University of Alberta

Mass Spectrometric Method Development and Applications for Comprehensive Proteome Analysis

by

Xiaoxia Ye

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

Spring 2012

Edmonton, Alberta

To my parents for their dedication

To my husband for his encouragement and support

Abstract

In the field of proteomics research, it is desirable to detect the entire proteome or all the proteins present in a sample. However, due to the complexity of most samples, this task is challenging. My thesis work is mainly focused on the development of new protein solubilization and fractionation techniques to increase the identification efficiency in shotgun proteomic studies, namely to detect as many proteins as possible at high sample handling throughput. Several techniques have been developed or optimized in this thesis. First, protein level fractionation by using sequential protein precipitation and solubilization was effective in simplifying a complex sample and enhancing the proteome coverage. Second, microwave-assisted sequential protein solubilization (MAPS) was developed to speed up the protein solubilization process, increase protein solubility and protein digestion efficiency. As a result, by using MAPS combined with two dimensional-liquid chromatography (2D-LC) tandem mass spectrometry (MS/MS) analysis, peptide and protein identification efficiency was improved. Third, a faster and better resolution in protein fractionation was achieved by a macro-porous C18 reversed-phase liquid chromatography column (mRP-C18). The mRP-C18 fractionation method was then optimized and applied to the analysis of the phosphoproteome of MDA-MB-231 cells. Finally, sequential phosphopeptide enrichment by metal ion affinity chromatography (IMAC) and metal oxide affinity chromatography (TiO_2) combined with strong cation exchange (SCX)-reverse phase liquid chromatography (RPLC) MS/MS method was applied to the analysis of human breast cancer tissues. 297 phosphoproteins were found possibly related to the metastasis of breast tumor and 875

phosphoproteins were found possibly related to the genesis of the breast tumor. The techniques developed or optimized in my thesis work improved sample preparation and fractionation efficiency, and therefore enhanced the overall efficiency of proteome identification. The newly identified phosphoproteins in MDA-MB-231 cells and human breast cancer tissues may generate novel insight into breast cancer biology. These approaches also hold great potential for profiling a wider range of proteomes in a more comprehensive and efficient way.

Acknowledgements

It is a great pleasure to thank those who made this thesis possible.

First and foremost, I would like to express my sincere appreciation to my supervisor, Dr. Liang Li, for giving me the precious opportunity to study in his research group and for his invaluable guidance, inspiration, encouragement and advice throughout my Ph.D studies. I have learnt a lot from him about how to do interesting and rigorous research, and I am sure it will continue to benefit my future career.

I also would like to thank my committee members, Dr. Charles Lucy, Dr. John Vederas, Dr. Michael Serpe, Dr. Michael Schultz and Dr. Ken Yeung, for their precious time and contriution to the examination process and their valuale advice on my research and career. I am very grateful to Dr. Joe Weiner from the Department of Biochemistry at the University of Alberta for his contributions to the bioinformatics characterization of the proteins identified from *E. coli* and helpful discussions in this research area.

My sincere gratitude also goes to Dr. Nan Wang, for her professional training at the beginning of my Ph.D studies. Thank Dr. Fang Wu and Dr. Peng Wang for their helpful advices throughout my Ph.D studies. My appreciation also extends to all the other members in Dr. Liang Li's research group.

I would like to express special thanks to Dr. Randy Whittal and Jing Zheng from the Mass Spectrometry Facility in the Department of Chemistry at the University of Alberta for their professional training and help on the QTOF instrument maintenance and trouble-shooting. Thank Dr. Sandra Marcus and Gareth Lambkin for culturing cells. I also would like to thank all the other academic and non-academic staff in the Department of Chemistry who gave their kind help to me during my Ph.D study.

Finally, I would like to thank my parents, Mr. Defu Ye and Mrs. Weijun Cao. They bore me, raised me, supported me, taught me, and loved me. I must also thank my husband, Mr. Jiadong Wang, for sharing the ups and downs of the entire Ph.D. process with me, for his long term support, encouragement and love. To them I dedicate this thesis.

