of other methods are available for variance stabilisation and are employed in DNA micro array work and are beginning to be employed within data analysis in proteomics due to the issues discussed in section 1.7. In order to assess data homogeneity, i.e. how equal variances between groups are, the Levene's statistical test was employed using SPSS 17.0.1. The variance of data can be thought of in two ways either how homogenous the data is (that is how equal it is) or how heterogeneous the data is (how different it is). The Levene's statistic tests the null hypothesis that the population or sample variances are equal. This was performed across all groups with a full set of repeats. It was then subsequently performed with data with missing values replaced using k-nearest neighbour (KNN). Table 3-4 displays the results of the Levene's test. The threshold for rejecting the null hypothesis and concluding that the data is not homogeneous was 0.05 or less. Only 2.34% or 3.09% of the spots across all groups were not considered to have homogeneous variation. Therefore it was found that the assumption of homogeneity is true. This is because with a p-value of 0.05 it would be expected that 5% of spots would fail due to random sampling alone, therefore a result of 2.34% is within an acceptable tolerance. This also showed that the process of taking the logarithm to the base 10 of the standardised abundance considerably aided the stabilization of data by increasing the homogeneity of the data from 7.45-2.34% or 8.89-3.09%.

Data Set	Data type	Number of Spots included	% of spots with p value <0.05
Denator Pilot Investigation	SA	685	7.45
	log ₁₀ SA	685	2.34
Denator Pilot investigation with missing values replaced	SA	1072	8.96
	log ₁₀ SA	1072	3.09

Spots included had complete set of repeats

The P-Value was generated using Levene's test across groups with each master spot

Table 3-4: Homogeneity of variance Levene's statistical test results for denator pilot study. This statistical test was performed with n=4.

The variance was also visually assessed using a graphical representation. An example can be seen in Figure 3-16, although graphs have been generated for all time points and treatments. The average variance was calculated for each spot and plotted on the x-axis. It was plotted against arbitrarily points on the y-axis in order to differentiate points on the graph. This allowed the visualisation of how much the variance has spread.

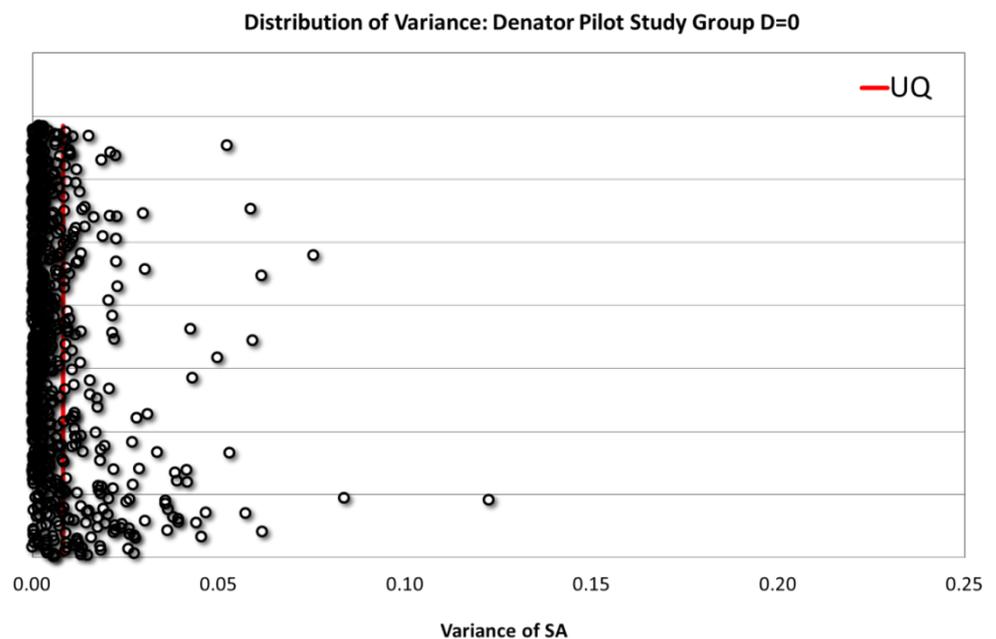


Figure 3-16: Example of the graphical representation of variance. For the denator group time point 0, for the standardised abundance using 683 different spots and n=4. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.

In addition to what is shown by using the Levene's test, the graphical representation shows tight clustering with the majority of data below the upper quartile close to zero. The outlier showing greater variance are thought to be a result of preferential labelling between Cy3 and Cy5 (Karp and Lilley, 2005). This helps to confirm the results seen using the Levene's test.

3.6.3.4 Profile analysis of pilot investigation and the distribution of global proteome.

After sorting on the basis of 1-way ANOVA, each of the 1072 significant spots was compared visually and manually against predicted profiles in Figure 3-5. This is essential, as a 1-way ANOVA is additionally complex compared to the universally used Student's t-test. Whereas the Student's t-test allows the user to search for differences between two different conditions, the 1-way ANOVA measures the difference between means in all experimental groups but not specifying which groups are different, making further classification necessary.

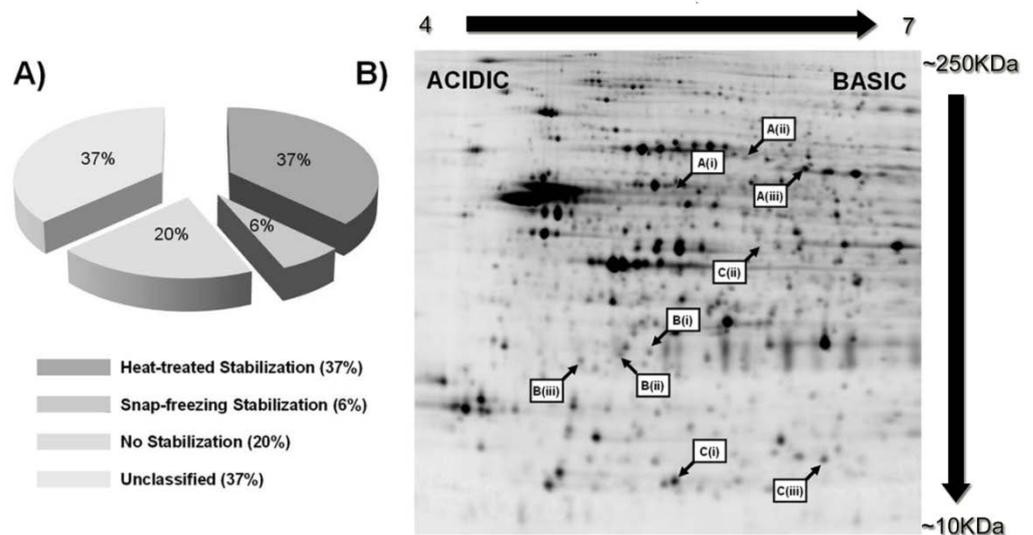


Figure 3-17: A) Shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 3-5 placed in the relevant categories following profile analysis of pilot investigation. A total of 1072 spots were included. B) Shows the positions of identified spots of examples given in Figure 3-18

Profiles were assigned to their relevant category and the distribution of spots can be seen in Figure 3-17 (A). The number of spots classified as treated stabilised (heat-treated) was over a third of the 1072 spots analysed at ~37%, with snap-frozen stabilisation accounting for just ~6%. A further ~20% of spots were classified as having profiles with no significant change with the remaining ~37% having intensity profiles showing very variable and unclassifiable profile behaviours within the representative profile groups. Thus from the initial profile analysis the heat-treatment appears to have a positive effect on protein degradation in comparison to snap-freezing on its own.

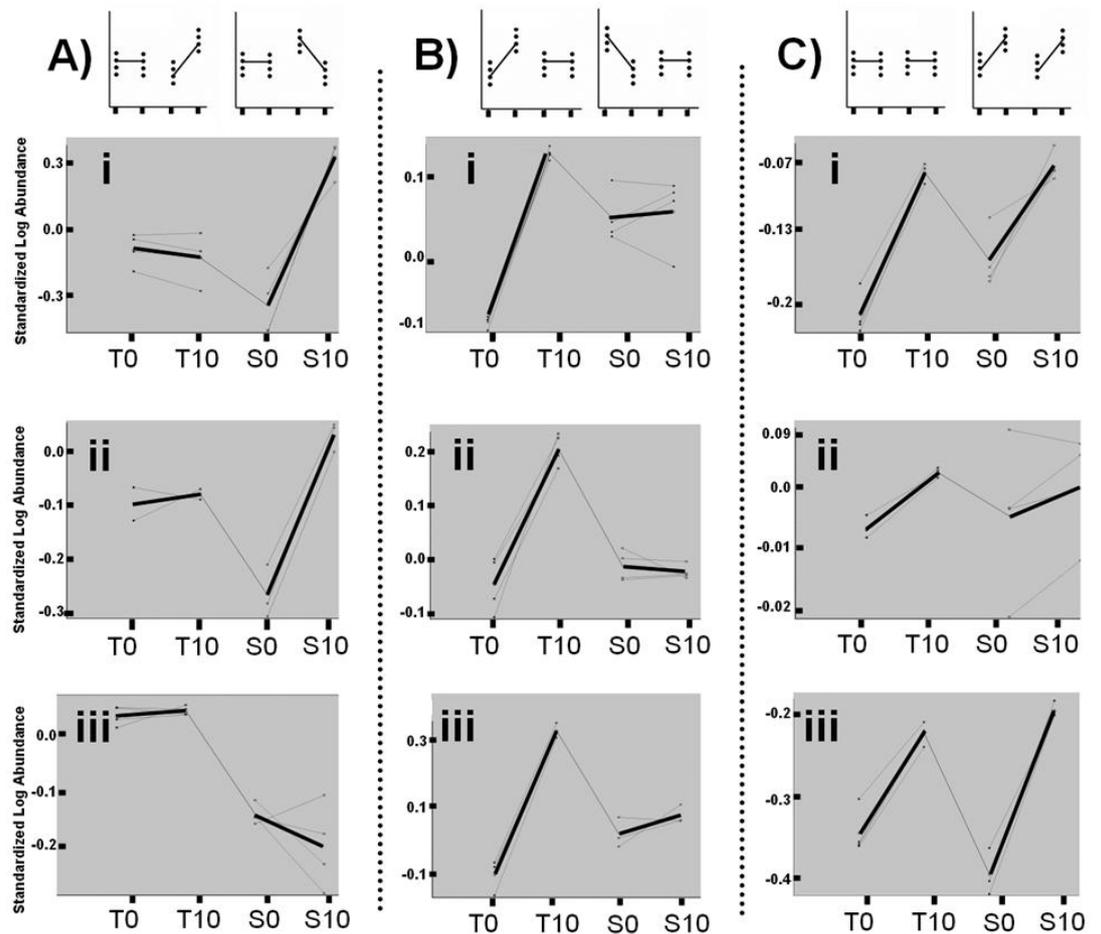


Figure 3-18: Example profiles for pilot investigation. Ai, ii and iii), Bi, ii and iii) and Ci, ii and iii) match predicted profiles given Figure 3-5. The positions on the spot map are given in Figure 3-17-B. All x-axis denoted time points (heat treated=0 min, 10 mins and snap-frozen=0 min, 10min respectively). A) Gives example profiles of heat-treated stabilised. B) Gives examples of snap-frozen stabilised and C) Gives examples of both stabilised or both changed.

A number of examples of experimentally obtained intensity profiles are given in Figure 3-18, with their position on the gel indicated in Figure 3-17.

Intensity profile Figure 3-18 (A) show examples of treated stabilisation, at position (Ai) and was identified as vacuolar H⁺ATPase 2B showing (t-test between snap-frozen=0 and 10 yielded a p-value=0.0002 and no significant p-value between T0 and T10, a required pattern for this profile to exhibit), Fig. A(ii), cardiotrophin like Cytokine factor 1 (t-test with p-value=0.0011 for S0 and S10 and no significant p-value between T0 and T10), A(iii) and T-complex protein, (t-test between snap-frozen=0 and 10 yielded a p-value=0.024 and no significant p-value between T0 and T10, it is also worth noting on this occasion there was a p value=7.8E-6 for treated = 0 snap-frozen =0 and there is not quite a significant t-tests of snap-frozen = 0 and 10 (p-value=0.064). This is a suspected case of rapid degradation, before the point of snap freezing but

not prior to heat treatment. In this last example the variance in intensity in the snap-frozen set is seen to be much greater than for the heat-treated samples, this kind of event is investigated more thoroughly in the main investigation.

The contrary result is shown in Figure 3-18 (Bi-iii), where snap-frozen exhibit stabilisation (this could equally, and importantly, be viewed as treatment affected). (Bii-iii) are unidentified; B (i) has been included as an example but warrants further explanation. It has been identified as Beta tubulin 2C, this is notable on a number of grounds. Firstly, this is an abundant and homologous set of proteins (i.e. the MS identification could easily be a number of different isoforms and validation would need to be performed to ascertain which) with structural function and therefore is notably in the saturated portion of the gel, so not too much attention is warranted regarding absolute quantitation but relatively the profile can be considered. Secondly, it is also notable that the black line in (Bi) denoting the average intensity masks the high level of variation less noticeable but snap-frozen for this profile has a substantial variation in comparison to treated. (Bii-iii) shows stabilisation with little variation between spots, with p-value = 1.6×10^{-8} , 0.00017 and 0.00015 respectively for t-test Treated = 0 and treated = 10 but showing no significant change between Snap-frozen = 0 and Snap-frozen = 10. These proteins are currently not identified. For intensity profiles in Figure 3-18 (Ci-iii), no change or difference is displayed, due to the treatment process, and the marker has not been affected over time. In (Ci and iii) marker intensities have increased in both experimental groups indicating that stabilisation has not occurred in either treated or snap-frozen samples, the slight difference in level of change in (Ci and iii) may indicate an effect of treatment. Peroxiredoxin 6 has been identified as (Ci) this is interesting as in Figure 3-41 Peroxiredoxin 6 was identified in another location exhibiting an intensity profile suggesting treated stabilisation, with substantial differences presented between snap-frozen = 0 and 10 Figure 3-41 (2). The fact that there are multiple locations suggests the possibility of degradation, and one profile matches. Circumspection is still necessary as Peroxiredoxin 6 is also homologous and is part of a set of proteins with many isoforms, meaning identification could be one of these variants.

C(ii), identified as lactate dehydrogenase B, at both time = 0 and time = 10 is greater in the snap-frozen processed samples indicates that the degradation is

not as predictable or consistent without heat-treatment, even though the averaged intensity profiles appear similar between the two treatment processes.

As a whole, the DiGE pilot investigation alludes towards that for significant spots (ANOVA score p value < 0.05) that were classified, heat-treatment does appear to beneficially stabilize the proteome (stabilising 37%) not affecting a further 37%, with little adverse effect exhibited in around 6%. A summary of the P-values obtained and Mass Spectrometry data are displayed below in Table 3-5 and Table 3-6.

3.6.3.5 Identifications of examples used in pilot investigation.

Below is a summary of the statistical and MS data obtained for the example markers discovered in the pilot investigation.

Spot master Number	Label on fig	Group	Protein name	t-test p value	1-ANOVA score	Average log standard abundance	Standard Deviation of log standard abundance	Number of experimental replicates	Degrees of Freedom
3149	Ai	T0	Vacuolar H+ATPase 2B	0.5	1.40E-05	-0.083	0.071	4	3
3149	Ai	T10				-0.130	0.104	4	3
3149	Ai	S0		0.0002		-0.330	0.136	4	3
3149	Ai	S10				0.293	0.078	4	3
2829	Aii	T0	Cardiotrophin like Cytokinase factor 1	0.78	0.00064	-0.140	0.150	4	3
2829	Aii	T10				-0.038	0.061	4	3
2829	Aii	S0		0.0011		-0.155	0.106	4	3
2829	Aii	S10				-0.158	0.139	4	3
2911	Aiii	T0	T-Complex Protein	0.96	5.50E-07	0.018	0.015	4	3
2911	Aiii	T10				0.015	0.006	4	3
2911	Aiii	S0		0.064		-0.155	0.015	4	3
2911	Aiii	S10				-0.235	0.070	4	3
4568	Bi	T0	Unidentified at this time	1.60E-08	2.40E-06	-0.163	0.013	4	3
4568	Bi	T10				0.178	0.013	4	3
4568	Bi	S0		0.8		0.048	0.050	4	3
4568	Bi	S10				0.060	0.074	4	3
4580	Bii	T0	Unidentified at this time	0.00017	4.80E-06	-0.038	0.052	4	3
4580	Bii	T10				0.213	0.028	4	3
4580	Bii	S0		0.06		-0.003	0.028	4	3
4580	Bii	S10				-0.040	0.008	4	3
4564	Biii	T0	Unidentified at this time	0.00015	7.90E-07	-0.130	0.013	4	3
4564	Biii	T10				0.225	0.013	4	3
4564	Biii	S0		0.6		-0.043	0.050	4	3
4564	Biii	S10				0.148	0.074	4	3
5430	Ci	T0	Peroxioredoxin 6	0.024	4.00E-06	-0.213	0.022	4	3
5430	Ci	S0				-0.158	0.026	4	3
5430	Ci	T10		0.21		-0.080	0.008	4	3
5430	Ci	S10				-0.093	0.013	4	3
3703	Cii	T0	L-Lactate dehydrogenase B	0.76	0.48	-0.068	0.013	4	3
3703	Cii	S0				-0.050	0.123	4	3
3703	Cii	T10		0.39		0.013	0.005	4	3
3703	Cii	S10				-0.025	0.087	4	3
5290	Ciii	T0	Unidentified at this time	0.06	7.40E-07	-0.230	0.123	4	3
5290	Ciii	S0				-0.230	0.215	4	3
5290	Ciii	T10		0.026		-0.125	0.091	4	3
5290	Ciii	S10				-0.195	0.078	4	3

Table 3-5: Table of p-values and standard deviations obtained for identified spots shown in Figure 3-17. The p-values given are from the Student's t-tests between the shown time points. This confirms the validity of the profile analysis by showing no significant differences between the treated = 0 and 10 minutes but showing difference between the snap-frozen = 0 and 10 minutes

Spot master Number	Label on fig	Group	Protein name	MOWSE Score	Score MS (PMF)	Machine identified on	P Value	Peptides MS/MS	Charge State	Peptides for PMF	Percentage coverage
1186	Ai	T0	Tpr [Mus musculus]	N/A	63	4700	0.05	N/A	1+	26	83
1186	Ai	T10									
1186	Ai	T20									
1186	Ai	S0									
1186	Ai	S10									
1186	Ai	S20									
1286	Aii	T0	nebulin-related anchoring protein isoform C [Mus musculus]	78	N/A	q star	0.01	4	2+	N/A	72
1286	Aii	T10									
1286	Aii	T20									
1286	Aii	S0									
1286	Aii	S10									
1286	Aii	S20									
1554	Bi	T0	Unidentified at this time	N/A							
1554	Bi	T10									
1554	Bi	T20									
1554	Bi	S0									
1554	Bi	S10									
1554	Bi	S20									
3068	Ci	T0	Unidentified at this time	N/A							
3068	Ci	T10									
3068	Ci	T20									
3068	Ci	S0									
3068	Ci	S10									
3068	Ci	S20									
1852	Cii	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	43	N/A	q star	0.05	1	2+	N/A	1
1852	Cii	T10									
1852	Cii	T20									
1852	Cii	S0									
1852	Cii	S10									
1852	Cii	S20									

Table 3-6: Table of identifications using Mass Spectrometry shown in Figure 3-17. Table gives details of the returned MS identification, p-value and percentage coverage of the identifications

3.6.3.6 Venn analysis for pilot investigation.

Venn analysis allows a more holistic approach, allowing the inclusion of a greater number of spots and a more targeted and specific approach than sorting on the basis of 1-way ANOVA as all spots are included.

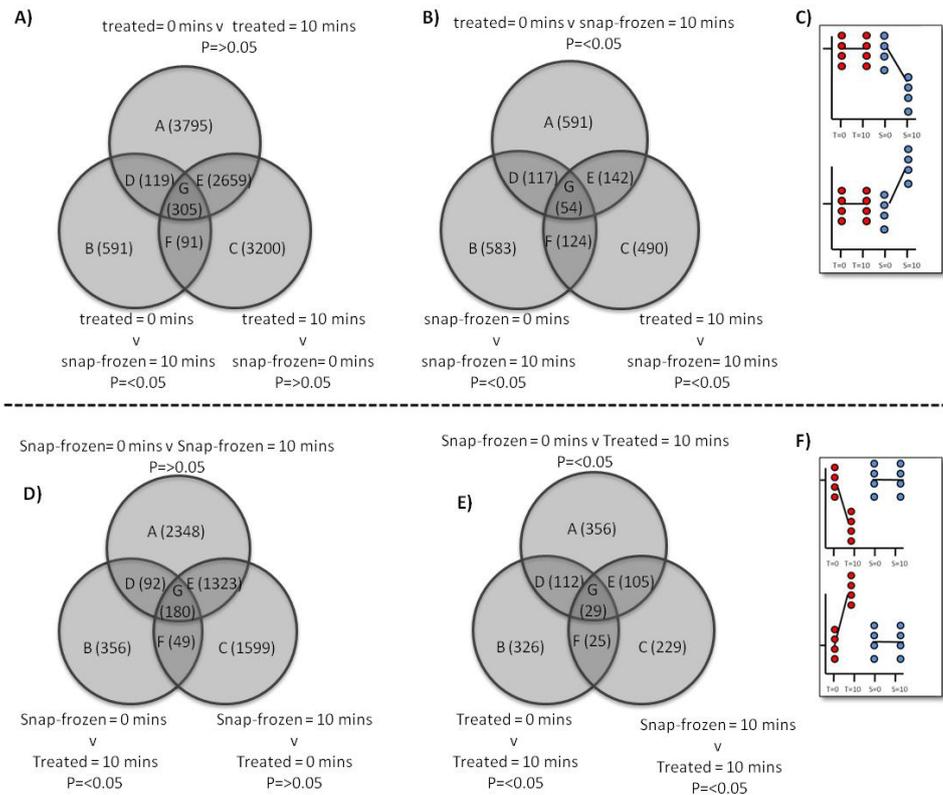


Figure 3-19: Venn analysis for pilot investigation: A) and B) give two independent but comparative strategies for matching profiles in c) searching for predicted profiles in sector G focusing on treated stabilisation. D) and E) give two independent but comparative strategies for matching profiles in F) searching for predicted profiles in sector G focusing on snap-frozen stabilisation. Profile shown in C) and F) are those shown in Figure 3-5. Red spots = heat treated and blue spots = snap-frozen.

In Figure 3-19 (A and B) show two different possible strategies for targeting treated stabilised profiles. (A) Shows an inclusive strategy. If a t-test p-value is ≥ 0.05 (not significant) between treated = 0 and 10, then this denotes no difference which is expected if stabilisation has taken place, for those spots that cross over between this group and treated = 10 and snap-frozen 0 with p-value ≥ 0.05 , and a significant p-value for snap-frozen = 0 and treated = 10 (i.e. significant difference) sector G should rule out intensity profiles in the no changes category while maintaining profiles showing differences between treated and snap-frozen with stability in the treated. In (B) a less inclusive strategy is employed, displayed by the reduction in spots considered, look for significant differences ($p \leq 0.05$) in the different experiment groups. In (A) this has allowed more focused attention of the data from 1072 (20% of total

number of spots) spots considered and manually checked using 1-way ANOVA to sort compared to 4412 spots narrowed to 305 of interest (6.9% of included spots and 5.15% of total spots to be manually checked), the manual work has been severally reduced whilst including more spots in total. In (B) using the more direct strategy a total number of spots included was 1227 narrowing to 54 for manual consideration (4.4% of included spots and 1% of total spots). This method still allows a greater number of included spots but limits manual checking to 54. For (D) and (E) which are targeting snap-frozen stabilisation (F) (or treated affected), 1830 and 540 total spots respectively, showing less are considered with respect to treated showing less spots meet the designated criteria for inclusion.

Comparing (A and B) and (C and D) we can see that the ratio of treated stabilised is $305/180 = 1.69$ and $54/29 = 1.86$ respectively. Both ratios are comparable and treatment appears to have a positive impact of the global degradative proteome. Venn analysis is shown in Figure 3-32 to Figure 3-27 to be successful in identifying targeted groups. In the pilot investigation shown in (A) and (B), this allows the user to reduce the amount of manual checking and lead therefore to more attention being given to particular spots while also maintaining a perspective on the overall global degradative proteome.

This Venn analysis allows a more inclusive approach and allows for a more objective angle using statistics, rather than a vast amount of manual checking. The strategy employed in Venn diagram A) is more likely to show false spots as it rules spots in on the basis of no difference. This means there may be no difference because there is nothing present in both tests. However, using the Venn process allows these to be ruled out and therefore reduces the complexity of the analysis from over 3000 down to about 300 spots which is more manageable manually. Then of course the spot can be checked and validated using profiles analysis.

Finally, it can be seen from this analysis that treatment seems to be twice as effective at stabilising degradation compared to snap-freezing alone.

3.6.3.7 Summary and conclusions of the pilot investigation

Initial evidence from the pilot investigation results suggest that heat-treating wild type mouse brain tissue samples immediately post excision has a positive effect in halting nonspecific proteomic degradation of protein or peptides markers; in comparison to traditional snap freezing alone. This has the prospect of increasing validity, particularly to quantitative studies. The possible inactivation of proteolytic enzymes could allow for the increased potential handling of tissue at room temperature without affecting proteomic profiles. This is particularly important where procedures such as dissection or preparation of tissue are required. The possibility of detrimental effects was not covered in the pilot study but will be further reviewed in the main investigation and validation sections. One barrier to the use of heat-treatment would be trying to compare the considerable wealth of data already obtained from tissue which has been snap-frozen. Any further investigation would need to take comparative differences into account.

Of course, having two time-points has proved limiting and does not allow for any conclusions to be drawn regarding the evolution of a marker. This is to be considered further. This is important as degradative processes may be delayed as shown in section 3.6.6.1 on page 237. Therefore it is useful to extend the time period for warming to tract how markers can change over the course of an extended time period. This is considered in the main investigation

3.6.4 Main Investigation

From the pilot study it was clear (see Figure 3-17 (A)) that the heat-treatment had changed the profile of the proteome in comparison to snap-freezing alone and had led to an increased number of stabilised proteins or peptides. It was also apparent that in order to see if the proteome had reached equilibrium or whether it would continue to change, another time-point would be advantageous. This was further considered in the main investigation.

3.6.4.1 Typical Gel in the Main investigation.

Gels for the main investigation were run as described in methods using pH 4-7 IPG strips from Ettan™ (GE Healthcare, Bucks, UK) for isoelectric focussing. Cropping of gels was performed prior to importing into DeCyder software in order to minimise errors in the spot matching. The algorithm was told to estimate 10,000 spots and filter out spot volume $\leq 29,999$, as dust falls into this range.

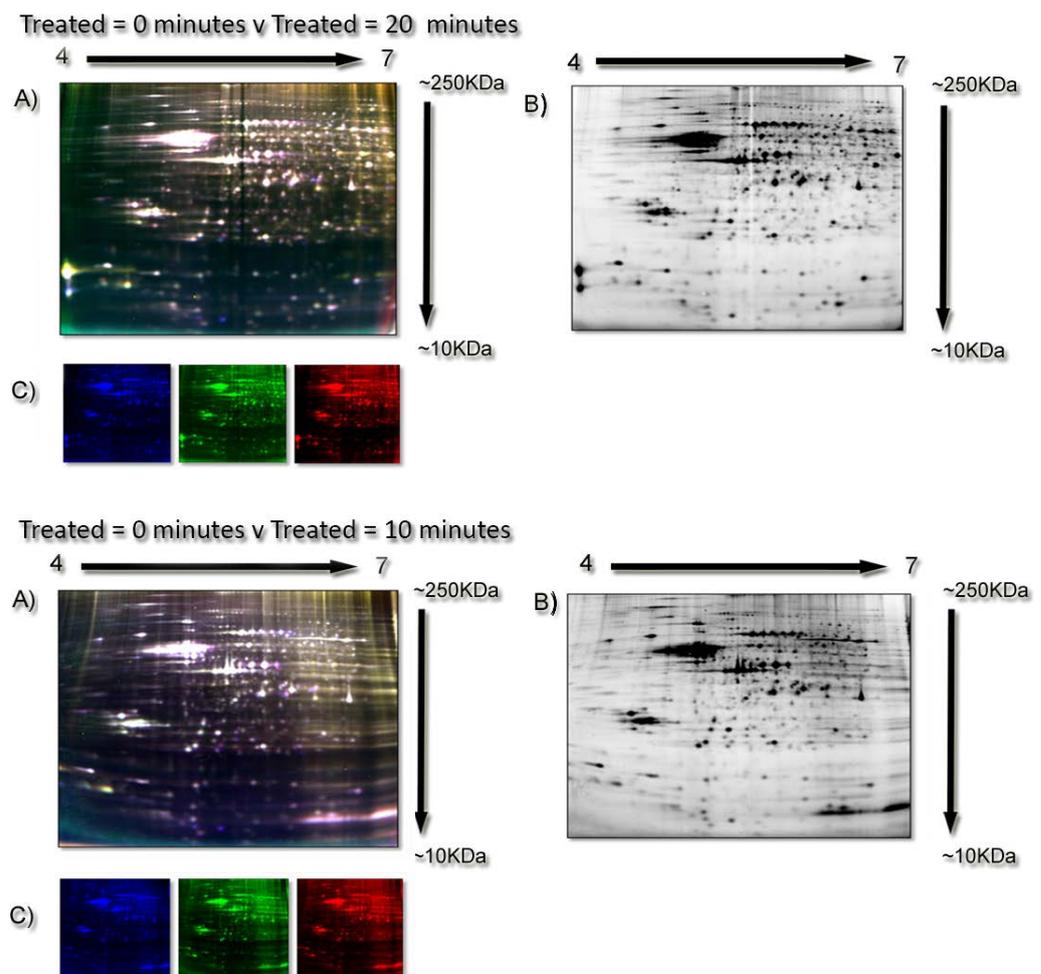


Figure 3-20: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels treated = 0, 20 and treated = 0,10 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) 0 minutes and Cy5 (red) 10 and 20 minutes respectively. Acidic to basic left to right.

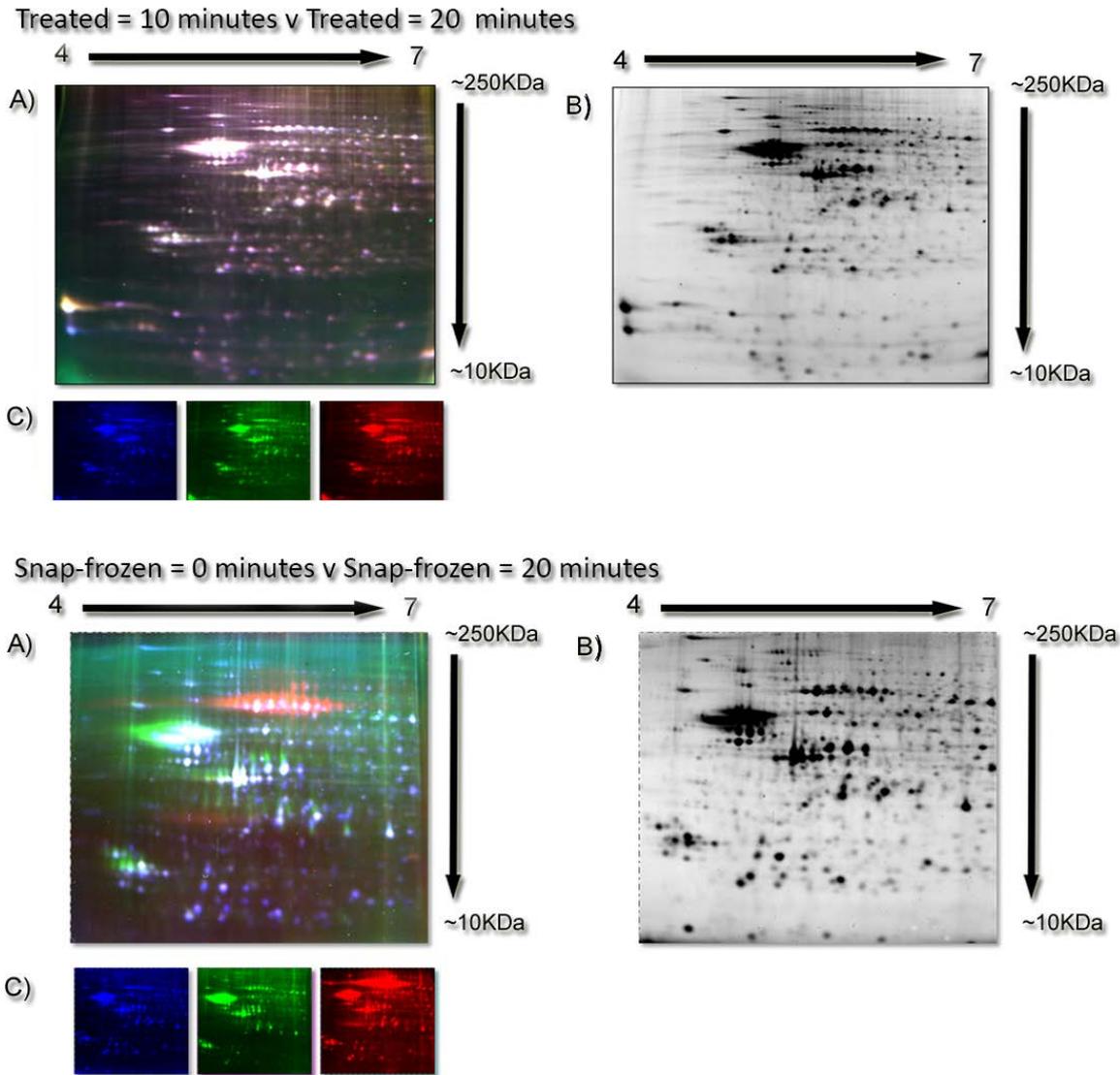


Figure 3-21: Set of typical DiGE gels showing A) Fluorescent image with Cy2, 3 and 5 for gels treated = 10, 20 and snap-frozen= 0, 20 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) treated = 10 minutes and snap-frozen = 0 minutes respectively and Cy5 (red) treated and snap-frozen 20 minutes respectively. Acidic to basic left to right.

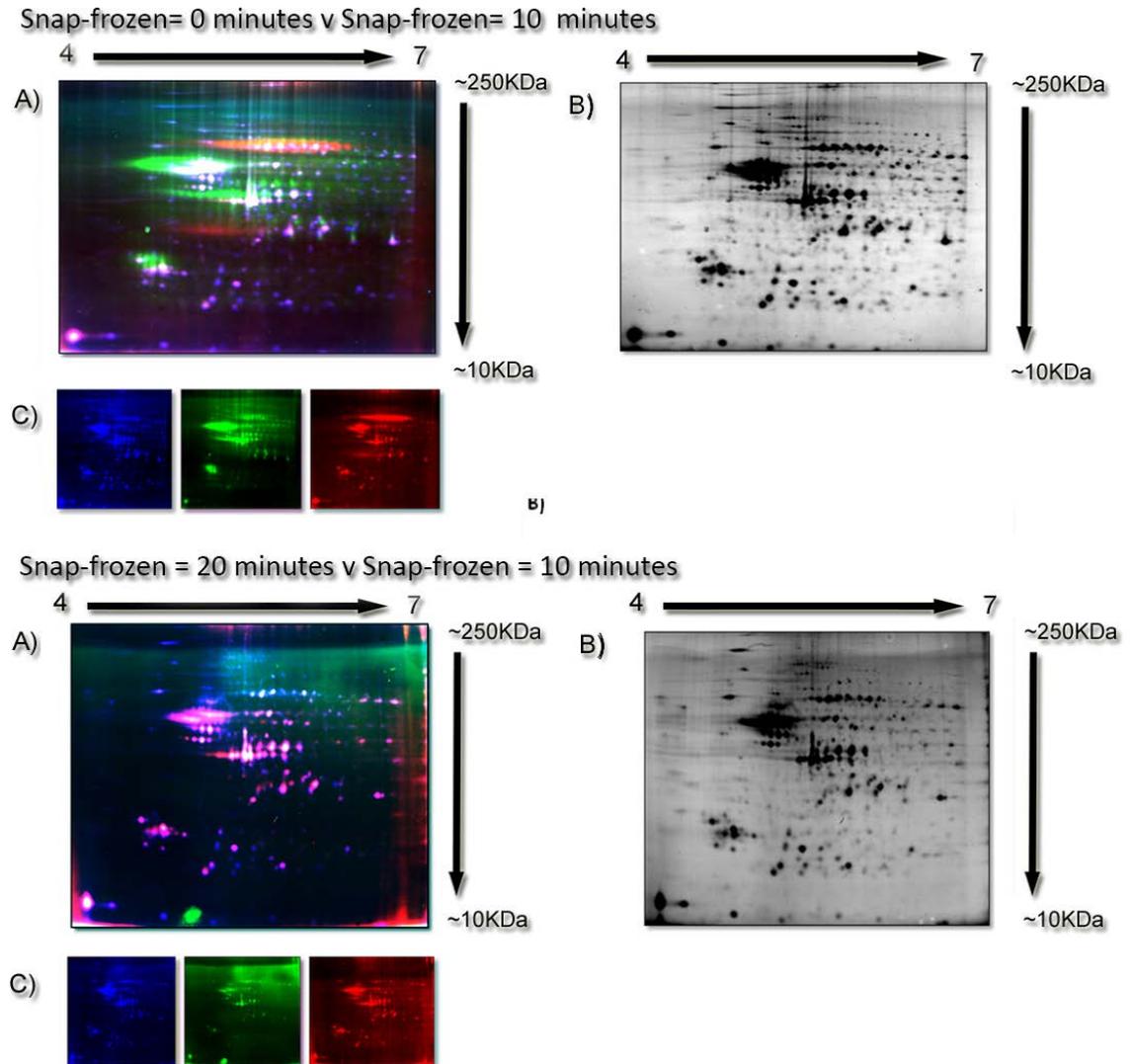


Figure 3-22: Set of typical DiGE gels showing A) Fluorescent image with Cy2,3 and 5 for gels snap-frozen = 0, 10 and snap-frozen= 20,10 respectively from the main investigation, B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) snap-frozen = 0 and 20 minutes respectively and Cy5 (red) treated and snap-frozen 10 minutes respectively. Acidic to basic left to right.

Typical gels for the main investigation are presented in Figure 3-20, Figure 3-21 and Figure 3-22. As can be seen, the gel maps have generally ran well showing clear and well resolved protein spots with, little to no smearing particularly in the treated samples and in the middle portion of the gel. Internal standard channels compare well across gel maps. The gels with snap-frozen samples are acceptable, but lack the same degree of resolution found in the treated gels. As before the average protein intensity spots (relative value of 80,000) fall within 15% of each other between all channels and across gels allowing for accurate quantitation and care must be taken when considering saturated areas.

3.6.4.2 Data Set considered for the main investigation

Following the schematic in Figure 3-23, data was sorted for profile analysis using both 1-way ANOVA (including 623 spots (20.0%)) and 2-way ANOVA (including 573 spots (18.4%)). The total number of detected spots was lower than in pilot investigation at 3110, but the proportion of spots with a significant ANOVA score is similar. This may be accounted for by the increased number of replicate gels for the main investigation and the slightly reduced resolution in some of the snap-frozen gel maps.

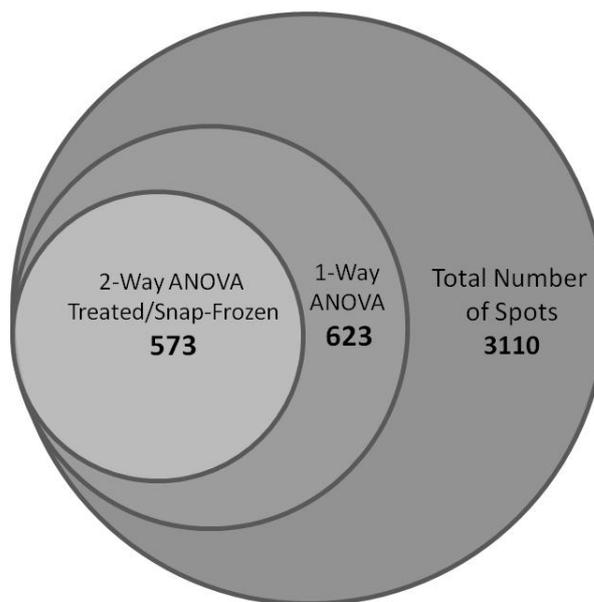


Figure 3-23: Showing total amount of spots detected in the main investigation and Total spots included in profile analysis using 1-way and 2-way ANOVA as a sorting method

3.6.4.3 Statistical results

3.6.4.3.1 Normality testing

As described in the pilot investigation, normality testing was performed and Table 3-7 shows the percentage of protein spots that were out with the Shapiro-Wilk p-value of 0.05 or less. As can be seen, scores range from 7.87 - 2.09%, the average being 4.40%. Therefore it was found that the assumption of normality is true.

Dataset	Number of Proteins spots tested	Data type	Percentage spot significance score <0.05
Snap-frozen = 0	151	log ₁₀ SA *	3.31
Snap-frozen = 0	556	log ₁₀ SA **	3.42
Snap-frozen = 0	623	log ₁₀ SA ***	2.25
Snap-frozen = 10	89	log ₁₀ SA *	7.87
Snap-frozen = 10	320	log ₁₀ SA **	6.56
Snap-frozen = 10	623	log ₁₀ SA ***	5.46
Snap-frozen = 20	94	log ₁₀ SA *	6.38
Snap-frozen = 20	322	log ₁₀ SA **	4.04
Snap-frozen = 20	623	log ₁₀ SA ***	4.01
Denator = 0	94	log ₁₀ SA *	4.26
Denator = 0	363	log ₁₀ SA **	3.86
Denator = 0	623	log ₁₀ SA ***	3.37
Denator = 10	73	log ₁₀ SA *	6.85
Denator = 10	283	log ₁₀ SA **	4.59
Denator = 10	623	log ₁₀ SA ***	2.09
Denator = 20	103	log ₁₀ SA *	4.85
Denator = 20	392	log ₁₀ SA **	3.32
Denator = 20	623	log ₁₀ SA ***	2.73

*Spots included had log₁₀SA for all 6 repeats and an 1-way ANOVA score of <0.05 or better

**Spots included had log₁₀SA for all 6 repeats

*** Spots included had log₁₀SA for all 6 repeats and an 1-way ANOVA score of <0.05 or better with missing values replaced

Assessing normality of the log₁₀SA using the Shapiro-Wilk goodness-of-fit test

Table 3-7: Assessing normality of log₁₀SA using the statistical test Shapiro-Wilk, results for the denator main study.

3.6.4.3.2 Homogeneity of Variance

As described in the pilot investigation, the Levene’s test was used to assess the homogeneity of variance. This was performed across all groups with a full set of repeats. It was then subsequently performed on all matched spot data with missing values replaced using k-nearest neighbour (KNN). Table 3-8 displays the results of the Levene’s test. Only 4.78% or 4.98% respectively of the spots across all groups were not considered to have homogeneous variation. Therefore it was found that the assumption of homogeneity is true. This also showed that the process of taking the logarithm to the base 10 of the standardised abundance considerably aided the stabilization of data by increasing the homogeneity of the data from 14.52-4.78% or 18.94-4.98%. Example graphs visually depicting the distribution of variance can be seen in Figure 3-24, as also described in the pilot investigation. This allowed the visualisation of how much the variance has spread.

Data Set	Data type	Number of Spots included	% of spots with p value <0.05
Denator Main Investigation	SA	62	14.52
	log ₁₀ SA	62	4.78
Denator Main Investigation with missing values replaced	SA	623	18.94
	log ₁₀ SA	623	4.98

Spots included had complete set of repeats

The P-Value was generated using Levene’s test across groups with each master spot

Table 3-8: Homogeneity of variance Levene’s statistical test results, for denator main study.

Distribution of Variance: Denator Pilot Study Group D=0

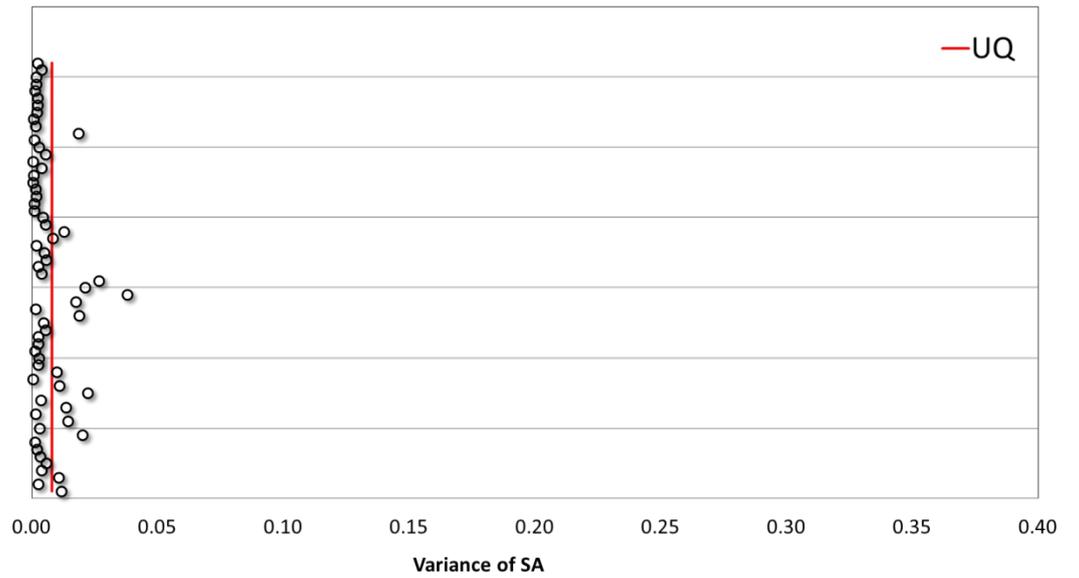


Figure 3-24: Graphical representation of variance for the denator group time point 0, for the standardised abundance using 62 different spots and n=6. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance is clustered close to zero with few outliers.

3.6.4.4 Profile analysis of the main investigation

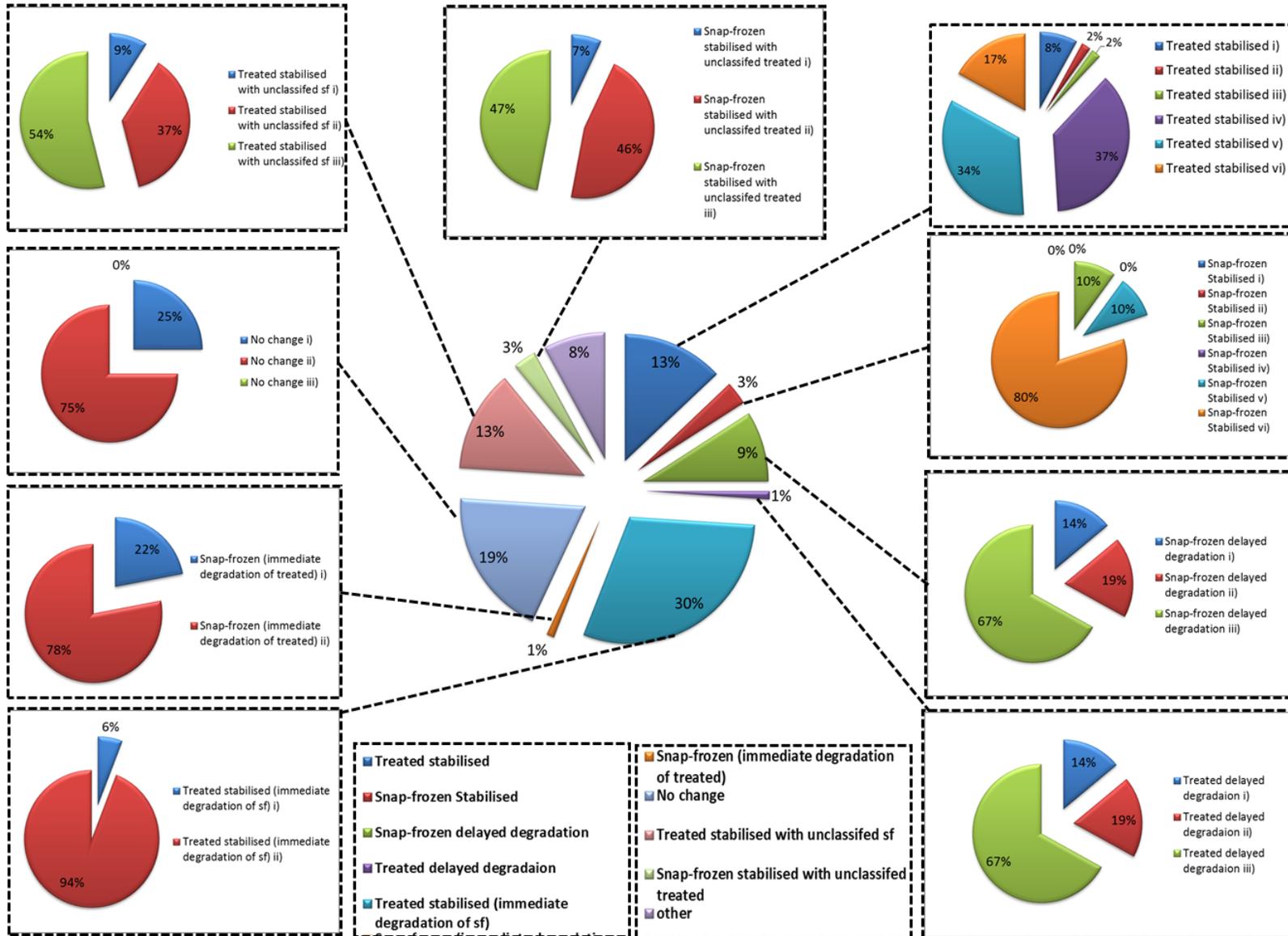
In completion of analysis in pilot investigation, limitations were apparent in using just two time point of 0 and 10 minutes of warming shown in section 3.6.6.1 on page 237. To visualise the evolution of markers an extended time course was performed.

The profile analysis was performed similarly to pilot investigation, but there were some noticeable differences. The categories of the predicted profiles were extended for completeness (as shown in Figure 3-8 to Figure 3-10). This was necessary as the addition of an extra time point considerably increased the permutations of possible profile, thus increasing the amount of data obtainable. The data was also assigned conditions allowing the clearer visualisation of changes using profile and the use of 2-way ANOVA sorting as well as 1-way ANOVA sorting. Having an extra time point increases the amount of Venn analysis that could be considered.

3.6.4.5 One-way ANOVA as sorting method.

The data was first sorted as in pilot investigation by 1-way ANOVA and revealed a subset of 623/3110 (20%, the same proportion as in pilot investigation), which showed a significant difference.

A)



B)

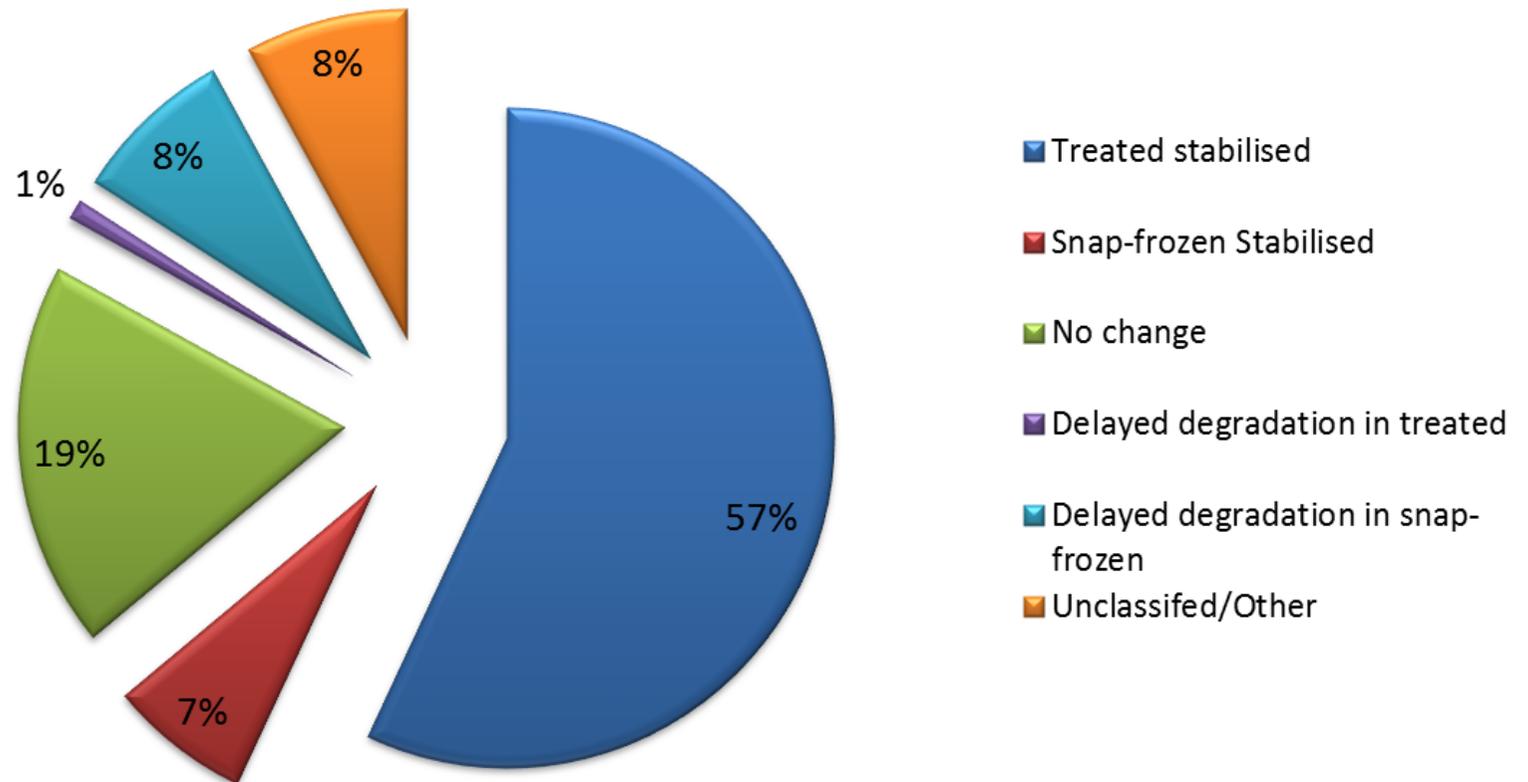


Figure 3-25: A) Profile distribution for the main investigation using 1-way ANOVA as a sorting method of categories shown in Figure 3-8, Figure 3-9 and Figure 3-10. A) Shows the proportions of various experimental group categories considered in the predicted profiles then further broken down into proportions assigned to individual profiles. B) The proportion of all stabilised groups compiled for both treated and snap-frozen. This gives an overview of relevant category amalgamation. As can be seen, the treated stabilised proportion is the greatest at 57% while the snap-frozen stabilisation is only 7%. 623 spot were included in total.

In the pilot investigation in Figure 3-17, the proportion shown to be treated stabilized was ~37%, this amount has increased to ~57% in Figure 3-25(B), in actual number terms (398 spots to 355 in pilot investigation and main investigation respectively) the spots are comparable between the two experiments. For snap-frozen stabilized (treated affected), proportions are similar at 6% and 7% but actual numbers have therefore increased. In Figure 3-8 in treated stabilized the highest proportions were seen in profiles (iv, v, vi) making up about 88%, showing the predicted appearance or accumulation of a degradation product in snap-frozen tissues. 30%/57% of treated stabilization was exhibited in Figure 3-9, showing the predicted profiles of immediate degradation in snap-frozen while stabilization in treated. This alludes to degradation being rapid after animal sacrifice but in excess of 60 seconds or occurring during storage at -80°C, this is in stark contrast in comparison to immediate degradation in treated tissues making up only 1% of included spots. With snap-frozen stabilisation (treated affected) accounting for 7% in total. In the main investigation, it appears treated samples have a lower proportion of spots delayed from degradation (1%) in comparison to snap-frozen (8%), alluding to treatment leading to stabilisation totally or not at all but having very small effect on delaying degradation; in contrast snap-freezing was more successful at this.

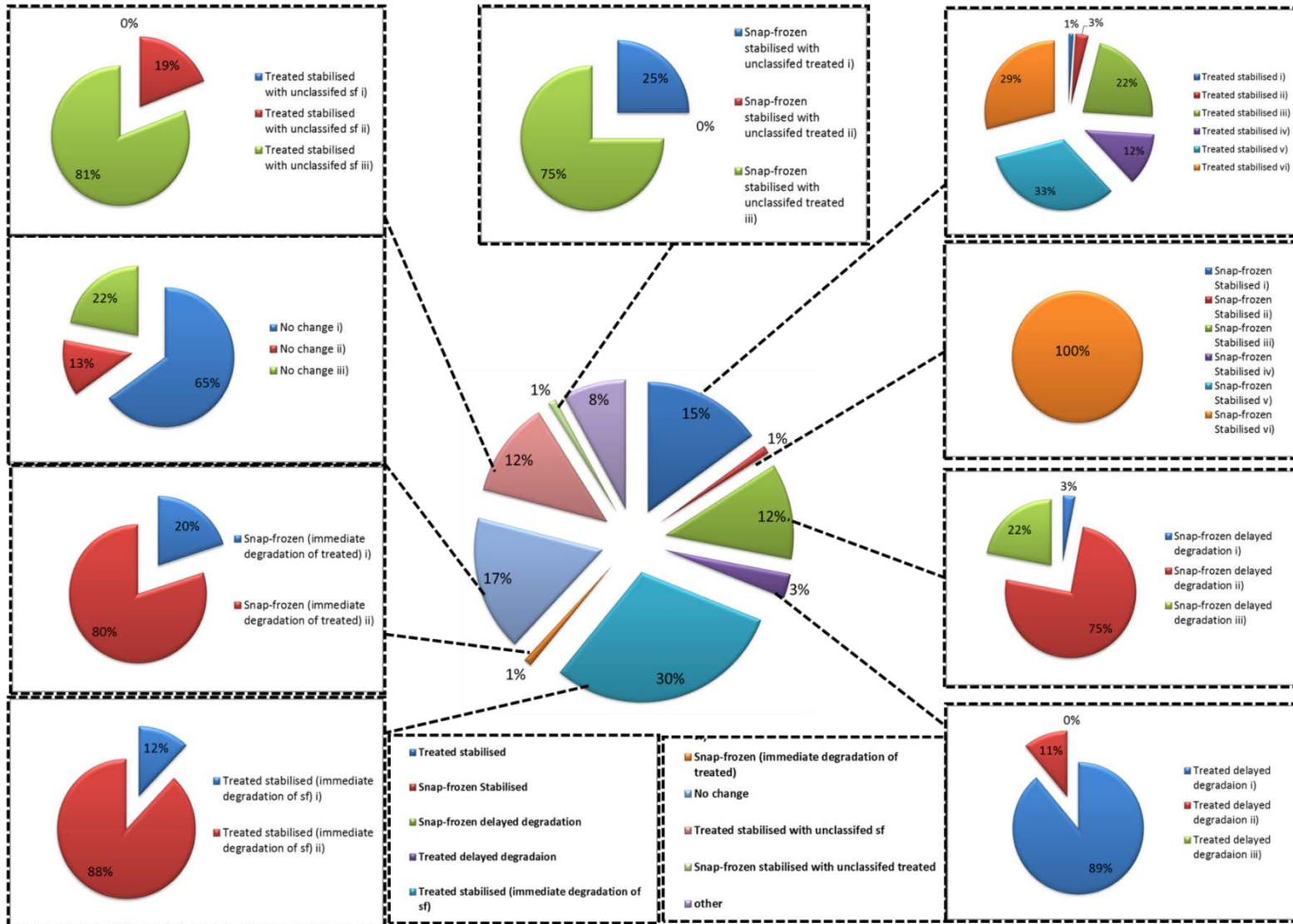
What has equally become apparent is, by running the additional time point delayed degradation could be visualised by viewing the profiles in parallel rather than serial, as in pilot investigation shown in section 3.6.6.1 on page 237. Therefore, a notable caveat is that profiles in pilot investigation being viewed serially, with one time point less, would be seen as treated stabilized (when it could possibly be snap-frozen delayed degradation) and vice versa, skewing results in pilot investigation towards treatment stabilisation.

The increase of stabilised protein or peptide markers in comparison to the pilot study shows that in the snap-frozen samples not all degradation is rapid and can continue to occur up to the 20 minutes time point.

3.6.4.6 Two-way ANOVA (condition 1) as sorting method

Profile analysis was further carried out using 2-way ANOVA as sorting method. This was seen to be a reasonable strategy as the assigned condition 1 was targeted between treated and non-treated groups.

A)



B)

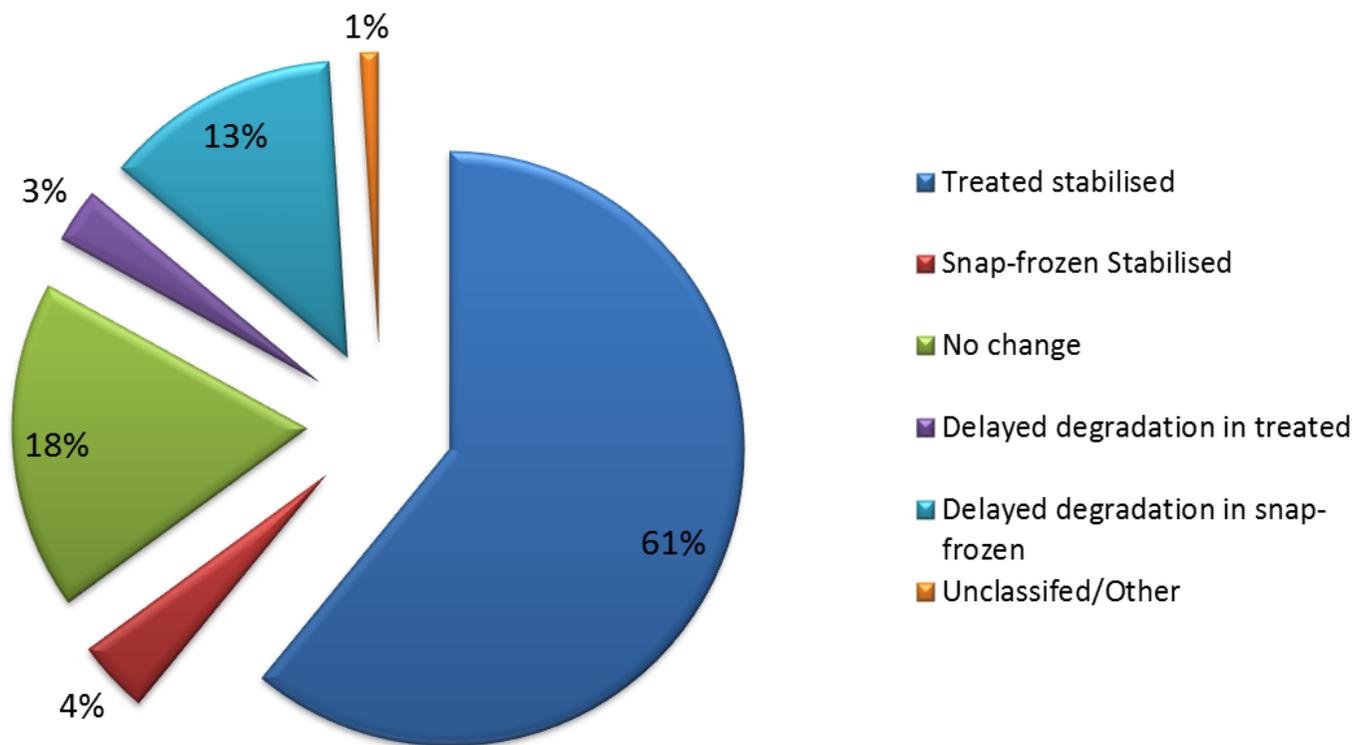


Figure 3-26: A) Profile distribution for the main investigation using 2-way ANOVA as a sorting method of categories shown in Figure 3-8, Figure 3-9 and Figure 3-10. A) Shows the proportions of various experimental group categories considered in the predicted profiles then further broken down into proportions assigned to individual profiles. B) The proportion of all stabilised groups compiled for both treated and snap-frozen. This gives an overview of relevant category amalgamation. As can be seen, the treated stabilised proportion is the greatest at 61% while the snap-frozen stabilisation is only 4%. 573 spot were included in total.

The proportions compared to 1-way ANOVA analysis in experiment two have changed slightly, but do not show massive deviation. With a 4% increase seen in treated stabilized and 3% reduction in snap-frozen stabilized. It seems using 2-way ANOVA for condition 1 (treated/snap-frozen) leads to a slight biased approach to illuminate a higher proportion of snap-frozen stabilisation spots in comparison with 1-way ANOVA. Other proportion follows a similar pattern, within the different intensity profile classes in the majority of cases, between the two methods of sorting. Therefore it would appear acceptable to use either method compared to one another.

3.6.4.7 Summary of profile analysis

Whilst being aware of the limitation of the profile analysis discussed the DiGE pilot investigation and main investigation profile analysis as a whole was successful, alluding towards:

- For protein/marker spots sorted on the basis of significance scores, treatment using the Stabilzor T1 denaturing device (Denator AB, Gothenburg, Sweden), has a positive effect on the majority of spots. Treated spots showed proportions of 37% and 57% (of stabilised markers) for pilot investigation and main investigation respectively, which is a considerably higher portion in comparison to the snap-frozen (or treated affected) quotient.
- Not all degradation can be considered rapid as protein markers have been seen to rise after 10 minutes, exhibiting the need for careful sample procedures.
- Deconvolution and manual check of data is comparatively easier when using 2-way ANOVA to sort due to the smaller subset of markers considered. However, a more inclusive approach is to use 1-way ANOVA as a sorting method.

3.6.4.8 Caveat

There are always limitations and assumptions present when trying to view proteomic data globally and simplification is generally necessary in order to make conclusions. Indeed, at its core, proteomics is a field of research which required separation of samples in an attempt to deconvolute data to a more comprehensible level. Specific examples of proteomic stabilization have equally been presented within the framework of a global view. It is understood that specific examples require more validation work in the future, but is beyond

what this study was trying to achieve and would be presented for future attention.

3.6.4.9 Example intensity profiles for the main investigation

In choosing specific profile example to include, all 623 and 573 spots included in both 1-way and 2-way ANOVA sorting were to be used (total inclusion of 932 spots with 264 spots in common as shown in 4-way Venn diagram in Figure 3-27). As before, each of these spots were divided into the relevant categories by manual visualisation against predicted profiles in Figure 3-8 and Figure 3-10. The innate nature of ANOVA makes visual comparison essential.

In addition, it is important to be aware that sorting by different forms of ANOVA may drastically change results as the Venn in Figure 3-27 shows by viewing the global proportions of spots. It allows the display of crossover of various ANOVA sorting of significant spots with $p\text{-values} \leq 0.05$. It displays considerable crossover out of 1049 spots only 40 actually appear in all 4 possible ways to that ANOVA can be performed in DeCyder Software.

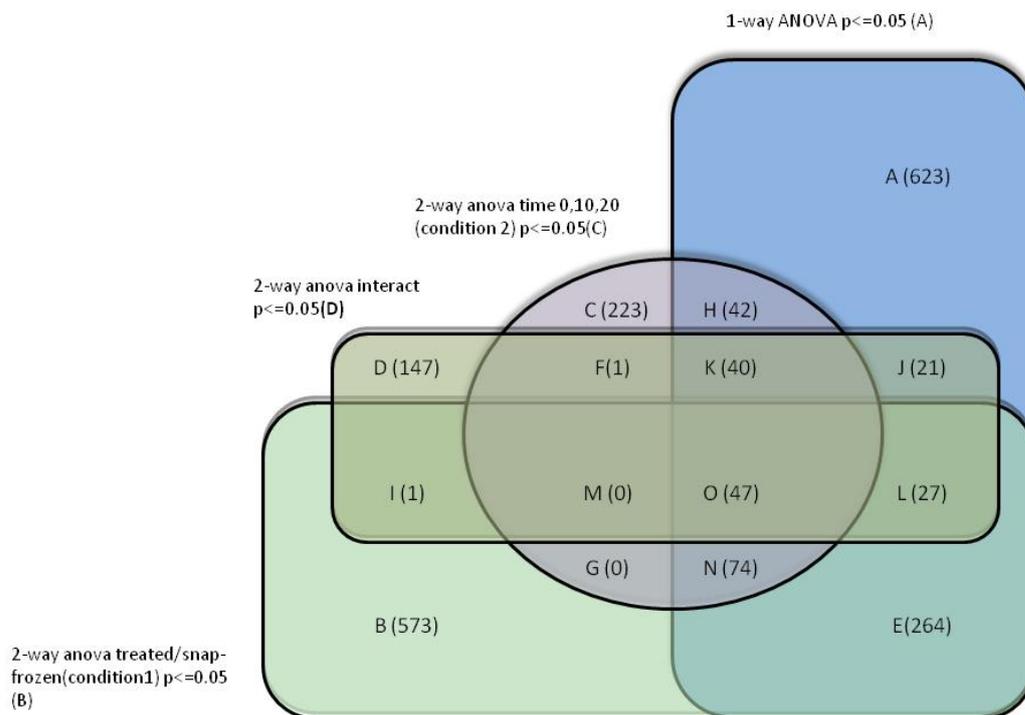


Figure 3-27: 4-Venn analysis for the main investigation: Shows cross-over of various ANOVA methods used to sort data. The low proportion of spots in common between all 4 methods of sorting is significant as data will be lost if only one is considered.

The conclusion from this dictates that one form of analysis will always leave a gap where possible markers might be missed. It is therefore essential to have multiple and complementary analytical methods to allow for comparison to

avoid missed data. It is therefore important to consider which analytical approach is most appropriate for the data and conclusion that wish to be drawn. This helps to validate the use of multiple forms of sorting methods and the use of a number of profiles to find examples of markers. Below are some examples of markers found for the various profiles considered.

3.6.4.10 Treated Stabilisation.

First, to be considered, are examples of treated stabilised intensity profiles, as outlined in Figure 3-8 (i-vi) and their positions on the gel map.

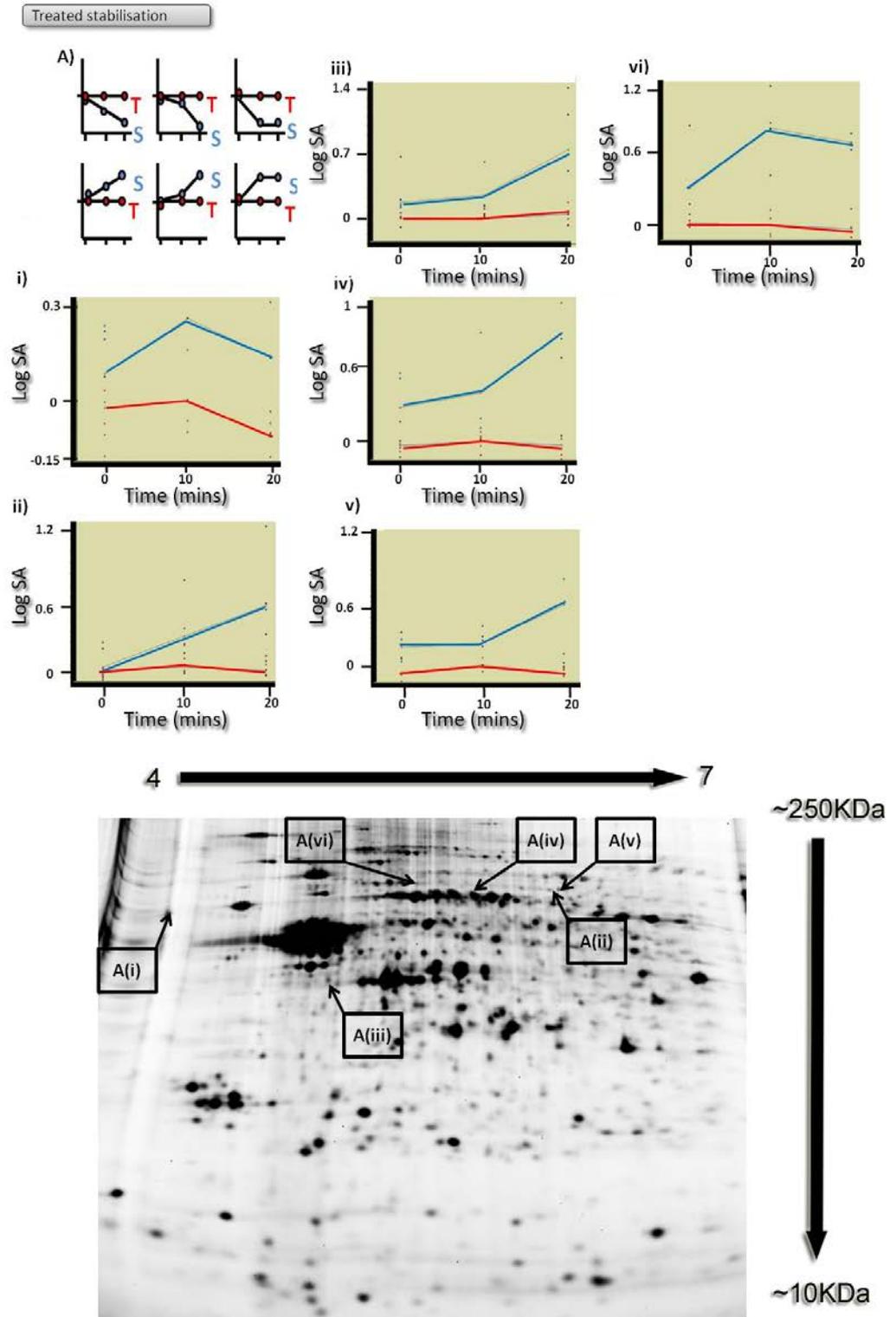


Figure 3-28: Example profiles from the main investigation matching predicted profiles given in Figure 3-8 for treated stabilised A(i-vi). Example profiles of actual spots and their location on the gel map below. Blue lines denote snap-frozen and red heat-treated

Intensity profiles in Figure 3-28 (Ai-vi) show examples of identified spots, exhibiting treated stabilisation and make up 13% of markers indicated through 1-way ANOVA profile analysis. Intensity profiles like (Ai) make up 8% of the treated stabilised group and has been identified as Sfi1 homolog; spindle assembly associated (yeast) [Mus musculus] showing t-test between treated/snap-frozen= 20 of p-value=0.035 and t-test of treated = (0, 10) and 20 of p-value = 0.13 and treated and snap-frozen = 0 of p=0.32 (no significant difference as a required pattern for this profile). The validation of profiles like this is essential, as a first glance at the pattern looks identical between the two groups but it is important to consider treated is at a log standardised abundance of 0. In (ii, iii and v) PREDICTED: similar to pORF2 [Mus musculus] has been identified. Multiple identifications are covered in greater detail in sections 3.5.2.9 on page 172. Both (ii and iii) follow similar intensity profile patterns showing t-test between treated/snap-frozen= 20 of p-value= 0.005, 0.0076 and 0.007 respectively and being no significantly different between treated = (0, 10) and 20 having p-value = 0.72, 0.69 and 0.31 respectively. Heat shock protein 9 [Mus musculus] or mortalin mot-1=hsp70 homolog Cytosolic form was identified as (iv) showing t-test between treated/snap-frozen= 20 of p-value=0.00017 and of treated = (0, 10) and 20 of p-value = 0.70 (vi) making up 37% of the treated stabilised group is an example where degradation may have occurred within the 10 minute time frame and has reached a plateau after that and has been identified as Albumin. It shows a t-test between treated/snap-frozen= 20 of p-value=0.00020 and of treated = (0, 10) and 20 of p-value = 0.56. It is interesting to note that many of the examples, and indeed a high proportion of these in general, would be miscategorised in an experiment. All the data is summarised in Table 3-9, Table 3-10 and Table 3-11.

As can be seen in Table 3-10, using profiles as a way to pull out valid data is robust, the exception being only one of the examples which is highlighted. This shows that collating the profile analysis with individual statistics is a good method and showed that a cross validating approach was required. This helps to show the validation of using a Venn analysis (which uses statistics) to help confirm valid protein/peptide marker spots. The exception shown in Table 3-10 is the highlighted cell. They show that the starting points at time point = 0 minutes differ. This could indicate that degradation may have occurred

between treatment (heating or snap freezing) or the heat-treatment itself. This requires further investigation. This trend is replicated in Table 3-13.

Spot master Number	Label on fig	Group	Protein name	Average log standard abundance	Standard Deviation of log standard abundance
694	Ai	T0	Sfi1 homolog, spindle assembly associated (yeast) [Mus musculus]	-0.020	0.119
694	Ai	T10		0.000	0.107
694	Ai	T20		-0.105	0.039
694	Ai	S0		0.080	0.179
694	Ai	S10		0.24	
694	Ai	S20		0.123	0.155
581	Aii	T0	PREDICTED: similar to pORF2 [Mus musculus]	0.013	0.097
581	Aii	T10		0.042	0.114
581	Aii	T20		0.010	0.056
581	Aii	S0		0.050	0.119
581	Aii	S10		0.388	0.308
581	Aii	S20		0.575	0.380
1316	Aiii	T0	ORF2 [Mus musculus domesticus]	0.030	0.143
1316	Aiii	T10		0.018	0.130
1316	Aiii	T20		0.073	0.085
1316	Aiii	S0		0.202	0.266
1316	Aiii	S10		0.288	0.242
1316	Aiii	S20		0.785	0.676
527	Aiv	T0	heat shock protein 9 [Mus musculus] OR mortalin mot-1=hsp70 homolog cytosolic form [mice, CD1-ICR embryonic fibroblasts, MEF,	-0.050	0.035
527	Aiv	T10		-0.010	0.067
527	Aiv	T20		-0.045	0.092
527	Aiv	S0		0.265	0.316
527	Aiv	S10		0.383	0.423
527	Aiv	S20		0.867	0.224
577	Av	T0	PREDICTED: similar to pORF2 [Mus musculus]	-0.065	0.089
577	Av	T10		0.000	0.054
577	Av	T20		-0.072	0.060
577	Av	S0		0.208	0.208
577	Av	S10		0.265	0.198
577	Av	S20		0.570	0.465
535	Avi	T0	albumin [Mus musculus]	0.020	0.136
535	Avi	T10		0.008	0.116
535	Avi	T20		-0.033	0.133
535	Avi	S0		0.337	0.490
535	Avi	S10		0.878	0.333
535	Avi	S20		0.750	0.075

Table 3-9: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-28. Highlighted cell indicated only 2 readings present and therefore a standard deviation cannot be performed.

Spot master Number	Label on fig	Group			P-value criteria	pvalue	1wayanova	2wayanova	Does the p-value meet the criteria	Number of experimental replicates	Degrees of Freedom
694	Ai	T0	vs	T20	p=>0.05	0.22	9.40E-10	1.20E-11	true	6	5
694	Ai	T0	vs	S0	p=>0.05	0.32			true	6	5
694	Ai	T20	vs	S20	p=<0.05	0.035			true	6	5
581	Aii	T0	vs	T20	p=>0.05	0.96	0.00028	0.00056	true	6	5
581	Aii	T0	vs	S0	p=>0.05	0.69			true	6	5
581	Aii	T20	vs	S20	p=<0.05	0.005			true	6	5
1316	Aiii	T0	vs	T20	p=>0.05	0.56	0.0042	0.00097	true	6	5
1316	Aiii	T0	vs	S0	p=>0.05	0.2			true	6	5
1316	Aiii	T20	vs	S20	p=<0.05	0.03			true	6	5
527	Aiv	T0	vs	T20	p=>0.05	0.94	3.60E-05	2.70E-06	true	6	5
527	Aiv	T0	vs	S0	p=>0.05	0.068			true	6	5
527	Aiv	T20	vs	S20	p=<0.05	0.00017			true	6	5
577	Av	T0	vs	T20	p=>0.05	0.91	0.00078	0.00028	true	6	5
577	Av	T0	vs	S0	p=>0.05	0.012			false	6	5
577	Av	T20	vs	S20	p=<0.05	0.0076			true	6	5
535	Avi	T0	vs	T20	p=>0.05	0.6	0.00017	0.000013	true	6	5
535	Avi	T0	vs	S0	p=>0.05	0.27			true	6	5
535	Avi	T20	vs	S20	p=<0.05	0.00026			true	6	5

Table 3-10: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-28. The p-values given are from the shown Student's t-tests. This confirms the validity of the profile analysis by showing no significant differences between the treated = 0 and 10 minutes but showing difference between the snap-frozen = 0 and 10 minutes. With the exception of the highlighted cell.

Spot master Number	Label on fig	Group	Protein name	MOWSE Score	Score MS (PMF)	Machine identified on	p-value	Peptides MS/MS	Charge State	Peptides for PMF	Percentage coverage
694	Ai	T0	Sfi1 homolog, spindle assembly associated (yeast) [Mus musculus]	N/A	68	4700	0.05	N/A	1+	6	72
694	Ai	T10									
694	Ai	T20									
694	Ai	S0									
694	Ai	S10									
694	Ai	S20									
581	Aii	T0	PREDICTED: similar to pORF2 [Mus musculus]	87	N/A	q star	0.05	3	2+	12	30
581	Aii	T10									
581	Aii	T20									
581	Aii	S0									
581	Aii	S10									
581	Aii	S20									
1316	Aiii	T0	ORF2 [Mus musculus domesticus]	74	N/A	q star	0.01	3	2+	14	22
1316	Aiii	T10									
1316	Aiii	T20									
1316	Aiii	S0									
1316	Aiii	S10									
1316	Aiii	S20									
527	Aiv	T0	heat shock protein 9 [Mus musculus] OR mortalin mot-1=hsp70 homolog cytosolic form [mice, CD1-ICR embryonic fibroblasts, MEF, Peptide, 679 aa]	130	N/A	q star	0.05	4	2+	16	5
527	Aiv	T10									
527	Aiv	T20									
527	Aiv	S0									
527	Aiv	S10									
527	Aiv	S20									
577	Av	T0	PREDICTED: similar to pORF2 [Mus musculus]	82	q star	q star	0.05	FIND	2+	9	26
577	Av	T10									
577	Av	T20									
577	Av	S0									
577	Av	S10									
577	Av	S20									
535	Avi	T0	albumin [Mus musculus]	86	N/A	q star	0.05	4	2+	13	6
535	Avi	T10									
535	Avi	T20									
535	Avi	S0									
535	Avi	S10									
535	Avi	S20									

Table 3-11: Table of identifications using Mass Spectrometry shown in Figure 3-28. Table gives details of the return MS identification and, p-value and percentage coverage of the identifications.

3.6.4.11 Treated Stabilised with immediate degradation and unclassified in snap-frozen

Secondly, to be considered, examples of treated stabilised profiles exhibiting and immediate degradation and unclassified response in snap-frozen intensity profiles, as outlined in Figure 3-9 and Figure 3-10 and their positions on the gel map.

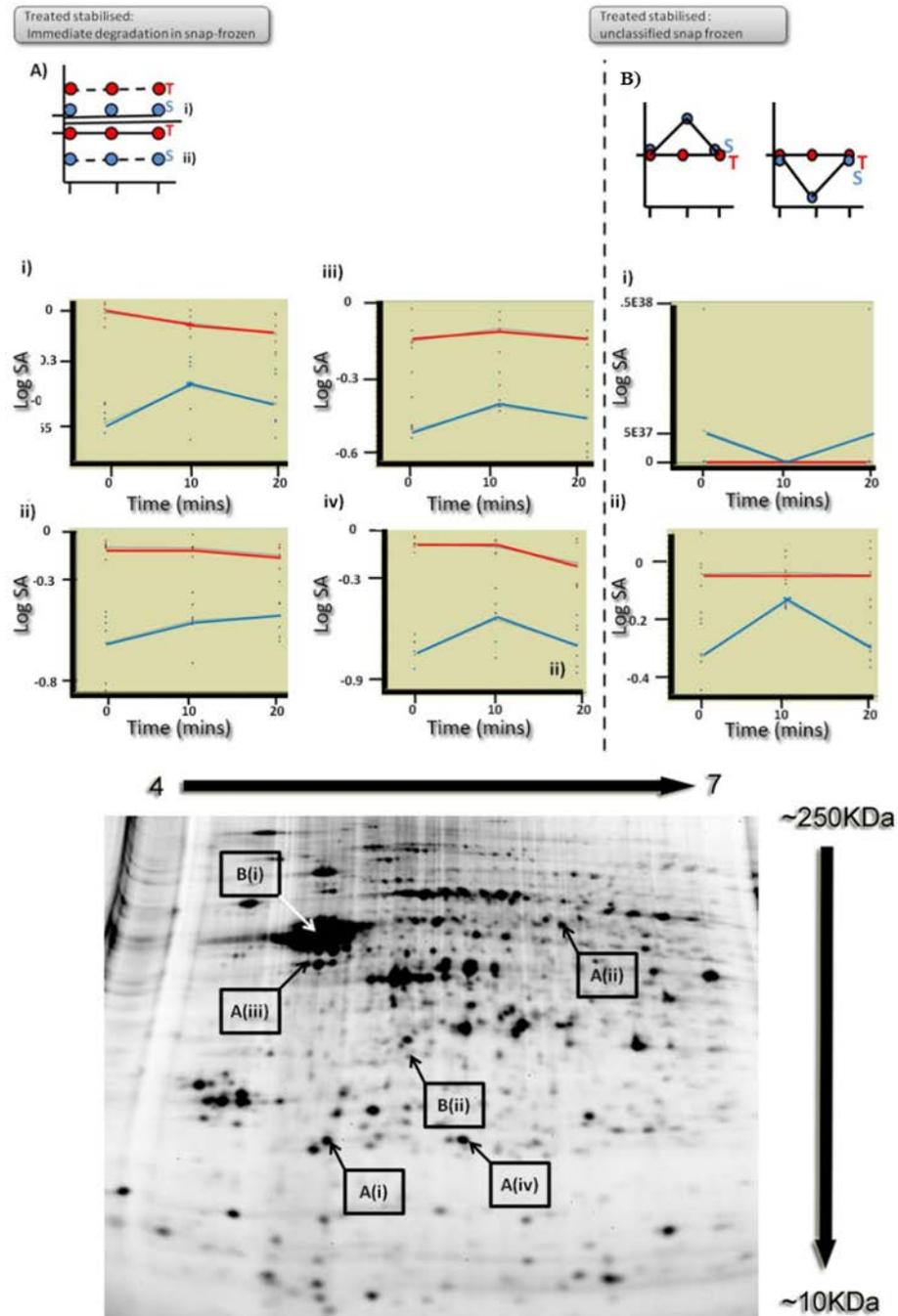


Figure 3-29: Example profiles from the main investigation matching predicted profiles given in Figure 3-9 and Figure 3-10 for treated stabilised Ai-iv) Example profiles of actual spots and their location on the gel map of treated stabilised with snap-frozen showing immediate degradation. Bi-ii) Example profiles of actual spots and their location on the gel map of treated stabilised with snap-frozen unclassified. Blue lines denote snap-frozen and red heat-treated.

Intensity profile in Figure 1-26 (Ai-iv) show examples of identified spots exhibiting treated stabilisation, whilst in snap-frozen displaying a predicted degradate effect prior to the warming time course. Additionally, intensity profile in Figure 1-26 (Bi-ii) shows treated stabilisation whilst in snap-frozen, displaying an unclassified response. (Ai) has been identified RIKEN cDNA 1810014F10; isoform CRA_i [Mus musculus] showing t-test between treated/snap-frozen = 20 of p-value=0.00045 and t-test of treated = 0 and 10 of p-value = 0.15 and treated and snap-frozen = 0 p-value=4.2E-6, showing complete separation of experimental groups treated and snap-frozen at the outset, suggestion degradation in the snap-frozen prior to warming or in storage. These profile classification make up a large proportion of treated stabilisation (30%/57% and 30%/64% in 1-way and 2-way ANOVAs respectively). (ii, iii and iv) show similar profiles to (i) and have been identified as; (ii) mKIAA1735 protein [Mus musculus]/Ccd1BbetaL [Mus musculus]/PREDICTED: similar to pORF2 [Mus musculus], (iii) Atp5b protein [Mus musculus]/mitochondrial ATP synthase beta subunit [Mus musculus] and (iv) 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]. With corresponding p-values = 0.00019, 0.0035 and 0.00018 for treated/snap-frozen= 20. (Bi-ii) gives examples of snap-frozen tissue exhibits unclassified behaviour with treated being stabilised. These examples are as yet unidentified.

As is shown in Table 3-13 the statistical criteria for all except one has been satisfied, showing that those markers can be placed in these categories and helps to validate the use of profile analysis.

Spot master Number	Label on fig	Group	Protein name	Average log standard abundance	Standard Deviation of log standard abundance
2208	Ai	T0	RIKEN cDNA 1810014F10; isoform CRA_i [Mus musculus]	-0.045	-0.048
2208	Ai	T10		-0.115	-0.100
2208	Ai	T20		-0.152	0.100
2208	Ai	S0		-0.642	0.128
2208	Ai	S10		-0.438	0.175
2208	Ai	S20		-0.532	0.150
868	Aii	T0	mKIAA1735 protein [Mus musculus]	-0.063	0.036
868	Aii	T10		-0.068	0.039
868	Aii	T20		-0.107	0.059
868	Aii	S0		-0.580	0.190
868	Aii	S10		-0.460	0.130
868	Aii	S20		-0.433	0.126
1043	Aiii	T0	Atp5b protein [Mus musculus]	-0.133	0.082
1043	Aiii	T10		-0.085	0.070
1043	Aiii	T20		-0.123	0.025
1043	Aiii	S0		-0.497	0.090
1043	Aiii	S10		-0.400	0.127
1043	Aiii	S20		-0.445	0.154
2231	Aiv	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	-0.007	0.037
2231	Aiv	T10		-0.025	0.042
2231	Aiv	T20		-0.123	0.150
2231	Aiv	S0		-0.680	0.128
2231	Aiv	S10		-0.468	0.154
2231	Aiv	S20		-0.617	5.000
857	Bi	T0	alpha-internexin [Mus musculus]	-0.030	-0.015
857	Bi	T10		0.020	0.013
857	Bi	T20		-0.062	0.053
857	Bi	S0		-0.308	-0.228
857	Bi	S10		-0.158	-0.200
857	Bi	S20		-0.490	0.395
1721	Bii	T0	Unidentified at this time	-0.063	-0.055
1721	Bii	T10		-0.055	-0.018
1721	Bii	T20		-0.053	0.113
1721	Bii	S0		-0.330	0.115
1721	Bii	S10		-0.180	-0.145
1721	Bii	S20		-0.370	-0.335

Table 3-12: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-29.

Spot master Number	Label on fig	Group			P-value criteria	pvalue	1wayanova	2wayanova	Does the p-value meet the criteria	Number of experimental replicates	Degrees of Freedom
2208	Ai	T0	vs	T20	p=>0.05	0.047	3.30E-09	4.00E-11	false	6	5
2208	Ai	T0	vs	S0	p=<0.05	4.2E-06			true	6	5
2208	Ai	T20	vs	S20	p=<0.05	4.50E-04			true	6	5
868	Aii	T0	vs	T20	p=>0.05	0.16	5.30E-10	4.30E-12	true	6	5
868	Aii	T0	vs	S0	p=<0.05	0.000065			true	6	5
868	Aii	T20	vs	S20	p=<0.05	1.90E-04			true	6	5
1043	Aiii	T0	vs	T20	p=>0.05	0.79	3.30E-07	3.20E-09	true	6	5
1043	Aiii	T0	vs	S0	p=<0.05	0.000029			true	6	5
1043	Aiii	T20	vs	S20	p=<0.05	3.50E-03			true	6	5
2231	Aiv	T0	vs	T20	p=>0.05	0.096	3.90E-12	3.80E-14	true	6	5
2231	Aiv	T0	vs	S0	p=<0.05	2.3E-07			true	6	5
2231	Aiv	T20	vs	S20	p=<0.05	1.80E-04			true	6	5
857	Bi	T0	vs	T20	p=>0.05	0.69	0.023	0.0015	true	6	5
857	Bi	T0	vs	S0	p=<0.05	0.00018			true	6	5
857	Bi	T20	vs	S20	p=<0.05	0.0097			true	6	5
1721	Bii	T0	vs	T20	p=>0.05	0.95	1.80E-05	1.90E-06	true	6	5
1721	Bii	T0	vs	S0	p=<0.05	0.0025			true	6	5
1721	Bii	T20	vs	S20	p=<0.05	0.0042			true	6	5

Table 3-13: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-29. The p-values given are from the shown Student's t-tests. This confirms the validity of the profile analysis by showing no significant differences between the treated = 0 and 10 minutes but showing difference between the snap-frozen = 0 and 10 minutes. With the exception of the highlighted cell.

Spot master Number	Label on fig	Group	Protein name	MOWSE Score	Score MS(PMF)	Machine identified on	p-value	Peptides MS/MS	Charge State	Peptides for PMF	Percentage coverage
2208	Ai	T0	RIKEN cDNA 1810014F10; isoform CRA_i [Mus musculus]	N/A	65	4700	0.05	N/A	1+	7	83
2208	Ai	T10									
2208	Ai	T20									
2208	Ai	S0									
2208	Ai	S10									
2208	Ai	S20									
868	Aii	T0	mKIAA1735 protein [Mus musculus]	N/A	67	4700	0.05	N/A	1+	13	23
868	Aii	T10									
868	Aii	T20									
868	Aii	S0									
868	Aii	S10									
868	Aii	S20									
1043	Aiii	T0	Atp5b protein [Mus musculus]	411	N/A	4700	0.05	8	1+	N/A	17
1043	Aiii	T10									
1043	Aiii	T20									
1043	Aiii	S0									
1043	Aiii	S10									
1043	Aiii	S20									
2231	Aiv	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	43	N/A	q star	0.05	1	2+	N/A	1
2231	Aiv	T10									
2231	Aiv	T20									
2231	Aiv	S0									
2231	Aiv	S10									
2231	Aiv	S20									
857	Bi	T0	alpha-internexin [Mus musculus]	82	N/A	4700	0.05	4	1+	N/A	7
857	Bi	T10									
857	Bi	T20									
857	Bi	S0									
857	Bi	S10									
857	Bi	S20									
1721	Bii	T0	Unidentified at this time	N/A							
1721	Bii	T10									
1721	Bii	T20									
1721	Bii	S0									
1721	Bii	S10									
1721	Bii	S20									

Table 3-14: Table of identifications using Mass Spectrometry shown in Figure 3-29. Table gives details of the return MS identification and, p-value and percentage coverage of the identifications.

