

THE UNIVERSITY OF HULL

**The Identification of Biomarkers of Chemotherapy
Resistance in Breast Cancer using Comparative
Proteomics**

Being a thesis submitted for the degree of

Doctor of Philosophy

at the University of Hull by

Victoria C Hodgkinson *MSc. BSc (Hons)*

September 2011

Acknowledgements

First of all I would like to thank my supervisors Dr Lynn Cawkwell and Professor Mike Lind, for all their support and guidance throughout, and for making this project possible.

I would also like to thank Dalia ELFadl for her involvement in the collection of the clinical tissue samples and clinical information. Also thanks to Bilal Elahi and all the surgeons at the Breast Unit for the collection of the clinical tumour samples. Thank you to Dr Vijay Agarwal for his clinical involvement in the project, and for his support and friendship throughout. Thank you also to Tasadooq Hussain for his help with core biopsy samples. I would also like to thank all the patients who agreed to take part in the study.

Thank you to Dr Gina Eagle and Dr James Bailey for the teaching of techniques and discussing methodologies, as well as their friendship. Thank you to Adam Dowle at the University of York for all his help with the MALDI-TOF/TOF MS. A special thank you must also go to all my colleagues and friends at the Daisy Building, especially Lucy Scaife, Sam Drennan, Katie Wraith and Sian Leech; for all their friendship and support, for our 12-o'clock lunches and for making the lab a happy place to work.

Thank you to my parents and sister for their never-ending support. To my friends, especially Nina, for always being there, for all their support and encouragement and for snapping me back into reality! Finally, a massive thank you must go to Paul, for the huge amount of support and encouragement he has given me – I cannot thank you enough.

Abstract

Background:

Chemotherapy resistance is a major obstacle in effective neoadjuvant treatment for locally advanced breast cancer. The ability to predict tumour response would allow chemotherapy administration to be directed towards only those patients who would benefit, thus maximising treatment efficiency. This project aimed to identify predictive protein biomarkers associated with chemotherapy resistance, using proteomic analysis of fresh breast cancer tissue samples.

Materials and Methods:

Chemotherapy-sensitive (CS) and chemotherapy-resistant (CR) tumour samples were collected from breast cancer patients who received neoadjuvant therapy consisting of epirubicin with cyclophosphamide followed by docetaxel. Comparative proteomic analysis was performed, to identify differentially expressed proteins (DEPs) between CS and CR invasive ductal carcinoma samples, using 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) with MALDI-TOF/TOF mass spectrometry and antibody microarray analysis. DEPs were submitted to Ingenuity Pathway Analysis (IPA) to identify any canonical pathway links, confirmed using western blotting and clinically validated in a pilot series of archival breast cancer samples, from patients treated with neoadjuvant chemotherapy.

Results:

Five datasets were generated by antibody microarray analysis, revealing 38 targets. Of these, 7 DEPs were identified in at least 2 datasets and these included 14-3-3 theta/tau, BID and Bcl-xL. Three datasets were generated using 2D-PAGE with MALDI-TOF/TOF MS, containing 132 unique DEPs. These included several isoforms of 14-3-3 proteins. The differential expression of 14-3-3, BID and Bcl-xL was confirmed by immunoblotting in samples used for the discovery phase. Clinical validation using immunohistochemical analysis of archival breast cancers revealed 14-3-3 theta/tau and tBID to be significantly associated with chemotherapy resistance.

Discussion:

The use of comparative proteomic techniques using fresh clinical tumour samples, for the search for putative biomarkers of chemotherapy resistance has been successful. Two DEPs; 14-3-3 theta/tau and tBID have passed through all stages of the biomarker discovery pipeline, and present themselves as putative predictive biomarkers of neoadjuvant chemotherapy resistance in breast cancer.

Publications

V Hodgkinson, D ELFadl, C Russell, V Agarwal, V Garimella, P Drew, M J Lind and L Cawkwell. Predictive biomarkers of chemotherapy resistance in breast cancer: a possible role for 14-3-3 tau and BID? – *revisions submitted*

L Scaife, **V C Hodgkinson**, P J Drew, M J Lind, L Cawkwell. (2011) Differential proteomics in the search for biomarkers of radiotherapy resistance. *Expert Review of Proteomics*, 8 (4) 535-52

D ELFadl, **VC Hodgkinson**, ED Long, L Scaife, PJ Drew, MJ Lind, L Cawkwell (2011). A pilot study to investigate the role of the 26S proteasome in radiotherapy resistance and loco-regional recurrence following breast conserving therapy for early breast cancer. *The Breast*, 20 (4) 334-7

V C Hodgkinson, D ELFadl, P J Drew, M J Lind, L Cawkwell. (2011) Repeatedly identified differentially expressed proteins (RIDEPs) from antibody microarray proteomic analysis. *Journal of Proteomics*; 74:698-703.

V C Hodgkinson, G L Eagle, P J Drew, M J Lind, L Cawkwell. (2010) Biomarkers of chemotherapy resistance in breast cancer identified by proteomics: current status. *Cancer Letters*; 294: 13-24

S Mehmood, **VC Hodgkinson**, L Scaife, GP Liney, MP Bush, AW Beavis, IA Hunter, L Cawkwell. Proteomic identification of putative biomarkers of radioresistance in rectal cancer. *British Journal of Surgery* 2011: 98 (S5) S42-43 – **Abstract** - Presented at the 46th Congress of the European Society for Surgical Research (May 2011, Germany)

V Agarwal, **V C Hodgkinson**, G L Eagle, L Scaife, M J Lind, L Cawkwell. Proteomic analysis of the mechanism of action of a COX2 inhibitor (DuP-697) in mesothelioma. *Lung Cancer*, 2011, 71 (suppl 1), S17 – **Abstract** - Presented at the 9th Annual British Thoracic Oncology Group Conference (Jan 2011, Ireland)

D ELFadl, **V Hodgkinson**, E Long, A W Beavis, P J Drew, M J Lind, L Cawkwell. Immunohistochemistry confirms downregulation of protein markers associated with resistance to radiotherapy. *British Journal of Surgery* 2011: 98 (suppl 2) 16-17 – **Abstract** - Presented at the Society of Academic & Research Surgery Meeting (Jan 2011, Ireland)

Conferences (attended and presented)

Hodgkinson VC, ELFadl D, Agarwal V, Garimella V, Drew PJ, Lind MJ, Cawkwell L. Proteomic identification of predictive biomarkers of resistance to neoadjuvant chemotherapy in luminal breast cancer: a possible role for 14-3-3 and BID? – accepted as a poster presentation at the San Antonio Breast Cancer Symposium (December 2011, Texas)

V Hodgkinson, D ELFadl, C Russell, V Agarwal, V Garimella, P Drew, M J Lind and L Cawkwell. Predictive biomarkers of chemotherapy resistance in breast cancer: a possible role for 14-3-3 tau and BID? Abstract for EMBO Cancer Proteomics (June 2011, Ireland)

V Agarwal, **V C Hodgkinson**, G L Eagle, L Scaife, M J Lind, L Cawkwell. Investigating the mechanism of action of DuP-697 in mesothelioma cells. Abstract for EMBO Cancer Proteomics (June 2011, Ireland)

D ELFadl, **V C Hodgkinson**, L Scaife, E D Long, P J Drew, M J Lind, L Cawkwell. DR4, identified by antibody microarray analysis, may be a novel biomarker of radioresistance. Abstract for EMBO Cancer Proteomics (June 2011, Ireland)

L Scaife, D ELFadl, **V Hodgkinson**, J T Murphy, A W Beavis, P J Drew, M J Lind and L Cawkwell. Antibody Microarray Identification of Putative Biomarkers associated with Radioresistance in Head and Neck Cancer. Abstract for EMBO Cancer Proteomics (June 2011, Ireland)

V Hodgkinson, D ELFadl, V Garimella, P Drew, M J Lind, Lynn Cawkwell. The identification of predictive biomarkers of resistance to neoadjuvant chemotherapy in breast cancer using antibody microarrays. Abstract for EMBO Cancer Proteomics (June 2009, Ireland)

Contents

List of Figures.....	xix
List of Tables.....	xxii
List of Abbreviations	xxiv
Chapter 1. Clinical Introduction to Breast Cancer	1
1.1 Breast Cancer Epidemiology.....	1
1.2 Breast Cancer Aetiology	1
1.3 Histological types of breast cancer	2
1.4 Diagnosis and Classification.....	4
1.5 Treatment	6
1.6 Prognostic and Predictive Factors.....	7
1.6.1 Prognostic factors.....	8
1.6.2 Predictive factors.....	9
1.7 Monitoring response to treatment	10
1.8 Resistance to Treatment	12
Chapter 2. Molecular Introduction to Breast Cancer	14
2.1 Normal Tissue Homeostasis.....	14
2.1.1 The Cell Cycle.....	14
2.1.2 DNA damage response pathway	15
2.1.3 DNA repair.....	18
2.1.4 Apoptosis	20

2.1.4.1 Intrinsic pathway	20
2.1.4.2 Extrinsic pathway	21
2.2 Breast Cancer	25
2.2.1 Hallmarks of Cancer	25
2.2.2 Breast cancer at a molecular level.....	26
2.3 Breast Cancer treatment using neoadjuvant chemotherapy	28
2.3.1 Anthracyclines.....	29
2.3.1.1 Topoisomerase II inhibitors	29
2.3.1.2 DNA intercalators	30
2.3.1.3 Other mechanisms of action.....	31
2.3.2 Cyclophosphamide	32
2.3.3 Taxanes	33
2.4 Chemotherapy resistance	34
2.4.1 History and basis of multi-drug resistance.....	34
2.4.2 Predictive markers of chemo-resistance.....	36
2.4.2.1 Adenosine triphosphate-binding cassette (ABC) transporters	36
2.4.2.2 p53 status.....	37
2.4.2.3 Topoisomerase II alterations	38
2.4.2.4 Tumour cell proliferation (Ki67 status)	39
2.4.2.5 DNA damage response and BRCA1 alterations	40
2.4.2.6 Breast cancer stem cell markers	41

2.4.2.7 Molecular Subtype (ER, PR and HER2 status).....	42
2.4.3 Predictive gene expression signatures.....	43
2.4.4 Overview of putative markers of response to neoadjuvant chemotherapy in breast cancer	44
Chapter 3. Proteomic techniques and the identification of biomarkers of chemotherapy resistance	49
3.1 Biomarker discovery pipeline	49
3.2 Proteomics.....	51
3.3 Gel-based MS methods	52
3.3.1 1D-PAGE separation.....	52
3.3.2 2D-PAGE separation.....	53
3.3.2.1 Sample Preparation	53
3.3.2.2 Separation in the 1 st Dimension: Isoelectric Focusing.....	54
3.3.2.3 Separation in the 2 nd dimension: by molecular weight	55
3.3.2.4 Visualisation of Proteins	57
3.3.2.5 Identification of differentially expressed proteins	57
3.3.2.6 In-gel digest.....	58
3.3.2.7 Protein identification by mass spectrometry	58
3.3.2.8 Protein identification.....	62
3.3.2.9 Tandem mass spectrometry.....	62
3.4 Gel-free methods.....	63

3.4.1 ESI MS	64
3.4.2 Quantitative Shotgun Proteomics.....	65
3.5 Microarray-based methods.....	66
3.6 Techniques for the confirmation and validation of putative biomarkers	70
3.6.1 Data mining.....	70
3.6.2 Western Blotting	70
3.6.3 Clinical Validation	72
3.7 Proteomics research to identify biomarkers of chemotherapy resistance in breast cancer	73
3.8 Project Aim	80
Chapter 4. Materials and Methods	83
4.1 Culture of Cell Lines.....	83
4.1.1 Human Caucasian Breast Adenocarcinoma ‘MCF7’ Cell Line.....	83
4.1.2 Thawing Cells	83
4.1.3 Culturing Cells	83
4.1.4 Freezing Cells.....	84
4.2 Collection of Clinical Samples	85
4.3 Panorama Antibody Microarray XPRESS725 Profiler.....	88
4.3.1 Protein Extraction.....	88
4.3.2 Protein Precipitation.....	89
4.3.3 Protein Quantification	90

4.3.4 Protein Labelling.....	91
4.3.5 Determination of dye-to-protein molar ratio.....	91
4.3.6 Antibody Incubation.....	92
4.3.7 Scanning and Analysis	92
4.4 Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled to matrix-assisted laser desorption/ionization time-of-flight/time of flight (MALDI-TOF/TOF) mass spectrometry.....	94
4.4.1 Protein Extraction.....	94
4.4.2 ReadyPrep 2-D Cleanup Kit.....	95
4.4.3 Protein Quantification	96
4.4.4 Isoelectric Focusing	97
4.4.5 Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	98
4.4.6 Staining of Proteins	99
4.4.7 PDQuest Analysis Software.....	99
4.4.8 Spot excision	100
4.4.9 In-gel digestion.....	100
4.4.9.1 De-stain	101
4.4.9.2 Digest	101
4.4.10 Preparation of MALDI matrix and plate-spotting.....	101
4.4.11 MALDI-TOF/TOF Mass Spectrometry	102
4.5 Ingenuity Pathway Analysis.....	103

4.6 Western Blotting	105
4.6.1 Protein Extraction.....	105
4.6.1.1 From cultured cells.....	105
4.6.1.2 From breast tumour tissue.....	105
4.6.2 Protein Quantification	106
4.6.3 One-dimensional gel electrophoresis	107
4.6.4 Transfer of proteins onto nitrocellulose membrane	108
4.6.5 Blocking of binding sites on the membrane.....	108
4.6.6 Immunoblotting.....	108
4.6.7 Loading controls.....	109
4.6.8 Detection of proteins	109
4.6.9 Densitometry	110
4.7 Immunohistochemistry.....	110
4.7.1 De-waxing and rehydration.....	111
4.7.2 Blocking of endogenous peroxidase	111
4.7.3 Antigenic site retrieval	111
4.7.4 Blocking of non-specific binding sites within sections.....	111
4.7.5 Incubation with primary antibody	112
4.7.6 Antibody detection.....	112
4.7.6.1 StreptABComplex/HRP Duet Kit	112
4.7.6.2 R.T.U VECTASTAIN Universal <i>Quick</i> Kit	113

4.7.7 Antibody Visualisation	113
4.7.8 Enhance, counterstain and differentiate	113
4.7.9 Rehydration, clearing and mounting	114
4.7.10 Antibodies used for Immunostaining	114
4.7.11 Histological scoring of immunostained tissue sections	114
4.7.12 Fishers Exact test for determination of statistical significance	115
4.7.13 Determination of inter-observer variability using the Kappa statistic	116
Chapter 5. Clinical tumour sample collection	119
5.1 Introduction	119
5.1.1 Ethical approval	119
5.1.2 Clinical samples for proteomic analysis	119
5.2 Methodology	121
5.2.1 Ethical approval for this study	121
5.2.2 Study design	121
5.2.3 Clinical tumour sample collection	122
5.2.4 Classification of response to neoadjuvant chemotherapy	123
5.3 Results	124
5.3.1 Clinical tumour sample series	124
5.4 Discussion	127
Chapter 6. Clinical tissue samples – optimisation of methods and preliminary work	130
6.1 Introduction	130

6.1.1 Cell lines versus clinical samples.....	130
6.1.1.1 Cell lines	130
6.1.1.2 Clinical Samples	132
6.1.2 Proteomics using clinical tissue samples	137
6.1.2.1 Sample preparation.....	137
6.1.3 Proteomic methods.....	138
6.1.3.1 Core biopsy samples	138
6.2 Methods.....	139
6.2.1 Antibody microarray – based methods	139
6.2.1.1 Protein extraction	139
6.2.1.2 Protein precipitation	139
6.2.1.3 Protein quantification.....	139
6.2.1.4 Determination of protein yield from core biopsy samples.....	140
6.2.2 One-dimensional gel electrophoresis	140
6.2.3 2D-PAGE MALDI-TOF/TOF MS – based methods.....	140
6.2.3.1 Protein extraction	140
6.2.3.2 Sample preparation.....	141
6.2.3.3 2D-PAGE	141
6.2.3.4 Excision of protein spots and in-gel tryptic digest.....	141
6.2.3.5 Peptide analysis by MALDI-TOF/TOF MS and protein identification	141
6.2.4 Western blot	141

6.3 Results	142
6.3.1 Antibody Microarray	142
6.3.1.1 Relationship between tissue weight and protein yield	143
6.3.1.2 Precipitation of protein using ProteoExtract Protein Precipitation Kit	144
6.3.2 Two-dimensional polyacrylamide gel electrophoresis (2D PAGE)	146
6.3.2.1 Optimisation of extraction method	146
6.3.3 Western Blot	150
6.3.3.1 Molecular weight markers	150
6.3.3.2 Western blotting with tissue samples	151
6.3.3.3 Antibody Optimisation	152
6.3.4 Determination of protein yield from core biopsy samples	152
6.4 Discussion	154
Chapter 7. Repeatedly identified differentially expressed proteins (RIDEPs) from antibody microarray proteomic analysis	158
7.1 Introduction	158
7.2 Materials and Methods	161
7.2.1 Panorama Antibody Microarray XPRESS Profiler725 Kit	161
7.2.2 Experimental sample group 1: Stimulation of the B-cell receptor in malignant B- cells from patients with chronic lymphocytic leukaemia (n=1)	161
7.2.3 Experimental sample group 2: Treatment of human lung cancer and mesothelioma cell lines with a COX-2 inhibitor (n=2)	162

7.2.4 Experimental sample group 3: Biomarkers of the radiotherapy-resistant phenotype in human breast cancer cell lines (n=3).....	163
7.2.5 Experimental sample group 4: Biomarkers of the radiotherapy-resistant phenotype in human oral cancer cell lines (n=2)	163
7.2.6 Experimental sample group 5: Biomarkers of chemotherapy response in human breast cancer tissue (n=5).....	163
7.2.7 Protein Labelling.....	164
7.2.8 Protein Binding	164
7.2.9 Image acquisition and analysis.....	165
7.3 Results.....	166
7.4 Discussion	168
7.4.1 Quality control	168
7.4.2 Fold change cut-off	168
7.4.3 RIDEPS.....	169
7.5 Conclusions.....	171
Chapter 8. Antibody microarray analysis for the identification of biomarkers of chemotherapy resistance	173
8.1 Introduction.....	173
8.1.1 Antibody microarrays.....	173
8.2 Methods.....	175
8.2.1 Fresh tumour samples	175

8.2.2 Antibody microarray analysis	177
8.3 Results.....	178
8.4 Discussion	179
Chapter 9. Two-dimensional gel electrophoresis and mass spectrometry for the identification of biomarkers of chemotherapy resistance	182
9.1 Introduction.....	182
9.1.1 MALDI TOF/TOF mass spectrometry.....	183
9.2 Methodology	186
9.2.1 Clinical Samples	186
9.2.2 Protein extraction	187
9.2.3 Protein clean-up and quantification	187
9.2.4 2D PAGE	188
9.2.4.1 1 st dimension (IEF).....	188
9.2.4.2 2 nd dimension (SDS-PAGE).....	188
9.2.5 PDQuest	188
9.2.6 Spot excision and in-gel digest	189
9.2.7 Identification by MALDI-TOF/TOF MS.....	189
9.3 Results.....	190
9.3.1 2D-PAGE.....	190
9.3.2 PDQuest	194
9.3.3 Selection and identification of DEPs	198

9.3.4	Examples of mass spectra	199
9.3.5	Example of a Mascot summary report page.....	199
9.3.6	A protein identification with a single peptide match	199
9.3.7	Differentially expressed proteins identified by 2D-PAGE MS.....	202
9.4	Discussion	206
Chapter 10. Data mining, confirmation and clinical validation		218
10.1	Introduction	218
10.1.1	Data mining.....	218
10.1.2	Confirmation	219
10.1.3	Clinical validation	219
10.2	Methodology	219
10.2.1	Protein selection for data-mining.....	219
10.2.1.1	Antibody microarray data	219
10.2.1.2	2D-PAGE/MS data	221
10.2.2	IPA	221
10.2.3	Confirmation: western blotting	224
10.2.4	Clinical validation: immunohistochemistry	225
10.2.4.1	Archival pre-treatment core biopsy samples.....	225
10.2.4.2	Immunohistochemistry.....	225
10.3	Results	226
10.3.1	Discovery-phase data.....	226

10.3.2 IPA	227
10.3.2.1 Antibody microarray data	227
10.3.2.2 2D-PAGE/MS data	227
10.3.2.3 Combined antibody microarray and 2D/MS data	232
10.3.3 Confirmation: western blotting	232
10.3.4 Clinical validation: immunohistochemistry	234
10.4 Discussion	237
10.4.1 Apoptosis-related proteins	238
10.4.2 14-3-3 proteins	240
Chapter 11. Conclusions	245
11.1 Antibody microarray analysis	246
11.2 2D-PAGE MALDI-TOF/TOF MS analysis	247
11.3 Future work	248
11.3.1 Discovery phase	248
11.3.2 Confirmation and clinical validation	249
11.3.3 Core biopsy samples	249
11.3.4 Establishment of chemotherapy-resistant cell lines	250
11.4 Concluding Remarks	252
REFERENCES	253
APPENDIX 1: 725 Antibodies (Panorama Antibody Microarray XPRESS Profiler	270
APPENDIX 2 Buffers and reagents	275

APPENDIX 3: REC Approval for the Study.....	276
APPENDIX 4: Patient Information Sheet	278
APPENDIX 5: All raw clinical data for clinical samples collected.....	282
APPENDIX 6: Annotated Spectra (ProteinScape).....	285
APPENDIX 7: Peptide View (MASCOT).....	286
APPENDIX 8: Prohossi MS/MS Report.....	287
APPENDIX 9: DEPs from 2D-PAGE/MS analysis (identified in 1/3 experiments)	288
APPENDIX 10: 2D-PAGE MALDI TOF/TOF MS raw data.....	290
APPENDIX 11: Peptide views for protein identifications with only a single peptide match in 2/3 and 3/3 2D-PAGE MALDI-TOF/TOF experiments	305

List of Figures

Figure 1: Incidence and mortality rates (female) for major cancers 2004-2006 in the UK.....	2
Figure 2: Breast cancer incidence continues to rise	3
Figure 3: Structure of the breast.....	4
Figure 4: Survival rates 0-5 years after diagnosis, according to breast cancer stage.....	6
Figure 5: Overview of the cell division cycle and DNA damage response	17
Figure 6: Apoptosis Pathways.....	22
Figure 7: Summary of response to DNA damage	24
Figure 8: Differences in clinical outcome between molecular subtypes of breast cancer	28
Figure 9: Stages of the biomarker discovery pipeline.....	49
Figure 10: Pie charts representing the 22 major proteins of the human serum proteome.....	50
Figure 11: Stages of protein separation by isoelectric focusing (IEF) – 1 st dimension	55
Figure 12: Stages of protein separation by SDS-PAGE – 2 nd dimension	56
Figure 13: MALDI-TOF MS	61
Figure 14: Antibody microarray workflow	69
Figure 15: Consort chart outlining where clinical tumour resection samples have been used for proteomic analysis.....	87
Figure 16: Layout of target plate for MALDI TOF/TOF MS.....	102
Figure 17: Flow diagram illustrating study design	123
Figure 18: Types of Invasive Breast Carcinoma within the clinical sample series	127
Figure 19: Representation of molecular subtypes within invasive carcinoma samples.....	127
Figure 20: Dynamic range of normal human plasma proteins.....	133
Figure 21: Sample preparation and protein extraction from clinical tissue samples for antibody microarray analysis	142

Figure 22: Correlation between breast tumour tissue weight and protein yield	144
Figure 23: Breast tumour lysate optimisation of protein precipitation	145
Figure 24: Graph showing protein concentration for precipitated and non-precipitated tumour extracts	146
Figure 25: 2D-PAGE with breast tumour tissue extract (insufficient protein)	147
Figure 26: Breast tumour tissue extract separated by 2D PAGE	148
Figure 27: Breast tumour tissue extract separated by 2D PAGE for protein identification using MALDI-TOF/TOF MS	149
Figure 28: Western Blot Molecular Weight Marker	151
Figure 29: Antibodies optimised for western blotting with clinical samples	153
Figure 30: Collection of core biopsy samples for determination of protein yield	154
Figure 31: Scanning of antibody microarray slides	175
Figure 32: 2D-PAGE/MS workflow	182
Figure 33: Schematic of the Ultraflex III MALDI-TOF/TOF mass spectrometer	185
Figure 34: Peptide fragmentation (Roepstorff and Fohlman notation)	186
Figure 35: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 1) ...	191
Figure 36: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 2) ..	192
Figure 37: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 3) ...	193
Figure 38: Histograms and DEP spots from PDQuest analysis (Experiment 1)	195
Figure 39: Histograms and DEP spots from PDQuest analysis (Experiment 2)	196
Figure 40: Histograms and DEP spots from PDQuest analysis (Experiment 3)	197
Figure 41: Examples of mass spectra	200
Figure 42: An example of a Mascot concise protein summary page	202

Figure 43: The biological processes of the putative biomarkers identified within both the clinical tissue experiments and the MCF7 breast cancer cell line models from the literature (n=11).....	214
Figure 44: Top canonical pathway identified by IPA from the antibody microarray data set: ERK5 signalling.....	228
Figure 45: Top canonical pathway identified by IPA from the 2D-PAGE/MS data set: Cell cycle. G2/M DNA damage checkpoint regulation.....	230
Figure 46: One of the top canonical pathway identified by IPA from combined antibody microarray 2D-PAGE/MS data set: 14-3-3 mediated signalling	233
Figure 47: Confirmation of DEPs using western blotting.....	234
Figure 48: Immunohistochemical analysis of 14-3-3 theta/tau and tBID expression in invasive breast carcinoma cells.....	236

List of Tables

Table 1: TNM Staging System for Breast Cancer Classification.....	5
Table 2: Summary of the Response Evaluation Criteria In Solid Tumours (RECIST) Guidelines.....	11
Table 3: A summary of putative predictive biomarkers of chemotherapy resistance in breast cancer.....	45
Table 4: Proteins demonstrating differential expression in chemotherapy-resistant MCF-7 cell sub-lines identified by 2D-PAGE/MS.....	78
Table 5: Table of primary antibodies used for western blotting.....	110
Table 6: Primary antibodies used for immunohistochemical staining.....	114
Table 7: Scoring method for cytoplasmic immunostaining.....	114
Table 8: Scoring method for immunostaining of the nuclear membrane.....	114
Table 9: Scoring method for nuclear staining.....	115
Table 10: A 2 x 2 contingency table for the Fishers exact test	115
Table 11: A 2 x 2 contingency table for the Kappa statistical test.....	116
Table 12: Criteria for interpretation of the kappa coefficient.....	117
Table 13: Summarised data for clinical tumour samples.....	126
Table 14: Comparison of cell line models with clinical tissue samples.....	130
Table 15: Extraction of protein from breast tumour resection samples.....	143
Table 16: Preliminary test for proteins identification by MALDI-TOF/TOF MS peptide analysis.....	150
Table 17: Protein yield from core biopsy samples.....	154
Table 18: ‘TOP 15’ RIDEPS identified from 2D-PAGE/MS experiments.....	160
Table 19: RIDEPS from antibody microarray analysis.....	167
Table 20: Clinical samples selected for antibody microarray analysis.....	176
Table 21: Pair-wise combinations of samples used for 5 antibody microarray experiments....	177
Table 22: DEPs in chemotherapy resistant tumour tissue identified by 5 antibody microarray experiments comparing CR and CS samples.....	178

Table 23: Clinical samples selected for 2D-PAGE/MS analysis.....187
Table 24: Pair-wise combinations of clinical samples for three 2D-PAGE/MS experiments..187
Table 25: Search criteria specified for protein identification190
Table 26: The number of DEPs identified for each experiment, with relation to the number initially identified by PDQuest and analysed by MALDI TOF/TOF MS.....198
Table 27: The number of unique DEPs identified from each experiment, with relation to the number initially identified by PDQuest and analysed by MALDI TOF/TOF MS.....203
Table 28: The number of DEPs which have been identified in 3/3, 2/3 and 1/3 experiments.....203
Table 29: Differentially expressed proteins associated with chemotherapy resistance, identified using 2D-PAGE MALDI-TOF/TOF MS.....204
Table 30: The 22 proteins which constitute ~99% of the plasma proteome....210
Table 31: DEPs identified which are included in the ‘TOP15’ RIDEPS.....211
Table 32: DEPs identified, which have also been associated with drug resistance in breast cancer (ER+) MCF7 cell lines, using 2D-PAGE MALDI-TOF MS and MS/MS analysis...213
Table 33: DEPs in chemotherapy resistant tumour tissue identified by 5 antibody microarray experiments comparing CR and CS samples.....220
Table 34: DEPs associated with chemotherapy resistance, identified using 2D-PAGE MALDI-TOF/TOF MS.....222
Table 35: Canonical pathways identified by IPA performed on DEPs generated from antibody microarray analysis.....229
Table 36: Canonical pathways identified by Ingenuity Pathway Analysis performed on DEPs generated from 2D-PAGE MALDI-TOF/TOF MS analysis.....231

List of Abbreviations

2D-DIGE	Two-dimensional Difference Gel Electrophoresis
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
ABC	ATP-Binding Cassette
AJCC	American Joint Committee on Cancer
ATM	Ataxia Telangiectasia Mutated
ATP	Adenosine Triphosphate
ATR	Ataxia Telangiectasia and Rad3-related kinase
BCRP	Breast Cancer Resistance Protein
BCS	Breast Conserving Surgery
BCT	Breast Conserving Therapy
BER	Base Excision Repair
BRISQ	Biospecimen Reporting for Improved Study Quality
BSA	Bovine Serum Albumin
Cdk	Cyclin-dependent kinase
CHCA	Alpha-cyano-4-hydroxycinnamic acid
CR	Chemotherapy-resistant
CS	Chemotherapy-sensitive
CSC	Cancer Stem Cell
DCE-MRI	Dynamic Contrast Enhanced –Magnetic Resonance Imaging
DCIS	Ductal Carcinoma In Situ
ddH ₂ O	Ultrapure water
DDR	DNA Damage Response

DEP	Differentially Expressed Protein
DFS	Disease Free Survival
DMSO	Dimethyl Sulphoxide
DSB	Double Strand Break
DTT	Dithiothreitol
EB	Equilibration Buffer
EC	Epirubicin and Cyclophosphamide
ELISA	Enzyme-Linked-Immunosorbent Assay
ER	Estrogen Receptor
ESI	Electrospray Ionisation
FA	Fanconi Anaemia
FEC	Fluorouracil (5-FU), Epirubicin and Cyclophosphamide
FFPE	Formalin-Fixed Paraffin Embedded
HER2	Human Epidermal Growth Factor Receptor 2
HR	Homologous Recombination
HRP	Horseradish Peroxidase
IAA	Iodoacetamide
ICL	Intra-strand Cross Link
IDC	Invasive Ductal Carcinoma
IEF	Isoelectric Focusing
IHC	Immunohistochemistry
ILC	Invasive Lobular Carcinoma
IPA	Ingenuity Pathway Analysis

IPI	International Protein Index
LABC	Locally Advanced Breast Cancer
LC	Liquid Chromatography
LCIS	Lobular Carcinoma In Situ
LCM	Laser Capture Microdissection
LID	Laser Induced Dissociation
MALDI	Matrix-Assisted Laser Desorption/Ionisation
MDR	Multidrug Resistance
MINDACT	Microarray In Node negative Disease may Avoid Chemotherapy
MMR	Mismatch repair
MS	Mass Spectrometry
NER	Nucleotide Excision Repair
NHEJ	Non Homologous End Joining
NPI	Nottingham Prognostic Index
OS	Overall Survival
PBS	Phosphate-Buffered Saline
pCR	Pathological Complete Response
PFF	Peptide Fragmentation Fingerprint
PMF	Peptide Mass Fingerprint
PR	Progesterone Receptor
REC	Research Ethics Committee
RECIST	Response Evaluation Criteria In Solid Tumours
RIDEP	Repeatedly Identified Differentially Expressed Protein

RT	Room Temperature
SDS	Sodium Dodecyl Sulphate
SSB	Single Strand Break
TAILORx	Trial Assigning Individualised Options for Treatment (Rx)
TBS	Tris-Buffered Saline
TFA	Trifluoroacetic Acid
TNM	Tumour-Node-Metastasis
TOF	Time-Of-Flight
US	Ultrasound
WB	Western blot

CHAPTER 1:
CLINICAL INTRODUCTION TO
BREAST CANCER

Chapter 1. Clinical Introduction to Breast Cancer

1.1 Breast Cancer Epidemiology

Breast cancer is the most commonly occurring cancer in the UK, despite it predominantly affecting females, and the second-leading cause of cancer-related deaths. In 2004, breast cancer incidence in females represented approximately one third of total malignancies in the UK (Westlake and Cooper 2008) and mortalities were the second highest after lung cancer, representing 17% of all female cancer deaths (Westlake and Cooper 2008) (Figure 1). Incidence rates for breast cancer in females were 24% higher than those for the cancer of highest incidence in males (Westlake 2009). Incidence rates are continuing to rise (Figure 2), with close to 47,700 new cases reported in the UK in 2008, of which 5,360 were reported within local Northern and Yorkshire Cancer networks (Northern and Yorkshire Cancer Registration Information Service [NYCRIS], Leeds). Currently, it is estimated that women in the UK have a 1 in 8 lifetime risk of developing breast cancer.

1.2 Breast Cancer Aetiology

Breast cancer results from the accumulation of genetic abnormalities and mutations which lead to malignant transformation of cells in the breast. There are several factors which may contribute to the development of breast cancer, of which female gender, old age and country of birth are the strongest disease risk factors (Hulka and Moorman 2001). Other factors include mutations in the cancer susceptibility genes (BRCA1, BRCA2, p53, PTEN, ATM) (section 2.2.1) family history, high mammographic density, early menarche, late menopause, high post-menopausal bone density, higher age at first full-term pregnancy, post-menopausal obesity, use of oral contraceptives, hormone replacement therapy and exposure to ionising radiation (Hulka and Moorman 2001; Key, Verkasalo et al. 2001;

Sakorafas, Krespis et al. 2002; Schwab, Claas et al. 2002). Alcohol consumption has also been linked to increase risk of breast cancer (Li, Baer et al. 2009). Models are available to estimate the risk of breast cancer development for an individual; namely the Gail Model and the Claus Model (McTiernan, Kuniyuki et al. 2001).

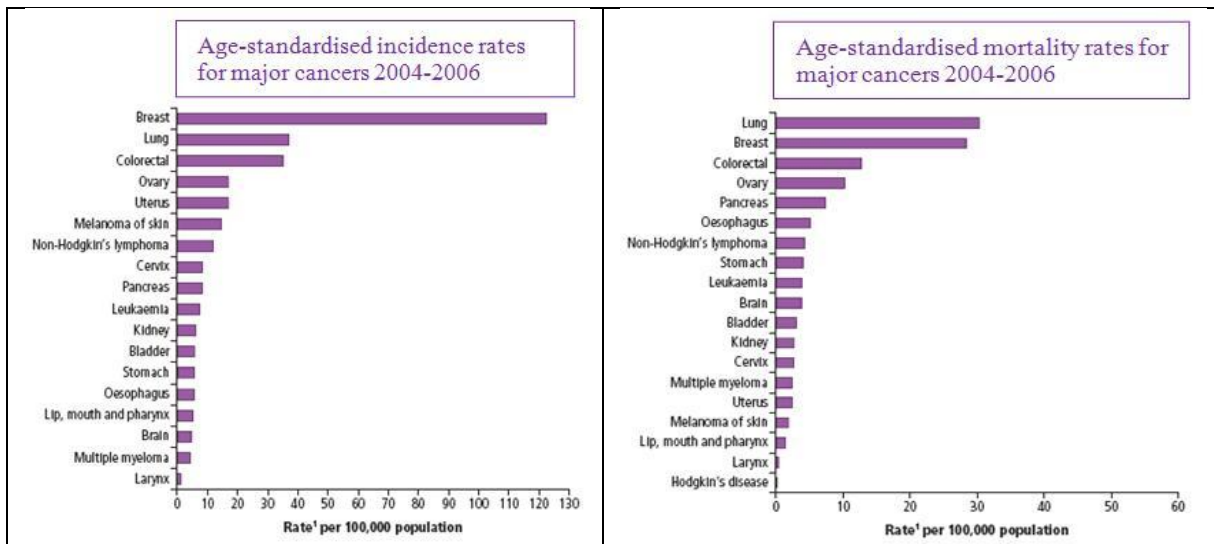


Figure 1: Incidence and mortality rates (female) for major cancers 2004-2006 in the UK

Breast cancer incidence rates, at 122 per 100,000 people, are currently the highest out of all malignancies. The mortality rates for breast cancer are slightly lower than lung cancer (28 and 31 per 100,000 people respectively), displaying the second-highest mortality rate (Westlake 2009)

1.3 Histological types of breast cancer

Breast tissue has many constituents including: *lobules*, which are milk-producing glands; *ducts*, which transport milk to the nipple; nerves; lymph vessels; pectoral muscle; adipose and connective tissue, which line the vessels (Figure 1). The two most common types of invasive breast cancer include invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) which represent 70-80% and 10% of cases respectively

(CancerResearchUK 2002). Other types of breast cancer are rare and include inflammatory, tubular, mucinous and medullary breast cancers.

Non-invasive pre-cancerous lesions which may develop into cancer are termed 'ductal carcinoma *in situ*' (DCIS) and 'lobular carcinoma *in situ*' (LCIS), depending on their location of origin. These lesions are not regarded as true cancer; however their presence increases the risk of developing invasive breast cancer.

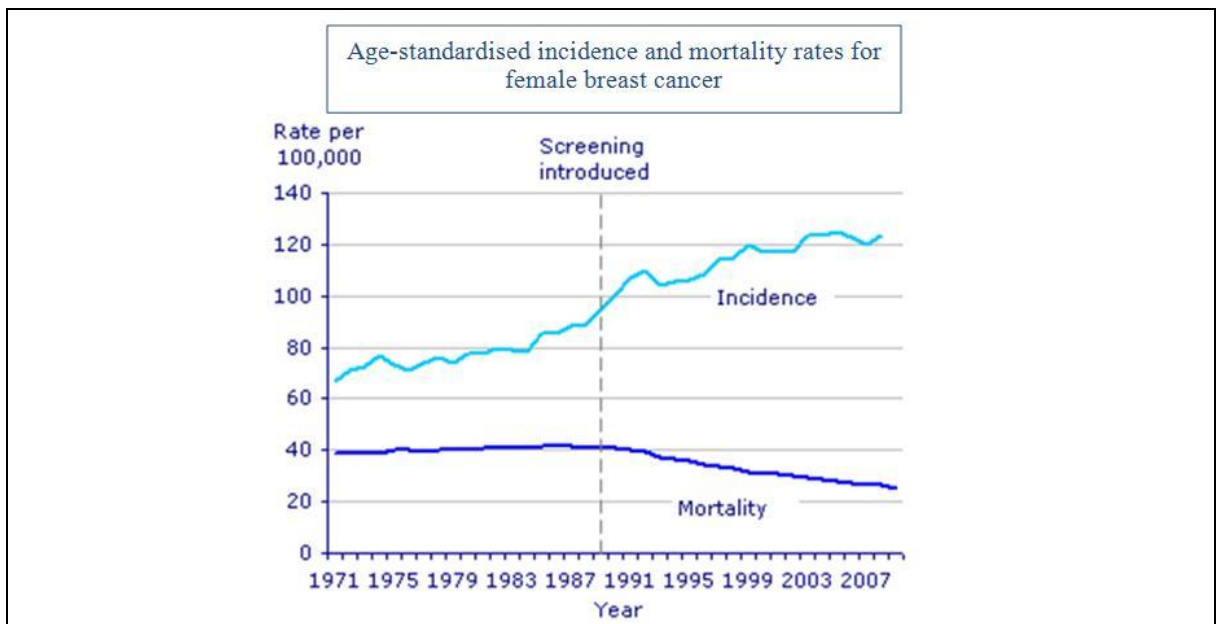


Figure 2: Breast cancer incidence continues to rise

Incidence rates of breast cancer continue to rise, whilst mortality rates decrease. This may be due to the introduction of the breast screening program as well as the introduction of the drug tamoxifen for the treatment of early breast cancer.

Office for National Statistics 2010.

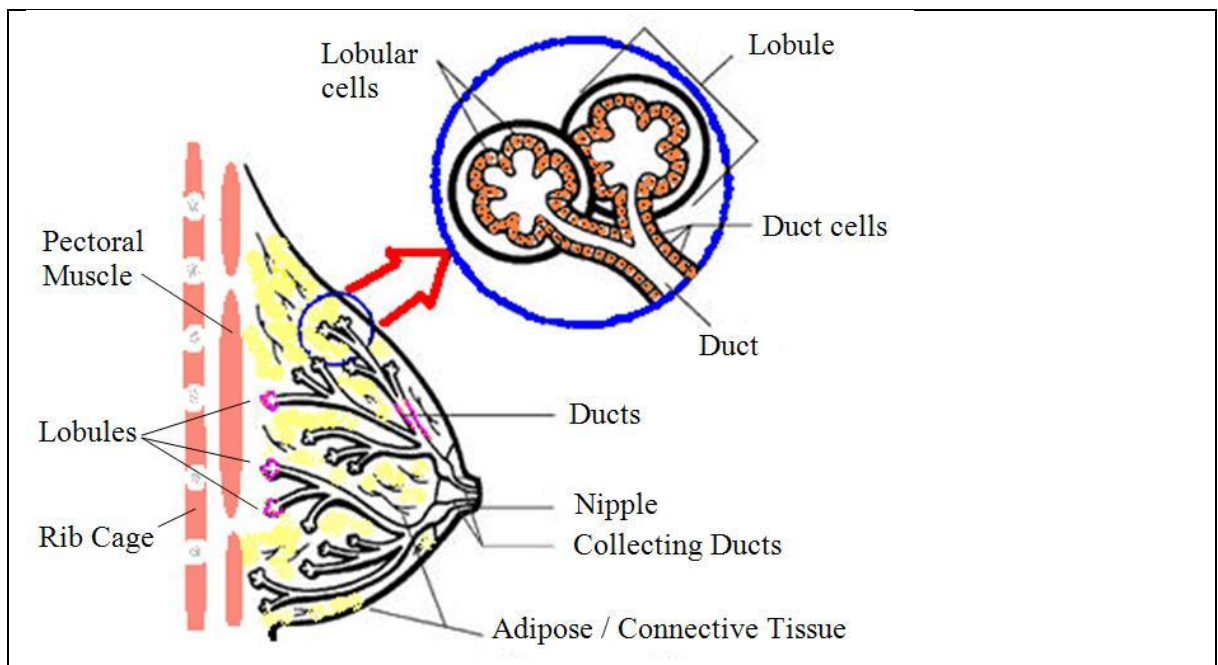


Figure 3: Structure of the breast

The structure of the breast, showing the location of the lobules (milk-producing glands) and ducts (which lead from the lobules and deliver milk to the nipple) representing the origin of 70-80% and 10% of breast cancers respectively.

Adapted from (AmericanCancerSociety 2008)

1.4 Diagnosis and Classification

Diagnosis of symptomatic breast disease is through triple assessment, which constitutes clinical (palpation), radiological (mammography and ultrasound) and histological assessment (fine needle aspirate or core needle biopsy).

Upon diagnosis, breast cancers are classified according to the AJCC (American Joint Committee on Cancer) Staging Manual, where factors assessed include tumour size, lymph node involvement and presence of metastases. This is referred to as the tumour-node-metastases (TNM) staging system (Singletary and Connolly 2006) (Table 1). These TNM categories can then be grouped into five stages (0 to IV) (Singletary, Allred et al. 2002), where higher stages are associated with poorer survival rates (Singletary and Connolly 2006) (Figure 4). Other factors which affect prognosis and management include tumour

grade, lymphovascular invasion, hormone receptor status (section 1.6.1) and presence or absence of multifocal disease.

Table 1: TNM Staging System for Breast Cancer Classification

The TNM staging system is used to classify breast cancer according to its size, lymph node involvement and presence of metastases. This aids the prognosis and management of the disease (Singletary and Connolly 2006).

T (tumour)	N (nodes)	M (metastases)
<p>T1 : the tumour is less than 2cm across</p> <p>T1 (a): tumour is > 0.1 cm but < 0.5 cm</p> <p>T1 (b): tumour is > 0.5 cm but < 1.0 cm</p> <p>T1 (c): tumour is > 1.0 cm but < 2.0 cm</p>	<p>N0: no cancer cells in lymph nodes</p>	<p>M0: no sign of cancer spread</p>
<p>T2: the tumour is > 2cm but < 5 cm</p>	<p>N1: cancer cells in lymph nodes in armpit but nodes not attached to other structures</p>	<p>M1: cancer has spread to another part of body not including breast and lymph nodes in the armpit</p>
<p>T3: the tumour is > 5 cm across</p>	<p>N2:N2 (a): cancer cells in lymph nodes in the armpit, which are stuck to each other and other structures</p> <p>N2 (b): cancer cells in lymph nodes under breast bone (seen on scan/felt by doctor) and no evidence in armpit.</p>	<p>The three stages are used together to classify a tumour to give an overall stage. E.g. T2 N0 M0</p>
<p>T4</p> <p>T4 (a): Tumour is fixed to chest wall</p> <p>T4 (b): Tumour is fixed to the skin</p> <p>T4 (c): Tumour is fixed to skin & chest wall</p> <p>T4 (d): Inflammatory carcinoma – erythema of overlying skin, swollen and painful to touch</p>	<p>N3:</p> <p>N3 (a): cancer in lymph nodes below collarbone</p> <p>N3 (b): cancer in lymph nodes in armpit and under breast bone</p> <p>N3 (c): cancer in lymph nodes above the collarbone</p>	

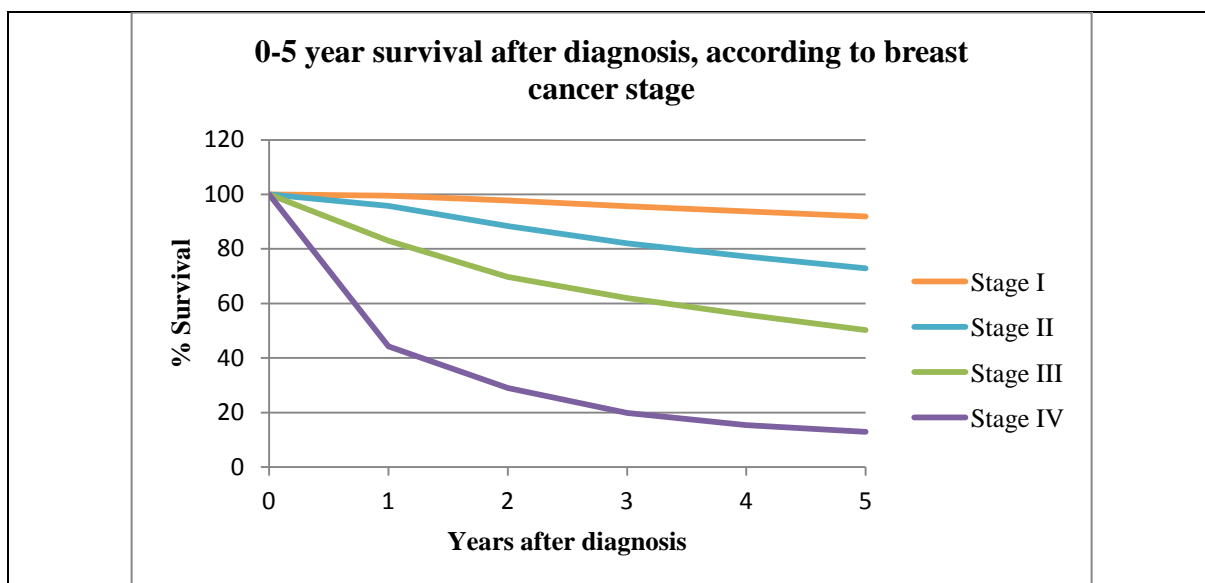


Figure 4: Survival rates 0-5 years after diagnosis, according to breast cancer stage
 This data was obtained from the West Midlands during 1990-1994, followed up to 2004 and published in 2008. The above graph shows the percentage survival 0-5 years after diagnosis according to breast cancer stage. Five year survival rates were 92% for Stage I, 73% for Stage II, 50% for stage III and 13% for Stage IV. (Cancer Research UK).

1.5 Treatment

From a clinical perspective, the treatment administered for breast cancer is dependent upon the stage, grade and type of disease, as well as tumour and patient characteristics. The definitive treatment for early breast cancer is surgery, but may also include chemotherapy, in the ‘neoadjuvant’ (pre-surgery) or ‘adjuvant’ (post-surgery) setting, radiotherapy, and hormone therapy. Treatment can also be tailored to suit the molecular status of the tumour and specific therapies can be administered. Trastuzumab (Herceptin[®]), for example, may be administered for the treatment of human epidermal growth factor receptor 2 (HER2) positive tumours, or Tamoxifen for tumours expressing the estrogen receptor (section 2.2.2). Treatment for locally advanced breast cancer (LABC), which is defined as a tumour > 5cm (T3) or with fixed skin or chest involvement (T4) and/or extensive axillary involvement (N2-N3), (Espinosa, Morales et al. 2004; Mathew, Asgeirsson et al. 2009)

(Table 1) may involve combination therapy, including chemotherapy, surgery and radiotherapy. However, increasing numbers of patients are choosing breast conserving therapy (BCT). This involves the use of neoadjuvant chemotherapy, which aims to shrink the tumour prior to resection, to increase the chance of breast conserving surgery (BCS), rather than a mastectomy. Neoadjuvant chemotherapy currently administered in Hull and East Yorkshire NHS Trust includes 4 cycles of EC [epirubicin (90 mg/m²), cyclophosphamide (600 mg/m²)] followed by 4 cycles of docetaxel (100 mg/m²). Following this, surgical intervention is applied to remove the remaining tumour mass; this may involve BCS or a full mastectomy, depending upon the response to neoadjuvant chemotherapy. Radiotherapy may then also be administered which aims to target any residual tumour cells thus maximising eradication of the tumour and reducing the chance of recurrence.

The European Organization for Research and Treatment of Cancer (EORTC) Trial 10902 determined that neoadjuvant chemotherapy does increase rates of BCS, however it shows no difference in overall survival (OS) or disease free survival (DFS) (van der Hage, van de Velde et al. 2001). The NSABP-B-18 randomized clinical trial demonstrated the effectiveness of doxorubicin / cyclophosphamide combination neoadjuvant chemotherapy, where 80% patients showed $\geq 50\%$ reduction in tumour size, of which 36% showed a complete clinical response (Fisher, Bryant et al. 1998). This is discussed further in section 2.3.

1.6 Prognostic and Predictive Factors

Prognostic and predictive factors in breast cancer are considered when decisions regarding disease management and therapeutic strategy are made. Prognostic factors are used to 'predict patient clinical outcome independently of treatment', whereas predictive factors

predict the ‘response of a patient to a specific therapeutic modality, as well as tumour sensitivity or resistance to the therapy’ (Liu, Huang et al. 2010; Weigel and Dowsett 2010). However, several factors have both *prognostic* and *predictive* relevance, thus enabling the clinician to tailor treatment accordingly, as well as having some knowledge of how the tumour may respond.

1.6.1 Prognostic factors

The main prognostic factors in breast cancer include histological type, grade, lymph node involvement, lymphovascular invasion, metastasis, hormone receptor status, proliferation rate of tumour cells and tumour size. The TNM staging system (Table 1) encompasses three of these factors, which are of greatest prognostic value; tumour size, lymph node involvement and metastasis. A useful prognostic tool, for invasive breast carcinoma patients, which combines nodal status, tumour size and histological grade, is the Nottingham Prognostic Index (NPI). These three factors were initially found to be independently associated with survival upon multivariate analysis, and were combined to form the NPI, which has now been validated in several studies (Lee and Ellis 2008). Patients are grouped into one of six prognostic groups, ranging from ‘excellent prognostic group’ to ‘very poor prognostic group’ (Blamey, Ellis et al. 2007).

On a molecular level, there are few biomarkers that have been transferred to the clinic as routine prognostic markers. To date, the only routine markers currently used include estrogen receptor (ER) and human epidermal growth factor 2 (HER2). A major adverse prognostic factor is *ERBB2* (*HER2*) gene amplification and protein over-expression; this is associated with increased risk of relapse and shorter overall survival (section 2.2.2) (Ross, Fletcher et al. 2003; Weigel and Dowsett 2010).

More recently, gene expression profiling has discovered signatures with prognostic abilities. The discovery of molecular subtypes within breast cancer (Perou, Sorlie et al. 2000) allows tumours to be classified according to their hormone receptor expression, and also presents a powerful prognostic tool, as each of these subtypes display very different prognoses (section 2.2.2). Gene expression signatures have also been discovered which predict disease outcome, the most significant of which is the 70-gene signature; Mammaprint[®], which is able to distinguish between good and bad prognosis tumours (van 't Veer, Dai et al. 2002; van de Vijver, He et al. 2002). This signature has been customised into the high-throughput diagnostic array 'Mammaprint[®]' (Glas, Floore et al. 2006) and is currently incorporated in the prospective randomised Phase III Clinical Trial 'MINDACT' (Microarray In Node negative Disease may Avoid ChemoTherapy) (Cardoso, Piccart-Gebhart et al. 2007). Another gene signature with prognostic ability is Oncotype DX[®]; a 21-gene signature, which is able to predict recurrence in ER positive tumours treated with tamoxifen (Paik, Shak et al. 2004; Albain, Barlow et al. 2010), and is currently being tested for its clinical use in the Phase III Clinical Trial 'TAILORx'. These, along with other gene expression signatures have been extensively reviewed (Bonnetoi, Underhill et al. 2009; de Snoo, Bender et al. 2009; Slodkowska and Ross 2009; Sotiriou and Pusztai 2009; Stadler and Come 2009; Espinosa, Vara et al. 2011).

1.6.2 Predictive factors

There are contradicting arguments regarding tumour size and therapy response, however it is largely believed to show inverse correlation; smaller tumours show a better response to treatment (Fernandez-Sanchez, Gamboa-Dominguez et al. 2006; Mieog, Hage et al. 2006). Different tumour types also respond differently to neoadjuvant chemotherapy. Invasive lobular carcinomas have been shown to have lower rates of complete response to

neoadjuvant anthracycline/taxane chemotherapy when compared to invasive ductal carcinomas; 3% and 15% response rates respectively (Cristofanilli, Gonzalez-Angulo et al. 2005). However, overall outcome was shown to be better for invasive lobular carcinoma patients, who showed longer rates of DFS and OS (Cristofanilli, Gonzalez-Angulo et al. 2005). On a molecular level, the expression of ER and HER2 has also been associated with response to neoadjuvant anthracycline chemotherapy in breast cancer. Tumours displaying over-expression of HER2 are associated with high rates of pathological complete response (pCR) to neoadjuvant chemotherapy (Weigel and Dowsett 2010) (section 2.4.2.7). In contrast, tumours showing ER expression have demonstrated poor rates of pCR to neoadjuvant chemotherapy (section 2.4.2.7) (Kim, Sohn et al. 2010). The above-mentioned gene expression profiling and the discovery of gene expression signatures have also been associated with the prediction of therapy response (section 2.4.3).

1.7 Monitoring response to treatment

Neoadjuvant chemotherapy is not beneficial for all patients; some tumours display resistance to chemotherapy, where progression may be seen with an increase in tumour size, and require early surgical intervention. It is not currently possible to distinguish between those patients who will benefit from neoadjuvant chemotherapy and those who will not. For this reason, tumour response to treatment is carefully monitored clinically and radiologically, to avoid continuation of neoadjuvant chemotherapy treatment in the event of tumour progression. Radiological imaging using dynamic contrast enhanced-magnetic resonance imaging has been shown to be more accurate and sensitive than traditional x-ray mammography and ultrasound methods for the assessment of breast tumours (Julius, Kemp et al. 2005; McLaughlin and Hylton 2011). Tumour response can be assessed radiologically, using RECIST (Response Evaluation Criteria In Solid Tumours) criteria

(Therasse, Arbuck et al. 2000; Eisenhauer, Therasse et al. 2009) by measuring the longest diameter of the tumour before, during and after the treatment course. RECIST guidelines were introduced to internationally standardise the reporting of tumour response to therapy when new chemotherapeutic agents are tested in clinical trials. These criteria can therefore be used to assess tumour response to therapy in the clinical setting. Tumour response is assessed as described in Table 2 (Therasse, Arbuck et al. 2000). Occurrence of new metastasis during neoadjuvant chemotherapy must also be considered as progression regardless of changes in tumour size. Histopathological reports from the tumour resection must also be acknowledged to fully assess response to neoadjuvant chemotherapy. It has been reported that after neoadjuvant chemotherapy, particularly in ER+/HER2- tumours, when assessing the extent of residual disease, there may be inaccuracies when using dynamic contrast enhanced-magnetic resonance imaging (DCE-MRI) measurements, in comparison to histopathological measurements (Chen, Feig et al. 2008; Wright, Zubovits et al. 2010; Loo, Straver et al. 2011). This should therefore be considered when determining therapy response; tumour measurements determined within the final pathology report may be the most accurate parameter.

When tumour response to treatment is determined, decisions regarding therapy and its continuation or termination can be made. In the event that tumours do not respond, or even progress during treatment, the course may be aborted and a different approach may be taken to treat the tumour. A pathological complete response (pCR) following neoadjuvant chemotherapy, is a strong indicator of survival and has been associated with highest rates of DFS (Kuerer, Newman et al. 1999; Smith, Heys et al. 2002; Ladoire, Arnould et al. 2008).

Table 2: Summary of the Response Evaluation Criteria In Solid Tumours (RECIST) Guidelines (Therasse, Arbuck et al. 2000; Eisenhauer, Therasse et al. 2009)

<i>Response</i>	<i>Criteria</i>
Complete Response (CR)	Complete disappearance of the tumour
Partial Response (PR)	≥ 30% decrease in tumour size
Stable Disease (SD)	< 30% decrease and < 20% increase in tumour size
Progressive Disease (PD)	≥ 20% increase in tumour size

1.8 Resistance to Treatment

Using neoadjuvant chemotherapy as an example, a major pitfall in effective treatment is tumours which display resistance to the chemotherapeutic agents targeting them. The desired effect when treating a tumour with neoadjuvant chemotherapy would be its complete disappearance. However, chemotherapeutic drugs may not have this desired effect; they may not have any effect and the tumour may not change in size. More detrimentally, the tumour may increase in size during therapy, allowing disease progression. In this situation, the patient will have been exposed to high levels of cytotoxic agents and suffered unpleasant side-effects for no therapeutic gain. The ability to predict whether an individual tumour will respond to neoadjuvant chemotherapy or display resistance to it would therefore be extremely valuable. Within a clinical setting, this may involve screening diagnostic core biopsy samples for the expression of predictive biomarkers of response. Treatment could then be tailored accordingly on an individual patient basis, selecting the most appropriate and effective treatment, with minimal side effects

CHAPTER 2:
MOLECULAR INTRODUCTION TO
BREAST CANCER

Chapter 2. Molecular Introduction to Breast Cancer

2.1 Normal Tissue Homeostasis

Efficient functioning of a system, organ or tissue relies on the maintenance of a carefully controlled and regulated, stable internal environment; *homeostasis*, which provides the perfect balance of cell proliferation and cell death (Vermeulen, Berneman et al. 2003). Tissue homeostasis is a complex process, involving a fine balance between several critical pathways, the principle ones being the cell division cycle, DNA damage response, DNA repair and apoptosis.

2.1.1 The Cell Cycle

The cycle of cell division is a four-stage process by which cells ultimately divide to produce two identical daughter cells, that is tightly controlled by cyclins and cyclin-dependent kinases (cdk's) (Giacinti and Giordano 2006). The two main phases are the S phase, where DNA is replicated, and the M phase, where the cell divides to produce two daughter cells. Between each of these two phases, there are 'gaps', where cells grow in response to growth signals; G_1 and G_2 (Figure 5) (Garrett 2001; Shah and Schwartz 2001). Where cell division is not appropriate, the cell is able to reversibly enter the G_0 phase. At critical points within the cell cycle, such as prior to transition to the subsequent phase, DNA damage checkpoints exist, which ensure that the cell cycle does not proceed in the presence of DNA damage. Four of the main checkpoints include the G_1/S checkpoint, the intra-S-phase checkpoint, the G_2/M checkpoint and the spindle checkpoint. The G_1/S and G_2/M checkpoints are able to induce cell cycle arrest in the presence of DNA damage; however the G_2/M checkpoint also monitors DNA replication. The intra-S-phase checkpoint recognises stalled replication forks and prevents mitosis of cells where DNA

replication is incomplete (Houtgraaf, Versmissen et al. 2006). The spindle checkpoint monitors the formation of the mitotic spindle, and prevents mitosis where spindles are incorrectly formed (Garrett 2001). An additional checkpoint is the restriction point, which differs from the above-mentioned checkpoints as it does not specifically assess the genome (Garrett 2001). The restriction point ensures that cells have received sufficient growth signals to be able to complete one cycle of cell division; where this is not shown, cells enter the G₀ phase (Garrett 2001) (Figure 5). The regulation of cell cycle progression is determined by a number of factors, including the association of cdk's with their respective cyclins, phosphorylation state, interaction with cdk inhibitors (INK4 and CIP/KIP families) and the specific proteolysis of cyclins via the ubiquitin-proteasome pathway (Nakayama and Nakayama 2005).

2.1.2 DNA damage response pathway

The DNA damage response (DDR) is a complex mechanism, involving a balance between the phosphorylation, ubiquitination, acetylation and sumoylation of the main promoters of the pathway, which is required to initiate the DNA damage signal (Huen and Chen 2008). Mediators involved in the response are also dependent upon the type of DNA damage; DNA single-strand breaks (SSBs), DNA double-strand breaks (DSBs) or formation of DNA adducts. Two of the most important proteins involved in the DNA damage response are Ataxia telangiectasia mutated (ATM) and ATM and RAD3-related (ATR). They respond to DSBs (ATM), SSBs (ATR) and stalled DNA replication forks (ATR), and their main substrates are Chk2, p53 (ATM) and Chk1 (ATR) (Hurley and Bunz 2007; Flynn and Zou 2011). The phosphorylation of their substrates leads to the activation of important regulators of the cell cycle, and the regulation of S phase (G₁/S checkpoint) and M phase (G₂/M checkpoint) progression through the cell cycle (section 2.1.1) (Houtgraaf,

Versmissen et al. 2006; Hurley and Bunz 2007; Huen and Chen 2008). The activation of several important proteins is dependent upon ATM/ATR. An example of such a protein is H2AX, which is one of the mediators of the initial signal, but more importantly is responsible for the accumulation of MDC1. This is an important regulator of damaged chromatin which is also responsible for the accumulation of other important mediators, such as BRCA1 (Huen and Chen 2008), which is important for DNA damage-induced cell cycle checkpoint activation (Wu, Lu et al. 2010). An important regulator of the cell cycle, which is able to respond to DNA damage, is the p53 tumour suppressor protein, encoded by the *TP53* gene. In the event of DNA damage, it is able to induce cell cycle arrest, via proteins such as p21 (CIP1/WAF1) (Figure 5), after which DNA repair or apoptosis may be initiated (sections 2.1.3 and 2.1.4). Upon detection of DNA damage, p53 is stabilised, which then up-regulates p21, which is able to bind to the cyclin D/cdk 4 complex and prevent G₁/S transition, as well as the cyclin B/cdk1 complex which prevents G₂/M transition. Up-regulation of p21 at the G₂/M checkpoint by p53 is accompanied by the up-regulation of 14-3-3 sigma which is able to sequester the cyclin B/cdk1 complex and prevent it reaching its nuclear targets thus preventing cell cycle progression (Garrett 2001). The effect is accentuated by the activity of ATM; upon DNA damage detection, ATM is up-regulated, which activates proteins including Chk2. Chk2 is able to block degradation of p53 by MDM2, which is an E3 protein ligase involved in the ubiquitin-proteasome degradation pathway targeting p53 for degradation by the 26S proteasome. Chk2 therefore acts by stabilising p53 levels thus promoting induction of cell cycle arrest (Garrett 2001).

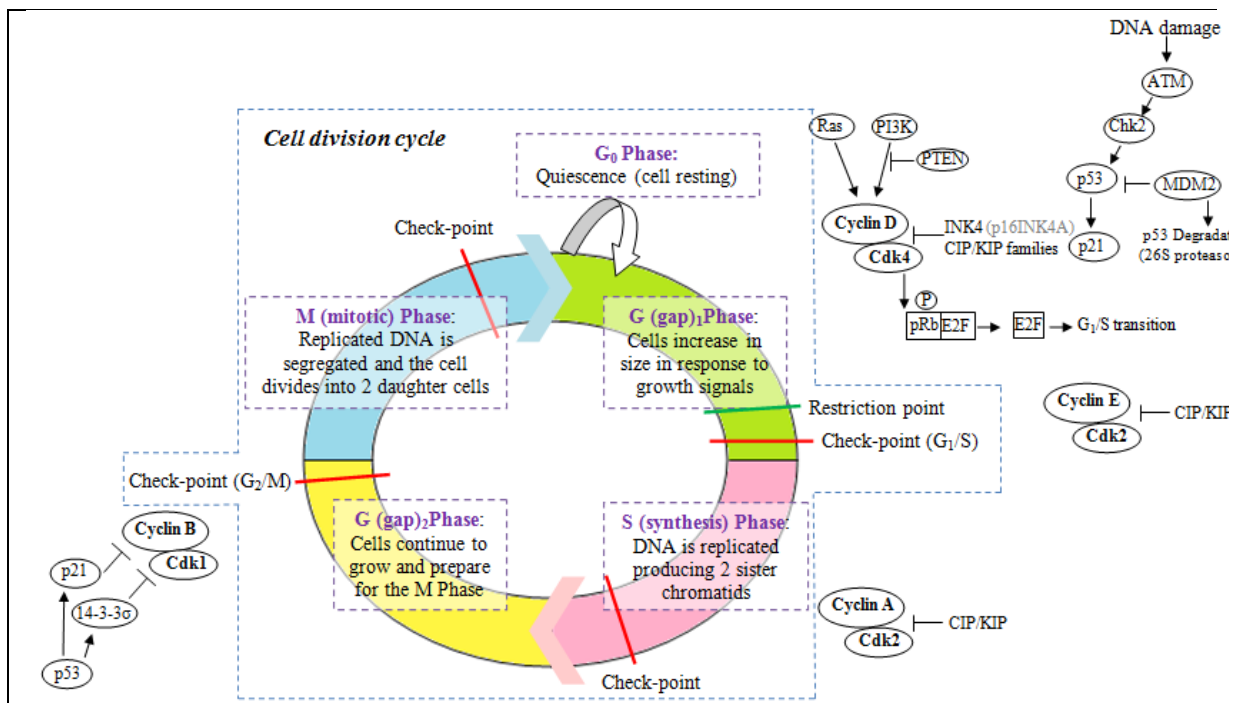


Figure 5: Overview of the cell division cycle and DNA damage response

The cycle of cell division consists of 4 sequential phases; G_1 , S, G_2 and M. Progression through the cell cycle involves cell growth in response to growth signals (G_1), replication of DNA to produce 2 sister chromatids (S), further cell growth in preparation for division (G_2) and the division the cell into two daughter cells, each containing identical DNA copies (M). Throughout the cycle, there are several DNA damage checkpoints, to ensure damaged cells do not proliferate. The restriction point is a different type of checkpoint, which assesses whether the cell has received sufficient growth signals to complete one round of the cell division cycle, rather than detecting DNA damage. The whole process is tightly controlled and regulated by cyclins and cyclin-dependent kinases (cdk's), which are themselves controlled by important proteins such as p53, the 'guardian of the genome', and ATM. In the presence of DNA damage, caused by reactive oxygen species for example, these proteins have the ability to arrest the cell cycle, and initiate DNA repair mechanisms or apoptosis (Garrett 2001; Nakayama and Nakayama 2005; Giacinti and Giordano 2006). Upon detection of DNA damage, p53 is up-regulated and initiates the expression of genes which cause cell cycle arrest (or apoptosis). An example of such a protein is p21 which, when initiated by p53, binds to the cyclin D/cdk4 complex, causing cycle arrest at the G_1/S transition. The p21 protein can also cause G_2/M arrest by binding to the cyclin B/cdk1 complex. At the G_2/M checkpoint, p53 also causes up-regulation of 14-3-3 σ , which is able to sequester the cyclin B/cdk1 complex and prevent it reaching its nuclear targets. Another protein sensitive to DNA damage, ATM, further stabilises p53 by initiating Chk2 expression which is able to block the action of MDM2, an E3 protein ligase promoting p53 ubiquitin-proteasomal degradation; therefore promoting cell cycle arrest (Garrett 2001).

2.1.3 DNA repair

There are several mechanisms of DNA repair, which are employed depending upon the type of damage incurred and the stage of the cell cycle (Branzei and Foiani 2008). The main types of DNA repair include mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), homologous recombination (HR) and non-homologous end joining (NHEJ) and repair via the Fanconi anemia (FA) pathway (Christmann, Tomicic et al. 2003; Branzei and Foiani 2008; Reed 2010). The most genotoxic type of DNA-damaging lesion is a double strand break (DSB), which is exerted by many chemotherapeutic agents (section 2.3) (Asakawa, Koizumi et al. 2010), and it has been reported that a single un-repaired DSB in an essential gene is enough to cause cell death by apoptosis (Christmann, Tomicic et al. 2003). In the event of a DSB, the DNA repair pathways utilised to repair the damage include HR and NHEJ (Houtgraaf, Versmissen et al. 2006), which are less error-prone and more error-prone respectively (Christmann, Tomicic et al. 2003). For agents which form DNA-adducts, or intra-strand cross-links (ICLs), such as bifunctional alkylating agents (section 2.3.2), a more complex DNA repair mechanism is required, involving a combination of NER (Houtgraaf, Versmissen et al. 2006) the FA pathway, translesion synthesis and HR (Kondo, Takahashi et al. 2010; Reed 2010), which is also dependent upon the stage in the cell cycle (Vasquez 2010). ICLs cause blocked replication forks, which lead to the activation of ATR and subsequent activation of members of the FA core complex (Kratz, Schopf et al. 2010). This FA complex catalyses the monoubiquitination of FANCD2 (MacKay, Declais et al. 2010), which then travels to the site of DNA damage, where it co-localises with a recently reported nuclease, FANL1 (Smogorzewska, Desetty et al. 2010). FANL1 has been reported as a repair nuclease, which exhibits exonuclease and endonuclease activity (Kratz, Schopf et al. 2010; MacKay,

Declais et al. 2010; Smogorzewska, Desetty et al. 2010), yet its exact role is not fully understood. Subsequently, incisions are made either side of the ICL, involving MUS81-EME1 and possibly the NER proteins ERCC1 and XPF, creating DSBs (Kondo, Takahashi et al. 2010), to ‘unhook’ the ICL (MacKay, Declais et al. 2010). Translesion synthesis then fills the gap, involving Rev3, Rev7 and Rev1 (Kondo, Takahashi et al. 2010), and the ‘unhooked’ ICL lesion is subsequently removed by excision repair (MacKay, Declais et al. 2010). HR is then initiated to complete the repair process and repair DSBs, but may also involve NHEJ (Kondo, Takahashi et al. 2010).

For the repair of DSB’s by HR, a template is required, in the form of the sister chromatid, so this repair mechanism can only function after DNA replication has occurred in the late S phase or G₂ phase of the cell cycle (Figure 5). Following recognition of the DSB and the recruitment of several required activation mediators (ATM, BRCA1, BRCA2, RAP80 etc), HR is mediated by the Rad52 epistasis group, involving Rad51, Rad52 and Rad54. Nucleolytic resection of the DSB is initiated by MRN (MRE-11-Rad50-NSB1) complex, after which Rad52 binds to the exposed 3’ ends. Interaction between Rad52, RPA and Rad51, which has located the intact copy of the section of the genome on the sister chromatid, leads to sister chromatid exchange and subsequent DNA synthesis and ligation (Christmann, Tomicic et al. 2003; Houtgraaf, Versmissen et al. 2006; Branzei and Foiani 2008; Asakawa, Koizumi et al. 2010; de Campos-Nebel, Larripa et al. 2010; Zou 2010). In contrast, the NHEJ DNA repair mechanism does not require a template and usually takes place in the G₀/G₁ phase of the cell cycle. NHEJ is performed by the Ku heterodimeric protein, comprised of Ku78 and Ku80 binding to the two ends of a DSB, following which the two ends are joined by DNA-PK and the DNA ligase IV/XRCC4 complex (Houtgraaf, Versmissen et al. 2006; Branzei and Foiani 2008). Efficient DNA repair is essential for the

maintenance of genomic stability, and is therefore a critical component of cell and tissue homeostasis.

2.1.4 Apoptosis

Apoptosis was initially described in 1972 by John Kerrs, and is a physiological form of cell death (Lawen 2003). It is a protective mechanism, also involved in normal growth and development, which acts to remove damaged or unwanted cells without raising an inflammatory response (Lawen 2003), thus maintaining tissue homeostasis. Several of the proteins involved in cell cycle control and proliferation are also involved in apoptosis, thus despite being two distinct processes, they are closely related (Shah and Schwartz 2001; Vermeulen, Berneman et al. 2003). There are two main pathways involved in the initiation of apoptosis, which are referred to as the 'intrinsic', or 'mitochondrial', and 'extrinsic' pathways however both converge at the activation of the executioner caspase, caspase 3 (Hengartner 2000) (Figure 6). The intrinsic pathway, which progresses via the mitochondria, is initiated following a variety of internal signals, such as the DNA damage pathway, and is thought to be the pathway most widely associated with cancer pathogenesis (Hanahan and Weinberg 2011).

2.1.4.1 Intrinsic pathway

The intrinsic pathway of apoptosis is mediated via the mitochondria and is tightly controlled by the balance between anti- and pro-apoptotic proteins, largely members of the Bcl-2 family (Ghobrial, Witzig et al. 2005). Two of the main proteins required for the initiation of apoptosis by the intrinsic pathway are cytochrome c, which is required for caspase activation, and apoptosis protease-activating factor-1 (Apaf-1) (Chowdhury, Tharakan et al. 2006). In the event of DNA damage, or following a death signal, BH3-only

domain proteins, such as Bid, Bad and PUMA, are activated and transfer signals to the mitochondria where they facilitate the accumulation of pro-apoptotic proteins, such as Bax and Bak, at the mitochondrial membrane (Ghobrial, Witzig et al. 2005; Chowdhury, Tharakan et al. 2006). These proteins associate with the outer mitochondrial membrane, and a change in membrane permeability results in the release of apoptosis-inducing factor (AIF) and cytochrome c (Hengartner 2000; Chowdhury, Tharakan et al. 2006). Smac/DIABLO is also released from the mitochondria during apoptosis initiation, which binds to inhibitors-of-apoptosis (IAP) proteins, preventing their anti-apoptotic activity (Hengartner 2000). Cytochrome c, along with pro-caspase 9, binds to Apaf-1 to form the apoptosome (Harrington, Ho et al. 2008). This then activates the caspase cascade, resulting in execution of apoptosis via nuclear breakdown, resulting in cell death (Ghobrial, Witzig et al. 2005) (Figure 6).

2.1.4.2 Extrinsic pathway

The extrinsic pathway involves death receptors from the tumour necrosis factor (TNF) superfamily, such as death receptor 4 (DR4) being activated by interaction with a death ligand, such as the tumour necrosis factor-related apoptosis inducing ligand (TRAIL). Upon activation, death receptors trimerize and form a death-inducing signalling complex (DISC), involving recruitment of Fas-associated death domain (FADD) protein and pro-caspase-8 (Fadeel and Orrenius 2005; Jin and El-Deiry 2005; Fulda and Debatin 2006; Falschlehner, Emmerich et al. 2007; Harrington, Ho et al. 2008; Indran, Tufo et al. 2011). Activation of caspase 8 subsequently leads to the activation of the main executioner caspase of apoptosis; caspase 3 (Walsh, Cullen et al. 2008). This can occur directly or indirectly, via the mitochondrial pathway and cleavage of Bid (Fadeel and Orrenius 2005; Fulda and Debatin 2006) (Figure 6).

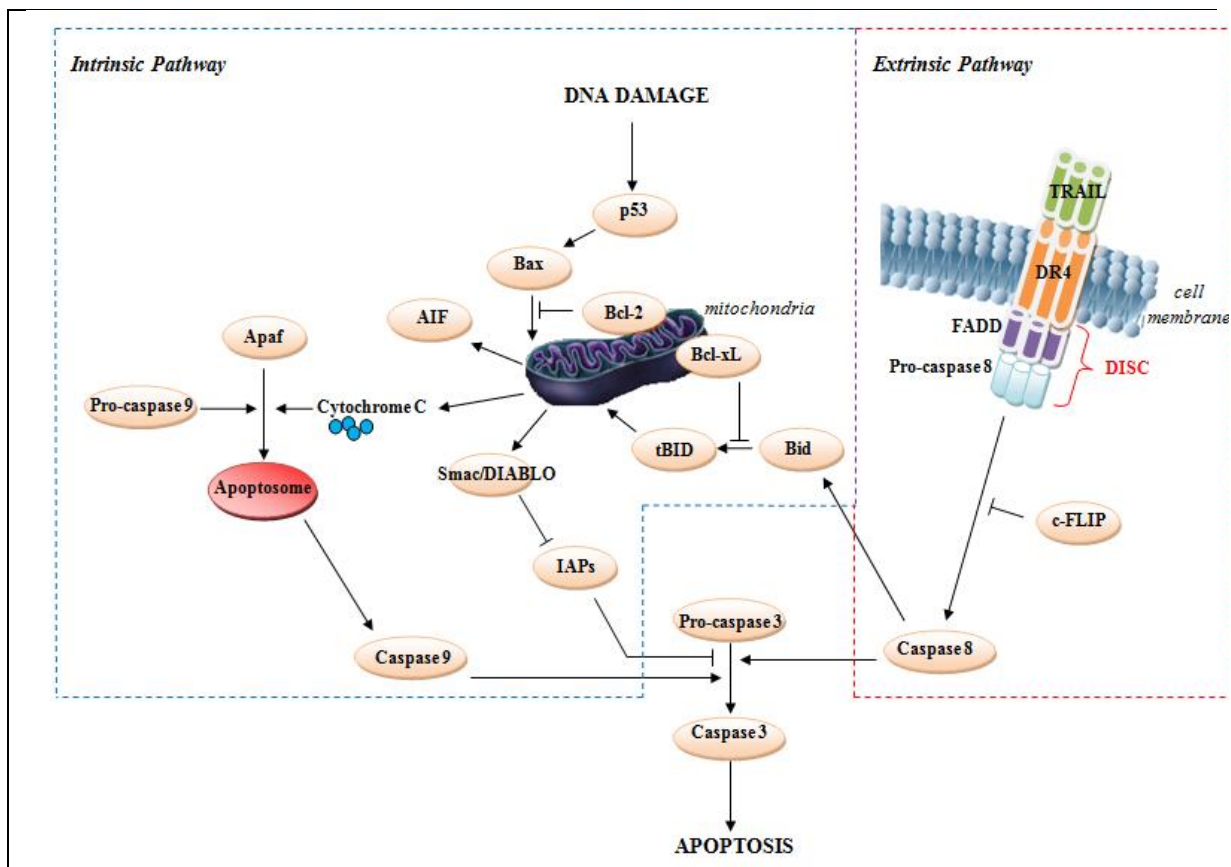


Figure 6: Apoptosis Pathways

The extrinsic pathway of apoptosis (right) is initiated by the binding of a death ligand, such as TRAIL, to a death receptor, such as DR4. Recruitment of FADD and pro-caspase 8 molecules form the death-inducing signalling complex (DISC). This leads to the activation of caspase 8, which can be inhibited by c-FLIP, and subsequent activation of caspase 3, leading to apoptosis. Caspase 8 can activate caspase 3 either directly or via the intrinsic, mitochondrial, pathway by interaction with Bid, which is activated when it is cleaved to truncated (t)Bid. The intrinsic pathway of apoptosis (left) is initiated in response to a death signal, such as DNA damage. Pro-apoptotic members of the Bcl-2 family (Bid, Bax) transfer the signal to the mitochondria, which can be inhibited by anti-apoptotic members (Bcl-2, Bcl-xL). Changes in mitochondrial membrane permeability cause the release of cytochrome c, along with apoptosis-inducing-factor (AIF) and Smac/DIABLO, which acts to neutralize the activity of inhibitors-of-apoptosis (IAP) proteins. Cytochrome c binds to Apaf-1, along with pro-caspase 9, to form the apoptosome. This then activates the caspase cascade, resulting in apoptosis. The two pathways converge at the executioner caspase, caspase 3, from which apoptosis follows by nuclear degradation (Hengartner 2000; Harrington, Ho et al. 2008).

To summarise, some of the main pathways involved in tissue homeostasis; cell division cycle, DNA damage response, DNA repair and apoptosis, each involve a plethora of

different mediators, which are tightly controlled and regulated, by processes including phosphorylation and degradation by the ubiquitin-proteasome pathway. There is a relevant degree of cross-talk between the pathways, and not all mediators are mutually exclusive to any one of the pathways; for example the p53 protein is able to induce cell cycle arrest as well as initiate apoptosis.

In the event of DNA damage, the DNA damage response pathway is initiated, and subsequently the cell cycle is arrested via activation of the cell cycle checkpoints. The cell is then given time to repair the damage, following which cell cycle progression may continue if the repair is successful. Where this is not possible, the cell should be directed towards apoptosis (Figure 7) or entry into a quiescent state. This process is reliant upon effective functioning and balance between tumour suppressor genes and oncogenes as well as control of the cell division cycle. Incorrect functioning of this process may leave unrepaired DNA damage, which may cause mutations leading to genomic instability and ultimately oncogenesis (Houtgraaf, Versmissen et al. 2006).

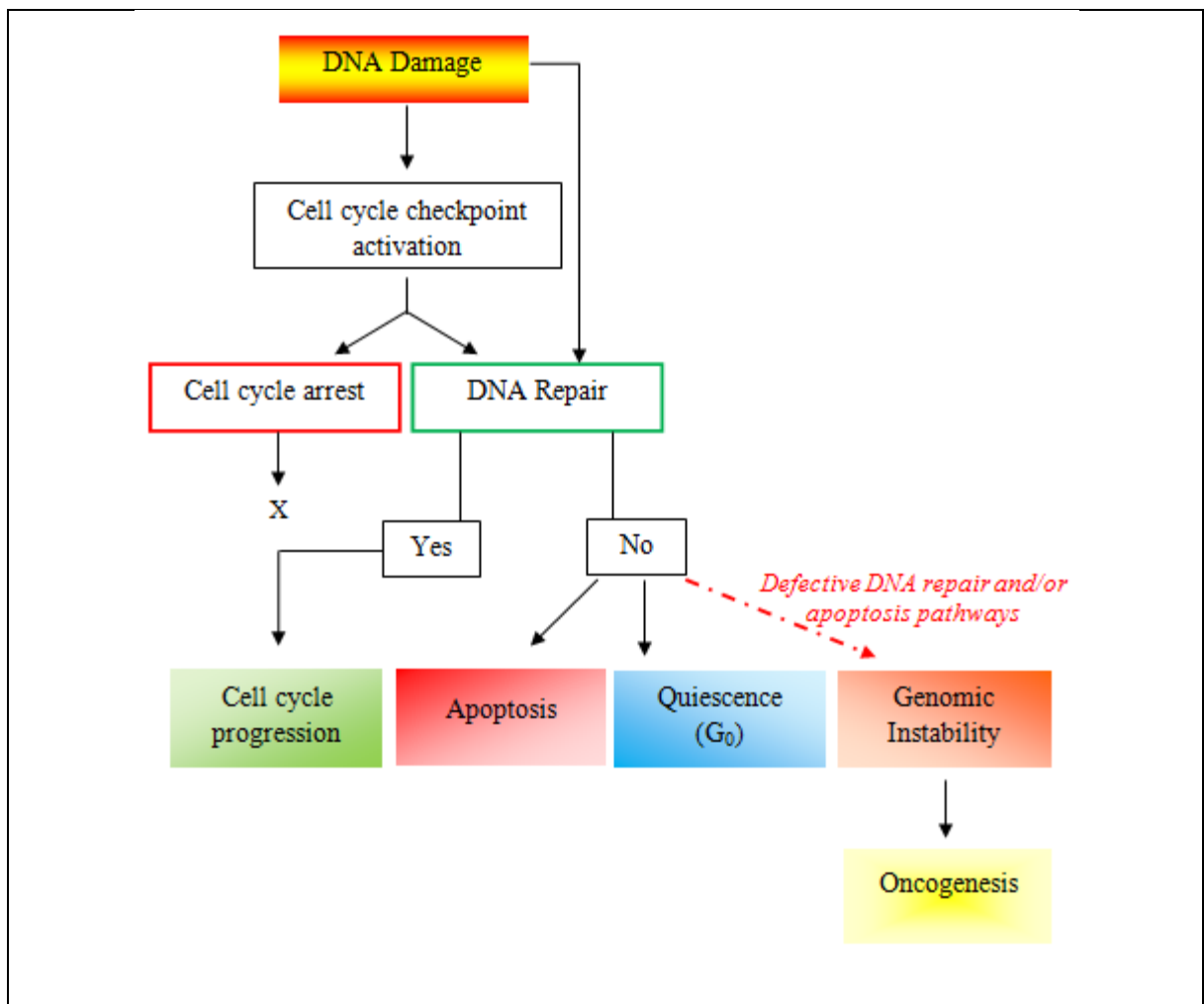


Figure 7: Summary of response to DNA damage

In the event of DNA damage, the DNA damage response initiates DNA repair either directly or via the activation of checkpoints within the cell cycle. Cell cycle arrest is also initiated to allow time for damaged DNA to be repaired. If DNA repair is successful, the cell cycle may progress. However if DNA repair is not successful it may either enter a quiescent state or may be directed towards apoptosis. Where DNA repair is unsuccessful, and apoptosis is not initiated, mutations may arise causing genomic instability, which may cause oncogenesis (Houtgraaf, Versmissen et al. 2006).

Adapted from (Houtgraaf, Versmissen et al. 2006).

2.2 Breast Cancer

2.2.1 Hallmarks of Cancer

Oncogenesis, or carcinogenesis, is a multi-factoral process by which a cell is transformed from a normal cell into a malignant cancer cell. The seminal paper by Hanahan and Weinberg, published in 2000, described the ‘hallmarks of cancer’; the alterations a cell has to undergo and characteristics it develops in order for it to be regarded as a malignant cell. These include *self-sufficiency in growth signals*, *insensitivity to growth inhibitory signals*, *evasion of programmed cell death (apoptosis)*, *limitless replicative potential*, *sustained angiogenesis* and *tissue invasion and metastasis*, with underlying *genomic instability* (Hanahan and Weinberg 2000). This has been recently reviewed following substantial progress in understanding over the last decade, and two emerging hallmarks have been added, in addition to the initial six hallmarks; *reprogramming of energy metabolism* and *evading immune destruction* (Hanahan and Weinberg 2011). The importance of the tumour microenvironment, and the heterotypic signalling within it which allow the development of complex malignancies, has also been emphasised (Hanahan and Weinberg 2011).

For each type of cancer, there are common genetic alterations, which may or may not be hereditary, which may act as a predisposition towards the changes in cell characteristics listed above and overall disruption of normal tissue homeostasis. Generally, these are likely to involve gain-of-function mutations in oncogenes, which promote survival and proliferation, and loss-of-function mutations in tumour suppressor genes, which normally protect against uncontrolled proliferation and promote DNA repair and cell cycle checkpoint control (Lee and Muller 2010). In breast cancer, frequently altered genes include *ERBB2 (HER2)*, *PI3K*, *TP53*, *BRCA1*, *BRCA2* and *PTEN*, as well as important regulators of the cell cycle; cyclins and cyclin-dependent kinases, which have been

extensively studied (Osborne, Wilson et al. 2004; Lee and Muller 2010). However, a popular approach is to obtain gene expression profile signatures, rather than the study of specific genes/proteins, which accommodates the complexity of the disease. There is an enormously high level of molecular heterogeneity within breast cancer, even between tumours of the same type, grade and stage, and is complicated by the fact that cells within the tissue are at different levels of differentiation and maturation (Keller, Lin et al. 2010). It is therefore imperative to understand the disease at the molecular level, in order to improve understanding of factors affecting prognosis and treatment response

2.2.2 Breast cancer at a molecular level

The molecular diversity within breast tumours has been demonstrated using global expression microarray analysis of mRNA (Perou, Sorlie et al. 2000; Sorlie, Perou et al. 2001). This highlighted variation in expression patterns between tumours, and revealed five molecular subtypes of breast cancer, which are biologically different and exhibit distinct clinical behaviour. The ERBB2+ subtype shows high expression of human epidermal growth factor receptor (HER2). Luminal subtypes (A and B) demonstrate estrogen receptor (ER) expression. Basal-like tumours are negative for ER, progesterone receptor (PR) and HER2, and are also known as 'triple-negative' tumours. These subtypes reflect the distinct luminal and basal epithelial cells found in the mammary gland, with basal cells (cytokeratin 5/6 positive) lying closest to the basement membrane and luminal cells forming the upper differentiated layer. The fifth subtype, normal-like tumours are less characterised, but resemble normal breast tissue. More recently, another molecular subtype was identified; claudin-low, the majority of which are negative for the ER, PR and HER2 (triple negative), but can be distinguished by their expression of tight junction proteins claudin 3,4 and 7 as well as E-cadherin (Prat, Parker et al. 2010; Prat and Perou 2011). Luminal A subtype

tumours, which have high ER expression, display the best prognosis, demonstrated by longest time to development of distant metastasis and longest overall survival, whereas ERBB2+ and triple negative tumours are associated with poor prognosis, demonstrated by shortest time to development of distant metastasis and shortest overall survival time (Figure 8) (Sorlie, Perou et al. 2001; Sorlie, Tibshirani et al. 2003; Cheang, Voduc et al. 2008). A panel of four immunohistochemical markers (PR, ER, HER2 and Ki67) has been identified which can distinguish between Luminal A and Luminal B tumours (Cheang, Chia et al. 2009). Luminal B tumours have been shown to have higher rates of proliferative gene expression (*MKI67*, *CCNB1* and *MYBL2*) and 30% of Luminal B tumours were found to express HER2 and associated genes (*ERBB2* and *GRB7*) (Cheang, Chia et al. 2009). *TP53* gene mutations were seen in 13% of Luminal A tumours, 71% of ERBB2+ and 82% of triple negative tumours (Sorlie, Perou et al. 2001). Each of these tumours clearly display very different prognoses and can therefore almost be regarded as different diseases (Figure 8). This highlights and emphasises the heterogeneity of breast cancer as a disease, and even within the above-mentioned molecular subtypes a high degree of heterogeneity is still observed, especially within the luminal subtypes.

Determination of molecular subtype, by routine testing of ER PR and HER2 receptor status, is therefore used as a prognostic marker, and to aid determination of therapeutic strategy. Molecular subtype has also been associated with response to neoadjuvant chemotherapy (section 2.4.2.7). This is confirmed from the core needle biopsy taken at the diagnostic stage, using immunohistochemistry and/or fluorescent-in-situ-hybridisation (FISH). The HercepTest™ (Dako) may also be used for the determination of HER2 status, however routine immunohistochemistry can also be performed using the A0485 antibody (Selvarajan, Bay et al. 2004).

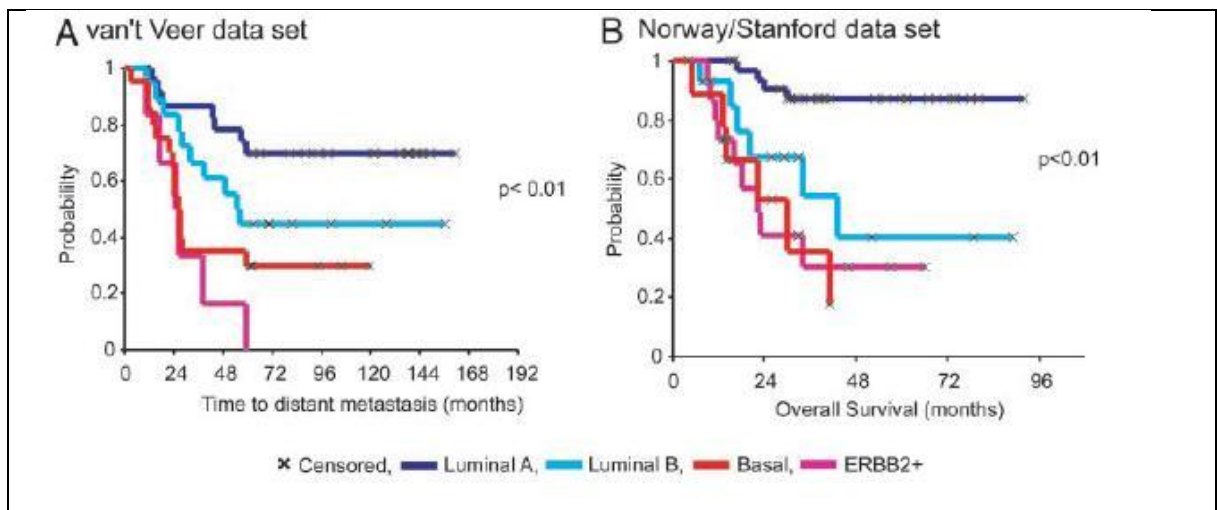


Figure 8: Differences in clinical outcome between molecular subtypes of breast cancer Kaplan-Meier analysis (Sorlie, Tibshirani et al. 2003), showing differences in clinical outcome between subtypes, identified by gene expression analysis of two patient cohorts; A: time to distant metastasis development, from analysis of previously published gene expression data by van't Veer *et al* (van 't Veer, Dai et al. 2002) and B: overall survival, from gene expression analysis of the Norway/Stanford cohort, containing 122 tissue samples (Diehn, Sherlock et al. 2003). Both sets of data show Luminal A tumours to have the greatest disease-free survival period as well as overall survival (Sorlie, Tibshirani et al. 2003).

2.3 Breast Cancer treatment using neoadjuvant chemotherapy

Neoadjuvant chemotherapy is used as the primary treatment for locally advanced breast cancer, and aims to reduce the volume of the tumour prior to surgical resection of the remaining tumour mass, thus increasing the likelihood of breast conserving surgery (BCS) rather than a full mastectomy. Chemotherapeutic agents ultimately aim to inhibit the proliferation of tumour cells (section 2.1.1) and cause cell death by induction of apoptosis, (section 2.1.4) potentially via the DNA damage response pathway (section 2.1.2). An anthracycline in combination with cyclophosphamide, followed by a taxane is currently the gold standard regimen for neoadjuvant chemotherapy (Chuthapisith, Eremin et al. 2006); a popular example of which comprises epirubicin/cyclophosphamide (4 cycles) and sequential docetaxel (4 cycles), which is currently used within the Hull and East Yorkshire

NHS Trust Hospitals. The National and Surgical Adjuvant Breast and Bowel Project (NSABP)-27 trial demonstrated increased rates of pathological complete response (pCR) when docetaxel was used following anthracycline/cyclophosphamide (Buzdar 2007).