Table of Contents

Chapter 1 Introduction to Proteome Analysis by Mass Spectrometry..... 1

1.1 Research objective and scope of the thesis work 1

1.2 Overview of protein identification by mass spectrometry 3

1.2.1 Protein identification strategies by mass spectrometry..... 3

1.2.2 Overview of sample preparation methods for shotgun proteome analysis4

1.2.3 Protein and peptide fractionation methods 6

1.2.4 Tandem mass spectrometry and instrumentation..... 10

1.2.4.1 General introduction of mass spectrometry…….10

1.2.4.2 Quadrupole time-of-flight (Q-TOF) MS…….12

1.2.5 Peptide ion fragmentation..... 16

1.2.6 Protein identification by database searching based on MS and MS/MS 18

1.3 Analysis of phosphoproteome 19

1.3.1 Enrichment strategies of phosphoproteome..... 19

1.3.1.1 Immobilized metal affinity chromatography (IMAC)…….20

1.3.1.2 Metal oxide affinity chromatography (MOAC)……..20

1.3.1.3 Strong cation exchange chromatography (SCX)…….23

1.3.1.4 Immunoaffinity chromatography…….23

1.3.1.5 Chemical transformation methods……...25

1.3.2 Analysis of phosphopeptides by mass spectrometry..... 24

1.4 References 25

Chapter 2 Comprehensive Analysis of MCF-7 Membrane Proteome by Sequential Protein Precipitation and Solubilization Combined with 2D-LC MS/MS 33

2.1 Introduction 33

2.2 Experimental 34

2.2.1 Chemicals and reagents 34
2.2.2 Membrane protein extraction 34
2.2.3 Sequential protein precipitation 35
2.2.4 Sequential protein solubilization 35
2.2.5 In-solution digestion 37
2.2.6 Strong cation exchange chromatography 37
2.2.7 Peptide desalting and quantification by RPLC 37
2.2.8 LC-ESI QTOF MS and MS/MS analysis 38
2.2.9 Protein database search 38
2.2.10 Hydropathy calculation and annotation of localization 39
2.3 Results and discussion 39
2.3.1 Protein and peptide identification 39
2.3.2 Sequential precipitation and solubilization validation 40
2.3.3 Protein level separation characterization 46
2.3.4 Subcellular location 49
2.3.5 Comparison to other work 54
2.4 Conclusions 55
2.5 References 56

Chapter 3 Comprehensive Proteome Profiling of *E. coli* K12 Cell Line by Shotgun Proteomic Strategy* 60
3.1 Introduction 60
3.2 Experimental 61
3.2.1 Chemicals and reagents 61
3.2.2 Cell culture and protein extraction 62
3.2.3 Sequential protein precipitation 62
3.2.4 Sequential protein solubilization 63

3.2.5 In-solution digestion 64
3.2.6 Strong cation exchange (SCX) liquid chromatography 64
3.2.7 Peptide desalting and quantification by RPLC 65
3.2.8 Mass spectrometric analysis 65
3.2.9 Protein database search 66
3.2.10 Hydropathy calculation and annotation of localization 66
3.3 Results and discussion 66
3.3.1 Protein identification results 67
3.3.2 Properties of identified proteins 74
3.4 Conclusions 81
3.5 References 82

Chapter 4 Microwave-assisted Protein Solubilization for Mass Spectrometry-based Shotgun Proteome Analysis 84
4.1 Introduction 84
4.2 Experimental 85
4.2.1 Chemicals and reagents 85
4.2.2 Cell culture and membrane preparation 85
4.2.3 Sequential microwave-assisted protein solubilization (MAPS) and vortex-assisted protein solubilization (VAPS). 86
4.2.4 In-solution digestion 87
4.2.5 Cation exchange chromatography 87
4.2.6 Peptide desalting and quantification by RPLC 87
4.2.7 LC-ESI QTOF MS and MS/MS analysis 88
4.2.8 Protein database search 88
4.2.9 Hydropathy calculation and annotation of localization 89
4.3 Results and discussion 89

4.3.1 Standard test 89
4.3.2 Solubilization time and solubility 90
4.3.3 BSA tryptic digestion 95
4.3.4 Sequential solubilization 97
4.4 Conclusions 114
4.5 References 115

Chapter 5 Microwave-assisted Protein Solubilization Combined with Three-dimensional Liquid Chromatography MS/MS for Improving the Efficiency of Shotgun Proteome Analysis 118
5.1 Introduction 118
5.2 Experimental 119
5.2.1 Chemicals and Reagents 119
5.2.2 Cell culture and protein extraction 119
5.2.3 Protein quantification and purification by acetone 120
5.2.4 Protein solubilization by SDS and microwave-assisted protein solubilization (MAPS) in urea 120
5.2.4 Macro-porous reversed-phase C18 (mRP-C18) separation 122
5.2.5 In-solution digestion 122
5.2.6 Strong cation exchange (SCX) liquid chromatography 122
5.2.7 Peptide desalting and quantification by RPLC 123
5.2.8 Mass spectrometric analysis 123
5.2.9 Protein database search 124
5.2.10 Hydropathy Calculation and Annotation of Localization 124
5.3 Results and discussion 125
5.3.1 Carbamylation rate 125
5.3.2 Protein and peptide identification 125