3.6.4.12 Snap-frozen Stabilisation, Snap-frozen stabilisation in immediate degradation and unclassified in treated

Thirdly, to be considered, examples of snap-frozen stabilised, snap-frozen stabilised with immediate degradation in treated and unclassified in treated intensity profiles, as outlined in Figure 3-9 and Figure 3-10 and their positions on the gel map.

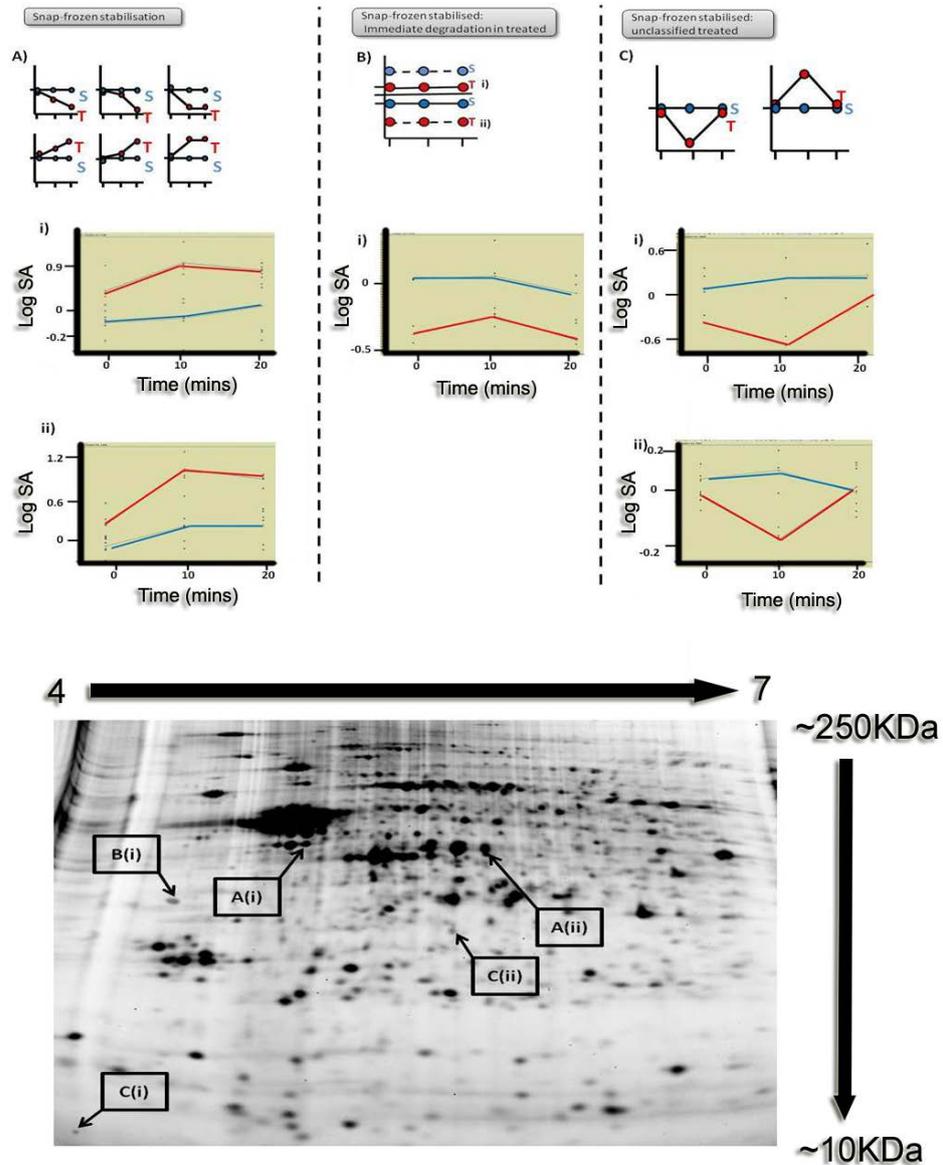


Figure 3-30: Example profiles from the main investigation matching predicted profiles given in Figure 3-8, Figure 3-9 and Figure 3-10 for snap-frozen stabilised Ai-ii) Example profiles of actual spots and their location on the gel map. Bi) Example profiles of actual spots and their location on the gel map of treated stabilised with treated showing immediate degradation. Ci-ii) Example profiles of actual spots and their location on the gel map of snap-frozen stabilised with treated unclassified. Blue line denotes snap-frozen and red heat treated.

Intensity profile in Figure 3-30(Ai-ii) show examples of spots exhibiting snap-frozen stabilisation intensity profiles identified as; (i) Tpr [Mus musculus] and (ii) nebulin-related anchoring protein isoform C [Mus musculus]. Validation using t-test p-value was used with p-value = 0.021 and 0.00031 respectively for t-test treated/snap-frozen = 20 showing significant difference between treated/snap-frozen = 20 minutes. Additionally, t-test for snap-frozen = (0, 10) and 20 have p-value = 0.24 for both (i and ii). Displayed in section (Bi) an example of an unidentified marker showing an intensity profile matching stabilisation in snap-frozen associated with immediate degradation in treated. (Ci-ii) are typical examples of profile intensity examples of snap-frozen stabilised linked with unclassified treated intensity profiles. This is an unusual profile, with an explanation of technical and biological variation causing the averaging of a low abundance for the 10 minute time point. (Ci) as yet is unidentified, however, (Cii) has been identified as 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus], which has also been identified in other spots at various locations on the gel including exhibiting behaviour seen in Figure 3-28 (Aiv) which is treated stabilised, homology issues need to be considered here and could lead to an explanation to multiple location demonstrating different profile intensities. Once again the need for cross-validation is required, as in this case the statistical criteria shown in Table 3-16 shows more contradictions.

Spot master Number	Label on fig	Group	Protein name	Average log standard abundance	Standard Deviation of log standard abundance
1186	Ai	T0	Tpr [Mus musculus]	-0.288	-0.223
1186	Ai	T10		0.028	0.048
1186	Ai	T20		-0.045	0.072
1186	Ai	S0		-0.648	-0.513
1186	Ai	S10		-0.598	-0.478
1186	Ai	S20		-0.37	0.284
1286	Aii	T0	nebulin-related anchoring protein isoform C [Mus musculus]	-0.840	-0.670
1286	Aii	T10		0.050	0.150
1286	Aii	T20		-0.103	-0.083
1286	Aii	S0		-0.952	0.117
1286	Aii	S10		-0.687	0.238
1286	Aii	S20		-0.683	0.226
1554	Bi	T0	Unidentified at this time	-0.07	
1554	Bi	T10		0.030	0.460
1554	Bi	T20		0.16	
1554	Bi	S0		0.345	0.163
1554	Bi	S10		0.323	0.191
1554	Bi	S20		0.213	0.000
3068	Ci	T0	Unidentified at this time	-0.27	
3068	Ci	T10		-0.373	0.156
3068	Ci	T20		0.05	
3068	Ci	S0		0.183	0.268
3068	Ci	S10		0.207	0.368
3068	Ci	S20		0.345	0.468
1852	Cii	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	-0.160	0.154
1852	Cii	T10		-0.255	0.197
1852	Cii	T20		-0.035	0.017
1852	Cii	S0		-0.035	0.037
1852	Cii	S10		0.223	0.047
1852	Cii	S20		0.093	0.117

Table 3-15: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-30. Highlighted cell indicated only 2 readings present and therefore a standard deviation cannot be performed.

Spot master Number	Label on fig	Group			P-value criteria	pvalue	1wayanova	2wayanova	Does the p-value meet the criteria	Number of experimental replicates	Degrees of Freedom
1186	Ai	S0	vs	S20	p=>0.05	0.29	2.90E-05	1.20E-06	true	6	5
1186	Ai	S0	vs	T0	p=>0.05	0.005			false	6	5
1186	Ai	S20	vs	T20	p=<0.05	0.021			true	6	5
1286	Aii	S0	vs	S20	p=>0.05	0.028	7.30E-09	7.40E-09	false	6	5
1286	Aii	S0	vs	T0	p=>0.05	0.016			false	6	5
1286	Aii	S20	vs	T20	p=<0.05	3.10E-04			true	6	5
1554	Bi	S0	vs	S20	p=>0.05	0.88	0.38	N/A	true	6	5
1554	Bi	S0	vs	T0	p=<0.05					6	5
1554	Bi	T20	vs	S20	p=<0.05					6	5
3068	Ci	S0	vs	S20	p=>0.05	0.57	0.088	N/A	true	6	5
3068	Ci	S0	vs	T0	p=>0.05					6	5
3068	Ci	T10	vs	S10	p=<0.05	0.01			true	6	5
1852	Cii	S0	vs	S20	p=>0.05	0.087	0.00081	0.00061	true	6	5
1852	Cii	S0	vs	T0	p=>0.05	0.16			true	6	5
1852	Cii	T10	vs	S10	p=<0.05	0.093			false	6	5

Table 3-16: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-30. The p-values given are from the shown Student's t-tests. This confirms the validity of the profile analysis by showing no significant differences between the treated = 0 and 10 minutes but showing difference between the snap-frozen = 0 and 10 minutes. With the exception of the highlighted cell. Missing values are shown where a t-test was not able to be performed due to lack of data. Contractions are shown by a false result.

Spot master Number	Label on fig	Group	Protein name	MOWSE Score	Score MS (PMF)	Machine identified on	P Value	Peptides MS/MS	Charge State	Peptides for PMF	Percentage coverage
1186	Ai	T0	Tpr [Mus musculus]	N/A	63	4700	0.05	N/A	1+	26	83
1186	Ai	T10									
1186	Ai	T20									
1186	Ai	S0									
1186	Ai	S10									
1186	Ai	S20									
1286	Aii	T0	nebulin-related anchoring protein isoform C [Mus musculus]	78	N/A	q star	0.01	4	2+	N/A	72
1286	Aii	T10									
1286	Aii	T20									
1286	Aii	S0									
1286	Aii	S10									
1286	Aii	S20									
1554	Bi	T0	Unidentified at this time	N/A							
1554	Bi	T10									
1554	Bi	T20									
1554	Bi	S0									
1554	Bi	S10									
1554	Bi	S20									
3068	Ci	T0	Unidentified at this time	N/A							
3068	Ci	T10									
3068	Ci	T20									
3068	Ci	S0									
3068	Ci	S10									
3068	Ci	S20									
1852	Cii	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	43	N/A	q star	0.05	1	2+	N/A	1
1852	Cii	T10									
1852	Cii	T20									
1852	Cii	S0									
1852	Cii	S10									
1852	Cii	S20									

Table 3-17: Table of identifications using Mass Spectrometry shown in Figure 3-30. Table gives details of the return MS identification and, p-value and percentage coverage of the identifications.

3.6.4.13 Delayed degradation in treated and snap-frozen, No changes and others.

Finally, to be considered, examples of delayed degradation in treated and snap-frozen samples and no change intensity profiles, as outlined in Figure 3-9 and Figure 3-10 and their positions on the gel map.

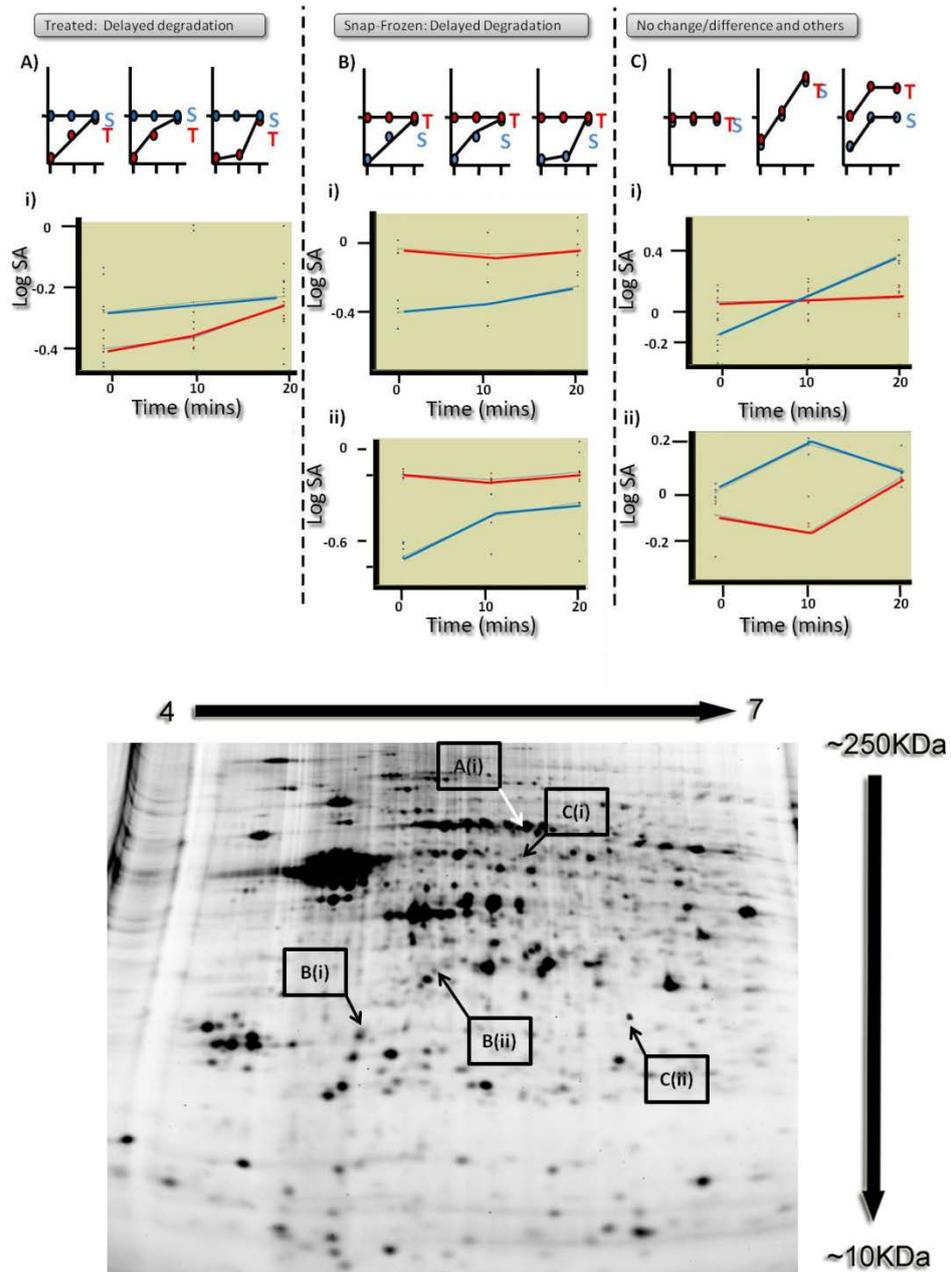


Figure 3-31: Example profiles from the main investigation matching predicted profiles given in Figure 3-8, Figure 3-9 and Figure 3-10 for delayed degradation and no changes Ai) Example profiles of actual spots and their location on the gel map for delayed degradation in treated. Bi-ii) Example profiles of actual spots and their location on the gel map of delayed degradation in snap-frozen. Ci-ii) Example profiles of actual spots and their locations on the gel map of spots classified as not changing or have no difference. Blue lines denote snap-frozen and red heat-treated

Intensity profiles in Figure 3-31 (A and B) show some typical example of how sample-frozen and treated samples have possibly delayed the degradation of that marker. (Bi-ii) demonstrated profiles, where degradation is delayed in snap-frozen exhibited in 9% of spot profile when using 1-way ANOVA to sort this is in comparison to 1% for treated samples exhibiting profiles typically represented by unidentified marker in (Ai). The chain of protein spots that runs in line with Ai) do not all exhibit the same profile. In comparison between heat-treated and snap-frozen, heat treated spots exhibit high intensities over all for this train of spots, also shown by Robinson et al, 2009. Although identifications were not obtained for this chain they are likely to be phosphorylated as this gives a train of proteins at various pIs. It is unlikely to be caused by glycosylations, as they often appear as a small mass shift or slight smear of the gel, with little change in pI. There are stains available and enzyme based approaches, however, characterising PTM on gels, although possible, can be practically difficult (Nelson et al., 2008). (Bi) has been identified as Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]. The profile possibly illustrated the degradation of treated tissue before the warming time course was performed, showing no change across time points in contrast to the evolution the snap frozen markers intensity falling at time point = 20 minutes. A further example is given in (Bii) mCG9572; isoform CRA_a [Mus musculus] demonstrating similar behaviour. In 26% (with 1-way ANOVA sorting) of the markers profiles similar to the examples in (ci and ii) occurs in either no change (19%) or other (7%). (Ci) has been identified as protein kinase [Mus musculus]/budding uninhibited by benzimidazoles 1 isoform 2 [Mus musculus] and displays at first glance a profile where snap-frozen looks to be changing with respect to treated by examining the variance across replicate gels reveals no change between each of the experimental groups. Interestingly (Cii) shows mirror profiles and is identified as 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus], one of the multiply identified markers exhibiting a number of different intensity profiles. The position on the spot map relative to other instances of this identification can be seen in section 3.5.2.9.

Spot master Number	Label on fig	Group	Protein name	Average log standard abundance	Standard Deviation of log standard abundance
612	Ai	T0	Unidentified at this time	-0.448	-0.363
612	Ai	T10		-0.375	-0.298
612	Ai	T20		-0.107	0.096
612	Ai	S0		-0.153	0.221
612	Ai	S10		-0.095	0.347
612	Ai	S20		-0.06	0.298
1892	Bi	T0	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	-0.027	0.040
1892	Bi	T10		-0.050	0.087
1892	Bi	T20		-0.045	0.127
1892	Bi	S0		-0.333	0.068
1892	Bi	S10		-0.197	0.148
1892	Bi	S20		-0.207	0.363
1645	Bii	T0	mCG9572, isoform CRA_a [Mus musculus]	-0.015	0.021
1645	Bii	T10		-0.055	0.021
1645	Bii	T20		-0.0125	-0.005
1645	Bii	S0		-0.618	-0.515
1645	Bii	S10		-0.265	-0.230
1645	Bii	S20		-0.186	-0.185
827	Ci	T0	protein kinase [Mus musculus] OR budding uninhibited by benzimidazoles 1 isoform 2 [Mus musculus]	-0.15	
827	Ci	T10		0.060	0.033
827	Ci	T20		0.08	
827	Ci	S0		0.340	0.055
827	Ci	S10		0.130	0.088
827	Ci	S20		0.250	0.200
1852	Cii	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	-0.160	-0.200
1852	Cii	T10		-0.255	-0.200
1852	Cii	T20		0.045	0.028
1852	Cii	S0		-0.035	-0.020
1852	Cii	S10		0.223	0.047
1852	Cii	S20		0.093	0.117

Table 3-18: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-31.

Spot master Number	Label on fig	Group			p-value criteria	p-value	1wayanova	2wayanova	p-value criteria check	Number of experimental replicates	Degrees of Freedom
612	Ai	S0	vs	S20	p=>0.05	0.56	0.19	0.055	true	6	5
612	Ai	S0	vs	T0	p=>0.05	0.083			true	6	5
612	Ai	S20	vs	T20	p=<0.05	0.71			false	6	5
1892	Bi	T0	vs	T20	p=>0.05	0.84	0.19	0.015	true	6	5
1892	Bi	S20	vs	T20	p=>0.05	0.43			true	6	5
1892	Bi	S0	vs	T0	p=<0.05	0.0026			true	6	5
1645	Bii	T0	vs	T20	p=>0.05	0.72	5.00E-05	8.20E-06	true	6	5
1645	Bii	S20	vs	T20	p=>0.05	0.21			true	6	5
1645	Bii	S0	vs	T0	p=<0.05	1.20E-05			true	6	5
827	Ci	S0	vs	T0	p=>0.05	0.02	0.001	0.55	false	6	5
827	Ci	S20	vs	T20	p=>0.05	0.00021			false	6	5
827	Ci										
1852	Cii	S0	vs	T0	p=>0.05	0.16	0.00081	0.00061	true	6	5
1852	Cii	S20	vs	T20	p=>0.05	0.46			true	6	5
1852	Cii										

Table 3-19: Table of average log abundance and standard deviations obtained for identified spots shown in Figure 3-31. The p-values given are from the shown Student's t-tests. This confirms the validity of the profile analysis by showing no significant differences between the treated = 0 and 10 minutes but showing difference between the snap-frozen = 0 and 10 minutes. With the exception of the highlighted cell.

Spot master Number	Label on fig	Group	Protein name	Score MS/MS	Score MS(PMF)	Machine identified on	P score	Peptides MS/MS	Charge State	Peptides for PMF	Percentage coverage
612	Ai	T0	Unidentified at this time					N/A			
612	Ai	T10									
612	Ai	T20									
612	Ai	S0									
612	Ai	S10									
612	Ai	S20									
1892	Bi	T0	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	N/A	226	4700	0.05	5	1+	N/A	33
1892	Bi	T10									
1892	Bi	T20									
1892	Bi	S0									
1892	Bi	S10									
1892	Bi	S20									
1645	Bii	T0	mCG9572, isoform CRA_a [Mus musculus]	64	N/A	4700	0.05	N/A	1+	15	20
1645	Bii	T10									
1645	Bii	T20									
1645	Bii	S0									
1645	Bii	S10									
1645	Bii	S20									
827	Ci	T0	protein kinase [Mus musculus] OR budding uninhibited by benzimidazoles 1 isoform 2 [Mus musculus]	N/A	36	4700	0.05	1	1+	N/A	1
827	Ci	T10									
827	Ci	T20									
827	Ci	S0									
827	Ci	S10									
827	Ci	S20									
1852	Cii	T0	5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]	43	N/A	q star	0.05	1	2+	N/A	20
1852	Cii	T10									
1852	Cii	T20									
1852	Cii	S0									
1852	Cii	S10									
1852	Cii	S20									

Table 3-20: Table of identifications using Mass Spectrometry shown in Figure 3-31. Table gives details of the return MS identification and, p-value and percentage coverage of the identifications.

3.6.4.14 Venn analysis from the main investigation.

As with the pilot investigation, Venn analysis was performed allowing for a validative comparison against profile analysis. In order to target treated and snap-frozen stabilisation, a variety of strategies were employed shown in Figure 3-32 to Figure 3-27.

In Figure 3-32 an inclusive strategy is utilised, including a total number of 1701 spots in contrast to 623 and 573 using the ANOVAs. The inclusion was achieved by using a 3-way Venn diagram using t-tests as a basis for selection. Treated stabilisation and snap-frozen stabilisation shown in (B), the red boxes, denote treated/snap frozen stabilisation for predicted profiles in Figure 3-8 (i,iii,iv,vi) and the green treated/snap- frozen stabilisation for Figure 3-8 (ii and v). In both experimental groups for these profiles treated/snap-frozen should show no difference, so a non-significant p-value of ≤ 0.05 is considered crossing over with treated/snap-frozen = 10 minutes and treated/snap-frozen = 20 minutes. Sector G in Venn diagram (A) is indicative of target profiles in the red box, while sector E considered profiles from the green box. In sector G, all of the 38 spots showed treated stabilisation split into the proportions shown in (Cii). In sector E out of 107 spots, 83% showed treated stabilisation intensity profiles with the remaining 17% showing snap-frozen stabilisation intensity profiles. Similar proportions were seen in Figure 3-34, using 4-way Venn analysis, targeting sectors O and L. This Venn had the addition of 2-way ANOVA of condition 1. Sector O being a close equivalence to sector G in Figure 3-32, and L to sector E. The cross-over of spots using the 4-way Venn and 3-way Venn is substantial.

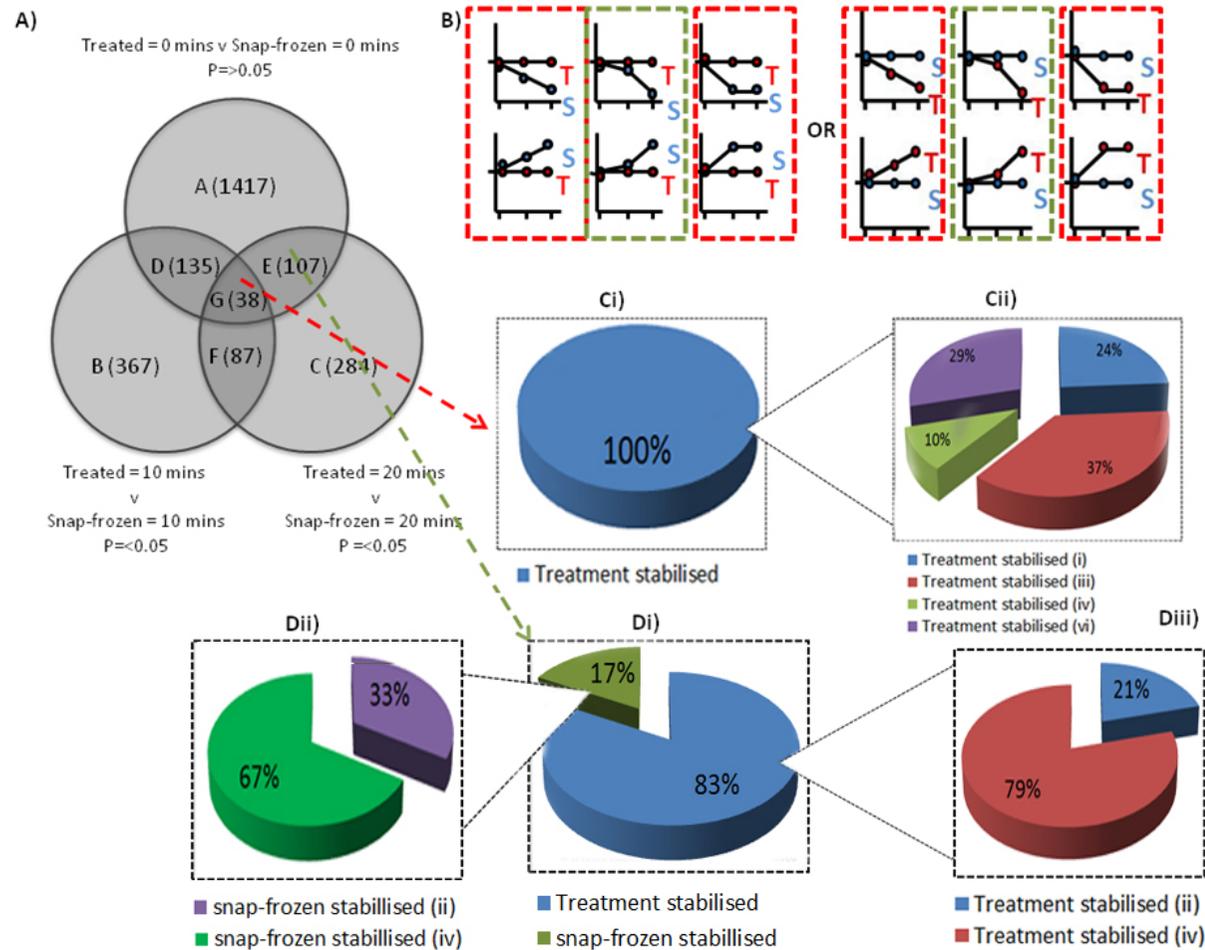


Figure 3-32: Venn analysis for the main investigation. A) Shows Venn analysis corresponding to red and green boxed profiles in B). Ci-ii-Di-iii) Show profile distribution of protein spots taken from areas G and E of the Venn diagram. The pie charts show how starting particular areas on the Venn diagram allow for the quick discovery of particular profile types. The one targeted here being treatment stabilised spots. This can significantly reduce manually sorting through profiles.

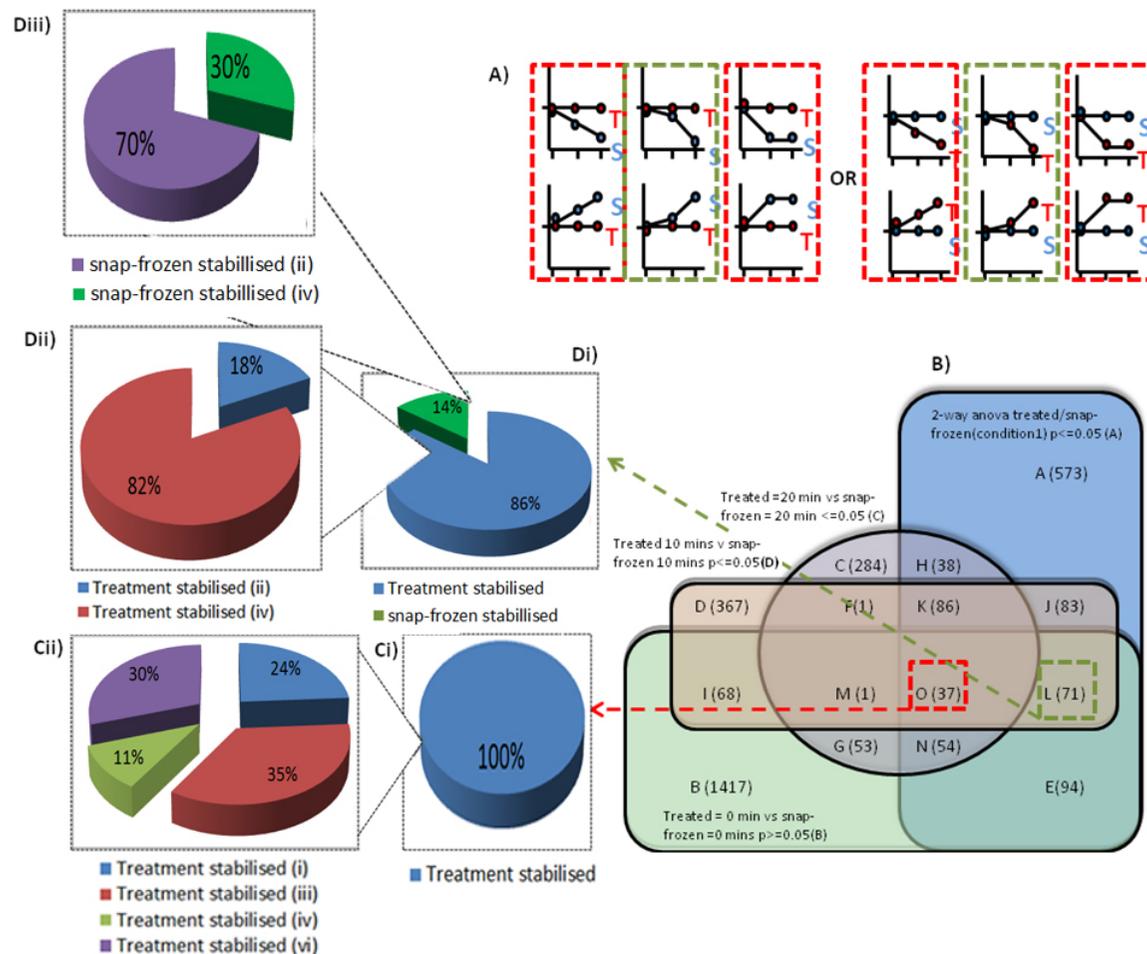


Figure 3-33: 4-Venn analysis for the main investigation: A) Red and Green boxes isolate profiles targeted in Sectors in B) relating to the profile distributions shown via pie chart in C) and D). Both Venn diagrams in Figure 3-32 and Figure 3-33 are successful at finding the predicted profiles indicated. The pie charts show how starting particular areas on the Venn diagram allow for the quick discovery of particular profile types. The one targeted here being treatment stabilised spots. This can significantly reduce manually sorting through profiles.

A combination of two strategies searching for treated frozen stabilisation is seen in Figure 3-34 and Figure 3-35. The cross-over between the two sector Gs of the Venn diagrams in (A and B) is performed and successfully identifies 94% treated stabilisation and 3% snap-frozen stabilisation in accordance with the profiles in (C).

A similar approach is used to find snap-frozen stabilisation in Figure 3-35. However, no cross over was found as can be seen in (D) despite in (A and B) the sector Gs have a higher number of identified spots in comparison to Figure 3-34.

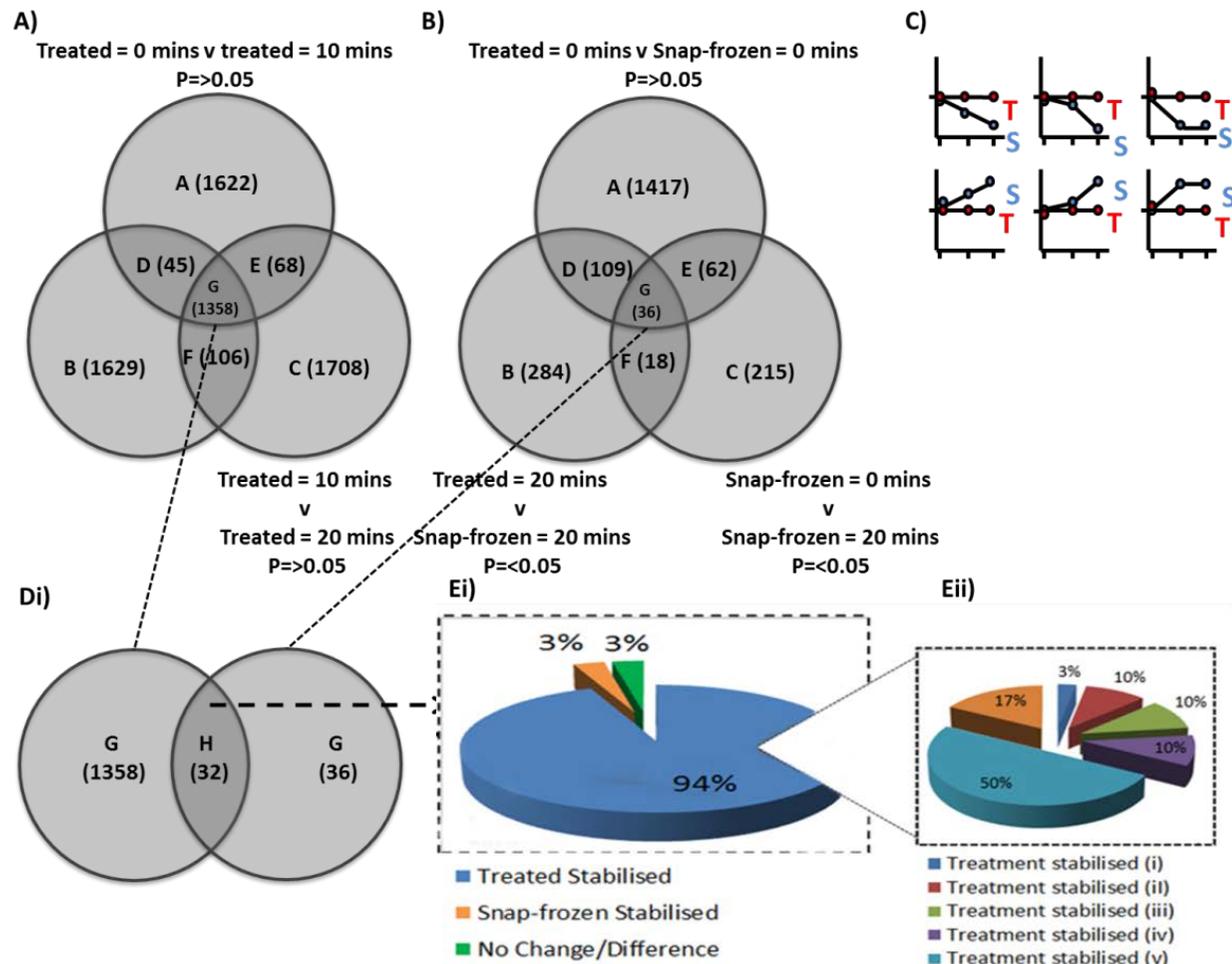


Figure 3-34: Venn analysis for the main investigation. A and B) give two independent but comparative strategies for matching profiles in C) searching for predicted profiles in sector G focusing on treated stabilisation. D) Gives crossover of the two G sectors. E(i-ii) Show profile distribution of protein spots taken from areas G Venn diagram.

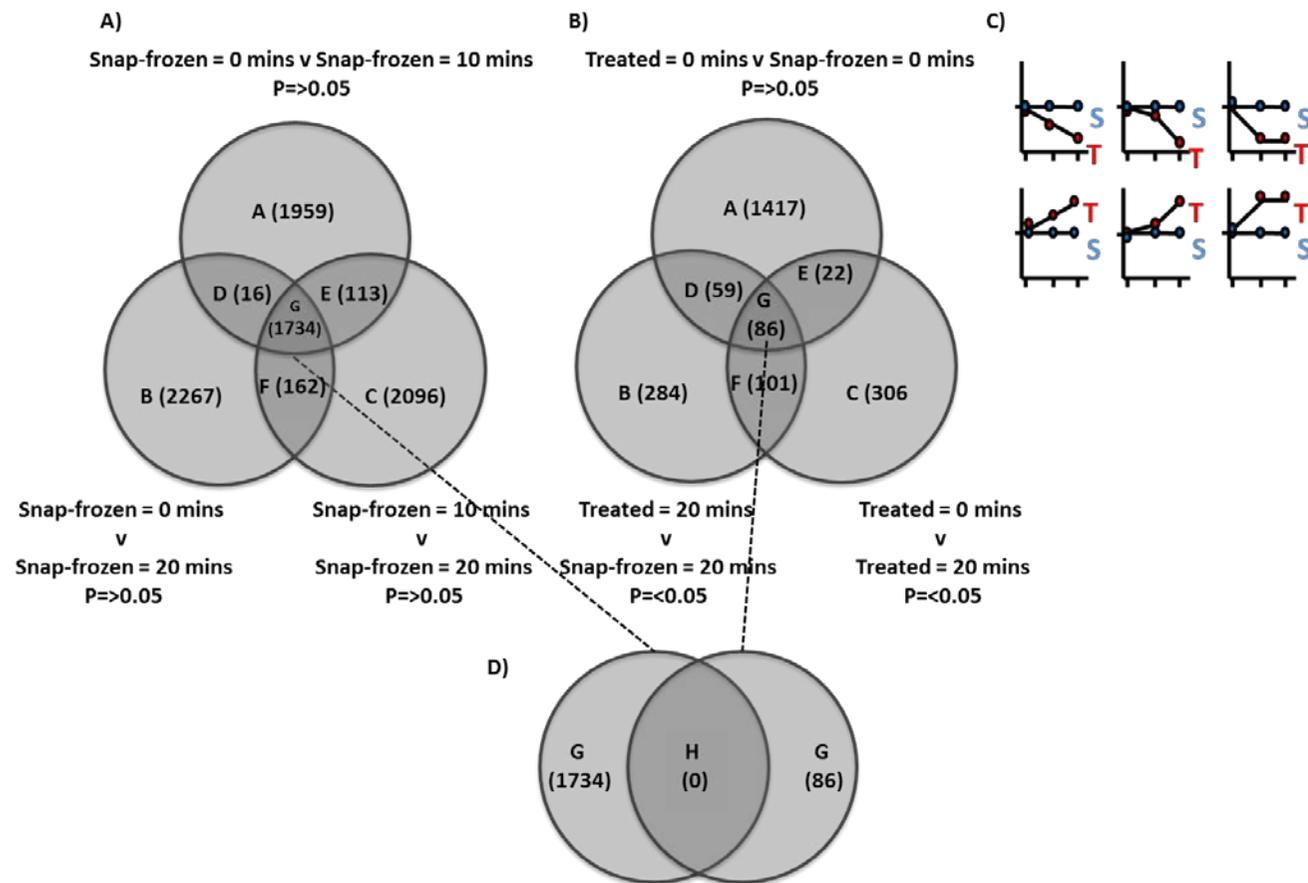


Figure 3-35: Venn analysis for the main investigation. A) and B) give two independent but comparative strategies for matching profiles in C), searching for predicted profiles in sector G focusing on snap-frozen stabilisation. D) Gives crossover of the two G sectors

3.6.5 Summary of Venn analysis for the main investigation

Overall, the analysis performed using Venn diagrams correlates well with the profile analysis, with similar proportions identified in the various experimental groups. Even using 2-way Venn diagrams analysis can be performed more inclusively by including more spots than using ANOVA sorting whilst reducing the number of profiles required to be viewed manually. This is displayed in Figure 3-36, where the number of spots included in (A-C) exceeds the ANOVA sorting performed of 623 and 573, but effectively reduced manual work substantially, with number of 158, 200, and 98 respectively.

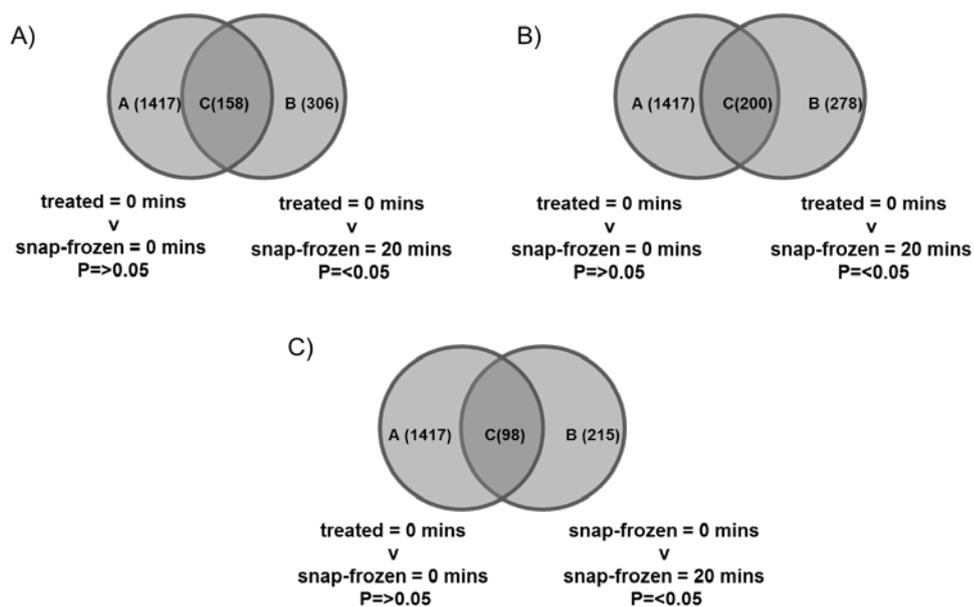


Figure 3-36: Venn analysis for the main investigation. Comparison of spot distribution between A), B) and C)

3.6.5.1 Summary of the main investigation

The main findings are summarised as follows:

- Not all degradation is rapid of onset and the use of a further time point shows how the stabilisation of markers as a percentage has increased in comparison to the pilot study. This shows that heat-treatment does aid stabilisation. It shows how the amount of time tissue can be handled at room temperature can be increased, if wild type mouse brain tissues are heat-treated. It also highlights the importance of keeping tissue chilled when preparing samples, particularly for non-heat-treated samples.
- The choice of sorting method is crucial. If one method is considered over another, valuable data could be lost. Shown in Figure 3-27. Only a small proportion of spot data is present in 4 different possible sorting methods. The necessity to deconvolute data, and be inclusive is hard to balance and should be considered on an individual basis for the study under consideration.
- The need for cross validation is critical. It is clear from the analysis that one method on its own is insufficient to draw conclusions on data. Using profile analysis alone would lead to false positive data as can be seen by comparing it to the highlighted cells in the tables above, showing contradictions with some of the statistical test which would not come out with ANOVA sorting.
- ANOVA sorting and profile analysis is a good method for scanning for prospective marker spots but should be validated with the correct statistical tests. Cross-validation with Venn diagrams are also useful.
- Venn analysis allows for a more inclusive strategy and eliminating some of the time issue associated with manual checking by using statistical tests. Therefore Venn analysis is an excellent strategy for visualising data relatively quickly and objectively.

3.6.6 Validative results

This section discusses the pertinent findings that reinforce the main findings of both the pilot and main investigation results outlined above.

3.6.6.1 Validation for including time point= 20 minutes.

With the addition of a further time point, it became apparent that data from the pilot investigation would be skewed for markers which degrade very rapidly within the first 10 minutes and not over an extended period. With retrospect analysis, it was observed that the addition of the further time point in the main investigation in comparison to the pilot investigation was validated. This is shown below in Figure 3-37, comparing spot maps from the pilot investigation and the main investigation. Markers that would be included in the no change category of pilot investigation are now shown to be included in treated stabilisation profiles for the main investigation. This is an interesting development and could possibly explain the increase in the percentage of degradation between snap-frozen and treated samples.

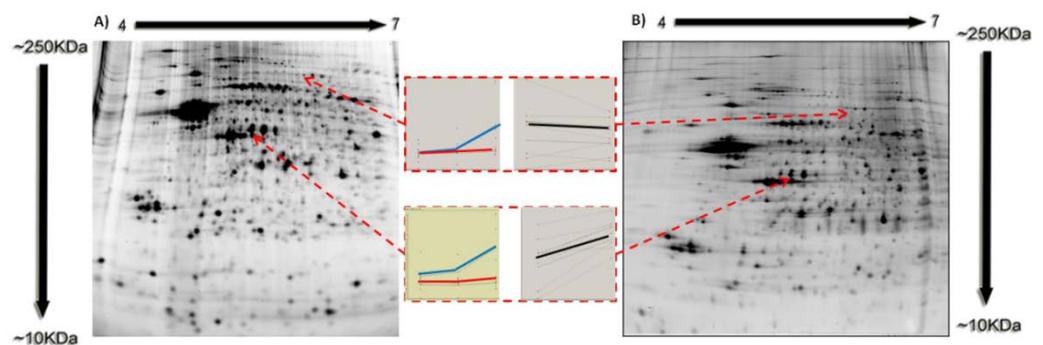


Figure 3-37: Spot maps A) main investigation and B) pilot investigation indicating two example spots that have been defined in pilot investigation as no change and in the main investigation as treated stabilised. The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance.

3.6.6.2 Analysis of presence vs. absence spots pilot investigation and main investigation.

An expected characteristic of degradation, if treatment had been successful, would be the appearance of different marker spots which appear on their own, without the treated group featuring them. Using traditional sorting methods for DiGE would not discriminate very accurately for such markers as so a different strategy was employed, as outline in section 3.5.2.8.

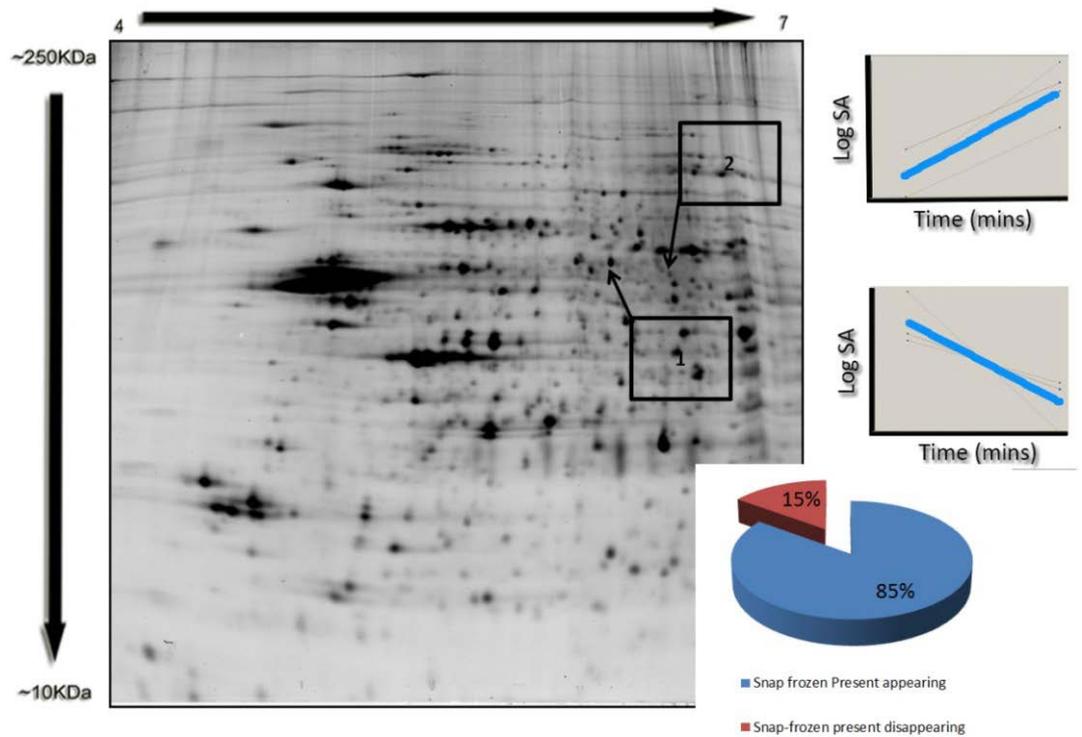


Figure 3-38: Examples and positions of gel map of protein appearing in 1 experimental group only (pilot investigation). Pie chart shows the distribution of these profiles with total number of 39 found. All were in snap-frozen = 0 min and 10 group. The blue lines denote snap-frozen samples.

In the pilot investigation this strategy discriminated very well and found 39 profiles in total with no average ratio but significant $p\text{-value} \leq 0.05$, where markers were seen to be appearing or disappearing for one experimental group only. As summarised in Figure 3-38 in this case 100% were intensity profiles of snap-frozen only, not containing any treated. 85% of which showed markers disappearing and 15% showing markers appearing over the time points from 0-10 in. This further validation in the main investigation summarised in Figure 3-39, where only snap-frozen markers were found too. In this case only 6 in total had no average ratio and had a significant 1-way ANOVA score. As can be seen by the spot maps in both Figure 3-38 and Figure 3-39, of the examples they were found high in the gel. Correlating the position and low numbers found there is a suggestion that most degradative markers are small and have simply run off the end of the gel. This is a limitation of DiGE in the small molecular mass area. The only identified spot from Figure 3-39 (1) is 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus].

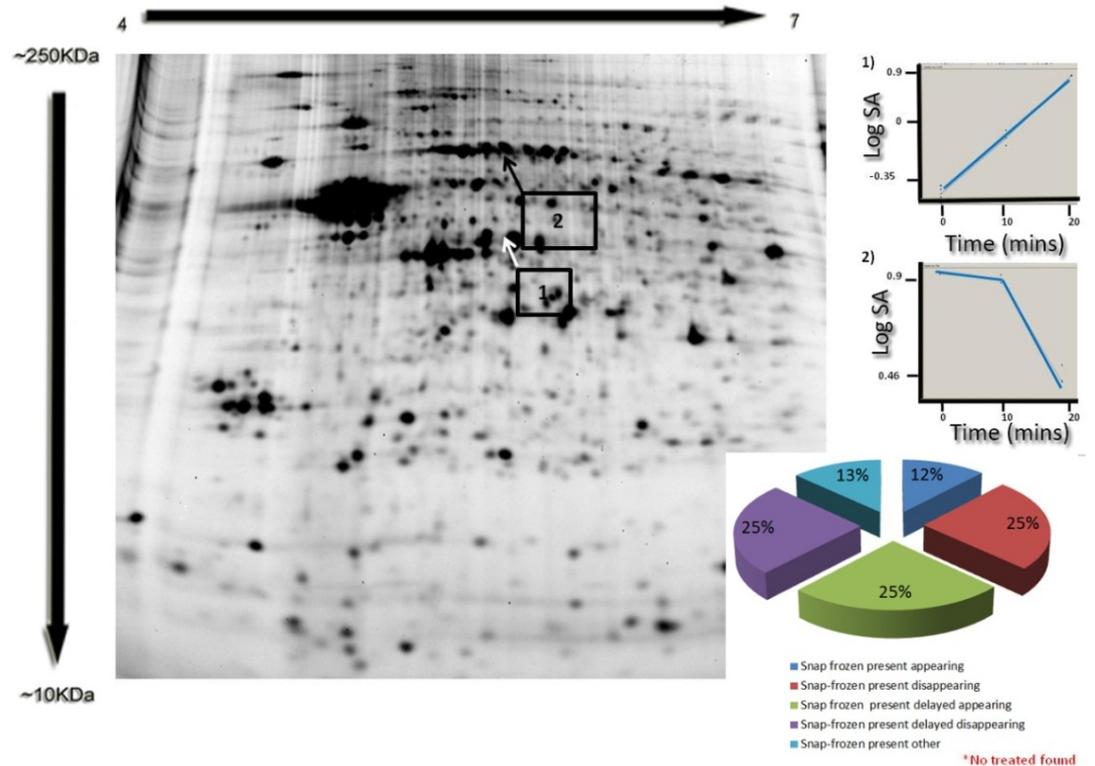


Figure 3-39: Examples and positions of gel map of protein appearing in 1 experimental group only in main investigation. Pie chart shows the distribution of these profiles with total number of 6 found. All were in snap-frozen = 0, 10 and 20 group. Profile number 1 has been identified as 5-hydroxytryptamine (serotonin) receptor 2C [Mus musculus]. The blue lines denoted snap-frozen samples.

3.6.6.3 Multiple identifications pilot investigation

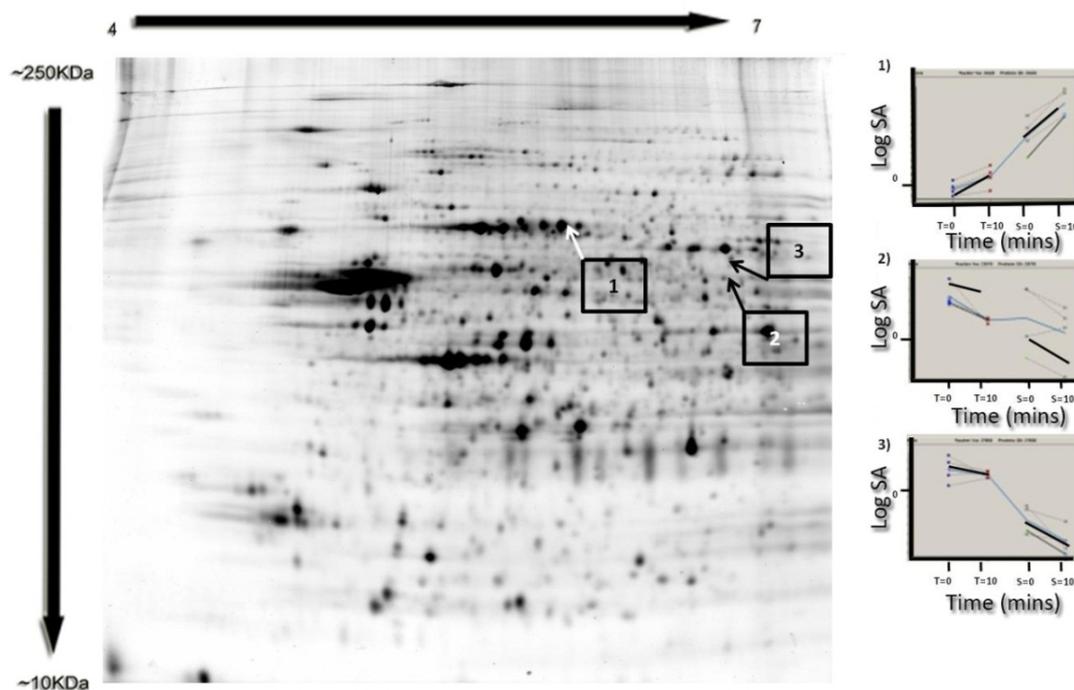


Figure 3-40: Gel map showing multiple identifications and spots shifts in pilot investigation of Dihydropyrimidinase-like 2 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

In Figure 3-40 where Dihydropyrimidinase-like 2 [Mus musculus] has been identified, it can be seen in the intensity profile (1) and (3) opposing profiles and in profile (2) no change. This would be suggestive of a degradation process, where there is a disappearing snap-frozen marker in (3), appearing and increasing markers in (1) and another site of no change (2). All locations of the gels are relatively close showing the marker is not necessarily changing in sequence.

In Figure 3-41 peroxiredoxin 6 has been identified, which is one of a family of many peroxiredoxin proteins. It is also exhibiting different intensity profiles (1) showing no change and (2) showing treated stabilisation.

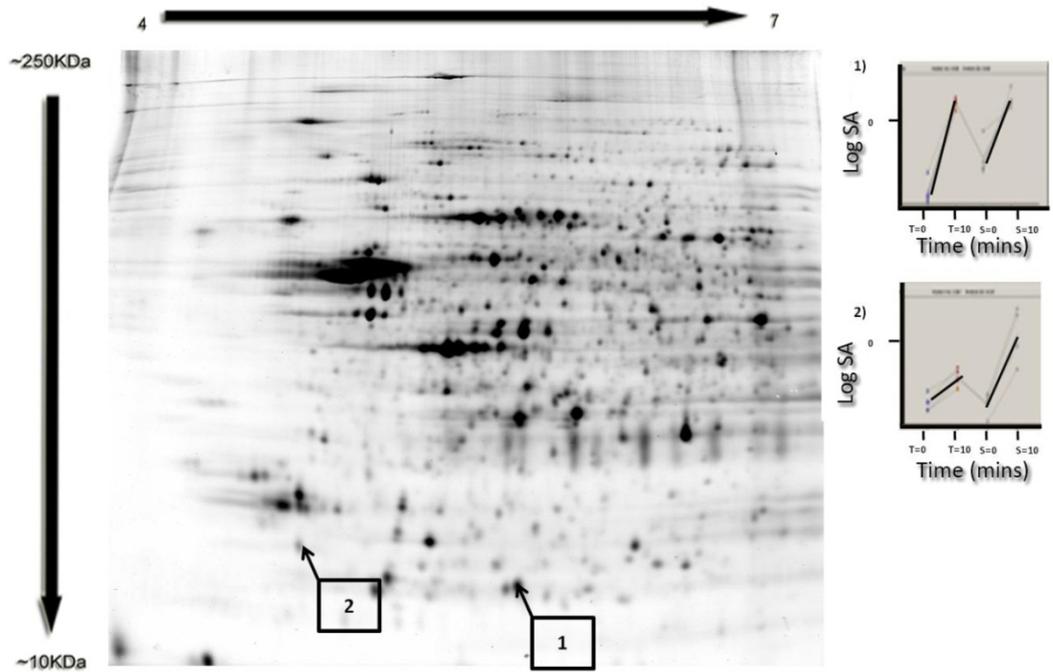


Figure 3-41: Gel map showing multiple identifications and spots shifts in pilot investigation of Peroxiredoxin 6 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

In Figure 3-42 a brain creatine kinase has been identified. From the gel map it can be seen that there has been only a small shift between positions (1 and 2). This potentially is showing the presence of small amounts of proteolytic activity or the modification of the markers sequence with PTMs. This is consistent with the intensity profiles (1 and 2).

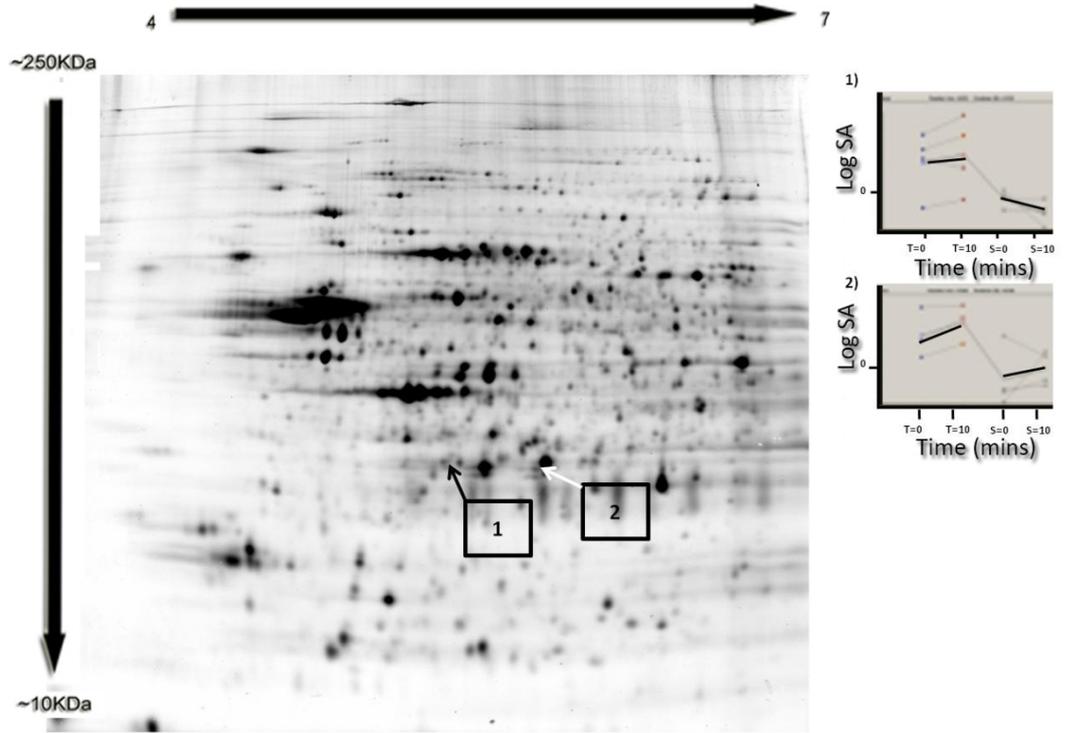


Figure 3-42: Gel map showing multiple identifications and spots shifts in pilot investigation of creatine kinase, brain [*Mus musculus*] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

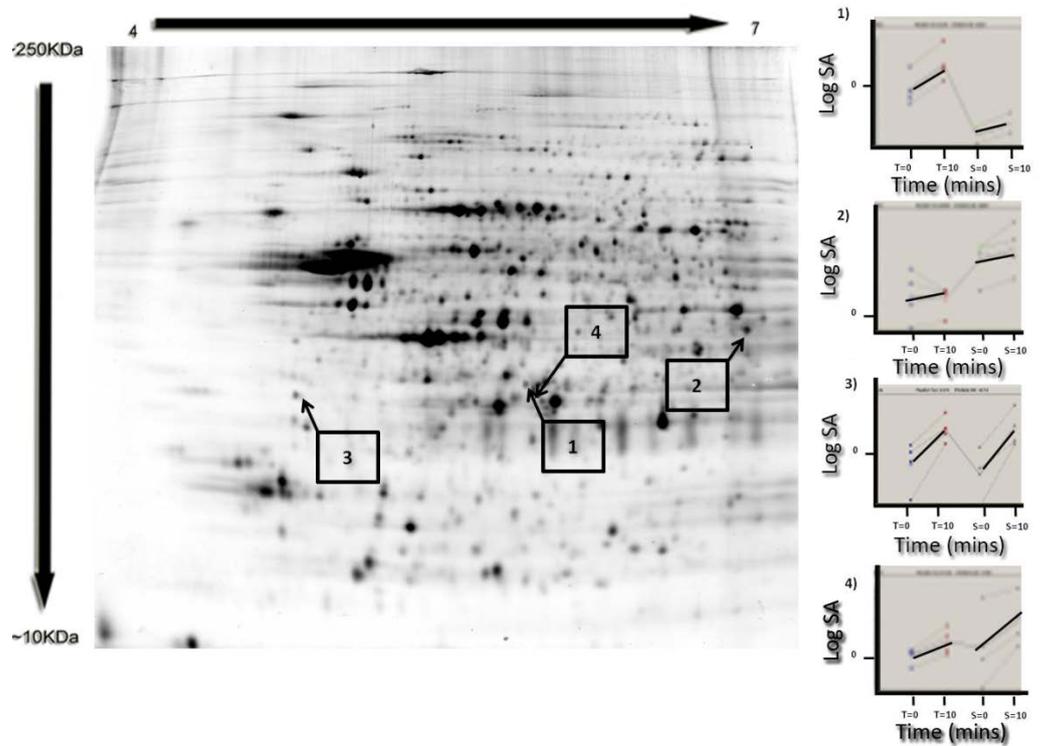


Figure 3-43: Gel map showing multiple identifications and spots shifts in pilot investigation of gamma-actin [*Mus musculus*] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

In Figure 3-43 four locations can be seen where gamma-actin has been identified. Gamma actin is a structural protein, which forms substantial elements of the cytoskeleton of a cell and is conserved in eukaryotes. It comes in many different isoforms, so is difficult to make too many inferences about it (Hennessey et al., 1993), this is equally true of a protein like tubulin, which has been identified in two different forms in multiple locations on the gel, shown in Figure 3-44 and Figure 3-45. Additionally, it has been identified in a portion of the gel known to contain high amounts of tubulin, it therefore again must be treated with trepidation, particularly as it is contained in a saturated portion of the gel.

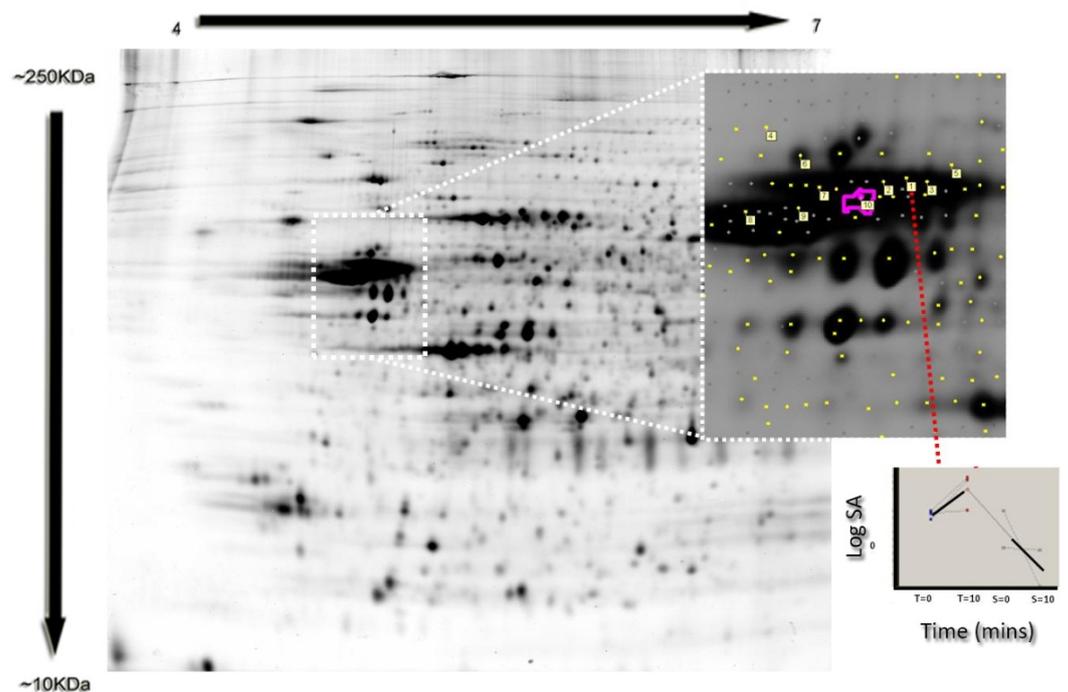


Figure 3-44: Gel map showing multiple identifications and spots shifts in pilot investigation of alpha-tubulin isotype M-alpha-2 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

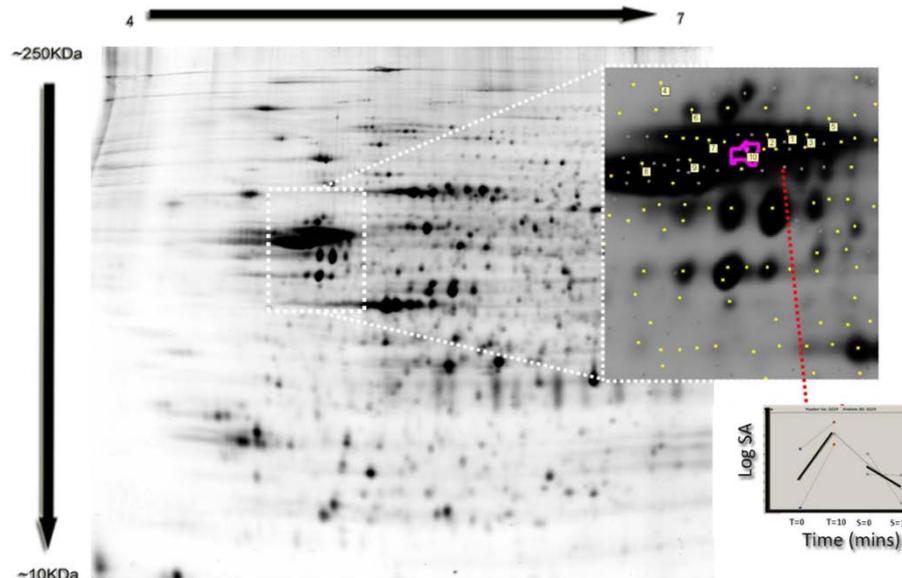


Figure 3-45: Gel map showing multiple identifications and spots shifts in pilot investigation of tubulin, beta [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

3.6.6.4 Multiple identifications main investigation.

Further multiply identified markers have been discovered in the main investigation as shown in Figure 3-46, Figure 3-47 and Figure 3-48. In Figure 3-46 an identification of a protein predicted: similar to pORF2 at six different locations on the gel map. Vertical and horizontal shifts are present. Intensity profiles (1, 2, 3, 4 and 6) matching profiles which indicate treated stabilisation. Further examples are given in Figure 3-47.

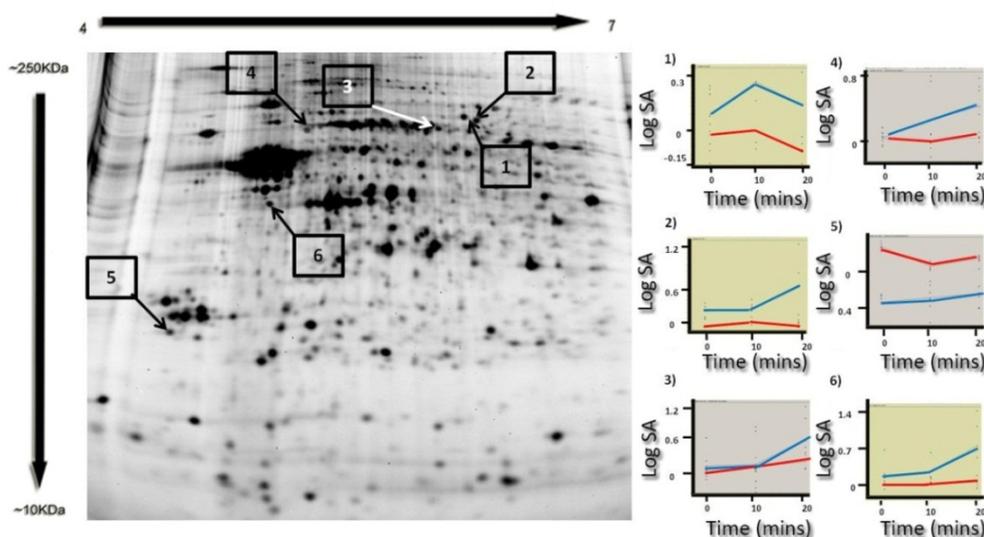


Figure 3-46: Gel map showing multiple identifications and spots shifts in the main investigation of PREDICTED: similar to pORF2 [Mus musculus], with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance

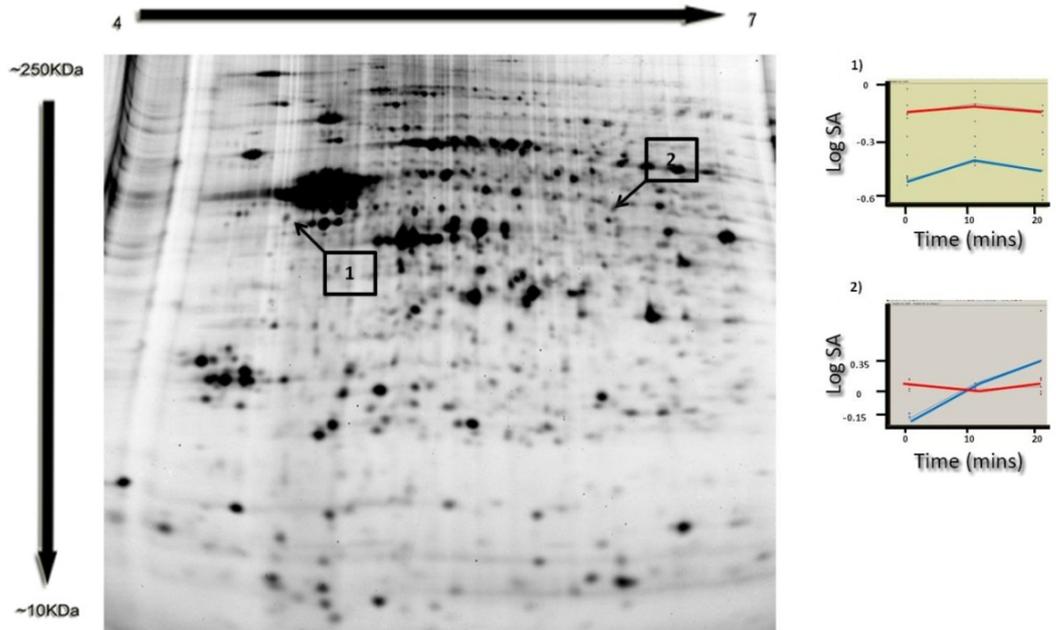


Figure 3-47: Gel map showing multiple identifications and spots shifts in the main investigation of ATP5b protein or Mitochondrial ATP synthase beta subunit [*Mus musculus*] with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance

Figure 3-48 demonstrates large vertical shifts indicting large size change in the identified marker, particularly between (1, 4, 3 and 2) interesting (4) appears in just the snap-frozen.

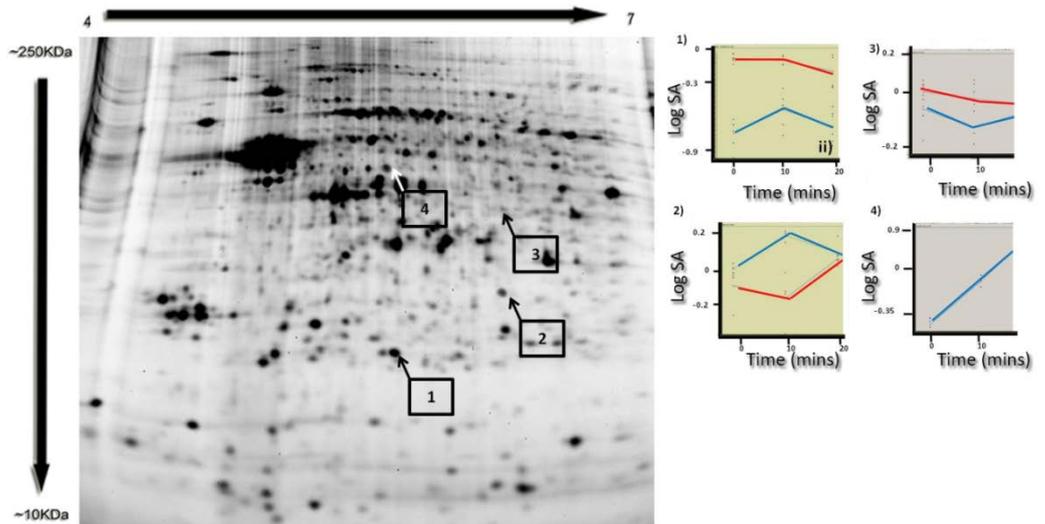


Figure 3-48: Gel map showing multiple identifications and spots shifts in the main investigation of 5-hydroxytryptamine (serotonin) receptor 2C [*Mus musculus*] with the corresponding profiles. Red line is heat-treated samples and blue line is snap-frozen samples. . The graphs have an x-axis have times point 0, 10 and 20 minutes and the blue lines show snap-frozen whilst red displays heat treated samples. The y-axis has the scale of log standardised abundance

3.6.6.5 Summary

The presence of multiple identification could be indicative of a degradative process. However, circumspection must be maintained, particularly when dealing with a homologous group of proteins. A further limitation of DiGE when trying to track degradation is with regard to mass. It could easily be the case that protein fragments are just running off the end of the gel. Thus it is important to have complementary methodologies such as the use of LC-MS and MSI to look at those markers in the smaller mass range. The search for spots of presence vs. absence was a time consuming process. However, the presence and absence of spots from treated and snap-frozen samples is strong evidence of degradative processes having taken place. Degradation studies would benefit from further development of this type of analysis.

3.6.6.6 Principal components analysis using EDA module for pilot investigation.

Principal component analysis is a robust statistical method which is useful for validating data by reducing the variables of a multidimensional space into a smaller number of dimensions, by correlating and grouping similarities. The validation is obtained by statistical analysis being performed without initially assigning groups until the data is displayed. The further apart experimental groups are, the larger the difference between those groups, with PC1 being the 1st principal component and therefore contains the largest differences followed by PC2. The PCA plots displayed in this thesis are called score plots. These show an overview of the spot maps and show similarities between groups. Another kind of plot, called a loading plot, can be used to visualise the difference between variables. The score plots of the spot maps are most useful in this case as it allows the visualisation of all groups compared to each other. This means, at a glance, it can be seen if the gels have run well and the extent of variation between groups. The closer the groups are together the more “alike” they are. PCA on pilot investigation is shown Figure 3-49. For both (A) and (B) all experimental groups can be clearly distinguished separately, meaning generally (all spot data complied) there is differences between all groups. The greatest differences are seen in PC1, meaning that the treated groups and snap-frozen groups have the greatest differences between them. This is equally true for PC2, were treated groups having close vertical proximity compared to the snap-frozen groups.

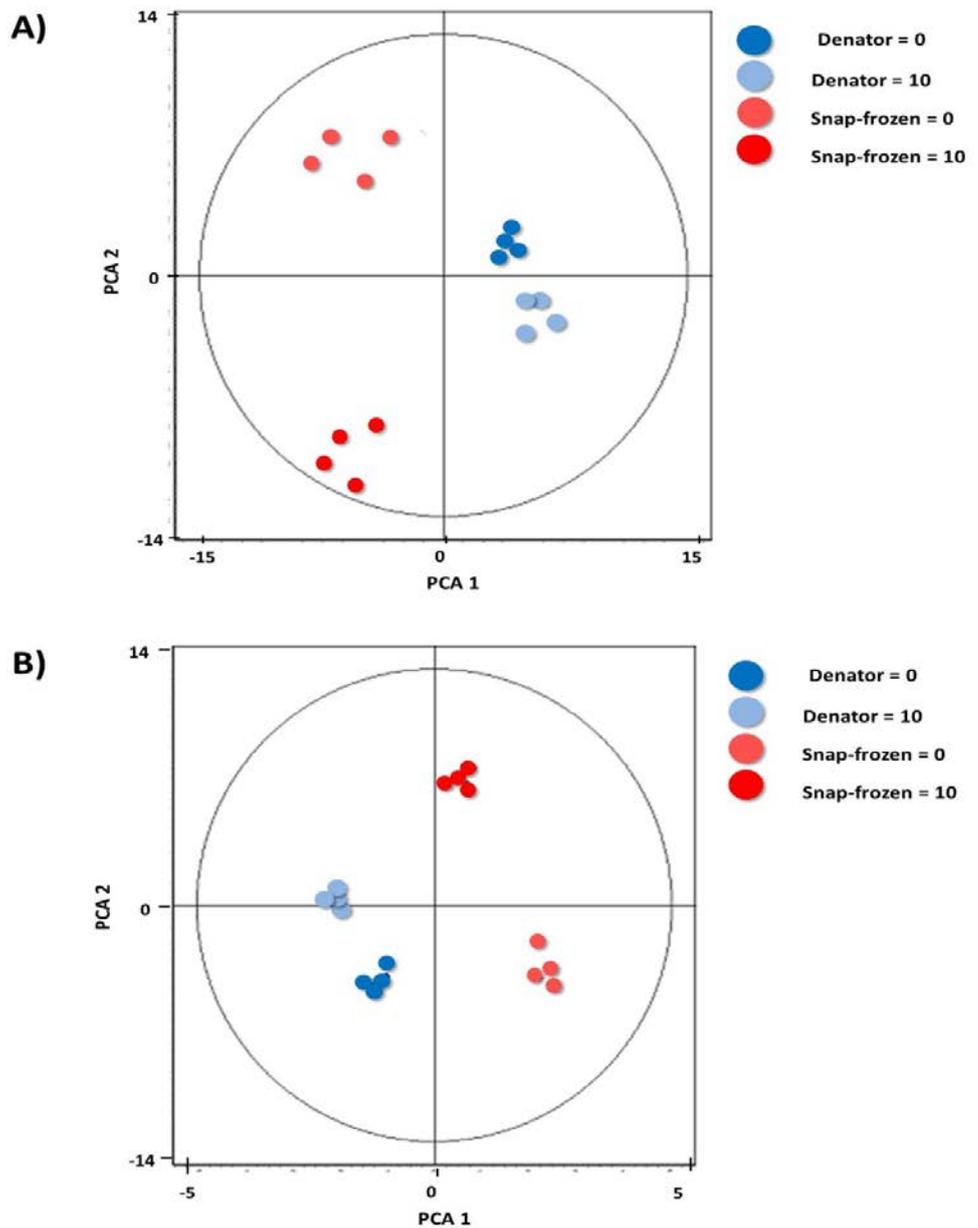


Figure 3-49: PCA score plots. A) Shows score plots including all detected spots for pilot investigation and B) Shows score plots including only 1-way ANOVA spots included

This reassures the relationship between treated groups being more closely related to each other than snap-frozen groups in both PC1 and PC2. This relationship is indicative of an association with treatment having a positive effect of degradative processes. As there is no discernable difference between the diagrams including all spots, Figure 3-49(A), in the PC analysis and just spots with 1-way ANOVA of $p \leq 0.05$ (B) it seems that in this case sorting the data by 1-way ANOVA does not have an adverse effect on the data set, but will however reduce the complexity of analysis by reducing the number of spots and moreover increasing the quality of spots data and the chance of seeing differences.

3.6.6.7 Principal components analysis using EDA module for the main investigation.

PCA on pilot investigation shown Figure 3-49, displayed reassuring validation regarding the separation of experimental groups as they were clearly distinguished. This is further validated for the main investigation displayed in Figure 3-50. The close relationship of the heat treated groups and snap-frozen group shows that the groups can be clearly defined from each other. Also, the denator group shows tight clustering, while the snap-frozen samples have a greater degree of separation from each other. This, very speculatively, is showing a greater difference between time points in the snap-frozen samples compared to heat-treated samples. Therefore, suggesting greater changes in the snap-frozen samples.

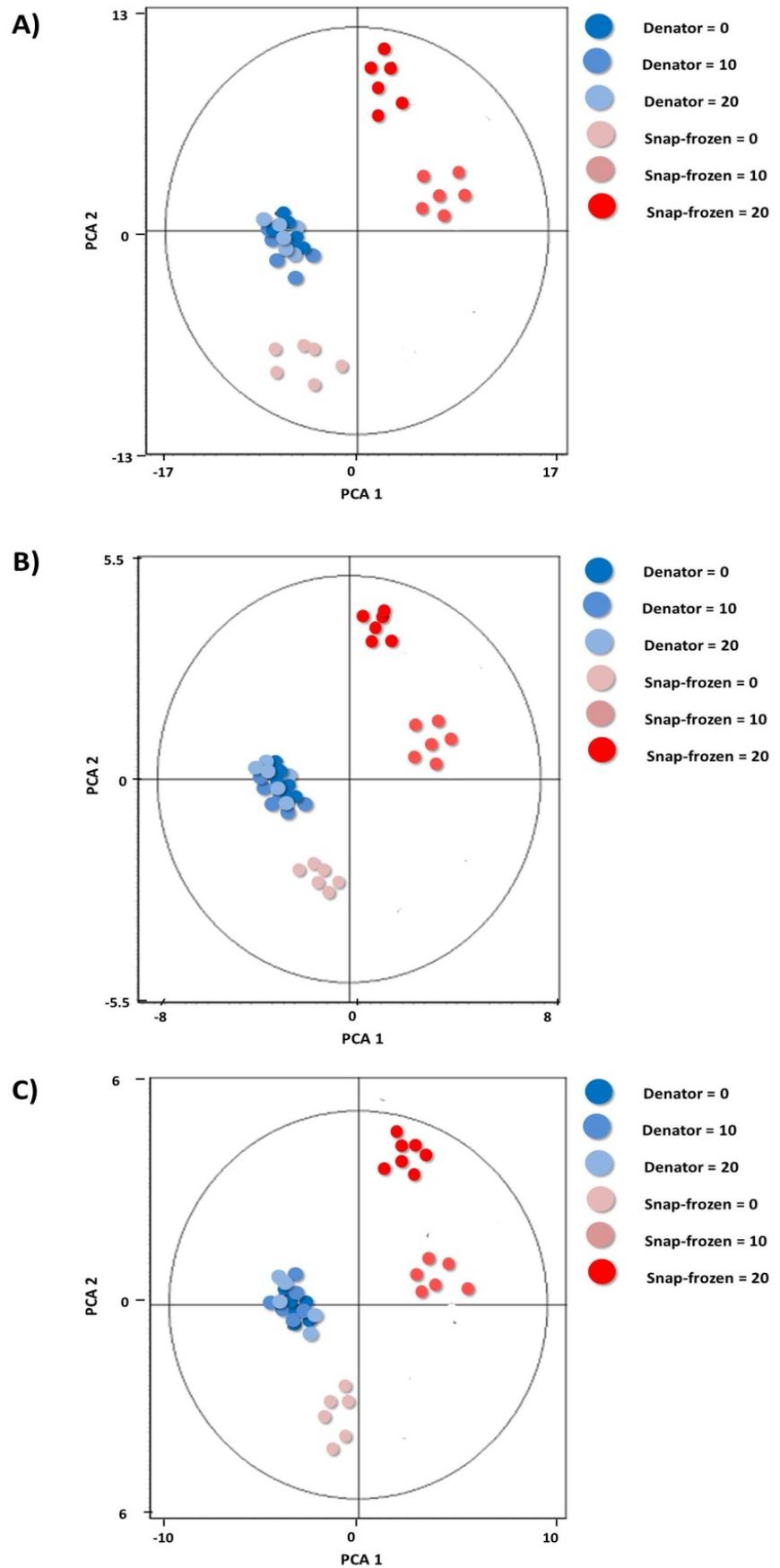


Figure 3-50: PCA score plots for the main investigation A) PCA score plots including all spots matched B) PCA score plots sorted by One-way ANOVA C) PCA score plots sorted by two-way ANOVA. As can be seen the different experimental groups can be clearly defined, with a greater variation in the snap-frozen samples.

PCA was performed on all spots, 1-way ANOVA and 2-way ANOVA respectively. All PC diagrams show consistent separation between experimental groups with treated groups showing definitive clustering whilst snap-frozen generally showing large degrees of separation. With the exception of snap-frozen = 10 and 20 minutes, in (B) and (C). There is therefore a suggestion that experimental groups snap-frozen = 10 and 20 minutes are more closely related, but interestingly snap-frozen = 0 is distinctly separated from all other experimental groups showing a similar relationship to snap-frozen = 10 and 20 minutes in PCA but a large divergence in PC2. When sorting the data using ANOVAs in (B) and (C), compared to including all spots in (A) it has little effect on the treated experimental groups but brings the relationship of snap-frozen = 10 and 20 minutes closer together, alluding to the fact that sorting the data using the statistical methods of 1-way and 2-way ANOVA shows you have a less varied and more reproducible set of data. This could be explained by the fact that, by using spots that have had a significant ANOVA it generally is removing any spots which are mismatched or outliers (i.e. using a more robust data set).

3.6.6.8 Western blot analysis

Western blotting was performed on whole tissue lysate preparation that had been used for the DiGE main investigation. The antibody used was anti-anti-peroxiredoxin 6 which was one of the multiply identified spots in the pilot investigations. The western blot would also have been performed on the samples from the pilot investigation, however, there was insufficient sample remaining. This was performed 3 times.

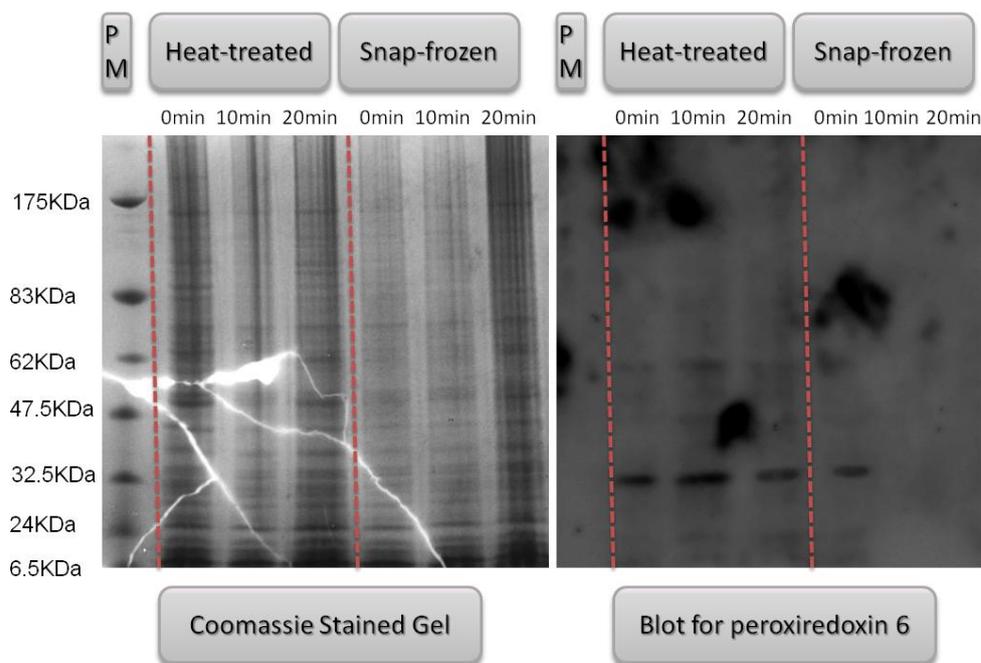


Figure 3-51: An example of one of the three repeats of 1-D Coomassie stain gel and western blot of heat-treated and snap-frozen samples blotted for peroxiredoxin 6. As can be seen heat-treated 0, 10 and 20 minutes and snap-frozen 0 minutes show the presence of a strong band at approximately between 24-32.5KDa and snap-frozen 10 and 20 minutes the band is absent. Note: the PM labels refer only to the Coomassie stain gel.

As can be seen in Figure 3-51, peroxiredoxin 6 has shown up in the western blot in treated = 0, 10 and 20 minutes and snap-frozen = 0 minutes. This is strong evidence when correlated with the predicted profiles of what is likely to occur in degradative process and that treatment has halted proteomic degradation in the case of peroxiredoxin 6 for the treated samples. Additionally, due to the high amount of homogeneity in the peroxiredoxin family, the antibody selected was a monoclonal antibody, so it is specific to only one epitope of peroxiredoxin 6. This allowed for only the intact version of the protein to be revealed and the fragments would run off the end of the gel. It should be pointed out that 1D western blot analysis gives no information regarding the different isoforms that exist of a protein. This is because isoforms of a protein do not vary greatly in molecular weight (as opposed to

slice variants), however can have different pIs and therefore they are revealed using 2D analysis.

A densitometry analysis was performed using Image J 1.46r software, National Institute of Health, USA. A set area was drawn around each band and including a portion of background. A curve was plotting using Image Js gel plot function. The area under the curves were found and averages taken as a relative comparison of the bands fold change by dividing the results by the first lane (i.e. giving a relative value compared to the denator=0 and snap frozen=0 lane). The results are shown in

Table 3-21 and Figure 3-52.

Sample (n=3)	Mean fold change relative to time point zero	integrated area under curve (repeat 1)	integrated area under curve (repeat 2)	integrated area under curve (repeat 3)	Mean integrated area under curve	Standard deviation	Standard deviation of fold change
Denator = 0	1.0	7816	7709	7789	7771	56	0.01
Denator = 10	1.7	12789	12981	12908	12893	97	0.01
Denator = 20	1.3	10087	10123	10001	10070	63	0.01
Snap-frozen = 0	1.0	6309	6259	6498	6355	126	0.02
Snap-frozen = 10	0.0	0	0	0	0	0	0.00
Snap-frozen =20	0.0	0	0	0	0	0	0.00

Table 3-21: The mean relative intensities of western blot for Peroxiredoxin 6. Integrated area is in arbitrary units. The fold change was relatively ascertained by indexing again time point 0 min for both the denator heat treated samples and snap-frozen samples. n=3.

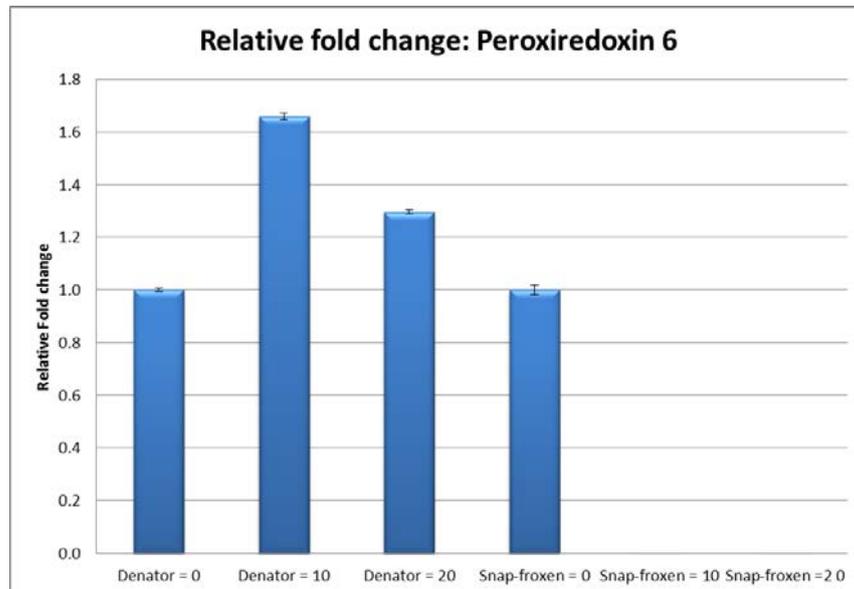


Figure 3-52: Graph displaying relative fold change of Peroxiredoxin 6 generated from Table 3-21. The error bars show +/- 1 standard deviation. n=3.