2.3.1 Anthracyclines

The first of the antibiotic group of anthracyclines to be introduced were doxorubicin and daunorubicin in the 1960's (Minotti, Menna et al. 2004), however problems with cardiotoxicities led to the development and introduction of epirubicin. Epirubicin is a stereoisomer of doxorubicin, showing only a difference in the orientation of the hydroxyl atom (axial-to-equatorial epimerization) on the C-4' of the hexapyranosyl sugar, which has little effect on the mode of action but shows increased therapeutic index, a shorter half life and less side-effects (Minotti, Menna et al. 2004; Charak, Jangir et al. 2011). Several mechanisms of action have been proposed for epirubicin; however the exact mechanism by which it exerts its cytotoxic effect remains unclear (Charak, Jangir et al. 2011). The mechanisms proposed, with the principal ones highlighted (*) include:

- DNA intercalation, alkylation and cross-linking*
- Generation of free radicals causing DNA damage and lipid peroxidation
- Interference with DNA unwinding and helicase activity
- Inhibition of topoisomerase II*
- Induction of apoptosis via topoisomerase II inhibition (Minotti, Menna et al. 2004)

2.3.1.1 Topoisomerase II inhibitors

One of the main mechanisms of action of epirubicin is thought to be the inhibition of topoisomerase II (Cleator, Parton et al. 2002). Topoisomerase II is a nuclear enzyme that creates transient DSBs in DNA to allow the cell to manipulate the topology of the DNA

(Hande 2008), after which the DNA is re-ligated and the structure is restored, without alteration of the sequence. This occurs during replication or repair processes (Cleator, Parton et al. 2002), and is closely linked to the cell cycle (Kellner, Sehested et al. 2002). Anthracyclines act by stabilising the covalent DNA-Topoisomerase II complexes and ultimately preventing re-ligation of the cleaved DNA (Hande 2008; de Campos-Nebel, Larripa et al. 2010), thus causing the most potent DNA lesion; a DSB. This type of lesion causes the initiation of the DNA damage response, and DNA repair by NHEJ and HR mechanisms (section 2.1.3), or may ultimately lead to apoptosis; the desired effect of the chemotherapeutic agent, resulting in the removal of the malignant cell from the system. If topoisomerase II causes a transient DSB in the gene of an important protein, such as p53 (a critical mediator of cell cycle control and apoptosis), in the presence of epirubicin, the inhibition of re-ligation by epirubicin of this important protein has huge ramifications upon normal control mechanisms (Kellner, Sehested et al. 2002). This may lead to the selection of therapy-resistant clones, where malignant cells survive by evasion of normal apoptotic or DNA repair pathways.

2.3.1.2 DNA intercalators

The other main mechanism of action of epirubicin is thought to be intercalation of the drug into DNA bases and prevention of synthesis and replication (Charak, Jangir et al. 2011). Charak and co-workers demonstrated the intercalation of epirubicin with DNA via guanine and cytosine bases and external binding through the phosphate backbone, using fourier transform infrared spectroscopy and UV-visible spectroscopy (Charak, Jangir et al. 2011). Following DNA intercalation at the 5'-G-C-3' site, doxorubicin, which is closely related to epirubicin, has been shown to form a formaldehyde-mediated DNA adduct, which is stabilised by hydrogen bonding to the complementary DNA strand (Spencer, Bilardi et al.

2008). This type of DNA adduct has been shown to initiate both NHEJ and HR (section 2.1.3) mechanisms of DNA repair, along with cell cycle arrest.

2.3.1.3 Other mechanisms of action

Several other mechanisms have been proposed, to describe the method by which anthracyclines may exert their effects (listed above). However, for some of these proposed mechanisms, including free radical formation, lipid peroxidation and types of DNA damage, there is both disproving and approving evidence; summarised by (Minotti, Menna et al. 2004).

One of the other less well-recognised mechanisms by which anthracyclines may act, describes the involvement of the proteasome. The 26S proteasome (comprised of a 20S core and two 19S capping structures) is an imperative mediator of critical processes controlling normal homeostasis; the function of many mediators required for the execution of several important pathways (cell cycle division, DNA damage response and DNA repair) relies on the activity of the ubiquitin-proteasome pathway. It has been proposed that following entry to the cell via passive diffusion, the translocation of anthracyclines to the nucleus is facilitated by the 26S proteasome (Kiyomiya, Matsuo et al. 2001). Research based on doxorubicin describes the binding of the drug to the 20S subunit of the 26S proteasome, forming a complex, and the subsequent translocation to the nucleus where doxorubicin dissociates from the proteasome and binds to DNA, for which it has higher affinity (Minotti, Menna et al. 2004). The nature of the complex formed between the 20S proteasome subunit and the drug molecule is said to be reversible and non-competitive, however other research has shown the action of the 26S proteasome to be inhibited by 4 different anthracyclines in a dose-dependent manner (Fekete, McBride et al. 2005).

2.3.2 Cyclophosphamide

Cyclophosphamide is a nitrogen mustard derivative and acts as a bischloroethylamine bifunctional DNA alkylating agent, of the oxazaphosphorine family (Middleton and Margison 2003). It is administered as a pro-drug and requires metabolic activation by cytochrome P450 enzymes in the liver (Murata, Suzuki et al. 2004), specifically CYP2B6, CYP3A4 and CYP2C9 (Zhang, Tian et al. 2005). Cyclophosphamide is activated to its cytotoxic metabolite by 4-hydroxylation of C-4 on the ring structure, producing 4-hydroxy-cyclophosphamide, which exists in equilibrium with its tautomer; aldophosphamide (Baumann and Preiss 2001). The metabolite 4-hydroxy-cyclophosphamide subsequently breaks down to produce reactive intermediates; phosphoramidate mustard (PM) and acrolein (Murata, Suzuki et al. 2004; Goldstein, Roos et al. 2008). PM is then converted into an aziridinium ion, which then alkylates guanine residues of DNA at the N7 position, forming N⁷G: N⁷G cross-links. Formation of these cross-links on opposite DNA strands (intrastrand cross-links) block replication forks and are regarded as the main cytotoxic lesion (Hansen, Ludeman et al. 2007; Kondo, Takahashi et al. 2010). This type of DNA lesion requires a complex DNA repair process, including a combination of NER, FA and HR DNA repair mechanisms (Kondo, Takahashi et al. 2010) (section 2.1.3). Using a cyclophosphamide analogue, mafosfamide, Goldstein and co-workers proposed that apoptosis may be induced by the blockage of replication caused by the DNA intrastrand cross-links (Goldstein, Roos et al. 2008). Blockage of replication can cause DNA DSBs, which then leads to the recruitment of ATM, subsequent activation of Chk1/Chk2 and stabilisation of p53, ultimately leading to apoptosis. Goldstein *et al*, demonstrated up-regulation of both intrinsic and extrinsic pathway mediators of apoptosis; PUMA and Fas, upon mafosfamide treatment (Goldstein, Roos et al. 2008).

2.3.3 Taxanes

The taxane family of anti-cancer agents includes paclitaxel and docetaxel, also known as Taxol[®] and Taxotere[®] respectively. Docetaxel was introduced after paclitaxel, and is said to be the more cytotoxic of the two agents, displaying a broad range of anti-tumour activity (Herbst and Khuri 2003), including inhibition of cell proliferation and induction of apoptosis. Docetaxel is a semi-synthetic agent derived from the needles of the European Yew tree (*Taxus baccata*) (Li, Hussain et al. 2005). The main mechanism of action of docetaxel is to cause cell death by 'mitotic catastrophe', where it prevents normal microtubule dynamics by binding to the main component of microtubules; the beta-tubulin heterodimer (Montero, Fossella et al. 2005; Morse, Gray et al. 2005). Taxanes actually promote the formation of microtubules during mitosis, unlike other anti-microtubule agents which prevent their assembly, however by binding to the beta-tubulin units they prevent the normal depolymerisation of microtubules, causing a state of hyperstabilisation (Herbst and Khuri 2003). The binding of docetaxel to beta-tubulin disrupts centrosome organisation in late S phase, G₂ and M phases, leading to incomplete mitosis, therefore preventing cell proliferation and initiation of apoptosis (Herbst and Khuri 2003; Montero, Fossella et al. 2005; McGrogan, Gilmartin et al. 2008). Another proposed method of cell death by docetaxel is by interaction with the anti-apoptotic protein, Bcl-2, and inhibition of its activity by phosphorylation (Herbst and Khuri 2003; Kraus, Samuel et al. 2003), thus promoting apoptosis of the malignant cell. Studies by Li *et al.*, involving gene expression profiling of cells treated with docetaxel identified a large range of differentially expressed genes related to microtubules, cell cycle arrest and apoptosis in prostate cancer cells, showing the broad range of cytotoxic influences initiated by docetaxel (Li, Li et al. 2004; Li, Hussain et al. 2005).

2.4 Chemotherapy resistance

Resistance to chemotherapy is responsible for a high percentage of cases where treatment fails to be effective. Resistance may be either ‘intrinsic’, where tumour cells are innately resistant to chemotherapy prior to treatment or ‘extrinsic’, where tumour cells acquire resistance during administration of chemotherapeutic agents.

The basic mechanisms of chemotherapy resistance include increased drug efflux / decreased influx, drug/target modification, drug detoxification, altered apoptosis and DNA repair (Coley 2009), and different chemotherapeutic agents may confer resistance in different ways. However, this does not describe mechanisms at a molecular level. It has been suggested by many that drug resistance is likely to be a complex multi-factorial phenomenon, involving several different mechanisms. This may explain why, as yet, no marker of chemotherapy-resistance has reached routine clinical use.

2.4.1 History and basis of multi-drug resistance

The phenomenon of multi-drug resistance (MDR) was first demonstrated by Biedler and Riehm in 1970, when they developed an *in vitro* model using hamster cells that displayed resistance to actinomycin D, as well as cross-resistance to a range of other compounds. Subsequently, the study of chemotherapy resistance rapidly emerged. One of the most studied proteins, now known to be associated with MDR in cancer is P-glycoprotein (P-gp). This protein was discovered using proteomics techniques in 1976, although the term ‘proteomics’ had not been coined. It was first identified in by Rudolph Juliano and Victor Ling, where it was recognised as a surface glycoprotein which could modulate drug permeability in hamster ovarian cell lines displaying colchicine resistance (Juliano 1976). Polyacrylamide gel electrophoresis was exploited to determine that the relative amount of surface labelled P-gp was proportional to the level of drug resistance exhibited (Juliano

1976). This original and pioneering work set the basis for understanding the mechanisms of drug resistance at a molecular level. P-gp is a member of the adenosine triphosphate (ATP)-Binding Cassette (ABC) transporter family of transmembrane proteins, which utilise ATP for the transfer of substances across membranes. Its mechanism of action involves the direct transport of hydrophobic compounds to the outside of the cell, using ATP and intrinsic ATPase (Ling 1997), thus eliminating substances from the cell by acting as an efflux pump. P-gp, and its ability to transport drugs across membranes, now characterises the 'classical' MDR phenotype, which is recognised by cross-resistance against natural anti-cancer agents including taxanes, anthracyclins and vinca alkaloids (Lage 2003) as well as sensitivity to chemosensitisers which reverse resistance, such as verapamil and cyclosporin (Lage 2003). It is important to note that MDR via a P-gp-independent mechanism, known as 'atypical' MDR, is also possible. This may involve altered drug targeting, by altered topoisomerase II for example, or increased detoxification by glutathione S-transferase (Gonzalez-Angulo 2007). There have been several clinical studies involving the use of P-gp inhibitors to reverse MDR. Early P-gp inhibitors included verapamil and cyclosporine, yet their use was limited by associated toxicities (Thomas and Coley 2003). The clinical use of subsequent P-gp inhibitors, such as dexverapamil and valsopodar, were also limited by toxicity and their interference with drug metabolism, by acting as competing substrates for CYP3A4 enzymes (Thomas and Coley 2003). These agents have been shown to affect drug distribution, where one study using murine breast cancer models, showed increased uptake of doxorubicin in cells proximal to blood vessels, yet decreased uptake in distal cells, showing limitations in therapeutic efficacy of doxorubicin when treated in combination with P-gp inhibitors (Patel and Tannock 2009). An example of a more recently-introduced P-gp inhibitor is tariquidar, which has shown to

overcome limitations associated with earlier P-gps, and inhibit both P-gp and Breast Cancer Resistance Protein (BCRP) (Kannan, Telu et al. 2011), and is being tested for its clinical use for the reversal of MDR.

2.4.2 Predictive markers of chemo-resistance

There are several markers which can be used to predict tumour response to a certain extent, giving the clinician an idea of how an individual breast tumour may respond to treatment. General factors affecting response include tumour size, tumour type (section 1.6.2) and molecular subtype, determined by hormone receptor expression (section 2.4.2.7). More specifically, a positive response to chemotherapy relies on efficient functioning of mediators of important pathways such as the cell cycle, DNA damage response, DNA repair and apoptosis; these mediators are responsible for delivering the desired effect of the chemotherapeutic agent, cell death. Therefore, alterations or defects in the mediators of these important pathways could result in the desired effect not being achieved; cell evasion of apoptosis and development of chemotherapy resistance. Putative biomarkers for the prediction of response to neoadjuvant chemotherapy have been widely studied and reviewed, however results have been inconsistent and sometimes contradictory (Kennedy, Quinn et al. 2004; Coley 2008; Tewari, Krishnamurthy et al. 2008), which explains why none of them have yet been transferred to the clinic. Some of the most widely studied putative biomarkers include:

2.4.2.1 Adenosine triphosphate-binding cassette (ABC) transporters

Adenosine triphosphate-binding cassette (ABC) transporters are a large family of cell membrane proteins, consisting of > 40 members, which are thought to play an important role in drug influx/efflux. The three main members of ABC transporters include P-

glycoprotein (P-gp/ABCB1/multi-drug resistance 1 (MDR1)), MDR-associated protein (MRP1/ABCC1) and breast cancer resistance protein (BCRP/ABCG2) (Chuthapisith, Eremin et al. 2006; Coley 2008). Overexpression of the most well-recognised ABC transporter, P-gp, has been associated with resistance to several anti-cancer agents, including anthracyclines and taxanes, in several cancer types, including ovarian, breast and head and neck cancers (Chuthapisith, Eremin et al. 2006), as a result of increased drug efflux. Expression of P-gp in tumour cells is thought to reduce the accumulation of intracellular xenobiotics, resulting in sub-optimal concentrations for exertion of cytotoxic effects (Coley 2008). Meta-analysis has shown that, on average, 40% of breast cancers are shown to express P-gp, and that its expression increases the chance of chemotherapy-resistance by 3-fold (Coley 2008).

2.4.2.2 p53 status

The tumour protein p53 plays a critical role in the response to DNA damage, by initiating cell cycle arrest and inducing apoptosis. It is therefore of little surprise that defective functioning of this fundamental protein is likely to have a variety of effects on the cellular response to chemotherapeutic agents. Point mutations or deletions in the p53 gene (*TP53*) are said to be present in over 50% of all cancers, and in 25% of sporadic breast cancers (Coley 2008). The relationship between pCR and p53 expression is slightly controversial, yet the general assumption is that p53 mutations increase the risk of chemotherapy resistance. An extensive study by Geisler and co-workers demonstrated *TP53* mutations within certain domains of the p53 protein to be associated with resistance to neoadjuvant doxorubicin in locally advanced breast cancer (Geisler, Lonning et al. 2001; Aas, Geisler et al. 2003), but acknowledged the controversy associated with p53 and therapy resistance (Lonning 2010). Mutations in the *TP53* gene have also been associated with resistance to

epirubicin (Chrisanthar, Knappskog et al. 2011) administered in the neoadjuvant setting in breast cancer patients. In contrast, *TP53* gene mutations were seen in 82% of triple negative breast tumours (Sorlie, Perou et al. 2001), yet it is this subtype which shows the highest response rates to neoadjuvant chemotherapy (section 2.4.2.7) (Rouzier, Perou et al. 2005). This controversy may be due to the difference between predicting response and prognosis; it is not necessarily true that a factor predictive of a positive response to treatment will predict a positive prognosis.

2.4.2.3 Topoisomerase II alterations

Topoisomerase II is the main target for anthracycline-based chemotherapy, however studies relating the expression of topoisomerase II with prediction of therapy response have been controversial (Tewari, Krishnamurthy et al. 2008; Oakman, Moretti et al. 2009; Lonning 2010). It would seem logical to presume that increased expression of topoisomerase II would yield a greater response to anthracyclines due to higher availability for drug-target interactions. This was demonstrated by Konecny and co-workers where amplification of the gene encoding topoisomerase II (*TOP2A*) was significantly associated with higher pCR to anthracycline-based neoadjuvant chemotherapy in breast cancer patients (Konecny, Pauletti et al. 2010). Brase and co-workers also demonstrated a significant association between increased *TOP2A* expression and increased rates of pCR in breast cancer patients who received neoadjuvant anthracycline chemotherapy (Brase, Schmidt et al. 2010). Studies have also suggested that the increased rate of pCR observed in HER2 positive tumours (section 2.4.2.7) may be due to the co-amplification of the *TOP2A* gene, which is located on the same chromosome (chromosome 17) (Brase, Schmidt et al. 2010; Miyoshi, Kurosumi et al. 2010). *TOP2A* gene amplification has been observed in ~40% of HER2 positive breast cancers (Jarvinen, Tanner et al. 2000). This was also demonstrated by

Konecny and co-workers, where they reported HER2 amplification to be significantly associated with increased pCR to anthracycline-based neoadjuvant chemotherapy, but only when the TOP2A gene was co-amplified, suggesting a putative role for TOP2A gene status as a predictive marker of response to chemotherapy in HER2 positive breast tumours (Konecny, Pauletti et al. 2010). Despite being a promising marker of response, meta-analysis and extensive reviews by Di Leo and co-workers has shown that further studies are required before topoisomerase II gene status can be transferred to the clinic as a predictive marker of response to anthracycline chemotherapy in breast cancer patients (Di Leo and Isola 2003; Di Leo, Biganzoli et al. 2008; Oakman, Moretti et al. 2009).

2.4.2.4 Tumour cell proliferation (Ki67 status)

Many studies have been carried out to determine the relationship between tumour cell proliferation, analysed by Ki67 expression, and chemotherapy response, yet this is another area where reports are controversial (Chuthapisith, Eremin et al. 2006; Tewari, Krishnamurthy et al. 2008). Gene expression profiling revealed that higher expression of cell proliferation gene clusters (including *CDC20*, *E2F1*, *MYBL2*, *FBXO5*, *MCM2*, *MCM6*, *CDC25B* and *TOP2A*) correlated with higher pCR to neoadjuvant anthracycline and paclitaxel chemotherapy in breast cancer patients (Gianni, Zambetti et al. 2005). In contrast, Geisler and co-workers have shown significant association between increased cell proliferation, (by high mitotic frequency) and resistance to doxorubicin, in locally advanced breast cancer (Aas, Geisler et al. 2003). A high Ki-67 score, which is a marker of cell proliferation, was also associated with lack of response to neoadjuvant anthracycline and/or taxane chemotherapy in breast cancer (Caudle, Gonzalez-Angulo et al. 2010).

2.4.2.5 DNA damage response and BRCA1 alterations

BRCA1 is involved in several pathways that respond to DNA damage, including cell cycle control, DNA repair (by HR) and transcriptional regulation. However, decreased expression of this protein is associated with ~ 30% of sporadic breast cancer cases (Kennedy, Quinn et al. 2004). BRCA1 mutations prevent the repair of DSBs, so the cell has to rely on the activity of poly(ADP-ribose)polymerase-1 (PARP-1), which is a critical component for the initial phases of the DNA damage response leading to repair of SSBs (Dianov and Parsons 2007; Eustermann, Videler et al. 2011). This has led to the investigation of PARP inhibitors such as Olaparib monotherapy, and so far studies have shown promising results, showing anti-tumour effects in breast tumours carrying BRCA1/BRCA2 mutations (Lonning 2010). A review by Kennedy and colleagues has described loss of BRCA1 function, through mutation, to be associated with sensitivity to DNA-damaging chemotherapeutic agents, such as anthracyclines (due to loss of effective DNA repair) (Kennedy, Quinn et al. 2004). However, they have also described BRCA1 to be involved in modulating the response to chemotherapeutic agents which act as spindle poisons, such as taxanes, (due to its involvement in the mitotic checkpoint), thus loss of BRCA1 function would actually confer resistance to this class of chemotherapeutic agents (Kennedy, Quinn et al. 2004). Data analysis carried out by Sorlie *et al*, found breast tumours with BRCA1 mutations to be mainly of triple negative subtype (Sorlie, Tibshirani et al. 2003), and it is this subtype which displays highest rates of pCR, to combination (anthracycline and taxane) neoadjuvant chemotherapy (Rouzier, Perou et al. 2005). A study by Asakawa and co-workers used clinical breast tissue samples to analyse the ability of various factors, including tumour size, nodal status, subtype and DNA damage response score to predict response to neoadjuvant EC plus docetaxel. The DNA damage response (DDR) score was

based on assessment of BRCA1, Rad51, H2AX and conjugated ubiquitin. Multivariate analysis with tumour size, nodal status, subtype and DDR revealed that only the DDR was significantly able to predict response to neoadjuvant chemotherapy ($p=0.0402$) (Asakawa, Koizumi et al. 2010). This therefore highlights the importance of the DDR pathway, and highlights its potential role as a predictive factor of chemotherapy response.

2.4.2.6 Breast cancer stem cell markers

More recently, the potential importance of cancer stem cells (CSCs), or tumour-initiating cells, has emerged, including their involvement in therapy resistance. Breast CSCs are classified by their high expression of CD44 and no or low expression of CD24 (CD44+/CD24-) (Nguyen, Almeida et al. 2010). Another characteristic of breast CSCs is their elevated expression of BCRP (ABCG2) (Chuthapisith, Eremin et al. 2010), thus increasing their ability to promote chemotherapy resistance via increased drug efflux. This is supported by their ability to cause rapid efflux of Hoechst 33342 dye from within the cell, demonstrated by flow cytometry (Hirschmann-Jax, Foster et al. 2005). Several studies have shown that the population of breast CSCs has increased following neoadjuvant treatment with several different agents including anthracyclines and taxanes (Li, Lewis et al. 2008; Nguyen, Almeida et al. 2010); the therapy has resulted in selection of this chemotherapy-resistant sub-population of cells. These cells have been associated with the poor-prognosis basal-like tumours, and more specifically with basal-like tumours with BRCA1 mutations, where one study found that 16/17 tumours with BRCA1 mutation contained cells positive for the CD44+/CD24- phenotype (Honeth, Bendahl et al. 2008). The most recently defined molecular subtype, claudin-low (section 2.2.2), which is a poor-prognosis subtype, has been shown to be the only molecular subtype which is enriched for tumour-initiating cells/ breast CSCs (Perou 2010).

2.4.2.7 Molecular Subtype (ER, PR and HER2 status)

As well as being a good prognostic factor, breast tumour molecular subtype, identified by gene expression analysis, can also be a good predictive marker of tumour response to neoadjuvant chemotherapy. Basal-like and ERBB2+ breast tumours have been associated with highest rates of pathological complete response (pCR); averaging 36% and 40% respectively, after neoadjuvant chemotherapy administration (Rouzier, Perou et al. 2005; Carey, Dees et al. 2007; Chen, Chang et al. 2008) whereas luminal tumours have shown lowest pCR rates at 6% (Rouzier, Perou et al. 2005; Carey, Dees et al. 2007). However, regardless of increased chemosensitivity, the basal-like and ERBB2+ molecular subtypes of breast cancer show the poorest disease-free survival (Carey, Dees et al. 2007). This demonstrates that tumours with a good-prognosis signature (luminal subtypes) are less sensitive to chemotherapy than tumours with a poor-prognosis signature (triple negative). The reason for the increased rates of pCR in the triple negative tumours is likely to be associated with the finding that *TP53* gene mutations were associated with 82% of triple negative tumours (Sorlie, Perou et al. 2001), but are also associated with BRCA1 mutation (defective DNA repair) status (Sorlie, Perou et al. 2001; Sorlie, Tibshirani et al. 2003); thus damaged DNA is not being repaired, so the cell would naturally be directed towards apoptosis. A defective DNA repair signature, by mutations in BRCA1 has shown to be able to predict response to anthracycline-based neoadjuvant chemotherapy (Asakawa, Koizumi et al. 2010; Rodriguez, Makris et al. 2010). The expression of ER, PR and HER2 proteins, which reveal the breast tumour molecular subtype (section 2.2.2), are routinely screened at the time of diagnosis, to aid prognosis and choice of therapeutic strategy. Research has been carried out to assess the individual ability of ER and HER2 proteins as single predictive markers of response to chemotherapy, however findings have been inconclusive

and their expression has mainly been significantly associated with prognosis (Tiezzi, Andrade et al. 2007; Oakman, Moretti et al. 2009; Asakawa, Koizumi et al. 2010).

2.4.3 Predictive gene expression signatures

Several gene signatures for predicting chemotherapy response in breast cancer have been reported, the principal one being the 70-gene signature (van 't Veer, Dai et al. 2002). This was originally discovered for predicting breast cancer outcome, and was customised into the high-throughput diagnostic array ‘Mammaprint[®]’ (Agendia) (Glas, Floore et al. 2006). Its prognostic value is currently being assessed in the prospective randomised phase III clinical trial ‘MINDACT’ (Microarray In Node negative Disease may Avoid ChemoTherapy) (Cardoso, Piccart-Gebhart et al. 2007), however it has also been specifically recognised as a predictor of response to several chemotherapeutic agents, including anthracyclines and taxanes, administered to breast cancer patients in the neoadjuvant setting (Straver, Glas et al. 2010). Breast tumours with a good-prognosis signature showed a 0% pCR rate (0/23), yet 20% (39/144) of tumours with a poor-prognosis signature achieved a pCR to neoadjuvant chemotherapy. This agrees with other studies (2.4.2.7), where breast tumours with a poor prognosis have shown to be more sensitive to chemotherapy. The biological function of the 70 genes associated with this signature have been interpreted using Ingenuity Pathway Analysis (section 10.1.1), and have been shown to reflect the six hallmarks of cancer, which were originally described by Hanahan and Weinberg (2000) (section 2.2.1) (Tian, Roepman et al. 2010). Another gene signature, the 21-gene recurrence score, which was customised into the ‘Oncotype DX[®]’ assay (Genomic Health), was originally reported to quantify risk of recurrence in tamoxifen-treated ER+ breast cancer patients (Paik, Shak et al. 2004). It is currently being evaluated within the TAILORx (Trial Assigning Individualised Options for Treatment

[Rx]) clinical trial (Sparano and Paik 2008)); however it has also been linked with prediction of pCR to neoadjuvant chemotherapy. The 21-gene recurrence score was positively associated with pCR, where breast cancer patients at greatest risk of recurrence showed the greatest rates of pCR to neoadjuvant paclitaxel and doxorubicin (Gianni, Zambetti et al. 2005). More recently, the importance of the tumour microenvironment, particularly tumour-associated stroma, is being recognised for its predictive values (Farmer, Bonnefoi et al. 2009). Farmer and co-workers identified a 50-gene stroma-related signature, which is able to predict anthracycline-based chemotherapy response to breast cancer in the neoadjuvant setting (Farmer, Bonnefoi et al. 2009). Several other multigene signatures have been proposed, which have been reported to be predictors of sensitivity to different anthracycline and taxane chemotherapeutic agents in breast cancer (Bonnefoi, Underhill et al. 2009; Colombo, Milanezi et al. 2011), however none have yet been incorporated into randomised trials or transferred to the clinic.

2.4.4 Overview of putative markers of response to neoadjuvant chemotherapy in breast cancer

Extensive research has been performed, as outlined above, which has presented and analysed a variety of different gene and protein markers and signatures, as markers predictive of response to neoadjuvant chemotherapy in breast cancer; some showing great promise. The above-mentioned putative biomarkers, and the molecules associated with them, which may be associated with the phenomenon of chemotherapy resistance are listed in Table 3.

Despite this extensive research, predictive biomarkers of chemotherapy resistance have not yet been transferred to the clinic. The only markers which are routinely tested in the clinic include ER, PR and HER2, which are used to give an indication of prognosis and potential

therapeutic strategy, but do not act as predictive markers of neoadjuvant chemotherapy response. Resistance to chemotherapeutic agents presents a large obstacle in effective tumour treatment. In this instance, patients receive cytotoxic drugs for no therapeutic gain, in which time tumour progression may occur. The ability to predict how a tumour may respond to therapy, by using a panel of predictive biomarkers as a screening tool at the time of diagnosis, would allow individualisation of treatment, thus maximising treatment efficacy on an individual patient basis. However, the phenomenon of chemotherapy resistance appears to be a complex multi-factoral process, which requires further research. The use of proteomics, as a tool to identify putative predictive biomarkers of resistance to neoadjuvant chemotherapy in breast cancer, will be discussed in Chapter 3.

Table 3: A summary of putative predictive biomarkers of chemotherapy resistance in breast cancer

A summary of gene signatures, pathways, families/groups of molecules and single molecules which have been described as putative predictive biomarkers of chemotherapy-resistance in breast cancer are listed. The pathways and molecules associated with the mechanism of action of chemotherapeutic agents (most relevant to this project) have also been listed, as alterations in these molecules may also potentially be associated with chemotherapy resistance. The ‘70-gene signature’, which comprises the MammaPrint® diagnostic array, has been shown to represent the original six hallmarks of cancer: (1) self-sufficiency in growth signals, (2) insensitivity to growth inhibitory signals, (3) evasion of apoptosis, (4) limitless replicative potential, (5) sustained angiogenesis and (6) tissue invasion and metastasis (Hanahan and Weinberg 2000). Genes within the ‘70-gene signature’ associated with these hallmarks are shown in the table. Genes which are present within >1 of the predictive gene signatures are highlighted (n=3).

Putative predictive biomarker	Associated molecules			
Gene signatures				
MammaPrint® 70-gene signature (van't Veer 2002)	<i>ALDH4A1</i> <i>AYTL2</i> <i>BBC3</i> <i>CCNE2</i> <i>CDC42BPA</i> <i>CDCA7</i>	<i>FLT1</i> <i>GMPS</i> <i>GNAZ</i> <i>GPR180</i> <i>GPR126</i> <i>GSTM3</i>	<i>PALM2</i> <i>PECI</i> <i>PITRM1</i> <i>PRC1</i> <i>QSCN6L1</i> <i>RASSF7</i>	Genes associated with the ‘6 Hallmarks of Cancer’

(Tian, 2010)	<i>CENPA</i> <i>COL4A2</i> <i>DCK</i> <i>DIAPH3</i> <i>DTL</i> <i>EBF4</i> <i>ECT2</i> <i>EGLN1</i> <i>ESM1</i> <i>EXT1</i> <i>FGF18</i>	<i>HRASLS</i> <i>IGFBP5</i> <i>KNTC2</i> <i>LIN9</i> <i>MCM6</i> <i>MELK</i> <i>MMP9</i> <i>MTDH</i> <i>NUSAP1</i> <i>ORC6L</i> <i>OXCT1</i>	<i>RECQL5</i> <i>RFC4</i> <i>RTN4RL1</i> <i>RUNDC1</i> <i>SCUBE2</i> <i>SLC2A3</i> <i>STK32B</i> <i>TGFB3</i> <i>TSPYL5</i> <i>WISP1</i>	
	<i>LGP2</i> <i>NMU UCHL5</i> <i>JHDM1D</i> <i>AP2B1</i> <i>MS4A7</i>	<i>RAB6B</i> <i>LOC100288906</i> <i>C9orf30</i> <i>ZNF533</i> <i>C16orf61</i>	<i>SERF1A</i> <i>C20orf46</i> <i>LOC730018</i> <i>LOC100131053</i> <i>AA555029_RC</i>	Genes <i>not</i> associated with the '6 Hallmarks of Cancer'
Oncotype DX [®] 21-gene signature (Paik, 2004)	<i>ACTB</i> <i>BAG1</i> <i>BCL2</i> <i>CCNB1</i> <i>CD68</i> <i>CTSL2</i>	<i>ER</i> <i>GAPDH</i> <i>GRB7</i> <i>GSTM1</i> <i>GUS</i>	<i>HER2</i> <i>Ki67</i> <i>MMP11</i> <i>MYBL2</i> <i>PGR</i>	<i>RPLPO</i> <i>SCUBE2</i> <i>STK15</i> <i>Survivin</i> <i>TFRC</i>
Stroma-related 50-gene signature (Farmer 2009)	<i>ADAM12</i> <i>AEBP1</i> <i>ANGPTL2</i> <i>ASPN</i> <i>C1QTNF3</i> <i>C1R</i> <i>CALD1</i> <i>CDH11</i> <i>COL10A1</i> <i>COL1A2</i> <i>COL3A1</i> <i>COL5A2</i> <i>COL6A1</i>	<i>COL6A3</i> <i>COPZ2</i> <i>CSPG2</i> <i>CTSK</i> <i>DACT1</i> <i>DCN</i> <i>DPYSL3</i> <i>ECM2</i> <i>FAP</i> <i>FBLN1</i> <i>FBN1</i> <i>GAS1</i> <i>HTRA1</i>	<i>ITGBL1</i> <i>LOXL1</i> <i>LRP1</i> <i>LRRC17</i> <i>MFAP2</i> <i>MGC3047</i> <i>MMP11</i> <i>MMP14</i> <i>MMP2</i> <i>NDN</i> <i>OLFML2B</i> <i>PCOLCE</i>	<i>PDGFRB</i> <i>PDGFRL</i> <i>PEDF</i> <i>PLAU</i> <i>POSTN</i> <i>RARRES2</i> <i>SFRP4</i> <i>SNAI2</i> <i>SPARC</i> <i>SPON1</i> <i>TGFB3</i> <i>THBS2</i>
Pathways, families and groups of molecules				
DNA Damage Response (DDR)	DDR score (Asakawa 2010): <i>BRCA1</i> , <i>Rad51</i> , <i>H2AX</i> and conjugated ubiquitin			
ABC drug transporters	P-glycoprotein (<i>P-gp/ABCB1</i> /multi-drug resistance 1 (<i>MDR1</i>)) MDR-associated protein (<i>MRP1/ABCC1</i>) Breast cancer resistance protein (<i>BCRP/ABCG2</i>)			
Tumour cell proliferation	<i>Ki67</i> Gene clusters <i>CDC20</i> , <i>E2F1</i> , <i>MYBL2</i> , <i>FBXO5</i> , <i>MCM2</i> , <i>MCM6</i> , <i>CDC25B</i> and <i>TOP2A</i>			
Molecular subtype	ER, PR and <i>HER2</i>			
Breast cancer stem cell markers	<i>CD44</i> , <i>CD24</i> , <i>BCRP/ABCG2</i> and <i>BRCA1</i> mutations			
Single molecules				
p53	<i>TP53</i>			
Topoisomerase II	<i>TOP2A</i>			

Chemotherapy drug	Pathways and molecules associated with the mechanism of action of chemotherapeutic agents
Anthracyclines (e.g. epirubicin, doxorubicin)	<p><i>Main drug targets:</i> Topoisomerase II, guanine and cytosine bases</p> <p><i>Effect:</i> DNA double strand breaks, initiation of DNA repair pathways: NHEJ and HR (ATM, BRCA1, BRCA2, RAP80, Rad51, Rad52, Rad54, Ku78, Ku80, DNA ligase IV/XRCC4 complex etc), induction of apoptosis</p> <p><i>Other associated molecules:</i> 26S Proteasome (20S subunit)</p>
Cyclophosphamide	<p><i>Main drug target:</i> Guanine bases</p> <p><i>Effect:</i> Blockage of replication, DNA double strand breaks, induction of apoptosis. ICL DNA repair: NER, HR, FA (as above, plus FANCD2, FANL1, MUS81-EME1, ERCC1, XPF, Rev1, Rev3, Ref7 etc)</p> <p><i>Other associated molecules:</i> PUMA and Fas</p>
Taxanes (e.g. docetaxel, paclitaxel)	<p><i>Main drug target:</i> beta-tubulin (microtubules)</p> <p><i>Effect:</i> Inhibition of mitosis, prevention of proliferation, induction of apoptosis</p> <p><i>Other associated molecules:</i> Bcl-2</p>

CHAPTER 3:
PROTEOMIC TECHNIQUES AND THE
IDENTIFICATION OF BIOMARKERS OF
CHEMOTHERAPY RESISTANCE

Hodgkinson, V.C., Eagle, G.L., Drew, P.J., Lind, M.J., Cawkwell, L. (2010) Biomarkers of chemotherapy resistance in breast cancer identified by proteomics: current status. *Cancer Letters*, **294**(1):13-24

Chapter 3. Proteomic techniques and the identification of biomarkers of chemotherapy resistance

3.1 Biomarker discovery pipeline

The search for biomarkers generally consists of four phases; *discovery*, where mass spectrometry (MS) - or microarray-based approaches may be used to generate large lists of differentially expressed proteins (DEPs); *data mining*, to analyse discovery-phase data using knowledge bases to prioritise candidates to carry forward; *confirmation*, where techniques such as western blotting or ELISA are used to confirm the differential expression of candidates from the discovery phase, and *validation* where putative biomarkers are evaluated in the clinical context, which may use immunohistochemistry. The number of samples used increases throughout this process, and the number of potential candidates decrease (Figure 9).

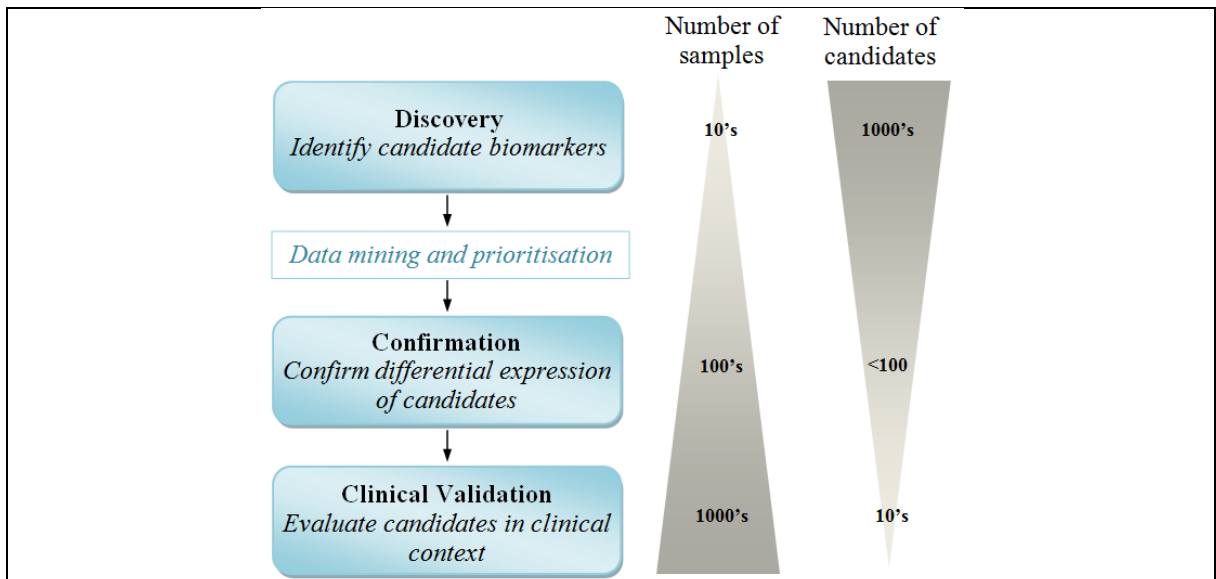


Figure 9: Stages of the biomarker discovery pipeline

The number of candidates decreases as false-positive and weaker candidates are removed. The number of samples increases to improve and challenge the strength of the candidates and to ensure extensive clinical validation before significant observations can be reported.

One of the main challenges encountered during biomarker discovery in clinical samples is accessing the low-abundant proteins of interest. This is made difficult by the huge dynamic range of protein concentrations in serum, covering 10 orders of magnitude (Anderson and Anderson 2002). The most abundant protein present in serum is albumin, which constitutes 55% of the serum proteome (Anderson and Anderson 2002). In fact, only 22 proteins account for 99% of serum (Tirumalai, Chan et al. 2003) (Figure 10). The remaining 1% therefore represents the low abundant proteins of interest. Depletion of the highly abundant proteins and accessing the proteins of interest presents challenges and therefore highlights potential limitations of such research.

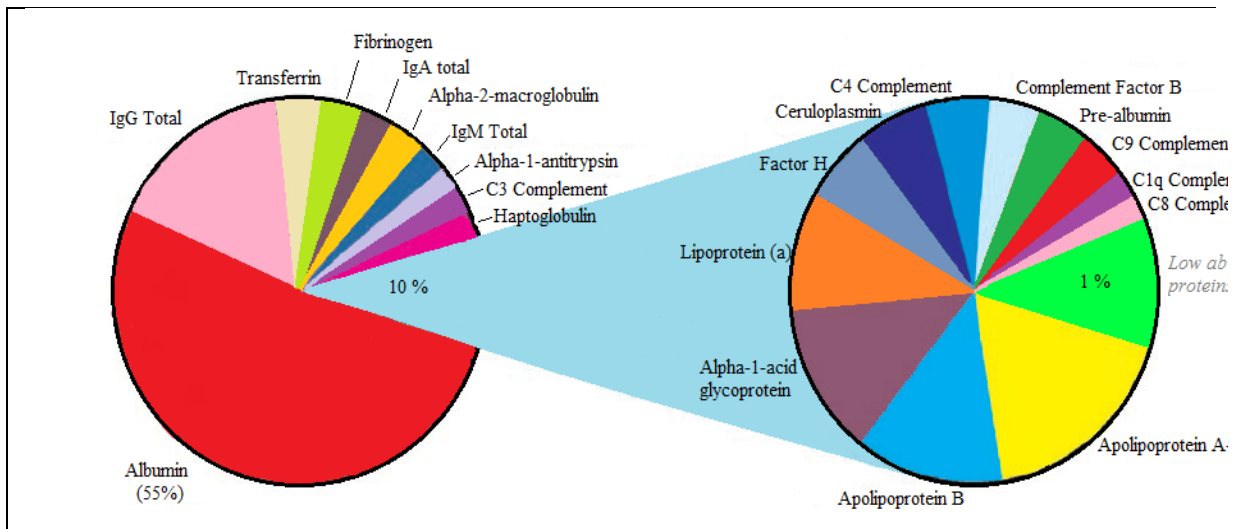


Figure 10: Pie charts representing the 22 major proteins of the human serum proteome

Ninety percent of the human serum proteome is composed of 10 proteins (pie chart to left). The remaining 10% (pie chart to right) is composed of a further 12 known proteins as well as the low abundant proteins, which are of interest in biomarker discovery, representing 1% of the serum proteome (not to scale). Accessing these low abundant proteins is made difficult by the huge dynamic range; 10 orders of magnitude separate these proteins from the most abundant protein, albumin (Anderson and Anderson 2002).

From (Tirumalai, Chan et al. 2003).

3.2 Proteomics

The term 'proteome' was first introduced in 1994 by Marc Wilkins and refers to the **proteins** expressed by the **genome**. The study of the proteome was coined 'proteomics', and describes the large-scale global study of the proteome, and is the most recently introduced global screening technique following transcriptomics and genomics.

The choice of molecular level for studying cancer-associated alterations is complicated by the different ways that genes may be transcribed into a variety of functionally distinct proteins, which can themselves undergo essential post-translational modifications. The human genome is relatively static and gene mutations may not have any functional relevance in the resulting mRNA or protein species. Due to alternative splicing mechanisms, one gene can produce more than one protein species. Levels of mRNA and protein expression are dynamic and constantly changing over time, according to the state and microenvironment of the individual cell. The global mRNA-based analysis of gene expression with microarrays (transcriptomics) does not address post-translation events and therefore may lead to lack of correlation between mRNA and the functional proteome. A lack of correlation may also occur due to protein expression levels lagging behind the peak in mRNA production or mRNA may undergo such high turn-over that no protein is expressed. Proteins can undergo a wide variety of post-translational modifications which affect protein stability, localisation and function (Hoffman, Sniatynski et al. 2008). Therefore the analysis of protein expression may provide the most realistic picture of the functional aberrations within a cancer cell, as it arguably provides access to the most accurate molecular repertoire.

Proteomics, the global analysis of protein expression in a given proteome, has been an important developing area of cancer research. The recent developments in mass

spectrometry (MS), which have enabled high sensitivity and automation of protein identification, facilitated an increased interest in global proteomics-based research. Although previously called ‘fishing expeditions’, today the term ‘discovery science’ is respectfully used for such large scale studies (Baak, Janssen et al. 2005). Initial experiments using established cell lines are now giving rise to the analysis of complex tissue and biological fluids to establish protein changes associated with disease (Aldred, Grant et al. 2004).

Proteomics can be used as a comparative tool to identify differences in protein expression between two samples, such as normal versus disease or chemotherapy-sensitive versus chemotherapy-resistant. This enables differentially expressed proteins (DEPs) to be identified which may be associated with a certain disease phenotype.

Current proteome analysis methods can be separated into gel-based MS methods (for example two-dimensional polyacrylamide gel electrophoresis with matrix assisted laser desorption ionisation time of flight MS; 2D-PAGE/MALDI-TOF-MS), gel-free MS methods (for example liquid chromatography with electrospray ionisation MS; LC/ESI-MS), and microarray-based methods (for example antibody microarrays).

3.3 Gel-based MS methods

This type of method is based upon achieving effective separation of proteins in a sample, using gel electrophoresis followed by identification of the protein using mass spectrometry.

3.3.1 1D-PAGE separation

One-dimensional PAGE (1D-PAGE) is used to separate proteins by molecular weight in a denaturing polyacrylamide gel. Protein samples must be extracted or resuspended in a suitable buffer (e.g. Laemmli buffer) composed of a detergent (e.g. sodium dodecyl

sulphate; SDS) to solubilise membrane proteins and disrupt protein-protein interactions; a reducing agent (e.g. β -mercaptoethanol) to reduce protein disulphide bonds prior to denaturing SDS-PAGE; protease inhibitors to block the function of protease enzymes which digest proteins; phosphatase inhibitors to block the action of phosphatase enzymes which remove phosphate groups from phosphorylated proteins by hydrolysing phosphoric acid esters; glycerol to increase the density of the sample so that it sinks into the well of the polyacrylamide gel; and dye (e.g. bromophenol blue) to visualise the protein sample during gel loading and electrophoresis. The protein mixture is loaded into the gel and separated out into individual protein bands; the molecular weights of which can be estimated from the co-electrophoresis of a ladder of molecular weight marker proteins.

3.3.2 2D-PAGE separation

Two-dimensional electrophoresis was first reported by O'Farrell in 1975, where the technique was developed to separate proteins in the 1st dimension according to their isoelectric point (pI), known as isoelectric focusing (IEF), and by molecular weight in the 2nd dimension, using sodium dodecyl sulphate PAGE electrophoresis (O'Farrell 1975). 2D-PAGE is a proteomic method widely used to achieve a higher level of separation by resolving proteins into individual spots prior to MS analysis, which has been extensively reviewed (Rabilloud 2002).

3.3.2.1 Sample Preparation

Proteins are extracted from their source; either cell line or tissue origin by chemical or mechanical means. Sample preparation is carried out to ensure the sample is in a suitable physicochemical state for IEF. The sample must therefore be solubilised in an appropriate buffer containing suitable components, which may involve reduction and denaturation of

the protein sample (Shaw and Riederer 2003). The buffer comprises a detergent (e.g. CHAPS) to solubilise membrane proteins and disrupt protein-protein interactions; a chaotrope (e.g. urea/thiourea) to disrupt hydrogen bonded structures in the protein sample and to increase solubility; protease inhibitors; phosphatase inhibitors; a reducing agent (e.g. dithiothreitol; DTT) to reduce disulphide bonds between cysteine residues so that all proteins in the sample have the same linear shape; ampholytes to establish a stable pH gradient for use in IEF; and dye (e.g. bromophenol blue) to visualise the protein sample.

3.3.2.2 Separation in the 1st Dimension: Isoelectric Focusing

IEF involves the horizontal separation of proteins in the sample according to their isoelectric point (pI); the point at which the proteins carry no net electrical charge (Figure 11) (Gorg, Drews et al. 2009). The proteins migrate towards the electrode with the opposite charge and gain or lose protons along the pH gradient. The net charge and mobility decrease and each protein will stop at the point in the pH gradient that is equal to its pI. Immobilised pH gradient (IPG) strips are used to create a stable pH gradient (Rabilloud, Valette et al. 1994; Rahimpour, Soheili et al. 2007). They are acrylamide derivatives of simple buffers known as ‘immobilines’ which do not exhibit amphoteric behaviour and co-polymerise with the gel matrix and form an appropriate pH gradient. These strips are ‘rehydrated’ with the sample and when an electric current is applied, the proteins in the sample migrate to their pI.

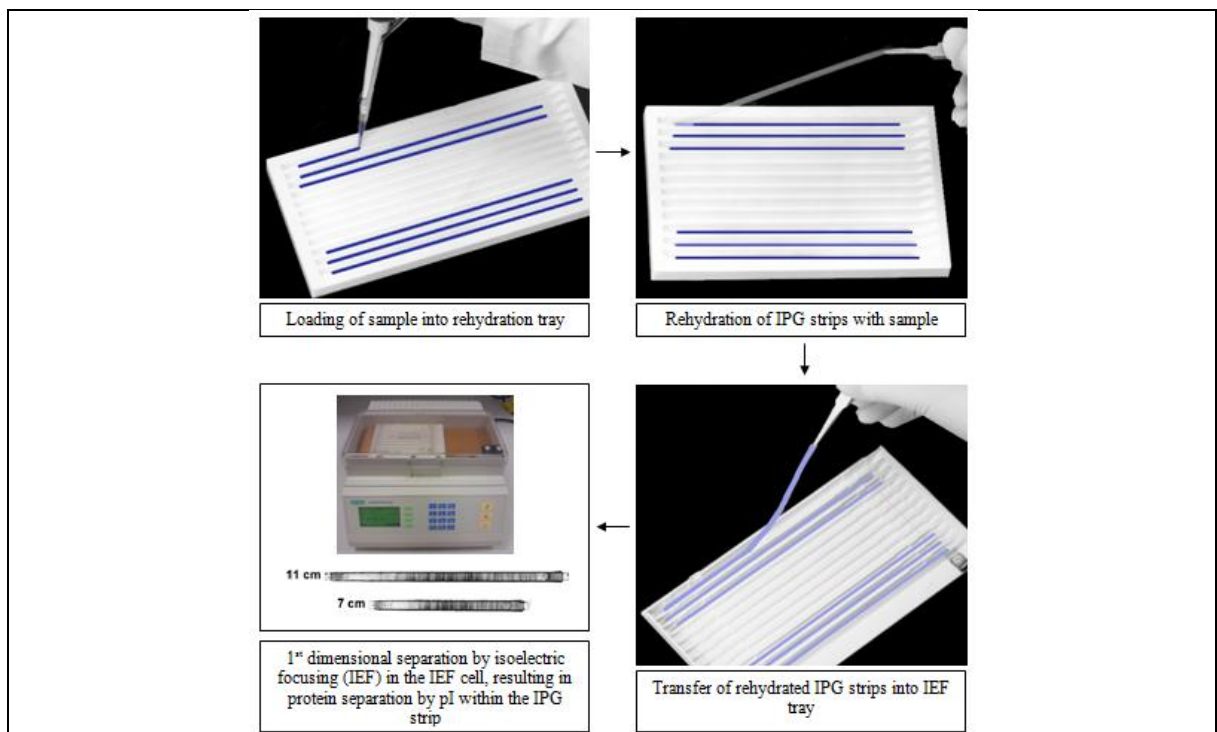


Figure 11: Stages of protein separation by isoelectric focusing (IEF) – 1st dimension

The sample is pipetted along the back of a well in a rehydration tray. The IPG strips are then rehydrated with the sample overnight. Rehydrated IPG strips are then transferred into the IEF tray, which is placed in the IEF cell for IEF.

3.3.2.3 Separation in the 2nd dimension: by molecular weight

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) is used to separate proteins vertically according to molecular weight, following separation by pI (Figure 12). SDS is an anionic detergent which is used to denature the proteins and give them a net negative charge, which ensures that they migrate towards the anode (Garfin 2003). It also disrupts hydrogen bonds, blocks hydrophobic interactions and unfolds the proteins. The IPG strips containing the protein sample must be equilibrated prior to the SDS-PAGE step to further denature proteins, maintain solubility, and establish an appropriate pH. DTT may be added to reduce any re-formed disulphide bonds and iodoacetamide (IAA) can be used for carbamidomethylation, by alkylating free thiol groups, thus preventing re-formation of disulphide bonds. The equilibrated IPG strip is placed at the top of the gel and proteins then migrate through the gel and are separated

according to molecular weight. The resolution of protein spots (and therefore the level of proteome interrogation) is dependent on the size of gel and the pH range analysed. Maximal proteome analysis requires the largest gel size combined with a series of ultra-narrow range pH strips but this approach has to be balanced against time and cost, as well as sample availability. The complexity of the protein sample can be reduced by conducting a pre-fractionation step. For example, methodologies are available to enrich for nuclear proteins (Fu and Fenselau 2005) membrane-associated proteins (Tan, Tan et al. 2008) and phosphorylated proteins (Morandell, Stasyk et al. 2006).

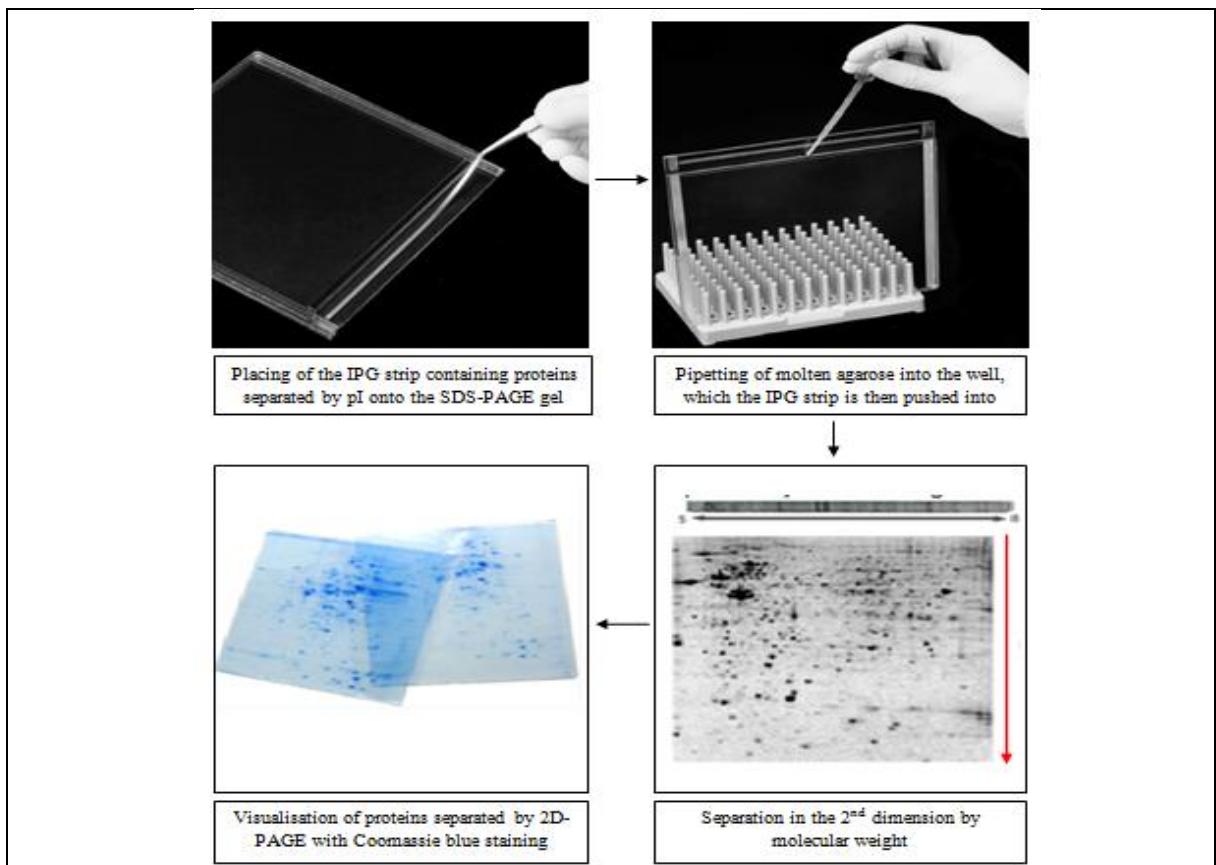


Figure 12: Stages of protein separation by SDS-PAGE – 2nd dimension

The IPG strip containing proteins separated horizontally by pI is placed at the top of the SDS-PAGE gel. Molten agarose is placed within the well at the top of the gel, and the IPG is set within it. Proteins are then separated vertically by molecular weight. Separated proteins can then be visualised using a stain such as Coomassie blue.

3.3.2.4 Visualisation of Proteins

Following 1D-PAGE or 2D-PAGE, proteins are visualised by staining which is required for protein excision and quantitative analysis. The stain of choice must be compatible with downstream mass spectrometry, and a commonly used stain is coomassie blue, which has a detection limit of approximately 10 ng of protein (Garfin 2003), and is relatively inexpensive and easy to use. Other available stains include silver and fluorescent stains, which are more sensitive than Coomassie staining, however if proteins are to be subsequently identified, the technique used for this purpose must have a matching level of sensitivity.

3.3.2.5 Identification of differentially expressed proteins

Two groups of samples (e.g. 'chemotherapy-sensitive' and 'chemotherapy-resistant') can be analysed to identify differentially expressed proteins between them. At least 3 technical replicates (3 individual gels per sample) should be performed, as well as multiple biological replicates to account for analytical and the high degree of sample variability observed in clinical samples, and to reduce false positive results (Smith, Qutob et al. 2009). The use of dual colour Cy3/Cy5 fluorescent labelling in two dimensional difference gel electrophoresis (2D-DIGE) allows the co-electrophoresis of differentially labelled samples in the same gel (Von Eggeling, Gawriljuk et al. 2001). This can improve the comparison of samples by removing gel to gel variability. After separation using 2D-PAGE, and protein visualisation, proteins that are differentially expressed (typically a 2-fold difference in expression is accepted as significant) between the two samples are determined using a software package (Marengo, Robotti et al. 2005), such as PDQuest (Bio-Rad). Differences in spot pattern and intensity between the gels are identified by relative quantification and individual DEP spots are highlighted. These can then be excised manually using a scalpel, or robotically.

3.3.2.6 In-gel digest

Protein bands separated by 1D-PAGE can be cut from the gel lane in thin slices of approximately 1mm width, each containing a mixture of proteins. Proteins from 2D-PAGE are cut out as individual protein spots. In order to release the proteins from the gel matrix, it is necessary to digest the proteins into peptides. This initially involves the de-staining and washing of protein spots with ammonium bicarbonate / acetonitrile solutions, prior to protein digestion. This is commonly achieved using an enzyme such as trypsin, which specifically cleaves proteins at the C-terminal side of lysine and arginine residues (Olsen, Ong et al. 2004). The peptides can then be analysed by mass spectrometry, and identified with the use of a database containing *in silico* tryptic peptides from known proteins. If necessary, ZipTip[®] pipette tips can be used to purify and concentrate the peptide samples.

3.3.2.7 Protein identification by mass spectrometry

Mass spectrometry is a complex and powerful analytical technique which is used to determine the chemical composition of a sample, and is used in an increasing number of applications, including analysis of biomolecules. Over the past decade, levels of sensitivity, detection, speed and analytical range have increased immensely, which has enabled new levels of application and the development of new methods. The basic principle of mass spectrometry is to produce ions and determine their mass. This is achieved by ionising the sample, separating ions according to their mass-to-charge ratio (m/z), detecting ions and determining their m/z ratio alongside relative abundance. This process may also include fragmentation of selected ions. A mass spectrometer is essentially comprised of three components; an ion source, a mass analyser and a detector. There may be more than one mass analyser, depending on the level of analysis required. A computer also forms part of the system, which is responsible for the processing of data generated.

There is a wide variety of ionisation methods, however the two main methods used for ionisation of peptides are matrix assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI) (Aebersold and Goodlett 2001; Aebersold and Mann 2003; Lin, Tabb et al. 2003). ESI is commonly coupled to liquid-based separation whereas MALDI is used for solid samples co-crystallised with matrix onto a rigid sample plate. MALDI is generally coupled with a time of flight (TOF) mass analyser and 2D-PAGE is usually combined with MALDI-TOF-MS for identification of the proteins contained in excised DEP spots (Aebersold and Goodlett 2001; Aebersold and Mann 2003; Lin, Tabb et al. 2003).

The introduction of MALDI by Karas and Hillenkamp in 1985 dramatically improved the analysis of proteins, peptides and other large molecules (Penque 2009). This ‘soft’ ionization technique allows energy from the laser to be transferred to the analyte indirectly, via an organic matrix, thus reducing sample damage and decomposition. For MALDI, there are several acidic organic matrix molecules the sample can be mixed with, examples of which include 2, 5, dihydroxybenzoic acid (DHB) and α -cyano-4-hydroxycinnamic acid (CHCA). The matrix solution also contains an organic solvent such as acetonitrile (ACN) to prevent sample aggregation, water, and may also contain trifluoroacetic acid (TFA). An example of a matrix solution is CHCA [5 mg/ml] in 50% ACN; 0.1 % TFA (aq). The peptides are mixed with matrix solution and spotted onto a MALDI target plate. The organic solvent evaporates to leave a homogenous co-crystallised mixture of peptides and organic matrix. The matrix molecules used have strong absorption at the laser wavelength, so that when the laser is fired, energy is absorbed by the matrix molecules. This accumulation of energy results in expansion of the matrix and co-release of peptide molecules and matrix molecules into the gas phase (Lin, Tabb et al. 2003). Peptide

molecules are ionized and enter the gas phase via indirect energy transfer from laser pulses via the matrix molecules, thus minimising sample damage and degradation (Figure 13).

There are different types of lasers which can be used for MALDI, the most common of which are nitrogen lasers and neodymium:yttrium aluminium garnet (Nd:YAG) lasers. In recent years, Bruker Daltonics introduced the Bruker smartbeam™ laser, which combines the best attributes of nitrogen and Nd:YAG lasers; laser performance and speed, respectively, ultimately delivering maximal peak intensity.

The role of the mass analyser is to separate and sort ions according to their mass-to-charge (m/z) ratio. There are several different types of mass analysers which are currently used for proteomics; quadrupole; ion trap; time-of-flight; fourier transform ion cyclotron resonance and fourier transform orbitrap (Aebersold and Mann 2003), within which differences are seen regarding type of field applied, method of ion transmission and kinetic energy rates. The performance of a mass analyser depends on the resolution, mass accuracy, analysis speed, transmission and mass range of the machine. A type of mass analyser commonly used for peptide analysis is the time-of-flight (TOF) mass analyser (Aebersold and Mann 2003), which is well suited to the pulsed nature of MALDI. MALDI-TOF-MS is a favoured technique for generating peptide mass fingerprint (PMF) information, which allows the identification of unknown proteins, as it is a robust, sensitive technique (to femtomole or attomole levels) and is capable of analysing a large mass range (Lin, Tabb et al. 2003).

The first type of TOF mass analyser to be introduced was a linear TOF analyser. Ions generated from the sample are accelerated by an electric field, gaining equal kinetic energy, and are subsequently separated according to the time taken to travel down a field-free flight tube, from which the m/z is determined. This is performed under a vacuum to ensure collisions do not occur before ions reach the detector. The flight time (t) is relative to the

mass (m) of the ion by the following equation: $t = (m/z)^{1/2}$ (Hillenkamp, Karas et al. 1991). The resolution of the TOF mass analyser was initially improved by the introduction of delayed pulsed extraction, rather than continuous extraction. This allowed differences in kinetic energy between ions of the same m/z to be corrected, therefore improving resolution by reducing peak broadening. Further improvements to resolution were seen when the reflectron was introduced. This corrects for ions of the same m/z arriving at the detector at different times due to differences in kinetic energy, by acting as an ion mirror. Ions with greater kinetic energy, and therefore velocity travel deeper into the reflectron than those with less kinetic energy, so that ions with the same m/z reach the detector at the same time. The detector then detects the ions upon their collision with it and produces a spectrum known as a PMF (Yates 2000) (Figure 13).

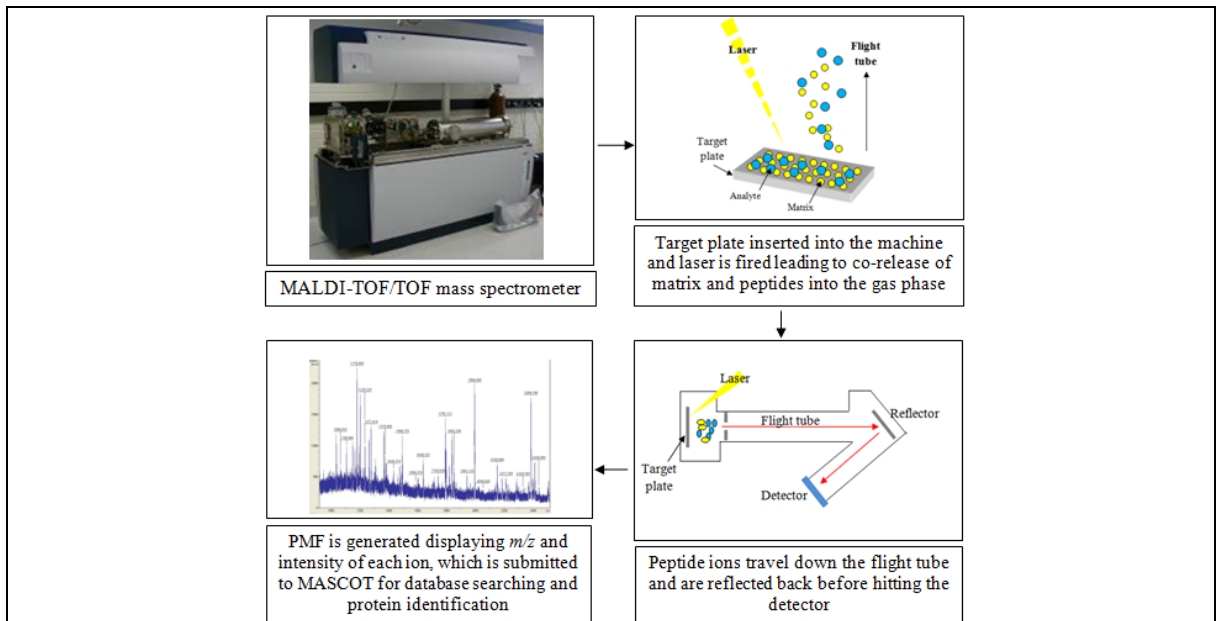


Figure 13: MALDI-TOF MS

The target plate, containing matrix and peptide molecules is inserted into the mass spectrometer. Peptide samples are ionised using MALDI, and subsequently travel down the flight tube where they are reflected back by the reflector towards the detector. The peptide ions hit the detector, which calculates the m/z of each ion and their intensity, which is displayed as a PMF. This is then submitted to a database such as MASCOT, for database searching and protein identification.

3.3.2.8 Protein identification

The list of selected PMF peaks is submitted to a database search (using a search engine such as Mascot from Matrix Science), which compares it to PMFs of theoretical protein tryptic digests, using databases such as the National Centre for Biotechnology Information non-redundant (NCBI nr) protein database or the International Protein Index (IPI) database. The latest version of the IPI Human database (released on 27th September 2011) contains 339,363 referenced entries in total.

3.3.2.9 Tandem mass spectrometry

Tandem mass spectrometry enables the analysis of biomolecules to be performed in much greater depth. For protein analysis, this method involves fragmentation of peptides which allows sequencing and structural information to be elucidated. There are several different types of fragmentation which require different energies and vary between different mass spectrometers, including post-source decay (PSD), laser induced dissociation (LID), collision-induced dissociation, (CID), surface induced dissociation (SID), electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) (Sleno and Volmer 2004). A common method of peptide fragmentation is collision-induced dissociation (CID). This involves the collision of ions with inert gas molecules, which generates vibrational internal energy within the ion, transforming it into an excited state. This results in decomposition of the activated ion and ultimately fragmentation by cleavage at the weakest bonds. Peptide chain cleavages may occur at either the C¹-C, C-N or N-C¹ bond, yielding six types of fragment ions; respectively a_n, b_n and c_n where the positive charge remains at the N-terminal fragment and x_n, y_n and z_n where the positive charge remains at the C-terminal fragment. Within peptides, the weakest bond is the amide bond, thus generating mainly b_n and y_n ions upon fragmentation (Wysocki, Resing et al. 2005). Protein identification by

tandem MS can be performed with the addition of a second mass analyser, which separates fragment ions further to give amino acid sequence data, thus increasing the accuracy and confidence of the identification. Different commercially available machines may have modified methods for analysis of molecules by tandem MS, such as the Ultraflex III (Bruker Daltonics), which will be discussed in chapter 9.

3.4 Gel-free methods

For many years 2D-PAGE/MS, which is a global, comparative, quantitative proteomic technique, was the gold standard for analysis of protein expression and biomarker discovery. However, there are several drawbacks associated with gel-based proteomic techniques. DEPs may not be recovered from the gel, whilst contamination with keratins can be introduced during the numerous experimental stages. Throughput is low and gel to gel reproducibility can be a challenge. Co-migration of proteins can cause problems during the excision and identification steps as there may be more than one protein present in the gel spot excised. This results in a mixed PMF from more than one protein, making a positive identification unfeasible. In addition 2D-PAGE may exclude extreme examples of hydrophobic, acidic, basic, small, large and low-abundance proteins. Membrane-spanning hydrophobic proteins and high molecular weight proteins may not dissolve in the IEF buffer and will therefore not enter the gel used for the 2nd dimension of separation (He, Liu et al. 2007). Also, low-abundant proteins may be masked in the gel by high-abundant proteins. Therefore, in more recent years, there has been a move towards gel-free MS methods for proteome analysis. These gel-free methods are based on the high-throughput “shotgun” analysis of peptides from a digested complex protein sample using a high performance liquid chromatography (HPLC) method for the separation of proteins within the sample, prior to identification using MS (Wu and MacCoss 2002). There are several

types of HPLC which can be used for the separation of complex protein samples, including ion exchange, reverse phase, affinity or a multidimensional approach which combines more than one separation method prior to MS (Yates, Ruse et al. 2009). An example of such an approach may involve two-dimensional cation exchange chromatography, where separation is based on charge, coupled to a reverse phase column, where separation is based on hydrophobicity. This is a popular approach, known as multidimensional protein identification technology (MudPIT) (Yates, Ruse et al. 2009). The separated sample may be eluted directly into the ionisation source, commonly ESI, in preparation for mass analysis.

3.4.1 ESI MS

ESI produces charged solvent droplets when a high electric potential is set between a capillary and the inlet to a mass spectrometer (Lin, Tabb et al. 2003). Typically peptides or proteins are analysed as positive ions using the capillary as an anode and the mass spectrometer inlet as the cathode. The sample to be analysed is digested in an appropriate solvent (such as 10:90 acetonitrile/water with 1% acetic acid) and injected directly into the instrument. ESI produces mainly doubly charged ions of tryptic peptides which allows the determination of m/z . Ions are generated directly from the solution by a fine spray. As the droplet size decreases, due to solvent evaporation, the electric charge density on the surface increases. Mutual repulsion between like charges on the surface break the surface tension and ions begin to leave the droplet and enter the gas phase (Ho, Lam et al. 2003). Ions are then accelerated into the mass analyser for determination of m/z and abundance. ESI is readily coupled to HPLC as it uses a steady stream of solvent to continuously produce ions. The peptide mixture is first separated by HPLC, which is coupled on-line to ESI-MS to gain product ion data. ESI sources are combined with tandem mass spectrometry (MS/MS), which comprises at least two stages of mass analysis for the generation of peptide sequence

data. An ESI source can be coupled to a variety of mass analysers including quadrupole, ion trap, orbitrap or Fourier transform ion cyclotron resonance systems (Ahmed 2008; Yates, Ruse et al. 2009), offering a variety of combinations which differ in cost, complexity, sensitivity, accuracy and function.

3.4.2 Quantitative Shotgun Proteomics

Unlike proteomic analysis based on 2D-PAGE/MS, traditional shotgun proteomics was not initially a comparative or quantitative approach and was mainly used for the identification of proteins in a given sample. However, recent methodological advances in MS have allowed the emergence of quantitative gel-free MS-based shotgun proteomic approaches (Chen and Yates 2007).

Stable isotopes can be used to differentially label protein samples, providing a comparative and quantitative LC-MS/MS analysis. Stable isotope labelling by amino acids in cell culture (SILAC) is employed to label proteins during cell culture prior to mass spectrometry analysis (Chen and Yates 2007). Isotope-coded affinity tagging (ICAT) reagents can be used to label paired protein extracts (Gygi, Rist et al. 1999) or the multiplex analysis of up to 8 samples can be achieved using isobaric tags for relative and absolute quantification (iTRAQ) technology (Ross, Huang et al. 2004; Aggarwal, Choe et al. 2006; Choe, D'Ascenzo et al. 2007). Label-free approaches have also been developed for quantitative shotgun proteomics. These include absolute quantification (AQUA), selected reaction monitoring (SRM) and multiple reaction monitoring (MRM) assays (Kim and Kim 2009; Pan, Aebersold et al. 2009).

3.5 Microarray-based methods

The use of conventional MS-based proteomic approaches relies on an initial sample separation or fractionation step, which may reduce the range of proteome interrogation, as well as the identification of proteins from a public database. However, some proteins may not generate a sufficient number of peptides via this technique to gain a significant identification, or the specific form of protein may not be represented in the database. Microarray-based proteomics offers a range of methods to complement traditional MS-based approaches. Microarray-based proteomic methods can be employed in ‘reverse-phase’ (where multiple test samples are immobilised for simultaneous screening with an antibody or probe) or ‘forward-phase’ (where multiple monoclonal antibodies are immobilised for simultaneous screening using protein lysates) (Caiazzo, Maher et al. 2009). Antibody microarrays are a relatively new and powerful tool, which offer high-throughput multiplex screening on non-fractionated complex proteomes from a variety of clinical and biological samples. Proteomic analysis of chosen samples can be performed simultaneously, with high sensitivity and specificity, detecting proteins within the ng/ml range, whilst overcoming some of the problems associated with gel-based and MS-based approaches, such as the dynamic range of complex proteomes. Antibody microarrays therefore offer a valuable complementary technique to traditional MS-based approaches (Kopf, Shnitzer et al. 2005; Kopf and Zharhary 2007). It is important to note that antibody microarrays cannot be considered as a ‘global’ proteomic approach, as they are only able to analyse the expression of proteins that correspond to the pre-selected antibodies printed on to the slide. Antibodies chosen can be related to proteins involved a specific signalling pathway (for example apoptosis), or a variety of different signalling pathways. An antibody microarray is a collection of antibodies spotted in an ordered pattern onto a nitrocellulose-

coated glass microscope slide. Key technological issues which are considered during the design of an antibody microarray platform include: (1) antibody content; (2) array design / format; (3) array fabrication (4) assay design (5) sample handling and (6) data handling (Wingren and Borrebaeck 2008; Borrebaeck and Wingren 2009). The antibodies printed onto the array should be highly-specific and well-characterised to minimise potential problems with lack of specificity or cross-reactivity (Borrebaeck and Wingren 2009). The antibody microarray can be used to compare protein expression profiles of two samples (for example, chemotherapy-sensitive versus chemotherapy-resistant). Protein samples can be directly labelled with fluorescent dyes (conventionally Cy3 or Cy5), which are mixed in equal amounts and co-incubated with the slide (Figure 14) The expression of a protein is detected when it competitively binds to its corresponding antibody spotted on the slide. A drawback associated with this type of assay is that the fluorescent label may interfere with the antigen-antibody interaction, which may limit detection (Sanchez-Carbayo 2006). It is therefore important to optimise parameters such as the dye-to-protein molar ratio to ensure that a balance is achieved, and a protein is not under- or over-labelled. The slide is analysed using a fluorescent scanner and the relative amount of each dye present on each antibody spot (levels of Cy3 versus Cy5), corresponding to the relative abundance of protein bound from each sample, is measured by signal intensity (Figure 14). Differentially expressed proteins between samples are then identified, by relative fluorescence of the dyes, and fold changes between samples can be calculated, where a 2-fold difference in expression is generally accepted as a significant finding (Smith, Watson et al. 2006). The expression of hundreds of proteins can be analysed simultaneously. An example of a large-scale commercially available antibody microarray kit is the Panorama[®] Antibody Microarray-XPRESS Profiler725 from Sigma Aldrich. This array contains 725 antibodies, (listed in

Appendix 1) spotted (robotically) in duplicate, onto a nitrocellulose-coated glass slide, and is capable of analysing the expression of a wide variety of proteins, including those involved in cell-signalling, apoptosis, cell cycle control and proliferation. The use of antibody microarrays is expanding, and has been recognised as a powerful tool in cancer proteomics (Kopf and Zharhary 2007), including breast cancer (Celis, Moreira et al. 2005; Celis, Moreira et al. 2005; Smith, Watson et al. 2006), as well as biomarker discovery and potential for the development of diagnostic assays (Brennan, O'Connor et al. 2010). However, due to the current high cost of commercially available kits, replicate antibody microarray experiments can become expensive and protein identifications will always be limited to those antibodies immobilised on the slide.

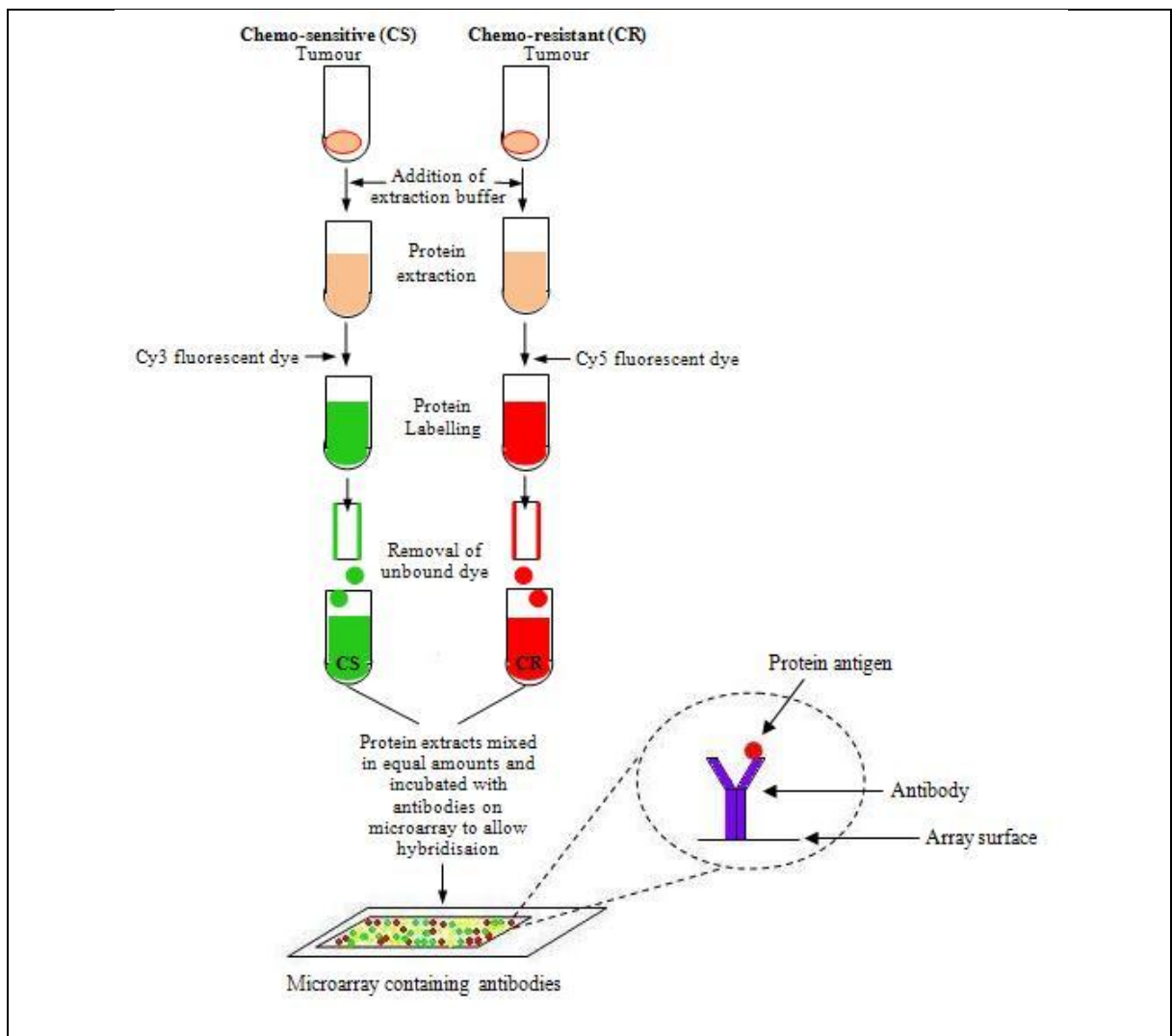


Figure 14: Antibody microarray workflow

Proteins are extracted from the two samples selected for comparative analysis, and labelled with different fluorescent dyes; Cy3 (control) and Cy5 (test). Unbound dye is removed to leave protein-dye complexes, and samples are co-incubated with the antibody microarray slide in equal amounts to allow competitive protein binding. The relative abundance of each protein in each sample is determined by measuring the relative fluorescence of each dye on each antibody spot. Differentially expressed proteins are identified where proteins show ≥ 2 -fold change in expression between each sample.

3.6 Techniques for the confirmation and validation of putative biomarkers

3.6.1 Data mining

The selected proteins identified within the discovery phase, to be carried forward to the confirmation phase, can be a long and difficult process as proteins must be selected in a logical and meaningful manner. A useful tool to aid the prioritisation of proteins, which are to be carried forward, is software which analyses and interprets the data using knowledge bases (section 10.1.1). An example of such software is Ingenuity Pathway Analysis (IPA) (Ingenuity Systems Inc., USA) (Jimenez-Marin, Collado-Romero et al. 2009). Protein lists can be uploaded into IPA software, where they are mapped against the Ingenuity Knowledge Base, to highlight direct relationships between candidate proteins using networks and canonical pathways. This highlights and prioritises the candidates of most interest, whilst aiding understanding and presenting potential hypotheses, giving researchers informative direction for downstream confirmation.

3.6.2 Western Blotting

Due to the high throughput, simultaneous analysis of large numbers of proteins when using proteomic techniques there is a high chance of false discovery (Qian and Huang 2005). Therefore, a second independent technique (for example western blotting) is used to confirm the identification and expression change of individual putative biomarkers which have been suggested from proteomic studies. Western blotting, also known as immunoblotting, was first introduced in 1979 (Towbin, Staehelin et al. 1979), yet it was not termed western blotting until a paper was published naming it this in 1981 (Burnette 1981). Firstly, protein extraction from cell line origin or tissue is achieved by chemical or mechanical lysis of the cells in the presence of Laemmli buffer, which contains reagents

such as sodium dodecyl sulphate (SDS) that induce the unfolding and reduction of proteins which is required for effective separation by molecular weight. Extracted proteins are then separated by 1D-PAGE, according to molecular weight and subsequently transferred onto a nitrocellulose membrane. Nitrocellulose membranes have a high affinity for proteins, which accommodates the transfer of proteins from the gel. Once the proteins from the extract are on the nitrocellulose membrane, it is necessary to 'block' the free sites on the membrane where proteins have not bound, using a blocking solution (consisting of bovine serum albumin or non-fat dried milk powder). This prevents binding of the antibodies to the membrane as opposed to the target protein of interest. It is then possible to 'probe' for specific proteins with antibodies to these proteins, by incubation with the 'primary antibody', which is specific to the protein under investigation. After washing steps to remove unbound antibody, a method widely used for the visualisation of protein expression is the use of chemiluminescent detection, using a horseradish peroxidase (HRP) conjugated 'secondary antibody. A chemiluminescent agent is used to initiate a reaction with the HRP, which produces luminescence in proportion to the amount of protein. The luminescence can be used to visualise the protein band by exposing the membrane to a photographic film. Densitometry can then be used to give a ratio of the level of protein between samples. To ensure that equal amounts of protein have been loaded onto the gel the membrane is probed with an anti-'housekeeping' antibody (for example anti-GAPDH, anti-beta-actin or anti-alpha tubulin) which should demonstrate constant levels of expression and act as a loading control (Aldridge, Podrebarac et al. 2008). Densitometry can be used to normalise gel loading differences and compare the intensity of the bands produced by the primary antibody from each sample, producing a quantitation of the fold-change in expression. Immunoblotting relies on the availability of a reliable specific primary antibody to the

precise protein species which has been identified by the proteomic analysis. Conventional immunoblotting is a low-throughput technique but, if the appropriate antibodies are represented, the method can be performed using high-throughput membrane-based techniques such as PowerBlot (Yoo, Piechocki et al. 2002) or Profiler array membranes (Oliveras-Ferraros, Vazquez-Martin et al. 2008). An alternative approach to immunoblotting formats is provided by the enzyme-linked immunosorbent assay (ELISA), where the protein lysate and test primary antibody are analysed in a microtitre plate-based experiment.

3.6.3 Clinical Validation

Putative protein biomarkers which successfully pass through technical confirmation experiments are then further analysed using methods which can provide clinical validation. This relies on the availability of a sufficient number of suitable clinical samples, each with the required clinical information. The proteomic methodologies described do not give any information regarding the cellular localisation of putative protein biomarkers. Immunohistochemistry (IHC) is a low through-put technique which can be used to validate the expression and localisation of a protein of interest in whole sections of formalin-fixed, paraffin-embedded (FFPE) clinical tissue samples on glass microscope slides. High throughput IHC can be achieved if effort is first invested in the production of a suitable tissue microarray (Camp, Neumeister et al. 2008). For this, cores of tissue are taken from hundreds of different donor FFPE samples and co-embedded into a single new recipient FFPE block. A single slide of the composite tissue section (consisting of up to 800 individual tissue core samples) can then be simultaneously screened for expression of one test protein using IHC. For non-tissue based assays an ELISA format, with an immobilised capture antibody, can be used to simultaneously screen protein samples from multiple

clinical samples. Alternatively MS-based assays (such as MRM analysis) could be designed for the quantitation and validation of putative biomarkers in clinical samples. After successful validation and rigorous testing of putative biomarkers from this discovery pipeline (McShane, Altman et al. 2005), a few may eventually find a role in clinical use.

3.7 Proteomics research to identify biomarkers of chemotherapy resistance in breast cancer

Several global methods have been used in the attempt to identify markers of response to chemotherapy in breast cancer, including genomic, transcriptomic and proteomic approaches. To date, proteomics-based studies are largely *in vitro* based, identifying differential expression between parental and chemotherapy-resistant cell sub-lines. The first requirement for proteomic analysis is sample acquisition and a relatively large quantity of sample is required for proteomics due to the inability to amplify at the protein level. The non-invasive collection of biofluids (such as blood serum, plasma, tumour interstitial fluid, nipple aspirate fluid etc) for proteomic identification of circulating predictive biomarkers is an attractive approach however there are a number of technical drawbacks (Schrohl, Wurtz et al. 2008; Tuck, Chan et al. 2009). A major challenge to the proteomic analysis of blood samples is based on the very broad dynamic range (approximately 10 orders of magnitude) (section 6.1.1.2) of protein expression levels in plasma (Anderson and Anderson 2002). The most abundant protein present in serum is albumin and, in fact, 95% of the protein content of serum is made up of around 20 highly abundant proteins, removal of which is possible prior to proteomic analysis in order to decrease masking the expression of low abundant putative biomarkers (Roche, Tiers et al. 2009). The extreme low abundance of protein biomarkers which may be of potential clinical relevance also necessitates the highest level of sensitivity in the mass spectrometry instrument. The proteomic profile of proteins which are released by cancer cells into the surrounding interstitial fluid or blood can also be

analysed for biomarker discovery. This type of cancer “secretome” analysis has yielded a number of putative biomarkers (Xue, Lu et al. 2008). Tumour interstitial fluid may carry the highest concentration of tumour-specific proteins and a proteomic analysis of breast cancer interstitial fluid was able to identify over 1000 proteins involved in a vast array of biological activities (Celis, Gromov et al. 2004). Predictive biomarkers of chemotherapy response in breast cancer patients may also be identified from the tumour interstitial fluid proteome (Cortesi, Barchetti et al. 2009). The type of biological fluid analysed also depends on the type of biomarker being searched for; this may be diagnostic, prognostic or predictive. For example, a diagnostic biomarker which could be identified from a blood sample would be more clinically accessible than a diagnostic biomarker present within the tumour interstitial fluid.

Cultured cell lines or clinical samples of tumour tissue can be analysed using proteomics, however tissue samples may be more technically challenging (section 6.1.1). Although the use of cell lines as models of breast cancer can not accurately mimic the tumour in its biological microenvironment (Lacroix and Leclercq 2004), research using cell lines has been prominent in the proteomics literature since they are much easier to handle within controlled conditions.

The majority of published proteomics-based studies for the identification of biomarkers of chemotherapy-resistance in breast cancer are based upon research using cell lines, which are used as *in vitro* models to simulate a clinical scenario. This requires the *in vitro* establishment of new chemotherapy-resistant cell sub-lines using a choice of strategies (Watson, Lind et al. 2007), so that their proteome can be compared to that of the parental cell line and DEPs can be identified. Many DEPs have been identified from chemotherapy-resistant breast cancer cell sub-lines, mainly derived from luminal-type (ER-positive)

MCF-7 parent cells (Lacroix and Leclercq 2004), using 2D-PAGE/MS as a global proteome screening technique. The DEPs (with at least a 2-fold change in expression) identified by these published studies are summarised in Table 4. An MCF-7 cell sub-line demonstrating low-level resistance to the alkylating agent cisplatin was established by treating MCF-7 cells with 7 cyclic, 24-hour treatments of 50 μ M cisplatin and 15 DEPs were identified in the resistant MCF-7 cells using 2D-PAGE/MALDI-TOF-MS (Smith, Welham et al. 2007) (Table 4). Semi-quantitative immunoblotting was used for technical confirmation of a subset of DEPs and independently identified the differential expression of the cytokeratin 17, glutathione-S-transferase mu 3 and peroxiredoxin 4 proteins in the resistant cells (Smith, Welham et al. 2007).

An MCF-7 sub-line which was 246 times more resistant to the taxane paclitaxel than parental cells was established by Wosikowski *et al* (Wosikowski, Regis et al. 1995) and subsequently analysed using 2D-PAGE/MALDI-TOF-MS (Chuthapisith 2007) (Table 4). The authors reported that immunoblotting of 14-3-3 epsilon, cytokeratin-19, HSP27, sorcin and stathmin provided results which were in agreement with the 2D-PAGE analysis. IHC analysis of HSP27, sorcin and stathmin in a small series of archival clinical samples from patients treated with neoadjuvant doxorubicin/cyclophosphamide followed by docetaxel did not reveal a statistical correlation with response (Chuthapisith, Bean et al. 2009).

A number of topoisomerase II poisons (eg doxorubicin, etoposide, mitoxantrone) (Nitiss 2009) have been analysed using drug-resistant MCF-7 sub-lines and proteomic approaches. Several studies have focused on the analysis of the anthracycline antibiotic doxorubicin (adriamycin; Adr) in drug-resistant MCF-7 cell sub-lines. The original long-established doxorubicin-resistant MCF-7/AdrR (or MCF-7/ADR) sub-line created in 1986 (Batist, Tulpule et al. 1986) has recently been re-classified since it was shown to be derived from

ovarian cancer cells (Liscovitch and Ravid 2007) and therefore proteomic studies which may have been performed on this cell line have been excluded from this discussion. The MCF-7/AdrVp sub-line was established by selecting MCF-7 cells which survived incremental doses of doxorubicin in the presence of the P-glycoprotein membrane pump inhibitor verapamil (Vp), resulting in 900-fold resistance to doxorubicin. Nuclear proteins were sub-fractionated prior to analysis by 2D-PAGE/MALDI-TOF-MS and MS/MS, revealing 7 differentially expressed proteins (Fu and Fenselau 2005) (Table 4). An alternative doxorubicin-resistant MCF-7 cell sub-line, MCF7/AdVp3000, again established in the presence of Vp, was analysed using 2D-PAGE/MALDI-TOF-MS (Liu, Liu et al. 2006). Sub-cellular fractionation was not carried out in this case and the proteins which demonstrated greater than 2-fold change in expression are shown in Table 4. Due to an association with the p53 pathway, the 14-3-3 sigma (stratifin) protein was selected for further *in vitro* functional analysis, which confirmed the association of 14-3-3 sigma over-expression with drug-resistance in this cell sub-line (Liu, Liu et al. 2006).

A sub-line of MCF-7 which was 28 times more resistant to etoposide (VP-16) than parental cells was established by selection in increasing drug concentrations. Pre-fractionation of nuclear proteins and analysis by 2D-PAGE/MALDI-TOF-MS and MS/MS revealed a number of differentially expressed proteins which were also identified in doxorubicin-resistant sub-lines (Fu and Fenselau 2005) (Table 4). Furthermore the analysis of nuclear proteins from an MCF-7 sub-line which was 4000-fold resistant to mitoxantrone revealed a number of differentially expressed proteins that were also seen in doxorubicin- and/or etoposide-resistant cells (Fu and Fenselau 2005) (Table 4). The down-regulation of the cytoskeletal proteins alpha tropomyosin, cytokeratin 8, cytokeratin 19 and septin 2 was

seen in all three sub-lines which were resistant to the topoisomerase II poisons doxorubicin, etoposide or mitoxantrone (Fu and Fenselau 2005) (Table 4).

Alternative approaches to 2D-PAGE have been used in a few studies of drug-resistant MCF-7 sub-lines. The mitoxantrone-resistant MCF-7 sub-line described above was further analysed by MS following pre-fractionation of plasma membrane proteins (Rahbar and Fenselau 2005). This approach identified 15 further proteins which demonstrated a difference in expression of at least 2-fold between resistant and parental cells. Antibody microarray analysis has not been utilised as yet in MCF-7 drug-resistant cell lines, however a doxorubicin-resistant sub-line derived from triple-negative MDA-MB-231 cells was analysed using a cell signalling microarray slide composed of 224 antibodies (Smith, Watson et al. 2006). Decreased expression (at least 2-fold) of cyclin B1, cyclin D2 and p-ERK was identified from the microarray and confirmed by immunoblotting in the drug-resistant cells.

Several proteomics-based studies have been carried out using fresh breast cancer tissue samples however these are mainly based on comparisons between the proteome of normal versus malignant tissue rather than investigations of chemotherapy response (Deng, Xing et al. 2006; Othman, Majid et al. 2008). These proteomic studies have addressed the technical issues regarding the heterogeneity of breast cancer tissue through the use of laser capture microdissection (Hudelist, Singer et al. 2006), the presence of high-abundance proteins from contaminating blood serum through the use of depletion strategies (Kim, Bae et al. 2009) and also successfully utilised the limited amount of tissue available in pre-treatment diagnostic biopsies (Bisca, D'Ambrosio et al. 2004). A search for biomarkers which predict response to neoadjuvant chemotherapy using proteomic methods in breast cancer tissue has not been published as yet. However a 2D-PAGE/MALDI-TOF-MS investigation of breast

cancer samples from patients treated post-operatively with cyclophosphamide/methotrexate/5-fluorouracil revealed a number of putative biomarkers which correlated with tumour recurrence (Nimeus, Malmstrom et al. 2007).

Table 4: Proteins demonstrating differential expression in chemotherapy-resistant MCF-7 cell sub-lines identified by 2D-PAGE/MS.

Putative biomarkers demonstrated an increase (↑) or decrease (↓) in expression of at least 2-fold in resistant cells. Drugs, to which resistance is being displayed, which are most relevant to this project; anthracyclines (doxorubicin) and taxanes (paclitaxel) are shown in bold.