5.3.3 Characterization of identified proteins and peptides 129
5.3.4 Resolution of the mRP-C18 separation 131
5.3.5 Characterization of the precipitation induced by acidification before mRP-C18 separation 131
5.4 Conclusions 134
5.5 References 136

Chapter 6 Multidimensional mRP-RPLC Separation of Proteins and Peptides Combined with ESI-MS/MS for Comprehensive Profiling of the Phosphoproteome of MDA-MB-231 cells 138
6.1 Introduction 138
6.2 Experimental 139
6.2.1 Chemicals and reagents 139
6.2.2 Cell culture and protein extraction 139
6.2.3 In-solution digestion 142
6.2.4 Sequential phosphopeptides enrichment by immobilized metal ion affinity chromatography (IMAC) and titanium dioxide 142
6.2.5 Strong cation exchange (SCX) liquid chromatography 143
6.2.6 mRP-C18 separation，in-solution digestion and sequential phosphopeptide enrichment by IMAC and TiO_2 beads 143
6.2.7 Mass spectrometric analysis 144
6.2.8 Database search and data analysis 144
6.3 Results and discussion 145
6.3.1 Identification of phosphopeptides 145
6.3.2 Sequential phosphopeptide enrichment 146
6.3.3 Characterization of the SCX-RPLC and mRP-RPLC methods 150
6.3.3 Bioinformatic analysis 152

6.4 Conclusions 156
6.5 References 185

Chapter 7 Large-scale Phosphoproteome Profiling of Human Breast Cancer Tissues ... 194
7.1 Introduction 194
7.2 Experimental 195
7.2.1 Chemicals and reagents 195
7.2.2 Protein extraction from uuman breast cancer tissue samples 196
7.2.3 In-solution digestion 196
7.2.4 Sequential phosphopeptides enrichment by immobilized metal ion affinity chromatography (IMAC) and titanium dioxide 197
7.2.5 Strong cation exchange (SCX) liquid chromatography 197
7.2.6 Mass spectrometric analysis 198
7.2.7 Database search and data analysis 198
7.3 Results and discussion 199
7.4 Conclusions 204
7.5 References 289

Chapter 8 Conclusions and Future Work 294
8.1 References 298

List of Tables

Table 1.1 Summary of the properties of LC methods commonly used for protein and peptide separation. 7

Table 1.2 List of PMF search engines. 18

Table 1.3 List of MS/MS search engines. 19

Table 2.1 Summary of the protein amounts, the numbers of SCX fractions and the numbers of proteins and peptides identified from the five protein fractions generated by sequential protein precipitation and solubilization. 40

Table 2.2 List of previously published MCF-7 profiling works. 54

Table 3.1 Summary of the 13 protein fractions generated from the three cellular compartments and subsequent sequential protein precipitation and solubilization. 64

Table 4.1 BSA test results: number of identified peptides and sequence coverage by different solubilization methods. (A) BSA solubilized in 25 mM NH_4HCO_3. (B) BSA solubilized in 8 M urea. 98

Table 4.2 Modification effects on the identification number of unique peptides. 108

Table 5.1 Comparison of search results by different searching parameters. 125

Table 6.1 Properties of phosphopeptides enriched by the IMAC and TiO_2 beads. 157

Table 6.2 List of previously reported phosphoproteins found to be functionally involved in the breast cancer that were identified in this work. 157

Table 6.3 List of new phosphoproteins found in this work. 170

Table 7.1 Clinical information of the patient samples. 196

Table 7.2 Summary of phosphoproteins and phosphopeptides identified from the human breast tissue samples. 201

Table 7.3 Summary of the 875 phosphoproteins possibly related to the genesis of breast cancer identified in this work. 206

Table 7.4 Summary of the 297 phosphoproteins possibly related to the genesis of breast cancer identified in this work. 268

Table 7.5 List of phosphoproteins previously reported to be associated with tumor genesis……....284

Table 7.6 List of phosphoproteins previously reported to be associated with tumor invasiveness……..……...……..287

List of Figures

Figure 1.1 Structures of RapiGest and PPS. 5

Figure 1.2 Schematic diagram of the (A) on-line and (B) off-line LC strategy. 9

Figure 1.3 Schematic diagram of electrospray ionization (ESI). 11

Figure 1.4 Schematic diagram of matrix-assisted laser desorption ionization (MALDI). 13

Figure 1.5 Schematic diagram of the Waters Quadrupole-time-of-flight mass spectrometric instrument (Q-TOF). 14