The spots showing the Peroxiredoxin 6 identification on the 2D-DiGE gel are shown in Figure 3-53.

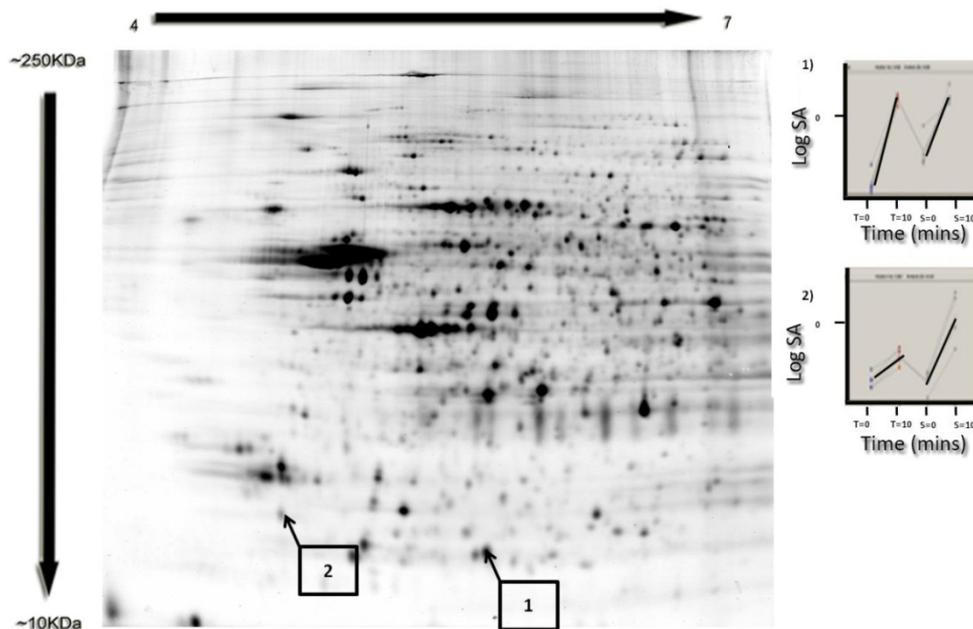


Figure 3-53: A repeat of Figure 3-54: Gel map showing multiple identifications and spots shifts in pilot investigation of Peroxiredoxin 6 [Mus musculus] with the corresponding profiles. The x-axis has time points treated 0 and 10 minutes and snap-frozen 1 and 10 minutes respectively.

As can be seen from both the 1D and 2D gels the molecular weights approximately match the quoted molecular weight of Peroxiredoxin 6 of 24,871Da (<http://www.uniprot.org/uniprot/O08709>). The protein spot labelled 2 can be seen to be heat-treated stabilized from the profile in comparison to the snap-frozen tissue.

3.6.7 LC-MS with Label free quantitation

A complementary quantitative strategy for looking at degradation is the use of label-free quantitation using Mass Spectrometry. This work was performed by Dr Heather Allingham, (a Ph.D Student at Glasgow University at the time). The project was in collaboration with her and the same samples were used.

<10kDa Intact Sample Results

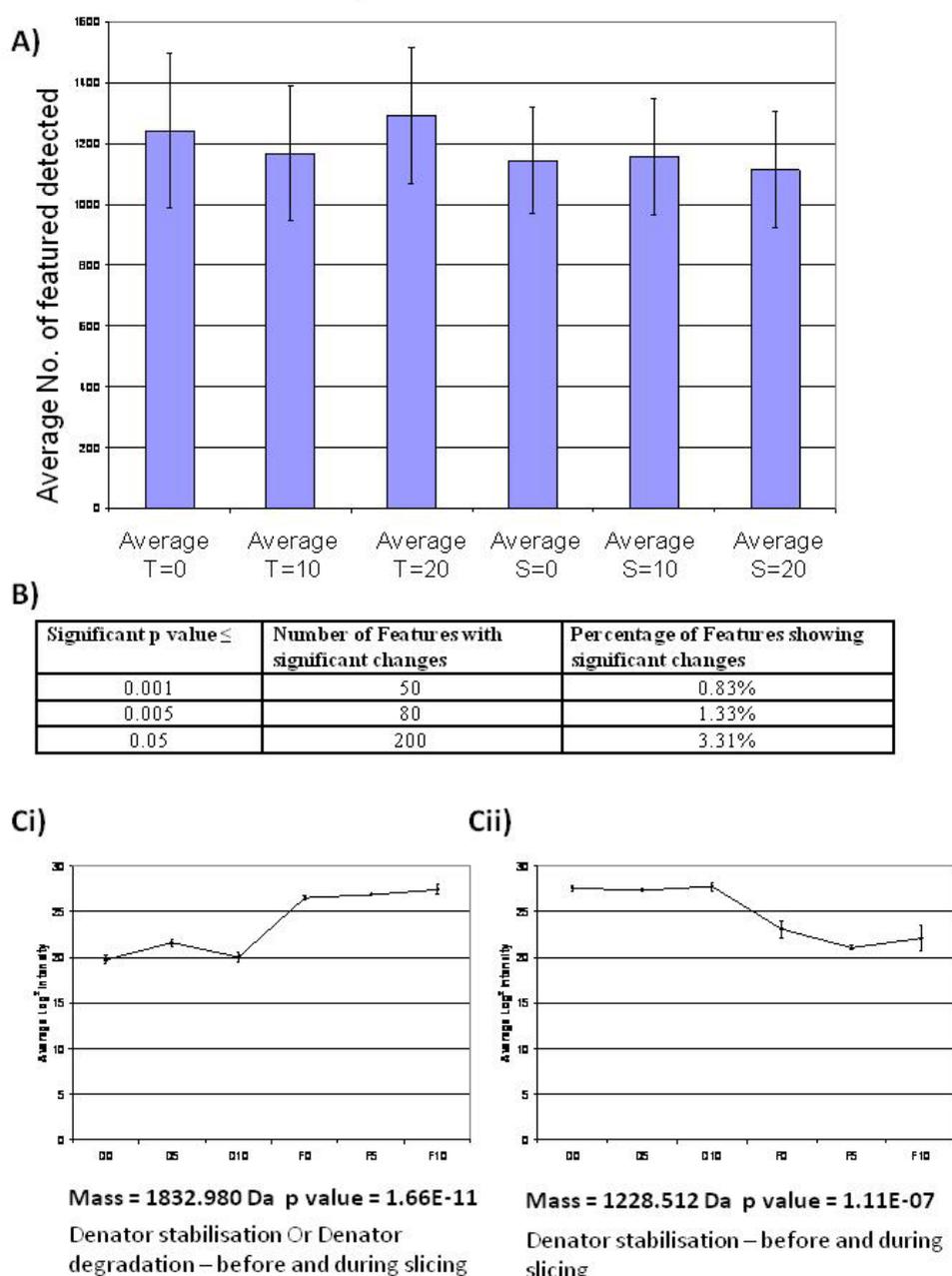


Figure 3-55: LC-MS using label free quantitation: performed by Miss H. Allingham, Ph.D Student's, University of Glasgow. A) Show the average number of identified features from intact sample results in T=0, 10 and 20 minutes and S=0, 10 and 20 minutes. The number of features across both groups is comparable. B) Give the percentage of features showing a significant change. Ci-ii) shows a profile consistent with the profiles observed in the two DiGE experiments in sections 3.6.3 and 3.6.4.

<10kDa Digested Sample Results

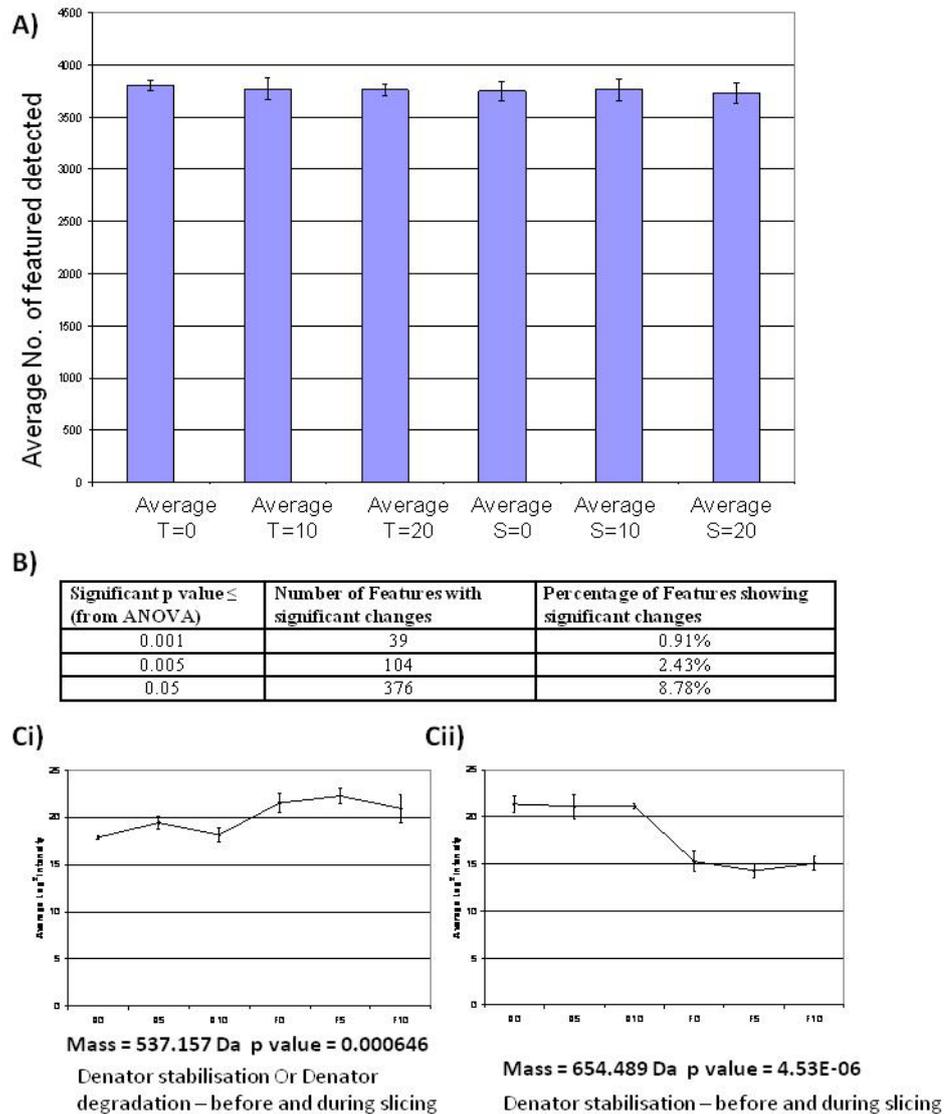


Figure 3-56: LC-MS using label free quantitation: performed by Miss H. Allingham, Ph.D student, University of Glasgow. A) Show the average number of identified features from digested sample results in T=0, 10 and 20 minutes and S=0, 10 and 20 minutes. The number of features across both groups is comparable. B) Give the percentage of features showing a significant change. Ci-ii) Shows a profile consistent with the profiles observed in the two DiGE experiments in sections 3.6.3 and 3.6.4.

Figure 3-55 and Figure 3-56 show intact and digested treated and snap-frozen samples. As can be seen in both MS experiments a similar number of features were detected in all groups. A collaboration of analytical techniques were employed regarding the use of profile analysis. The profile changes observed in sections Ci and Cii of both figures are similar to those seen in the pilot investigation and helps as a quid pro quo validation of the concept. Interestingly Dr Allingham discovered that in the lower sub 10kDa fraction the heat treatment only had an effect of stabilisation of about 1% of total identified feature a contrast to the larger intact proteins of the DiGE investigation.

3.6.8 Mass Spectrometry Imaging

Following treatment with the Stabilzor T1 denaturing device (Denator AB, Gothenburg, Sweden) mouse brain tissue maintained its overall integrity with some small noticeable morphological changes in terms of slight darkening and rigid texture of the tissue, but commonly noticeable structures are still apparent. This can be seen in Figure 3-58(F). Consequentially, with a rigid texture sectioning was made easier with a reduced level of cracking, but the appearance of small holes and increased hindrance in thaw mounting was apparent and noted (by Dr R. J. A. Goodwin). These structural changes had an adverse effect on the amount of MS data attainable.

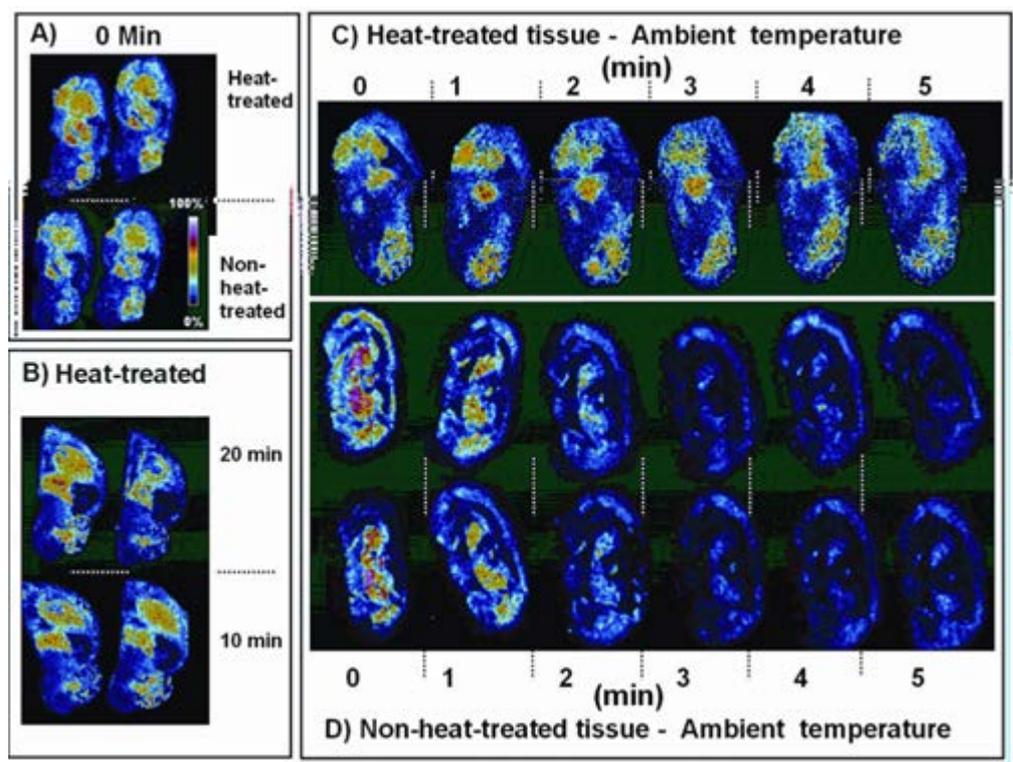


Figure 3-57: MSI experiment, Performed by Dr R. J. A. Goodwin, Research Assistant, University of Glasgow. A) Treated (heat-treated) and snap-frozen (non-heat-treated) sections at 0 minutes for an example marker. B) Treated (heat-treated) and snap-frozen (non-heat-treated) sections at 20 and 10 minutes respectively for the same example marker in (A). C) Shows time course series of sections warmed consecutively from 0-5 minutes for treated and D) Snap-frozen tissue have been duplicated. Difficulties in cutting treated samples precluded duplicates.

Despite the reduction in the amount of MS data obtained, it has yielded some very intriguing data. A strategy to determine the effectiveness rapid heat-treatment was to watch the evolution of markers accumulate over a time course of treated and snap-frozen tissue. Consecutive slices were allowed to warm at 1 minutes intervals for 5 minutes this can be seen in Figure 3-57 (C) and (D). Difficulties in sectioning treated tissue precluding duplicate sections (Figure

3-57 (C)) but was performed in snap-frozen tissue samples (Figure 3-57 (D)). Figure 3-57 (A) shows a spatial distribution of marker for 6723.5 m/z across both the treated and snap-frozen at 0 minutes (i.e. not warmed to room temperature). Furthermore, the same marker is seen from different biological replicates with a similar distribution, in treated and = 10 and 20 minutes suggesting little to no degradation (Figure 3-57 (B)). This marker's evolution was tracked in Figure 3-57C and D over a 5 minute time course and can be seen to disappear rapidly in the snap-frozen samples (Figure 3-57 (D)) but maintained in the treated samples (Figure 3-57 (C)) suggesting degradation has been abated or even prevented.

Gaining information off-tissue is increased by performing tryptic digestion on-tissue to increase the number of fragments that fall into the effective dynamic mass range (≤ 30 KDa) of MSI. In Figure 3-58 treated and snap-frozen (non-treated) tissue has been sprayed with a trypsin solution of 50% methanol in 20 mM ammonium bicarbonate, using a standard TLC sprayer and on-tissue digestion performed to assess the quantity of data obtainable from the different experimental groups. A blank solution containing no trypsin was used on the snap-frozen and treated samples. This allowed a comparison of auto digestion caused by degradation and on-tissue tryptic digestion to be assessed. In Figure 3-58 (A-G) marker displays a variety of activity. At this point is important to be aware that by wetting tissue to apply trypsin or indeed the blank, this can cause the reactivation or reconstitution of any native enzyme action This makes on-tissue digestion and MSI particularly difficult, particularly in fresh tissue as opposed to formalin fixed which suffer less of such problems, although it is unlikely to eclipse the signal from tryptically digested proteins. Figure 3-58 (G) where all 4 sections have been sprayed with trypsin, demonstrates the need for circumspect on these grounds. There is a marker 842.5m/z, a known auto-digested tryptic peak. It has been detected in the treated section but not in the snap-frozen section, where an alternative process has occurred. An identification of the tryptic peak 842.5m/z was obtained by MS/MS and the spectra are shown in (H).

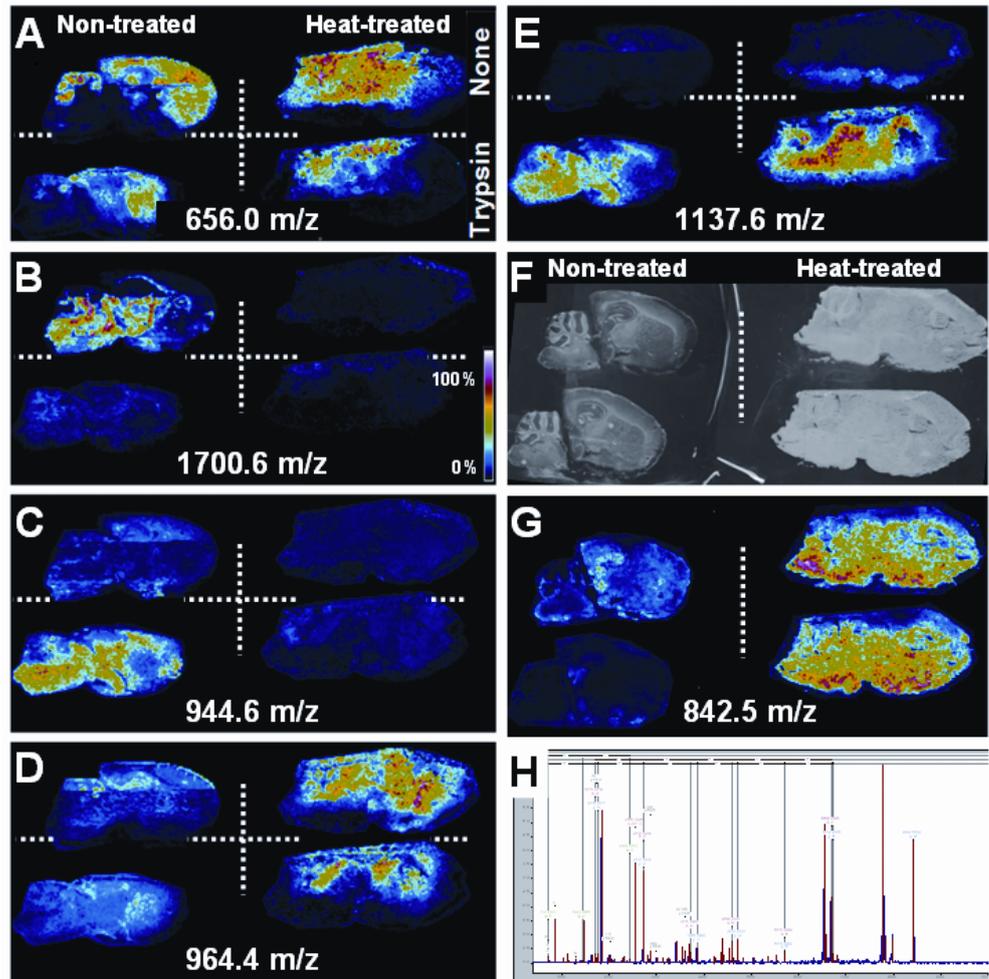


Figure 3-58: MSI experiment, Performed by Dr R. J. A. Goodwin, Research Assistant, and University of Glasgow. A-G) Shows snap frozen and treated sections typically digested and non-digested for different marker masses. F) Shows the optical images of those slices

In Figure 3-58 (A) the marker is not showing signs of tryptic activity in either of the experimental groups. In (B) expression of this particular marker is intense in the snap-frozen section and therefore possibly stabilised in the treated section where intensity is extremely low. There is also the suggestion of endogenous enzyme activity as the marker is not detected in the tryptically digested slice. This is in contrast to (C) where enzyme activity has not been prevented in the treated sections and therefore there appears to be no association with tryptic activity. In (D) the marker displays an unusual pattern as it appears to be associated both with snap-frozen and tryptic digestion. A possible explanation comes from a combination of endogenous enzyme activity occurring in the snap-frozen section and the tryptic activity occurring in the products, leaving the treated sections unaffected. An example of tryptic activity is associated with (E), where both the sections that were sprayed are affected, but the non-sprayed sections are not, and nor has any auto-digestion been prevented.

It is clear that MSI has demonstrated that degradation can be rapid and significant within a 5 minute time interval with markers degrading within as little as 2-3minutes. Meaning the need for a process to slow or halt degradation would be valuable alongside good rigorous lab practice. The treatment using the Stabilzor T1 denaturing device (Denator AB, Gothenburg, Sweden) has a significant effect at reducing proteomic degradation of several markers, and still gives spatial information for which MSI is known to provide. Further to this, enzymatic digestion can and does increase the level of markers obtainable and exponentially increase the usefulness of the information by providing the possibility of gaining off-tissue identifications. However, it should not be forgotten that treatment has a negative effect on the quality of tissue and the degree of information that was obtainable from it, further enzyme treatment can be used to reduce this effect and increase the number of detectable features.

3.7 Summary and Conclusions

In this chapter two broad areas were considered:

- The effect of rapid heat treatment on proteomic degradation in wild type mouse in comparison to snap-freezing alone
- The use of a multifaceted approach to using DiGE data in the assessment of proteomic degradation.

It has been shown that heat treatment using the Stabilzor T1 denaturing device immediately post excision does appear to reduce the percentage of nonspecific degradation of the proteome of wild type mouse brain tissue and therefore has the potential of being used to help in the sample handling and preservation of tissue for quantitative proteomic analysis. This preservation would allow for procedures at room temperature to be performed without the need for cooling. This would have significant impact in the laboratory for performing procedure which cannot be cooled, or where cooling hinders the operator's performance and progression, such as dissection. From the results it can be seen that both rapid degradation occurs post excision which is consistent to literature (Fountoulakis, 2001), but it was also seen how markers can continue to degrade over time, when exposed to warming to room temperature.

The usefulness of data could be significant in obtaining higher quality, closer to life information, however it is also noted that snap-freezing has been the standard choice for much of the tissue work in research and therefore holds a wealth of information that could not be used as a direct comparison if heat-treatment was performed. The MALDIMS1 help to validate and added to the DiGE finding that degradation was rapid and occurred during a 5 minute time course. It is however also noted that the heat-treatment also had a detrimental effect of the visual appearance of the tissue. This morphological effect has not been assessed, however, with form and function of central importance *in vivo* it is unlikely a visual change is not associated with a biological change. Having said that, the MSI should show an increased amount of spectral data collected in comparison to snap-freezing alone.

In terms of the analytical method used: Just as in a proteomic workflow there is no "silver bullet" to tackle all issues arising; the analytical strategy is not

different. The use of multiple and complementary analytical methods to mine data and to view the range of data was as a whole successful. Using profile analysis combined with Venn analysis allowed for cross-validation and finding spots otherwise missed by using ANOVA sorting alone. Venn analysis also reduced the need for manual checking, which plagues DiGE operators and take inordinate amounts of time, despite a well-developed and thought out standard workflow. It is, of course, essential to still use manual checking and statistical significance to support any markers found as profile analysis alone does allow some false positives to occur. The use of the EDA module and PCA analysis is a useful statistical validative process, however if biochemical significance is to be proven then western blots are required. Using this form of analysis however can reduce the amount of costly and time consuming western blot experiments to a necessary minimum.

4 Biomarker discovery and the assessment of variation in the proteomic profiles of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model.

4.1 Aims

To assess and investigate the variation of proteomic profile changes which occur in kidney tissue in hypertension using a Wistar Kyoto rat (WKY), congenic and spontaneously hypertensive rat-stroke prone model (SHRSP) with a view for future work as a strategy of quantitative trait loci. Specifically the aims of the experiments detailed in this chapter are:

- To assess proteomic changes in
 - Salt stressed and non-salt stressed rat kidney tissue between WKY, Congenic and SHRSP rats.
 - Cortex and Medulla regions of kidney tissue between WKY, Congenic and SHRSP rats.
- To assess macro dissection methods as a means of separating cortex and medulla regions.

To find potential examples of candidate proteins, peptide or biomarkers in hypertension potentially related to the region of chromosome 2 using proteomic data. Some of the approaches used in chapter 3 will be employed.

4.2 Structure of the chapter

Once again it is prudent to lay out a structure to help with the clarity of the chapter. This chapter has been broken up from one large chapter to 5 sub chapters. The following allows for a point of reference to aid the reading of this investigation.

4.5 Rationale of analysis

This outlines the methods taken to analyse the data giving an overview of what has guided the approach to analysing the DiGE gels, the profile analysis which was undertaken and the logic used to construct the targeted profiles.

4.6 Results and discussion

The results section is split into 3 subsections:

4.6.1 Pilot study

The results of the pilot study help to inform and guide the strategies used for the main investigation.

4.6.2 Main investigation

This section extends the investigation from the pilot study. It looks more deeply into the use of profile analysis, Venn analysis and the benefit of the additional time point.

4.6.3 Validative results

The validative section corroborates the results discussed in the main investigation and pilot study.

4.7 Summary and conclusion

4.3 Introduction

The maintenance and supply of blood to the heart, brain and major organs in the body is essential in order to deliver oxygen, nutrients and hormones at the required levels and rates. Therefore, systemic blood pressure is of great importance and central to the maintenance of the crucial life process of respiration. The control of systemic blood pressure therefore is crucial. In order to achieve this, there is a set of complex physiological mechanisms that are fundamental to the control of mean arterial blood pressure in both the long and short term. As with many areas in biology, our understanding of normative processes comes from looking at pathological problems. Understanding these mechanisms in hypertension and identifying possible markers would aid understanding of the control mechanism and give possible diagnostic or even therapeutic targets for further research to study.

4.3.1 Epidemiology of hypertension and control of blood pressure

Systemic hypertension is of ever increasing interest, particularly in western society, due to the large and increasing numbers of affected people (Burt et al., 1995, Ostchega et al., 2007, Berglund et al., 1976). This increase in the number of affected people creates a substantial burden on healthcare services worldwide (Monica et al., 2009). Hypertension affects 1/6th of the world population and 25% of adults (Kearney et al., 2005, Cheung et al., 2006) and is a major risk factor in a number of cardiovascular diseases from heart disease and atherosclerosis, Type II diabetes (Whelton, 2009), renal disease and stroke, to mention but a few (Kannel, 1996). It is therefore of intrinsic interest to understand the factors and mechanisms that affect the control of hypertension and identify and indicate biomarkers for early intervention. The issue of hypertension is predicted to be on the increase and cardiovascular disease is predicted to be the most common cause of death in the developed world by 2025 with hypertension being a common cause and symptom (Kearney et al., 2005).

Systemic hypertension is a quantifiable phenotype with a continuous range throughout populations but it has been clinically defined by a number of sources including the world health organisation as a systolic pressure in excess of 140mmHg and/or a diastolic pressure of over 90mmHg. With a normal acceptable range of between >90-119/60-790mmHg (Chobanian et al., 2003a, Elliott, 2007, Mancia, 2007).

In recent years there has been an increasing understanding of the risk factors that are associated with hypertension and cardiovascular disease. However, receiving a diagnosis of systemic hypertension does not exactly pinpoint the epidemiology and explain what physiological processes have been affected meaning treatments are difficult to administer. This has been characterised by the fact that 95% of hypertension is idiopathic or essential hypertension (Carretero and Oparil, 2000), with only 5% defined as secondary hypertension caused by various issues such as vascular, hormonal or known genetic diseases. Factors that influence essential hypertension have been studied and reviewed from numerous sources and include both environmental and genetic causes and risk factors such as; diet (Lee et al., 2008) and weight (Goodwin, 2010), age (Fischer and O'Hare, 2010, Kosugi et al., 2009) and sodium (Kyrou et al., 2006) intake, stress, smoking, alcohol consumption and obesity (Wofford and Hall, 2004) which is often linked to socio-economic class (Wenge et al., 2008, Monica et al., 2009, Kannel, 2000). The current most effective strategy for reducing hypertension is to control diet and sodium intake and reduce other risk factors including alcohol consumption (Annest, 1983, Annest, 1979) however the estimated effect of environmental to genetic is between 1:1.5 to 1:3 in terms of systolic blood pressure variance. The idiopathic nature of hypertension is indicative of the complex control of blood pressure and as such has led to the focus of preventions rather than cure, by looking at risk factors (Chobanian et al., 2003b).

In the short term blood pressure is controlled by a complex interplay between nervous and endocrine control and in the long term there is an osmotic control performed by the kidneys. Due to this complex interplay understanding, these mechanisms are a difficult undertaking. Despite the complex nature of its control, blood pressure is dictated by only two factors and simple physics. $\text{Blood Pressure} = \text{Cardiac Output} \times \text{Vascular resistance}$, with the vascular resistance being determined by the bore of the blood vessels and the continuity of the internal surface. Therefore much of the body's control focuses on the cardiac output and changing the internal diameter of the arterioles and this is closely regulated (Lifton et al., 2001). This control of the peripheral arterioles is essential for evening out fluctuation caused by the beating heart.

The organ with greatest responsibility in the control of long term blood pressure is the kidney (Guyton, 1991), achieving this by controlling the

osmotic balance using the re-absorption of water by controlling the level of salts in the blood. It achieves this by using hormones such as Anti-diuretic hormone (ADH) and pathways such as the rennin-Angiotensin aldosterone system (Laragh and al, 1972) in conjunction with systemic controls of arteriole diameter from arterial baroreceptor reflex (Drummond et al., 2001, Heymans and Neil, 1958, Ling et al., 1998, Lohmeier et al., 2005), chemoreceptors, kidney kinin-kallikrein (Sharma et al., 1994) and sympathetic nervous control (Janssen and Smits, 2002). The importance of the RAAS in the control of blood pressure and in the treatment of blood pressure through angiotensin-converting enzyme (ACE) inhibitors has been shown. A summary of the RAAS and the use of inhibitor are shown in Figure 4-1 below (Brewster and Perazella, 2004).

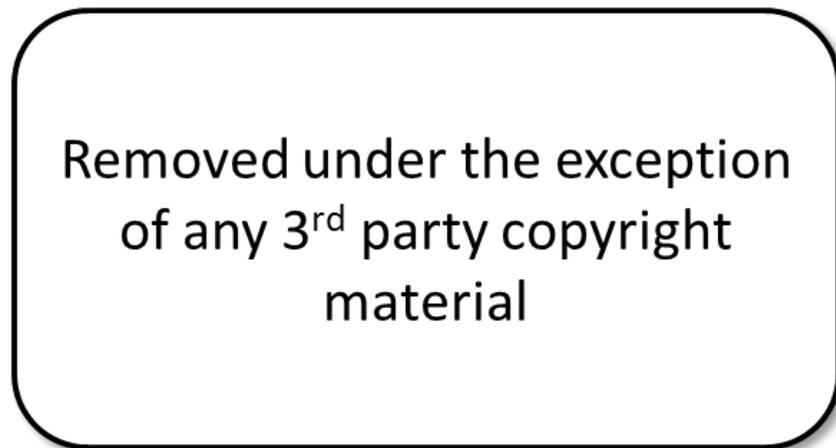


Figure 4-1: Summary of the Renin Angiotensin Aldosterone System and the use of ACE inhibitors. The system allows for control of blood pressure by measuring increase and decrease glomerular blood flow. This is one of the main mechanisms for controlling blood pressure in mammals. This figure was reproduced from (Brewster and Perazella, 2004).

The importance of the kidney in hypertension has been shown by a number of studies (Lohmeier et al., 2005, Blaustein et al., 2006), however it was beautifully illustrated by the transplantation of normotensive rat kidneys into hypertensive and vice versa (Bianchi, 1974). Originally, normotensive rats, which had the transplanted kidneys, were shown to develop hypertension. Equally, the hypertensive rats' blood pressure normalised. This has also been demonstrated in human transplantations (Guidi et al., 1996, Rettig R, 2005, Kopf et al., 1993, Opelz et al., 1998).

4.3.2 Genetic factors: Monogenic

The causes of hypertension are undoubtedly multi-factorial in nature, with factors such as obesity and diet playing at least some role, however, there is a growing bank of genetic evidence connected with hypertension. The majority of work looking at understanding the genetics of hypertension involves studies to find the causes of monogenic forms of the disease. Monogenic forms of hypertension are caused by very rare mutations, however, identification of these genes could provide candidates for essential hypertension and give a wealth of information regarding blood pressure control for the identity of drug targets. Most of the monogenic causes affect the function of the kidney in handling salt and water, underlying the importance of the kidney in homeostatic control of water and salt. This kind of research has been fruitful at identifying genes that cause both hypertension and hypotension. Lifton et al, 1993 reported 10 genes responsible for hypertension and 7 for hypotension. Monogenic causes tend to affect the kidney's ability to keep osmotic control. Often affecting the Na⁺ channels in some form, preventing the kidneys ability to control the osmotic balance of the blood. There are several different kinds of monogenic forms of hypertension that have been identified. To go into detail is out with the scope of this introduction, however briefly;

- Liddle Syndrome: Caused by a deletion or missense of the Cytoplasmic C-terminal domain of the ENaC β or γ subunits. This causes a reduced capacity for clearing of ENaC from the tubules membrane. This leads to an increase in the levels of salt and leads to hypertension and hypokalaemic alkalosis (Palmer and Alpern, 1999, Shimkets, 1997).
- Pseudohypoaldosteronism type II: This is an autosomal dominant disorder where serine-threonine kinases are affected by the lack of a conserved lysine (K). The genes affected are WNK1 and WNK 4. Sufferers present with a number of phenotypes including hypertension, hyperkalaemia and hyperchloraemia. The deletions and missense of WNK1 and WNK4 lead to over expression. Once again this leads to an outcome of salt retention and loss of osmotic control by affecting sodium and potassium levels (Wilson, 2001, Wilson, 2003).
- Glucocorticoid-Remediable Aldosteronism: This disorder causes the sufferers to present with hypertension from a young age due to an autosomal dominant disorder which affects the rennin-Angiotensin

aldosterone system by causing an imbalance in aldosterone levels and affecting the activity of rennin. The disease is caused by unequal crossover between two genes (CYP11B1 and CYP11B2), which lead to aldosterone synthase activity being controlled by adrenocorticotrophic hormone rather than angiotensin II. Increased activity of ENaC subsequently causes the increased retention and reabsorption of Na⁺ increasing plasma volume, leading to hypertension. Fortunately, this type of disorder can be treated with glucocorticoids to restore normal blood pressure (Comiter et al., 1995, Dluhy and Lifton, 1995, Lifton and Dluhy, 1993).

- Other monogenic disorders include; Apparent Mineralocorticoid Excess (Stowasser, 2006, Stewart, 1999), Hypertension with Brachydactyly, mutations of Peroxisome proliferator-activated receptor gamma and Mineralocorticoid Receptor mutations and hypertension in pregnancy. Most of which affect the regulation of salts in the nephron of the kidney leading to hypertension and a variety of other phenotypes (Lifton et al., 2001, Burke et al., 2005). This is of interest in hypertension as exacerbation during pregnancy called preeclampsia, which occurs in approximately 5-6% of pregnancies and can be fatal. A number of studies have investigated predictive markers for this disease (Kenny et al., 2005, Carty et al., 2008).

The biomarkers associated with hypertension have been covered in section 1.6.1.2.3.

4.3.3 Rat models of human hypertension: The use of stroke-prone spontaneously hypertensive Rat (SHRSP)

In this investigation, the effect of salt is initially evaluated and the changes in proteomic profiles with tissue types (medulla and cortex) are considered. It is hoped that in future studies the data obtained may be used to help link genomic and proteomic data. In order to link the proteomic and genomic data at a later date, the animals that have been selected are SHRSP (as discussed below), SP.WKYGla.2a and WKY. These are well characterised by the Cardiovascular Research Unit at Glasgow University (BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA). The strategy in place being that for using Quantitative Trait Loci (QTL), congenic and sub-congenic strains in order to tally genes and proteins together (Graham D, 2007b).

Rattus Rattus has been an essential tool in the study of many diseases including hypertension, on both a physiological and genetic level. They allow relatively large scale breeding experiments, with the capacity to control different experimental factors including medication, diet, simulation and exercise. The laboratory at Glasgow maintains a number of different strains of hypertensive rats and genetic variants. These include stroke-prone spontaneously hypertensive SHRSP, inbred normotensive Wistar Kyoto (WKY) rats. The SHRSP rats are selectively bred, on the basis of displaying higher blood pressure than the spontaneous hypertensive rats (SHR). The SHRSP are, however, predisposed to cerebrovascular accidents (Yamori, 1994). The SHRSP and the link to salt intake has been well characterised in a number of studies and shows hypersensitivity to salt, left ventricular hypertrophy and endothelial dysfunction (Clark et al., 1996; , Kerr et al., 1999, Koga et al., 2008) and show a contrasting phenotypical response to salt treatment in comparison to WKY. This was shown when SHRSP had backcrosses performed to generate populations with increasing SHRSP genetic compositions. When salt treated the increasing genetic composition showed a positive correlation with the increased instance of stroke recorded in each group (Nagaoka, 1976). There are a number of studies implicating increased levels of salt (or more accurately sodium from NaCl), with a higher instance of hypertension (De Wardener and Macgregor, 2002, Fountoulakis, 2001, Koga et al., 2008, Meneton et al., 2005). In the case of excess salt causing the renal capacity to be reduced, the reduced renal function may lead to hypertension caused by an increase plasma volume retention (Hamlyn and Blaustein, 1986).

As part of work done previously, by the cardiovascular research unit, QTL have been identified in this rat model at chromosome 2 (Jeffs et al., 2000) and validated using congenic and sub-congenic mapping with candidate genes identified and micro array analysis (McBride et al., 2003). The congenic gene map can be seen in Figure 4-2 below taken from (Graham D, 2007a)

Removed under the exception of
any 3rd party copyright material

Figure 4-2: Chromosome 2 genetic map in SP.WKYGla2a, SP.WKYGla2k, and SP.WKYGla2c* strains. This shows the genetic makeup of the congenic strains in relation to Wky and SHRSP. Showing the congenic strain SP.WKYGla2k employed in this study, has a largely SHRSP genetic makeup. Taken from (Graham D, 2007a).

-  = WKY homozygosity.
-  = SHRSP homozygosity
-  = recombination/heterozygosity

Various congenic strains and sub-strains are kept by Glasgow University. Congenic strains are used as a mode of mining down into the relatively complex QTL. In QTLs normotensive and hypertensive rats are crossbred to create heterozygous first generation progeny. Sibling mating follows to generate F2 animals, which are then genotyped and phenotyped for continuous traits. Trait probability is then calculated using polymorphic markers across the genome and linked with phenotype across the genome. Congenic strains allow for a QTL of importance to be studied by selecting a region of the chromosome and replacing one strain (normotensive) with another strain (hypertensive). If a phenotypical change is noticed then that phenotype can be linked with that QTL. Congenic strain take time to develop, due to a series of backcrosses being necessary to ensure the donor's genetic background is fully replaced.

The samples used in this study are congenic strains which have been created using speed congenic breeding by repeating screening for a polymorphic marker locus, allowing for specific selection of the donor rat. A rat fast gestation period is one advantage of using rat models. Although some work has just started in assessing the candidate genes from human clinical samples, most information held pertains to a rat model. Clinical samples tend to have a high heterogeneity, leading to larger sample variance and are more complicated to analyse in comparison to a rat model. Additionally, human studies take longer time and it is difficult to control all variables. It is recognised that increasing sample size will reduce type I and type II errors when looking for candidate markers, genes or proteins, but this needs to be balanced with expense and logistics.

4.3.4 Proteomics and Hypertension

With the wealth of information provided by the completion of the human genome project, the focus is now linking functional information regarding proteins to the genomic information. By identifying proteins and peptide markers for hypertension, it is possible to gain structural information for diagnostic markers for early detection and eventually therapeutic targets for treatments.

Differential in Gel electrophoresis in theory is an ideal mode for revealing possible biomarkers for disease. A biomarker is a biological molecule such as a protein or peptide, which might be used to classify, diagnose or monitor a disease and would ideally be present in sufferers and absent in healthy patients. However, in reality, it may be a collection of biomolecules which may be subtle quantitative changes in expression. DiGE is ideally placed to look for changes in patterns of abundance of proteins with a relatively wide dynamic range. In addition DiGE can also look at 1000s of proteins at one time, which out performs most other proteomic techniques. However proteomic techniques are somewhat under developed in comparison to their genomic counterparts and require substantial work (Jones MB, 2002, Listgarten and Emili, 2005, Hilario et al., 2004, Zhou, 2005). This is often due to the intrinsically complex nature of proteins and the lack of amplification technology such as PCR that revolutionised genomic research. In addition, the deconvolution of data possesses a massive issue. DiGE can give abundance information for 1000s of

proteins in one experiment, this poses issues of operator time and data mining, particularly if spots are to be identified and validated.

The use of DiGE to discover biomarkers is reasonably well represented in the literature as either a main method or complementary to other proteomic techniques. Most of these utilise a body fluid rather than tissue directly. Biomarker studies in proteomics are wide spread and well described, but often look for markers in urine (Pisitkun et al., 2006, Gozal et al., 2009, Vivanco et al., 2005). Often studies are relatively focused on one pathway, rather than a global approach undertaken in this investigation. It is recognised that each strategy has its relatively advantages and disadvantages. DiGE analysis allows for a greater degree of an “unguided” approach, where no previous information is necessarily required about a given set of markers and therefore it acts like a “fishing expedition”. Therefore, being that DiGE can map 1000s of proteins at a time, this strategy is ideally placed to search for biomarkers. The disadvantage of going into the study blindly is that vital information can be missed. However, by using a well characterised rat model it is hoped that this disadvantage can be minimised and any candidate markers can be further investigated by tying in genomic and proteomic data at a later date.

This investigation attempts to address some strategies for deconvolution and data mining, while at the same time presenting some possible candidate protein markers to be further investigated with respect to hypertension in SHRSP rat model.

4.4 Methods

A detailed account of general methodology is outlined in General material and methods chapter of this thesis. Methodology specific to this chapter is given below.

It was decided that using the Denator Stabilisation instrument in this study was not appropriate for a number of reasons;

- The tissue tested in the first study was mouse brain due to the original intention of investigating biomarkers in stroke using mouse brain. Unfortunately, due to logistical problems with the supply of stroke mouse brain tissue in this project was suspended.
- It has been shown that using the Denator instrument causes detrimental effects on RNA levels. It can lead to its break up or increase and decrease in expression (Kultima et al., 2011).
- The animals harvested were being utilised in a number of different studies, so it was decided that the same upstream treatment needed to be applied across all tissue types, if fair comparisons are to be made later.
- The kidney was chosen as opposed to brain, as it has a closer association to the disease of hypertension than brain and the model animal being employed was a rat not mouse as in the previous study.
- The Denator instrument has been heavily studied using brain tissue and there is evidence from a number of sources that it has a beneficial effect on stabilising brain tissue. In the previous chapter it was decided to try and validate this using a different method (DiGE). Also some other tissue types have been shown to have either no effect so there it is something that needs greater investigation across a multitude of tissue types.
- Finally, the instrument was on loan for a limited period of time and was not available for this study.

4.4.1 Sample collection, extraction and processing

Tissue for this study was provided by the Cardiovascular Research Group at Glasgow University. Animals were euthanized inline with the U.K. Animals (Scientific Procedures) Act, 1986 and local ethical guidelines. Male WKY (Wistar Kyoto Rat), SP.WKYGl2a (congenic strain) and SHRSP (Spontaneously Hypertensive Rat Stroke Prone) rats were euthanized (21 weeks) by cervical dislocation and dissection of the brain, heart, kidney and liver were rapidly performed. During the 21 weeks plasma and urine samples were taken. Tissue was snap frozen in liquid nitrogen following dissection. Heat treatment using the Denator heat stabilisation process was ruled out in order to maintain cross group compatibility when comparing any proteomic and genomic data between genomic and proteomic groups. The genomic group had been using traditional snap-freezing methods for the duration of the project, so it seemed expedient to maintain this status quo. Samples were then stored at -80°C in anticipation of use for downstream processing. Samples were always handled with care and never above 4°C, as all processing was performed on ice. In addition, protease inhibitors were added at the earliest possible time point. Blood pressure measurements were taken at regular intervals throughout growth and plasma and urine samples taken. A salt treatment was applied to half of the cohorts with inception at 18 weeks using a 1% NaCl solution as drinking water. Blood pressure was measured using Dataquest IV telemetry system. The genetic models used for this study are given in Figure 4-3 and the phenotypical blood pressure readings are given in Figure 4-4, which distinguish the three types of rats. The SHRSP rats have a higher average systemic blood pressure compared to the Wky and congenic strains. Also the SHRSP are particularly prone to sudden cerebral infarct events. They are raised in very low light environments as the sudden shock can induce stroke. Wky have the lowest phenotypical blood pressure as being in a normal range for rats. The congenic rats have an intermediate phenotype between the SHRSP and Wky rats. Therefore, the samples used in this study are from three distinct versions of disease states; normal, intermediate and diseased. It should be noted that the intermediate disease state may therefore hold certain characteristics closer to one group or the other. The blood pressure of the congenic animals is closer to the SHRSP animals than the Wky. Therefore it is of value to see the pattern of how the agents of phenotype (the proteins) compare across the three strains.

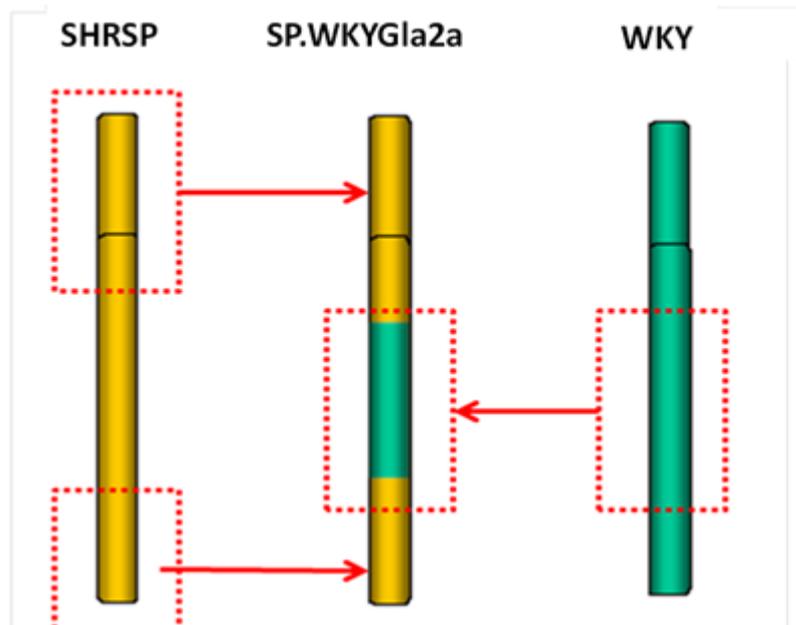


Figure 4-3: Schematic of genetic model employed. A congenic strain has been developed with a SHRSP background with sections of Chromosome 2 from WKY. Figure provided by Dr Martin McBride, Cardiovascular research unit.

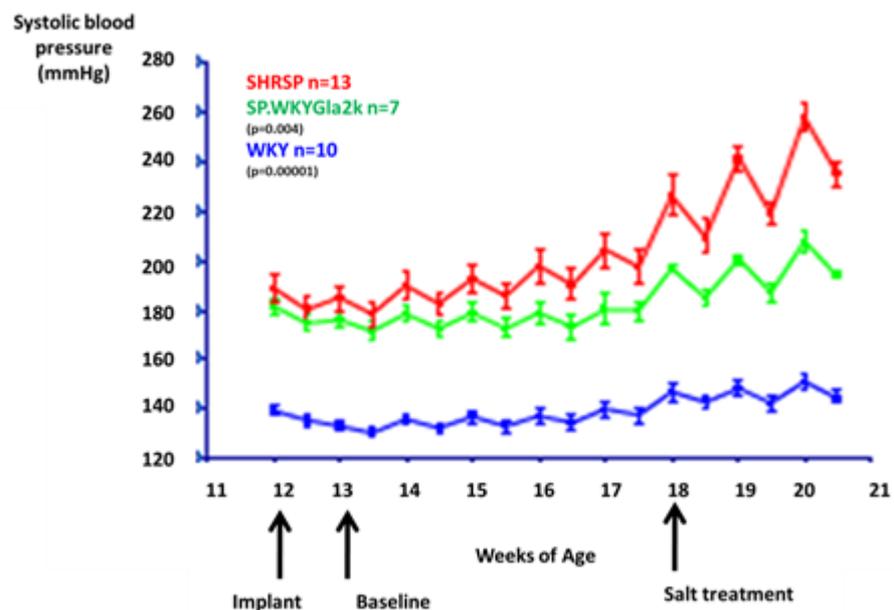


Figure 4-4: Systolic blood pressure phenotyping using radiotelemetry. Systolic blood pressure was significantly increased in the salt-loaded SHRSP compared to salt-loaded SP.WKYGla2a and WKY rats. Figure provided by Dr Martin McBride, Cardiovascular research unit.

The study was divided into a pilot investigation and main investigation. The pilot study used 3 samples from the same animal for each condition. The main investigation had sample n=3 (three separate animals). It is recognized that sample size is low but these animals are expensive to raise and material is

limited. A summary schematic of the two experiments is given in Figure 4-5 and Figure 4-6.

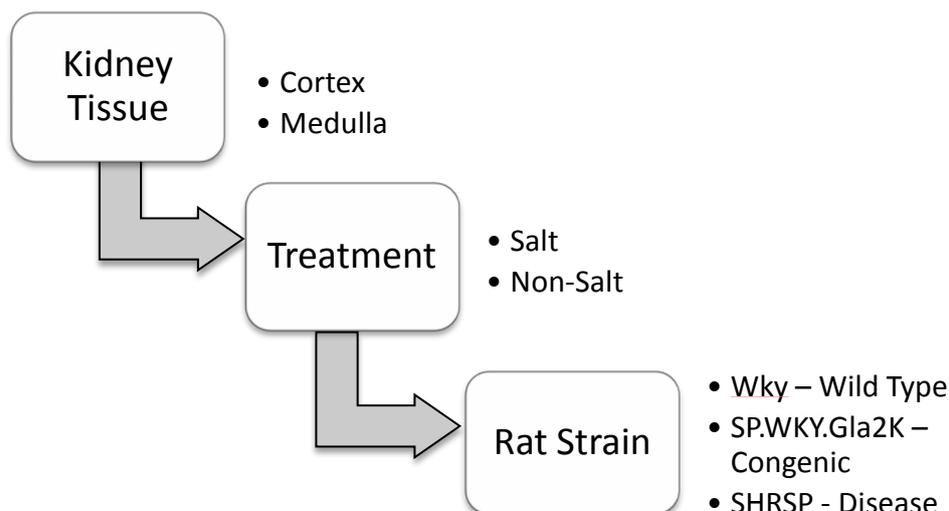


Figure 4-5: Sample work flow for the pilot investigation replicates. Showing that both cortex and medulla tissue was used in addition to salt and non-salt treatment.

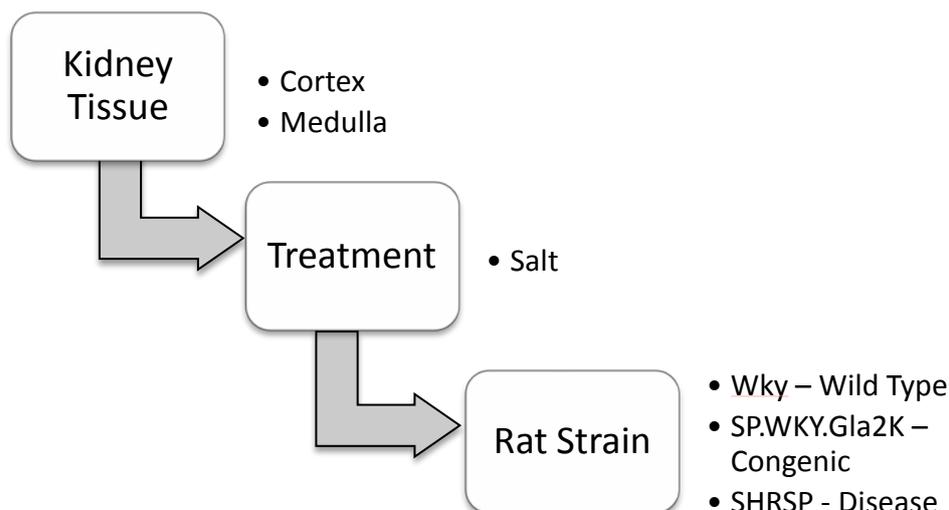


Figure 4-6: Sample workflow for main investigation. Showing that only salt treatment samples were used but tissue segregation into cortex and medulla regions was still employed.

Sectioning of kidney tissue was performed using a cryostat microtome (Leica Microsystems CM 1900UV, UK). Multiple sections were taken and divided up between eppendorfs and MALDI ITO coated glass slides for this study and collaborators. Tissue sections were cut at a thickness of 14µm in a pre-chilled chamber of -20°C and a sample stage temperature of -19°C. Eppendorfs used

for storage were kept and allowed to pre-chill in the chamber as well, before being transferred to dry ice.

Sections used in MALDI-Mass Spectrometry Imaging (MSI), were thaw mounted onto ITO slides and stored on dry ice further to long term storage at -80°C . Sections were taken from between $\frac{1}{4}$ and $\frac{3}{4}$ the coronal depth of the kidney. Sections for DiGE minimal labelling were placed onto glass slides in order for macro dissection to take place. The sequence of slices was kept strict and regular, in order to minimize aberration downstream and allow collaborators to compare data at a later date; the order of slicing can be seen in Figure 4-7.

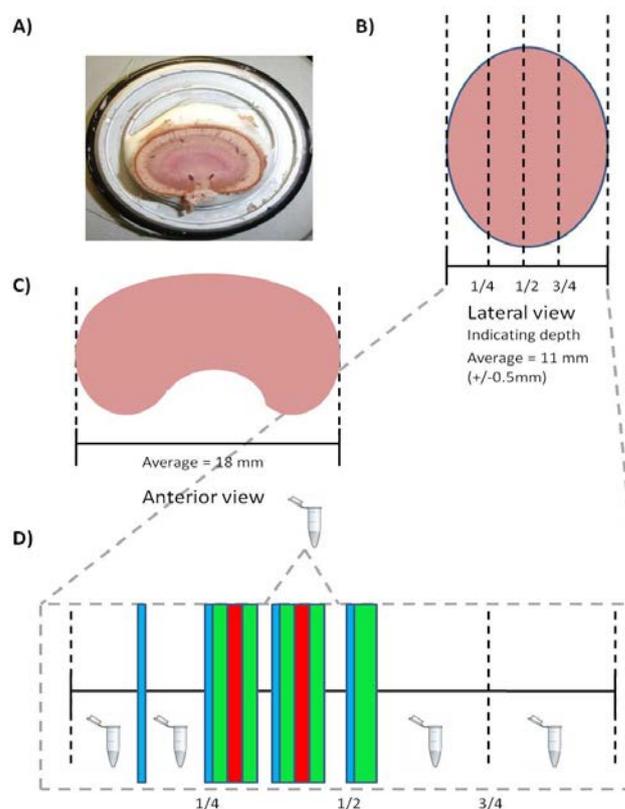


Figure 4-7: Schematic of section sequence. Kidney tissue was mounted (A) after measurements of the dimensions were taken (B and C). Sections were taken at regular intervals. (D) Blue: ITO slides for MSI (collaborator), Green: Glass slides for macro dissecting for DiGE minimal labelling, Red: Glass slides for Saturation labelling, Laser Micro dissection slides and Glass slides for IHC (future validation of MSI). The first 100 slices taken into Eppendorfs, 6 for MSI then up to 200 slices into Eppendorfs ($\frac{1}{4}$ depth). Then 6 sections for MSI, 74 sections for DiGE, LMD and IHC. Up to 300th section into Eppendorfs. Following this pattern till $\frac{1}{2}$ depth. Then a further 40 sections for DiGE.

Sections for collaborators were stored at -80°C . Sections on glass slides for DiGE minimal labelling were taken through for macro-dissection. Sections were placed and illuminated on a light box to allow definition of defined structures to be seen. A scalpel blade was used to remove the cortex regions of

the kidneys into an Eppendorf, which had been prepared by using an anti-static gun to avoid loss of sample. In order to obtain enough tissue, 50 slices were used from each kidney. Following macro-dissection of the cortex, the remaining medulla region was then scraped and pooled. The regions of the kidney have been defined but it is recognized that such definitions are complex and convoluted so the definition of cortex and medulla (corresponding to macro-dissected regions), for the purpose of this study are given visually in Figure 4-8. It is recognized the undulations and convolutions mean regions are difficult to assign when viewed histologically, while the definition visually is distinct which allowed for accurate dissection between the two.

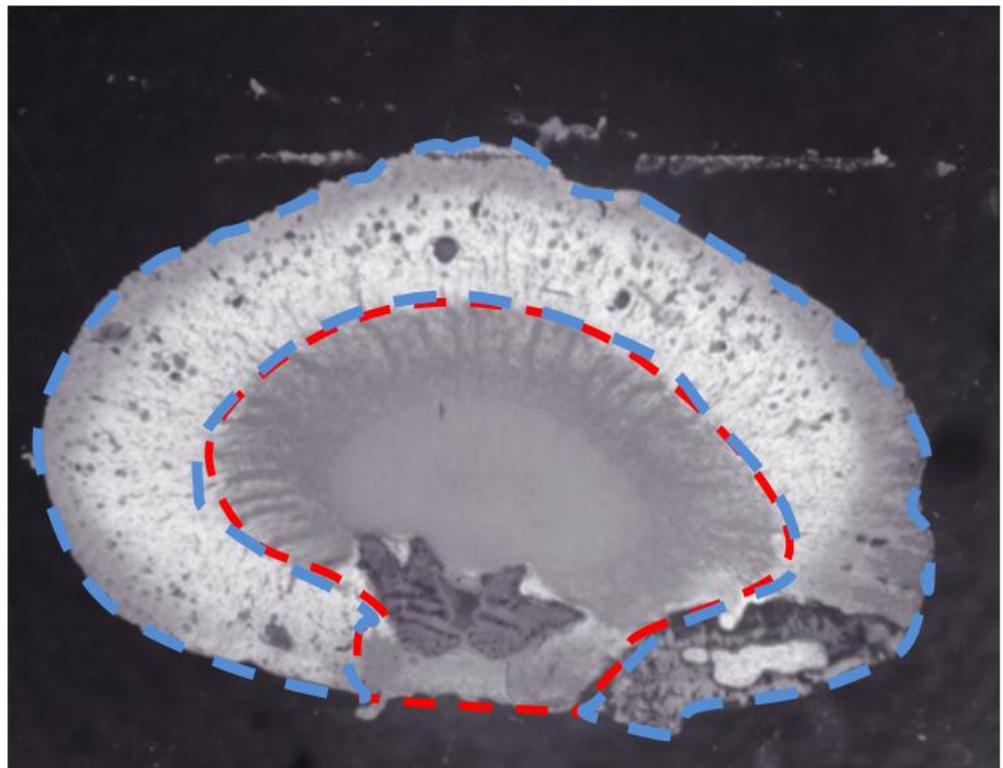


Figure 4-8: Optical image of kidney section defining regions of macro-dissection. The region within the red dotted line gives the “medulla” and blue dotted line “cortex.” It is recognised the undulations and convolutions mean regions are difficult to assign, while the definition visually is distinct.

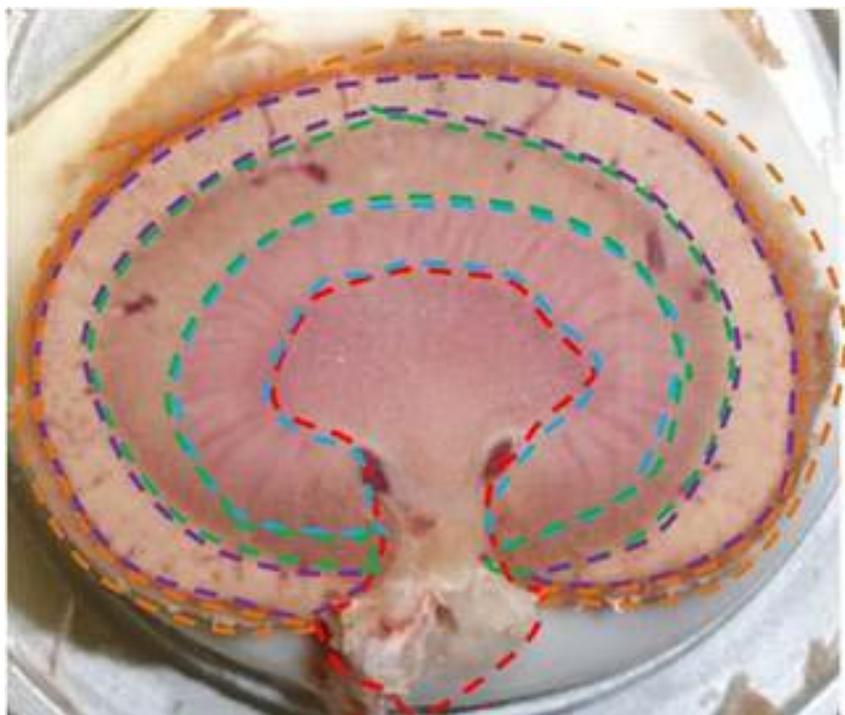


Figure 4-9: Visible regions of the kidney with the eye: These regions of the kidney are noticeable as having a colour difference and show the dynamic and complex tissue of the kidney organ. This is only as eye level and the complexity deepen when using light microscopes. Therefore there intrinsically and intuitively is a need to try and separate tissue types in order to fully understand their functional significance.

Protein lysis and extraction was performed immediately after treatment in a DiGE compatible lysis buffer, with the addition of protease inhibitor and DNase; 7M Urea, 2 M thioruea, 4% CHAPS [w/v], 30 mM Tris base (solution in section 2.2.7). Following the addition of lysis buffer, the samples were subject to 3 Cycles of snap freezing, thawing and 4 x 5 minute Cycles in an iced sonication bath with 1 minute cooling on ice between sonication.

Extractions were followed by protein precipitation and clean-up using Ettan™ 2D Clean-up Kit (GE Healthcare, Bucks, UK cat #80-8484-51). Minimal labelling reactions were performed with Cyanine dyes 2, 3 and 5 using protocol specified in the product booklet (GE Healthcare, Bucks, UK #25-8009-83/84.) and is in the general methods. In short, 50µg of protein was used for analytical gels and 500µg for preparative gels, which were stained using Sypro Orange (Sigma Aldrich, Dorset, UK, cat #S5692-500UL). The reaction was performed at pH 8.5, using 400 pmol of CyDye incubated on ice for 30 minutes in the dark. The reaction was stopped using the addition of 10 mM lysine. They were loaded (as shown in Table 3-1 and Table 3-2) for 1st dimension separation which was carried out using IEF on 24 cm IPG strips of pH 4-7 [GE Healthcare, Bucks, UK cat # 17-6002-46]) with a minimum of 75000 Volt

hours at 20°C with a program of 30 V step and hold for 12 hours, 300 volts step and hold for 2 hours, 1000 volts gradient for 2 hours, 8000 volt gradient for 5 hours, 8000 volts step and hold for 8 hours and 1000 volts step and hold for up to 24 hours to avoid strip diffusion prior to next step. Prior to 2nd dimension separation by SDS-PAGE, an equilibration was performed using SDS equilibration buffer (50 mM Tris-HCL, pH 8.8, 6 M urea, 30% glycerol [w/v] 2% SDS [w/v] and 0.002% BPB [w/v]) with 10 mL/gel followed by reaction with DTT (10 mg/mL) for 15 minutes, then with iodoacetamide (25 mg/mL) for 15 minutes to prevent reduction/alkylation. 2nd dimension separation was performed at 1-2 watts per gel for approximately 12 – 15 hours or until the dye front reached the bottom of the gel. Gels were imaged using GE Healthcare Typhoon 9400 Series Variable Imager at 100µm resolution after optimization of photomultiplier voltages, using a pre-scan at a resolution of 1000µm. Gels were then loaded, (see chapter 2) matched and analysed, (DeCyder Version 5.01.01, GE Healthcare, Bucks, UK), and spots selected for picking using Ettan™ Gel Handling Work Station and MS identification. A schematic of the workflow can be seen in Figure 3-2.

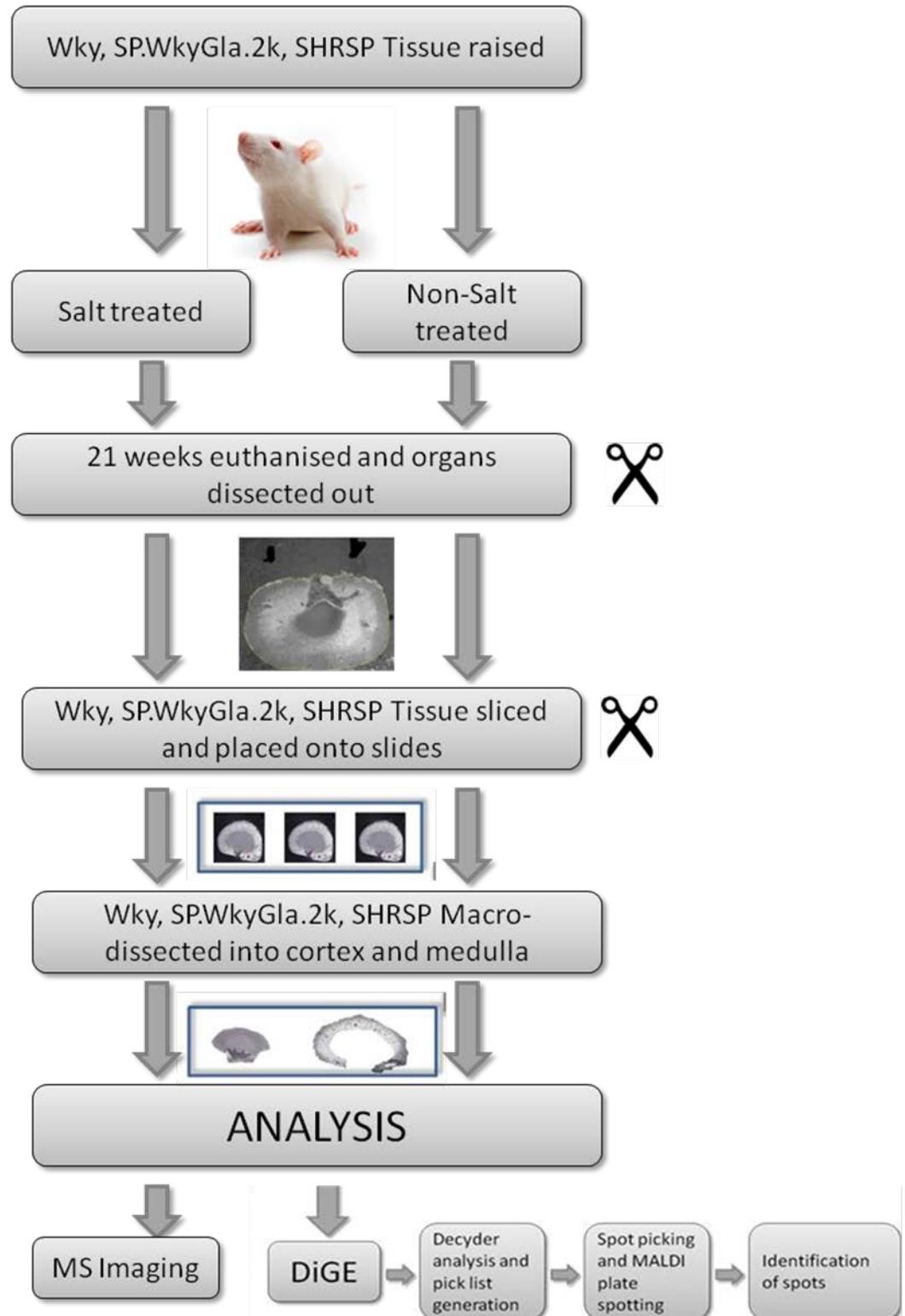


Figure 4-10: Schematic of experimental work flow. Kidney tissue was processed for DiGE analysis. MALDI-MSI was performed separately by Dr R. J. A. Goodwin, Research Associate, University of Glasgow. The larger proteins considered using DiGE and the smaller markers considered using MALDI-MSI.

4.4.2 Identification of proteins from gel spots

Spots were picked and tryptically digested, as describe in the general materials and methods and spotted with a-Cyano-4-hydroxycinnamic acid in 50% acetonitrile/0.5% trifluoroacetic TFA using a GE Healthcares Ettan Spot Handling Work Station and then analysed on 4700 Proteomics Analyser (Applied Biosystems, Cheshire, UK) MALDI-ToF-ToF-MS using standard settings (outlined in chapter 2). MS/MS was performed on the top 10 precursor ions in each spot. Any unidentified spots of particular interest were further analysed by LC-MS/MS on a Dionex Ultimate+ LC system coupled to a QStar Pulsar I (Applied Biosystems, Cheshire, UK). GPS Explorer and MACOT Daemon Software was used to automate submission of collected data to MASCOT database searching software for searching with fixed modification of carbamidomethyl (C) and variable modification of oxidation (M), peptide and MS/MS +/- tolerances of 0.8 Da searching NCBI database on Mus musculus taxonomy with 1 missed cleavage allowed. Spots identified as keratins were excluded.

4.4.3 Macro-dissection of kidney tissue

Macro dissection was used as a method to segregate medulla and cortex of kidney tissue. In order to assess the validity of macro dissection as a mode for the segregation of the Kidney, a validation of the technique was performed using Wild type OCR male rats scarified at 16 weeks. The kidney tissue was sliced and either thaw mounted onto glass slides, or placed in 1.5mm eppendorf tubes in accordance with methods above and Figure 4-7, after which the 6 slices were dissected as described separating medulla and cortex, 6 slices were scraped as a whole and 6 slices were used straight from the tube for protein concentration analysis using a Bradford assay. Results can be seen in section 4.6.4.1. As consecutive slices were used, an assessment of average area was performed using 20 slices. A schematic of the experiment can be seen below in Figure 4-11.

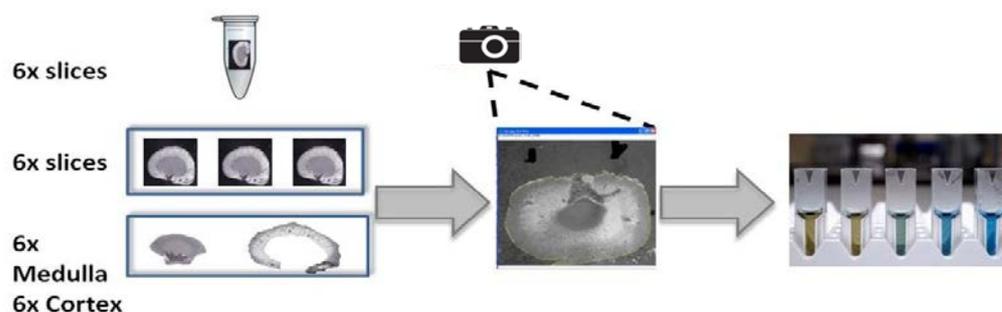


Figure 4-11: Schematic for the assessment of macro dissection. Kidney tissue was sliced and placed in eppendorfs and on glass slides. The area of the tissue was calculated using image j. An assessment of protein per unit area was made using a Bradford assay with a BSA standard curve of whole slices from eppendorfs, whole dissection from a glass slide and macro dissection of medulla and cortex from a glass slide. Results are shown in 4.6.4.1 on page 346.

4.4.4 Experimental Design

The experimental design used for DiGE gels were as follows:

4.4.4.1 Pilot investigation: Experimental replicates of salt treated and non-salt-treated WKY, congenic and SHRSP. Cortex and medulla.

	Cy5	Cy3	Cy2
gel 1	Cortex WKY salt	Medulla WKY salt	internal standard
gel 2	Cortex WKY non-salt	Medulla WKY non-salt	internal standard
gel 3	Cortex Congenic salt	Medulla Congenic salt	internal standard
gel 4	Cortex Congenic non-salt	Medulla Congenic non-salt	internal standard
gel 5	Cortex SHRSP salt	Medulla SHRSP salt	internal standard
gel 6	Cortex SHRSP non-salt	Medulla SHRSP non-salt	internal standard
gel 7	Medulla WKY salt	Cortex WKY salt	internal standard
gel 8	Medulla WKY non-salt	Cortex WKY non-salt	internal standard
gel 9	Medulla Congenic salt	Cortex Congenic salt	internal standard
gel 10	Medulla Congenic non-salt	Cortex Congenic non-salt	internal standard
gel 11	Medulla SHRSP salt	Cortex SHRSP salt	internal standard
gel 12	Medulla SHRSP non-salt	Cortex SHRSP non-salt	internal standard
gel 13	Cortex WKY salt	Medulla WKY salt	internal standard
gel 14	Cortex WKY non-salt	Medulla WKY non-salt	internal standard
gel 15	Cortex Congenic salt	Medulla Congenic salt	internal standard
gel 16	Cortex Congenic non-salt	Medulla Congenic non-salt	internal standard
gel 17	Cortex SHRSP salt	Medulla SHRSP salt	internal standard
gel 18	Cortex SHRSP non-salt	Medulla SHRSP non-salt	internal standard
			n=3

Table 4-1: Experimental design of gels ran for experiment 1: Experimental replicates for pilot investigation of salt treated and non-salt-treated WKY, congenic and SHRSP. Cortex and medulla.

4.4.4.2 Main Investigation: Biological replicates of salt treated WKY, congenic and SHRSP. Cortex and medulla.

	Cy5	Cy3	Cy2
gel 1	Cortex WKY Salt (A4642)	Medulla WKY Salt (A4642)	internal standard
gel 2	Cortex SHRSP Salt (C5850)	Cortex WKY Salt (A4631)	internal standard
gel 3	Cortex WKY Salt (A4737)	Cortex Congenic Salt (N6718)	internal standard
gel 4	Medulla WKY Salt (A4631)	Medulla SHRSP Salt (C992)	internal standard
gel 5	Medulla WKY Salt (A4737)	Medulla Congenic Salt (N6718)	internal standard
gel 6	Cortex Congenic Salt (N6841)	Cortex SHRSP Salt (C5847)	internal standard
gel 7	Cortex Congenic Salt (N6842)	Medulla Congenic Salt (N6841)	internal standard
gel 8	Medulla Congenic Salt (N6842)	Medulla SHRSP Salt (C5847)	internal standard
gel 9	Cortex SHRSP Salt (C5992)	Medulla SHRSP Salt (C5850)	internal standard
			n=3

Table 4-2: Experimental design of gels ran for experiment 2: Biological replicates of main investigation of Salt treated WKY, congenic and SHRSP. Cortex and medulla.

4.4.6 Statistical Methods

As in section 3.4.5, all statistical calculations have been performed in either the DeCyder 2D 7.0, GE Healthcare software directly with further statistical analysis done using SPSS (Statistical Package for the Social Sciences) version 17.0.1 (2008).

In summary, the data was exported from the DeCyder work space. Normality testing was then performed in SPSS 17.0.1 on $\log_{10}SA$ using the Shapiro-Wilk statistical test. A spot was considered of non-normal distribution if a p-value was returned of 0.05 or less. This was performed on spots that had a full set of repeats across all gels, those that had a full set of repeats across all gels and an ANOVA score of 0.05 or less and those that had an ANOVA score of 0.05. Missing values replaced using k-Nearest Neighbour (KNN) algorithm in SPSS 17.0.1 to replace values.

To assess the homogeneity of variance, the SA and $\log_{10}SA$ data across all groups was analysed using the Levene's test for homogeneity. This was performed on spots that had a full set of repeats in every group for each master spot and those with replaced values. A spot was considered not to have homogeneity of variance when it has a p-value of 0.01 or less. Additionally graphs were produced to assess visually the spread of data.

The ANOVA calculation and Student's t-test was performed within DeCyder 2D 7.0, GE Healthcare in the BVA module. In addition to testing these assumptions, the issue of multiple testing of data has been addressed by applying a p-value correction. This has been applied directly in DeCyder via the BVA module option for the application of a false discovery rate.

4.5 Rationale for Analysis

In this section the rationale behind the analytical approaches is laid out and discussed.

In order to assess the proteomic expression of kidney tissue in hypertension using a WKY, congenic and SHRSP rat model and the search for candidate markers towards the assessment of chromosome 2, it is important first to build a complete profile of hypertension proteomic expression using the following criteria:

- Differences and similarities between strains
- Differences and similarities across tissues
- Differences and similarities due to treatment.

All three of these will be considered and addressed in the pilot investigation. In the main investigation the difference in treatments has been omitted. This is discussed in section 4.6 and 4.7.

The three main strategies for assessing the proteomic data are:

- Profile analysis, as piloted in Chapter 3.
- Coupling targeted Venn and profile analysis
- Validation using gel map data and principal component analysis.

The strategies employed to discuss these three key areas are outlined below.

4.5.1 Pilot investigation and Main investigation

The first part of the study to be summarised is the pilot investigation. The three areas to be looked at, mentioned above, are outlined below.

4.5.1.1 Differences and similarities between strains.

As outlined in section 4.3.3) the use of QTL and congenic strains is a useful means for comparing gene information and phenotype. The use of proteomic technique can be employed to view the proteomic expression and therefore phenotype. In using profile analysis, the relationship between protein expressions in the different strains can be viewed easily and be related to possible biomarkers within each strain. Eventually proteomic and genomic data could be pooled and combined to assess the similarities and differences in

expression at the genomic and proteomic level. Further pathway analysis can be employed to try and map the path from gene to protein.

This study places its limits within this work flow by mining DiGE data for possible proteomic biomarkers that maybe of interest in future genomic studies, those of which are being undertaken by Dr Martin McBride in the Cardiovascular Research Unit at Glasgow University.

Profile analysis for differences and similarities between strains

The pattern of the profile therefore gives information about how the three strains exhibit different or similar phenotypes (protein expression) and DiGE is perfectly placed to view this profile pattern.

The predicted profile patterns considered for searching for similarities and differences are shown in Figure 4-12. Of particular interest with respect to relating hypertensive biomarkers to chromosome 2 are A), B) and C) but all possible patterns potentially hold information that may inform research scientists looking for genetic and phenotypical links with these strains.

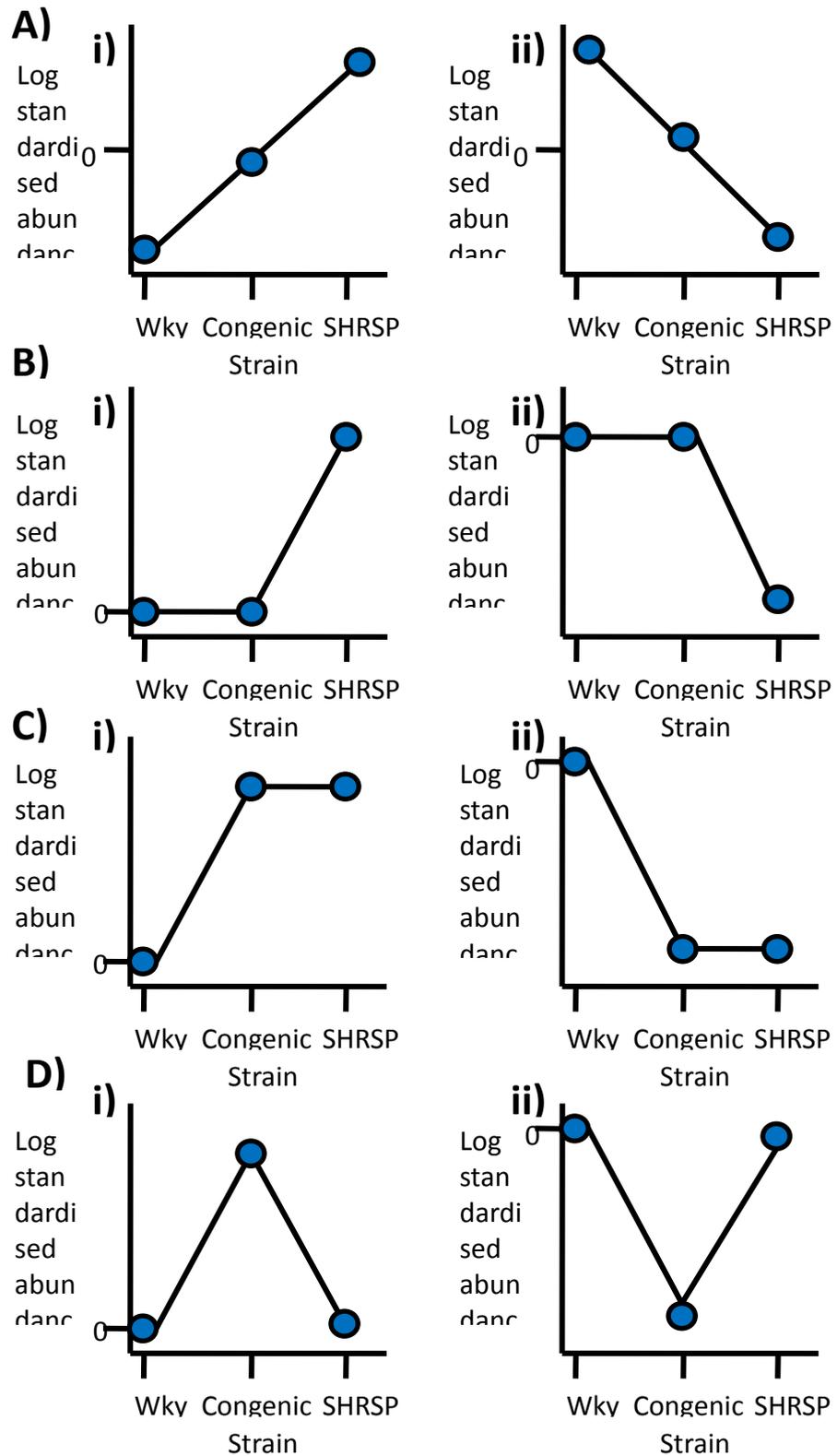


Figure 4-12: Predicted Profiles of spot intensities. Shows the predicted relationship between WKY, congenic and SHRSP strains. A i-ii) Shows possible intermediate effect in the congenic strain compared to WKY and SHRSP (Intermediate effect i-ii). B i-ii) No change between the WKY and Congenic strains (WKY maintained i-ii). C i-ii) No change between the congenic and SHRSP strains (SHRSP maintained i-ii). D i-ii) Congenic strains presenting possible different spot intensity in comparison to WKY and SHRSP strains (no change in terms of WKY and SHRSP). Profiles can be overlaid to compare two different tissue types or salt treatment.

4.5.1.2 Targeted Venn analysis using the predicted profiles for differences in strains

An example of how the Venn diagrams are used to target and mine the data is shown below in Figure 4-13. This will be extended for all three sections of stains, tissues and salt and no-salt for the pilot investigation and just strains and tissues for the main investigation.

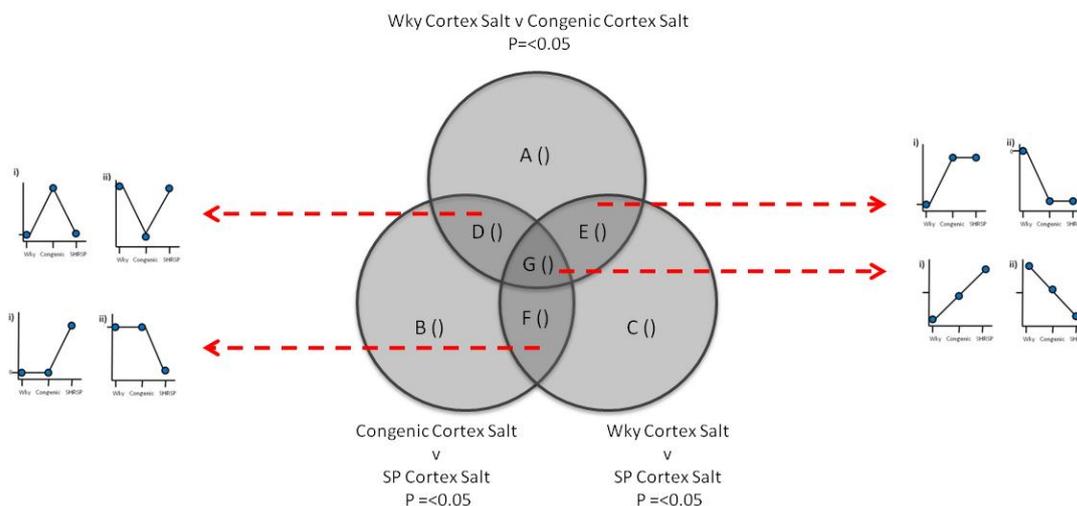


Figure 4-13: Targeted Venn analysis. Different crossover regions of the Venn diagram can be used to mine data by correlating them with predicted profiles. Additionally, statistical tests have already been run and cross checking is not required.

4.5.1.3 Differences and similarities across tissues using Venn analysis with profile validation

In order to explore the significant differences and similarities across tissue types, an approach of Venn analysis and profile analysis will be employed. An example of the Venn strategy is shown in Figure 4-14.

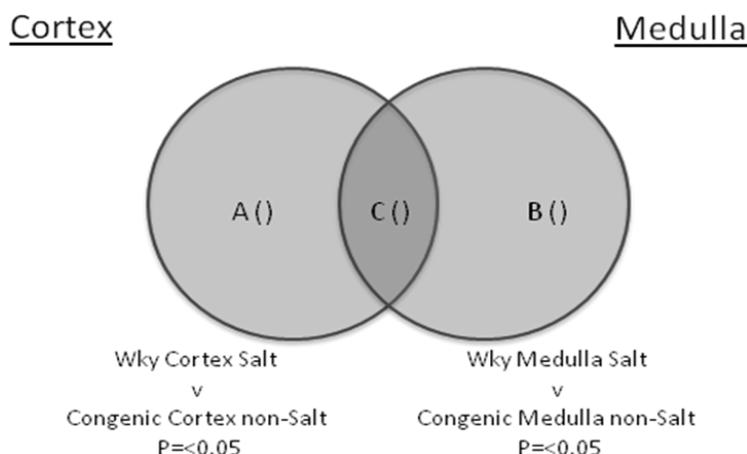


Figure 4-14: Example of targeted Venn analysis to illicit profiles showing differences and similarities in tissue types. C shows the changes which cortex and medulla have in common.

4.5.1.4 Differences and similarities due to treatment using Venn Analysis with profile validation (pilot investigation only).

To explore the significant differences and similarities across salt and non-salt treated tissues, once again, a combination of Venn and profile analysis is employed to search for possible candidate markers. An example of the Venn diagram used can be seen below in Figure 4-15.

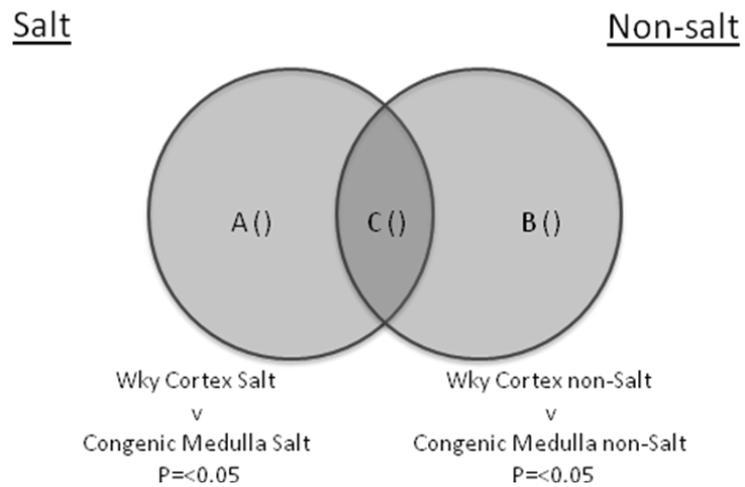


Figure 4-15: Example of targeted Venn analysis to illicit profiles showing differences and similarities between treatment with salt and no salt treatment. C) Shows any similarities between treatment of salt and non-salt.

4.6 Results and discussion

The integration of genomic and proteomic data is and will become increasingly more essential, if genotype and phenotype are to be tied together. By bringing both pieces of biological information together, it will increasingly become possible to understand biological processes and pathways at a more detailed level. Unfortunately, due to the speedy evolution of genomics and the comparative slow progression in proteomics, researchers have tended to be focused in one particular area and integrative approaches have not been as speedy as they might have been (Thongboonkerd, 2005). On the positive side, scientific collaboration is greater than ever and the two fields are starting to be knitted together. The use of DiGE as a mode for looking for candidate markers is not novel but will not suffice on its own, an integrative approach where multiple methods are employed is essential. In order for the research to become robust collaboration and translation between the genome and proteome is essential.

Hypertension itself is reasonably well described (see section 4.3) but the use of DiGE with the prospect of matching up genomic data in hypertension is novel. 2DE and DiGE are well described in proteomic research and as such this forms a solid foundation for use with genomic data. In addition, with DiGE's ability to quantify and display 1000s of proteins, the chances of being able to match data with RNA micro array data (available through Cardiovascular Research Lab, Glasgow University) is improved in comparison with many other proteomic techniques. It is of course recognised that a lot more work is required in order to marry both sets of data, much of which is outside the scope of this thesis. However, it is intended for this thesis to be a starting block for any future research.

It is the hope that the following results allow for a starting point by giving candidate markers and a set of analytical methods using DiGE to sort and mine data. In addition, it is hoped that these methods might be extended by moving towards small sample proteomics, to gain better resolution to help reduce the averaging disadvantages that arise in much of tissue proteomics research, due to the homogenising of tissue.

4.6.1 Pilot investigations

Firstly, a pilot investigation was undertaken in order to assess whether a salt treated strain or not was to be used and also if separation into medulla and cortex was to be investigated. As importantly, a pilot study was prudent in order not to euthanise animals unnecessarily before embarking on the main study.

4.6.1.1 Typical gels

Gels for pilot investigation were ran as described in section 4.4 using pH 4-7 IPG strips from Ettan™ (GE Healthcare, Bucks, UK) for isoelectric focusing. Cropping of gels was performed prior to importing into DeCyder software in order to minimise errors in the spot matching and allow the algorithm to function optimally. To gain maximum sensitivity, the algorithm was told to estimate 10,000 spots and filter out spot volume $\leq 29,999$, as dust falls into this range.

Typical gels for pilot investigation are presented in Figure 3-22. As can be seen the gel maps show clear and well resolved protein spots, with little to no smearing particularly in the middle portion of the gel, indicative of using a pH4-7 strip. As expected, a small amount of precipitation and smear has occurred at the extreme fringes. Internal standard channels compare well to each other, and average protein intensity spots (relative value of 80,000) fall within 15% of each other between all channels and across gels allowing for accurate quantitation. Some small areas of the gel, corresponding to known areas of structural proteins, have saturated portions so quantitation will not be accurate for those regions. This is normal for DiGE gels as a compromise must be made between being able to visualise low intensity spots without losing more abundant spots.

The overlay of colours for the different channels allows for an initial view of the spread of change over the proteome. The gels were then taken forward for quantitative analysis. The data obtainable from well-run DiGE gels, such as those above, is considerable, regarding the ability to multiplex and quantitate. About 3000-5000 spots across all the gels are quantifiable, a significant number in a proteomic workflow.

Cortex WKY Salt treated V Medulla WKY Salt treated

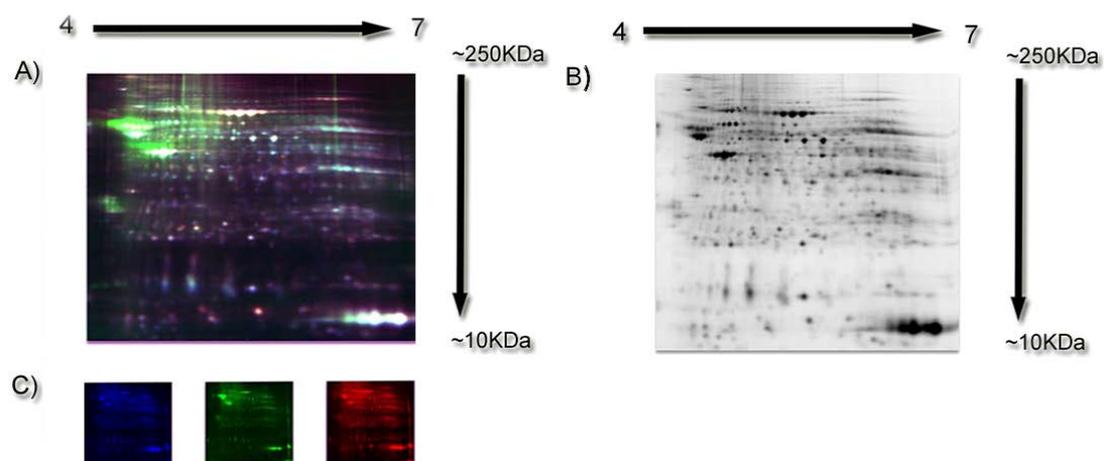


Figure 4-16: Set of typical DiGE gels showing. This particular gel shows A) Fluorescent image with Cy2,3 and 5 for gels Cortex WKY Salt treated and Medulla WKY Salt treated respectively from the pilot investigation B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) Cortex WKY Salt treated (red) Medulla WKY Salt treated. Acidic to basic left to right.