Drug resistance	Drug mechanism	Putative biomarker	Reference
Cisplatin	DNA damaging agent	Beta-tubulin type 3 (↓)	Smith, 2007
Cisplatin		Cytokeratin 17 (↓)	Smith, 2007
Cisplatin		Electron transfer flavoprotein beta (↑)	Smith, 2007
Cisplatin		Glutathione-S-transferase mu 3 (↓)	Smith, 2007
Cisplatin		Heterogeneous nuclear ribonucleoprotein A3 (↑)	Smith, 2007
Cisplatin		HSP 27 (↓)	Smith, 2007
Cisplatin		Hydroxyprostaglandin dehydrogenase-15 (NAD) (↑)	Smith, 2007
Cisplatin		Isocitrate dehydrogenase 3 (↓)	Smith, 2007
Cisplatin		Matrix metalloproteinase 9 (↑)	Smith, 2007
Cisplatin		Peptidyl-prolyl isomerase	Smith, 2007
Cisplatin		Peptidyl-prolyl isomerase B (↑)	Smith, 2007
Cisplatin		Peroxiredoxin 4 (↓)	Smith, 2007
Cisplatin		Proteasome beta 1 subunit (↑)	Smith, 2007
Cisplatin		Ribosomal protein P0 (↓)	Smith, 2007
Cisplatin		Tropomyosin 1-alpha (↓)	Smith, 2007
Paclitaxel	Anti-microtubule agent	14-3-3 epsilon (↑)	Chuthapisith, 2007
Paclitaxel		Cytokeratin 19 (↓)	Chuthapisith, 2007
Paclitaxel		HSP 27 (↓)	Chuthapisith, 2007
Paclitaxel		Phosphoglycerate kinase-1 (↓)	Chuthapisith, 2007
Paclitaxel		Proliferating cell nuclear antigen (↑)	Chuthapisith, 2007
Paclitaxel		Sorcin (↑)	Chuthapisith, 2007
Paclitaxel		Stathmin (↑)	Chuthapisith, 2007
Doxorubicin	Topoisomerase II poison	14-3-3 sigma (stratifin) (↑)	Liu, 2006
Mitoxantrone		40S ribosomal protein SA (↑)	Fu, 2005
Etoposide		40S ribosomal protein SA (↑)	Fu, 2005
Doxorubicin		Alpha tropomyosin (↓)	Fu, 2005
Mitoxantrone		Alpha tropomyosin (↓)	Fu, 2005
Etoposide		Alpha tropomyosin (↓)	Fu, 2005
Doxorubicin		ATP Synthase β (↑)	Chuthapisith, 2007
Doxorubicin		Cathepsin D, chain B (↑)	Chuthapisith, 2007
Doxorubicin		Cyclophilin B (↑)	Fu, 2005
Etoposide		Cyclophilin B (↑)	Fu, 2005
Doxorubicin		Cytokeratin 8 (↓)	Fu, 2005

Mitoxantrone		Cytokeratin 8 (↓)	Fu, 2005
Etoposide		Cytokeratin 8 (↓)	Fu, 2005
Doxorubicin		Cytokeratin 19 (↓)	Fu, 2005
Doxorubicin		Cytokeratin 19 (↑)	Chuthapisith, 2007
Mitoxantrone		Cytokeratin 19 (↓)	Fu, 2005
Etoposide		Cytokeratin 19 (↓)	Fu, 2005
Mitoxantrone		Glucose-regulated protein 78K (GRP78) (↑)	Fu, 2005
Mitoxantrone		HMG-1 (↑)	Fu, 2005
Etoposide		HMG 1 (↑)	Fu, 2005
Doxorubicin		Mitotic checkpoint protein BUB 3 (↑)	Fu, 2005
Mitoxantrone		Nucleolin (↑)	Fu, 2005
Etoposide		Nucleolin (↑)	Fu, 2005
Mitoxantrone		PARP-1 (↓)	Fu, 2005
Etoposide		PARP-1 (↓)	Fu, 2005
Doxorubicin		Peroxiredoxin 2 (↓)	Chuthapisith, 2007
Doxorubicin		Peroxiredoxin 6 (↑)	Chuthapisith, 2007
Mitoxantrone		Prohibitin (↑)	Fu, 2005
Doxorubicin		Protein disulphide isomerase (↑)	Chuthapisith, 2007
Doxorubicin		Septin 2 (↓)	Fu, 2005
Mitoxantrone		Septin 2 (↓)	Fu, 2005
Etoposide		Septin 2 (↓)	Fu, 2005
Doxorubicin		Septin 7 (↓)	Fu, 2005
Doxorubicin		Triose-phosphate isomerase (↑),(↓)	Chuthapisith, 2007

To summarise, the use of clinical proteomics is potentially an excellent approach for the discovery of predictive biomarkers that can be used in the future for individualisation of treatment for breast cancer patients. Complementary techniques such as 2D-PAGE/MS and antibody microarrays allow the simultaneous analysis of many proteins in a single sample, which is required for profiling complex cellular changes in cancer.

The above-mentioned studies present a list of putative biomarkers of chemotherapy resistance, identified by 2D-PAGE/MS in breast cancer MCF7 cell lines. When comparing these putative biomarkers of anthracycline / taxane chemotherapy resistance to those presented in Table 3 (section 2.4.4), there appears to be little overlap, or obvious common theme. There is therefore a need for further research into the search for putative biomarkers of neoadjuvant chemotherapy resistance in breast cancer using proteomics techniques, where methods utilising clinical tumour tissue have not yet been reported.

Protein markers which are already in routine clinical use include ER, PR and HER2. These markers provide information on molecular subtype and prognosis, and are used to determine the type of therapy administered. However, despite this partially-tailored therapy, successful treatment can still not be guaranteed. In order to maximise treatment efficiency, greater understanding of the tumour proteome is required, facilitating the ability to predict therapy response on an individual tumour basis. This is an area which would benefit greatly from further research as a major obstacle in effective tumour treatment is the occurrence of tumour resistance to therapeutic agents being administered. In the event of tumour resistance to therapy not only are resources wasted but, more importantly, patients are exposed to cytotoxic drugs which cause unpleasant side-effects unnecessarily, and the tumour still remains. As discussed, using global approaches several proteins have been identified as putative biomarkers of therapy response from cell line models, however as yet none have been validated for routine use in the clinical setting.

3.8 Project Aim

The aim of this project is to use proteomics technologies for the detection of proteins associated with chemotherapy resistance in breast cancer, specifically with the use of clinical tumour tissue obtained from locally advanced breast cancer (LABC) patients who received standard (epirubicin/cyclophosphamide plus docetaxel) neoadjuvant chemotherapy. The specific aims of the project are:

- Collection of clinical samples and clinical information
- Optimisation of proteomics methods for breast cancer tissue, based on existing methods for breast cancer cell lines and overcoming problems associated with the use of tissue.

- To identify biomarkers of chemotherapy resistance using the biomarker discovery pipeline;
 - The generation of lists of differentially expressed proteins (DEPs) using 2D-PAGE coupled to MALDI TOF/TOF MS and antibody microarray analysis
 - Performance of data-mining with the use of Ingenuity Pathway Analysis software to identify potential relationships between DEPs and significant canonical pathways the DEPs may be involved in.
 - The confirmation of the DEPs using western blotting
 - Clinical validation of DEPs using immunohistochemistry, to assess their clinical relevance and predictive ability using pre-treatment samples, in order to identify putative markers of chemotherapy resistance in breast cancer.

The identification of protein biomarkers which predict tumour response to chemotherapy will be of great value to both the patient and the clinician, as occurrence of tumour resistance to therapy is currently a major obstacle in effective tumour treatment. In the event of tumour resistance to therapy, patients are exposed to strong cytotoxic drugs for no therapeutic gain, during which time the tumour may even progress. The ability to predict response to chemotherapy at the diagnostic stage would allow treatment to be tailored to the individual patient, and administration of chemotherapeutic agents to only those who are likely to show a positive response.

CHAPTER 4:
MATERIALS AND METHODS

Chapter 4. Materials and Methods

4.1 Culture of Cell Lines

4.1.1 Human Caucasian Breast Adenocarcinoma ‘MCF7’ Cell Line

This adherent cell line was originally established from the pleural effusion of a breast adenocarcinoma in a Caucasian 69-year old female. The cells exhibit epithelial-like morphology, are ER and PR positive and are therefore of Luminal subtype (section 2.2.2).

4.1.2 Thawing Cells

Cells which had previously been stored at minus 80 °C were thawed quickly, inside a sealed plastic bag, in a water bath at 37 °C. Once thawed, the cells were transferred to a 30 ml screw-cap universal tube and were then diluted with 9 ml of culture media (1:10 dilution) over 1-2 min, to allow the cells to acclimatise to their new environment. The cell suspension was then spun at 1600 rpm for 3 min, which pelleted the cells thus removing them from the DMSO (section 4.1.4) which they had been stored in whilst frozen. This was necessary as DMSO is toxic to cells when they are not frozen. The supernatant was discarded and the cells were re-suspended in an appropriate volume of fresh tissue culture media (Appendix 2), pre-heated to 37 °C, in either a T25 (25 cm² area on largest side) or T75 (75 cm² area on largest side) flask depending upon the size of the pellet. The flask of cells was then placed in an incubator in a humid atmosphere at 37 °C with 5% CO₂.

4.1.3 Culturing Cells

All equipment, including the tissue culture hood, water bath and incubator was cleaned thoroughly at regular intervals with virkon disinfectant and 70% alcohol to ensure the area

was contaminant-free before cells were removed from the incubator. To maintain this clean environment, and reduce the risk of infection, sterile technique was adopted at all times during cell culture and all equipment was sprayed thoroughly with 70% alcohol before placing in the Class II tissue culture hood. Cells were cultured in RPMI tissue culture media (Appendix 2) at 37 °C, in an atmosphere of 5% CO₂, to represent the conditions in the human body and to maintain the pH of the media.

Cells were cultured in T75 flasks, which were changed along with the media, 3 times each week. The medium was pre-warmed in the water bath at 37 °C for 30 min before use to ensure the cells experienced a minimal amount of stress when they were transferred into their new environment. Adherent cells were removed from their flasks by trypsinisation using pre-warmed TrypLE Select (#12563, Invitrogen). Trypsin was added to the flask to remove adherent cells, at a volume of 3 ml (for T25 and T75 flasks), gently agitated to ensure complete coverage of cells, and incubated at 37 °C for 3-5 min. The flasks were lightly tapped to ensure the cells were no longer adhered, and 7 ml media was added to saturate the action of the trypsin. The cell suspension was centrifuged at 1600 rpm for 3 min and the pellet of cells was re-suspended in a suitable volume of medium and transferred to a new flask.

4.1.4 Freezing Cells

Cells were frozen when they reached 80% confluence. They were frozen in ‘freezing media’, which consists of tissue culture media containing 10% dimethyl sulphoxide (DMSO) (#D2650, Sigma Aldrich). Cells from each flask were pelleted and slowly re-suspended in 1 ml of freezing medium and transferred into a cryovial. The cells were then stored at minus 80 °C or in liquid nitrogen.

4.2 Collection of Clinical Samples

Patient selection for fresh tissue samples:

This study was approved by the South Humber Local Research Ethics Committee (ref 07/Q1105/43), and included patients receiving neoadjuvant chemotherapy for locally advanced breast cancer from 2007 onwards at Hull and East Riding Hospitals NHS Trust, Hull. The treatment regimen consisted of 4 cycles of EC [epirubicin (90 mg/m^2) + cyclophosphamide (600 mg/m^2)] followed by 4 cycles of docetaxel (100 mg/m^2), given at 3-weekly intervals. Five of the patients included in the study were assigned to the Neo-AnGo randomised phase III clinical trial of sequential epirubicin + cyclophosphamide and paclitaxel \pm gemcitabine. This was followed by resection of the residual tumour. Patient consent was obtained for a tumour sample to be taken at the time of definitive surgery, which was snap-frozen in liquid nitrogen and stored at minus $80 \text{ }^\circ\text{C}$ until required. Tumour samples collected varied in size from 2 mm^3 to $> 2 \text{ cm}^3$ and the number of pieces of solid tumour provided (in separate microcentrifuge tubes) by the surgeon ranged from 1 to 3. Consent was also obtained to allow access to relevant patient clinical information which included chemotherapy details, radiological and pathological results and reports for determination of response as well as molecular typing (ER/PR/HER2 status). The 8 tumour samples that were used for proteomic analysis (Figure 15) were all ductal tumours of luminal subtype. A consort chart is shown (Figure 15) which outlines which clinical samples were used for proteomic analysis and optimisation of methods. All samples are listed in Table 13 (section 5.3.1) and full clinical information is given in Appendix 5.

Patient selection for archival pre-treatment core biopsy samples:

Pre-treatment core biopsy samples from a previously characterised sample group (Garimella 2007) were used for this study. Ethical approval had previously been granted for the study entitled ‘monitoring the effects of chemotherapy in breast cancer patients using magnetic resonance imaging and molecular markers’ from the Hull and East Riding Research Ethics Committee (ref 03/00/038). All patients in this cohort were recruited between 2000 and 2002 from the Hull and East Riding NHS trust, and had histologically-proven breast cancer with a primary tumour of ≥ 3 cm. In total, 35 archival tissue samples were obtained from 36 locally advanced breast cancers (one patient had bilateral breast cancer). These comprised 75% ER-positive tumours and 69% PR-positive tumours. Patient consent was obtained to allow access to pre-treatment core biopsy samples and to perform serial DCE-MRI scans (pre-treatment, after 2nd cycle of chemotherapy, and post-treatment) so that tumour response to therapy could be monitored. Patients were treated with 6 cycles of 5-fluorouracil (200 mg/m²), epirubicin (60 mg/m²) and cyclophosphamide (600 mg/m²) (infusional FEC), administered at 3-weekly intervals. Tumour response was assessed after the 2nd cycle using DCE-MRI scans. Patients who showed a response continued with the full course of treatment, and where no response was observed, chemotherapy was terminated. Following this, definitive surgery was performed to remove residual tumour.

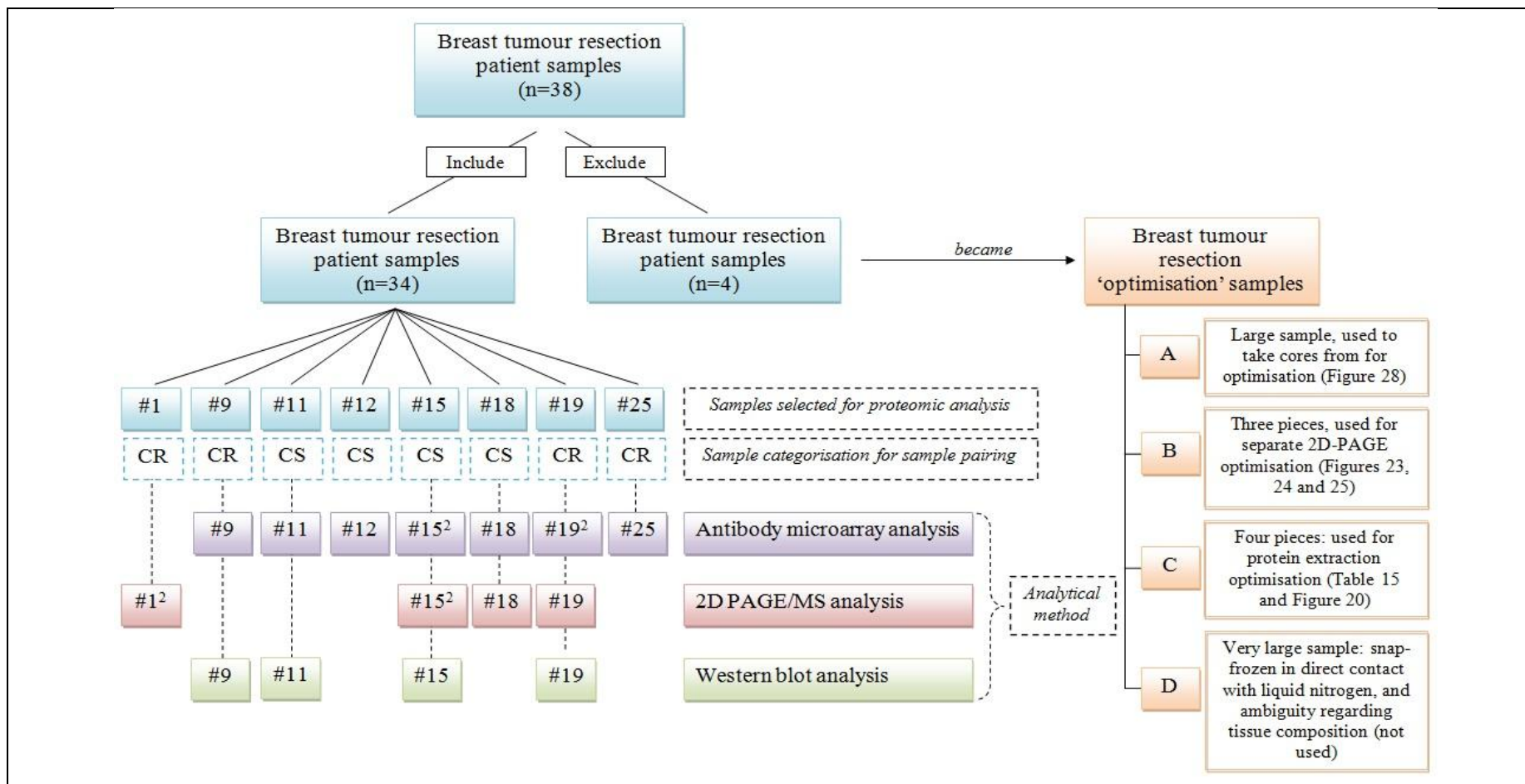


Figure 15: Consort chart outlining where clinical tumour resection samples have been used for proteomic analysis

Of the 38 breast tumour samples collected, 34 were carried forward to be considered for proteomic analysis (listed in Table 13). For samples which were included in more than one experiment (²) is shown. Four samples were excluded from the study and became 'optimisation samples', due to ambiguity regarding sample composition (A), insufficient information to accurately determine tumour response to treatment (B, C) and incorrect sample handling (D).

4.3 Panorama Antibody Microarray XPRESS725 Profiler

The Panorama Antibody Microarray XPress Profiler725 (#XP725, Sigma Aldrich) was used to compare protein expression in two different samples. Clinical breast tumour samples selected for antibody microarray analysis are shown in Figure 15.

Solutions which required preparation included:

- Protease Inhibitor Cocktail: 0.3 ml of dH₂O was added to the vial provided (#P4495, Sigma Aldrich) and the reconstituted solution was then stored at -20 °C.
- Benzonase Working Solution: 2 µl of Benzonase Ultrapure (#B8309, Sigma Aldrich) was added to 18 µl of Extraction/Labelling Buffer to give a solution of 5 units/µl.

4.3.1 Protein Extraction

For all proteomics work, polypropylene microcentrifuge tubes were used, to minimise contaminations from plastics and to prevent proteins/peptides being retained on the surface of the tubes. Proteomics-grade distilled water was also used throughout. Proteins were extracted from the tissue using the Antibody Microarray Extraction/Labelling buffer supplied in the kit. To each 10 ml of Extraction/Labelling buffer, 50 µl of the reconstituted Protease Inhibitor Cocktail, 100 µl of Phosphatase Inhibitor Cocktail II and 1.2 µl of the Benzonase Working Solution was added and kept on ice until required. This was then referred to as Buffer A. The addition of these reagents helps to maintain protein composition by inhibiting their breakdown in the sample and the Benzonase is added to remove nucleic acid present in the sample. Breast tumour tissue was removed from the minus 80 °C freezer and weighed. Work following this was carried out as quickly as possible in a Class II Tissue Culture Hood, using sterile technique to avoid contamination,

and the tumour tissue was kept on ice to ensure minimal protein degradation. The tissue was then placed in a sterile petri dish on ice, cut into small pieces using a disposable scalpel and washed by bathing and agitating gently in chilled PBS to remove any residual blood. Four volumes (w/v) of Buffer A was added to the tissue and the tissue was then homogenised using a TissueRuptor (#9001273, Qiagen) with a disposable probe (#990890, Qiagen) on ice in the tissue culture hood. The sample was then centrifuged for 10 s at 10,000 x g in a microcentrifuge tube. If the sample contained fat, as a layer at the top of the supernatant, the supernatant below this was carefully removed using a pipette and transferred into a fresh microcentrifuge tube. Centrifugation and discarding the layer of fat was repeated until as much of the fat as possible had been removed from the sample and the supernatant was stored in a clean microcentrifuge tube. The protein concentration of the sample was then determined by the Bradford assay (section 4.3.3).

4.3.2 Protein Precipitation

Due to the heterogeneity of breast tissue and the nature of working with tissue rather than cell lines, it was thought it may be beneficial to precipitate the protein out of solution and re-suspend it in fresh buffer, thus cleaning and purifying the sample. This was achieved using the ProteoExtract® Protein Precipitation Kit (#539180, Calbiochem), which is suitable for use with 2D-PAGE. Before use, the Precipitation Agent was prepared using reagents supplied in the kit. To one bottle (29 ml) of Precipitant 1, 1.7 ml of Precipitant 2 was added as well as 1.7 ml Precipitant 3 and 1.7 ml Precipitant 4 and the solution was mixed. This was referred to as Precipitation Agent and was stored at -20 °C for 1 hour before use. The Wash Solution was also prepared in advance; 150 ml of high quality ethanol was added to the bottle (65 ml) of Wash Solution to reconstitute, mixed and stored at -20 °C. To precipitate the proteins, 800 µl of cold Precipitation Agent (-20 °C) was

added to 200 μ l of sample in a microcentrifuge tube. This was vortexed and incubated at -20 $^{\circ}$ C for 60 min. The sample was then centrifuged for 10 min at 10,000 \times g to pellet the protein. The supernatant was carefully removed. The pellet was washed by adding 500 μ l of cold (-20 $^{\circ}$ C) Wash Solution and vortexing briefly. The sample was then centrifuged for 2 min at 10,000 \times g to pellet the protein and the Wash Solution was carefully removed. The pellet was allowed to dry for 5 min at *RT* and subsequently re-suspended in Buffer A.

4.3.3 Protein Quantification

The protein concentration of the tissue extract was determined using the Bradford protein assay, which is compatible with reagents used in Antibody Microarray analysis. The Bradford Reagent (#B6916, Sigma Aldrich) consists of Brilliant Blue G in phosphoric acid and methanol. Brilliant Blue G is a dye which forms a complex with the proteins in the solution, and the formation of the complex causes a shift in the absorbance of the dye from 465 to 595 nm. The absorbance of the sample is therefore proportional to the amount of protein present in the sample. The linear protein concentration range is 0.1 to 1.4 mg/ml, where bovine serum albumin (BSA) was used as the standard. Eight BSA protein standards were prepared, diluted in Buffer A, in microcentrifuge tubes, ranging from concentrations of 0.1 to 1.4 mg/ml and 5 μ l of each was placed in separate wells in a 96-well plate. Tissue extracts of unknown protein concentration were also placed into separate wells in the 96-well plate at a volume of 5 μ l. Bradford Reagent was mixed gently and brought to room temperature and 250 μ l was added to each standard and each sample. The 96-well plate was then mixed for 30 s on the spectrophotometer (Multiscan MS plate reader, Labsystems) and incubated at room temperature for 5 min. Absorbance was subsequently measured at 595 nm. The protein concentration of each known protein standard was plotted against the

absorbance at 595 nm to produce a standard curve. The protein concentration of the tissue extracts was then calculated using the equation of the line.

4.3.4 Protein Labelling

This procedure was carried out in a darkened room, as the fluorescent dyes are sensitive to light. Proteins extracted from chemotherapy-sensitive tumour samples were labelled with Cy3 (#PA23001, GE Healthcare) fluorescent dye and proteins extracted from chemotherapy-resistant tumour samples were labelled with Cy5 (#PA25001, GE Healthcare) fluorescent dye. The extract was diluted to a protein concentration of 1 mg/ml in Buffer A. Labelling required the addition of 1 ml of tissue extract to the respective dye vials. The vial was then mixed by vortexing and incubated at *RT* for 30 min. During this incubation the vial was vortexed every 10 min.

Sigma Spin Columns (#S0185-8EA, Sigma Aldrich), supplied in the Antibody Microarray kit, were used to remove any unbound dye from the sample. The storage buffer contained in the column was removed by centrifugation for 2 min at 750 x *g* and discarded. A volume of 150 μ l of each of the labelled samples was passed through the columns by centrifugation for 4 min at 750 x *g* and the eluates were retained. The eluate obtained is the labelled protein sample, which is light-sensitive. The protein concentration was determined again using the Bradford protein assay (section 4.3.3), which was expected to be close to 1 mg/ml.

4.3.5 Determination of dye-to-protein molar ratio

The Dye to Protein Molar Ratio (D:P ratio) was determined by measuring the absorbance of the Cy3-labelled and Cy5-labelled protein samples at 552 nm and 650 nm respectively, with Buffer A as a blank. This was calculated as shown below, which was specified in the

Antibody Microarray kit, and as recommended, samples were only used if their D:P ratio was > 2

$$Cy3 \text{ concentration } (\mu\text{M}) = \frac{A552}{0.15} \times 10$$

$$Cy5 \text{ concentration } (\mu\text{M}) = \frac{A650}{0.25} \times 10$$

Y (mg/ml) = protein concentration after labelling with fluorescent dyes

$$\text{Protein concentration } (\mu\text{M}) = \frac{Y}{60,000} \times 1\,000\,000$$

$$D:P \text{ ratio} = \frac{Cy3 \text{ or } Cy5 \text{ concentration } (\mu\text{M})}{\text{Protein concentration of sample } (\mu\text{M})}$$

4.3.6 Antibody Incubation

This procedure was carried out in a darkened room. Equal amounts of protein (90 μg) from each sample were mixed with 5 ml Array Incubation Buffer (supplied in the Antibody Microarray kit), which was then placed in the first well of the quadriPERM Cell Culture Vessel supplied in the kit. The Antibody Microarray slide supplied in the kit was washed briefly in PBS before incubation with the samples in the well. The slide was incubated with the protein samples for 40 min on an orbital shaker at low speed, protected from the light. After this period, the slide was washed 3 times in Wash Buffer (supplied in the kit) for 5 min on an orbital shaker, protected from the light, followed by a 2 min wash with ultrapure distilled water under the same conditions. The slide was then allowed to air-dry for 30 min, protected from the light before scanning.

4.3.7 Scanning and Analysis

The antibody microarray slide containing antibody-protein complexes was scanned using a GenePix Personal 4100A Microarray scanner (Axon Instruments) with 532 and 635 nm

lasers. GenePix Pro software (Axon Instruments) was used to grid the antibody microarray slide and to apply protein names in the form of a list with their respective location on the array slide. Negative controls on the antibody microarray slide were flagged as negative and all antibody-protein spots were manually edited to ensure accurate analysis. This involved manually editing and re-positioning the circle generated by the computer which defines the area of the spot (feature), to ensure correct representation of the spot for analysis and to ensure background pixels are not included in the analysis of the spot. Acuity software (Axon Instruments) was used to identify differentially expressed proteins between the two samples (chemotherapy-sensitive and chemotherapy-resistant tumours). Normalization was carried out based on the Lowess method, and spot criteria were applied to only include spots which contained < 3% saturated pixels, spots with 'relatively' uniform intensity and background, those which were detectable above the background and those which were not flagged (negative controls), as a form of quality control. Log ratios were given based on the relative intensity of each Cy3 / Cy5 dye-labelled sample protein. Fold changes ≥ 1.8 were considered significant, and fold changes ≥ 1.5 were also recorded for each experiment, as supporting data. Experiments were considered successful when the percentage of 'substances matched', provided by the software during analysis, was ≥ 90 . This ensured that only slides of good quality were carried forward for data interpretation. The direction of fold change, showing an increase or decrease in the expression of a particular protein in chemotherapy resistant samples was given at the analysis stage. However, this was not expressed in the results, as dye-swap experiments were not performed to confirm this. The reason for this is that the advantage of performing a dye-swap experiment did not out-weigh the cost of repeating the experiment, and more importantly, sufficient clinical sample was not available for labelling with an additional

fluorescent dye. The value of differential expression has therefore been given, as a fold change in expression, but the direction of change was elucidated at the confirmation stage using western blotting and following further clinical validation using immunohistochemistry.

4.4 Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled to matrix-assisted laser desorption/ionization time-of-flight/time of flight (MALDI-TOF/TOF) mass spectrometry

All 2D PAGE / MS work was carried out using high-grade chemicals, which were personal to the user, in a PCR hood using polypropylene plastic-ware. nitrile gloves, hair protection and a lab coat were worn at all times, and extra care was taken throughout to reduce the risk of keratin contamination.

4.4.1 Protein Extraction

Clinical breast tumour samples selected for 2D-PAGE/MALDI TOF/TOF MS analysis are outlined in Figure 15. 2D extraction buffer (Appendix 2) was prepared immediately before it was required; it was made fresh for each use and could not be stored on ice due to the precipitation of urea out of solution. It is important to use fresh buffer each time as a major factor affecting the accuracy of this technique is carbamylation, which results from the breakdown of urea to cyanate which can react with amino groups of proteins (Garfin 2003) and affect IEF. Breast tumour tissue was removed from the minus 80 °C freezer and weighed. Work following this was carried out as quickly as possible in a Class II Tissue Culture Hood, using sterile technique to avoid contamination, and the tumour tissue was kept on ice to ensure minimal protein degradation. The tissue was then placed in a sterile

petri dish on ice, cut into small pieces using a disposable scalpel and washed by bathing and agitating gently in chilled PBS to remove any residual blood. The tissue was divided into pieces ≤ 0.1 g and transferred to microcentrifuge tubes containing 1 ml of 2D extraction buffer. Tissue > 0.1 g had previously been tested, but there was insufficient 2D extraction buffer to extract sufficient protein from this amount of tissue. The microcentrifuge tubes containing sample were sonicated for 15 min, with 5 min incubation on ice every 5 min to prevent excess heating of the sample. The tubes were then vortexed for 5 min (30 s on followed by 30 s off throughout to prevent excess foaming) and incubated for 16 hours at 4 °C with end-over-end rotation. Centrifugation for 20 min at 10,000 \times g at 4 °C was used to remove tissue debris and fat from the sample, followed by further centrifugation for 5 min under the same conditions, twice to ensure complete removal of fat. The supernatant was transferred into a clean chilled microcentrifuge tube and stored at minus 80 °C until required.

4.4.2 ReadyPrep 2-D Cleanup Kit

The ReadyPrep 2-D Cleanup Kit (#163-2130, Bio-Rad) was used to prepare the samples for isoelectric focusing (IEF) by concentrating the protein in the sample and removing components which interfere with IEF, such as lipids, salts and nucleic acids. The kit was able to clean up 200 μ l of sample per 1.5 ml microcentrifuge tube; the sample was therefore divided between microcentrifuge tubes before commencing. All reagents used were supplied in the kit, excluding the dH₂O (proteomic-grade). Wash Reagent 2 was stored at – 20 °C for 1 hour before use. Six hundred μ l of Precipitating Agent 1 was added to each tube, mixed thoroughly by vortexing and incubated on ice for 15 min. Six hundred μ l of Precipitating Agent 2 was also then added to each tube, and mixed by vortexing. The tubes were then centrifuged at maximum speed (\sim 12,000 \times g) for 5 min to form a pellet. The

tubes were removed promptly and the supernatant was carefully removed by pipetting. The tubes were re-centrifuged at maximum speed for 10 s and any remaining supernatant was carefully removed by pipetting. Forty μl of Wash Reagent 1 was then added to each tube on top of the pellet. The tubes were then centrifuged at maximum speed ($\sim 12,000 \times g$) for 5 min and the supernatant was removed by careful pipetting. Twenty five μl of dH_2O (proteomic-grade) was then added to the tubes containing the pellet, and the tubes were vortexed. One ml of pre-chilled ($-20\text{ }^\circ\text{C}$) Wash Reagent 2 was added to each tube along with 5 μl of Wash 2 Additive. The tubes were vortexed and incubated at $-20\text{ }^\circ\text{C}$ for 30 min. During the 30-min incubation, the tubes were vortexed every 10 min for 30 s. After this incubation, the tubes were centrifuged at maximum speed for 5 min to form a tight pellet. The supernatant was carefully removed by pipetting and centrifuged again to ensure all remaining liquid was removed. The pellet was air-dried for < 5 min until translucent. The pellets were then re-suspended in 200 μl of fresh 2D extraction buffer by pipetting and vortexing.

The samples, which had been cleaned up by the ReadyPrep 2-D Cleanup Kit were then quantified using the 2-D Quant Kit (#80-6483-56, GE Healthcare) (section 4.4.3) to ensure accurate loading of the sample (200 μg protein per gel) was achieved.

4.4.3 Protein Quantification

The 2-D Quant Kit (#80-6483-56, GE Healthcare) was used to determine the concentration of protein extracts which are to be used for IEF and 2D-PAGE. This quantification kit was chosen based upon reagent compatibility. The assay is based upon the binding of copper to the proteins, and unbound copper is measured by absorbance. The colour intensity is inversely proportional to the protein concentration. Working Colour Reagent was prepared by mixing 100 parts of Colour Reagent A to 1 part of Colour Reagent B, as stated in the kit

protocol. Six BSA protein standards were prepared, ranging from 0-50 μg by adding increasing volumes of 2 mg/ml BSA solution to microcentrifuge tubes. Samples to be quantified were added in duplicate to microcentrifuge tubes at volumes of 2 μl and 5 μl . To each tube, 500 μl Precipitant was added and the tube was vortexed and incubated for 3 min at *RT*. The same amount (500 μl) of Co-Precipitant was added to each tube and vortexed. The tubes were centrifuged at 10,000 \times g for 5 min to pellet the protein. The supernatant was carefully removed by pipetting. To each tube, 100 μl of Copper Solution and 400 μl of dH_2O were added and the tube was vortexed to re-suspend the protein. One ml of Working Colour Reagent was added to each tube, mixed by inversion and incubated at *RT* for 15 min. The absorbance of each sample and standard was read at 480 nm, on a Multiscan plate reader (Labsystems) with dH_2O as a blank in a 96 well plate. The protein concentration of the samples was then calculated from the equation of the line produced from the standard curve.

4.4.4 Isoelectric Focusing

Two hundred micrograms of protein sample was pipetted along the back of an 11 cm disposable Rehydration/Equilibration Tray (#165-4025, Bio-Rad), at a volume of 185 μl (Figure 11). This was performed in triplicate for each sample. ReadyStrips IPG strips (pH 4-7; 11 cm) (#163-2015, Bio-Rad) were rehydrated with the sample by placing them gel-side down into the sample in the rehydration tray, ensuring equal coverage of sample along the strip. The sample was left to absorb into the strip for 1 hour before 3 ml of mineral oil (#163-2129, Bio-Rad) was added into the well. This prevented evaporation of the sample during the rehydration process, where the IPG strip was incubated with the sample for 16 hours at *RT*.

The Protean[®] IEF tray (#165-4020, Bio-Rad) was cleaned and dried thoroughly prior to use. Paper electrode wicks (#165-4071, Bio-Rad) were placed over the electrodes in the tray and 8 µl of dH₂O was pipetted onto each. Rehydrated IPG strips were transferred to corresponding wells in the focusing tray gel-side down and covered with 3 ml of mineral oil. The Protean[®] IEF tray was transferred to the Protean[®] IEF Cell (#165-4001) and IEF was performed using a 3-step program consisting of 20 min at 250 V (linear); 150 min at 8000 V (linear); 20 000 V-hours at 8000 V (rapid), which lasted 5.5 hours in total. The focused IPG strips were transferred into a clean Rehydration/Equilibration tray gel-side up and stored at minus 80 °C until required for SDS-PAGE (for no longer than 1 month).

4.4.5 Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

The IPG strips were defrosted and equilibrated in preparation for SDS-PAGE. Stock equilibration buffer (EB) was used to prepare EB-1 and EB2 (Appendix 2) which contained DTT and IAA respectively. These agents ensure the effective separation of proteins in the 2nd dimension by ensuring they are in the right conditions; they are saturated with SDS which gives the proteins a negative charge ensuring their migration to the anode during separation; DTT and IAA prevent reformation of disulphide bonds by reduction and alkylation. The IPG strips were transferred into a clean rehydration tray and incubated with 4 ml of EB-1 for 10 min on an orbital shaker. This was discarded and the IPG strips were then incubated with 4 ml of EB-2 in the same way but also covered with foil as IAA is sensitive to light. During this period, 1% overlay agarose solution (Appendix 2) was repeatedly heated (on a medium heat) to melt it and maintain it in a liquid state. Criterion[™] pre-cast gels (8-16% Tris-HCl polyacrylamide gel; 11 cm) (#345-0105, Bio-Rad) were prepared by washing wells with dH₂O and blotted dry immediately before required with filter paper. IPG strips were washed briefly in Tris-glycine running buffer (#161-0772, Bio-

Rad) before being blotted and placed at the top of the gel. Molten agarose was transferred into the main well of the gels and the IPG strips were pushed into it, ensuring that no air-bubbles were present. The agarose was allowed to set for 5 min. Both chambers of the tank were filled with Tris-glycine running buffer and Precision Plus Protein Standards Dual Colour Marker (#161-0374, Bio-Rad) was added to its designated well. The gel was run for 65 min at a constant voltage of 200 V, 500 mA and 300 W.

4.4.6 Staining of Proteins

After running, gels were removed from their casing and washed 3 times for 5 min in dH₂O in a nalgene staining pot on an orbital shaker. Bio-safe Coomassie Stain (#161-0787, Bio-Rad) was used to stain the gels for 1 hour on an orbital shaker. The gels were de-stained for 16 hours on an orbital shaker at *RT* in dH₂O. Following this, the gels were washed again 3 x 5 min before scanning with a GS800 calibrated densitometer (Bio-Rad) and imaging with Quantity One (Bio-Rad) software.

4.4.7 PDQuest Analysis Software

PDQuest Analysis Software is a complex tool which is used to detect and analyse protein spots on and between gels and identify differentially expressed protein spots between groups of gels (e.g 'test' and 'control'). Gels were 'test' (chemotherapy-resistant sample) and 'control' (chemotherapy-sensitive samples), in triplicate. Spot detection parameters were set by identifying faint, weak and clusters of protein spots. This software automatically detected and matched spots, however all spots were edited and defined manually and re-matched to ensure spots had been identified and matched correctly across all gels, which took 4-5 days for an experiment consisting of 6 gels. Spots which contained more than one protein, were part of a cluster or could not be matched with confidence were

not included. Matched spots were then normalised using ‘total quantity in valid spot’ parameters. A dataset was then automatically created and the criteria for differentially expressed protein spots was applied; only spots with a fold change ≥ 2 , of 95% significance were identified. For clusters of spots, or large spots where it was difficult to determine how many spots were present, a 3-D viewer tool was used. Analysis, which identified and quantified differentially expressed protein spots, used Boolean quantification and the Students *t*-test. Differentially expressed protein spots were highlighted on gels and histograms were provided to demonstrate the difference in protein expression between chemotherapy-sensitive and chemotherapy-resistant samples.

4.4.8 Spot excision

Protein spots to be excised were identified and excised using a disposable scalpel (which was washed in ddH₂O between uses) on ProteoWorks Plus Gel Cutting Sheets (#165-7057, Bio-Rad). Spots were excised from 2-3 respective gels of the same sample type only (e.g chemotherapy-resistant gels) and transferred into 0.5 ml Protein LoBind microcentrifuge tubes (#022431064, Eppendorf). Spots were only excised when they contained a single protein, were not part of a streak or a cluster and were not adjoining another protein spot.

4.4.9 In-gel digestion

This procedure involved the tryptic digestion of proteins into peptides within the gel in order to release them from the gel. It is essential to minimise sample loss and contamination by following a basic protocol which does not include an excessive number of steps.

Ammonium bicarbonate 100 mM stock solution was prepared by dissolving 0.395 g in 50 ml ddH₂O. From this, 25 mM ammonium bicarbonate (50% ACN) and 25 mM ammonium bicarbonate (aq) were prepared.

4.4.9.1 De-stain

Gel pieces were de-stained by incubating with 100 μ l of 25 mM ammonium bicarbonate (50% ACN) for 20 min at room temperature. The supernatant was removed and this step was repeated. Following this, gel pieces were washed by incubating with 100 μ l of acetonitrile (ACN) for 5 min at room temperature. Gel pieces were then dried by vacuum centrifugation for 20 min.

4.4.9.2 Digest

Trypsin Gold (#V5280, Promega) was reconstituted with 50 mM acetic acid to a concentration of 0.1 mg/ml (stock). Twenty microlitres (2 μ g) of stock trypsin was diluted with 80 microlitres of 25 mM ammonium bicarbonate (aq) [0.02 μ g/ μ l], of which 10 μ l was added to each eppendorf containing gel pieces. Gels were given 5-10 min to re-hydrate, after which they were covered with 5-15 μ l of 25 mM ammonium bicarbonate (aq), whilst keeping the volume as low as possible. The gel pieces were then incubated for 16 hours at 37 °C, during which the proteins were digested into peptides.

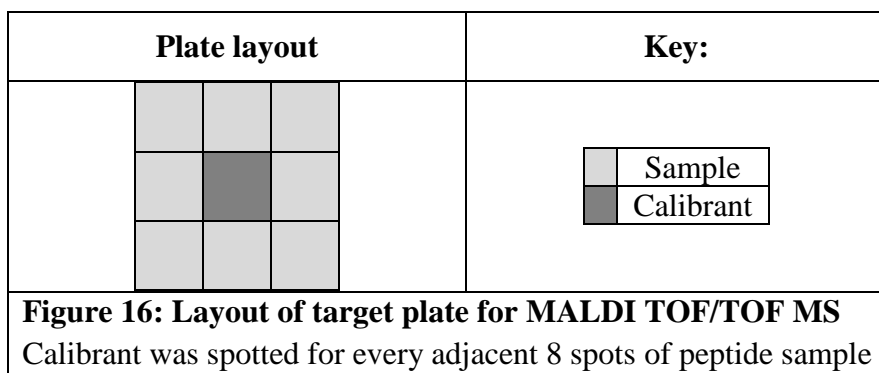
4.4.10 Preparation of MALDI matrix and plate-spotting

The MTP384 target plate polished steel TF (#209520, Bruker Daltonics) was cleaned by wiping with 2-propanol and ddH₂O and sonicating in 2-propanol followed by a 70% ddH₂O: 30% ACN and 0.1% Trifluoroacetic acid (TFA) solution, as recommended by Bruker. The matrix consisted of a freshly-prepared 5 mg/ml solution of 4-hydroxy- α -cyanocinnamic acid (CHCA) (#70990, Fluka) in 50% ACN and 0.1% TFA (aq) (v/v). One microlitre of each peptide sample was spotted directly onto the target plate, immediately followed by 1 μ l of matrix solution. Calibrant consisting of six known peptides (section

4.4.11) was also spotted onto designated locations on the plate (1 μ l) followed by 1 μ l of matrix solution.

4.4.11 MALDI-TOF/TOF Mass Spectrometry

The MALDI-TOF/TOF MS which was used was an Ultraflex III (Bruker Daltonics). Mass spectra were obtained in reflectron mode from positive ions generated by a Nd:YAG smartbeam laser. Data were acquired using FlexControl (version 3.3, Bruker Daltonics) in AutoXecute mode (comprising AutoXMethods and AutoXSequences) to enable automation of MS calibration and sample data acquisition. Mass Spectra were acquired over a mass range of m/z 800-4000. Final mass spectra were externally calibrated using an adjacent spot containing 6 known peptides (des-Arg¹-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu¹-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929). Calibrant was spotted for every 8 sample spots (Figure 16)



For acquisition of MS spectra, 50 laser shots were fired at 16 random positions to yield a sum of 800 shots. For acquisition of MS/MS spectra, 500 shots were used to check the precursor ion and 2500 laser shots were used for fragment ions. Monoisotopic masses were obtained using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a signal-to-noise threshold of 2. From the PMF generated for each spot, the 10

highest peaks, with a signal-to-noise threshold >30 , were automatically selected for MS/MS fragmentation. Fuzzy control was used to optimise the laser power for MS/MS acquisitions. Fragmentation was performed in LIFT mode without addition of collision gas. The default calibration method was used for MS/MS spectra, which involved base-line subtraction and smoothing (Savitsky-Golay, width $0.15 m/z$, cycles 4). Monoisotopic peaks were detected using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a signal-to-noise threshold ≥ 6 . Flex Analysis software (version 3.3, Bruker Daltonics) was used to perform the spectral processing and peak list generation for both MS and MS/MS spectra. Processed peak lists were submitted to Mascot (version 2.1, Matrix Science Ltd) for database searching (IPI Human) via the ProteinScope interface (version 2.3, Bruker Daltonics). Search criteria were specified, which included; enzyme, trypsin; fixed modifications, carbamidomethyl (C); variable modifications, oxidation (M); peptide tolerance, 250 ppm; MS/MS tolerance, 0.5 Da; Instrument, MALDI-TOF/TOF.

4.5 Ingenuity Pathway Analysis

Data generated by antibody microarray analysis and 2D-PAGE MALDI-TOF/TOF MS was analysed using IPA (Ingenuity Systems, www.ingenuity.com). Each set of data, containing a list of gene symbols, which had been checked against the IPI and NCBI databases, was uploaded into IPA software online.

The Ingenuity Knowledge Base is the core facility and repository behind IPA, holding all the biological and chemical information, functional annotations and modelled relationships for genes, proteins, complexes, disease states, cells, tissues etc in a well-structured and accessible manner. The Ingenuity Knowledge Base is a comprehensive database containing manually reviewed, accurate information. Within the Ingenuity Knowledge Base there are four types of information, including both experimental- and literature-based

sources, which is all manually reviewed: (1) Ingenuity® Expert Findings, which contains experimentally-demonstrated information; (2) Ingenuity® ExpertAssist Findings, from recently published journal abstracts; (3) Ingenuity® Expert Knowledge, containing signalling and metabolic pathway information, which is curated from a team of Ingenuity experts ; (4) Ingenuity® Supported Third Party Information, which is selected from a range of specific sources and databases including Entrez Gene, Gene Ontology and RefSeq.

For network generation, each gene was mapped to the corresponding gene within the Ingenuity Knowledge Base, and an ‘annotated dataset’ was generated. Genes which were successfully mapped into the Ingenuity Knowledge Base were called ‘network eligible’ molecules, and were subsequently overlaid onto a global molecular network developed from the information contained within the Ingenuity Knowledge Base. During analysis of data, networks of ‘network eligible’ molecules were then algorithmically generated based on their connectivity. The general settings allowed the maximum number of ‘molecules per network’ and ‘networks per analysis’ to be included, to highlight direct relationships between human molecules which had been reported in both tissues and cell lines.

Canonical pathway analysis of the dataset involved the identification of pathways within the IPA library of canonical pathways that were most significant to the dataset. All molecules mapped within the dataset were considered for canonical pathway analysis. The significance of the association between the dataset and the canonical pathway was measured by two factors; 1) A ratio of the number of molecules within the dataset that can be mapped into a pathway, divided by the total number of molecules involved in that pathway, 2) Fisher’s exact test was performed to determine the probability that the association between the dataset and the canonical pathway identified had occurred by chance, which was displayed as a p-value.

4.6 Western Blotting

4.6.1 Protein Extraction

4.6.1.1 From cultured cells

Western blot (WB) extraction buffer was prepared (Appendix 2) and to each ml, 10 μ l each of Phosphatase Inhibitor Cocktail 1 (#P2850, Sigma Aldrich), Phosphatase Inhibitor Cocktail 2 (#P5726, Sigma Aldrich) and Protease Inhibitor (#80-6501-23, Amersham Biosciences) was added as well as 50 μ l 2-Mercaptoethanol (#M-7522, Sigma Aldrich). Cells were grown to 80% confluence, pelleted by centrifugation at 1600 rpm for 3 min and re-suspended in PBS 3 times to wash and ensure removal of the media from the cells. The cell pellet was then re-suspended in 250 μ l of WB extraction buffer containing inhibitors and vortexed for 5 min. WB extraction buffer contained Tris-HCl buffer; glycerol, which helped to weigh down the protein sample enabling it to run efficiently; SDS, which denatured proteins to polypeptides and applied a net negative charge to the proteins; bromophenol blue, which is the standard dye used for visualisation of the protein samples as they run down the gel; beta-mercaptoethanol, which ensured proteins were fully denatured. It was then placed on an end-over-end rotator for 16 hours at 4 °C, to aim for maximum extraction of protein, and then centrifuged at 10,000 x g for 15 min at 4 °C to remove cell debris. The supernatant was transferred into a fresh chilled microcentrifuge tube and stored at minus 80 °C.

4.6.1.2 From breast tumour tissue

Clinical breast tumour samples selected for western blotting are shown in Figure 15. WB extraction buffer was prepared (Appendix 2), containing 10 μ l each of Phosphatase Inhibitor Cocktail 1 (#P2850, Sigma Aldrich), Phosphatase Inhibitor Cocktail 2 (#P5726,

Sigma Aldrich) and Protease Inhibitor (#80-6501-23, Amersham Biosciences) and 50 μ l 2-Mercaptoethanol (#M-7522, Sigma Aldrich) per ml.

Breast tumour tissue was removed from the minus 80 °C freezer and weighed. Work following this was carried out as quickly as possible in a Class II Tissue Culture Hood, using sterile technique to avoid contamination, and the tumour tissue was kept on ice to ensure minimal protein degradation. The tissue was then placed in a sterile petri dish on ice, cut into small pieces using a disposable scalpel and washed by bathing and agitating gently in chilled PBS to remove any residual blood. The tumour was then transferred into 800 μ l WB extraction buffer containing inhibitors in a universal tube and homogenised on ice using a TissueRuptor (#9001273, Qiagen) with a disposable probe (#990890, Qiagen). The tissue was homogenised for 3 x 20 s, with 20-second gaps between each to prevent excess heating. After homogenisation the tissue extract was transferred into a chilled microcentrifuge tube and centrifuged at 13,000 \times g for 5 min at 4 °C. The supernatant was transferred into a clean chilled microcentrifuge tube and the pellet and layer of fat (due to the high proportion of adipose tissue present in breast tissue) was discarded. This was repeated 3-5 times until there was no fat remaining and the tissue extract was stored at minus 80 °C.

4.6.2 Protein Quantification

The *RCDC* (Reducing agent Compatible, Detergent Compatible) Protein Quantification kit (#500-0119 to -0122, Bio-Rad) was used, which is a colourimetric assay based upon the Lowry Assay was used due to its compatibility with components used in this procedure. Five BSA protein standards were prepared ranging from 0.25 to 1.5 mg/ml, as recommended in the assay protocol, by diluting a 2 mg/ml stock of BSA with dH₂O in microcentrifuge tubes. Sample to be quantified was diluted to 1:2, 1:5 and 1:10 dilutions to

ensure its concentration fell within the assay range. *RC* Reagent I was added to each tube at a volume of 125 μ l, vortexed and incubated at *RT* for 1 min. *RC* Reagent II was added to each tube, at a volume of 125 μ l, vortexed and centrifuged at 15,000 \times g for 5 min. This precipitated the protein out of solution so a pellet remained in the tube. The supernatant was discarded and the tube was inverted on absorbent paper to ensure maximum removal of the liquid. Working Reagent A was prepared, which included 20 μ l of Reagent S for every 1 ml of Reagent A. This was added to each tube at a volume of 127 μ l and vortexed to re-suspend the protein. One ml of Reagent B was added to each tube and incubated at *RT* for 15 min. Standards and samples to be quantified were transferred to a 96-well plate where their absorbances were read at 690 nm using a Multiscan plate reader (Labsystems).

4.6.3 One-dimensional gel electrophoresis

Protein extracts were diluted with WB extraction buffer (Appendix 2) containing 5% 2-Mercaptoethanol (#M-7522, Sigma Aldrich) to achieve 25 μ l of 20 μ g of protein, depending on the optimised conditions of the antibody to be probed for. The proteins in the extracts were then denatured by heating at 95 $^{\circ}$ C in a thermocycler for 5 min. They were then placed on ice to prevent reversal of protein denaturation, vortexed and centrifuged at maximum speed (\sim 12,000 \times g) for 30 s. Twenty μ l of extract was then loaded into appropriate wells in a 12% Precise Protein Gel (#25222, Thermo Scientific) with Tris-HEPES-SDS running buffer (#28368), alongside 10 μ l of Precision Plus Protein WesternC Standard (#161-0376, Bio-Rad), as a marker of molecular weight. The gel was run at a constant voltage of 140V for 40 min.

4.6.4 Transfer of proteins onto nitrocellulose membrane

Proteins that had been separated by molecular weight were transferred onto a nitrocellulose membrane. This was achieved using the iBlot dry transfer system (Invitrogen). When using the iBlot, 'iBlot gel transfer stacks, nitrocellulose' (#IB3010-01, Invitrogen) were used. Firstly, the 'bottom' disposable transfer stacks containing membranes were placed in the machine. Following this, gels were placed on top of the membrane in the required orientation. Filter paper soaked with dH₂O was placed on top of the gels and air bubbles were removed using a roller. The 'top' disposable pack containing the anode was placed on top of the membrane followed by a sponge containing an electrode. The standard transfer, as recommended by the manufacturer was used, which ran for 7 min, and transferred the proteins from the gel onto the nitrocellulose membrane.

4.6.5 Blocking of binding sites on the membrane

Once the proteins had been transferred onto the membrane, the free binding sites on the membrane were blocked. This was achieved by incubating the membrane with 'blocking solution' (5% low-fat milk powder (Marvel), diluted in TBS-Tween20) in a Nalgene staining box on an orbital shaker for 1 hour at *RT* or 16 hours at 4 °C. This was necessary to prevent unwanted binding of antibodies to the membrane when probing for a specific protein. This may occur due to the high affinity membranes have for proteins.

4.6.6 Immunoblotting

The primary antibody to the protein of interest was optimised and diluted to its optimum concentration in blocking solution. It was incubated with the membrane for 2 hours at *RT* on an orbital shaker. Following this, the membrane was washed 3 times with TBS-Tween20 (5 min per wash) on an orbital shaker, to remove any unbound antibody. The membrane

was then incubated with a HRP-conjugated secondary antibody to the animal the primary antibody was raised in (Table 5). This was diluted to its optimum concentration in blocking solution and incubated with the membrane for 1 hour at *RT* on an orbital shaker. For visualisation of the Precision Plus Protein WesternC Standard molecular weight marker, 1 μ l of Precision Protein StrepTactin-HRP conjugate (#161-0381, Bio-Rad) was also added to the blocking solution containing secondary antibody. Three washes of 5 min each with TBS-Tween20 were carried out on an orbital shaker.

4.6.7 Loading controls

To test for accurate loading of proteins into the gel, thus allowing fair comparisons to be made between samples, proteins which should be present in all cells at equal concentrations are probed for. These are known as ‘housekeeping proteins’; here β -actin was used as the loading control (Table 5).

4.6.8 Detection of proteins

In order to detect the proteins, the membrane was incubated with equal amounts Supersignal West Pico Stable Peroxide Solution and Supersignal West Pico Luminol Enhancer Solution from the Supersignal West Pico Chemiluminescent Substrate Kit (#34078, Thermo Scientific) for 5 min in the dark with frequent gentle manual agitation. The membrane was then placed between transparent plastic sheets and placed in an intensifying cassette with CL-XPosure Film (#34090, Thermo Scientific) and developed using 250 ml each of GBX Developer (#P7042, Sigma Aldrich), by gentle manual agitation in a developer tray until bands appeared, followed by 30 s incubation in 250 ml 5% Acetic Acid and then 250 ml GBX Fixer (#P7167), with gentle manual agitation in a plastic tray. The developed films were then allowed to air-dry before scanning and densitometry.

4.6.9 Densitometry

Densitometry was used to quantify the density of bands on films, representing expression of a particular protein in the chosen protein extract. The film was scanned using a GS800 Calibrated Densitometer (Bio-Rad) and Quantity One software (Bio-Rad) was used to normalise the protein of interest against the loading control and quantify relative densities of bands in order to calculate fold changes between samples.

Table 5: Table of primary antibodies used for western blotting

The table lists the primary antibodies used for analysis of protein expression by western blotting. The secondary antibody used at all times was goat anti-rabbit IgG-HRP (#SC-2030, Santa Cruz), which was applied at 1:1000 in 5% milk for 1 hour at room temperature,

Antibodies	Concentration and blocking agent	Incubation period	Details
Beta-actin	1:1000 in 5% milk	2 hours	Rabbit polyclonal (#ab8227, Abcam)
14-3-3 (beta, eta, tau and sigma)	1:1000 in 5% milk	2 hours	Rabbit polyclonal (#ab9063, Abcam)
BCL2L1 (Bcl-xL)	1:400 in 5% milk	16 hours	Rabbit monoclonal (#ab32370, Abcam)
BID	1:200 in 5% milk	16 hours	Rabbit monoclonal (#ab32060, Abcam)
14-3-3 epsilon	1:2500 in 5% milk	2 hours	Rabbit polyclonal (#ab43057, Abcam)
14-3-3 zeta	1:1000 in 5% milk	2 hours	Rabbit polyclonal (#ab51129, Abcam)

4.7 Immunohistochemistry

Formalin-fixed, paraffin-embedded pre-treatment core biopsy tissue samples were retrieved for the 36 samples (from 35 patients) from the previously described sample cohort (Garimella 2007) (section 4.2) (REC 03/00/038). The sections were cut to a thickness of 4 μm , mounted onto Superfrost Plus microscope slides (#00594, Menzel-Glaser) and incubated at 37 °C overnight. Each immunohistochemical staining experiment contained a negative control, from which primary antibody was omitted.

4.7.1 De-waxing and rehydration

Tissue sections were de-waxed by incubating in warm (~50 °C) HistoClear II (#HS-200, National Diagnostics) for 10 min, followed by two 10 s incubations (with gentle agitation) in separate solutions of HistoClear II (#HS-200, National Diagnostics). Sections were rehydrated by incubating (with gentle agitation) for 10 s in 100% ethanol. This was repeated three times, using 3 separate ethanol solutions. Sections were then rinsed in running tap water for 1 min.

4.7.2 Blocking of endogenous peroxidase

The endogenous peroxidase of red blood cells was blocked by incubating with methanol containing 30% hydrogen peroxide for 20 min.

4.7.3 Antigenic site retrieval

Antigenic site retrieval was achieved by boiling slides in a stainless steel pressure cooker (Prestige) containing 1500 ml of 1:100 Antigen Unmasking Solution (#H-3300, Vector Laboratories) at full pressure (103 kPa) for 3 min. Slides were then transferred into 1 x Tris Buffered Saline (TBS).

4.7.4 Blocking of non-specific binding sites within sections

Slides were assembled onto a sequenza system (Shandon, Basingstoke, UK) for immunohistochemical staining, using TBS-washes to ensure accurate assembly. Non-specific binding sites within sections were blocked by incubating slides with 100 µl of 1 x casein (#SP-5020) in TBS for 10 min, where the StreptABCComplex/HRP Duet Kit (#K0492, DakoCytomation Ltd) was used downstream. Where the R.T.U VECTASTAIN Universal *Quick* Kit (#PK-7800, Vector Laboratores Ltd) was used downstream, blocking

was achieved with 100 µl of pre-diluted normal horse serum, provided in the kit, for 10 min. Slides were then washed twice for 5 min in TBS.

4.7.5 Incubation with primary antibody

Depending upon the visualisation kit used downstream, primary antibody was diluted in either:

- 0.2 x casein (#SP-5020, Vector Laboratories) in TBS (StreptABComplex/HRP Duet Kit (#K-0492, DakoCytomation Ltd))
- 1.5% normal horse serum (provided) in TBS (R.T.U VECTASTAIN Universal *Quick* Kit (#PK-7800, Vector Laboratores Ltd))

Antibody details and dilutions used are listed in Table 6. One hundred microlitres of diluted antibody was incubated with the tissue sections for 2 hours at room temperate. The negative control was incubated with 100 µl of either 0.2 x casein or 1.5% normal horse serum.

4.7.6 Antibody detection

Antibody detection was achieved using one of two kits (the first has now been discontinued).

4.7.6.1 StreptABComplex/HRP Duet Kit

Antibody detection was achieved using the StreptABComplex/HRP Duet Kit (#K-0492, DakoCytomation Ltd) kit following manufacturers' instructions. Reagent C (biotinylated goat anti-mouse/rabbit secondary antibody) was diluted 1:100 in TBS, of which 100 µl was applied to each slide and incubated for 30 min. The slides were then washed in TBS for 5 min. Reagents A (streptavidin) and B (biotinylated peroxidase) were diluted together with TBS, each at 1:100, of which 100 µl was incubated with each slide for 30 min. The slides

were rinsed again for 5 min in TBS before the slides were dismantled from the sequenza system into fresh TBS.

4.7.6.2 R.T.U VECTASTAIN Universal *Quick* Kit

The R.T.U VECTASTAIN Universal Quick Kit (#PK-7800, Vector Laboratories Ltd) procedure was carried out according to manufacturers' instructions, however incubation times described were found to be sub-optimal. Optimisation of the protocol resulted in doubling the incubation times described by the manufacturer. One hundred microlitres of pre-diluted biotinylated pan-specific universal secondary antibody was incubated with each slide for 20 min. The slides were then washed in TBS for 5 min. Slides were then incubated with pre-prepared streptavidin/peroxidase complex reagent; 100 µl for 10 min. Slides were then washed again for 5 min in TBS, dismantled from the sequenza system and transferred into fresh TBS.

4.7.7 Antibody Visualisation

Antibody visualisation was achieved using 0.02% diaminobenzidine (DAB) in TBS containing 0.125% hydrogen peroxide (30% w/w solution). Slides were incubated in the solution until brown staining of the sections could be seen under a light microscope; this did not exceed 30 min, due to precipitation of DAB.

4.7.8 Enhance, counterstain and differentiate

The contrast of the staining was enhanced by incubating in 0.5% copper sulphate in 0.9% saline for 5 min. Sections were then counterstained using filtered Harris' Haematoxylin (#HHS32, Sigma Aldrich), by incubating (with gentle agitation) for 20 s. Excess haematoxylin was removed by washing slides in running tap water. The counterstain was

differentiated by incubating (with gentle agitation) for 10 s in acid alcohol (70% alcohol, 1% HCl (conc)), followed by washing slides in running tap water.

4.7.9 Rehydration, clearing and mounting

Tissue sections were rehydrated by taking slides through 3 solutions of 100% ethanol, with gentle agitation for 10 s in each. Sections were cleared in HistoClear II (#HS-200, National Diagnostics), by taking slides through 3 solutions with gentle agitation for 10 s in each. Slides were mounted onto cover-slips using Histomount (#HS-103, National Diagnostics) and allowed to dry overnight.

4.7.10 Antibodies used for Immunostaining

Table 6: Primary antibodies used for immunohistochemical staining

Antibody	Details	Dilution	Detection kit
14-3-3 tau	Mouse monoclonal (#T5942, Sigma Aldrich)	1:50	Dakocytomation
tBID	Rabbit polyclonal (#ab10640, Abcam)	1:50	Vector Laboratories
Bcl-xL	Mouse monoclonal (#B9429, Sigma Aldrich)	1:35	Vector Laboratories

4.7.11 Histological scoring of immunostained tissue sections

Scoring systems were developed after observation of all slides across the sample series, and were unique for each staining localisation. Scoring was performed by two observers, independently, after consultation with a consultant histopathologist. In the event of disagreement between the two observers, a third observer was introduced, so a majority-based score could be assigned.

Table 7: Scoring method for cytoplasmic immunostaining

% Coverage within invasive tumour	>50%	>50%	>50%	>50%
Symbol	-	-/+	+	++
Intensity	Weak	Weak-moderate	Moderate	Strong
Score	Negative		Positive	

Table 8: Scoring method for immunostaining of the nuclear membrane

Staining in $\geq 20\%$ of tumour cells	Staining in $< 20\%$ of tumour cells
Positive	Negative

Table 9: Scoring method for nuclear staining

Intensity	Weak (0 points)	Moderate (1 point)	Strong (2 points)
Nuclear staining coverage in $\geq 50\%$ tumour	Patchy (0 points)		Solid (1 point)
Points	Score		
0	Negative		
1	Negative		
2	Positive		
3	Positive		

4.7.12 Fishers Exact test for determination of statistical significance

In order to assess the statistical significance between histological scores and chemo-resistance in breast cancer, two-tailed Fishers exact tests were performed from 2×2 contingency tables, which generated exact probability (P) values. Where P values were ≤ 0.05 , the null hypothesis was rejected and the association between histological scores and chemo-resistance was deemed significant. Calculations were performed using GraphPad

Software Inc (USA) (at <http://www.graphpad.com/quickcalcs/contingency1.cfm>), with data entered as shown in Table 10.

Table 10: A 2 x 2 contingency table for the Fishers exact test

A 2 x2 contingency table for calculation of statistical significance using the Fishers exact test. Values were entered (x) for numbers of chemotherapy-sensitive and chemotherapy-resistant negative and positive scores.

	<i>Negative</i>	<i>Positive</i>
<i>Chemotherapy-sensitive</i>	x	x
<i>Chemotherapy-resistant</i>	x	x

4.7.13 Determination of inter-observer variability using the Kappa statistic

The Kappa statistic was used to determine the variability between observers for histological scoring, which not only determines the agreement between observers but also takes into account agreements that could have occurred by chance.

Table 11: A 2 x 2 contingency table for the Kappa statistical test

A 2 x 2 contingency table for determination of inter-observer variability using the kappa statistic. Agreement between observers is seen at a and d , and disagreement is seen at b and c . If there is 100 % agreement, values at b and c would be equal to 0.

		<i>Observer 1</i>		
		<u>Positive</u>	<u>Negative</u>	<u>Total</u>
<i>Observer 2</i>	<u>Positive</u>	a	b	m_1
	<u>Negative</u>	c	d	m_0
	<u>Total</u>	n_1	n_0	n

In order to calculate the kappa coefficient, two parameters were initially determined:

Observed agreement (P_o): This was the percentage agreement seen between the two observers

$$P_o = \frac{a + d}{n}$$

Chance agreement (P_c): This was calculated to determine how much agreement would have occurred by chance alone

$$P_c = \left[\left(\frac{n_1}{n} \right) \times \left(\frac{m_1}{n} \right) \right] + \left[\left(\frac{n_0}{n} \right) \times \left(\frac{m_0}{n} \right) \right]$$

The Kappa coefficient (k) was then determined using the equation:

$$k = \frac{P_o - P_c}{1 - P_c}$$

The calculated value for the Kappa coefficient was then interpreted using the criteria listed in Table 12.

Table 12: Criteria for interpretation of the kappa coefficient

Criteria for determination of inter-observer variability in histological scoring using the kappa coefficient

Kappa coefficient	Inter-observer variability
0	Agreement occurred by chance
0.01 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

CHAPTER 5:

CLINICAL TUMOUR SAMPLE COLLECTION

Chapter Aim:

The collection of clinical breast tumour tissue samples from locally advanced breast cancer patients treated with neoadjuvant chemotherapy, for proteomic analysis

Chapter 5. Clinical tumour sample collection

5.1 Introduction

5.1.1 Ethical approval

In order for clinical research studies to be carried out within the NHS, ethical approval must be obtained from a Research Ethics Committee (REC) and the NHS Trust Department for Research Governance. Before applications can be made, it is necessary to produce a project document, which includes all information regarding the background of the project, its aims and objectives, study design, protocol and methods, as well as details of those involved in the study and any sponsorship, along with any necessary documentation such as Patient Information Sheets and Patient Consent Forms. Once the project document is complete, the application process can commence, which is carried out via the Integrated Research Application System (IRAS). IRAS amalgamates the process of determining which relevant bodies approval is required from, for the specific project, and generates all forms required from the 'full set of project data'. Once the project has been submitted to the relevant bodies, all details and information about the study are considered. The REC then provides their opinion regarding the project, which may be 'favourable with/without conditions', 'provisional', where more information may be required, or 'unfavourable'. Once approval from both the Trust and REC has been gained, the study may commence.

5.1.2 Clinical samples for proteomic analysis

The use of clinical samples for identification of protein biomarkers has many advantages over the use of cell lines (section 6.1.1). However, there is a strong requirement for pre-analytical standardisation encompassing standardised specimen acquisition, handling and

storage, as well as sample preparation (Apweiler, Aslanidis et al. 2009). This presents a major challenge, but is a requirement for reproducible and accurate results, as variability in these processes can alter molecular composition and therefore experimental data (Moore, Kelly et al. 2011). It is also important that studies using clinical samples report details of the handling of the samples to improve consistency and reproducibility. The Biospecimen Reporting for Improved Study Quality (BRISQ) recommendations highlight criteria which should be reported and applied to any study using tissue samples (Moore, Kelly et al. 2011). Details include, and are not limited to, biospecimen type, anatomical site, disease state of the patient, relevant clinical information, sample collection mechanism, stabilisation mechanism, type of long-term preservation, storage temperature and shipping temperature (Moore, Kelly et al. 2011).

The quality of data generated from clinical samples is heavily reliant upon sample conditions immediately following excision from the patient. It is important to recognise tissue excised from patients is still alive. It will therefore be suddenly exposed to *ex vivo* stress, and will begin to adapt to its new conditions; hypoxia, temperature change, ischaemia, accumulation of cellular waste, lack of electrolytes and lack of vascular perfusion (Espina, Mueller et al. 2009). These conditions will induce changes in cellular signal-transduction pathways related to the cellular stress response and wound repair, thus distorting the tumour proteome. The level of distortion is dependent upon the time-delay between tissue excision and stabilization, either by immersion in fixative or snap-freezing in liquid nitrogen (Espina, Mueller et al. 2009). It is therefore paramount that excised tissue is stabilized as soon as possible (< 20 minutes), to regulate the process of clinical sample collection and to avoid the introduction of pre-analytical variables (Apweiler, Aslanidis et al. 2009; Espina, Mueller et al. 2009).

5.2 Methodology

5.2.1 Ethical approval for this study

The study upon which this thesis is based, entitled ‘the identification of biomarkers associated with therapy response in breast cancer’, was approved by the South Humber Research Ethics Committee (ref 07/Q1105/43) in 2007. Letters confirming ethical approval from the REC can be seen in Appendix 3. The study recruited patients with locally advanced breast cancer who had received standard neoadjuvant chemotherapy prior to surgical resection of the remaining tumour, at Castle Hill Hospital (Hull and East Yorkshire Hospitals NHS Trust) from 2007 onwards. Approval was granted to allow:

- Locally advanced patients receiving standard neoadjuvant chemotherapy to be recruited to the study, after being given a Patient Information Sheet (Appendix 4), discussion with a clinician, and signed recorded consent
- A sample of surplus residual tumour, to be taken during routine surgery for resection of residual tumour
- Storage of the tumour sample for future research
- Access to patients’ medical notes for information relevant to the study
- Access to the diagnostic core biopsy specimen, stored in the Histopathology archive

5.2.2 Study design

An outline of the study is depicted in Figure 17. Initially, patients who matched the requirements of the study were identified by the clinicians in the Breast Unit at Castle Hill Hospital. They were informed of the research study, and if patients wished to participate they were provided with an information sheet and a signed record of consent was taken. This may have been towards the end of the chemotherapy treatment, prior to surgery.

During surgical resection of the residual tumour, breast surgeons took a small sample of macroscopic tumour, which was immediately stabilised by snap-freezing in liquid nitrogen. Tumour samples provided by the surgeon varied in size from 2 mm³ to 2 cm³, and the number of pieces of tumour tissue provided ranged from 1 to 3. Tumour samples were then stored at minus 80 °C until required. After allowing time for histopathological tests and reports to be completed, patient notes were accessed by a clinician involved in the study, and relevant data was recorded. This included details regarding the type of tumour, the molecular subtype of the tumour (section 2.2.2), the chemotherapy administered, tumour sizes pre- and post-treatment from dynamic contrast enhanced – magnetic resonance imaging (DCE-MRI) and ultrasound (US) scans, as well as the pathology reports from both the core biopsy specimen taken upon diagnosis and the tumour resection.

5.2.3 Clinical tumour sample collection

Taking into consideration the influence sample acquisition has upon downstream analysis (section 5.1.2), tumour samples were obtained for research from the tumour after macroscopic examination by a surgeon, to ensure the quality of the sample. Samples were snap-frozen in liquid nitrogen as soon as possible, where they remained until they reached the laboratory where they were transferred to a minus 80 °C freezer for long-term storage. Samples were processed as they were required, which was performed on ice, with the addition of phosphatase and protease inhibitors to minimise sample degradation.

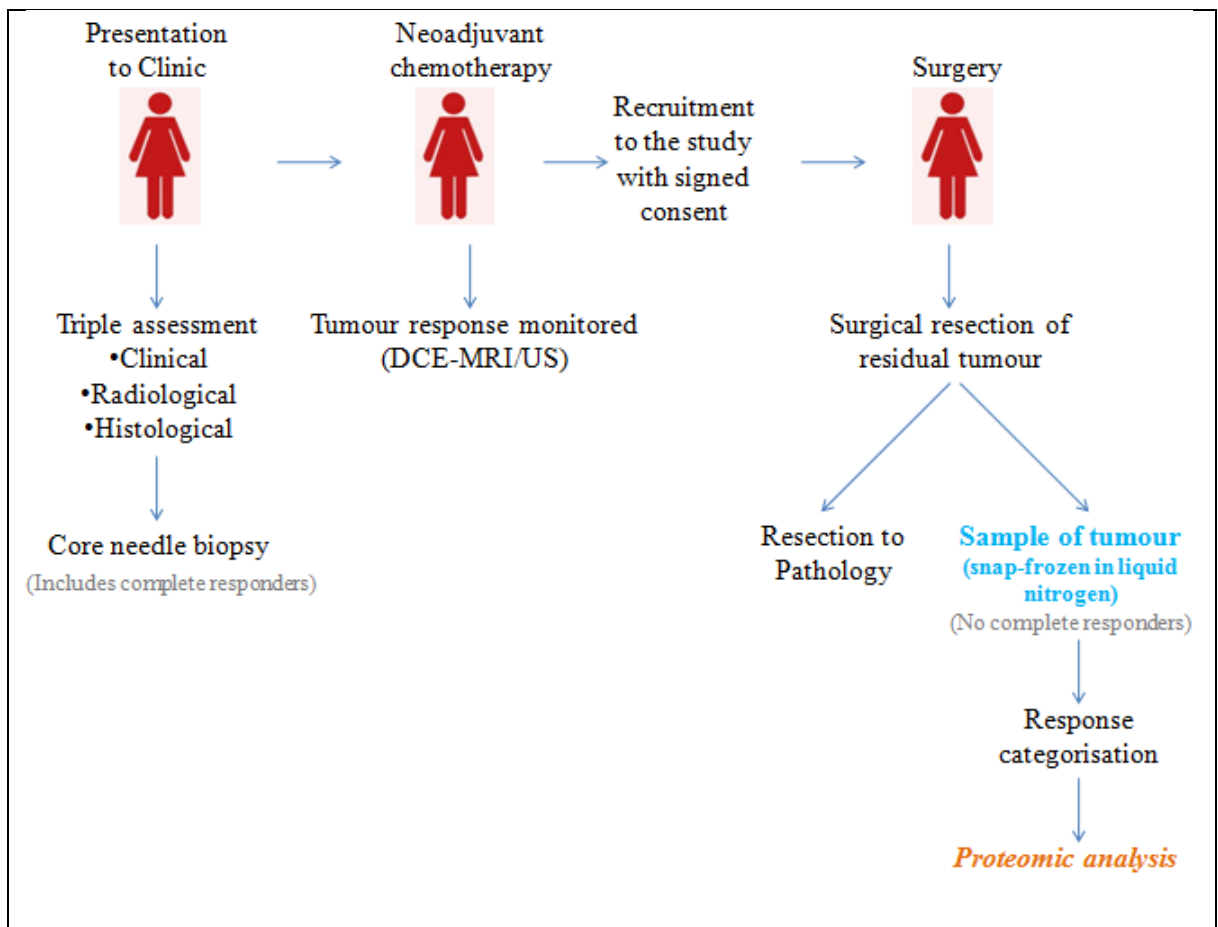


Figure 17: Flow diagram illustrating study design

Study workflow from patient presentation at the clinic through to proteomic analysis in the laboratory. Where tumours have shown a complete response to neoadjuvant chemotherapy, it is not possible to obtain a sample for the study from the tumour resection. These patients could be included in future archival studies to validate putative biomarkers.

5.2.4 Classification of response to neoadjuvant chemotherapy

Tumour response to treatment was calculated, for all samples collected, by a specialist oncologist and patients were assigned to one of two groups; (1) responders (chemotherapy-sensitive) or (2) non-responders (chemotherapy-resistant). To determine response, data from pre-treatment imaging (DCE-MRI and US) was compared to data from post-treatment scans (DCE-MRI and US) as well as the pathological report of the tumour resection, whilst considering RECIST guidelines (Therasse, Arbuck et al. 2000; Eisenhauer, Therasse et al.

2009), where tumour size was determined by measuring the longest diameter. Pre-treatment DCE-MRI data was compared to the post-treatment pathological report, as inaccuracies in DCE-MRI measurements after neoadjuvant chemotherapy have been reported, in comparison to histopathological measurements (Wright, Zubovits et al. 2010; Loo, Straver et al. 2011). Where only US data was available pre-treatment, US data post-treatment was also considered. Response calculations were determined in agreement with a specialist oncologist and breast surgeon. For the purpose of this study, patients were divided into one of two groups, as described previously (Garimella 2007):

- Responders (chemotherapy-sensitive)
 - These include patients who showed a partial response ($\geq 30\%$ decrease in tumour size)
- Non-responders (chemotherapy-resistant)
 - These included patients who showed either stable disease ($< 30\%$ decrease and $< 20\%$ increase in tumour size) or progressive disease ($\geq 20\%$ increase in tumour size or development of metastases during therapy)

For patients receiving the combined anthracyclin (EC) and taxane (docetaxel) regimen, to be classified as a 'responder' the patient must have responded to both components. Where progression was seen during any one of the components, the patient was categorised as a 'non-responder'.

5.3 Results

5.3.1 Clinical tumour sample series

In total, tumour samples were obtained from 38 patients. Of the 38 tumour samples some had to be excluded from the study, for reasons such as the inability to confidently determine

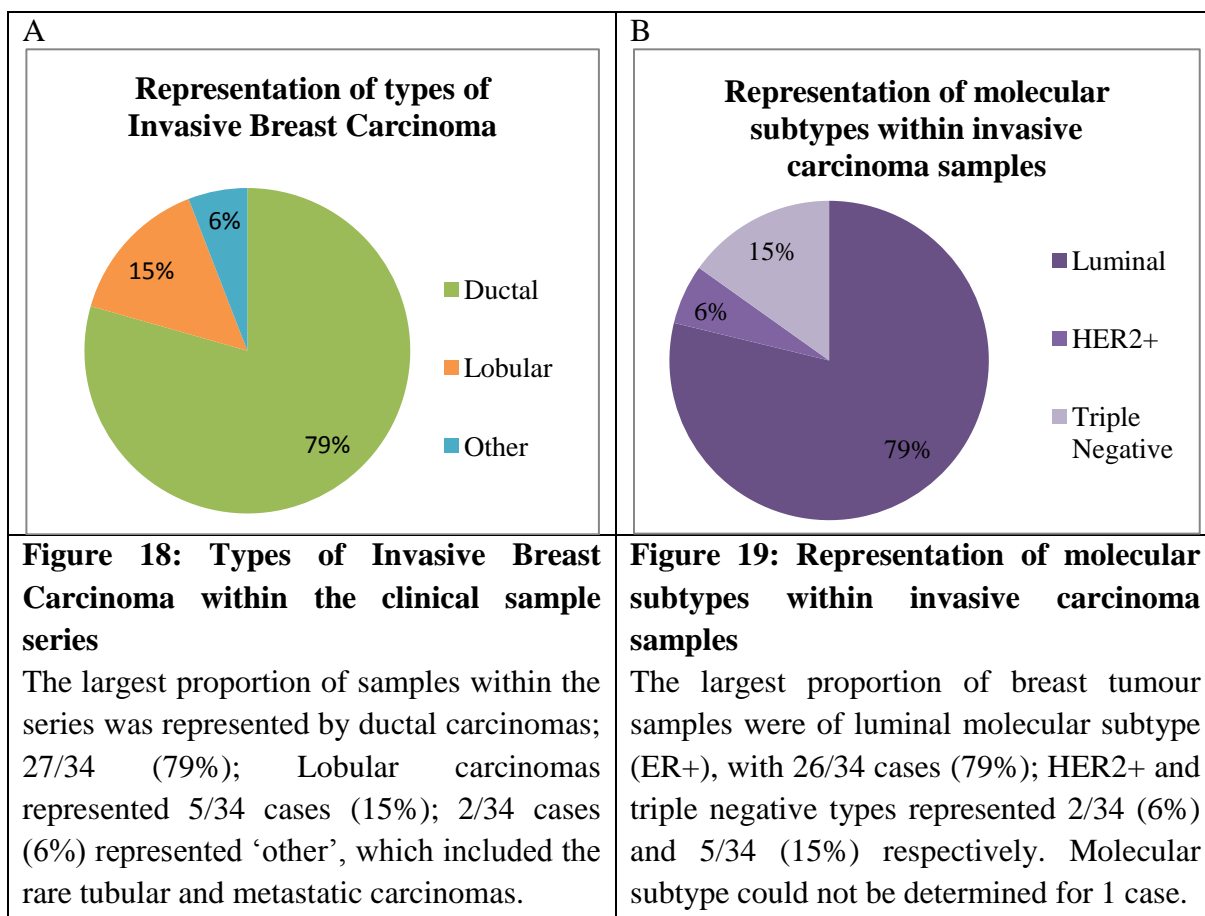
tumour response to therapy, and incorrect storage/handling, leaving a total of 34 clinical tumour samples. A summarised table of clinical tumour samples involved in the study is shown in Table 13, and the full version of the table including measurements used for response determination can be found in Appendix 5.

Of the 34 clinical tumour samples, 27 (79%) were invasive *ductal* carcinoma (IDC), 5 (14%) were invasive *lobular* carcinoma (ILC) and 2 (5%) were other rare types (metaplastic and tubular) (Figure 18). Within the breast tumour samples, 26/33 (78%) were luminal (ER+), 2/33 (6%) were HER2+ and 5/33 (15%) were triple negative molecular subtypes (section 2.2.2) (Figure 19). Molecular subtype could not be determined for one of the samples due to lack of HER2 status. All ILC samples within the series were of luminal molecular subtype. Overall, IDCs of luminal molecular subtype (ER+) represented the largest proportion of the samples, representing 20/34 cases (58%). The majority of patients received standard neoadjuvant chemotherapy for Hull and East Yorkshire NHS Trust, comprising 4 cycles of EC [epirubicin (90 mg/m^2) + cyclophosphamide (600 mg/m^2)] followed by 4 cycles of docetaxel (100 mg/m^2), given at 3-weekly intervals. Not all patients completed the full 8 cycles (Table 13), which may have been due to tumour progression during the chemotherapy, poor tolerance of associated side-effects or previous exposure to the chemotherapeutic agents. Five out of the 34 patients were assigned to the Neo-tAnGo randomised phase III clinical trial of sequential epirubicin + cyclophosphamide and paclitaxel \pm gemcitabine doses, with 4 treatment arms (Earl, Vallier et al. 2009), which ended in 2007 and aimed to identify the best drug combination and order of administration. Of the 34 tumour samples, 18/34 (52%) were classified as ‘chemotherapy-sensitive’ and 16/34 (47%) were classified as ‘chemotherapy-resistant’.

Table 13: Summarised data for clinical tumour samples

Table listing clinical tumour samples (n=34), showing the type of breast carcinoma, the receptor status (ER/PR/HER2), molecular subtype and therapy administered, represented by EC: epirubicin + cyclophosphamide, D: docetaxel, P: paclitaxel, G: gemcitabine, with the number of cycles. Reduced doses are indicated* (also see Appendix 5) Response was classified as chemo-sensitive (CS) or chemo-resistant (CR).

Sample #	Type	Receptor status			Molecular Subtype	Therapy administered	Response (CS/CR)
		ER	PR	HER2			
#1	Ductal	+	-	-	Luminal	EC x 4, D x 4	CR
#3	Ductal	-	-	+	HER2+	Neo-tAnGo: EC x 6	CR
#4	Ductal	+	+	-	Luminal	Neo-tAnGo: EC x 4, P x 4	CR
#5	Ductal	-	-	-	Triple Negative	Neo-tAnGo: PG x 4, EC x 4	CR
#6	Ductal	+	+	-	Luminal	Neo-tAnGo:EC x 4, PG x 4	CS
#7	Ductal	+	+	-	Luminal	EC x 4, D x 2*	CS
#8	Ductal	-	-	-	Triple Negative	Neo-tAnGo: EC x 4, P x 4	CS
#9	Ductal	+	+	+	Luminal	EC x 4	CR
#10	Metaplastic	-	-	?		EC x 4, D x 4	CS
#11	Ductal	+	+	-	Luminal	EC x 4, D x 4	CS
#12	Ductal	+	+	?	Luminal	EC x 4, D x 2*	CS
#13	Ductal	-	-	+	HER2+	EC x 4, D x 1	CR
#15	Ductal	+	+	-	Luminal	EC x 4, D x 4	CS
#16	Ductal	+	+	?	Luminal	EC x 4, D x 4	CR
#17	Lobular	+	+	-	Luminal	EC x 3	CR
#18	Ductal	+	+	-	Luminal	EC x 4, D x 2	CS
#19	Ductal	+	-	-	Luminal	EC x 2	CR
#20	Ductal	-	-	-	Triple Negative	EC x 4	CR
#22	Ductal	+	-	-	Luminal	EC x 4, D x 4	CS
#23	Ductal	+	-	-	Luminal	EC x 4, D x 3*	CS
#24	Lobular	+	+	-	Luminal	EC x 4, D x 4	CR
#25	Ductal	+	+	+	Luminal	EC x 2, D x 4*	CR
#27	Tubular	+	+	-	Luminal	EC x 4, D x 4	CS
#28	Ductal	+	+	+	Luminal	EC x 4, D x 2*	CS
#29	Ductal	+	-	-	Luminal	EC x 4, D x 4*	CS
#30	Ductal	+	-	+	Luminal	EC x 4, D x 4	CS
#31	Ductal	+	+	+	Luminal	EC x 4, D x 3*	CS
#32	Lobular	+	+	-	Luminal	EC x 6	CR
#33	Lobular	+	+	-	Luminal	EC x 4, D x 4	CR
#34	Lobular	+	+	-	Luminal	EC x 4, D x 4	CS
#35	Ductal	-	-	-	Triple Negative	EC x 5	CR
#36	Ductal	-	-	-	Triple Negative	EC x 4, D x 1	CR
#37	Ductal	+	+	-	Luminal	EC x 4, D x 4*	CS
#38	Ductal	+	+	-	Luminal	EC x 4, D x 4	CS



5.4 Discussion

In total, clinical tumours samples were obtained from 34 locally advanced breast cancer patients who had received neoadjuvant chemotherapy treatment. This involved a sample of macroscopic tumour being taken by a surgeon during standard surgical resection of the remaining tumour mass. The relevant factors outlined in the BRISQ guidelines, including all relevant clinical information, were recorded for each patient, and factors relating to sample collection method, stabilisation mechanism, type of long term storage, and shipping and storage temperatures were kept constant throughout.

The largest proportion of the clinical breast tumour samples collected was represented by IDC samples, which comprised 79% (27/34) of samples. This could be expected, as IDC is

the most common type of invasive cancer, and represents 70-80% of cases clinically. Two recent reports from large study cohorts have described the prevalence of each molecular subtype of breast cancer; (1) Out of 1487 patients diagnosed with invasive breast cancer over 2004-2005 in Italy, 70.3% were classified as luminal A, 15.6% luminal B, 6.0% HER2+ and 8.1% triple negative (Caldarella, Crocetti et al. 2011); (2) Out of 1945 patients diagnosed with invasive breast cancer between 1976-1997 in America, 65.8% were classified as luminal A, 14.3% luminal B, 4.9% HER2+ and 10.4% triple negative (Dawood, Hu et al. 2011). Therefore, on average, invasive breast tumours of luminal subtype represent 83% of cases, HER2+: 5.5% and triple negative: 9.3% of cases. The breast tumour clinical samples collected within this study, show a similar representation of each molecular subtype; luminal: 79% (26/34), HER2+: 6% (2/34) and triple negative: 15% (5/34).

Of the 34 tumour samples collected, 18/34 (52%) were classified as 'chemotherapy-sensitive' and 16/34 (47%) were classified as 'chemotherapy-resistant'. These figures do not reflect the clinical situation, as samples were not available for the tumours which showed a pathological complete response to neoadjuvant chemotherapy, and therefore not represented within the collection of 34 tumour samples.

Clinical breast tumour samples which have been collected for the study include the main histological types and molecular subtypes of invasive breast cancer. Differences in tumour response to neoadjuvant chemotherapy were observed, and the response of each individual tumour was determined by a specialist oncologist. This allowed the categorisation of clinical breast tumour samples into chemotherapy-sensitive and chemotherapy-resistant types. These clinical samples may now therefore be used for comparative proteomic analysis, to identify biomarkers of chemotherapy resistance in breast cancer.

CHAPTER 6:
CLINICAL TISSUE SAMPLES –
OPTIMISATION OF METHODS AND
PRELIMINARY WORK

Chapter Aim:

Optimisation of proteomic methodologies for use with clinical breast tumour tissue samples, prior to comparative proteomic analysis for the search for putative biomarkers of chemotherapy resistance in breast cancer

Chapter 6. Clinical tissue samples – optimisation of methods and preliminary work

6.1 Introduction

6.1.1 Cell lines versus clinical samples

The choice of samples currently available, for the identification of biomarkers of chemotherapy resistance in breast cancer using proteomic analysis, within the laboratory are cell lines and clinical tissue samples. Both of these have their own limitations, which are briefly outlined and compared in Table 14.

Table 14: Comparison of cell line models with clinical tissue samples

Cell lines	Clinical tissue samples
Easy to handle in controlled conditions	Difficult to handle and process in controlled conditions
Homogeneous sample	Heterogeneous sample
Unlimited amount	Limited amount
Inaccurate representation of a tumour cell	Actual representation of a cell from a tumour microenvironment

6.1.1.1 Cell lines

Current methodologies and proteomics platforms within the laboratory have previously been established for use with cell line models (Smith, Watson et al. 2006; Smith, Welham et al. 2007). Cell lines are standardised homogenous collections of cells, which present an *in vitro* model of the disease. Their use has several advantages; they can be cultured to unlimited amounts, are homogeneous, and are easy to store and handle. However, the use of cell lines also carries several disadvantages; the principle of which being that continuous culturing over long periods of time can cause alterations in the genotypic and phenotypic

characteristics of the cell (Bahia, Ashman et al. 2002; Burdall, Hanby et al. 2003; Watson, Greenman et al. 2004). It has been shown that alterations exist in the characteristics of commonly used cell lines, such as MCF7 (Bahia, Ashman et al. 2002) and MDA-MB-231 (Watson, Greenman et al. 2004) between laboratories and cultures of different time points, thus making the accuracy of these *in vitro* models questionable. Cross-contamination and mis-identification of cell lines has also been observed, the most well known example of which is the cross-contamination of cell lines with HeLa cells. This wide-spread problem was initially identified by Stanley Gartler, where he discovered that 18 cell lines of ‘independent origin’ were all HeLa cells (Gartler 1967). Further studies were carried out by the late Walter Nelson-Rees (Masters 2010), and it was suggested at one point that all cell lines should be regarded as HeLa cells unless proved otherwise. Initially, these findings appeared to have little impact, yet more recently, disclosure of cell line origin and their authentication is increasingly required for publication. Microbial contamination of cell lines, with microorganisms such as mycoplasma is also a major, but largely underrated problem (Masters 2000). Mycoplasma infections are thought to be present in 15-35% of cell lines in continuous culture, and have the potential to cause several different types of effects on eukaryotic cells; alterations in protein synthesis, receptor and surface antigen expression, cell morphology and proliferation characteristics, for example (Drexler and Uphoff 2002). However, solutions are available to the above-mentioned shortcomings regarding the use of cell lines; validation of the cell line identity, by short-tandem-repeat profiling for example; purchase of cell lines from reputable repositories rather than other laboratories; routine testing for infections such as mycoplasma; avoiding over-culturing and high passaging of cells by regularly returning to original frozen stocks of low passage number, and careful laboratory aseptic technique (Masters 2000; Drexler and Uphoff 2002;

Burdall, Hanby et al. 2003; Buehring, Eby et al. 2004; Capes-Davis, Theodosopoulos et al. 2010; Masters 2010). In summary, high quality-control is necessary when using cell lines in order to be confident of their accuracy as an *in vitro* model and confident in the data generated from them.

6.1.1.2 Clinical Samples

Clinical samples give researchers the opportunity to perform analysis on samples from more clinically relevant sources. Clinical samples may include biological fluids (serum, blood and tumour interstitial fluid, for example) and tissue (normal or diseased). However, the complexity of these samples is reflected by the complexity of the approach required to gain access to the relevant information hidden within. As described previously (section 3.1), 22 proteins constitute 99% of serum, thus making low-abundant proteins of interest difficult to access. The well-recognised graph by Anderson and Anderson (Figure 20) (Anderson and Anderson 2002), clearly demonstrates the dynamic range of the serum proteome, spanning 10 orders of magnitude between the most abundant protein (albumin; normal level $35\text{-}50 \times 10^9$ pg/ml) and a low abundant protein (e.g. interleukin-6; normal level 0-5 pg/ml). This therefore explains why it is difficult to identify such a low abundant protein within the serum; as described by Anderson and Anderson, 'it would be like trying to identify one person after screening the whole worldwide population: 1 in 6.2 billion' (Anderson and Anderson 2002). In order to improve access to the low-abundant proteins, and remove the high-abundant proteins which mask them, pre-fractionation steps such as depletion strategies may be employed. Different depletion strategies may be used, using different commercially-available products, which have been compared for their efficiency, specificity and reproducibility (Bjorhall, Miliotis et al. 2005). The approach used also depends on the number of highly abundant proteins to be depleted. This may include

albumin ± immunoglobulin (Ahmed, Barker et al. 2003; Steel, Trotter et al. 2003; Kim, Bae et al. 2009), or the top six, twelve or twenty serum proteins (Bjorhall, Miliotis et al. 2005; Echan, Tang et al. 2005; Roche, Tiers et al. 2009). Depletion strategies seem popular, and successful, but should be tailored to the specific method being applied.

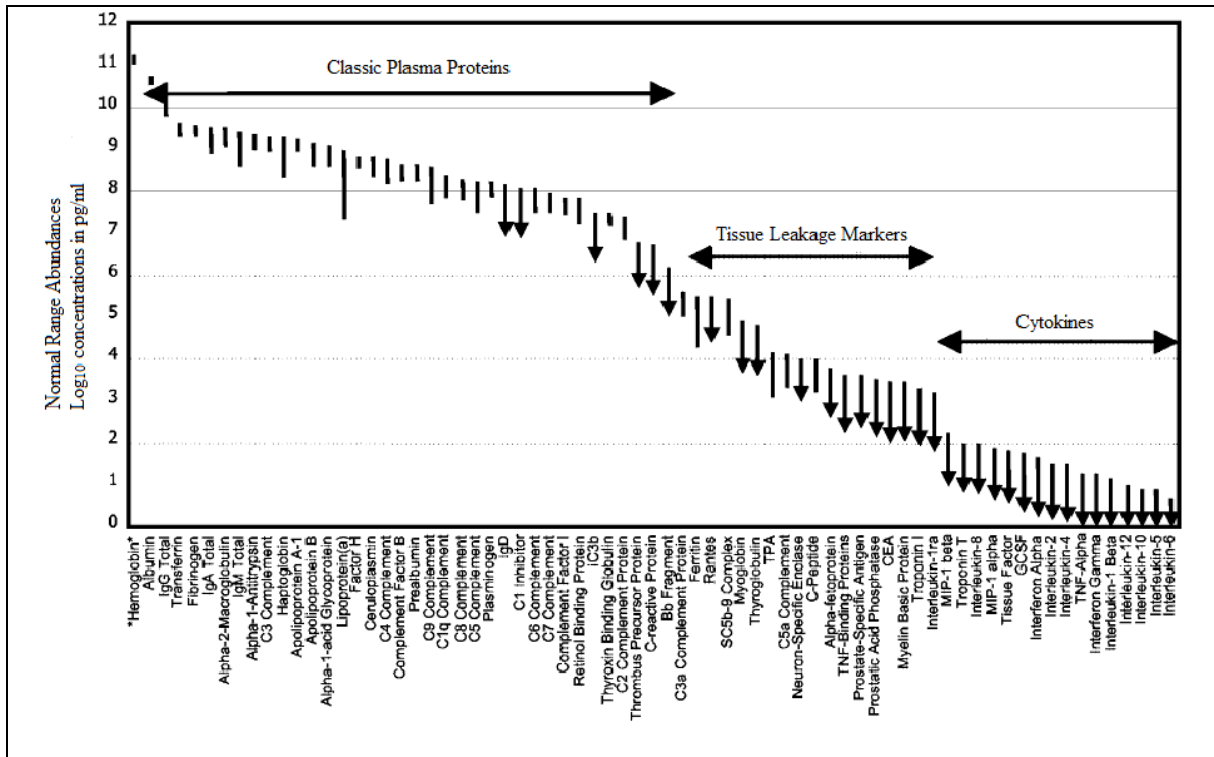


Figure 20: Dynamic range of normal human plasma proteins
 The classic log-scale graph by Anderson and Anderson (2002), giving reference intervals for 70 proteins within the plasma. The difference between the most abundant protein (albumin) and the least abundant protein (interleukin-6), spans 10 orders of magnitude. Haemoglobin is shown (far left) as a reference point. The classic plasma proteins are shown to the left (high abundance), tissue leakage markers in the centre and cytokines to the right (low abundance).
 Abbreviations: TPA: tissue plasminogen activator, CEA: carcinoembryonic antigen, G-CSF: granulocyte colony-stimulating factor, TNF: tumour necrosis factor
 From (Anderson and Anderson 2002)

When working with clinical samples such as tumour tissue, one of the main challenges is that it is more difficult to store and handle in a controlled manner (section 5.1.2), is usually of limited supply, and access requires ethical approval and patient consent. Another

technical issue, associated with breast tumour for example, is the heterogeneity of the tissue. Macroscopically dissected tumour contains a dense area of tumour, yet it may also contain microscopic areas of normal cells, stroma, adipose tissue and inflammatory cells. Depending on experimental design, approaches such as laser capture microdissection (LCM), which is most effective on archival tissue, may be used to enrich for tumour cells within the sample to be analysed (Curran, McKay et al. 2000; Craven, Totty et al. 2002; Ball and Hunt 2004; Morrogh, Olvera et al. 2007). However, this has to be balanced by sample availability, sample degradation and contamination, and the downstream analytical technique(s) being applied.