Figure 1.6 Cleavage patterns of the precursor ion. 17

Figure 1.7 Structures of (A) phosphoserine. (B) phosphothreonine. (C) phosphotyrosine. 21

Figure 1.8 Structures of metal ion chelating stationary phase in IMAC (A) iminodiacetic acid (IDA). (B) nitrilotriacetic acid (NTA). 22

Figure 2.1 Work flow for sample preparation, separation and analysis. 36

Figure 2.2 (A) Venn diagram of the proteins identified from reduction/alkylation fraction. A total number of 3812 proteins were identified: 3153 proteins from the NH_4HCO_3-assisted solubilization fraction, 2717 proteins from the SDS-assisted solubilization fraction. (B) Venn diagram of the peptides identified from reduction/alkylation fraction. A total number of 16476 peptides were identified: 11369 peptides from the NH_4HCO_3-assisted solubilization fraction, 7820 peptides from the SDS-assisted solubilization fraction. 41

Figure 2.3 (A) Venn diagram of the proteins identified from acetone fraction. A total number of 3784 proteins were identified: 2943 proteins from the NH_4HCO_3-assisted solubilization fraction, 1041 proteins from the

methanol-assisted fraction, 2391 proteins from the SDS-assisted solubilization fraction. (B) Venn diagram of the peptides identified from acetone fraction. A total number of 14674 peptides were identified: 8437 peptides from the NH_4HCO_3-assisted solubilization fraction, 2677 peptides from the methanol-assisted fraction, 8175 peptides from the SDS-assisted solubilization fraction. 42

Figure 2.4 (A) Venn diagram of the proteins identified from two sequential precipitation techniques. A total number of 5011 proteins were identified, with the overlap of 51% between two seuqential precipitation fractions. (B) Venn diagram of the peptides identified from two sequential precipitation techniques. A total number of 24261 peptides were identified, with the overlap of 28% between two sequential precipitation fractions. 43

Figure 2.5 (A) Venn diagram of the proteins identified from sequential solubilization techniques. MeOH and SDS-assisted solubilization fractions were combined together. The overlap between two sequential solubilization fractions was 55%. (B) Venn diagram of the peptides identified from sequential solubilization techniques. MeOH and SDS-assisted solubilization fractions were combined together. The overlap between two sequential solubilization fractions was 21%. 45

Figure 2.6 Comparison of cystein residues in one protein from two sequential precipitation techniques. 47

Figure 2.7 Comparison of molecular mass (MW) of identified proteins from two sequential precipitation techniques. 48

Figure 2.8 Distribution of GRAVY (Grand Average of Hydropathy) indices of identified proteins. 50

Figure 2.9 Subcellular localizations of the MCF-7 membrane extract. 52

Figure 2.10 Transmembrane domain (TMD) distribution of identified integral membrane

proteins and membrane proteins. 53

Figure 3.1 Workflow of cellular compartments fractionation for *E. coli* K12 cell line. . 68

Figure 3.2 Workflow of sequential protein precipitation and solubilization followed by 2D-LC MS/MS analysis for *E. coli* K12 cell line. 69

Figure 3.3 (A) Venn diagram of the 2595 unique proteins identified from the cytoplasm_reduction/alkylation fraction. (B) Venn diagram of the 2276 unique proteins identified from the cytoplasm_acetone fraction, with a total number of 2913 protein identified from the cytoplasm. 70

Figure 3.4 (A) Venn diagram of the 2069 unique proteins identified from the peripheral membrane extract_reduction/alkylation fraction. (B) Venn diagram of the 1961 unique proteins identified from the peripheral membrane extract_acetone fraction, with a total number of 2579 protein identified from the cytoplasm. 71

Figure 3.5 Venn diagram of the 1684 unique proteins identified from the integral membrane extract. 72

Figure 3.6 Venn diagram of the 3325 unique proteins identified from the *E. coli* K12 cell line. 73

Figure 3.7 (A) GRAVY and (B) TMD distribution of identified proteins. 75

Figure 3.8 (A) molecular weight distribution of identified proteins. (B) the percentage of proteins identified comparing to the genome as a function of molecular weight. 77

Figure 3.9 Workflow of enrichment of low molecular weight proteins by molecular weight cut off filters. 78

Figure 3.10 Venn diagram of the unique proteins identified from the low molecular weight fractions. 2241 unique proteins were identified. 79

Figure 3.11 (A) Distribution of the percentage of unique proteins identified in the low

molecular weight fractions. (B) Molecular weight distribution of all the identified proteins and the genome-predicted proteins. 80