4.6.1.2 Statistical results

4.6.1.2.1 Normality testing

As described in section 3.6.3.3, Table 4-3 shows the percentage of protein spots that were out with the Shapiro-Wilk p-value of 0.05 or less. As can be seen, scores range from 7.86 - 2.99%, the average being 5.06%. Therefore it was found that the assumption of normality is true. This is because with a p-value of 0.05 it would be expected that 5% of spots would fail due to random sampling alone, therefore a result of 5.06% is within an acceptable tolerance.

Dataset	Number of Proteins spots tested	Data type	Percentage spot significance score <0.05
Cortex Wky Salt	233	log ₁₀ SA *	4.72
Cortex Wky Salt	1117	log ₁₀ SA **	6.00
Cortex Wky Salt	806	log ₁₀ SA ***	3.35
Medulla Wky Salt	233	log ₁₀ SA *	5.15
Medulla Wky Salt	1117	log ₁₀ SA **	5.82
Medulla Wky Salt	806	log ₁₀ SA ***	4.96
Cortex Wky Non-salt	278	log ₁₀ SA *	5.04
Cortex Wky Non-salt	1236	log ₁₀ SA **	4.53
Cortex Wky Non-salt	806	log ₁₀ SA ***	4.09
Medulla Wky Non-salt	278	log ₁₀ SA *	6.12
Medulla Wky Non-salt	1236	log ₁₀ SA **	6.31
Medulla Wky Non-salt	806	log ₁₀ SA ***	4.47
Cortex Cn Salt	201	log ₁₀ SA *	4.48
Cortex Cn Salt	986	log ₁₀ SA **	4.46
Cortex Cn Salt	806	log ₁₀ SA ***	3.72
Medulla Cn Salt	201	log ₁₀ SA *	2.99
Medulla Cn Salt	986	log ₁₀ SA **	3.45
Medulla Cn Salt	806	log ₁₀ SA ***	3.97
Cortex Cn Non-salt	324	log ₁₀ SA *	7.41
Cortex Cn Non-salt	1333	log ₁₀ SA **	6.53
Cortex Cn Non-salt	806	log ₁₀ SA ***	5.58
Medulla Cn Non-salt	324	log ₁₀ SA *	4.01
Medulla Cn Non-salt	1333	log ₁₀ SA **	5.25

Medulla Cn Non-salt	806	log ₁₀ SA ***	4.59
Cortex SHRSP Salt	465	log ₁₀ SA *	4.30
Cortex SHRSP Salt	2040	log ₁₀ SA **	6.03
Cortex SHRSP Salt	806	log ₁₀ SA ***	5.33
Medulla SHRSP Salt	465	log ₁₀ SA *	3.01
Medulla SHRSP Salt	2040	log ₁₀ SA **	5.15
Medulla SHRSP Salt	806	log ₁₀ SA ***	4.22
Cortex SHRSP Non-salt	219	log ₁₀ SA *	5.48
Cortex SHRSP Non-salt	1247	log ₁₀ SA **	7.86
Cortex SHRSP Non-salt	806	log ₁₀ SA ***	5.58
Medulla SHRSP Non-salt	219	log ₁₀ SA *	6.85
Medulla SHRSP Non-salt	1247	log ₁₀ SA **	6.26
Medulla SHRSP Non-salt	806	log ₁₀ SA ***	5.21

*Spots included had log₁₀SA for all 3 repeats and an 1-way ANOVA score of <0.05 or better

**Spots included had log₁₀SA for all 3 repeats

*** Spots included had log₁₀SA for all 3 repeats and an 1-way ANOVA score of <0.05 or better With missing values replaced

Assessing normality of the log₁₀SA using the Shapiro-Wilk

goodness-of-fit test

Table 4-3 Assessing normality of log₁₀SA using the statistical test Shapiro-Wilk results for the hypertension pilot study

4.6.1.2.2 Homogeneity of Variance

As described in section 3.6.3.3, the Levene’s statistical test was employed using SPSS 17.0.1. This tested the null hypothesis that the sample variances are equal. This was performed across all groups with a full set of repeats. It was then subsequently performed with missing values replaced using k-nearest neighbour (KNN). Table 4-4 displays the results of the Levene’s test. The threshold for rejecting the null hypothesis and concluding that the data is not homogeneous was 0.05 or less. Only 5.97% or 3.97% respectively, of the spots across all groups were not considered to have homogeneous variation. Graphs visually depicting the distribution of variance can be seen in Figure 4-17.

Data Set	Data type	Number of Spots included	% of spots with p value <0.05
Hypertension Pilot Investigation	SA	67	14.93
	log ₁₀ SA	67	5.97
Hypertension Pilot Investigation with missing values replaced.	SA	806	16.00
	log ₁₀ SA	806	3.97

Spots included had complete set of repeats

The P-Value was generated using Levene’s test across groups with each master spot

Table 4-4: Showing the percentage of spots which failed the homogeneity of variance Levene’s test, for the pilot investigation.

Distribution of Variance: Congenic Cortex Non-Salt-treated

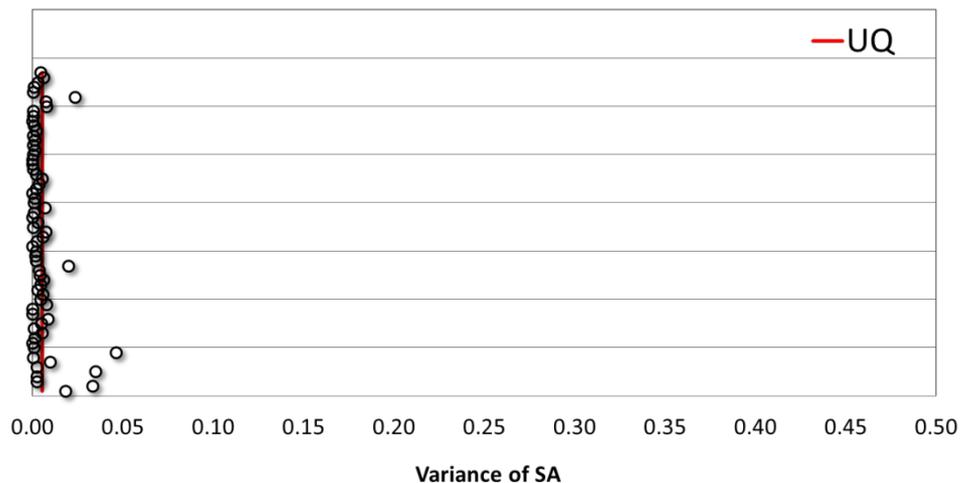


Figure 4-17: Graphical representation of variance for the Congenic Cortex non-salt treated group, for the standardised abundance using 67 different spots and n=3. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.

4.6.1.3 Profile, pie chart and Venn analysis

In order to ascertain the effect of tissues and salt treatment across the three different strains, the profile analysis was divided up into, tissue type and treatment.

4.6.1.3.1 Candidate markers and profile analysis

The first to be considered is the distribution of profiles for cortex non-salt treated. The results can be seen in Figure 4-18 below. The profiles were assigned to their relevant category as per Figure 4-12. The spots were sorted on the basis of significant ANOVA with p-value ≤ 0.05 . This gave 355 spots to be checked manually. Firstly, approximately 36% of spots showed an intermediate effect, relative to the congenic strain possibly alluding to the WKY part of the congenic interval compensating for any hypertensive effect that the SHRSP section of the congenic's genome may be having, giving it the intermediate phenotype. Secondly, approximately 38% of detected spots show maintenance of the WKY feature in the congenic strain and 7% show maintenance of the SHRSP in the congenic strain. It should be noted that, although there is a percentage of 7% for the maintenance of SHRSP phenotype in the congenics, the WKY sections of the genome is significantly smaller in the congenic strain in comparison to the SHRSP section of the genome. Therefore this does allude to the fact that the section of WKY genome in the congenic strain is having an impact on the phenotype of the congenic animal. Whether or not this particular complement of markers is specifically linked to hypertension is not apparent from this data alone, however the 48% of the spots shown show some tentative links between WKY and Congenic strain, which may lead towards a link to hypertension. Equally the other 19%, that elicits towards the section of genes that originated from SHRSP gives information of a profile pattern that could be used to help compared to other treatment of tissue type. It is also interesting to note that 13/19% shows an exaggeration in the congenic phenotype in comparison to both the WKY and SHRSP strains. The identified examples of the profiles given in Figure 4-18 are shown below the pie chart and their positions on the gels are given in Figure 4-19 and the identifications are given in the table in Table 4-5.

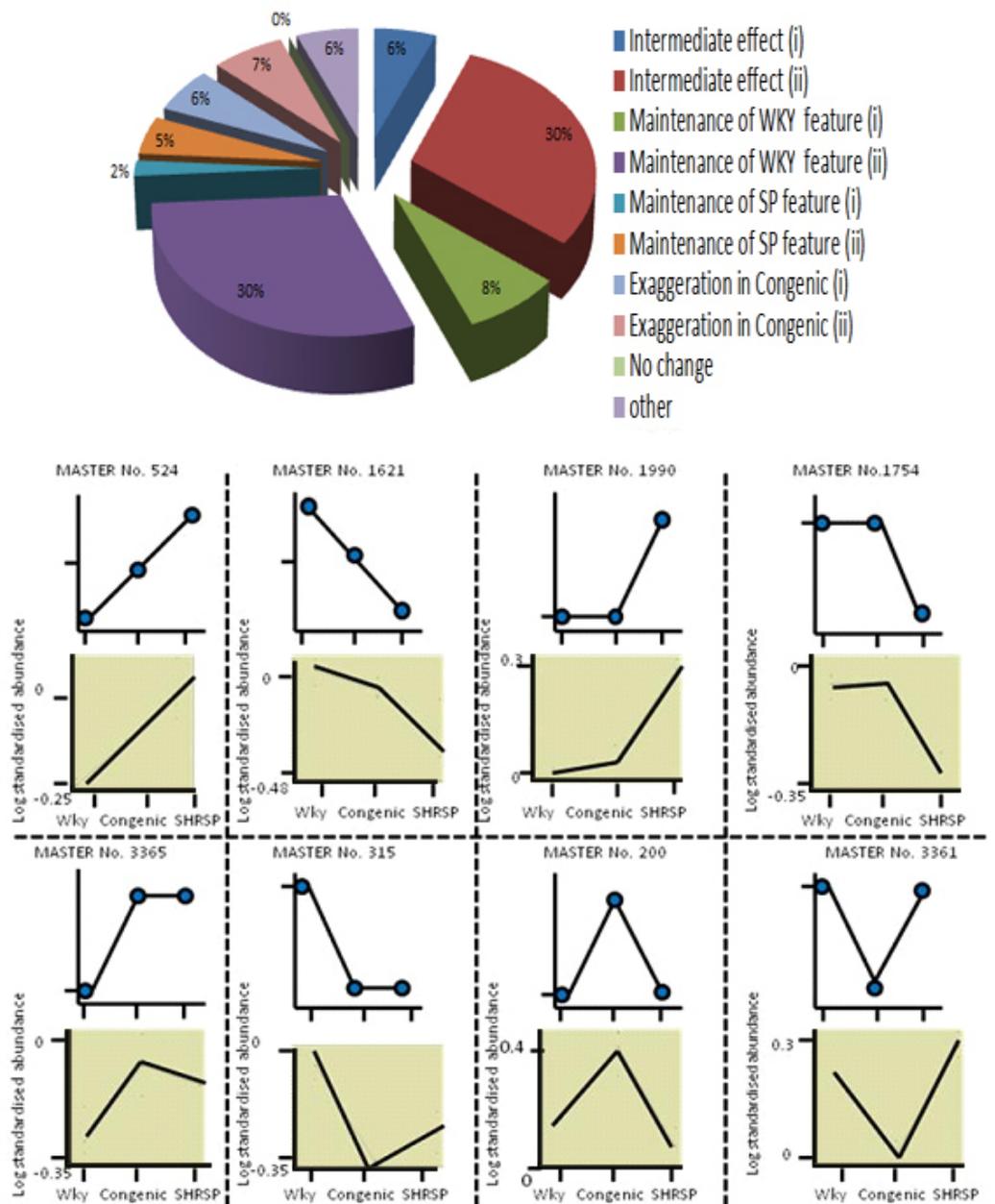


Figure 4-18: Shows profile analysis for Cortex tissue and non-salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12, placed in the relevant categories following profile analysis of pilot investigation. A total of 355 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

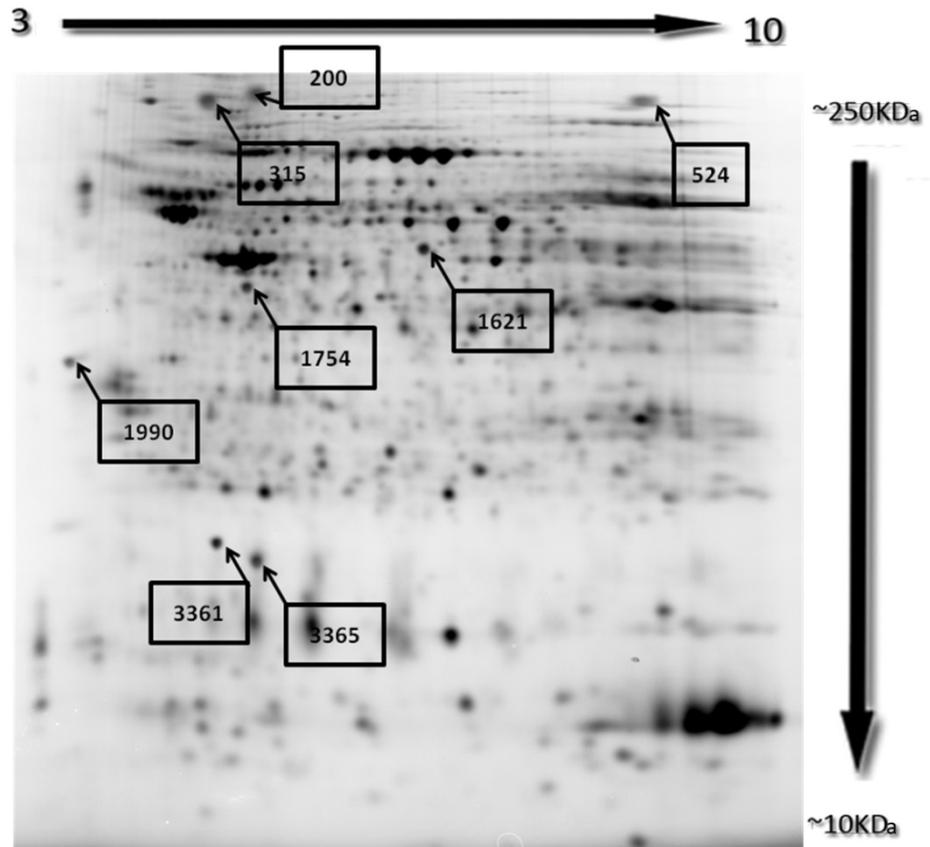


Figure 4-19: Gel map of example spot given in Figure 4-18 for Cortex non-salt treated.

The mitochondrial subunit 5A of Cytochrome C oxidase was identified. This enzyme is part of a large family of enzymes which is located in the inner membrane of the mitochondrion and is crucial in donating electrons to Cytochrome oxidase, which plays a vital role as the terminal oxidase in the electron transport chain in the mitochondria. Deficiency has been shown to affect a number of tissues (Glerum, 2006), including the heart and is linked to cardiomyopathy (Antonicka, 2003).

Spot Master	Identified Protein	Mowse Score	MS/MS Peptides
315	alpha-globin [Rattus sp.]	93	2
315	rCG34342, isoform CRA_b [Rattus norvegicus]	93	2
315	hemoglobin alpha 1 chain [Rattus norvegicus]	93	2
315	hemoglobin alpha 2 chain [Rattus norvegicus]	93	2
524	cytochrome c oxidase, subunit Va [Rattus norvegicus]	45	3
1621	subunit d of mitochondrial H-ATP synthase [Rattus norvegicus]	42	2
1621	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d [Rattus norvegicus]	42	2
1621	rCG33654, isoform CRA_b [Rattus norvegicus]	42	2
1754	PREDICTED: similar to Glutathione S-transferase alpha-4 (Glutathione S-transferase Yk) (GST Yk) (GST 8-8) (GST K) (GST A4-4) [Rattus norvegicus]	63	2
1754	rCG25753, isoform CRA_b [Rattus norvegicus]	63	2
1754	glutathione S-transferase A4 [Rattus norvegicus]	63	2
3361	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus] - Various isoforms identified	87	25
3365	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_g [Rattus norvegicus]	64	2

Table 4-5: Identifications for example spots given in Figure 4-18 and Figure 4-19 and there corresponding mascot scores (MOWSE score). All scores have a p-value < 0.05.

The next profile pattern to be considered is in cortex tissue with salt treatment. The percentage distribution of the profile analysis and example profiles are shown in Figure 4-20, the position of the examples on the gel map are given in Figure 4-21 and the identification found for this set of examples are shown in Table 4-6. As can be seen, the salt treatment has led to a change in the profile distribution. The spots were sorted on the basis of significant ANOVA with p-value ≤ 0.05 . A similar number of spots were included (359 compared to 355 in the cortex non-salt treated), however, the percentages in the salt treated cortex are somewhat different compared to non-salt treated. Only 8% showed an intermediate effect in the congenic (compared 36% in non-salt treated). In

addition, salt treatment has had the effect of exaggerating the congenic strains phenotype in 37% of those markers where the WKY and SHRSP have the same phenotypical response. Compared to the non-salt treatment that is a 3 fold increase. The maintenance of the WKY phenotype holds a similar percentage compared to non-salt treated at 34%, possibly indicating that WKY phenotype is being maintained in a 1/3 of significant markers in the congenic strain. This maintenance of phenotype is interesting, as it has the potential to be mapped back towards the characterised section of genes in the congenic strain. Also, at a similar percentage to the non-salt treatment is the maintenance of phenotype with the SHRSP and congenic strain with 11%. The decision to use salt treatment or non-salt treatment for the main investigation is discussed in section 4.6.1.7.

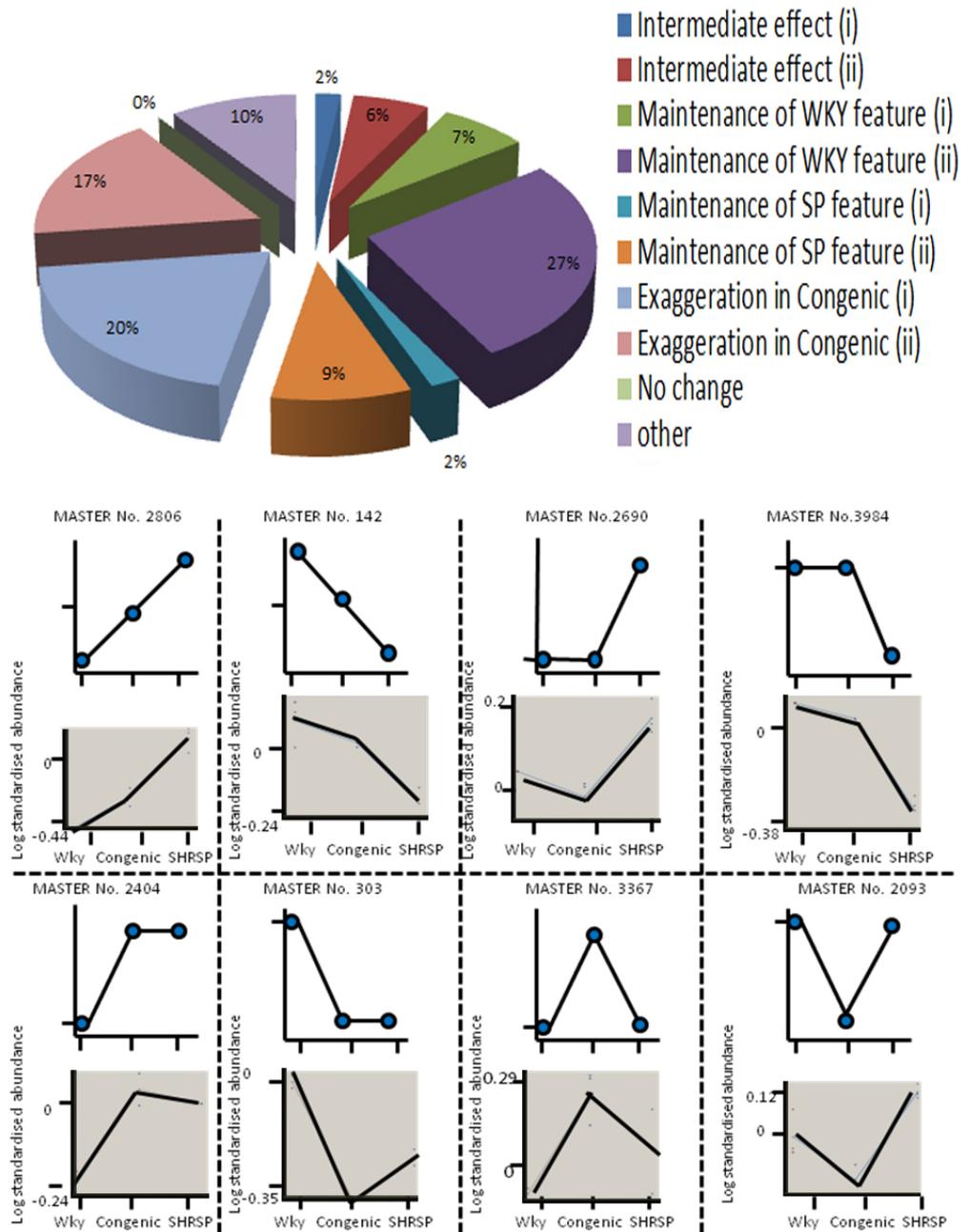


Figure 4-20: Shows profile analysis for Cortex tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of pilot investigation. A total of 359 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

3  10

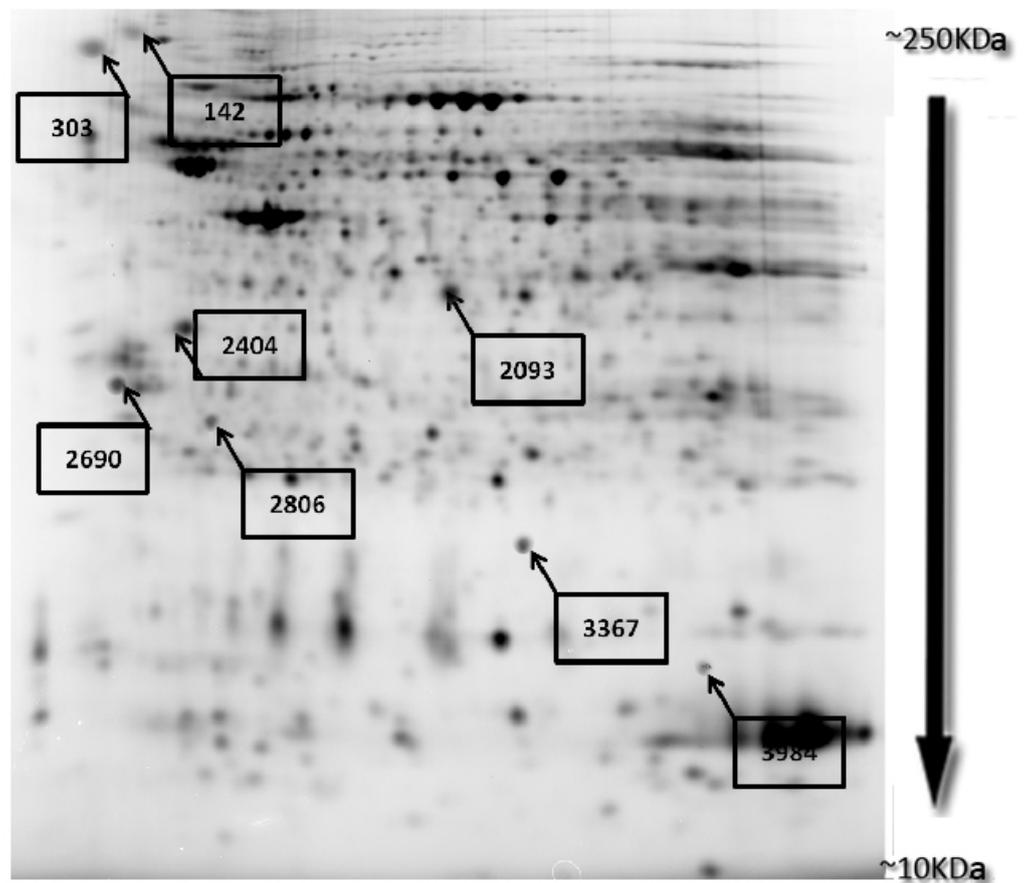


Figure 4-21: Gel map of example spot given in Figure 4-20 for Cortex salt treated.

Spot Master	Identified Protein	Mowse Score	MS Score	MS/MS Peptides	Percentage Coverage
2806	aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus] - various isoforms	194	N/A	4	12
2690	crystallin, zeta [Rattus norvegicus]	65	N/A	3	13
2093	rCG23467, isoform CRA_a [Rattus norvegicus]	64	N/A	52	22
303	major beta-hemoglobin	116	N/A	4	34
303	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	130	N/A	4	34
303	beta-globin [Rattus norvegicus]	130	N/A	4	34
303	hemoglobin beta chain complex [Rattus norvegicus]	130	N/A	4	34
303	rCG39881, isoform CRA_a [Rattus norvegicus]	130	N/A	4	34
142	hemoglobin alpha 1 chain [Rattus norvegicus]	208	N/A	3	37
142	hemoglobin alpha 2 chain [Rattus norvegicus]	208	N/A	3	37
2404	rCG23467, isoform CRA_a [Rattus norvegicus]	67	N/A	52	24
3367	liver annexin-like protein [Rattus norvegicus]	59	N/A	1	2
3367	plasma glutamate carboxypeptidase [Rattus norvegicus]	59	N/A	1	2

Table 4-6: Identifications for example spots given in Figure 4-20 and Figure 4-21 and their corresponding mascot scores (MOWSE score). All scores have a p-value < 0.05.

The next profile distribution to be considered is medulla tissue with non-salt treatment. The percentage distribution of the profile analysis and example profiles are shown in Figure 4-22, the position of the examples on the gel map are given in Figure 4-23 and the identification found for this set of examples are shown in Table 4-7. Once again the data was sorted on the basis of significant ANOVA with p-value ≤ 0.05 . This time approximately 1/3 fewer spots were included for manual profile analysis with only 238 spots satisfying the criteria for selection. In comparison to the salt treatment in the medulla tissue (discussed below), the percentages are somewhat different. Additionally, they also are different in comparison to the cortex tissue equivalent of non-salt treatment. The intermediate effect, maintenance of Congenic with WKY, maintenance of SHRSP and congenic and exaggeration of congenic are evenly spread with 22%, 20%, 26% and 26% respectively. This is different to both medulla salt treated and cortex non-salt treated, however, it is closer present by cortex non-salt treated than the other groups. It therefore appears that the profile distribution of the strains gets more affected by salt treatment than that of tissue type.

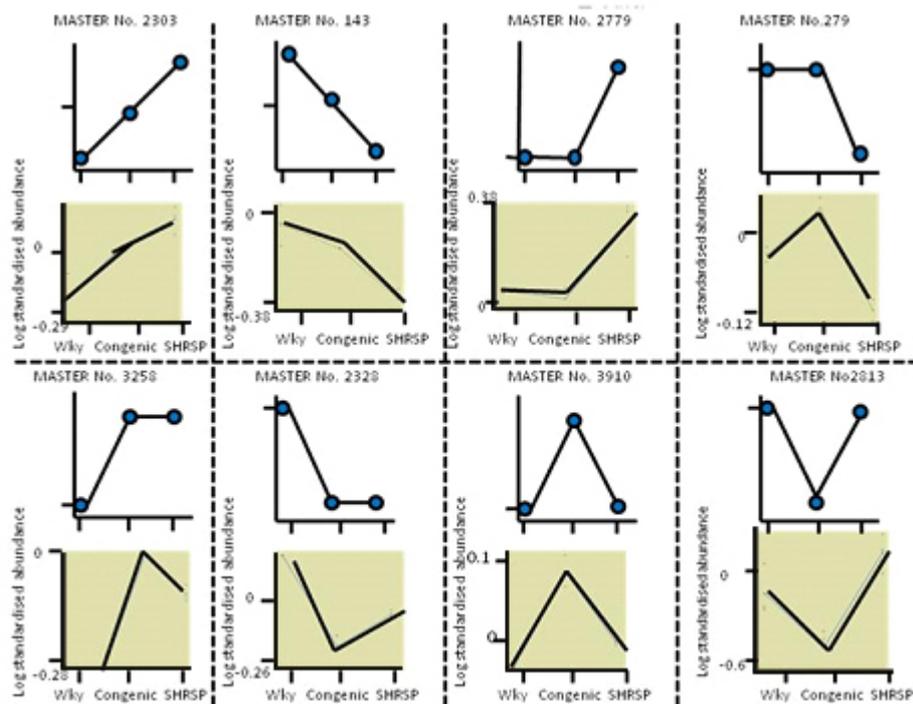
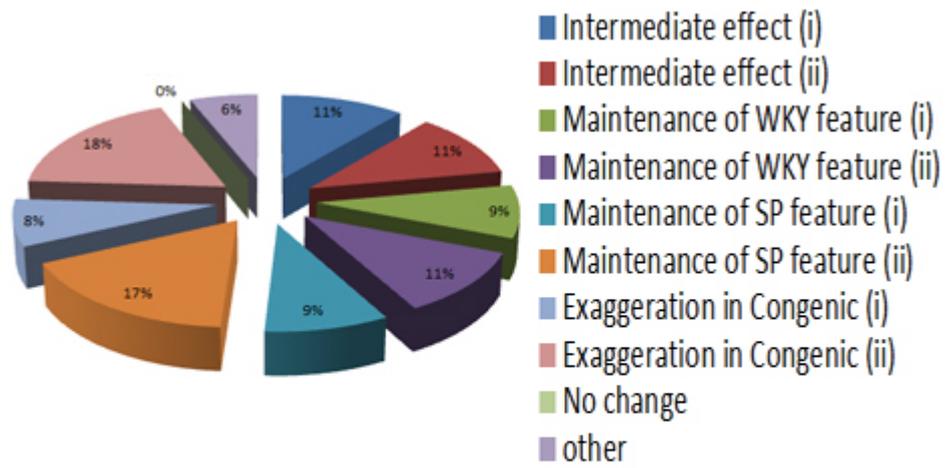


Figure 4-22: Shows profile analysis for medulla tissue and non-salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. Figure 3-5 placed in the relevant categories following profile analysis of pilot investigation. A total of 238 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

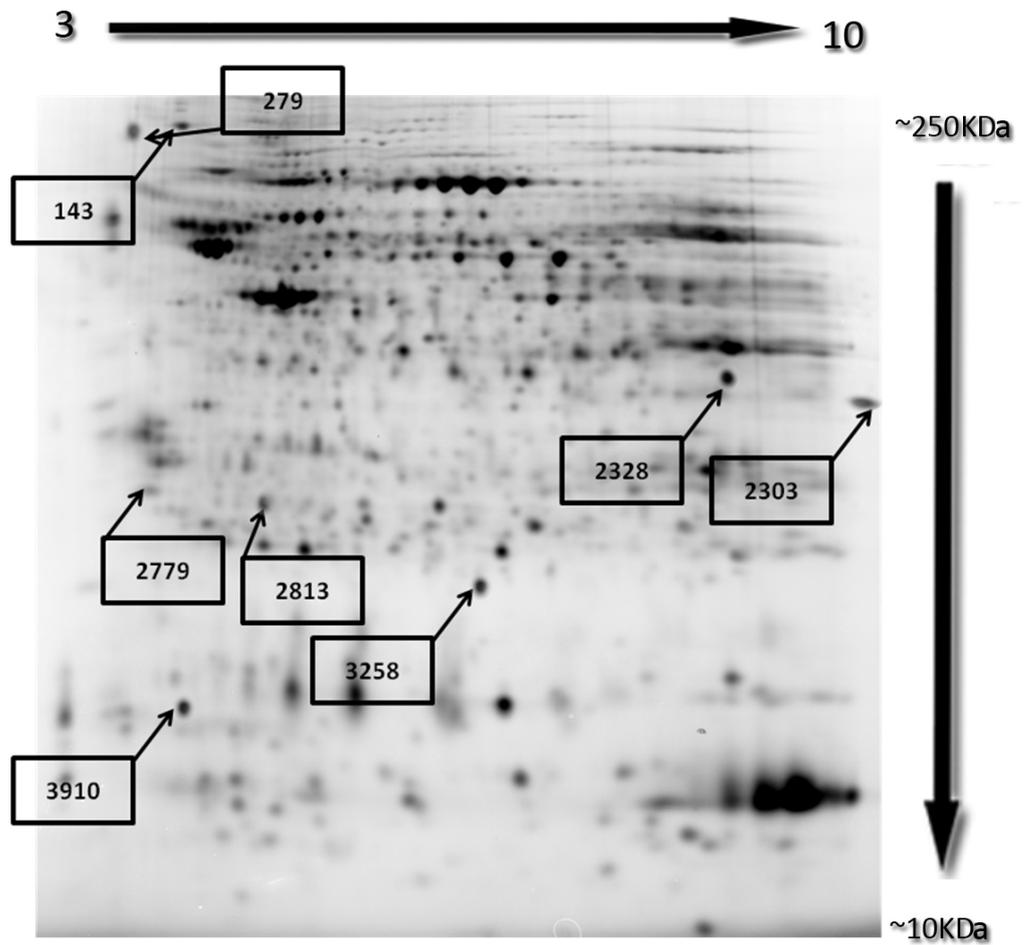


Figure 4-23: Gel map of example spot given in Figure 4-22 for medulla non-salt treated

Spot Master	Identified Protein	Mowse Score	MS/MS Peptides	Percentage Coverage
2303	rCG23467, isoform CRA_a [Rattus norvegicus]	63	51	24
279	alpha-globin [Rattus sp.]	60	2	34
279	rCG34342, isoform CRA_b [Rattus norvegicus]	60	2	21
279	hemoglobin alpha 1 chain [Rattus norvegicus]	60	2	21
279	hemoglobin alpha 2 chain [Rattus norvegicus]	60	2	21
3258	mitochondrial aldehyde dehydrogenase 2 [Rattus norvegicus]- various forms	52	1	2
143	hemoglobin alpha 1 chain [Rattus norvegicus]	246	3	37
143	hemoglobin alpha 2 chain [Rattus norvegicus]	246	3	37
2813	Chain A, Crystal Structure Analysis Of Recombinant Rat Kidney Long- Chain Hydroxy Acid Oxidase	64	13	29

Table 4-7: Identifications for example spots given in Figure 4-22 and Figure 4-23 and there corresponding mascot scores (MOWSE score). All scores have a p-value < 0.05.

The next profile distribution to be considered is medulla tissue with salt treatment. The percentage distribution of the profile analysis and example profiles are shown in Figure 4-24, the position of the examples on the gel map are given in Figure 4-25 and the identification found for this set of examples are shown in Table 4-8. Once again the data was sorted on the basis of significant ANOVA with p-value ≤ 0.05 . This meant 219 spots were included in a similar number to medulla non-salt above. Only 8% showed an intermediate effect in the congenic this is consistent with the cortex salt treated result of 8%. In fact, except for the total number of spots included, the percentage profile distribution is very similar to the cortex salt treated profile distribution.

The maintenance of the congenic phenotype with the WKY phenotype is 31% compared with 34%, indicating that medulla and cortex behave similarly in this respect. Also, at a similar percentage is the maintenance of phenotype with the SHRSP and congenic strain being 7% compared with 11%. The exaggeration

of the congenic feature is even higher at 47%. The similarities in percentage express the possibility of a similar process occurring in the two tissue types described.

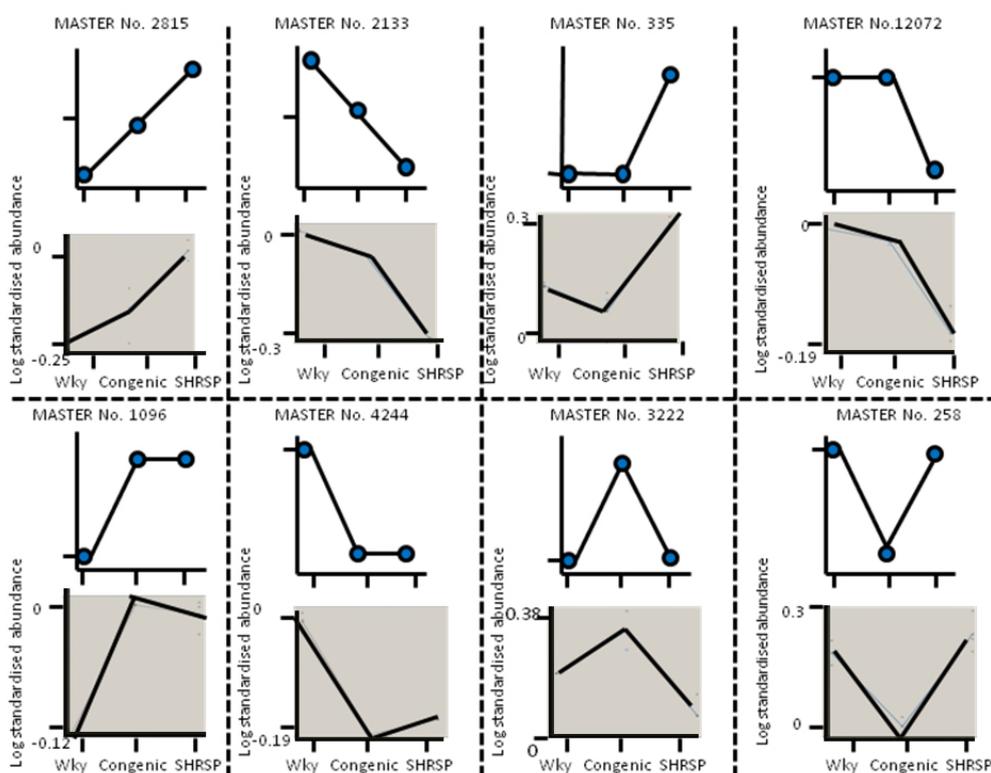
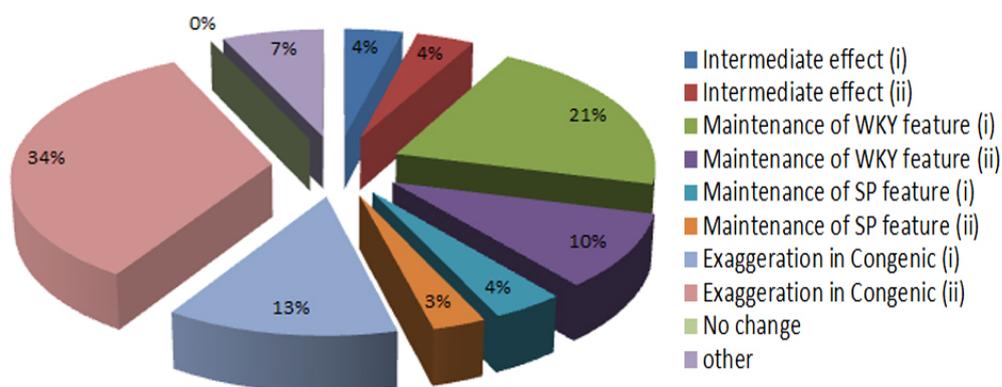


Figure 4-24: Shows profile analysis for medulla tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of pilot investigation. A total of 219 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

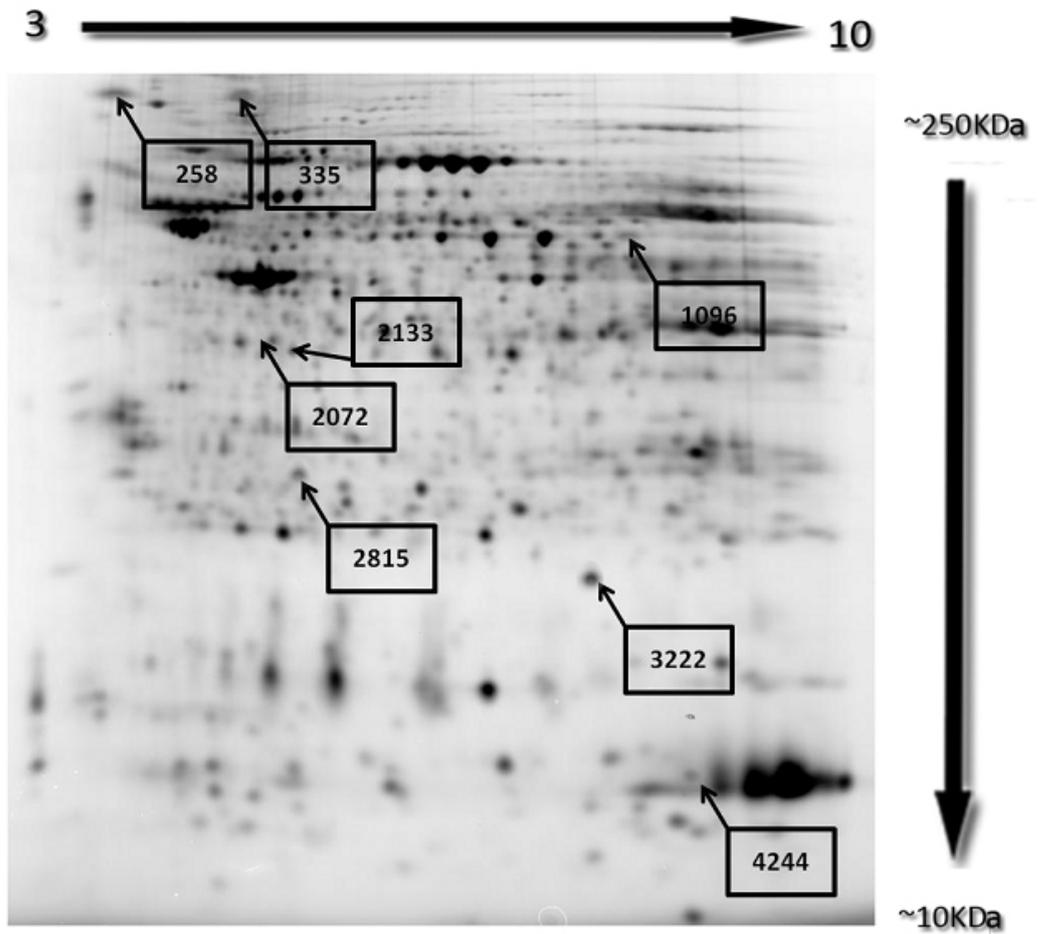


Figure 4-25: Gel map of example spot given in Figure 4-24 for medulla salt treated.

Spot Master	Identified Protein	Mowse Score	MS/MS Peptides	Percentage Coverage
2072	phosphoglycerate mutase 1 [Mus musculus]	79	4	22
335	alpha-globin [Rattus sp.]	37	2	34
335	rCG34342, isoform CRA_b [Rattus norvegicus]	37	2	21
335	hemoglobin alpha 1 chain [Rattus norvegicus]	37	2	21
335	hemoglobin alpha 2 chain [Rattus norvegicus]	37	2	21
2815	aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus]	68	25	12
258	alpha-globin [Rattus sp.]	104	2	34
258	rCG34342, isoform CRA_b [Rattus norvegicus]	104	2	21
258	hemoglobin alpha 1 chain [Rattus norvegicus]	104	2	21
258	hemoglobin alpha 2 chain [Rattus norvegicus]	104	2	21
3222	PREDICTED: similar to sarcoma antigen NY-SAR-41 [Rattus norvegicus]	64	43	27

Table 4-8: Identifications for example spots given in Figure 4-24 and Figure 4-25 and their corresponding mascot scores (MOWSE score). All scores have a p-value < 0.05.

To summarise, it appears that the profile distributions are more affected by salt treatment, in comparison to tissue type as the profile percentage distribution changes when salt treatment applies and has a greater effect on those percentages as tissue types do not maintain the same pattern within the cortex and medulla groups. Additionally, it appears that within the cortex, an area that contains the major components and the majority of the functional unit, the nephron, the WKY portion in the congenic interval is causing the maintenance of the WKY phenotype more predominantly than the SHRSP phenotypes.

4.6.1.4 Pie chart analysis

Pie charts have been used to visualise the proportion and direction of changes occurring between tissue type and treatment of salt or non-salt.

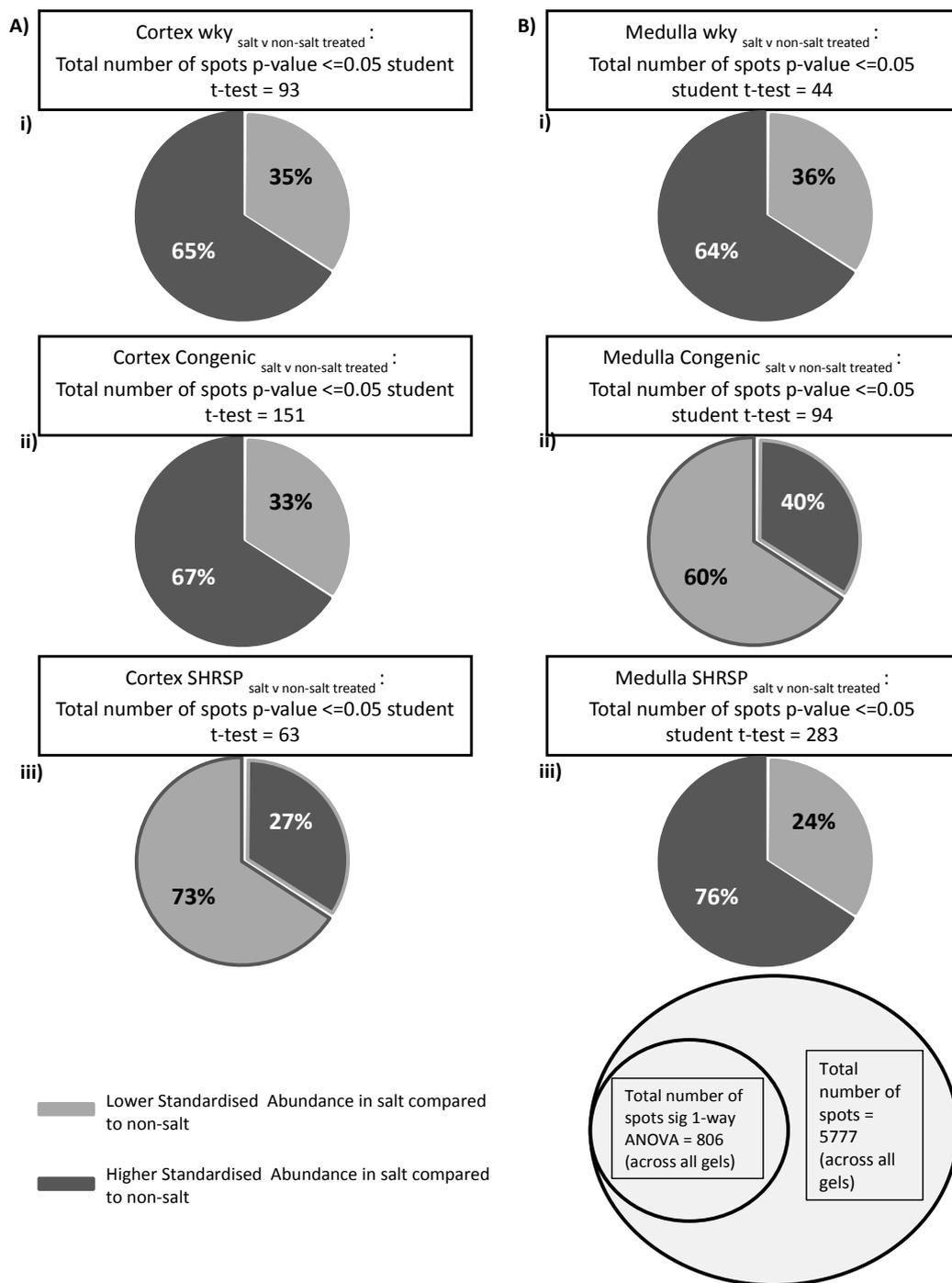


Figure 4-26: Comparison of salt and non-salt treatment in both cortex and medulla within the same strain. Pie charts show the direction of change.

Shown in Figure 4-26 is a pie chart analysis of the Student's t-tests run to show significant differences between salt and non-salt treatment. First spots were sorted using 1-way ANOVA with p-value = 0.05 as a criteria. This included a total of 806 spots for analysis. Considering cortex tissue type first, it can be

seen that a higher proportion of markers in salt treatment caused a higher standardised abundance in salt compared to no salt in both WKY and Congenic strain but the vice versa in SHRSP. This means that the treatment of salt caused higher abundance of markers in WKY and Congenic. Although it cannot be said that these markers are connected to hypertension, the increase in abundance of these possible markers is consistent and in line with other studies finding into the effect of salt on hypertension (Koga et al., 2008, Fountoulakis, 2001, Meneton et al., 2005, Graham et al., 2007, Blaustein et al., 2006). It is interesting, possibly even counterintuitive, that the majority of the significant difference is abundance of markers in SHRSP strains actually are lower in salt compared to non-salt. In regards to medulla tissue type, the proportions do not exactly follow the same pattern as in cortex. WKY strain follows the same pattern in medulla and cortex, but is opposite with regards to congenic. This is a possible indication that completely different phenotypical processes in common with WKY are occurring in the different tissue types. Seeing as the congenic strain is made up from sections of WKY and SHRSP, it would tentatively appear the genes in common with WKY are being more activated in the cortex region of the kidney.

A similar approach was used to look at the difference between cortex and medulla tissue types. This is shown in Figure 4-27. Salt and non-salt will be considered separately. In the salt treatment there appears to be no real pattern between strains for differences in tissue. Both WKY and Congenic have similar proportions, with the significant changes occurring in opposite directions, with congenics showing a higher standardised abundance in cortex tissue as opposed to medulla. It is notable that the number of significant changes is higher between tissues than between treatments, this is contradictory to what was noted in the profile analysis. This shows how important a multi-faceted approach to analysis is in order to try and build a clear picture of relationships. Looking at changes in tissue type in non-salt treatment, the changes seem to be having an approximately similar spread between all strains with higher/lower abundance changes being roughly equal in all three strains.

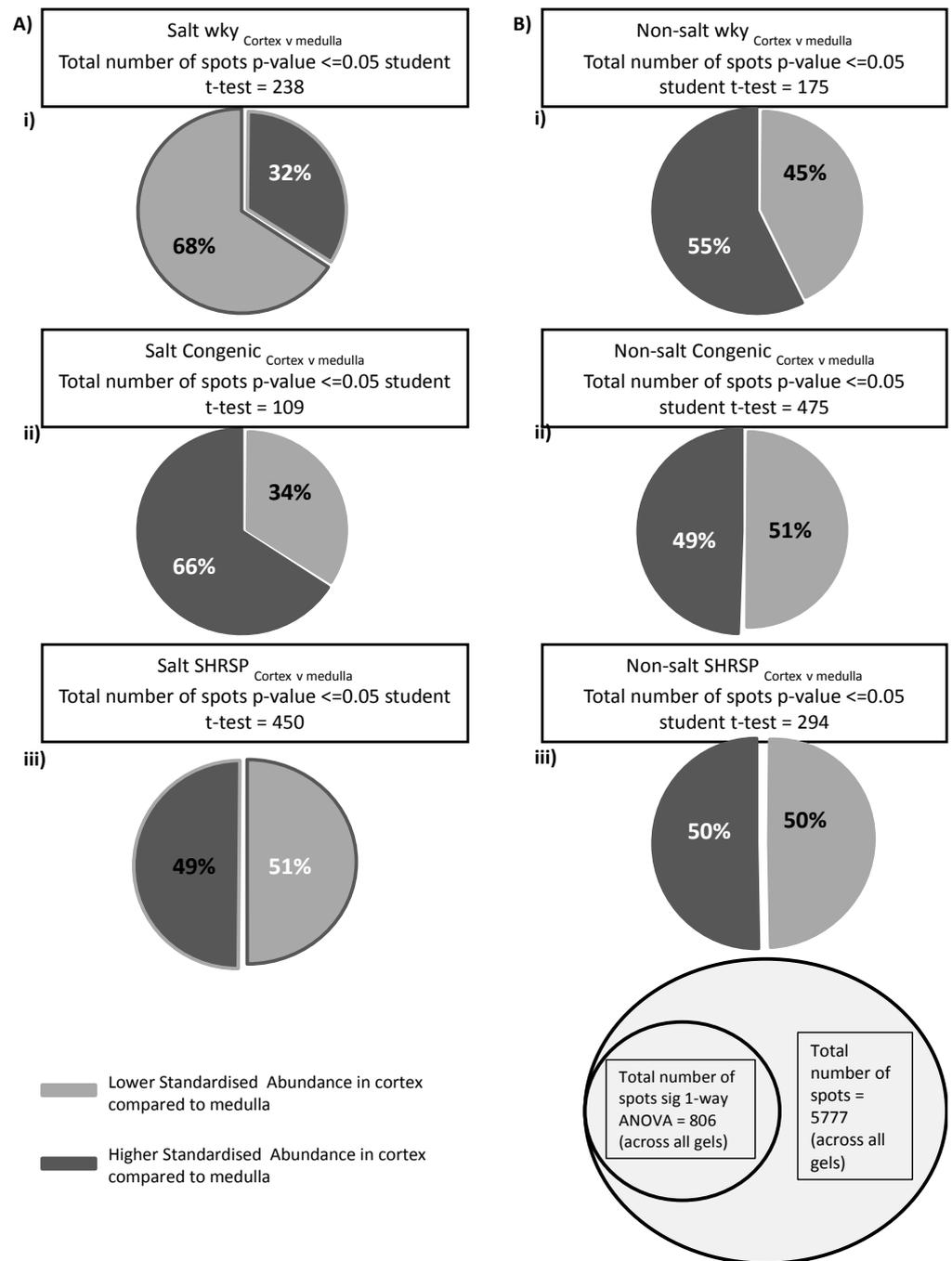


Figure 4-27: Comparison of tissue types in both salt and non-salt treatment within the same strain. Pie charts show the proportion and direction of change.

Now that the proportion changes have been considered it would be useful to scrutinise what changes the tissue types and the treatments have in common. Venn analysis can be performed to achieve this.

4.6.1.5 Venn analysis

Venn analysis has been performed to look at the crossover for the presented statistical t-tests. Venn analysis will be further elaborated in the main investigation as a mode for looking for further candidate markers.

Looking for Significant Changes in Common with salt/no-salt between tissue types:

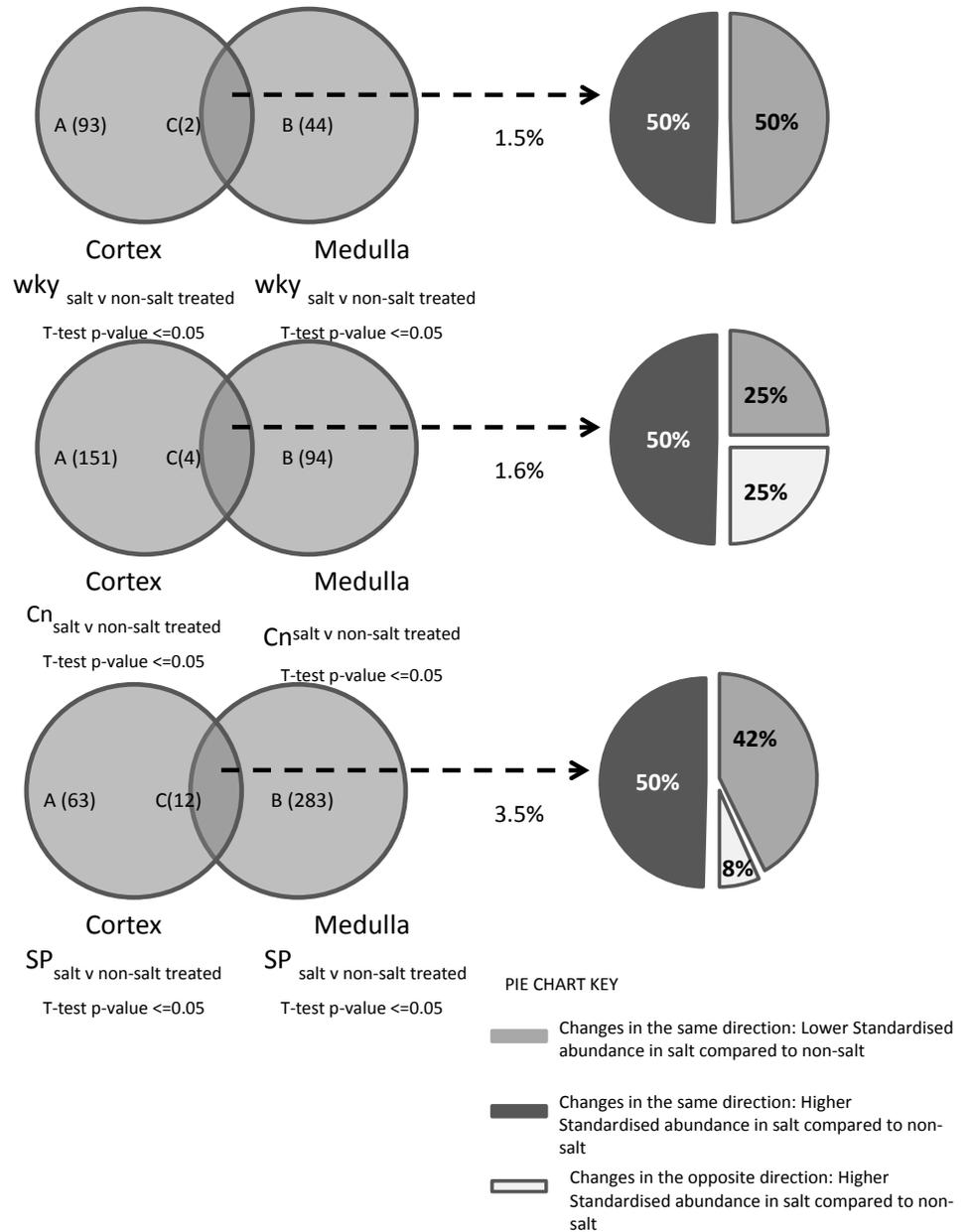


Figure 4-28: Venn analysis showing spots in common between salt/non-salt treatment across tissue within strains. The proportion and direction of the change is shown by pie charts. It can be seen that the crossover of markers between tissue types is small.

In Figure 4-28 it is clearly apparent that there is little cross over, showing that the changes in the abundances in the markers due to salt-treatment are different in the two different tissues. This is extremely interesting as if the markers are different they must be being transcribed from different genes. Of course the gene may come from the same chromosome but the gene must be different in

all the different markers except those that appear in the cross over. This helps to show that the medulla and cortex are quite distinctly different in their tissue type and therefore function. Perhaps even more interesting is that the greatest degree of cross over occurs in the SHRSP strain, with over double the amount of cross over at 3.5%.

The differences between the tissues are further displayed in Figure 4-29. This is shown by the lack of markers in the cross-over section of both three way Venn diagrams. Intriguingly, there seems to be more cross-over in the non-salt treated Venn diagram than the salt treated. The greatest cross-over in both (region F) represents the profiles showing the maintenance of WKY phenotype in the congenic strain. This shows, that in the main investigation, there may be candidate markers of interest with regards to using this congenic model to map genomic and proteomic data.

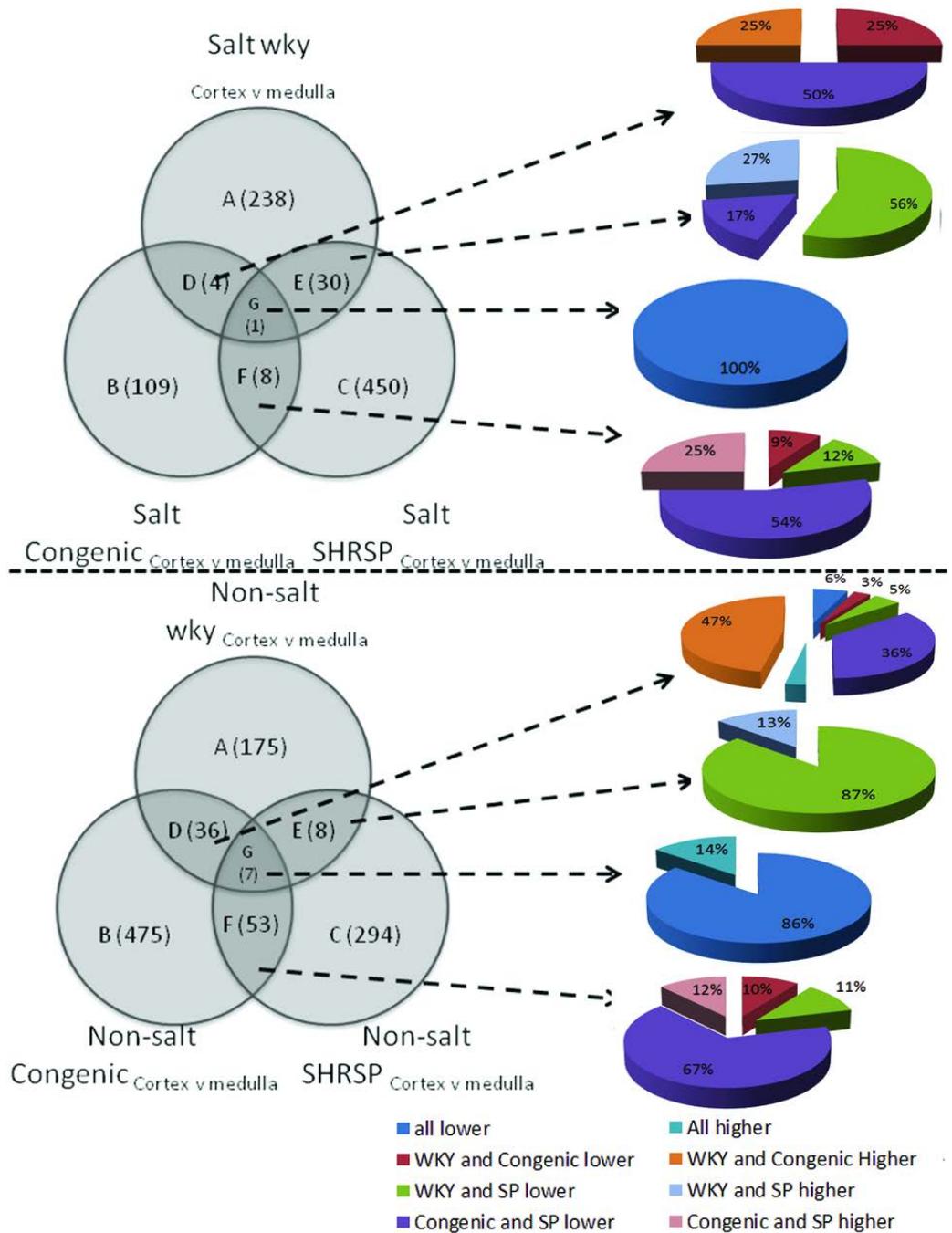


Figure 4-29: Venn analysis comparing cortex and medulla when p-value ≤ 0.05 . The proportion of change in each sector is shown by pie charts. Salt and non-salt are compared separately.

4.6.1.6 Identification in the pilot study

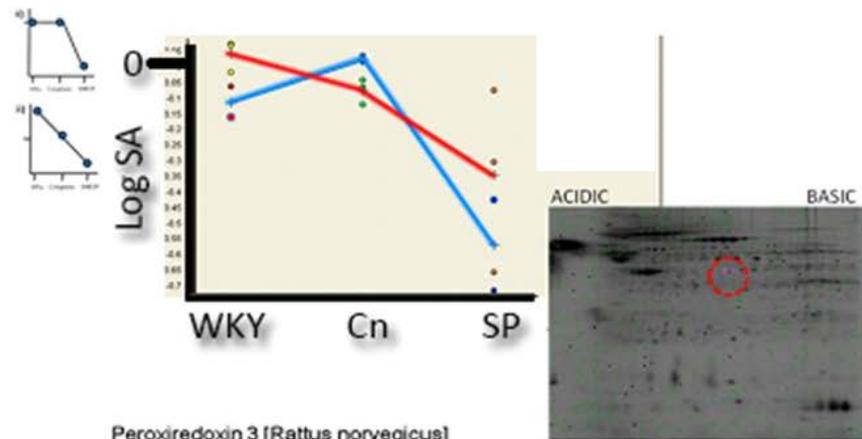
Due to the nature of the pilot study, the analysis lacks the depth shown in the main investigation, as it is simply being used to inform the main study. However, the table given in Table 7-1, in the appendix, is a complete list of all the identifications made in the pilot study. Not all the identifications are linked to the different profiles discussed above, but they have been included as a reference for any future work.

4.6.1.7 Validating the use of salt treatment only in the main investigation

One of the main purposes of the pilot study was to ascertain whether to investigate further salt or non-salt treated samples. It was thought that it might be advantageous in terms of maximising the hypertensive effect and therefore increase the number of markers present that it may be expedient to choose salt treatment over non-salt treatment. Upon data mining, it started to become apparent that the salt treatment was starting to change and exaggerate the profiles seen. Identified examples of these exaggerations can be seen below in Figure 4-30.

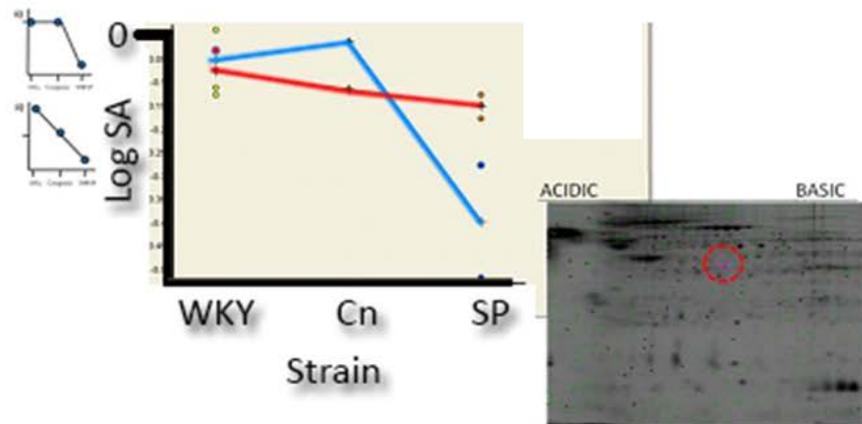
A)

Spot 1637: Subunit of mitochondrial H-ATP synthase [Rattus norvegicus]



B)

Peroxisredoxin 3 [Rattus norvegicus]



C)

aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus]

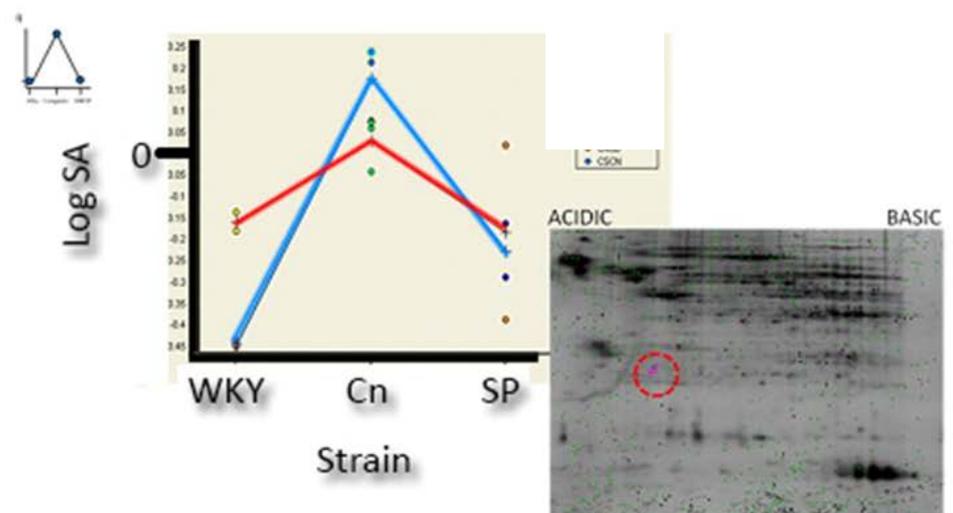


Figure 4-30: Example markers from the pilot study where salt treatments lead to a change in profile shape. A)-C) show how profiles have been changed due to salt treatment Blue line shows non-salt treated samples and red line shows salt treated samples. 1, 2 and 3 on the x-axis are WKY, congenic and SHRSP respectively.

The profile shown in A) has been identified at a subunit of mitochondrial H-ATP synthase. As can be seen, the non-salt treated profile shows that WKY and congenic strains are exhibiting profiles that may show a possible connection to chromosome 2, as the congenic is displaying the same phenotypical response as WKY. This indicates that it may be linked to the same region of genetic code that has been selected for in the congenic strain. Without any further validation, it is noted that this assertion is speculative. A subunit of mitochondrial H-ATP synthase has been implicated in a number of disease processes, including neurotoxicity in the brain and Alzheimer's disease (Sergeant et al., 2003) , and renal cell carcinomas (Yusenko et al., 2010). It can also be seen from Figure 4-30 A) that the effect of salt treatment has been to exaggerate the phenotype and bring it closer to a SHRSP response in comparison to WKY and therefore the congenic interval is not maintaining the WKY phenotypical feature. This would not have become apparent, had salt treatment not been performed. A similar situation occurs in B) which was identified (spot 524) at Peroxiredoxin 3. Peroxiredoxin three has also been linked to a number of conditions and knock out of the gene has been shown to render the heart vulnerable to reperfusion injuries as peroxiredoxins are proteins which help control the level of hydrogen peroxide and hydroperoxides which are produced in reperfusion injuries (Nagy et al., 2006). Once again, the exaggeration of the salt treatment reveals the change in relationship between WKY, congenic and SHRSP strains. The case presented in C) is somewhat different. The profile maintains its shape and therefore relationship. However, once again shows the exaggeration that salt treatment has caused, identified as part of the Aldo-Ketoreducace family. This class of protein has been shown to play a role in vascular dysfunction and the oxidation and reduction of other enzymes (Hwang et al., 2002).

The decision was therefore made, in order to increase the number of markers available and to see the profile in as close to hypertensive phenotype as possible, to use salt-treated animals in the main investigation.

4.6.3 Main investigation

The main investigation is going to focus on the following:

- Looking for candidate markers in hypertension using true biological replicates.
- Look at the differences found in medulla and cortex in hypertension.
- Use multifaceted analytical approach.

4.6.3.1 Typical gels

As with the pilot investigation Gels, the main investigation gels were run as described in methods using pH 4-7 IPG strips from Ettan™ (GE Healthcare, Bucks, UK) for isoelectric focusing. Cropping of gels was performed prior to importing into DeCyder software, in order to minimise errors in the spot matching and allow the algorithm to function optimally. To gain maximum sensitivity, the algorithm was told to estimate 10,000 spots and filter out spot volume $\leq 29,999$, as dust falls into this range.

Typical gels for pilot investigation are presented in Figure 4-31. As can be seen the gel maps show clear and well resolved protein spots with little to no smearing, particularly in the middle portion of the gel, indicative of using a pH4-7 strip. As expected, a small amount of precipitation and smear has occurred at the extreme fringes. Internal standard channels compare well to each other, and average protein intensity spots (relative value of 80,000) fall within 15% of each other between all channels and across gels allowing for accurate quantitation. Some small areas of the gel, corresponding to known areas of structural proteins, have saturated portions so quantitation will not be accurate for those regions. This is normal for DiGE gels as a compromise must be made between being able to visualise low intensity spots without losing more abundant spots.

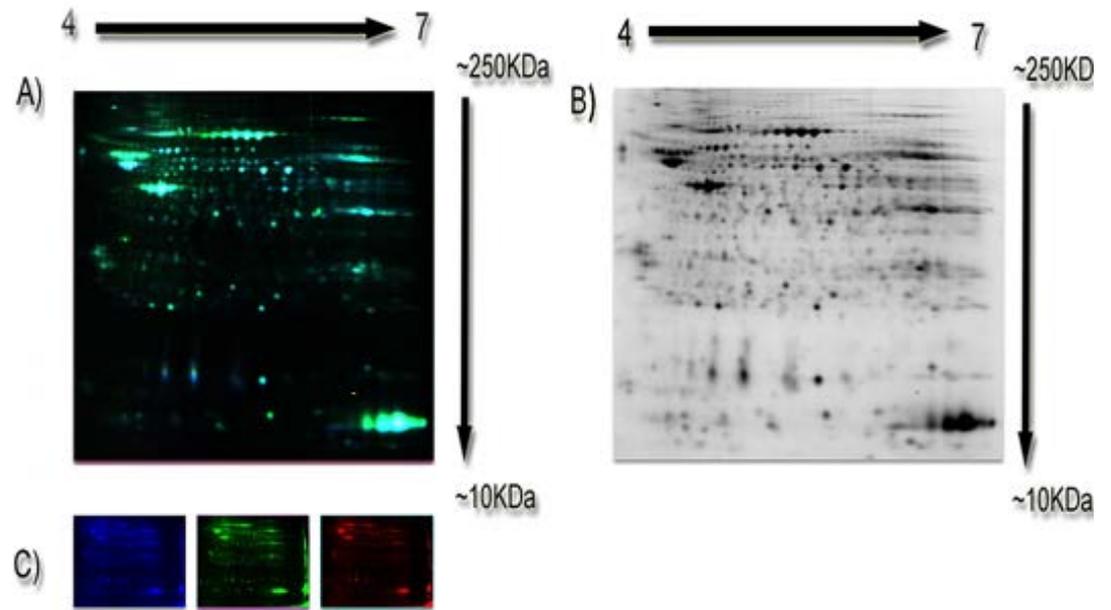


Figure 4-31: Set of typical DiGE gels showing. This particular gel shows A) Fluorescent image with Cy2,3 and 5 for gels Cortex SHRSP Salt treated and Medulla SHRSP Salt treated respectively from the Main investigation B) Cy2 Internal standard C) 3 Cy channels; Cy2 (blue) internal standard, Cy3 (green) Cortex SHRSP Non-Salt treated (red) Medulla SHRSP Non-Salt treated. Acidic to basic left to right.

4.6.3.2 Statistical methods

4.6.3.2.1 Normality Testing

As described in section 3.6.3.3, Table 4-9 shows the percentage of protein spots that were out with the Shapiro-Wilk p-value of 0.05 or less. As can be seen scores range from 6.19 - 2.11%, the average being 4.07%. Therefore it was found that the assumption of normality is true. This is because with a p-value of 0.05 it would be expected that 5% of spots would fail due to random sampling alone, therefore a result of 4.07% is within an acceptable tolerance.

Dataset	Number of Proteins spots tested	Data type	Percentage spot significance score <0.05
Cortex Wky Salt	99	log ₁₀ SA *	5.05
Cortex Wky Salt	1576	log ₁₀ SA **	3.11
Cortex Wky Salt	190	log ₁₀ SA ***	3.68
Medulla Wky Salt	113	log ₁₀ SA *	6.19
Medulla Wky Salt	1609	log ₁₀ SA **	2.67
Medulla Wky Salt	190	log ₁₀ SA ***	2.11
Cortex Cn Salt	76	log ₁₀ SA *	5.26
Cortex Cn Salt	1149	log ₁₀ SA **	4.44
Cortex Cn Salt	190	log ₁₀ SA ***	3.16
Medulla Cn Salt	84	log ₁₀ SA *	3.57
Medulla Cn Salt	1150	log ₁₀ SA **	3.39
Medulla Cn Salt	190	log ₁₀ SA ***	4.21
Cortex SHRSP Salt	84	log ₁₀ SA *	5.95
Cortex SHRSP Salt	1345	log ₁₀ SA **	4.54
Cortex SHRSP Salt	190	log ₁₀ SA ***	3.16
Medulla SHRSP Salt	98	log ₁₀ SA *	4.08
Medulla SHRSP Salt	1431	log ₁₀ SA **	4.96
Medulla SHRSP Salt	190	log ₁₀ SA ***	3.68

*Spots included had log₁₀SA for all 3 repeats and an 1-way ANOVA score of <0.05 or better

**Spots included had log₁₀SA for all 3 repeats

*** Spots included had log₁₀SA for all 3 repeats and an 1-way ANOVA score of <0.05 or better With missing values replaced

Assessing normality of the log₁₀SA using the Shapiro-Wilk goodness-of-fit test

Table 4-9: Normality statistical test results, for main investigation Assessing normality of the log₁₀SA using the Shapiro-Wilk statistical test.

4.6.3.2.2 Homogeneity of variance

As described in section 3.6.3.3, The Levene’s statistical test was employed using SPSS 17.0.1. This was performed across all groups with a full set of repeats. It was then subsequently performed with data with missing values replaced using k-nearest neighbour (KNN). Table 4-10 displays the results of the Levene’s test. The threshold for rejecting the null hypothesis and concluding that the data is not homogeneous was 0.05 or less. Only 3.07% or 4.21% respectively of the spots across all groups were not considered to have homogeneous variation. This also showed that the process of taking the logarithm to the base 10 of the standardised abundance considerably aided the stabilization of data by increasing the homogeneity of the data from 13.41-3.07% or 14.21-4.21%. An example of the graphical visualisation employed in depicting the distribution of variance can be seen in **Error! Reference source not found.**

Data Set	Data type	Number of Spots included	% of spots with p value <0.05
Hypertension Main Investigation	SA	522	13.41
	log ₁₀ SA	522	3.07
Hypertension Main Investigation with missing values replaced	SA	190	14.21
	log ₁₀ SA	190	4.21

Spots included had complete set of repeats

The P-Value was generated using Levene’s test across groups with each master spot

Table 4-10: Showing the percentage of spots which failed the homogeneity of variance Levene’s test, for main study.

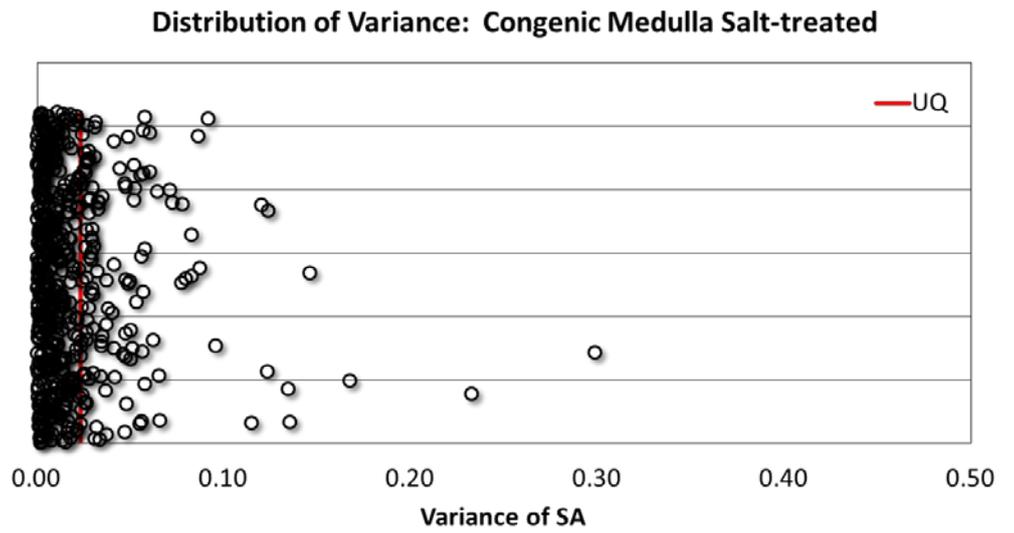


Figure 4-32: Graphical representation of variance for the Congenic medulla salt treated group, for the standardised abundance using 522 different spots and n=3. This is a typical distribution of variance generated across all treatments and time points. This shows how the variance s clustered close to zero with few outliers.

4.6.3.3 Profile and Candidate analysis

As with the pilot investigation, the first area to be considered was to review the profile distribution and display some example markers found when manually checking through the 1-way ANOVA sorted data. A total of 100 spots were included in the profile classifications. This is somewhat lower than the pilot study using technical replicates; however biological replicates often show greater variation. The results for cortex tissue salt treated are shown in Figure 4-33. As can be seen, 35% of markers showed an intermediate effect between strains from WKY, Congenic and SHRSP respectively. 42% of markers exhibit maintenance of WKY features in the congenic strains, the majority of which show a fall in expression in the SHRSP strain (35%). Only 15% of the markers show maintenance of the SHRSP features in the congenic strain about 3 fold less than for WKY, despite the majority of the congenic strains genome being of origin. This profile distribution pattern does not match that of the pilot investigation (cortex salt treated). It is recognised that further replicate will be required in any future work in order to validate patterns. A further 8% of markers showed an exaggeration in the congenic strains response.

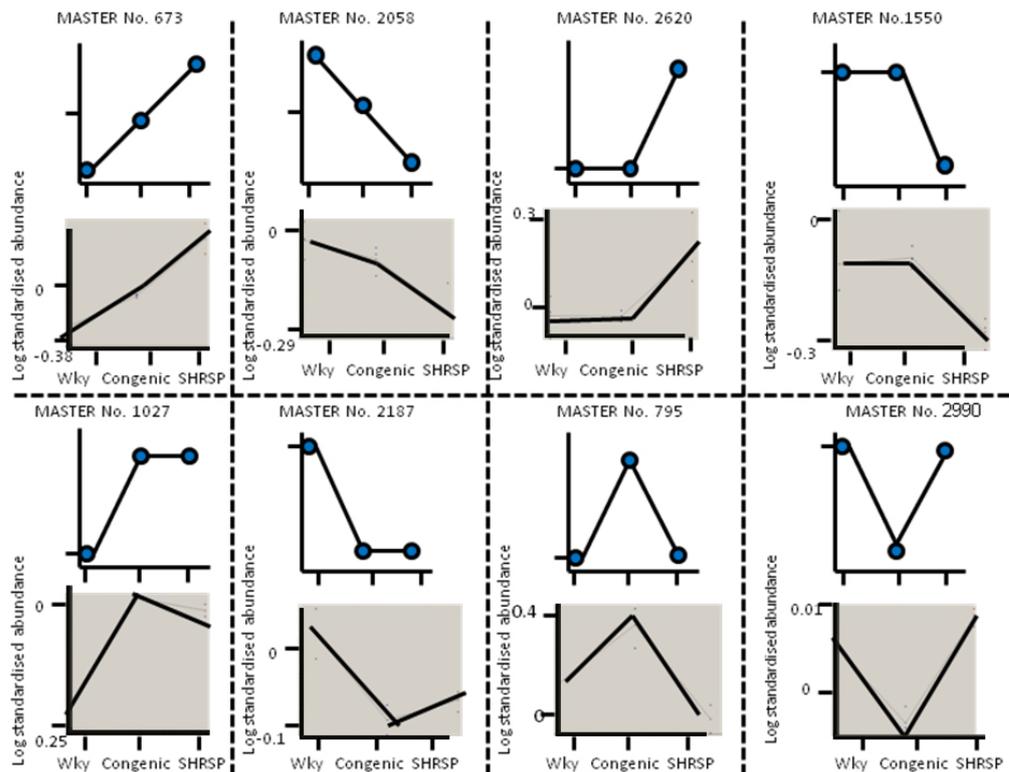
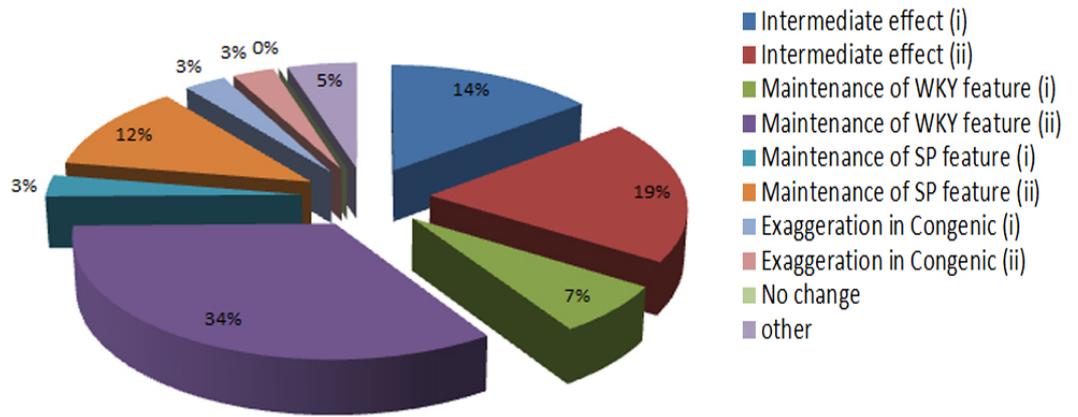


Figure 4-33: Shows profile analysis for cortex tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12 placed in the relevant categories following profile analysis of main investigation. A total of 100 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

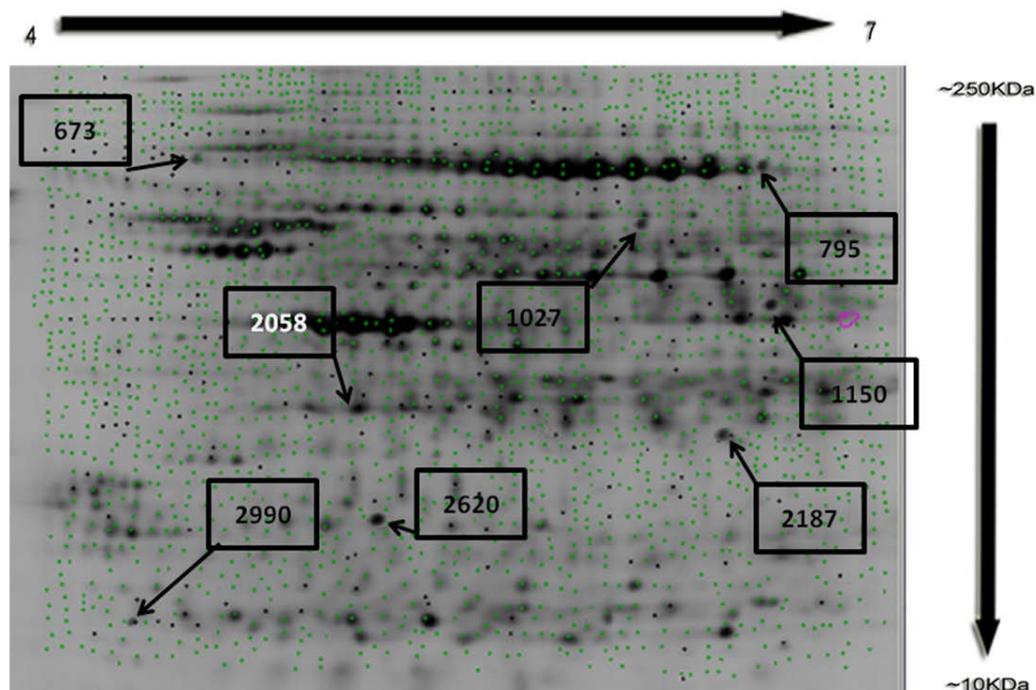


Figure 4-34: Gel map of example spot given in Figure 4-33 for cortex salt treated. Proteins identified are; 1550. rCG25777, isoforms CRA_a / aminoacylase 1. 2058. Regucalcin. 2187. Mercaptopyruvate sulfurtransferase and 2620. Isoamyl acetate-hydrolyzing esterase 1 homolog.

The next profile pattern to be considered is medulla salt treated. Once again spots were sorted using 1-way ANOVA and a criteria of $p\text{-value} = 0.05$. 174 spots fell within this criterion. This is a greater number of spots compared to cortex salted treated above and is the opposite of the pilot study. Where the medulla tissue had lower number of markers in comparison to cortex tissue type. The results are shown in Figure 4-35. A similar pattern is displayed in medulla compared to cortex tissue type. 39% (compared to 35%) showed an intermediate effect in the congenics, 54% (compared to 42%) showed maintenance of WKY phenotype in the congenic strain (with the majority of markers also of lower abundance in the SHRSP strain) and 3% (compared to 15%) for the maintenance of the SHRSP phenotype in the congenic strain. This is where the major difference in the pattern is. Finally only 4% showed an exaggerated effect in the congenic strain.

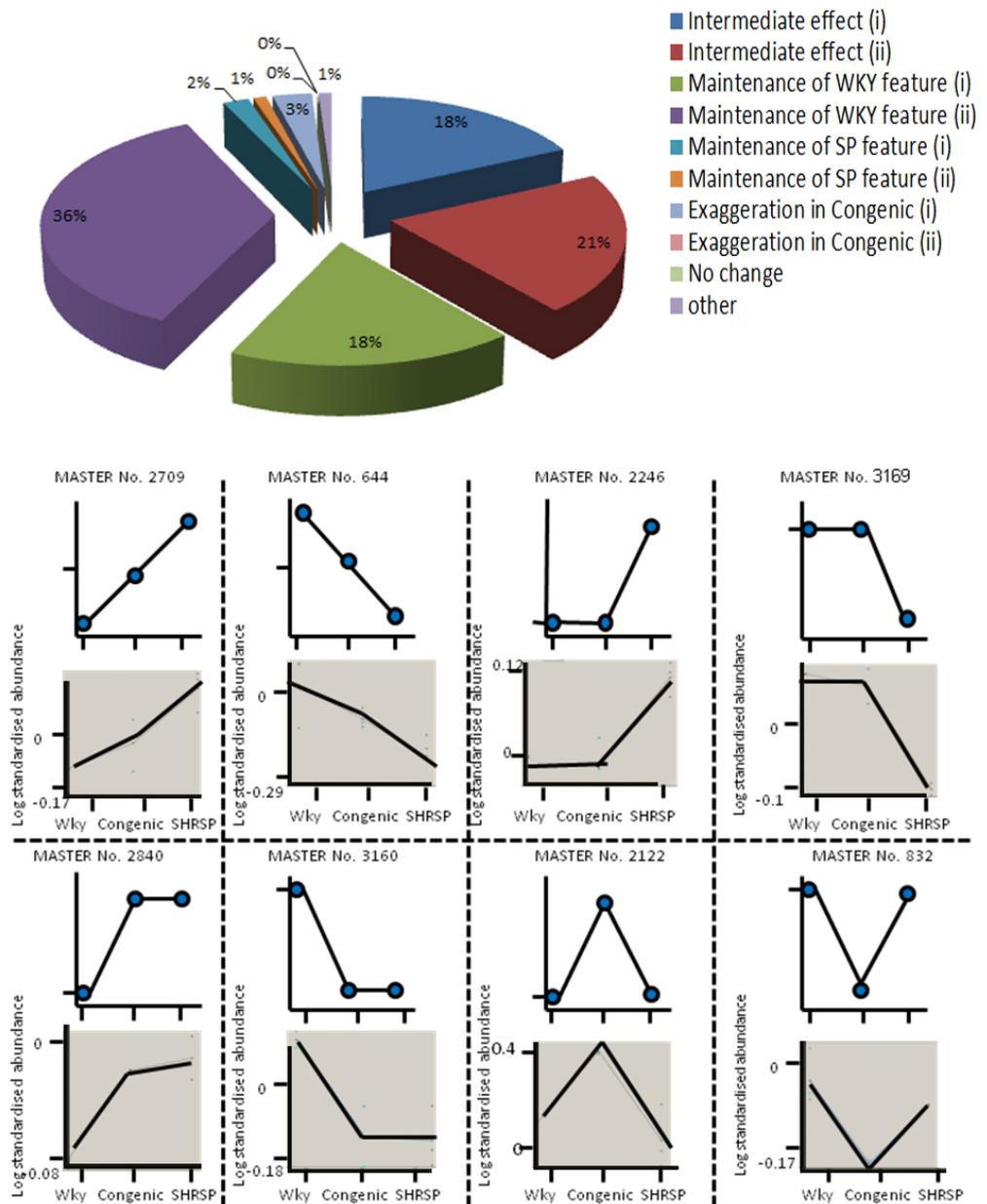


Figure 4-35: Shows profile analysis for medulla tissue and salt treated. The pie chart shows the distribution of spots manually matched experimental profiles with predicted profiles shown in Figure 4-12. **Figure 3-5** placed in the relevant categories following profile analysis of main investigation. A total of 174 spots were included on the basis of a significant 1 way ANOVA score. Example profiles are given under the pie chart.

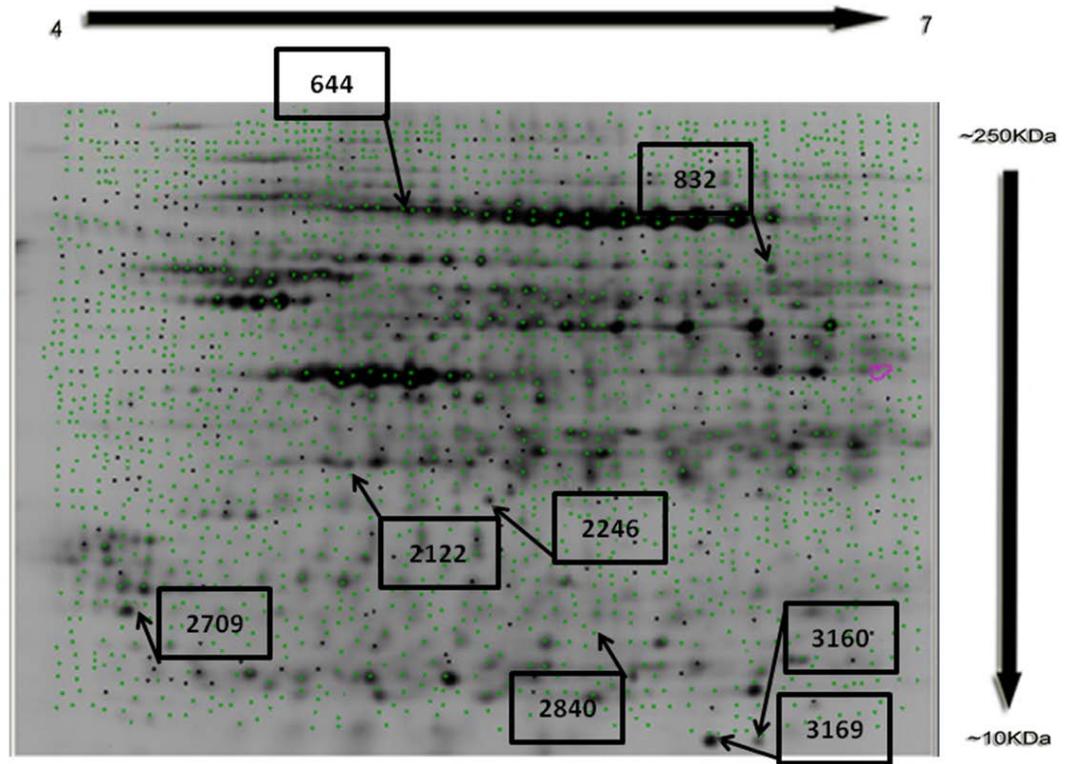


Figure 4-36: Gel map of example spot given in Figure 4-35 for medulla salt treated. Identification of spots are; 644. dnaK-type molecular chaperone hsp72-ps1 / Heat shock protein 8. 2709. 14-3-3 zeta isoform / typtohan 5-monooxygenase activation protein. 3169. Mitochondrial ribosomal protein L51.