Laser capture microdissection (LCM) is a powerful technology which was introduced to overcome the problem of tissue heterogeneity, by allowing the procurement of enriched populations of cells from tissue sections under direct microscopic visualisation (Curran, McKay et al. 2000) for the study of DNA, RNA and protein (Ball and Hunt 2004; Morrogh, Olvera et al. 2007). The main advantage of using LCM prior to molecular evaluation is that the sample under investigation contains a highly enriched population of the cells of interest, thus increasing the accuracy of the data produced. This approach acts as an alternative to the 'averaging out' approach employed by un-microdissected samples. Although the use of LCM seems very appealing, many aspects have to be considered. Firstly considerations must be made regarding the pre-LCM procedures of tissue fixation and staining, and their compatibility with downstream applications, and the risk of artefacts (Craven, Totty et al. 2002).

LCM requires the staining of tissue sections, for visualisation of cell populations, and this must be compatible with downstream applications. Different staining methods, including the use of haematoxylin and eosin, methyl green and toluidine blue, have shown to have

different effects upon protein recovery, IEF and protein visualisation (Craven, Totty et al. 2002; Ball and Hunt 2004). Moreover, the success of LCM is said to be tissue-dependent (Craven, Totty et al. 2002).

For proteomic analysis, fresh frozen tissue is often required. The use of LCM with fresh frozen tissue and the subsequent proteomic analysis by LC MS/MS and western blotting has been achieved (Johann, Rodriguez-Canales et al. 2009). However, the main concern regarding the use of fresh frozen tissue is protein degradation; at this stage protease inhibitors cannot be employed, which may result in loss of proteome integrity, leading to sample bias, and inaccurate data. In order to overcome the issue of sample degradation, the use of formalin-fixed paraffin-embedded (FFPE) tissue would be required. However, this complicates downstream proteomic analysis due to the cross-linking of proteins (Ball and Hunt 2004). Different fixation methods have different effects upon proteins, depending upon the type of reagents used, and non-cross-linking reagents such as methanol have been preferred for proteomic research (Gutstein and Morris 2007). However, FFPE is the standard method of fixation used for pathological specimens, so the ability to use FFPE for proteomic analysis may make the use of clinical samples more accessible. Within recent years, the use of FFPE for proteomic analysis has developed dramatically, with the introduction of a method involving the heat-mediated reversal of cross-links (Nirmalan, Harnden et al. 2008), allowing LC MS/MS (Nirmalan, Hughes et al. 2011) as well as western blotting (Nirmalan, Harnden et al. 2009).

The most critical limitation of performing LCM, on any sample type, prior to proteomic analysis is sample loss, and the ability to retain sufficient sample for downstream applications (Craven, Totty et al. 2002). This is due to the selection of specific regions of the tissue section only, not the whole tissue section. Sufficient sample must be obtained to

run 2D-PAGE (e.g. 200 µg of protein per gel), including technical replicates as well as antibody microarray analysis and western blotting. It is possible to increase the sensitivity of 2D-PAGE by using fluorescent dyes, which can also be run on one gel, therefore decreasing the amount of sample required (Ball and Hunt 2004). Another option would be to pool clinical samples, yet this is largely dependent on the range and type of sample availability, and may introduce sample bias. The reduced sample availability following LCM would also demand optimal protein extraction techniques, of which there are several options, however this still needs to be tailored to the downstream application, the type of tissue and the amount of tissue (Gutstein and Morris 2007). A movement towards microproteomics has also been suggested to overcome issues with small sample sizes (Gutstein, Morris et al. 2008). When performing comparative proteomics analysis, care must also be taken not to introduce bias as a result of LCM (Gutstein and Morris 2007); for example, changes in protein abundance, or protein modification/degradation, by fixation/staining methods could introduce false differential expression.

LCM is a very useful and powerful technique, the efficacy of which can be determined by protein yield and the percentage of enrichment (Craven, Totty et al. 2002); however its use is very dependent upon the nature of the experiment, the aim of the research and the sample type. A balance must be achieved between time, expense, protein modifications and protein degradation, as well as contamination (Curran, McKay et al. 2000; Craven, Totty et al. 2002).

Despite the technical hurdles associated with the use of clinical samples, their use potentially has a great advantage over *in vitro* cell line models. The cells / molecules being analysed from clinical samples originate from the true biological microenvironment; they have been influenced by their surroundings and other cells in contact with them, so

arguably provide a far more accurate representation of the clinical scenario. More recently, the importance of the tumour microenvironment is being recognised (Liotta and Kohn 2001), particularly the tumour-associated stroma (Farmer, Bonnefoi et al. 2009). Therefore, arguably, the use of clinical tumour samples is more beneficial than using cultured cell lines, as cultured cells do not truly represent the tumour in its biological microenvironment (Geho 2004).

6.1.2 Proteomics using clinical tissue samples

In order for proteomic analysis to be carried out using clinical tissue samples, protein extraction methods first had to be established and optimised for each of the analytical techniques being applied. A different approach is required, in comparison to cells lines; the most obvious reason for this being that tissue samples are a tangible collection of structurally organised cells, and not a single-cell suspension.

6.1.2.1 Sample preparation

Sample preparation of tissue for proteomic analysis generally consists of two phases: (1) tissue disintegration and (2) tissue solubilisation in an appropriate lysis buffer for protein extraction (Gromov, Celis et al. 2008). Tissue disintegration involves the breaking-up of a solid piece of tissue into smaller pieces, to increase the surface area, and solubilisation involves cell suspension within a lysis buffer and subsequent protein extraction. This can be achieved using several methods, which are dependent upon downstream applications and the possibility of sample contamination should also be considered. Methods include mechanical homogenisation by instruments such as the TissueRuptor (Qiagen), sonication, grinding in liquid nitrogen, cryostat sectioning (Gromov, Celis et al. 2008) enzymatic digestion and pressure cycling technology (Smejkal, Witzmann et al. 2007). For

proteomics, grinding in liquid nitrogen and mechanical homogenisation appear to be most popular.

The choice of extraction/lysis buffer is dependent upon the downstream application, due to reagent compatibilities and the required physicochemical state of the protein. Generally, different lysis buffers are used for different applications, which complicate the sample preparation process. However, Gromov *et al* recently reported a single lysis solution which is suitable for both array-based proteomics and 2D-PAGE protein separation, extracting proteins from a variety of clinical tissues (Gromov, Celis et al. 2008).

6.1.3 Proteomic methods

The proteomic methods used to analyse the samples include antibody microarrays (section 3.5), 2D-PAGE MALDI-TOF/TOF MS (section 3.3) as well as western blotting (section 3.6.2). These all require a minimum protein concentration of 1 mg/ml, and a total volume of 1 ml, with proteins suspended in their own respective buffers.

6.1.3.1 Core biopsy samples

If predictive biomarkers of chemotherapy resistance were transferred to the clinic, they would be used as a screening tool at the diagnostic stage. This may involve screening core biopsy samples, and predicting tumour response to treatment based on the expression of selected protein biomarkers. Therefore, in order to identify *predictive* markers of resistance the proteomic analysis of the pre-treatment core biopsy samples would be desirable. However, this would only be possible where proteomics techniques for use with small amounts of clinical tissue (of core biopsy size) were established. Due to the invasive procedure involved with taking core biopsy samples, the number of core biopsies taken for proteomic analysis would have to be kept to a minimum. In order to assess the feasibility of

performing proteomic analysis on core biopsy samples, protein yield must initially be determined. This would determine whether sufficient protein could be obtained from an ethically-reasonable number of cores. The first study which used core biopsies samples for proteomic analysis was carried out in Italy, and involved taking at least 5 core biopsy samples from each patient (Bisca, D'Ambrosio et al. 2004); this is a large number of core biopsy samples, which may not currently be approved by the Ethics Committee or gain patient consent.

6.2 Methods

6.2.1 Antibody microarray – based methods

6.2.1.1 Protein extraction

Protein was extracted, using ‘Buffer A’ from the antibody microarray kit (#XP725, Sigma Aldrich), as described in section 4.3.1, from tumour resection samples (Figure 15).

6.2.1.2 Protein precipitation

Protein was precipitated from the sample using the ProteoExtract® Protein Precipitation Kit (#539180, Calbiochem), as described in section 4.3.2.

6.2.1.3 Protein quantification

Protein quantification for proteins extracted in ‘Buffer A’, from the antibody microarray kit (#XP725, Sigma Aldrich), were quantified using the Bradford protein assay, as described in section 4.3.3.

6.2.1.4 Determination of protein yield from core biopsy samples

This was achieved by taking an ‘optimisation sample’ of tumour resection, collected as part of the existing study (Figure 15) and taking core biopsy samples from it with the use of a core biopsy gun. This was performed by a surgeon, which involved securing the tissue sample to an ‘Easi pad’ using a stitch, and the core biopsy gun was used to obtain a series of core biopsy samples from the tumour mass. Due to the size of the tissue, protein yield could only be determined from 1, 3 and 6 core biopsy samples. Protein was then extracted using 500 µl ‘Buffer A’, provided in the antibody microarray kit and subsequent sonication for 15 minutes (with 5 mins on ice every 5 mins) and incubation for 16 hours at 4 °C on an end-over-end rotator. This was followed by centrifugation to clarify the sample, and the protein concentration was determined using the Bradford Assay (section 4.3.3).

6.2.2 One-dimensional gel electrophoresis

One-dimensional gel electrophoresis was performed as described in section 4.6.3. Proteins separated on the gel were visualised using Bio-safe Coomassie Stain (#161-0787, Bio-Rad), as described in section 4.4.6.

6.2.3 2D-PAGE MALDI-TOF/TOF MS – based methods

6.2.3.1 Protein extraction

Protein extraction for 2D-PAGE was performed using the extraction method described in section 4.4.1 upon ‘optimisation samples’ (Figure 15).

6.2.3.2 Sample preparation

The ReadyPrep 2-D Cleanup Kit (#163-2130, Bio-Rad) was used to clean-up the samples in preparation for IEF (section 4.4.2). The protein in the sample was quantified using the 2-D Quant Kit (#80-6483-56, GE Healthcare) (section 4.4.3).

6.2.3.3 2D-PAGE

Separation in the first dimension by pI; isoelectric focusing (IEF), was performed as described in section 4.4.4. Subsequently, separation in the second dimension by molecular weight, using SDS-PAGE, was performed as described previously (section 4.4.5). Proteins were visualised using Bio-safe Coomassie Stain (#161-0787, Bio-Rad), as described in section 4.4.6.

6.2.3.4 Excision of protein spots and in-gel tryptic digest

Protein spots were excised from the gel as described in section 4.4.8. Proteins were then digested into peptide using the method described in section 4.4.9.

6.2.3.5 Peptide analysis by MALDI-TOF/TOF MS and protein identification

Peptide samples were analysed by MALDI-TOF/TOF and data was submitted to MASCOT using specified search parameters for protein identification via the IPI human database, as described in section 4.4.11.

6.2.4 Western blot

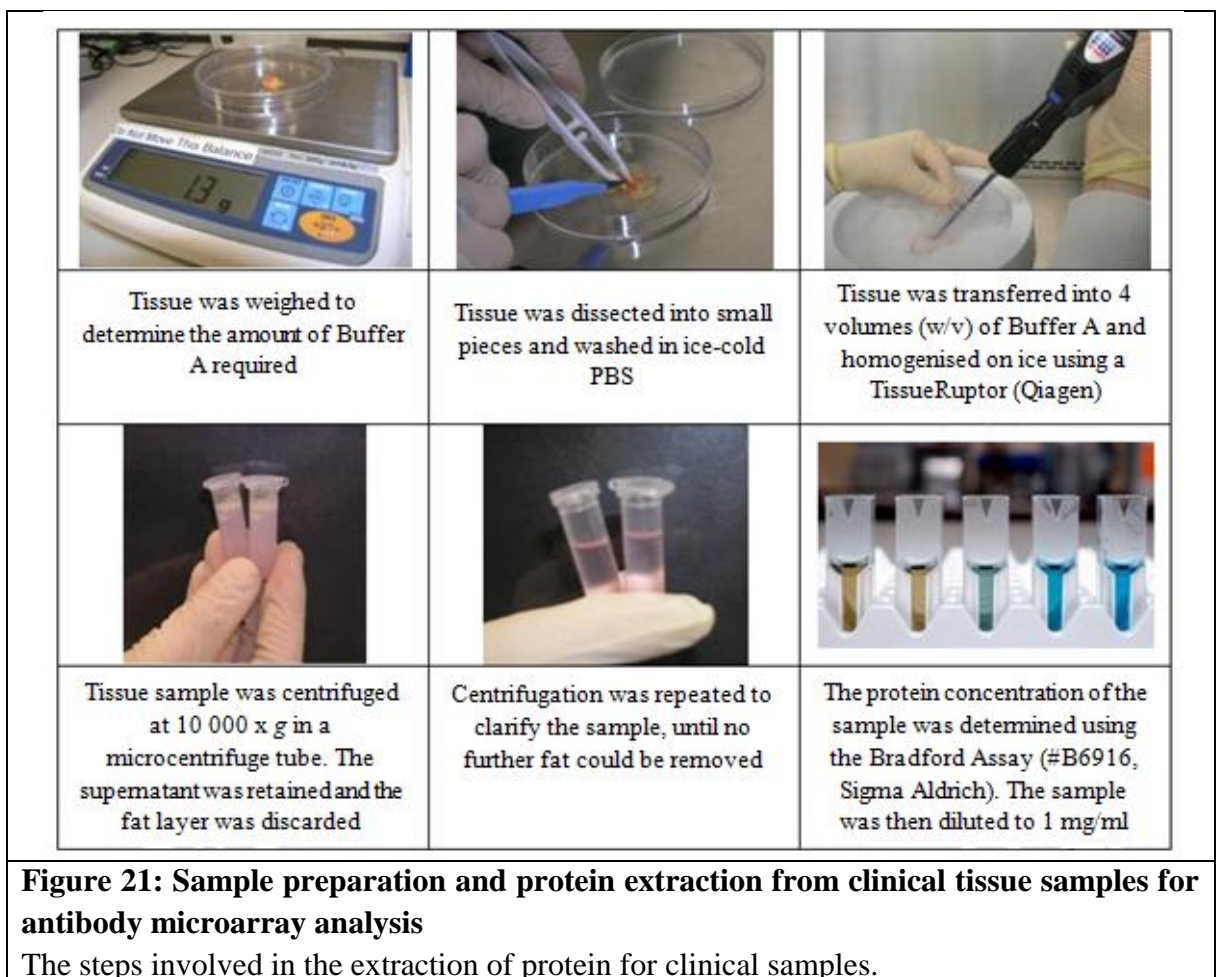
One dimensional gel electrophoresis was performed, as described previously (section 4.6.3). Proteins were then transferred onto a nitrocellulose membrane using the iBlot system (section 4.6.4), which was subsequently blocked for 1 hour with 5% non-fat milk and immunoblotting was performed as previously described (section 4.6.6). Protein

expression was visualised using the Supersignal West Pico Chemiluminescent Substrate Kit (#34078, Thermo Scientific) and CL-XPosure Film (#34090, Thermo Scientific), as described previously (section 4.6.8).

6.3 Results

6.3.1 Antibody Microarray

Protein extraction for antibody microarray analysis involved the mechanical homogenisation of the tumour resection sample in lysis 'Buffer A', as illustrated in Figure 21.



6.3.1.1 Relationship between tissue weight and protein yield

Several protein extractions using antibody microarray extraction buffer ('buffer A') were performed (n=11), with the aim to correlate tumour sample size (by eye) with tumour mass (in grams) and protein yield (Table 15). However, there was no correlation between tumour sample size (by eye) or mass (g) and protein yield (mg/ml) (Figure 22). Tumour samples which were approximately the same size by eye showed great variability in weight and protein yield. This was due to the high amount of tissue heterogeneity observed between tumour samples, despite carefully removing macroscopic fat from the sample prior to extraction.

Table 15: Extraction of protein from breast tumour resection samples

Protein was extracted from 'optimisation samples', using the method described within the antibody microarray protocol, and clinical samples prior to antibody microarray analysis (Figure 15). Tumour samples were weighed, protein was extracted and the protein concentration was determined using the Bradford Assay. This data is displayed as a graph in Figure 22.

Tumour resection sample	Weight (g)	Protein concentration of extract (mg/ml)
Optimisation sample	0.5	2.03
Optimisation sample	0.4	1.23
Optimisation sample	0.7	0.78
Optimisation sample	0.4	0.67
#9	0.8	3.45
#11	0.4	4.99
#12	0.16	2.64
#15	0.5	3.11
#18	0.3	2.43
#19	1.3	4.53
#25	0.3	3.05

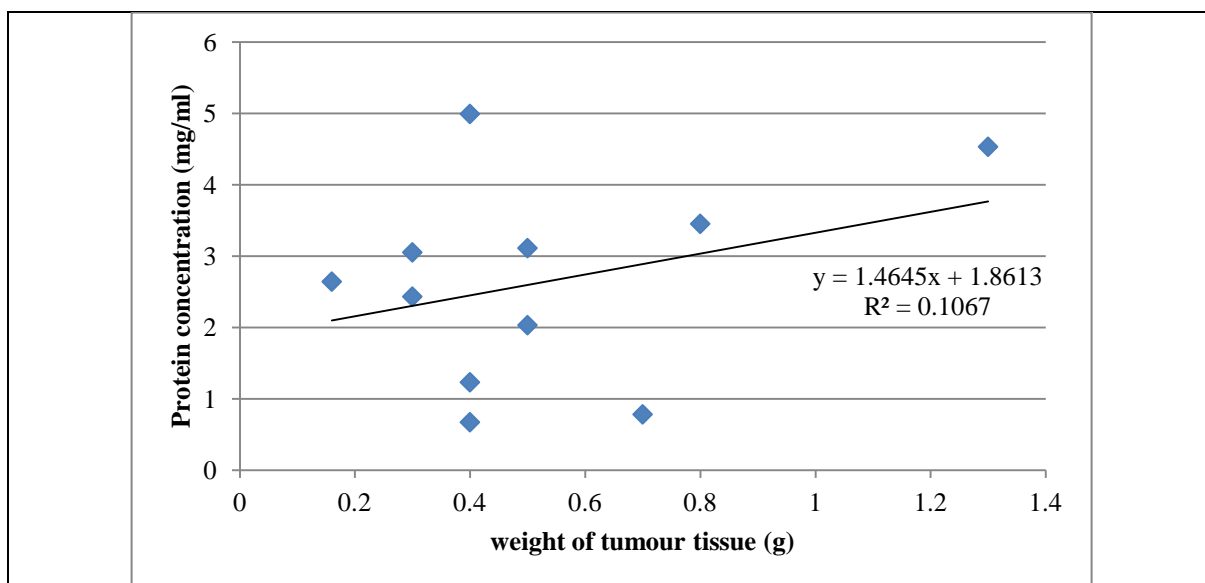


Figure 22: Correlation between breast tumour tissue weight and protein yield
 Scatter graph showing poor relationship between breast tumour resection sample weight (g) and protein yield, calculated by extracting protein from breast tumour sample with the method described for antibody microarray analysis and protein quantification using the Bradford assay (#B6916, Sigma Aldrich) (Table 15). The R^2 value of 0.1, calculated by the trend-line shows poor linear reliability.

6.3.1.2 Precipitation of protein using ProteoExtract Protein Precipitation Kit

As illustrated in Figure 21 a layer of fat was observed during the protein extraction method. Removal of this fat was achieved by repeated centrifugation. However, this also resulted in the reduction of sample volume. As an alternative approach, for the purification of the sample, the protein was precipitated out of the solution using the ProteoExtract Protein Precipitation Kit (#539180, Calbiochem) and re-suspended in fresh Buffer A. This was tested using 200 μ l of the protein extracts from samples #11 and #19. Ten μ g each of: #11 (precipitated and non-precipitated); #19 (precipitated and non-precipitated) (n=4) were loaded in duplicate and run on a 1D gel which was subsequently stained with Bio-safe Coomassie Stain (#161-0787, Bio-Rad) to visualise the proteins present on the gel. Samples where the protein had been precipitated appeared to have lost protein, especially proteins of

high molecular weight. This was shown by darker bands in the non-precipitated samples in the 250 kDa protein region (Figure 23). Protein concentration was determined using the Bradford Assay, in both non-precipitated and precipitated protein samples from tumours #11 and #19 (Figure 24).

Due to the loss of high molecular weight proteins (Figure 23), as well as a reduction in protein concentration (Figure 24), this strategy for the purification of samples was not used for further work. The initial approach, involving repeated centrifugation and careful removal of the fat layer by pipetting was preferred. The associated reduction in sample volume was not apparent enough to affect subsequent experiments.

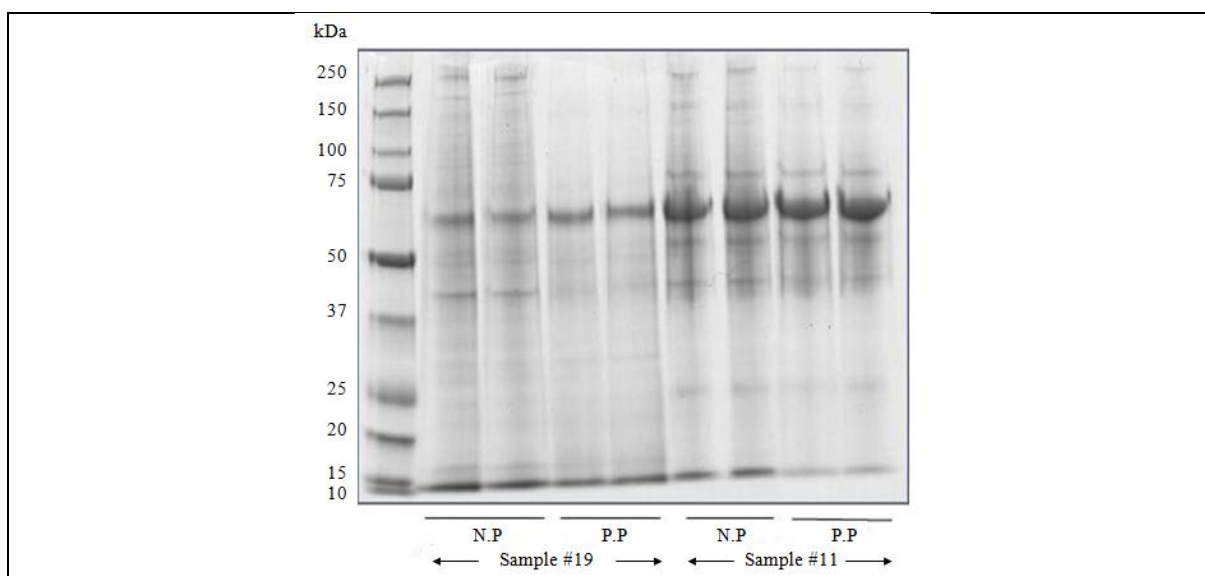


Figure 23: Breast tumour lysate optimisation of protein precipitation

Protein (10 µg) extracted from breast tumour resection tissue (samples #19 and #11; Figure 15) was run on a 1D gel and stained with Bio-safe Coomassie Stain (#161-0787, Bio-Rad). For 200 µl of each protein extract, protein was precipitated (P.P) using the ProteoExtract Protein Precipitation Kit (#539180, Calbiochem) and the rest of the extract remained (N.P). Protein samples which had been precipitated particularly lost high molecular weight proteins, shown by darker bands at 250 kDa in non-precipitated samples. This is quantified in Figure 24.

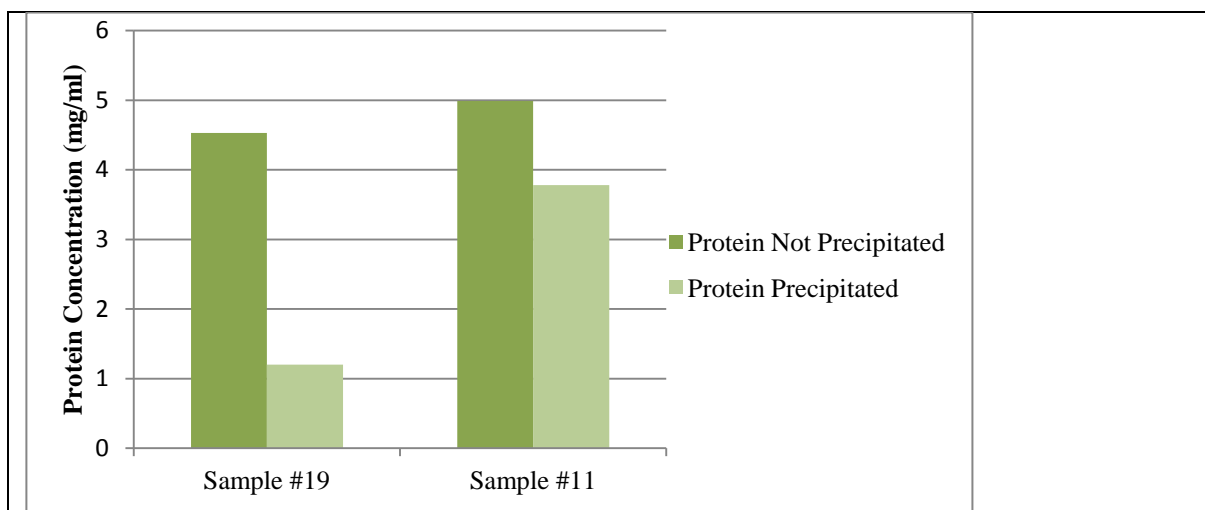


Figure 24: Graph showing protein concentration for precipitated and non-precipitated tumour extracts

Protein concentration (samples from Figure 23) was determined using the Bradford Assay (#B6916, Sigma Aldrich) and calculated for extracts where the protein was precipitated using the ProteoExtract Protein Precipitation Kit (#539180, Calbiochem) and where protein was not precipitated. Precipitation of protein from sample # 19 showed a 73.5% loss of protein. Protein precipitation of sample #11 showed a 24.3% decrease in protein concentration.

6.3.2 Two-dimensional polyacrylamide gel electrophoresis (2D PAGE)

6.3.2.1 Optimisation of extraction method

The tissue extraction method used for antibody microarray analysis involved the mechanical homogenisation of the tissue in ‘Buffer A’ extraction buffer, using a TissueRuptor (Qiagen) (section 4.3.1). However, when using this method with 2D extraction buffer, excess foam was produced due to the presence of detergent, which resulted in sample loss. As an alternative method, samples in 2D extraction buffer were sonicated, in a sonicator water bath, for 15 minutes in total (section 4.4.1). Initially, an ‘optimisation sample’ (Figure 15) was used to determine how much protein could be extracted from a 0.3 g piece of tumour. The protein extraction procedure was performed as described (section 4.4.1), however the 0.3 g tissue sample was extracted with 1 ml of 2D

extraction buffer, in a single microcentrifuge tube. This was then quantified, showing a protein yield of 2.2 mg/ml. Subsequently, 200 µg of protein, for each 2D-PAGE replicate, was cleaned-up using the 2D Cleanup Kit (Bio-Rad). This involved only cleaning up the amount of protein required for 2D-PAGE, thus only using the necessary amount of clean-up kit reagents. This protein extract was separated by 2D PAGE in duplicate and visualised with Bio-Safe Coomassie Stain (#161-0787, Bio-Rad). The resulting images were poor quality due to a lack of protein. The 2 gels were identical, one of which is shown in Figure 25.

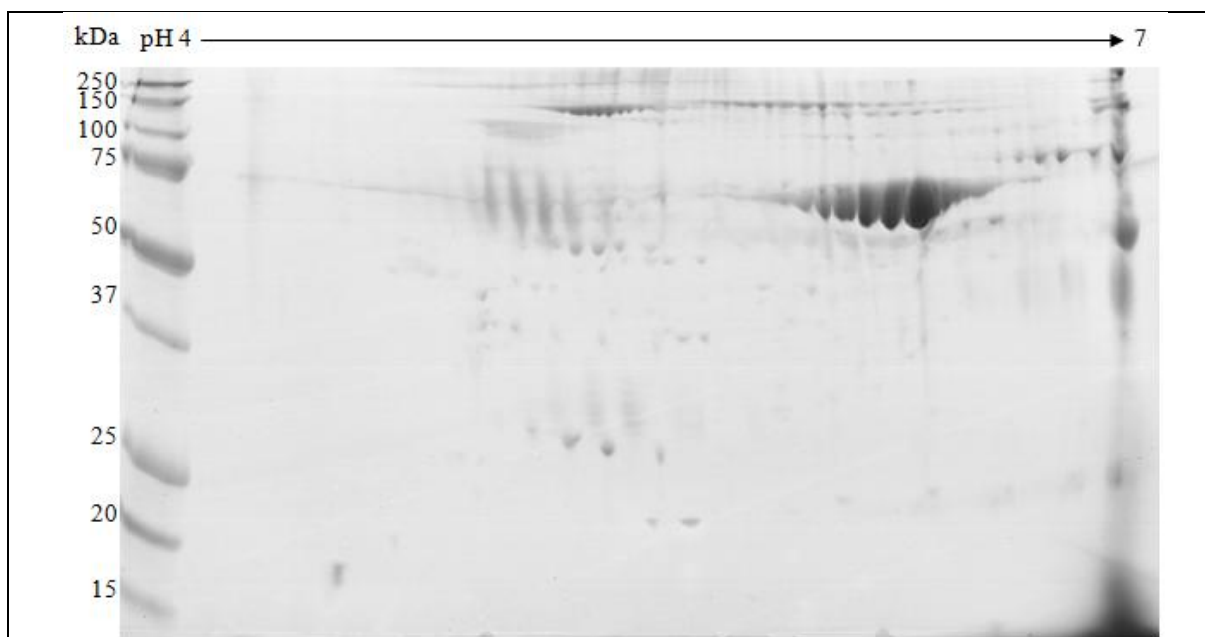


Figure 25: 2D-PAGE with breast tumour tissue extract (insufficient protein)

Breast tumour resection ‘optimisation sample’ (Figure 15) protein extract (200 µg) separated by 2D-PAGE and stained with Biosafe Coomassie Stain (#161-0374, Bio-Rad). This gel is poor quality as insufficient protein was loaded. The method of sample preparation and protein extraction therefore requires further optimisation.

To improve the quality of 2D PAGE gels and ensure sufficient protein load, samples were quantified after clean-up instead of before. However, this required cleaning-up the whole sample, which is a more expensive approach as it requires larger volumes of clean-up kit

reagents. Despite this, protein quantification after cleanup would ensure that 200 µg of protein is accurately loaded for 2D PAGE.

As an attempt to achieve maximal protein yield, during the protein extraction procedure, another ‘optimisation sample’ (Figure 15) was divided between 4 microcentrifuge tubes (containing <0.1 g tissue), and each extracted in 1 ml 2D extraction buffer, with sonication and a 16-hour incubation on an end-over-end rotator at 4 °C (section 4.4.1). The total protein extracted from each sample was re-suspended into 1 microcentrifuge tube with a final volume of 1 ml during the clean-up procedure (section 4.4.2). The sample was quantified (section 4.4.3), and 200 µg was separated by 2D PAGE and stained with Bio-Safe Coomassie Stain (#161-0374, Bio-Rad) (Figure 26).



Figure 26: Breast tumour tissue extract separated by 2D PAGE

Breast tumour resection ‘optimisation sample’ (Figure 15) protein extract (200 µg) separated by 2D-PAGE and stained with Biosafe Coomassie Stain (#161-0374, Bio-Rad).

The successful identification of proteins, from clinical samples separated by 2D PAGE, using MALDI TOF/TOF MS for peptide analysis was confirmed. As a preliminary test, proteins were extracted from a clinical breast tumour resection ‘optimisation sample’ (Figure 15) and were separated by 2D-PAGE (Figure 27). A selection of protein spots (Figure 27) were excised, digested into peptides using trypsin and analysed by MALDI-TOF/TOF MS for subsequent protein identification. The protein identifications, with associated data for the spots shown in Figure 27 are shown in Table 16 (n=5).

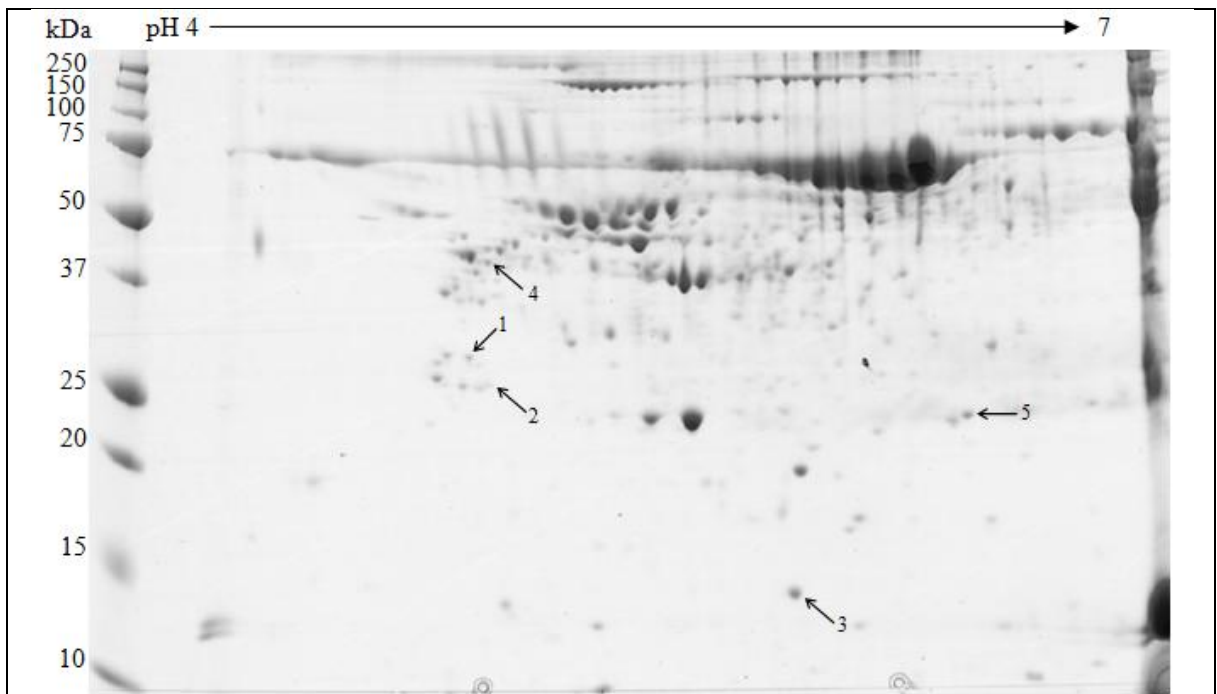


Figure 27: Breast tumour tissue extract separated by 2D PAGE for protein identification using MALDI-TOF/TOF MS

Breast tumour resection ‘optimisation sample’ (Figure 15) protein extract (200 µg) separated by 2D-PAGE and stained with Biosafe Coomassie Stain (#161-0374, Bio-Rad). A selection of protein spots (1-5) were excised from the gel for protein identification by MALDI-TOF/TOF MS.

Table 16: Preliminary test for proteins identification by MALDI-TOF/TOF MS peptide analysis

Peptide samples were analysed using MALDI-TOF/TOF MS analysis.

Spot #	Protein	Estimated mass	Actual mass	pI	Score	# matched peptides	% sequence coverage
1	Topoisomerase alpha-3 chain	27000-30000	29243	4.8	285	6	25
2	14-3-3 gamma	26000-29000	28456	4.8	101	1	5
3	Transthyretin	14000-17000	15991	5.5	290	4	48
4	Vimentin	45000-55000	53676	5.1	212	6	15
5	Heat shock protein beta-1	22000-25000	22826	6.0	286	5	29

6.3.3 Western Blot

6.3.3.1 Molecular weight markers

Molecular weight markers which were in use in the laboratory were Precision Plus Protein Standards Dual Colour (#161-0374, Bio-Rad) and Cruz Marker MW Standards (#SC-2035, SantaCruz). The former appeared on the membrane and did not appear upon film development and the latter did not appear on the membrane yet it appeared after film development. They were therefore used in combination and used two wells on the gel when running a western blot. Cruz Marker Molecular Weight Standard did not appear consistently, which caused inaccuracies when predicting molecular weight of bands and band intensity was too high (Figure 28). Precision Plus Protein WesternC Standards (#161-0376) is a marker that is visible on the gel during electrophoresis, on the nitrocellulose membrane and on the film after development (Figure 28). It is beneficial as it has both properties, which would only use 1 well of the gel when running a western blot.

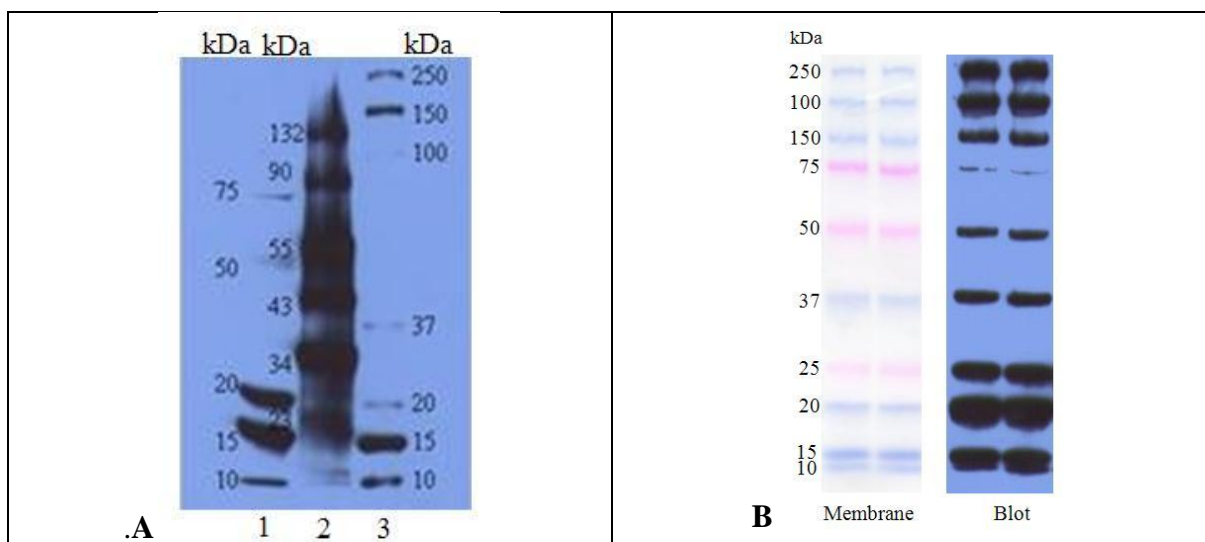


Figure 28: Western Blot Molecular Weight Marker

A Comparison of markers; 1: Precision Plus Protein Standards Dual Color (#161-0374, Bio-Rad) This marker is visible on the membrane; 2: Cruz Marker MW Standard (# SC-2035) Santa Cruz; 3: Precision Plus Protein WesternC Standards (#161-0376, Bio-Rad). Molecular weights are shown in kDa. **B** Precision Plus Protein WesternC Standards (# 161-0376, Bio-Rad). This marker was visible on the gel, the membrane when transferred and on the film when developed, after labelling with Precision Protein StrepTactin-HRP conjugate (#161-0381, Bio-Rad), with a 10 μ l load.

6.3.3.2 Western blotting with tissue samples

The extraction of proteins using Laemmli buffer was limited by the detergent constituent, which caused excess froth to be produced during homogenisation. Another limiting factor was the limited supply of clinical tissue samples available for the different extraction methods. In order to overcome this, protein extracts in antibody microarray buffer (section 4.3.1) were separated by one-dimensional gel electrophoresis, diluted in excess Laemmli buffer, to obtain the required physicochemical state of the protein, for western blotting (section 4.6.3).

6.3.3.3 Antibody Optimisation

In order to analyse protein expression within clinical samples using western blotting, antibodies were tested for their suitability, and optimised. These antibodies included 14-3-3 (specific to beta, eta, tau and sigma isoforms), 14-3-3 epsilon, 14-3-3 zeta, Bcl-xL, BID and beta-actin, to be used as a loading control, which were optimised according to the conditions described in Table 5 (section 4.6.6) (Figure 29).

6.3.4 Determination of protein yield from core biopsy samples

To determine the number of core biopsy samples required to obtain 1 mg/ml of protein, which is the amount required for all proteomic analysis, a series of core biopsy samples were obtained from an 'optimisation sample' of tumour resection (Figure 15). The tissue sample was secured to an 'Easi pad' using a stitch, and a core biopsy gun was used to core biopsy samples from the tumour mass (Figure 30). This was performed by a surgeon.

Protein was extracted from 1, 3 and 6 core biopsy samples, using the antibody microarray extraction buffer 'Buffer A', and the method described (section 4.3.1), and quantified using the Bradford Assay (section 4.3.3). This is shown in Table 17. Based on this single preliminary test, the minimum number of core biopsies required for 1 ml of 1 mg/ml protein would be six core biopsy samples.

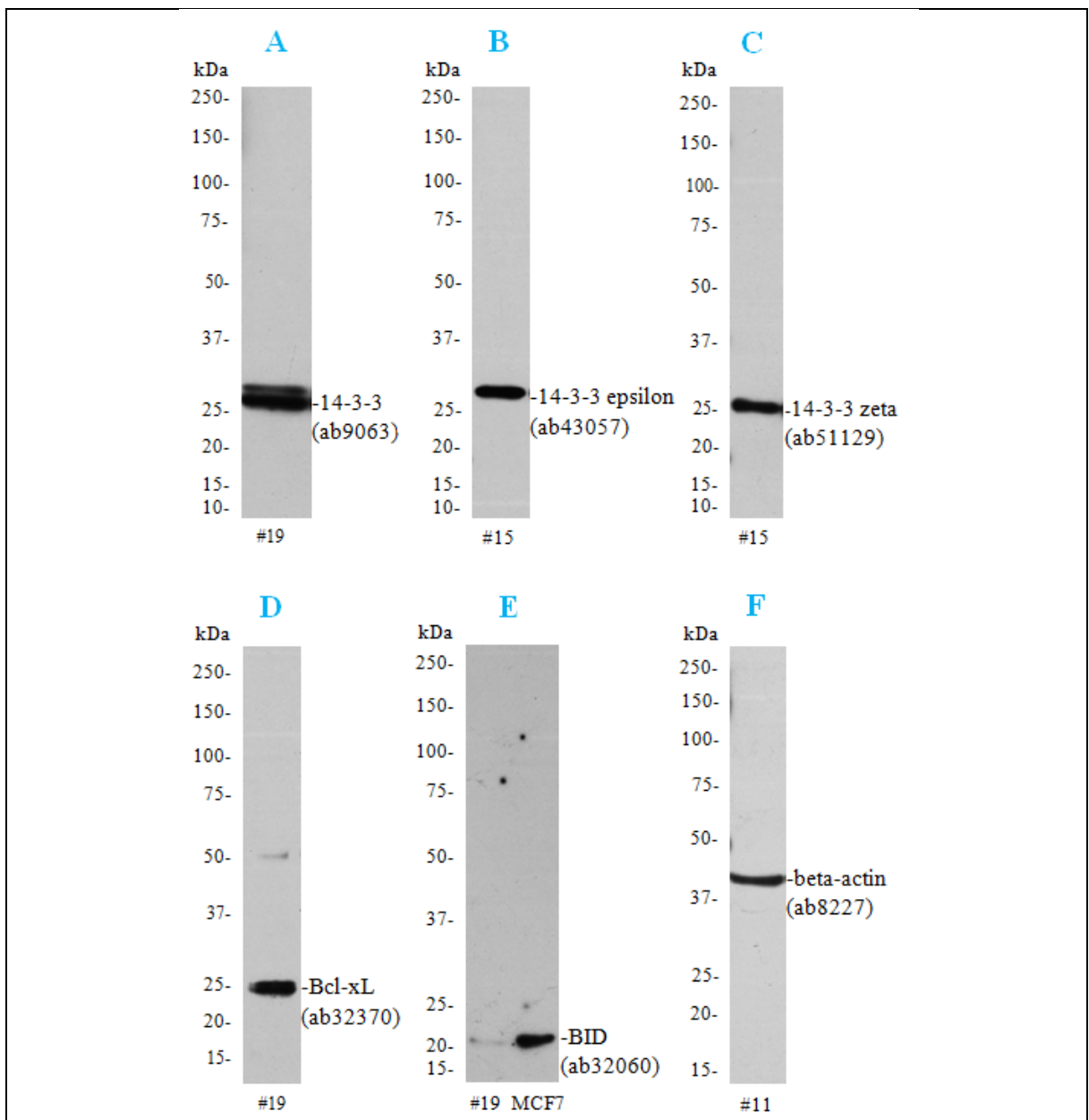


Figure 29: Antibodies optimised for western blotting with clinical samples

All antibodies optimised show a clear band at their expected molecular weights. The concentrations and incubation times for each antibody are shown Table 5 (section 4.6.6). A: 14-3-3 (ab9063) (this is specific to 14-3-3 beta, eta, tau and sigma isoforms, explaining why a single band is not observed) B: 14-3-3 epsilon (ab43057). C: 14-3-3 zeta (ab51129). D: Bcl-xL (ab32370). E: BID (ab32060). F: beta-Actin (ab8227), as a loading control.



Figure 30: Collection of core biopsy samples for determination of protein yield

Core biopsy samples were taken from a tumour resection ‘optimisation sample’ (Figure 15) in order to determine protein yield. The sample was secured to an ‘Easi pad’ using a stitch (A). The core biopsy gun (B), was then used to obtain a core biopsy sample from the tumour mass (C). This was performed by a surgeon.

Table 17: Protein yield from core biopsy samples

Protein was extracted from 1, 3 and 6 core biopsies, which had been taken from a tumour resection ‘optimisation sample’ (Figure 15), in a total volume of 0.5 ml. Samples were quantified using the Bradford Assay, shown as protein concentration in mg/ml. The actual amount of protein obtained is shown (mg). The amount of protein required for all proteomic techniques is 1 mg, at a concentration of 1mg/ml.

Number of core biopsy samples	Protein concentration	Actual amount of protein obtained	Amount of protein required	Sufficient?
1	0.81 mg/ml	0.405 mg	1.00 mg	X
3	1.1 mg/ml	0.550 mg		X
6	2.03 mg/ml	1.015 mg		√

6.4 Discussion

This chapter aimed to achieve the optimisation of established proteomic methodologies for use with clinical tissue samples, for the first time within the laboratory. The sample preparation methods have now been optimised for clinical tumour tissue (approximately 5-10 mm³ in size), for each of the proteomics techniques, including antibody microarray analysis, 2D-PAGE/MS and western blotting.

The chosen sample preparation methods include mechanical homogenisation of tissue using a TissueRuptor, as well as sonication. Protein precipitation during sample preparation, as an attempt to clarify the sample, resulted in loss of protein, especially proteins of high molecular weight; therefore this was not selected for future work. Instead, repeated centrifugation was performed to clarify the sample by removal of fat. The lack of correlation between sample size and weight is a result of tissue heterogeneity, so can be expected for the type of clinical tissue samples being used.

Depletion strategies have not been tested at this stage; however this is an option to consider for future work. Other techniques such as LCM have been considered, however due to low sample availability, and the potential for detrimental effects such as protein degradation or modification, as well as effects on downstream proteomic techniques being employed, such as IEF, this was considered to be inappropriate.

These proteomic platforms are therefore now ready for comparative proteomic experiments using clinical tissue samples (chapter 8, chapter 9 and section 10.3.3).

As a consideration for future work, the protein yield from core biopsy samples was also assessed. If predictive biomarkers of chemotherapy resistance were transferred to the clinic, screening would be required at the diagnosis stage, where core biopsy samples are taken. In order to identify predictive biomarkers, the most clinically relevant sample to analyse would therefore be core biopsy samples. However, obtaining core biopsy samples is an invasive procedure, which causes discomfort and is therefore unpleasant for patients. It is therefore unethical to ask for multiple cores to be taken for research purposes at this stage. In order to obtain ethical approval for the collection of core biopsy samples, optimisation of methodologies is required so that analysis could be performed on the lowest number of cores possible. In order to assess the feasibility of performing proteomic analysis on core

biopsy samples, and to determine how many core biopsy samples would be required, a single preliminary test has been performed. This demonstrated that to obtain sufficient protein for the current proteomics methods, 6 core biopsy samples would be required, which would be regarded as unethical if requested from a single patient. In this situation, sample pooling could be considered, as well as advancing methodologies and increasing their sensitivity, therefore requiring a fewer number of core biopsy samples. However, in order to fully assess this, further preliminary testing, with replication would be required, yet the possibility of performing proteomic analysis on core biopsy samples does show promise. The number of core biopsies required could be reduced by making appropriate alterations to proteomics methods; such as the use of 2D-DIGE, which would require smaller protein samples (section 3.3.2). The amount of protein required for antibody microarray analysis is small (0.15 mg), yet the current manufacturer's recommended process for labelling the proteins with fluorescent dyes requires 1 ml of protein extract at 1 mg/ml (1 mg). The labelling procedure is therefore more demanding of protein concentration than the antibody microarray platform. For future work, the possibility of reducing the amount of protein required for the labelling process should therefore be considered.

CHAPTER 7:
REPEATEDLY IDENTIFIED
DIFFERENTIALLY EXPRESSED PROTEINS
(RIDEPS) FROM ANTIBODY MICROARRAY
ANALYSIS

Chapter Aim:

To analyse multiple antibody microarray experiments, to identify any RIDEPS in the data
and to establish guidelines for quality control and fold change cut-off values

Hodgkinson, V.C., ELFadl, D., Drew, P.J., Lind, M.J., Cawkwell, L. (2011) Repeatedly identified differentially expressed proteins (RIDEPS) from antibody microarray proteomic analysis. *Journal of proteomics*, 74: 698-703

Chapter 7. Repeatedly identified differentially expressed proteins (RIDEPs) from antibody microarray proteomic analysis

7.1 Introduction

Forward-phase antibody microarrays are a powerful new tool in the field of comparative proteomics and the search for clinical biomarkers (Kopf, Shnitzer et al. 2005; Kopf and Zharhary 2007). They offer rapid expression analysis of multiple proteins in a sample simultaneously, via incubation with antibodies immobilised on a glass slide. Some of the problems associated with gel-based and mass spectrometry (MS)-based proteomics methods can be overcome using antibody microarray analysis and this may therefore be considered a valuable complementary proteomic technique (Alhamdani, Schroder et al. 2009). The search for protein biomarkers generally consists of three phases; *discovery*, where MS- or microarray-based comparative proteomic approaches may be used to generate long lists of differentially expressed proteins (DEPs); *confirmation*, where additional techniques such as western blotting or ELISA are used to confirm the differential expression of candidate proteins from the discovery phase, and *validation* where putative biomarkers are evaluated in the clinical context (Rifai, Gillette et al. 2006; Paulovich, Whiteaker et al. 2008). The success of this biomarker discovery pipeline therefore relies heavily on the quality of the data generated in the discovery phase and the selection of DEPs carried forward to the confirmation phase. This is complicated by the fact that a number of candidate proteins generated by any proteomics approach may be false-positive results (Rifai, Gillette et al. 2006; Paulovich, Whiteaker et al. 2008). It is therefore

important that data generated in the discovery phase is interpreted carefully, and confirmation and validation of differential protein expression is carried out rigorously.

Recent meta-analysis studies have reported a number of repeatedly identified differentially expressed proteins (RIDEPs) from MS-based proteomics research carried out in a range of different species and tissues (Petрак, Ivanek et al. 2008; Mariman 2009; Ponomarenko, Lisitsa et al. 2009; Wang, Bouwman et al. 2009). Petрак *et al* produced a dataset containing the identity of DEPs identified from 2D-PAGE/MS studies using human and rodent tissue, published within 3 volumes of 'Proteomics' (volumes 4-6; 2004-2006). The dataset contained proteins identified from experiments using total cellular homogenates, and excluded experiments which used bodily fluids, subfractionated tissues and tissue culture supernatants (Petрак, Ivanek et al. 2008). The appearance of each protein within the dataset was calculated, and this did not include the direction of differential expression of each protein (increase or decrease). The total number of proteins within the dataset was 4700, from 169 published articles comprising 186 experiments, which each reported an average of 25 DEPs. Ninety-nine of the 186 articles involved the study of human cells, comprising 108 individual experiments, 70 of which were cancer-related (Petрак, Ivanek et al. 2008). Based on the 186 sets of data generated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled to MS methods Petрак *et al* (2008) identified the top 15 RIDEPs from studies using human and rodent samples (Petрак, Ivanek et al. 2008). The most frequently identified RIDEPs from human tissue were heat-shock protein 27 and enolase 1, which were identified in 31% and 29% of experiments respectively (Table 18). Other protein families identified repeatedly were the keratins, annexins and peroxiredoxins (Petрак, Ivanek et al. 2008).

Table 18: ‘TOP 15’ RIDEPs identified from 2D-PAGE/MS experiments

The table lists the ‘TOP 15’ RIDEPs identified from 2D-PAGE/MS data using human samples (Petrak, Ivanek et al. 2008)

TOP 15 Human RIDEPs			
Protein	Identified in percentage of experiments (%)	Protein	Identified in percentage of experiments (%)
HSP27	31	HSC71	17
Enolase 1	29	Peptidyl-prolyl isomerase A	16
Triosephosphate isomerase	20	Cytokeratin 8	16
Pyruvate kinase M1/M2	19	Cathepsin D	15
Peroxiredoxin 1	19	ATP synthase beta subunit	14
Peroxiredoxin 2	19	Grp78/Bip	13
Vimentin	19	RhoGDI 1	13
Annexin A4	18		

These findings were supported by Wang *et al* (2009) using data from 66 biologically different studies encompassing *in vivo* and *in vitro* experiments on 20 tissue types from 5 species where DEPs were identified using 2D-PAGE coupled with MS analysis (Wang, Bouwman et al. 2009). A list of 44 RIDEPs was generated and this included 73% of the top 15 RIDEPs from both human and rodent tissue previously identified by Petrak *et al* (2008) (Petrak, Ivanek et al. 2008; Mariman 2009; Wang, Bouwman et al. 2009). More recently, a meta-analysis to identify colorectal cancer-associated proteins from published proteomics studies, largely based on 2D-PAGE/MS, generated a list of RIDEPs (Jimenez, Knol et al. 2010). This list of putative colorectal cancer-associated proteins included a number of RIDEPs previously reported by Petrak *et al* (2008) and Wang *et al* (2009) (Petrak, Ivanek et al. 2008; Mariman 2009; Wang, Bouwman et al. 2009). It has been hypothesised that RIDEPs may be related to the cellular stress response and should be treated with caution in

the selection of proteins for the confirmation stage of the biomarker discovery pipeline (Mariman 2009; Wang, Bouwman et al. 2009).

The studies presented above have generated lists of RIDEPs from MS-based proteomic studies and, currently, there are no reports of such proteins from non MS-based proteomic studies. The identification of any RIDEPs based on data from antibody microarray experiments, which provide a valid complementary approach to MS-based proteomics, is essential for the informed decision-making process in the verification phase of putative biomarkers revealed by this method.

7.2 Materials and Methods

7.2.1 Panorama Antibody Microarray XPRESS Profiler725 Kit

Five experimental sample groups encompassing a range of oncology-related research on human tissue, cells or cell lines were analysed using the Panorama Antibody Microarray XPRESS Profiler725 Kit (XP725, Sigma Aldrich), which is able to identify the differential expression of a wide range of proteins including those involved in apoptosis, cell signalling, cell cycle, cell adhesion and proliferation. The 725 antibodies are listed in Appendix 1. Biological replicates within these 5 sample groups generated a total of 13 experiments. Some of the experimental work, in experimental sample groups 1-4, was performed by others, from 2007 onwards using the same protocol, but all analysis was retrospectively performed personally (section 7.2.9).

7.2.2 Experimental sample group 1: Stimulation of the B-cell receptor in malignant B-cells from patients with chronic lymphocytic leukaemia (n=1)

Approval was obtained from the Hull and the East Riding Local Research Ethics Committee (ref 05/Q1104/33) for the study entitled ‘chronic lymphoproliferative disorders,

factors regulating the survival of mature malignant B cells' and a blood sample was collected with informed consent from a patient with chronic lymphocytic leukaemia (CLL). In order to identify DEPs following activation of the B-cell receptor, peripheral blood mononuclear cells were isolated from whole blood and the B-cell receptor was artificially stimulated by cross-linking the BCR for 5.5 hours with goat anti-human IgM antibody (#109-006-129, Jackson ImmunoResearch Laboratories Inc) at a final concentration of 10ug/ml. Control (unstimulated) cells for comparison were produced using the relevant isotype control (#005-000-006, Jackson ImmunoResearch Laboratories Inc). Cells were washed twice in 5 ml cold PBS before protein extraction was achieved by suspending cells in 1ml of 'Buffer A' (section 4.3.1). Samples were incubated on an end-over-end rotator for 5 min at 4°C and subsequently centrifuged at 10,000 x g for 2 min, after which the supernatant was retained. This experimental work was performed by Dr Gina Eagle.

7.2.3 Experimental sample group 2: Treatment of human lung cancer and mesothelioma cell lines with a COX-2 inhibitor (n=2)

In order to identify DEPs following treatment of human cells with the specific COX-2 inhibitor DuP-697 two cell lines were selected. COX-2 positive lung cancer cells (A549) and malignant pleural mesothelioma cells (MSTO-211H) were exposed to DuP-697 (#1430, Tocris Bioscience) for 72 hours using a dose based on the IC₅₀ which was previously determined for each cell line (O'Kane, Eagle et al. 2010). Control (untreated) cells for comparison were produced using the drug carrier dimethyl sulfoxide alone. Cells were washed twice in 5 ml of cold PBS before protein extraction was performed as above (section 4.3.1). This experimental work was performed by Dr Gina Eagle.

7.2.4 Experimental sample group 3: Biomarkers of the radiotherapy-resistant phenotype in human breast cancer cell lines (n=3)

In order to identify DEPs associated with a radiotherapy resistant phenotype in human breast cancer cells three novel radio-resistant cell line derivatives MCF7RR, MDA-MB-231RR and T47DRR (Smith, Qutob et al. 2009) were compared with the respective parental (radio-sensitive) cells. Protein extraction was performed as above (section 4.3.1). The experimental work was performed by Dalia ELFadl.

7.2.5 Experimental sample group 4: Biomarkers of the radiotherapy-resistant phenotype in human oral cancer cell lines (n=2)

In order to identify DEPs associated with a radiotherapy resistant phenotype in human oral cancer cells two novel radio-resistant cell line derivatives were established using PE/CA-PJ41 and PE/CA-PJ49 parent cells following the protocol described previously (Smith, Qutob et al. 2009). Protein extraction was performed for radio-resistant and parental (radio-sensitive) cells as above (section 4.3.1). Experimental work was performed alongside Dalia ELFadl.

7.2.6 Experimental sample group 5: Biomarkers of chemotherapy response in human breast cancer tissue (n=5)

Ethical approval was obtained (REC 07/Q1105/43) and breast cancer samples were collected with informed consent from patients who had received neoadjuvant chemotherapy (epirubicin/cyclophosphamide followed by docetaxel) for locally advanced breast cancer (section 4.2). Samples were snap-frozen in liquid nitrogen and stored at minus 80°C until required. Chemotherapy-sensitive and chemotherapy-resistant samples were identified by calculation of the extent of tumour response to chemotherapy using the Response

Evaluation Criteria In Solid Tumours guidelines (Therasse, Arbuck et al. 2000; Eisenhauer, Therasse et al. 2009). In order to identify DEPs associated with chemotherapy response in ER-positive ductal breast cancers 5 sample pairs (Table 21) were analysed, each comparing a chemotherapy-sensitive and a chemotherapy-resistant sample. For protein extraction the tissue was weighed, dissected into small pieces and washed in cold phosphate buffered saline. It was then transferred to 4 volumes (w/v) of Buffer A and mechanically homogenised on ice using a TissueRuptor (Qiagen Ltd). The extract was then centrifuged until a transparent supernatant was obtained (section 6.3.1) (Figure 21).

7.2.7 Protein Labelling

Protein Concentrations were determined using the Bradford Assay (#B6916, Sigma Aldrich) (section 4.3.3). In all experiments, proteins from control (untreated/sensitive) samples were labelled with Cy3 (#PA23001, GE Healthcare) and proteins from test (treated/resistant) samples were labelled with Cy5 (#PA25001, GE Healthcare) fluorescent dyes according to the manufacturers protocol (section 4.3.4). Protein extracts were diluted to 1 mg/ml in 'Buffer A', and labelled by incubating 1 ml of the extract with the respective dye vials for 30 min at room temperature in the dark. During this time, vials were vortexed every 10 min. Un-bound dye was removed from the samples using SigmaSpin columns, leaving only dye-protein complexes.

7.2.8 Protein Binding

Prior to protein binding, dye-to-protein (D/P) molar ratios were calculated for each sample, and as recommended in the kit protocol, only samples with a D/P ratio >2 were carried forward (section 4.3.5). In a darkened room, equal amounts (50-150 μ g) of labelled protein from each sample pair was mixed with array incubation buffer (provided), applied to the

array slide and incubated 45 min in a quadriPERM cell culture vessel on an orbital shaker at low speed. Following this, the slide was washed three times in wash buffer (provided) for 5 min and once in ultrapure water for 2 min under the same conditions. The slide was then allowed to air-dry for 30 min (section 4.3.6).

7.2.9 Image acquisition and analysis

The hybridised antibody microarray slides were scanned using a GenePix Personal 4100A Microarray Scanner (Axon Instruments) with 532 and 635 nm lasers. GenePix Pro software (Axon Instruments) was used to grid the antibody microarray slide and to apply protein names in the form of a list with their respective location on the slide. Negative controls on the antibody microarray slide were flagged as 'absent', and all antibody/protein spots were manually edited (section 4.3.7). This manual editing process took approximately 6 hours for each antibody microarray slide. Acuity software (Axon Instruments) was then used to identify differentially expressed proteins. Editing and analysis of all 13 antibody microarray slides, from the 5 experimental sample groups described was carried out personally, to produce 13 sets of data which had been analysed by the same individual. Data normalisation was carried out based on the Lowess method, and spot quality control criteria were applied to only include spots with <3% saturated pixels, those which were not flagged as absent and those that had relatively uniform intensity and were detectable above the background (section 4.3.7). Experiments which showed 'percentage substances matched' values of ≥ 90 were carried forward (section 7.4), which indicates how many of the pairs of antibody spots were recognised and therefore reflects the quality of the slide and the quality of the experiment. Calculated log ratios for the relative expression of each protein were converted into fold changes.

7.3 Results

Antibody microarray data was obtained from 13 individual experiments using the XPRESS Profiler725 assay to identify DEPs from a range of projects utilising human tissue, cells or established cell lines. Analysis of all antibody microarray slides was carried out by the same individual (VH) to reduce inter-observer variability. Experiments were considered successful only when the percentage of 'substances matched', provided by the software during analysis, was $\geq 90\%$ thus ensuring that good quality data had been produced for a high proportion of the 725 antibodies on the microarray in each experiment. Based on the data generated from all 13 experiments it was decided that a fold change of ≥ 1.8 would be considered significant. In addition, fold changes ≥ 1.5 were also recorded for each experiment but this data was only considered as supporting evidence if there was other data indicating that the protein was a significant DEP. A total of 13 RIDEPs were seen, each appearing in at least 4/13 (30%) antibody microarray analyses from at least 2 experimental sample groups (Table 19).

Table 19: RIDEPs from antibody microarray analysis

A total of 13 different antibody microarray assays were performed using 13 samples derived from 5 experimental sample groups. The Table shows RIDEPs seen in at least 4/13 (30%) antibody microarray analyses from at least 2 experimental sample groups. Significant expression fold-change (≥ 1.8) is indicated in bold. For proteins which show ≥ 1.8 -fold change in expression, supporting data from other experiments is shown upward of 1.5-fold. Values below 1.5 were considered to be not significant (---). Any proteins which did not pass the analysis criteria for experimental quality control are indicated as \otimes .

Protein (Ab #)	Experimental sample group #												
	#1 BCR stimulation (CLL cells)	#2 COX-2 inhibitor (lung cancer and mesothelioma cell lines)		#3 Radio-resistance (breast cancer cell lines)			#4 Radio-resistance (oral cancer cell lines)		#5 Chemo-resistance (breast cancer tissue)				
	n=1	n=2		n=3			n=2		n=5				
	003	A549	MSTO- 211H	MDA- MB- 231RR	T47DRR	MCF7RR	PE/CA- PJ41	PE/CA- PJ49	11/19	15/9	15/19	12/25	18/25
Zyxin (#Z0377)	3.14	4.39	4.74	3.07	2.99	1.70	---	1.54	7.80	2.01	2.21	2.02	2.63
BID (#B3183)	3.05	2.42	2.54	---	---	---	2.05	2.33	---	2.16	1.55	1.97	1.96
MyD88 (#M9934)	4.53	---	---	---	2.08	2.02	---	---	---	---	2.08	2.18	---
IKKa (#I6139)	---	2.37	1.99	---	---	---	---	2.29	1.61	---	---	2.03	---
Bcl-xL (#B9429)	---	2.13	2.24	---	---	---	---	---	---	2.26	---	1.57	2.62
Chondroitin sulphate (#C8035)	---	1.85	1.54	---	1.79	---	2.30	1.98	1.56	---	2.00	---	---
14-3-3 theta/tau (#T5942)	---	1.64	2.22	---	---	---	---	---	1.54	1.90	2.29	2.55	1.52
Centrin (#C7736)	2.00	2.51	2.04	---	---	---	---	---	1.51	---	---	1.90	---
SLIPR MAGI3 (#S4191/#S1190)	3.22	2.15	1.92	---	---	---	1.84	1.77	---	---	---	---	---
Pinin (#P0084)	\otimes	7.67	7.32	\otimes	\otimes	\otimes	---	---	\otimes	2.54	2.39	1.53	---
Protein kinase C (#P5704)	2.10	---	---	1.92	---	1.91	2.10	---	---	---	---	---	---
Smad4 (#S3934)	3.56	1.87	1.61	1.81	---	1.96	---	---	---	---	---	---	---
Siah2 (#S7945)	5.22	---	---	1.92	2.05	1.50	2.10	\otimes	---	---	1.66	---	---

7.4 Discussion

7.4.1 Quality control

The percentage of ‘substances matched’ by the software on the microarray slide is indicated following analysis. This represents how many of the ‘substances’ (pairs of antibodies spotted on the slide) were detected after passing the quality control criteria which were applied, thus determining the overall quality of the experiment. There are currently no recommendations in the literature regarding thresholds for quality control for antibody microarray analysis, and no evidence of other studies comprising this number of experiments (n=13), which have been analysed by the same individual. Therefore, based on personal experience, due to the extent of antibody microarray data analysis which has been performed, a quality control threshold is being proposed. After analysing multiple antibody microarray experiments, the proposed level of ‘% substances matched’ is $\geq 90\%$. This would ensure that only slides which are of high quality are considered for result interpretation. Any slides which fall below this bench-mark should be carefully investigated, especially for spot quality, spot morphology and non-specific dye background problems.

7.4.2 Fold change cut-off

Following the convention in expression microarray analysis, a 2-fold change in expression has previously been utilised to indicate a significant DEP using antibody microarray platforms (Ghobrial, McCormick et al. 2005; Smith, Qutob et al. 2009; Wu, Wang et al. 2010). The decision to pass a DEP into the verification stage of the biomarker discovery pipeline is usually based on reaching a threshold level of significant fold change. The fold change cut-off employed when using the XPRESS Profiler725 antibody microarray has

ranged from ≥ 1.5 -fold to ≥ 2.5 -fold (Mohri, Mohri et al. 2009; Uemura, Nakanishi et al. 2009; Wu, Wang et al. 2010). Here, 13 individual data-sets have been assessed and it was noticed that proteins which were appearing in multiple experiments within a sample group, were occasionally falling short of the 2-fold threshold. Fold changes in expression ≥ 1.8 have therefore been considered to be significant, to ensure potentially interesting proteins are not overlooked. In addition fold changes ≥ 1.5 were also recorded, as supporting evidence, for each protein once a value of ≥ 1.8 was reached in one experiment. This ensures that all fold change data of ≥ 1.5 is considered for each set of biological replicates before prioritisation decisions are made for the confirmation stage.

7.4.3 RIDEPS

A cluster of approximately 900 genes has been reported in yeast which responds to stressful environmental changes as a protective mechanism, usually where conditions become sub-optimal (Gasch, Spellman et al. 2000). The aim of the cellular stress response is to protect against adverse environmental conditions that may perturb cell homeostasis and proteins which are responsible for the cellular stress response are highly conserved across the three superkingdoms *Archaea*, *Bacteria* and *Eukarya* (Kültz 2003). Approximately 300 conserved proteins constitute the ‘minimal stress proteome’ (Kültz 2005). Of the 300 proteins, 44 known functional proteins have been identified which are ubiquitously conserved in all 3 superkingdoms (Kültz 2005). The functional roles of these proteins include redox regulation, energy metabolism and the DNA damage/repair response, processes which are all essential for management of stress and maintenance of cell homeostasis. The enolase, GAPDH and peroxiredoxin proteins feature amongst the 44 which have been listed as part of the minimal stress proteome and this formed the basis of the hypothesis that the RIDEPS identified by Petrak *et al* (2008), which included these

proteins, appeared due to a cellular stress response (Petрак, Ivanek et al. 2008; Mariman 2009). These lists of RIDEPs should be considered when interpreting proteomics data from the discovery phase.

Antibody microarray analysis has identified 13 RIDEPs which were seen in at least 4/13 (30%) XPRESS Profiler725 antibody microarray analyses from at least 2 experimental sample groups. None of these proteins have previously been reported as RIDEPs in 2D-PAGE/MS-based experiments (Petрак, Ivanek et al. 2008; Mariman 2009; Wang, Bouwman et al. 2009) and this may be due to the technical differences between the platforms, which produce the complementary nature of proteomic techniques, (Smith, Qutob et al. 2009), sample differences or cut-off values employed.

The most frequently identified protein was zyxin which demonstrated significant differential expression in 10/13 (76%) experiments from 4/5 sample groups. Zyxin is a focal adhesion LIM domain protein involved in maintenance of actin stress fibres and apoptotic signalling (Hervy, Hoffman et al. 2010; Smith, Blankman et al. 2010). Its possible role in homeostasis and the position as the top RIDEP in these antibody microarray experiments has resulted in this putative biomarker being treated with caution since the differential expression of zyxin may be a result of cellular stress. Other proteins identified frequently across all experiments include BID, MyD88, BclxL, 14 3 3 theta/tau, SLIPR MAGI3, Protein Kinase C, Smad4 and Siah2 which have all been linked with a functional role in cell apoptosis. There are many antibodies against apoptosis-related proteins on the XPRESS Profiler725 antibody microarray slide and this family of proteins may be anticipated to produce candidate biomarkers for the types of experimental sample groups which have been presented, however these RIDEPs would need careful verification. In recent publications which have also used the XPRESS Profiler725 antibody microarray

for human cell or tissue research (Mohri, Mohri et al. 2009; Uemura, Nakanishi et al. 2009; Wu, Wang et al. 2010) 3 of these RIDEPs (BID, MyD88 and BclxL) were also reported in at least 1 study.

7.5 Conclusions

The phenomenon of RIDEPs may exist not only in mass spectrometry-based proteomic experiments, but also in antibody microarray proteomics and a preliminary list of 13 RIDEPs has been produced from the XPRESS Profiler725 antibody microarray platform for the first time. This information will be useful when interpreting experimental data and considering which DEPs should be prioritised for verification.

CHAPTER 8:

ANTIBODY MICROARRAY ANALYSIS FOR THE IDENTIFICATION OF BIOMARKERS OF CHEMOTHERAPY RESISTANCE

Chapter Aim:

To use the antibody microarray proteomics platform for the identification of putative biomarkers of neoadjuvant chemotherapy resistance in breast cancer using clinical tumour tissue samples, forming part of the discovery phase of the biomarker discovery pipeline.

Hodgkinson, V.C., ELFadl, D., Russell, C., Agarwal, V., Garimella, V., Drew, P., Lind, M.J., Cawkwell, C. Predictive biomarkers of chemotherapy resistance in breast cancer: a possible role for 14-3-3 tau and Bid? – *submitted*

Chapter 8. Antibody microarray analysis for the identification of biomarkers of chemotherapy resistance

8.1 Introduction

This chapter involves the examination of luminal (ER+) breast cancer samples from patients treated with neoadjuvant chemotherapy in order to identify proteins which may be associated with therapy resistance. In order to achieve this, the antibody microarray platform was employed to assess the differential expression of 725 protein targets simultaneously per experiment, forming part of the ‘discovery’ phase of the biomarker discovery pipeline (section 3.1 and section 3.5). The quality-control thresholds, fold-change cut-off values and RIDEPs associated with the antibody microarray platform were outlined in Chapter 7. The results presented in this chapter include the same five pairs of breast cancer samples which were described in Chapter 7 for the identification of RIDEPs.

8.1.1 Antibody microarrays

Antibody microarrays offer a complementary approach to techniques such as 2D-PAGE / MS for the discovery of biomarkers, and have the ability to identify the differential expression of multiple proteins simultaneously, between two samples (section 3.5). They consist of small amounts of antibodies, arranged on a solid support such as a glass microscope slide, which is coated with a substrate such as nitrocellulose (Alhamdani, Schroder et al. 2009). There are several commercially available antibody microarrays, an example of which is the Panorama[®] Antibody Array – XPRESS Profiler725, by Sigma Aldrich. This platform contains 725 antibodies (listed in Appendix 1) specific to a variety of proteins involved in important biological pathways, such as cell signalling, gene regulation and

apoptosis, and is able to detect protein levels as low as nanograms per millilitre (Kopf, Shnitzer et al. 2005). Antibodies on the Panorama[®] Antibody Array platform are robotically spotted onto nitrocellulose-coated slides under controlled conditions, after which the slide is blocked with a specific proprietary blocking buffer to minimise background staining (Kopf, Shnitzer et al. 2005). The most common detection method currently used for antibody microarrays is the fluorescence-based detection method (Pavlickova, Schneider et al. 2004; Borrebaeck and Wingren 2009). This involves the labelling of proteins with fluorescent dyes, thus allowing protein expression to be determined when it binds to the corresponding antibody on the slide. Each sample is labelled with a different fluorescent dye, commonly Cy3 and Cy5 dyes (Gu, Sivanandam et al. 2006). An example of such dyes are the cyanine Cy3 and Cy5 mono-functional N-hydroxysuccinimidyl ester (NHS-ester) dyes (#PA23001 and #PA25001, GE Healthcare) which label proteins by binding to free amino groups on lysine amino acid residues. The array slide is scanned at two wavelengths, typically 532 nm and 635 nm, to analysis Cy3- and Cy5-labelled samples respectively. These two images are then combined to produce a ratio image (Figure 31). The relative intensity of each dye on each antibody spot can be used to determine relative expression of a specific protein in each sample. The differential expression of a protein is determined when a significant fold change in expression is observed between the two samples; in this chapter, a fold change of ≥ 1.8 is deemed significant (section 7.4.2). Following analysis of each spot on the antibody microarray, a list of differentially expressed proteins is generated.

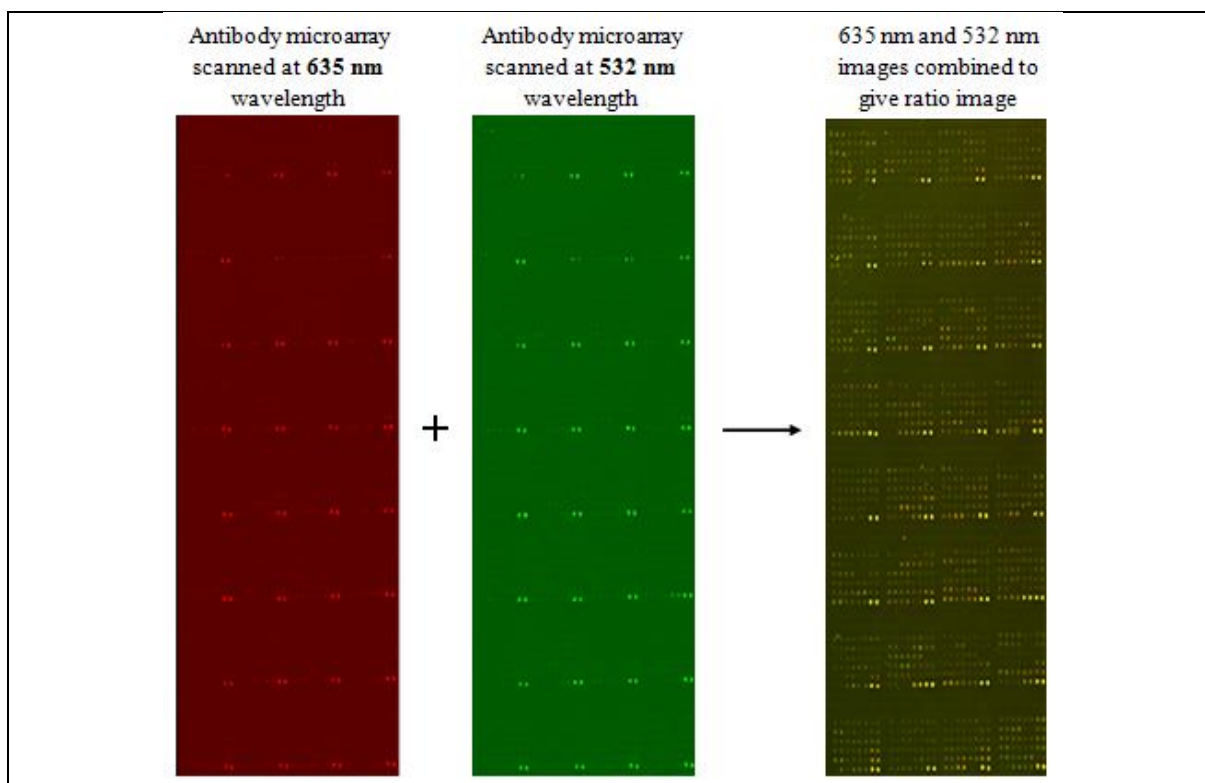


Figure 31: Scanning of antibody microarray slides

When using the fluorescence-based detection method of protein expression, where proteins are labelled with Cy3 and Cy5 dyes, the array slide is scanned at two different wavelengths (532 nm and 653 nm). The two images are then combined, to give a ratio image from which the relative intensity of each dye can be determined for each antibody. Where the difference in relative intensity of each dye is ≥ 1.8 -fold, this represents the significant differential expression of this protein.

8.2 Methods

8.2.1 Fresh tumour samples

Following Research Ethics Committee approval (ref 07/Q1105/43) patients receiving neoadjuvant chemotherapy for locally advanced breast cancer at Castle Hill Hospital in Hull were identified. The standard treatment regimen consisted of 4 cycles of 90 mg/m² epirubicin with 600 mg/m² cyclophosphamide (EC) followed by 4 cycles of 100 mg/m² docetaxel given at three-weekly intervals prior to surgical resection. All patients completed at least 6 cycles of chemotherapy unless early surgical intervention was indicated due to

disease progression. Following surgical resection a fresh sample of macroscopic residual breast cancer was immediately snap-frozen in liquid nitrogen and stored at minus 80°C until required. All samples selected for analysis were ductal, ER positive (luminal) tumours. Considering Response Evaluation Criteria In Solid Tumours guidelines (Therasse, Arbuck et al. 2000; Eisenhauer, Therasse et al. 2009), clinical response was determined by a specialist in Oncology using pre-treatment dynamic contrast enhanced-magnetic resonance imaging (DCE-MRI) data, post-treatment histopathological measurement of residual primary tumour and clinical observation of metastasis during therapy. For one case where DCE-MRI data was not available, pre-treatment and post-treatment ultrasound scans were compared for tumour measurements. For the purpose of this study, samples were designated as ‘chemotherapy sensitive’ (CS; samples #11^{CS}, #12^{CS}, #15^{CS} and #18^{CS}) where a partial response (at least 30% decrease in primary tumour size) was seen. Samples were designated as ‘chemotherapy resistant’ (CR; samples #9^{CR}, #19^{CR} and #25^{CR}) where disease progression was observed during therapy and/or where there was no significant decrease in primary tumour size. These 7 clinical samples (Table 20) were selected based upon sample availability, and according to the amount of protein each sample yielded, they were used to create five pair-wise combinations (Table 21), in order to maximise the data generated.

Table 20: Clinical samples selected for antibody microarray analysis

Seven clinical samples were selected, which were all ductal tumours of luminal subtype (ER+). Tumour response was determined based on tumour size pre-and post-neoadjuvant chemotherapy treatment and categorised as chemotherapy-sensitive^{CS} or chemotherapy-resistant^{CR} by a specialist oncologist.

<i>Sample #</i>	<i>Hormone Status</i>	<i>Tumour response</i>
9^{CR}	ER+ PR+ HER2+	8% size increase. Stable Disease
11^{CS}	ER+ PR+ HER2-	40.5% size reduction. Partial Response
15^{CS}	ER+ PR+ HER2-	69.2% size reduction. Partial Response
18^{CS}	ER+ PR+ HER2-	78.1% size reduction. Partial response
19^{CR}	ER+ PR- HER2-	71.4% size increase. Progressive disease
25^{CR}	ER+PR+HER2+	26.9% size reduction. Stable Disease

Table 21: Pair-wise combinations of samples used for 5 antibody microarray experiments

Five experiments were performed, where protein expression was compared between chemotherapy-sensitive samples^{CS} and chemotherapy-resistant samples^{CR}.

<i>Experiment #</i>	<i>Chemotherapy-sensitive^{CS} sample</i>	<i>Chemotherapy-resistant^{CR} sample</i>
1	#11 ^{CS}	#19 ^{CR}
2	#15 ^{CS}	#9 ^{CR}
3	#15 ^{CS}	#19 ^{CR}
4	#12 ^{CS}	#25 ^{CR}
5	#18 ^{CS}	#25 ^{CR}

8.2.2 Antibody microarray analysis

The Panorama Antibody Microarray XPRESS Profiler725 Kit (#XP725, Sigma Aldrich) which comprises 725 antibodies spotted in duplicate onto a nitrocellulose-coated glass microscope slide was used as previously described (section 7.2). In brief, total protein lysates from CS samples were fluorescently labelled with Cy3 (#PA23001, GE Healthcare) and lysates from CR samples were labelled with Cy5 (#PA25001, GE Healthcare). Prior to protein binding, dye-to-protein molar ratios were determined for each sample to ensure that the ratio was ≥ 2 . A total of 5 comparative experiments were performed, each comparing a CR and a CS sample. Equal amounts of protein from each sample (90 μ g) were incubated with the slide for 45 minutes on an orbital shaker at low speed. Data normalisation and analysis was performed as described previously and experiments were considered successful when the percentage of ‘substances matched’ was ≥ 90 (section 7.4). To denote differentially expressed proteins (DEPs), fold changes ≥ 1.8 were considered significant, with fold changes ≥ 1.5 also recorded for each experiment for use as supporting data. As previously described (section 4.3.7), the direction of fold change (increase or decrease in expression of a protein) was not shown, as dye-swap experiments had not been performed (section 4.3.7). Subsequent western blotting and clinical validation using immunohistochemistry were used to determine the direction of fold change.

8.3 Results

Five antibody microarray experiments were performed, which involved the comparison of protein expression in chemotherapy-sensitive samples with protein expression in chemotherapy-resistant samples, in order to identify differentially expressed proteins (DEPs) associated with chemo-resistance. This generated 5 sets of data, with a total of 38 DEPs. Seven of these DEPs were significantly (≥ 1.8 -fold) identified in at least two experiments (zyxin, 14-3-3 theta/tau, tBID, pinin, Bcl-xL, RIP and MyD88) (Table 22). The RIDEPs identified in chapter 7, which are present within this list of 38 DEPs include zyxin, BID, MyD88, IKKa, Bcl-xL, chondroitin sulphate, 14-3-3 theta/tau, centrin, pinin and siah2 (n=10) (Table 19) (section 7.3).

Table 22: DEPs in chemotherapy resistant tumour tissue identified by 5 antibody microarray experiments comparing CR and CS samples.

Significant expression fold change (≥ 1.8) is indicated in bold. For proteins which show ≥ 1.8 -fold change in expression, supporting data from other experiments is shown upward of 1.5-fold. Values considered to be not significant (---) and antibody spots which did not pass the analysis criteria for experimental quality control (\otimes) are also indicated.

Ab #	Protein	Gene	11 ^{CS} /19 ^{CR}	15 ^{CS} /9 ^{CR}	15 ^{CS} /19 ^{CR}	12 ^{CS} /25 ^{CR}	18 ^{CS} /25 ^{CR}
Z0377	Zyxin	ZYX	7.80	2.01	2.21	2.02	2.63
T5942	14-3-3 theta/tau	YWHAQ	1.54	1.90	2.29	2.55	1.52
B3183	tBID	BID	---	2.16	1.55	1.97	1.96
P0084	Pinin	PNN	\otimes	2.54	2.39	1.53	---
B9429	Bcl-xL	BCL2L1	---	2.26	---	1.57	2.62
R8274	RIP	RIPK1	2.07	---	---	2.56	---
M9934	MyD88	MYD88	---	---	2.08	2.18	---
P3203	Protein Kinase Cb2	PRKCB	3.20	---	---	1.70	---
P3078	Protein Kinase Cb1		---	---	---	2.06	---
T5530	Tau	MAPT	1.63	\otimes	2.04	\otimes	---
I6139	IKKa	CHUK	1.61	---	---	2.03	---
C8035	Chondroitin sulfate	ACAN	1.56	---	2.00	---	---
F7926	FAK (pTyr397)	PTK2	---	---	1.51	---	1.95

Ab #	Protein	Gene	11 ^{CS} /19 ^{CR}	15 ^{CS} /9 ^{CR}	15 ^{CS} /19 ^{CR}	12 ^{CS} /25 ^{CR}	18 ^{CS} /25 ^{CR}
F8926	FAK (pTyr577)		---	⊗	---	2.20	---
C7736	Centrin	CETN1	<i>1.51</i>	---	---	1.90	---
P9371	PINCH 1	LIMS1	---	---	<i>1.51</i>	1.90	---
T9191	TRAIL	TNFSF10	---	---	1.83	<i>1.62</i>	---
S0315	SAPK3	MAPK12	---	3.76	⊗	---	---
R8529	RALAR	RALA	---	⊗	3.70	⊗	---
C8854	Caspase 13	CASP13	---	---	2.69	---	---
T2780	Tropomyosin	TPM1	---	2.67	---	⊗	---
S4047	S6 Kinase	RPS6KB1	---	---	---	2.44	---
R5145	Rsk1	RPS6KA1	---	---	---	2.33	---
H9286	Acetyl Histone H3 AcLys9	H3F3A	---	---	---	2.21	---
D5567	Dimethyl Histone H3		2.14	---	---	---	---
D1314	DRAK1	STK17A	---	2.18	---	---	---
C3956	cMyc	MYC	---	---	---	2.17	---
M9317	MeCP2	MECP2	---	---	---	2.17	---
Ab #	Protein	Gene	11 ^{CS} /19 ^{CR}	15 ^{CS} /9 ^{CR}	15 ^{CS} /19 ^{CR}	12 ^{CS} /25 ^{CR}	18 ^{CS} /25 ^{CR}
E2520	Epidermal Growth Factor	EGF	---	---	---	2.13	---
M6194	Munc13 1	UNC13A	---	---	---	2.08	---
T1827	TBP	TBP	---	---	---	2.08	---
S5313	Sir2	SIRT1	---	---	---	2.08	---
H9912	hSNF5 INI1	SMARCB1	---	---	---	2.03	---
A8604	Annexin V	ANXA5	---	---	---	2.02	---
P1601	Protein Kinase Ba	AKT1	2.10	---	---	---	---
R4904	Reelin	RELN	1.95	⊗	---	⊗	---
P6834	Ki-67	MKI67	1.92	---	---	---	---
D1286	Desmosomal protein	DSC1	---	---	1.92	---	---
H9411	HDAC4	HDAC4	---	---	---	---	1.89
T0825	Transportin 1	TNPO1	---	---	---	1.84	---
S9809	Sp1	SP1	---	---	---	1.82	---

8.4 Discussion

A list of 38 differentially expressed proteins (DEPs) associated with chemo-resistance has been identified, using antibody microarray analysis on fresh tumour tissue. Seven DEPs

were significantly identified in more than one experiment (zyxin, 14-3-3 theta/tau, tBID, Pinin, Bcl-xL, RIP and MyD88). Following the identification of several RIDEPs (section 7.4.3), of which zyxin was the most common, the selection of proteins for further analysis must be carefully considered. This data will be further analysed in chapter 10, to include the data mining, confirmation and clinical validation phases of the biomarker discovery pipeline.

The 14-3-3 theta/tau isoform of the 14-3-3 family of proteins, which is the only isoform present on the array slide, was found to be differentially expressed by at least 1.5 fold in 5/5 experiments. The 14-3-3 family of proteins have previously been implicated in resistance to anthracycline or taxane therapy in breast cancer cells (section 3.7) (Liu, Liu et al. 2006). This will be discussed further in chapter 10. Another factor which has a well-established putative role in chemotherapy response, and is a critical biological process responsible for the execution of cell death, is the apoptosis pathway (Pommier, Sordet et al. 2004; Chuthapisith 2007) (section 2.1.4). Proteins identified, which are associated with this pathway therefore warrant further investigation; this includes tBID, Bcl-xL and MyD88.

In order to analyse and interpret the data, and to aid the selection of candidates for confirmation and validation phases, the list of DEPs identified by antibody microarray must now be carried forward to the data mining stage. This involves the use of Ingenuity Pathway Analysis (IPA) software, which identifies relationships between protein candidates and highlights canonical pathways these candidates may be associated with. This will be described in chapter 10.

CHAPTER 9:

**TWO-DIMENSIONAL GEL ELECTROPHORESIS
AND MASS SPECTROMETRY FOR THE
IDENTIFICATION OF BIOMARKERS OF
CHEMOTHERAPY RESISTANCE**

Chapter Aim:

To use 2D-PAGE MALDI-TOF/TOF MS analysis for the identification of putative biomarkers of neoadjuvant chemotherapy resistance in breast cancer using clinical tumour tissue. This will form part of the discovery phase of the biomarker discovery pipeline.

Chapter 9. Two-dimensional gel electrophoresis and mass spectrometry for the identification of biomarkers of chemotherapy resistance

9.1 Introduction

This chapter involves the use of comparative two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled with matrix-assisted laser desorption/ionisation time-of-flight tandem mass spectrometry (MALDI-TOF/TOF MS) (section 3.3), to identify potential protein biomarkers associated with chemo-resistance, using clinical tumour tissue samples. This forms part of the ‘discovery’ phase of the biomarker discovery pipeline, aiming to generate a list of differentially expressed proteins (DEPs) between chemotherapy-sensitive and chemotherapy-resistant tumour samples. There are several stages involved in the 2D-PAGE/MS process, which are outlined in Figure 32.

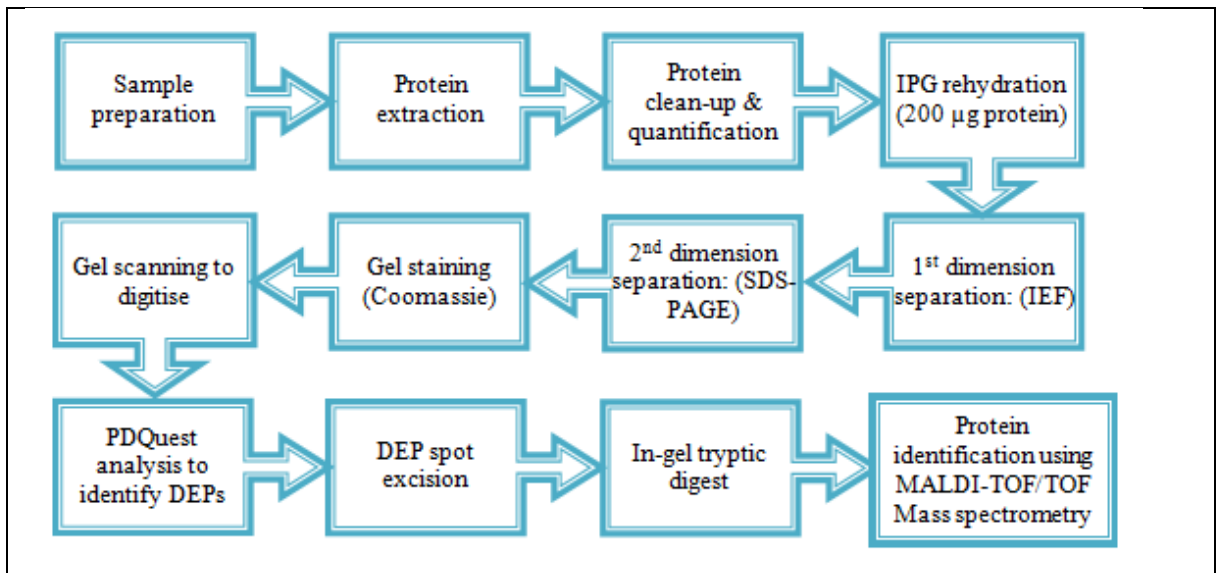


Figure 32: 2D-PAGE/MS workflow

An outline of the workflow and steps involved in the discovery of differentially expressed proteins (DEPs) using 2D-PAGE coupled to mass spectrometry.

9.1.1 MALDI TOF/TOF mass spectrometry

The mass spectrometer being used for the analysis of peptides for protein identification is a Bruker Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonics). As previously discussed (section 3.3.2.7), a mass spectrometer is composed of three components; an ion source, a mass analyser and a detector.