4.6.3.4 Pie chart analysis

A pie chart analysis was performed to look for the proportion and direction of changes occurring between cortex and medulla tissue types.

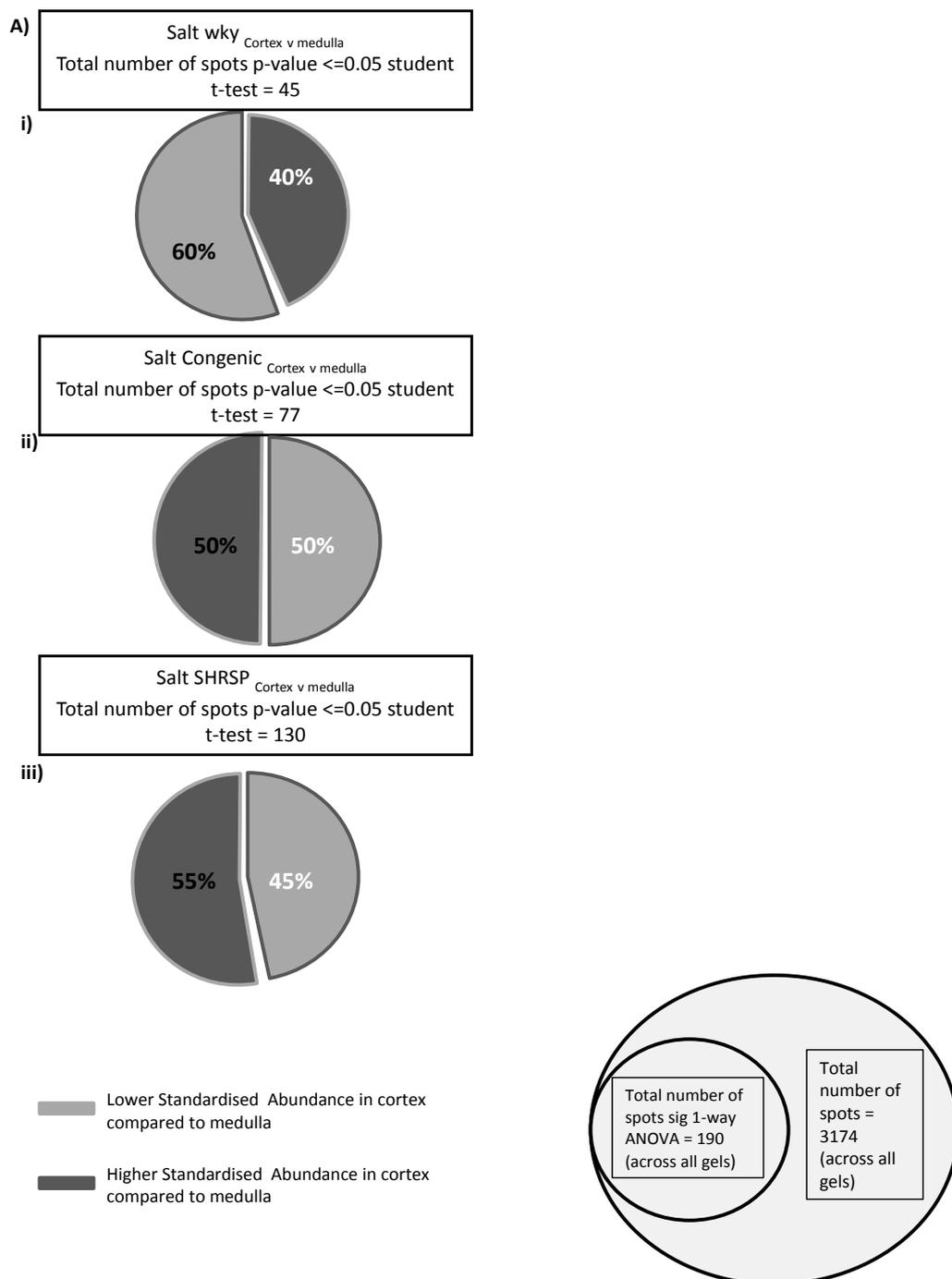


Figure 4-37: Comparison of tissue types within the same strain. Pie charts show both the proportion of the direction of change.

Shown in Figure 4-37 is a pie chart analysis of the Student's t-tests run to show significant differences between salt and non-salt treatment. First spots were sorted using 1-way ANOVA with p-value = 0.05 as a criteria. This included a total of 190 spots for analysis. This shows that the direction in standardised abundance between the two tissues was reasonable, even in the markers with

significant differences. It is interesting to note that the number of markers with a significant difference is higher in the SHRSP strain (approximately 2x) than that of the WKY and Congenic strains.

4.6.3.5 Venn analysis and candidate markers

The Venn diagrams can be used to mine down into the data and to see where any differences between tissues cross over in relation to strains. In Figure 4-38 a three way Venn looks at the cross over of differences between the tissue types in the salt treated samples.

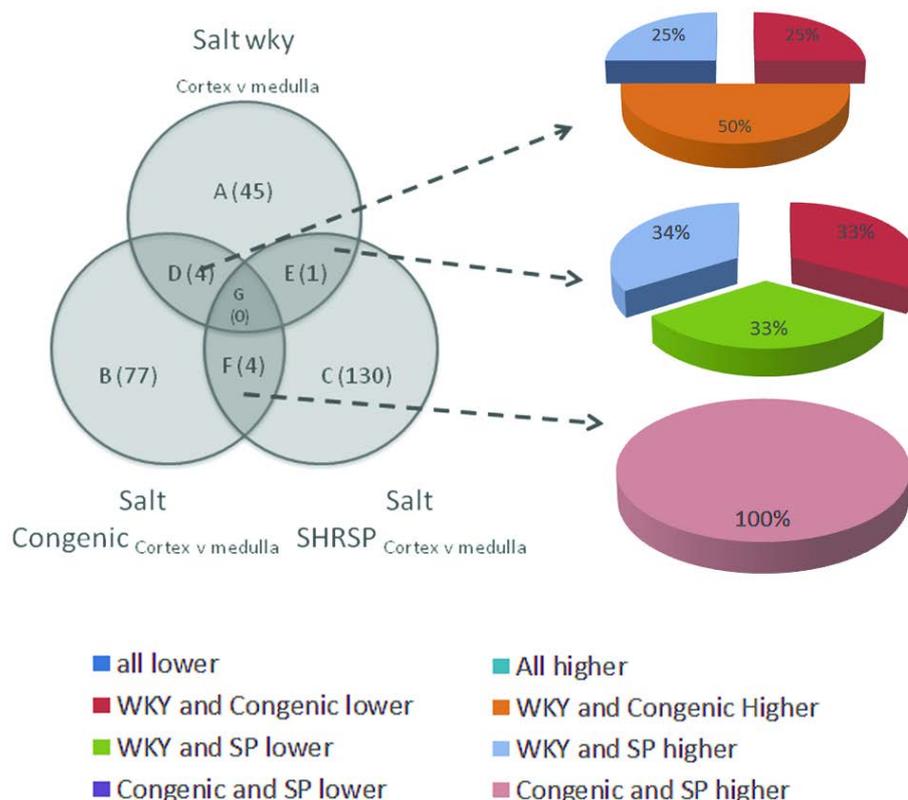


Figure 4-38: Venn analysis comparing cortex and medulla and shown strains when p-value ≤ 0.05 . The proportion of change in each sector is shown by pie charts.

As can be seen, the main investigation displays a similar cross-over to the equivalent diagram in the pilot study. This Venn diagram analysis is extended in Figure 4-39 to show an example of identified markers. In region (F) of Figure 4-38 it is shown that the markers are common between the medulla and cortex and that all are WKY and Congenic strains are higher in abundance in comparison to SHRSP. This is consistent with the Profile analysis above. It appears that the SHRSP causes the drop off in abundance in a number of markers in comparison to the WKY and Congenic strains and therefore the WKY phenotype is maintained in the congenic strains. This would allude to the fact that those markers are possibly linked to a section of the genome that

could be involved in hypertension. One of the markers has been identified as; Predicted: Similar to Actin Cytoplasmic 2 (gamma Actin). This protein comes from a family with many different isoforms and therefore has some highly homologous sections of amino acid coding, therefore identification would need to be validated to find exactly which variant it is. However, Gamma actin has been identified as a possible marker for pulmonary artery remodelling and hypertension (Thakker-Varia et al., 1999). The shape of the profile confirms the statistics and the Venn diagram cross-over as the profile is identical in shape and abundance for both tissue types.

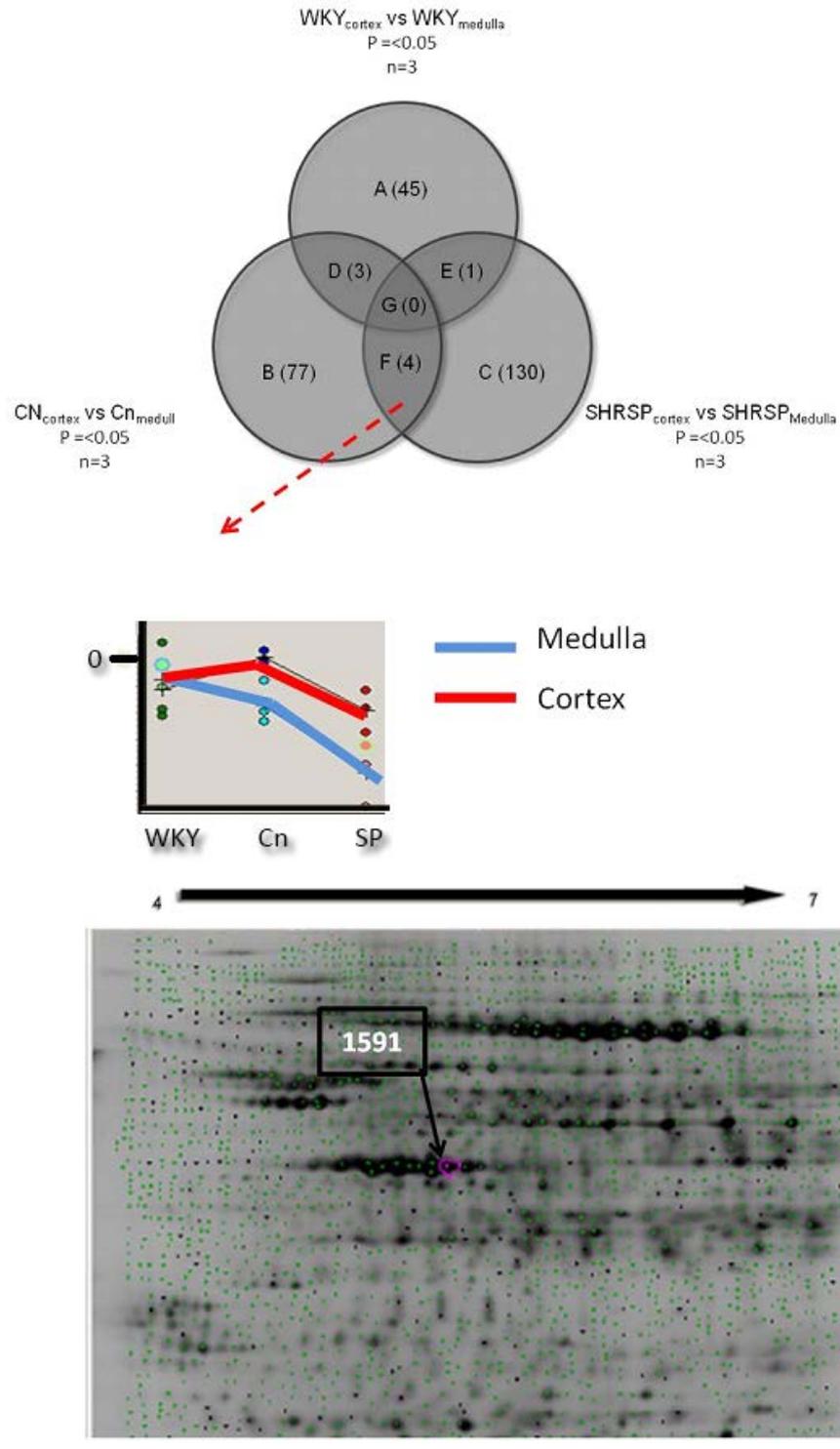


Figure 4-39: Combination of Venn and profile analysis in mining down of data between tissue type. Example and of candidate marker profile is given for sector F that match predicted profile in Figure 4-12. Gel map shows the position and id of possible marker. Identified candidate marker is Predicted: Similar to Actin Cytoplasmic 2 (gamma Actin).

The Venn analysis continues in Figure 4-40. This shows the markers in common between the Student's t-test performed between each strain in cortex tissue type. Possible example markers are given from the cross-over points which have been linked to the profiles given in Figure 4-12. Unfortunately,

none of the markers were identified in this case. It is interesting to note that most of the expression profiles shown for cortex (in the example graphs) do not usually match that for medulla, which confirms much of the cross-over analysis between tissue types.

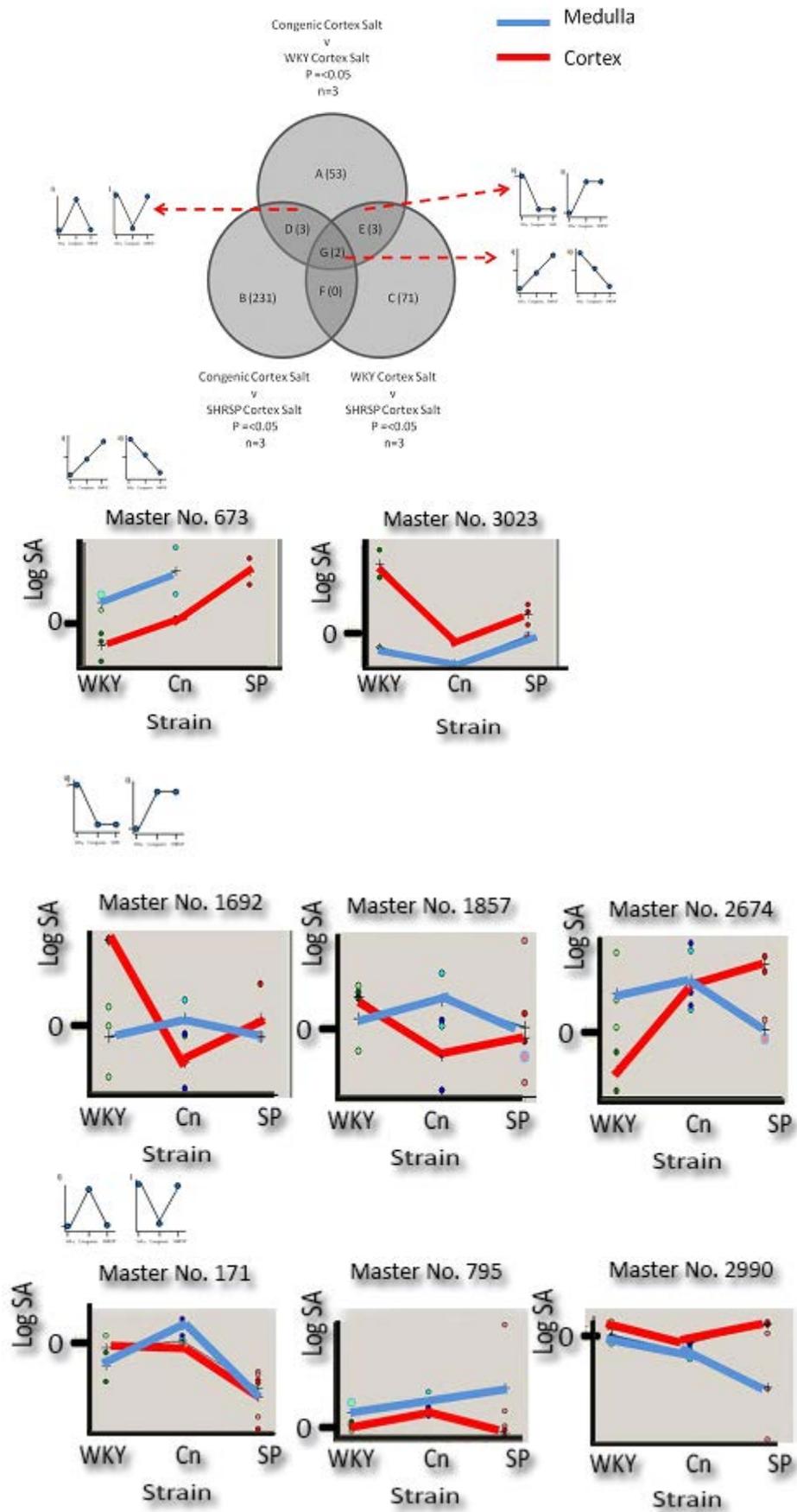


Figure 4-40: Combination of Venn and profile analysis in mining down of data. Example and candidate marker profiles are given for each sector that matches predicted profiles in Figure 4-12. Venn analysis is with cortex tissue however medulla tissue profiles have been overlaid for comparison.

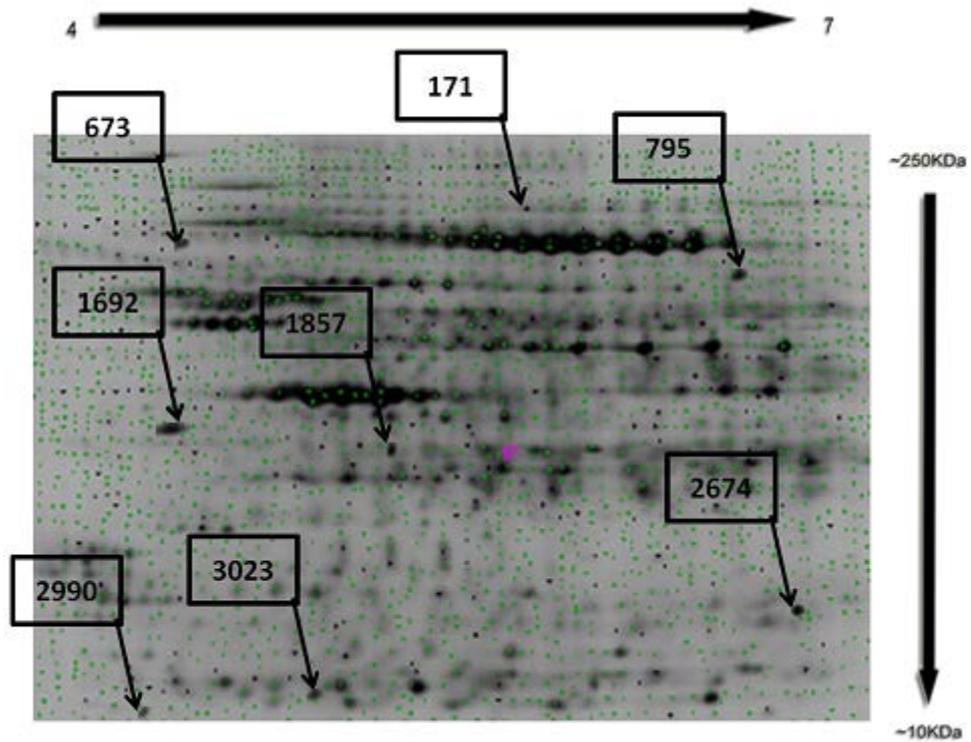


Figure 4-41: Gel map of example candidate markers given in Figure 4-40.

This shows that those markers are possibly playing significantly different roles within each of the tissues types. The positions of the markers are given on the gel spot map in Figure 4-41 shown below.

Looking at the equivalent three-way Venn diagram with medulla tissue in Figure 4-42, there is some success with the identification of the possible markers. Starting with sector (G), spot 2635 was identified as Uracil-DNA Glycosylase, isoforms CRA_a, this proteins function is to eliminate non-required uracil from DNA molecules by cleaving the N-glycosylic bond and initiating the base-excision repair (Haug et al., 1994, Haug et al., 1997), this marker is showing an intermediate effect in the congenic strain between that of WKY and SHRSP. This suggests a polygenic relationship (where the combination of several genes lead to a given phenotype), where the WKY phenotypes is suppressing the effects of the SHRSP genes expression in the congenic strain. The other identification from the markers found using this Venn analysis was spot 2698 Hypothetical Protein LOC619574. This marker is shown to preserve the SHRSP traits in the congenic strain. As can also be seen, the cortex and medulla tissue for this marker hold the opposite profile pattern and therefore are likely to have different functions within the different tissue types.

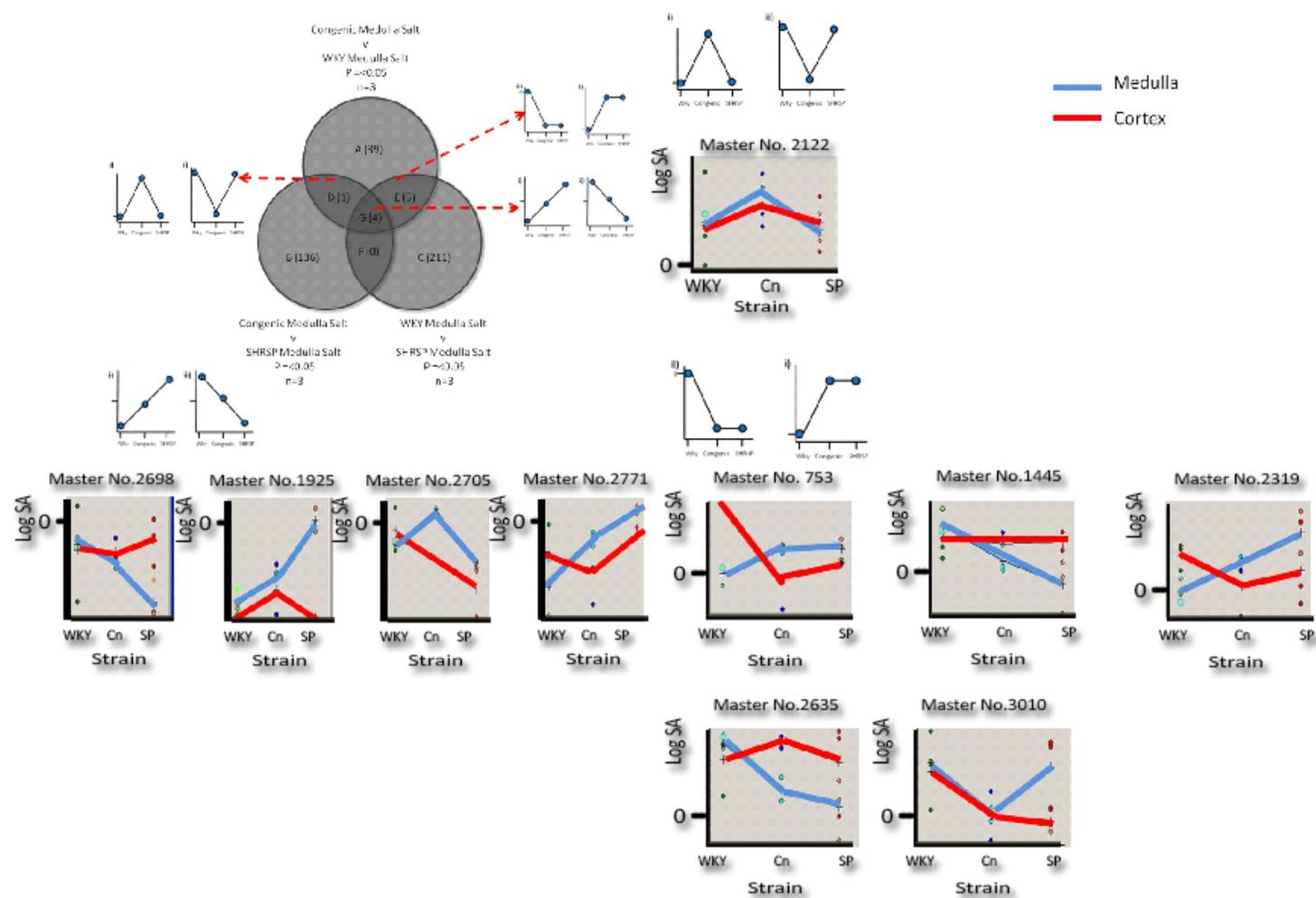


Figure 4-42: Combination of Venn and profile analysis in mining down of data. Example and candidate marker profiles are given for each sector that matches predicted profiles in Figure 4-12. Venn analysis is with medulla tissue, however cortex tissue profiles have been overlaid for comparison.

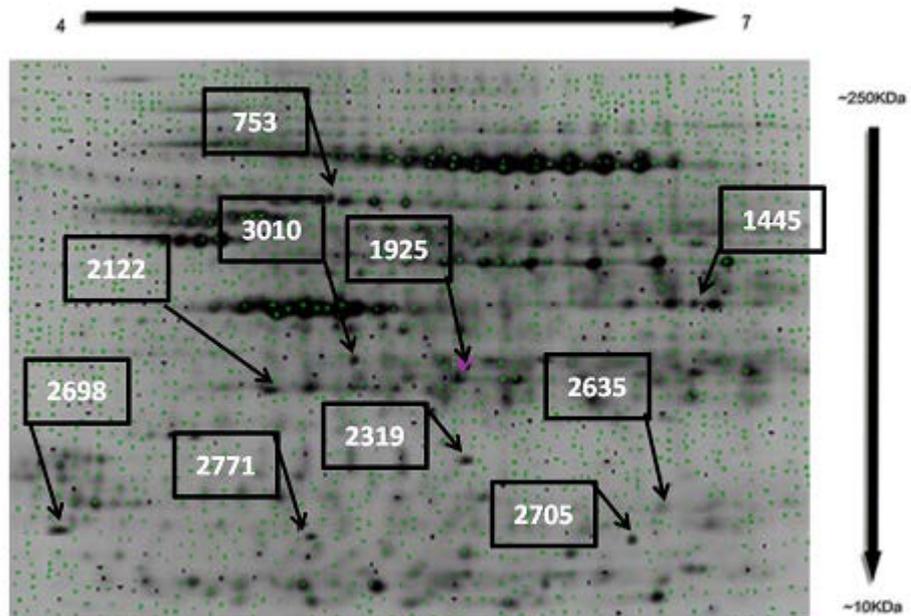


Figure 4-43: Gel map of example candidate markers given in Figure 4-42. The two identified markers are; 2635. Uracil-DNA Glycosylase, isoforms CRA_a and 2698. Hypothetical Protein LOC619574

With regards to the identified spots in the main investigation, there was no cross-over between tissue types found. This is shown in Figure 4-44 below. This goes to further validate the need for smaller sample proteomic discussed in section 5.5 regarding future work. If the two tissues had just been homogenised then this morphological information would be lost. This was also demonstrated by MALDI-MS imagining shown in section 4.6.5.

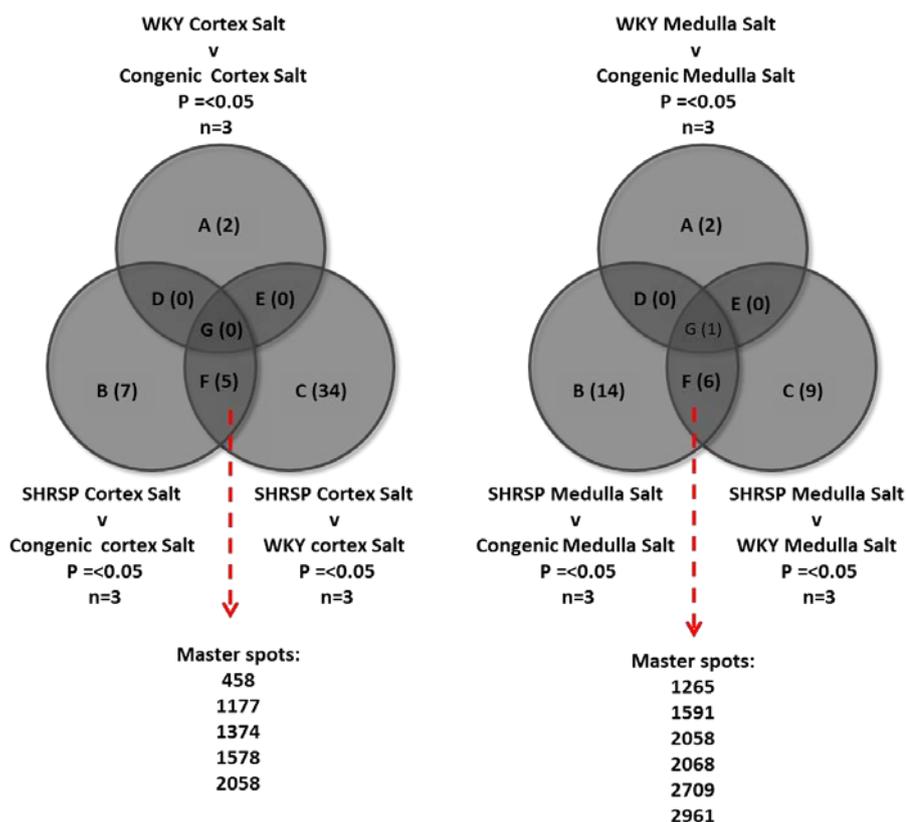


Figure 4-44: Venn analysis for identified spots only. As can be seen there is little crossover of identified candidate markers between tissue types as the master spot numbers are different (with the exception of one). This helps to indicate the need for segregation of tissue types.

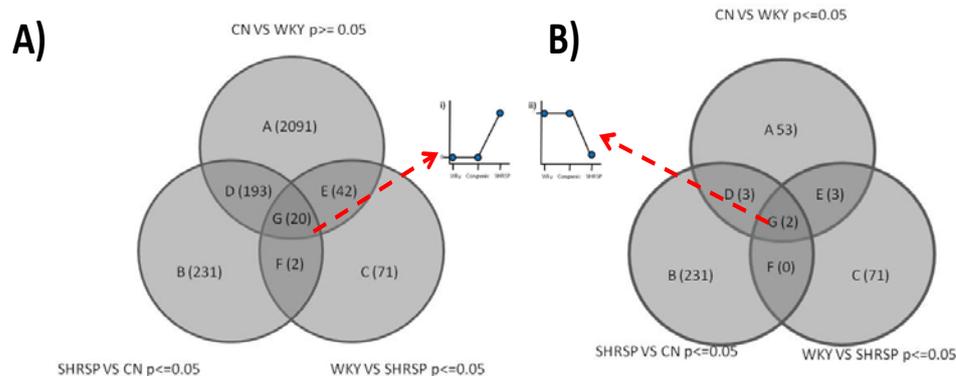


Figure 4-45: Venn diagrams comparing strains using an alternative strategy. An assumption of no significant difference is used between WKY and Congenic strains in order to target Sector (G) as being the intensity profile shown. A) Cortex and B) medulla. A possible strategy for increasing the number of candidate markers discovered

An alternative possibility to increase the number of possible candidate markers discovered is to use the strategy employed in Chapter 3, by using an approach to target the specific profile in sector G. This could be used if the number of markers is very low. As seen in Figure 4-45, the number of markers in sector G is 20, the equivalent to 0 in Figure 4-40 and 2 compared to 0 in Figure 4-41. All of the markers identified during the course of the main investigation are given in Table 7-2, in the appendix.

4.6.4 Validation

The following section includes results which help to confirm and validate that which has been discussed in the sections above.

4.6.4.1 The effectiveness of Macro-dissection

In order to validate if macro-dissection was a valid way to divide tissue into the medulla and cortex, an assessment of total protein content in a given area was performed.

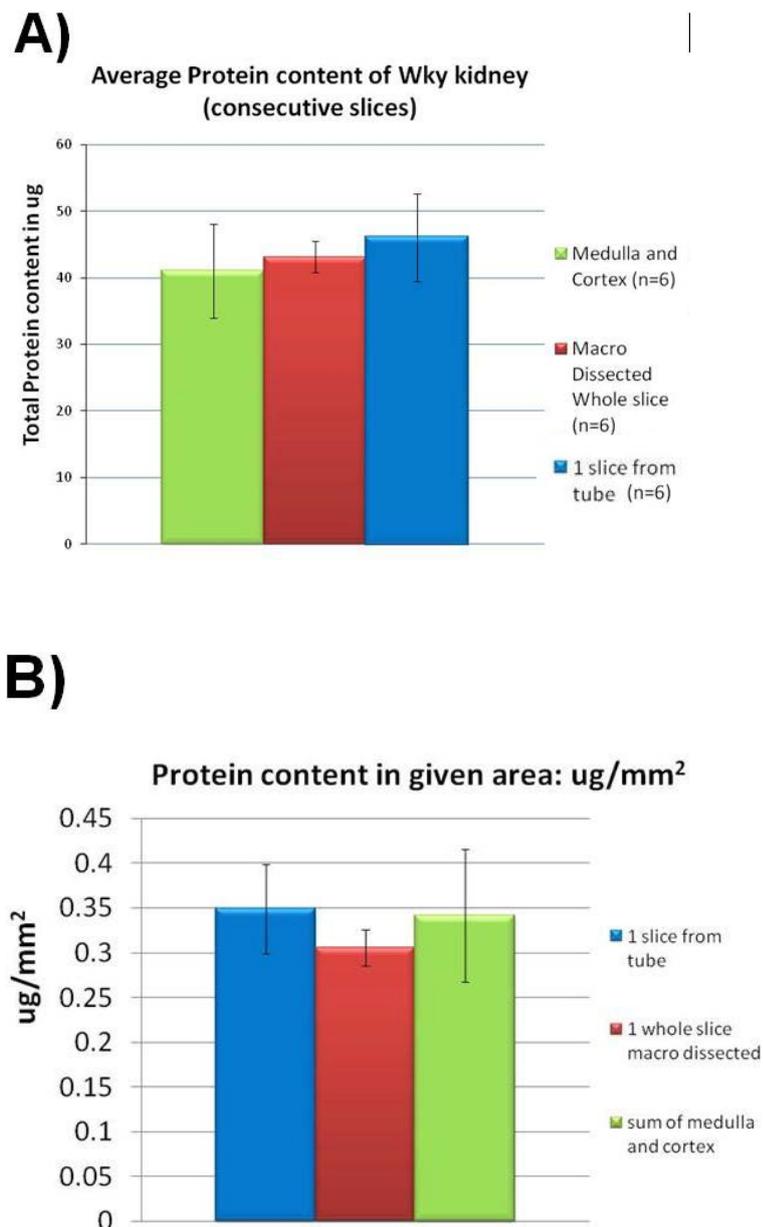


Figure 4-46: Validation of macro-dissection A) shows average protein content per 14um slice of tissue from a WKY rat. n=6 by comparing macro dissection cortex and medulla (green), a whole slice macro dissected (red) and a slice placed directly into lysis buffer from an eppendorf tube B) Shows the protein content per given area from a WKY rat. Error bars in both A) and B) are of two standard errors. n=6 for both A) and B).

As can be seen in Figure 4-46, the protein content was compared in two different ways. Firstly, by comparing protein content (using a Bradford assay) per a slice and secondly, photographing the slice and comparing the protein content per given area. In both cases, the variation within two standard error crosses over for each method of dissection/lysis, showing that macro-dissected medulla and cortex gives a comparable amount of tissue per a given slice or area in comparison to a whole slice, or whole macro-dissected slice. The use of Laser Micro-dissection was considered, as an accurate and precise means of selecting regions of tissue. However LMD presented a number of logistical difficulties. Firstly the protein recovered from a LMD excised disc was extremely variable from disc to disc, as can be seen in Figure 4-47.

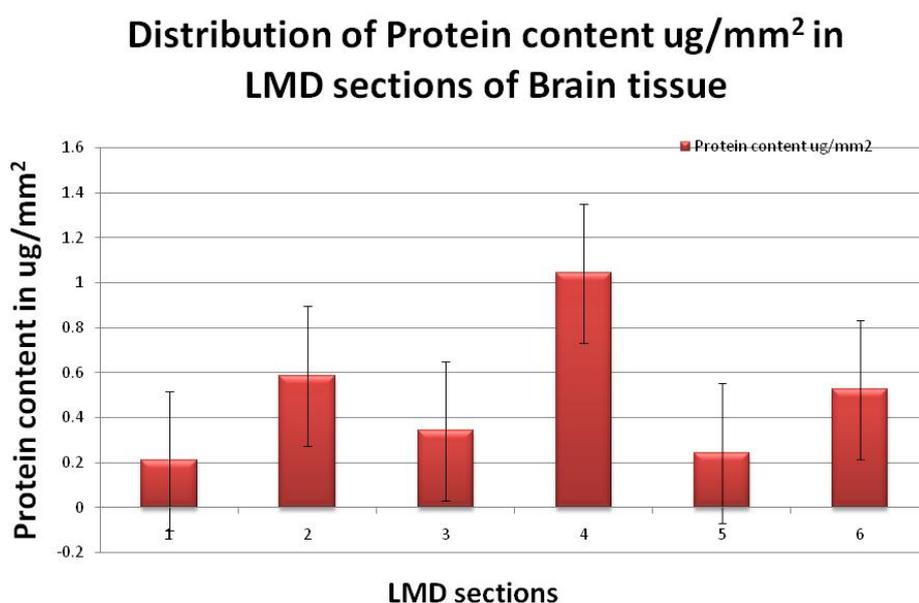


Figure 4-47: Variation of protein concentration from Kidney tissue from LMD tissue discs. As can be seen the variation is considerable but still overlapping. The error bars displayed are two standard errors. There were 6 slices in total and they are compared against each other individually and not averaged.

Also using LMD meant there were two options for quantifying the protein content. One method employed by Meyer *et al* 2005 was to pool samples from multiple slices to get enough material for Bradford assaying and then to use saturation labelled DiGE for analysis. This method does not really get around any disadvantage of averaging samples. The second option was to use a bioanalyser chip (often used in DNA work) to quantify the concentration of protein. This was considered and the results are shown in Figure 5-3.

Additionally, even with the use of an anti-static gun, LMD was an extremely time consuming activity, where sample was frequently lost as it did not always enter the eppendorfs. It was considered due to the limited and expensive nature of the kidney tissue not to perform this experiment using LMD.

4.6.4.2 Principal components analysis of groups in the pilot investigation.

PCA is an effective way of validating experimental groups. The validation is obtained by statistical analysis being performed without initially assigning groups until the data is displayed. The further apart experimental groups are, the larger the difference between those groups, with PC1 being the 1st principal component and therefore contains the largest differences followed by PC2. In this section the PCA for the different tissue types in the pilot investigation are considered. In Figure 4-48 all spots are considered so may contain anomalous spots.

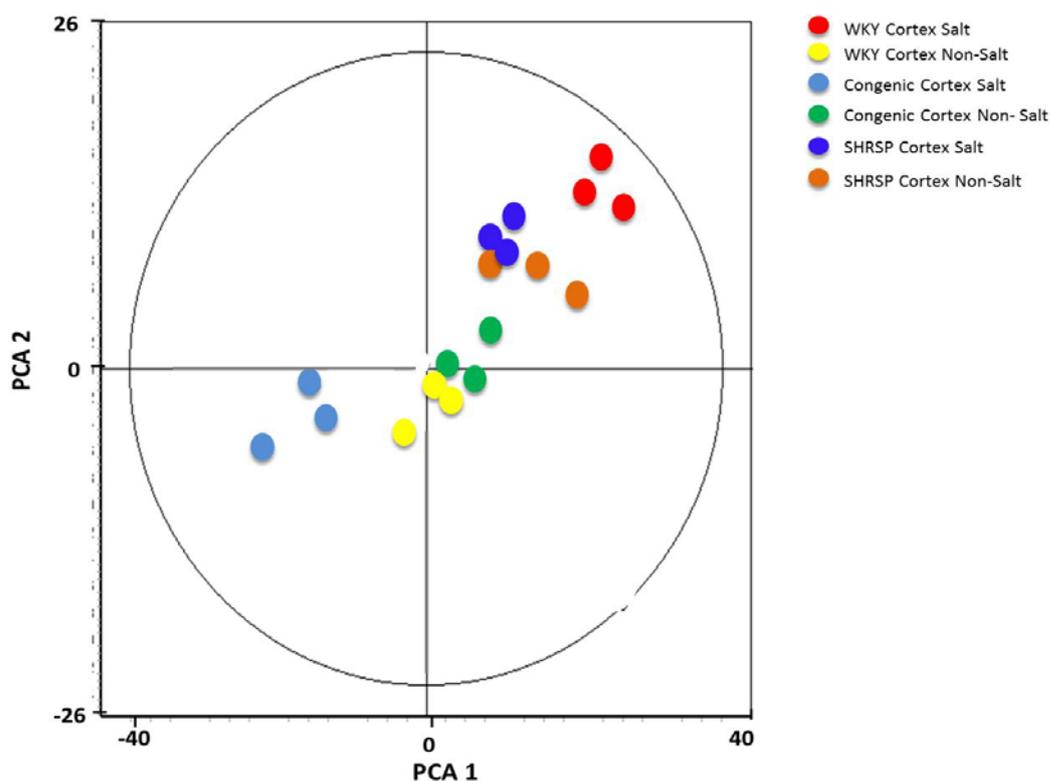


Figure 4-48: PCA score plot containing all 5777 protein spot in the pilot investigation for Cortex Tissue

As can be seen SHRSP salt and non-salt are clustered together, as are WKY and congenic non-salt. This suggests a close relationship to each other.

The WKY Salt and congenic salt are furthest in both PC1 and PC2 suggesting salt treatment has caused the experimental group to diverge from each other. This is consistent with the idea that salt treatment causes a more pronounced

phenotypical response (hypertensive) as has been confirmed in a number of studies already mentioned (Koga et al., 2008, Fountoulakis, 2001, Meneton et al., 2005, Blaustein et al., 2006). Therefore it is possible that the salt treatment has caused a divergent response from non-salt treated WKY/Congenic in comparison to the salt treated WKY/Congenic. Equally interesting is that salt treatment has cause WKY strain to be more closely related to SHRSP strain.

This pattern continues in a similar vein in Figure 4-49, where the spots are sorted into significant 1-way ANOVA scores ($p\text{-value} \leq 0.05$). By doing this, a significant number of spots have been eliminated. SHRSP still have a reasonably close relationship (in the PC1 dimension, which has the greatest effect). As do Congenic and WKY strains for non-salt treatment, while once again congenic and WKY salt treatment exhibits the greatest difference between groups.

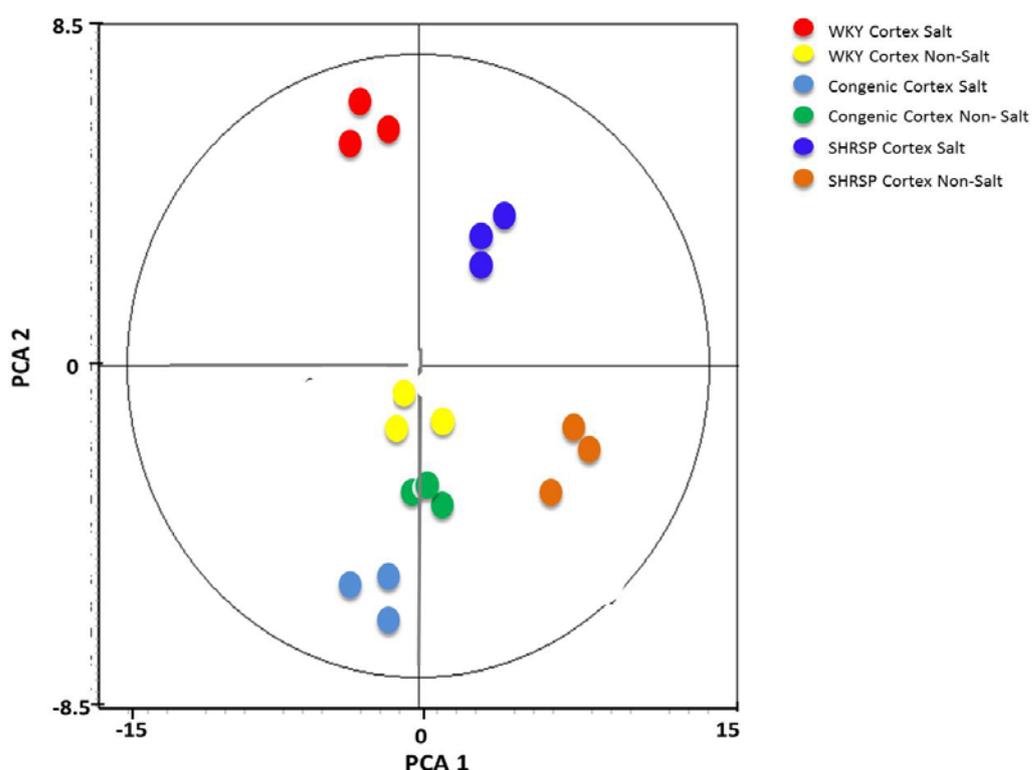


Figure 4-49: PCA score plot of all one way anova spots 802 protein spot in the pilot investigation for Cortex Tissue

With regards to medulla tissue type in the pilot investigation, the PCA analysis can be seen in a Figure 4-50 and Figure 4-51. Once again Figure 4-50 contains all spots detected on the gel map. In contrast to cortex tissue, medulla tissue

shows different relationships. All strains (both salt treated and non-salt treated), with the exception of Congenic salt treated, reveal a very close relationship with each other. This could suggest that cortex tissue is the tissue type that is predominantly affected when it comes to hypertension, as cortex tissue seems to exhibit greater differences according to PCA analysis. This is of course a very speculative assertion, however, the PCA analysis is showing very little differences between strains in the medulla tissue type. Relating this to the structure of the kidney, most salt and water control occurs in the cortex, as this is where the majority of the function unit, the nephron, is situated. It is still very intriguing that the congenic salt treated strain has been distinguished significantly from the other strains. It appears that salt treatment has a significant effect on the congenic strain in the medulla region of the kidney.

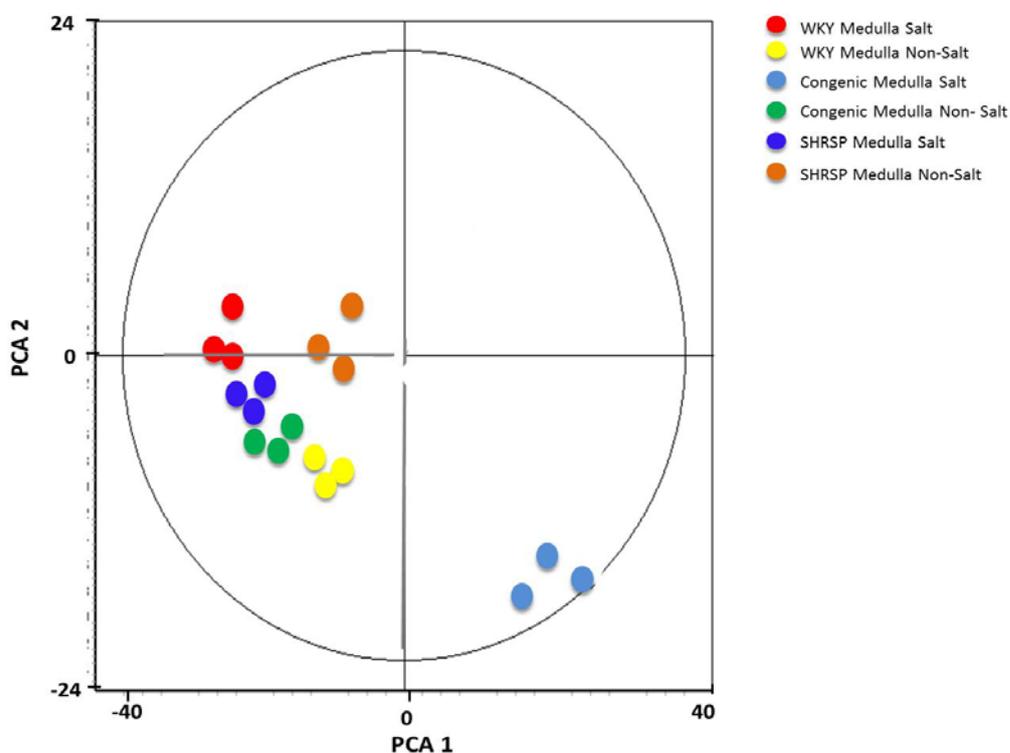


Figure 4-50: PCA score plot of all 5777 protein spot in the pilot investigation for Medulla Tissue

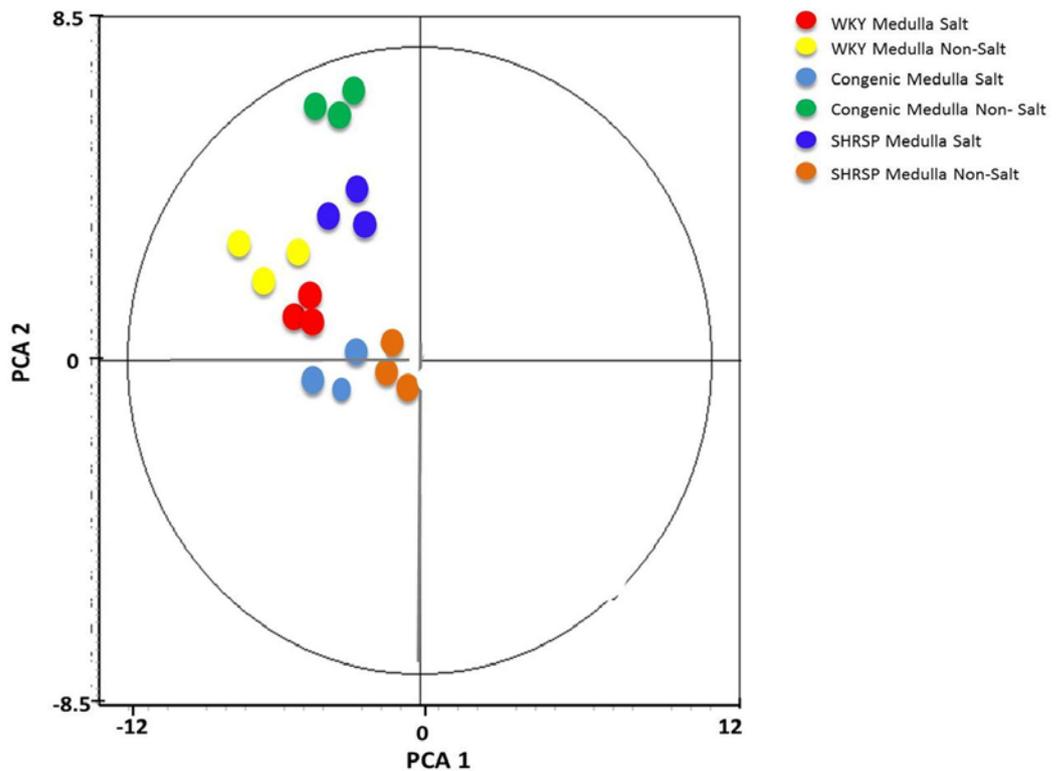


Figure 4-51: PCA score plot all one way anova spots 802 protein spot in the pilot investigation for Medulla Tissue.

The pattern does however change slightly with regards to spots sorted by 1-way ANOVA, as can be seen in Figure 4-51. The difference between the main group and the congenic salt treated is not as significantly pronounced. Additionally the SHRSP non-salt treated group is now discernable from the main cluster. This is an interesting shift and not readily explainable.

4.6.4.3 Principal components analysis of groups in the main investigation.

PCA analysis was further conducted for the main investigation. In the main experiment the variation will be, more greatly, down to biological differences than technical variance. This time both medulla and cortex have been included into the same analysis as only salt treated samples were used.

A similar strategy using PCA has been employed by performing it both on the whole data set, Figure 4-52, and then by reducing the data to contain only spots with a 1-way ANOVA score of $p\text{-value} \leq 0.05$ as can be seen in Figure 4-53.

Starting with the whole data set shown in Figure 4-52, it can be seen that in both tissue types the WKY and Congenic seem to have a closer relationship in comparison to the SHRSP strains. Equally, the relationship between tissues appears to be stronger than the relationship between strains, as the cortex spots

are more closely related to each other than that of the same strains. This is intrinsically reasonable as tissue type by its nature is visually and structurally different, while WKY and Congenic strains differ by only a few sections of their genome. What is interesting, however is that the SHRSP strains are significantly different as they are further from the WKY and Congenic strains and equally intriguing is that they are responsibly far apart from each other too.

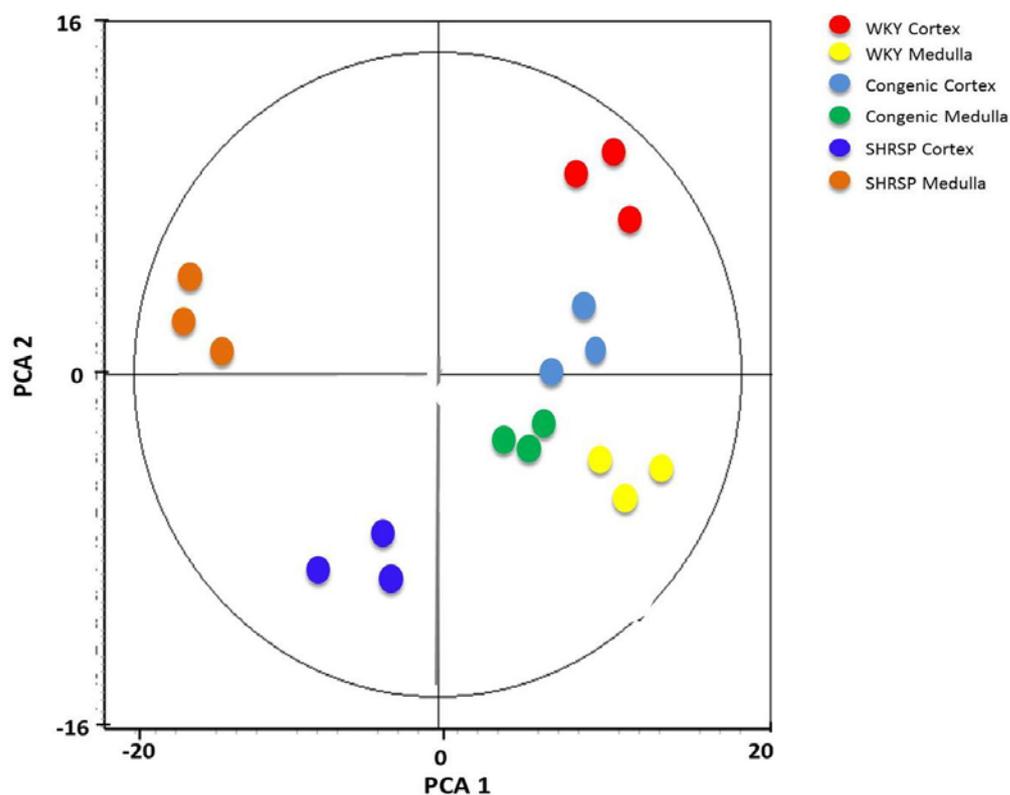


Figure 4-52: PCA score plot of the main investigation. All 3174 spots included

The pattern is closely matched in Figure 4-53.

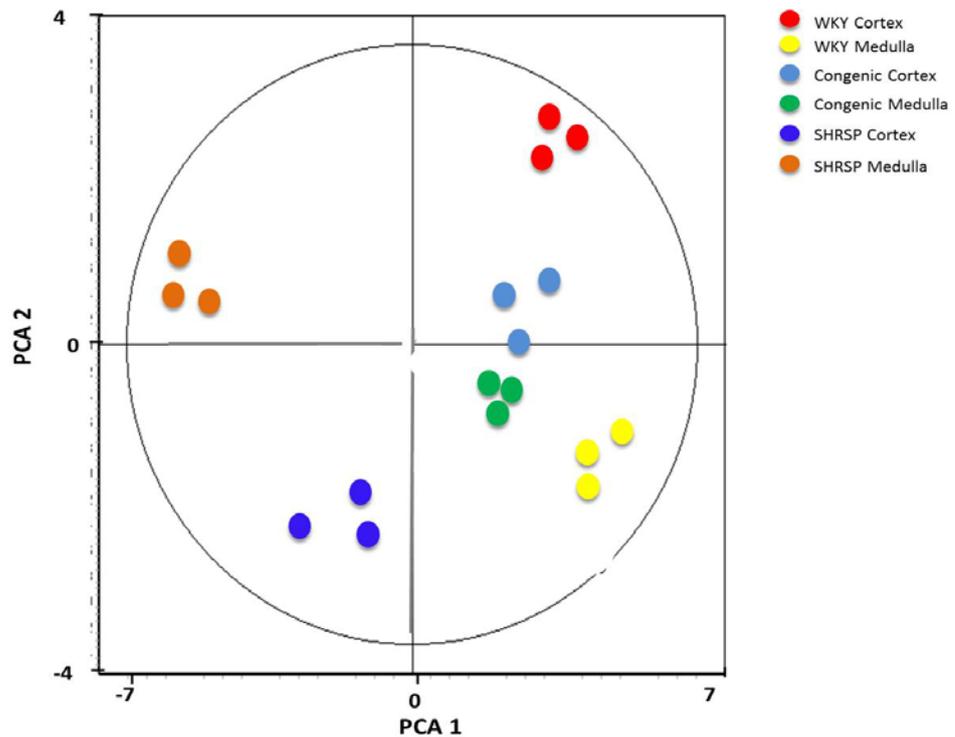


Figure 4-53: PCA score plot of main investigation. Only spots with one-way anova $p < 0.05$ considered. No of spots = 189.

4.6.4.4 Example of genomic and proteomic integration

The intention was that the markers found during this investigation could be linked to genomic data and eventually mapped back to the correct portion of the rat genome in order to further our understanding of hypertensive pathways. The cardiovascular research unit has performed a number of analyses using microarray technology allowing for the measurement of thousands of mRNA in the kidney tissue, the strains studied in this investigation. The microarray technology allows for parental strains to be compared to congenic strains for gene mapping. Difference in gene expression can then be tallied with known gene function and further correlated to identified proteins from proteomic work flow to be compared with direction of expression and further pathway analysis using software such as Ingenuity Systems Pathway Analysis software. Work from Jeff et al (2000) and further extended by McBride et al (2003) (Jeffs et al., 2000, McBride et al., 2003) has used subcongenic breeding and microarray expression techniques. The congenic strains investigation (including SP.WKYGla.2a) showed a significant reduction in baseline blood pressure in comparison to the SHRSP parental strain. Comparison of the 3 strain using the Pathway analysis software was performed by Dr Martin McBride of the Cardiovascular Research Unit. A full pathway analysis of the correlation between the proteomic data from this investigation and the genomic data is

beyond the scope of this investigation, but is there for use in future work. An example of how this data is integrated is shown in Figure 4-54. The proteomics data from the identified possible markers were tabulated including only those with t-test p-values of ≤ 0.05 . The t-test performed was salt treated from the main experiment congenic vs. SHRSP strains. Only identified proteins were considered to compare against the mRNA expression data. The genomic data was analysed on the same basis (Congenic vs. SHRSP strains with p-value of ≤ 0.05). It should be noted that the microarray expression analysis was not done separately for tissue types.

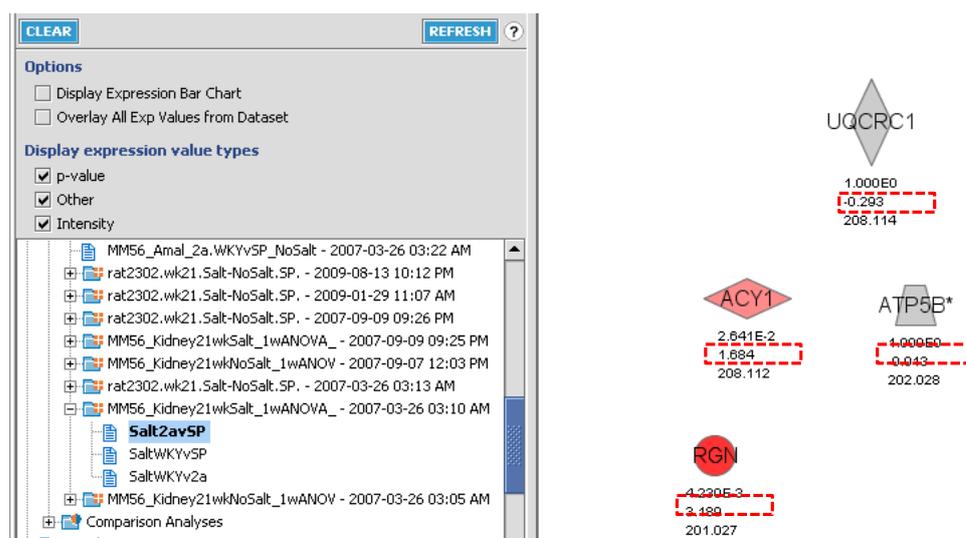


Figure 4-54: figure showing matches between micro array data and proteomic data. Red indicates upward expression in SHRSP compared to congenic and grey indicates downward expression. The highlight number is the fold change in expression.

The genomic data shows the increase in expression of the mRNA for ACY1 (Aminoacylase 1) and RGN (regucalcin) of 1.684 and 3.189 respectively and a downward expression of the mRNA for UQCRC1 (ubiquinol-Cytochrome c reductase core protein I) and ATP5B (ATP synthase, H⁺ transporting, mitochondrial F1 complex, beta polypeptide, isoform CRA_a). The genes and corresponding proteins were searched using ingenuity systems pathway analysis software and the rat genome database. This software and database give information on gene ontology, sequence data and associated diseases. ACY1 is associated with chromosome 8 in rats and is associated with a number of liver conditions including cancer of the bile ducts and hepatocellular carcinoma. It forms a zinc binding enzyme, which is homodimeric and functions in the Cytosol and catalyses the hydrolysis of acylated L-amino acids to L-amino acids and is postulated to salvage acylated amino acids. Its expression has been shown to be non-existent in certain kinds of lung cancer

but has yet not been affiliated with hypertension or kidney tissue. Regucalcin holds an apoptotic role in cells but as yet has been assigned to the X chromosome location. It is a conserved calcium binding protein that is expressed in the liver and kidney and is thought to hold an important role in calcium regulation and homeostasis. Although no association to hypertension has been made it has been shown to be down regulated in aging rats. ATP5B has been located on chromosome 7 in the rat and catalyses the synthesis of ATP in the mitochondria. It does, however, have homologous and conserved regions present in many ATPases. UQCRC1 is also associated with chromosome 8 and its function in humans is to catalyse the transfer of electrons from a coenzyme called QH₂ in ferriCytochrome C in the electron transport chain of mitochondria.

If this is compared to the proteomic data shown in Figure 4-55 to Figure 4-57 and Table 4-11 and Table 4-12, all the candidate markers show reduced proteomic expression; this is consistent to mRNA expression in both UQCRC1 and ATP5B but opposite to RGN and ACY1. This opposition requires further analysis to determine the cause. The only marker, from both the genomic and proteomic data sets, with a common proteomic profile across both tissue types, is Regucalcin as shown in the validative Venn diagram in Figure 4-57.

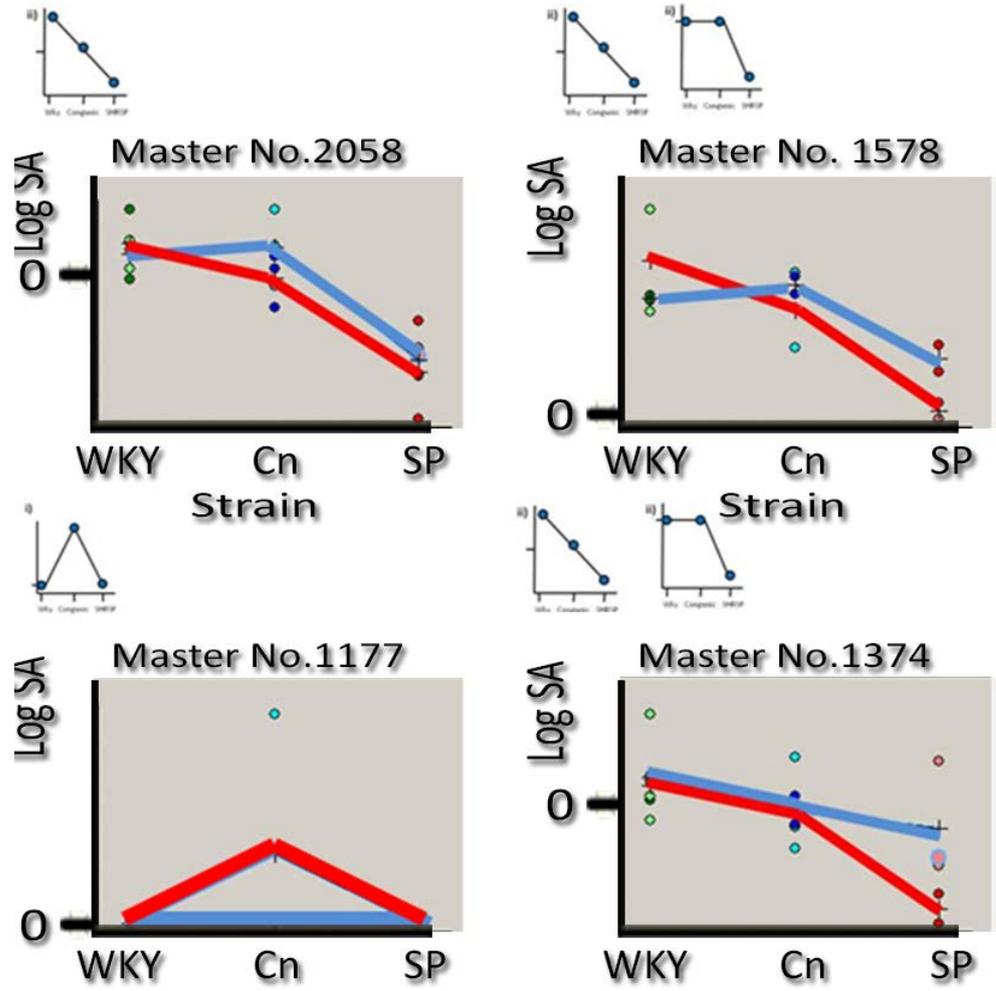


Figure 4-55: Profiles of spot intensities in the main investigation, which match with RNA micro array data from cardiovascular research unit.

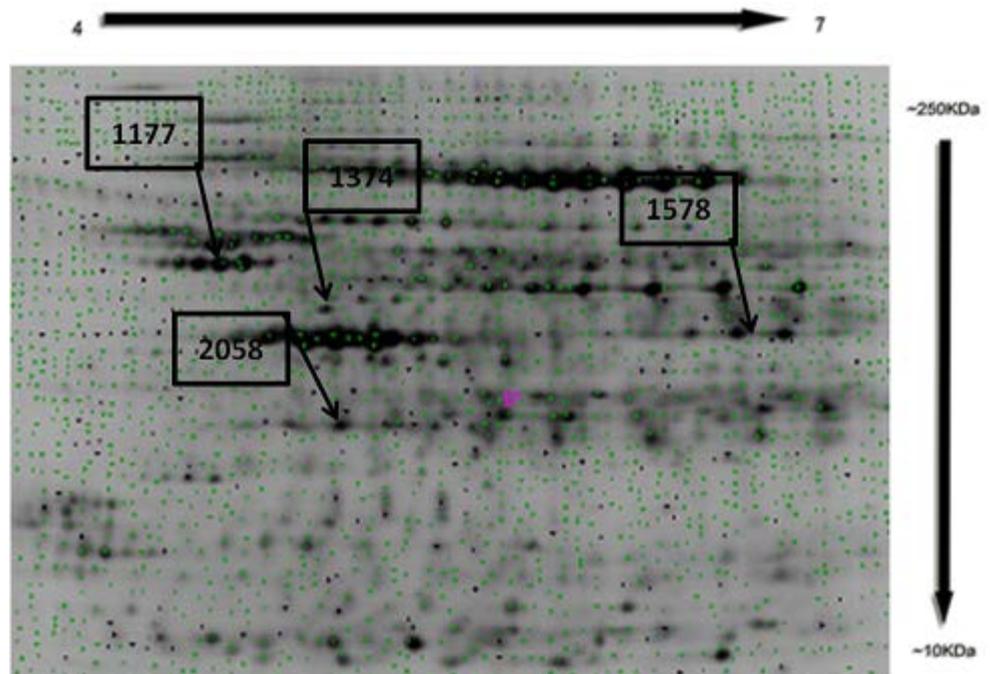


Figure 4-56: Gel map giving position for possible candidate markers from Figure 4-55.

master spot no	name	MOWSWE Score	gi number	t-test p-value for Cnvsp	fold change (from average ratio) of CnvSP
1177	Chain B, Rat Liver F1- Atpase	122	gi 6729935	0.0018	-2.51
1177	mitochondrial ATP synthase beta subunit [Rattus norvegicus]	122	gi 54792127	0.0018	-2.51
1177	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide, isoform CRA_a [Rattus norvegicus]	122	gi 54792127	0.0018	-2.51
1374	ubiquinol-Cytochrome c reductase core protein I [Rattus norvegicus]	73	gi 51948476	0.04	-1.43
2058	regucalcin [Rattus norvegicus]	77	gi 408807	0.043	-1.5
1578	rCG25777, isoform CRA_a [Rattus norvegicus]	121	gi 149018671	0.044	-1.44
1578	aminoacylase 1 [Rattus norvegicus]	105	gi 52851387	0.044	-1.44

Table 4-11: Identified markers from Figure 4-55 and Figure 4-56 present in cortex tissue. Table gives information, the MOWSE score from the mascot database search, p-value and fold change for congenic vs SHRSP (CnvSP).

master spot no	name	MOWSWE Score	gi number	t-test p-value for Cnvsp	fold change (from average ratio) of CnvSP
2058	regucalcin [Rattus norvegicus]	77	gi 408807	0.009	-1.66

Table 4-12: Identified markers from Figure 4-55 and Figure 4-56 present in medulla tissue. Table gives information, the MOWSE score from the mascot database search, p-value and fold change for congenic vs SHRSP (CnvSP).

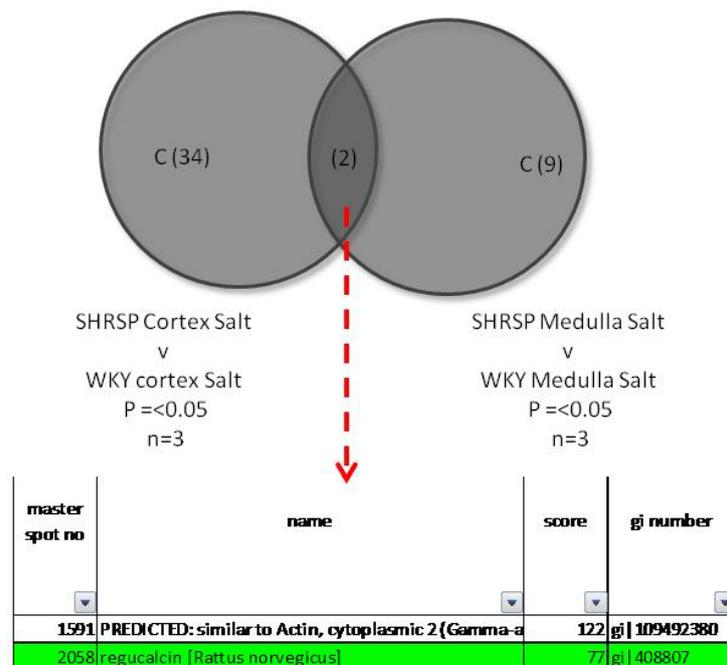


Figure 4-57: Comparison of tissue types and matches with microarray expression data.

4.6.5 Mass spectrometry Imaging

MADLI Mass spectrometry imaging was also performed to help gain more spatial and morphological information and to try and compensate for some of the issues with DiGE. This work is currently still underway and was performed by Dr Richard Goodwin, Research Associate at Glasgow University. MSI was performed on both salt treated and non-salt treated kidneys from the same WKY, Congenic and SHRSP strains as above (not the same animals). The results are shown in Figure 4-58. It was shown that there was a considerable number of markers across the three strains with no difference, however it was also shown that there was evidence of varying morphological expression differences for a number of different mass filters. An example is shown in (C) for protein identified as Histone H1 at mass filter 2039 m/z. For this particular marker the cortex region is the main region affected. Using imaging affords a great advantage of being able to tract these morphological changes through the organ and also over time. This change in cortex region reflects result described in section 4.6.4.2.

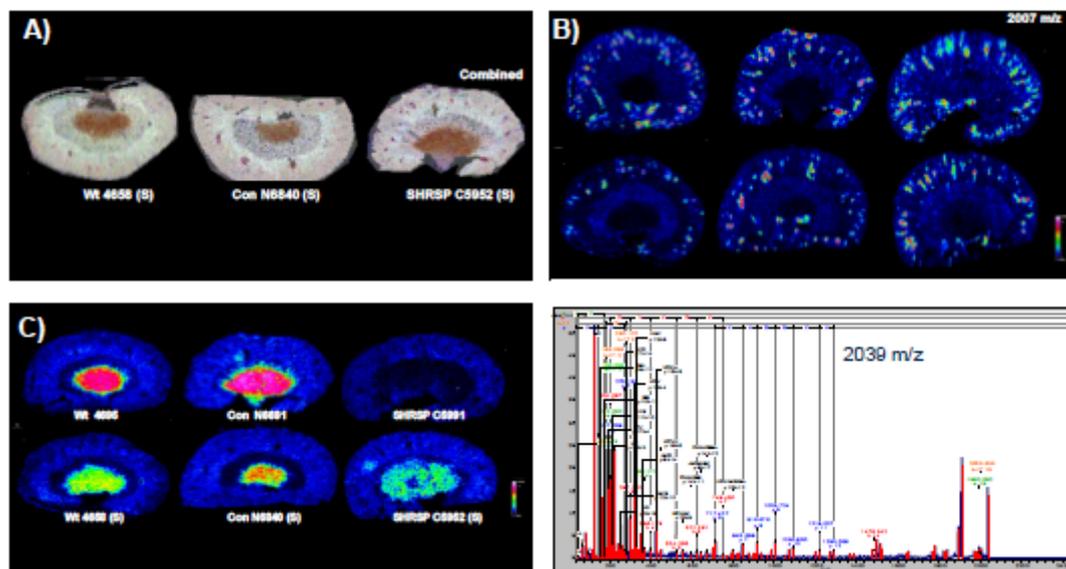


Figure 4-58: MALDI-mass spectrometry imaging (MSI) in non-salt and salt treated WKY (WT), Congenic (Con) and SHRSP strains. A) This is composite images showing all the mass filters giving an overview of the distribution of markers across the three strains. B) Mass filter of (2007 m/z) with heat-map display showing regions of highest intensity. Mass distribution is even across all kidneys. C) Mass filter of (2039 m/z) Mass seen to vary significantly between the SHRSP kidney and that of wild type and congenic Salt treatment affects the intensities across the kidneys. D) ms/ms identification of Histone H1 direct from tissue for 2039 m/z. Figure provided by Dr Richard Goodwin, Research Associate, Glasgow University.

4.6.6 Label free quantitation using LC-MS of rat plasma

In addition to the work performed in this investigation, a complementary methodology of label free quantitation was performed by Heather Allingham investigating plasma samples taken from the three strains of rats used in my investigation. Once again, this investigation helps to try and compensate for any disadvantage in other proteomic technique and if fortunate, could lead to cross validation of techniques. Although it did not lead to any matched identifications as hoped, it does demonstrate that the analysis technique can be applied for biomarker discovery across different technologies.

A strategy of searching for profiles using Venn analysis was employed. Two examples, also validated using western blot exhibit, similarly predicted profiles from both the label free quantitation and the western blot analysis. The example profiles are shown in Figure 4-60, Figure 4-61 and Figure 4-61.

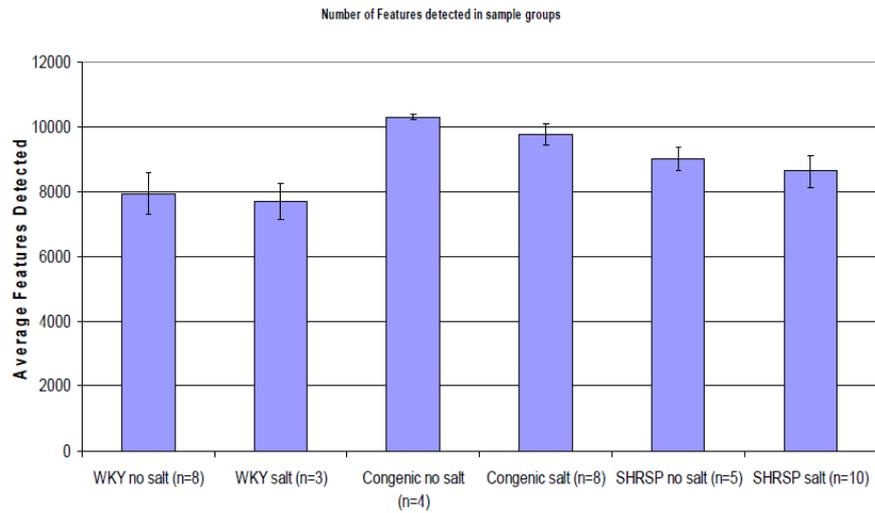


Figure 4-59: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Show the average number of identified features from intact sample results across all groups

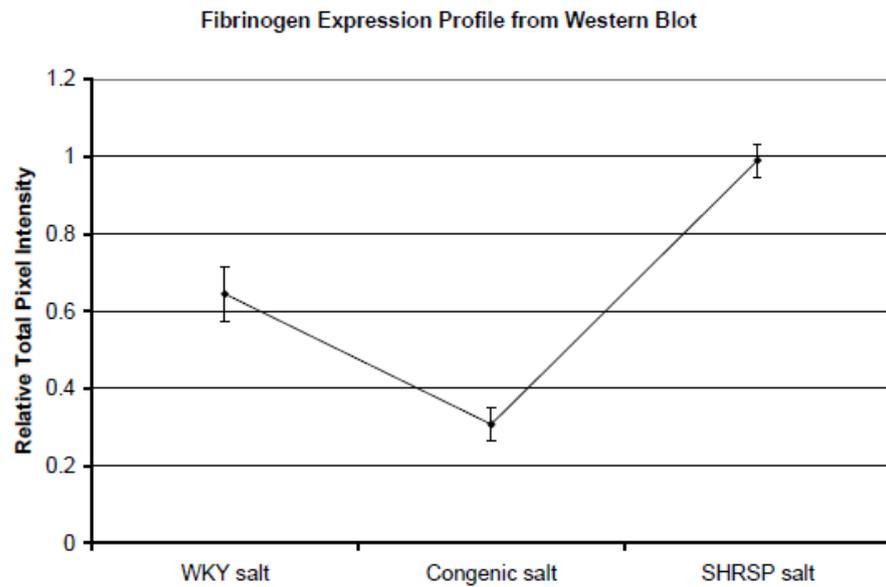


Figure 4-60: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Expression profile from quantitative western blot for Fibrinogen matching profiles from the label free quantitation LC-MS.

Hemopexin Expression Profile from Western Blot

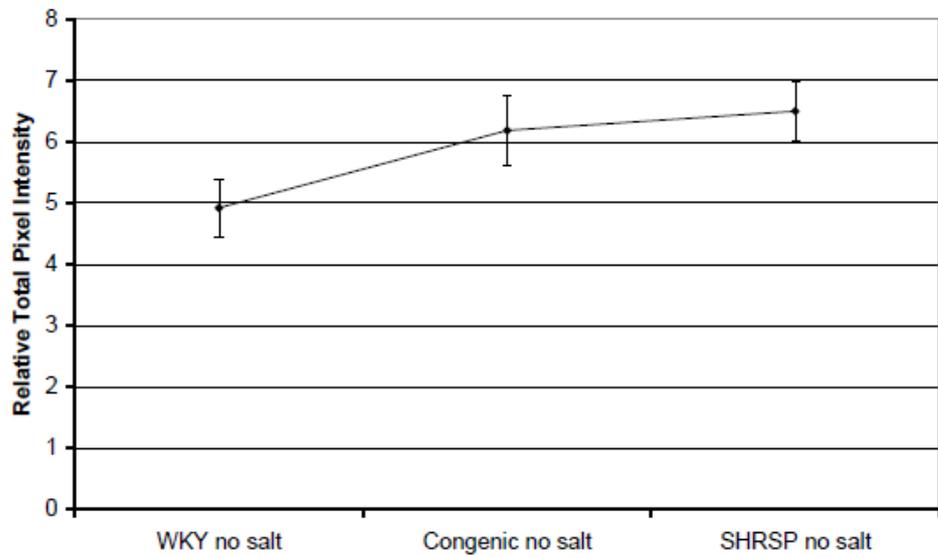


Figure 4-61: This figure was taken from the Thesis of Dr Heather Allingham, Glasgow University. Expression profile from quantitative western blot for Hemopexin matching profiles from the label free quantitation LC-MS.

4.6.7 Biological significance of potential biomarkers discovered

A number of putative markers were discovered during this study. It is accepted that they are putative in nature and therefore they need further validation, but could be a focus for future studies. In this section, an account of their known biological role in both humans and animal studies is reviewed and whether there are any known associations to hypertension or cardiovascular disease. The initial source for information in the search for biological function is the curated database UniProtKB/Swissatp-Prot (<http://www.uniprot.org/>). This examination was then followed up with a search of the literature for publications associated with the protein identification.

Below is a summary of some of the most robust identifications in the main and pilot investigation and their biological roles.

- Peroxiredoxin 3 or thioredoxin-dependent peroxide reductase, mitochondrial is a unique class of peroxiredoxins, as it is the only one isolated in the mitochondria. The peroxiredoxins are a class of antioxidant enzymes that control the cytokine-induced peroxide levels and therefore are involved in signal transduction pathways. Peroxiredoxin 3 is a protein which responds to situations of oxidative stress by dealing with excesses in hydrogen peroxide and protects against oxidative stress (Zhang et al., 2007a, Cox et al., 2009). From Figure 4-30 it can be seen that salt-treatment has no effect on peroxiredoxin 3 levels in both the Wky and congenic strains of rat, but it is considerably higher in the SHRSP strains. This tentatively suggests that SHRSP rat strain produce lower levels of peroxiredoxin 3 and this would leave them more open to damage via oxidative stress. It has been shown, in mice, that increases in the levels of peroxiredoxin (by causing overexpression) can prevent left ventricular remodelling after myocardial infarction due to their anti-oxidative function (Matsushima et al., 2006). The thioredoxin research in the literature has a strong link in both hypertension and cardiovascular processes (Ebrahimian and Touyz, 2008). It has also been shown that thioredoxin expression is increased in cardiac fibroblasts (Lijnen et al., 2012). It also has been shown to protect neuronal cells (Cox et al., 2009). Additionally, overexpression has been shown to protect cancer cells by prevention of hydrogen peroxide induced apoptosis (Nonn et al., 2003, Li et al., 2012b, Newick et al., 2012). It is thought to have a protective

characteristic in neurological disorders (Zhu et al., 2012). With its association to a number of different disease processes it is unlikely that peroxiredoxin 3 is going to be a specific enough biomarker for hypertensive processes.

- Regucalcin or otherwise known as gluconolactonase or senescence marker protein 30 is a protein found in many different cell types and plays a key role in intracellular Ca^{2+} homeostasis by binding to and activating Ca^{2+} pumps on various membranes, including the plasma membrane. It also has an inhibitory effect on protein kinase and phosphatase activity. It has been shown to play a role in transcriptional regulation, when it migrates to the nucleus, thus it can have a large influence on phenotype by controlling transcription (Yamaguchi, 2005). In this study, in cortex, the regucalcin is mostly expressed in the Wky strain and has the lowest expression in the SHRSP strain with an intermediate expression in the congenic. The presence of regucalcin aid the resistance of renal tubules against injury in the presence of high Ca^{2+} (Inoue et al., 1999) but it has been described as being at its highest abundance in liver tissue and is suspected to play a role in liver regeneration (Yamaguchi, 2000). A combination of hypertension and a high salt intake is known to aggravate the loss of Ca^{2+} from the renal tubules causing damage (Cappuccio et al., 2000, Timio et al., 2003). In healthy individuals this is mediated by regucalcin. Within this putative assumption, the reduction in regucalcin in congenic and SHRSP rat strains could be partly playing a role in the hypertensive symptoms, due to loss of Ca^{2+} and renal tubule damage. The preliminary nature of this proposition is acknowledged.
- Aminoacylase 1 (ACY1) is an enzyme which is known to be expressed at high levels in the kidney (Uttamsingh et al., 2000). It catalyses deacylation of N-acyl-L-amino acids to give fatty acids and amino acids as products. It has been thought to be linked to degradation, antioxidant defence and redox sensitive reactions in the kidney. An investigation, using a spontaneous hypertensive rat model, showed that a number of metabolic enzymes among the carbonylated proteins had increased levels of carbonylations compared to the Wky rat strain. Among these enzymes was ACY1. These changes in expression has been shown to

be associated with oxidative stress (Tyther et al., 2009). They found that there were more changes between SHR strains compared to Wky strains of ACY1 (and others) in medulla compared to cortex. In this thesis, in cortex tissue, the ACY1 protein was expressed most highly in the Wky and congenic strain and was lowest in SHRSP strain. It was identified in a number of spots; these are shown in Figure 4-62 with varying profile characteristics.

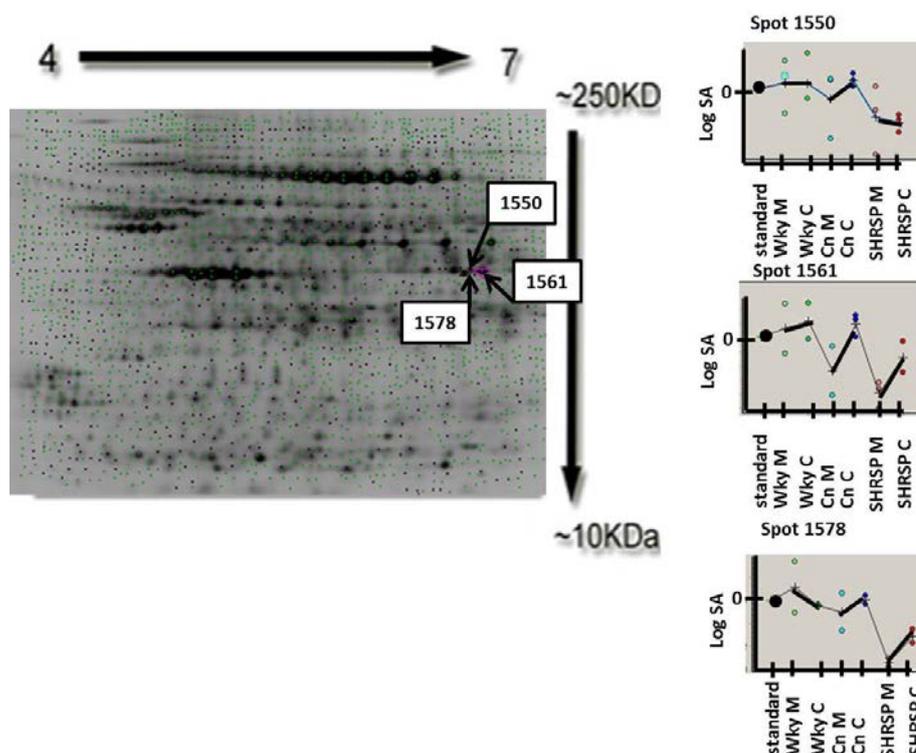


Figure 4-62: Positions and profiles of multiple identifications of protein Aminoacylase 1, from the main investigation. . The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.

- Enolase 1 (Eno1) protein or α -enolase is a 434 amino acid long and 47,128 Da glycolytic enzyme, which is abundant in the majority of tissues and cells. In addition, to its role in glycolysis it has been shown to have an effect in growth control, hypoxia (hypoxia induces the transcription of eno1) and severe allergies including severe asthma (Semenza, 2001, Rey and Semenza, 2010, Gracey et al., 2001). It was also one of the differentially expressed proteins described by Tyther et al, 2009 associated with oxidative stress and hypertension in SPR, showing an increase in carbonylation. Eno1 is present in most cells in the kidney and is particularly prevalent in the cortex. It has been found

to be differentially expressed in nephrotoxicity, including antibiotic toxicity. Thus, nephrotoxicity causes *eno1* to be down regulated affecting glycolysis in the kidney (Charlwood et al., 2002b). The renal medulla is the area where interstitial concentration gradients are maintained and it plays an important part in water balance in the body and therefore hypertension. It is also known to be under greater oxidative and hypoxic stress. Hypoxia in the kidney is exacerbated in hypertension sufferers and this is equally true of SHR, which have been shown to have increased carbonylation of *eno1* protein in the medulla and to a lesser extent cortex (Johns et al., 2010). It has also been shown to be integral in the pathways involved in alanine-glyoxylate aminotransferase gene (*AGXT*) knockout mice, which is a model for hyperoxaluria type I. Along with other proteins, it was substantially down regulated in these animals (Hernández-Fernaud and Salido, 2010). It has also been associated with diabetic rat kidney (Chougale et al., 2012), neuroblastoma and Wilms' tumour (Odelstad et al., 1982) and other kidney tumours (Kuroda et al., 2000) therefore, it is unlikely to act as a specific biomarker for hypertension due to its associations with multiple diseases at different sites around the body. In this study it is down regulated in SHRSP but in Wky and congenic strains, the level is maintained. This can be seen in the profiles in Figure 4-63.

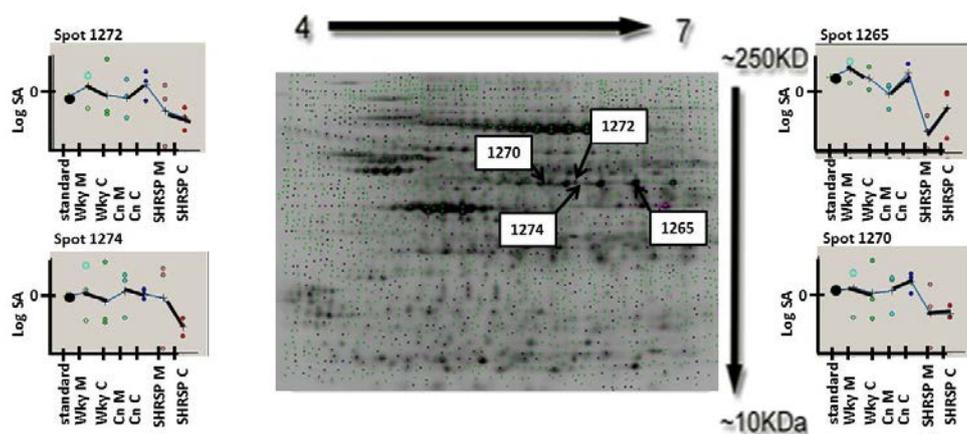


Figure 4-63: Positions and profiles of multiple identifications of protein *Eno1*, from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.

- Aldo-Keto reductase family 1, member A1: The aldo-keto reductases are a set of enzyme superfamilies' which contain 14 different families. They perform oxidoreduction on a wide variety of substances in the body. The Aldo-Keto reductase family 1 is the largest of the family groups and contains aldose reductases, the aldehyde reductases, the hydroxysteroid dehydrogenases and steroid 5b-reductases (Hyndman et al., 2003b). The A1 member identified in this study, is part of the group commonly referred to as alcohol dehydrogenase or aldehyde reductase (EC 1.1.1.2) (Hyndman et al., 2003a). This enzyme is expressed in the highest quantities in the liver but is expressed in numerous tissues to eliminate toxins. A member of this family ALDH2 has been found to protect cardiac tissue from oxidative stress and it has been linked to a possible treatment and protection against cardiac disease (Chen et al., 2010a). Although the increase in ALDH2 and other forms of ALDH may help to protect tissue against oxidative stress, such as the ingestion of alcohol, the consumption of alcohol and other toxin has been associated with increased hypertension, (paradoxically a little alcohol reduces the risk of many CVDs) (Eapen et al., 2011). The association is not completely understood and is likely to be complex with ALDH group playing a role. However, as yet, no direct link between hypertension and any of the AKR family has been stated in the literature. Although, they are synthesisers, along with prostaglandin F₂ α (PGF₂ α) synthases, for molecules such as NADPH-dependent reduction of PGH₂Aldolase, which have been shown to be linked with essential hypertension and high salt intake (Nagata et al., 2011, Weber, 1980). Aldo-Keto reductase family 1 member A1 was found to have differential expression between Wky and SHRSP strains in both the pilot and main study. The main investigation showed one identification for this protein, with a down regulation in SHRSP strains compared to Wky and maintained in the congenic strain.
- ATP synthase, H⁺ transporting, mitochondrial F1 complex, beta polypeptide (ATPB5), is a part of the mitochondrial ATP synthase F1 catalytic core, from the enzyme, which catalyzes ATP synthesis. This was identified in this investigation multiple times in both the main and pilot investigation. This is an abundance polypeptide as it is required in every cell in the process of ATP production (Runswick and Walker,

1983, Zhou et al., 2006). It was identified at three positions on the gel given in Figure 4-64. They are all at the same level indicating a pI change, as opposed to a molecular weight change. Spot 1177 is particularly interesting due to the rise in just the medulla congenic strain. In the other spots there is a down expression given in the cortex and medulla SHRSP strain as compared to the others. From the literature it appears that there is no known association with hypertension.

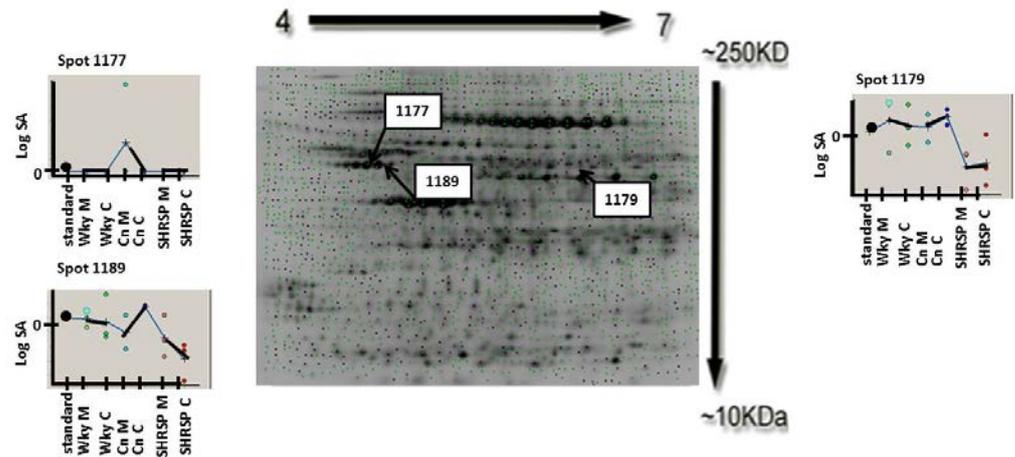


Figure 4-64: Positions and profiles of multiple identifications of ATP synthase, H⁺ transporting, mitochondrial F1 complex, beta polypeptide from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.

- Ubiquinol-Cytochrome c reductase core protein I is a component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex). Its function is part of the respiratory transport chain and therefore is important in ATP production. It has been shown to be differentially express in a number of CVDs (Rai et al., 2005, Hsieh et al., 2006, Kondo, 2008a). Including, in brain ischemia and hypertensive stroke in a SHR rat model being treated with PZH, a Chinese herb. This down regulation was associated with 12 other proteins and an improved prognosis. It was proposed that these down regulations reduced the number of reactive oxygen species (Goodwin et al., 2008b). Within this study the expression showed a down regulation between the Wky and the SHRSP strains with an intermediate expression in the congenic strain.

- Heat shock proteins: A number of heat shock proteins were discovered during this investigation. These are a ubiquitous family of proteins, which aids the folding and unfolding of proteins around the cell. They are also known to be up regulated during oxidative stress. Hypertension increases the stress on vessels and tissues throughout the body and therefore heat shock proteins would be expected to increase in the SHRSP strain as they protect against conformational change (Sitek et al., 2005b). Examining the profiles in Figure 4-65 shows, heat shock proteins both up and down regulated. Due to their ubiquitous nature it is unlikely that the heat shock proteins are going to make very good biomarkers. However, due to the multiple varieties, it is possible that a pattern could be generated with a greater number of identifications. Heat shock proteins have been described in a number of disease states; Cardiomyopathy (Jan Eriksson, 2007), hypertension (Knepper, 2002, Wu et al., 2006a), brain and spinal cord ischemia (Brownridge et al., 2011) and atherosclerosis (Delles et al., 2012)

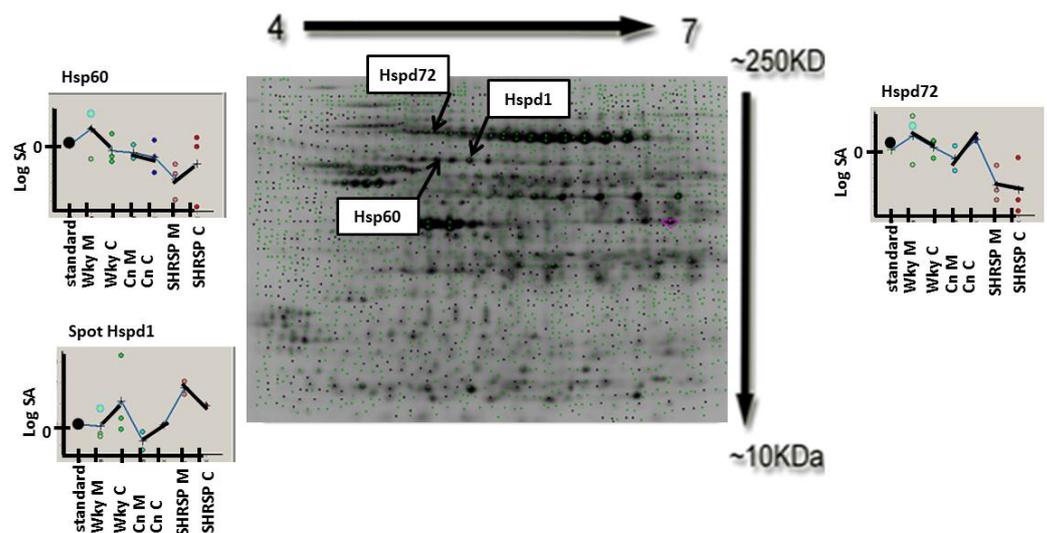


Figure 4-65: Positions and profiles of 3 identified heat shock proteins, from the main investigation. The x-axis represents the rat strain in the order of Standard, Wky medulla, Wky cortex, congenic medulla, congenic cortex, SHRSP medulla and SHRSP cortex. The y-axis displays log standardised abundance.

- Actin, Cytoplasmic 2 otherwise known as Gamma-actin is a ubiquitously expressed protein in eukaryotic cells involved in various processes in motility and forms cytoskeletal structures (Thongboonkerd and Malasit, 2005). They also can have numerous PTM attached to them, which makes them appear on multiple positions on a 2D gel.

They tend to form in a heavy band and due to their abundance are distinct. They are important in maintaining cell and muscle mass and therefore, deficiencies have been associated with myopathy (Nielsen et al., 2002a). Although there is no strong evidence in the literature that directly links gamma-actin to hypertension it has been linked to defects in nephrogenesis in the kidney, which leads to the restriction of glomeruli and leads to increased blood pressure (Pisitkun et al., 2004). A predicted protein similar to cytoplasmic 2 was discovered in multiple spots (suspected due to PTMs) with one interesting spot in particular that was only expressed in congenic medulla.

- Serum albumin is an abundant protein in blood plasma. It has a number of functions in the blood, including controlling osmotic pressure, carriage of hormones and transport of ions (Siew et al., 2011, Pesce et al., 2013, Cutillas et al., 2004a). Albumin has long been a marker for hypertension by its presence in urine (Fliser et al., 2007) and reduces in abundance upon treatment with antihypertensive drugs (Janech et al., 2007). Also, albumin overload has been associated in the activation of intrarenal renin–angiotensin system through protein kinase C and NADPH oxidase-dependent pathway, thus has an important role to play in hypertension (Xu et al., 2005a). In this thesis it was shown that albumin exhibited various profiles in multiple spots. A very high expression in comparison to Wky and congenic was displayed in spot 682 in medulla SHRSP tissue and also up regulation in another two spots in cortex SHRSP tissue. However, it is unlikely to be a specific marker, due to its abundant nature. It may be possible, with validation, to show a pattern of regulation of various albumin isoforms to act as a marker, but this would require a great deal of future work.
- Fructose-1,6-bisphosphatase 1 is an enzyme involved in gluconeogenesis and the calvin cycle by converting fructose-1,6-bisphosphate to fructose 6-phosphate (Kirtley and McKay, 1977, Marcus et al., 1982, Gottschalk et al., 1982). As a protein involved in these key processes, it is present in all cells and the kidney is no exemption. It has shown to be affected by toxin treatment on kidney cortex by the antibiotic Gentamicin, which can cause renal toxicity. It was shown to be down-regulated in high dose samples but up regulated

in low dose samples, the reasons behind this remain unclear (Charlwood et al., 2002b). In this study it was down regulated in SHRSP strains compared to Wky and congenic strains in cortex tissue. This is compared to medulla tissue, where the putative marker was down regulated in congenic by the greatest amount.

- Haemoglobins: Various chain of haemoglobin was discovered in the pilot investigations. This is not surprising, due to the biological relationship of kidney tissue with the blood. The relatively high vascularisation of the kidney tissue is likely to have led to varied blood contamination across different slices. It is cautiously noted that most of the haemoglobin fragments identified were expressed more in the SHRSP strains, speculatively, consistent with blood proteins being forced through glomerular filtration into the nephron. They also were positioned at various spots around the gel, this may have skewed analysis by masking other proteins during identification in the mass spectrometer.

It is unlikely that these identifications alone will be specific biomarkers for hypertension. This is because many of these biomarkers are proteins, which show oxidative stress and are abundant and present in many tissue types.

4.7 Conclusion and Summary

It is clear that understanding and treating hypertension is of great importance in the western world. The complexity of its idiopathic aetiology means investigating causes is difficult as best. In this investigation the need for segregation of tissue types is highlighted and demonstrated by the profile differences exhibited between the cortex and medulla tissue profiles. It appears that much of the changes shown from the pilot study indicated that the cortex was exhibiting a great deal of the WKY maintenance in the congenic interval causing the maintained phenotype, with greater regularity than the SHRSP. In addition, DiGE analysis has confirmed that salt treatment appears to exaggerate phenotypical response and profile changes. An aid in validation was shown in the MSI and label free quantitation.

One major issue and limitation in this investigation was gaining identification for the required spots. In addition, Data mining is obviously essential to pull out the highest quality and quantity of possible markers, but without identification the marker becomes just another spot on a gel map. Therefore a major limitation of this proteomic work flow was the ability to gain identification of the required spots. Therefore future attention needs to be paid to improving and developing the downstream method protocol for gaining identification back from large format DiGE gels. Positively, this study revealed a number of putative biomarkers for hypertension and has highlighted a number of analytical approaches to DiGE data, including the usefulness of Venn diagrams to efficiently visualise data.

5 General Conclusions.

5.1 Introduction to general discussion

The main aim of this thesis was to investigate the use of DiGE in examining two areas of tissue proteomics; protein degradation of mouse brain tissue and biomarker discovery in kidney tissue in hypertension.

- To employ and develop the quantitative method of DiGE to look at tissue profiles.
- To develop analytical strategies to visualise quantitative proteomic data from tissue samples gained from DiGE.
- To assess the proportion of change in proteomic degradation in heat-treated and snap frozen samples.
- To find candidate examples of proteins or markers of degradation of wild type mouse brain tissue.
- To find putative biomarkers in hypertension from kidney tissue

It was hoped that investigating global degradation might have wider implications for not just brain tissue, but start to raise questions about what is a less than glamorous research topic. Initially, from the protein degradation investigation, an expectation of gaining an insight into how degradation may be halted using heat-treatment would be gained, however like many experiments, it conjured more questions, specifically relating to data visualisation and handling. The project thus evolved into the use of a multi-faceted approach to looking at DiGE data. This proved a useful optimisation step for the subsequent investigation into hypertension. In some respects this investigation has surpassed the original scope of the investigation regarding just degradation and yielded some interesting approach to examining DiGE data. Indeed, the large amount of data gained is necessary to make such experiments economical due to their considerable expense. One such experiment costs in the region of £1000s so the most needs to be gained from such an investment. It is acknowledged that the data is wide in nature but shallow and needs a more

focused approach in order to be validated. However, the aim was to assess global degradation and to this extent this study has been relatively successful.

The investigation into hypertension and kidney tissue profiles also yielded some interesting preliminary results regarding putative biomarkers discovered. It also allowed, as a proof of concept, the start of bridging a gap and correlating genomic and proteomic data and making a connection between kidney tissue and hypertension. It also has provided a list of a number of proteins that may have an association with hypertension. This may prove useful in future studies in hypertension, as a starting point for a more targeted approach.

5.2 Major findings

The major findings that are presented in this thesis are as follows;

5.2.1 Proteomic degradation in mouse brain tissue

- There is some evidence, via the use of profiles analysis, that the use of rapid heat treatment on wild type homogenised OCR male mouse brain tissue halts non-specific proteomic degradation for a number of protein spots. Therefore, this could prove useful for some proteins but detrimental for others, where snap-freezing showed stabilisation. However, a greater proportion of spots were stabilised in the heat-treated samples compared to snap-freezing.
- The use of a multi-faceted approach helps in the confirmation that rapid heat-treatment halts global proteomic degradation in wild type homogenised OCR male mouse brain tissue.
- Multi-faceted approaches to analysis would allow for the more targeted approach and more efficient data visualisation required to reduce some of the limitations of DiGE analysis
- Additionally, a complementary proteomic technique (performed and analysed by Dr R.J.A.Goodwin, Research Associate, Glasgow University) supports these conclusions, by showing that heat-treatment halted global proteomic degradation when viewed with MS imaging. The parallel study performed by

Dr Heather Allingham showed that heat-treatment had little effect on the sub 10 kDa fraction of the OCR Male mouse proteome, although some features showed stabilisation with heat-treatment (Allingham, 2012).

- The use of this system is not universally recommended for a number of reasons:
 - The stabilisation is not universal across all spots, so the usefulness of this system depends on what specific proteins of interest to be examined.
 - It was also found that the morphological quality of tissue sections in MALDI imaging limiting its applicability across all proteomic methodologies.
 - In the sub 10 KDa fraction analysis by Allingham, 2012 it was shown that only 1% of features showed any difference between snap-frozen and heat-treated samples.
- However, it was shown that more features, in the post 10KDa region, were stabilised using the system in comparison to the traditional method of snap freezing in OCR mice brain tissue. Therefore, it warrants further investigation in both brain tissue (to confirm the results here) and in other tissues to extend its application.

5.2.1.1 Proteomic degradation in mouse brain tissue

The use of tissue in research, particularly with clinical samples and biopsies, is an increasing field and the need for improvements to sample integrity and upstream processing is paramount to the quality of data downstream. Therefore, the assessment of this technique for the prevention of protein degradation in mouse brain tissue is of importance, particularly to clinical neuroscience. In addition, to fresh tissue extracted from animals, there is a multitude of tissue samples preserved in freezers that may benefit from the use of heat-treatment. This technique is also applicable and complementary to the increasing number of researchers who are now looking into how to reverse the

process of fixation of tissue. If this is to be successful, there will be a need for preservation of this tissue. It is however recognized, that heat treating tissue itself may have long lasting effects that were not tested in this investigation and with much of the proteomic and genomic information to current date obtained from tissue that is preserved by freezing there may be a valid reluctance to break from this tradition. Indeed, it may not be applicable to break from this tradition, as comparison may be extremely difficult. This investigation, however, has opened this avenue for other scientific researchers to have another alternative to freezing or fixing tissue.

In addition to assessing the degradation, by using a multi-faceted approach to the DiGE analysis, it is hoped that some of the methods and ideas pursued in this investigation may be of use in various and wide ranging fields in proteomics, when using DiGE gels. The findings of this study have been echoed in other investigations to study the degradation and heat treatment of neuroproteins and peptides (Smejkal et al., 2011, Scholz et al., 2011, Kultima et al., 2011, Kokkat et al., 2011, Svensson et al., 2009b, Robinson et al., 2009). An interesting consideration for researchers considering this as an alternative to snap-freezing, is that it was shown by Robinson et al, 2009, that although heat treatment stabilised a significant subset of proteins in mouse brain, it showed no significant differences for the treatment of mouse heart tissue. This shows that heat treatment will not necessarily be of benefit to every tissue. Additionally, Robinson et al, 2009 showed that within the heat treated mouse brain that the proteins preserved often showed higher abundances and maintained PTM such as phosphorylations. They also concluded that this may be of more interest in certain disease states. Therefore, at best, this is likely to be a complementary technique to be used in conjunction with traditional methods of storage and stabilisation (Robinson et al., 2009). It is clear that a great deal more research is required in degradative processes, in order to improve sample stability and gain maximum accuracy and impact from tissue samples.

5.2.2 Proteomic Profiling of Kidney tissue

The proteomic profiles of Kidney tissue in WKY, congenic and SHRSP rat model were accessed and allude to;

- Medulla and cortex, whilst having markers in common and similarities, possesses a different proteomic profile, validating the need for segregation.

In future experiments, in order to gain the most detailed information needed for biomarker discovery.

- Genomic and proteomic linkage of data, segregation is required.
- The use of salt-treatment appears to exaggerate the effect on proteomic expression or phenotype. This is validated by the phenotypical response given by spot profiles and the measured raised blood pressure from the animal strains.
- Presented in chapter 4 are some candidate markers which may be of interest in a future investigation and some of these have had prior biological association with hypertension and numerous other diseases.
- The potential to marry proteomic and genomic data is preliminarily displayed and with careful collaboration and experimental design, genotype and phenotype can be linked using DiGE and genomic data.

5.2.2.1 Proteomic Profiling of Kidney tissue

The use of this approach in the analysis of kidney tissue had several advantages. This study further demonstrated the need for segregation of tissue types and that indeed homogenisation of tissue does not allow for an accurate and targeted approach. In need of further investigation, is the requirement for assessing where the boundary of tissue types may be. A large degree of effort, therefore, needs to be placed into mining down into smaller and smaller samples, as proteomics lacks this crucial amplification process that benefits genomics.

Additionally, by using a recognized model with micro array data associated with the strains of rat used, it was hoped that there would be the possibility of future studies, particularly in respect to associating genomic and proteomic data which would start to fill any gaps between genotype and phenotypes.

Once again, the analytical approach could be used for any number of studies of different diseases and tissue types and therefore it was felt that this impact is wide and large.

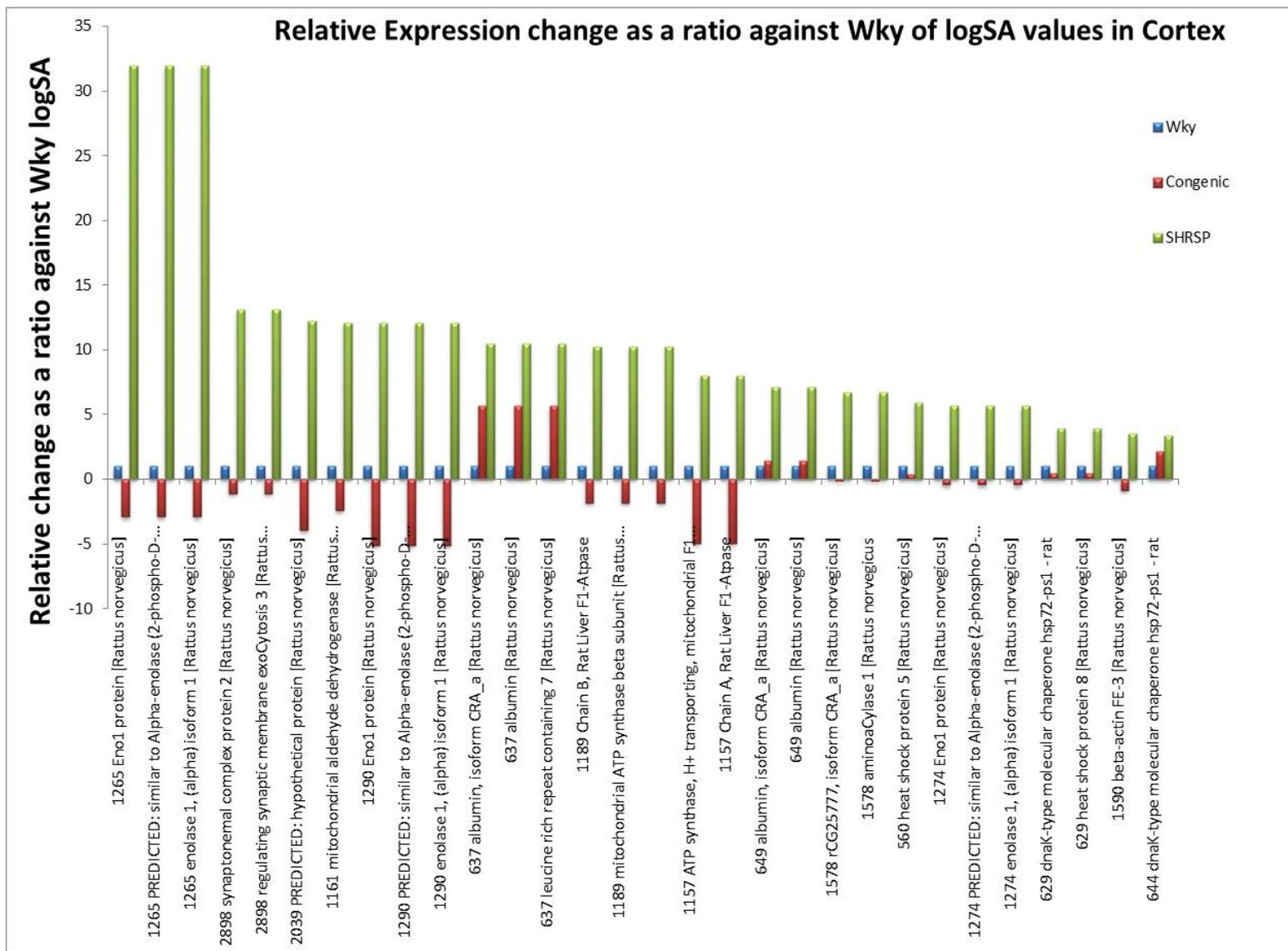
5.2.2.1.1 Biological significance of the putative biomarkers

There have been a number of identifications made in this investigation, which have associated biologically significant processes and links to hypertension. Some of the most significant and abundant proteins have been discussed in section 4.6.7.

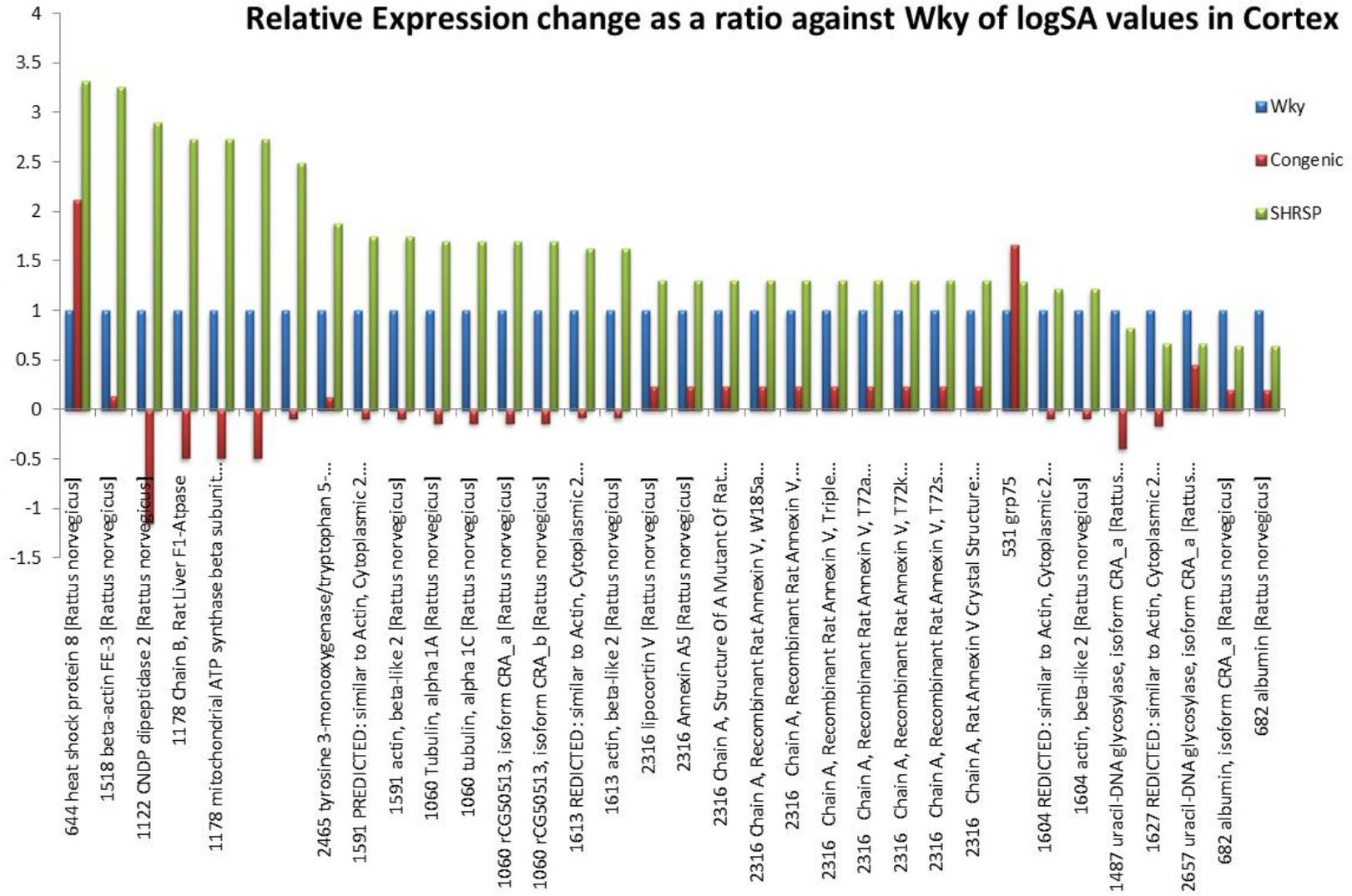
It is unlikely that any of the identified putative markers would be specific markers for hypertension, as they are abundant proteins linked with numerous other states of pathology. Additionally, they would require validation with techniques such as those described in 1.3.4.2 with antibodies specific to those markers. After this sort of validation, it may be possible to compile a list of markers associated with hypertension and use an approach where multiple markers could be used in a more specific diagnosis. A biomarker profile, if you will. In order that further research can be done, a relative profile of putative markers has been created from the identified proteins in this study in Figure 5-1 and Figure 5-2 in section 5.2.2.1.1.1.

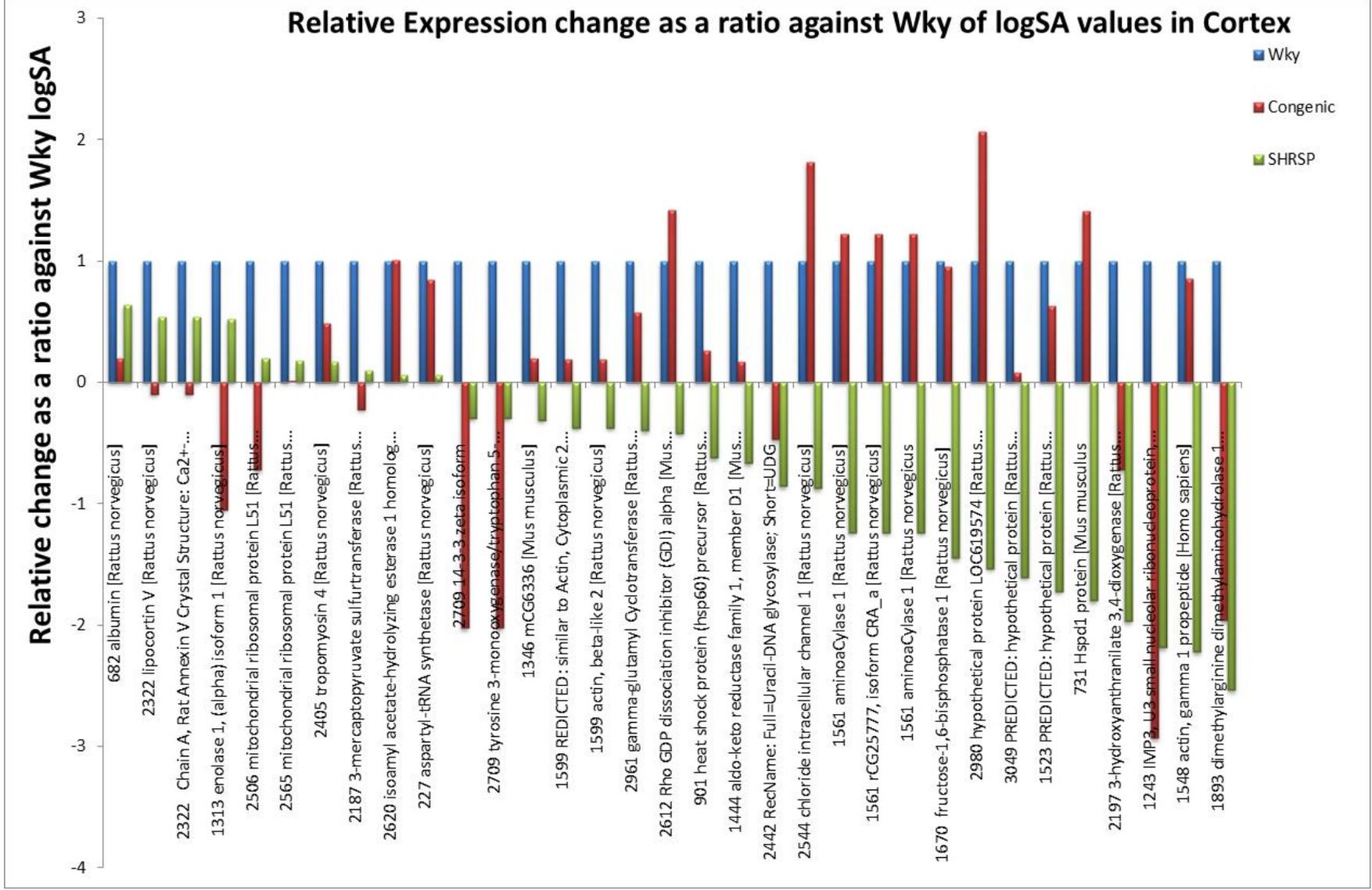
5.2.2.1.1.1 *Putative biomarker quantitative profile pattern*

In order to provide a way that future researcher may compare the gel spot identifications and aid the validation of these putative markers, the quantitative values have been displayed graphically by applying the log standardised abundance as a ratio against Wky readings. Wky therefore is always represented by 1.00 (=Wky/Wky) and then an up or down regulation in congenic or SHRSP (=Congenic/Wky or =SHRSP/Wky). This allows the visualisation of all the identifications given in Table 7-2 and their relative expression in comparison to Wky. They are shown for cortex tissue in the graphs in Figure 5-1 and medulla in Figure 5-2. This, in conjunction with the gel spot images, allows for a comparison for any future investigations using a similar model or checking it for similarities in clinical samples.



Relative change as a ratio against Wky logSA





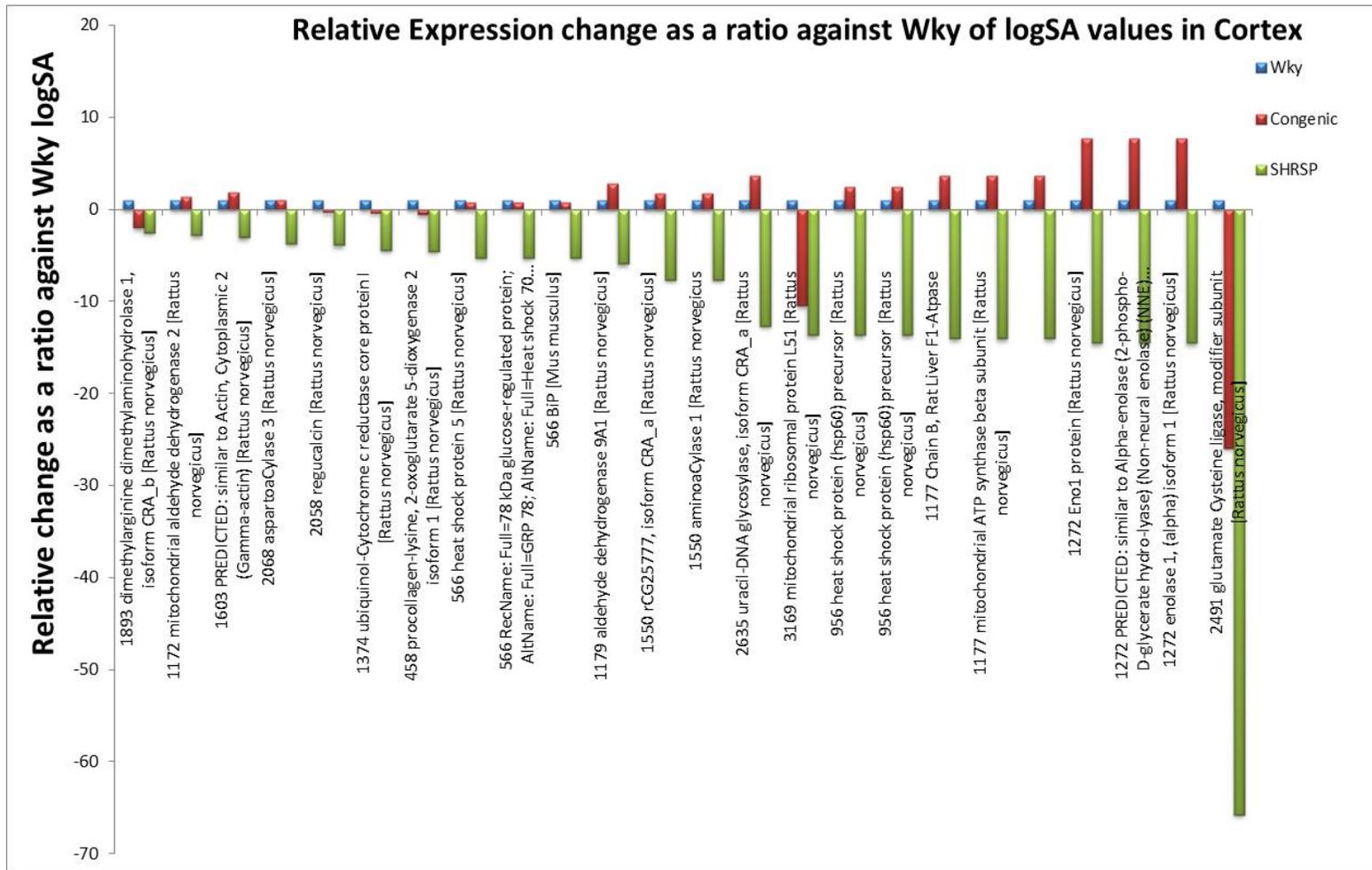
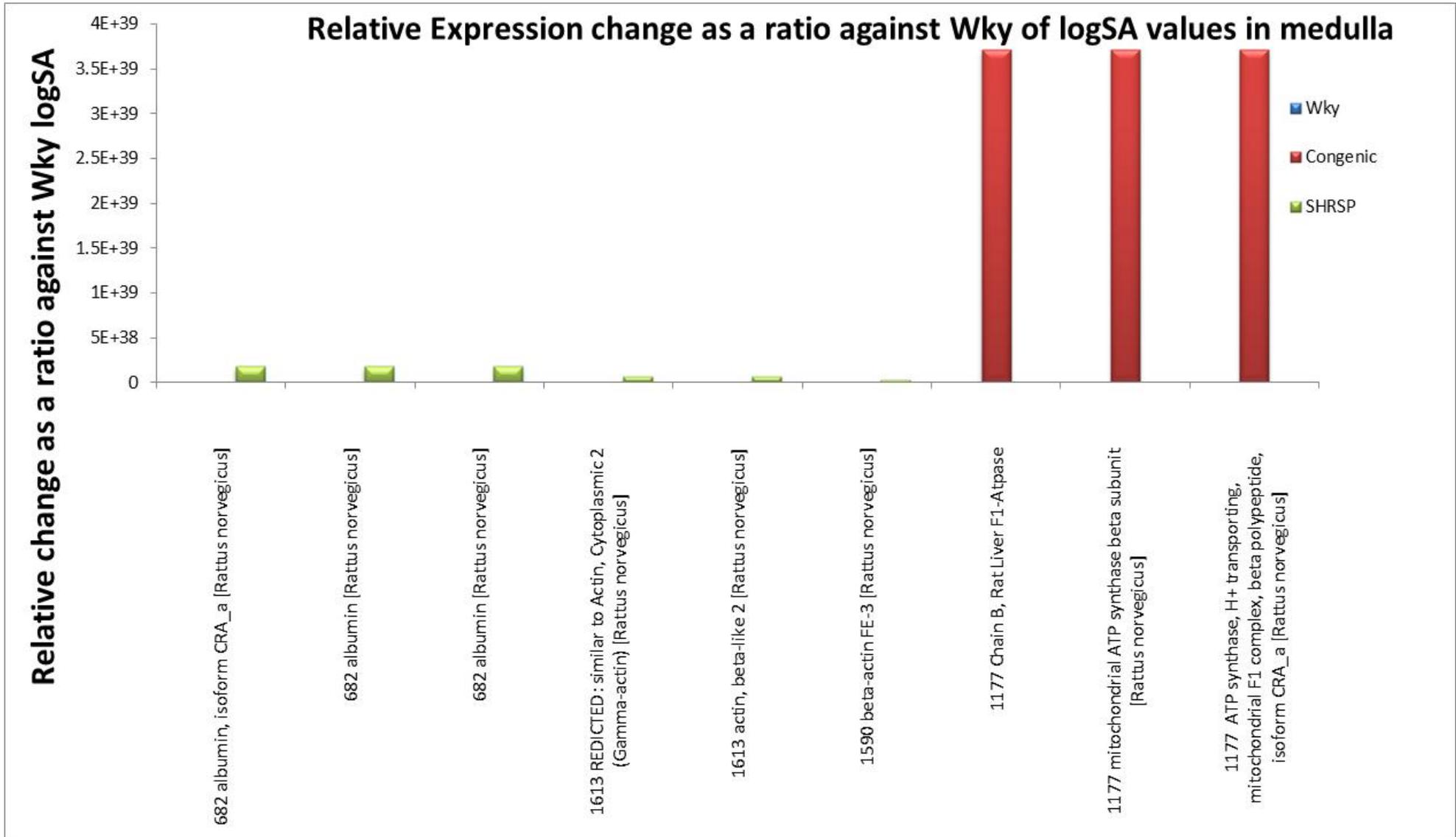
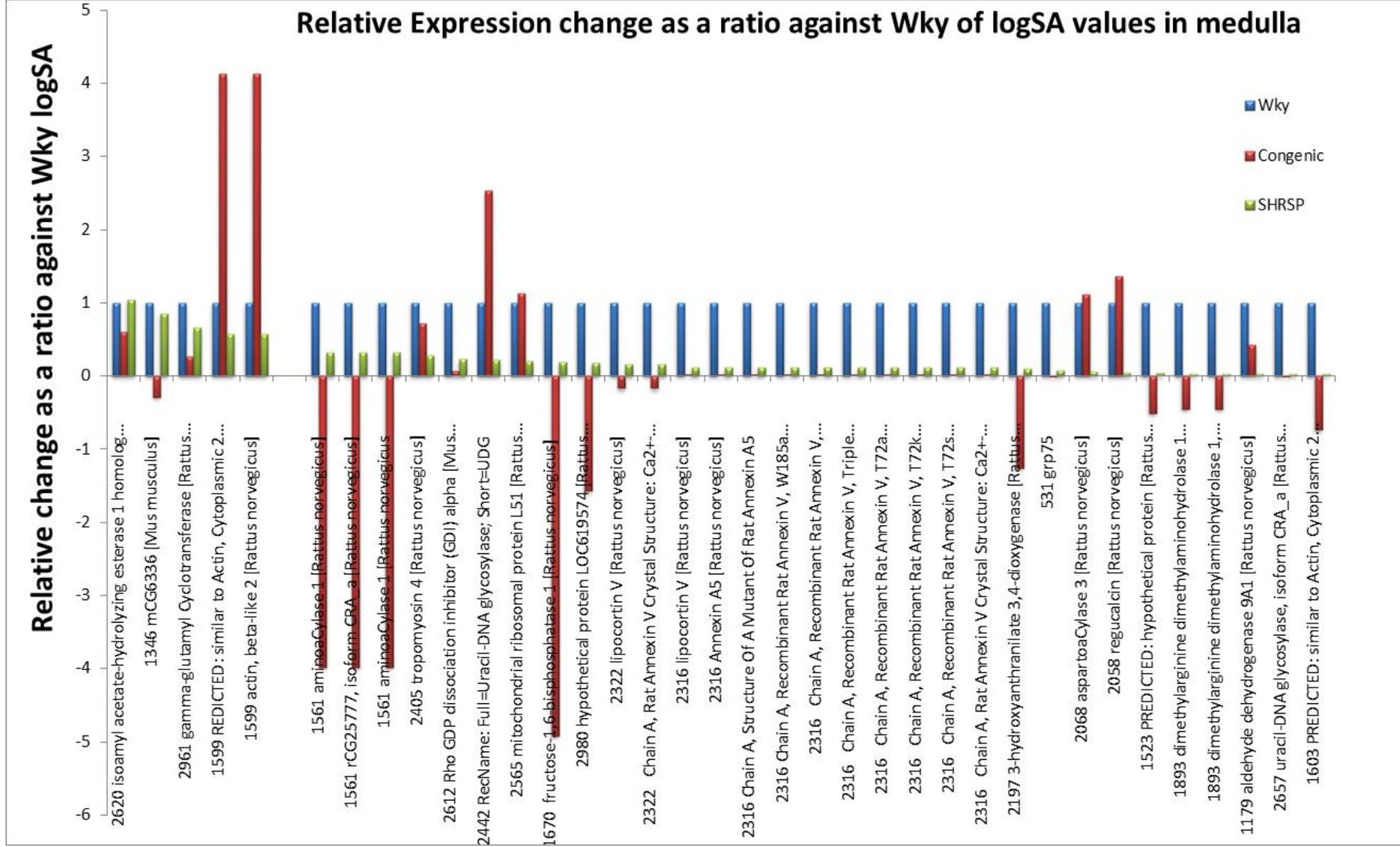
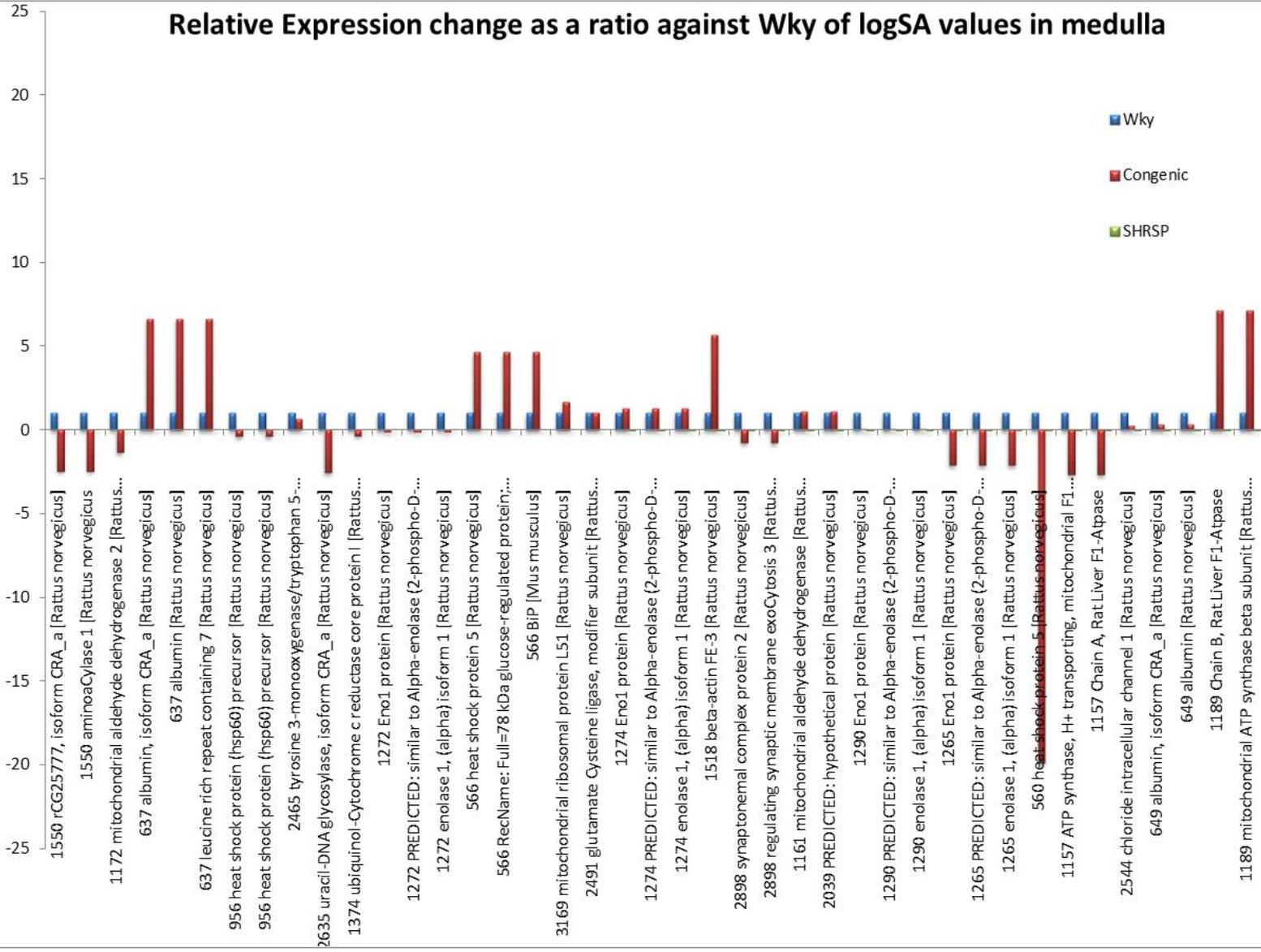


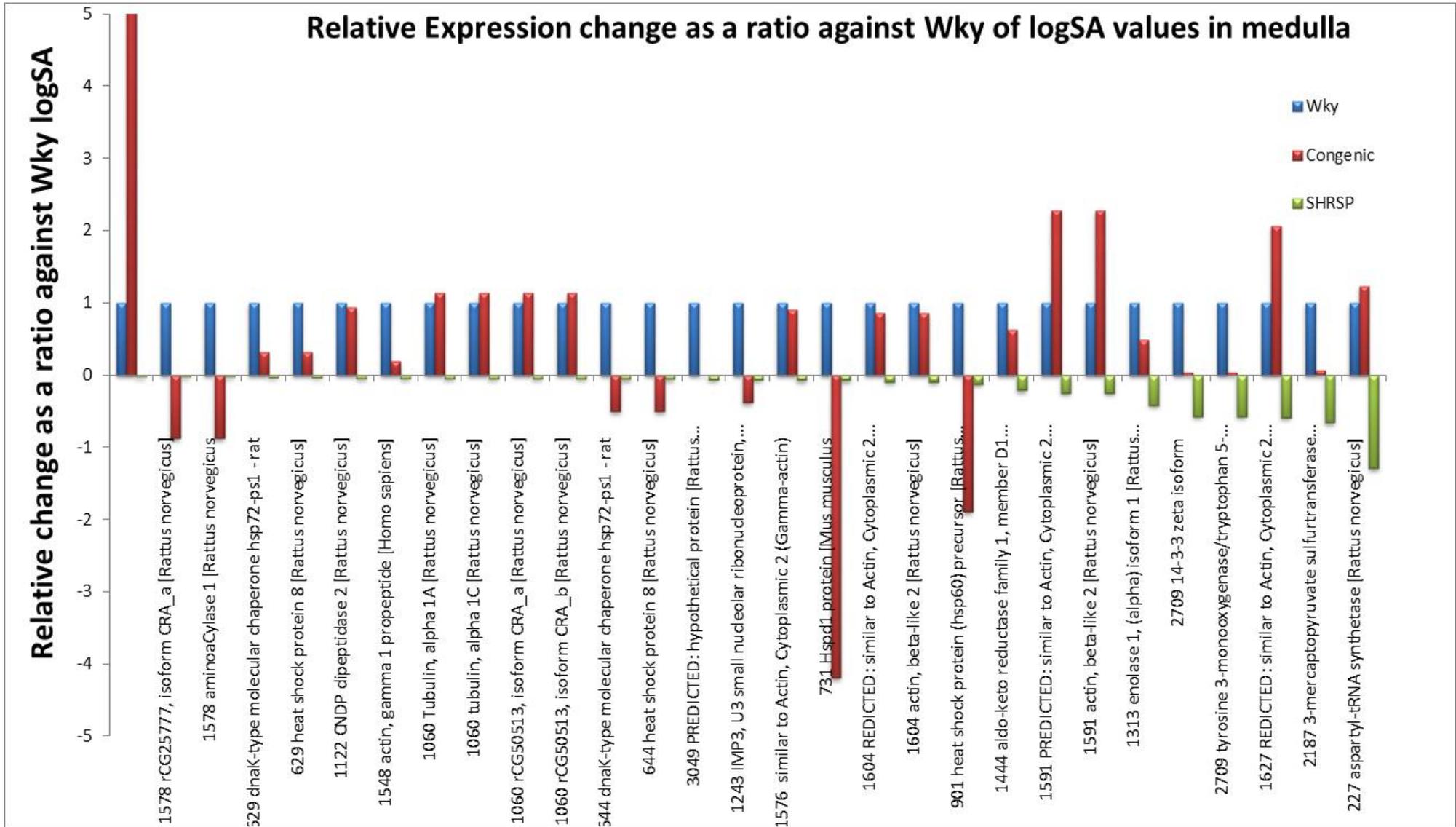
Figure 5-1: Set of graphs showing the relative expression of Wky, Congenic and SHRSP strains. Columns are an expression of the ratio of themselves against Wky stain. It is proposed that these identifications might serve to aid future work into looking for a profile pattern from 2D-DiGE gels. These are for Cortex tissue.





Relative change as a ratio against Wky of logSA values in medulla





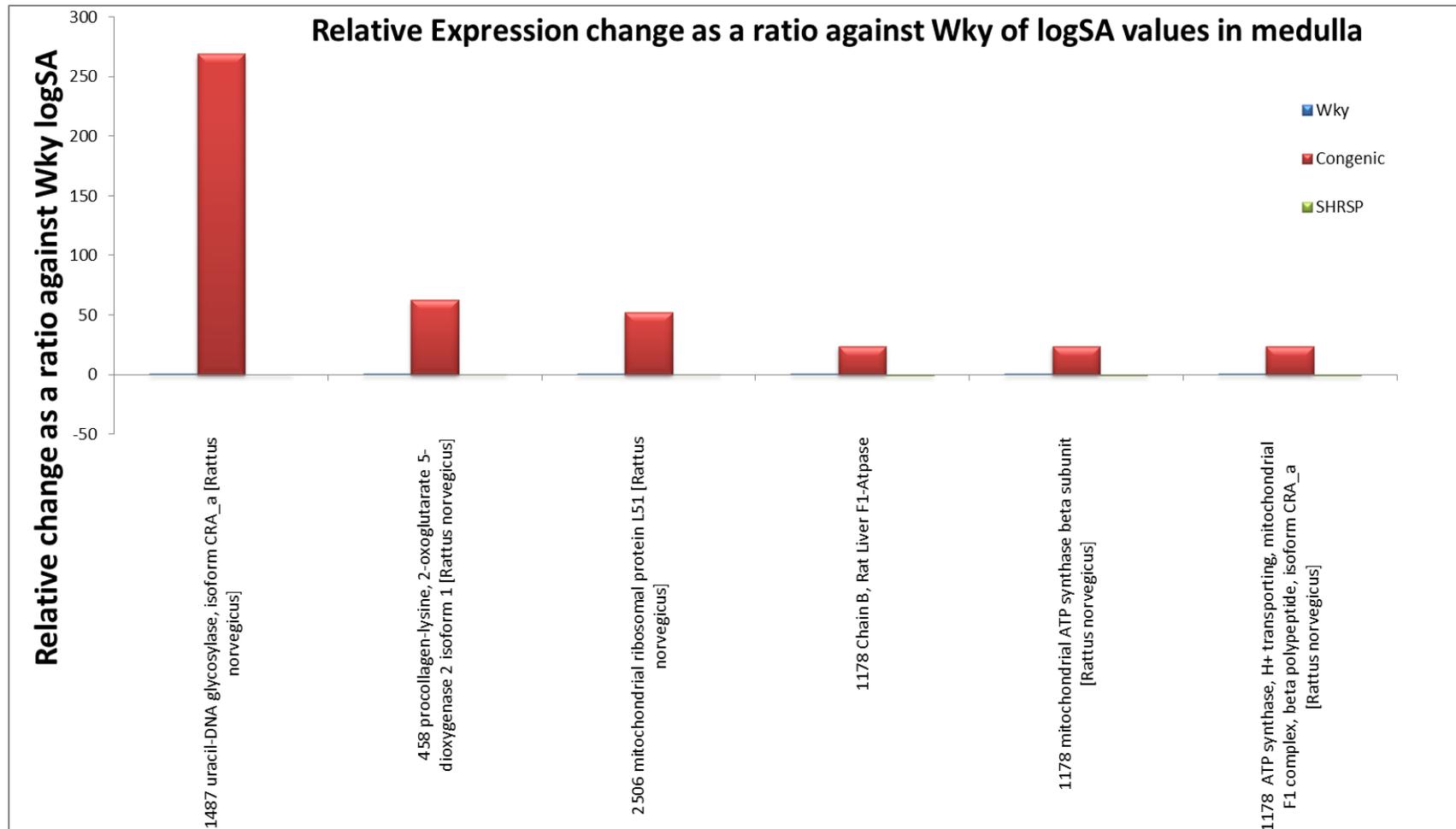


Figure 5-2: Set of graphs showing the relative expression of Wky, Congenic and SHRSP strains. Columns are an expression of the ratio of themselves against Wky stain. It is proposed that these identifications might serve to aid future work into looking for a profile pattern from 2D-DiGE gels. These are for medulla tissue.

5.3 Limitations of DiGE

Any scientific investigation has limitations. Although DiGE itself manages to surpass the standard 2DE in its ease of multiplexing gels and getting meaningful quantitation, it is somewhat shadowed by the expense of experiments. Running large format gels is a relatively time and labour intensive process, coupled with the expense of the proprietary Cydyes. This expense means optimisation is also costly and many labs shy away from performing DiGE. Additionally, due to the number of protocol steps, DiGE is a skilled process with many stages which can go wrong. However, it can be argued in the long term, that these costs fade in the substantial information gained in return. The cost can equally be mitigated by careful experimental design and multiplexing thus gaining high quality quantifiable and robust data.

The quantity of data makes analysis difficult, hence much of this investigation was spent trying to sort and optimise an approach to deal with the enormous quantities of data.

The next problem encountered in DiGE is gaining identification. After a lot of time, effort and money has been expended, gaining identifications can be somewhat “hit and miss” with difficulties in spots picking (if done by hand very time consuming and there are technical issues with automation) and mass spectrometry.

However, DiGE itself should never be expected to be a magic bullet due to its limitations in resolving hydrophobic proteins or difficulties with the low molecular weight or low abundant proteins. Many of these limitations, however, are an intrinsic product of the dynamic field of proteins and require a large jump in theory or technology to breach these difficulties. Small steps are the likely immediate future, hence the requirement for close collaboration and complementary methodologies.

5.4 DiGE as a biomarker discovery platform

It is clear that DiGE can provide a vast amount of quantitative information with regards to biomarker discovery, depicting a snap shot in time for proteomic analysis on a global scale. However, it is labour intensive to run such gels and more so to analyze the results. Therefore, the use of DiGE in biomarker discovery is a very time intensive prospect. At one time, 2DE and DiGE were the gold standard for proteomic analysis. Even with their limitations, there was nothing that could revival the range and magnitude of data. This is almost certainly no longer the case. Recent developments in mass spectrometry, coupled with liquid chromatography mean a new era of high dynamic range and quantitative techniques, to analyse high numbers of proteins and peptides, in a more automated manner. Isotopic labelling has brought mass spectrometry into the realms of quantitation, once reserved for DiGE. Biomarker discovery and the complexity and dynamic range of samples requires high throughput, high resolution and high sensitivity in order gain the specificity that clinicians and mass spectrometry has the potential to deliver this. However, DiGE still has a place for biomarker discovery as a complementary technique in order to cross validate biomarkers from MS, visualise isoforms and PTMs.

The major barrier to biomarker discovery is the integration of researchers and clinicians. According to Rifai et al, 2006 the way to increase the number of clinically relevant biomarkers is to have a predefined “pipeline” to make a coherent connection between discovery and clinical assay. This is certainly necessary, as there is a high degree of proteomics focused on biomarker discovery but very little focus on validation, with even fewer in direct collaboration with clinicians in the development of assays for clinical use. This is partly a collaboration issue and the pipeline would help, however, I believe it is also an issue with expense (discovery is less expensive as fewer samples are utilised), dynamic complexity and technical limitations. In many respects it is acknowledged that, although this study has made impact in the areas discussed, it is a biomarker discovery based investigation (chapter 4) and has not bridged the gap between discovery – validation – clinical applications. However, the future is hopeful, as when this biomarker pipeline is coherent and there will be a huge bank of data from proteomic discovery experiments. In addition, there is a

prospect in the future where mass spectrometry could be directly used in clinical diagnosis by the building of biomarker MS profiles data banks to compare clinical samples. There is no real possibility that DiGE will be employed in this fashion.

5.5 Future Investigation

The original vision of the project was to work towards small sample proteomics but the data visualising element of DiGE proved vast, complicated and time expensive. However, some of the future work, which I believe is required for mining down into the high information, was started.

The need for going smaller and more sensitive comes from the limitations of proteomics compared to genomics. Genomics is a field of study that is not plagued with the same issues that surrounds Proteomics. With the inception of Polymerase Chain Reaction (PCR) problems with sensitivity were banished overnight. Proteomics on the other hand suffers greatly by not having a means of amplification and therefore there is a major need and requirement to increase the levels of sensitivity and detection required. With the lack of an amplification method, it is often necessary to pool samples to get enough material to work with. However, this leads to a lot of issues and loosing information in background noise and signals. This is particularly the case in tissue proteomics, where by homogenising tissue causes an averaging effect across what might be very different types of cells. As noted earlier; what exactly is “medulla” and “cortex” and is there something inbetween? Whilst doing this investigation, the next step in the proteomics work flow was considered: Small sample proteomics. Below is an outline of some work that was started to help the move towards getting to a small sample work flow, and would benefit many different investigations in proteomics, particularly with regards to working with tissue.

5.5.1 Saturation labelling of kidney tissue

In order to work towards a small sample work flow, more sensitive techniques are required. One such technique is saturation labelling. Unlike minimal labelling, it is essential to determine the correct concentration of dye to use, as under labelling and multiple labelling becomes an issue. In working towards a small sample work flow in this investigation, a dye determination experiment was conducted to ascertain the correct concentration of dye for brain tissue. The results are shown in Figure 5-3. All scanner settings were kept constant and performed consecutively. Most tissue types require 2nM- 4nM concentration. Those tissues with a large amount of Cysteines sometime require additional dye. For kidney tissue it is shown that 4nM is sufficient for efficient labelling with little smearing

and a dye concentration of up to 8nM can be performed without over labelling occurring.

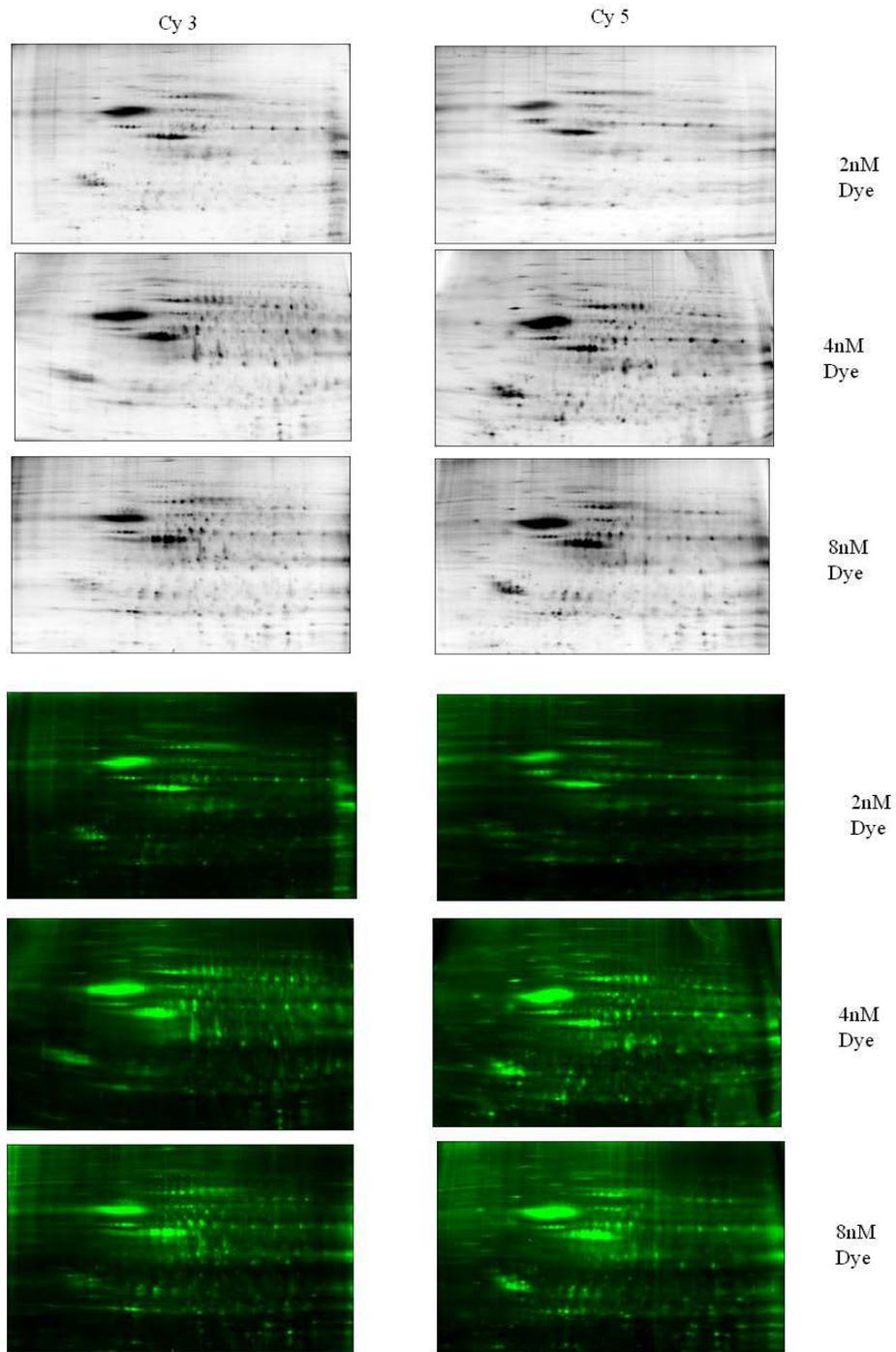


Figure 5-3: Figure showing Dye concentration determination for DiGE saturation labelling (labelling for scarce samples). All images were scanned using GE Healthcare Typhoon Scanner at the same scan settings. Fluorescent and traditional gel displays are shown.

5.5.2 Determination of Protein concentration

If proteins are to be accurately quantified using DiGE, then a protein concentration must be obtained in order to load the correct amounts on the DiGE gels. This becomes increasingly difficult as sample size gets smaller, meaning that a high percentage of samples might be used in order to gain the required measurement for protein concentration. The standard methods of protein concentration determination are to construct a standard curve, using an assay such as a Bradford assay or BSA. However, in order to miniaturise, the possibility of using Amersham's, Bioanlayers Chip reader was investigated. It is predominantly used for DNA quantitation but is increasingly being used to run small proteomic samples. The chip is effectively 10 small micro fluidic capillary and it acts like a miniature capillary electrophoresis machine, separating the sample. A standard curve can be built from the separation and protein concentration of samples can be determined. Initial investigation was promising, but as can be seen in Figure 5-4, the lysis buffer used in DiGE contains chaps which reduces the ability of the chip reader to resolve the signal and therefore is not compatible with a DiGE application.

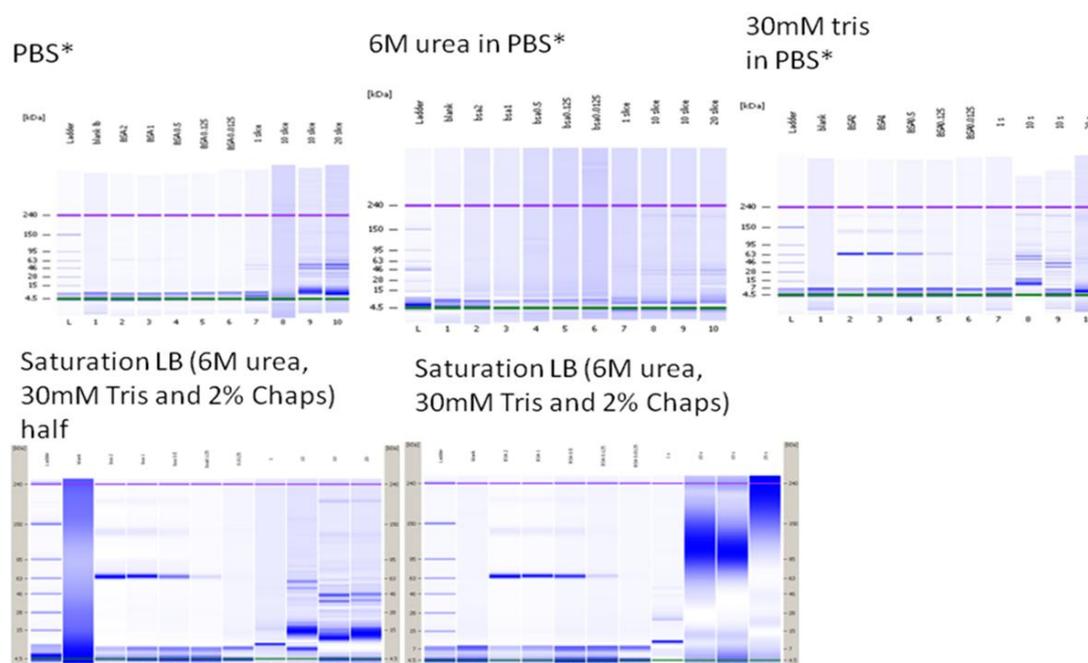


Figure 5-4: Figure showing gel representation of the effect of DiGE lysis buffer on the resolution of Amersham Bioscience, Bioanalyser Chip reader. If true miniaturisation and quantitation is to be achieved a reliable form of protein concentration determination is required for small samples. The lysis Buffer used in DiGE labelling contains a number of components and is not compatible with the Bioanalyser Chip reader, as the concentration of Chaps rises the resolution is lost making calibration curves too inaccurate for protein concentration determination. Lane 1. Protein marker, 2-7 is BSA standard in water. 8-10 wild type mouse brain samples in solution labelled above.

5.5.3 Laser Micro-dissection of tissue and Mass spectrometry

The logistical difficulties of obtaining small samples are high. One well documented strategy is to use laser micro-dissection. The compatibility of LMD with MALDI-MS was assessed. Wild type OCR 21 week rat brain tissue of 14µm thick was sliced using a Lieca Cryostat and the thaw mounted onto Lieca LMD slides. The slide is made from polyethylene. LMD was performed by cutting out relatively large discs of 1mm in diameter onto an ITO slide ready to be tested on a Bruker Ultraflex III MALDI-TOF-MS. The tissue Discs were immediately spotted with 5µl of sinapinic acid (matrix) in acetonitrile (ACN) solution as shown in Figure 5-5. Initially, MS signal was poorly resolved. It was thought that the polyethylene was acting as a barrier and was preventing co-crystallisation with the matrix solution. It was discovered necessary to disturb the tissue discs in order to get resolved signal from the tissue. An example spectrum is shown in Figure 5-5.

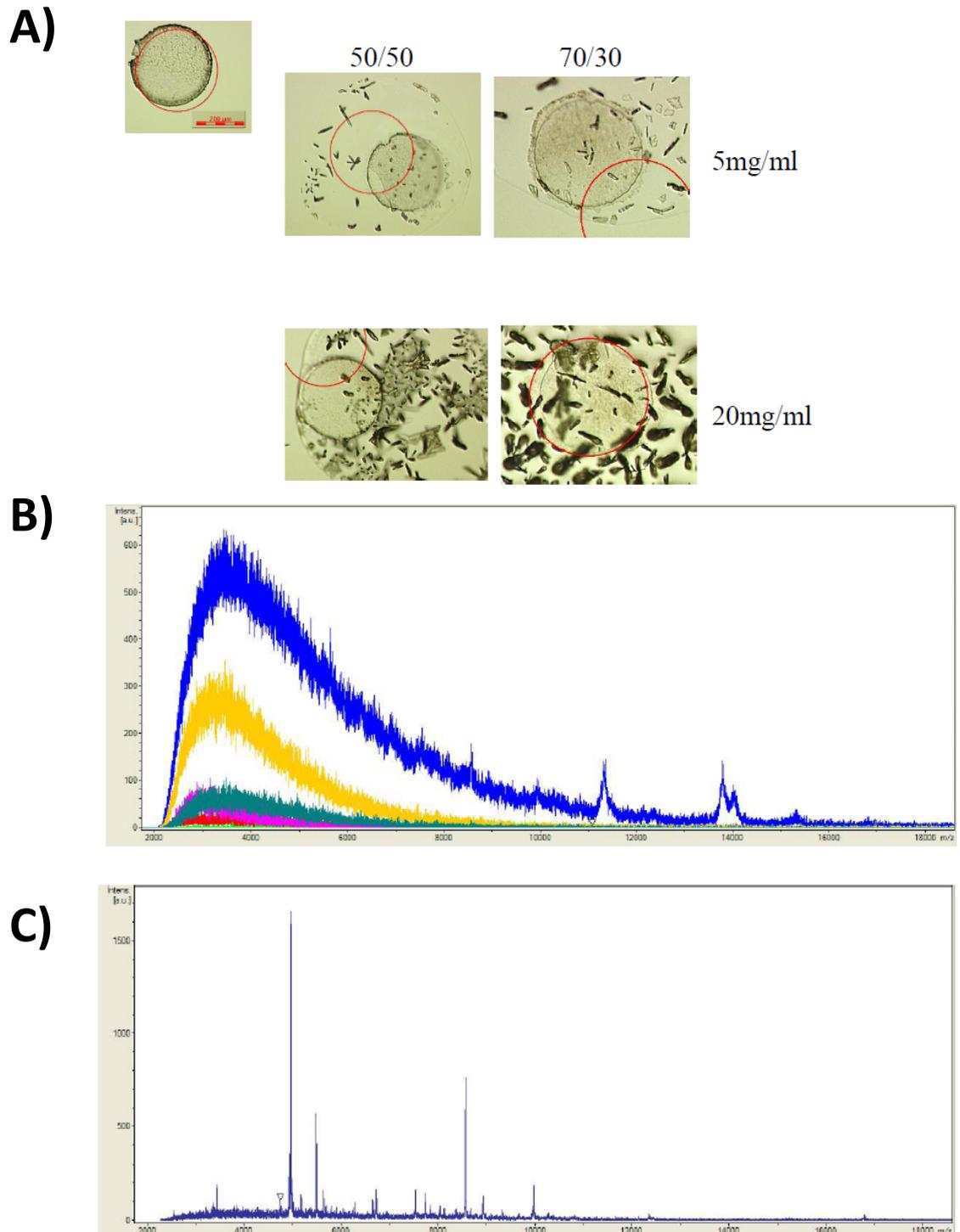


Figure 5-5: Laser Micro-Dissection. The effect of LMD slide plastic on downstream mass spectrometry. A) Discs of Tissue adhered to LMD slide Polyethylene film, and spotted in various concentrations of matrix on an ITO slide ready for MALDI-MS. 50/50 ACN Solution B) Mass spectra at various laser intensities. As can be seen resolution is very poor. C) Mass spectra of tissue after it was disturbed using Sterile Gilson tip.

5.6 Summary

This investigation into DiGE and tissue proteomic profiling has been successful in some aspects. The investigation has had some preliminary impact in two main areas;

Specific impact;

- That using rapid heat can lead to the halting of proteomic degradation, probably due to the inactivation of proteolytic enzymes, thus maintaining the integrity and state of the given tissue (mouse brain).
- Profiling of kidney tissue in a WKY, congenic and SHRSP rat model has led to the discovery of candidate markers for hypertension.

Wider implications;

- Shows a number of different methods for the analysis of DiGE data that may be applicable to numerous applications.
- It displays a number of candidate markers that could be used in proteomic degradation research.
- It displays a number of candidate makers that may be linked to hypertension ready for future investigation and validation
- Shows methods for data visualisation in DiGE experiments for proteomic profiling of tissues.
- Shows that the marrying of proteomic and genomic data has potential in bridging the genotypical and phenotypical gap.
- Additionally, some general optimisation work regarding small sample proteomics and saturation labelling is displayed and is presented as future aspects to this thesis.

6 References

1. ABDUL-SALAM, V. B., PAUL, G. A., ALI, J. O., GIBBS, S. R., RAHMAN, D., TAYLOR, G. W., WILKINS, M. R. & EDWARDS, R. J. 2006. Identification of plasma protein biomarkers associated with idiopathic pulmonary arterial hypertension. *PROTEOMICS*, 6, 2286-2294.
2. ABONNENC, M. & MAYR, M. 2012. *Proteomics of Atherosclerosis*.
3. ADAM, B.-L., QU, Y., DAVIS, J. W., WARD, M. D., CLEMENTS, M. A., CAZARES, L. H., SEMMES, O. J., SCHELLHAMMER, P. F., YASUI, Y. & FENG, Z. 2002. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Research*, 62, 3609-3614.
4. ADAM, B.-L., VLAHOU, A., SEMMES, O. J. & WRIGHT JR, G. L. 2001. Proteomic approaches to biomarker discovery in prostate and bladder cancers. *PROTEOMICS*, 1, 1264-1270.
5. ADDIS, T., BARRETT, E., POO, L. & YUEN, D. 1947. The relation between the serum urea concentration and the protein consumption of normal individuals. *Journal of Clinical Investigation*, 26, 869.
6. ADDONA, T. A., ABBATIELLO, S. E., SCHILLING, B., SKATES, S. J., MANI, D., BUNK, D. M., SPIEGELMAN, C. H., ZIMMERMAN, L. J., HAM, A. J. L. & KESHISHIAN, H. 2009. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nature Biotechnology*, 27, 633-641.
7. ADDONA, T. A., SHI, X., KESHISHIAN, H., MANI, D., BURGESS, M., GILLETTE, M. A., CLAUSER, K. R., SHEN, D., LEWIS, G. D. & FARRELL, L. A. 2011. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nature Biotechnology*, 29, 635-643.
8. AFKARIAN, M., BHASIN, M., DILLON, S. T., GUERRERO, M. C., NELSON, R. G., KNOWLER, W. C., THADHANI, R. & LIBERMANN, T. A. 2010. Optimizing a proteomics platform for urine biomarker discovery. *Molecular & Cellular Proteomics*, 9, 2195-2204.
9. AHMED, M. M. & GARDINER, K. J. 2011. Preserving protein profiles in tissue samples: Differing outcomes with and without heat stabilization. *Journal of Neuroscience Methods*, 196, 99-106.
10. AHMED, M. M., STURGEON, X., ELLISON, M., DAVISSON, M. T. & GARDINER, K. J. 2012. Loss of Correlations among Proteins in Brains of the Ts65Dn Mouse Model of Down Syndrome. *Journal of Proteome Research*, 11, 1251-1263.
11. AHOLA, H., HEIKKILA, E., ASTROM, E., INAGAKI, M., IZAWA, I., PAVENSTADT, H., KERJASCHKI, D. & HOLTHOFER, H. 2003. A novel protein, densin, expressed by glomerular podocytes. *J Am Soc Nephrol* 14, 1731-1737.
12. AHRENDTS, R., PIEPER, S., KÜHN, A., WEISSHOFF, H., HAMESTER, M., LINDEMANN, T., SCHELER, C., LEHMANN, K., TAUBNER, K. & LINSCHIED, M. W. 2007. A Metal-coded Affinity Tag Approach to Quantitative Proteomics. *Molecular & Cellular Proteomics*, 6, 1907-1916.
13. AITOKALLIO, T. 2010. Dealing with missing values in large-scale studies: microarray data imputation and beyond. *Briefings in bioinformatics*, 11, 253-264.

14. ALBRECHT, S., SCHOLS, H. A., KLARENBEK, B., VORAGEN, A. G. & GRUPPEN, H. 2010. Introducing Capillary Electrophoresis with Laser-Induced Fluorescence (CE–LIF) as a Potential Analysis and Quantification Tool for Galactooligosaccharides Extracted from Complex Food Matrices. *Journal of agricultural and food chemistry*, 58, 2787–2794.
15. ALKHALAF, A., ZÜRBIG, P., BAKKER, S. J. L., BILO, H. J. G., CERNA, M., FISCHER, C., FUCHS, S., JANSSEN, B., MEDEK, K., MISCHAK, H., ROOB, J. M., ROSSING, K., ROSSING, P., RYCHLÍK, I., SOURIJ, H., TIRAN, B., WINKLHOFFER-ROOB, B. M., NAVIS, G. J. & FOR THE, P. G. 2010. Multicentric Validation of Proteomic Biomarkers in Urine Specific for Diabetic Nephropathy. *PLoS ONE*, 5, e13421.
16. ALLINGHAM, H. 2012. *Development of proteomic techniques for biomarker discovery*. PhD, University of Glasgow.
17. ALLISON, D. B., CUI, X., PAGE, G. P. & SABRIPOUR, M. 2006. Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews Genetics*, 7, 55–65.
18. ALTELAAR, A. I. M., ET AL 2006. High-resolution MALDI imaging mass spectrometry allows localization of peptide distributions at cellular lengths scales in pituitary tissue. *International Journal of Mass Spectrometry* 260, 203–211.
19. ALTRIA, K. 1999. Overview of capillary electrophoresis and capillary electrochromatography. *Journal of Chromatography A* 000–000.
20. ANDERSON, F. & MURPHY, R. 1976. Isocratic separation of some purine nucleotide, nucleoside, and base metabolites from biological extracts by high-performance liquid chromatography. *J Chromatogr.*, 121(2), 251–62.
21. ANDERSON, L. 2005. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J Physiol*, 563, 23–60.
22. ANDERSON, N., ANDERSON, N. & TOLLAKSEN, S. 1979. Proteins of human urine. I. Concentration and analysis by two-dimensional. *electrophoresis. Clin Chem* 25, 1199–210.
23. ANDERSON, N. L. 2010. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin Chem*, 56, 177–185.
24. ANDREASEN, N., MINTHON, L., CLARBERG, A., DAVIDSSON, P., GOTTFRIES, J., VANMECHELEN, E., VANDERSTICHELE, H., WINBLAD, B. & BLENNOW, K. 1999. Sensitivity, specificity, and stability of CSF-tau in AD in a community-based patient sample. *Neurology*, 53, 1488–1488.
25. ANNEST, J. L., ET AL 1979. Familial affreftation of blood pressure and weight in adoptive families. II. Estimation of the relative contributions of genetic and common environmental factors to blood pressure correlation between family members. *American Journal of Epidemiology*, 110, 492–503.
26. ANNEST, J. L., ET AL 1983. Familial aggregation of blood pressure and weight in adoptive families. III. analysis of the role of shared genes and shred household environment in explaining family resemblance for height, weight and selected weight/height indeices. *American Journal of Epidemiology*, 117, 492–506.
27. ANTONICKA, H., MATTMAN, A., CARLSON, C.G., GLERUM, D.M., HOFFBUHR, K.C., LEARY, S.C., KENNAWAY, N.G., AND SHOUBRIDGE, E.A. 2003. Mutations in COX15 produce a defect in the mitochondrial heme biosynthetic pathway, causing early-onset fatal hypertrophic cardiomyopathy. *Am. J. Hum. Genet.*, 72, 101–114.

28. ARAKAWA, T., PRESTRELSKI, S. J., KENNEY, W. C. & CARPENTER, J. F. 2001. Factors affecting short-term and long-term stabilities of proteins. *Advanced drug delivery reviews*, 46, 307-326.
29. ARIKE, L., VALGEPEA, K., PEIL, L., NAHKU, R., ADAMBERG, K. & VILU, R. 2012. Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *Journal of Proteomics*.
30. ARTHUR, J. M., THONGBOONKERD, V., SCHERZER, J. A., CAI, J., PIERCE, W. M. & KLEIN, J. B. 2002. Differential expression of proteins in renal cortex and medulla: A proteomic approach. *Kidney International*, 62, 1314–1321.
31. ASTON, F. W. 1919. The Mass-Spectra of Chemical Elements. *Phil. Mag.*, XXXVIII, 707.
32. ASTON, J., BALL, R., POPE, J., JONES, K. & COCKER, J. 2002. Development and validation of a competitive immunoassay for urinary S-phenylmercapturic acid and its application in benzene biological monitoring. *Biomarkers*, 7, 103-112.
33. ATTARD, B., AITKEN, J. & BAKER, M. 2004. TWO-DIMENSIONAL DIFFERENTIAL IN GEL ELECTROPHORESIS (DIGE): A NOVEL METHOD FOR HIGH THROUGHPUT PROTEOMICS. *AUSTRALIAN BIOCHEMIST*, 35.
34. AZIMZADEH, O., BARJAKTAROVIC, Z., AUBELE, M., CALZADA-WACK, J., SARIOGLU, H., ATKINSON, M. J. & TAPIO, S. 2010. Formalin-fixed paraffin-embedded (FFPE) proteome analysis using gel-free and gel-based proteomics. *Journal of Proteome Research*, 9, 4710-4720.
35. BACHMAIR, A., FINLEY, D. AND VARSHAVSKY, A. 1986. In vivo Half-Life of a Protein Is a Function of Its Amino-Terminal Residue. *Science*, 234, 179-186.
36. BALLARD, J. L., PEEVA, V. K., DESILVA, C. J. S., LYNCH, J. L. & SWANSON, N. R. 2007. Comparison of Alexa Fluor® and CyDye™ for practical DNA microarray use. *Molecular biotechnology*, 36, 175-183.
37. BANKS, R. E., DUNN, M. J., FORBES, M. A., STANLEY, A., PAPPIN, D., NAVEN, T., GOUGH, M., HARNDEN, P. & SELBY, P. J. 1999. The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis — Preliminary findings. *Electrophoresis*, 20, 689-700.
38. BAÑÓN-MANEUS, E., DIEKMANN, F., CARRASCAL, M., QUINTANA, L., MOYA-RULL, D., BESCÓS, M., RAMÍREZ-BAJO, M., ROVIRA, J., GUTIERREZ-DALMAU, A., SOLÉ-GONZÁLEZ, A., ABIÁN, J. & CAMPISTOL, J. 2010. Two-dimensional difference gel electrophoresis urinary proteomic profile in the search of nonimmune chronic allograft dysfunction biomarkers. *Transplantation*, 15, 548-58.
39. BANTSCHIEFF, M., DÜMPPELFELD, B. & KUSTER, B. 2004. Femtomol sensitivity post-digest 18O labeling for relative quantification of differential protein complex composition. *Rapid Communications in Mass Spectrometry*, 18, 869-876.
40. BANTSCHIEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry*, 389, 1017-1031.
41. BAO, Y. & ZUO, L. 2009. Effect of repeated freeze-thaw cycles on urinary albumin-to-creatinine ratio. *Scandinavian Journal of Clinical & Laboratory Investigation*, 69, 886-888.
42. BEER, L. A., TANG, H. Y., BARNHART, K. T. & SPEICHER, D. W. 2011a. Plasma biomarker discovery using 3D protein profiling coupled with label-free quantitation. *Methods Mol. Biol*, 728, 3-27.

43. BEER, L. A., TANG, H. Y., SRISWASDI, S., BARNHART, K. T. & SPEICHER, D. W. 2011b. Systematic discovery of ectopic pregnancy serum biomarkers using 3-D protein profiling coupled with label-free quantitation. *J. Proteome Res*, 10, 1126-1138.
44. BELLOMO, R., RONCO, C., KELLUM, J. A., MEHTA, R. L. & PALEVSKY, P. 2004. Acute Dialysis Quality Initiative workgroup. Acute renal failure-definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care*, 8, R204-R212.
45. BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.
46. BENJAMINI, Y. & HOCHBERG, Y. 2000. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, 25, 60-83.
47. BENSON, M., WHANG, I., PANTUCK, A., RING, K., KAPLAN, S., OLSSON, C. & COONER, W. 1992. Prostate specific antigen density: a means of distinguishing benign prostatic hypertrophy and prostate cancer. *The Journal of urology*, 147, 815.
48. BERANOVA-GIORGIANNI, S. 2003. Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations *Trends in Analytical Chemistry*, 22, 273-281
49. BERGLUND, G., ANDERSSON, O. & WILHELMSEN, L. 1976. Prevalence of primary and secondary hypertension: studies in a random population sample. *British Medical Journal*, , 2, 554-556.
50. BERGMANN, U., AHRENDTS, R., NEUMANN, B., SCHELER, C. & LINSCHIED, M. 2012. Application of Metal-Coded Affinity Tags (MeCAT): Absolute Protein Quantification with Top-Down and Bottom-Up Workflows by Metal-Coded Tagging. *Anal Chem*, 84, 5268-5275.
51. BERHANE, B. T., ZONG, C., LIEM, D. A., HUANG, A., LE, S., EDMONDSON, R. D., JONES, R. C., QIAO, X., WHITELEGGE, J. P. & PING, P. 2005. Cardiovascular-related proteins identified in human plasma by the HUPO Plasma Proteome Project Pilot Phase. *PROTEOMICS*, 5, 3520-3530.
52. BERLIER, J. E., ROTHE, A., BULLER, G., BRADFORD, J., GRAY, D. R., FILANOSKI, B. J., TELFORD, W. G., YUE, S., LIU, J., CHEUNG, C.-Y., CHANG, W., HIRSCH, J. D., BEECHEM ROSARIA P. HAUGLAND, J. M. & HAUGLAND, R. P. 2003. Quantitative Comparison of Long-wavelength Alexa Fluor Dyes to Cy Dyes: Fluorescence of the Dyes and Their Bioconjugates. *Journal of Histochemistry & Cytochemistry*, 51, 1699-1712.
53. BERSON, S., YALOW, R., BAUMAN, A., ROTHSCHILD, M. & NEWERLY, K. 1956. Insulin-I metabolism in human subjects: demonstration of insulin binding globulin in the circulation of insulin-treated subjects. . *Journal of Clinical Investigation* 35, 170-90.
54. BEYNON, R. J., DOHERTY, M. K., PRATT, J. M. & GASKELL, S. J. 2005. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nature methods*, 2, 587-589.
55. BEYNON, R. J. & PRATT, J. M. 2005. Metabolic labeling of proteins for proteomics. *Molecular & Cellular Proteomics*, 4, 857-872.
56. BIANCHI, B., FOX, U., ET AL 1974. Blood pressure changes produced by kidney corss-transplantation between Spontaneously hypertensive rates and normotensive rats. *Clinical Science of Molecular Medicine*, 47y, 435-448.

57. BINZ, P. M., M. ET AL 1999. A molecular scanner to automate proteomic research and to display proteomic images. *Analytical Chemistry*, 71, 4981-4988.
58. BIRON, D. G., BRUN, C., LEFEVRE, T., LEBARBENCHON, C., LOXDALE, H. D., CHEVENET, F., BRIZARD, J. P. & THOMAS, F. 2006. The pitfalls of proteomics experiments without the correct use of bioinformatics tools. *PROTEOMICS*, 6, 5577-5596.
59. BLAUSTEIN, M., ZHANG, J., CHEN, L. & HAMILTON, B. 2006. How does salt retention raise blood pressure? . *American Journal of Physiology-Regulatory Integrative and Comparative Physiology*, 290, R514-R523.
60. BLENNOW, K., HAMPEL, H., WEINER, M. & ZETTERBERG, H. 2010. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nature Reviews Neurology*, 6, 131-144.
61. BLUMENSTEIN, M., MCMASTER, M. T., BLACK, M. A., WU, S., PRAKASH, R., COONEY, J., MCCOWAN, L. M., COOPER, G. J. & NORTH, R. A. 2009. A proteomic approach identifies early pregnancy biomarkers for preeclampsia: novel linkages between a predisposition to preeclampsia and cardiovascular disease. *PROTEOMICS*, 9, 2929-2945.
62. BOERSEMA, P. J., AYE, T. T., VAN VEEN, T. A., HECK, A. J. & MOHAMMED, S. 2008. Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *PROTEOMICS*, 8, 4624-4632.
63. BOERSEMA, P. J., RAIJMAKERS, R., LEMEER, S., MOHAMMED, S. & HECK, A. J. 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature protocols*, 4, 484-494.
64. BOSSO, N., CHINELLO, C., PICOZZI, S., GIANAZZA, E., MAININI, V. & GALBUSERA, C. 2008. Human urine biomarkers of renal cell carcinoma evaluated by ClinProt. . *Proteomics Clin Appl* 2, 1036-46.
65. BOTTARI, P., AEBERSOLD, R., TURECEK, F. & GELB, M. 2003. Design and Synthesis of Visible Isotope-Coded Affinity Tags for the Absolute Quantification of Specific Proteins in Complex Mixtures. *Bioconjugate Chemistry*, 2, 380-388.
66. BREWSTER, U. & PERAZELLA, M. 2004. The Renin-Angiotensin-Aldosterone System and the Kidney: Effects on Kidney Disease. *Am J Med.* , 116 263-272.
67. BROWNRIDGE, P., HOLMAN, S. W., GASKELL, S. J., GRANT, C. M., HARMAN, V. M., HUBBARD, S. J., LANTHALER, K., LAWLESS, C., O'CUALAIN, R., SIMS, P., WATKINS, R. & BEYNON, R. J. 2011. Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy. *PROTEOMICS*, 11, 2957-2970.
68. BRUN, V., DUPUIS, A., ADRAIT, A., MARCELLIN, M., THOMAS, D., COURT, M., VANDENESCH, F. & GARIN, J. 2007. Isotope-labeled Protein Standards: Toward Absolute Quantitative Proteomics. *Molecular & Cellular Proteomics*, 6, 2139-2149.
69. BRUN, V., MASSELON, C., GARIN, J. & DUPUIS, A. 2009. Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, 72, 740-749.
70. BRUNET, S., THIBAUT, P., GAGNON, E., KEARNEY, P., BERGERON, J. J. & DESJARDINS, M. 2003. Organelle proteomics: looking at less to see more. *Trends in cell biology*, 13, 629-638.
71. BUGGER, H., CHEN, D., RIEHLE, C., SOTO, J., THEOBALD, H. A., HU, X. X., GANESAN, B., WEIMER, B. C. & ABEL, E. D. 2009. Tissue-specific remodeling of the mitochondrial proteome in type 1 diabetic akita mice. *Diabetes*, 58, 1986-1997.

72. BURKE, B., GUNGADOO, J., MARÇANO, A., NEWHOUSE, S., SHIEL, J., CAULFIELD, M. & MUNROE, P. 2005. Monogenic Forms of Human Hypertension. *Comprehensive Hypertension (First Edition)*, 417-428.
73. BURT, V., CUTLER, J., HIGGINS, M. & AL, E. 1995. Trends in the prevalence, awareness, treatment, and control of hypertension in the adult US population. Data from the health examination surveys, 1960 to 1991. *Hypertension* 26, 60-9.
74. CADIEUX, P. A., BEIKO, D. T., WATTERSON, J. D., BURTON, J. P., HOWARD, J. C., KNUDSEN, B. E., GAN, B. S., MCCORMICK, J. K., CHAMBERS, A. F. & DENSTEDT, J. D. 2004. Surface-enhanced laser desorption/ionization, time of flight, mass spectrometry (SELDI_TOF_MS): A new proteomic urinary test for patients with urolithiasis. *Journal of clinical laboratory analysis*, 18, 170-175.
75. CAFFREY, R. E. 2010. A review of experimental design best practices for proteomics based biomarker discovery: focus on SELDI-TOF. *Methods Mol Biol*, 641, 167-183.
76. CALVO, K. R., LIOTTA, L. A. & PETRICOIN, E. F. 2005. Clinical proteomics: from biomarker discovery and cell signaling profiles to individualized personal therapy. *Bioscience reports*, 25, 107-125.
77. CAÑAS, B., PIÑEIRO, C., CALVO, E., LÓPEZ-FERRER, D. & GALLARDO, J. M. 2007. Trends in sample preparation for classical and second generation proteomics. *Journal of Chromatography A*, 1153, 235-258.
78. CANTUTI-CASTELVETRI, I. & AL, E. 2002. Differential Gene expression in dopaminergic neurons isolated from post mortem human parkinson disease brain. *Society of Neuroscience*
79. CAPPUCCIO, F. P., KALAITZIDIS, R., DUNECLIFT, S. & EASTWOOD, J. B. 2000. Unravelling the links between calcium excretion, salt intake, hypertension, kidney stones and bone metabolism. *Journal of nephrology*, 13, 169-177.
80. CAPRIOLI, R. M., T.B. FARMER, AND J. GILE, 1997. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. . *Analytical Chemistry*, 69, 4751-60.
81. CARGILE, B. J., BUNDY, J. L. & STEPHENSON JR, J. L. 2004. Potential for false positive identifications from large databases through tandem mass spectrometry. *Journal of Proteome Research*, 3, 1082-1085.
82. CARR, S. A. & ANDERSON, L. 2008. Protein Quantitation through Targeted Mass Spectrometry: The Way Out of Biomarker Purgatory? *Clin Chem*, 54, 1749-1752.
83. CARRETERO, O. & OPARIL, S. 2000. Essential hypertension. Part I: definition and etiology *Circulation* 101 329–35.
84. CARTY, D. M., DELLES, C. & DOMINICZAK, A. F. 2008. Novel biomarkers for predicting preeclampsia. *Trends in cardiovascular medicine*, 18, 186.
85. CASTAÑO, E., ROHER, A., ESH, C., KOKJOHN, T. & BEACH, T. 2006. Comparative proteomics of cerebrospinal fluid in neuropathologically-confirmed Alzheimer's disease and non-demented elderly subjects. *Neurological research*, 28, 155-163.
86. CASTRONOVO, V., WALTREGNY, D., KISCHEL, P., ROESLI, C., ELIA, G., RYBAK, J.-N. & NERI, D. 2006. A chemical proteomics approach for the identification of accessible antigens expressed in human kidney cancer. *Molecular & Cellular Proteomics*, 5, 2083-2091.
87. CEDAZO-MINGUEZ, A. & WINBLAD, B. 2010. Biomarkers for Alzheimer's disease and other forms of dementia: clinical needs, limitations and future aspects. *Experimental gerontology*, 45, 5-14.