The target plate, containing sample/matrix is inserted into the machine, after which it is transferred via the vacuum lock into the system. Ions are formed using the MALDI technique, with a pulsed ion extraction (PIE) method. This method reduces differences in kinetic energy, which is due to differences in energy distribution, between ions of the same mass. If ions of the same mass arrive at the detector at different times due to differences in kinetic energy, it results in peak broadening and decreases resolution. The PIE method is composed of three components; P1 (target plate), P2 (electrode) and ground potential (electrode) (Figure 33), where ions are exposed to different electric potentials, which results in ions with higher initial energy (faster) being exposed to less electric potential, and ions with lower initial energy (slower) being exposed to more electric potential. This corrects for differences in kinetic energy, and ensures ions of the same mass arrive at the detector at the same time. Peak broadening and loss of resolution is also minimised by the use of the reflectron, which also corrects for differences in velocity between ions of the same mass. Following ionisation, ions travel down the flight tube, are reflected by the reflector, and subsequently reach the micro-channel-plate detector. Ions with higher mass will reach the detector after ions with a lower mass. The detector determines the m/z of each ion and its relative abundance in the sample, and presents this data in the form of a peptide mass fingerprint (PMF). This can be submitted to MASCOT where the masses observed can be

compared to theoretical tryptic digests within a protein sequence database, in order to achieve protein identification.

Tandem mass spectrometry involves further fragmentation of ions within the initial PMF, and provides a more confident and accurate protein identification. The Ultraflex III offers both laser-induced dissociation (LID) and the high energy collision-induced dissociation (CID), the former of which is most commonly used for protein identification (Suckau, Resemann et al. 2003), and in combination with the patented LIFT device, offers high resolution, mass accuracy and sensitivity. For MS/MS, the laser intensity is increased, which increases the yield of precursor ions per shot (Suckau, Resemann et al. 2003). Increasing voltage and laser intensity, gives ions more internal energy and upon collision with nitrogen ions, fragmentation occurs. The most intense peaks from the PMF, for example the highest 10 peaks, are selected for MS/MS one by one. Following fragmentation by LID, ion 'families', consisting of the precursor ion and its fragments travelling together, are selected by the precursor ion selector (PCIS) (Suckau, Resemann et al. 2003) and directed towards the LIFT cell. Ions enter the LIFT cell, where they are accelerated and given amounts of kinetic energy proportional to their mass, so that ions only of the same mass arrive at the detector at the same time.

The main types of peptide fragments obtained from MS/MS, produced by the cleavage of the C¹-C, C-N or N-C¹ bond. Where the charge remains on the N-terminus, *a*, *b* and *c* fragments are produced and where the charge remains the C-terminus, *x*, *y* and *z* fragments are produced (Figure 34). The most frequent cleavage site is the C-N bond, the cleavage of which yields *b* and *y* fragments. In these experiments, using the Ultraflex III in LIFT mode without the addition of a collision-gas, we can expect to see mainly *b* and *y* fragments (Figure 34) (Shenar, Sommerer et al. 2009). The amino acid residue can be determined by

calculating the mass difference between two consecutive fragments (e.g. b_1 and b_2). The spectra generated by MS/MS illustrate the fragmentation pattern of each precursor ion, by the mass of each ion produced; this is similar to a PMF, and may be referred to as a peptide fragmentation fingerprint (PFF). The masses observed can be compared to the theoretical peptide fragment ion masses in a database, using MASCOT, and protein identification can be proposed. For confident and stringent protein identification, matching of two peptides is desirable. One peptide-match is commonly accepted, however for publication-purposes extra information is sometimes required (section 9.3.6).

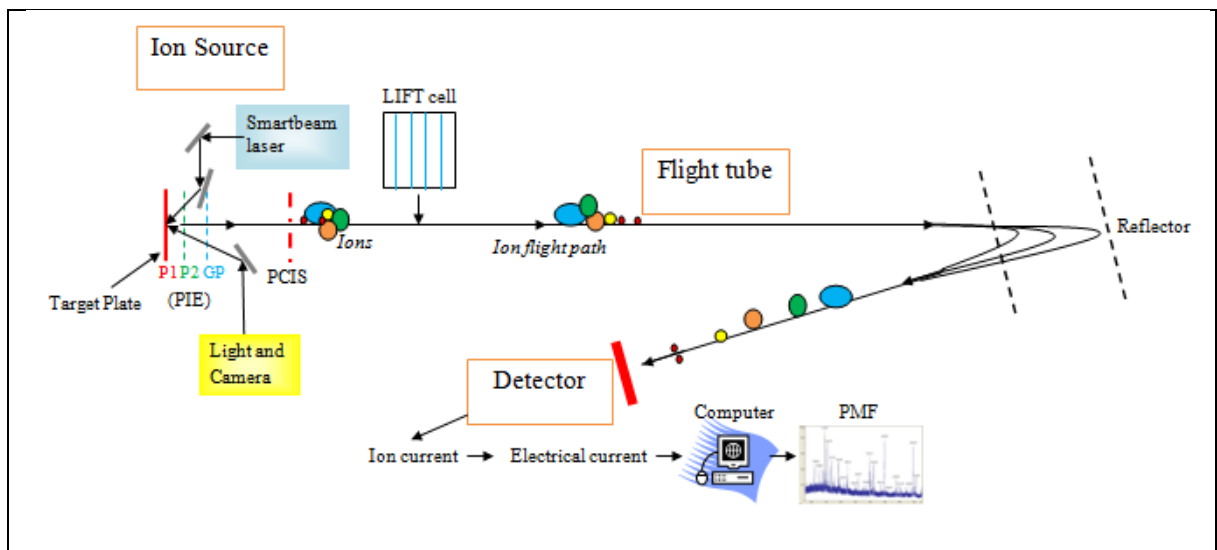


Figure 33: Schematic of the Ultraflex III MALDI-TOF/TOF mass spectrometer

An Ultraflex III MALDI-TOF/TOF mass spectrometer, equipped with a Smartbeam laser (designed to combine the best attributes from nitrogen and Nd:YAG lasers), was used to analyse peptide samples. A peptide mass fingerprint (PMF) is generated initially, by ionisation of the sample and determination of its m/z ratio upon reaching the detector, by calculation of the time taken to travel down the flight tube. Peaks can then be automatically selected for MS/MS, fragmented and analysed using the LID-LIFT method. The submission of the spectra to a sequence database, via MASCOT, then allows protein identification upon comparison with known sequences in the database, by knowledge of digest and fragmentation chemistry.

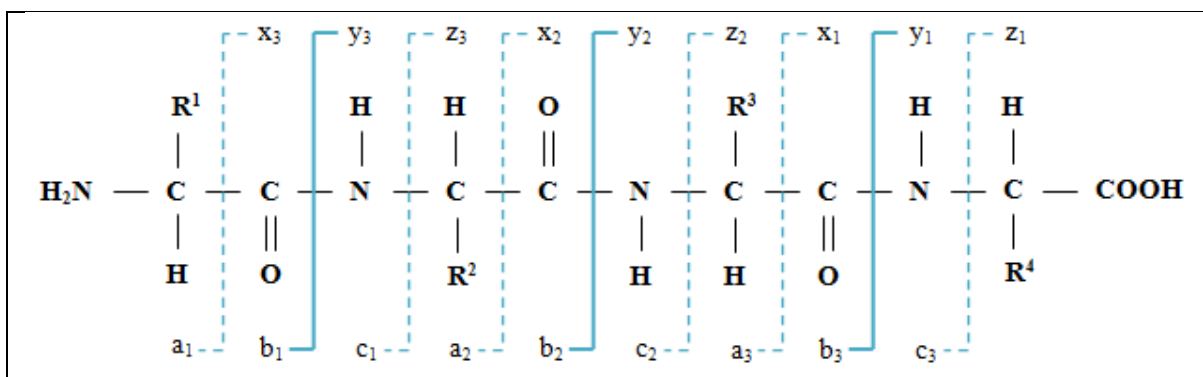


Figure 34: Peptide fragmentation (Roepstorff and Fohlman notation)

The main types of peptide fragments obtained from MS/MS, produced by the cleavage of the C¹-C, C-N or N-C¹ bond. Where the charge remains on the N-terminus, *a*, *b* and *c* fragments are produced and where the charge remains the C-terminus, *x*, *y* and *z* fragments are produced. The number (subscript) represents the number of amino acid residues within the fragment. The weakest bond, and therefore most frequent cleavage site, is the C-N bond; the cleavage of which yields *b* and *y* fragments. For experiments presented within this chapter, using the Ultraflex III in LIFT mode without the addition of a collision-gas, the type of fragments expected would be mainly *b* and *y* fragments (shown as solid coloured line). The mass difference between two adjacent ions (e.g. between *b*₂ and *b*₃) will reveal the amino acid residue, thus uncovering the peptide sequence.

9.2 Methodology

9.2.1 Clinical Samples

Clinical samples were selected based upon sample availability and paired according to tumour type, molecular subtype, chemotherapy treatment regimen and response to chemotherapy. All tumour samples used were ductal tumours of luminal (ER+) subtype. Four clinical tumour samples (#15^{CS}, #19^{CR}, #1^{CR} and #18^{CS}) were selected (Table 23) and used to create three pair-wise combinations, depending upon protein yield from each, to maximise the data generated from the samples (Table 24). Three of these samples (#15^{CS}, #19^{CR} and #18^{CS}) were also analysed by antibody microarray analysis (section 8.2.1).

Table 23: Clinical samples selected for 2D-PAGE/MS analysis

Four clinical samples were selected, which were all ductal tumours of luminal subtype (ER+) and were all negative for the expression of HER2. Tumour response was determined, based on the change in size pre- and post-treatment as well as consideration of tumour progression in the event of metastasis. Based upon this, tumours were classified as chemotherapy-sensitive (CS) or chemotherapy-resistant (CR)

<i>Sample #</i>	<i>Hormone Status</i>	<i>Tumour response</i>
15^{CS}	ER+ PR+ HER2-	69.2% size reduction. Partial Response
18^{CS}	ER+ PR+ HER2-	78.1% size reduction. Partial response
19^{CR}	ER+ PR- HER2-	71.4% size increase. Progressive disease
1^{CR}	ER+ PR- HER2-	Progressive disease (metastasis during therapy)

Table 24: Pair-wise combinations of clinical samples for three 2D-PAGE/MS experiments

Due to sample availability, four clinical tumour samples were used to generate three pair-wise combinations for three 2D-PAGE/MS experiments

<i>Experiment #</i>	<i>Chemotherapy-sensitive (CS) Sample</i>	<i>Chemotherapy-resistant (CR) sample</i>
1	#15^{CS}	#19^{CR}
2	#15^{CS}	#1^{CR}
3	#18^{CS}	#1^{CR}

9.2.2 Protein extraction

Following optimisation of methods in section 6.3.2, proteins were extracted from each tumour sample in 2D extraction buffer with sonication, as described previously (section 4.4.1). Samples were stored at minus 80 °C in polypropylene microcentrifuge tubes.

9.2.3 Protein clean-up and quantification

The ReadyPrep Cleanup Kit (#163-2130, Bio-Rad) was used to prepare samples for isoelectric focusing (IEF) by removing contaminants such as lipids, salts and nucleic acids. The procedure was carried out according to manufacturers' instructions, as described in section 4.4.2. Following the clean-up of samples, samples were quantified using the 2-D

Quant Kit (#80-6483-56, GE Healthcare). The procedure was performed according to manufacturers' instructions, as described in section 4.4.3.

9.2.4 2D PAGE

9.2.4.1 1st dimension (IEF)

ReadyStrips IPG strips (pH 4-7; 11cm) (#163-2015, Bio-Rad) were rehydrated with 200 µg of protein sample, for 16 hours. This was performed for each sample, in triplicate. IEF was performed using a 3-step program, consisting of 20 min at 250 V (linear); 150 min at 8000 V (linear) and 20 000 V-hours at 8000 V (rapid), as described in section 4.4.4.

9.2.4.2 2nd dimension (SDS-PAGE)

Following IEF, proteins within the IPG strip were equilibrated with DTT and IAA, which involved the reduction and alkylation of the proteins in preparation for SDS-PAGE, as described in section 4.4.5. IPG strips were then placed at the top of a Criterion™ pre-cast gel (11 cm 8-16% Tris-HCl polyacrylamide gel) (#354-0105, Bio-Rad), embedded in 1% overlay agarose. Proteins were separated by mass at 200 V, 500 mA and 300 W for 65 min, as previously described (section 4.4.5).

Proteins were visualised by staining with Bio-safe Coomassie Stain (#161-0787, Bio-Rad) for 1 hour, following de-staining for 16 hours in ddH₂O, on an orbital shaker. Gels were scanned using a GS800 calibrated densitometer (Bio-Rad) and Quantity One software (Bio-Rad) (section 4.4.6).

9.2.5 PDQuest

For each experiment, three gel images for each sample (chemotherapy-sensitive and chemotherapy-resistant) were up-loaded into PDQuest software. Spots were manually

detected and matched across all six gels, and PDQuest software was used to identify the significant (fold change ≥ 2 , $p < 0.05$) difference in expression of a protein spot between chemotherapy-sensitive and chemotherapy-resistant samples, using Boolean quantification and the Students *t*-test, as described in section 4.4.7. Histograms showing relative abundance of each spot within the six gels were provided for each DEP. Within PDQuest, chemotherapy-sensitive gels were coded red, and chemotherapy-resistant gels were coded green, as represented by the histograms.

9.2.6 Spot excision and in-gel digest

DEP spots selected by PDQuest were manually excised and transferred into LoBind eppendorf tubes (#022431064, Eppendorf) (section 4.4.8). Following this, gel pieces were washed and de-stained in 25 mM ammonium bicarbonate solution and proteins were digested overnight into peptides using Trypsin Gold (#V5280, Promega), as described in section 4.4.9.2.

9.2.7 Identification by MALDI-TOF/TOF MS

Peptides from each spot were mixed with an equal amount of 5 mg/ml 4-hydroxy- α -cyanocinnamic acid (CHCA) matrix, in 50% ACN and 0.1% TFA (aq), and spotted onto an MTP384 polished steel target plate (Bruker Daltonics). Mass spectra were obtained using the Ultraflex III MALDI-TOF/TOF MS (Bruker Daltonics) in reflectron mode over a mass range of 800-4000 *m/z*, from positive ions generated by a Nd:YAG Smartbeam laser, as described in section 4.4.11. From the PMF generated, the 10 highest peaks, with a signal-to-noise threshold >30 , were automatically selected for MS/MS fragmentation. Fragmentation was performed in LIFT mode without addition of collision gas. The default calibration method was used for MS/MS spectra, as described in section 4.4.11. Flex

Analysis software (version 3.3, Bruker Daltonics) was used to process the spectra and generate peak lists for both MS and MS/MS spectra. MS/MS data was submitted to Mascot (version 2.1, Matrix Science Ltd) for searching of the IPI Human database, via the ProteinScape interface (version 2.3, Bruker Daltonics). The search criteria that were specified are listed in Table 25.

Table 25: Search criteria specified for protein identification

Enzyme:	Trypsin
Missed cleavages	1
Fixed modifications:	Carbamidomethyl (C)
Variable modifications:	Oxidation (M)
Peptide tolerance:	250 ppm
MS/MS tolerance:	0.5 Da
Instrument:	MALDI-TOF-TOF

9.3 Results

9.3.1 2D-PAGE

Three experiments were performed, which involved the direct comparison of proteins extracted from a chemotherapy-sensitive (CS) tumour to proteins extracted from a chemotherapy-resistant (CR) tumour, in order to identify DEPs between the two disease phenotypes (Table 24); Experiment 1: #15^{CS} versus #19^{CR}; Experiment 2: #15^{CS} versus #1^{CR}; Experiment 3: #18^{CS} versus #1^{CR}. Within each experiment, each sample (CS and CR) was separated by 2D-PAGE in triplicate, to provide 3 technical replicate gels. One example of a CS gel and a CR gel stained with Coomassie stain is shown for each experiment; experiment 1 (Figure 35); experiment 2 (Figure 36) and experiment 3 (Figure 37).

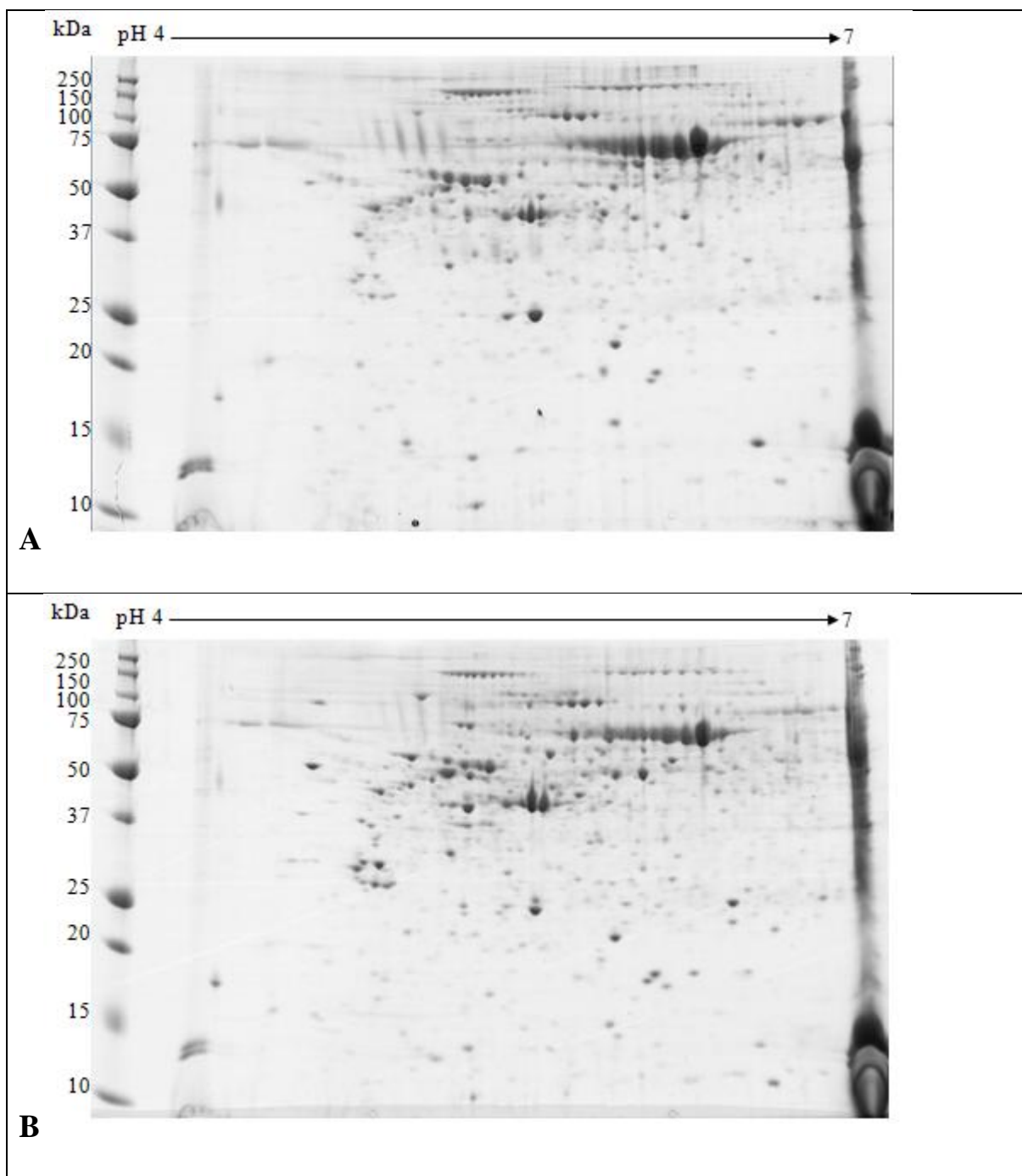


Figure 35: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 1)

Protein extracts were separated by 2D PAGE; horizontal separation by pI in the pH range 4-7, and vertical separation by size on an 11 cm polyacrylamide gel, run with a molecular weight marker (kDa). Proteins were visualised using Coomassie blue protein stain. **A:** 2D-PAGE separation of sample #15^{CS}. **B:** 2D-PAGE separation of sample #19^{CR}. The separation of each of the two samples was performed in triplicate, giving 6 gels in total.

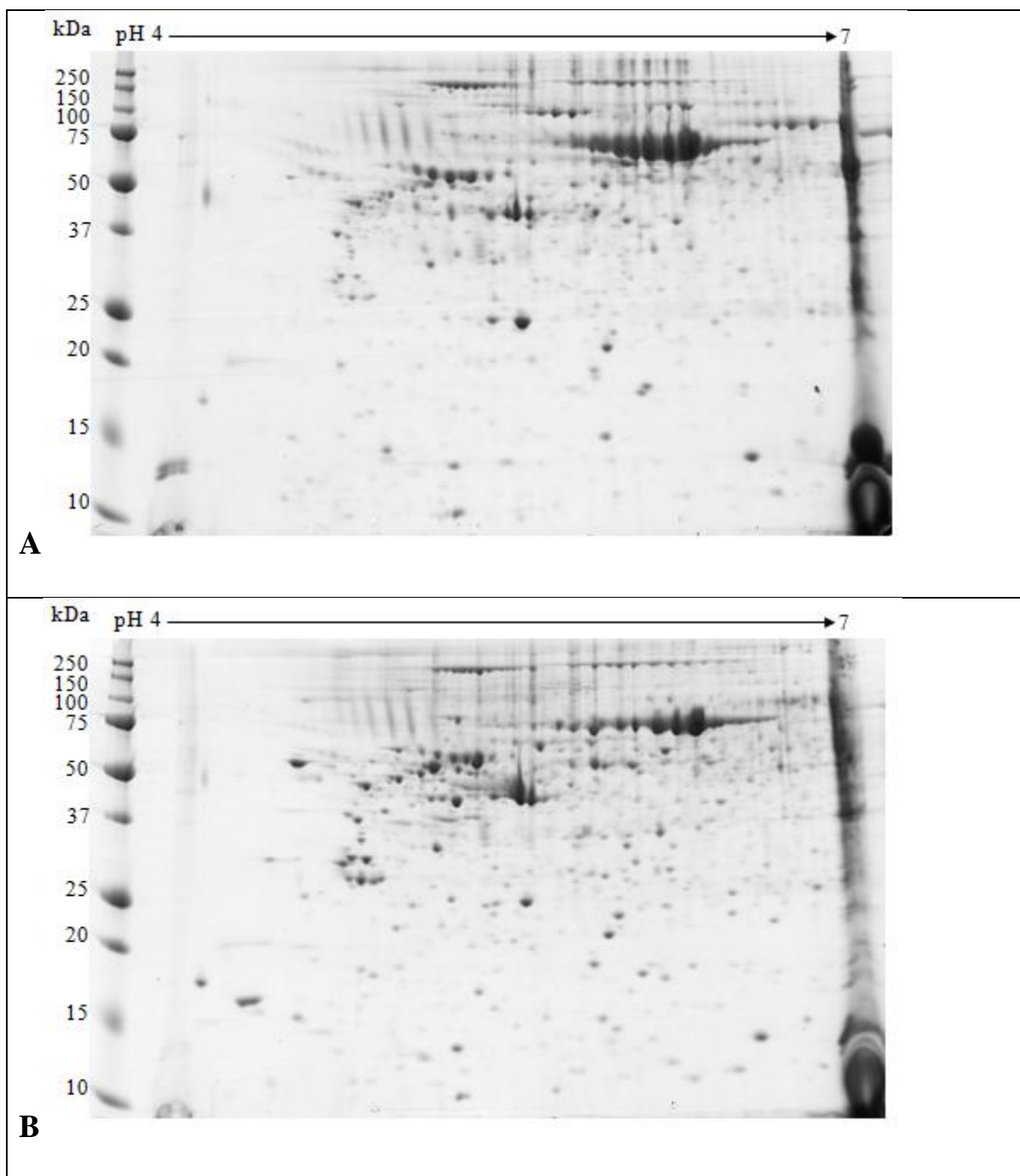


Figure 36: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 2)

Protein extracts were separated by 2D PAGE; horizontal separation by pI in the pH range 4-7, and vertical separation by size on an 11 cm polyacrylamide gel, run with a molecular weight marker (kDa). Proteins were visualised using Coomassie blue protein stain. **A:** 2D-PAGE separation of sample #15^{CS}. **B:** 2D-PAGE separation of sample #1^{CR}. The separation of each of the two samples was performed in triplicate, giving 6 gels in total.

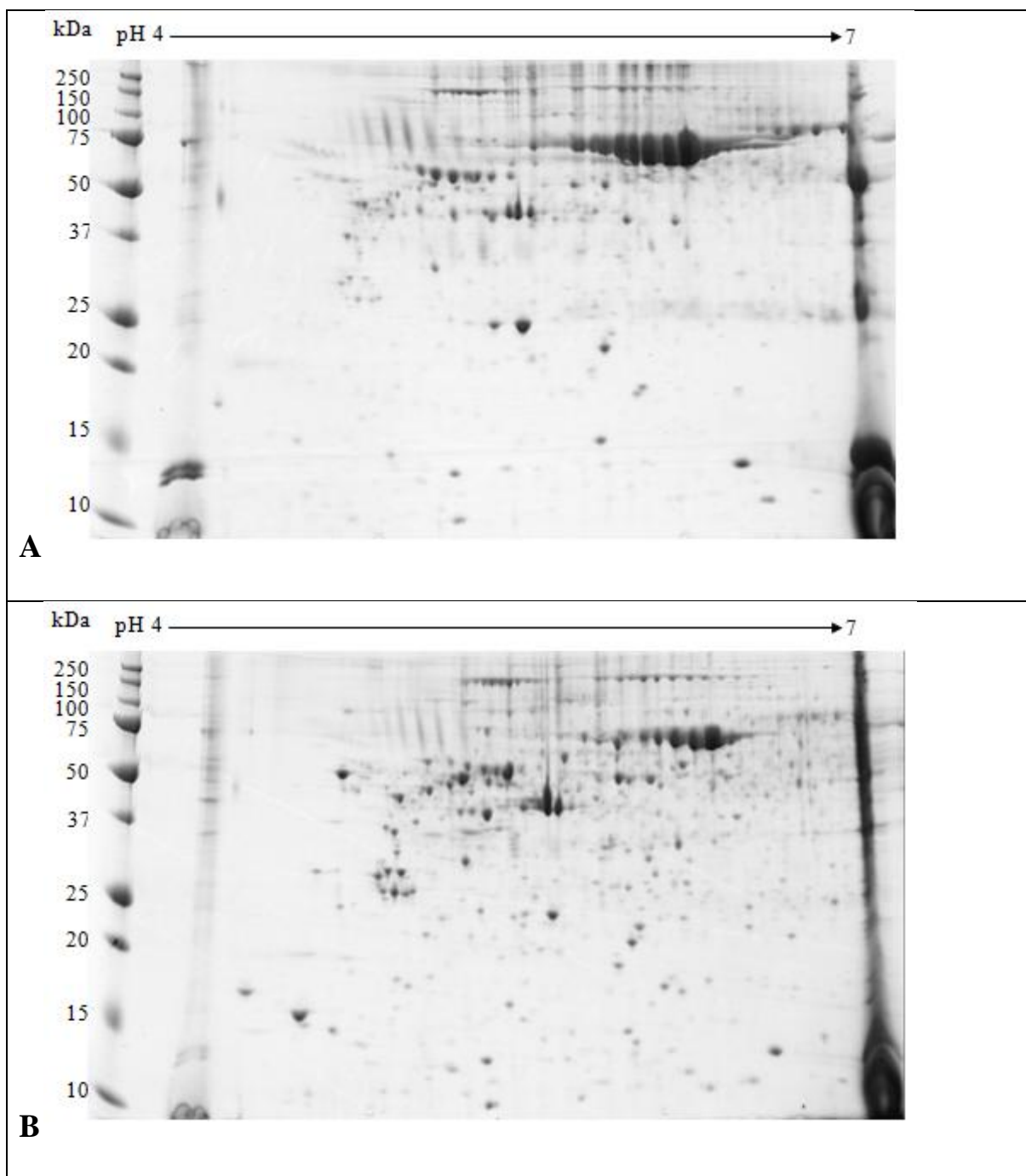


Figure 37: 2D-PAGE gel images stained with Coomassie blue protein stain (Experiment 3)

Protein extracts were separated by 2D PAGE; horizontal separation by pI in the pH range 4-7, and vertical separation by size on an 11 cm polyacrylamide gel, run with a molecular weight marker (kDa). Proteins were visualised using Coomassie blue protein stain. **A:** 2D-PAGE separation of sample #18^{CS}. **B:** 2D-PAGE separation of sample #1^{CR}. The separation of each of the two samples was performed in triplicate, giving 6 gels in total.

9.3.2 PDQuest

Analysis was performed using PDQuest software to identify DEPs (≥ 2 -fold; $p < 0.05$) between CS and CR samples. For each experiment, different numbers of DEPs were identified by PDQuest; experiment 1, 250 DEPs; experiment 2, 308 DEPs; experiment 3, 179 DEPs. As an example, the first 10 consecutive DEPs that were carried forward are shown for each experiment. This includes the histogram for each DEP, which shows the intensity of the spot in 3 x CS gels (red) and 3 x CR gels (green), as well as an example of a spot from a CS gel and an example of a spot CR gel. These are shown for experiment 1 (Figure 38), experiment 2 (Figure 39) and experiment 3 (Figure 40).

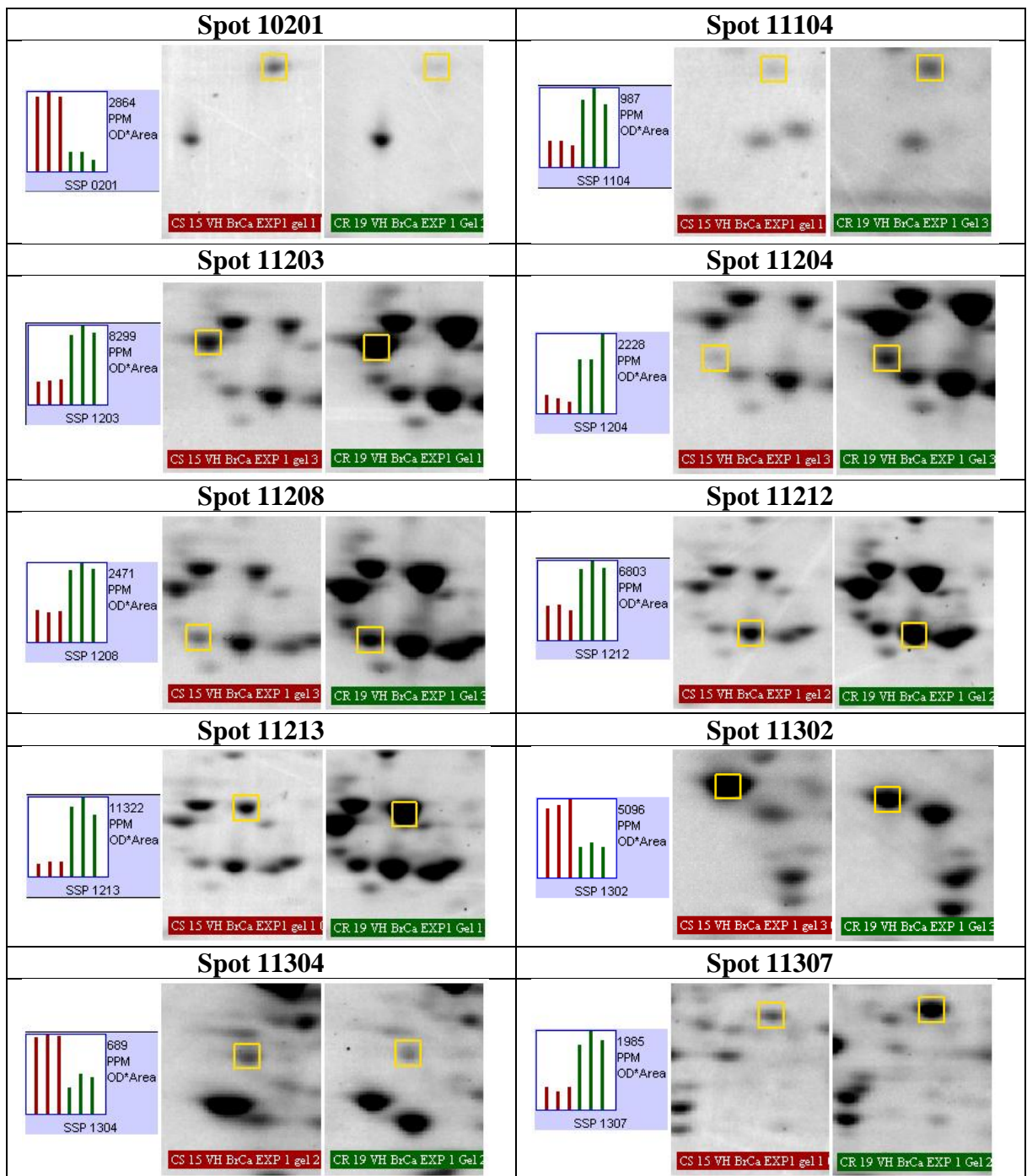


Figure 38: Histograms and DEP spots from PDQuest analysis (Experiment 1)

The first 10 spots identified by PDQuest as DEPs (≥ 2 fold; $p < 0.05$) have been selected as examples. The chemotherapy-sensitive (sample #15^{CS}) is shown in red, and the chemotherapy-resistant (sample #19^{CR}) is shown in green. Histograms show differences in spot density between each sample, representing fold change in expression, in triplicate. DEPs are highlighted with the yellow box, and the DEP is shown in respective CS and CR gels.

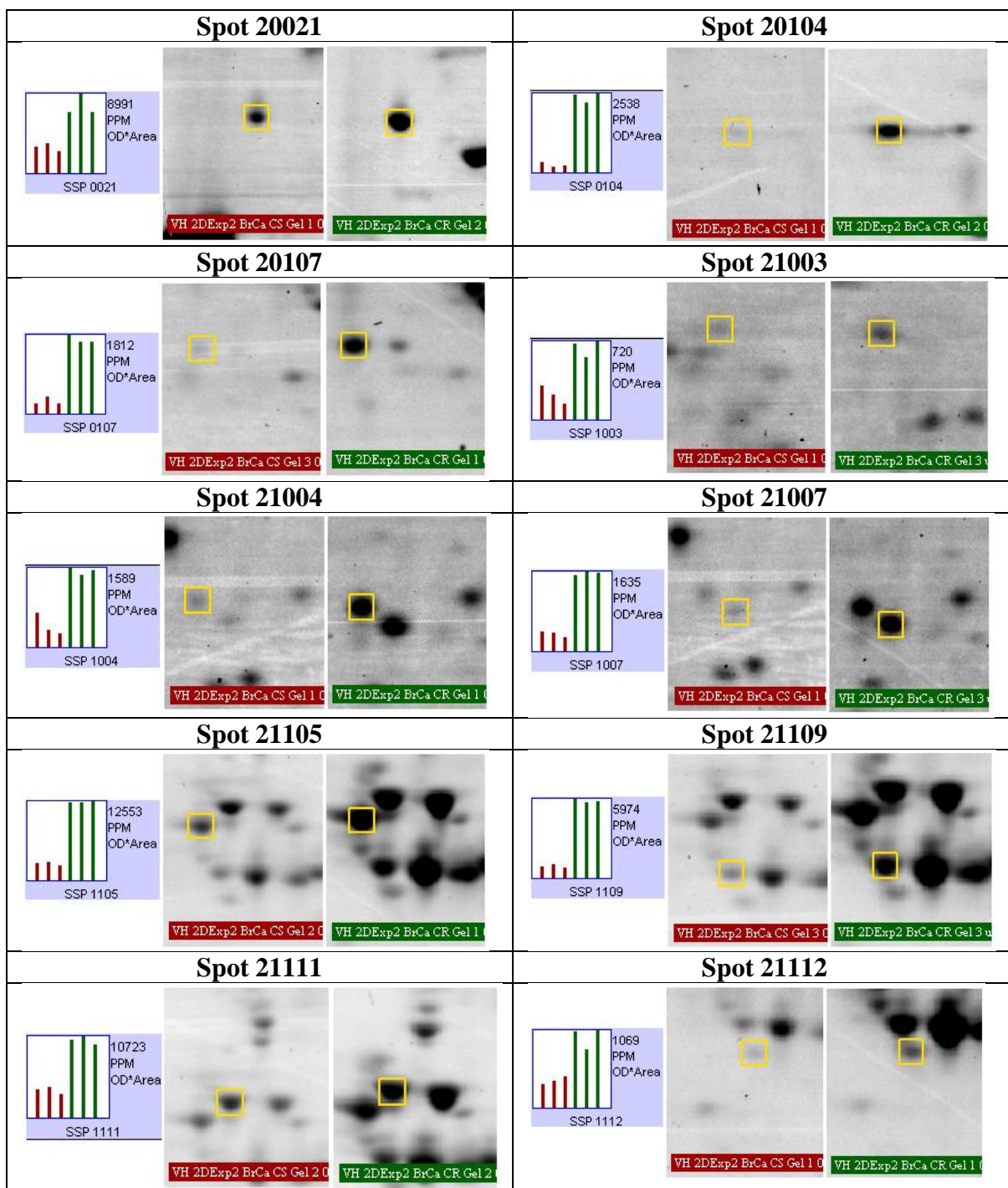


Figure 39: Histograms and DEP spots from PDQuest analysis (Experiment 2)

The first 10 spots identified by PDQuest as DEPs (≥ 2 fold; $p < 0.05$) have been selected as examples. The chemotherapy-sensitive (sample #15^{CS}) is shown in red, and the chemotherapy-resistant (sample #1^{CR}) is shown in green. Histograms show differences in spot density between each sample, representing fold change in expression, in triplicate. DEPs are highlighted with the yellow box, and the DEP is shown in respective CS and CR gels.

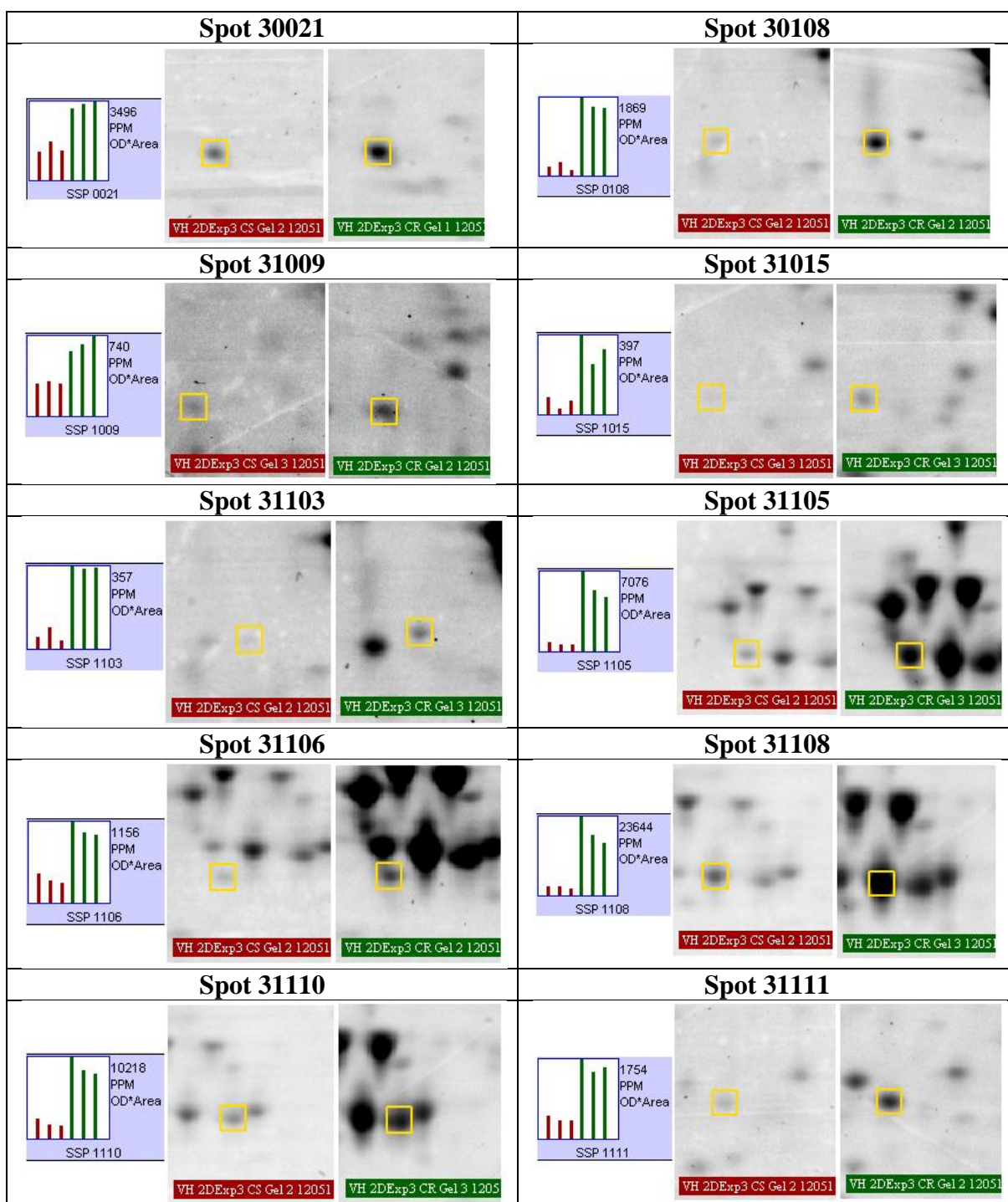


Figure 40: Histograms and DEP spots from PDQuest analysis (Experiment 3)

The first 10 spots identified by PDQuest as DEPs (≥ 2 fold; $p < 0.05$) have been selected as examples. The chemotherapy-sensitive (sample #18^{CS}) is shown in red, and the chemotherapy-resistant (sample #1^{CR}) is shown in green. Histograms show differences in spot density between each sample, representing fold change in expression, in triplicate. DEPs are highlighted with the yellow box, and the DEP is shown in respective CS and CR gels.

9.3.3 Selection and identification of DEPs

Following identification of DEPs by PDQuest analysis, selected spots were manually excised from the gels, digested into peptides by trypsin and analysed using the Ultraflex III MALDI-TOF/TOF MS (Bruker Daltonics). The spots which were selected for excision included spots which contained 1 protein (checked by 3D viewer within PDQuest), spots which were not part of a cluster or overlapping other protein spots, and were not part of a streak.

The number of DEPs identified in each of the 3 experiments, along with the number that were excised from the gel for protein identification and the number for which a protein identification was achieved, is listed in Table 26. All the proteins identified from all experiments (n=250), with their corresponding data, are listed in Appendix 10.

Table 26: The number of DEPs identified for each experiment, with relation to the number initially identified by PDQuest and analysed by MALDI TOF/TOF MS.

The number of protein identifications achieved following MALDI-TOF/TOF peptide analysis is also shown, for each experiment, to determine the identification rate. On average, the identification rate following MALDI-TOF/TOF analysis was 66%.

Experiment	Total number of matched spots in PDQuest	Number of DEPs identified by PDQuest	Number of DEP spots excised, digested and analysed by MALDI TOF/TOF MS	Number of protein identifications following MALDI-TOF/TOF MS
1	737	250	120 (48%)	68 (56%)
2	573	308	139 (45%)	99 (71%)
3	336	179	118 (65%)	83 (70%)
<i>total</i>	<i>1646</i>	<i>737</i>	<i>377 (51%)</i>	<i>250 (66%)</i>
<i>average</i>	<i>548</i>	<i>245</i>	<i>125 (51%)</i>	<i>83 (66%)</i>

9.3.4 Examples of mass spectra

An example of a peptide mass fingerprint (PMF) and an example of a fragmented peptide mass spectrum (PFF), obtained with the Ultraflex in LIFT mode is shown in Figure 41. An annotated fragment ion spectrum is also shown (Appendix 6).

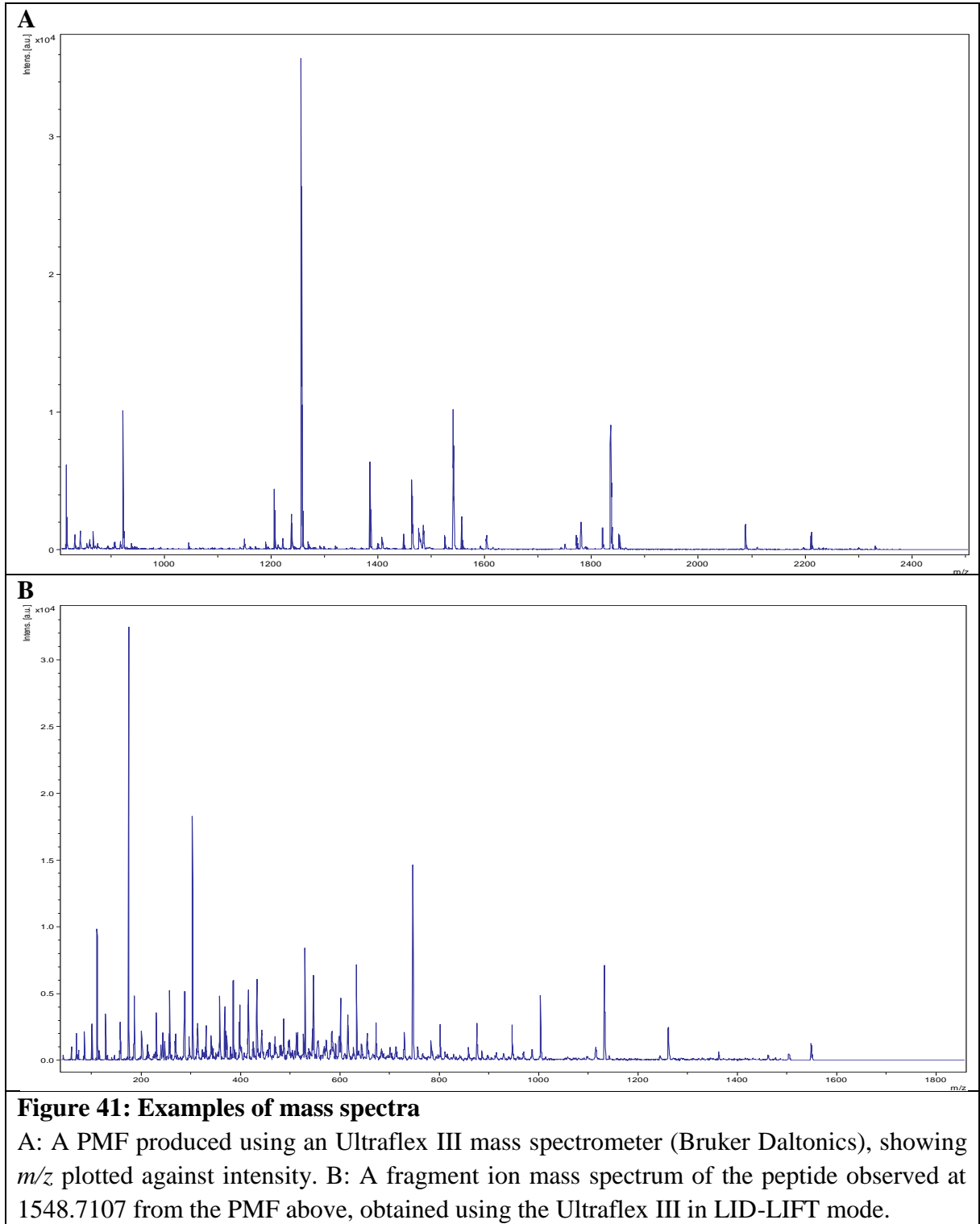
9.3.5 Example of a Mascot summary report page

An example of a Mascot summary report page is shown in Figure 42. The most relevant information (highlighted in Figure 42) includes the protein name, its gene symbol, accession number, protein mass, the score, the number of peptides matched ('expect' value ≤ 0.05), the sequence of the peptides matched, their scores, 'expect' values and mass error. By clicking on the accession number, the 'protein view' is given, from which the pI of the protein and the percentage sequence coverage can be obtained. By clicking on the query number for a peptide, the 'peptide view' is given (Appendix 7), which shows a mass spectrum with labelled fragment ions, and a table containing the matched fragment ions, highlighting those used for scoring.

9.3.6 A protein identification with a single peptide match

Protein identification may be presented as a single peptide match, or may show several matched peptides. The greater number of peptides matched represents a more confident protein identification, yet the 'expect' value, which states the likelihood that the peptide match occurred by chance, must also be considered. If protein identification with a single peptide match is not sufficient, extra information can be given as supporting evidence. This may include annotated spectra and/or lists of fragment ions, which can be obtained from ProteinScape (Appendix 6) 'peptide view' within Mascot (Appendix 7) or using spectrum annotation software such as Prophossi, which was designed to automate the validation of

phosphopeptide spectra generated by tandem mass spectrometry (Martin, 2010), but can also be used for the annotation of non-phosphopeptide mass spectra (<http://www.compbio.dundee.ac.uk/prophossi/bin/prophossi-cgi.pl>) (Appendix 8).

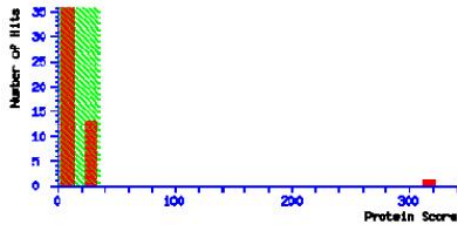


MATRIX SCIENCE Mascot Search Results

User : hull
 Email :
 Search title : Hull UP IPIHuman
 MS data file : 13229323905412293.mgf
 Database : IPI human (86845 sequences; 35122444 residues)
 Timestamp : 3 Jun 2011 at 13:24:41 GMT
 Protein hits : [IPI00328319](#) Tax_Id=9606 Gene_Symbol=RBBP4 Histone-binding protein RBBP4

Mascot Score Histogram

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Individual ions scores > 35 indicate identity or extensive homology ($p < 0.05$) Protein ions are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Peptide Summary Report

Format As: Peptide Summary [Help](#)

Significance threshold $p < 0.05$ Max. number of hits: AUTO

Standard scoring @ MudPIT scoring Ions score or expect cut off: 0 Show sub sets: 0

Show pop-ups @ Suppress pop-ups Sort unassigned: Decreasing Score Require bold red

Select All Select None Search Selected Error tolerant Archive Report

1. [IPI00328319](#) Mass: 47511 Score: 315 Matches: 6(4) Sequences: 6(4)
 Tax_Id=9606 Gene_Symbol=RBBP4 Histone-binding protein RBBP4

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 1	950.4841	949.4768	949.4691	8.07	0	23	0.83	1	U	K.LIIVDTR.S + Oxidation (M)
<input checked="" type="checkbox"/> 2	973.5582	972.5510	972.5393	12.0	0	54	0.00091	1	U	K.IVALNDLR.N
<input checked="" type="checkbox"/> 4	1067.5355	1066.5282	1065.5155	11.9	0	44	0.006	1	U	K.INHEGEVNR.A
<input checked="" type="checkbox"/> 5	1130.6335	1129.6263	1129.6244	1.68	1	20	1.6	1	U	R.RLNVDLSE.I
<input checked="" type="checkbox"/> 7	1251.5507	1250.5434	1250.5415	1.51	0	82	1.1e-36	1	U	K.EAIPDVAEER.V
<input checked="" type="checkbox"/> 10	1805.7911	1804.7839	1804.8162	-17.94	0	91	8.3e-38	1	U	K.HSKEDPSGEHPDLR.L

Proteins matching the same set of peptides:
[IPI00645329](#) Mass: 46415 Score: 315 Matches: 6(4) Sequences: 6(4)
 Tax_Id=9606 Gene_Symbol=RBBP4 46 kDa protein

Peptide matches not assigned to protein hits: (no details means no match)

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 1	1005.5353	1004.5260	1004.5291	-1.08	0	23	0.99	1		ISPFASDIR
<input checked="" type="checkbox"/> 2	1607.6971	1606.6899	1605.6933	-2.14	0	19	1.7	1		WLMUGIESPAESER
<input checked="" type="checkbox"/> 3	1297.6860	1296.6788	1296.7401	-47.34	0	11	12	1		T.LTGGIVVNR
<input checked="" type="checkbox"/> 6	1162.6212	1161.6139	1161.6618	-41.22	0	9	21	1		QAALLHVSAPK

Search Parameters

Type of search : MS/MS Ion Search
 Enzyme : Trypsin
 Fixed modifications : Carbamidomethyl (C)
 Variable modifications : Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 250 ppm
 Fragment Mass Tolerance : ± 0.5 Da
 Max Missed Cleavages : 1
 Instrument type : MALDI-TOF-TOF
 Number of queries : 10

Mascot: <http://www.matrixscience.com/>

Figure 42: An example of a Mascot concise protein summary page

This is an example of a Mascot concise protein summary page, where the information which needs to be extracted is highlighted. This includes the accession number, the mass of the protein matched, the score, the number of peptides significantly matched ('expect' value ≤ 0.05), the protein name, the sequence of the peptides matched and their scores, 'expect' values and mass error values. By clicking on the accession number, a separate page; 'protein view' is given, from which the pI and percentage sequence coverage is required. The values given for pI and protein mass should be compared to the 2D-PAGE gel when identifying a protein, to ensure these parameters are where they are expected. The Mascot search results may also give 'proteins matching the same set of peptides', or may give other protein matches with lower scores or number of peptides matched. The most appropriate protein identification can be elucidated by comparing protein masses, pI and the number of peptides in the sample matched to that protein.

9.3.7 Differentially expressed proteins identified by 2D-PAGE MS

Each of the three 2D-PAGE MS experiments produced a list of protein identifications following peptide analysis of excised DEP spots; experiment 1: 68 protein identifications; experiment 2, 99 protein identifications; experiment 3, 83 protein identifications, giving a combined list of 250 protein identifications (Appendix 10) However, within each of these experiments, several of the excised DEP spots yielded the same protein identification, due to different post-translational modifications (PTMs). The total list of DEPs for each experiment, after removal of replicate protein identifications for each experiment is: experiment 1, 53 DEPs; experiment 2, 85 DEPs; experiment 3, 77 DEPs, as shown in Table 27.

Combining the lists of DEPs generated from each experiment, highlighted the unique DEPs which had been identified overall, and those which had been identified across multiple experiments. This combined list included a total of 132 unique DEPs, which had been identified by performing three 2D-PAGE MALDI-TOF/TOF MS experiments. This is summarised in Table 28, showing the number of proteins which were identified in 3/3, 2/3 and 1/3 experiments. Proteins which were identified in at least two experiments (n=57) are

listed in Table 29. Supplementary data is given for the protein identifications which only have one peptide match (n=3) (Appendix 11). Proteins identified in only one experiment (n=75) are listed in an additional table (Appendix 9).

Table 27: The number of unique DEPs identified from each experiment, with relation to the number initially identified by PDQuest and analysed by MALDI TOF/TOF MS. The number of peptide samples identified by MALDI-TOF/TOF analysis is also shown, for each experiment, to determine the identification rate. On average, the identification rate following MALDI-TOF/TOF analysis was 66%. This is different to the total number of unique DEPs identified in each experiment (last column) due to presence of duplicate identifications as a result of PTM.

Experiment	Total number of matched spots in PDQuest	Number of DEPs identified by PDQuest	Number of DEP spots excised and analysed by MALDI TOF/TOF MS	Number of peptide samples identified following MALDI-TOF/TOF MS	Number of DEPs identified (after removal of duplicates)
1	737	250	120 (48%)	68 (56%)	53
2	573	308	139 (45%)	99 (71%)	85
3	336	179	118 (65%)	83 (70%)	77
<i>total</i>	<i>1646</i>	<i>737</i>	<i>377 (51%)</i>	<i>250 (66%)</i>	<i>215</i>
<i>average</i>	<i>548</i>	<i>245</i>	<i>125 (51%)</i>	<i>83 (66%)</i>	<i>71</i>

Table 28: The number of DEPs which have been identified in 3/3, 2/3 and 1/3 experiments

Number of experiments a DEP was identified in:				Number of DEPs
3/3	√	√	√	26
2/3		√	√	31
1/3			√	75
	<i>Total number of DEPs</i>			<i>132</i>

Where a protein was not identified as a DEP in an experiment, this may have been due to one of several reasons; (1) the protein may not have been differentially expressed; (2) the protein may have been differentially expressed but not excised from the gel following PDQuest (due to poor spot quality, for example); (3) the protein may not have been

identified during peptide analysis by MS. Proteins which are not represented as DEPs in an experiment are therefore referred to as ‘status unknown’ (Table 29). Highly abundant serum proteins (section 3.1 and section 5.1.2) that have been identified (n=3) and ‘Top 15’ RIDEP proteins (section 7.1) (n=4) that have been identified as DEPs are also shown in Table 29.

Table 29: Differentially expressed proteins associated with chemotherapy resistance, identified using 2D-PAGE MALDI-TOF/TOF MS

Three comparative 2D-PAGE MALDI-TOF/TOF MS experiments were performed to identify differentially expressed proteins (DEPs) associated with chemotherapy resistance. The table lists (alphabetically by gene symbol, from the IPI database) those DEPs identified in at least two experiments (n=57), showing ≥ 2 -fold change in expression, along with the direction of change ($\downarrow\uparrow$). Protein identifications with 1 peptide match are indicated (¹). Where a protein is not identified as a DEP, --- is shown, to represent status unknown. Proteins within the ‘22 proteins comprising ~99% of the serum proteome’ are indicated* (n=3), as well as those present in the TOP15 (human) RIDEP list* (n=4).

Protein	Gene Symbol	#15 ^{CS} v #19 ^{CR}	#15 ^{CS} v #1 ^{CR}	#18 ^{CS} v #1 ^{CR}
Activator of 90 kDa heat shock protein ATPase homolog 1	<i>AHSA1</i>	\uparrow^1	\uparrow	---
Annexin A3	<i>ANXA3</i>	---	\uparrow	\uparrow
Serum amyloid P-component	<i>APCS</i>	\downarrow	\downarrow	\downarrow^1
Apolipoprotein A1*	<i>APOA1</i>	\downarrow	\downarrow	\downarrow^1
Adenine phosphoribosyltransferase	<i>APRT</i>	\uparrow^1	---	\uparrow^1
Rho GDP-dissociation inhibitor 1*	<i>ARHGDI1</i>	---	\uparrow	\uparrow
Rho GDP-dissociation inhibitor 2	<i>ARHGDI2</i>	\uparrow	\uparrow	\uparrow^1
ATP synthase subunit beta, mitochondrial*	<i>ATP5B</i>	\uparrow	---	\uparrow^1
Barrier-to-autointegration factor	<i>BANF1</i>	---	\uparrow	\uparrow^1
Macrophage-capping protein	<i>CAPG</i>	---	\uparrow	\uparrow
Isoform 2 of F-actin-capping protein subunit beta	<i>CAPZB</i>	\uparrow	\uparrow	\uparrow^1
T-complex protein 1 subunit beta	<i>CCT2</i>	\downarrow^1	\uparrow	---
Creatine kinase B-type	<i>CKB</i>	\downarrow	\downarrow	\downarrow
Chloride intracellular channel protein 1	<i>CLIC1</i>	\uparrow	\uparrow	\uparrow^1
Coactosin-like protein	<i>COTL1</i>	---	\uparrow^1	\uparrow
Cellular retinoic acid-binding protein 2	<i>CRABP2</i>	\uparrow^1	\uparrow	\uparrow

Protein	Gene Symbol	#15^{CS} v #19^{CR}	#15^{CS} v #1^{CR}	#18^{CS} v #1^{CR}
Isoform 1 of Eukaryotic translation initiation factor 5A-1	<i>EIF5A</i>	---	↑	↑
Ferritin light chain	<i>FTL</i>	↓	↓ ¹	↑
Glycerol-3-phosphate dehydrogenase [NAD+] cytoplasmic	<i>GPD1</i>	↓	↓	---
Glutathione S-transferase omega-1	<i>GSTO1</i>	---	↑	↑
Glutathione S-transferase P	<i>GSTP1</i>	↓	↑	↑
HEBP2 protein (fragment)	<i>HEBP2</i>	---	↑	↑
highly similar to Heat-shock protein beta-6	<i>HSPB6</i>	↓	↓	---
Keratin, type I cytoskeletal 19	<i>KRT19</i>	↑/↓	↑	↑ ¹
Keratin, type II cytoskeletal 8*	<i>KRT8</i>	↑	↓	---
Isoform 1 of Acyl-protein thioesterase 1	<i>LYPLA1</i>	↑ ¹	↑ ¹	---
Microtubule-associated protein RP/EB family member 1	<i>MAPRE1</i>	---	↑	↑
Microfibril-associated glycoprotein 4	<i>MFAP4</i>	↓	↓ ¹	---
Myosin regulatory light chain 12B	<i>MYL12B</i>	↑	↑	↑
Isoform 1 of Nucleoside diphosphate kinase A	<i>NME1</i>	↑	↑	↑
Protein disulfide-isomerase	<i>P4HB</i>	↑	---	↑
Platelet-activating factor acetylhydrolase IB subunit beta	<i>PAFAH1B2</i>	---	↑ ¹	↑
Prohibitin	<i>PHB</i>	---	↑	↑ ¹
Inorganic pyrophosphatase	<i>PPA1</i>	---	↑	↑
Peroxiredoxin 3 isoform b	<i>PRDX3</i>	↑	↑	↑
Proteasome subunit alpha type-1(isoform long)	<i>PSMA1</i>	---	↑	↑ ¹
Proteasome subunit beta type-3	<i>PSMB3</i>	---	↑	↑
Proteasome activator complex subunit 1	<i>PSME1</i>	↑	↑	↑
Proteasome activator subunit 2	<i>PSME2</i>	---	↑	↑
Histone-binding protein RBBP4	<i>RBBP4</i>	---	↑	↑
Ribonuclease inhibitor	<i>RNH1</i>	↑	↑	↑
RPSA 40S ribosomal protein SA	<i>RPSAP15</i>	↑	↑	↑
Protein SEC13 homolog	<i>SEC13</i>	↑ ¹	↑	↑
Isoform 1 of Alpha-1-antitrypsin*	<i>SERPINA1</i>	↓	↓ ¹	---
Stathmin	<i>STMN1</i>	↑	↑ ¹	↑
Tubulin-specific chaperone A	<i>TBCA</i>	↑ ¹	↑ ¹	↑ ¹
Isoform 3 of Tropomyosin alpha-1 chain	<i>TPM1</i>	↓	↑	↑

Protein	Gene Symbol	#15 ^{CS} v #19 ^{CR}	#15 ^{CS} v #1 ^{CR}	#18 ^{CS} v #1 ^{CR}
Isoform 2 of Tropomyosin alpha-3 chain	<i>TPM3</i>	↑	↑	↑
Isoform 2 of Tropomyosin alpha-4 chain	<i>TPM4</i>	---	↑	↑
Tumor protein, translationally-controlled 1	<i>TPT1</i>	↑	↑ ¹	↑
Transthyretin*	<i>TTR</i>	---	↓	↓ ¹
Vimentin*	<i>VIM</i>	↓	↓	↑
14-3-3 protein beta/alpha	<i>YWHAB</i>	---	↑	↑
14-3-3 protein epsilon	<i>YWHAE</i>	↑	↑	↑ ¹
14-3-3 protein gamma	<i>YWHAG</i>	---	↑	↑
14-3-3 protein theta	<i>YWHAQ</i>	↑	↑	↑
14-3-3 protein zeta/delta	<i>YWHAZ</i>	↑	↑	↑

9.4 Discussion

The successful extraction of protein from a clinical breast tumour tissue sample, and the separation of these proteins by 2D-PAGE (section 6.3.2), as well as the successful identification of a selection of these proteins using MALDI-TOF/TOF (section 6.3.2) paved the way for comparative proteomic analysis using clinical tumour tissue. Sufficient protein for proteomic analysis was extracted from approximately 5-10 mm³ of tissue, which allowed 3-6 technical replicates to be performed, as some samples were used more than once within the pair-wise combinations for the three experiments (#15^{CS} and #1^{CR}).

Previous 2D-PAGE experiments based upon samples from cell lines, performed within the laboratory, using the same experimental equipment, have shown much lower numbers of DEPs identified by PDQuest. One study used 2D-PAGE/MS to compare protein expression between 3 parental breast cancer cell lines (MCF7, MDA-MB-231 and T47D) and their radio-resistant sublines (MCF7^{RR}, MDA-MB-231^{RR} and T47D^{RR}). From the three experiments, the average number of DEPs identified by PDQuest from experiments performed within the same range (pH range 4-7; 11 cm gels), was 17 (Smith, Qutob et al. 2009). Whereas, within this project based upon clinical tissue samples, the average number

of DEPs identified by PDQuest across the three experiments was 245. Therefore, whilst allowing for user-variability within PDQuest software, there is > 10-fold increase in the number of DEPs identified from breast cancer clinical tissue samples compared to breast cancer cell lines. However, this is understandable as cell lines are a homogenous collection of cells, so proteins extracted are from only one type of cell which should only exhibit subtle differences in protein expression. In contrast, breast clinical tumour samples contain a variety of different cells (cancerous cells, fibroblasts, pre-malignant cells, inflammatory cells, adipocytes etc), as well as serum which diffuses into the tissue, thus showing a greater variety of protein expression, whilst also increasing the risk of false discovery. Also, the clinical samples to be comparatively analysed cannot be matched exactly; clinical samples used within these experiments were all matched according to histological type (invasive ductal carcinoma) and molecular subtype (ER+, luminal), yet other differences will exist between individuals, as well as expected breast tumour heterogeneity.

Working with clinical samples is a well-recognised challenge, due to their limited supply, their complexity and the dynamic range of serum proteome (section 5.1.2 and section 6.1.1.2). As previously discussed, approaches such as depletion strategies can be employed to improve access to the low abundant proteins of interest. When working with tumour tissue, techniques such as laser capture microdissection (LCM) can be employed, which are able to select tumour cells, thus increasing the proportion of tumour cells within the sample (section 6.1.1.2). However this needs to be balanced with time, sample interference, sample degradation, downstream applications and their sensitivity, as well as tissue availability.

Table 27 shows the number of DEPs identified from each experiment, the number of those that were excised and carried forward for MALDI-TOF/TOF MS analysis, and the number of those that were identified. On average, the number of DEPs identified by PDQuest was

245. This varied depending on the quality of the gels, and the ability to match the spots across CS and CR gels. The average proportion of DEPs excised and carried forward to peptide analysis by MS was 51%. Due to the heterogeneous nature of the samples being used, a high level of stringency was adopted, to minimise false discovery rates. Spots were only excised from gels when not present within a streak or a cluster of spots, not overlapping another spot, and only when the spot was composed of a single protein (checked by 3D viewer within PDQuest). The average number of identifications obtained by MALDI-TOF/TOF peptide analysis was 66%. When the Ultraflex MALDI-TOF/TOF instrument was introduced, the rate of protein identification from 2D-PAGE gels (pH range 5-10), of human endothelial cell line lysates, with standard sample preparation and automated processing was reported to be 77% (74 identifications were obtained from 96 excised and digested protein spots) (Suckau, Resemann et al. 2003).

In these experiments, where protein identification was not achieved from a peptide sample, this may have been due to poor quality or low sample concentration. Some peptide samples were 'mis'-identified as albumin; in this instance, peptides within the sample were matched to the albumin protein, yet the molecular weight of the identification did not match the molecular weight of the spot excised from the 2D-PAGE gel. If depletion strategies, for the removal of albumin and other highly-abundant proteins, had been employed during sample preparation, this may have been avoided, and identification rates may have been higher. This is something that could be considered for future work.

The contamination of 2D-PAGE gel protein spots with keratins from the operative and the laboratory environment is a common problem, and such contamination may hinder the identification of proteins at the mass spectrometry stage. Differentially expressed proteins identified as keratins within an experiment may be a true finding and originate from the

experimental sample, or may have been introduced as a contaminant by the laboratory environment (Lyngholm, Vorum et al. 2011). This was previously a problem within this laboratory, and common keratin peaks had to be excluded from MS data for protein identification. This thesis describes a far more simple in-gel digest method, which reduces the chance for contamination and as a result keratin contamination was not an issue at the protein identification stage, and exclusion of keratin peaks was not required. Several keratins were identified as DEPs within the data, namely keratins 7, 8, 17, and 19 (Appendix 10). These are all epithelial keratins, and the expression of keratins 7, 8 and 19 have all been associated with invasive ductal carcinoma of the breast (Moll, Divo et al. 2008).

The total number of unique DEPs identified from each 2D-PAGE MALDI-TOF/TOF experiment was 53 DEPs (experiment 1), 85 DEPs (experiment 2) and 77 DEPs (experiment 3). Combining the data from the 3 experiments generated a list of 132 unique DEPs, and highlighted those which had been identified by more than one experiment. There were 26 DEPs which had been identified in 3/3 experiments, 31 DEPs which had been identified in 2/3 experiments and 75 DEPs which had been identified in 1/3 experiments (Table 28). Due to the nature of the experiments performed, it is necessary to be stringent when carrying DEPs forward to subsequent data mining and confirmation phases. In order to prioritise the DEPs, and reduce the number of DEPs in the dataset which are likely to be false discoveries, only those which have been identified in more than one experiment (n=57) will be carried forward. This will be reported in Chapter 10.

Another aspect to consider is the identification of the highly-abundant proteins which comprise the plasma proteome, which diffuse into the tissue. The identification of these proteins as DEPs associated with chemotherapy resistance should be interpreted with

caution, as it is usually the low-abundant proteins that are of interest as biomarkers. Therefore, the proteins identified as DEPs, which belong to the group of 22 proteins that make up ~99% of the plasma proteome (Tirumalai, Chan et al. 2003) (section 3.1), have been highlighted (Table 30). Out of the 22 major plasma proteins, 3 have been identified as DEPs; apolipoprotein A-1, alpha-1-antitrypsin and prealbumin (transthyretin).

Table 30: The 22 proteins which constitute ~99% of the plasma proteome

The 22 major plasma proteins (Tirumalai, Chan et al. 2003) are listed, and those which been identified as DEPs in 2D-PAGE MALDI-TOF/TOF experiments are highlighted (red) along with the number of experiments ^(1-3/3) they were identified in.

Albumin	Alpha-2-macroglobulin	Apolipoprotein A-1^{3/3}	Ceruloplasmin
IgG total	IgM total	Apolipoprotein B	C4 Complement
Transferrin	Alpha-1-antitrypsin^{2/3}	Alpha-1-acid glycoprotein	Complement Factor B
Fibrinogen	C3 Complement	Lipoprotein (a)	C1q Complement
IgA total	Haptoglobin	Factor H	C9 Complement
C8 complement	Prealbumin (transthyretin)^{2/3}		

The phenomenon of repeatedly-identified differentially expressed proteins (RIDEPs) within 2D-PAGE-based experiments, reported by Petrak *et al.*, (Petrak, Ivanek et al. 2008) must also be considered when interpreting data. Within their study, Petrak and co-workers compiled a list of DEPs which had been published within 3 volumes of the *Proteomics* journal, from 2004 to 2006, based upon 2D-PAGE data using human tissues. From this, they assembled a list of the ‘TOP15’ proteins most frequently identified and advised ‘extreme caution’ when interpreting the differential expression of these RIDEPs (Petrak, Ivanek et al. 2008). The ‘TOP15’ proteins are listed in Table 18 (section 7.1), and those have been identified within the three experiments are highlighted, along with the number of experiments they were identified in. In total, 6 of the 15 RIDEPs have been identified as DEPs (Table 31).

Table 31: DEPs identified which are included in the ‘TOP15’ RIDEPs

The ‘TOP15’ RIDEPs from 2D-PAGE studies are listed. Those which have been identified within the three 2D-PAGE MS experiments are highlighted (red), including the number of experiments^(1-3/3) they were identified in.

HSP27 (HSPB1)^{1/3}	Enolase 1	Triosephosphate isomerase	Pyruvate kinase M1/M2	Peroxiredoxin 1
Peroxiredoxin 2	Vimentin^{3/3}	Annexin A4^{1/3}	HSC7 1 (HSPA8)	Peptidyl-prolyl isomerase A
Cytokeratin 8 (KRT8)^{2/3}	Cathepsin D	ATP synthase beta subunit^{2/3}	Grp78/Bip (HSPA5)	Rho GDI 1 (ARHGDI A)^{2/3}

When assigning a protein identification, the ‘expect’ value, which states the likelihood that the peptide has been matched by chance is considered, where a lower score indicates a more confident match. It is a way of determining the quality of a match and also considers the score and identity threshold. The number of peptides matched is also considered, where a higher number of matched peptides indicate more confidence in the matched protein. Other factors such as number of missed cleavages and the significance threshold should also be considered. Where a protein identification is obtained with a single peptide match, the Molecular and Cellular Proteomics (MCP) journal stipulate in their guidelines that a MS/MS spectral overlay, showing labelled fragment ions and their masses, is provided for publication purposes. Two or more matched peptides are desirable for confident protein identification. Several of the DEPs identified show protein identifications with 1 peptide match, which have been highlighted in Table 29, and this will be considered for proteins being carried forward through the biomarker discovery pipeline. Of the DEPs which are being carried forward to the data mining phase (identified in at least two experiments), three of these show only one peptide match in each experiment; adenine phosphoribosyl transferase (*APRT*), isoform 1 of acyl-protein thioesterase 1 (*LYPLA1*) and tubulin-specific chaperone A (*TBCA*). Supplementary data, showing MS/MS annotated spectra and

fragment ions, is given for each of these protein identifications (Appendix 11). Although these three identifications show only one peptide match in each experiment, confidence in the identification is also increased by the fact that they have been identified in at least two experiments.

Table 4 (section 3.7), lists DEPs from chemotherapy-resistant breast cancer (ER+) MCF7 sublines identified by 2D-PAGE MALDI-TOF MS and MS/MS analysis. When comparing this list of DEPs to the list of DEPs generated during these 3 experiments using breast cancer (ER+) clinical samples, there are several similarities. These are listed in

Table 32, showing the DEPs identified within these 3 experiments, and the drug(s) to which they have been associated with as a putative markers of resistance (n=11). The DEPs which are present in the 'TOP15' list of RIDEPs in 2D-PAGE studies are also highlighted, as they should be interpreted with caution (Petrak, Ivanek et al. 2008). The main biological functions of the 11 proteins (Figure 43) include cell growth and/or maintenance (4/11) and cell communication (4/11). Other biological processes included protein metabolism, metabolism (energy pathways) and DNA repair. Effective cell growth and maintenance, as well as effective cell communication are paramount for normal tissue homeostasis (section 2.1). In order for cell death to be induced in cancerous cells by chemotherapeutic agents, these pathways also need to be effective in order for important processes such as cell cycle arrest and induction of apoptosis to be executed (section 2.1.2 and section 2.1.4). Resistance to chemotherapy occurs when chemotherapeutic agents fail to achieve their desired effect, which is ultimately the removal of the malignant cell from the system by the induction of apoptosis. Therefore alterations, by the differential expression, in critical mediators of important cellular processes, such as cell communication and signal transduction, have the potential to affect the efficacy of chemotherapeutic agents.

Table 32: DEPs identified, which have also been associated with drug resistance in breast cancer (ER+) MCF7 cell lines, using 2D-PAGE MALDI-TOF MS and MS/MS analysis

The DEPs identified within this chapter, as putative biomarkers of chemo-resistance, which have also been identified within other 2D-PAGE/MS experiments based upon MCF7 cell lines are listed (n=11). The number of experiments they were identified in^{1-3/3} and the direction of expression (≥ 2 fold change) (\uparrow/\downarrow) is also shown. Where these DEPs were identified in MCF7 cell line studies, the direction of expression (≥ 2 fold change) (\uparrow/\downarrow) is shown, as well as the drug the resistance was associated with. The proteins which also listed on the ‘TOP15’ list of RIDEPs are highlighted*(n=3). The drugs which are most relevant to the study; anthracyclines (doxorubicin) and taxanes (paclitaxel) are shown in bold.

Putative biomarkers ^{1-3/3} from clinical tissue	Putative biomarkers from MCF7 chemo-resistant cell lines
Heat shock protein beta -1 ^{1/3} (HSP27) (\uparrow) *	\downarrow Cisplatin (Smith, 2007) \downarrow Paclitaxel (Chuthapisith, 2007)
Tropomyosin alpha-1 chain ^{3/3} ($\uparrow/\uparrow/\downarrow$)	\downarrow Cisplatin (Smith, 2007) \downarrow Doxorubicin (Fu, 2005) \downarrow Mitoxantrone (Fu, 2005) \downarrow Etoposide (Fu, 2005)
14-3-3 epsilon ^{3/3} (\uparrow)	\uparrow Paclitaxel (Chuthapisith, 2007)
Keratin 19 ^{3/3} (cytokeratin 19) (\uparrow/\downarrow)	\downarrow Paclitaxel (Chuthapisith, 2007) \uparrow Doxorubicin (Chuthapisith, 2007) \downarrow Doxorubicin (Fu, 2005) \downarrow Mitoxantrone (Fu, 2005) \downarrow Etoposide (Fu, 2005)
Proliferating cell nuclear antigen ^{1/3} (\uparrow)	\uparrow Paclitaxel (Chuthapisith, 2007)
Stathmin ^{3/3} (\uparrow)	\uparrow Paclitaxel (Chuthapisith, 2007)
14-3-3 sigma ^{1/3} (\uparrow)	\uparrow Doxorubicin (Liu, 2006)
40S ribosomal protein SA ^{3/3} (\uparrow)	\uparrow Mitoxantrone (Fu, 2005) \uparrow Etoposide (Fu, 2005)
ATP synthase beta ^{2/3} (\uparrow)*	\uparrow Doxorubicin (Chuthapisith, 2007)
Keratin 8 ^{2/3} (cytokeratin 8) (\uparrow/\downarrow)*	\downarrow Doxorubicin (Fu, 2005) \downarrow Mitoxantrone (Fu, 2005) \downarrow Etoposide (Fu, 2005)
Prohibitin ^{2/3} (\uparrow)	\uparrow Mitoxantrone (Fu, 2005).

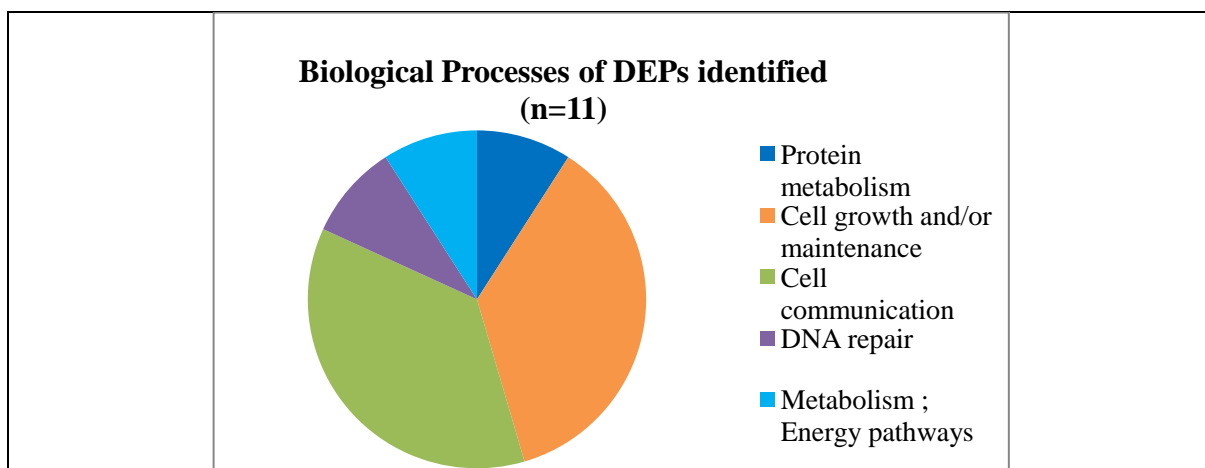


Figure 43: The biological processes of the putative biomarkers identified within both the clinical tissue experiments and the MCF7 breast cancer cell line models from the literature (n=11)

The main biological processes of the putative biomarkers of chemotherapy resistance include cell growth and/or maintenance (4/11) and cell communication (4/11). Other processes include protein metabolism, DNA repair and metabolism (energy pathways).

Within the DEPs generated within this chapter, by the discovery phase of the biomarker discovery pipeline, the proteins which attract attention are the 14-3-3 family of proteins. There are seven mammalian isoforms of 14-3-3 (beta/alpha, epsilon, gamma, eta, theta/tau, sigma and zeta/delta). Five of these (beta/alpha, epsilon, gamma, theta/tau and zeta/delta) have been identified in at least 2 out of the 3 2D-PAGE MALDI-TOF/TOF experiments. 14-3-3 proteins have previously been associated with doxorubicin and paclitaxel chemotherapy resistance in breast cancer cells (Liu, Liu et al. 2006; Chuthapisith 2007), and the isoform which was present on the antibody microarray slide (14-3-3 theta/tau) was shown to be differentially expressed by at least 1.5-fold in 5/5 antibody microarray experiments, as described in chapter 8 (Table 22). These proteins are ubiquitously expressed, and have been shown to have over 300 protein targets (Sluchanko and Gusev 2010), related to apoptosis, cell cycle control, proliferation, transcription and regulation of the cytoskeleton (Sluchanko and Gusev 2010), which is also associated with cytokinesis during the mitosis phase of the cell cycle (section 2.1.1) via interaction of 14-3-3 with

microtubules (Robinson 2010; Zhou, Kee et al. 2010). One of the main roles of 14-3-3 proteins is as apoptosis-related proteins, where they possess anti-apoptotic properties via interaction with pro-apoptotic mediators, such as Bad, Bim and FOXO3a (Sunayama, Tsuruta et al. 2005; Memos, Kataki et al. 2011; Tzivion, Dobson et al. 2011). 14-3-3 proteins have even been shown to interact with the ‘guardian of the genome’; p53 (Rajagopalan, Sade et al. 2010), by increasing its stability via interaction with MDM2 resulting in the prevention of p53 degradation via the ubiquitin-proteasome pathway (section 2.1.2) (Rajagopalan, Sade et al. 2010). 14-3-3 gamma and epsilon have also been associated with increased DNA-binding by p53, leading to cell cycle arrest and apoptosis (Rajagopalan, Sade et al. 2010). It has been shown that the ability of 14-3-3 to sequester Bad, and prevent translocation to the mitochondria, may be regulated by the phosphorylation of 14-3-3 by JNK (Sunayama, Tsuruta et al. 2005). It is thought that 14-3-3 is able to sequester pro-apoptotic proteins in response to pro-survival signals, including those mediated by Akt (Sunayama, Tsuruta et al. 2005; Tzivion, Dobson et al. 2011). An interesting study by Choi and colleagues demonstrated over-expression of 14-3-3 sigma in hepatocellular carcinoma, and showed subsequent silencing of 14-3-3 sigma, which resulted in increased chemotherapy-sensitivity to cisplatin in hepatocellular carcinoma (Choi, Hur et al. 2011), by causing cell cycle arrest and inhibition of tumour cell growth. The 14-3-3 proteins therefore offer a plethora of different mechanisms by which resistance to chemotherapy may be conferred, and therefore warrant further research as potential putative biomarkers of chemotherapy resistance. This will be analysed further in chapter 10.

The nature of the three 2D-PAGE MALDI-TOF/TOF experiments performed in this chapter has provided a list of DEPs, which may be potential biomarkers of chemotherapy resistance. However, the amount and quality of DEPs generated may be improved by the addition of other alternative methods in the sample preparation stage, including pre-fractionation methods, depletion strategies to remove highly abundant proteins, as well the 2D-PAGE stage, where analysis could be performed within different pH ranges with larger polyacrylamide gels to obtain a high degree of separation. Nevertheless, this also has to be balanced by sample availability, as when working with clinical samples there is often a limited supply.

The use of antibody microarray analysis (Chapter 8) as an alternative, yet complementary approach, to the discovery of DEPs as potential biomarkers is extremely valuable. DEPs identified by antibody microarray analysis may not be the same as those identified by 2D-PAGE MALDI-TOF/TOF; data from both platforms will be compared and analysed further in Chapter 10. This has been apparent within the laboratory, where 2D-PAGE MALDI-TOF/TOF and antibody microarray analysis has been performed on the same breast cancer cell line samples, and different lists of DEPs have been generated (Smith, Qutob et al. 2009). This emphasises the complementary approach of the two methods, which together provide the opportunity to identify a wider range of DEPs.

The DEPs identified by 2D-PAGE MALDI-TOF/TOF which will be carried forward to the data mining phase of the biomarker discovery pipeline, which will be described in Chapter 10, will only include those which have been identified in at least two experiments, as listed in Table 29

CHAPTER 10: DATA MINING, CONFIRMATION AND CLINICAL VALIDATION

Chapter Aim:

To analyse the data generated within the discovery phase of the biomarker discovery pipeline, from antibody microarray and 2D-PAGE/MS platforms, for the prioritisation of putative biomarkers to be carried forward. To carry a selection of DEPs identified forward through the confirmation and clinical validation phases

Hodgkinson, V.C., ELFadl, D., Russell, C., Agarwal, V., Garimella, V., Drew, P., Lind, M.J., Cawkwell, C. Predictive biomarkers of chemotherapy resistance in breast cancer: a possible role for 14-3-3 tau and Bid? – *submitted*

Chapter 10. Data mining, confirmation and clinical validation

10.1 Introduction

The importance of data mining has been increasingly reinforced over recent years, to encourage thorough interpretation and determination of the biological relevance behind the large lists of proteins generated within the discovery phase of the biomarker discovery pipeline. Following this, greater understanding can be applied during the selection of protein candidates to be carried forward to confirmation and clinical validation phases.

10.1.1 Data mining

There are several different approaches and a variety of different tools available for data mining. This may involve grouping proteins by biological function, pathway analysis, protein-protein interaction networks, using knowledge bases and applications such as Gene Ontology, PPI Spider, PANTHER, DAVID, STRING, Reactome, IntAct, ArrayUnlock and Ingenuity Pathway Analysis (Deutsch, Lam et al. 2008; Viswanathan, Seto et al. 2008; Antonov, Dietmann et al. 2009; Jimenez-Marin, Collado-Romero et al. 2009; Deighton, Kerr et al. 2010; Malik, Dulla et al. 2010; Croft, O'Kelly et al. 2011; Szklarczyk, Franceschini et al. 2011). Ingenuity Pathway Analysis (IPA) was launched in 2003, and uses the manually-curated Ingenuity Knowledge Base to provide a wide range of high-quality detailed information, including direct and indirect protein interaction networks, common biological functions and canonical pathways present within a dataset. It provides great insight by allowing data to be modelled within complex biological pathways and networks (Jimenez-Marin, Collado-Romero et al. 2009; Deighton, Kerr et al. 2010), thus aiding generation of hypotheses and selection of targets to be carried forward to clinical validation.

10.1.2 Confirmation

The confirmation phase of the biomarker discovery pipeline is used to confirm the differential expression of the protein targets carried forward following data mining, which were identified during the discovery phase. An example of a technique which can be used to assess the expression of a single protein in chemotherapy-sensitive and chemotherapy-resistant samples is western blotting, as discussed previously (section 3.6.2).

10.1.3 Clinical validation

Clinical validation represents the final phase of the biomarker pipeline, where the clinical relevance of each putative biomarker is assessed, to analyse their role as potential prognostic or predictive biomarkers. Proteins which were confirmed to be differentially expressed during the confirmation phase are carried forward to the validation phase, where their differential expression is assessed again, but in a clinical context, commonly using archival tissue samples. This may be performed using immunohistochemistry, as previously discussed (section 3.6.3), which is ideally suited to archival samples. A large sample cohort should be used at this stage, in order to fully assess the strength of each putative biomarker.

10.2 Methodology

10.2.1 Protein selection for data-mining

10.2.1.1 Antibody microarray data

Protein targets identified by antibody microarray analysis have been carried forward to the data mining stage from chapter 8 (section 8.3, Table 21). In total, 38 proteins were identified in the discovery phase, and 37 of these were mapped into the Ingenuity Knowledge Base and included in the analysis (Table 33).

Table 33: DEPs in chemotherapy resistant tumour tissue identified by 5 antibody microarray experiments comparing CR and CS samples.

Significant expression fold change (≥ 1.8) is indicated in bold. For proteins which show ≥ 1.8 -fold change in expression, supporting data from other experiments is shown upward of 1.5-fold. Values considered to be not significant (---) and antibody spots which did not pass the analysis criteria for experimental quality control (\otimes) are also indicated. One gene identifier (indicated *) was not mapped in the Ingenuity® Knowledge Base. Highlighted protein targets were selected for confirmation and pilot validation experiments.

Ab #	Protein	Gene identifier	11 ^{CS} /19 ^{CR}	15 ^{CS} /9 ^{CR}	15 ^{CS} /19 ^{CR}	12 ^{CS} /25 ^{CR}	18 ^{CS} /25 ^{CR}
C8035	Chondroitin sulfate	<i>ACAN</i>	1.56	---	2.00	---	---
P1601	Protein Kinase Ba	<i>AKT1</i>	2.10	---	---	---	---
A8604	Annexin V	<i>ANXA5</i>	---	---	---	2.02	---
B9429	Bcl-xL	<i>BCL2L1</i>	---	2.26	---	1.57	2.62
B3183	tBID	<i>BID</i>	---	2.16	1.55	1.97	1.96
C8854	Caspase 13	<i>CASP13*</i>	---	---	2.69	---	---
C7736	Centrin	<i>CETN1</i>	1.51	---	---	1.90	---
I6139	IKKa	<i>CHUK</i>	1.61	---	---	2.03	---
D1286	Desmosomal protein	<i>DSC1</i>	---	---	1.92	---	---
E2520	Epidermal Growth Factor	<i>EGF</i>	---	---	---	2.13	---
H9286	Acetyl Histone H3 AcLys9	<i>H3F3A</i>	---	---	---	2.21	---
D5567	Dimethyl Histone H3		2.14	---	---	---	---
H9411	HDAC4	<i>HDAC4</i>	---	---	---	---	1.89
P9371	PINCH 1	<i>LIMS1</i>	---	---	1.51	1.90	---
S0315	SAPK3	<i>MAPK12</i>	---	3.76	\otimes	---	---
T5530	Tau	<i>MAPT</i>	1.63	\otimes	2.04	\otimes	---
M9317	MeCP2	<i>MECP2</i>	---	---	---	2.17	---
P6834	Ki-67	<i>MKI67</i>	1.92	---	---	---	---
C3956	cMyc	<i>MYC</i>	---	---	---	2.17	---
M9934	MyD88	<i>MYD88</i>	---	---	2.08	2.18	---
P0084	Pinin	<i>PNN</i>	\otimes	2.54	2.39	1.53	---
P3203	Protein Kinase Cb2	<i>PRKCB</i>	3.20	---	---	1.70	---
P3078	Protein Kinase Cb1		---	---	---	2.06	---
F7926	FAK (pTyr397)	<i>PTK2</i>	---	---	1.51	---	1.95
F8926	FAK (pTyr577)		---	\otimes	---	2.20	---
R8529	RALAR	<i>RALA</i>	---	\otimes	3.70	\otimes	---

Ab #	Protein	Gene identifier	11 ^{CS} /19 ^{CR}	15 ^{CS} /9 ^{CR}	15 ^{CS} /19 ^{CR}	12 ^{CS} /25 ^{CR}	18 ^{CS} /25 ^{CR}
R4904	Reelin	<i>RELN</i>	1.95	⊗	---	⊗	---
R8274	RIP	<i>RIPK1</i>	2.07	---	---	2.56	---
R5145	Rsk1	<i>RPS6KA1</i>	---	---	---	2.33	---
S4047	S6 Kinase	<i>RPS6KB1</i>	---	---	---	2.44	---
S5313	Sir2	<i>SIRT1</i>	---	---	---	2.08	---
H9912	hSNF5 INI1	<i>SMARCB1</i>	---	---	---	2.03	---
S9809	Sp1	<i>SP1</i>	---	---	---	1.82	---
D1314	DRAK1	<i>STK17A</i>	---	2.18	---	---	---
T1827	TBP	<i>TBP</i>	---	---	---	2.08	---
T9191	TRAIL	<i>TNFSF10</i>	---	---	1.83	1.62	---
T0825	Transportin 1	<i>TNPO1</i>	---	---	---	1.84	---
T2780	Tropomyosin	<i>TPM1</i>	---	2.67	---	⊗	---
M6194	Munc13 1	<i>UNC13A</i>	---	---	---	2.08	---
T5942	14-3-3 theta/tau	<i>YWHAQ</i>	1.54	1.90	2.29	2.55	1.52
Z0377	Zyxin	<i>ZYX</i>	7.80	2.01	2.21	2.02	2.63

10.2.1.2 2D-PAGE/MS data

Protein targets identified by 2D-PAGE MALDI-TOF/TOF MS analysis have been carried forward from chapter 9 (section 9.3.7 Table 29) for data mining. This includes proteins which were identified in at least two experiments (n=57). In total, 55 of these proteins were mapped into the Ingenuity Knowledge Base and included in the analysis (Table 34).

10.2.2 IPA

Data was analysed using IPA (Ingenuity Systems, www.ingenuity.com). Each set of data, containing a list of gene symbols, which had been checked against the IPI and NCBI databases, was uploaded into IPA software online.

Genes which were successfully mapped into the Ingenuity Knowledge Base were referred to as ‘network eligible’ molecules, as described in section 4.5. Networks of ‘network eligible’ molecules were then algorithmically generated based on their connectivity.

Canonical pathway analysis of the dataset involved the identification of pathways within the IPA library of canonical pathways that were most significant to the dataset. All molecules mapped within the dataset were considered for canonical pathway analysis, as described in section 4.5.

Table 34: DEPs associated with chemotherapy resistance, identified using 2D-PAGE MALDI-TOF/TOF MS

Three comparative 2D-PAGE MALDI-TOF/TOF MS experiments were performed to identify differentially expressed proteins (DEPs) associated with chemotherapy resistance. The table lists (alphabetically by gene symbol, from the IPI database) those DEPs identified in at least two experiments (n=57), showing ≥ 2 -fold change in expression, along with the direction of change ($\downarrow\uparrow$). Two gene identifiers (indicated*) were not mapped into the Ingenuity Knowledge Base. Highlighted protein targets were selected for confirmation and/or pilot validation experiments.