88. CEUPPENS, R., DUMONT, D., VAN BRUSSEL, L., VAN DE PLAS, B., DANIELS, R., NOBEN, J.-P., VERHAERT, P., VAN DER GUCHT, E., ROBBEN, J., CLERENS, S. & ARCKENS, L. 2007. Direct profiling of myelinated and demyelinated regions in mouse brain by imaging mass spectrometry. *International Journal of Mass Spectrometry*, 260, 185.
89. CHANG, W. C., HUANG, L. C. L., WANG, Y. S., PENG, W. P., CHANG, H. C., HSU, N. Y., YANG, W. B. AND CHEN, C. H. 2007. Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. . *Analytica Chimica Acta* 582, 1-9.
90. CHARLWOOD, J., SKEHEL, J., KING, N., CAMILLERI, P., LORD, P., BUGELSKI, P. & ATIF, U. 2002a. Proteomic analysis of rat kidney cortex following treatment with gentamicin. . *J Proteome Res* 1, 73–82.
91. CHARLWOOD, J., SKEHEL, J. M., KING, N., CAMILLERI, P., LORD, P., BUGELSKI, P. & ATIF, U. 2002b. Proteomic Analysis of Rat Kidney Cortex Following Treatment with Gentamicin. *Journal of Proteome Research*, 1, 73-82.
92. CHAURAND, P., CORNETT, D. S. & CAPRIOLI, R. M. 2006. Molecular imaging of thin mammalian tissue sections by mass spectrometry. *Current opinion in biotechnology*, 17, 431-436.
93. CHAURAND, P., ET AL., 2005. Imaging Mass Spectrometry: Principles and Potentials. . *Toxicologic Pathology*, , 33, 92.
94. CHAURAND, P., RAHMAN, M. A., HUNT, T., MOBLEY, J. A., GU, G., LATHAM, J. C., CAPRIOLI, R. M. & KASPER, S. 2008. Monitoring mouse prostate development by profiling and imaging mass spectrometry. *Molecular & Cellular Proteomics*, 7, 411-423.
95. CHAURAND, P., SCHWARTZ, S. A. & CAPRIOLI, R. M. 2002. Imaging mass spectrometry: a new tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Current Opinion in Chemical Biology*, 6, 676-681.
96. CHE, F.-Y., LIM, J., PAN, H., BISWAS, R. ET AL., 2005. Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. . *Mol. Cell. Proteomics* 4, 1391-1405.
97. CHEN, C.-H., SUN, L. & MOCHLY-ROSEN, D. 2010a. Mitochondrial aldehyde dehydrogenase and cardiac diseases. *Cardiovascular research*, 88, 51-57.
98. CHEN, R., PAN, S., BRETNALL, T. A. & AEBERSOLD, R. 2005. Proteomic profiling of pancreatic cancer for biomarker discovery. *Molecular & Cellular Proteomics*, 4, 523-533.
99. CHEN, Y.-T., CHEN, C.-L., CHEN, H.-W., CHUNG, T., WU, C.-C., CHEN, C.-D., HSU, C.-W., CHEN, M.-C., TSUI, K.-H., CHANG, P.-L., CHANG, Y.-S. & YU, J.-S. 2010b. Discovery of Novel Bladder Cancer Biomarkers by Comparative Urine Proteomics Using iTRAQ Technology. *Journal of Proteome Research*, 9, 5803-5815.
100. CHENG, A.-L., HUANG, W.-G., CHEN, Z.-C., PENG, F., ZHANG, P.-F., LI, M.-Y., LI, F., LI, J.-L., LI, C. & YI, H. 2008. Identification of novel nasopharyngeal carcinoma biomarkers by laser capture microdissection and proteomic analysis. *Clinical Cancer Research*, 14, 435-445.
101. CHERTOW, G., BURDICK, E., HONOUR, M., BONVENTRE, J. & BATES, D. 2005a. Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. *J Am Soc Nephrol*, 16, 3365 - 3370.
102. CHERTOW, G. M., BURDICK, E., HONOUR, M., BONVENTRE, J. V. & BATES, D. W. 2005b. Acute Kidney Injury, Mortality, Length of Stay, and Costs in Hospitalized Patients. *Journal of the American Society of Nephrology*, 16, 3365-3370.

103. CHEUNG, B., ONG, K., MAN, Y., LAM, K. & LAU, C. 2006. Prevalence, awareness, treatment, and control of hypertension: United States National Health and Nutrition Examination Survey 2001-2002. *J Clin Hypertens (Greenwich)* 8, 93-98.
104. CHICH, J.-F., DAVID, O., VILLERS, F., SCHAEFFER, B., LUTOMSKI, D. & HUET, S. 2007. Statistics for proteomics: Experimental design and 2-DE differential analysis. *Journal of Chromatography B*, 849, 261-272.
105. CHOBANIAN, A., BAKRIS, G., BLACK, H. & AL, E. 2003a. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension*, 42, 1206-52.
106. CHOBANIAN, A., BAKRIS, G., BLACK, H., CUSHMAN, W., GREEN, L., IZZO, J. & AL., E. 2003b. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. . *Journal of American Medical Association*. .
107. CHOUGALE, A. D., BHAT, S. P., BHUJBAL, S. V., ZAMBARE, M. R., PUNTAMBEKAR, S., SOMANI, R. S., BOPANA, R., GIRI, A. P. & KULKARNI, M. J. 2012. Proteomic analysis of glycosylated proteins from streptozotocin-induced diabetic rat kidney. *Molecular biotechnology*, 50, 28-38.
108. CHUGHTAI, K. & HEEREN, R. M. A. 2010. Mass Spectrometric Imaging for Biomedical Tissue Analysis. *Chemical Reviews*, 110, 3237-3277.
109. CLARK, J., JEFFS, B., DAVIDSON, A., LEE, W., ANDERSON, N., BIHOREAU, M. & AL, E. 1996; . Quantitative trait loci in genetically hypertensive rats - Possible sex specificity. . *Hypertension* 28, 898-906.
110. COCA, A. 2008. Economic benefits of treating high-risk hypertension with angiotensin II receptor antagonists (blockers). *Clinical Drug Investigation* 28 211-20.
111. COCA, S., YALAVARTHY, R., CONCATO, J. & PARIKH, C. 2007. Biomarkers for the diagnosis and risk stratification of acute kidney injury: A systematic review. *International Society of Nephrology*, 73, 1008-1016.
112. COLGRAVE, M. L., XI, L., LEHNERT, S. A., FLATSCHER-BADER, T., WADENSTEN, H., NILSSON, A., ANDREN, P. E. & WIJFFELS, G. 2011. Neuropeptide profiling of the bovine hypothalamus: Thermal stabilization is an effective tool in inhibiting post-mortem degradation. *PROTEOMICS*, 11, 1264-1276.
113. COMITER, C., CAPELOUTO, C., TOBIN, M., DLUHY, R. & RICHIE, J. 1995. GLUCOCORTICOID REMEDIABLE ALDOSTERONISM: A RARE HEREDITARY FORM OF ADRENOCORTICOTROPIC HORMONE REGULATED MINERALOCORTICOID HYPERTENSION. *JOURNAL OF UROLOGY*, 154, 510-312.
114. COON, J. J., ZÜRBIG, P., DAKNA, M., DOMINICZAK, A. F., DECRAMER, S., FLISER, D., FROMMBERGER, M., GOLOVKO, I., GOOD, D. M. & HERGET-ROSENTHAL, S. 2008. CE_MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. *PROTEOMICS-Clinical Applications*, 2, 964-973.
115. COONS, A. H. 1961. The beginnings of immunofluorescence. *Journal of Immunology*, 87, 499-503.
116. CORESH, J., ASTOR, B. C., GREENE, T., EKNOYAN, G. & LEVEY, A. S. 2003. Prevalence of chronic kidney disease and decreased kidney function in the adult US population: Third National Health and Nutrition Examination Survey. *American Journal of Kidney Diseases*, 41, 1-12.
117. CORNETT, D. C., RM 2006. A novel histology directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in Human breast cancer. *MCP*, 5, 1975-83.

118. COX, A. G., PESKIN, A. V., PATON, L. N., WINTERBOURN, C. C. & HAMPTON, M. B. 2009. Redox potential and peroxide reactivity of human peroxiredoxin 3. *Biochemistry*, 48, 6495-6501.
119. COX, B. & EMILI, A. 2006. Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nature protocols*, 1, 1872-1878.
120. CRAVEN, R. A., STANLEY, A. J., HANRAHAN, S., DODS, J., UNWIN, R., TOTTY, N., HARNDEN, P., EARDLEY, I., SELBY, P. J. & BANKS, R. E. 2006. Proteomic analysis of primary cell lines identifies protein changes present in renal cell carcinoma. *PROTEOMICS*, 6, 2853-2864.
121. CRAVEN, R. A., VASUDEV, N. S. & BANKS, R. E. 2013. Proteomics and the search for biomarkers for renal cancer. *Clinical Biochemistry*.
122. CRISTONI, S. B., LUIGI ROSSI 2004. Bioinformatics in mass spectrometry data analysis for proteomics studies *Expert Review of Proteomics*, 1, 469-483(15).
123. CSERHÁTI, T. 2010. Data evaluation in chromatography by principal component analysis. *Biomed Chromatogr.* , 24, 20-8.
124. CUTILLAS, P., BURLINGAME, A. & UNWIN, R. 2004a. Proteomic Strategies and Their Application in Studies of Renal Function. *Physiology*, 19, 114-119.
125. CUTILLAS, P. R., BIBER, J., MARKS, J., JACOB, R., STIEGER, B., CRAMER, R., WATERFIELD, M., BURLINGAME, A. L. & UNWIN, R. J. 2004b. Proteomic analysis of plasma membrane vesicles isolated from the rat renal cortex. *PROTEOMICS*, 5, 101-112.
126. DANESH, J., WHEELER, J. G., HIRSCHFELD, G. M., EDA, S., EIRIKSDOTTIR, G., RUMLEY, A., LOWE, G. D., PEPYS, M. B. & GUDNASON, V. 2004. C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *New England Journal of Medicine*, 350, 1387-1397.
127. DARDÉ, V. M., DE LA CUESTA, F., GIL DONES, F., ALVAREZ-LLAMAS, G., BARDERAS, M. G. & VIVANCO, F. 2010. Analysis of the plasma proteome associated with acute coronary syndrome: does a permanent protein signature exist in the plasma of ACS patients? *Journal of Proteome Research*, 9, 4420-4432.
128. DAUTEL, F., KALKHOF, S., TRUMP, S., LEHMANN, I., BEYER, A. & VON BERGEN, M. 2010. Large-scale 2-D DIGE studies-guidelines to overcome pitfalls and challenges along the experimental procedure. *Journal of Integrated OMICS*, 1, 170-179.
129. DE LA CUESTA, F., ALVAREZ-LLAMAS, G., MAROTO, A. S., DONADO, A., ZUBIRI, I., POSADA, M., PADIAL, L. R., PINTO, A. G., BARDERAS, M. G. & VIVANCO, F. 2011. A Proteomic Focus on the Alterations Occurring at the Human Atherosclerotic Coronary Intima. *Molecular & Cellular Proteomics*, 10.
130. DE ROOS, B. 2008. Proteomic analysis of human plasma and blood cells in nutritional studies: development of biomarkers to aid disease prevention. *Expert Review of Proteomics*, 5, 819-826.
131. DE SOUZA, L. C., LAMARI, F., BELLIARD, S., JARDEL, C., HOUILLIER, C., DE PAZ, R., DUBOIS, B. & SARAZIN, M. 2011. Cerebrospinal fluid biomarkers in the differential diagnosis of Alzheimer's disease from other cortical dementias. *Journal of Neurology, Neurosurgery & Psychiatry*, 82, 240-246.
132. DE WARDENER, H. & MACGREGOR, G. 2002. Sodium and blood pressure. *Current Opinion in Cardiology* 17, 360-367.

133. DECRAMER, S., WITTKER, S., MISCHAK, H., ZÜRBIG, P., WALDEN, M., BOUISSOU, F., BASCANDS, J. L. & SCHANSTRA, J. P. 2006. Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis. *Nature Medicine*, 12, 398-400.
134. DELLES, C., NEISIUS, U. & CARTY, D. M. 2012. Proteomics in hypertension and other cardiovascular diseases. *Annals of Medicine*, 44, S55-S64.
135. DESOUSA, L. V., KRAKOVSKA, O., DARFLER, M. M., KRIZMAN, D. B., ROMASCHIN, A. D., COLGAN, T. J. & SIU, K. W. M. 2010. mTRAQ-based quantification of potential endometrial carcinoma biomarkers from archived formalin-fixed paraffin-embedded tissues. *PROTEOMICS*, 10, 3108-3116.
136. DIAMANDIS, E. 2004a. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool opportunities and potential limitations. *Molecular & Cellular Proteomics*, 3, 367-378.
137. DIAMANDIS, E. P. 2004b. How are we going to discover new cancer biomarkers? A proteomic approach for bladder cancer. *Clin Chem*, 50, 793-795.
138. DIAMANDIS, E. P. & VAN DER MERWE, D.-E. 2005. Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. *Clinical Cancer Research*, 11, 963-965.
139. DIHAZI, H., ASIF, A. R., AGARWAL, N. K., DONCHEVA, Y. & MÜLLER, G. A. 2005. Proteomic analysis of cellular response to osmotic stress in thick ascending limb of Henle's loop (TALH) cells. *Molecular & Cellular Proteomics*, 4, 1445-1458.
140. DIUHY, R. & LIFTON, R. 1995. Glucocorticoid-remediable aldosteronism (GRA): Diagnosis, variability of phenotype and regulation of potassium homeostasis. , 1995, vol. 60, 60, 48-51.
141. DOLE, M., MACK, L. L. AND HINES, R. L. 1968. Molecular Beams of Macroions. . *Journal of Chemical Physics* 49, 2240.
142. DRUMMOND, H., WELSH, M. & ABOUD, F. 2001. ENaC subunits are molecular components of the arterial baroreceptor complex. . *Ann NY Acad Sci* 940, 42-47.
143. EAPEN, D. J., MANOCHA, P., VALIANI, K., MANTINI, N., SPERLING, L. & MCGORISK, G. M. 2011. Alcohol and the Heart: An Ounce of Prevention. *Current treatment options in cardiovascular medicine*, 13, 313-325.
144. EBRAHIMIAN, T. & TOUYZ, R. M. 2008. Thioredoxin in vascular biology: role in hypertension. *Antioxidants & redox signaling*, 10, 1127-1136.
145. ECHAN, L. A., TANG, H. Y., ALIKHAN, N., LEE, K. & SPEICHER, D. W. 2005. Depletion of multiple high_abundance proteins improves protein profiling capacities of human serum and plasma. *PROTEOMICS*, 5, 3292-3303.
146. ELLIOTT, M. H., SMITH, D. S., PARKER, C. E. & BORCHERS, C. 2009. Current trends in quantitative proteomics. *Journal of Mass Spectrometry*, 44, 1637-1660.
147. ELLIOTT, P. & PEAKMAN, T. C. 2008. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *International Journal of Epidemiology*, 37, 234-244.
148. ELLIOTT, W. 2007. Systemic hypertension. . *Current Problems in Cardiology*, 32, 201-259.

149. ENGELEN, K., SIFRIM, A., VAN DE PLAS, B., LAUKENS, K., ARCKENS, L. & MARCHAL, K. 2010. Alternative experimental design with an applied normalization scheme can improve statistical power in 2D-DIGE experiments. *Journal of Proteome Research*, 9, 4919.
150. ESCHELBACH, J. W. & JORGENSON, J. W. 2006. Improved Protein Recovery in Reversed-Phase Liquid Chromatography by the Use of Ultrahigh Pressures. *Analytical Chemistry*, 78, 1697-1706.
151. FENN, J. B. 2002. Electrospray ionization mass spectrometry: How it all began. *J. Biomol Tech*, 13, 13, 101-118.
152. FENSELAU, C. 2007. A review of quantitative methods for proteomic studies. *Journal of Chromatography B*, 855, 14-20.
153. FENTZ, J. S., ZORNIG, C., JUHL, H. H. & DAVID, K. A. 2004. Tissue ischemia time affects gene and protein expression patterns within minutes following surgical tumor excision. *Biotechniques*, 36, 1030-1037.
154. FEY, S. J., LARSEN, P. M., 2001. 2D or not 2D. *Curr. Opin. Chem. Biol.*, 5, 26-33.
155. FIELDS, S. 2001. The interplay of biology and technology. *Proceedings of the National Academy of Sciences*, 98, 10051-10054.
156. FINOULST, I., PINKSE, M., VAN DONGEN, W. & VERHAERT, P. 2011. Sample Preparation Techniques for the Untargeted LC-MS-Based Discovery of Peptides in Complex Biological Matrices. *Journal of Biomedicine and Biotechnology*, 2011.
157. FISCHER, M. & O'HARE, A. 2010. Epidemiology of hypertension in the elderly with chronic kidney disease. *Adv Chronic Kidney Dis*, 17, 329-40.
158. FISHER, R. 1918. The Correlation Between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399-433
159. FISKERSTRAND, T., REFSUM, H., KVALHEIM, G. & UELAND, P. M. 1993. Homocysteine and other thiols in plasma and urine: automated determination and sample stability. *Clin Chem*, 39, 263-271.
160. FITZGIBBONS, P. L. & COOPER, K. 2009. Immunohistochemistry of Biomarkers. *Basic Concepts of Molecular Pathology*, 133-137.
161. FLISER, D., NOVAK, J., THONGBOONKERD, V., ARGILÉS, À., JANKOWSKI, V., GIROLAMI, M. A., JANKOWSKI, J. & MISCHAK, H. 2007. Advances in Urinary Proteome Analysis and Biomarker Discovery. *Journal of the American Society of Nephrology*, 18, 1057-1071.
162. FLOWER, L., AHUJA, R. H., HUMPHRIES, S. E. & MOHAMED-ALI, V. 2000. Effects of sample handling on the stability of interleukin 6, tumour necrosis factor- α and leptin. *Cytokine*, 12, 1712-1716.
163. FODOR, I. & NELSON, D. 2005. Leveraging Genomics Software to Improve Proteomics Results. Lawrence Livermore National Laboratory (LLNL), Livermore, CA.
164. FODOR, I. K., NELSON, D. O., ALEGRIA-HARTMAN, M., ROBBINS, K., LANGLOIS, R. G., TURTELTAUB, K. W., CORZETT, T. H. & MCCUTCHEN-MALONEY, S. L. 2005. Statistical challenges in the analysis of two-dimensional difference gel electrophoresis experiments using DeCyder™. *Bioinformatics*, 21, 3733-3740.
165. FOERCH, C., MONTANER, J., FURIE, K., NING, M. & LO, E. 2009. Invited article: searching for oracles? Blood biomarkers in acute stroke. *Neurology*, 73, 393-399.

166. FOUNTOULAKIS, M., HARDMEIER, R., HOGER, H., LUBEC, G., 2001. Postmortem changes in the level of brain proteins. . *Exp. Neurol.*, 167, 86–94.
167. 167., 86–94.
168. FREDRIKSSON, S., DIXON, W., JI, H., KOONG, A. C., MINDRINOS, M. & DAVIS, R. W. 2007. Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nature methods*, 4, 327-329.
169. FRÖBEL, J., HARTWIG, S., PAßLACK, W., ECKEL, J., HAAS, R., CZIBERE, A. & LEHR, S. 2010. ProteoMiner™ and SELDI-TOF-MS: A robust and highly reproducible combination for biomarker discovery from whole blood serum. *Archives of Physiology and Biochemistry*, 116, 174-180.
170. FRÖHLICH, T. & ARNOLD, G. 2006. Proteome research based on modern liquid chromatography--tandem mass spectrometry: separation, identification and quantification. *J Neural Transm.* , 113, 973-994.
171. FUNG, E. T. 2010. A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. *Clin Chem*, 56, 327-329.
172. GAO, Y. Â. C., YUAN, Z. Â. B., YANG, Y. Â. D. & LU, H. Â. K. 2007. Effect of freeze-thaw cycles on serum measurements of AFP, CEA, CA125 and CA19â€š. *Scandinavian Journal of Clinical & Laboratory Investigation*, 67, 741-747.
173. GARBIS, S., LUBEC, G. & FOUNTOULAKIS, M. 2005. Limitations of current proteomics technologies. . *Journal of Chromatography A* 1077:1-18.
174. GEIGER, T., WISNIEWSKI, J. R., COX, J., ZANIVAN, S., KRUGER, M., ISHIHAMA, Y. & MANN, M. 2011. Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nature protocols*, 6, 147-157.
175. GERBER, S. A., RUSH, J., STEMMAN, O., KIRSCHNER, M. W. & GYGI, S. P. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences*, 100, 6940-6945.
176. GERSZTEN, R. E., ACCURSO, F., BERNARD, G. R., CAPRIOLI, R. M., KLEE, E. W., KLEE, G. G., KULLO, I., LAGUNA, T. A., ROTH, F. P., SABATINE, M., SRINIVAS, P., WANG, T. J., WARE, L. B. & PROGRAMS, F. T. N. C. P. 2008. Challenges in translating plasma proteomics from bench to bedside: update from the NHLBI Clinical Proteomics Programs. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 295, L16-L22.
177. GERSZTEN, R. E., ASNANI, A. & CARR, S. A. 2011. Status and Prospects for Discovery and Verification of New Biomarkers of Cardiovascular Disease by Proteomics. *Circulation Research*, 109, 463-474.
178. GLEN, A., EVANS, C. A., GAN, C. S., CROSS, S. S., HAMDY, F. C., GIBBINS, J., LIPPITT, J., EATON, C. L., NOIREL, J., WRIGHT, P. C. & REHMAN, I. 2010. Eight-plex iTRAQ analysis of variant metastatic human prostate cancer cells identifies candidate biomarkers of progression: An exploratory study. *The Prostate*, 70, 1313-1332.
179. GLERUM, J. M. Z. A. D. M. 2006. Defects in cytochrome oxidase assembly in humans: lessons from yeast. *Biochem. Cell Biol.* , 84, 859–869.
180. GOO, Y. A. & GOODLETT, D. R. 2010. Advances in proteomic prostate cancer biomarker discovery. *Journal of Proteomics*, 73, 1839-1850.

181. GOOD, D., ZURBIG, P., ARGILES, A., BAUER, H. & AL, E. 2010. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. . *Mol. Cell. Proteomics*.
182. GOODWIN, R. J., DUNGWORTH, J. C., COBB, S. R. & PITT, A. R. 2008a. Time-dependent evolution of tissue markers by MALDI-MS imaging. *PROTEOMICS*, 8, 3801-3808.
183. GOODWIN, R. J. A., DUNGWORTH, J. C., COBB, S. R. & PITT, A. R. 2008b. Time-dependent evolution of tissue markers by MALDI-MS imaging. *PROTEOMICS*, 8, 3801-3808.
184. GOODWIN, R. J. A., LANG, A. M., ALLINGHAM, H., BORÉN, M. & PITT, A. R. 2010. Stopping the clock on proteomic degradation by heat treatment at the point of tissue excision. *PROTEOMICS*, 10, 1751-1761.
185. GOODWIN, R. J. A., LANG, ALASTAIR, M., ALLINGHAM, HEATHER, BORÉN, MATS, PITT, ANDREW, R. 2010. Stopping the clock on proteomic degradation by heat treatment at the point of tissue excision. *PROTEOMICS*, 10, 1751-1761.
186. GOODWIN, R. J. A., NILSSON, A., BORG, D., LANGRIDGE-SMITH, P. R. R., HARRISON, D. J., MACKAY, C. L., IVERSON, S. L. & ANDRÉN, P. E. 2012. Conductive carbon tape used for support and mounting of both whole animal and fragile heat-treated tissue sections for MALDI MS imaging and quantitation. *Journal of Proteomics*, 75, 4912-4920.
187. GOODWIN, R. J. A., PENNINGTON, S. R. & PITT, A. R. 2008c. Protein and peptides in pictures: Imaging with MALDI mass spectrometry. *PROTEOMICS*, 8, 3785-3800.
188. GORG, A., WEISS, W. AND DUNN, M. J. 2004. Current two-dimensional electrophoresis technology for proteomics. . *PROTEOMICS* 4, 3665-3685.
189. GOTTSCHALK, M. E., CHATTERJEE, T., EDELSTEIN, I. & MARCUS, F. 1982. Studies on the mechanism of interaction of fructose 2, 6-bisphosphate with fructose-1, 6-bisphosphatase. *Journal of Biological chemistry*, 257, 8016-8020.
190. GOUDSMIT, S. A. 1948. A Time-of-Flight Mass Spectrometer. *Physical Review*, 74, 622-623.
191. GOUW, J. W., KRIJGSVELD, J. & HECK, A. J. R. 2010. Quantitative Proteomics by Metabolic Labeling of Model Organisms. *Molecular & Cellular Proteomics*, 9, 11-24.
192. GOZAL, D., JORTANI, S., SNOW, A., KHEIRANDISH-GOZAL, L., BHATTACHARJEE, R., KIM, J. & CAPDEVILA, O. 2009. Two-Dimensional Differential In-Gel Electrophoresis Proteomic Approaches Reveal Urine Candidate Biomarkers in Pediatric Obstructive Sleep Apnea *American Journal of Respiratory and Critical Care Medicine* 180, 1253-1261.
193. GRACEY, A. Y., TROLL, J. V. & SOMERO, G. N. 2001. Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proceedings of the National Academy of Sciences*, 98, 1993-1998.
194. GRAHAM, D., MCBRIDE, M., GAASENBEEK, M., GILDAY, K., BEATTIE, E., MILLER, W., MCCLURE, J., POLKE, J., MONTEZANO, A., TOUYZ, R. & DOMINICZAK, A. 2007. Candidate Genes That Determine Response to Salt in the Stroke-Prone Spontaneously Hypertensive Rat Congenic Analysis. *Hypertension Journal of the american heart association*, 50, 1134-1141.
195. GRAHAM D, M. M., GAASENBEEK M, GILDAY K, BEATTIE E, MILLER WH ET AL. 2007a. Candidate genes that determine response to salt in the stroke-prone spontaneously hypertensive rat - Congenic analysis. . *Hypertension* 50, 1134-1141.

196. GRAHAM D, M. M., GAASENBEEK M, GILDAY K, BEATTIE E, MILLER WH ET AL. 2007b. Candidate genes that determine response to salt in the stroke-prone spontaneously hypertensive rat - Congenic analysis. . *Hypertension* 50, 1134-1141.
197. GRASSL, J., WESTBROOK, J. A., ROBINSON, A., BORÉN, M., DUNN, M. J. & CLYNE, R. K. 2009. Preserving the yeast proteome from sample degradation. *PROTEOMICS*, 9, 4616-4626.
198. GRAVETT, M. G., NOVY, M. J., ROSENFELD, R. G., REDDY, A. P., JACOB, T., TURNER, M., MCCORMACK, A., LAPIDUS, J. A., HITTI, J. & ESCHENBACH, D. A. 2004. Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers. *JAMA: the journal of the American Medical Association*, 292, 462-469.
199. GRØNBORG, M., KRISTIANSEN, T. Z., IWAHORI, A., CHANG, R., REDDY, R., SATO, N., MOLINA, H., JENSEN, O. N., HRUBAN, R. H. & GOGGINS, M. G. 2006. Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. *Molecular & Cellular Proteomics*, 5, 157-171.
200. GROSSMANN, J., ROSCHITZKI, B., PANSE, C., FORTES, C., BARKOW-OESTERREICHER, S., RUTISHAUSER, D. & SCHLAPBACH, R. 2010. Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *Journal of Proteomics*, 73, 1740-1746.
201. GUERRINI, U., SIRONI, L., TREMOLI, E., CIMINO, M., POLLO, B., CALVIO, A. M., PAOLETTI, R. & ASDENTE, M. 2002. New insights into brain damage in stroke-prone rats a nuclear magnetic imaging study. *Stroke*, 33, 825-830.
202. GUIDI, E., MENGHETTI, D., MILANI, S., MONTAGNINO, G., PALAZZI, P. & BIANCHI, G. 1996. Hypertension may be transplanted with the kidney in humans: A long-term historical prospective follow-up of recipients grafted with kidneys coming from donors with or without hypertension in their families. . *Journal of the American Society of Nephrology* 7, 1131-1138.
203. GUO, T., WANG, W., RUDNICK, P. A., SONG, T., LI, J., ZHUANG, Z., WEIL, R. J., DEVOE, D. L., LEE, C. S. & BALGLEY, B. M. 2007. Proteome analysis of microdissected formalin-fixed and paraffin-embedded tissue specimens. *Journal of Histochemistry & Cytochemistry*, 55, 763-772.
204. GUYTON, A. 1991. Blood Pressure Control Special Role of the Kidneys and Body Fluids. *Science*, 252, 1813-1816.
205. GYGI, S. P., RIST, B., GERBER, S. A., TURECEK, F., GELB, M. H. & AEBERSOLD, R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology*, 17, 994-999.
206. HAMDAN, M. & RIGHETTI, P. G. 2003. Modern strategies for protein quantification in proteome analysis: Advantages and limitations. *Mass Spectrometry Reviews*, 21, 287-302.
207. HAMLIN, J. & BLAUSTEIN, M. 1986. Sodium-Chloride, Extracellular Fluid Volume, and Blood-Pressure Regulation. . *American Journal of Physiology* 251, F563-F575.
208. HAN, C.-L., CHIEN, C.-W., CHEN, W.-C., CHEN, Y.-R., WU, C.-P., LI, H. & CHEN, Y.-J. 2008. A Multiplexed quantitative strategy for membrane proteomics opportunities for mining therapeutic targets for autosomal dominant polycystic kidney disease. *Molecular & Cellular Proteomics*, 7, 1983-1997.
209. HAN, D. K., ENG, J., ZHOU, H. & AEBERSOLD, R. 2001. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nature Biotechnology*, 19, 946-951.
210. HAN, W. K. & BONVENTRE, J. V. 2004. Biologic markers for the early detection of acute kidney injury. *Current Opinion in Critical Care*, 10, 476-482.

211. HANCOCK, W. S., WU, S. L., STANLEY, R. R. & GOMBOCZ, E. A. 2002. Publishing large proteome datasets: scientific policy meets emerging technologies. *Trends in biotechnology*, 20, s39-s44.
212. HARIHARAN, D., WEEKS, M. & CRNOGORAC-JURCEVIC, T. 2010. Application of proteomics in cancer gene profiling: two-dimensional difference in gel electrophoresis (2D-DIGE). *Methods Mol Biol.* , 576, 197-211.
213. HASEGAWA, H., KOHNO, M., SASAKI, M. & AL, E. 2003. "Antagonist of monocyte chemoattractant protein 1 ameliorates the initiation and progression of lupus nephritis and renal vasculitis in MRL/lpr mice," *Arthritis and Rheumatism*, 48, 2555–2566.
214. HAUG, T., SKORPEN, F., KVALØY, K. & AL, E. 1997. Human uracil-DNA glycosylase gene: sequence organization, methylation pattern, and mapping to chromosome 12q23-q24.1. *Genomics* 36, 408-16.
215. HAUG, T., SKORPEN, F., LUND, H. & KROKAN, H. 1994. Structure of the gene for human uracil-DNA glycosylase and analysis of the promoter function. *FEBS Lett.* , 353, 180–4.
216. HELLER, M., MATTOU, H., MENZEL, C. & YAO, X. 2003. Trypsin catalyzed¹⁶O-to-¹⁸O exchange for comparative proteomics: tandem mass spectrometry comparison using MALDI-TOF, ESI-QTOF, and ESI-ion trap mass spectrometers. *Journal of the American Society for Mass Spectrometry*, 14, 704-718.
217. HENNESSEY, E. S., DRUMMOND, D. & SPARROW, J. 1993. Molecular genetics of actin function. *Biochemical Journal*, 291, 657.
218. HERBERT, C., ET AL 2001. Reduction and alkylation of proteins in preparation of two-dimensional map analysis: Why, when, and how? *Electrophoresis.*, 22, 2046-2057.
219. HERNÁNDEZ-FERNAUD, J. R. & SALIDO, E. 2010. Differential expression of liver and kidney proteins in a mouse model for primary hyperoxaluria type I. *FEBS Journal*, 277, 4766-4774.
220. HEYMANS, C. & NEIL, E. 1958. *Reflexogenic Areas of the Cardiovascular System*, Little Brown, Boston, .
221. HEYWOOD, W., MADGETT, T., WANG, D., WALLINGTON, A., HOGG, J., MILLS, K. & AVENT, N. 2011. 2D DIGE analysis of maternal plasma for potential biomarkers of Down Syndrome. *Proteome Science*, 9, 56.
222. HILARIO, M., KALOUSIS, A., PRADOS, J. & BINZ, P.-A. 2004. Data mining for mass-spectra based diagnosis and biomarker discovery. *Drug Discovery Today*, 2, 214-222.
223. HOCHSTRASSER, D. F. 1998. Proteome in perspective. *Clinical Chemistry and Laboratory Medicine* 36, 825-836.
224. HOFFERT, J. D., WANG, G., PISITKUN, T., SHEN, R.-F. & KNEPPER, M. A. 2007. An automated platform for analysis of phosphoproteomic datasets: application to kidney collecting duct phosphoproteins. *Journal of Proteome Research*, 6, 3501-3508.
225. HOFFMAN, E. 2007. *Mass Spectrometry: principle and Applications.* , WILEY.
226. HOLLY, M. K., DEAR, J. W., HU, X., SCHECHTER, A. N., GLADWIN, M. T., HEWITT, S. M., YUEN, P. S. T. & STAR, R. A. 2006. Biomarker and drug-target discovery using proteomics in a new rat model of sepsis-induced acute renal failure. *Kidney International*, 70, 496-506.

227. HOLTEN-ANDERSEN, M., SCHROHL, A., BRÜNNER, N., NIELSEN, H., HØGDALL, C. & HØGDALL, E. 2003. Evaluation of sample handling in relation to levels of tissue inhibitor of metalloproteinases-1 measured in blood by immunoassay. *The International journal of biological markers*, 18, 170.
228. HONDA, K., HAYASHIDA, Y., UMAKI, T., OKUSAKA, T., KOSUGE, T., KIKUCHI, S., ENDO, M., TSUCHIDA, A., AOKI, T. & ITOI, T. 2005. Possible detection of pancreatic cancer by plasma protein profiling. *Cancer Research*, 65, 10613-10622.
229. HONG, M. L., ET AL., 2006. Proteomics technology and therapeutics. *Clinical & Experimental Pharmacology & Physiology*, 33, 563-8.
230. HOORN, E. J., HOFFERT, J. D. & KNEPPER, M. A. 2006. The application of DIGE-based proteomics to renal physiology. *Nephron Physiology*, 104, p61-p72.
231. HOORN, E. J., PISITKUN, T., ZIETSE, R., GROSS, P., FROKIAER, J., WANG, N. S., GONZALES, P. A., STAR, R. A. & KNEPPER, M. A. 2005. Prospects for urinary proteomics: Exosomes as a source of urinary biomarkers (Review Article). *Nephrology*, 10, 283-290.
232. HSICH, G., KENNEY, K., GIBBS, C. J., LEE, K. H. & HARRINGTON, M. G. 1996. The 14-3-3 Brain Protein in Cerebrospinal Fluid as a Marker for Transmissible Spongiform Encephalopathies. *New England Journal of Medicine*, 335, 924-930.
233. HSIEH, S.-Y., CHEN, R.-K., PAN, Y.-H. & LEE, H.-L. 2006. Systematical evaluation of the effects of sample collection procedures on low-molecular-weight serum/plasma proteome profiling. *PROTEOMICS*, 6, 3189-3198.
234. HSU, J.-L., HUANG, S.-Y., CHOW, N.-H. & CHEN, S.-H. 2003. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem*, 75, 6843-6852.
235. HUANG, H.-L., STASYK, T., MORANDELL, S., DIEPLINGER, H., FALKENSAMMER, G., GRIESMACHER, A., MOGG, M., SCHREIBER, M., FEUERSTEIN, I., HUCK, C. W., STECHER, G., BONN, G. K. & HUBER, L. A. 2006a. Biomarker discovery in breast cancer serum using 2-D differential gel electrophoresis/ MALDI-TOF/TOF and data validation by routine clinical assays. *Electrophoresis*, 27, 1641-1650.
236. HUANG, S. Y., TSAI, M. L., WU, C. J., HSU, J. L., HO, S. H. & CHEN, S. H. 2006b. Quantitation of protein phosphorylation in pregnant rat uteri using stable isotope dimethyl labeling coupled with IMAC. *PROTEOMICS*, 6, 1722-1734.
237. HUBER, L. A., PFALLER, K. & VIETOR, I. 2003. Organelle Proteomics Implications for Subcellular Fractionation in Proteomics. *Circulation Research*, 92, 962-968.
238. HUCK, C. W., R. BAKRY, AND G.K. BONN,X 2006. Progress in capillary electrophoresis of biomarkers and metabolites between 2002 and 2005. *Electrophoresis*, 27, 111-125.
239. HUMPEL, C. 2011. Identifying and validating biomarkers for Alzheimer's disease. *Trends in biotechnology*, 29, 26-32.
240. HUZAREWICH, R. L., SIEMENS, C. G. & BOOTH, S. A. 2010. Application of "omics" to prion biomarker discovery. *Journal of Biomedicine and Biotechnology*, 2010.
241. HWANG, S., THUMAR, J., LUNDGREN, D., REZAUL, K., MAYYA, V., WU, L., ENG, J., WRIGHT, M. & HAN, D. 2006. Direct cancer tissue proteomics: a method to identify candidate cancer biomarkers from formalin-fixed paraffin-embedded archival tissues. *Oncogene*, 26, 65-76.

242. HWANG, Y., SATO, S., TSAI, J., YAN, S., BAKR, S., ZHANG, H., OATES, P. & RAMASAMY, R. 2002. Aldose reductase activation is a key component of myocardial response to ischemia. *FASEB J.*, 16, 243-5.
243. HYNDMAN, D., BAUMAN, D. R., HEREDIA, V. V. & PENNING, T. M. 2003a. The aldo-keto reductase superfamily homepage. *Chemico-Biological Interactions*, 143, 621-631.
244. HYNDMAN, D., BAUMAN, D. R., HEREDIA, V. V. & PENNING, T. M. 2003b. The aldo-keto reductase superfamily homepage. *Chemico-Biological Interactions*, 143-144, 621-631.
245. ICHIMURA, T., ASSELDONK, E., HUMPHREYS, B., GUNARATNAM, L., DUFFIELD, J. & BONVENTRE, J. 2008. Kidney injury molecule-1 is a phosphatidylserine receptor that confers a phagocytic phenotype on epithelial cells. *J Clin Invest* 118, 1657-68.
246. ILYIN, S. E., BELKOWSKI, S. M. & PLATA-SALAMÁN, C. R. 2004. Biomarker discovery and validation: technologies and integrative approaches. *Trends in biotechnology*, 22, 411-416.
247. INOUE, H., FUJITA, T., KITAMURA, T., SHIMOSAWA, T., NAGASAWA, R., INOUE, R., MARUYAMA, N. & NAGASAWA, T. 1999. Senescence marker protein-30 (SMP30) enhances the calcium efflux from renal tubular epithelial cells. *Clinical and Experimental Nephrology*, 3, 261-267.
248. IRIBARNE, J. & THOMSON, B. 1976. On the evaporation of small ions from charged droplets *Journal of Chemical Physics* 64 2287-2294.
249. ISABEL PADRÃO, A., FERREIRA, R., VITORINO, R. & AMADO, F. 2012. Proteome-base biomarkers in diabetes mellitus: Progress on biofluids' protein profiling using mass spectrometry. *PROTEOMICS – Clinical Applications*, 6, 447-466.
250. ISHIHAMA, Y., ODA, Y., TABATA, T., SATO, T., NAGASU, T., RAPPSILBER, J. & MANN, M. 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics*, 4, 1265-1272.
251. ISHIHARA, T., FUKUDA, I., MORITA, A., TAKINAMI, Y., OKAMOTO, H., NISHIMURA, S. I. & NUMATA, Y. 2011. Development of quantitative plasma N-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma. *Journal of Proteomics*, 74, 2159-2168.
252. ISSAQ, H. J., FOX, S. D., CHAN, K. C. & VEENSTRA, T. D. 2011. Global proteomics and metabolomics in cancer biomarker discovery. *Journal of Separation Science*, 34, 3484-3492.
253. ISSAQ, H. J., VEENSTRA, T. D., CONRADS, T. P. & FELSCHOW, D. 2002. The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochemical and biophysical research communications*, 292, 587-592.
254. IWANAGA, Y. & MIYAZAKI, S. 2010. Heart Failure, Chronic Kidney Disease, and Biomarkers
255. —; An Integrated Viewpoint —. *Circulation Journal*, 74, 1274-1282.
256. JACKSON, D., CRAVEN, R., HUTSON, R., GRAZE, I., LUETH, P., TONGE, R., HARTLEY, J., NICKSON, J., RAYNER, S., JOHNSTON, C., DIEPLINGER, B., HUBALEK, M., WILKINSON, N., PERREN, T., KEHOE, S., HALL, G., DAXENBICHLER, G., DIEPLINGER, H., SELBY, P. & BANKS, R. 2007. Proteomic profiling identifies afamin as a potential biomarker for ovarian cancer. *Clin Cancer Res* 13, 7370-7379.

257. JACOBS, J. M., ADKINS, J. N., QIAN, W.-J., LIU, T., SHEN, Y., CAMP, D. G. & SMITH, R. D. 2005. Utilizing Human Blood Plasma for Proteomic Biomarker Discovery†. *Journal of Proteome Research*, 4, 1073-1085.
258. JAIN, K. 2004. Role of pharmacoproteomics in the development of personalized medicine. *Pharmacogenomics*, 5, 331-336.
259. JAN ERIKSSON, D. F. 2007. Improving the success rate of proteome analysis by modeling protein-abundance distributions and experimental designs. *Nature Biotechnology*, 25, 651-655.
260. JANECH, M. G., RAYMOND, J. R. & ARTHUR, J. M. 2007. Proteomics in renal research. *American Journal of Physiology - Renal Physiology*, 292, F501-F512.
261. JANINI, G. M. & VEENSTRA, T. D. 2002. Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis*, 23, 3048-3061.
262. JANSSEN, B. & SMITS, J. 2002. Autonomic control of blood pressure in mice: basic physiology and effects of genetic modification. *Am J Physiol Regulatory Integrative Comp Physiol* 282, R1545–R1564.
263. JANTOS-SIWY, J., SCHIFFER, E., BRAND, K., SCHUMANN, G., ROSSING, K., DELLES, C., MISCHAK, H. & METZGER, J. 2008. Quantitative Urinary Proteome Analysis for Biomarker Evaluation in Chronic Kidney Disease. *Journal of Proteome Research*, 8, 268-281.
264. JEFFS, B., NEGRIN, C., GRAHAM, D., CLARK, J., ANDERSON, N., GAUGUIER, D. & AL, E. 2000. Applicability of a "speed" congenic strategy to dissect blood pressure quantitative trait loci on rat chromosome 2. *Hypertension* 35, 79-187.
265. JIANG, L., HE, L. & FOUNTOULAKIS, M. 2004a. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *Journal of Chromatography A*, 1023, 317-320.
266. JIANG, X.-S., ZHOU, H., ZHANG, L., SHENG, Q.-H., LI S.-J., LI L., HAO, P., LI Y.-X., XIA, Q.-C. & WU, J.-R. 2004b. A high-throughput approach for subcellular proteome identification of rat liver proteins using subcellular fractionation coupled with two-dimensional liquid chromatography tandem mass spectrometry and bioinformatic analysis. *Molecular & Cellular Proteomics*, 3, 441-455.
267. JOHNS, E. J., O'SHAUGHNESSY, B., O'NEILL, S., LANE, B. & HEALY, V. 2010. Impact of elevated dietary sodium intake on NAD(P)H oxidase and SOD in the cortex and medulla of the rat kidney. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 299, R234-R240.
268. JOHNSON, R. S., MARTIN, S. A., BIEMANN, K., STULTS, J. T. AND WATSON, J. T. 1987. Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass-Spectrometer - Differentiation of Leucine and Isoleucine. . *Analytical Chemistry* 59, 2621-2625.
269. JOLLIFFE, I. 1986. *Jolliffe I.T. Principal Component Analysis*, NY, Springer.
270. JONES MB, K. H., SHU HJ, ZHAO YM, LIOTTA LA, ET AL. 2002. Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. *Proteomics*, 2, 76-84.
271. JONES, M. B., KRUTZSCH, H., SHU, H., ZHAO, Y., LIOTTA, L. A., KOHN, E. C. & PETRICOIN, E. F. 2002. Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. *PROTEOMICS*, 2, 76-84.

272. JOOS, T. & BACHMANN, J. 2009. Protein microarrays: potentials and limitations. *Frontiers in bioscience: a journal and virtual library*, 14, 4376.
273. JUNGBAUER, A. & HAHN, R. 2009. Ion-exchange chromatography. *Methods Enzymol.*, 463, 349-71.
274. KAISER, T., HERMANN, A., KIELSTEIN, J., WITTKE, S., BARTEL, S., KREBS, R., HAUSADEL, F., HILLMANN, M., GOLOVKO, I., KOESTER, P., HALLER, H., WEISSINGER, E., FLISER, D. & MISCHAK, H. 2003. Capillary electrophoresis coupled to mass spectrometry to establish polypeptide patterns in dialysis fluids. *J Chromatogr A* 1013, 157-171.
275. KALETAŞ, B. K., VAN DER WIEL, I. M., STAUBER, J., LENNARD, J. D., GÜZEL, C., KROS, J. M., LUIDER, T. M. & HEEREN, R. M. A. 2009. Sample preparation issues for tissue imaging by imaging MS. *PROTEOMICS*, 9, 2622-2633.
277. KANG, U. B., AHN, Y., LEE, J. W., KIM, Y. H., KIM, J., YU, M. H., NOH, D. Y. & LEE, C. 2010a. Differential profiling of breast cancer plasma proteome by isotope-coded affinity tagging method reveals biotinidase as a breast cancer biomarker. *BMC cancer*, 10, 114.
278. KANG, X., SUN, L., GUO, K., SHU, H., YAO, J., QIN, X. & LIU, Y. 2010b. Serum protein biomarkers screening in HCC patients with liver cirrhosis by ICAT-LC-MS/MS. *Journal of cancer research and clinical oncology*, 136, 1151-1159.
279. KANNEL, W. 1996. Blood pressure as a cardiovascular risk factor - Prevention and treatment. *Jama-Journal of the American Medical Association* 275, 1571-1576.
280. KANNEL, W. 2000. Risk stratification in hypertension: New insights from the Framingham Study. *American Journal of Hypertension* 13, 3S-10S.
281. KARP, N. A. & LILLEY, K. S. 2005. Maximising sensitivity for detecting changes in protein expression: Experimental design using minimal CyDyes. *PROTEOMICS*, 5, 3105-3115.
282. KARP, N. A. & LILLEY, K. S. 2009. Investigating sample pooling strategies for DIGE experiments to address biological variability. *PROTEOMICS*, 9, 388-397.
283. KARP, N. A., MCCORMICK, P. S., RUSSELL, M. R. & LILLEY, K. S. 2007. Experimental and Statistical Considerations to Avoid False Conclusions in Proteomics Studies Using Differential In-gel Electrophoresis. *Molecular & Cellular Proteomics*, 6, 1354-1364.
284. KASPER, S., SHEPPARD, P., YAN, Y., PETTIGREW, N., BOROWSKY, A., PRINS, G., DODD, J., DUCKWORTH, M. & MATUSIK, R. 1998. Development, progression, and androgen-dependence of prostate tumors in probasin large T antigen transgenic mice: a model for prostate cancer. *Lab Invest*, 78, 319-333.
285. KAVALLARIS, M. A. G. M. M. 2005. Proteomics and disease: opportunities and challenges. *Medical Journal of Australia*, 182, 575-9.
286. KAWAMURA, T., NOMURA, M., TOJO, H., FUJII, K., HAMASAKI, H., MIKAMI, S., BANDO, Y., KATO, H. & NISHIMURA, T. 2010. Proteomic analysis of laser-microdissected paraffin-embedded tissues:(1) Stage-related protein candidates upon non-metastatic lung adenocarcinoma. *Journal of Proteomics*, 73, 1089-1099.

287. KAWANO, M., THET, M., MAKINO, T., KUSHIDA, T. & SAKAGAMI, H. 2010. DNA microarray analysis of signaling pathway in macrophages stimulated by lignin-carbohydrate complex from *lentinus edodes* mycelia (LEM) extract. *Anticancer Res.*, 30, 2567-76.
288. KEARNEY, P., WHELTON, M., REYNOLDS, K., MUNTNER, P., WHELTON, P. & HE, J. 2005. Global burden of hypertension: analysis of worldwide data. *Lancet* 365, 217-223.
289. KEBARLE, P. V., UH. 2009. Electrospray: From ions in solution to ions in the gas phase, what we know now *Mass Spectrometry Reviews*, 28 898-917
290. KEEPING, A. J. & COLLINS, R. A. 2010. Data Variance and Statistical Significance in 2D-Gel Electrophoresis and DIGE Experiments: Comparison of the Effects of Normalization Methods. *Journal of Proteome Research*, 10, 1353-1360.
291. KENNEDY, S. A., SCAIFE, C., DUNN, M. J., WOOD, A. E. & WATSON, R. W. G. 2011. Benefits of heat treatment to the protease packed neutrophil for proteome analysis: Halting protein degradation. *PROTEOMICS*, 11, 2560-2564.
292. KENNY, L. C., DUNN, W. B., ELLIS, D. I., MYERS, J., BAKER, P. N. & KELL, D. B. 2005. Novel biomarkers for pre-eclampsia detected using metabolomics and machine learning. *Metabolomics*, 1, 227-234.
293. KENTISIS, A. 2011. Challenges and opportunities for discovery of disease biomarkers using urine proteomics. *Pediatrics International*, 53, 1-6.
294. KENTISIS, A., MONIGATTI, F., DORFF, K., CAMPAGNE, F., BACHUR, R. & STEEN, H. 2009. Urine proteomics for profiling of human disease using high accuracy mass spectrometry. *PROTEOMICS-Clinical Applications*, 3, 1052-1061.
295. KERR, S., BROSINAN, M., MCINTYRE, M., REID, J., DOMINICZAK, A. & HAMILTON, C. 1999. Superoxide anion production is increased in a model of genetic hypertension - Role of the endothelium. *Hypertension* 33, 1353-1358.
296. KESHISHIAN, H., ADDONA, T., BURGESS, M., KUHN, E. & CARR, S. A. 2007. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Molecular & Cellular Proteomics*, 6, 2212-2229.
297. KIANI, A., JOHNSON, K., CHEN, C. & AL, E. 2009. Urine osteoprotegerin and monocyte chemoattractant protein-1 in lupus nephritis. *Journal of Rheumatology*, 36, 2224-2230.
298. KIKUCHI, T., CARBONE, D. P., 2007. Proteomics analysis in lung cancer: challenges and opportunities. *Respirology* 12, 22-28.
299. KIKUTA, K., GOTOH, M., KANDA, T., TOCHIGI, N., SHIMODA, T., HASEGAWA, T., KATAI, H., SHIMADA, Y., SUEHARA, Y. & KAWAI, A. 2010. Pfetin as a prognostic biomarker in gastrointestinal stromal tumor: novel monoclonal antibody and external validation study in multiple clinical facilities. *Japanese journal of clinical oncology*, 40, 60-72.
300. KIM, G., ECELBARGER, C., MITCHELL, C., PACKER, R., WADE, J. & KNEPPER, M. A. 1999. Vasopressin increases Na-K-2Cl cotransporter expression in thick ascending limb of Henle's loop. *Am J Physiol Renal Physiol*, 276, F96-F103.
301. KIM, H., GOLUB, G. H. & PARK, H. 2005. Missing value estimation for DNA microarray gene expression data: local least squares imputation. *Bioinformatics*, 21, 187-198.

302. KIM, K., ARONOV, P., ZAKHARKIN, S. O., ANDERSON, D., PERROUD, B., THOMPSON, I. M. & WEISS, R. H. 2009. Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Molecular & Cellular Proteomics*, 8, 558-570.
303. KING, K. M. & RUBIN, G. 2003. A history of diabetes: from antiquity to discovering insulin. *British journal of nursing*, 12, 1091-1095.
304. KIRTLEY, M. E. & MCKAY, M. 1977. Fructose-1, 6-bisphosphate, a regulator of metabolism. *Molecular and cellular biochemistry*, 18, 141-149.
305. KISTLER, A. D., MISCHAK, H., POSTER, D., DAKNA, M., WÜTHRICH, R. P. & SERRA, A. L. 2009. Identification of a unique urinary biomarker profile in patients with autosomal dominant polycystic kidney disease. *Kidney International*, 76, 89-96.
306. KLOSE, J. 1999. Fractionated extraction of total tissue proteins from mouse and human for 2-D electrophoresis. *METHODS IN MOLECULAR BIOLOGY-CLIFTON THEN TOTOWA-*, 112, 67-86.
307. KNEPPER, M. A. 1994. The aquaporin family of molecular water channels. *Proc. Natd. Acad. Sci.*, 91, 6255-6258.
308. KNEPPER, M. A. 1997. Molecular physiology of urinary concentrating mechanism: regulation of aquaporin water channels by vasopressin. *Am J Physiol Renal Physiol*, 272, F3-F12.
309. KNEPPER, M. A. 2002. Proteomics and the Kidney. *Journal of the American Society of Nephrology*, 13, 1398-1408.
310. KNEPPER, M. A. & BROOKS, H. L. 2001. Regulation of the sodium transporters NHE3, NKCC2 and NCC in the kidney. *Current Opinion in Nephrology and Hypertension*, 10, 655-659.
311. KNEPPER, M. A. & MASILAMANI, S. 2001. Targeted proteomics in the kidney using ensembles of antibodies. *Acta Physiologica Scandinavica*, 173, 11-21.
312. KOENIG, W., TWARDELLA, D., BRENNER, H. & ROTHENBACHER, D. 2005. Plasma concentrations of cystatin C in patients with coronary heart disease and risk for secondary cardiovascular events: more than simply a marker of glomerular filtration rate. *Clin Chem*, 51, 321-327.
313. KOGA, Y., HIROOKA, Y., ARAKI, S., NOZOE, M., KISHI, T. & SUNAGAWA, K. 2008. High Salt Intake Enhances Blood Pressure Increase during Development of Hypertension via Oxidative Stress in Rostral Ventrolateral Medulla of Spontaneously Hypertensive Rats. *Hypertension Research*, 31, 2075-2083.
314. KOKKAT, T. J., MCGARVEY, D., LOVECCHIO, L. C. & LIVOLSI, V. A. 2011. Effect of Thaw Temperatures in Reducing Enzyme Activity in Human Thyroid Tissues. *Biopreservation and Biobanking*, 9, 349-354.
315. KOLCH, W., H. MISCHAK, AND A.R. PITT, 2005. The molecular make-up of a tumour: proteomics in cancer research. *Clinical Science*, 108, 369-83.
316. KOMMU, S., SHARIFI, R., EDWARDS, S. & EELES, R. 2004. Proteomics and urine analysis: a potential promising new tool in urology. *BJU international*, 93, 1172-1173.
317. KONDO, T. 2008a. Tissue proteomics for cancer biomarker development - Laser microdissection and 2D-DIGE -. *BMB Reports*.
318. KONDO, T. 2008b. Tissue proteomics for cancer biomarker development: laser microdissection and 2D-DIGE. *Bmb Rep*, 41, 626-34.

319. KONDO, T., SEIKE, M., MORI, Y., FUJII, K., YAMADA, T. & HIROHASHI, S. 2003. Application of sensitive fluorescent dyes in linkage of laser microdissection and two-dimensional gel electrophoresis as a cancer proteomic study tool. *PROTEOMICS*, 3, 1758-1766.
320. KOOMEN, J. M., LI, D., XIAO, L.-C., LIU, T. C., COOMBES, K. R., ABBRUZZESE, J. & KOBAYASHI, R. 2005. Direct Tandem Mass Spectrometry Reveals Limitations in Protein Profiling Experiments for Plasma Biomarker Discovery. *Journal of Proteome Research*, 4, 972-981.
321. KOPF, D., WALDHERR, R. & RETTIG, R. 1993. Source of kidney determines blood pressure in young renal transplanted rats. *Am J Physiol Renal Physiol* 265 F104-F111.
322. KOSUGI, T., NAKAGAWA, T., KAMATH, D. & JOHNSON, R. 2009. Uric acid and hypertension: an age-related relationship? *J Hum Hypertens* 23 75-6.
323. KOYNER, J., BENNETT, M., WORCESTER, E., MA, Q., RAMAN, J., JEEVANANDAM, V., KASZA, K., O'CONNOR, M., KONCZAL, D., TREVINO, S., DEVARAJAN, P. & MURRAY, P. 2008. Urinary cystatin C as an early biomarker of acute kidney injury following adult cardiothoracic surgery. *Kidney International* 74, 1059-1069.
324. KREIL, D. P., KARP, N. A. & LILLEY, K. S. 2004. DNA microarray normalization methods can remove bias from differential protein expression analysis of 2D difference gel electrophoresis results. *Bioinformatics*, 20, 2026-2034.
325. KUBOTA, D., ORITA, H., YOSHIDA, A., GOTOH, M., KANDA, T., TSUDA, H., HASEGAWA, T., KATAI, H., SHIMADA, Y. & KANEKO, K. 2011. Pfetin as a prognostic biomarker for gastrointestinal stromal tumor: validation study in multiple clinical facilities. *Japanese journal of clinical oncology*, 41, 1194-1202.
326. KUHR, W. G. & MONNIG, C. A. 1992. Fundamental Reviews: Capillary electrophoresis. *Anal. Chem.*, 64, 389-407
327. KUKLINSKA, A. M., MROCKO, B., MUSIAL, W. J., USOWICZ-SZARYNSKA, M., SAWICKI, R., BOROWSKA, H., KNAPP, M. & SZMITKOWSKI, M. 2009. Diagnostic Biomarkers of Essential Arterial Hypertension
328. The Value of Prostacyclin, Nitric Oxide, Oxidized-LDL, and Peroxide Measurements. *International Heart Journal*, 50, 341-351.
329. KULASINGAM, V. & DIAMANDIS, E. P. 2008. Tissue culture-based breast cancer biomarker discovery platform. *International Journal of Cancer*, 123, 2007-2012.
330. KULTIMA, K., NILSSON, A., SCHOLZ, B., ROSSBACH, U. L., FÄLTH, M. & ANDRÉN, P. E. 2009. Development and Evaluation of Normalization Methods for Label-free Relative Quantification of Endogenous Peptides. *Molecular & Cellular Proteomics*, 8, 2285-2295.
331. KULTIMA, K., SKÖLD, K. & BORÉN, M. 2011. Biomarkers of disease and post-mortem changes — Heat stabilization, a necessary tool for measurement of protein regulation. *Journal of Proteomics*, 75, 145-159.
332. KUNTUMALLA, S., BRAISTED, J. C., HUANG, S.-T., PARMAR, P. P., CLARK, D. J., ALAMI, H., ZHANG, Q., DONOHUE-ROLFE, A., TZIPORI, S. & FLEISCHMANN, R. D. 2009. Comparison of two label-free global quantitation methods, APEX and 2D gel electrophoresis, applied to the *Shigella dysenteriae* proteome. *Proteome Science*, 7, 22.
333. KURODA, M., URANO, M., ABE, M., MIZOGUCHI, Y., HORIBE, Y., MURAKAMI, M., TASHIRO, K. & KASAHARA, M. 2000. Primary primitive neuroectodermal tumor of the kidney. *Pathology International*, 50, 967-972.

334. KUZYSK, M., SMITH, D., YANG, J. & AL., E. 2009. MRM-based, Multiplexed, Absolute Quantitation of 45 proteins in human plasma. *Molecular and Cellular Proteomics*, 8, 1860.
335. KWAPISZEWSKA, G., MEYER, M., BOGUMIL, R., BOHLE, R. M., SEEGER, W., WEISSMANN, N. & FINK, L. 2004. Identification of proteins in laser-microdissected small cell numbers by SELDI-TOF and Tandem MS. *BMC biotechnology*, 4, 30.
336. KYROU, I., CHROUSOS, G. & TSIGOS, C. 2006. Stress, visceral obesity, and metabolic complications. *Annals of the New York Academy of Sciences* 1083, 77–110.
337. LANGE, V., PICOTTI, P., DOMON, B. & AEBERSOLD, R. 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular systems biology*, 4.
338. LARAGH, J. & AL, E. 1972. Renin, Angiotensin and Aldosterone System in Pathogenesis and Management of Hypertensive Vascular Disease. *The american journal of med.*, 52.
339. LEE, J., O'KEEFE, J., BELL, D., HENSRUD, D. & HOLICK, M. 2008. Vitamin D deficiency an important, common, and easily treatable cardiovascular risk factor? . *J. Am. Coll. Cardiol.* , 52, 1949–56.
340. LEE, Y. H., HAN, H., CHANG, S.-B. & LEE, S.-W. 2004. Isotope-coded N-terminal sulfonation of peptides allows quantitative proteomic analysis with increased de novo peptide sequencing capability. *Rapid Communications in Mass Spectrometry*, 18, 3019-3027.
341. LEI, R., WU, C., YANG, B., MA, H., SHI, C., WANG, Q., WANG, Q., YUAN, Y. & LIAO, M. 2008. Integrated metabolomic analysis of the nano-sized copper particle-induced hepatotoxicity and nephrotoxicity in rats: A rapid *in vivo* screening method for nanotoxicity. *Toxicology and applied pharmacology*, 232, 292-301.
342. LEMAIRE, R., AIT MENGUELLET, S., STAUBER, J., MARCHAUDON, V. ET AL., 2007. Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, reg alpha fragment, is a new potential ovary cancer biomarker. . *J. Proteome Res.*, 6, 4127–4134.
343. LEMAIRE, R., DESMONS, A., TABEL, J., DAY, R., SALZET, M. & FOURNIER, I. 2007. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *Journal of Proteome Research*, 6, 1295-1305.
344. LESCUYER, P., ALLARD, L., ZIMMERMANN-IVOL, C. G., BURGESS, J. A., HUGHES-FRUTIGER, S., BURKHARD, P. R., SANCHEZ, J. C. & HOCHSTRASSER, D. F. 2004. Identification of post-mortem cerebrospinal fluid proteins as potential biomarkers of ischemia and neurodegeneration. *Proteomics*, 4, 2234-41.
345. LESCUYER, P., HOCHSTRASSER, D. & RABILLOUD, T. 2007. How Shall We Use the Proteomics Toolbox for Biomarker Discovery? *Journal of Proteome Research*, 6, 3371-3376.
346. LI, C., HONG, Y., TAN, Y.-X., ZHOU, H., AI, J.-H., LI, S.-J., ZHANG, L., XIA, Q.-C., WU, J.-R. & WANG, H.-Y. 2004. Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. *Molecular & Cellular Proteomics*, 3, 399-409.

347. LI, F., CHEN, D., HE, C., ZHOU, Y., OLKKONEN, V. M., WAN, P., CHEN, S., ZHU, Y., LAN, K. & TAN, W. 2012a. Identification of urinary Gc-globulin as a novel biomarker for bladder cancer by two-dimensional fluorescent differential gel electrophoresis (2D-DIGE). *Journal of Proteomics*.
348. LI, J., LERICHE, T., TREMBLAY, T.-L., WANG, C., BONNEIL, E., HARRISON, D. J. & THIBAUT, P. 2002a. Application of Microfluidic Devices to Proteomics Research Identification of Trace-level Protein Digests and Affinity Capture of Target Peptides. *Molecular & Cellular Proteomics*, 1, 157-168.
349. LI, J., ZHANG, Z., ROSENZWEIG, J., WANG, Y. & CHAN, D. 2002b. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem*, 48, 1296-1304.
350. LI, K. K.-W., PANG, J. C.-S., NG, H.-K., MASSIMINO, M., GANDOLA, L., BIASSONI, V., SPREAFICO, F., SCHIAVELLO, E., POGGI, G. & CASANOVA, M. 2012b. MB-01. Involvements of hsa-miR-383 and its target peroxiredoxin 3 (PRDX3) in controls of medulloblastoma cell growth. *Neuro-Oncology*, 14, i82-i105.
351. LIFTON, R. & DLUHY, R. 1993. The Molecular Basis of a Hereditary Form of Hypertension, Glucocorticoid-Remediable Aldosteronism. *Trends Endocrinol Metab* 4, 57-61.
352. LIFTON, R., GHARAVI, A. & GELLER, D. 2001. Molecular Mechanisms of Human Hypertension. *Cell*, 104, 545-556.
353. LIFTON, R., GHARAVI, A. & GELLER, D. 2001. Molecular mechanisms of human hypertension. *Cell* 104, 545-556.
354. LIJNEN, P. J., PICCART, Y., COENEN, T. & PRIHADI, J. S. 2012. Angiotensin II-induced mitochondrial reactive oxygen species and peroxiredoxin-3 expression in cardiac fibroblasts. *Journal of hypertension*, 30, 1986-1991
10.1097/HJH.0b013e32835726c1.
355. LILLEY, K. S. & FRIEDMAN, D. B. 2004. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Review of Proteomics*, 1, 401-409.
356. LING, G., CAO, W., ONODERA, M., JU, K., KURIHARA, H., KURIHARA, Y., YAZAKI, Y., KUMADA, M., FUKUDA, Y. & KUWAKI, T. 1998. Renal sympathetic nerve activity in mice: comparison between mice and rats and between normal and endothelin-1 deficient mice. *Brain Res* 808, 238-249.
357. LISTGARTEN, J. & EMILI, A. 2005. Practical proteomic biomarker discovery: taking a step back to leap forward. *Drug Discovery Today*, 10, 1697-1702.
358. LIU, H. B., LIN, D. Y. AND YATES, J. R. 2002. Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques* 32.
359. LIU, P. & HWANG, J. G. 2007. Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics*, 23, 739-746.
360. LOGIN, G. & DVORAK, A. 1988. Microwave fixation provides excellent preservation of tissue, cells and antigens for light and electron microscopy. *Histochem J.*, 20, 373-87.
361. LOHMEIER, T., HILDEBRANDT, D., WARREN, S., MAY, P. & CUNNINGHAM, J. 2005. Recent insights into the interactions between the baroreflex and the kidneys in hypertension. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology* 288, R828-R836.

362. LOO, J., YAN, W., RAMACHANDRAN, P. & WONG, D. 2010. Comparative human salivary and plasma proteomes. *Journal of dental research*, 89, 1016-1023.
363. LU, P., VOGEL, C., WANG, R., YAO, X. & MARCOTTE, E. M. 2006. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology*, 25, 117-124.
364. LULL, M. E., FREEMAN, W. M., VANGUILDER, H. D. & VRANA, K. E. 2010. The use of neuroproteomics in drug abuse research. *Drug and Alcohol Dependence*, 107, 11-22.
365. LUNDBY, A., LAGE, K., WEINERT, BRIAN T., BEKKER-JENSEN, DORTE B., SECHER, A., SKOVGAARD, T., KELSTRUP, CHRISTIAN D., DMYTRIYEV, A., CHOUDHARY, C., LUNDBY, C. & OLSEN, JESPER V. 2012a. Proteomic Analysis of Lysine Acetylation Sites in Rat Tissues Reveals Organ Specificity and Subcellular Patterns. *Cell Reports*, 2, 419-431.
366. LUNDBY, A., SECHER, A., LAGE, K., NORDSBORG, N. B., DMYTRIYEV, A., LUNDBY, C. & OLSEN, J. V. 2012b. Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nature Communications*, 3, 876.
367. LUNDGREN, D., HWANG, S., WU, L. & HAN, D. 2010. Role of spectral counting in quantitative proteomics. *Expert Rev Proteomics*, 7, 39-53.
368. MÁJEK, P., REICHELTOVÁ, Z., SUTNAR, J., MALÝ, M., ORAVEC, M., PEČÁNKOVÁ, K. & DYR, J. E. 2011. Plasma proteome changes in cardiovascular disease patients: novel isoforms of apolipoprotein A1. *Journal of Translational Medicine*, 9, 84.
369. MAKAWITA, S. & DIAMANDIS, E. P. 2010. The Bottleneck in the Cancer Biomarker Pipeline and Protein Quantification through Mass Spectrometry–Based Approaches: Current Strategies for Candidate Verification. *Clin Chem*, 56, 212-222.
370. MAMYRIN, B. A., KARATAEV, V. I., SHMIKK, D. V. AND ZAGULIN, V. A. 1973. Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer. . *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki* 64, v.
371. MANCIA, G., DE BACKER, G., DOMINICZAK, A., ET AL 2007. 2007 ESH-ESC Practice Guidelines for the Management of Arterial hypertension: ESH-ESC Task Force on the Management of arterial hypertension. *Journal of hypertension*, 25, 1751-1762.
372. MARCUS, F., EDELSTEIN, I., REARDON, I. & HEINRIKSON, R. L. 1982. Complete amino acid sequence of pig kidney fructose-1, 6-bisphosphatase. *Proceedings of the National Academy of Sciences*, 79, 7161-7165.
373. MAROUGA, R., DAVID, S. & HAWKINS, E. 2005. The development of the DIGE system: 2D fluorescence difference gel analysis technology. . *Analytical and Bioanalytical Chemistry*, 382, 669-78.
374. MARPLES, D., KNEPPER, M. A., ERIK, C. & NIELSEN, S. 1995. Redistribution of aquaporin-2 water channels induced by vasopressin in rat kidney inner medullary collecting duct. *Am J Physiol Cell Physiol*, 269, C655-C664.
375. MÅRTENSSON, J., MARTLING, C.-R. & BELL, M. 2012. Novel biomarkers of acute kidney injury and failure: clinical applicability. *British Journal of Anaesthesia*, 109, 843-850.
376. MARTIN, A. J. A. S., R. L. 1941. A new form of chromatogram employing two liquid phases: A theory of chromatography. 2. Application to the microdetermination of the higher monoamino-acids in proteins. *Biochem. J.* , 35, 1358-1368.

377. MARTYNIUK, C. J., ALVAREZ, S. & DENSLow, N. D. 2011. DIGE and iTRAQ as biomarker discovery tools in aquatic toxicology. *Ecotoxicology and Environmental Safety*.
378. MASILAMANI, S., KIM, G., MITCHELL, C., WADE, J. & KNEPPER, M. A. 1999. Aldosterone-mediated regulation of ENaC a, b, and g subunit proteins in rat kidney. *The Journal of Clinical Investigation*, 104, R19-R23.
379. MATSUSHIMA, S., IDE, T., YAMATO, M., MATSUSAKA, H., HATTORI, F., IKEUCHI, M., KUBOTA, T., SUNAGAWA, K., HASEGAWA, Y., KURIHARA, T., OIKAWA, S., KINUGAWA, S. & TSUTSUI, H. 2006. Overexpression of Mitochondrial Peroxiredoxin-3 Prevents Left Ventricular Remodeling and Failure After Myocardial Infarction in Mice. *Circulation*, 113, 1779-1786.
380. MATT, P., FU, Z., FU, Q. & VAN EYK, J. E. 2008. Biomarker discovery: proteome fractionation and separation in biological samples. *Physiological Genomics*, 33, 12-17.
381. MATTSSON, N., ZETTERBERG, H., HANSSON, O. & ET AL. 2009. CSF biomarkers and incipient alzheimer disease in patients with mild cognitive impairment. *JAMA*, 302, 385-393.
382. MAYR, M., ZHANG, J., GREENE, A. S., GUTTERMAN, D., PERLOFF, J. & PING, P. 2006. Proteomics-based Development of Biomarkers in Cardiovascular Disease Mechanistic, Clinical, and Therapeutic Insights. *Molecular & Cellular Proteomics*, 5, 1853-1864.
383. MCBRIDE, M., CARR, F., GRAHAM, D., ANDERSON, N., CLARK, J., LEE, W. & AL, E. 2003. Microarray analysis of rat chromosome 2 congenic strains. *Hypertension* 41, 847-853.
384. MCKHANN, G. M., KNOPMAN, D. S., CHERTKOW, H., HYMAN, B. T., JACK, C. R., KAWAS, C. H., KLUNK, W. E., KOROSHETZ, W. J., MANLY, J. J. & MAYEUX, R. 2011. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia*, 7, 263-269.
385. MEHTA, R., KELLUM, J., SHAH, S., MOLITORIS, B., RONCO, C., WARNOCK, D., LEVIN, A. & NETWORK, T. A. K. I. 2007. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Critical Care*, 11, R31.
386. MEISTERMANN, H., NORRIS, J. L., AERNI, H.-R., CORNETT, D. S., FRIEDLEIN, A., ERSKINE, A. R., AUGUSTIN, A., MUDRY, M. C. D. V., RUEPP, S. & SUTER, L. 2006. Biomarker Discovery by Imaging Mass Spectrometry Transthyretin is a Biomarker for Gentamicin-induced Nephrotoxicity in Rat. *Molecular & Cellular Proteomics*, 5, 1876-1886.
387. MENETON, P., JEUNEMAITRE, X., DE WARDENER, H. & MACGREGOR, G. 2005. Links between dietary salt intake, renal salt handling, blood pressure, and cardiovascular diseases. *Physiological Reviews* 85, 679-715.
388. MERTINS, P., UDESHI, N. D., CLAUSER, K. R., MANI, D., PATEL, J., ONG, S.-E., JAFFE, J. D. & CARR, S. A. 2012. iTRAQ Labeling is Superior to mTRAQ for Quantitative Global Proteomics and Phosphoproteomics. *Molecular & Cellular Proteomics*, 11.
389. METZGER, J., KIRSCH, T., SCHIFFER, E., ULGER, P., MENTES, E., BRAND, K., WEISSINGER, E. M., HAUBITZ, M., MISCHAK, H. & HERGET-ROSENTHAL, S. 2010. Urinary excretion of twenty peptides forms an early and accurate diagnostic pattern of acute kidney injury. *Kidney International*, 78, 1252-1262.

390. MIRANDA, K., BOND, D., MCKEE, M., SKOG, J., PAUNESCU, T., DA SILVA, N., BROWN, D. & RUSSO, L. 2010. Nucleic acids within urinary exosomes/microvesicles are potential biomarkers for renal disease. *Kidney International*, 78, 191–199.
391. MIRZAEI, H., MCBEE, J. K., WATTS, J. & AEBERSOLD, R. 2008. Comparative Evaluation of Current Peptide Production Platforms Used in Absolute Quantification in Proteomics. *Molecular & Cellular Proteomics*, 7, 813–823.
392. MISCHAK, H., APWEILER, R., BANKS, R., M, C., COON, J., DOMINICZAK, A., EHRICH, J., FLISER, D., GIROLAMI, M., HERMJAKOB, H., HOCHSTRASSER, D., JANKOWSKI, J., JULIAN, B., KOLCH, W., MASSY, Z., NEUSUESS, C., NOVAK, J., PETER, K., ROSSING, K., SCHANSTRA, J., SEMMES, O., THEODORESCU, D., THONGBOONKERD, V., WEISSINGER, E., VAN EYK, J. & YAMAMOTO, T. 2007. Clinical proteomics: A need to define the field and to begin to set adequate standards. *Proteomics Clin. Appl.*, 1, 148–156.
393. MISCHAK, H. & ROSSING, P. 2010. Proteomic biomarkers in diabetic nephropathy—reality or future promise? *Nephrology Dialysis Transplantation*, 25, 2843–2845.
394. MISCHAK, H. & SCHANSTRA, J. P. 2011. CE_MS in biomarker discovery, validation, and clinical application. *PROTEOMICS-Clinical Applications*, 5, 9–23.
395. MISHRA, J., MA, Q., PRADA, A. & AL., E. 2003. Identification of neutrophil gelatinase-associated lipocalin as a novel urinary biomarker for ischemic injury. *J Am Soc Nephrol* 4, 2534–2543.
396. MISHRA, J., MORI, K., MA, Q. & AL., E. 2004. Neutrophil Gelatinase-Associated Lipocalin (NGAL): a novel urinary biomarker for cisplatin nephrotoxicity. *Am J Nephrol.*, 24, 307–315.
397. MIZUHIRA V, H. H. 1996. Microwave fixation method for cytochemistry. For conventional electron microscopy, enzyme-immunocytochemistry, autoradiography elemental distribution studies and staining methods. *Eur J Morphol.* , 34, 385–91.
398. MOLLOY, M. 2000. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. *Anal Biochem*, 280, 1–10.
399. MONICA, H., SMITH, D., PARKERA, C. & BORCHERS, C. 2009. Current trends in quantitative proteomics. *J. Mass. Spectrom*, 44, 1637–1660.
400. MOON, P.-G., YOU, S., LEE, J.-E., HWANG, D. & BAEK, M.-C. 2011. Urinary exosomes and proteomics. *Mass Spectrometry Reviews*, 30, 1185–1202.
401. MOUTON-BARBOSA, E., ROUX-DALVAI, F., BOUYSSIÉ, D., BERGER, F., SCHMIDT, E., RIGHETTI, P. G., GUERRIER, L., BOSCHETTI, E., BURLET-SCHILTZ, O. & MONSARRAT, B. 2010. In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification. *Molecular & Cellular Proteomics*, 9, 1006–1021.
402. MUNDY, C. 2001. The human genome project: a historical perspective. *Pharmacogenomics* 2, 37–49.
403. MUSANTE, L., CANDIANO, G., BRUSCHI, M., ZENNARO, C., CARRARO, M., ARTERO, M., GIUFFRIDA, M., CONTI, A., SANTUCCI, A. & GHIGGERI, G. 2002. Characterization of plasma factors that alter the permeability to albumin within isolated glomeruli. *Proteomics*, 2, 197–205.

404. NABITY, M. B., LEES, G. E., DANGOTT, L. J., CIANCIOLO, R., SUCHODOLSKI, J. S. & STEINER, J. M. 2011. Proteomic analysis of urine from male dogs during early stages of tubulointerstitial injury in a canine model of progressive glomerular disease. *Veterinary Clinical Pathology*, 40, 222-236.
405. NADON, R. & SHOEMAKER, J. 2002. Statistical issues with microarrays: processing and analysis. *TRENDS in Genetics*, 18, 265-271.
406. NAGAOKA, A., IWATSUKA, H., SUZUOKI, Z. ET AL 1976. Genetic predisposition to stroke in spontaneously hypertensive rats. *American journal of physiology*, 230, 1354-1359.
407. NAGATA, N., KUSAKARI, Y., FUKUNISHI, Y., INOUE, T. & URADE, Y. 2011. Catalytic mechanism of the primary human prostaglandin F₂ α synthase, aldo-keto reductase 1B1 – prostaglandin D₂ synthase activity in the absence of NADP(H). *FEBS Journal*, 278, 1288-1298.
408. NAGY, N., MALIK, G., FISHER, A. & DAS, D. 2006. Targeted disruption of peroxiredoxin 6 gene renders the heart vulnerable to ischemia-reperfusion injury. *Am J Physiol Heart Circ Physiol.*, 291., H2636-40.
409. NELSON, M., JONES, A., CARMEN, J., SINAI, A., BURCHMORE, R. & WASTLING, J. 2008. Modulation of the host cell proteome by the intracellular apicomplexan parasite *Toxoplasma gondii*. *Infection and immunity*, 76, 828-844.
410. NESVIZHSHKII, A. I. & AEBERSOLD, R. 2004. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discovery Today*, 9, 173-181.
411. NESVIZHSHKII, A. I., VITEK, O. & AEBERSOLD, R. 2007. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nature methods*, 4, 787-797.
412. NEWICK, K., CUNNIFF, B., PRESTON, K., HELD, P., ARBISER, J., PASS, H., MOSSMAN, B., SHUKLA, A. & HEINTZ, N. 2012. Peroxiredoxin 3 Is a Redox-Dependent Target of Thiostrepton in Malignant Mesothelioma Cells. *PLoS ONE*, 7, e39404.
413. NGUYEN, M., ROSS, G., DENT, C. & AL., E. 2005. Early prediction of acute renal injury using urinary proteomics. *Am J Nephrol*, 2005, 25:318-326.
414. NIELSEN, S., CHOU, C., MARPLES, D., CHRISTIANSEN, E., KISHORE, B. & KNEPPER, M. A. 1995. Vasopressin increases water permeability of kidney collecting duct by inducing translocation of aquaporin-CD water channels to plasma membrane. *Proc. Natl. Acad. Sci.*, 92, 1013-1017.
415. NIELSEN, S., DIGIOVANNI, S., CHRISTIANSEN, E., KNEPPER, M. A. & HARRIS, H. 1993a. Cellular and subcellular immunolocalization of vasopressin-regulated water channel in rat kidney. *Proc. Natl. Acad. Sci.*, 90, 11663-11667.
416. NIELSEN, S., FRØKJÆR, J., MARPLES, D., KWON, T.-H., AGRE, P. & KNEPPER, M. A. 2002a. Aquaporins in the Kidney: From Molecules to Medicine. *Physiological Reviews*, 82, 205-244.
417. NIELSEN, S., FROKLER, J., MARPLES, D., KWON, T.-H., AGRE, P. & KNEPPER, M. A. 2002b. Aquaporins in the Kidney: From Molecules to Medicine. *Physiol Rev*, 82, 205-244.
418. NIELSEN, S., SMITH, B., CHRISTIANSEN, E., KNEPPER, M. A. & AGRE, P. 1993b. CHIP28 water channels are localized in constitutively water-permeable segments of the nephron. *J Cell Biol* 120, 120, 371-383.

419. NIRMALAN, N., HARNDEN, P., SELBY, P. & BANKS, R. 2009a. Development and validation of a novel protein extraction methodology for quantitation of protein expression in formalin-fixed paraffin-embedded tissues using western blotting. *J Pathol.* 2009 Mar;x, 217, 497-506.
420. NIRMALAN, N. J., HARNDEN, P., SELBY, P. J. & BANKS, R. E. 2008. Mining the archival formalin-fixed paraffin-embedded tissue proteome: opportunities and challenges. *Molecular BioSystems*, 4, 712-720.
421. NIRMALAN, N. J., HARNDEN, P., SELBY, P. J. & BANKS, R. E. 2009b. Development and validation of a novel protein extraction methodology for quantitation of protein expression in formalin_fixed paraffin_embedded tissues using western blotting. *The Journal of pathology*, 217, 497-506.
422. NIRMALAN, N. J., HUGHES, C., PENG, J., MCKENNA, T., LANGRIDGE, J., CAIRNS, D. A., HARNDEN, P., SELBY, P. J. & BANKS, R. E. 2010. Initial Development and Validation of a Novel Extraction Method for Quantitative Mining of the Formalin-Fixed, Paraffin-Embedded Tissue Proteome for Biomarker Investigations. *Journal of Proteome Research*, 10, 896-906.
423. NISHIMORI, T., TOMONAGA, T., MATSUSHITA, K., KODERA, Y., MAEDA, T., NOMURA, F., MATSUBARA, H., SHIMADA, H. & T. O. 2006. Proteomic analysis of primary esophageal squamous cell carcinoma reveals downregulation of a cell adhesion protein, periplakin. *Proteomics*, 6, 1011-1018.
424. NONN, L., BERGGREN, M. & POWIS, G. 2003. Increased Expression of Mitochondrial Peroxiredoxin-3 (Thioredoxin Peroxidase-2) Protects Cancer Cells Against Hypoxia and Drug-Induced Hydrogen Peroxide-Dependent Apoptosis 11CA52995 and CA772049. *Molecular Cancer Research*, 1, 682-689.
425. NORDON, I., BRAR, R., HINCHLIFFE, R., COCKERILL, G. & THOMPSON, M. 2010. Proteomics and pitfalls in the search for potential biomarkers of abdominal aortic aneurysms. *Vascular.* , 18, 264-8.
426. O'FARRELL, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* , 250, 4007-4021.
427. ODA, Y., HUANG, K., CROSS, F. R., COWBURN, D. & CHAIT, B. T. 1999. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences*, 96, 6591-6596.
428. ODELSTAD, L., PHLMAN, S., LÄCKGREN, G., LARSSON, E., GROTTTE, G. & NILSSON, K. 1982. Neuron specific enolase: A marker for differential diagnosis of neuroblastoma and Wilms' tumor. *Journal of Pediatric Surgery*, 17, 381-385.
429. OH-ISHI, M. & MAEDA, T. 2002. Separation techniques for high-molecular-mass proteins. *J Chromatogr B Analyt Technol Biomed Life Sci.* , 771, 49-66.
430. OH-ISHI, M. & MAEDA, T. 2007. Disease proteomics of high-molecular-mass proteins by two-dimensional gel electrophoresis with agarose gels in the first dimension (Agarose 2-DE). *J Chromatogr B Analyt Technol Biomed Life Sci.* , 849, 211-22.
431. ONG, S.-E., BLAGOEV, B., KRATCHMAROVA, I., KRISTENSEN, D. B., STEEN, H., PANDEY, A. & MANN, M. 2002. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics*, 1, 376-386.
432. ONG, S.-E. & MANN, M. 2005. Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology*, 1, 252-262.
433. OPELZ, G., WUJCIAK, T. & RITZ, E. 1998 Association of chronic kidney graft failure with recipient blood pressure. *Kidney International* 53, 217-222.

434. OSTCHEGA, Y., DILLON, C., HUGHES, J., CARROLL, M. & YOON, S. 2007. Trends in hypertension prevalence, awareness, treatment, and control in older U.S. adults: data from the National Health and Nutrition Examination Survey 1988 to 2004. *Journal of the American Geriatrics Society* 55 1056–65.
435. OVERGAARD, A., THINGHOLM, T., LARSEN, M., TARNOW, L., ROSSING, P., MCGUIRE, J. & POCIOT, F. 2010. Quantitative iTRAQ-Based Proteomic Identification of Candidate Biomarkers for Diabetic Nephropathy in Plasma of Type 1 Diabetic Patients. *Clinical Proteomics*, 6, 105-114.
436. PALMER-TOY, D. E., KRASSTINS, B., SARRACINO, D. A., NADOL JR, J. B. & MERCHANT, S. N. 2005. Efficient method for the proteomic analysis of fixed and embedded tissues. *Journal of Proteome Research*, 4, 2404-2411.
437. PALMER, A. M., LOWE, S. L., FRANCIS, P. T., BOWEN, D. M., 1988. Are post-mortem biochemical studies of human brain worthwhile? *Biochem. Soc. Trans.* , 6, 472–475.
438. PALMER, B. & ALPERN, R. 1999. Liddle's Syndrome. *Am J Med.* , 104, 301–309.
439. PAN, L. & LI, J. 2010. K-Nearest Neighbor Based Missing Data Estimation Algorithm in Wireless Sensor Networks. *Wireless Sensor Network*, 2, 115-122.
440. PAN, S., CHEN, R., CRISPIN, D. A., MAY, D., STEVENS, T., MCINTOSH, M. W., BRONNER, M. P., ZIOGAS, A., ANTON-CULVER, H. & BRENTNALL, T. A. 2011. Protein alterations associated with pancreatic cancer and chronic pancreatitis found in human plasma using global quantitative proteomics profiling. *Journal of Proteome Research*, 10, 2359-2376.
441. PAN, S., ZHANG, H., RUSH, J., ENG, J., ZHANG, N., PATTERSON, D., COMB, M. J. & AEBERSOLD, R. 2005. High throughput proteome screening for biomarker detection. *Molecular & Cellular Proteomics*, 4, 182-190.
442. PAPAIE, M., DI PAOLO, S., MAGISTRONI, R., LAMACCHIA, O., DI PALMA, A. M., DE MATTIA, A., TERESA ROCCHETTI, M., FURCI, L., PASQUALI, S., DE COSMO, S., CIGNARELLI, M. & GESUALDO, L. 2010. Urine Proteome Analysis May Allow Noninvasive Differential Diagnosis of Diabetic Nephropathy. *Diabetes Care*, 33, 2409-2415.
443. PAPAIE, M., PEDICILLO, M. C., THATCHER, B. J., DI PAOLO, S., MUZIO, L. L., BUFO, P., ROCCHETTI, M. T., CENTRA, M., RANIERI, E. & GESUALDO, L. 2007. Urine profiling by SELDI-TOF/MS: monitoring of the critical steps in sample collection, handling and analysis. *Journal of Chromatography B*, 856, 205-213.
444. PAPPIN, D. J., P. HOJRUP, AND A.J. BLEASBY, 1993. Rapid identification of proteins by peptide-mass fingerprinting. . *Current Biology.* , 3, 327-32.
445. PARIKH, C., MISHRA, J., THIESSEN-PHILBROOK, H., DURSUN, B., MA, Q., KELLY, C., DENT, C., DEVARAJAN, P. & EDELSTEIN, C. 2006. Urinary IL-18 is an early predictive biomarker of acute kidney injury after cardiac surgery. *Kidney International* 70 199–203.
446. PARTIN, A. W., CATALONA, W. J., SOUTHWICK, P. C., SUBONG, E. N., GASIOR, G. H. & CHAN, D. W. 1996. Analysis of percent free prostate-specific antigen (PSA) for prostate cancer detection: influence of total PSA, prostate volume, and age. *Urology*, 48, 55-61.
447. PASQUALI, C., FIALKA, I. & HUBER, L. A. 1999. Subcellular fractionation, electromigration analysis and mapping of organelles. *Journal of Chromatography B: Biomedical Sciences and Applications*, 722, 89-102.

448. PAULOVICH, A. G., WHITEAKER, J. R., HOOFNAGLE, A. N. & WANG, P. 2008. The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline. *PROTEOMICS – Clinical Applications*, 2, 1386-1402.
449. PAWELETZ, C. P., LIOTTA, L. A. & PETRICOIN, E. F. 2001. New technologies for biomarker analysis of prostate cancer progression: Laser capture microdissection and tissue proteomics. *Urology*, 57, 160-163.
450. PAWITAN, Y., MICHIELS, S., KOSCIELNY, S., GUSNANTO, A. & PLONER, A. 2005a. False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics*, 21, 3017-3024.
451. PAWITAN, Y., MURTHY, K. R. K., MICHIELS, S. & PLONER, A. 2005b. Bias in the estimation of false discovery rate in microarray studies. *Bioinformatics*, 21, 3865-3872.
452. PEARSON, K. 1901. On Lines and Planes of Closest Fit to Systems of Points in Space *Philosophical Magazine* 2, 559–572. .
453. PEDRESCHI, R., HERTOOG, M. L., CARPENTIER, S. C., LAMMERTYN, J., ROBBEN, J., NOBEN, J. P., PANIS, B., SWENNEN, R. & NICOLAÏ, B. M. 2008a. Treatment of missing values for multivariate statistical analysis of gel-based proteomics data. *PROTEOMICS*, 8, 1371-1383.
454. PEDRESCHI, R., HERTOOG, M. L. A. T. M., CARPENTIER, S. C., LAMMERTYN, J., ROBBEN, J., NOBEN, J.-P., PANIS, B., SWENNEN, R. & NICOLAÏ, B. M. 2008b. Treatment of missing values for multivariate statistical analysis of gel-based proteomics data. *PROTEOMICS*, 8, 1371-1383.
455. PERCO, P., PLEBAN, C., KAINZ, A., LUKAS, A., MAYER, G., MAYER, B. & OBERBAUER, R. 2006. Protein biomarkers associated with acute renal failure and chronic kidney disease. *European journal of clinical investigation*, 36, 753-763.
456. PERKINS, D. N., PAPPIN, D.J., CREASY, D.M., COTTRELL, J.S., 1999. Probability based protein identification by searching sequence databases using mass spectrometry data. . *Electrophoresis*, 20, 3551–3567.
457. PERRONE, R., MADIAS, N. & LEVEY, A. 1992. Serum creatinine as an index of renal function: new insights into old concepts. *Clin Chem*, 38, 1933 - 1953.
458. PESCE, F., PATHAN, S. & SCHENA, F. P. 2013. From -omics to personalized medicine in nephrology: integration is the key. *Nephrology Dialysis Transplantation*, 28, 24-28.
459. PETRI, A., SIMONSEN, A., YIP, T., HOGDALL, E., FUNG, E., LUNDVALL, L. & HOGDALL, C. 2009. Three new potential ovarian cancer biomarkers detected in human urine with equalizer bead technology. *Acta Obstet Gynecol Scand*, 88, 18-26.
460. PETRICOIN, E. F., ZOON, K. C., KOHN, E. C., BARRETT, J. C. & LIOTTA, L. A. 2002. Clinical proteomics: translating benchside promise into bedside reality. *Nature Reviews Drug Discovery*, 1, 683-695.
461. PETRICOIN III, E. F., ORNSTEIN, D. K., PAWELETZ, C. P., ARDEKANI, A., HACKETT, P. S., HITT, B. A., VELASSCO, A., TRUCCO, C., WIEGAND, L. & WOOD, K. 2002. Serum proteomic patterns for detection of prostate cancer. *Journal of the National Cancer Institute*, 94, 1576-1578.
462. PETZOLD, A., BRETTSCHEIDER, J., JIN, K., KEIR, G., MURRAY, N. M. F., HIRSCH, N. P., ITOYAMA, Y., REILLY, M. M., TAKEDA, A. & TUMANI, H. 2009. CSF protein biomarkers for proximal axonal damage improve prognostic accuracy in the acute phase of Guillain-Barré syndrome. *Muscle & Nerve*, 40, 42-49.