Protein	Gene identifier	#15 ^{CS} v #19 ^{CR}	#15 ^{CS} v #1 ^{CR}	#18 ^{CS} v #1 ^{CR}
Activator of 90 kDa heat shock protein ATPase homolog 1	<i>AHSA1</i>	↑	↑	---
Annexin A3	<i>ANXA3</i>	---	↑	↑
Serum amyloid P-component	<i>APCS</i>	↓	↓	↓
Apolipoprotein A1	<i>APOA1</i>	↓	↓	↓
Adenine phosphoribosyltransferase	<i>APRT</i>	↑	---	↑
Rho GDP-dissociation inhibitor 1	<i>ARHGDI1</i>	---	↑	↑
Rho GDP-dissociation inhibitor 2	<i>ARHGDI2</i>	↑	↑	↑
ATP synthase subunit beta, mitochondrial	<i>ATP5B</i>	↑	---	↑
Barrier-to-autointegration factor	<i>BANF1</i>	---	↑	↑
Macrophage-capping protein	<i>CAPG</i>	---	↑	↑
Isoform 2 of F-actin-capping protein subunit beta	<i>CAPZB</i>	↑	↑	↑
T-complex protein 1 subunit beta	<i>CCT2</i>	↓	↑	---
Creatine kinase B-type	<i>CKB</i>	↓	↓	↓
Chloride intracellular channel protein 1	<i>CLIC1</i>	↑	↑	↑
Coactosin-like protein	<i>COTL1</i>	---	↑	↑
Cellular retinoic acid-binding protein 2	<i>CRABP2</i>	↑	↑	↑
Isoform 1 of Eukaryotic translation initiation factor 5A-1	<i>EIF5A</i>	---	↑	↑

Protein	Gene identifier	#15^{CS} v #19^{CR}	#15^{CS} v #1^{CR}	#18^{CS} v #1^{CR}
Ferritin light chain	<i>FTL</i>	↓	↓	↑
Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	<i>GPD1</i>	↓	↓	---
Glutathione S-transferase omega-1	<i>GSTO1</i>	---	↑	↑
Glutathione S-transferase P	<i>GSTP1</i>	↓	↑	↑
HEBP2 protein (fragment)	<i>HEBP2*</i>	---	↑	↑
highly similar to Heat-shock protein beta-6	<i>HSPB6</i>	↓	↓	---
Keratin, type I cytoskeletal 19	<i>KRT19</i>	↑/↓	↑	↑
Keratin, type II cytoskeletal 8	<i>KRT8</i>	↑	↓	---
Isoform 1 of Acyl-protein thioesterase 1	<i>LYPLA1</i>	↑	↑	---
Microtubule-associated protein RP/EB family member 1	<i>MAPRE1</i>	---	↑	↑
Microfibril-associated glycoprotein 4	<i>MFAP4</i>	↓	↓	---
Myosin regulatory light chain 12B	<i>MYL12B</i>	↑	↑	↑
Isoform 1 of Nucleoside diphosphate kinase A	<i>NME1</i>	↑	↑	↑
Protein disulfide-isomerase	<i>P4HB</i>	↑	---	↑
Platelet-activating factor acetylhydrolase IB subunit beta	<i>PAFAH1B2</i>	---	↑	↑
Prohibitin	<i>PHB</i>	---	↑	↑
Inorganic pyrophosphatase	<i>PPA1</i>	---	↑	↑
Peroxiredoxin 3 isoform b	<i>PRDX3</i>	↑	↑	↑
Proteasome subunit alpha type-1(isoform long)	<i>PSMA1</i>	---	↑	↑
Proteasome subunit beta type-3	<i>PSMB3</i>	---	↑	↑
Proteasome activator complex subunit 1	<i>PSME1</i>	↑	↑	↑
Proteasome activator subunit 2	<i>PSME2</i>	---	↑	↑
Histone-binding protein RBBP4	<i>RBBP4</i>	---	↑	↑
Ribonuclease inhibitor	<i>RNH1</i>	↑	↑	↑
RPSA 40S ribosomal protein SA	<i>RPSAP15*</i>	↑	↑	↑
Protein SEC13 homolog	<i>SEC13</i>		↑	↑
Isoform 1 of Alpha-1-antitrypsin	<i>SERPINA1</i>	↓	↓	---
Stathmin	<i>STMN1</i>	↑	↑	↑
Tubulin-specific chaperone A	<i>TBCA</i>	↑	↑	↑
Isoform 3 of Tropomyosin alpha-1 chain	<i>TPM1</i>	↓	↑	↑
Isoform 2 of Tropomyosin alpha-3 chain	<i>TPM3</i>	↑	↑	↑

Protein	Gene identifier	#15^{CS} v #19^{CR}	#15^{CS} v #1^{CR}	#18^{CS} v #1^{CR}
Isoform 2 of Tropomyosin alpha-4 chain	<i>TPM4</i>	---	↑	↑
Tumor protein, translationally-controlled 1	<i>TPT1</i>	↑	↑	↑
Transthyretin	<i>TTR</i>	---	↓	↓
Vimentin	<i>VIM</i>	↓	↓	↑
14-3-3 protein beta/alpha	<i>YWHA B</i>	---	↑	↑
14-3-3 protein epsilon	<i>YWHA E</i>	↑	↑	↑
14-3-3 protein gamma	<i>YWHA G</i>	---	↑	↑
14-3-3 protein theta	<i>YWHA Q</i>	↑	↑	↑
14-3-3 protein zeta/delta	<i>YWHA Z</i>	↑	↑	↑

10.2.3 Confirmation: western blotting

Western blotting was performed using the method previously described, following optimisation of antibodies (section 6.3.3.3), using the samples shown in Figure 15. Briefly, protein lysates were diluted with Laemmli buffer (62.5 mM Tris-HCl [pH 6.8], 2% SDS, 5% β -mercaptoethanol, 1% protease inhibitor mix and 0.00125% bromophenol blue). Twenty micrograms of protein was separated by molecular weight using one-dimensional gel electrophoresis on a pre-cast 12% ‘Precise Protein’ polyacrylamide gel (#25222, Pierce) at a constant voltage of 140 V for 40 min. Proteins were transferred onto nitrocellulose membrane using the iBlot dry transfer system (#IB3010-01, Invitrogen). The membrane was blocked in 5% non-fat milk on an orbital shaker, followed by incubation with primary and subsequently secondary antibodies (diluted in blocking buffer), as described in Table 5. Bands were detected using the Supersignal West Pico Chemiluminescent Substrate Kit (#34078, Pierce). Films were scanned using a GS800 calibrated densitometer (Bio-Rad) with Quantity One software (Bio-Rad), which was also used for data normalisation of proteins of interest against a loading control. Fold changes ≥ 2 were considered significant.

10.2.4 Clinical validation: immunohistochemistry

10.2.4.1 Archival pre-treatment core biopsy samples

Pre-treatment archival core biopsy samples were obtained from a previously characterised cohort (Garimella 2007) (section 4.2) (REC 03/00/038). Patient consent was obtained to allow access to pre-treatment core biopsy samples from patients with locally advanced breast cancer who were treated with neoadjuvant anthracycline based chemotherapy (Garimella 2007). All patients in this cohort had histologically proven breast cancer with a primary tumour of at least 3cm and were treated with 6 cycles of 60 mg/m² epirubicin and 600 mg/m² cyclophosphamide with 200 mg/m² infusional 5-fluorouracil (infusional FEC), administered at 3-weekly intervals. Serial DCE-MRI scans were performed in order to assess tumour response. In total, 35 archival tissue samples were obtained with 36 locally advanced breast cancers (one patient had bilateral breast cancer). These comprised 75% ER-positive tumours.

10.2.4.2 Immunohistochemistry

Immunohistochemistry was performed as previously described (section 4.7). Following blocking of endogenous peroxidase, 4 µm tissue sections were boiled for 3 min in a pressure cooker containing Antigen Unmasking Solution Low pH (#H-3300, Vector Laboratories). Blocking serum (#PK-7800, Vector Laboratories) was applied for 10 min. Following incubation with primary antibody a biotinylated secondary antibody from the Vectastain Universal Quick kit (#PK-7800, Vector Laboratories) was applied for 20 min. A streptavidin/peroxidase complex reagent from the Vectastain Universal Quick kit (#PK-7800, Vector Laboratories) was applied for 10 min and visualisation was achieved with 3', 3'-diaminobenzidine. Each batch of slides included a negative control (primary antibody

omitted). Slides were scored blindly and independently by at least two observers and discussed with a consultant histopathologist to give a consensus result, using the scoring methods described in section 4.7.11. The Kappa test was also performed to assess inter-observer variability, using the method described (section 4.7.13). Statistical analysis was performed using SPSS (version 14).

10.3 Results

10.3.1 Discovery-phase data

Lists of DEPs were generated within the discovery phase of the biomarker discovery pipeline by antibody microarray analysis (chapter 8) (Table 22) and 2D-PAGE MALDI-TOF/TOF MS analysis (chapter 9) (Table 29). Five experiments were performed using antibody microarray analysis and three experiments were performed using 2D-PAGE MALDI-TOF/TOF MS analysis. Only two DEPs were identified by both proteomic platforms; tropomyosin alpha-1 (*TPM1*) and 14-3-3 theta/tau (*YWHAQ*). The annexin protein family was also common across both proteomic platforms where antibody microarray analysis identified annexin a5 (*ANXA5*) and 2D-PAGE MALDI-TOF/TOF MS analysis identified annexin a3 (*ANXA3*). Only one pair of samples (#15^{CS} versus #19^{CR}) was analysed by both proteomic platforms, and only one DEP was commonly identified; 14-3-3 theta/tau (*YWHAQ*). Analysis performed by 2D-PAGE MALDI-TOF/TOF MS identified tropomyosin alpha-1 (*TPM1*) as a DEP in this sample pair, yet the differential expression value of this protein by antibody microarray analysis was not regarded as significant. Analysis using IPA software for both sets of data, from the two proteomic platforms, will reveal canonical pathways within each dataset, as well as pathways which may be common to both datasets.

10.3.2 IPA

Both sets of data, from antibody microarray analysis and 2D-PAGE/MS analysis were individually analysed using IPA software. The two datasets were also combined for further analysis using IPA software.

10.3.2.1 Antibody microarray data

In total, 37/38 of the molecules within the data set were mapped into the Ingenuity Knowledge Base. The top network identified from the data was ‘cellular assembly and organisation, cell-to-cell signalling and interaction, cell death’, with a score of 77, which contained all 37 of the focus molecules. The top canonical pathway identified was ‘ERK5 signalling’, which included 6 of the DEPs from the data set (Figure 44). Other top canonical pathways include ‘IL-8 signalling’ and ‘myc-mediated apoptosis’, which included 8 DEPs and 5 DEPs respectively. Other canonical pathways included 14-3-3 mediated signalling (6 DEPs), apoptosis signalling (4 DEPs), death receptor signalling (4 DEPs), ERK MAPK signalling (6 DEPs), molecular mechanisms of cancer (8 DEPs), mTOR signalling (4 DEPs), NF-kB signalling (6 DEPs), p70S6K signalling (5 DEPs), PI3K/AKT signalling (6 DEPs) and PTEN signalling (5 DEPs). These are all listed in Table 35, along with the DEPs associated with each pathway.

10.3.2.2 2D-PAGE/MS data

In total, 55/57 of the DEPs uploaded into IPA were mapped into the Ingenuity Knowledge Base. The top networks identified included ‘drug metabolism, glutathione depletion in liver, lipid metabolism’, with a score of 81 and 44 focus molecules, and ‘nucleic acid metabolism, small molecule biochemistry and protein synthesis’, with a score of 13 and 12 focus molecules. The top canonical pathway identified by IPA was ‘cell cycle: G2/M DNA

damage checkpoint regulation, which contained 5 of the DEPs (Figure 45) Other top canonical pathways included myc-mediated apoptosis signalling (5 DEPs), ERK5 signalling (5 DEPs), 14-3-3 mediated signalling (6 DEPs) and IGF-1 signalling (5 DEPs). Other canonical pathways included ERK MAPK signalling (4 DEPs), p70S6K signalling (5 DEPs), PI3K/AKT signalling (5 DEPs), Protein Kinase A signalling (6 DEPs) and the protein ubiquitination pathway (5 DEPs). These are listed in Table 36, along with the DEPs associated with each pathway

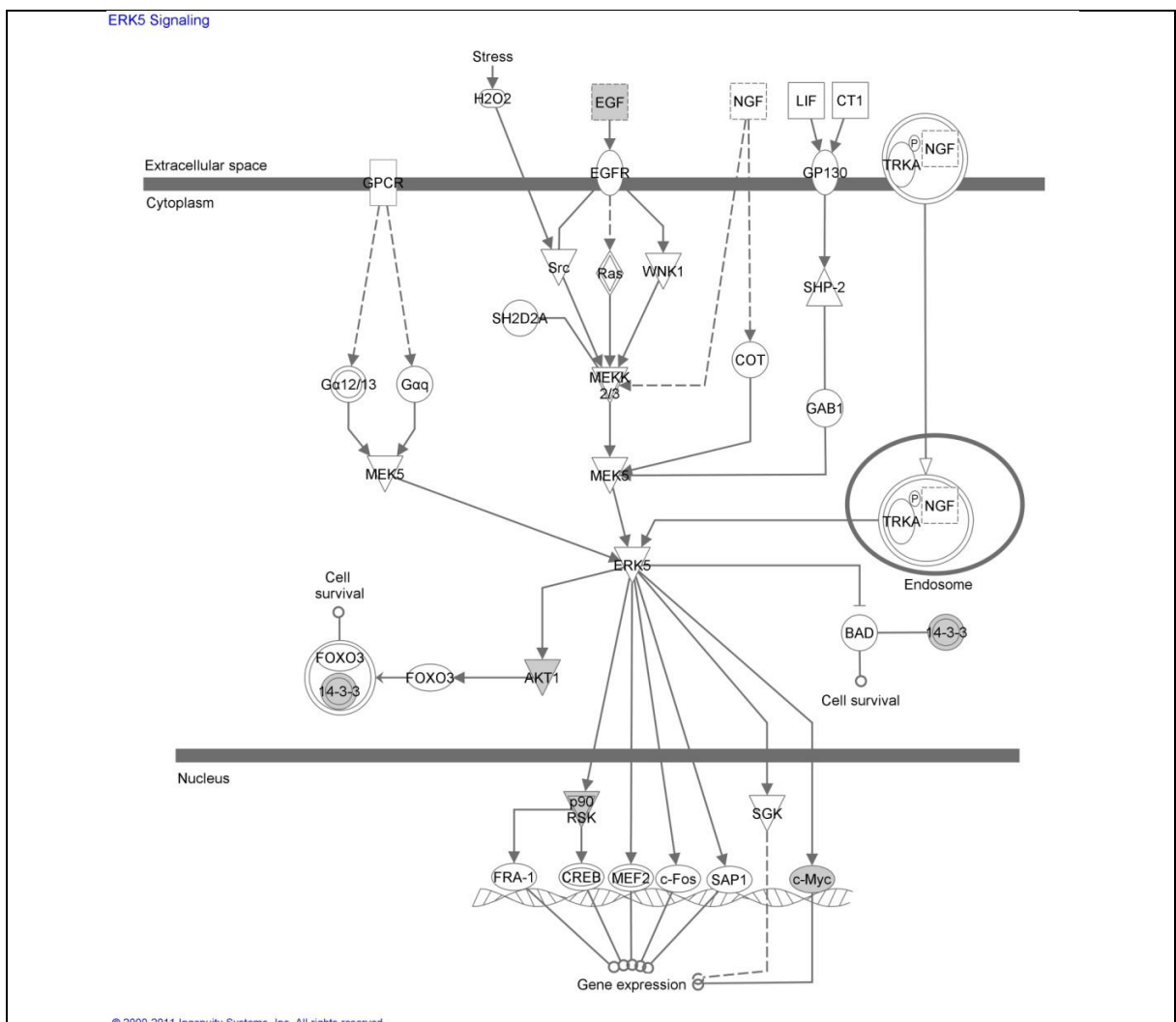


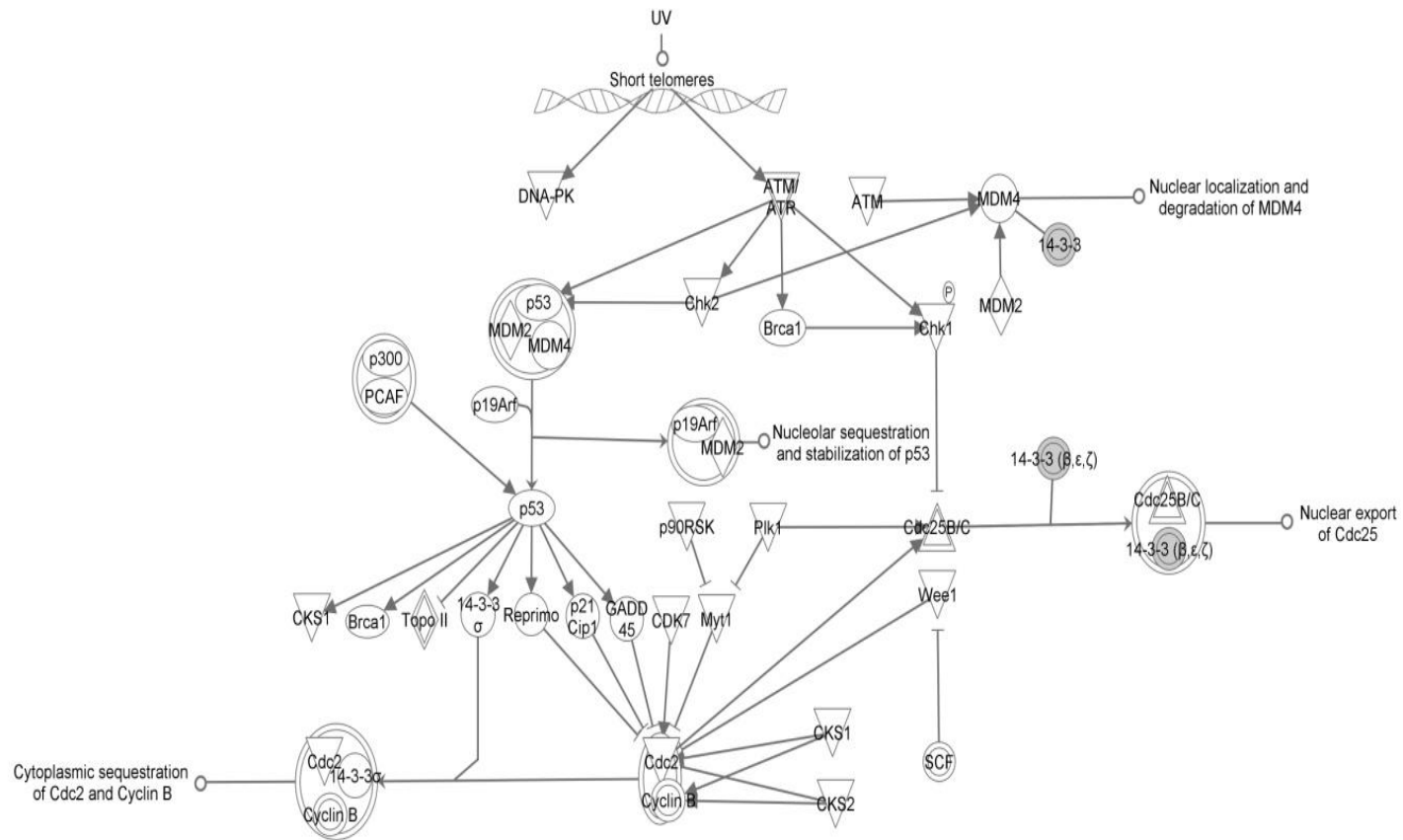
Figure 44: Top canonical pathway identified by IPA from the antibody microarray data set: ERK5 signalling

Six of the DEPs identified by microarray analysis are included in the ERK5 signalling pathway; AKT1, EGF, MYC, RPS6KA1, RPS6KB1 and YWHAQ (14-3-3 theta/tau) (also see Table 35)

Table 35: Canonical pathways identified by IPA performed on DEPs generated from antibody microarray analysis.

The table shows the most relevant canonical pathways identified by IPA (n=13), that contain at least 4 molecules from the data. The molecules (gene identifiers) associated with each of the canonical pathways are listed, along with the number of canonical pathways ⁽⁻⁾ each of them appeared in. For each canonical pathway, the number of matched molecules from the data is shown along with the ratio.

Gene identifier	Canonical pathways identified by IPA												
	ERK5 signalling	IL-8 signalling	Myc-mediated apoptosis	14-3-3 mediated signalling	Apoptosis signalling	Death receptor signalling	ERK MAPK signalling	mTOR signalling	NF-κB signalling	P70S6K signalling	PI3K/AKT signalling	PTEN signalling	Molecular mechanisms of cancer
YWHAQ (14-3-theta/tau) ⁽⁶⁾	✓		✓	✓			✓			✓	✓		
MYC ⁽⁴⁾	✓		✓				✓						✓
RPS6KB1 ⁽⁶⁾	✓	✓						✓		✓	✓	✓	
AKT1 ⁽¹⁰⁾	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓
EGF ⁽³⁾	✓	✓							✓				
RPS6KA1 ⁽⁵⁾	✓			✓	✓		✓	✓					
PTK2 ⁽³⁾		✓										✓	✓
BCL2L1 ⁽⁵⁾		✓			✓						✓	✓	✓
CHUK ⁽⁶⁾		✓			✓	✓			✓		✓	✓	
MAPK12 ⁽⁴⁾		✓	✓	✓									✓
PRKCB ⁽⁷⁾		✓		✓			✓	✓	✓	✓			✓
BID ⁽⁴⁾			✓		✓	✓							✓
MAPT ⁽²⁾				✓						✓			
RIPK1 ⁽²⁾						✓			✓				
TNFSF10 ⁽¹⁾						✓							
PTK2 ⁽¹⁾							✓						
H3F3A/H3F3B ⁽¹⁾							✓						
MYD88 ⁽¹⁾									✓				
LIMS1 ⁽¹⁾											✓		
RALA ⁽¹⁾													✓
# molecules	6	8	5	6	4	4	6	4	6	5	6	5	8
ratio	9.68E-02	4.49E-02	8.33E-02	5.22E-02	4.35E-02	6.45E-02	3.03E-02	2.74E-02	3.53E-02	4.1E-02	4.65E-02	4.17E-02	2.23E-02



© 2000-2011 Ingenuity Systems, Inc. All rights reserved.

Figure 45: Top canonical pathway identified by IPA from the 2D-PAGE/MS data set: Cell cycle. G2/M DNA damage checkpoint regulation

Five of the DEPs identified by 2D-PAGE/MS analysis are included in cell cycle: G2/M DNA damage checkpoint regulation; YWHAQ (14-3-3 theta/tau (θ/τ)), YWHAG (14-3-3 gamma (γ)), YWHAB (14-3-3 beta/alpha (β/α)), YWHAE (14-3-3 epsilon (ϵ)) and YWHAZ (14-3-3 zeta/delta (ζ/δ)). Also see Table 36.

Table 36: Canonical pathways identified by Ingenuity Pathway Analysis performed on DEPs generated from 2D-PAGE MALDI-TOF/TOF MS analysis.

The table shows the most relevant canonical pathways identified by IPA (n=10), that contain at least 4 molecules from the data. The molecules (gene identifiers) within the data that are associated with each of the canonical pathways are listed, along with the number of canonical pathways ^(c) each of them appeared in. For each canonical pathway, the number of matched molecules from the data is shown along with the ratio.

Gene identifier	Canonical pathways identified by IPA									
	Cell cycle: G2/M DNA damage checkpoint regulation	Myc-mediated apoptosis	ERK5 signalling	14-3-3 mediated signalling	IGF-1 signalling	PI3K/AKT signalling	P70S6K signalling	Protein Kinase A signalling	Protein ubiquitination pathway	ERK/MAPK signalling
YWHAQ (14-3-3 theta/tau) ⁽⁹⁾	√	√	√	√	√	√	√	√		√
YWHAG (14-3-3 gamma) ⁽⁹⁾	√	√	√	√	√	√	√	√		√
YWHAE (14-3-3 epsilon) ⁽⁸⁾	√	√	√	√	√	√	√	√		
YWHAB (14-3-3 beta/alpha) ⁽⁹⁾	√	√	√	√	√	√	√	√		√
YWHAZ (14-3-3 zeta/delta) ⁽⁹⁾	√	√	√	√	√	√	√	√		√
VIM ⁽¹⁾				√						
MYL12B ⁽¹⁾								√		
PSMB3 ⁽¹⁾									√	
PSME1 ⁽¹⁾									√	
PSME2 ⁽¹⁾									√	
PSMA1 ⁽¹⁾									√	
HSPB6 ⁽¹⁾									√	
# molecules	5	5	5	6	5	5	5	6	5	4
ratio	1.04E-01	8.33E-02	8.06E-02	5.22E-02	4.90E-02	3.88E-02	4.1E-02	1.96E-02	1.86E-02	2.02E-02

10.3.2.3 Combined antibody microarray and 2D/MS data

Data was combined from both antibody microarray and 2D/MS analysis, which generated a total of 93 molecules (38 from antibody microarray plus 57 from 2D/MS (95), minus two duplicate entries (*TPM1* and *YWHAQ* (14-3-3 theta/tau)). In total, 90/93 molecules were mapped into the Ingenuity Knowledge Base. The top networks identified included ‘cellular assembly and organisation’, with a score of 117 containing 65 focus molecules, and ‘cellular development, cellular growth and proliferation’, with a score of 25 containing 22 focus molecules. The top canonical pathway was ‘ERK 5 signalling’, which contained 10 DEPs. Other top canonical pathways included ‘myc-mediated signalling’ (9 DEPs), ‘14-3-3-mediated signalling’ (11 DEPs) (Figure 46) and ‘PI3K/AKT signalling’ (10 DEPs).

10.3.3 Confirmation: western blotting

Protein candidates selected for confirmation using western blotting include 14-3-3 theta/tau, 14-3-3 epsilon, Bcl-xL and tBID. Western blotting for 14-3-3 tau and tBID were unsuccessful, as a result of poor quality non-specific secondary and primary antibodies. However an antibody which recognised four 14-3-3 isoforms (14-3-3 eta, 14-3-3 beta/alpha, 14-3-3 theta/tau and 14-3-3 sigma), and an antibody specific to full-length BID were successful. Western blotting demonstrated significant (≥ 2 -fold) up-regulation of 14-3-3 isoforms (including 14-3-3 epsilon), Bcl-xL and full-length BID expression in chemotherapy-resistant samples (Figure 47).

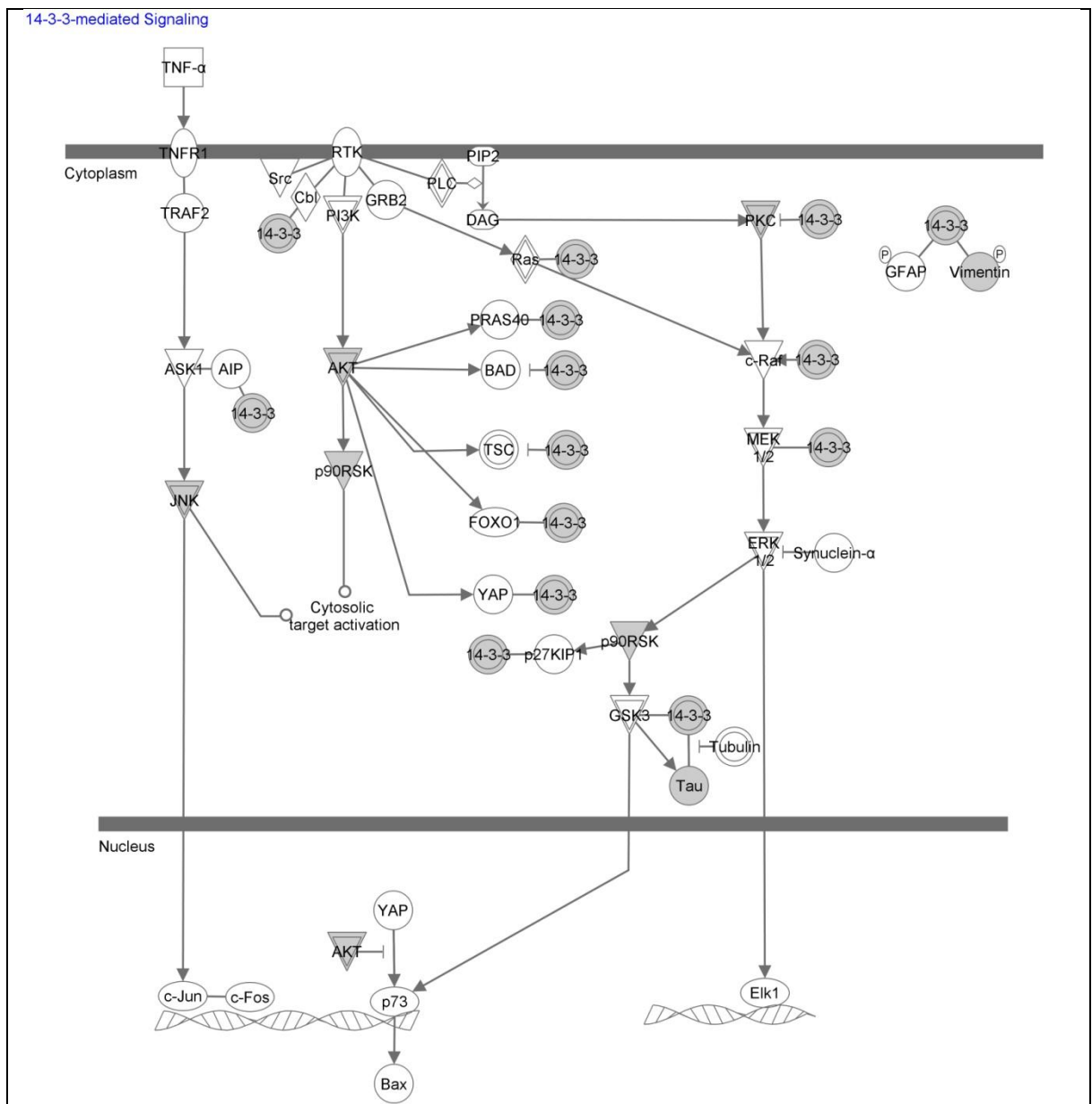


Figure 46: One of the top canonical pathway identified by IPA from combined antibody microarray 2D-PAGE/MS data set: 14-3-3 mediated signalling

Eleven of the DEPs identified by combined antibody microarray and 2D-PAGE/MS analysis are involved in 14-3-3 mediated signalling. The molecules listed in red originate from antibody microarray data and the molecules listed in blue originate from the 2D-PAGE/MS data. Molecules which were present in both antibody microarray and 2D-PAGE/MS data are listed in purple. These molecules include YWHAQ (14-3-3 theta/tau), YWHAG (14-3-3 gamma), AKT1, YWHAE (14-3-3 epsilon), YWHAB (14-3-3 beta/alpha), YWHAZ (14-3-3 zeta/delta), VIM, MAPT, RPS6KA1, MAPK12 and PRKCB.

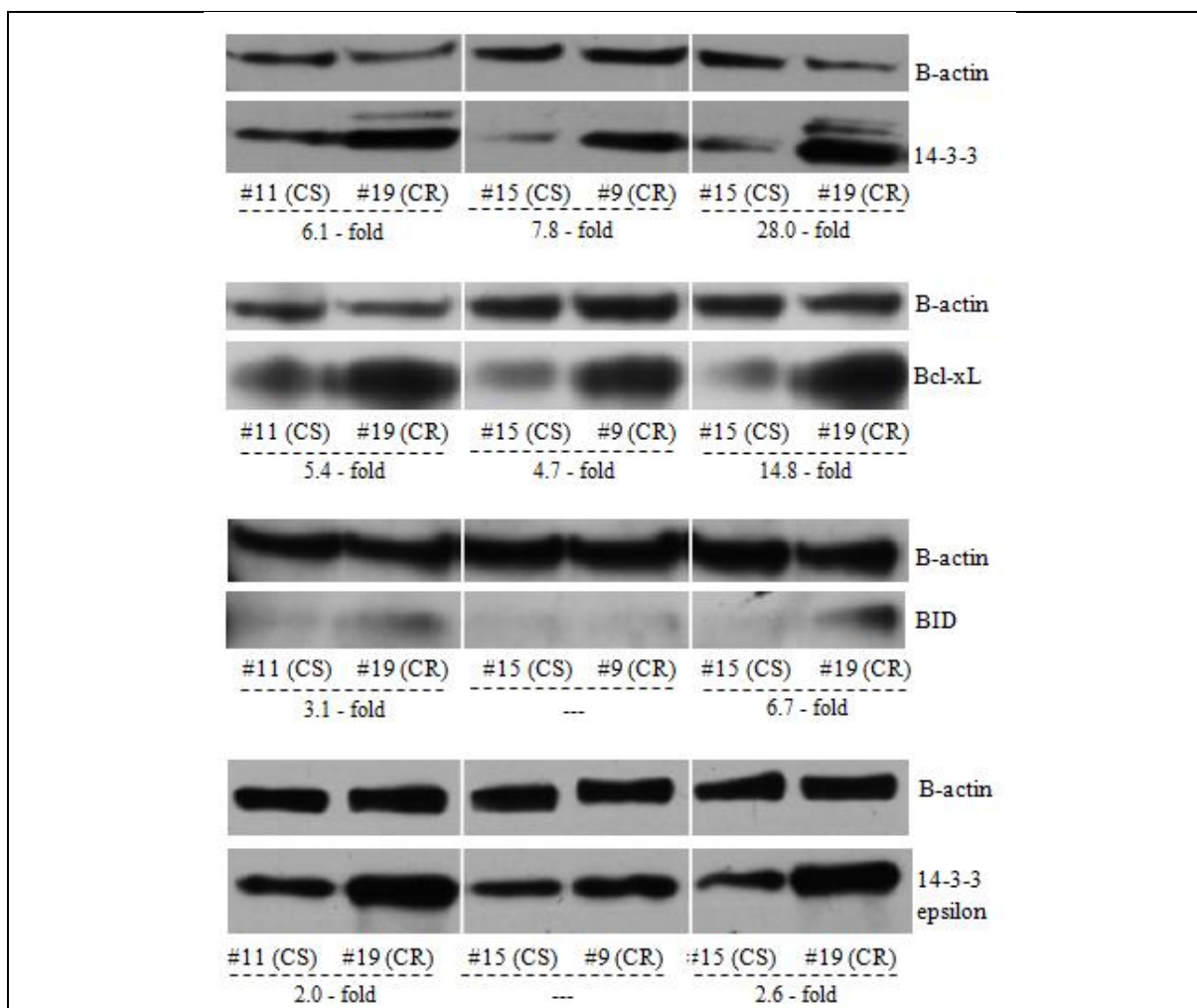


Figure 47: Confirmation of DEPs using western blotting

The up-regulation of 14-3-3 isoforms, (including 14-3-3 epsilon), Bcl-xL and BID was significantly (≥ 2 -fold change in expression) associated with chemotherapy-resistance. Beta-actin was used as loading control. All antibodies were used at optimised concentrations (section 6.3.3.3) as outlined in Table 5.

10.3.4 Clinical validation: immunohistochemistry

Immunohistochemistry was used to assess the DEPs in a clinical context, as a small pilot study. Archival pre-treatment biopsy samples from a series of breast cancer patients who later received neoadjuvant chemotherapy were analysed for expression of 14-3-3 theta/tau, Bcl-xL and tBID. Following assessment of slides, positive staining of 14-3-3 theta/tau was recorded when strong nuclear membrane positivity was seen in at least 20% of invasive carcinoma cells (Figure 48 (A)). Positive staining was seen in 8/9 (88%) chemotherapy-

resistant (CR) samples, compared with 9/22 (40%) of chemotherapy-sensitive (CS) samples ($p=0.020$, Fisher's exact test). Positive staining of the apoptotic protein tBID was recorded when moderate to strong cytoplasmic staining was observed in at least 50% invasive carcinoma cells (Figure 48 (C)). Positive staining of tBID was seen in 13/19 (68%) of CS samples, compared with 2/9 (22%) of CR samples ($p=0.041$, Fisher's exact test). Positive staining of the anti-apoptotic protein Bcl-xL was recorded when moderate to strong cytoplasmic staining was observed in at least 50% invasive carcinoma cells. There was no significant difference in expression of Bcl-xL between CS and CR samples. Histological scoring for both 14-3-3 theta/tau and tBID by two independent observers showed 100% agreement. This represented 'perfect agreement' by Kappa statistic criteria, which is used to assess inter-observer variability, with a value of 1.00 (section 4.7.13, Table 12). In order to fully assess the roles of 14-3-3 theta/tau and tBID as putative biomarkers of resistance to neoadjuvant chemotherapy in breast cancer, more extensive clinical validation in a larger sample cohort would be required.

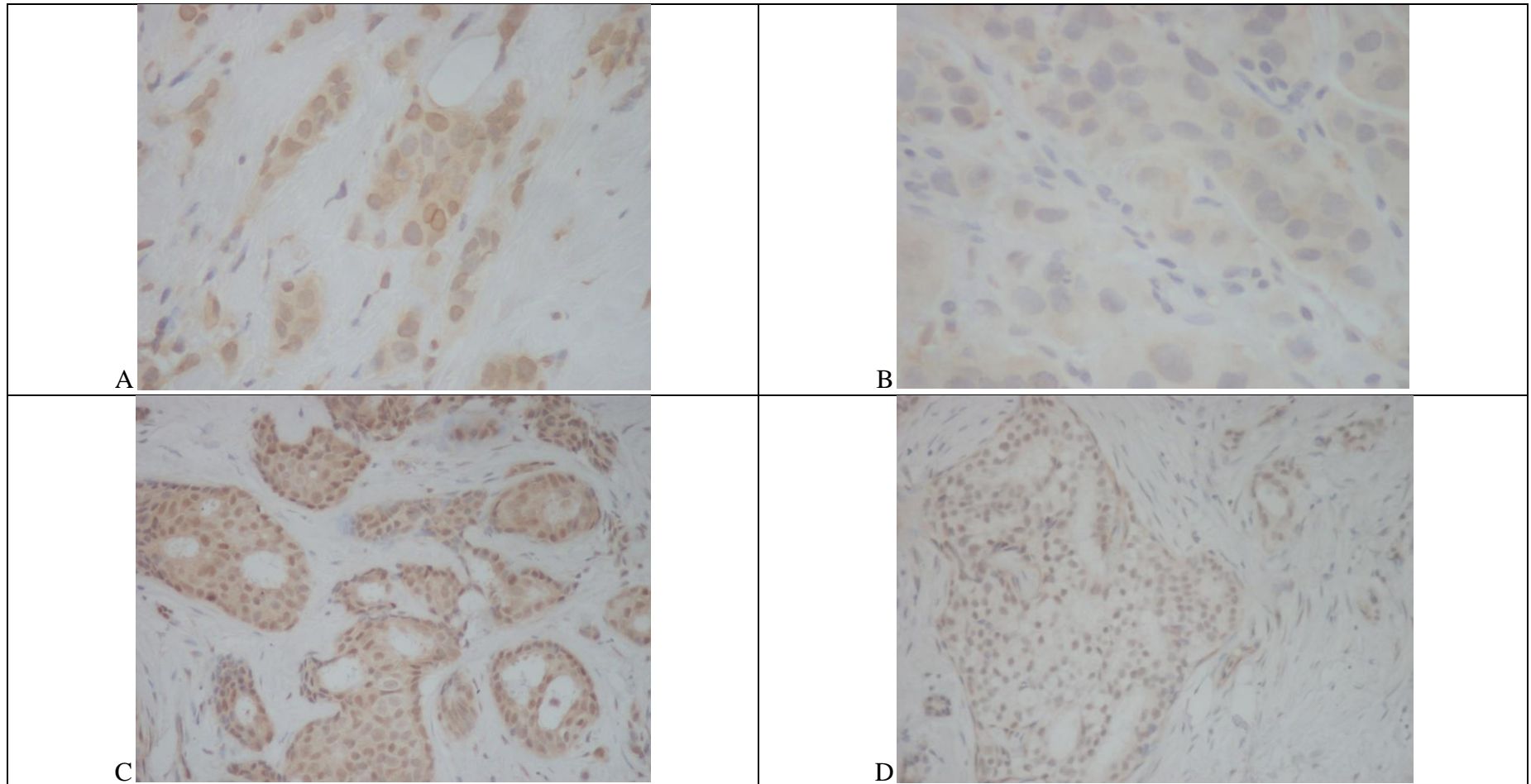


Figure 48: Immunohistochemical analysis of 14-3-3 theta/tau and tBID expression in invasive breast carcinoma cells

A: Positive staining of 14-3-3 theta/tau was recorded when cytoplasmic staining with and strong nuclear membrane positivity was seen in at least 20% of invasive carcinoma cells. B: Weak staining of 14-3-3 theta/tau, which was recorded as negative. C: Positive staining of tBID was recorded when moderately strong cytoplasmic staining was observed in at least 50% invasive carcinoma cells. D: Weak staining of tBID, which was recorded as negative. Magnification: A and B, x 400; C and D, x 200.

10.4 Discussion

During the discovery-phase of the biomarker discovery pipeline, antibody microarray analysis (chapter 8) and 2D-PAGE MALDI-TOF/TOF analysis (chapter 9) was used to identify DEPs associated with chemotherapy resistance in breast cancer. When comparing the lists of DEPs generated from each proteomic platform, only two proteins had been commonly identified by both methods; 14-3-3 theta/tau (*YWHAQ*) and tropomyosin alpha-1 (*TPM1*). One protein family had been identified across both proteomic platforms; annexins, where annexin a5 (*ANXA5*) was identified by antibody microarray analysis and annexin a3 (*ANXA3*) was identified by 2D-PAGE MALDI-TOF/TOF analysis. IPA was then used to analyse each of the two datasets, to highlight any canonical pathways present within individual datasets and to also highlight any canonical pathways common to both datasets.

A total of 37 out of 38 DEPs identified by antibody microarray analysis were mapped into the Ingenuity Knowledge Base and included in the analysis. A total of 55 out of 57 DEPs identified by 2D-PAGE/MS analysis were mapped into the Ingenuity Knowledge Base and analysed using IPA. Analysis using this software revealed canonical pathways, which were found to be associated with the DEPs within each of the data sets. These are listed in Table 35 and Table 36. The top canonical pathways included ‘ERK5 signalling’ (from the antibody microarray data set) and ‘cell cycle: G2/M DNA damage checkpoint regulation’ (from the 2D-PAGE/MS data set), which contained 6 DEPs and 5 DEPs respectively. These tables (Table 35 and Table 36) also highlight the DEPs which appeared in several canonical pathways, and therefore warrant further research, due to their roles in multiple significantly-identified canonical pathways. When the datasets were combined for IPA, the top canonical pathway was ERK5 signalling, which contained 10 DEPs. This canonical pathway had already been identified as a ‘top canonical pathway’ by individual analysis of

both datasets by IPA. Other top canonical pathways identified from the combined dataset included myc-mediated signalling, 14-3-3-mediated signalling and PI3K/AKT signalling, which had all been identified previously during both sets of individual analysis. Therefore, combining the datasets did not identify any new canonical pathways; it only increased the number of molecules associated with each of the top canonical pathways. From the antibody microarray data set (Table 33), the top 5 molecules involved in multiple pathways include YWHAQ (14-3-3 theta/tau)⁽ⁿ⁼⁶⁾, RPS6KB1⁽ⁿ⁼⁶⁾, AKT1⁽ⁿ⁼¹⁰⁾, CHUK⁽ⁿ⁼⁶⁾ and PRKCB⁽ⁿ⁼⁷⁾. From the 2D-PAGE/MS data set (Table 34), the top 5 molecules involved in multiple pathways include YWHAQ (14-3-3 theta/tau)⁽ⁿ⁼⁹⁾, YWHAG (14-3-3 gamma)⁽ⁿ⁼⁹⁾, YWHAE (14-3-3 epsilon)⁽ⁿ⁼⁸⁾, YWHAB (14-3-3 beta/alpha)⁽ⁿ⁼⁹⁾ and YWHAZ (14-3-3 zeta/delta)⁽ⁿ⁼⁹⁾. It is important to note that the 725 antibodies spotted onto the antibody microarray slide are pre-selected from canonical cell signalling pathways, so some clustering of DEPs can be expected, and therefore possible bias towards selected canonical pathways.

10.4.1 Apoptosis-related proteins

Aberrant apoptotic signalling has a well established putative role in tumour response to chemotherapy (Pommier, Sordet et al. 2004; Yip and Reed 2008; Indran, Tufo et al. 2011). It is this critical pathway that delivers the desired effect of the chemotherapeutic agent; death of the cancerous cell by apoptosis, thus acting as an indirect drug mechanism of action. Therefore it is understandable that alterations in the mediators of this pathway, may present possible mechanisms of drug resistance. Therefore candidates initially selected for confirmation phases included the apoptosis-related proteins, Bcl-xL and tBID. Western blotting revealed significant up-regulation of Bcl-xL in chemotherapy-resistant tumours. The expression of tBID could not be assessed using western blotting, however an antibody

against full-length BID demonstrated the significant up-regulation of this protein in chemotherapy-resistant tumour samples. To assess the clinical relevance of these proteins as putative markers of chemotherapy resistance, a pilot immunohistochemical study was performed using archival samples from a well-characterised patient cohort (Garimella 2007). The expression of tBID and Bcl-xL was assessed at this stage. Positive cytoplasmic staining of the cleavage product of BID (tBID) was significantly associated with chemotherapy-resistant tumours ($p=0.04$). The expression of Bcl-xL was not shown to be associated with chemotherapy-resistance in the clinical setting at this stage. The differential expression of the Bcl-2-like apoptosis-related proteins BID (pro-apoptotic) and Bcl-xL (anti-apoptotic) highlights the potential involvement of the apoptotic pathway in chemotherapy-resistance. Western blot data showed increased expression of BID (full-length) in chemotherapy-resistant samples, and immunohistochemical analysis revealed decreased expression of BID in its active cleaved form (tBID). This data suggests that the reduced levels of apoptosis seen in chemotherapy-resistant tumours may be associated with reduced cleavage of BID into its active form, hence why full-length BID remains. Studies have shown that BID may also be involved in cell proliferation, and possess a pro-proliferative function, which have been well reviewed (Yin 2006). This may also support the increased survival of cells, and resistance to chemotherapy, where BID expression is shown to be increased. Bcl-xL was shown to be over-expressed in chemotherapy-resistant tumours. Here, this increased Bcl-xL is able to bind to and sequester tBID (Lovell, Billen et al. 2008) and prevent apoptosis of tumour cells, hence they are surviving chemotherapy and displaying resistance.

10.4.2 14-3-3 proteins

One of the proteins which was identified by both of the discovery-phase platforms; antibody microarray and 2D-PAGE MALDI-TOF/TOF MS analysis, was 14-3-3 theta/tau (*YWHAQ*). This was the only member of the 14-3-3 family of proteins present on the antibody microarray, yet 2D-PAGE/MS analysis identified 4 other isoforms in at least 2/3 experiments; 14-3-3 epsilon (*YWHAE*), 14-3-3 zeta/delta (*YWHAZ*), 14-3-3 gamma (*YWHAG*) and 14-3-3 beta/alpha (*YWHAB*). Of the 55 DEPs in the 2D-PAGE/MS data set, the five 14-3-3 isoforms are present in 8 of the 10 canonical pathways listed in Table 36. The potential role of 14-3-3 theta/tau in chemotherapy-resistance has previously been reported (section 9.4). This protein was therefore also included in those initially selected for the confirmation phase of the biomarker discovery pipeline, as well as other 14-3-3 isoforms.

The expression of 14-3-3 theta/tau could not be assessed using western blotting, however an antibody against four isoforms of 14-3-3 (beta, eta, tau and sigma) demonstrated the significant up-regulation of these proteins in chemotherapy-resistant tumour samples. The up-regulation of 14-3-3 epsilon was also shown to be significantly associated with chemotherapy-resistance. To assess the clinical relevance of expression of 14-3-3 theta/tau, a pilot immunohistochemical study was performed using archival samples from a well-characterised patient cohort (Garimella 2007). Immunohistochemical analysis revealed a significant association between high expression of 14-3-3 theta/tau within the nuclear membrane and chemotherapy-resistant tumours ($p=0.02$). In order to fully assess the strength of this putative biomarker, further screening should be performed in a larger patient cohort. The other isoforms of 14-3-3 should also be carried forward to confirmation and clinical validation phases.

The association between doxorubicin and paclitaxel chemotherapy resistance in breast cancer and expression of 14-3-3 proteins has also previously been made (Liu, Liu et al. 2006; Chuthapisith 2007). They have also been found to be associated with cisplatin resistance in cervix squamous cell carcinoma cell lines (Castagna, Antonioli et al. 2004), *vinca* alkaloid resistance in acute lymphoblastic leukaemia (Verrills, Walsh et al. 2003), as well as mitoxantrone chemotherapy-resistance in pancreatic cancer (Sinha, Hütter et al. 1999). There are seven mammalian isoforms of 14-3-3; beta/alpha, gamma, epsilon, eta, sigma, theta/tau and zeta/delta, which are reported to associate with proteins involved in critical processes including cell cycle regulation, intracellular signalling and apoptosis (Tzivion, Gupta et al. 2006). Due to the nature of their protein targets, 14-3-3 proteins have been widely associated with cancer, including response to therapeutic agents. Overall, it is thought that 14-3-3 proteins promote cell survival by inhibition of apoptosis (Masters, Subramanian et al. 2002). It has been reported that 14-3-3 theta/tau inhibits tamoxifen-induced apoptosis in MCF7 breast cancer cells via interaction with p21, which is required for tamoxifen to generate a response (Wang, Liu et al. 2010). The overexpression of 14-3-3 tau has therefore been associated with tamoxifen resistance in MCF7 breast cancer cell lines. Another study also found 14-3-3 theta/tau to be associated with response to chemotherapeutic agents, where single nucleotide polymorphisms in the gene encoding 14-3-3 theta/tau (*YWHAQ*) were shown to have a significant effect upon cellular response to chemotherapeutic agents (Vazquez 2010). As discussed in chapter 9 (section 9.4), 14-3-3 proteins have been shown to have over 300 protein targets (Sluchanko and Gusev 2010), which are associated with apoptosis, cell cycle control, proliferation, cytoskeleton regulation and transcription. These are all critical pathways, which are involved either directly or indirectly with the deliverance of the desired effect of chemotherapeutic agents,

or mechanism of action, ultimately leading to death of the cancerous cell and its removal from the system. Aberrations within the 14-3-3 proteins may therefore present an array of mechanisms by which resistance to chemotherapeutic agents may arise, which warrants further research into their role as putative biomarkers of chemotherapy resistance in breast cancer.

So far, two proteins have passed through all the phases of the discovery pipeline, where their differential expression was first recognised during discovery-phase experiments, was confirmed by western blotting and validated in clinical samples in small pilot series. They have also shown to be involved in important cellular processes which have established putative roles in response to chemotherapy. These proteins include 14-3-3 theta/tau and tBID. Further clinical validation, in a larger sample cohort, would be required to fully assess the role of these proteins as putative biomarkers of chemotherapy resistance in breast cancer. There are also many other proteins within each of the discovery-phase data sets which have been highlighted during the data mining phase, using IPA, and warrant further research. DEPs identified by antibody microarray analysis which have not yet been carried forward, but were associated with several of the canonical pathways identified by IPA, including PI3K/AKT, mTOR, and P70S6K signalling, which have been widely studied for their association with breast cancer tumourigenesis, prognosis and therapy resistance, as well as potential therapeutic targets (LoPiccolo, Blumenthal et al. 2008; Bartlett 2010; Ghayad and Cohen 2010; Wallin, Guan et al. 2010; Kim, Kim et al. 2011; Wang, Yi et al. 2011; Xiang, Jia et al. 2011). These molecules include Protein Kinase Ba (*AKT1*), IKKa (*CHUK*), Protein kinase Cb1/2 (*PRKCB*) and Rsk1 (*RPS6KB1*), which are members of the serine/threonine kinase family, and their involvement in the above-mentioned canonical

pathways warrants further research into their role as putative biomarkers of chemotherapy resistance in breast cancer.

CHAPTER 11:
CONCLUSIONS

Chapter 11. Conclusions

One of the main aims of this project was to identify biomarkers associated with resistance to neoadjuvant chemotherapy in locally advanced breast cancer, using comparative proteomic techniques. The other main aim of this project was to perform the transition from proteomics research based upon cell line samples to clinical tissue samples.

Currently, response to neoadjuvant chemotherapy cannot be predicted; some tumours respond well, and show a decrease in size, whilst others show no response or more detrimentally, progressive disease during the treatment course. In this instance, patients receive cytotoxic drugs, with unpleasant side-effects, for no therapeutic gain. Resistance to neoadjuvant chemotherapy is therefore a major obstacle in achieving effective tumour treatment. The ability to predict tumour response at the time of diagnosis would benefit both the patient and clinician, allowing the individualisation of treatment and the administration of chemotherapy to only those who are mostly likely to benefit, thus maximising treatment efficacy.

In order to achieve this, clinical samples had to be collected alongside corresponding relevant clinical information. These clinical samples were from tumour resections, which had been treated with standard neoadjuvant chemotherapy. Comparative proteomic analysis performed on these clinical samples would allow the identification of putative biomarkers associated with both the 'intrinsic', (where cancer cells are innately resistant to chemotherapy), and 'acquired' (where cancer cells develop resistance during treatment) mechanisms of chemotherapy resistance.

Proteomic techniques, which were established for cell line samples, then had to be optimised for use with clinical tissue samples. These included antibody microarrays, 2D-PAGE MALDI-TOF/TOF MS and western blotting. Once optimised, these

techniques could then be applied as comparative proteomics methods, using the biomarker discovery pipeline, for the identification of putative biomarkers of chemotherapy resistance.

11.1 Antibody microarray analysis

Antibody microarray analysis encompassing five experiments, comparing the expression of a chemotherapy-sensitive tumour with a chemotherapy-resistant tumour, generated a list of 38 differentially expressed proteins (DEPs). Data mining was subsequently performed, using IPA, which highlighted important canonical pathways within the data set. The top, most significant, canonical pathway identified was the ERK5 signalling pathway. The ERK5 protein has been described as a pro-survival factor which has a role in the regulation of cell proliferation, mainly during G₁/S transition of the cell cycle (Girio, Montero et al. 2007). It has also been shown to be activated during mitosis, where it contributes to cell survival by sequestering the pro-apoptotic protein Bim and preventing induction of caspase activation and cell death by apoptosis (Girio, Montero et al. 2007). The ERK5 signalling pathway has also been associated with the invasive phenotype of prostate cancer, where the expression of ERK5 was shown to be up-regulated in metastatic prostate cancer (Ramsay, McCracken et al. 2011), which also highlights its role as pro-survival factor.

Protein candidates selected for the confirmation phase included the apoptosis-related proteins tBID and Bcl-xL, and 14-3-3 theta/tau which was identified in 5/5 antibody microarray experiments and was also present in the ERK5 signalling pathway. The differential expression of 14-3-3 proteins, full length BID and Bcl-xL was confirmed by western blotting. The differential expression of 14-3-3 theta/tau and tBID was also clinically validated using immunohistochemistry in a small pilot study. The increased expression of 14-3-3 theta/tau, which is an anti-apoptotic protein, was found to be associated with chemotherapy-resistance. Decreased expression of tBID was observed,

showing reduced cleavage of BID into its active form (tBID), thus reduced levels of apoptosis. Both of these factors support increased survival of the cancer cell, by evasion of apoptosis, which is essential for the chemotherapy-resistant phenotype.

Antibody microarray analysis was also performed on 13 data sets from 5 different sample groups, which produced a list of repeatedly-identified DEPs (RIDEPs) from the antibody microarray platform. The phenomenon of RIDEPs had been reported for 2D-PAGE/MS experiments (Petrak, Ivanek et al. 2008; Mariman 2009; Wang, Bouwman et al. 2009), but not for data generated by antibody microarray analysis. The list of RIDEPs associated with the antibody microarray platform has therefore been produced and published for the first time, as well as recommendations for quality-control thresholds. If proteins present on the RIDEP list are identified as DEPs, confirmation of their differential expression and clinical validation should be performed rigorously.

11.2 2D-PAGE MALDI-TOF/TOF MS analysis

2D-PAGE MALDI-TOF/TOF analysis for the identification of biomarkers of chemotherapy resistance was performed, and included 3 experiments comparing protein expression in chemotherapy-sensitive tumours versus chemotherapy-resistant tumours. In total, 132 DEPs were identified across the three experiments, and 57 of these were identified in at least two experiments. Data mining, using IPA software, revealed the top canonical pathway associated with the data set to be 'cell cycle: G2/M DNA damage checkpoint control', which involved 5 isoforms of the 14-3-3 family of proteins (14-3-3 theta/tau, 14-3-3 epsilon, 14-3-3 gamma, 14-3-3 zeta/delta and 14-3-3 beta/alpha). These proteins were also present in several other canonical pathways identified, which indicates the extent and range of their involvement in critical biological processes. Their role in chemotherapy resistance to anthracyclines and taxanes in breast cancer has also previously been suggested (Liu, Liu et al. 2006; Chuthapisith 2007), and overall pro-survival function via inhibition of apoptosis. The observed overexpression of 14-3-3

proteins therefore supports increased the cell survival and evasion of apoptosis associated with the chemotherapy-resistant phenotype.

The differential expression of 14-3-3 theta/tau and 14-3-3 epsilon isoforms was confirmed using western blotting, and 14-3-3 theta/tau was carried forward for clinical validation, where its differential expression was successfully confirmed in a clinical context using immunohistochemistry, in a small pilot study. Further confirmatory testing and clinical validation is required for the other 14-3-3 isoforms, and this should be considered for future work.

Overall there has been little overlap between the DEPs identified by antibody microarray analysis and 2D-PAGE MALDI-TOF/TOF MS. This was also noticed by Smith *et al*, where different lists of DEPs were produced from antibody microarray and MS-based methods (Smith, Qutob *et al.* 2009). The two proteomic platforms therefore provide a complementary approach to the discovery of DEPs, and when used in combination, increase the range of DEPs which can be identified.

11.3 Future work

11.3.1 Discovery phase

There are many different options for future work. One of these may include increasing the lists of DEPs, by performing additional discovery-phase experiments, using antibody microarray analysis and 2D-PAGE MALDI-TOF/TOF analysis on additional samples. This may also include performing 2D-PAGE MALDI-TOF/TOF in different pH ranges or on larger gels, to allow additional DEPs to be identified. Depletion strategies may also be employed, to remove the highly-abundant proteins present in the samples, such as albumin, which may improve access to low abundant proteins and increase the chance of biomarker identification. So far, discovery-phase experiments

have been performed using luminal (ER+) invasive ductal carcinoma samples, as this is the most common subtype of breast tumours and the type which displays poorest rates of response to neoadjuvant chemotherapy. It may also be interesting to perform the search for biomarkers using tumour samples of different molecular subtype, such as HER2, however these are less common and have higher rates of pathological complete response to neoadjuvant chemotherapy regardless.

11.3.2 Confirmation and clinical validation

There are many other DEPs within the discovery-phase data, which were highlighted by IPA, which warrant further research, as so far only a few have been carried forward to the confirmation phase. These include the 14-3-3 family of proteins, which are involved in several critical biological processes and overall promote cell survival by inhibition of apoptosis. The relationship between aberrant apoptosis and chemotherapy response is already well-established (Pommier, Sordet et al. 2004), and the expression of these proteins has already been associated with chemotherapy resistance in breast cancer. Therefore these proteins may have roles as potential predictive biomarkers of chemotherapy resistance, however extensive clinical validation, in a large cohort of pre-treatment samples, possibly within a prospective study, is required.

11.3.3 Core biopsy samples

If predictive biomarkers of chemotherapy resistance were transferred to the clinic, screening would be performed at the time of diagnosis, to allow subsequent treatment to be tailored accordingly. This may involve the routine screening of formalin-fixed paraffin-embedded (FFPE) core biopsy samples, with an established panel of predictive biomarkers. This is a clinically accepted approach, currently used for routine ER PR and HER2 screening, therefore if a panel of predictive biomarkers was identified, it would

be relatively simple to incorporate this into an existing routine protocol, without requiring extra patient samples.

The most clinically relevant sample type to analyse using proteomics techniques for the identification of predictive biomarkers of chemotherapy-resistance is the pre-treatment core biopsy sample, which is taken at the diagnostic stage. This type of sample provides cells from the core of the tumour, and therefore allows direct analysis and characterisation. Biomarkers accessible within alternative samples, such as serum, allow less invasive diagnostic testing, however proteomic changes responsible for the potential innate resistance to chemotherapeutic agents may not be reflected in the serum. Identification of predictive biomarkers of innate resistance to neoadjuvant chemotherapy would require proteomic analysis to be performed, in the discovery phase of the biomarker discovery pipeline, using fresh core biopsy tissue samples. The quality of the sample may also be higher at this stage, with a larger proportion of tumour cells, in comparison to the resection samples where tumour cells may be more diffuse as a result of chemotherapy treatment. However, the procedure of obtaining core biopsy samples is invasive and unpleasant for the patient. Therefore, ethical approval would only allow, if at all, a minimum number of extra fresh core biopsy samples to be taken for research purposes. Currently, proteomic methods are not established for such small samples. A single preliminary test was performed, which showed that 6 core biopsy samples would be required for proteomic analysis using current methods. Further optimisation of methods would therefore be required before proteomic analysis could be performed on an ethically-suitable number of cores biopsy samples, and before an ethics application could be considered.

11.3.4 Establishment of chemotherapy-resistant cell lines

Another consideration for future work, may involve the development of chemotherapy resistant cell line models (Watson, Lind et al. 2007). This would present a

complementary approach to the discovery of biomarkers already performed using clinical samples. DEPs generated using this method would also require confirmation and validation, so despite controversies regarding their clinical relevance, data generated from cell-line samples would ultimately be tested within the clinical context. In order to represent the standard anthracycline plus taxane neoadjuvant chemotherapy treatment regimen, cell lines could be generated displaying resistance to anthracycline and taxane chemotherapeutic drugs, such as epirubicin and docetaxel. These are currently available to purchase from Sigma Aldrich (Epirubicin (#E9406, Sigma Aldrich) and Docetaxel (#01885, Sigma Aldrich).

Another consideration for future work could be the use of drug inhibitors or small molecule inhibitors, which target the putative biomarkers which have already been identified, whose expression has shown to be up-regulated in the chemotherapy-resistant phenotype (14-3-3 proteins and Bcl-xL). If chemotherapy-resistant cell line models were established, the inhibition of these proteins could be tested in combination with neoadjuvant chemotherapy, for drug sensitisation. One such example may involve the inhibition of Bcl-xL with Navitoclax, which is a Bcl-2 family inhibitor. This inhibitor has been shown to accelerate apoptosis during drug-induced mitotic arrest, by taxanes, thus increasing the therapeutic efficacy of taxanes in an epithelial cell line model (Shi, Zhou et al. 2011) and a non-small cell lung cancer cell line model (Tan, Malek et al. 2011). The inhibition of 14-3-3 proteins, as a potential therapeutic target has also been discussed. An example of a 14-3-3 inhibitor is difopein; a 14-3-3-binding dipeptide which has been shown to increase rates of apoptosis in tumour cells being treated with anti-neoplastic agents, thus acting as a sensitising agent (Hermeking 2003). The use of such agents, in established chemotherapy resistant models may therefore provide further information regarding the role of these putative biomarkers of chemotherapy resistance and also as potential therapeutic targets.

11.4 Concluding Remarks

The use of proteomic methods for the identification of predictive biomarkers of chemotherapy resistance in breast cancer, following the stages outlined in the biomarker discovery pipeline, and most importantly using fresh tumour tissue, has shown great potential.

Within this study, methods have been established for the use of clinical tissue with two proteomic platforms, including antibody microarray analysis and 2D-PAGE MALDI-TOF/TOF MS, for the discovery of putative biomarkers of chemotherapy resistance. Guidelines associated with the use of the antibody microarray platform, including quality-control thresholds, fold-change cut-off values and RIDEPs, have also been reported and published for the first time. These guidelines will be useful for any researcher using this proteomic platform. Preliminary work for the use of fresh pre-treatment core biopsy samples has also been performed, providing information regarding the feasibility of the use of these samples for discovery-phase proteomic analysis.

Proteomic analysis within the discovery phase of the biomarker discovery pipeline produced large lists of DEPs which present themselves as putative biomarkers of neoadjuvant chemotherapy resistance in breast cancer. To the best of my knowledge, these are the first reported putative biomarkers of neoadjuvant chemotherapy resistance in breast cancer which have been discovered using proteomic analysis of fresh clinical tumour tissue samples. Two of the DEPs identified have been followed through all the stages of the biomarker discovery pipeline, and show clinical potential as putative predictive biomarkers of neoadjuvant chemotherapy resistance in breast cancer.

References

- Aas, T., S. Geisler, et al. (2003). "Predictive value of tumour cell proliferation in locally advanced breast cancer treated with neoadjuvant chemotherapy." Eur J Cancer **39**(4): 438-446.
- Aebersold, R. and D. R. Goodlett (2001). "Mass spectrometry in proteomics." Chem Rev **101**(2): 269-295.
- Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.
- Aggarwal, K., L. H. Choe, et al. (2006). "Shotgun proteomics using the iTRAQ isobaric tags." Brief Funct Genomic Proteomic **5**(2): 112-120.
- Ahmed, F. E. (2008). "Utility of mass spectrometry for proteome analysis: part I. Conceptual and experimental approaches." Expert Rev Proteomics **5**(6): 841-864.
- Ahmed, N., G. Barker, et al. (2003). "An approach to remove albumin for the proteomic analysis of low abundance biomarkers in human serum." Proteomics **3**(10): 1980-1987.
- Albain, K. S., W. E. Barlow, et al. (2010). "Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial." Lancet Oncol **11**(1): 55-65.
- Aldred, S., M. M. Grant, et al. (2004). "The use of proteomics for the assessment of clinical samples in research." Clin Biochem **37**(11): 943-952.
- Aldridge, G. M., D. M. Podrebarac, et al. (2008). "The use of total protein stains as loading controls: an alternative to high-abundance single-protein controls in semi-quantitative immunoblotting." J Neurosci Methods **172**(2): 250-254.
- Alhamdani, M., C. Schroder, et al. (2009). "Oncoproteomic profiling with antibody microarrays." Genome Med **1**(7): 68.
- AmericanCancerSociety. (2008, 13.05.2009). "What is breast cancer?" Retrieved 13.08.09, 2009, from http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_is_breast_cancer_5.asp.
- Anderson, N. L. and N. G. Anderson (2002). "The human plasma proteome: history, character, and diagnostic prospects." Mol Cell Proteomics **1**(11): 845-867.
- Antonov, A. V., S. Dietmann, et al. (2009). "PPI spider: a tool for the interpretation of proteomics data in the context of protein-protein interaction networks." Proteomics **9**(10): 2740-2749.
- Apweiler, R., C. Aslanidis, et al. (2009). "Approaching clinical proteomics: current state and future fields of application in cellular proteomics." Cytometry A **75**(10): 816-832.
- Asakawa, H., H. Koizumi, et al. (2010). "Prediction of breast cancer sensitivity to neoadjuvant chemotherapy based on status of DNA damage repair proteins." Breast Cancer Res **12**(2): R17.
- Baak, J. P., E. A. Janssen, et al. (2005). "Genomics and proteomics--the way forward." Ann Oncol **16 Suppl 2**: ii30-44.
- Bahia, H., J. N. Ashman, et al. (2002). "Karyotypic variation between independently cultured strains of the cell line MCF-7 identified by multicolour fluorescence in situ hybridization." Int J Oncol **20**(3): 489-494.
- Ball, H. J. and N. H. Hunt (2004). "Needle in a haystack: microdissecting the proteome of a tissue." Amino Acids **27**(1): 1-7.

- Bartlett, J. M. (2010). "Biomarkers and patient selection for PI3K/Akt/mTOR targeted therapies: current status and future directions." Clin Breast Cancer **10 Suppl 3**: S86-95.
- Batist, G., A. Tulpule, et al. (1986). "Overexpression of a novel anionic glutathione transferase in multidrug-resistant human breast cancer cells." J Biol Chem **261**(33): 15544-15549.
- Baumann, F. and R. Preiss (2001). "Cyclophosphamide and related anticancer drugs." J Chromatogr B Biomed Sci Appl **764**(1-2): 173-192.
- Bisca, A., C. D'Ambrosio, et al. (2004). "Proteomic evaluation of core biopsy specimens from breast lesions." Cancer Lett **204**(1): 79-86.
- Bjorhall, K., T. Miliotis, et al. (2005). "Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples." Proteomics **5**(1): 307-317.
- Blamey, R. W., I. O. Ellis, et al. (2007). "Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999." Eur J Cancer **43**(10): 1548-1555.
- Bonnefoi, H., C. Underhill, et al. (2009). "Predictive signatures for chemotherapy sensitivity in breast cancer: are they ready for use in the clinic?" Eur J Cancer **45**(10): 1733-1743.
- Borrebaeck, C. A. and C. Wingren (2009). "Design of high-density antibody microarrays for disease proteomics: key technological issues." J Proteomics **72**(6): 928-935.
- Branzei, D. and M. Foiani (2008). "Regulation of DNA repair throughout the cell cycle." Nat Rev Mol Cell Biol **9**(4): 297-308.
- Brase, J. C., M. Schmidt, et al. (2010). "ERBB2 and TOP2A in breast cancer: a comprehensive analysis of gene amplification, RNA levels, and protein expression and their influence on prognosis and prediction." Clin Cancer Res **16**(8): 2391-2401.
- Brennan, D. J., D. P. O'Connor, et al. (2010). "Antibody-based proteomics: fast-tracking molecular diagnostics in oncology." Nat Rev Cancer **10**(9): 605-617.
- Buehring, G. C., E. A. Eby, et al. (2004). "Cell line cross-contamination: how aware are Mammalian cell culturists of the problem and how to monitor it?" In Vitro Cell Dev Biol Anim **40**(7): 211-215.
- Burdall, S. E., A. M. Hanby, et al. (2003). "Breast cancer cell lines: friend or foe?" Breast Cancer Res **5**(2): 89-95.
- Burnette, W. N. (1981). "'Western Blotting': Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A." Analytical Biochemistry **112**(2): 195-203.
- Buzdar, A. U. (2007). "Preoperative chemotherapy treatment of breast cancer--a review." Cancer **110**(11): 2394-2407.
- Caiazzo, R. J., A. J. Maher, et al. (2009). "Protein microarrays as an application for disease biomarkers." Proteomics Clin Appl **3**: 138-147.
- Caldarella, A., E. Crocetti, et al. (2011). "Female Breast Cancer Status According to ER, PR and HER2 Expression: A Population Based Analysis." Pathol Oncol Res **17**(3): 753-758.
- Camp, R. L., V. Neumeister, et al. (2008). "A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers." J Clin Oncol **26**(34): 5630-5637.
- CancerResearchUK. (2002, May 2009). "Breast cancer: types of breast cancer." Retrieved 13.08.2009, 2009, from <http://www.cancerhelp.org.uk/help/default.asp?page=3287>.

- Capes-Davis, A., G. Theodosopoulos, et al. (2010). "Check your cultures! A list of cross-contaminated or misidentified cell lines." *Int J Cancer* **127**(1): 1-8.
- Cardoso, F., M. Piccart-Gebhart, et al. (2007). "The MINDACT trial: The first prospective clinical validation of a genomic tool." *Mol Oncol* **1**(3): 246-251.
- Carey, L. A., E. C. Dees, et al. (2007). "The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes." *Clin Cancer Res* **13**(8): 2329-2334.
- Castagna, A., P. Antonioli, et al. (2004). "A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431." *Proteomics* **4**(10): 3246-3267.
- Caudle, A. S., A. M. Gonzalez-Angulo, et al. (2010). "Predictors of tumor progression during neoadjuvant chemotherapy in breast cancer." *J Clin Oncol* **28**(11): 1821-1828.
- Celis, J. E., P. Gromov, et al. (2004). "Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery." *Mol Cell Proteomics* **3**(4): 327-344.
- Celis, J. E., J. M. A. Moreira, et al. (2005). "Identification of Extracellular and Intracellular Signaling Components of the Mammary Adipose Tissue and Its Interstitial Fluid in High Risk Breast Cancer Patients." *Mol Cell Proteomics* **4**(4): 492-522.
- Celis, J. E., J. M. A. Moreira, et al. (2005). "Towards discovery-driven translational research in breast cancer." *FEBS Journal* **272**(1): 2-15.
- Charak, S., D. Jangir, et al. (2011). "Interaction studies of Epirubicin with DNA using spectroscopic techniques." *Journal of Molecular Structure* **1000**: 150-154.
- Cheang, M., D. Voduc, et al. (2008). "Basal-like Breast Cancer Defined by Five Biomarkers Has Superior Prognostic Value than Triple-Negative Phenotype." *Clin Cancer Res* **14**(5): 1368-1376.
- Cheang, M. C. U., S. K. Chia, et al. (2009). "Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer." *J. Natl. Cancer Inst.* **101**(10): 736-750.
- Chen, E. I. and J. R. Yates, 3rd (2007). "Cancer proteomics by quantitative shotgun proteomics." *Mol Oncol* **1**(2): 144-159.
- Chen, J. H., B. Feig, et al. (2008). "MRI evaluation of pathologically complete response and residual tumors in breast cancer after neoadjuvant chemotherapy." *Cancer* **112**(1): 17-26.
- Chen, S.-C., H.-K. Chang, et al. (2008). "High Pathologic Complete Response in HER2-positive Locally Advanced Breast Cancer after Primary Systemic Chemotherapy with Weekly Docetaxel and Epirubicin." *Jpn. J. Clin. Oncol.* **38**(2): 99-105.
- Choe, L., M. D'Ascenzo, et al. (2007). "8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease." *Proteomics* **7**(20): 3651-3660.
- Choi, J. E., W. Hur, et al. (2011). "Silencing of 14-3-3zeta over-expression in hepatocellular carcinoma inhibits tumor growth and enhances chemosensitivity to cis-diammined dichloridoplatinum." *Cancer Lett* **303**(2): 99-107.
- Chowdhury, I., B. Tharakan, et al. (2006). "Current concepts in apoptosis: the physiological suicide program revisited." *Cell Mol Biol Lett* **11**(4): 506-525.
- Chrisanthar, R., S. Knappskog, et al. (2011). "Predictive and prognostic impact of TP53 mutations and MDM2 promoter genotype in primary breast cancer patients treated with epirubicin or paclitaxel." *PLoS One* **6**(4): e19249.
- Christmann, M., M. T. Tomicic, et al. (2003). "Mechanisms of human DNA repair: an update." *Toxicology* **193**(1-2): 3-34.

- Chuthapisith, S., B. E. Bean, et al. (2009). "Annexins in human breast cancer: Possible predictors of pathological response to neoadjuvant chemotherapy." Eur J Cancer **45**(7): 1274-1281.
- Chuthapisith, S., J. Eremin, et al. (2010). "Breast cancer chemoresistance: emerging importance of cancer stem cells." Surg Oncol **19**(1): 27-32.
- Chuthapisith, S., J. M. Eremin, et al. (2006). "Neoadjuvant chemotherapy in women with large and locally advanced breast cancer: chemoresistance and prediction of response to drug therapy." Surgeon **4**(4): 211-219.
- Chuthapisith, S., Layfield, R., Kerr, I., Hughes, C., Eremin, O. (2007). "Proteomic profiling of MCF-7 breast cancer cells with chemoresistance to different types of anti-cancer drugs." Int J Oncol **30**(6): 1545-1551.
- Cleator, S., M. Parton, et al. (2002). "The biology of neoadjuvant chemotherapy for breast cancer." Endocr Relat Cancer **9**(3): 183-195.
- Coley, H. M. (2008). "Mechanisms and strategies to overcome chemotherapy resistance in metastatic breast cancer." Cancer Treat Rev **34**(4): 378-390.
- Coley, H. M. (2009). "Mechanisms and consequences of chemotherapy resistance in breast cancer." European Journal of Cancer Supplements **7**(1): 3-7.
- Colombo, P. E., F. Milanezi, et al. (2011). "Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction." Breast Cancer Res **13**(3): 212.
- Cortesi, L., A. Barchetti, et al. (2009). "Identification of protein clusters predictive of response to chemotherapy in breast cancer patients." J Proteome Res **8**(11): 4916-4933.
- Craven, R. A., N. Totty, et al. (2002). "Laser capture microdissection and two-dimensional polyacrylamide gel electrophoresis: evaluation of tissue preparation and sample limitations." Am J Pathol **160**(3): 815-822.
- Cristofanilli, M., A. Gonzalez-Angulo, et al. (2005). "Invasive Lobular Carcinoma Classic Type: Response to Primary Chemotherapy and Survival Outcomes." J Clin Oncol **23**(1): 41-48.
- Croft, D., G. O'Kelly, et al. (2011). "Reactome: a database of reactions, pathways and biological processes." Nucleic Acids Res **39**(Database issue): D691-697.
- Curran, S., J. A. McKay, et al. (2000). "Laser capture microscopy." Mol Pathol **53**(2): 64-68.
- Dawood, S., R. Hu, et al. (2011). "Defining breast cancer prognosis based on molecular phenotypes: results from a large cohort study." Breast Cancer Res Treat **126**(1): 185-192.
- de Campos-Nebel, M., I. Larripa, et al. (2010). "Topoisomerase II-mediated DNA damage is differently repaired during the cell cycle by non-homologous end joining and homologous recombination." PLoS One **5**(9).
- de Snoo, F., R. Bender, et al. (2009). "Gene expression profiling: Decoding breast cancer." Surg Oncol **18**(4): 366-378.
- Deighton, R. F., L. E. Kerr, et al. (2010). "Network generation enhances interpretation of proteomic data from induced apoptosis." Proteomics **10**(6): 1307-1315.
- Deng, S. S., T. Y. Xing, et al. (2006). "Comparative proteome analysis of breast cancer and adjacent normal breast tissues in human." Genomics Proteomics Bioinformatics **4**(3): 165-172.
- Deutsch, E. W., H. Lam, et al. (2008). "Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics." Physiol Genomics **33**(1): 18-25.
- Di Leo, A., L. Biganzoli, et al. (2008). "Topoisomerase II alpha as a marker predicting anthracyclines' activity in early breast cancer patients: ready for the primetime?" Eur J Cancer **44**(18): 2791-2798.

- Di Leo, A. and J. Isola (2003). "Topoisomerase II alpha as a marker predicting the efficacy of anthracyclines in breast cancer: are we at the end of the beginning?" Clin Breast Cancer **4**(3): 179-186.
- Dianov, G. L. and J. L. Parsons (2007). "Co-ordination of DNA single strand break repair." DNA Repair (Amst) **6**(4): 454-460.
- Diehn, M., G. Sherlock, et al. (2003). "SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data." Nucleic Acids Res **31**(1): 219-223.
- Drexler, H. G. and C. C. Uphoff (2002). "Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention." Cytotechnology **39**(2): 75-90.
- Earl, H., A. Vallier, et al. (2009). "Neo-tAnGo: a neoadjuvant randomised phase III trial of epirubicin/cyclophosphamide and paclitaxel +/- gemcitabine in the treatment of women with high-risk early breast cancer (EBC): First report of the primary endpoint, pathological complete response (pCR)." J Clin Oncol (Meeting Abstracts) **27**(15S): 522.
- Echan, L. A., H. Y. Tang, et al. (2005). "Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma." Proteomics **5**(13): 3292-3303.
- Eisenhauer, E. A., P. Therasse, et al. (2009). "New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)." Eur J Cancer **45**(2): 228-247.
- Espina, V., C. Mueller, et al. (2009). "Tissue is alive: New technologies are needed to address the problems of protein biomarker pre-analytical variability." Proteomics Clin Appl **3**(8): 874-882.
- Espinosa, E., S. Morales, et al. (2004). "Docetaxel and high-dose epirubicin as neoadjuvant chemotherapy in locally advanced breast cancer." Cancer Chemother Pharmacol **54**(6): 546-552.
- Espinosa, E., J. A. Vara, et al. (2011). "Gene profiling in breast cancer: Time to move forward." Cancer Treat Rev.
- Eustermann, S., H. Videler, et al. (2011). "The DNA-binding domain of human PARP-1 interacts with DNA single-strand breaks as a monomer through its second zinc finger." J Mol Biol **407**(1): 149-170.
- Fadeel, B. and S. Orrenius (2005). "Apoptosis: a basic biological phenomenon with wide-ranging implications in human disease." J Intern Med **258**(6): 479-517.
- Falschlehner, C., C. H. Emmerich, et al. (2007). "TRAIL signalling: decisions between life and death." Int J Biochem Cell Biol **39**(7-8): 1462-1475.
- Farmer, P., H. Bonnefoi, et al. (2009). "A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer." Nat Med **15**(1): 68-74.
- Fekete, M. R., W. H. McBride, et al. (2005). "Anthracyclines, proteasome activity and multi-drug-resistance." BMC Cancer **5**: 114.
- Fernandez-Sanchez, M., A. Gamboa-Dominguez, et al. (2006). "Clinical and pathological predictors of the response to neoadjuvant anthracycline chemotherapy in locally advanced breast cancer." Medical Oncology **23**(2).
- Fisher, B., J. Bryant, et al. (1998). "Effect of preoperative chemotherapy on the outcome of women with operable breast cancer." J Clin Oncol **16**(8): 2672-2685.
- Flynn, R. L. and L. Zou (2011). "ATR: a master conductor of cellular responses to DNA replication stress." Trends Biochem Sci **36**(3): 133-140.
- Fu, Z. and C. Fenselau (2005). "Proteomic evidence for roles for nucleolin and poly[ADP-ribosyl] transferase in drug resistance." J Proteome Res **4**(5): 1583-1591.

- Fulda, S. and K. M. Debatin (2006). "Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy." *Oncogene* **25**(34): 4798-4811.
- Garfin, D. E. (2003). "Two-dimensional gel electrophoresis: an overview." *TrAC Trends in Analytical Chemistry* **22**(5): 263-272.
- Garimella, V., Watson, M., Cairns, H., Chatuverdi, A., Drew, P., Hubbard, A., Lind, M., Maraveyas, A., Turnbull, L., Cawkwell, L (2007). "Assessment of apoptotic markers as predictors of response to neoadjuvant FEC chemotherapy in locally advanced breast cancer." *Cancer Therapy* **5**: 239-242.
- Garrett, M. (2001). "Cell cycle control and cancer." *Current Science* **81**(5): 515-522.
- Gartler, S. M. (1967). "Genetic markers as tracers in cell culture." *Natl Cancer Inst Monogr* **26**: 167-195.
- Gasch, A. P., P. T. Spellman, et al. (2000). "Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes." *Mol. Biol. Cell* **11**(12): 4241-4257.
- Geho, D., Petricoin, E., Liotta, L (2004). "Blasting into the microworld of tissue proteomics: a new window on cancer." *Clin Cancer Res* **10**: 825-827.
- Geisler, S., P. E. Lonning, et al. (2001). "Influence of TP53 gene alterations and c-erbB-2 expression on the response to treatment with doxorubicin in locally advanced breast cancer." *Cancer Res* **61**(6): 2505-2512.
- Ghayad, S. E. and P. A. Cohen (2010). "Inhibitors of the PI3K/Akt/mTOR pathway: new hope for breast cancer patients." *Recent Pat Anticancer Drug Discov* **5**(1): 29-57.
- Ghobrial, I. M., D. J. McCormick, et al. (2005). "Proteomic analysis of mantle-cell lymphoma by protein microarray." *Blood* **105**(9): 3722-3730.
- Ghobrial, I. M., T. E. Witzig, et al. (2005). "Targeting apoptosis pathways in cancer therapy." *CA Cancer J Clin* **55**(3): 178-194.
- Giacinti, C. and A. Giordano (2006). "RB and cell cycle progression." *Oncogene* **25**(38): 5220-5227.
- Gianni, L., M. Zambetti, et al. (2005). "Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer." *J Clin Oncol* **23**(29): 7265-7277.
- Girio, A., J. C. Montero, et al. (2007). "Erk5 is activated and acts as a survival factor in mitosis." *Cell Signal* **19**(9): 1964-1972.
- Glas, A., A. Floore, et al. (2006). "Converting a breast cancer microarray signature into a high-throughput diagnostic test." *BMC Genomics* **7**(1): 278.
- Goldstein, M., W. P. Roos, et al. (2008). "Apoptotic death induced by the cyclophosphamide analogue mafosfamide in human lymphoblastoid cells: contribution of DNA replication, transcription inhibition and Chk/p53 signaling." *Toxicol Appl Pharmacol* **229**(1): 20-32.
- Gonzalez-Angulo, A., Morales-Vasquez, F., Hortobagyi, G. (2007). Overview of Resistance to Systemic Therapy in Patients with Breast Cancer. *Breast Cancer Chemosensitivity*. D. Yu, Hung, MC. New York, Landes Bioscience and Springer Science and Business Media. Springer Series: Advances in Experimental Medicine and Biology. **608**.
- Gorg, A., O. Drews, et al. (2009). "2-DE with IPGs." *Electrophoresis* **30 Suppl 1**: S122-132.
- Gromov, P., J. E. Celis, et al. (2008). "A single lysis solution for the analysis of tissue samples by different proteomic technologies." *Mol Oncol* **2**(4): 368-379.
- Gu, Q., T. M. Sivanandam, et al. (2006). "Signal stability of Cy3 and Cy5 on antibody microarrays." *Proteome Sci* **4**: 21.
- Gutstein, H. B. and J. S. Morris (2007). "Laser capture sampling and analytical issues in proteomics." *Expert Rev Proteomics* **4**(5): 627-637.

- Gutstein, H. B., J. S. Morris, et al. (2008). "Microproteomics: analysis of protein diversity in small samples." Mass Spectrom Rev **27**(4): 316-330.
- Gygi, S. P., B. Rist, et al. (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." Nat Biotechnol **17**(10): 994-999.
- Hanahan, D. and R. A. Weinberg (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.
- Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.
- Hande, K. (2008). "Topoisomerase II inhibitors." Update on Cancer Therapeutics **3**: 13-26.
- Hansen, R. J., S. M. Ludeman, et al. (2007). "Role of MGMT in protecting against cyclophosphamide-induced toxicity in cells and animals." DNA Repair (Amst) **6**(8): 1145-1154.
- Harrington, H. A., K. L. Ho, et al. (2008). "Construction and analysis of a modular model of caspase activation in apoptosis." Theor Biol Med Model **5**: 26.
- He, J., Y. Liu, et al. (2007). "Comparison of two-dimensional gel electrophoresis based and shotgun strategies in the study of plasma membrane proteome." Proteomics Clin Appl **1**(2): 231-241.
- Hengartner, M. O. (2000). "The biochemistry of apoptosis." Nature **407**(6805): 770-776.
- Herbst, R. S. and F. R. Khuri (2003). "Mode of action of docetaxel - a basis for combination with novel anticancer agents." Cancer Treat Rev **29**(5): 407-415.
- Hermeking, H. (2003). "The 14-3-3 cancer connection." Nat Rev Cancer **3**(12): 931-943.
- Hervy, M., L. M. Hoffman, et al. (2010). "The LIM Protein Zyxin Binds CARP-1 and Promotes Apoptosis." Genes & Cancer **1**(5): 506-515.
- Hillenkamp, F., M. Karas, et al. (1991). "Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers." Anal Chem **63**(24): 1193A-1203A.
- Hirschmann-Jax, C., A. E. Foster, et al. (2005). "A distinct "side population" of cells in human tumor cells: implications for tumor biology and therapy." Cell Cycle **4**(2): 203-205.
- Ho, C. S., C. W. Lam, et al. (2003). "Electrospray ionisation mass spectrometry: principles and clinical applications." Clin Biochem Rev **24**(1): 3-12.
- Hoffman, M. D., M. J. Sniatynski, et al. (2008). "Current approaches for global post-translational modification discovery and mass spectrometric analysis." Anal Chim Acta **627**(1): 50-61.
- Honeth, G., P. O. Bendahl, et al. (2008). "The CD44+/CD24- phenotype is enriched in basal-like breast tumors." Breast Cancer Res **10**(3): R53.
- Houtgraaf, J. H., J. Versmissen, et al. (2006). "A concise review of DNA damage checkpoints and repair in mammalian cells." Cardiovasc Revasc Med **7**(3): 165-172.
- Hudelist, G., C. F. Singer, et al. (2006). "Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium." Proteomics **6**(6): 1989-2002.
- Huen, M. S. and J. Chen (2008). "The DNA damage response pathways: at the crossroad of protein modifications." Cell Res **18**(1): 8-16.
- Hulka, B. S. and P. G. Moorman (2001). "Breast cancer: hormones and other risk factors." Maturitas **38**(1): 103-113; discussion 113-106.
- Hurley, P. J. and F. Bunz (2007). "ATM and ATR: components of an integrated circuit." Cell Cycle **6**(4): 414-417.
- Indran, I. R., G. Tufo, et al. (2011). "Recent advances in apoptosis, mitochondria and drug resistance in cancer cells." Biochim Biophys Acta **1807**(6): 735-745.

- Jarvinen, T. A., M. Tanner, et al. (2000). "Amplification and deletion of topoisomerase II α associate with ErbB-2 amplification and affect sensitivity to topoisomerase II inhibitor doxorubicin in breast cancer." *Am J Pathol* **156**(3): 839-847.
- Jimenez-Marin, A., M. Collado-Romero, et al. (2009). "Biological pathway analysis by ArrayUnlock and Ingenuity Pathway Analysis." *BMC Proc* **3 Suppl 4**: S6.
- Jimenez, C. R., J. C. Knol, et al. (2010). "Proteomics of colorectal cancer: Overview of discovery studies and identification of commonly identified cancer-associated proteins and candidate CRC serum markers." *J Proteomics* **73**(10): 1873-1895.
- Jin, Z. and W. S. El-Deiry (2005). "Overview of cell death signaling pathways." *Cancer Biol Ther* **4**(2): 139-163.
- Johann, D. J., J. Rodriguez-Canales, et al. (2009). "Approaching solid tumor heterogeneity on a cellular basis by tissue proteomics using laser capture microdissection and biological mass spectrometry." *J Proteome Res* **8**(5): 2310-2318.
- Juliano, R., Ling, V. (1976). "A Surface Glycoprotein Modulating Drug Permeability in Chinese Hamster Ovary Cell Mutants." *Biochem Biophys Acta* **455**: 152-162.
- Julius, T., S. E. Kemp, et al. (2005). "MRI and conservative treatment of locally advanced breast cancer." *Eur J Surg Oncol* **31**(10): 1129-1134.
- Kannan, P., S. Telu, et al. (2011). "The 'specific' P-glycoprotein inhibitor Tariquidar is also a substrate and an inhibitor for Breast Cancer Resistance Protein (BCRP/ABCG2)." *ACS Chemical Neuroscience* **2**: 82-89.
- Keller, P. J., A. F. Lin, et al. (2010). "Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines." *Breast Cancer Res* **12**(5): R87.
- Kellner, U., M. Sehested, et al. (2002). "Culprit and victim -- DNA topoisomerase II." *Lancet Oncol* **3**(4): 235-243.
- Kennedy, R. D., J. E. Quinn, et al. (2004). "The role of BRCA1 in the cellular response to chemotherapy." *J Natl Cancer Inst* **96**(22): 1659-1668.
- Key, T. J., P. K. Verkasalo, et al. (2001). "Epidemiology of breast cancer." *Lancet Oncol* **2**(3): 133-140.
- Kim, D. H., J. Bae, et al. (2009). "Proteomic analysis of breast cancer tissue reveals upregulation of actin-remodeling proteins and its relevance to cancer invasiveness." *Proteomics Clin Appl* **3**(1): 30-40.
- Kim, E. K., H. A. Kim, et al. (2011). "Phosphorylated S6K1 is a possible marker for endocrine therapy resistance in hormone receptor-positive breast cancer." *Breast Cancer Res Treat* **126**(1): 93-99.
- Kim, K. and Y. Kim (2009). "Preparing multiple-reaction monitoring for quantitative clinical proteomics." *Expert Rev Proteomics* **6**(3): 225-229.
- Kim, S. I., J. Sohn, et al. (2010). "Molecular subtypes and tumor response to neoadjuvant chemotherapy in patients with locally advanced breast cancer." *Oncology* **79**(5-6): 324-330.
- Kiyomiya, K., S. Matsuo, et al. (2001). "Mechanism of specific nuclear transport of adriamycin: the mode of nuclear translocation of adriamycin-proteasome complex." *Cancer Res* **61**(6): 2467-2471.
- Kondo, N., A. Takahashi, et al. (2010). "DNA damage induced by alkylating agents and repair pathways." *J Nucleic Acids* **2010**: 543531.
- Konecny, G. E., G. Pauletti, et al. (2010). "Association between HER2, TOP2A, and response to anthracycline-based preoperative chemotherapy in high-risk primary breast cancer." *Breast Cancer Res Treat* **120**(2): 481-489.
- Kopf, E., D. Shnitzer, et al. (2005). "Panorama™ Ab Microarray Cell Signaling kit: A unique tool for protein expression analysis." *Proteomics* **5**(9): 2412-2416.

- Kopf, E. and D. Zharhary (2007). "Antibody arrays--An emerging tool in cancer proteomics." Int J Biochem Cell Biol **39**(7-8): 1305-1317.
- Kratz, K., B. Schopf, et al. (2010). "Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents." Cell **142**(1): 77-88.
- Kraus, L. A., S. K. Samuel, et al. (2003). "The mechanism of action of docetaxel (Taxotere) in xenograft models is not limited to bcl-2 phosphorylation." Invest New Drugs **21**(3): 259-268.
- Kuerer, H. M., L. A. Newman, et al. (1999). "Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy." J Clin Oncol **17**(2): 460-469.
- Kultz, D. (2003). "Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function." J Exp Biol **206**(18): 3119-3124.
- Kültz, D. (2005). "Molecular and evolutionary basis of the cellular stress response." Annu Rev Physiol. **67**: 225-257.
- Lacroix, M. and G. Leclercq (2004). "Relevance of breast cancer cell lines as models for breast tumours: an update." Breast Cancer Res Treat **83**(3): 249-289.
- Ladoire, S., L. Arnould, et al. (2008). "Pathologic complete response to neoadjuvant chemotherapy of breast carcinoma is associated with the disappearance of tumor-infiltrating foxp3+ regulatory T cells." Clin Cancer Res **14**(8): 2413-2420.
- Lage, H. (2003). "ABC-Transporters: implications on drug resistance from microorganisms to human cancers." International Journal of Antimicrobial Agents **22**: 188-199.
- Lawen, A. (2003). "Apoptosis-an introduction." Bioessays **25**(9): 888-896.
- Lee, A. H. and I. O. Ellis (2008). "The Nottingham prognostic index for invasive carcinoma of the breast." Pathol Oncol Res **14**(2): 113-115.
- Lee, E. Y. and W. J. Muller (2010). "Oncogenes and tumor suppressor genes." Cold Spring Harb Perspect Biol **2**(10): a003236.
- Li, X., M. T. Lewis, et al. (2008). "Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy." J Natl Cancer Inst **100**(9): 672-679.
- Li, Y., D. Baer, et al. (2009). "Wine, liquor, beer and risk of breast cancer in a large population." Eur J Cancer **45**(5): 843-850.
- Li, Y., M. Hussain, et al. (2005). "Gene expression profiling revealed novel mechanism of action of Taxotere and Furtulon in prostate cancer cells." BMC Cancer **5**: 7.
- Li, Y., X. Li, et al. (2004). "Regulation of microtubule, apoptosis, and cell cycle-related genes by taxotere in prostate cancer cells analyzed by microarray." Neoplasia **6**(2): 158-167.
- Lin, D., D. L. Tabb, et al. (2003). "Large-scale protein identification using mass spectrometry." Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics **1646**(1-2): 1-10.
- Ling, V. (1997). "Multidrug resistance: molecular mechanisms and clinical relevance." Caner Chemother Pharmacol **40**: Suppl 3-8.
- Liotta, L. A. and E. C. Kohn (2001). "The microenvironment of the tumour-host interface." Nature **411**(6835): 375-379.
- Liscovitch, M. and D. Ravid (2007). "A case study in misidentification of cancer cell lines: MCF-7/AdrR cells (re-designated NCI/ADR-RES) are derived from OVCAR-8 human ovarian carcinoma cells." Cancer Lett **245**(1-2): 350-352.
- Liu, M. T., W. T. Huang, et al. (2010). "Prediction of outcome of patients with metastatic breast cancer: evaluation with prognostic factors and Nottingham prognostic index." Support Care Cancer **18**(12): 1553-1564.