463. PHILLIPS, T. M. & WELLNER, E. F. 2007. Analysis of inflammatory biomarkers from tissue biopsies by chip_based immunoaffinity CE. *Electrophoresis*, 28, 3041-3048.
464. PICARD, G., LEBERT, D., LOUWAGIE, M., ADRAIT, A., HUILLET, C., VANDENESCH, F., BRULEY, C., GARIN, J., JAQUINOD, M. & BRUN, V. 2012. PSAQ™ standards for accurate MS-based quantification of proteins: from the concept to biomedical applications. *Journal of Mass Spectrometry*, 47, 1353-1363.
465. PIERSMA, S. R., FIEDLER, U., SPAN, S., LINGNAU, A., PHAM, T. V., HOFFMANN, S., KUBBUTAT, M. H. & JIMÉNEZ, C. R. 2010. Workflow comparison for label-free, quantitative secretome proteomics for cancer biomarker discovery: method evaluation, differential analysis, and verification in serum. *Journal of Proteome Research*, 9, 1913-1922.
466. PIERSON, J., ET AL., 2004. Molecular Profiling of Experimental Parkinson's Disease: Direct Analysis of Peptides and Proteins on Brain Tissue Sections by MALDI Mass Spectrometry. *J. Proteome Res*, 3, 289-295.
467. PISITKUN, T., JOHNSTONE, R. & KNEPPER, M. A. 2006. Discovery of Urinary Biomarkers. *Molecular and Cellular Proteomics*, 5, 1760-1771.
468. PISITKUN, T., SHEN, R.-F. & KNEPPER, M. A. 2004. Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci U S A*, 101, 13368-13373.
469. PLAVINA, T., WAKSHULL, E., HANCOCK, W. S. & HINCAPIE, M. 2006. Combination of Abundant Protein Depletion and Multi-Lectin Affinity Chromatography (M-LAC) for Plasma Protein Biomarker Discovery. *Journal of Proteome Research*, 6, 662-671.
470. PLEBANI, M., PROTEOMICS: 2005. The next revolution in laboratory medicine? . *Clinica Chimica Acta*, 357, 113-122.
471. PLOUSSARD, G. & DE LA TAILLE, A. 2010. Urine biomarkers in prostate cancer. *Nature Reviews Urology*, 7, 101-109.
472. POLMAN, C. H., REINGOLD, S. C., BANWELL, B., CLANET, M., COHEN, J. A., FILIPPI, M., FUJIHARA, K., HAVRDOVA, E., HUTCHINSON, M. & KAPPOS, L. 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology*, 69, 292-302.
473. POON, T. C., YIP, T.-T., CHAN, A. T., YIP, C., YIP, V., MOK, T. S., LEE, C. C., LEUNG, T. W., HO, S. K. & JOHNSON, P. J. 2003. Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes. *Clin Chem*, 49, 752-760.
474. POTTHOFF, S. A., SITEK, B., STEGBAUER, J., SCHULENBORG, T., MARCUS, K., QUACK, I., RUMP, L. C., MEYER, H. E., STÜHLER, K. & VONEND, O. 2008. The glomerular proteome in a model of chronic kidney disease. *PROTEOMICS – Clinical Applications*, 2, 1127-1139.
475. PRATT, J. M., SIMPSON, D. M., DOHERTY, M. K., RIVERS, J., GASKELL, S. J. & BEYNON, R. J. 2006. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature protocols*, 1, 1029-1043.
476. PUCHADES, M., HANSSON, S. F., NILSSON, C. L., ANDREASEN, N., BLENNOW, K. & DAVIDSSON, P. 2003. Proteomic studies of potential cerebrospinal fluid protein markers for Alzheimer's disease. *Molecular brain research*, 118, 140-146.

477. QIAN, W.-J., MONROE, M. E., LIU, T., JACOBS, J. M., ANDERSON, G. A., SHEN, Y., MOORE, R. J., ANDERSON, D. J., ZHANG, R., CALVANO, S. E., LOWRY, S. F., XIAO, W., MOLDAWER, L. L., DAVIS, R. W., TOMPKINS, R. G., CAMP, D. G., SMITH, R. D., INFLAMMATION, T. & PROGRAM, T. H. R. T. I. L. S. C. R. 2005. Quantitative Proteome Analysis of Human Plasma following in Vivo Lipopolysaccharide Administration Using 16O/18O Labeling and the Accurate Mass and Time Tag Approach. *Molecular & Cellular Proteomics*, 4, 700-709.
478. RABEK, J. P., HAFER-MACKO, C. E., AMANING, J. K., DEFORD, J. H., DIMAYUGA, V. L., MADSEN, M. A., MACKO, R. F. & PAPACONSTANTINO, J. 2009. A Proteomics Analysis of the Effects of Chronic Hemiparetic Stroke on Troponin T Expression in Human Vastus Lateralis. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 64, 839.
479. RAI, A. J., GELFAND, C. A., HAYWOOD, B. C., WARUNEK, D. J., YI, J., SCHUCHARD, M. D., MEHIGH, R. J., COCKRILL, S. L., SCOTT, G. B. I., TAMMEN, H., SCHULZ-KNAPPE, P., SPEICHER, D. W., VITZTHUM, F., HAAB, B. B., SIEST, G. & CHAN, D. W. 2005. HUPO Plasma Proteome Project specimen collection and handling: Towards the standardization of parameters for plasma proteome samples. *PROTEOMICS*, 5, 3262-3277.
480. RAIMONDO, F., SALEMI, C., CHINELLO, C., FUMAGALLI, D., MOROSI, L., ROCCO, F. & AL., E. 2012a. Proteomic analysis in clear cell renal cell carcinoma: identification of differentially expressed protein by 2-D DIGE. *Mol Biosyst* 2012;8; 8, 1040-51.
481. RAIMONDO, F., SALEMI, C., CHINELLO, C., FUMAGALLI, D., MOROSI, L., ROCCO, F., FERRERO, S., PEREGO, R., BIANCHI, C., SARTO, C., PITTO, M., BRAMBILLA, P. & MAGNI, F. 2012b. Proteomic analysis in clear cell renal cell carcinoma: identification of differentially expressed protein by 2-D DIGE. *Molecular BioSystems*, 8, 1040-1051.
482. RANDA, B. A., LAURENCE, M., CAPUCINE, B., CHAMSEDDINE, K., FAYÇAL, J., HAMMADI, A., FRANCK, M. & CLAUDE, G. 2010. Proteomic approaches for discovering biomarkers of diabetic nephropathy. *Nephrology Dialysis Transplantation*, 25, 2866-2875.
483. RAPPILBER, J., MANN, M. & ISHIHAMA, Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols*, 2, 1896-1906.
484. RASMUSSEN, H., ORNTOFT, T., WOLF, H. & CELIS, J. 1996. Towards a comprehensive database of proteins from the urine of patients with bladder cancer. *J Urol* 155, 2113-2119.
485. RAY, C. A., BOWSER, R. R., SMITH, W. C., DEVANARAYAN, V., WILLEY, M. B., BRANDT, J. T. & DEAN, R. A. 2005. Development, validation, and implementation of a multiplex immunoassay for the simultaneous determination of five cytokines in human serum. *Journal of pharmaceutical and biomedical analysis*, 36, 1037-1044.
486. RAYLEIGH, L. 1882. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 14, 184-186.
487. REEVES, W. B., KWON, O. & RAMESH, G. 2008. Netrin-1 and kidney injury. II. Netrin-1 is an early biomarker of acute kidney injury. *American Journal of Physiology-Renal Physiology*, 294, F731-F738.
488. REN, F., WU, H., LEI, Y., ZHANG, H., LIU, R., ZHAO, Y., CHEN, X., ZENG, D., TONG, A. & CHEN, L. 2010. Research Quantitative proteomics identification of phosphoglycerate mutase 1 as a novel therapeutic target in hepatocellular carcinoma.

489. RETTIG R, G. O. 2005. The kidney as a determinant of genetic hypertension - Evidence from renal transplantation studies. . *Hypertension* 46, 463-468.
490. REY, S. & SEMENZA, G. L. 2010. Hypoxia-inducible factor-1-dependent mechanisms of vascularization and vascular remodelling. *Cardiovascular research*, 86, 236-242.
491. REYZER, M. L. & CAPRIOLI, R. M. 2005. MALDI mass spectrometry for direct tissue analysis: a new tool for biomarker discovery. *Journal of Proteome Research*, 4, 1138-1142.
492. RICHARD, J. A. G., JESSICA, C. DUNGWORTH ,STUART, R. COBB, ANDREW, R. PITT 2008. Time-dependent evolution of tissue markers by MALDI-MS imaging. *PROTEOMICS*, 8, 3801-3808.
493. RIDKER, P. M., HENNEKENS, C. H., BURING, J. E. & RIFAI, N. 2000. C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. *New England Journal of Medicine*, 342, 836-843.
494. RIFAI, N. & GERSZTEN, R. E. 2006. Biomarker Discovery and Validation. *Clin Chem*, 52, 1635-1637.
495. RIFAI, N., GILLETTE, M. A. & CARR, S. A. 2006a. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature Biotechnology*, 24, 971.
496. RIFAI, N., GILLETTE, M. A. & CARR, S. A. 2006b. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature Biotechnology*, 24, 971-983.
497. RINNE, J. O., BROOKS, D. J., ROSSOR, M. N., FOX, N. C., BULLOCK, R., KLUNK, W. E., MATHIS, C. A., BLENNOW, K., BARAKOS, J. & OKELLO, A. A. 2010. ¹¹C-PiB PET assessment of change in fibrillar amyloid- β load in patients with Alzheimer's disease treated with bapineuzumab: a phase 2, double-blind, placebo-controlled, ascending-dose study. *The Lancet Neurology*, 9, 363-372.
498. ROBINSON, A. A., WESTBROOK, J. A., ENGLISH, J. A., BORÉN, M. & DUNN, M. J. 2009. Assessing the use of thermal treatment to preserve the intact proteomes of post-mortem heart and brain tissue. *PROTEOMICS*, 9, 4433-4444.
499. RODRIGUEZ-PINEIRO, A. M., DE LA CADENA, M. P., LOPEZ-SACO, A. AND RODRIGUEZ-BERROCAL, F. J. 2006. Differential Expression of Serum Clusterin Isoforms in Colorectal Cancer. *Mol Cell Proteomics* 5, 1647-1657.
500. ROHNER, T. C., STAAB, D. & STOECKLI, M. 2005. MALDI mass spectrometric imaging of biological tissue sections. *Mechanisms of Ageing & Development*, 126, 177-85.
501. ROHRER, L., HERSBERGER, M. & VON ECKARDSTEIN, A. 2004. High density lipoproteins in the intersection of diabetes mellitus, inflammation and cardiovascular disease. *Current opinion in lipidology*, 15, 269-278.
502. ROMANUIK, T. L., UEDA, T., LE, N., HAILE, S., YONG, T. M., THOMSON, T., VESSELLA, R. L. & SADAR, M. D. 2009. Novel biomarkers for prostate cancer including noncoding transcripts. *The American journal of pathology*, 175, 2264-2276.
503. ROMIG, T. S., BELL, C. & DROLET, D. W. 1999. Aptamer affinity chromatography: combinatorial chemistry applied to protein purification. *Journal of chromatography. B, Biomedical sciences and applications*, 731, 275.
504. ROSS, P. L., HUANG, Y. N., MARCHESE, J. N., WILLIAMSON, B., PARKER, K., HATTAN, S., KHAINOVSKI, N., PILLAI, S., DEY, S., DANIELS, S., PURKAYASTHA, S., JUHASZ, P., MARTIN, S., BARTLET-JONES, M., HE, F.,

- JACOBSON, A. & PAPPIN, D. J. 2004. Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics*, 3, 1154-1169.
505. ROSSING, K., MISCHAK, H., DAKNA, M., ZÜRBIG, P., NOVAK, J., JULIAN, B. A., GOOD, D. M., COON, J. J., TARNOW, L., ROSSING, P. & NETWORK, O. B. O. T. P. 2008. Urinary Proteomics in Diabetes and CKD. *Journal of the American Society of Nephrology*, 19, 1283-1290.
506. ROSTY, C., CHRISTA, L., KUZDZAL, S., BALDWIN, W., ZAHURAK, M., CARNOT, F., CHAN, D., CANTO, M., LILLEMOE, K., CAMERON, J., YEO, C., HRUBAN, R. & GOGGINS, M. 2002. Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. *Cancer Res*, 62, 1868-1875.
507. ROUNTREE, C. B., VAN KIRK, C. A., YOU, H., DING, W., DANG, H., VANGUILDER, H. D. & FREEMAN, W. M. 2010. Clinical application for the preservation of phospho-proteins through in-situ tissue stabilization. *Proteome Sci*, 8, 61.
508. RUIJTER ET, M. G., AALDERS TW, VAN DE KAA CA, SCHALKEN JA, DEBRUYNE FM, BOON ME. 1997. Rapid microwave-stimulated fixation of entire prostatectomy specimens. Biomed-II MPC Study Group. *J Pathol.*, 183, 369-75.
509. RUNSWICK, M. J. & WALKER, J. E. 1983. The amino acid sequence of the beta-subunit of ATP synthase from bovine heart mitochondria. *Journal of Biological chemistry*, 258, 3081-3089.
510. SAIKI, R. K., ET AL., 1985. . Enzymatic amplification of b-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia *Science* 230,
511. SANDERS, M. E., DIAS, E. C., XU, B. J., MOBLEY, J. A., BILLHEIMER, D., RODER, H., GRIGORIEVA, J., DOWSETT, M., ARTEAGA, C. L. & CAPRIOLI, R. M. 2008. Differentiating proteomic biomarkers in breast cancer by laser capture microdissection and MALDI MS. *Journal of Proteome Research*, 7, 1500.
512. SARTO, C., DEON, C., DORO, G., HOCHSTRASSER, D., MOCARELLI, P. & SANCHEZ, C. 2001. Contribution of proteomics to the molecular analysis of renal cell carcinoma with an emphasis on manganese superoxide dismutase. *Proteomics* 1, 1288-1294.
513. SAWICKI, G. & JUGDUTT, B. I. 2004. Detection of regional changes in protein levels in the in vivo canine model of acute heart failure following ischemia_reperfusion injury: Functional proteomics studies. *PROTEOMICS*, 4, 2195-2202.
514. SCHAUB, S., WILKINS, J. & NICKERSON, P. 2005. Proteomics in Renal Transplantation: Opportunities and Challenges. *Clinical transplants*, 2005, 253.
515. SCHAUB, S., WILKINS, J. & NICKERSON, P. 2008. Proteomics and renal transplantation: searching for novel biomarkers and therapeutic targets.
516. SCHAUB, S., WILKINS, J., WEILER, T., SANGSTER, K., RUSH, D. & NICKERSON, P. 2004. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney International*, 65, 323-332.
517. SCHIFFMANN, R., WALDEK, S., BENIGNI, A. & AURAY-BLAIS, C. 2010. Biomarkers of Fabry Disease Nephropathy. *Clinical Journal of the American Society of Nephrology*, 5, 360-364.
518. SCHNERMANN, J., CHOW, C., TONGHUI, M., TRAYNOR, T., KNEPPER, M. A. & VERKMAN, A. 1998. Defective proximal tubular fluid reabsorption in transgenic aquaporin-1 null mice. *Proc. Natl. Acad. Sci.*, 96, 9660-9664.

519. SCHOLLER, N., GROSS, J., GARVIK, B., WELLS, L., LIU, Y., LOCH, C., RAMIREZ, A., MCINTOSH, M., LAMPE, P. & URBAN, N. 2008. Use of cancer-specific yeast-secreted in vivo biotinylated recombinant antibodies for serum biomarker discovery. . *J Transl Med* 6.
520. SCHOLZ, B., ALM, H., MATTSSON, A., NILSSON, A., KULTIMA, K., SAVITSKI, M. M., FÄLTH, M., SKÖLD, K., BRUNSTRÖM, B. & ANDREN, P. E. 2010a. Neuropeptidomic analysis of the embryonic Japanese quail diencephalon. *BMC developmental biology*, 10, 30.
521. SCHOLZ, B., SKÄLD, K., KULTIMA, K., FERNANDEZ, C., WALDEMARSON, S., SAVITSKI, M. M., SVENSSON, M., BOREN, M., STELLA, R., ANDREN, P. E., ZUBAREV, R. & JAMES, P. 2010b. Impact of temperature dependent sampling procedures in proteomics and peptidomics – A characterization of the liver and pancreas post mortem degradome. *Molecular & Cellular Proteomics*, -.
522. SCHOLZ, B., SKÖLD, K., KULTIMA, K., FERNANDEZ, C., WALDEMARSON, S., SAVITSKI, M. M., SÖDERQUIST, M., BORÉN, M., STELLA, R., ANDRÉN, P., ZUBAREV, R. & JAMES, P. 2011. Impact of Temperature Dependent Sampling Procedures in Proteomics and Peptidomics – A Characterization of the Liver and Pancreas Post Mortem Degradome. *Molecular & Cellular Proteomics*, 10.
523. SCHULZE WX, U. B. 2010. Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol.* , 61, 491-516.
524. SCHULZE, W. X. & USADEL, B. 2010. Quantitation in mass-spectrometry-based proteomics. *Annual review of plant biology*, 61, 491-516.
525. SCHWANHÄUSSER, B., BUSSE, D., LI, N., DITTMAR, G., SCHUCHHARDT, J., WOLF, J., CHEN, W. & SELBACH, M. 2011. Global quantification of mammalian gene expression control. *Nature*, 473, 337-342.
526. SCHWARTZ, S. A., REYZER, M. L. & CAPRIOLI, R. M. 2003. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *Journal of Mass Spectrometry*, 38, 699-708.
527. SCHWARZ, G., BECK, S., WELLER, M. G. & LINSCHIED, M. W. 2011. MeCAT—new iodoacetamide reagents for metal labeling of proteins and peptides. *Analytical and Bioanalytical Chemistry*, 401, 1203-1209.
528. SCOTT, R. P. W. 1992. Modern Liquid-Chromatography. . *Chemical Society Reviews*, 21 137-145.
529. SEMENZA, G. L. 2001. Hypoxia-inducible factor 1: oxygen homeostasis and disease pathophysiology. *Trends in Molecular Medicine*, 7, 345-350.
530. SERGEANT, N., WATTEZ, A., GALVÁN-VALENCIA, M., GHESTEM, A., DAVID, J., LEMOINE, J., SAUTIÉRE, P., DACHARY, J., MAZAT, J., MICHALSKI, J., VELOURS, J., MENA-LÓPEZ, R. & DELACOURTE, A. 2003. Association of ATP synthase alpha-chain with neurofibrillary degeneration in Alzheimer's disease. *Neuroscience.*, 117, 293-303.
531. SESSO, H., BURING, J., RIFAI, N., BLAKE, G., GAZIANO, J. & RIDKER, P. 2003. C-reactive protein and the risk of developing hypertension. *JAMA*. 2003, 290:2945–2951.
532. SHAPIRO, S. S. & WILK, M. B. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591-611.
533. SHARMA, J., UMAA, K., NOORA, A. & RAHMANA, A. 1994. Blood Pressure Regulation By The Kallikrein-Kinin System *General Pharmacology*, 27, 55-63.

534. SHARMA, K., LEE, S., HAN, S., LEE, S., FRANCOS, B., MCCUE, P., WASSELL, R., SHAW, M. A. & RAMACHANDRARAO, S. P. 2005. Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy. *PROTEOMICS*, 5, 2648-2655.
535. SHEN, Y., SENZER, N. & NEMUNAITIS, J. 2008. Use of Proteomics Analysis for Molecular Precision Approaches in Cancer Therapy. *Drug Target Insights*, 3, v.
536. SHIMADA Y, S. F., SHIMIZU K, TSUJIMOTO G, TSUKADA K. 2009 cDNA microarray analysis of esophageal cancer: discoveries and prospects. *Gen Thorac Cardiovasc Surg.*, 57, 347-56.
537. SHIMKETS, R. A., LIFTON, R.P., ET AL 1997. The activity of the epithelial sodium channels is regulated by clathrin-mediated endocytosis. *Journal of Biological chemistry*, 272, 25537-25541.
538. SIEW, E. D., WARE, L. B. & IKIZLER, T. A. 2011. Biological Markers of Acute Kidney Injury. *Journal of the American Society of Nephrology*, 22, 810-820.
539. SIRONI, L., GUERRINI, U., TREMOLI, E., MILLER, I., GELOSA, P., LASCIALFARI, A., ZUCCA, I., EBERINI, I., GEMEINER, M. & PAOLETTI, R. 2004. Analysis of pathological events at the onset of brain damage in stroke-prone rats: A proteomics and magnetic resonance imaging approach. *Journal of neuroscience research*, 78, 115-122.
540. SIRONI, L., TREMOLI, E., MILLER, I., GUERRINI, U., CALVIO, A. M., EBERINI, I., GEMEINER, M., ASDENTE, M., PAOLETTI, R. & GIANAZZA, E. 2001. Acute-phase proteins before cerebral ischemia in stroke-prone rats identification by proteomics. *Stroke*, 32, 753-760.
541. SITEK, B., LÜTTGES, J., MARCUS, K., KLÖPPEL, G., SCHMIEGEL, W., MEYER, H. E. & AL, E. 2005a. Application of fluorescence difference gel electrophoresis saturation labelling for the analysis of microdissected precursor lesions of pancreatic ductal adenocarcinoma. *Proteomics* 5, 2665–2679.
542. SITEK, B., LÜTTGES, J., MARCUS, K., KLÖPPEL, G., SCHMIEGEL, W., MEYER, H. E., HAHN, S. A. & STÜHLER, K. 2005b. Application of fluorescence difference gel electrophoresis saturation labelling for the analysis of microdissected precursor lesions of pancreatic ductal adenocarcinoma. *PROTEOMICS*, 5, 2665-2679.
543. SKÖLD, K., SVENSSON, M., NORRMAN, M., SJÖGREN, B., SVENNINGSSON, P. & ANDRÉN, P. E. 2007a. The significance of biochemical and molecular sample integrity in brain proteomics and peptidomics: Stathmin 2-20 and peptides as sample quality indicators. *PROTEOMICS*, 7, 4445-4456.
544. SKÖLD, K., SVENSSON, M., NORRMAN, M., SJÖGREN, B., SVENNINGSSON, P. & ANDRÉN, P. E. 2007b. The significance of biochemical and molecular sample integrity in brain proteomics and peptidomics: Stathmin 2-20 and peptides as sample quality indicators. *Proteomics*, 7, 4445–4456.
545. SLENO, L. & VOLMER, D. 2004. Ion activation methods for tandem mass spectrometry. *J Mass Spectrom.*, 10, 1091-112.
546. SMEJKAL, G. B., RIVAS-MORELLO, C., CHANG, J.-H. R., FREEMAN, E., TRACHTENBERG, A. J., LAZAREV, A., IVANOV, A. R. & KUO, W. P. 2011. Thermal stabilization of tissues and the preservation of protein phosphorylation states for two-dimensional gel electrophoresis. *Electrophoresis*, 32, 2206-2215.
547. SPELLMAN, C., AHMED, M. M., DUBACH, D. & GARDINER, K. J. 2013. Expression of trisomic proteins in Down syndrome model systems. *Gene*, 512, 219-225.

548. SPICKETT, C. M., PITT, A. R., MORRICE, N. AND KOLCH, W. 2006. Proteomic analysis of phosphorylation, oxidation and nitrosylation in signal transduction. . *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1764, 1823-1841.
549. SPRUNG, R. W., BROCK, J. W., TANKSLEY, J. P., LI, M., WASHINGTON, M. K., SLEBOS, R. J. & LIEBLER, D. C. 2009. Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis. *Molecular & Cellular Proteomics*, 8, 1988-1998.
550. SRINIVAS, P. R., VERMA, M., ZHAO, Y. & SRIVASTAVA, S. 2002. Proteomics for cancer biomarker discovery. *Clin Chem*, 48, 1160-1169.
551. STAES, A., DEMOL, H., VAN DAMME, J., MARTENS, L., VANDEKERCKHOVE, J. & GEVAERT, K. 2004. Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. *Journal of Proteome Research*, 3, 786-791.
552. STASIAK J, K. M., BOBER L, BACZEK T. 2010. Principal component analysis of HPLC retention data and molecular modeling structural parameters of cardiovascular system drugs in view of their pharmacological activity. *Int J Mol Sci*, 11, 2681-98.
553. STASYK, T. & HUBER, L. A. 2004. Zooming in: fractionation strategies in proteomics. *PROTEOMICS*, 4, 3704-3716.
554. STENVINKEL, P., CARRERO, J. J., AXELSSON, J., LINDHOLM, B., HEIMBÜRGER, O. & MASSY, Z. 2008. Emerging Biomarkers for Evaluating Cardiovascular Risk in the Chronic Kidney Disease Patient: How Do New Pieces Fit into the Uremic Puzzle? *Clinical Journal of the American Society of Nephrology*, 3, 505-521.
555. STEPHENS, W. 1946. Pulsed Mass Spectrometer with Time Dispersion *Bull. Am. Phys. Soc*, 21, 22.
556. STERN, M. P., WILLIAMS, K. & HAFFNER, S. M. 2002. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Annals of Internal Medicine*, 136, 575.
557. STEWART, P. M. 1999. Mineralocorticoid Hypertension. *The Lancet*, 358, 1341-1347.
558. STOECKLI, M., CHAURAND, P., HALLAHAN, D. E., CAPRIOLI, R. M., 2001. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. . *Nat. Med.* , 7, 493-496.
559. STOECKLI, M., ET AL., 2001. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. . *Nature Medicine*, 7, 493-6.
560. STOECKLI, M., STAAB, D., STAUFENBIEL, M., WIEDERHOLD, K.-H. & SIGNOR, L. 2002. Molecular imaging of amyloid [beta] peptides in mouse brain sections using mass spectrometry. *Analytical Biochemistry*, 311, 33.
561. STOJNEV, S., PEJICIC, M., DOLICANIN, Z., VELICKOVIC, L. J., DIMOV, I. & STEFANOVIC, V. 2009. Challenges of Genomics and Proteomics in Nephrology. *Renal Failure*, 31, 765-772.
562. STOWASSER, M. 2006. Monogenic mineralocorticoid hypertension. *Best Practice & Research Clinical Endocrinology & Metabolism*, 20, 401-420.
563. STUDENT, A. 1908. THE PROBABLE ERROR OF A MEAN. *BIOMETRIKA*, 6, 1-25.
564. SUPAVEKIN, S., ZHANG, W., KUCHERLAPATI, R. & AL, E. 2003. Differential gene expression following early renal ischemia-reperfusion. *Kidney Int* 63, 1714-1724.

565. SUSZTAK, K. & BÖTTINGER, E. P. 2006. Diabetic nephropathy: a frontier for personalized medicine. *Journal of the American Society of Nephrology*, 17, 361-367.
566. SVENSSON, M., BOREN, M., SKÖLD, K., FÄLTH, M., SJÖGREN, B., ANDERSSON, M., SVENNINGSSON, P. & ANDRÉN, P. E. 2009a. Heat Stabilization of the Tissue Proteome: A New Technology for Improved Proteomics. *Journal of Proteome Research*, 8, 974-981.
567. SVENSSON, M., BORÉN, M., SKÖLD, K., FÄLTH, M., SJÖGREN, B., ANDERSSON, M., SVENNINGSSON, P. & ANDRÉN, P. E. 2009b. Heat Stabilization of the Tissue Proteome: A New Technology for Improved Proteomics. *Journal of Proteome Research*, 8, 974-981.
568. SVENSSON, M., SKOLD, K., SVENNINGSSON, P., ANDREN, P. E., 2003. Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.*, 2, 213-219.
569. SVENSSON, M., SKVLD, K., NILSSON, A., FDLTH, M., NYDAHL, K., SVENNINGSSON, P. & ANDRIN, P. E. 2007. Neuropeptidomics: MS Applied to the Discovery of Novel Peptides from the Brain. *Anal Chem*, 79, 14-21.
570. SZÖKŐ, É. & TÁBI, T. 2010. Analysis of biological samples by capillary electrophoresis with laser induced fluorescence detection. *Journal of pharmaceutical and biomedical analysis*, 53, 1180-1192.
571. TAGUCHI, A., POLITI, K., PITTERI, S. J., LOCKWOOD, W. W., FAÇA, V. M., KELLY-SPRATT, K., WONG, C.-H., ZHANG, Q., CHIN, A. & PARK, K.-S. 2011. Lung cancer signatures in plasma based on proteome profiling of mouse tumor models. *Cancer cell*, 20, 289-299.
572. TANAKA, K., WAKI, H, IDO, Y, AKITA, S, YOSHIDA, Y, AND YOSHIDA, T., 1988a. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2, 151-153.
573. TANAKA, K., WAKI, H, IDO, Y, AKITA, S, YOSHIDA, Y, AND YOSHIDA, T., MONTGOMERY SB, DERMITZAKIS ET. 2009. he resolution of the genetics of gene expression. *Hum Mol Genet* 15, R211-5.
574. TANAKA, K. W., H.; IDO Y.; AKITA, S.; YOSHIDA, Y.; TAMIO, Y.; MATSUO, Y.T. 1988b. Protein and polymer analyses up to m/z 100 000 by laser ionization timeof-flight mass spectrometry.
575. *Rapid Communications in Mass Spectrometry* 2, , 151-153.
576. TAYLOR, C. F., PATON, N. W., GARWOOD, K. L., KIRBY, P. D., STEAD, D. A., YIN, Z., DEUTSCH, E. W., SELWAY, L., WALKER, J. & RIBA-GARCIA, I. 2003. A systematic approach to modeling, capturing, and disseminating proteomics experimental data. *Nature Biotechnology*, 21, 247-254.
577. TERRIS, J., ECELBARGER, C., NIELSEN, S., MARPLES, D. & KNEPPER, M. A. 1995. Distribution of aquaporin-4 water channel expression within rat kidney. *Am J Physiol Renal Physiol*, 269, F775-F785.
578. TERRIS, J., ECELBARGER, C., NIELSEN, S., MARPLES, D., TAKASKI, M., BAKER, E., VERBALIS, J. & KNEPPER, M. A. 1997. Role of Renal Aquaporins in Escape from Vasopressin-induced Antidiuresis in Rat. *The Journal of Clinical Investigation*, 99, 1852-1863.
579. TERRIS, J., ECELBARGER, C., NIELSEN, W. & KNEPPER, M. A. 1996. Long-term regulation of four renal aquaporins in rats. *American Journal of Physiology - Renal Physiology*, 271, F414-F422.

580. THAKKER-VARIA, S., TOZZI, C., CHARI, S., TIKU, K. & RILEY, D. 1999. Isolation of differentially expressed genes in hypertensive pulmonary artery of rats. *Exp Lung Res.*, 25, 689-99.
581. THOMAS, C. E., SEXTON, W., BENSON, K., SUTPHEN, R. & KOOMEN, J. 2010. Urine collection and processing for protein biomarker discovery and quantification. *Cancer Epidemiology Biomarkers & Prevention*, 19, 953-959.
582. THOMPSON, A., SCHÄFER, J., KUHN, K., KIENLE, S., SCHWARZ, J., SCHMIDT, G., NEUMANN, T. & HAMON, C. 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem*, 75, 1895-1904.
583. THOMSON, J. J. 1897. Cathode Rays. *Phil. Mag.*, 44, 293.
584. THONGBOONKERD, V. 2004. Proteomics in nephrology: current status and future directions. *American journal of nephrology*, 24, 360-378.
585. THONGBOONKERD, V. 2005. Genomics, proteomics and integrative 'omics' in hypertension research. *Current Opinion in Nephrology and Hypertension*, 14, 133-139.
586. THONGBOONKERD, V. 2007. Practical points in urinary proteomics. *Journal of Proteome Research*, 6, 3881-3890.
587. THONGBOONKERD, V. 2010. Current status of renal and urinary proteomics: ready for routine clinical application? *Nephrology Dialysis Transplantation*, 25, 11-16.
588. THONGBOONKERD, V., BARATI, M. T., MCLEISH, K. R., BENARAFI, C., REMOLD-O'DONNELL, E., ZHENG, S., ROVIN, B. H., PIERCE, W. M., EPSTEIN, P. N. & KLEIN, J. B. 2004. Alterations in the Renal Elastin-Elastase System in Type 1 Diabetic Nephropathy Identified by Proteomic Analysis. *Journal of the American Society of Nephrology*, 15, 650-662.
589. THONGBOONKERD, V., GOZAL, E., SACHLEBEN, L. R., ARTHUR, J. M., PIERCE, W. M., CAI, J., CHAO, J., BADER, M., PESQUERO, J. B., GOZAL, D. & KLEIN, J. B. 2002. Proteomic Analysis Reveals Alterations in the Renal Kallikrein Pathway during Hypoxia-Induced Hypertension. *Journal of Biological Chemistry*, 277, 34708-34716.
590. THONGBOONKERD, V. & MALASIT, P. 2005. Renal and urinary proteomics: Current applications and challenges. *PROTEOMICS*, 5, 1033-1042.
591. TIMIO, F., KERRY, S. M., ANSON, K. M., EASTWOOD, J. B. & CAPPUCIO, F. P. 2003. Calcium urolithiasis, blood pressure and salt intake. *Blood pressure*, 12, 122-127.
592. TONGE, R., SHAW, J., MIDDLETON, B., ROWLINSON, R., RAYNER, S., YOUNG, J. & AL, E. 2001. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics*, 1, 377-96.
593. TROIANO, N., CIOVACCO, W. & KACENA, M. 2009. The Effects of Fixation and Dehydration on the Histological Quality of Undecalcified Murine Bone Specimens Embedded in Methylmethacrylate. *J Histotechnol.*, 1, 27-31.
594. TROYANSKAYA, O., CANTOR, M., SHERLOCK, G., BROWN, P., HASTIE, T., TIBSHIRANI, R., BOTSTEIN, D. & ALTMAN, R. B. 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17, 520-525.
595. TSWETT, M. S. 1905. O novoy kategorii adsorbtsionnykh yavleny i o primenenii ikh k biokhimiicheskomu analizu" (On a new category of adsorption phenomena and on its application to biochemical analysis). *Trudy Varhavskago Obshchestva Estestvoispytatelei, Otdelenie Biologii (Proceedings of the Warsaw Society of Naturalists)*, 14, 20-39.

596. TU, C., RUDNICK, P. A., MARTINEZ, M. Y., CHEEK, K. L., STEIN, S. E., SLEBOS, R. J. & LIEBLER, D. C. 2010. Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of Proteome Research*, 9, 4982.
597. TURTOI, A., MAZZUCHELLI, G. & DE PAUW, E. 2010. Isotope coded protein label quantification of serum proteins— Comparison with the label-free LC–MS and validation using the MRM approach. *Talanta*, 80, 1487.
598. TYTHER, R., AHMEDA, A., JOHNS, E. & SHEEHAN, D. 2009. Protein carbonylation in kidney medulla of the spontaneously hypertensive rat. *PROTEOMICS – Clinical Applications*, 3, 338-346.
599. UHLÉN, M., BJÖRLING, E., AGATON, C., SZIGYARTO, C. A.-K., AMINI, B., ANDERSEN, E., ANDERSSON, A.-C., ANGELIDOU, P., ASPLUND, A. & ASPLUND, C. 2005. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Molecular & Cellular Proteomics*, 4, 1920-1932.
600. UMMANNI, R., MUNDT, F., POSPISIL, H., VENZ, S., SCHARF, C., BARETT, C., FÄLTH, M., KÖLLERMANN, J., WALTHER, R. & SCHLOMM, T. 2011. Identification of clinically relevant protein targets in prostate cancer with 2D-DIGE coupled mass spectrometry and systems biology network platform. *PLoS ONE*, 6, e16833.
601. UNLÜ, M., MORGAN, M. & MINDEN, J. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*, 18, 2071-7.
602. ÜNLÜ, M., MORGAN, M. & MINDEN, J. 1997. Difference Gel Electrophoresis: A single method for detecting changes in protein extracts. *Electrophoresis* 18, 2071-2077.
603. UTTAMSINGH, V., BAGGS, R. B., KRENITSKY, D. M. & ANDERS, M. 2000. Immunohistochemical localization of the acylases that catalyze the deacetylation of N-acetyl-L-cysteine and haloalkene-derived mercapturates. *Drug Metabolism and Disposition*, 28, 625-632.
604. VAILAYA, A. 2005. Fundamentals of reversed phase chromatography: Thermodynamic and exothermodynamic treatment. *Journal of Liquid Chromatography & Related Technologies* 28, 965-1054.
605. VAISAR, T., MAYER, P., NILSSON, E., ZHAO, X.-Q., KNOPP, R. & PRAZEN, B. J. 2010. HDL in humans with cardiovascular disease exhibits a proteomic signature. *Clinica Chimica Acta*, 411, 972-979.
606. VALAFAR, F. 2002. Pattern recognition techniques in microarray data analysis. *Annals of the New York Academy of Sciences*, 980, 41-64.
607. VAN DEEMTER, J. J., ZUIDERWEG, F. J. AND KLINKENBERG, A. 1959. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science* 5, 271-289.
608. VAN HOOF, D., PINKSE, M. W., WARD-VAN OOSTWAARD, D., MUMMERY, C. L., HECK, A. J. & KRIJGSVELD, J. 2007. An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. *Nature methods*, 4, 677-678.
609. VAN OUDENHOVE, L. & DEVREESE, B. 2013. A review on recent developments in mass spectrometry instrumentation and quantitative tools advancing bacterial proteomics. *Applied microbiology and biotechnology*, 97, 4749-4762.
610. VARGHESE, S., POWELL, T., JANECH, M., BUDISAVLJEVIC, M., STANISLAUS, R., ALMEIDA, J. & ARTHUR, J. 2010. Identification of diagnostic urinary biomarkers for acute kidney injury. *J Investig Med*, 58, 612-20.

611. VASSALOTTI, J., STEVENS, L. & LEVEY, A. 2007. Testing for chronic kidney disease: A position statement from the National Kidney Foundation. *Am J Kidney Dis.*, 50, 169-80.
612. VBLOKESHWAR, V. & SOLOWAY, M. 2001. Current bladder tumor tests: does their projected utility fulfill clinical necessity? *The Journal of urology*, 165, 1067-1077.
613. VEENSTRA, T. D. 2006. Global and targeted quantitative proteomics for biomarker discovery. *Journal of Chromatography B*, 847, 3-11.
614. VEENSTRA, T. D., CONRADS, T. P., HOOD, B. L., AVELLINO, A. M., ELLENBOGEN, R. G. & MORRISON, R. S. 2005. Biomarkers: mining the biofluid proteome. *Molecular & Cellular Proteomics*, 4, 409-418.
615. VENTER, J. C., ET AL., 2001. The sequence of the human genome. *Science*, 291, 1304-51.
616. VIVANCO, F., MARTÍN-VENTURA, J. L., DURAN, M. C., BARDERAS, M. G., BLANCO-COLIO, L., DARDÁ, V. N. M., MAS, S. N., MEILHAC, O., MICHEL, J. B., TUÑÓN, J. & EGIDO, J. S. 2005. Quest for Novel Cardiovascular Biomarkers by Proteomic Analysis. *Journal of Proteome Research*, 4, 1181-1191.
617. VLAHOU, A., CHARONIS, A. & BENIGNI, A. 2009. Report on the first combined working group and management committee meeting of EuroKUP (Urine and Kidney Proteomics cost action). *Journal of Proteomics*, 71, 682.
618. VOGEL, C. & MARCOTTE, E. M. 2008. Calculating absolute and relative protein abundance from mass spectrometry-based protein expression data. *Nature protocols*, 3, 1444-1451.
619. WAGGONER, A. 2006. Fluorescent labels for proteomics and genomics. *Current Opinion in Chemical Biology*, 10, 62-66.
620. WANG L, N. Z., XIE Z, YANG F, HE B, LIU J, DAI H, QIAN J, JIA M. 2010. Analysis of the urine proteome of human contrast-induced kidney injury using two-dimensional fluorescence differential gel electrophoresis/matrix-assisted laser desorption time-of-flight mass spectrometry/liquid chromatography mass spectrometry. *Am J Nephrol.*, 31, 45-52.
621. WANG, T., EVANS, J., MEIGS, J., RIFAI, N., FOX, C. & D'AGOSTINO, R. 2005. Low-grade albuminuria and the risks of hypertension and blood pressure progression - Response. *Circulation*, 112, E121.
622. WANG, T. J., GONA, P., LARSON, M. G., LEVY, D., BENJAMIN, E. J., TOFLER, G. H., JACQUES, P. F., MEIGS, J. B., RIFAI, N., SELHUB, J., ROBINS, S. J., NEWTON-CHEH, C. & VASAN, R. S. 2007. Multiple Biomarkers and the Risk of Incident Hypertension. *Hypertension*, 49, 432-438.
623. WANG, Y. T., TSAI, C. F., HONG, T. C., TSOU, C. C., LIN, P. Y., PAN, S. H., HONG, T. M., YANG, P. C., SUNG, T. Y. & HSU, W. L. 2010. An informatics-assisted label-free quantitation strategy that depicts phosphoproteomic profiles in lung cancer cell invasion. *Journal of Proteome Research*, 9, 5582-5597.
624. WANG, Z., ZAHEDI, K., BARONE, S., TEHRANI, K., RABB, H., MATLIN, K., CASERO, R. A. & SOLEIMANI, M. 2004. Overexpression of SSAT in Kidney Cells Recapitulates Various Phenotypic Aspects of Kidney Ischemia-reperfusion Injury. *Journal of the American Society of Nephrology*, 15, 1844-1852.
625. WARWICK, G., THOMAS, P. & YATES, D. 2008. Biomarkers in pulmonary hypertension. *Eur Respir J*, 32, 503-512.
626. WASHBURN, A., LUCHANSKY, M., BOWMAN, A. & BAILEY, R. 2010. Quantitative, label-free detection of five protein biomarkers using multiplexed arrays of silicon photonic microring resonators. *Anal Chem.*, 1, 69-72.

627. WASHBURN AL, G. L., BAILEY RC. 2009. Label-free quantitation of a cancer biomarker in complex media using silicon photonic microring resonators. *Anal Chem.* , 15, 9499-506.
628. WASHBURN, A. L., GUNN, L. C. & BAILEY, R. C. 2009. Label-free quantitation of a cancer biomarker in complex media using silicon photonic microring resonators. *Anal Chem*, 81, 9499-9506.
629. WATSON, J. D. A. F. H. C. 1953. . Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. . *Nature*, 171, 737-8.
630. WEBER, K. & OSBORN, M. 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J Biol Chem* 244, 4406-12.
631. WEBER, P. 1980. Renal prostaglandins, kidney function and essential hypertension. *Contributions to nephrology*, 23, 83.
632. WEISSINGER, E. M., KAISER, T., MEERT, N., DE SMET, R., WALDEN, M., MISCHAK, H. & VANHOLDER, R. C. 2004. Proteomics: a novel tool to unravel the patho-physiology of uraemia. *Nephrology Dialysis Transplantation*, 19, 3068-3077.
633. WENGE, B., BÖNISCH, H., GRABITZKI, J., LOCHNIT, G., SCHMITZ, B. & AHREND, M. 2008. Separation of membrane proteins by two-dimensional electrophoresis using cationic rehydrated strips. *Electrophoresis* 29, 1511-7.
634. WERNER DUBITZKY, M. G., DANIEL P. BERRAR` 2007. *Fundamentals of data mining in genomics and proteomics*, New York, Springer
635. WEST-NIELSEN, M., HØGDALL, E. V., MARCHIORI, E., HØGDALL, C. K., SCHOU, C. & HEEGAARD, N. H. 2005. Sample handling for mass spectrometric proteomic investigations of human sera. *Anal Chem*, 77, 5114-5123.
636. WHEELOCK AM, M. D., BARTOSIEWICZ M, BUCKPITT AR. 2006. Use of a fluorescent internal protein standard to achieve quantitative two-dimensional gel electrophoresis. *Proteomics* 6.
637. WHEELOCK, Å. M., MORIN, D., BARTOSIEWICZ, M. & BUCKPITT, A. R. 2006a. Use of a fluorescent internal protein standard to achieve quantitative two-dimensional gel electrophoresis. *PROTEOMICS*, 6, 1385-1398.
638. WHEELOCK, Å. M., MORIN, D., BARTOSIEWICZ, M. & BUCKPITT, A. R. 2006b. Use of a fluorescent internal protein standard to achieve quantitative two-dimensional gel electrophoresis. *PROTEOMICS*, 6, 1385-1398.
639. WHELTON, P. 2009. Hypertension in diabetes mellitus. *Endocrinol Nutr.*, 56, 63-6.
640. WHITE, M. Y., TCHEN, A. S., MCCARRON, H. C., HAMBLY, B. D., JEREMY, R. W. & CORDWELL, S. J. 2006. Proteomics of ischemia and reperfusion injuries in rabbit myocardium with and without intervention by an oxygen-free radical scavenger. *PROTEOMICS*, 6, 6221-6233.
641. WHITEAKER, J. R., LIN, C., KENNEDY, J., HOU, L., TRUTE, M., SOKAL, I., YAN, P., SCHOENHERR, R. M., ZHAO, L. & VOYTOVICH, U. J. 2011. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nature Biotechnology*, 29, 625-634.
642. WHITEAKER, J. R., ZHAO, L., ANDERSON, L. & PAULOVICH, A. G. 2010. An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. *Molecular & Cellular Proteomics*, 9, 184-196.

643. WIENKOOP S, L. E., NIEMANN M, GONZALEZ EM, LEHMANN U, WECKWERTH W. 2006. Stable isotope-free quantitative shotgun proteomics combined with sample pattern recognition for rapid diagnostics. *Journal of Separation Science* 29, 2793-801.
644. WILDGRUBER R, H. A., OBERMAIER C, BOGUTH G, WEISS W, FEY SJ, LARSEN PM, GORG A: 2000. Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. *Electrophoresis* 21, 2610-2616.
645. . 21, 2610-2616.
646. WILKINS, M. 2009. Hares and tortoises: the high- versus low-throughput proteomic race. *Electrophoresis* , 30, S150-5.
647. WILKINS, M., APPEL, R., HOCHSTRASSER, D. & AL, E. 1997. *Proteome research: new frontiers in functional genomics* , Berlin: Springer. .
648. WILSON, F. H., ET AL 2003. Molecular pathogenesis of inherited hypertension with hyperkalemia: the Na-Cl cotransporter is inhibited by wild-type but not mutant WNK4. *Proc.Natl.Acad.Sci.U.S.A*, 100, 680-684.
649. WILSON, F. H., ET AL 2001. Human hypertension caused by mutation in WNK kinases. *Science*, 293, 1107-1112.
650. WITZMANN, F., CLACK, J., FULTZ, C. & JARNOT, B. 1995. Two-dimensional electrophoretic mapping of hepatic and renal stress proteins. *Electrophoresis*, 16, 451-459.
651. WITZMANN FA, F. C., GRANT RA, WRIGHT & LS, K. S., SIEGEL FL: 1999. Regional protein alterations in rat kidneys induced by lead exposure. *Electrophoresis* 20, 943-951.
652. WOFFORD, M. & HALL, J. 2004. Pathophysiology and treatment of obesity hypertension. *Current Pharmaceutical Design* 10 3621-37.
653. WOLF-YADLIN, A., HAUTANIEMI, S., LAUFFENBURGER, D. A. & WHITE, F. M. 2007. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proceedings of the National Academy of Sciences*, 104, 5860-5865.
654. WU L, H. D. 2006. Overcoming the dynamic range problem in mass spectrometry-based shotgun proteomics. *Expert Rev Proteomics* , 3(611-9).
655. WU, W., HU, W. & KAVANAGH, J. 2002. Proteomics in cancer research. *International Journal of Gynecological Cancer*, 12, 409-423.
656. WU, W., W. HU, AND J.J. KAVANAGH, 2002. Proteomics in cancer research. *International Journal of Gynecological Cancer*, 12, 409-23.
657. WU, W. W., WANG, G., BAEK, S. J. & SHEN, R.-F. 2006a. Comparative Study of Three Proteomic Quantitative Methods, DIGE, cICAT, and iTRAQ, Using 2D Gel- or LC-MALDI TOF/TOF. *Journal of Proteome Research*, 5, 651-658.
658. WU, W. W., WANG, G., BAEK, S. J. & SHEN, R.-F. 2006b. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel-or LC-MALDI TOF/TOF. *Journal of Proteome Research*, 5, 651-658.

659. XIAO, H., ZHANG, L., ZHOU, H., LEE, J. M., GARON, E. B. & WONG, D. T. W. 2012. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Molecular & Cellular Proteomics*, 11.
660. XU, B., YOSHIDA, Y., ZHANG, Y., YAOITA, E., OSAWA, T. & YAMAMOTO, T. 2005a. Two-dimensional electrophoretic profiling of normal human kidney: differential protein expression in glomerulus, cortex and medulla. *Journal of Electrophoresis*, 49, 5-13.
661. XU, H., HU, L.-S., CHANG, M., JING, L., ZHANG, X.-Y. & LI, G. S. 2005b. Proteomic analysis of kidney in fluoride-treated rat. *Toxicology letters*, 160, 69-75.
662. XUE, J. L., DANIELS, F., STAR, R. A., KIMMEL, P. L., EGGERS, P. W., MOLITORIS, B. A., HIMMELFARB, J. & COLLINS, A. J. 2006. Incidence and Mortality of Acute Renal Failure in Medicare Beneficiaries, 1992 to 2001. *Journal of the American Society of Nephrology*, 17, 1135-1142.
663. YAMAGUCHI, M. 2000. The role of regucalcin in nuclear regulation of regenerating liver. *Biochemical and biophysical research communications*, 276, 1-6.
664. YAMAGUCHI, M. 2005. Role of regucalcin in maintaining cell homeostasis and function (review). *International journal of molecular medicine*, 15, 371.
665. YAMAMOTO, T., LANGHAM, R. G., RONCO, P., KNEPPER, M. A. & THONGBOONKERD, V. 2008. Towards standard protocols and guidelines for urine proteomics: a report on the Human Kidney and Urine Proteome Project (HKUPP) Symposium and Workshop 6 October 2007, Seoul, Korea and 1 November 2007, San Francisco, CA, USA. *PROTEOMICS*, 8, 2156-2159.
666. YAMORI, Y. 1994. *Development of the spontaneously hypertensive rat (SHR), the stroke-prone SHR (SHRSP) and their various substrain models for hypertension-related cardiovascular disease*, Elsevier, New York.
667. YANAGISAWA, K., SHYR, Y., XU, B., MASSION, P., LARSEN, P., WHITE, B., ROBERTS, J., EDGERTON, M., GONZALEZ, A., NADAF, S., MOORE, J., CAPRIOLI, R. & CARBONE, D. 2003. Proteomic patterns of tumor subsets in non-smallcell lung cancer. *Lancet* 362, 433-439.
668. YANG, L., RUDSER, K. D., HIGGINS, L. A., ROSEN, H. R., ZAMAN, A., CORLESS, C. L., DAVID, L. & GOURLEY, G. R. 2011a. Novel biomarker candidates to predict hepatic fibrosis in hepatitis c identified by serum proteomics. *Digestive diseases and sciences*, 56, 3305-3315.
669. YANG, N., FENG, S., SHEDDEN, K., XIE, X., LIU, Y., ROSSER, C. J., LUBMAN, D. M. & GOODISON, S. 2011b. Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification. *Clinical Cancer Research*, 17, 3349-3359.
670. YAO, X., AFONSO, C. & FENSELAU, C. 2003. Dissection of proteolytic 18O labeling: endoprotease-catalyzed 16O-to-18O exchange of truncated peptide substrates. *Journal of Proteome Research*, 2, 147-152.
671. YAO, X., BAJRAMI, B. & SHI, Y. 2010. Ultrathroughput Multiple Reaction Monitoring Mass Spectrometry. *Anal Chem*, 82, 794-797.
672. YE, H., GEMPERLINE, E. & LI, L. 2012a. A vision for better health: Mass spectrometry imaging for clinical diagnostics. *Clinica Chimica Acta*.

673. YE, H., GREER, T. & LI, L. 2012b. Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *Journal of Proteomics*, 75, 5014-5026.
674. YOSHIDA, Y., MIYAZAKI, K., KAMIIE, J., SATO, M., OKUIZUMI, S., KENMOCHI, A., KAMIJO, K. I., NABETANI, T., TSUGITA, A. & XU, B. 2005. Two-dimensional electrophoretic profiling of normal human kidney glomerulus proteome and construction of an extensible markup language (XML)-based database. *PROTEOMICS*, 5, 1083-1096.
675. YUSENKO, M., RUPPERT, T. & KOVACS, G. 2010. Analysis of differentially expressed mitochondrial proteins in chromophobe renal cell carcinomas and renal oncocytomas by 2-D gel electrophoresis. *Int J Biol Sci.*, 6, 213-24.
676. ZECH, H., ECHTERMEYER, C., WÖHLBRAND, L., BLASIUS, B. & RABUS, R. 2011. Biological versus technical variability in 2-D DIGE experiments with environmental bacteria. *PROTEOMICS*, 11, 3380-3389.
677. ZELLNER, M., VEITINGER, M. & UMLAUF, E. 2009. The role of proteomics in dementia and Alzheimer's disease. *Acta neuropathologica*, 118, 181-195.
678. ZHANG, B., BAREKATI, Z., KOHLER, C., RADPOUR, R., ASADOLLAHI, R., HOLZGREVE, W. & ZHONG, X. Y. 2010. Proteomics and Biomarkers for Ovarian Cancer Diagnosis. *Annals of Clinical & Laboratory Science*, 40, 218-225.
679. ZHANG, H., GO, Y.-M. & JONES, D. P. 2007a. Mitochondrial thioredoxin-2/peroxiredoxin-3 system functions in parallel with mitochondrial GSH system in protection against oxidative stress. *Archives of biochemistry and biophysics*, 465, 119-126.
680. ZHANG, J., GOODLETT, D. R. & MONTINE, T. J. 2005. Proteomic biomarker discovery in cerebrospinal fluid for neurodegenerative diseases. *Journal of Alzheimer's Disease*, 8, 377-386.
681. ZHANG, Q., TANG, N., BROCK, J. W., MOTTAZ, H. M., AMES, J. M., BAYNES, J. W., SMITH, R. D. & METZ, T. O. 2007b. Enrichment and analysis of nonenzymatically glycosylated peptides: boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry. *Journal of Proteome Research*, 6, 2323-2330.
682. ZHANG X, F. A., RILEY CP, WANG M, REGNIER FE, BUCK C. 2010. Multi-dimensional liquid chromatography in proteomics. *Anal Chim Acta.*, 664, 101-113.
683. ZHANG, X., LEUNG, S.-M., MORRIS, C. R. & SHIGENAGA, M. K. 2004a. Evaluation of a novel, integrated approach using functionalized magnetic beads, bench-top MALDI-TOF-MS with prestructured sample supports, and pattern recognition software for profiling potential biomarkers in human plasma. *Journal of biomolecular techniques: JBT*, 15, 167.
684. ZHANG, X., PETRUZZIELLO, F., ZANI, F., FOUILLEN, L., ANDREN, P. E., SOLINAS, G. & RAINER, G. 2012. High Identification Rates of Endogenous Neuropeptides from Mouse Brain. *Journal of Proteome Research*, 11, 2819-2827.
685. ZHANG, Y., WU, D., GUAN, M., LIU, W., WU, Z., CHEN, Y., ZHANG, W. & LU, Y. 2004b. Tree analysis of mass spectral urine profiles discriminates transitional cell carcinoma of the bladder from noncancer patient. *Clin Biochem.*, 37, 772-779.
686. ZHANG, Z., BAST, R. C., YU, Y., LI, J., SOKOLL, L. J., RAI, A. J., ROSENZWEIG, J. M., CAMERON, B., WANG, Y. Y. & MENG, X.-Y. 2004c. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Research*, 64, 5882-5890.
687. ZHAO, Y., LEE, W. N. P. & XIAO, G. G. 2009. Quantitative proteomics and biomarker discovery in human cancer. *Expert Review of Proteomics*, 6, 115.

688. ZHENG, Y., XU, Y., YE, B., LEI, J., WEINSTEIN, M. H., O'LEARY, M. P., RICHIE, J. P., MOK, S. C. & LIU, B. C. S. 2003. Prostate carcinoma tissue proteomics for biomarker discovery. *Cancer*, 98, 2576-2582.
689. ZHOU, H., YUEN, P., PISITKUN, T., GONZALES, P., YASUDA, H., DEAR, J., GROSS, P., KNEPPER, M. & STAR, R. 2006. Collection, storage, preservation, and normalization of human urinary exosomes for biomarker discovery. *Kidney International*, 69, 1471-1476.
690. ZHOU, M., T.P. CONRADS, AND T.D. VEENSTRA, 2005. Proteomics approaches to biomarker detection. . *Briefings in Functional Genomics & Proteomics*, 4, 69-75.
691. ZHU, H., SANTO, A. & LI, Y. 2012. The antioxidant enzyme peroxiredoxin and its protective role in neurological disorders. *Experimental Biology and Medicine*, 237, 143-149.
692. ZIMMERLI, L. U., SCHIFFER, E., ZÜRBIG, P., GOOD, D. M., KELLMANN, M., MOULS, L., PITT, A. R., COON, J. J., SCHMIEDER, R. E. & PETER, K. H. 2008a. Urinary proteomic biomarkers in coronary artery disease. *Molecular & Cellular Proteomics*, 7, 290-298.
693. ZIMMERLI, L. U., SCHIFFER, E., ZÜRBIG, P., GOOD, D. M., KELLMANN, M., MOULS, L., PITT, A. R., COON, J. J., SCHMIEDER, R. E., PETER, K. H., MISCHAK, H., KOLCH, W., DELLES, C. & DOMINICZAK, A. F. 2008b. Urinary Proteomic Biomarkers in Coronary Artery Disease. *Molecular & Cellular Proteomics*, 7, 290-298.
694. ZORIY, M., MATUSCH, A., SPRUSS, T. & BECKER, J. S. 2007. Laser ablation inductively coupled plasma mass spectrometry for imaging of copper, zinc, and platinum in thin sections of a kidney from a mouse treated with cis-platin. *International Journal of Mass Spectrometry*, 260, 102.
695. ZÜRBIG, P., DECRAMER, S., DAKNA, M., JANTOS, J., GOOD, D. M., COON, J. J., BANDIN, F., MISCHAK, H., BASCANDS, J.-L. & SCHANSTRA, J. P. 2009. The human urinary proteome reveals high similarity between kidney aging and chronic kidney disease. *PROTEOMICS*, 9, 2108-2117.

7 Appendices

7.1 Tables of Identifications

7.1.1 Hypertension pilot study identifications

Spot Master Number	Protein name	score	coverage %	matched ions
67	hemoglobin alpha 1 or 2 chain [Rattus norvegicus]	158	43	4
130	alpha-globin [Rattus sp.]	77	34	2
130	rCG34342, isoform CRA_b [Rattus norvegicus]	77	21	2
130	hemoglobin alpha 1 chain [Rattus norvegicus]	77	21	2
130	hemoglobin alpha 2 chain [Rattus norvegicus]	77	21	2
137	hemoglobin alpha 1 chain [Rattus norvegicus]	277	37	3
137	hemoglobin alpha 2 chain [Rattus norvegicus]	277	37	3
142	hemoglobin alpha 1 chain [Rattus norvegicus]	208	37	3
142	hemoglobin alpha 2 chain [Rattus norvegicus]	208	37	3
143	hemoglobin alpha 1 chain [Rattus norvegicus]	246	37	3
143	hemoglobin alpha 2 chain [Rattus norvegicus]	246	37	3
146	hemoglobin alpha 1 chain [Rattus norvegicus]	206	26	2
146	hemoglobin alpha 2 chain [Rattus norvegicus]	206	26	2
146	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric acid soluble protein	120	18	2
146	heat-responsive protein 12 [Rattus	120	18	2

	norvegicus			
146	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	120	18	2
146	hemoglobin alpha 1 chain [Rattus norvegicus]	244	37	3
146	hemoglobin alpha 2 chain [Rattus norvegicus]	244	37	3
149	alpha-globin [Rattus sp.]	109	34	2
149	rCG34342, isoform CRA_b [Rattus norvegicus]	109	21	2
149	hemoglobin alpha 1 chain [Rattus norvegicus]	109	21	2
149	hemoglobin alpha 2 chain [Rattus norvegicus]	109	21	2
154	hemoglobin alpha 1 chain [Rattus norvegicus]	264	37	3
154	hemoglobin alpha 2 chain [Rattus norvegicus]	264	37	3
188	alpha-globin [Rattus sp.]	134	34	2
188	rCG34342, isoform CRA_b [Rattus norvegicus]	134	21	2
188	hemoglobin alpha 1 chain [Rattus norvegicus]	134	21	2
188	hemoglobin alpha 2 chain [Rattus norvegicus]	134	21	2
188	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric acid soluble protein	125	40	4
188	heat-responsive protein 12 [Rattus norvegicus]	125	40	4
188	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	125	40	4
188	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric	223	59	5

	acid soluble protein			
188	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	223	59	5
201	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric acid soluble protein	143	40	2
201	heat-responsive protein 12 [Rattus norvegicus]	143	40	2
201	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	143	40	2
212	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric acid soluble protein	47	18	2
212	heat-responsive protein 12 [Rattus norvegicus]	47	18	2
212	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	47	18	2
228	hemoglobin alpha 1 chain [Rattus norvegicus]	135	37	3
228	hemoglobin alpha 2 chain [Rattus norvegicus]	135	37	3
234	D-dopachrome tautomerase [Rattus norvegicus]	51	35	3
234	D-dopachrome tautomerase, isoform CRA_b [Rattus norvegicus]	51	35	3
236	alpha-globin [Rattus sp.]	97	34	2
236	rCG34342, isoform CRA_b [Rattus norvegicus]	97	21	2
236	hemoglobin alpha 1 chain [Rattus norvegicus]	97	21	2
236	hemoglobin alpha 2 chain [Rattus norvegicus]	97	21	2
243	hemoglobin alpha 1 chain [Rattus norvegicus]	152	37	3

243	hemoglobin alpha 2 chain [Rattus norvegicus]	152	37	3
244	hemoglobin alpha 1 chain [Rattus norvegicus]	128	37	3
244	hemoglobin alpha 2 chain [Rattus norvegicus]	128	37	3
245	hemoglobin alpha 1 chain [Rattus norvegicus]	139	37	3
245	hemoglobin alpha 2 chain [Rattus norvegicus]	139	37	3
246	major beta-hemoglobin	100	34	4
246	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	100	34	4
246	beta-globin [Rattus norvegicus]	100	34	4
246	hemoglobin beta chain complex [Rattus norvegicus]	100	34	4
246	rCG39881, isoform CRA_a [Rattus norvegicus]	100	34	4
248	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	134	43	5
248	beta-globin [Rattus norvegicus]	134	42	5
250	alpha-globin [Rattus sp.]	156	34	2
250	rCG34342, isoform CRA_b [Rattus norvegicus]	156	21	2
250	hemoglobin alpha 1 chain [Rattus norvegicus]	156	21	2
250	hemoglobin alpha 2 chain [Rattus norvegicus]	156	21	2
258	alpha-globin [Rattus sp.]	104	34	2
258	rCG34342, isoform CRA_b [Rattus norvegicus]	104	21	2
258	hemoglobin alpha 1 chain [Rattus norvegicus]	104	21	2
258	hemoglobin alpha 2 chain [Rattus norvegicus]	104	21	2
267	major beta-hemoglobin	117	34	4

267	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	117	34	4
267	beta-globin [Rattus norvegicus]	117	34	4
267	hemoglobin beta chain complex [Rattus norvegicus]	117	34	4
267	rCG39881, isoform CRA_a [Rattus norvegicus]	117	34	4
278	alpha-globin [Rattus sp.]	102	34	2
278	rCG34342, isoform CRA_b [Rattus norvegicus]	102	21	2
278	hemoglobin alpha 1 chain [Rattus norvegicus]	102	21	2
278	hemoglobin alpha 2 chain [Rattus norvegicus]	102	21	2
279	alpha-globin [Rattus sp.]	60	34	2
279	rCG34342, isoform CRA_b [Rattus norvegicus]	60	21	2
279	hemoglobin alpha 1 chain [Rattus norvegicus]	60	21	2
279	hemoglobin alpha 2 chain [Rattus norvegicus]	60	21	2
283	histidine triad nucleotide binding protein 1 [Mus musculus]	110	35	3
283	PREDICTED: similar to Histidine triad nucleotide-binding protein 1 (Adenosine 5-monophosphoramidase) (Protein kinase C inhibitor 1) (Protein kinase C-interacting protein 1) (PKCI-1) [Rattus norvegicus]	110	23	3
295	major beta-hemoglobin	111	44	6
295	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	111	44	6
295	beta-globin [Rattus norvegicus]	111	44	6
295	hemoglobin beta chain complex [Rattus norvegicus]	111	44	6
295	rCG39881, isoform CRA_a [Rattus norvegicus]	111	44	6

296	major beta-hemoglobin	130	34	4
296	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	130	34	4
296	beta-globin [Rattus norvegicus]	130	34	4
296	hemoglobin beta chain complex [Rattus norvegicus]	130	34	4
296	rCG39881, isoform CRA_a [Rattus norvegicus]	130	34	4
302	alpha-globin [Rattus sp.]	87	34	2
302	rCG34342, isoform CRA_b [Rattus norvegicus]	87	21	2
302	hemoglobin alpha 1 chain [Rattus norvegicus]	87	21	2
302	hemoglobin alpha 2 chain [Rattus norvegicus]	87	21	2
303	major beta-hemoglobin	116	34	4
303	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	130	34	4
303	beta-globin [Rattus norvegicus]	130	34	4
303	hemoglobin beta chain complex [Rattus norvegicus]	130	34	4
303	rCG39881, isoform CRA_a [Rattus norvegicus]	130	34	4
315	alpha-globin [Rattus sp.]	93	34	2
315	rCG34342, isoform CRA_b [Rattus norvegicus]	93	21	2
315	hemoglobin alpha 1 chain [Rattus norvegicus]	93	21	2
315	hemoglobin alpha 2 chain [Rattus norvegicus]	93	21	2
319	PREDICTED: similar to germinal histone H4 gene [Rattus norvegicus]	66	28	3
324	alpha-globin [Rattus sp.]	140	34	2
324	rCG34342, isoform CRA_b [Rattus norvegicus]	140	21	2

324	hemoglobin alpha 1 chain [Rattus norvegicus]	140	21	2
324	hemoglobin alpha 2 chain [Rattus norvegicus]	140	21	2
325	alpha-globin [Rattus sp.]	148	34	2
325	rCG34342, isoform CRA_b [Rattus norvegicus]	148	21	2
325	hemoglobin alpha 1 chain [Rattus norvegicus]	148	21	2
325	hemoglobin alpha 2 chain [Rattus norvegicus]	148	21	2
335	alpha-globin [Rattus sp.]	37	34	2
335	rCG34342, isoform CRA_b [Rattus norvegicus]	37	21	2
335	hemoglobin alpha 1 chain [Rattus norvegicus]	37	21	2
335	hemoglobin alpha 2 chain [Rattus norvegicus]	37	21	2
337	hemoglobin alpha 1 chain [Rattus norvegicus]	47	37	3
337	hemoglobin alpha 2 chain [Rattus norvegicus]	47	37	3
342	alpha-globin [Rattus sp.]	66	34	2
342	rCG34342, isoform CRA_b [Rattus norvegicus]	66	21	2
342	hemoglobin alpha 1 chain [Rattus norvegicus]	66	21	2
342	hemoglobin alpha 2 chain [Rattus norvegicus]	66	21	2
352	PREDICTED: similar to Histidine triad nucleotide-binding protein 1 (Adenosine 5-monophosphoramidase) (Protein kinase C inhibitor 1) (Protein kinase C-interacting protein 1) (PKCI-1) [Rattus norvegicus]	77	10	2
355	hemoglobin alpha 1 chain [Rattus norvegicus]	56	37	3
355	hemoglobin alpha 2 chain [Rattus	56	37	3

	norvegicus]			
395	alpha-globin [Rattus sp.]	55	34	2
395	rCG34342, isoform CRA_b [Rattus norvegicus]	55	21	2
395	hemoglobin alpha 1 chain [Rattus norvegicus]	55	21	2
395	hemoglobin alpha 2 chain [Rattus norvegicus]	55	21	2
396	alpha-globin [Rattus sp.]	45	34	2
396	rCG34342, isoform CRA_b [Rattus norvegicus]	45	21	2
396	hemoglobin alpha 1 chain [Rattus norvegicus]	45	21	2
396	hemoglobin alpha 2 chain [Rattus norvegicus]	45	21	2
397	PREDICTED: similar to Histidine triad nucleotide-binding protein 1 (Adenosine 5-monophosphoramidase) (Protein kinase C inhibitor 1) (Protein kinase C-interacting protein 1) (PKCI-1) [Rattus norvegicus]	63	28	4
403	major beta-hemoglobin	91	43	5
403	beta 1 globin [rats, Sprague-Dawley, Peptide, 146 aa]	91	43	5
403	beta-globin [Rattus norvegicus]	91	43	5
403	hemoglobin beta chain complex [Rattus norvegicus]	91	43	5
403	rCG39881, isoform CRA_a [Rattus norvegicus]	91	43	5
404	alpha-globin [Rattus sp.]	78	44	3
404	rCG34342, isoform CRA_b [Rattus norvegicus]	78	28	3
404	hemoglobin alpha 1 chain [Rattus norvegicus]	78	28	3
404	hemoglobin alpha 2 chain [Rattus norvegicus]	78	28	3
432	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational	109	18	2

	inhibitor protein; AltName: Full=Perchloric acid soluble protein			
432	heat-responsive protein 12 [Rattus norvegicus]	109	18	2
432	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	109	18	2
433	RecName: Full=Ribonuclease UK114; AltName: Full=14.5 kDa translational inhibitor protein; AltName: Full=Perchloric acid soluble protein	47	18	2
433	heat-responsive protein 12 [Rattus norvegicus]	47	18	2
433	Chain A, Crystal Structure Of Perchloric Acid Soluble Protein-A Translational Inhibitor	47	18	2
473	alpha-globin [Rattus sp.]	60	34	2
473	rCG34342, isoform CRA_b [Rattus norvegicus]	60	21	2
473	hemoglobin alpha 1 chain [Rattus norvegicus]	60	21	2
473	hemoglobin alpha 2 chain [Rattus norvegicus]	60	21	2
486	RecName: Full=Transthyretin; AltName: Full=Prealbumin; AltName: Full=TBPA; Flags: Precursor	47	18	1
486	Chain A, Rat Transthyretin	47	18	1
486	transthyretin [Rattus norvegicus]	47	18	1
486	Chain A, Rat Transthyretin Complex With Thyroxine (T4)	68	18	1
510	fatty acid binding protein 3, muscle and heart [Rattus norvegicus]	64	60	14
515	PREDICTED: similar to SH3 domain-binding glutamic acid-rich-like protein [Rattus norvegicus]	46	2	1
515	SH3-binding domain glutamic acid-rich protein like (predicted), isoform CRA_a [Rattus norvegicus]	46	2	1

515	Cytochrome c oxidase, subunit Va [Rattus norvegicus]	45	17	3
524	Cytochrome c oxidase, subunit Va [Rattus norvegicus]	38	17	3
627				
627	aldolase 3, C isoform, isoform CRA_e [Mus musculus]	64	68	20
726	Cu/Zn-superoxide dismutase	78	8	1
727	myosin, light polypeptide 6, alkali, smooth muscle and non-muscle [Mus musculus]	56	20	3
731	mCG140959, isoform CRA_g [Mus musculus]	63	48	15
739	Cu/Zn-superoxide dismutase	110	8	15
752	myosin, light polypeptide 6, alkali, smooth muscle and non-muscle [Mus musculus]	57	20	3
759	peroxisomal membrane protein 20 [Mus musculus]	55	8	1
1618	peroxiredoxin 1 [Mus musculus]	59	10	2
1600	superoxide dismutase 2 [Rattus norvegicus]	60	6	1
1621	subunit d of mitochondrial H-ATP synthase [Rattus norvegicus]	42	15	2
1621	ATP synthase, H ⁺ transporting, mitochondrial F ₀ complex, subunit d [Rattus norvegicus]	42	15	2
1621	rCG33654, isoform CRA_b [Rattus norvegicus]	42	15	2
1832	subunit d of mitochondrial H-ATP synthase [Rattus norvegicus]	128	24	3
1832	ATP synthase, H ⁺ transporting, mitochondrial F ₀ complex, subunit d [Rattus norvegicus]	128	24	3
1637	subunit d of mitochondrial H-ATP synthase [Rattus norvegicus]	99	55	13
1637	ATP synthase, H ⁺ transporting, mitochondrial F ₀ complex, subunit d [Rattus norvegicus]	93	55	12

1647	adenine phosphoribosyl transferase [Rattus norvegicus]	56	28	4
1647	adenine phosphoribosyl transferase (predicted), isoform CRA_a [Rattus norvegicus]	56	28	4
1649	phosphatidylethanolamine binding protein [Rattus norvegicus]	69	18	2
1649	phosphatidylethanolamine binding protein 1, isoform CRA_b [Rattus norvegicus]	69	18	2
1649	Chain A, Rat Phosphatidylethanolamine-Binding Protein Containing The S153e Mutation In The Complex With O-Phosphorylethanolamine	69	18	2
1649	Chain A, Rat Phosphatidylethanolamine-Binding Protein	69	18	2
1650	NADH dehydrogenase (ubiquinone) Fe-S protein 8 [Rattus norvegicus]	36	17	3
1721	catechol-O-methyltransferase [Rattus norvegicus]	67	35	13
1776	plasma glutathione peroxidase precursor [Rattus norvegicus]	41	11	2
1776	glutathione peroxidase 3 precursor [Rattus norvegicus]	41	11	2
1734	PREDICTED: similar to Glutathione S-transferase alpha-4 (Glutathione S-transferase Yk) (GST Yk) (GST 8-8) (GST K) (GST A4-4) [Rattus norvegicus	49	12	3
1734	rCG25753, isoform CRA_b [Rattus norvegicus]	49	12	3
1734	glutathione S-transferase A4 [Rattus norvegicus]	49	12	3
1754	PREDICTED: similar to Glutathione S-transferase alpha-4 (Glutathione S-transferase Yk) (GST Yk) (GST 8-8) (GST K) (GST A4-4) [Rattus norvegicus	63	9	2
1754	rCG25753, isoform CRA_b [Rattus norvegicus]	63	9	2
1754	glutathione S-transferase A4 [Rattus norvegicus]	63	9	2

1811	methionine sulfoxide reductase A [Rattus norvegicus]	51	18	3
1811	methionine sulfoxide reductase A splice variant 2a [Rattus norvegicus]	51	18	3
1839	peroxiredoxin 3 [Rattus norvegicus]	40	14	3
1839	RecName: Full=Thioredoxin-dependent peroxide reductase, mitochondrial; AltName: Full=Peroxiredoxin-3; Short=PRX-3; AltName: Full=PRx III; Flags: Precursor	40	14	3
1839	peroxiredoxin 3 [Rattus norvegicus]	40	14	3
1849	rCG42432, isoform CRA_d [Rattus norvegicus]	62	42	21
1884	glutathione S-transferase A2	55	8	2
2049	rCG23467, isoform CRA_a [Rattus norvegicus]	60	22	52
2052	rCG23467, isoform CRA_a [Rattus norvegicus]	63	22	53
2054	rCG23467, isoform CRA_a [Rattus norvegicus]	61	22	53
2055	rCG23467, isoform CRA_a [Rattus norvegicus]	66	22	47
2057	rCG23467, isoform CRA_a [Rattus norvegicus]	76	25	48
2060	rCG23467, isoform CRA_a [Rattus norvegicus]	83	26	56
2072	phosphoglycerate mutase 1 [Mus musculus]	79	22	4
2080	rCG23467, isoform CRA_a [Rattus norvegicus]	65	24	51
2093	rCG23467, isoform CRA_a [Rattus norvegicus]	64	22	52
2106	electron-transfer-flavoprotein, beta polypeptide [Rattus norvegicus]	30	22	4
2118	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	51	33	5
2120	carbonic anhydrase 1 [Rattus norvegicus]	82	67	15

2126	glutathione S-transferase A5 [Rattus norvegicus]	64	20	11
2137	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	46	19	3
2156	rCG23467, isoform CRA_a [Rattus norvegicus]	67	24	53
2164	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide [Rattus norvegicus]	70	9	2
2168	glutathione S-transferase A5 [Rattus norvegicus]	63	17	4
2174	glutathione S-transferase omega 1 [Rattus norvegicus]	63	12	3
2216	PREDICTED: similar to ATP-binding cassette, sub-family A, member 12 isoform a [Rattus norvegicus]	38	0	1
2216	rCG25036 [Rattus norvegicus]	38	0	1
2226	cathepsin A [Rattus norvegicus]	69	2	1
2226	rCG32401, isoform CRA_a [Rattus norvegicus]	69	2	1
2226	rCG32401, isoform CRA_a [Rattus norvegicus]	69	2	1
2241	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein [Rattus norvegicus]	39	11	3
2251	rCG23467, isoform CRA_a [Rattus norvegicus]	66	24	54
2252	PREDICTED: similar to ATP-binding cassette, sub-family A, member 12 isoform a [Rattus norvegicus]	38	0	1
2252	rCG25036 [Rattus norvegicus]	38	0	1
2254	rCG23467, isoform CRA_a [Rattus norvegicus]	67	25	51
2262	rCG23467, isoform CRA_a [Rattus norvegicus]	71	23	52
2298	glutamate Cysteine ligase, modifier subunit [Rattus norvegicus]	82	9	2

2303	rCG23467, isoform CRA_a [Rattus norvegicus]	63	24	51
2307	rCG23467, isoform CRA_a [Rattus norvegicus]	67	23	54
2309	rCG23467, isoform CRA_a [Rattus norvegicus]	63	21	52
2374	rCG23467, isoform CRA_a [Rattus norvegicus]	63	21	53
2404	rCG23467, isoform CRA_a [Rattus norvegicus]	67	24	52
2407	3-hydroxyanthranilate 3,4-dioxygenase [Rattus norvegicus]	43	12	2
2423	3-hydroxyanthranilate 3,4-dioxygenase, isoform CRA_b [Rattus norvegicus]	104	58	25
2444	3-hydroxyanthranilate 3,4-dioxygenase, isoform CRA_b [Rattus norvegicus]	74	49	22
2472	3-mercaptopyruvate sulfurtransferase [Rattus norvegicus]	182	19	5
2475	rCG23467, isoform CRA_a [Rattus norvegicus]	66	24	50
2502	3-hydroxyanthranilate 3,4-dioxygenase [Rattus norvegicus]	45	12	2
2685	aldo-keto reductase family 7, member A2 (aflatoxin aldehyde reductase) [Rattus norvegicus]	86	16	4
2685	aldehyde reductase AFAR2 subunit [Rattus norvegicus]	86	16	4
2685	aflatoxin B1 aldehyde reductase [Rattus norvegicus]	86	16	4
2690	crystallin, zeta [Rattus norvegicus]	65	13	3
2709	crystallin, zeta [Rattus norvegicus]	43	13	3
2710	hypothetical protein LOC293949 [Rattus norvegicus]	42	4	1
2715	regucalcin [Rattus norvegicus]	62	15	3
2730	aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus]	252	32	9

2734	L-lactate dehydrogenase B [Rattus norvegicus]	34	16	5
2742	L-lactate dehydrogenase B [Rattus norvegicus]	55	12	4
2746	crystallin, zeta [Rattus norvegicus]	82	13	3
2751	rCG23467, isoform CRA_a [Rattus norvegicus]	62	22	47
2796	dimethylarginine dimethylaminohydrolase 1 [Rattus norvegicus]	69	56	22
2772	glyceraldehyde-3-phosphate dehydrogenase [Rattus norvegicus]	157	20	5
2779	glyceraldehyde-3-phosphate dehydrogenase [Rattus norvegicus]	116	12	3
2781	glyceraldehyde-3-phosphate dehydrogenase [Rattus norvegicus]	64	12	3
2788	glyceraldehyde-3-phosphate dehydrogenase [Rattus norvegicus]	41	12	3
2789	hydroxyacid oxidase 3 (medium-chain) [Rattus norvegicus]	96	25	8
2806	aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus]	194	12	4
2813	Chain A, Crystal Structure Analysis Of Recombinant Rat Kidney Long- Chain Hydroxy Acid Oxidase	64	29	13
2815	aldo-keto reductase family 1, member A1 (aldehyde reductase) [Rattus norvegicus]	68	12	25
2816	hydroxyacid oxidase 3 (medium-chain) [Rattus norvegicus]	44	15	4
2824	PREDICTED: hypothetical protein [Rattus norvegicus]	62	76	9
2841	hydroxyacid oxidase 3 (medium-chain) [Rattus norvegicus]	68	15	4
2948	aldolase A	52	6	2
2954	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	109	12	5
2958	aldolase A	68	6	2

2962	aldolase B [Rattus norvegicus]	105	42	23
2967	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	147	18	7
2974	rCG25777, isoform CRA_a [Rattus norvegicus]	56	19	6
2986	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	89	10	4
2992	aldolase A	86	35	16
3002	aminoacylase 1 [Rattus norvegicus]	45	14	5
3009	RecName: Full=Aminoacylase-1A; AltName: Full=N-acyl-L-amino-acid amidohydrolase; AltName: Full=ACY-1A; AltName: Full=ACY IA	122	47	24
3191	Enolase 1, alpha non-neuron [Rattus norvegicus]	41	13	4
3191	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	41	13	4
3191	enolase 1, (alpha) [Rattus norvegicus]	41	13	4
3191	Eno1 protein [Rattus norvegicus]	41	13	4
3195	L-arginine:glycine amidinotransferase [Rattus norvegicus]	64	32	23
3213	Eno1 protein [Rattus norvegicus]	50	12	4
3213	Enolase 1, alpha non-neuron [Rattus norvegicus]	50	12	4
3213	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	50	12	4
3213	enolase 1, (alpha) [Rattus norvegicus]	50	12	4
3214	mitochondrial aldehyde dehydrogenase [Rattus norvegicus]	75	28	17
3222	PREDICTED: similar to sarcoma antigen NY-SAR-41 [Rattus norvegicus]	64	27	43
3231	ATP synthase beta subunit	249	53	35

3240	ATP synthase beta subunit	188	53	10
3258	mitochondrial aldehyde dehydrogenase 2 [Rattus norvegicus]	52	2	1
3259	Enolase 1, alpha non-neuron [Rattus norvegicus]	42	4	1
3261	mitochondrial aldehyde dehydrogenase 2 [Rattus norvegicus]	65	31	17
3262	Eno1 protein [Rattus norvegicus]	55	3	1
3315	ERM-binding phosphoprotein [Rattus norvegicus]	82	64	21
3319	vacuolar H ⁺ ATPase B2 [Rattus norvegicus]	52	7	3
3333	ATP synthase beta subunit	188	22	9
3341	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	53	8	4
3343	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	74	36	31
3353	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	74	36	31
3356	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	52	8	3
3360	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	74	36	31
3361	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	87	33	25
3365	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_g [Rattus norvegicus]	64	11	2

3367	liver annexin-like protein [Rattus norvegicus]	59	2	1
3367	plasma glutamate carboxypeptidase [Rattus norvegicus]	59	2	1
3371	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_g [Rattus norvegicus]	71	11	2
3372	selenium binding protein 1 [Rattus norvegicus]	113	37	23
3383	aldehyde dehydrogenase family 6, subfamily A1 [Rattus norvegicus]	50	3	1
3424	ATP synthase beta subunit	86	12	4
3426	tubulin, alpha 1C [Rattus norvegicus]	82	14	19
3433	glucose-6-phosphate dehydrogenase [Rattus norvegicus]	52	8	4
3437	tubulin, alpha 1C [Rattus norvegicus]	94	14	4
3440	tubulin T beta15	57	9	4
3453	tubulin, alpha 1C [Rattus norvegicus]	56	13	4
3459	heat shock protein (hsp60) precursor [Rattus norvegicus]	45	6	2
3461	heat shock protein (hsp60) precursor [Rattus norvegicus]	45	6	2
3470	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_f [Rattus norvegicus]	65	41	27
4908	PREDICTED: similar to IQ motif containing GTPase activating protein 3 [Rattus norvegicus]	65	15	27

Table 7-1: Complete table of identification made for the pilot investigation.

7.1.2 Hypertension main study identifications

master spot no	name	score	gi number
531	grp75	36	gi 1000439
566	heat shock protein 5 [Rattus norvegicus]	59	gi 25742763
566	RecName: Full=78 kDa glucose-regulated protein; AltName: Full=GRP 78; AltName: Full=Heat shock 70 kDa protein 5; AltName: Full=Immunoglobulin heavy chain-binding protein; Short=BiP; Flags: Precursor	59	gi 25742763
566	BiP [Mus musculus]	59	gi 25742763
629	dnaK-type molecular chaperone hsp72-ps1 - rat	81	gi 347019
629	heat shock protein 8 [Rattus norvegicus]	81	gi 13242237
637	albumin, isoform CRA_a [Rattus norvegicus]	148	gi 149033753
637	albumin [Rattus norvegicus]	146	gi 158138568
637	leucine rich repeat containing 7 [Rattus norvegicus]	33	gi 16924000
644	dnaK-type molecular chaperone hsp72-ps1 - rat	67	gi 347019
644	heat shock protein 8 [Rattus norvegicus]	67	gi 13242237
649	albumin, isoform CRA_a [Rattus norvegicus]	198	gi 149033753
649	albumin [Rattus norvegicus]	198	gi 158138568
682	albumin, isoform CRA_a [Rattus norvegicus]	252	gi 149033753

682	albumin [Rattus norvegicus]	252	gi 158138568
901	heat shock protein (hsp60) precursor [Rattus norvegicus]	56	gi 56383
731	Hspd1 protein [Mus musculus]	84	gi 76779273
956	heat shock protein (hsp60) precursor [Rattus norvegicus]	83	gi 56383
227	aspartyl-tRNA synthetase [Rattus norvegicus]	86	gi 16758642
566	plate 1 1st run	103	gi 25742763
956	heat shock protein (hsp60) precursor [Rattus norvegicus]	40	gi 56383
458	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 isoform 1 [Rattus norvegicus]	60	gi 218931161
560	heat shock protein 5 [Rattus norvegicus]	102	gi 25742763
682	albumin [Rattus norvegicus]	84	gi 158138568

1060	Tubulin, alpha 1A [Rattus norvegicus]	43	gi 38328248
1060	tubulin, alpha 1C [Rattus norvegicus]		gi 58865558
1060	rCG50513, isoform CRA_a [Rattus norvegicus]		gi 149032103
1060	rCG50513, isoform CRA_b [Rattus norvegicus]		gi 149032104
1122	CNDP dipeptidase 2 [Rattus norvegicus]	71	gi 58219062
1157	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, isoform CRA_d [Rattus norvegicus]	62	gi 149029483
1157	Chain A, Rat Liver F1-Atpase	61	gi 6729934
1161	mitochondrial aldehyde dehydrogenase [Rattus norvegicus]	69	gi 25990263
1172	mitochondrial aldehyde dehydrogenase 2 [Rattus norvegicus]	34	gi 14192933
1177	Chain B, Rat Liver F1-Atpase	122	gi 6729935
1177	mitochondrial ATP synthase beta subunit [Rattus norvegicus]	122	gi 54792127
1177	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide, isoform CRA_a [Rattus norvegicus]		gi 54792127
1178	Chain B, Rat Liver F1-Atpase	122	gi 6729935
1178	mitochondrial ATP synthase beta subunit [Rattus norvegicus]	122	gi 54792127
1178	ATP synthase, H+ transporting, mitochondrial F1		gi 54792127

	complex, beta polypeptide, isoform CRA_a [Rattus norvegicus]		
1179	aldehyde dehydrogenase 9A1 [Rattus norvegicus]	62	gi 75905479
1189	Chain B, Rat Liver F1-Atpase	122	gi 6729935
1189	mitochondrial ATP synthase beta subunit [Rattus norvegicus]	122	gi 54792127
1189	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide, isoform CRA_a [Rattus norvegicus]		gi 54792127
1243	IMP3, U3 small nucleolar ribonucleoprotein, homolog [Rattus norvegicus]	66	gi 157822893
1272	Eno1 protein [Rattus norvegicus]	63	gi 38649320
1272	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	63	gi 109468300
1272	enolase 1, (alpha) isoform 1 [Rattus norvegicus]	63	gi 158186649
1274	Eno1 protein [Rattus norvegicus]	68	gi 38649320
1274	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	68	gi 109468300
1274	enolase 1, (alpha) isoform 1 [Rattus norvegicus]	68	gi 158186649
1290	Eno1 protein [Rattus norvegicus]	60	gi 38649320
1290	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	60	gi 109468300

1290	enolase 1, (alpha) isoform 1 [Rattus norvegicus]	60	gi 158186649
1313	enolase 1, (alpha) isoform 1 [Rattus norvegicus]	138	gi 158186649
1523	PREDICTED: hypothetical protein [Rattus norvegicus]	63	gi 109457572
1548	actin, gamma 1 propeptide [Homo sapiens]	62	gi 4501887
1550	rCG25777, isoform CRA_a [Rattus norvegicus]	267	gi 149018671
1550	aminoacylase 1 [Rattus norvegicus]	238	gi 52851387
1561	aminoacylase 1 [Rattus norvegicus]	31	gi 52851387
1561	rCG25777, isoform CRA_a [Rattus norvegicus]	265	gi 149018671
1561	aminoacylase 1 [Rattus norvegicus]	245	gi 52851387
1576	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	77	gi 109492380
1578	rCG25777, isoform CRA_a [Rattus norvegicus]	121	gi 149018671
1578	aminoacylase 1 [Rattus norvegicus]	105	gi 52851387
1590	beta-actin FE-3 [Rattus norvegicus]	61	gi 13516471
1591	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	122	gi 109492380
1591	actin, beta-like 2 [Rattus norvegicus]	86	gi 157823033

1599	REDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	60	gi 109492380
1599	actin, beta-like 2 [Rattus norvegicus]	59	gi 157823033
1603	PREDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	72	gi 109492380
1604	REDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	105	gi 109492380
1604	actin, beta-like 2 [Rattus norvegicus]	60	gi 157823033
1613	REDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	132	gi 109492380
1613	actin, beta-like 2 [Rattus norvegicus]	70	gi 157823033
1627	REDICTED: similar to Actin, Cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]	60	gi 109492380
1670	fructose-1,6-bisphosphatase 1 [Rattus norvegicus]	132	gi 51036635
1893	dimethylarginine dimethylaminohydrolase 1 [Rattus norvegicus]	110	gi 11560131
1893	dimethylarginine dimethylaminohydrolase 1, isoform CRA_b [Rattus norvegicus]	93	gi 149026166
1346	mCG6336 [Mus musculus]	63	gi 148706088
1444	aldo-keto reductase family 1, member D1 [Mus musculus]	65	gi 21703734

1265	Eno1 protein [Rattus norvegicus]	79	gi 38649320
1265	PREDICTED: similar to Alpha-enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) [Rattus norvegicus]	79	gi 109468300
1265	enolase 1, (alpha) isoform 1 [Rattus norvegicus]	79	gi 158186649
1374	ubiquinol-Cytochrome c reductase core protein I [Rattus norvegicus]	73	gi 51948476
1487	uracil-DNA glycosylase, isoform CRA_a [Rattus norvegicus]	61	gi 149063625
1518	beta-actin FE-3 [Rattus norvegicus]	33	gi 13516471
3049	PREDICTED: hypothetical protein [Rattus norvegicus]	64	gi 109457572
3169	mitochondrial ribosomal protein L51 [Rattus norvegicus]	60	gi 157820005
2058	regucalcin [Rattus norvegicus]	77	gi 408807
			gi 13928740
2068	aspartoacylase 3 [Rattus norvegicus]	72	gi 57526957
2187	3-mercaptopyruvate sulfurtransferase [Rattus norvegicus]	97	gi 20304123
2197	3-hydroxyanthranilate 3,4-dioxygenase [Rattus norvegicus]	36	gi 4433351

2316	lipocortin V [Rattus norvegicus]	80	gi 2981437
2316	Annexin A5 [Rattus norvegicus]	80	gi 51858950
2316	Chain A, Structure Of A Mutant Of Rat Annexin A5	80	gi 150261262
2316	Chain A, Recombinant Rat Annexin V, W185a Mutant	80	gi 157830216
2316	Chain A, Recombinant Rat Annexin V, Quadruple Mutant (T72k, S144k, S228k, S303k)	80	gi 157830217
2316	Chain A, Recombinant Rat Annexin V, Triple Mutant (T72k, S144k, S228k)	80	gi 157830218
2316	Chain A, Recombinant Rat Annexin V, T72a Mutant	80	gi 157830229
2316	Chain A, Recombinant Rat Annexin V, T72k Mutant	80	gi 157830231
2316	Chain A, Recombinant Rat Annexin V, T72s Mutant	80	gi 157830232
2316	Chain A, Rat Annexin V Crystal Structure: Ca ²⁺ -Induced Conformational Changes	80	gi 157836327
2322	Chain A, Rat Annexin V Crystal Structure: Ca ²⁺ -Induced Conformational Changes	210	gi 157836327
2322	lipocortin V [Rattus norvegicus]	209	gi 2981437
2405	tropomyosin 4 [Rattus norvegicus]	64	gi 6981672
2465	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein [Rattus norvegicus]	43	gi 13928824
2544	chloride intracellular channel 1 [Rattus norvegicus]	107	gi 50657380
2565	mitochondrial ribosomal protein L51 [Rattus	75	gi 157820005

	norvegicus]		
2442	RecName: Full=Uracil-DNA glycosylase; Short=UDG	62	gi 45593577
2491	glutamate Cysteine ligase, modifier subunit [Rattus norvegicus]	65	gi 8393446
2506	mitochondrial ribosomal protein L51 [Rattus norvegicus]	63	gi 157820005
2039	PREDICTED: hypothetical protein [Rattus norvegicus]	61	gi 109496540
2612	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	70	gi 31982030
2709	14-3-3 zeta isoform	40	gi 1051270
2709	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide [Mus musculus]	40	gi 6756041
2898	synaptonemal complex protein 2 [Rattus norvegicus]	57	gi 18543333
2898	regulating synaptic membrane exoCytosis 3 [Rattus norvegicus]	57	gi 12621090
2961	gamma-glutamyl Cyclotransferase [Rattus norvegicus]	72	gi 157820337

2620	isoamyl acetate-hydrolyzing esterase 1 homolog [Rattus norvegicus]	62	gi 198278545
2635	uracil-DNA glycosylase, isoform CRA_a [Rattus norvegicus]	62	gi 149063625
2657	uracil-DNA glycosylase, isoform CRA_a [Rattus norvegicus]	60	gi 149063625
2980	hypothetical protein LOC619574 [Rattus norvegicus]	62	gi 77993361

Table 7-2: Table of all identified spots in main experiment

7.3 Poster Presentation: HUPO 2008, Amsterdam