- Liu, Y., H. Liu, et al. (2006). "Identification of 14-3-3 σ as a Contributor to Drug Resistance in Human Breast Cancer Cells Using Functional Proteomic Analysis." *Cancer Res* **66**(6): 3248-3255.
- Liu, Y., H. Liu, et al. (2006). "Identification of 14-3-3 σ as a contributor to drug resistance in human breast cancer cells using functional proteomic analysis." *Cancer Res* **66**(6): 3248-3255.
- Lonning, P. E. (2010). "Molecular basis for therapy resistance." *Mol Oncol* **4**(3): 284-300.
- Loo, C. E., M. E. Straver, et al. (2011). "Magnetic resonance imaging response monitoring of breast cancer during neoadjuvant chemotherapy: relevance of breast cancer subtype." *J Clin Oncol* **29**(6): 660-666.
- LoPiccolo, J., G. M. Blumenthal, et al. (2008). "Targeting the PI3K/Akt/mTOR pathway: effective combinations and clinical considerations." *Drug Resist Updat* **11**(1-2): 32-50.
- Lovell, J. F., L. P. Billen, et al. (2008). "Membrane binding by tBid initiates an ordered series of events culminating in membrane permeabilization by Bax." *Cell* **135**(6): 1074-1084.
- Lyngholm, M., H. Vorum, et al. (2011). "Attempting to distinguish between endogenous and contaminating cyokeratins in a corneal proteomic study." *BMC Ophthalmol* **11**: 3.
- MacKay, C., A. C. Declais, et al. (2010). "Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2." *Cell* **142**(1): 65-76.
- Malik, R., K. Dulla, et al. (2010). "From proteome lists to biological impact--tools and strategies for the analysis of large MS data sets." *Proteomics* **10**(6): 1270-1283.
- Marengo, E., E. Robotti, et al. (2005). "Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems." *Proteomics* **5**(3): 654-666.
- Mariman, E. C. (2009). "2DE-proteomics meta-data indicate the existence of distinct cellular stress-responsive mechanisms." *Expert Rev Proteomics* **6**(4): 337-339.
- Masters, J. (2010). "Cell line misidentification: the beginning of the end." *Nat Rev Cancer* **10**(6): 441-448.
- Masters, J. R. (2000). "Human cancer cell lines: fact and fantasy." *Nat Rev Mol Cell Biol* **1**(3): 233-236.
- Masters, S. C., R. R. Subramanian, et al. (2002). "Survival-promoting functions of 14-3-3 proteins." *Biochem. Soc. Trans.* **30**(4): 360-365.
- Mathew, J., K. S. Asgeirsson, et al. (2009). "Neoadjuvant chemotherapy for locally advanced breast cancer: a review of the literature and future directions." *Eur J Surg Oncol* **35**(2): 113-122.
- McGrogan, B. T., B. Gilmartin, et al. (2008). "Taxanes, microtubules and chemoresistant breast cancer." *Biochim Biophys Acta* **1785**(2): 96-132.
- McLaughlin, R. and N. Hylton (2011). "MRI in breast cancer therapy monitoring." *NMR Biomed* **24**(6): 712-720.
- McShane, L. M., D. G. Altman, et al. (2005). "Reporting recommendations for tumor marker prognostic studies." *J Clin Oncol* **23**(36): 9067-9072.
- McTiernan, A., A. Kuniyuki, et al. (2001). "Comparisons of two breast cancer risk estimates in women with a family history of breast cancer." *Cancer Epidemiol Biomarkers Prev* **10**(4): 333-338.
- Memos, N., A. Kataki, et al. (2011). "Alternations of 14-3-3 theta and beta protein levels in brain during experimental sepsis." *J Neurosci Res* **89**(9): 1409-1418.
- Middleton, M. and G. Margison (2003). "Improvement of chemotherapy efficacy by inactivation of a DNA-repair pathway." *Lancet Oncol* **4**: 37-44.

- Mieog, J. S. D., J. A. v. d. Hage, et al. (2006). "Tumour response to preoperative anthracycline-based chemotherapy in operable breast cancer: the predictive role of p53 expression." European journal of cancer (Oxford, England : 1990) **42**(10): 1369-1379.
- Minotti, G., P. Menna, et al. (2004). "Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity." Pharmacol Rev **56**(2): 185-229.
- Miyoshi, Y., M. Kurosumi, et al. (2010). "Predictive factors for anthracycline-based chemotherapy for human breast cancer." Breast Cancer **17**(2): 103-109.
- Mohri, Y., T. Mohri, et al. (2009). "Identification of macrophage migration inhibitory factor and human neutrophil peptides 1-3 as potential biomarkers for gastric cancer." Br J Cancer **101**(2): 295-302.
- Moll, R., M. Divo, et al. (2008). "The human keratins: biology and pathology." Histochem Cell Biol **129**(6): 705-733.
- Montero, A., F. Fossella, et al. (2005). "Docetaxel for treatment of solid tumours: a systematic review of clinical data." Lancet Oncol **6**(4): 229-239.
- Moore, H. M., A. B. Kelly, et al. (2011). "Biospecimen Reporting for Improved Study Quality (BRISQ)." J Proteome Res **10**(8): 3429-3438.
- Morandell, S., T. Stasyk, et al. (2006). "Phosphoproteomics strategies for the functional analysis of signal transduction." Proteomics **6**(14): 4047-4056.
- Morrogh, M., N. Olvera, et al. (2007). "Tissue preparation for laser capture microdissection and RNA extraction from fresh frozen breast tissue." Biotechniques **43**(1): 41-42, 44, 46 passim.
- Morse, D. L., H. Gray, et al. (2005). "Docetaxel induces cell death through mitotic catastrophe in human breast cancer cells." Mol Cancer Ther **4**(10): 1495-1504.
- Murata, M., T. Suzuki, et al. (2004). "Oxidative DNA damage induced by a hydroperoxide derivative of cyclophosphamide." Free Radic Biol Med **37**(6): 793-802.
- Nakayama, K. I. and K. Nakayama (2005). "Regulation of the cell cycle by SCF-type ubiquitin ligases." Semin Cell Dev Biol **16**(3): 323-333.
- Nguyen, N. P., F. S. Almeida, et al. (2010). "Molecular biology of breast cancer stem cells: potential clinical applications." Cancer Treat Rev **36**(6): 485-491.
- Nimeus, E., J. Malmstrom, et al. (2007). "Proteomic analysis identifies candidate proteins associated with distant recurrences in breast cancer after adjuvant chemotherapy." J Pharm Biomed Anal **43**(3): 1086-1093.
- Nirmalan, N. J., P. Harnden, et al. (2008). "Mining the archival formalin-fixed paraffin-embedded tissue proteome: opportunities and challenges." Mol Biosyst **4**(7): 712-720.
- Nirmalan, N. J., P. Harnden, et al. (2009). "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 **217**(4): 497-506.
- Nirmalan, N. J., C. Hughes, et al. (2011). "Initial development and validation of a novel extraction method for quantitative mining of the formalin-fixed, paraffin-embedded tissue proteome for biomarker investigations." J Proteome Res **10**(2): 896-906.
- Nitiss, J. L. (2009). "Targeting DNA topoisomerase II in cancer chemotherapy." Nat Rev Cancer **9**(5): 338-350.
- O'Farrell, P. H. (1975). "High resolution two-dimensional electrophoresis of proteins." J. Biol. Chem. **250**(10): 4007-4021.

- O'Kane, S. L., G. L. Eagle, et al. (2010). "COX-2 specific inhibitors enhance the cytotoxic effects of pemetrexed in mesothelioma cell lines." Lung Cancer **67**(2): 160-165.
- Oakman, C., E. Moretti, et al. (2009). "The role of topoisomerase IIalpha and HER-2 in predicting sensitivity to anthracyclines in breast cancer patients." Cancer Treat Rev **35**(8): 662-667.
- Oliveras-Ferraras, C., A. Vazquez-Martin, et al. (2008). "Growth and molecular interactions of the anti-EGFR antibody cetuximab and the DNA cross-linking agent cisplatin in gefitinib-resistant MDA-MB-468 cells: new prospects in the treatment of triple-negative/basal-like breast cancer." Int J Oncol **33**(6): 1165-1176.
- Olsen, J. V., S.-E. Ong, et al. (2004). "Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues." Mol Cell Proteomics **3**(6): 608-614.
- Osborne, C., P. Wilson, et al. (2004). "Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications." Oncologist **9**(4): 361-377.
- Othman, M. I., M. I. Majid, et al. (2008). "Isolation, identification and quantification of differentially expressed proteins from cancerous and normal breast tissues." Ann Clin Biochem **45**(Pt 3): 299-306.
- Paik, S., S. Shak, et al. (2004). "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer." N Engl J Med **351**(27): 2817-2826.
- Pan, S., R. Aebersold, et al. (2009). "Mass spectrometry based targeted protein quantification: methods and applications." J Proteome Res **8**(2): 787-797.
- Patel, K. J. and I. F. Tannock (2009). "The influence of P-glycoprotein expression and its inhibitors on the distribution of doxorubicin in breast tumors." BMC Cancer **9**: 356.
- Paulovich, A. G., J. R. Whiteaker, et al. (2008). "The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline." Proteomics Clin Appl **2**(10-11): 1386-1402.
- Pavlickova, P., E. M. Schneider, et al. (2004). "Advances in recombinant antibody microarrays." Clin Chim Acta **343**(1-2): 17-35.
- Penque, D. (2009). "Two-dimensional gel electrophoresis and mass spectrometry for biomarker discovery." Proteomics Clin Appl **3**: 155-172.
- Perou, C. M. (2010). "Molecular stratification of triple-negative breast cancers." Oncologist **15** Suppl 5: 39-48.
- Perou, C. M., T. Sorlie, et al. (2000). "Molecular portraits of human breast tumours." Nature **406**(6797): 747-752.
- Perou, C. M., T. Sorlie, et al. (2000). "Molecular portraits of human breast tumours." **406**(6797): 747-752.
- Petrak, J., R. Ivanek, et al. (2008). "Déjà vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins." Proteomics **8**(9): 1744-1749.
- Pommier, Y., O. Sordet, et al. (2004). "Apoptosis defects and chemotherapy resistance: molecular interaction maps and networks." Oncogene **23**(16): 2934-2949.
- Pommier, Y., O. Sordet, et al. (2004). "Apoptosis defects and chemotherapy resistance: molecular interaction maps and networks." Oncogene **23**(16): 2934-2949.
- Ponomarenko, E., A. Lisitsa, et al. (2009). "Identification of differentially expressed proteins using automatic meta-analysis of proteomics-related articles." Biomed Khim. **55**(1): 5-14.
- Prat, A., J. S. Parker, et al. (2010). "Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer." Breast Cancer Res **12**(5): R68.
- Prat, A. and C. M. Perou (2011). "Deconstructing the molecular portraits of breast cancer." Mol Oncol **5**(1): 5-23.

- Qian, H. R. and S. Huang (2005). "Comparison of false discovery rate methods in identifying genes with differential expression." *Genomics* **86**(4): 495-503.
- Rabilloud, T. (2002). "Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains." *Proteomics* **2**(1): 3-10.
- Rabilloud, T., C. Valette, et al. (1994). "Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension." *Electrophoresis* **15**(12): 1552-1558.
- Rahbar, A. M. and C. Fenselau (2005). "Unbiased examination of changes in plasma membrane proteins in drug resistant cancer cells." *J Proteome Res* **4**(6): 2148-2153.
- Rahimpour, M., M. Soheili, et al. (2007). "Carrier Ampholyte Isoelectric Focusing Based Two-Dimensional Electrophoresis in Preliminary Screening of Differential Proteomics Analysis." *Chromatographia* **66**: 133-136.
- Rajagopalan, S., R. S. Sade, et al. (2010). "Mechanistic differences in the transcriptional activation of p53 by 14-3-3 isoforms." *Nucleic Acids Res* **38**(3): 893-906.
- Ramsay, A. K., S. R. McCracken, et al. (2011). "ERK5 signalling in prostate cancer promotes an invasive phenotype." *Br J Cancer* **104**(4): 664-672.
- Reed, E. (2010). "DNA damage and repair in translational oncology: an overview." *Clin Cancer Res* **16**(18): 4511-4516.
- Rifai, N., M. A. Gillette, et al. (2006). "Protein biomarker discovery and validation: the long and uncertain path to clinical utility." *Nat Biotechnol* **24**(8): 971-983.
- Robinson, D. N. (2010). "14-3-3, an integrator of cell mechanics and cytokinesis." *Small Gtpases* **1**(3): 165-169.
- Roche, S., L. Tiers, et al. (2009). "Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: the more the better?" *J Proteomics* **72**(6): 945-951.
- Rodriguez, A. A., A. Makris, et al. (2010). "DNA repair signature is associated with anthracycline response in triple negative breast cancer patients." *Breast Cancer Res Treat* **123**(1): 189-196.
- Ross, J. S., J. A. Fletcher, et al. (2003). "The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy." *Oncologist* **8**(4): 307-325.
- Ross, P. L., Y. N. Huang, et al. (2004). "Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents." *Mol Cell Proteomics* **3**(12): 1154-1169.
- Rouzier, R., C. M. Perou, et al. (2005). "Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy." *Clin Cancer Res* **11**(16): 5678-5685.
- Rouzier, R., C. M. Perou, et al. (2005). "Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy." *Clinical Cancer Research* **11**(16): 5678-5685.
- Sakorafas, G. H., E. Krespis, et al. (2002). "Risk estimation for breast cancer development; a clinical perspective." *Surg Oncol* **10**(4): 183-192.
- Sanchez-Carbayo, M. (2006). "Antibody arrays: technical considerations and clinical applications in cancer." *Clin Chem* **52**(9): 1651-1659.
- Schrohl, A. S., S. Wurtz, et al. (2008). "Banking of biological fluids for studies of disease-associated protein biomarkers." *Mol Cell Proteomics* **7**(10): 2061-2066.
- Schwab, M., A. Claas, et al. (2002). "BRCA2: a genetic risk factor for breast cancer." *Cancer Lett* **175**(1): 1-8.
- Selvarajan, S., B. Bay, et al. (2004). "The HercepTest and Routine C-erbB2 Immunohistochemistry in Breast Cancer; any difference?" *Ann Acad Med Singapore* **33**(4): 473-476.
- Shah, M. A. and G. K. Schwartz (2001). "Cell cycle-mediated drug resistance: an emerging concept in cancer therapy." *Clin Cancer Res* **7**(8): 2168-2181.

- Shaw, M. M. and B. M. Riederer (2003). "Sample preparation for two-dimensional gel electrophoresis." Proteomics **3**(8): 1408-1417.
- Shenar, N., N. Sommerer, et al. (2009). "Comparison of LID versus CID activation modes in tandem mass spectrometry of peptides." J Mass Spectrom **44**(5): 621-632.
- Shi, J., Y. Zhou, et al. (2011). "Navitoclax (ABT-263) accelerates apoptosis during drug-induced mitotic arrest by antagonizing Bcl-xL." Cancer Res **71**(13): 4518-4526.
- Singletary, S. E., C. Allred, et al. (2002). "Revision of the American Joint Committee on Cancer staging system for breast cancer." J Clin Oncol **20**(17): 3628-3636.
- Singletary, S. E. and J. L. Connolly (2006). "Breast Cancer Staging: Working With the Sixth Edition of the AJCC Cancer Staging Manual." CA Cancer J Clin **56**(1): 37-47.
- Sinha, P., G. Hütter, et al. (1999). "Increased expression of epidermal fatty acid binding protein, cofilin, and 14-3-3- σ (stratifin) detected by two-dimensional gel electrophoresis, mass spectrometry and microsequencing of drug-resistant human adenocarcinoma of the pancreas." Electrophoresis **20**(14): 2952-2960.
- Sleno, L. and D. A. Volmer (2004). "Ion activation methods for tandem mass spectrometry." J Mass Spectrom **39**(10): 1091-1112.
- Slodkowska, E. A. and J. S. Ross (2009). "MammaPrint 70-gene signature: another milestone in personalized medical care for breast cancer patients." Expert Review of Molecular Diagnostics **9**(5): 417-422.
- Sluchanko, N. N. and N. B. Gusev (2010). "14-3-3 proteins and regulation of cytoskeleton." Biochemistry (Mosc) **75**(13): 1528-1546.
- Smejkal, G. B., F. A. Witzmann, et al. (2007). "Sample preparation for two-dimensional gel electrophoresis using pressure cycling technology." Anal Biochem **363**(2): 309-311.
- Smith, I. C., S. D. Heys, et al. (2002). "Neoadjuvant chemotherapy in breast cancer: significantly enhanced response with docetaxel." J Clin Oncol **20**(6): 1456-1466.
- Smith, L., O. Qutob, et al. (2009). "Proteomic identification of putative biomarkers of radiotherapy resistance: a possible role for the 26S proteasome?" Neoplasia **11**(11): 1194-1207.
- Smith, L., M. B. Watson, et al. (2006). "The analysis of doxorubicin resistance in human breast cancer cells using antibody microarrays." Mol Cancer Ther **5**(8): 2115-2120.
- Smith, L., M. B. Watson, et al. (2006). "The analysis of doxorubicin resistance in human breast cancer cells using antibody microarrays." Mol Cancer Ther **5**(8): 2115-2120.
- Smith, L., K. J. Welham, et al. (2007). "The proteomic analysis of cisplatin resistance in breast cancer cells." Oncol Res **16**(11): 497-506.
- Smith, M. A., E. Blankman, et al. (2010). "A Zyxin-Mediated Mechanism for Actin Stress Fiber Maintenance and Repair." Developmental Cell **19**(3): 365-376.
- Smogorzewska, A., R. Desetty, et al. (2010). "A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair." Mol Cell **39**(1): 36-47.
- Sorlie, T., C. M. Perou, et al. (2001). "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications." Proc Natl Acad Sci U S A **98**(19): 10869-10874.
- Sorlie, T., R. Tibshirani, et al. (2003). "Repeated observation of breast tumor subtypes in independent gene expression data sets." Proc Natl Acad Sci U S A **100**(14): 8418-8423.

- Sotiriou, C. and L. Pusztai (2009). "Gene-Expression Signatures in Breast Cancer." New England Journal of Medicine **360**(8): 790-800.
- Sparano, J. A. and S. Paik (2008). "Development of the 21-gene assay and its application in clinical practice and clinical trials." J Clin Oncol **26**(5): 721-728.
- Spencer, D. M., R. A. Bilardi, et al. (2008). "DNA repair in response to anthracycline-DNA adducts: a role for both homologous recombination and nucleotide excision repair." Mutat Res **638**(1-2): 110-121.
- Stadler, Z. K. and S. E. Come (2009). "Review of gene-expression profiling and its clinical use in breast cancer." Critical Reviews in Oncology/Hematology **69**(1): 1-11.
- Steel, L. F., M. G. Trotter, et al. (2003). "Efficient and specific removal of albumin from human serum samples." Mol Cell Proteomics **2**(4): 262-270.
- Straver, M., A. Glas, et al. (2010). "The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer." Breast Cancer Res Treat **119**(3): 551-558.
- Suckau, D., A. Resemann, et al. (2003). "A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics." Anal Bioanal Chem **376**(7): 952-965.
- Sunayama, J., F. Tsuruta, et al. (2005). "JNK antagonizes Akt-mediated survival signals by phosphorylating 14-3-3." J Cell Biol **170**(2): 295-304.
- Szklarczyk, D., A. Franceschini, et al. (2011). "The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored." Nucleic Acids Res **39**(Database issue): D561-568.
- Tan, N., M. Malek, et al. (2011). "Navitoclax enhances the efficacy of taxanes in non-small cell lung cancer models." Clin Cancer Res **17**(6): 1394-1404.
- Tan, S., H. T. Tan, et al. (2008). "Membrane proteins and membrane proteomics." Proteomics **8**(19): 3924-3932.
- Tewari, M., A. Krishnamurthy, et al. (2008). "Predictive markers of response to neoadjuvant chemotherapy in breast cancer." Surg Oncol **17**(4): 301-311.
- Therasse, P., S. G. Arbuck, et al. (2000). "New Guidelines to Evaluate the Response to Treatment in Solid Tumors." Journal of the National Cancer Institute **92**(3): 205-216.
- Thomas, H. and H. M. Coley (2003). "Overcoming multidrug resistance in cancer: an update on the clinical strategy of inhibiting p-glycoprotein." Cancer Control **10**(2): 159-165.
- Tian, S., P. Roepman, et al. (2010). "Biological functions of the genes in the mammaprint breast cancer profile reflect the hallmarks of cancer." Biomark Insights **5**: 129-138.
- Tiezzi, D. G., J. M. Andrade, et al. (2007). "HER-2, p53, p21 and hormonal receptors proteins expression as predictive factors of response and prognosis in locally advanced breast cancer treated with neoadjuvant docetaxel plus epirubicin combination." BMC Cancer **7**: 36.
- Tirumalai, R. S., K. C. Chan, et al. (2003). "Characterization of the low molecular weight human serum proteome." Mol Cell Proteomics **2**(10): 1096-1103.
- Towbin, H., T. Staehelin, et al. (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications" Proc Natl Acad Sci U S A **76**(9): 4350-4354.
- Tuck, M. K., D. W. Chan, et al. (2009). "Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group." J Proteome Res **8**(1): 113-117.
- Tzivion, G., M. Dobson, et al. (2011). "FoxO transcription factors; Regulation by AKT and 14-3-3 proteins." Biochim Biophys Acta.

- Tzivion, G., V. S. Gupta, et al. (2006). "14-3-3 proteins as potential oncogenes." Seminars in Cancer Biology **16**(3): 203-213.
- Uemura, N., Y. Nakanishi, et al. (2009). "Antibody-based proteomics for esophageal cancer: Identification of proteins in the nuclear factor- κ B pathway and mitotic checkpoint." Cancer Science **100**(9): 1612-1622.
- van 't Veer, L. J., H. Dai, et al. (2002). "Gene expression profiling predicts clinical outcome of breast cancer." Nature **415**(6871): 530-536.
- van de Vijver, M. J., Y. D. He, et al. (2002). "A gene-expression signature as a predictor of survival in breast cancer." N Engl J Med **347**(25): 1999-2009.
- van der Hage, J. A., C. J. van de Velde, et al. (2001). "Preoperative chemotherapy in primary operable breast cancer: results from the European Organization for Research and Treatment of Cancer trial 10902." J Clin Oncol **19**(22): 4224-4237.
- Vasquez, K. M. (2010). "Targeting and processing of site-specific DNA interstrand crosslinks." Environ Mol Mutagen **51**(6): 527-539.
- Vazquez, A., Grochola, L., Bond, E., Levine, A., Taubert, H., Muller, T., Wurl, P., Bond, G. (2010). "Chemosensitivity Profiles Identify Polymorphisms in the p53 Network Genes 14-3-3 tau and CD44 That Affect Sarcoma Incidence and Survival." Molecular and Cellular Pathobiology **70**(1): 172-180.
- Vermeulen, K., Z. N. Berneman, et al. (2003). "Cell cycle and apoptosis." Cell Prolif **36**(3): 165-175.
- Verrills, N. M., B. J. Walsh, et al. (2003). "Proteome Analysis of Vinca Alkaloid Response and Resistance in Acute Lymphoblastic Leukemia Reveals Novel Cytoskeletal Alterations." Journal of Biological Chemistry **278**(46): 45082-45093.
- Viswanathan, G. A., J. Seto, et al. (2008). "Getting started in biological pathway construction and analysis." PLoS Comput Biol **4**(2): e16.
- Von Eggeling, F., A. Gawriljuk, et al. (2001). "Fluorescent dual colour 2D-protein gel electrophoresis for rapid detection of differences in protein pattern with standard image analysis software." Int J Mol Med **8**(4): 373-377.
- Wallin, J. J., J. Guan, et al. (2010). "Nuclear phospho-Akt increase predicts synergy of PI3K inhibition and doxorubicin in breast and ovarian cancer." Sci Transl Med **2**(48): 48ra66.
- Walsh, J. G., S. P. Cullen, et al. (2008). "Executioner caspase-3 and caspase-7 are functionally distinct proteases." Proc Natl Acad Sci U S A **105**(35): 12815-12819.
- Wang, B., K. Liu, et al. (2010). "14-3-3{tau} Regulates Ubiquitin-Independent Proteasomal Degradation of p21, a Novel Mechanism of p21 Downregulation in Breast Cancer." Mol. Cell. Biol. **30**(6): 1508-1527.
- Wang, P., F. G. Bouwman, et al. (2009). "Generally detected proteins in comparative proteomics – A matter of cellular stress response?" Proteomics **9**(11): 2955-2966.
- Wang, X., L. Yi, et al. (2011). "AKT signaling pathway in invasive ductal carcinoma of the breast: correlation with ER α , ER β and HER-2 expression." Tumori **97**(2): 185-190.
- Watson, M., J. Greenman, et al. (2004). "Variation between independently cultured strains of the MDA-MB-231 breast cancer cell line identified by multicolour fluorescence in situ hybridisation." Cancer Therapy **2**: 167-172.
- Watson, M. B., M. J. Lind, et al. (2007). "Establishment of in-vitro models of chemotherapy resistance." Anticancer Drugs **18**(7): 749-754.
- Weigel, M. T. and M. Dowsett (2010). "Current and emerging biomarkers in breast cancer: prognosis and prediction." Endocr Relat Cancer **17**(4): R245-262.

- Westlake, S. (2009). "Cancer incidence and mortality in the United Kingdom and constituent countries, 2004-06." Health Stat Q(43): 56-62.
- Westlake, S. and N. Cooper (2008). "Cancer incidence and mortality: trends in the United Kingdom and constituent countries, 1993 to 2004." Health Stat Q(38): 33-46.
- Wingren, C. and C. A. Borrebaeck (2008). "Antibody microarray analysis of directly labelled complex proteomes." Curr Opin Biotechnol **19**(1): 55-61.
- Wosikowski, K., J. T. Regis, et al. (1995). "Normal p53 status and function despite the development of drug resistance in human breast cancer cells." Cell Growth Differ **6**(11): 1395-1403.
- Wright, F. C., J. Zubovits, et al. (2010). "Optimal assessment of residual disease after neo-adjuvant therapy for locally advanced and inflammatory breast cancer--clinical examination, mammography, or magnetic resonance imaging?" J Surg Oncol **101**(7): 604-610.
- Wu, C. C. and M. J. MacCoss (2002). "Shotgun proteomics: tools for the analysis of complex biological systems." Curr Opin Mol Ther **4**(3): 242-250.
- Wu, F., P. Wang, et al. (2010). "Studies of Phosphoproteomic Changes Induced by Nucleophosmin-Anaplastic Lymphoma Kinase (ALK) Highlight Deregulation of Tumor Necrosis Factor (TNF)/Fas/TNF-related Apoptosis-induced Ligand Signaling Pathway in ALK-positive Anaplastic Large Cell Lymphoma." Molecular & Cellular Proteomics **9**(7): 1616-1632.
- Wu, J., L. Y. Lu, et al. (2010). "The role of BRCA1 in DNA damage response." Protein Cell **1**(2): 117-123.
- Wysocki, V. H., K. A. Resing, et al. (2005). "Mass spectrometry of peptides and proteins." Methods **35**(3): 211-222.
- Xiang, T., Y. Jia, et al. (2011). "Targeting the Akt/mTOR pathway in Brca1-deficient cancers." Oncogene **30**(21): 2443-2450.
- Xue, H., B. Lu, et al. (2008). "The cancer secretome: a reservoir of biomarkers." J Transl Med **6**: 52.
- Yates, J. R., 3rd (2000). "Mass spectrometry. From genomics to proteomics." Trends Genet **16**(1): 5-8.
- Yates, J. R., C. I. Ruse, et al. (2009). "Proteomics by mass spectrometry: approaches, advances, and applications." Annu Rev Biomed Eng **11**: 49-79.
- Yin, X. M. (2006). "Bid, a BH3-only multi-functional molecule, is at the cross road of life and death." Gene **369**: 7-19.
- Yip, K. W. and J. C. Reed (2008). "Bcl-2 family proteins and cancer." Oncogene **27**(50): 6398-6406.
- Yoo, G. H., M. P. Piechocki, et al. (2002). "Docetaxel induced gene expression patterns in head and neck squamous cell carcinoma using cDNA microarray and PowerBlot." Clin Cancer Res **8**(12): 3910-3921.
- Zhang, J., Q. Tian, et al. (2005). "Insights into oxazaphosphorine resistance and possible approaches to its circumvention." Drug Resist Updat **8**(5): 271-297.
- Zhou, Q., Y. S. Kee, et al. (2010). "14-3-3 coordinates microtubules, Rac, and myosin II to control cell mechanics and cytokinesis." Curr Biol **20**(21): 1881-1889.
- Zou, L. (2010). "DNA repair: A protein giant in its entirety." Nature **467**(7316): 667-668.

APPENDIX 1: 725 Antibodies (Panorama Antibody Microarray XPRESS Profiler)

Antibody	Product number	Antibody	Product number	Antibody	Product number
14-3-3 q/t	T5942	BAD	B0559	Claspin	C7867
Acetylated Protein	A5463	BAF57	B0436	CaM Kinase IV (CaMKIV)	C2851
Actin	A5060	BAK	B5897	CaM Kinase Kinase a (CaMKKa)	C7099
Actin	A3853	BAP1	B9303	CaM Kinase II α (CaMKII α)	C6974
Actin, α -Smooth Muscle	A5228	Bax	B3428	CaM Kinase IV (CaMKIV)	C9973
β -Actin	A1978	Bax	B8429	CASK/LIN2	C4856
β -Actin	A2228	Bax	B8554	Casein Kinase 2 β	C3617
α -Actinin	A5044	Bax	B9054	Caspase 2	C7349
Actopaxin	A1226	Bcl-10	B7806	Caspase 3	C9598
AP2	A7107	Prion protein	P0110	Caspase 3, Active	C8487
β 1 and β 2-Adaptins	A4450	Bcl-10	B0431	Caspase 4	C4481
I-Afadin	A0349	Seladin	S4697	Caspase 4	C3392
AFX	A8975	Bcl-2	B9804	Caspase 5	C6979
AFX (FOXO4)	A5854	Bcl-2	B3170	Caspase 6	C7599
AKR1C3	A6229	Bcl-x	B9304	Caspase 7	C7724
Aly	A9979	Bcl-x _L	B9429	Caspase 7	C1104
β -Amyloid	A8354	BID	B4305	Caspase 8	C3101
Amyloid Precursor Protein, C-Terminal	A8717	BID	B3183	Caspase 8	C2976
Amyloid Precursor Protein, N-Terminal	A8967	Bim	B7929	Caspase 8	C4106
Amyloid Precursor Protein, KPI Domain	A8842	CDK5	C6118	Pro-Caspase 8	C7849
Androgen Receptor	A9853	Bmf, N-Terminal	B1684	Caspase 9	C7729
Annexin V	A8604	Bmf, C-Terminal	B1559	Caspase 9	C4356
Annexin VII	A4475	BNIP3	B7931	Caspase 10	C8351
AOP1	A7674	BOB.1/OBF.1	B7810	Caspase 10	C1229
AP-1	A5968	Brg1/hSNF2 β	B8184	Caspase 11	C1354
AP-2a	A0844	BTK, C-Terminal	B0811	Caspase 12	C7611
AP Endonuclease	A2105	BTK, N-Terminal	B0686	Caspase 13 (ERICE)	C8854
Apafl, N-Terminal	A8469	BUB1	B0561	Catalase	C0979
Apoptosis Inducing Factor (AIF)	A7549	BUBR1	B9310	α -E-Catenin	C8114
APRIL, Extracellular Domain	A1726	c-Abl	A5844	a-N-Catenin	C8239
APRIL, Extracellular Domain 2	A1851	c-Cbl	C9603	a-Catenin	C2081
ARC, C-Terminal	A8344	c-erbB-2	E2777	β -Catenin	C7207
ARNO (Cytohesin-2)	A4721	c-erbB-3	E8767	β -Catenin	C7082
Arp1a/Contractin	A5601	c-erbB-4	E5900	phospho- β -Catenin (pThr ⁴¹)	C8616
ARP2	A6104	phospho-c-Jun (pSer ⁶⁵)	J2128	phospho- β -Catenin (pSer ³³ /pSer ³⁷)	C4231
ARP3	A5979	phospho-c-Jun (pSer ⁷⁵)	J2253	phospho- β -Catenin (pSer ⁴⁵)	C5615
ARTS	A3720	c-Myc	M4439	phospho- β -Catenin (pSer ³³)	C2363
ARTS	A4471	c-Myc	C3956	δ -Catenin/NPRAP	C4864
ASAP1/Centaurin β 4	A4227	Uvomorulin/E-Cadherin	U3254	Cathepsin D	C0715
ASC-2	A5355	N-Cadherin	C2542	Cathepsin L	C2970
ASPP1	A4355	N-Cadherin	C2667	Caveolin-1	C3237
ASPP2	A4480	Pan Cadherin	C1821	CD40	C5987
ATF-1	A7833	Calbindin-D-28K	C7354	Cdc14A	C2238
ATF2	A4086	Calcineurin (a-Subunit)	C1956	Cdc25c	C0349
phospho-ATF-2 (pThr ^{69,71})	A4095	Caldesmon	C6542	Cystatin B	C5243
ATM	A6093	Calmodulin	C7055	Cdc27	C7104
ATM	A6218	Calnexin	C4731	Cdc6	C0224
Aurora-B	A5102	Calponin	C2687	Cdc7 Kinase	C6613
BACE-1	B0806	Calreticulin	C4606	Cdh1	C7855
BACH1	B1310	Calretinin	C7479	Cdk1 ^{p34cdc2}	C4973

Antibody	Product number	Antibody	Product number	Antibody	Product number
Cdk4	C8218	Cytokeratin peptide 19	C6930	E2F6	E1532
Cdk6	C8343	Pan Cytokeratin	C2931	phospho-FAK (pTyr ³⁹⁷)	F7926
Cdk-7/cak	C7089	DAPK	D2178	phospho-FAK (pTyr ³⁷⁵)	F8926
TBP	T1827	phospho-DAPK (pSer ³⁰⁸)	D4941	Falkor/PHD1	F5303
CENP-E	C7488	DAP Kinase 2	D3191	Fas (CD95/Apo-1)	F4424
Centrin	C7736	Daxx	D7810	Fas Ligand	F2051
Chk1	C9358	DcR1	D3566	Fas Ligand	F1926
Chk2	C9108	DcR2	D3188	FBI-1/PAKEMON	F9429
Chk2	C9233	DcR3	D1814	Fibroblast Growth Factor-9	F1672
Chondroitin Sulfate	C8035	DEDAF	D3316	Fibronectin	F0791
Ciliated Cell Marker	C5867	Desmin	D1033	Fibronectin	F3648
CIN85	C8116	Desmosomal Protein	D1286	Fibronectin	F7387
Casein Kinase 2 α	C5367	Destrin/ADF	D8940	Filamin	F1888
Clathrin Light Chain	C1985	Dnase I	D0188	Filensin	F1043
Clathrin Heavy Chain	C1860	Dnase II	D1689	FKHR (FOXO1a)	F6928
CNPase	C5922	DNMT1	D4567	FKHRL1 (FOXO3a)	F2178
Cofilin	C8736	DNMT1	D4692	FKHRL1 (FOXO3a)	F1304
Coilin	C1862	DOPA Decarboxylase	D0180	FLIP γ/δ , C-Terminal	F9925
Collagen, Type IV	C1926	DP2	D7438	FOXC2	F1054
Connexin 32	C3470	DR3	D3563	FOXP2	F6304
Connexin- 32	C6344	DR4	D3813	FANCD2	F0305
Connexin- 43	C8093	DR5	D3938	FXR2	F1554
Connexin- 43	C6219	DR6	D1564	FRS2 (SNT-1)	F9052
β -COP	G6160	DRAK1	D1314	G9a Methyltransferase	G6919
Cortactin	C6987	Dystrophin	D8168	Glutamic Acid Decarboxylase 65 (GAD 65)	G4913
Corticotropin Releasing Factor	C5348	Dystrophin	D8043	Glutamic Acid Decarboxylase 65 (GAD 65)	G5038
COX II	C9354	E2F1	E9026	Glutamic Acid Decarboxylase (GAD65/67)	G5163
Crk-L	C0978	E2F1	E8901	GADD 153 (CHOP-10)	G6916
Crk II	C0853	E2F2	E8776	GAP1 ^{IP4BP}	G6666
Csk	C7863	E2F3	E8651	GAPDH	G8795
CtBP1, N-Terminal	C9491	E2F4	E8526	GATA-1	G0290
CtBP1, C-Terminal	C8741	E6AP	E8655	Gelsolin	G4896
CUG-BP1	C5112	EGF receptor	E3138	Gemin 2	G6669
Cyclin A	C4710	ERK5 (Big MAPK-BMK1)	E1523	Gemin 3	G6544
Cyclin B ₁	C8831	Elastin	E4013	GFAP (Glial Fibrillary Acidic Protein)	G9269
Cyclin D ₁	C5588	ELKS	E4531	GFAP (Glial Fibrillary Acidic Protein)	G3893
Cyclin D ₁	C7464	Endothelial Cell Protein C Receptor	E6280	Growth Factor Independence-1 (GFI)	G6670
Cyclin D ₂	C7339	Endothelial Cells	E9653	Glutamate receptor NMDAR 2a	G9038
Cyclin D ₃	C7214	Endothelin	E0771	Glutamine Synthase	G2781
Cyclin H	C5351	Epidermal Growth Factor	E2520	Glycogen Synthase Kinase-3 β (GSK-3 β)	G7914
Cystatin A	C3095	Episialin (EMA)	E0143	Glycogen Synthase Kinase-3 (GSK-3)	G4414
Cytohesin-1	C8979	ERP57	E5031	Glycogen Synthase Kinase-3 (GSK-3)	G6414
Cytokeratin peptide 4	C5176	Estrogen Receptor	E0521	Granzyme B	G1044
Cytokeratin CK5	C7785	Estrogen Receptor	E1396	Grb-2	G2791
Cytokeratin peptide 7	C6417	Exportin T	E1531	GRK 2	G7670
Cytokeratin 8.12	C7034	Ezrin	E8897	GRP1	G6541
Cytokeratin 8.13	C6909	F1A	F3428	GRP 75	G4170
Cytokeratin peptide 13	C0791	FADD	F8053	GRP78/BiP	G8918
Cytokeratin Peptide 17	C9179	Focal Adhesion Kinase (pp125 ^{FAK})	F2918	GRP94	G4420
Cytokeratin peptide 18	C1399	FAK Phospho (pSer ⁷⁷³)	F9051	hABH1	A8103

Antibody	Product number	Antibody	Product number	Antibody	Product number
hABH2	A8228	hnRNP-Q	R5653	MAP Kinase, Monophosphorylated Threonine	M3557
hABH3	A8353	hnRNP-U	R6278	MAP Kinase (ERK-1)	M7927
hBRM/hSNF2 α	H9787	hnRNP M3-M4	R3777	MAP Kinase (ERK1+ERK2)	M5670
HAT1 (Histone acetyltransferase 1)	H7161	hPlk1	P5998	MAP Kinase Activated Protein Kinase-2 (MAPKAPK-2)	M3550
HDAC-1	H3284	hPlk1	P6123	MAP Kinase Phosphatase-1 (MKP-1)	M3787
HDAC-1	H6287	hSNF5/INI1	H9912	MAPK non phosphorylated ERK	M3807
HDAC-2	H3159	iASPP	A4605	MAP Kinase 2 (ERK-2)	M7431
HDAC-2	H2663	IFI-16	I1659	MAP Kinase Kinase (MEK, MAPKK)	M5795
HDAC-3	H6537	IkBa	I0505	MAP2 (2a+2b)	M2320
HDAC-3	H3034	IKKa	I6139	MAP1	M4278
HDAC-4	H9411	ILK	I0783	MAP1 (Light Chain)	M6783
HDAC-4	H9536	ILK	I1907	MAP1b	M4528
HDAC-5	H4538	ILP2	I4782	MAP2	M9942
HDAC-5	H8163	Importin-a1	I9658	MBD1	M6569
HDAC-6	H2287	Importin-a3	I9783	MBD2a	M7568
HDAC-7	H2537	Importin-a5/7	I9908	MBD2a,b	M7318
HDAC-7	H6663	INCENP	I5283	MBD4	M9817
HDAC-8	H6412	ING1	I3659	MBDin/XAB1	M1944
HDAC-10	H3413	a-Interneixin	I0282	MBNL 1	M3320
HDAC-11	H2913	JAB 1	J3395	MCH	M8440
HDRP/MITR	H9163	JAB 1	J3020	Mcl-1	M8434
Heat Shock Factor 1	H4163	JAK 1	J3774	MDC1	M2444
Heat Shock Factor 2	H6788	c-Jun N-Terminal Kinase JNK, Activated (Diphosphorylated JNK)	J4500	MDM2	M8558
Heat Shock Protein 25	H0148	KCNK9 (TASK-3)	K0514	MDM2	M4308
Heat Shock Protein 27	P1498	Kaiso	K4263	MDMX	M7815
Heat Shock Protein 27/25	H2289	KIF17	K3638	MeCP2	M0445
Heat Shock Protein 70	H5147	KIF3A	K3513	MeCP2	M9317
Heat Shock Protein 90	H1775	KSR	K4261	MeCP2	M7443
Heat Shock Protein 110	H7412	Ku Antigen	K2882	MEKK4	M6818
Heat Shock Protein 110	H7287	L1CAM	L4543	Melanocortin-3 Receptor	M7194
Acetyl Histone H3 (Ac-Lys ⁹)	H9286	l/s-Afadin	A0224	MGMT	M4937
Acetyl Histone H3 (Ac-Lys ⁹)	H0913	Laminin	L9393	Mint2	M3068
Acetyl- & phospho-Histone H3 (Ac-Lys ⁹ , Ser ¹⁰)	H9161	Laminin-2 (a-2 Chain)	L0663	LRRK2 (PARK8)	M3319
Acetyl- & phospho-Histone H3 (Ac-Lys ⁹ , Ser ¹⁰)	H0788	LAP2 (TMPO)	L3414	MRP1	L3044
Dimethyl Histone H3 (diMe-Lys ⁴)	D5692	Leptin	L3410	MRP2	M9192
Dimethyl Histone H3 (diMe-Lys ⁹)	D5567	LIM Kinase 1	L2290	a-MSH	M3692
methyl-Histone H3 (Me-Lys ³)	H7162	LIN-7	L1538	MSH6	M0939
phospho-Histone H2AX (pSer ¹³⁹)	H5912	LIS1	L7391	MSH6	M2445
phospho-Histone H3 (pSer ¹⁰)	H6409	LKB1	L7917	MSK-1	M2820
phospho-Histone H3 (pSer ²⁸)	H9908	LDS1	L4793	MTA 2	M5437
phospho-Histone H3 (pSer ¹⁰)	H0412	Mad1	M8069	MTA1	M7569
SUV39H1 Histone Methyl Transferase	S8316	Mad2	M8694	MTA1	M1320
HMG-1	H9537	MADD	M5683	MTA2/MTA1L	M7693
hMps1	M5818	MAFF	M8194	MTA3L	M7818
hnRNP-A1	R4528	MAGI-1	M5691	MTBP	M0819
hnRNP-A1	R9778	MAGI-2	M2441	mTOR	M3566
hnRNP-A2/B1	R4653	MAP Kinase, Activated/ Monophosphorylated (Phosphothreonine ERK-1&2)	M7802	Munc-18-1	T2949
hnRNP-C1/C2	R5028	MAP Kinase, Monophosphorylated Tyrosine	M3682	Munc-13/1	M2694
hnRNP-K/J	R8903	MAP Kinase, Activated (Diphosphorylated ERK-1&2)	M9692	MyD88	M6194
hnRNP-L	R4903				M9934

Antibody	Product number	Antibody	Product number	Antibody	Product number
Myosin	M1570	p19 ^{INK4d}	P4354	Protein Kinase C β_2	P2584
Myosin I β (Nuclear)	M3567	p21WAF1/Cip1	P1484	Protein Kinase C γ	P8083
Myosin IIA	M8064	p300/CBP	P2859	Protein Kinase C δ	P8333
Myosin IX/Myr5	M5566	p34 ^{cdc2}	C3085	Protein Kinase C ϵ	P8458
Myosin Light Chain Kinase	M7905	p35 (Cdk5 Regulator)	P9489	Protein Kinase C ζ	P0713
Myosin Va	M4812	p38 MAP Kinase, Non-Activated	M8432	Protein Kinase C η	P8090
Myosin Va	M5062	p38 MAPK	M0800	Protein Kinase D	P3987
Myosin VI	M0691	p38 MAPK activated (diphosphorylated p38)	M8177	PKR	P0244
Myosin VI	M5187	p53	P5813	Plakoglobin (Catenin g)	P8087
NBS1 (Nibrin)	N9287	p53	P6874	Platelet-Derived Growth Factor Receptor β	P7679
NBS1 (Nibrin)	N3037	phospho-p53 (pSer ³⁹³)	P8982	Plectin	P9318
NBS1 (Nibrin)	N3162	p53DINP1/SIP	P4868	PML	P6746
Nck-2	N2911	p53R21	P4993	Presenilin-1 (S182)	P7854
Nedd 8	N2786	p53 BP1	B4561	Prion Protein	P5999
Nerve Growth Factor- β	N3279	p53 BP1	B4436	PRMT1	P6871
Nerve Growth Factor Receptor	N5408	p57 ^{kip2}	P2735	PRMT1	P6996
Nerve growth factor receptor (NGFR p75)	N3908	p63	P3362	PRMT2	P0748
Neurabin I	N4412	p63	P3737	PRMT3	P9370
Neurabin II (C-terminal)	N5037	PABP	P6246	PRMT4	P4995
Neurabin-II	N5162	PAD14	P4749	PRMT5	P0493
Neurofibromin	N3662	phospho-PAK1 (pThr ²¹²)	P3237	PRMT6	P6495
Neurofilament 160	N2787	Par-4 (Prostate Apoptosis Response-4)	P5367	PRMT6	P2996
Neurofilament 200	N4142	γ Parvin	P5746	Proliferating Cell Protein Ki-67	P6834
Neurofilament 200	N0142	Parkin	P6248	Protein Phosphatase 1a	P7979
Neurofilament 200	N5389	PARP	P7605	Protein Phosphatase 1a	P7607
Neurofilament 68	N5139	Paxillin	P1093	Protein Phosphatase 2Aa (PP2Aa)	P8998
Neurofilament 160/200	N2912	PCAF	P7493	Protein S	P4555
NF- κ B	N8523	Proliferating Cell Nuclear Antigen (PCNA)	P8825	Protein Tyrosine Phosphatase PEST	P9109
NAK (NF κ B-Activating Kinase)	N2661	PDK 1	P3110	PSF	P2860
NG2	N8912	Pen-2	P5622	PTEN	P7482
Nicastrin	N1660	Peripherin	P5117	PTEN	P3487
Nitric Oxide Synthase, Brain (b-NOS)	N2280	Peroxiredoxin 3	P1247	PUMA/bbc3, C-Terminal	P4618
Nitric Oxide Synthase, Brain (b-NOS)	N7155	PERP	P5243	PUMA/bbc3, N-Terminal	P4743
Nitric Oxide Synthase, Endothelial (e-NOS)	N9532	Phospholipase A2 group V	P5242	Pyk2	P3902
Nitric Oxide Synthase, Endothelial (e-NOS)	N3893	Phosphoserine	P5747	AP2 beta	A9856
Nitric Oxide Synthase, Endothelial (e-NOS)	N2643	Phosphothreonine	P6623	phospho-Pyk2 (pTyr ^{579/580})	P6989
Nitric Oxide Synthase, Inducible (i-NOS)	N7782	Phosphotyrosine	P1869	AP2 alpha	A9981
Nitric Oxide Synthase, Inducible (i-NOS)	N9657	Phospholipase C γ 1 (PLC γ 1)	P8104	Rab5	R7904
Notch1	N6786	PhosphatidylSerine Receptor (PSR)	P1495	Rab 7	R8779
Nitrotyrosin	N0409	PIAS-x	P9498	Rab9	R5404
NTF2	N9527	PINCH-1	P9371	RAD1	R5029
Nuf2	N5287	Protein Kinase Ba /Akt1	P2482	Rad17 (C-terminal)	R8029
O-GlcNAc Transferase	O6264	Protein Kinase Ba /Akt1	P1601	Raf-1/c-Raf	R2404
OP-18/Stathmin	O0138	phospho-PKB (pSer ⁴⁷³)	P4112	Raf-1	R5773
Ornithine Decarboxylase (ODC)	O1136	phospho-PKB (pThr ³⁰⁸)	P3862	Sorting Nexin 6 (SNX6)	S6324
p115/TAP	P3118	Protein Kinase C (PKC)	P5704	RAIDD, Internal Domain	R9775
p120 ^{cm}	P1870	Protein Kinase Ca	P4334	RAIDD	R5275
p130 ^{CAS}	C0354	Protein Kinase C β_1	P3078	RALAR	R8529
p14 ^{arf}	P2610	Protein Kinase C β_1	P6959	Ran	R4777
p16 ^{INK4a/CDKN2}	P0968	Protein Kinase C β_2	P3203	PIASy	P0104

Antibody	Product number	Antibody	Product number	Antibody	Product number
RAP1	R8154	SMAC/Diablo	S0941	Tumor Necrosis Factor-a	T2824
RbAp48/RbAp46	R3779	SUMO-1	S8070	Nanog	N3038
Reelin	R4904	SUMO-1 (C-terminal)	S5446	TWEAK Receptor/Fn-14	T9700
Retinoblastoma	R6775	Survivin	S8191	Tyrosin hydroxylase	T2928
phospho-Retinoblastoma (pSer ⁷⁹⁵)	R6878	Synaptotagmin	S2177	U2AF ⁶⁵	U4758
RhoE	R6153	Synaptopodin	S9442	Ubiquitin	U0508
RICK, C-Terminal	R9650	Synaptopodin	S9567	Ubiquitin C-terminal Hydrolase L1	U5133
RIP (Receptor Interacting Protein)	R8274	SynCAM	S4945	Ubiquitin C-terminal Hydrolase L1	U5258
RNase L	R3529	a1 Syntrophin	S4688	Pinin	P0084
ROCK-1	R6028	a1 Syntrophin	S4813	Vanilloid Receptor-1	V2764
ROCK-2	R8653	Syntaxin	S0664	VDAC/Porin	V2139
Rsk1	R5145	Syntaxin 6	S9067	Vascular Endothelial Growth Factor Receptor-1 (VEGFR-1)	V4762
S-100	S2644	Syntaxin 8	S8945	Vesicular GABA Transporter	V5764
S-100 (a-Subunit)	S2407	a-Synuclein	S3062	VGLUT 1	V0389
S-100 (β-Subunit)	S2532	Tal	T1075	VGLUT 2	V2639
S-Nitrosocysteine	N5411	Tal	T1200	Vimentin	V6389
S6 Kinase	S4047	TAP	T1076	Vinculin	V4505
SAPK3	S0315	Tau	T9450	Vitronectin	V7881
Spectrin (a and β)	S3396	phospho-Tau (pSer ^{199/202})	T6819	WAVE	W0392
Serine/Threonine Protein Phosphatase 2 A/A	P8109	Tau	T5530	WSTF	W3516
Qa-1b	Q4962	Tenascin	T2551	Y14	Y1253
Serine/Threonine Protein Phosphatase 1g1	P7609	Thimet Oligopeptidase 1	T7076	ZAP-70	Z0627
Serine/Threonine Protein Phosphatase 2 A/Bg	P5359	TIS7	T2576	Zip Kinase	Z0134
Serine/Threonine Protein Phosphatase 2 A/B' pan2	P8359	Tumor Necrosis Factor Soluble Receptor II	T1815	Zyxin	Z0377
Serine/Threonine Protein Phosphatase 2C a/b	P8609	Tob	T2948	GAPDH	G8795
AP2 gamma	A3108	TOM22	T6319		
SGK	S5188	Topoisomerase-I	T8573		
SH-PTP2 (SHP-2)	S3056	TRAIL	T3067		
Siah2	S7945	TRAIL	T9191		
Sin3A, N-terminal	S4445	Transforming Growth Factor-β, pan	T9429		
Sin3A, C-Terminal	S6695	Transportin 1	T0825		
Sir2	S5313	TRF1	T1948		
SIRPa1 (SHPS-1)	S1311	Tropomyosin	T2780		
Sirt1	S5196	Tropomyosin (Sarcomeric)	T9283		
SKM1 (Skeletal Muscle Type 1)	S9568	Tryptophane Hydroxylase	T0678		
Beta tubulin III (neuronal)	T8578	TSG101	T5826		
SLIPR/MAGI-3	S1190	a-Tubulin	T6074		
SLIPR/MAGI-3	S4191	a-Tubulin	T6199		
Smad4 (DPC4)	S3934	b-Tubulin	T5201		
SMC1L1	S6446	β-Tubulin I	T7816		
SMN	S2944	b-Tubulin I+II	T8535		
a/b -SNAP, C-terminus	S9444	b -Tubulin III	T5076		
SNAP-23	S2194	β-Tubulin IV	T7941		
SNAP-25	S9684	g-Tubulin	T5326		
SNAP- 29	S2069	g-Tubulin	T3559		
Sos1	S2937	g -Tubulin	T3320		
Sp1	S9809	ε-Tubulin	T1323		
Spred-2	S7320	Tubulin, Polyglutamylated	T9822		
Striatin	S0696	Tubulin, Tyrosine	T9028		
Substance P Receptor	S8305	Tumor Necrosis Factor-a	T8300		

APPENDIX 2: Buffers and reagents

Tissue culture media

1 bottle of RPMI 1640 culture media (#31870, Invitrogen)
50 ml Fetal Bovine Serum (#10106, Invitrogen)
5 ml L-glutamine (#25030, Invitrogen)
5 ml Fungizone – Amphotericin B (#15290, Invitrogen)
5 ml Penicillin / Streptomycin (PenStrep) (#15140, Invitrogen)

Western blot (WB) extraction buffer

4 ml dH₂O
1 ml 0.5 M Tris:HCl pH 6.8
0.8 ml glycerol
1.6 ml 10 % SDS
200 µl 0.05 % Bromophenol Blue

TBSTween-20

<i>TBS Stock (concentrated)</i>	250 ml of TBS Stock (concentrated)
121 g Trizma Base (#93304, Fluka)	4750 ml dH ₂ O
170 g Sodium chloride (#S3014, Sigma Aldrich)	2.5 ml Tween20 (#P5927, Sigma Aldrich)

Made to 1 litre with dH₂O
Adjusted to pH 7.6 with concentrated HCl

2D extraction buffer

1.26 g Urea
0.456 g Thiourea
0.12 g CHAPS
0.0231 g Dithiothreitol (DDT)
30 µl Bio-Lyte 3/10 Ampholyte (#163-1113, Bio-Rad)
6 µl 1% Bromophenol Blue
1.65 ml dH₂O
30 µl Protease Inhibitor (#80-6501-23, Amersham Biosciences)
30 µl Phosphatase Inhibitor Cocktail 1 (#P2850, Sigma Aldrich)
30 µl Phosphatase Inhibitor Cocktail 2 (#P5726, Sigma Aldrich)

Equilibration buffer

<i>Stock</i>	<i>Equilibration Buffer I</i>
6.7 ml 1.5 M Tris-HCl pH 8.8	0.1 g DTT to every 10 ml of stock
72.07 g Urea	<i>Equilibration Buffer II</i>
69 ml 87% Glycerol	0.25 g IAA to every 10 ml of stock
4.0 g SDS	
Trace Bromophenol Blue Salt	
Made up to 200 ml with dH ₂ O	

1% Overlay Agarose

1 g Agarose
100 ml 1 x Tris-glycine running buffer (#161-0772, Bio-Rad)
Trace Bromophenol Blue

APPENDIX 3: REC Approval for the study



National Research Ethics Service

South Humber Local Research Ethics Committee

Room FC27
Coniston House
Trust Headquarters
Wilderby Hill Business Park
Wilderby
HULL
HU10 6NS

Telephone: 01482 380157
Facsimile: 01482 303918

30 April 2007

Dr Lynn Cawkwell
Senior Lecturer
University of Hull/HEY NHS Trust
R&D Building
Castle Hill Hospital
Hull
HU16 5JQ

Dear Dr Cawkwell

Full title of study: The identification of biomarkers associated with therapy response in breast cancer.
REC reference number: 07/Q1105/43

The Research Ethics Committee reviewed the above application at the meeting held on 25 April 2007. Thank you for attending to discuss the study.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

The Committee agreed that all sites in this study should be exempt from site-specific assessment (SSA). There is no need to submit the Site-Specific Information Form to any Research Ethics Committee. The favourable opinion for the study applies to all sites involved in the research.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Application	Version 5.3	12 April 2007
Investigator CV	Version 2	14 April 2007
Protocol	Version 1	13 April 2007

This Research Ethics Committee is an advisory committee to Yorkshire and The Humber Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within
the National Patient Safety Agency and Research Ethics Committees in England

Covering Letter	Version 1	12 April 2007
Participant Information Sheet	Version 1	13 April 2007
Participant Consent Form	Version 1	13 April 2007

R&D approval

All researchers and research collaborators who will be participating in the research at NHS sites should apply for R&D approval from the relevant care organisation, if they have not yet done so. R&D approval is required, whether or not the study is exempt from SSA. You should advise researchers and local collaborators accordingly.

Guidance on applying for R&D approval is available from
<http://www.rdforum.nhs.uk/rdform.htm>

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Feedback on the application process

Now that you have completed the application process you are invited to give your view of the service you received from the National Research Ethics Service. If you wish to make your views known please use the feedback form available on the NRES website at:

<https://www.nresform.org.uk/AppForm/Modules/Feedback/EthicalReview.aspx>

We value your views and comments and will use them to inform the operational process and further improve our service.

07/Q1105/43	Please quote this number on all correspondence
--------------------	---

With the Committee's best wishes for the success of this project

Yours sincerely

pp *K. Waltham*

Dr Stefan Herber
Chair

Email: karen.waltham@nres.nhs.uk

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
Standard approval conditions [SL-AC1 for CTIMPs, SL-AC2 for other studies]

Patient Information Sheet

Study Title: **Biomarkers in Breast Cancer**

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish.

- Part 1 tells you the purpose of this study and what will happen to you if you take part.
- Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

We would like to understand why some breast cancers respond well to chemotherapy whilst others respond less well. We will analyse proteins from a series of breast cancers to see if we can find any which correlate with how well the cancer responded to therapy. In the future we may be able improve breast cancer therapy by tailoring treatment according to the protein profile of the individual tumour.

Why have I been chosen?

We are inviting any women who have been undergoing chemotherapy for breast cancer and are now going to have an operation to remove any remaining cancer. We would like to take a sample of the surplus cancer tissue after your operation. We would like to study tissue samples from approximately 25-50 women.

Do I have to take part?

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to change your mind and withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?

You do not have to do anything. During your operation the surgeons will remove any remaining breast cancer tissue. Once this has been done they will take a small bit of surplus cancer tissue from what they have removed and send it to the Research Laboratory. The sample will be coded to protect your identity. **[1]**

If we have tumour tissue left over after the laboratory research we would like to store it for possible use in future breast cancer research projects. If this is the case further Research Ethics Committee Approvals will be sought. **[2]**

We would like to look at some parts of your medical records which relate to your treatment for breast cancer. **[3]**

We may also like to run some additional research tests on the biopsy sample which you had taken a few months ago when your breast cancer was first diagnosed. All

biopsy samples are routinely stored by the HistoPathologists and we would like your consent to use some of the tissue for research purposes. **[4]**

What are the possible disadvantages and risks of taking part?

None. You do not have to do anything.

What are the possible benefits of taking part?

There is no clinical benefit for you if you take part but the information we get might help improve the treatment for women with breast cancer in the future.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?

Yes. All the information about your participation in this study will be kept confidential. The details are included in Part 2.

Contact Details:

Breast Unit Doctors	Research Scientist
Mr Veera Garimella Tel 01482 875875 ext 2638 Mob 07970 517076	Dr Lynn Cawkwell Tel 01482 875875 ext 3617 L.Cawkwell@hull.ac.uk
Prof Philip Drew Tel 01482 463299	
Prof Mike Lind Tel 01482 676807	

This completes Part 1 of the Information Sheet.

If the information in Part 1 has interested you and you are considering participation, please continue to read the additional information in Part 2 before making any decision.

Part 2

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak with the researchers who will do their best to answer your questions (see phone numbers above). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital.

In the event that something does go wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for

compensation against Hull & East Yorkshire NHS Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you (if appropriate).

Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. The tissue sample will be coded before it is sent to the Research Laboratory so it is not readily identifiable and your personal details will be kept separately locked away in a secure office. We need to keep these details so that members of the research team can find your hospital notes and collect information about your breast cancer treatment and details of your previous biopsy.

What will happen to any samples I give?

The tissue sample you give for research will be used in the laboratory to find protein markers which correlate with therapy response. If any sample is left we will store it for possible use in future breast cancer projects- if this is the case further Research Ethics Committee Approvals will be sought.

Will any genetic tests be done?

No.

What will happen to the results of the research study?

The results will be published in scientific journals and presented at national and international conferences. You will not be identified in any publication or presentation. If you would like a lay-persons summary of the findings please contact Dr Cawkwell (details above).

Who is organising and funding the research?

The study is funded by the Hull York Medical School and the University of Hull. It is sponsored by Hull & East Yorkshire NHS Trust.

Who has reviewed the study?

This study was given a favourable ethical opinion for conduct in the NHS by the South Humber Research Ethics Committee.

You will be given a copy of the information sheet and a signed consent form to keep.

Thank you for considering taking part or taking time to read this sheet.

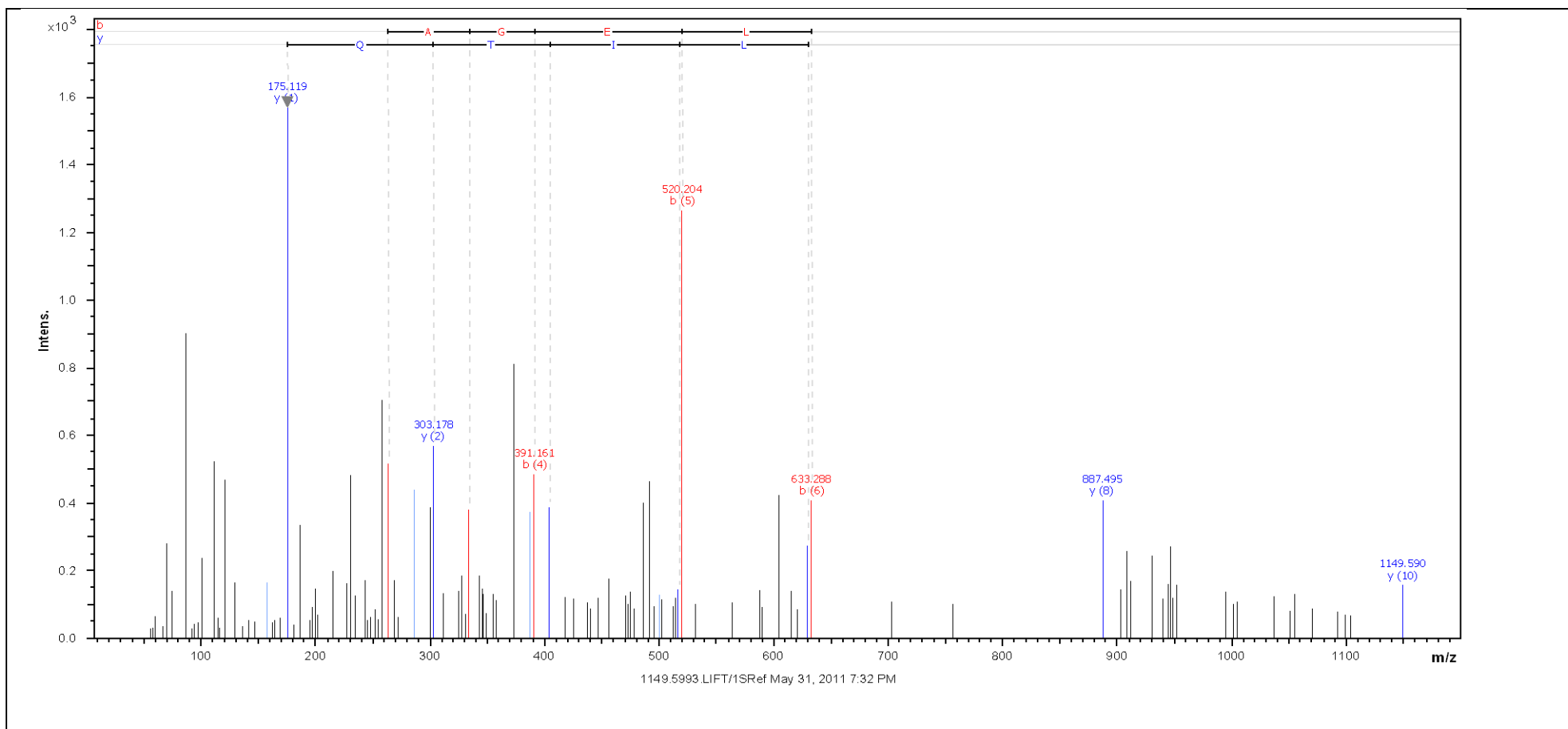
APPENDIX 5: All raw clinical data for clinical samples collected

Sample number	Ductal / Lobular	Grade	Molecular Subtype	Tumour size pre-chemotherapy		Therapy administered	Tumour size post-chemotherapy		Tumour size (pathology)	Response (% increase/reduction)	Comments
				MRI	US		MRI	US			
#1	Ductal	Grade 3 (poor)	ER+ PR- HER2-	-	24mm mass + diffuse area	ECx4 Docetaxel x 4	-	many nodules max 4mm. Diffuse area of 90mm	max overall 120mm composed of nodules/ foci of varying size	Progression (based on metastasis) Non-responder	No pre-MRI and US size is ambiguous Patient developed left side pleural effusion. Also increasing axillary lymph nodes. Metastatic carcinoma
#3	Ductal	Grade 3 (poor)	ER- PR- HER2+	-	30mm	neoTANGO. EC x 6 only	-	29mm	97mm	Stable disease non-responder	
#4	Ductal	Grade 2 (mod)	ER+ PR+ HER2-	26mm	16mm	neoTANGO ECx4 and Paclitaxel x4	17mm	2 patches with overall area 30mm	27mm	3.8% increase Non-responder	
#5	Ductal	Grade 2 (mod)	ER- (+ in core) PR- HER2- CK5/6+ CK14-	40mm	-	neoTANGO arm 2b Pac +Gem.x4, EC x 4	27mm and a new lesion of 10mm. Fragmented tumour spread over 50mm	-	38mm	5% reduction non-responder progressive disease (due to new lesion)	opposite neoTANGO. Change in subtype?? Hull reviewed Grimsby's core & confirmed ER+ (5/8). Mastectomy showed ER- (2/8). Proposed either poor fixation of mastectomy tissue to give false negative OR receptor status change during therapy
#6	Ductal	Grade 2 (mod)	ER+ PR+ HER2-	70mm	35mm	neoTANGO arm1B ECx4 then PAC+GEM x4 (20% reduced PAC for cycles 3&4)	34mm	-	25mm	64.3% reduction responder	
#7	Ductal	Grade 3 (poor)	ER+ PR+ HER2-	40mm	33mm	EC x 4, D x 2 (reduced by 20%) unwell	-	24mm	18mm	55% reduction responder	
#8	Ductal	Grade 3 (poor)	ER- PR- HER2-	55mm	13.7mm	neoTANGO arm 1a EC x4 and Paclitaxelx4	-	8mm & 'something else' at 38mm	no. of foci in residual tumour; largest at 15mm & 11 more foci of 1-2mm. = 30mm altogether	45.5% reduction Responder	
#9	Ductal	Grade 3 (poor)	ER+ PR+ HER2+	25mm	37mm	EC x 4 (no response, size increase - early surgery)	-	during EC - 28mm	27mm 'minimal effects of chemotherapy' reported	8% increase non-responder	
#10	Ductal (core) Metaplastic (resection)	Grade 3 (poor)	ER - PR- HER2 ?	46mm	33mm	EC x 4 Docetaxel x 4	25mm	-	chemotherapy changes', poorly differentiated, metaplastic type 5mm	89.1% reduction responder	Nothing to see/ feel. Fatty sample
#11	Ductal	Grade 1 (well) from resection	ER+ PR+ HER2-	42mm	9mm	EC x 4 Docetaxel x 4	33mm	-	max 25mm effects of chemotherapy reported	40.5% reduction responder	
#12	Ductal	Grade 3 (poor)	ER+ PR+ HER2? (HER2 not tested)	30mm and a 2nd lesion of 10mm	29mm	EC x 4, D x 2 (2nd at 20% reduced dose) could not tolerate side effects	21mm	16mm	main lesion 18mm 'apparent chemotherapy effects'	40% reduction responder	

Sample number	Ductal / Lobular	Grade	Molecular Subtype	Tumour size pre-chemotherapy		Therapy administered	Tumour size post-chemotherapy		Tumour size (pathology)	Response (% increase/reduction)	Comments
				MRI	US		MRI	US			
#13	Ductal	Grade 3 (poor)	ER- (core 3/8 +ve) PR- HER2 + CK5/6 +	28mm	23mm	EC x 4, D x 1 (no response)	-	30mm	22mm	21.4% reduction non-responder	
#15	Ductal	Grade 1 (well)	ER+ PR+ HER2 -	39mm	35mm	EC x 4 Docetaxel x 4	22mm	-	12mm	69.2% reduction responder	
#16	Ductal	Grade 1 (well)	ER+ PR+ HER2?	31mm	-	EC x 4 Docetaxel x 4	after EC 22mm	-	26mm	16.1% reduction non-responder	
#17	Lobular	Grade 2 (mod)	ER+ PR+ HER2-	48mm	45mm	EC x 3, no response so abandoned	-	41mm	at least 60mm 'no evidence of chemotherapy effects'	25% increase non-responder	
#18	Ductal	Grade 3 (poor)	ER+ PR+ HER2-	32mm	-	EC x 4, D x 2 (side effects)	16mm	-	7mm	78.1% reduction responder	clip to mark
#19	Ductal	Grade 3 (poor)	ER+ PR- HER2-	28mm nodule whole area 75mm	24mm	EC x 2, progressed - early surgery	nodule 38mm. 62mm overall new axillary and intra-mammary malignant lymph nodes (progression)	26mm	max 48mm	71.4% increase non-responder progression based on metastasis rather than size	taken 28mm pre-MRI as progressed on treatment
#20	Ductal	Grade 3 (poor)	ER- PR- HER2-CK 5/6+ CK14-	-	35 mm 'solid mass'	EC x 4 increase in size	-	54mm	52mm	54.3% increase non-responder	compare US to US
#22	Ductal	Grade 2 (mod)	ER+ PR- HER2-	50mm (mixed invasive and DCIS)	40mm	EC x 4 Docetaxel x 4	40mm	25mm	20mm inc. DCIS (high grade)	60% reduction responder	both measurements have DCIS included
#23	Ductal	Grade 2 (mod)	ER+ PR- HER2-	26mm max but 2 other areas 15mm away (multifocal)	27mm	EC x 4, D x 3 but 2 and 3 were at 25% dose (ill) but good response seen	13mm	-	14mm	46.2% reduction responder	
#24	Lobular	Grade 2 (mod)	ER+ PR+ HER2-	66mm	-	EC x 4 Docetaxel x 4	decreased in volume by 65% after EC. No further red. After D 50mm	-	multifocal. Largest focus 7mm. Extensive LCIS. Involves all 4 quadrants	24.2% reduction non-responder. Stable disease	no size from pathology - implies tumour throughout breast, so 7mm is under-estimate. Take post-treatment MRI
#25	Ductal	Grade 3 (poor)	ER+ PR+ HER2+	Only have Pre-treatment CT scan. 29mm	T1 14mm T2 12mm total 26mm	had surgery & chemo previously on other breast. ECx2 (already received). Doc x 4 (side effects so last 2 doses reduced by 25%)	T1 16mm T2 19mm take total = 35mm	After EC: 11.5 and 7 (total 18.5) After D: 13mm and 6mm (total 19)	3 lesions T1 10mm T2 10mm T3 6mm + 20mm DCIS. Multifocal	26.9 % reduction. Stable disease	Ultrasound shows stable disease throughout. can't use US for RECIST but should compare like for like. Shouldn't use MRI-post due to taxane. CT doesn't ID multifocality. We have sample from T1
#27	Tubular (core was ductal)	Grade 2 (mod)	ER+ PR+ HER2-	extensive malignancy 78mm	2 x foci of diffuse change. Whole area 27mm	EC x 4 Docetaxel x 4	fragmented. 24mm for largest	-	Grade 1 Tubular. 20mm	74.4% reduction responder	
#28	Ductal	Grade 2 (mod)	ER+ PR+ HER2+ luminal B	80mm diffuse thickening extensive malignancy	multifocal lesion largest 33mm	EC x 4. Docetaxel x 2 at full dose & docetaxel x 2 at 20% dose	57mm	-	partial response to chemotherapy. 33 mm tumour	58.8% reduction Responder	

Sample number	Ductal / Lobular	Grade	Molecular Subtype	Tumour size pre-chemotherapy		Therapy administered	Tumour size post-chemotherapy		Tumour size (pathology)	Response (% increase/reduction)	Comments
				MRI	US		MRI	US			
#29	Ductal	Grade 3 (poor)	ER+ PR- HER2- Luminal	max 47mm + another of 8mm. + suspicious lesion in other breast	multifocal. 43mm	EC x 4, docetaxel X 1 at full dose. Docetaxel ~ by 20% (severe reaction) for 2,3&4	good response' 21mm	-	5mm residual cancer	89.4% reduction Responder	
#30	Ductal	Grade 2 (mod)	ER+ PR+ (core) HER2+ (PR in resection was negative)	43mm multifocal	40mm possibly another focus	EC x 4 Docetaxel x 4	max diameter of 30mm	-	15mm ER+ PR-	65.1% reduction Responder	PR status changed
#31	Ductal	Grade 3 (poor)	ER+ PR+ HER2+	30mm	25mm	ECx4. Dtx2. 3rd cycle reduced by 20% (toxicities). no cycle 4	20mm	-	no residual cancer 0mm. Complete regression	100% reduction responder	
#32	Lobular	Grade 2 (mod)	ER+ PR+ HER2-	60mm max diameter	-	EC x 6 (had bone metastasis at start - palliative chemo)	no significant alteration' at least 54mm	-	47mm minimum.	21.7% reduction Non-responder	Palliative chemotherapy (bone metastasis at start)
#33	Lobular	Grade 2 (mod)	ER+ PR+ HER2-	max 56mm fragmented	diffuse area max 34mm	EC x 4 Docetaxel x 4	20mm	-	partial regression & response to chemo'. 3 ops; 1 (wide local) 26mm ; 2 (surgery to remove more tumour) 22mm; 3 (mastectomy). total of 48mm	14.3% reduction non-responder	
#34	Lobular	Grade 2 (mod)	ER+ PR+ HER2-	2 lesions. T1 23mm & T2 24mm. 47mm total	2 lesions; 23mm and 16mm. Area of 50mm total	EC x 4 Docetaxel x 4	-	-	max 32 mm	31.9% reduction Responder	
#35	Ductal	Grade 3 (poor)	ER- PR- HER2- E-cadherin +	extensive disease' 47mm	'diffuse'	EC x 5 (#6 abandoned - no response)	-	After 2/3 cycles 40mm. After 4th 'no change'	mastectomy. Large foci 24mm. Extensive overall 40mm malignancy	14.9% reduction non-responder	
#36	Ductal	Grade 3 (poor)	ER- PR- HER2- CK5/6/14+	max 82mm	2 separate lesions. Largest 42mm	ECx4. Docetaxel x 1 (tumour progressed and developed new tumour in other breast)	overall volume unchanged. 'not responded'	'no change'	(bilateral mastectomy) RHS: 2 lesions. Largest 47mm overall area 115mm	40.2% increase non-responder	
#37	Ductal	Grade 1 (well)	ER+ PR+ HER2-	40mm	6mm	EC x 4 Docetaxel x 1 then Docetaxel 2-4 at 20% reduced dose	good response' 2 lesions of 10mm each, but overall 30mm	-	15mm	62.5% reduction responder	
#38	Ductal	Grade 2 (mod)	ER+ PR+ HER2-	93mm	43mm	EC x 4 Docetaxel x 4	fragmented disease over max area of 87mm	-	60mm	35.5% reduction responder	


APPENDIX 6: Annotated Spectrum (Protein Scape)



Annotated fragment ion mass spectrum

This is an example of an annotated fragment ion mass spectrum, generated by the ProteinScape interface, showing the calculated amino acid residues (top). This is the single peptide matched to spot number 35212.

APPENDIX 7: Peptide View (MASCOT)



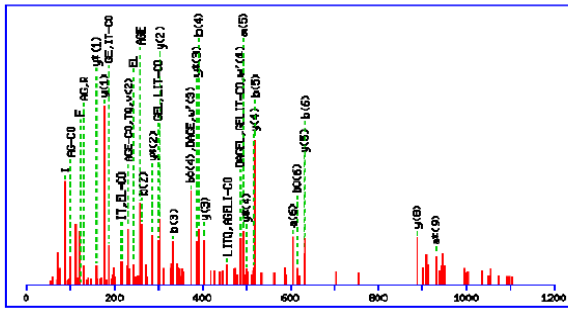
Mascot Search Results

Peptide View

MS/MS Fragmentation of **FDAGELITQR**
 Found in **IP100017334**, Tax_Id=9606 Gene_Symbol=PHB Prohibitin

Match to Query 4: 1148.592089 from(1149.599365,1+) index(2)
 Title: 1149.5993 LIFT/ISRref (13229323922229849)
 Local Instrument: MALDI-TOF-TOF
 Data file 13229323905412175.mgf

Click mouse within plot area to zoom in by factor of two about that point
 Or, to Da
 Label all possible matches Label matches used for scoring



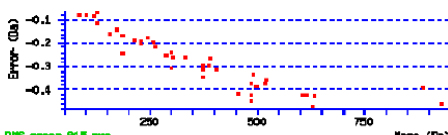
Bold italic red means the series contributed to the score. **Red** means that the number of matches in the ion series is greater than would be expected by chance, indicating that the ion series is present. **Non-bold red** means that the number of matches in the ion series is no greater than would be expected by chance, so that the matches themselves may be by chance

(taken from Matrix Science at http://www.matrixscience.com/help/results_help.html#PEP)

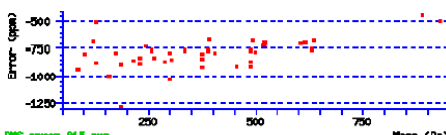
Monoisotopic mass of neutral peptide Mr (calc): 1148.5826
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 36 Expect: 0.039
 Matches : 47/140 fragment ions using 54 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	Seq.	v	w	w'	y	y*	y ⁰	#
1	120.0808	120.0808			148.0757			F							10
2	88.0393	235.1077		217.0972	263.1026		245.0921	D	942.5003	941.5051		1002.5215	985.4949	984.5109	9
3	44.0495	306.1448		288.1343	334.1397		316.1292	A	871.4632			887.4945	870.4680	869.4839	8
4	30.0338	363.1663		345.1557	391.1612		373.1506	G				816.4574	799.4308	798.4468	7
5	102.0550	492.2089		474.1983	520.2038		502.1932	E	685.3992	684.4039		759.4359	742.4094	741.4254	6
6	86.0964	605.2930		587.2824	633.2879		615.2773	L	572.3151	571.3198		630.3933	613.3668	612.3828	5
7	86.0964	718.3770		700.3665	746.3719		728.3614	I	459.2310	472.2514	486.2671	517.3093	500.2827	499.2987	4
8	74.0600	819.4247		801.4141	847.4196		829.4090	T	358.1833	371.2037	373.1830	404.2252	387.1987	386.2146	3
9	101.0709	947.4833	930.4567	929.4727	975.4782	958.4516	957.4676	Q	230.1248	229.1295		303.1775	286.1510		2
10	129.1135							R	74.0237	73.0284		175.1190	158.0924		1

Seq	ya	yb	Seq	ya	yb	Seq	ya	yb
DA	159.0764	187.0713	DAG	216.0979	244.0928	DAGE	345.1405	373.1354
DAGEL	458.2245	486.2195	DAGELI	571.3086	599.3035	DAGELIT	672.3563	700.3512
AG	101.0709	129.0659	AGE	230.1135	258.1084	AGEL	343.1976	371.1925
AGELI	456.2817	484.2766	AGELII	557.3293	585.3243	AGELITQ	685.3879	713.3828
GE	159.0764	187.0713	GEL	272.1605	300.1554	GELI	385.2445	413.2395
GELIT	486.2922	514.2871	GELITQ	614.3508	642.3457	EL	215.1390	243.1339
ELI	328.2231	356.2180	ELII	429.2708	457.2657	ELIIQ	557.3293	585.3243
LI	199.1805	227.1754	LIT	300.2282	328.2231	LITQ	428.2867	456.2817
IT	187.1441	215.1390	ITQ	315.2027	343.1976	TQ	202.1186	230.1135



RHS error 0.5 ppm



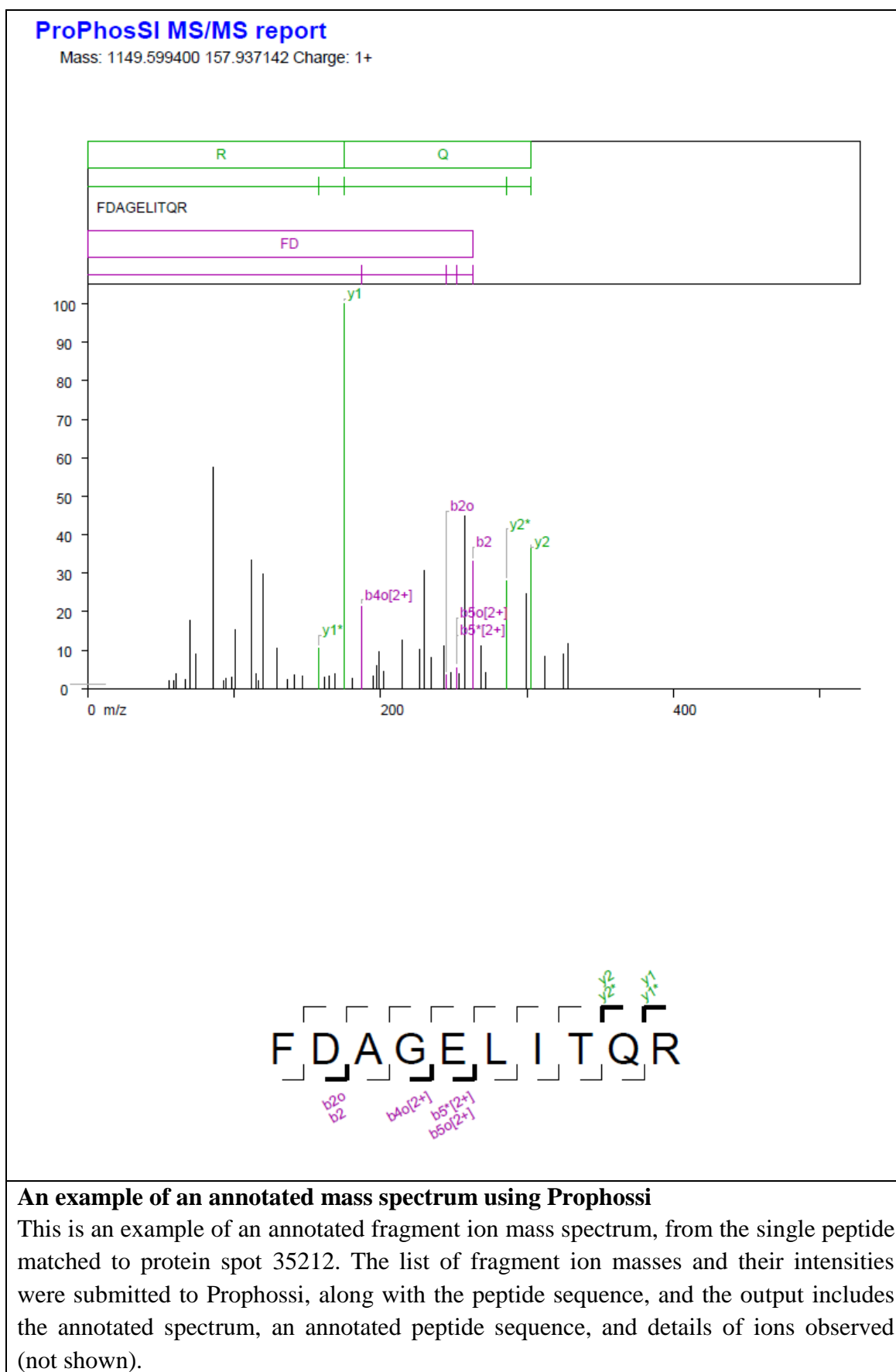
RHS error 0.5 ppm

NCBI BLAST search of **FDAGELITQR**

An example of a peptide view result page from Mascot

This is an example of a peptide view, of the single peptide matched to spot 35212. It includes an annotated mass spectrum and a table containing the matched fragment ions, where italic bold red indicates those that have been used for scoring. The most important ions for the experiments in this project include the 'b' and 'y' ions (highlighted), which show a good selection of italic bold red ions, used to elucidate the peptide sequence.

APPENDIX 8: ProPhossi MS/MS Report



APPENDIX 9: DEPs from 2D-PAGE/MS analysis (identified in 1/3 experiments)

Differentially expressed proteins associated with chemotherapy resistance, identified using 2D-PAGE MALDI-TOF/TOF MS

Three comparative 2D-PAGE MALDI-TOF/TOF MS experiments were performed to identify differentially expressed proteins (DEPs) associated with chemotherapy resistance. The table lists (alphabetically by gene symbol, from the IPI database) those DEPs identified in 1/3 experiments, showing ≥ 2 -fold change in expression, along with the direction of change (\downarrow / \uparrow). Protein identifications with 1 peptide match are indicated (\uparrow^1). Where a protein is not identified as a DEP, --- is shown, to represent status unknown. Proteins listed within the TOP15 (human) RIDEF list are shown* (n=1).

Protein	Gene Symbol	#15 vs #19	#15 vs #1	#18 vs #1
Alpha-1B-glycoprotein	A1BG	\downarrow	---	---
Isoform 1 of Low molecular weight phosphotyrosine protein phosphatase	ACP1	---	---	\uparrow
Putative uncharacterized protein ACTB	ACTB	---	---	\uparrow
highly similar to Actin, cytoplasmic 1	ACTG1 cDNA FLJ52842	---	---	\uparrow
Protein AMBP	AMBP	---	---	\downarrow
Annexin IV	ANXA4	---	---	\uparrow^1
Isoform 2 of Actin-related protein 2/3 complex subunit 5	ARPC5	---	\uparrow	---
ATPase ASNA1	ASNA1	---	\uparrow	---
Isoform 1 of ATP synthase subunit d, mitochondrial	ATP5H	---	\uparrow	---
Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP	---	\uparrow	---
protein C6orf115	C6orf115 UPF0727	---	---	\uparrow^1
Calmodulin	CALM1; CALM2; CALM3	---	\uparrow	---
F-actin-capping protein subunit alpha-1	CAPZA1	---	\uparrow	---
F-actin-capping protein subunit alpha-2	CAPZA2	---	---	\uparrow
Chromobox protein homolog 1	CBX1	---	\uparrow^1	---
Charged multivesicular body protein 4b	CHMP4B	---	---	\uparrow^1
Chloride intracellular channel protein 4	CLIC4	---	---	\uparrow
Isoform Soluble of Catechol O-methyltransferase	COMT	\uparrow	---	---
Highly similar to CATHEPSIN B (clone TESOP2000400)	CTSB cDNA FLJ40065 fis	---	---	\uparrow^1
N(G),N(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1	---	---	\uparrow
Dermatopontin	DPT	\downarrow	---	---
TUT1 Elongation factor 1-gamma	EEF1G	---	\uparrow	---
Eukaryotic translation initiation factor 1A, Y-chromosomal	EIF1AY	---	---	\uparrow^1
Eukaryotic translation initiation factor 6	EIF6	---	---	\uparrow
FK506-binding protein 4	FKBP4	\uparrow	---	---
Ferritin heavy chain	FTH1	---	---	\downarrow
Vitamin D-binding protein	GC	\uparrow^1/\downarrow^1	---	---
Rab GDP dissociation inhibitor beta	GDI2	---	\uparrow	---
PDZ domain-containing protein GIPC1	GIPC1	---	---	\uparrow
Lactoylglutathione lyase	GLO1	---	---	\uparrow

Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	---	↑	---
Heat shock protein beta-1*	HSPB1	↑	---	---
60 kDa heat shock protein, mitochondrial	HSPD1	↑	---	---
Isoform 1 of Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	IDH3A	---	↑	---
Ig kappa chain C region	IGKC	↓ ¹	---	---
Ig lambda-2 chain c regions	IGLC2	---	---	↓ ¹
Inosine triphosphate pyrophosphatase	ITPA	---	↑	---
Keratin, type I cytoskeletal 17	KRT17	↓	---	---
keratin 7	KRT7	---	↑	---
Isoform 2 of LIM and SH3 domain protein 1	LASP1	---	---	↑ ¹
Isoform A of Lamin-A/C	LMNA	---	↑ ¹	---
similar to complement component 3	LOC653879	---	---	↓
Myotrophin	MTPN	---	↑ ¹	---
Myosin regulatory light chain 12A	MYL12A	---	---	↑
Isoform Non-muscle of Myosin light polypeptide 6	MYL6B	---	---	↑ ¹
Putative uncharacterized protein NAP1L4	NAP1L4	---	↑ ¹	---
Alpha-soluble NSF attachment protein	NAPA	---	↓	---
Nicotinamide N-methyltransferase	NNMT	---	↑ ¹	---
ADP-sugar pyrophosphatase	NUDT5	---	↑	---
Nuclear transport factor 2	NUTF2	---	↑ ¹	---
Proliferation-associated protein 2G4	PA2G4	---	↑	---
Protein DJ-1	PARK7	---	---	↑
Proliferating cell nuclear antigen	PCNA	---	---	↑ ¹
highly similar to Glucosidase 2 subunit beta	PRKCSH cDNA FLJ59211	↑ ¹	---	---
Isoform 2 of Proteasome subunit alpha type-3	PSMA3	---	↑ ¹	---
Proteasome subunit alpha type-5	PSMA5	---	↑	---
Proteasome subunit beta type-4	PSMB4	---	↑	---
UV excision repair protein RAD23 homolog B	RAD23B	---	↑	---
Isoform 1 of RuvB-like 1	RUVBL1	---	↓ ¹	---
Protein S100-A11	S100A11	---	↑	---
Protein S100-A13	S100A13	↑ ¹	---	---
Protein S100-A16	S100A16	↑ ¹	---	---
Serum amyloid A protein	SAA2	---	---	↓
Isoform 1 of Septin-2	SEPT2	---	---	↑
Isoform 1 of 14-3-3 protein sigma	SFN	↑ ¹	---	---
Isoform 1 of Calcium-binding mitochondrial carrier protein SCaMC-1	SLC25A24	↑	---	---
Small nuclear ribonucleoprotein F	SNRPF	---	---	↑
TDP43	TARDBP	---	↑ ¹	---
Tubulin-folding cofactor B	TBCB	---	↑	---
Transcription elongation factor B polypeptide 1	TCEB1	---	↑	---
Transcription elongation factor B polypeptide 2	TCEB2	---	---	↑ ¹
tumor protein D52 isoform 2	TPD52	---	---	↑ ¹
Tropomyosin 2	TPM2	↓ ¹	---	---
GDP-L-fucose synthetase	TSTA3	↑ ¹	---	---
Thioredoxin	TXN	---	↑ ¹	---

APPENDIX 10: 2D-PAGE MALDI TOF/TOF MS Raw Data

All protein identification data obtained from 2D-PAGE/MALDI-TOF/TOF analysis

Protein identification data obtained from three 2D-PAGE MALDI-TOF/TOF experiments, where spectra were submitted to MASCOT and searched against the IPI human database. Accession numbers are also shown for the UniprotKB/Swiss-Prot database. Spot numbers with the prefix '1' were identified from experiment 1; prefix '2', experiment 2 and prefix '3', experiment 3.

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
13702	Alpha-1B-glycoprotein	A1BG	76600	54809	5.58	IPI00022895	P04217	182	3	K.HQFLLTGDTQGR.Y R.CEGPIPDVTFELLR.E K.VLILCVAPLSGVDFQLR.R	46 34 58	0.0041 0.049 0.00019	64.0 54.9 46.6	15	D
38107	Isoform 1 of Low molecular weight phosphotyrosine protein phosphatase	ACP1	16900	18487	6.30	IPI00219861	P24666	259	3	R.SPIAEAVFR.K K.LVTDQNIENWR.V R.VDSAATSGYEIGNPPDYR.G	46 75 104	0.0049 4.1e-06 4e-09	54.5 51.1 27.1	31	I
32011	Putative uncharacterized protein ACTB	ACTB	13400	13907	5.24	IPI00942659	n/a	91	2	K.AGFAGDDAPR.A K.IWHHTFYNELR.V	56 34	0.00045 0.052	1.27 -2.50	16	I
35209	highly similar to Actin, cytoplasmic 1	ACTG1 cDNA FLJ52842	26800	39543	5.40	IPI00894365	n/a	140	2	K.QEYDESGPSIVHR.K K.SYELPDGQVITIGNER.F	52 68	0.00091 1.8e-05	34.4 24.2	11	I
15303	Activator of 90 kDa heat shock protein ATPase homolog 1	AHSA1	41200	38421	5.41	IPI00030706	Q95433	79	1	R.ADATNVNWHWTER.D	37	0.022	43.1	11	I
25306	Activator of 90 kDa heat shock protein ATPase homolog 1	AHSA1	41000	38421	5.41	IPI00030706	Q95433	169	3	R.GIPAPEER.T K.ETFLTSPEELYR.V R.ADATNVNWHWTER.D	38 61 40	0.024 0.00013 0.013	-1.28 8.44 1.23	12	I
33213	Protein AMBP	AMBP	28600	39886	5.95	IPI00022426	P02760	287	4	R.HHGPTITAK.L R.ETLLQDFR.V R.GECVPGEQPEPELIPR.V R.VVAQQGVGPEDSIFTMADR.G + Oxidation (M)	50 55 111 72	0.002 0.00062 9.2e-10 7.5e-06	1.75 10.0 -24.24 -27.91	15	D
26203	Annexin A3	ANXA3	32500	36524	5.63	IPI00024095	P12429	687	8	K.ALLTLADGR.R K.LTFDEYR.N R.NTPAFLAER.L R.SEIDLLDIR.T K.GIGTDEFTLNR.I K.SLGDDISSETSGDFR.K K.GAGTNEDALIEILTTR.T R.WGTDEDKFTTEILCLR.S	42 62 68 57 93 133 134 100	0.012 0.00012 2.9e-05 0.00043 8e-08 7.1e-12 5.3e-12 1e-08	-37.40 -7.83 2.78 1.30 14.8 6.39 3.90 -5.24	28	I
36204	Annexin A3	ANXA3	30600	36524	5.63	IPI00024095	P12429	400	4	K.GIGTDEFTLNR.I K.SLGDDISSETSGDFR.K K.GAGTNEDALIEILTTR.T R.WGTDEDKFTTEILCLR.S	52 61 73 100	0.00096 0.00011 6.3e-06 1.2e-08	22.7 11.3 14.3 14.0	28	I
36213	Annexin IV	ANXA4	30300	36290	5.84	IPI00793199	P09525	132	1	K.GLGTDEDAIISVLAYR.N	54	0.00044	34.5	20	I
15205	Serum amyloid P-component	APCS	27300	25485	6.10	IPI00022391	P02743	200	3	R.VGEYSLYIGR.H K.IVLGQEQDSYGGK.F R.AYSLFSYNTQGR.D	58 61 50	0.00028 0.00013 0.0015	61.3 61.6 49.3	23	D
25106	Serum amyloid P-component	APCS	26600	25485	6.10	IPI00022391	P02743	161	2	K.IVLGQEQDSYGGK.F R.AYSLFSYNTQGR.D	51 43	0.001 0.0069	-0.86 -3.71	23	D
35104	Serum amyloid P-component	APCS	24900	25485	6.10	IPI00022391	P02743	85	1	R.AYSLFSYNTQGR.D	42	0.0088	8.79	18	D

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
13105	Apolipoprotein A1	APOA1	25200	28005	5.80	IPI00853525	n/a	352	4	K AKPALEDLR.Q R.THLAPYSDEL.R.Q K.VSFL.SALEEYTK.K R.EQLGPTVQEFWDNLEK.E	71 82 69 75	1.2e-05 9.2e-07 2e-05 3.8e-06	60.8 88.5 96.2 63.5	27	D
13107	Apolipoprotein A1	APOA1	25200	28005	5.80	IPI00853525	n/a	42	1	R.THLAPYSDEL.R.Q	42	0.0084	24.6	4	D
15105	Apolipoprotein A1	APOA1	24300	28005	5.80	IPI00853525	n/a	36	1	R.THLAPYSDEL.R.Q	36	0.036	33.5	4	D
22009	Apolipoprotein A-I	APOA1	20100	30759	5.56	IPI00021841	P02647	98	1	R.DYVSQFEFGSALGK.Q	66	3.8e-05	38.3	11	D
24103	Apolipoprotein A1	APOA1	24900	28005	5.80	IPI00853525	n/a	140	2	K AKPALEDLR.Q R.THLAPYSDEL.R.Q	54 86	0.0006 3.5e-07	-5.65 8.65	8	D
25109	Apolipoprotein A1	APOA1	24000	28005	5.80	IPI00853525	n/a	85	1	R.THLAPYSDEL.R.Q	52	0.00089	-1.11	8	D
25112	Apolipoprotein A1	APOA1	25600	28005	5.80	IPI00853525	n/a	123	2	K AKPALEDLR.Q R.THLAPYSDEL.R.Q	44 79	0.0062 1.9e-06	5.02 21.0	8	D
25116	Apolipoprotein A1	APOA1	25100	28005	5.80	IPI00853525	n/a	141	2	K AKPALEDLR.Q R.THLAPYSDEL.R.Q	43 70	0.0088 1.5e-05	-0.22 7.71	12	D
33109	Apolipoprotein A1	APOA1	23500	28005	5.80	IPI00853525	n/a	103	1	R.THLAPYSDEL.R.Q	69	2.00E-05	17.9	8	D
35111	Apolipoprotein A-I	APOA1	23200	30759	5.80	IPI00021841	P02647	83	1	R.THLAPYSDEL.R.Q	83	7.9e-07	-4.68	4	D
35113	Apolipoprotein A1	APOA1	23700	28005	5.80	IPI00853525	n/a	63	1	R.THLAPYSDEL.R.Q	42	0.0096	8.84	8	D
15102	Adenine phosphoribosyltransferase	APRT	20500	19766	5.78	IPI00218693	P07741	83	1	R.SFPDFPTPGVVFR.D	51	0.001	43.9	12	I
35102	Adenine phosphoribosyltransferase	APRT	19300	19766	5.78	IPI00218693	P07741	69	1	R.SFPDFPTPGVVFR.D	69	1.8e-05	39.9	7	I
23101	Rho GDP-dissociation inhibitor 1	ARHGDI A	25500	23250	5.02	IPI00003815	P52565	513	5	K.YIQHTYR.K K.IDKTDYIMVGSYGPR.A + Oxidation (M) R.AEEYEF.LTPVEEAPK.G K.SIQEIQLDKDDESLR.K R.FTDDDKTDHLSWEWNLTIK.K	54 69 99 141 116	0.00063 1.7e-05 1.6e-08 8.3e-13 2.1e-10	10.6 26.4 24.3 21.5 2.58	34	I
33101	Rho GDP-dissociation inhibitor 1	ARHGDI A	23700	23250	5.02	IPI00003815	P52565	147	2	R.AEEYEF.LTPVEEAPK.G K.SIQEIQLDKDDESLR.K	40 49	0.012 0.0014	-7.52 -7.71	24	I
13104	Rho GDP-dissociation inhibitor 2	ARHGDI B	24500	23031	5.10	IPI00003817	P52566	112	2	K.YVQHTYR.T K.APEPHVEEDDDDELDSK.L	45 67	0.005 2.3e-05	0.95 -12.39	11	I
23104	Rho GDP-dissociation inhibitor 2	ARHGDI B	24700	23031	5.10	IPI00003817	P52566	100	2	K.APNVVVTR.L K.YVQHTYR.T	56 44	0.0004 0.0065	-2.55 11.3	7	I
33104	Rho GDP-dissociation inhibitor 2	ARHGDI B	23000	23031	5.10	IPI00003817	P52566	60	1	K.APNVVVTR.L	35	0.049	21.2	7	I
26003	Isoform 2 of Actin-related protein 2/3 complex subunit 5	ARPC5	16800	16677	5.71	IPI00007280	Q15511	272	3	K.ALAAGGVGSIVR.V R.KVDVDEYDENK.F K.FVDEEDGGDQAGPDEGEVDSCLR.H	58 67 146	0.00024 2.6e-05 2.3e-13	-7.94 8.01 -38.3	30	I
22307	ATPase ASNA1	ASNA1	40200	39224	4.81	IPI00013466	Q43681	138	2	K.LPLLPHEVR.G R.LLNFPIIVR.G	46 65	0.0055 5.4e-05	1.51 6.85	8	I
12507	ATP synthase subunit beta, mitochondrial	ATP5B	50400	56525	5.26	IPI00303476	P06576	764	9	K.IPVGPETLGR.I R.IMNVIGEPIDER.G + Oxidation (M) K.AHGGYSVFAGVGER.T R.FTQAGSEVSALLGR.I R.VALTGLTVAEYFR.D K.VALVYQQMNEPPGAR.A + Oxidation (M) R.LVLEVAQHLGESTVR.T R.DQEGQDVLLFIDNIFR.F R.AIALGIVPAVDPLDSTSR.I	79 56 81 104 75 52 101 133 84	1.8e-06 0.00033 1.2e-06 6.2e-09 3.7e-06 0.00086 1.1e-08 5.7e-12 3.5e-07	30.0 33.0 23.0 39.3 26.1 39.4 38.4 31.9 37.9	24	I
33401	ATP synthase subunit beta, mitochondrial	ATP5B	48100	56525	5.26	IPI00303476	P06576	158	1	R.LVLEVAQHLGESTVR.T	44	0.0045	20.3	24	I
24106	Isoform 1 of ATP synthase subunit d, mitochondrial	ATP5H	21000	18537	5.21	IPI00220487	Q75947	212	3	K.SWNETLTSR.L R.IVYEKEMEK.M + Oxidation (M) K.YTAQVDAEEKEDVK.S	72 49 91	9.6e-06 0.0019 1e-07	-1.81 4.07 -1.82	20	I
25001	Barrier-to-autointegration factor	BANF1	10500	10280	5.81	IPI00026087	Q75531	194	3	K.KDEDLFR.E K.AYVVLGQFLVLK.K R.DFVAEPMGEKPVGSLAGIVEVLGK.K + Oxidation (M)	54 61 57	0.00064 0.00015 0.00018	-15.40 7.57 -29.48	56	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/ Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
35001	Barrier-to-autointegration factor	BANF1	10000	10280	5.81	IPI00026087	Q75531	36	1	KKDEDLFR E	36	0.043	28.5	7	I
20104	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP	30400	31742	4.74	IPI00014230	Q07021	535	5	K AFVDFLSDEIK E R.EVFSQSTGESEWK.D K.MSGGWLELNGTEAK.L + Oxidation (M) K.AFVDFLSDEIKEER K K.ALVLDCHYPEDEVGQEAEESDIFSIR E	80 72 94 115 173	1.4e-06 8.6e-06 4.4e-08 3.7e-10 2.8e-16	26.9 24.5 18.6 21.7 1.21	24	I
36007	Protein C6orf115	C6orf115 UPF0727	11500	9108	5.86	IPI00855846	Q9P1F3	73	1	K.CANLFEALVGTLK.A	73	7.8e-06	4.17	16	I
20021	Calmodulin	CALM1; CALM2; CALM3	17300	16827	4.09	IPI00075248	P62158	251	3	K.DGNGYTSAAELR.H K.MKDTDSEEEIR.E + Oxidation (M) K.EAFSLFDKDGDTITTK.E	75 103 73	4.5e-06 7.3e-09 5.7e-06	23.6 35.3 5.74	26	I
27303	Macrophage-capping protein	CAPG	41400	38779	5.88	IPI00027341	P40121	126	2	K.YQEGGVESAFHK.T R.QAALQVAEGFISRM	52 68	0.00084 2.7e-05	-8.05 -7.07	11	I
27305	Macrophage-capping protein	CAPG	41200	38779	5.88	IPI00027341	P40121	272	4	K.YQEGGVESAFHK.T R.QAALQVAEGFISRM R.MQYAPNTQVEILPQGR.E + Oxidation (M) R.EVQGNESDLFMSYFPR.G + Oxidation (M)	66 77 72 57	3.4e-05 3.1e-06 7.3e-06 0.00022	1.53 3.65 -12.86 -18.88	16	I
37304	Macrophage-capping protein	CAPG	37700	38779	5.88	IPI00027341	P40121	215	3	K.YQEGGVESAFHK.T R.QAALQVAEGFISRM R.EVQGNESDLFMSYFPR.G + Oxidation (M)	63 102 51	7.8e-05 1e-08 0.0009	42.1 44.3 12.6	11	I
25204	F-actin-capping protein subunit alpha-1	CAPZA1	34300	33073	5.45	IPI00005969	P52907	681	7	R.LLLNNDNLLR.E K.FITTPPTAQVVGVLK.I K.IQVHYVEDGNVQLVSHK.D K.FITHAPPGEFNEVFNDVRL R.EGAAHAFQYNMDQFTPVK.I + Oxidation (M) K.IEGYEDQVLITEHGDLGNSR.F K.TIDGQQTILAGIESHQFQPK.N	79 62 139 118 67 155 62	2e-06 9.6e-05 1.4e-12 1.3e-10 2.1e-05 3e-14 4.7e-05	8.86 6.19 -17.38 -12.09 -16.69 -16.26 -17.71	41	I
36205	F-actin-capping protein subunit alpha-2	CAPZA2	32300	33157	5.57	IPI00026182	P47755	411	6	R.LLLNNDNLLR.E K.IQVHYVEDGNVQLVSHK.D K.FITHAPPGEFNEVFNDVRL R.EGAAHAFQYNLDQFTPVK.I K.EATDPRPCVENAVESWR.T K.IEGYEDQVLITEHGDLGNGKF	44 104 83 51 58 72	0.0061 3.9e-09 4.7e-07 0.00077 0.00015 5.4e-06	28.4 -2.63 -3.89 -5.64 -1.47 -0.56	35	I
16204	Isoform 1 of F-actin-capping protein subunit beta	CAPZB	30700	31616	5.36	IPI00026185	P47756	214	3	R.RLPPQIEK.N K.LEVEANNAFDQYR.D R.KLEVEANNAFDQYR.D	36 57 42	0.038 0.00029 0.0072	73.1 75.4 71.6	25	I
26106	Isoform 1 of F-actin-capping protein subunit beta	CAPZB	30300	31616	5.36	IPI00026185	P47756	457	7	R.RLPPQIEK.N R.STLNEIYFGK.T K.DYLLCDYNRD K.YDPPLEDGAMP SAR.L + Oxidation (M) K.LEVEANNAFDQYR.D K.GCWDSIHVVEVQEK.S R.KLEVEANNAFDQYR.D	47 69 53 37 91 52 111	0.0034 2.4e-05 0.00098 0.027 1.1e-07 0.0008 9.9e-10	-2.60 -0.40 4.83 3.02 1.53 -5.31 -2.68	25	I
36206	Isoform 2 of F-actin-capping protein subunit beta	CAPZB	28400	30952	5.69	IPI00642256	P47756	125	1	R.RLPPQIEK.N	35	0.045	12.7	22	I
22105	Chromobox protein homolog 1	CBX1	24500	21519	4.85	IPI00010320	n/a	44	1	K.CPQVVISFYEER.L	44	0.0046	24.1	6	I
18711	T-complex protein 1 subunit beta	CCT2	59100	57794	6.01	IPI00297779	P78371	69	1	R.QVLLSAAEAEEVILR.V	42	0.0083	43.2	5	D
28517	T-complex protein 1 subunit beta	CCT2	58900	57794	6.01	IPI00297779	P78371	429	6	K.HGNCDFNR.Q R.GATQQILDEAER.S R.QVLLSAAEAEEVILR.V K.LIHFSGVALGEACTIVLR.G R.LALVTGGELASTFDHPELVK.L R.VQDDEVGDGITSVTVLAELLR.E	39 37 82 73 40 102	0.021 0.031 7.6e-07 4.9e-06 0.011 3.1e-09	11.1 9.70 4.02 -14.01 -17.40 -23.28	21	I
32204	Charged multivesicular body protein 4b	CHMP4B	30700	24935	4.76	IPI00025974	Q9H444	85	1	K.QLAQIDGTLSTIEFQR.E	51	0.001	2.07	12	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
15406	Creatine kinase B-type	CKB	45700	42902	5.34	IPI00022977	P12277	131	2	K.DLFDPIEDR.H K.VLTPELYAELR.A	37 45	0.032 0.0054	-3.38 5.33	21	D
25308	Creatine kinase B-type	CKB	45400	42902	5.34	IPI00022977	P12277	105	2	K.DLFDPIEDR.H K.VLTPELYAELR.A	42 38	0.01 0.024	25.6 23.6	8	D
35404	Creatine kinase B-type	CKB	41600	42902	5.34	IPI00022977	P12277	438	5	R.HGGYKPSDEHK.T K.VLTPELYAELR.A K.LAVEALSSLDGDLAGR.Y K.TFLVWVNEEDHLR.V R.GTGGVDTAAVGGVFDVSNADR.L	49 42 97 81 132	0.002 0.009 2.6e-08 1.1e-06 6.5e-12	-7.29 -8.16 -20.72 -26.39 -31.62	24	D
13202	Chloride intracellular channel protein 1	CLIC1	29600	27248	5.09	IPI00010896	Q00299	188	3	R.YLSNAYARE R.GFTIPEAFR.G K.IIEFLEAVLCPPR.Y	46 50 60	0.0048 0.0017 0.00014	41.8 50.3 48.5	16	I
24102	Chloride intracellular channel protein 1	CLIC1	29500	27248	5.09	IPI00010896	Q00299	344	5	R.YLSNAYARE R.GFTIPEAFR.G K.IGNCPFSQR.L K.IIEFLEAVLCPPR.Y R.EEFASTCPDDEIIELAYEQVAK.A	50 50 38 82 114	0.0019 0.0017 0.028 8.2e-07 2.6e-10	6.64 18.1 21.8 33.2 15.6	36	I
24108	Chloride intracellular channel protein 1	CLIC1	29200	27248	5.09	IPI00010896	Q00299	243	4	R.YLSNAYARE R.GFTIPEAFR.G K.IIEFLEAVLCPPR.Y K.VLDNYLTSPLPEEVDSETSAEDEGVSQR.K	45 47 62 65	0.0057 0.0036 8.7e-05 2.3e-05	-11.55 -0.07 10.5 -19.07	27	I
33209	Chloride intracellular channel protein 1	CLIC1	27800	27248	5.09	IPI00010896	Q00299	107	1	K.VLDNYLTSPLPEEVDSETSAEDEGVSQR.K	59	7.8e-05	28.1	27	I
35208	Chloride intracellular channel protein 4	CLIC4	26600	28982	5.45	IPI00001960	Q9Y696	268	4	R.YLTNAYS.RD K.NSRPEANEALER.G K.AGSDGESIGNCPFSQR.L R.KPADLQNLAPGTHPPFITFSEVK.T	41 70 102 54	0.016 1.4e-05 7.1e-09 0.00026	-13.32 -5.27 -21.41 -63.69	23	I
12105	Isoform Soluble of Catechol O-methyltransferase	COMT	24200	24833	5.15	IPI00375513	P21964	190	3	R.YLPDITLLEECGLLR.K R.LITIEINPDCAAITQR.M K.GTVLLADNVICPGAPDFLAHVR.G	99 59 32	1.5e-08 0.00014 0.049	9.93 16.0 9.07	23	I
24007	Coactosin-like protein	COTL1	15100	16049	5.54	IPI00017704	Q14019	54	1	R.AAYNLVLR.D	38	0.033	-23.09	9	I
34004	Coactosin-like protein	COTL1	14000	16049	5.54	IPI00017704	Q14019	129	2	K.FALITWIGENVSGLQR.A K.YDGSTIIVGEQGAEYQHFQQCTDDVLR.L	46 44	0.0032 0.0037	1.71 -48.61	40	I
16102	Cellular retinoic acid-binding protein 2	CRABP2	15000	15854	5.42	IPI00216088	P29373	137	1	K.VGEEFEEQTVDRPCK.S	93	5.8e-08	29.9	33	I
25006	Cellular retinoic acid-binding protein 2	CRABP2	14200	15854	5.42	IPI00216088	P29373	344	4	K.QEGDTFYIK.T R.SENFEELK.V K.VGEEFEEQTVDRPCK.S R.ELTNDGELLITMTADDVVCTR.V + Oxidation (M)	43 44 103 90	0.0086 0.0068 5.1e-09 7.6e-08	7.34 6.68 -3.22 -22.61	52	I
36001	Cellular retinoic acid-binding protein 2	CRABP2	13300	15854	5.42	IPI00216088	P29373	344	4	K.QEGDTFYIK.T R.SENFEELK.V K.VGEEFEEQTVDRPCK.S R.ELTNDGELLITMTADDVVCTR.V + Oxidation (M)	51 46 89 106	0.0014 0.0041 1.5e-07 1.9e-09	19.9 21.3 7.24 -15.02	52	I
33105	highly similar to CATHEPSIN B (clone TESOP2000400)	CTSB cDNA FLJ40065 fis	22000	30781	5.89	IPI00941296	n/a	83	1	R.GQDHCIESEVVAGIPR.T	83	5.6e-07	27.2	6	I
35210	N(G).N(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1	33000	31444	5.53	IPI00220342	Q94760	91	2	K.GHVLH.R.T R.ALPESLGQHALR.S	41 50	0.02 0.0015	-34.80 -12.78	6	I
10201	Dematopontin	DPT	20500	24559	4.70	IPI00292130	Q07507	330	6	R.YFESVLDR.E R.GATTTFSAVER.D R.EWQFYCCR.Y R.QGFSYQCQGGQVIVAVR.S R.AGMEWYQTCNNGLVAGFQSR.Y R.AGMEWYQTCNNGLVAGFQSR.Y + Oxidation (M)	55 52 42 107 75 40	0.00061 0.00087 0.011 1.9e-09 2.6e-06 0.007	48.6 43.5 49.5 19.3 0.51 1.10	32	D
28412	TUT1 Elongation factor 1-gamma	EEF1G	51100	50429	6.25	IPI00937615	P26641	105	2	K.ALIAAQYSGAQVR.V R.VLSAPPHHFHQITNR.T	53 37	0.00072 0.029	18.5 16.4	8	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
34005	Eukaryotic translation initiation factor 1A, Y-chromosomal	EIF1AY	17500	16546	5.07	IPI00023004	Q14602	51	1	K.VVWINTSDIILVGLR.D	51	0.001	-30.36	9	I
23005	Isoform 1 of Eukaryotic translation initiation factor 5A-1	EIF5A	16800	17049	5.08	IPI00411704	P63241	277	2	K.VHLVGDIFTGK.K R.NDFQLIGQDGYLSLLQDSGEVRE	93 184	7.2e-08 3e-17	26.5 -5.61	22	I
33005	Isoform 1 of Eukaryotic translation initiation factor 5A-1	EIF5A	15900	17049	5.08	IPI00411704	P63241	93	2	K.VHLVGDIFTGK.K R.NDFQLIGQDGYLSLLQDSGEVRE	54 39	0.00063 0.0086	37.1 5.08	22	I
31103	Eukaryotic translation initiation factor 6	EIF6	24500	27095	4.56	IPI00010105	P56537	267	4	R.NSLPDTVQIR.R K.LNEAQPSTIATSMR.D + Oxidation (M) R.ASFENNCEIGCFK.L K.TSIEDQDESSLQVPLVAGTVNR.G	63 34 68 89	8.4e-05 0.051 2.2e-05 8.6e-08	-15.46 -12.75 -11.88 -43.86	34	I
15602	FK506-binding protein 4	FKBP4	59400	52057	5.35	IPI00219005	Q02790	259	4	K.GEDLTFEEDGGIIR.R R.GEAHLAVNDFELAR.A R.RGEAHLAVNDFELAR.A R.FEIGEGENLDLPYGLER.A	50 65 48 89	0.0012 4.4e-05 0.002 1.4e-07	89.3 82.2 71.9 60.4	11	I
34109	Ferritin heavy chain	FTH1	19500	21383	5.30	IPI00554521	P02794	235	3	K.ELGDHVTNLR.K K.YFLHQSHHEER.E R.QNYHQDSEAAINR.Q	69 73 93	2e-05 8.1e-06 6.4e-08	2.39 11.6 7.95	18	D
15103	Ferritin light chain	FTL	19500	20064	5.51	IPI00375676	Q6S4P3	536	7	K.KPAEDEWVK.T K.MGDHLTNLHR.L + Oxidation (M) K.KMGDHLTNLHR.L + Oxidation (M) R.DDVVALEGVSHFFR.E K.LNQALLDLHALGSAR.T R.LGGPEAGLGEYLFER.L K.KLNQALLDLHALGSAR.T	48 76 35 92 98 116 70	0.0028 3.9e-06 0.05 8.9e-08 1.9e-08 3.5e-10 1.1e-05	66.3 79.6 64.0 62.5 57.4 68.8 57.4	47	D
25008	Ferritin light chain	FTL	19000	20064	5.51	IPI00375676	Q6S4P3	36	1	R.LGGPEAGLGEYLFER.L	36	0.037	19	8	D
35103	Ferritin light chain	FTL	18500	20064	5.51	IPI00375676	Q6S4P3	473	8	K.KPAEDEWVK.T K.MGDHLTNLHR.L K.MGDHLTNLHR.L + Oxidation (M) K.KMGDHLTNLHR.L + Oxidation (M) R.DDVVALEGVSHFFR.E K.LNQALLDLHALGSAR.T R.LGGPEAGLGEYLFER.L K.KLNQALLDLHALGSAR.T	36 43 65 41 116 59 86 70	0.049 0.0079 5.1e-05 0.011 3.2e-10 0.00016 3e-07 1.2e-05	18.9 16.0 23.7 15.7 13.4 10.2 14.4 9.34	36	I
14501	Vitamin D-binding protein	GC	56800	54526	5.40	IPI00555812	P02774	95	1	K.HLSLLTTLNLR.V	44	0.0058	0.67	6	D
14502	Vitamin D-binding protein	GC	57000	54526	5.40	IPI00555812	P02774	84	1	K.HLSLLTTLNLR.V	41	0.011	39.7	6	I
28405	Rab GDP dissociation inhibitor beta	GDI2	51300	51087	6.11	IPI00940148	P50395	122	2	K.DLGTESQIFISR.T R.IDDYLDQPCYETINR.I	69 38	2e-05 0.016	24.5 4.21	7	I
28519	Rab GDP dissociation inhibitor beta	GDI2	54200	51087	6.11	IPI00940148	P50395	126	2	K.DLGTESQIFISR.T K.EIRPALELLEPIEQK.F	56 70	0.00039 1.2e-05	53.8 25.1	6	I
38302	PDZ domain-containing protein GIPC1	GIPC1	37500	36141	5.90	IPI00024705	Q14908	140	2	K.APPLVENEAEPPGR.G R.LVFHTQLAHGSPTR.I	76 66	3.4e-06 3.5e-05	11.2 7.06	8	I
32106	Lactoylglutathione lyase	GLO1	20900	20992	5.12	IPI00220766	Q04760	154	2	K.DFLQQTMLR.V + Oxidation (M) R.GFGHIGLAVPDVYSACK.R	38 64	0.026 4.4e-05	-44.53 -66.31	22	I
17410	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	GPD1	37100	38171	5.81	IPI00295777	P21695	215	3	K.LISEVIGER.L R.ELYSILQHK.G K.IVGGNAAQLAQFDPR.V	43 47 61	0.01 0.0035 0.0001	35.3 33.1 23.2	23	D
27201	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	GPD1	35500	38171	5.81	IPI00295777	P21695	109	2	K.IVGGNAAQLAQFDPR.V R.ITVVQEVDVVEICGALK.N	36 73	0.033 4.9e-06	1.12 -13.52	9	D
28102	Glutathione S-transferase omega-1	GSTO1	29400	27833	6.23	IPI00019755	P78417	150	2	K.VPSLVGVSFR.S K.GSAPPQVPEGSIR.I	48 48	0.0031 0.0027	-5.57 -2.84	17	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
38202	Glutathione S-transferase omega-1	GSTO1	27500	27833	6.23	IPI00019755	P78417	375	5	R.FCPFAER.T K.VPSLVGSFIR.S R.HEVININL.K.N K.GSAPPGVPVEGSIR.I K.EDYAGLKEEFR.K	35 74 56 96 86	0.048 7.4e-06 0.00038 4.1e-08 4.1e-07	30.6 40.4 44.9 44.9 43.6	24	I
16103	Glutathione S-transferase P	GSTP1	24000	23569	5.43	IPI00219757	P09211	187	2	M.PPYTVVYFPVR.G K.FQDGDLLTYQSNTILR.H	57 116	0.0003 3.1e-10	36.4 27.1	21	D
26102	Glutathione S-transferase P	GSTP1	23400	23569	5.43	IPI00219757	P09211	317	4	M.PPYTVVYFPVR.G K.EEVVTVETWQEGSLK.A K.FQDGDLLTYQSNTILR.H K.DQQAALVDMVNDGVEDLR.C + Oxidation (M)	63 62 132 54	7.7e-05 8e-05 6.6e-12 0.00044	13.8 -3.79 -6.71 -17.13	38	I
36101	Glutathione S-transferase P	GSTP1	21800	23569	5.43	IPI00219757	P09211	231	4	M.PPYTVVYFPVR.G K.FQDGDLLTYQSNTILR.H K.FQDGDLLTYQSNTILR.H K.ALPGQLKPFETLLSQNGGK.T	68 46 129 34	2.4e-05 0.0029 1.4e-11 0.035	15.4 1.52 1.52 1.49	22	I
20107	HEBP2 protein (fragment)	HEBP2	25900	24014	4.30	IPI00644697	Q9YSZ4	387	4	R.NNEVWLIQK.N K.VVYTAGYNSPVK.L K.NQEQLLTLASILR.E K.APEDAGPQPGSYEIR.H	45 81 107 126	0.0054 1.3e-06 2.5e-09 3.8e-11	-7.21 -5.65 -7.31 -10.31	26	I
30108	HEBP2 protein (Fragment)	HEBP2	24200	24014	5.30	IPI00644697	Q9YSZ4	197	3	R.NNEVWLIQK.N K.NQEQLLTLASILR.E K.APEDAGPQPGSYEIR.H	40 74 83	0.018 5e-06 6.9e-07	22.6 12.4 13.9	17	I
27521	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	54500	49484	5.89	IPI00013881	P31943	352	5	R.VHIEIGPDGR.V R.GLPWSCSADEVQR.F K.HTGNPSPDTANDGFVRL.L R.ATENDIYNFFSPLNPNR.V R.YVELFLNSTAGASGGAYEHR.Y	75 43 104 49 80	5.4e-06 0.0065 4.4e-09 0.0013 9.1e-07	3.65 8.71 -0.91 -10.69 -14.77	16	I
16104	Heat shock protein beta-1	HSPB1	25900	22826	5.98	IPI00025512	P04792	144	2	R.QDEHGYSR.C R.LFDQAFGLPRL.L	43 52	0.0094 0.0011	76.6 80.5	24	I
16107	Heat shock protein beta-1	HSPB1	25800	22826	5.98	IPI00025512	P04792	193	3	R.QDEHGYSR.C R.LFDQAFGLPRL.L K.LATQSNITIPVTFESR.A	50 50 53	0.0019 0.0017 0.00056	43.4 34.7 34.8	36	I
17105	highly similar to Heat-shock protein beta-6	HSPB6 cDNA FLJ51906	18500	21276	5.88	IPI00908768	n/a	104	2	K.VVGEHVEVHAR.H R.HEERPDEHGFVAR.E	58 45	0.00025 0.0036	40.3 38.0	12	D
28010	highly similar to Heat-shock protein beta-6	HSPB6 cDNA FLJ51906	18200	21276	5.88	IPI00908768	n/a	130	2	K.VVGEHVEVHAR.H R.HEERPDEHGFVAR.E	65 65	0.000054 4.3e-05	7.22 5.47	12	D
14604	60 kDa heat shock protein, mitochondrial	HSPD1	61400	61187	5.70	IPI00784154	P10809	397	5	K.APGFGDNR.K K.GANPVEIR.R R.AAVEEGVLGGGCALLR.C K.ISSISQIVPALAIANHR.K R.IQEIEQLDVTSEYEK.E	43 37 101 126 42	0.0091 0.026 9e-09 2.6e-11 0.0052	4.29 11.1 42.0 36.0 23.7	25	I
26212	Isoform 1 of Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	IDH3A	37300	40022	6.47	IPI00030702	P50213	160	2	R.IAEFAFEYAR.N K.IPYTDVNVITR.E	53 74	0.00081 5.1e-06	16.0 14.4	8	I
18216	Ig kappa chain C region	IGKC	26600	25721	6.30	IPI00940069		53	1	K.SGTASVVCLLNFFYPRE	53	0.00054	0.49	6	D
34111	Ig lambda-2 chain c regions	IGLC2	25900	25119	5.93	IPI00154742	P0CG05	67	1	R.SYSQVTHEGSTVEK.T	67	2.1e-05	42	6	D
25120	Inosine triphosphate pyrophosphatase	ITPA	22200	21831	5.50	IPI00018783	Q9BY32	242	3	K.LKPEGLHQLLAGFEDK.S K.IDLPEYQGEPEDEISIQK.C K.SAYALCTFALSTGDPSPQVRL	78 104 60	2e-06 3.8e-09 9e-05	-0.19 -10.94 -14.37	27	I
12407	Keratin, type I cytoskeletal 17	KRT17	46700	48361	4.97	IPI00450768	Q04695	206	4	R.LEQEIATYR.R R.LSVEADINGLR.R R.TKFETEQLRL.L K.NHEEEMNALR.G + Oxidation (M)	37 45 35 41	0.035 0.006 0.051 0.012	35.1 23.5 21.0 28.2	16	D
21316	Keratin, type I cytoskeletal 19	KRT19	38900	44065	5.04	IPI00479145	P08727	180	2	K.NHEEESTLR.G R.DYSHYYITQDLR.D	50 69	0.0014 1.5e-05	15.4 11.1	12	D

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
12308	Keratin, type I cytoskeletal 19	KRT19	42000	44065	5.04	IPI00479145	P08727	396	4	R.TKFETEQLR.M K.NHEEEISTLR.G K.AALEDTLAETEAR.F R.DYSHYYTTIQDLR.D	45 37 56 101	0.0048 0.0026 0.00037 9e-09	86.3 72.8 94.1 78.1	22	I
22311	Keratin, type I cytoskeletal 19	KRT19	42000	44065	5.04	IPI00479145	P08727	547	9	R.QNQEYQR.L R.IVLQIDNAR.L K.DAEAWFTSR.T R.LEQEIATYR.S R.TKFETEQLR.M K.NHEEEISTLR.G K.SRLEQEIATYR.S R.EVAGHTEQLQMSR.S + Oxidation (M) R.DYSHYYTTIQDLR.D	59 46 41 41 46 56 53 68 108	0.00028 0.0037 0.011 0.013 0.0045 0.00033 0.00074 2.1e-05 1.8e-09	16.6 23.7 28.5 29.9 36.1 35.4 40.7 38.4 34.6	22	I
22313	Keratin, type I cytoskeletal 19	KRT19	42100	44065	5.04	IPI00479145	P08727	585	9	R.QNQEYQR.L R.IVLQIDNAR.L K.DAEAWFTSR.T R.LEQEIATYR.S R.TKFETEQLR.M K.NHEEEISTLR.G K.SRLEQEIATYR.S K.AALEDTLAETEAR.F R.DYSHYYTTIQDLR.D	52 59 48 47 50 50 50 95 108	0.0012 0.00019 0.0027 0.0036 0.0018 0.0016 0.0015 4.9e-08 1.8e-09	5.73 13.5 15.7 16.5 21.7 21.1 21.1 27.3 20.9	22	I
32311	Keratin, type I cytoskeletal 19	KRT19	39100	44065	5.04	IPI00479145	P08727	147	1	R.DYSHYYTTIQDLR.D	35	0.035	16.3	18	I
22312	Keratin 7	KRT7	44900	51411	5.40	IPI00847342	n/a	393	5	R.GQLEALQVDGGR.L K.LALDIEIATYR.K R.AKQEELEAALQR.G K.VDALNDEINFLR.T R.LPDIFEAQIAGLR.G	49 50 83 88 106	0.002 0.0014 7.1e-07 2.1e-07 3.3e-09	10.0 23.2 15.0 16.5 15.8	14	I
13408	Keratin, type II cytoskeletal 8	KRT8	46000	53671	5.52	IPI00554648	P05787	312	5	K.LSELEAALQR.A K.LALDIEIATYR.K R.ASLEAAIADAEQR.G R.LEGLTDEINFLR.Q R.ELQSQISDTSVVLSMDNSR.S + Oxidation (M)	40 38 52 64 76	0.014 0.022 0.00095 5.3e-05 2.4e-06	125 119 131 134 87.6	16	I
15504	Keratin, type II cytoskeletal 8	KRT8	50500	53671	5.52	IPI00554648	P05787	230	3	K.LSELEAALQR.A R.ASLEAAIADAEQR.G R.LEGLTDEINFLR.Q	37 42 57	0.037 0.0087 0.00032	24.4 23.5 21.9	16	I
22315	Keratin, type II cytoskeletal 8	KRT8	39400	53671	5.52	IPI00554648	P05787	238	3	K.LALDIEIATYR.K R.ASLEAAIADAEQR.G R.LEGLTDEINFLR.Q	55 75 60	0.0005 4.2e-06 0.00016	51.7 63.9 42.1	11	D
38305	Isoform 2 of LLM and SH3 domain protein 1	LASP1	34600	30097	6.61	IPI0000861	Q14847	53	1	K.GFSVVADTPELQRI	53	0.00076	11.2	4	I
28626	Isoform A of Lamin-A/C	LMNA	72300	74380	6.57	IPI00021405	P02545	116	1	R.NSNLVGAHFEELQQSRI	41	0.0091	4.73	6	I
31309	similar to complement component 3	LOC653879	40300	45642	4.94	IPI00739237	n/a	402	5	R.FYHPEKEDGK.L K.SGSDEVQVGGQR.T K.VSHSEDDCLAFK.V K.VYAYYNLEESCTR.F K.VHQYFNVLIQPGAVK.V	36 70 66 87 89	0.038 1.5e-05 3.1e-05 2.5e-07 1.3e-07	17.3 10.0 9.16 -1.19 -7.42	18	D
18213	Isoform 1 of Acyl-protein thioesterase 1	LYPLA1	24400	24996	6.29	IPI00939508	n/a	49	1	R.ASFPPQGPIGGANR.D	49	0.002	85	5	I
29101	Isoform 1 of Acyl-protein thioesterase 1	LYPLA1	24200	24996	6.29	IPI00939508	n/a	48	1	R.ASFPPQGPIGGANR.D	36	0.041	15.9	11	I
23103	Microtubule-associated protein RP/EB family member 1	MAPRE1	31100	30151	5.02	IPI00017596	Q15691	277	3	K.LEHEYIQNFK.I K.KPLTSSSAAPQPISTQR.T R.NIELICQENEGENDPVLQRI	92 64 121	1e-07 4.5e-05 7.1e-11	15.0 2.73 -10.52	17	I
33208	Microtubule-associated protein RP/EB family member 1	MAPRE1	29400	30151	5.02	IPI00017596	Q15691	247	3	K.LEHEYIQNFK.I K.KPLTSSSAAPQPISTQR.T R.NIELICQENEGENDPVLQRI	69 68 111	2e-05 2e-05 6.7e-10	17.7 -9.96 -18.16	17	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/ Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
12318	Microfibril-associated glycoprotein 4	MFAP4	35800	28972	5.38	IPI00022792	P55083	58	1	R.DQDLFVQNCALSSGAFWFR.S	58	0.00011	-15.6	7	D
13201	Microfibril-associated glycoprotein 4	MFAP4	34400	28972	5.38	IPI00022792	P55083	147	2	R.ADG EWYWLGLQNMHLLTLK.Q + Oxidation (M) R.DQDLFVQNCALSSGAFWFR.S	40 106	0.0082 1.9e-09	-6.90 -10.26	14	D
22207	Microfibril-associated glycoprotein 4	MFAP4	34700	28972	5.38	IPI00022792	P55083	76	1	R.DQDLFVQNCALSSGAFWFR.S	76	2.2e-06	-12.46	7	D
23003	Myotrophin	MITPN	11600	13058	5.27	IPI00924816	P58546	125	1	K.GPDGLTAFEATDNQAIK.A	125	3.50E-11	15.5	14	I
31111	Myosin regulatory light chain 12A	MYL12A	17100	19839	4.67	IPI00220573	P19105	254	2	R.FTDEEVDELRY.E R.NAFACFDEEATGTIQEDYLR.E	85 87	4.9e-07 1.7e-07	17.3 -20.95	57	I
11104	Myosin regulatory light chain 12B	MYL12B	18200	19824	4.71	IPI00033494	O14950	109	1	R.FTDEEVDELRY.E	64	5.2e-05	64.6	29	I
11110	Myosin regulatory light chain 12B	MYL12B	18700	19824	4.71	IPI00033494	O14950	218	2	R.FTDEEVDELRY.E R.NAFACFDEEATGTIQEDYLR.E	57 124	0.00027 2.8e-11	55.7 11.5	29	I
21004	Myosin regulatory light chain 12B	MYL12B	18400	19824	4.71	IPI00033494	O14950	465	6	K.LNGTDPEDVIR.N K.EAFNMIDQNR.D + Oxidation (M) K.GNFNYIEFTRI R.FTDEEVDELRY.E R.ATSNVVFAMFDQSQIQEFK.E + Oxidation (M) R.NAFACFDEEATGTIQEDYLR.E	57 35 36 91 85 160	0.00035 0.053 0.041 1.1e-07 3e-07 7.2e-15	25.8 30.0 25.2 27.3 -2.95 -8.38	46	I
21007	Myosin regulatory light chain 12B	MYL12B	17900	19824	4.71	IPI00033494	O14950	556	7	K.LNGTDPEDVIR.N R.FTDEEVDELRY.E R.DGFIDKEDLHDMLASLGK.N + Oxidation (M) R.ATSNVVFAMFDQSQIQEFK.E + Oxidation (M) R.NAFACFDEEATGTIQEDYLR.E K.EAFNMIDQNR.D + Oxidation (M) K.GNFNYIEFTRI	65 37 37 94 72 76 175	5.9e-05 0.033 0.036 5.2e-08 5.8e-06 2.7e-06 2.4e-16	10.8 14.8 15.1 17.0 -4.81 -8.05 -11.81	56	I
31109	Myosin regulatory light chain 12B	MYL12B	17500	19824	4.71	IPI00033494	O14950	273	2	R.FTDEEVDELRY.E R.NAFACFDEEATGTIQEDYLR.E	83 137	6.7e-07 1.6e-12	26.3 19.2	40	I
30021	Isoform Non-muscle of Myosin light polypeptide 6	MYL6B	14200	17090	4.56	IPI00335168	P60660	44	1	K.EAFQLFDR.T	39	0.022	28.7	24	I
21512	Putative uncharacterized protein NAP1L4	NAP1L4	52300	44280	4.62	IPI00017763	n/a	45	1	K.FYEEVHDLER.K	45	0.0048	33.9	2	I
24205	Alpha-soluble NSF attachment protein	NAPA	32700	33667	5.23	IPI00009253	P54920	382	5	K.IEEACEIYAR.A K.EAEAMALLAEAEK + Oxidation (M) K.YEELPFAFSDSR.E K.AIAHYEQSADYYK.G K.LLEAHEEQNVDSYTESVK.E	48 63 77 51 104	0.0026 7.1e-05 3.2e-06 0.0012 4.3e-09	27.3 27.6 26.8 20.6 2.29	34	D
17101	Isoform 2 of Nucleoside diphosphate kinase A	NME1	18900	19869	5.42	IPI00375531	P15531	396	6	K.EHYVDLK.D R.GDFCIQVGR.N K.DRPFFAGLVK.Y R.TFIAIKPDGVQR.G R.NIHGSDSVESAEK.E R.VMLGETNPADSKPGTIR.G + Oxidation (M)	41 39 61 85 57 113	0.017 0.021 0.00013 4.4e-07 0.00031 6.7e-10	39.6 42.4 52.3 64.9 62.9 56.2	38	I
27001	Isoform 1 of Nucleoside diphosphate kinase A	NME1	18500	17309	5.83	IPI00012048	P15531	417	5	R.GDFCIQVGR.N K.DRPFFAGLVK.Y R.TFIAIKPDGVQR.G R.NIHGSDSVESAEK.E R.VMLGETNPADSKPGTIR.G + Oxidation (M)	52 64 86 127 88	0.0012 6.8e-05 3.6e-07 2.9e-11 1.8e-07	13.4 20.9 25.4 20.0 6.30	40	I
37103	Isoform 1 of Nucleoside diphosphate kinase A	NME1	17500	17309	5.83	IPI00012048	P15531	459	7	K.EHYVDLK.D R.GDFCIQVGR.N K.DRPFFAGLVK.Y R.TFIAIKPDGVQR.G R.NIHGSDSVESAEK.E R.VMLGETNPADSKPGTIR.G R.VMLGETNPADSKPGTIR.G + Oxidation (M)	45 50 67 77 108 62 113	0.0074 0.0018 3.3e-05 2.9e-06 2.1e-09 7.3e-05 6.7e-10	42.1 44.1 42.4 47.8 39.7 20.7 24.9	45	I
25102	Nicotinamide N-methyltransferase	NNMT	27600	30011	5.56	IPI00027681	P40261	90	1	K.DTYLSHFNPR.D	67	3.1e-05	11.5	6	I
22103	ADP-sugar pyrophosphatase	NUDT5	30200	24597	4.87	IPI00296913	Q9UJKK9	215	2	K.EQTADGVAVIPVLR.T R.LDALVAEEHLTVDAR.V	106 110	3.4e-09 1.3e-09	5.11 4.23	13	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
22003	Nuclear transport factor 2	NUTF2	11800	14640	5.10	IPI00009901	P61970	61	1	K.NINDAWVCTNDMFR.L + Oxidation (M)	61	9.3e-05	0.15	11	I
12601	Protein disulfide-isomerase	P4HB	58500	57480	4.76	IPI00010796	P07237	366	4	K.FFPASADR.T K.YKPESEELTAERI K.VDATEESDLAQQYGV.R K.ILFIFIDSDHTDNQR.I	39 61 93 84	0.02 0.00012 5.7e-08 4.9e-07	40.2 92.1 77.6 71.6	19	I
32502	Protein disulfide-isomerase	P4HB	56300	57480	4.76	IPI00010796	P07237	287	3	K.YKPESEELTAERI K.VDATEESDLAQQYGV.R K.HNQLPLVIEFTQAPKI	61 74 38	0.00011 5.1e-06 0.014	39.4 26.8 14.5	18	I
28410	Proliferation-associated protein 2G4	PA2G4	49600	44101	6.13	IPI00299000	Q9UQ80	84	2	R.AFFSEVER.R K.AAHLCAEAALR.L	38 46	0.028 0.0039	-0.80 -1.88	4	I
26105	Platelet-activating factor acetylhydrolase IB subunit beta	PAFAH1B2	28200	25724	5.57	IPI00026546	P68402	131	1	R.ELFSPLHALNFGIGGDTTR.H	76	2.7e-06	-16.8	15	I
36104	Platelet-activating factor acetylhydrolase IB subunit beta	PAFAH1B2	26200	25724	5.57	IPI00026546	P68402	145	2	K.IIVLGLLPR.G R.ELFSPLHALNFGIGGDTTR.H	41 104	0.014 3.9e-09	7.86 5.96	12	I
38105	Protein DJ-1	PARK7	20900	20050	6.33	IPI00298547	Q99497	441	4	R.DVVICPDASLEDAKKE K.GAEEMETVIPVDVMR.R + 2 Oxidation (M) K.GLIAAICAGPTALLAHEIGFGSK.V K.EGPDYDVVVLPGGNLGAQNLSESAAVK.E	82 81 127 151	8.1e-07 9.6e-07 1.5e-11 6.2e-14	1.03 7.90 -21.09 -24.81	41	I
31203	Proliferating cell nuclear antigen	PCNA	30100	29092	4.57	IPI00021700	P12004	73	1	K.MPSGEFARI + Oxidation (M)	35	0.044	22.4	13	I
25104	Prohibitin	PHB	28300	29843	5.57	IPI00017334	P35232	785	8	K.QVAQQEAER.A R.FDAGELITQR.E K.DLQNVNITLRI R.IIFRPVASQLPRI R.IFTSIGEDYDER.V R.KLEAAEDIAYQLSR.S K.AAELIANSLATAGDGLIELR.K K.FGLALAVAGGVVNSALYNVDAGHR.A	52 76 84 45 105 113 132 178	0.0013 4e-06 6.9e-07 0.0048 4.2e-09 6e-10 6.9e-12 1.2e-16	16.7 25.7 21.3 20.4 27.4 20.6 -1.71 -14.32	40	I
35212	Prohibitin	PHB	26500	29843	5.57	IPI00017334	P35232	175	1	R.FDAGELITQR.E	36	0.039	8.28	31	I
25209	Inorganic pyrophosphatase	PPA1	33800	33095	5.54	IPI00015018	Q15181	476	7	R.AAPFSLRYR.V R.YVANLFPYK.G K.DVFMVVEVPR.W K.DVFMVVEVPR.W + Oxidation (M) R.LKPGYLEATVDWFR.R K.GQYISPFHDIPIYADK.D K.VPDGKPENEFNAFEFK.D	68 49 44 36 82 114 89	2.7e-05 0.0021 0.0066 0.039 7.6e-07 4.3e-10 1.4e-07	2.47 9.32 10.9 15.7 6.89 -2.14 -6.35	28	I
35211	Inorganic pyrophosphatase	PPA1	31800	33095	5.54	IPI00015018	Q15181	208	2	R.LKPGYLEATVDWFR.R K.VPDGKPENEFNAFEFK.D	65 86	3.6e-05 2.5e-07	-17.47 -30.23	14	I
17102	Peroxiredoxin 3 isoform b	PRDX3	23900	26107	7.04	IPI00374151	n/a	231	2	K.HLSVNDLPVGR.S R.DYGVLLLEGSLALR.G	74 115	7.6e-06 4.1e-10	55.4 56.5	18	I
18208	Peroxiredoxin 3 isoform b	PRDX3	24900	26107	7.04	IPI00374151	n/a	122	2	K.HLSVNDLPVGR.S R.DYGVLLLEGSLALR.G	57 65	0.00034 4.9e-05	26.9 20.1	10	I
27105	Peroxiredoxin 3 isoform b	PRDX3	23700	26107	7.04	IPI00374151	n/a	268	3	K.HLSVNDLPVGR.S R.GLFIDPNGVIK.H R.DYGVLLLEGSLALR.G	80 46 122	1.6e-06 0.0038 8.9e-11	17.0 16.2 21.2	18	I
37105	Peroxiredoxin 3 isoform b	PRDX3	21900	26107	7.04	IPI00374151	n/a	111	2	K.HLSVNDLPVGR.S R.DYGVLLLEGSLALR.G	58 41	0.00025 0.012	19.9 15.0	13	I
11802	highly similar to Glucosidase 2 subunit beta	PRKCSH	96000	61066	4.35	IPI00026154	P14314	98	1	R.NKFEAEER.S	42	0.0097	9.56	5	I
29102	Isoform Long of Proteasome subunit alpha type-1	PSMA1	31100	30505	6.51	IPI00472442	P25786	133	2	R.FVFDRLPVSRL R.NQYDNDVTVWSPQGR.I	38 95	0.027 3.6e-08	18.3 7.38	9	I
39201	Isoform Short of Proteasome subunit alpha type-1	PSMA1	29600	29822	6.15	IPI00016832	P25786	61	1	R.NQYDNDVTVWSPQGR.I	39	0.015	-13.01	9	I
24105	Isoform 2 of Proteasome subunit alpha type-3	PSMA3	28200	27858	5.19	IPI00171199	P25788	95	1	K.AFELELSWVGELTNGR.H	70	1.3e-05	11.1	10	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
21112	Proteasome subunit alpha type-5	PSMA5	26600	26565	4.74	IPI00291922	P28066	210	3	R.GVNTFSPEGR.L R.LFQVEYAIEAIK.L R.AIGSASEGAQSSLQEVYHK.S	48 61 100	0.0028 9.9e-05 9.3e-09	22.0 27.3 0.51	17	I
28105	Proteasome subunit beta type-3	PSMB3	24600	23219	6.14	IPI00028004	P49720	223	2	R.FGPHYTEPVIAGLDPK.T R.LYIGLAGLATDVQIVAQR.L	86 137	2.6e-07 2.1e-12	-15.66 -20.05	16	I
38113	Proteasome subunit beta type-3	PSMB3	22900	23219	6.14	IPI00028004	P49720	208	2	R.FGPHYTEPVIAGLDPK.T R.LYIGLAGLATDVQIVAQR.L	75 120	3.8e-06 9.7e-11	4.53 1.8	20	I
26108	Proteasome subunit beta type-4	PSMB4	24600	29242	5.72	IPI00555956	P28070	312	4	R.AIHSLWLR.A K.QPVLSTQTEAR.D K.FEGGVVIAADMLGSGYSLAR.F + Oxidation (M) K.QVLGQMVIDEELLGDGHSYSPR.A + Oxidation (M)	40 51 104 117	0.016 0.0014 4.4e-09 1.5e-10	-7.27 9.16 -9.59 -21.40	22	I
16206	Proteasome activator complex subunit 1	PSME1	29500	28876	5.78	IPI00479722	Q06323	266	4	K.QPHVGDYR.Q K.LEGFHTQISK.Y R.QLVHELDEAEYR.D R.IEDGNNGFVAVQEK.V	51 50 74 37	0.0014 0.0016 4.8e-06 0.027	69.9 76.8 65.5 71.7	26	I
26114	Proteasome activator complex subunit 1	PSME1	29000	28876	5.78	IPI00479722	Q06323	445	6	K.QPHVGDYR.Q K.ISELDALFK.E K.LEGFHTQISK.Y R.NAYAVLYDIILKN R.QLVHELDEAEYR.D R.IEDGNNGFVAVQEK.V	56 61 62 53 95 86	0.00044 0.00015 0.00011 0.00079 4.2e-08 3.7e-07	5.73 15.4 24.7 22.2 30.3 25.4	30	I
36211	Proteasome activator complex subunit 1	PSME1	27200	28876	5.78	IPI00479722	Q06323	131	2	K.LEGFHTQISK.Y R.QLVHELDEAEYR.D	35 43	0.051 0.0065	21.2 17.8	19	I
25101	Proteasome activator subunit 2	PSME2	29000	27555	5.54	IPI00746205	n/a	292	5	R.KQVEVFR.Q R.DEAAYGELR.A K.IEDGNDFGVAVQEK.V R.QNLFQEAEEFLYR.F K.IIYLNQLIQEDSLNVDLTSR.A	44 51 79 66 32	0.0072 0.0012 1.7e-06 2.9e-05 0.046	-23.01 -10.52 0.58 -3.65 -35.09	30	I
35202	Proteasome activator subunit 2	PSME2	27200	27555	5.54	IPI00746205	n/a	128	2	R.KQVEVFR.Q R.QNLFQEAEEFLYR.F	37 37	0.038 0.025	0.27 18.1	24	I
22502	UV excision repair protein RAD23 homolog B	RAD23B	55500	43202	4.79	IPI00008223	P54727	284	5	R.ASFNNDPDR.A R.AVEYLLMGIPGDRE R.AVEYLLMGIPGDRE + Oxidation (M) R.QIIQQNPSSLPALLQQIGRE K.QEKPAEKPAETPVATSPATDSTSGDSSR.S	49 45 53 98 60	0.0026 0.005 0.00073 1.4e-08 6.9e-05	12.9 32.1 35.5 8.23 -3.83	19	I
21409	Histone-binding protein RBBP4	RBBP4	54200	47911	4.74	IPI00328319	Q09028	421	6	K.LMIWWDTR.S + Oxidation (M) K.TVALWDLR.N K.INHEGEVNR.A K.EAAFDDAVEER.V K.HPSKPDPSGECNPDLR.L K.IGEEQSPEDAEDGPELFIHGGHTAKI	36 56 61 87 98 57	0.046 0.00049 0.00013 3.3e-07 1.8e-08 0.00013	-3.76 -2.67 6.87 15.5 -1.84 -30.40	20	I
31511	Histone-binding protein RBBP4	RBBP4	51300	47911	4.74	IPI00328319	Q09028	315	4	K.TVALWDLR.N K.INHEGEVNR.A K.EAAFDDAVEER.V K.HPSKPDPSGECNPDLR.L	54 44 82 91	0.00081 0.006 1.1e-06 8.3e-08	12.0 11.9 1.51 -17.94	14	I
11402	Ribonuclease inhibitor	RNH1	48500	51766	4.71	IPI00550069	P13489	104	2	R.VNPALAEELNLR.S R.FLLELQISNNR.L	54 50	0.00053 0.0018	32.7 24.8	4	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
21405	Ribonuclease inhibitor	RNH1	48800	51766	4.71	IPI00550069	P13489	796	9	R.VNPALAEI.NLR.S R.FLLELQISNNR.L K.LESCGVTS.DNCR.D R.ELCQGLGQPGSVLR.V K.ELSLAGNELGDEGAR.L K.ELTVSNN.DINEAGVR.V R.WAELLPLLQQCQVVR.L K.LSLQNCCLTGAGCGVLSSTLR.T K.LQLEYCSLSAASCEPLASVLR.A	80 41 78 70 117 132 50 87 142	0.0000013 0.013 2.2e-06 1.4e-05 2.4e-10 8.1e-12 0.001 1.8e-07 5.1e-13	41.3 41.6 41.8 35.9 38.7 35.3 24.1 13.6 15.0	29	I
31415	Ribonuclease inhibitor	RNH1	46200	51766	4.71	IPI00550069	P13489	857	8	R.VNPALAEI.NLR.S R.FLLELQISNNR.L R.ELCQGLGQPGSVLR.V K.ELSLAGNELGDEGAR.L K.ELTVSNN.DINEAGVR.V R.WAELLPLLQQCQVVR.L K.LQLEYCSLSAASCEPLASVLR.A R.ELDSLNNCLGDAGILQLVESVR.Q	90 59 92 122 129 92 133 140	1.4e-07 0.00019 9.2e-08 7.7e-11 1.7e-11 6.4e-08 4.4e-12 9.8e-13	-8.90 -15.03 -14.50 -9.12 -12.50 -37.01 -31.90 -29.48	26	I
11307	RPSA 40S ribosomal protein SA	RPSAP15	39200	32947	4.79	IPI00553164	P08865	211	3	K.FAAATGATPIAGRF R.FTPGTFTNQIQAAFR.E R.AIVAIENPADVSVSSR.N	68 63 73	2.6e-05 5.4e-05 5.5e-06	15.6 13.3 13.4	24	I
21210	RPSA 40S ribosomal protein SA	RPSAP15	39100	32947	4.79	IPI00553164	P08865	430	5	K.FAAATGATPIAGRF R.FTPGTFTNQIQAAFR.E R.AIVAIENPADVSVSSR.N R.EHPWEVMPDLYFYR.D + Oxidation (M) R.ADHQPLTEASYVNLPTIALCNTDSPLR.Y	86 95 121 35 62	4.1e-07 4.2e-08 1.1e-10 0.034 3.8e-05	31.9 33.2 31.4 20.8 -2.58	31	I
31308	RPSA 33 kDa protein	RPSAP15	36600	33464	4.79	IPI00413108	n/a	236	3	K.FAAATGATPIAGRF R.FTPGTFTNQIQAAFR.E R.AIVAIENPADVSVSSR.N	59 65 93	0.00023 3.8e-05 6.2e-08	23.4 13.8 19.5	17	I
29502	Isoform 1 of RuvB-like 1	RUVBL1	57000	50538	6.02	IPI00021187	Q9Y265	41	1	R.ALESSIAPVIFASNR.G	41	0.0088	-4.83	3	D
27005	Protein S100-A11	S100A11	11700	11847	6.56	IPI00013895	P31949	184	2	R.CIESLIAVFKQ.Y K.TEFLSFMNTELA.AAFK.N + Oxidation (M)	57 93	0.00031 5.7e-08	8.58 -6.05	34	I
17005	Protein S100-A13	S100A13	12500	11464	5.91	IPI00016179	Q99584	102	1	K.ELVTQQLPHLLK.D	77	2.6e-06	95	18	I
18002	Protein S100-A16	S100A16	12200	11851	6.28	IPI00062120	Q96FQ6	37	1	K.LIQNL.DANHDGR.I	37	0.028	27.7	11	I
36009	Serum amyloid A protein	SAA2	11100	13581	6.28	IPI00552578	P02735	144	2	R.SFFSFLGEAFDGDAR.D R.FFGHGAEDSLADQAANEWGR.S	42 72	0.008 6.6e-06	12.3 2.18	40	D
13306	Protein SEC13 homolog	SEC13	36200	36347	5.31	IPI00375370	n/a	36	1	K.LEAHS.DWVR.D	36	0.034	57.7	2	I
23314	Protein SEC13 homolog	SEC13	36000	36347	5.31	IPI00375370	n/a	104	2	K.LEAHS.DWVR.D R.DVAWAPSIGLPTSTIASCSQDGR.V	53 51	0.00083 0.00061	6.72 -18.60	9	I
33313	Protein SEC13 homolog	SEC13	34300	36347	5.31	IPI00375370	n/a	176	2	K.LEAHS.DWVR.D R.DVAWAPSIGLPTSTIASCSQDGR.V	68 108	2.1e-05 1.3e-09	18.2 -12.98	9	I
38409	Isoform 1 of Septin-2	SEPT2	42200	41689	6.15	IPI00014177	Q15019	302	4	R.YLHDESGLNR.R K.TIISVIDEQFER.Y R.TVQIEASTVEIEER.G K.STLINSLFL.TDLYPER.V	60 59 81 101	0.00018 0.00017 9.7e-07 7.9e-09	18.4 24.1 47.1 11.1	14	I
12505	Isoform 1 of Alpha-1-antitrypsin	SERPINA1	57100	46878	5.37	IPI00553177	P01009	268	2	K.TDTS.HHDQDHP.TFNK.I K.LQHLENL.THDIITK.F	111 71	9.7e-10 1.1e-05	39 35.7	11	D
12604	Isoform 1 of Alpha-1-antitrypsin	SERPINA1	59100	46878	5.37	IPI00553177	P01009	84	1	K.TDTS.HHDQDHP.TFNK.I	53	0.00058	50.8	7	D
12607	Isoform 1 of Alpha-1-antitrypsin	SERPINA1	59200	46878	5.37	IPI00553177	P01009	195	2	K.TDTS.HHDQDHP.TFNK.I K.LQHLENL.THDIITK.F	83 45	5.2e-07 0.004	1.67 -1.87	15	D
17308	Isoform 3 of Alpha-1-antitrypsin	SERPINA1	33300	34905	5.04	IPI00869004	P01009	133	2	K.LQHLENL.THDIITK.F K.DTEEDFHVDQV.TTVK.V	50 57	0.0013 0.00021	46.1 58.0	16	D
27206	Isoform 3 of Alpha-1-antitrypsin	SERPINA1	33000	34905	5.04	IPI00869004	P01009	111	1	K.LQHLENL.THDIITK.F	59	0.00017	-4.65	10	D
11204	Isoform 1 of 14-3-3 protein sigma	SFN	28600	27871	4.68	IPI00013890	P31947	65	1	K.SNEEGSE.EKGP.EVR.E	65	4e-05	107	5	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
17401	Isoform 1 of Calcium-binding mitochondrial carrier protein SCaMC-1	SLC25A24	47300	53548	6.00	IPI00337494	Q6NUK1	373	4	KHEGLGAFYK G R.YETLFQALDRN R.NGDGVVDIGELQEGLRN R.DYFLFNPVTDIEEIR.F	46 53 90 73	0.0043 0.00072 1.2e-07 5.7e-06	78.0 85.6 72.0 51.8	20	I
31009	Small nuclear ribonucleoprotein F	SNRPF	10300	9776	4.70	IPI00220528	P62306	77	2	R.CNNVLYIR.G R.GVEEEEDGEMRE. - + Oxidation (M)	39 38	0.02 0.021	8.92 14.4	24	I
16109	Stathmin	STMN1	18100	17292	5.76	IPI00479997	P16949	100	2	K.DKHIEEVR.K R.ASGQAFELLSPR.S	35 56	0.05 0.0004	45.6 65.4	14	I
26007	Stathmin	STMN1	17600	17292	5.76	IPI00479997	P16949	58	1	R.ASGQAFELLSPR.S	58	0.00023	24	8	I
36006	Stathmin	STMN1	16700	17292	5.76	IPI00479997	P16949	199	3	K.KLEAAEER.R K.DKHIEEVR.K R.ASGQAFELLSPR.S	50 49 84	0.0022 0.0019 6.3e-07	7.56 7.09 16.7	20	I
27304	TDP43	TARDBP	45900	45305	5.94	IPI00025815	Q13148	36	1	R.FGGNPGGFGNQGGFGNSR.G	36	0.029	-15.93	4	I
14101	Tubulin-specific chaperone A	TBCA	15000	12904	5.25	IPI00217236	Q75347	55	1	R.RLEAAYLDLQRI	55	0.00047	27.3	10	I
24004	Tubulin-specific chaperone A	TBCA	14400	12904	5.25	IPI00217236	Q75347	50	1	R.RLEAAYLDLQRI	50	0.0015	13.1	10	I
34001	Tubulin-specific chaperone A	TBCA	13500	12904	5.25	IPI00217236	Q75347	84	1	R.RLEAAYLDLQRI	52	0.00098	3.20	10	I
23102	Tubulin-folding cofactor B	TBCB	31300	27594	5.06	IPI00293126	Q99426	243	3	R.IHVIDHSGAR.L K.YTISQEAYDQR.Q K.LDQEDALLGSYPVDDGCR.I	60 66 85	0.00018 3.6e-05 2.7e-07	23.5 32.6 9.37	19	I
21003	Transcription elongation factor B polypeptide 1	TCEB1	12700	12636	4.74	IPI00300341	Q15369	117	2	K.LISSDGHFIVK.R K.AMLSPGQFAENETNEVNR.E + Oxidation (M)	85 32	4.2e-07 0.052	6.38 -22.31	28	I
31015	Transcription elongation factor B polypeptide 2	TCEB2	15100	13239	4.73	IPI00026670	Q15370	34	1	K.RPPDEQR.L	34	0.058	16.0	5	I
32104	Tumor protein D52 isoform 2	TPD52	22800	22464	5.30	IPI00619951	n/a	78	1	K.VEEIQTLSQVLAAKE	67	2.3e-05	-4.51	22	I
11312	Troponin 1 alpha chain isoform 2	TPM1	36800	32715	4.70	IPI00000230	n/a	170	2	R.IQLVEEELDR.A R.KLVIESDLER.A	35 54	0.044 0.00061	68.6 64.8	13	D
21205	Troponin 1 alpha chain isoform 2	TPM1	33100	32715	4.70	IPI00000230	n/a	309	6	R.KYEEVAR.K K.HIAEDADR.K K.LVIESDLER.A K.EDRYEEI.K V R.IQLVEEELDR.A R.KLVIESDLER.A	43 35 46 50 64 73	0.0083 0.054 0.0045 0.0016 6.1e-05 7.2e-06	2.22 11.0 23.9 28.8 26.7 22.7	15	I
31208	Isoform 3 of Troponin alpha-1 chain	TPM1	31500	32856	4.72	IPI00216135	P09493	251	4	R.KYEEVAR.K K.EDRYEEI.K V R.IQLVEEELDR.A R.KLVIESDLER.A	49 35 55 70	0.0024 0.051 0.00051 1.6e-05	2.91 20.1 19.3 22.8	18	I
31303	Troponin 1 alpha chain isoform 2	TPM1	34800	32715	4.70	IPI00000230	n/a	105	2	R.KYEEVAR.K R.IQLVEEELDR.A	42 64	0.012 5.4e-05	-5.43 39.9	5	I
11302	Troponin 2	TPM2	37800	37010	4.73	IPI00513698	n/a	81	1	R.KLVILEGELER.A	62	9.3e-05	32.4	5	D
11213	Troponin alpha-3 chain isoform 2	TPM3	30400	29243	4.75	IPI00218319	P06753	587	9	R.KYEEVAR.K K.HIAEEADR.K K.LVIEEGDLER.T R.IQLVEEELDR.A R.KLVIEGDLER.T R.EQAEAEVASLNR.R R.RIQLVEEELDR.A K.IQVLQQQADDAEER.A R.KIQVLQQQADDAEER.A	46 46 35 62 61 44 62 99 110	0.005 0.0037 0.049 9.1e-05 9.7e-05 0.0066 9.4e-05 1.6e-08 1.2e-09	100 114 131 133 137 139 136 133 128	29	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
21115	Tropomyosin alpha-3 chain isoform 2	TPM3	30100	29243	4.75	IPI00218319	P06753	781	10	K.HIAEEADR.K K.HIAEEADR.K.Y K.EEHLCTQRM K.LVIEGDLER.T R.IQLVEEELDR.A R.KLVIEGDLER.T R.EQAEAEVASLNR.R R.RIQLVEEELDR.A K.IQVLQQQADDAEER.A R.KIQVLQQQADDAEER.A	67 45 62 70 70 82 67 69 114 136	3.2e-05 0.0051 0.00015 1.7e-05 1.7e-05 8.7e-07 3.2e-05 1.9e-05 4.6e-10 2.9e-12	-15.22 -3.77 -3.19 3.41 6.68 12.7 10.7 11.8 8.56 4.50	26	I
31210	Tropomyosin 3 isoform 5	TPM3	28500	28890	4.76	IPI00477649	n/a	228	3	R.IQLVEEELDR.A R.KLVIEGDLER.T K.IQVLQQQADDAEER.A	40 53 103	0.016 0.00078 5.8e-09	31.2 31.8 29.6	14	I
21111	Tropomyosin alpha-4 chain isoform 1	TPM4	30300	28619	4.67	IPI00010779	P67936	761	9	K.HIAEEADR.K K.HIAEEADR.K.Y K.LVILEGELER.A R.IQLVEEELDR.A R.EKAEGDVAALNR.R R.KLVILEGELER.A R.RIQLVEEELDR.A K.IQALQQQADEAEDR.A R.KIQALQQQADEAEDR.A	70 48 71 81 96 80 70 121 125	1.6e-05 0.0026 1.1e-05 1.2e-06 4.1e-08 1.6e-06 1.4e-05 1.1e-10 3.4e-11	11.9 19.1 32.2 34.0 29.5 30.1 30.4 26.3 27.0	23	I
21203	Tropomyosin alpha-4 chain isoform 1	TPM4	36400	32874	4.69	IPI00216975	P67936	477	7	K.HIAEEADR.K K.LVILEGELER.A R.IQLVEEELDR.A K.ASDAEGDVAALNR.R R.KLVILEGELER.A K.QVEELTHLQK.K R.RIQLVEEELDR.A	58 68 67 47 49 81 57	0.00024 2.3e-05 3.1e-05 0.003 0.0018 1.2e-06 0.0003	27.7 44.9 50.0 46.3 48.4 52.6 46.9	22	I
31205	Tropomyosin alpha-4 chain isoform 1	TPM4	28700	28619	4.67	IPI00010779	P67936	603	9	K.HIAEEADR.K K.HIAEEADR.K.Y K.LVILEGELER.A R.IQLVEEELDR.A R.EKAEGDVAALNR.R R.KLVILEGELER.A R.RIQLVEEELDR.A K.IQALQQQADEAEDR.A R.KIQALQQQADEAEDR.A	65 44 66 63 62 73 57 88 87	5.4e-05 0.0069 4e-05 7.3e-05 0.00011 7.1e-06 0.00032 2.3e-07 2.1e-07	-6.38 2.52 10.3 10.8 6.21 11.2 11.3 7.08 3.00	23	I
12102	Tumor protein, translationally-controlled 1	TPT1	22400	21626	5.34	IPI00009943	n/a	134	2	R.VKPFMTGAAEQIK.H + Oxidation (M) R.DLISHDEMFSDIYK.I + Oxidation (M)	37 55	0.029 0.00036	52.2 44.6	19	I
22101	Tumor protein, translationally-controlled 1	TPT1	22500	21626	5.34	IPI00009943	n/a	133	1	R.DLISHDEMFSDIYK.I + Oxidation (M)	105	3.6e-09	39.5	11	I
32101	Tumor protein, translationally-controlled 1	TPT1	21100	21626	5.34	IPI00009943	n/a	128	2	R.VKPFMTGAAEQIK.H + Oxidation (M) R.DLISHDEMFSDIYK.I + Oxidation (M)	36 57	0.032 0.00024	7.74 -1.81	30	I
18301	GDP-L-fucose synthetase	TSTA3	41900	36098	6.12	IPI00014361	Q13630	42	1	K.ETCAWFTDNYEQAR.K	42	0.0069	35.5	4	I
25005	Transthyretin	TTR	15100	15991	5.52	IPI00022432	P02766	372	5	R.GSPAINVAVHVF.R.K K.AADDTWEPFASGK.T R.KAADDTWEPFASGK.T K.ALGISPFHEHAEVVFTANDSGP.R.R K.TSESGLHGLITTEEFVEGIYK.V	110 48 42 107 65	1.2e-09 0.0022 0.0088 1.5e-09 2.3e-05	20.0 20.3 13.6 -14.88 -16.70	48	D
34008	Transthyretin	TTR	14400	15991	5.52	IPI00022432	P02766	82	1	R.GSPAINVAVHVF.R.K	50	0.0012	17.7	24	D
35003	Transthyretin	TTR	14100	15991	5.52	IPI00022432	P02766	65	1	K.ALGISPFHEHAEVVFTANDSGP.R.R	47	0.0017	6.94	39	D
22005	Thioredoxin	TXN	12300	12015	4.82	IPI00216298	P10599	80	1	K.TAFQEAALDAAGDK.L	80	1.4e-06	22.5	12	I
11304	Vimentin	VIM	40700	53676	5.06	IPI00418471	P08670	52	1	R.EEAENTLQSF.R.Q	48	0.0022	55.0	6	D

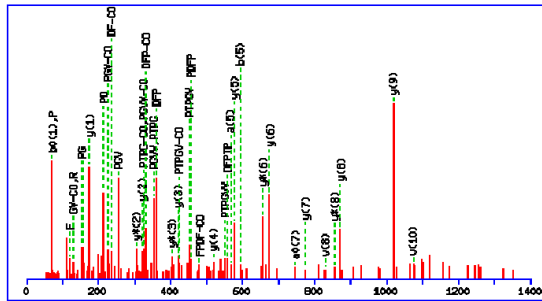
Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
11405	Vimentin	VIM	42800	53676	5.06	IPI00418471	P08670	160	2	R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E	51 50	0.0011 0.0015	34.8 27.1	13	D
11407	Vimentin	VIM	44800	53676	5.06	IPI00418471	P08670	193	3	R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E R.EMEEFAVEAANYQDTIGRL + Oxidation (M)	61 52 61	0.0001 0.0008 6.3e-05	30.2 31.0 -2.81	12	D
11408	Vimentin	VIM	47200	53676	5.06	IPI00418471	P08670	206	3	R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E R.EMEEFAVEAANYQDTIGRL + Oxidation (M)	54 74 33	0.00059 5.1e-06 0.041	88.2 77.6 51.9	13	D
12401	Vimentin	VIM	48300	53676	5.06	IPI00418471	P08670	148	2	R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E	41 53	0.01 0.00074	94.5 88.0	9	D
21311	Vimentin	VIM	47200	53676	5.06	IPI00418471	P08670	426	5	K.LLEGEESR.I K.FADLSEAANR.N K.VELQELNDR.F R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E	38 53 61 78 81	0.032 0.00092 0.00013 2.1e-06 9.9e-07	11.5 36.0 42.5 50.5 39.4	16	D
21410	Vimentin	VIM	47500	53676	5.06	IPI00418471	P08670	299	3	R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E R.EMEEFAVEAANYQDTIGRL + Oxidation (M)	45 53 83	0.0048 0.0007 4.1e-07	43.2 33.6 6.18	20	D
21416	Vimentin	VIM	41800	53676	5.06	IPI00418471	P08670	366	5	K.FADLSEAANR.N R.LGDLYEEEMR.E + Oxidation (M) R.EEAENTLQSF.R R.ISLPLPNFSSLNLR.E R.EMEEFAVEAANYQDTIGRL + Oxidation (M)	35 37 74 59 107	0.047 0.026 6.3e-06 0.00016 1.7e-09	32.9 43.0 39.6 31.0 6.62	18	D
32407	Vimentin	VIM	46900	53676	5.06	IPI00418471	P08670	330	4	R.SLYASSPGGVYATR.S R.ISLPLPNFSSLNLR.E R.DGQVINETSQHDDLE.- R.EMEEFAVEAANYQDTIGRL + Oxidation (M)	42 37 57 101	0.0085 0.03 0.00025 6.6e-09	28.2 27.1 25.3 20.3	24	I
21118	14-3-3 protein beta/alpha isoform long	YWHAB	27400	28179	4.76	IPI00216318	P31946	288	2	K.AVTEQGHELSNEER.N K.TAFDEAIAELDTLNEESYK.D	108 66	2e-09 2.6e-05	31.9 10.7	24	I
31110	14-3-3 protein beta/alpha isoform short	YWHAB	25800	27947	4.76	IPI00759832	P31946	226	2	K.AVTEQGHELSNEER.N K.TAFDEAIAELDTLNEESYK.D	98 66	1.8e-08 2.2e-05	21.0 14.6	24	I
11203	14-3-3 protein epsilon	YWHAE	29900	29326	4.63	IPI00000816	P62258	150	3	K.HLIPAANTGESK.V R.YLAEFATGNDR.K K.VAGMDVELTVEER.N + Oxidation (M)	44 47 48	0.0064 0.0036 0.0022	46.7 41.5 49.5	17	I
21105	14-3-3 protein epsilon	YWHAE	29500	29326	4.63	IPI00000816	P62258	389	5	K.DSTLJMQLLR.D + Oxidation (M) K.HLIPAANTGESK.V R.YLAEFATGNDR.K K.VAGMDVELTVEER.N + Oxidation (M) K.AASDIAMTELPPTHPIRL + Oxidation (M)	42 66 75 70 81	0.0098 3.8e-05 4.7e-06 1.3e-05 9.8e-07	28.4 28.2 34.4 33.7 20.1	27	I
21201	14-3-3 protein epsilon	YWHAE	31700	29326	4.63	IPI00000816	P62258	107	2	R.YLAEFATGNDR.K K.AASDIAMTELPPTHPIRL + Oxidation (M)	55 52	0.00051 0.00083	30.3 18.9	10	I
31201	14-3-3 protein epsilon	YWHAE	28000	29326	4.63	IPI00000816	P62258	68	1	R.YLAEFATGNDR.K	44	0.0059	12.7	15	I
21120	14-3-3 protein gamma	YWHAG	27700	28456	4.80	IPI00220642	P61981	294	3	R.YLAEVATGEK.R K.AYSEAEHSK.E K.NVTELNEPLSNEER.N	49 47 120	0.002 0.004 1.1e-10	31.5 39.8 28.7	24	I
31112	14-3-3 protein gamma	YWHAG	26100	28456	4.80	IPI00220642	P61981	204	2	K.EHMQPETHPIRL + Oxidation (M) K.NVTELNEPLSNEER.N	35 134	0.054 5.3e-12	25.4 33.5	21	I
11208	14-3-3 protein theta	YWHAQ	28100	28032	4.68	IPI00018146	P27348	237	2	R.YLAEVACGDDR.K K.AVTEQGAELSNEER.N	38 87	0.024 2.9e-07	70.2 77.2	29	I

Spot #	Protein name	Gene symbol	Predicted mass	Actual mass	pI	Accession #	Accession # UniprotKB/Swiss-Prot	Score	Number of peptides matched	Peptide sequence	Peptide score	Expect value	Mass error (ppm)	Sequence coverage (%)	Increased (I) or Decreased (D) expression in CR
21109	14-3-3 protein theta	YWHAQ	27600	28032	4.68	IPI00018146	P27348	496	7	K.MKGDYFR.Y + Oxidation (M) R.EKVESEL.R.S K.EMQPTHPIR.L + Oxidation (M) K.DSTLIMQLLR.D + Oxidation (M) R.YLAEVACGDDR.K R.YLAEVACGDDR.K.Q K.AVTEQGAELSNEER.N	38 62 45 43 73 67 135	0.029 0.00012 0.0047 0.0077 7.3e-06 3e-05 3.8e-12	-6.44 -0.78 12.8 14.2 17.6 13.4 16.0	24	I
31105	14-3-3 protein theta	YWHAQ	26100	28032	4.68	IPI00018146	P27348	241	3	R.YLAEVACGDDR.K R.YLAEVACGDDR.K.Q K.AVTEQGAELSNEER.N	45 37 116	0.0051 0.03 3.7e-10	24.1 18.1 26.0	18	I
31106	14-3-3 protein theta	YWHAQ	25100	28032	4.68	IPI00018146	P27348	138	2	R.YLAEVACGDDR.K K.AVTEQGAELSNEER.N	34 104	0.058 5.4e-09	26.4 19.8	10	I
31202	14-3-3 protein theta	YWHAQ	26800	28032	4.68	IPI00018146	P27348	78	1	K.AVTEQGAELSNEER.N	62	7.5e-05	14.0	9	I
11212	14-3-3 protein zeta/delta	YWHAZ	27800	27899	4.73	IPI00021263	P63104	250	2	K.SVTEQGAELSNEER.N K.GVVDQSQQAYQEAFAEISK.K	91 81	1e-07 7.9e-07	64.5 39.8	21	I
21113	14-3-3 protein zeta/delta	YWHAZ	27300	27899	4.73	IPI00021263	P63104	562	7	K.MKGDYYR.Y + Oxidation (M) R.EKIETELR.D K.EMQPTHPIR.L + Oxidation (M) K.DSTLIMQLLR.D + Oxidation (M) R.YLAEVAAGDDK.K.G K.SVTEQGAELSNEER.N K.GVVDQSQQAYQEAFAEISK.K	41 59 35 42 73 143 137	0.015 0.00023 0.053 0.012 7.8e-06 6.1e-13 1.9e-12	-18.73 -14.20 -2.66 0.05 -0.86 -1.87 -21.64	32	I
31108	14-3-3 protein zeta/delta	YWHAZ	25900	27899	4.73	IPI00021263	P63104	391	4	R.YLAEVAAGDDK.K.G K.SVTEQGAELSNEER.N K.GVVDQSQQAYQEAFAEISK.K K.TAFDEAIAELDTLSEESYK.D	55 88 108 43	0.00051 2.2e-07 1.3e-09 0.0047	-2.19 -1.95 -12.24 -10.43	36	I

APPENDIX 11: Peptide views for protein identifications with a only single peptide matches in 2/3 and 3/3 2D-PAGE/MS experiments (n=3)

Peptide View: Adenine phosphoribosyl transferase (APRT)

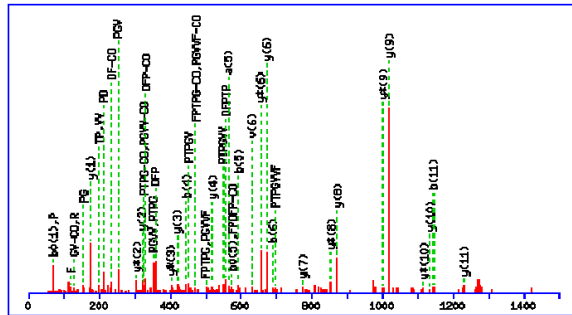
Experiment 1 (spot 15102)



Monoisotopic mass of neutral peptide Mr(calc): 1464.7402
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 51 Expect: 0.001
 Matches : 51/196 fragment ions using 61 most intense peaks ([help](#))

#	Immon.	a	a ⁰	b	b ⁰	Seq.	v	w	w'	y	y*	y ⁰	#
1	60.0444	60.0444	42.0338	88.0393	70.0287	S							13
2	120.0808	207.1128	189.1022	235.1077	217.0972	F	1286.6528			1378.7154	1361.6889	1360.7048	12
3	70.0651	304.1656	286.1550	332.1605	314.1499	P	1189.6000	1188.6048		1231.6470	1214.6204	1213.6364	11
4	88.0393	419.1925	401.1819	447.1874	429.1769	D	1074.5731	1073.5778		1134.5942	1117.5677	1116.5837	10
5	120.0808	566.2609	548.2504	594.2558	576.2453	F	927.5047			1019.5673	1002.5407	1001.5567	9
6	70.0651	663.3137	645.3031	691.3086	673.2980	P	830.4519	829.4567		872.4989	855.4723	854.4883	8
7	74.0600	764.3614	746.3508	792.3563	774.3457	T	729.4042	742.4246	744.4039	775.4461	758.4196	757.4355	7
8	70.0651	861.4141	843.4036	889.4090	871.3985	P	632.3515	631.3562		674.3984	657.3719		6
9	30.0338	918.4356	900.4250	946.4305	928.4199	G				577.3457	560.3191		5
10	72.0808	1017.5040	999.4934	1045.4989	1027.4884	V	476.2616	489.2820		520.3242	503.2976		4
11	72.0808	1116.5724	1098.5619	1144.5673	1126.5568	V	377.1932	390.2136		421.2558	404.2292		3
12	120.0808	1263.6408	1245.6303	1291.6358	1273.6252	F	230.1248			322.1874	305.1608		2
13	129.1135					R	74.0237	73.0284		175.1190	158.0924		1

Experiment 3 (spot 35102)

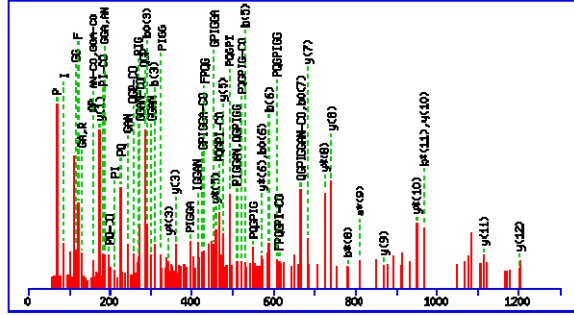


Monoisotopic mass of neutral peptide Mr(calc): 1464.7402
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 69 Expect: 1.8e-05
 Matches : 62/196 fragment ions using 61 most intense peaks ([help](#))

#	Immon.	a	a ⁰	b	b ⁰	Seq.	v	w	w'	y	y*	y ⁰	#
1	60.0444	60.0444	42.0338	88.0393	70.0287	S							13
2	120.0808	207.1128	189.1022	235.1077	217.0972	F	1286.6528			1378.7154	1361.6889	1360.7048	12
3	70.0651	304.1656	286.1550	332.1605	314.1499	P	1189.6000	1188.6048		1231.6470	1214.6204	1213.6364	11
4	88.0393	419.1925	401.1819	447.1874	429.1769	D	1074.5731	1073.5778		1134.5942	1117.5677	1116.5837	10
5	120.0808	566.2609	548.2504	594.2558	576.2453	F	927.5047			1019.5673	1002.5407	1001.5567	9
6	70.0651	663.3137	645.3031	691.3086	673.2980	P	830.4519	829.4567		872.4989	855.4723	854.4883	8
7	74.0600	764.3614	746.3508	792.3563	774.3457	T	729.4042	742.4246	744.4039	775.4461	758.4196	757.4355	7
8	70.0651	861.4141	843.4036	889.4090	871.3985	P	632.3515	631.3562		674.3984	657.3719		6
9	30.0338	918.4356	900.4250	946.4305	928.4199	G				577.3457	560.3191		5
10	72.0808	1017.5040	999.4934	1045.4989	1027.4884	V	476.2616	489.2820		520.3242	503.2976		4
11	72.0808	1116.5724	1098.5619	1144.5673	1126.5568	V	377.1932	390.2136		421.2558	404.2292		3
12	120.0808	1263.6408	1245.6303	1291.6358	1273.6252	F	230.1248			322.1874	305.1608		2
13	129.1135					R	74.0237	73.0284		175.1190	158.0924		1

Peptide View: Acyl-protein thioesterase 1, isoform 1 (LYPLAI)

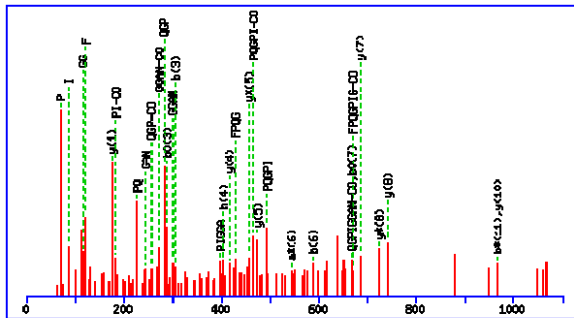
Experiment 1 (spot 18213)



Monoisotopic mass of neutral peptide Mr(calc): 1270.6418
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 49 Expect: 0.002
 Matches : 72/210 fragment ions using 89 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	Seq.	v	w	w'	y	y*	y ⁰	#
1	44.0495	44.0495			72.0444			A							13
2	60.0444	131.0815		113.0709	159.0764		141.0659	S	1168.5858	1167.5905		1200.6120	1183.5854	1182.6014	12
3	120.0808	278.1499		260.1394	306.1448		288.1343	F	1021.5174			1113.5800	1096.5534		11
4	70.0651	375.2027		357.1921	403.1976		385.1870	P	924.4646	923.4694		966.5116	949.4850		10
5	101.0709	503.2613	486.2347	485.2507	531.2562	514.2296	513.2456	Q	796.4060	795.4108		869.4588	852.4322		9
6	30.0338	560.2827	543.2562	542.2722	588.2776	571.2511	570.2671	G				741.4002	724.3737		8
7	70.0651	657.3355	640.3089	639.3249	685.3304	668.3039	667.3198	P	642.3318	641.3366		684.3787	667.3522		7
8	86.0964	770.4196	753.3930	752.4090	798.4145	781.3879	780.4039	I	529.2477	542.2681	556.2838	587.3260	570.2994		6
9	30.0338	827.4410	810.4145	809.4305	855.4359	838.4094	837.4254	G				474.2419	457.2154		5
10	30.0338	884.4625	867.4359	866.4519	912.4574	895.4308	894.4468	G				417.2205	400.1939		4
11	44.0495	955.4996	938.4730	937.4890	983.4945	966.4680	965.4839	A	344.1677			360.1990	343.1724		3
12	87.0553	1069.5425	1052.5160	1051.5320	1097.5374	1080.5109	1079.5269	N	230.1248	229.1295		289.1619	272.1353		2
13	129.1135							R	74.0237	73.0284		175.1190	158.0924		1

Experiment 2 (spot 29101)

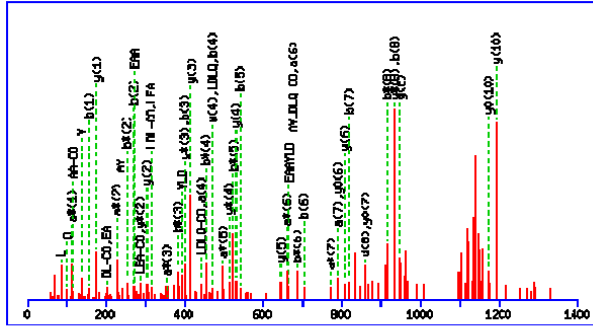


Monoisotopic mass of neutral peptide Mr(calc): 1270.6418
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 36 Expect: 0.041
 Matches : 38/210 fragment ions using 41 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	Seq.	v	w	w'	y	y*	y ⁰	#
1	44.0495	44.0495			72.0444			A							13
2	60.0444	131.0815		113.0709	159.0764		141.0659	S	1168.5858	1167.5905		1200.6120	1183.5854	1182.6014	12
3	120.0808	278.1499		260.1394	306.1448		288.1343	F	1021.5174			1113.5800	1096.5534		11
4	70.0651	375.2027		357.1921	403.1976		385.1870	P	924.4646	923.4694		966.5116	949.4850		10
5	101.0709	503.2613	486.2347	485.2507	531.2562	514.2296	513.2456	Q	796.4060	795.4108		869.4588	852.4322		9
6	30.0338	560.2827	543.2562	542.2722	588.2776	571.2511	570.2671	G				741.4002	724.3737		8
7	70.0651	657.3355	640.3089	639.3249	685.3304	668.3039	667.3198	P	642.3318	641.3366		684.3787	667.3522		7
8	86.0964	770.4196	753.3930	752.4090	798.4145	781.3879	780.4039	I	529.2477	542.2681	556.2838	587.3260	570.2994		6
9	30.0338	827.4410	810.4145	809.4305	855.4359	838.4094	837.4254	G				474.2419	457.2154		5
10	30.0338	884.4625	867.4359	866.4519	912.4574	895.4308	894.4468	G				417.2205	400.1939		4
11	44.0495	955.4996	938.4730	937.4890	983.4945	966.4680	965.4839	A	344.1677			360.1990	343.1724		3
12	87.0553	1069.5425	1052.5160	1051.5320	1097.5374	1080.5109	1079.5269	N	230.1248	229.1295		289.1619	272.1353		2
13	129.1135							R	74.0237	73.0284		175.1190	158.0924		1

Peptide View: Tubulin-specific chaperone A (TBCA)

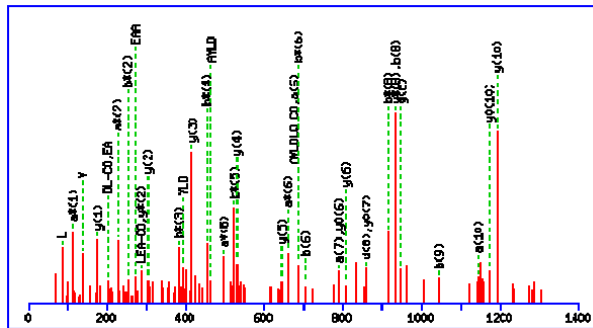
Experiment 1 (spot 14101)



Monoisotopic mass of neutral peptide Mr(calc): 1346.7306
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 55 Expect: 0.00047
 Matches : 60/177 fragment ions using 73 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	d	Seq.	v	w	y	y*	y ⁰	#
1	129.1135	129.1135	112.0869		157.1084	140.0818		44.0495	R						11
2	86.0964	242.1975	225.1710		270.1925	253.1659		200.1506	L	1133.5586	1132.5633	1191.6368	1174.6103	1173.6263	10
3	102.0550	371.2401	354.2136	353.2296	399.2350	382.2085	381.2245	313.2347	E	1004.5160	1003.5207	1078.5528	1061.5262	1060.5422	9
4	44.0495	442.2772	425.2507	424.2667	470.2722	453.2456	452.2616		A	933.4789		949.5102	932.4836	931.4996	8
5	44.0495	513.3144	496.2878	495.3038	541.3093	524.2827	523.2987		A	862.4417		878.4730	861.4465	860.4625	7
6	136.0757	676.3777	659.3511	658.3671	704.3726	687.3461	686.3620		Y	699.3784		807.4359	790.4094	789.4254	6
7	86.0964	789.4618	772.4352	771.4512	817.4567	800.4301	799.4461	747.4148	L	586.2944	585.2991	644.3726	627.3461	626.3620	5
8	88.0393	904.4887	887.4621	886.4781	932.4836	915.4571	914.4730	860.4989	D	471.2674	470.2722	531.2885	514.2620	513.2780	4
9	86.0964	1017.5728	1000.5462	999.5622	1045.5677	1028.5411	1027.5571	975.5258	L	358.1833	357.1881	416.2616	399.2350		3
10	101.0709	1145.6313	1128.6048	1127.6208	1173.6263	1156.5997	1155.6157	1088.6099	Q	230.1248	229.1295	303.1775	286.1510		2
11	129.1135								R	74.0237	73.0284	175.1190	158.0924		1

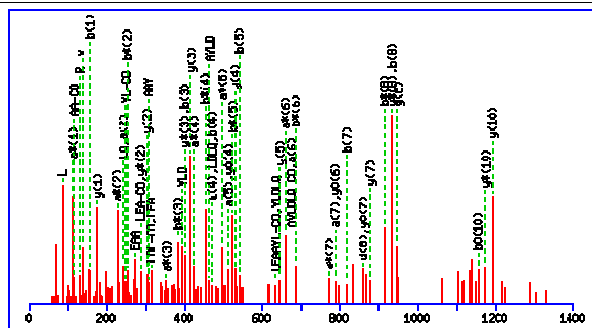
Experiment 2 (spot 24004)



Monoisotopic mass of neutral peptide Mr(calc): 1346.7306
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 50 Expect: 0.0015
 Matches : 43/177 fragment ions using 51 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	d	Seq.	v	w	y	y*	y ⁰	#
1	129.1135	129.1135	112.0869		157.1084	140.0818		44.0495	R						11
2	86.0964	242.1975	225.1710		270.1925	253.1659		200.1506	L	1133.5586	1132.5633	1191.6368	1174.6103	1173.6263	10
3	102.0550	371.2401	354.2136	353.2296	399.2350	382.2085	381.2245	313.2347	E	1004.5160	1003.5207	1078.5528	1061.5262	1060.5422	9
4	44.0495	442.2772	425.2507	424.2667	470.2722	453.2456	452.2616		A	933.4789		949.5102	932.4836	931.4996	8
5	44.0495	513.3144	496.2878	495.3038	541.3093	524.2827	523.2987		A	862.4417		878.4730	861.4465	860.4625	7
6	136.0757	676.3777	659.3511	658.3671	704.3726	687.3461	686.3620		Y	699.3784		807.4359	790.4094	789.4254	6
7	86.0964	789.4618	772.4352	771.4512	817.4567	800.4301	799.4461	747.4148	L	586.2944	585.2991	644.3726	627.3461	626.3620	5
8	88.0393	904.4887	887.4621	886.4781	932.4836	915.4571	914.4730	860.4989	D	471.2674	470.2722	531.2885	514.2620	513.2780	4
9	86.0964	1017.5728	1000.5462	999.5622	1045.5677	1028.5411	1027.5571	975.5258	L	358.1833	357.1881	416.2616	399.2350		3
10	101.0709	1145.6313	1128.6048	1127.6208	1173.6263	1156.5997	1155.6157	1088.6099	Q	230.1248	229.1295	303.1775	286.1510		2
11	129.1135								R	74.0237	73.0284	175.1190	158.0924		1

Experiment 3 (spot 34001)



Monoisotopic mass of neutral peptide Mr(calc): 1346.7306
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 52 Expect: 0.00098
 Matches : 62/177 fragment ions using 67 most intense peaks ([help](#))

#	Immon.	a	a*	a ⁰	b	b*	b ⁰	d	Seq.	v	w	y	y*	y ⁰	#
1	129.1135	129.1135	112.0869		157.1084	140.0818		44.0495	R						11
2	86.0964	242.1975	225.1710		270.1925	253.1659		200.1506	L	1133.5586	1132.5633	1191.6368	1174.6103	1173.6263	10
3	102.0550	371.2401	354.2136	353.2296	399.2350	382.2085	381.2245	313.2347	E	1004.5160	1003.5207	1078.5528	1061.5262	1060.5422	9
4	44.0495	442.2772	425.2507	424.2667	470.2722	453.2456	452.2616		A	933.4789		949.5102	932.4836	931.4996	8
5	44.0495	513.3144	496.2878	495.3038	541.3093	524.2827	523.2987		A	862.4417		878.4730	861.4465	860.4625	7
6	136.0757	676.3777	659.3511	658.3671	704.3726	687.3461	686.3620		Y	699.3784		807.4359	790.4094	789.4254	6
7	86.0964	789.4618	772.4352	771.4512	817.4567	800.4301	799.4461	747.4148	L	586.2944	585.2991	644.3726	627.3461	626.3620	5
8	88.0393	904.4887	887.4621	886.4781	932.4836	915.4571	914.4730	860.4989	D	471.2674	470.2722	531.2885	514.2620	513.2780	4
9	86.0964	1017.5728	1000.5462	999.5622	1045.5677	1028.5411	1027.5571	975.5258	L	358.1833	357.1881	416.2616	399.2350		3
10	101.0709	1145.6313	1128.6048	1127.6208	1173.6263	1156.5997	1155.6157	1088.6099	Q	230.1248	229.1295	303.1775	286.1510		2
11	129.1135								R	74.0237	73.0284	175.1190	158.0924		1