Figure 4.1 (A) SDS-PAGE images of acetone-precipitated bacteriorhodopsin (BR) dissolved in 2% SDS (lane 0), 25 mM NH_4HCO_3 with vortex (lane 1) or microwave (lane 2), 60% methanol with vortex (lane 3) or microwave (lane 4), and 8 M urea with vortex (lane 5) or microwave (lane 6). (B) SDS-PAGE images of the *E. coli* integral membrane protein extract solubilized in 2% SDS (lane 0), 25 mM NH_4HCO_3 with microwave (lane 1), and 25 mM NH_4HCO_3 with vortex (lane 3). The remaining protein pellets from both cases are dissolved in 8 M urea with microwave (lane 2) and vortex (lane 4), respectively. After solubilization by 8 M urea with microwave, the remaining proteins were solubilized completely in 2% SDS (lane 5). After solubilization by 8 M urea with vortex, the remaining proteins were also completely solubilized in 2% SDS (lane 6). 91

Figure 4.2 The concentration of dissolved proteins from 650 μg of the *E. coli* integral. 92

Figure 4.3 The amount of dissolved proteins from different rounds of solubilization in 8M urea with (A) vortex and (B) MAPS. (C) comparison of the vortex and MAPS results. 94

Figure 4.4 The amount of dissolved proteins from different rounds of solubilization in 25 mM NH_4HCO_3 with (A) vortex and (B) MAPS (C) comparison of the vortex and MAPS results..... 96

Figure 4.5 Venn diagram of the proteins identified from the two methods (replicate one): (A) vortex and (B) MAPS. 100

Figure 4.6 Venn diagram of the proteins identified from the two methods (replicate two): (A) vortex and (B) MAPS. 101

Figure 4.7 Venn diagram of the identified proteins by the MAPS and vortex methods: (A)

replicate one. (B) replicate two. (C) merged data from the two replicates. 102

Figure 4.8 Venn diagram of the peptides identified from the two methods (replicate one): (A) vortex and (B) MAPS. 104

Figure 4.9 Venn diagram of the peptides identified from the two methods (replicate two): (A) vortex and (B) MAPS. 105

Figure 4.10 Venn diagram of the identified peptides by the MAPS and vortex methods: (A) replicate one. (B) replicate two. (C) merged data from the two replicates. 106

Figure 4.11 Comparison of identified peptides from MAPS and VAPS (replicate one) (A) NH_4HCO_3-solubilized fraction. (B) urea-solubilized fraction. (C) SDS-solubilized fraction. 109

Figure 4.12 Comparison of identified peptides from MAPS and VAPS (replicate two) (A) NH_4HCO_3-solubilized fraction. (B) urea-solubilized fraction. (C) SDS-solubilized fraction. 110

Figure 4.13 (A) Molecular weight distribution of the proteins identified by the vortex and MAPS methods. (B) Protein distribution as a function of protein hydrophobicity, gauged by the GRAVY indices grouped into four bins..... 112

Figure 4.14 (A) Peptide distribution as a function of peptide hydrophobicity, gauged by the GRAVY indices grouped into eight bins. (B) Distribution of the number of transmembrane domains (TMDs) of identified proteins by the vortex and microwave methods. 113

Figure 5.1 Workflow for sample preparation, separation and analysis. 121

Figure 5.2 (A) Venn diagram of the proteins identified by 2D- and 3D-LC methods. A total number of 2333 unique proteins were identified. (B) Venn diagram of the peptides identified by 2D- and 3D-LC methods. 17158 different peptides were identified in total. (C) Venn diagram of the membrane proteins identified by

2D- and 3D-LC methods. 528 unique membrane proteins were identified in total. 127

Figure 5.3 (A) The molecular weight distribution of proteins uniquely identified by 2D- and 3D-LC methods. (B) Protein GRAVY distribution of identified proteins by the 2D-and 3D-LC methods and in Chapter 3. (C) Peptide GRAVY distribution of identified peptides by the 2D- and 3D-LC methods. 128

Figure 5.4 (A) Venn diagram of the membrane proteins identified by 2D and 3D-LC methods. A total number of 528 unique membrane proteins were identified. (B) Transmembrane domain (TMD) distribution of identified proteins by 2D and 3D-LC methods. The percentage is calculated based on the number of proteins with at least one TMD. 130

Figure 5.5 Chromatogram generated from the separation of the *E.coli* whole cell lysate by mRP-C18 column. Flow rate: 0.75 mL/min; Detection wavelength: 214 nm. 132

Figure 5.6 Distribution of number of identified proteins' appearance times in different mRP-C18 fractions by the 3D-LC method. 133

Figure 5.7 Venn diagram of the unique (A) peptides and (B) proteins identified from the eluate of mRPLC and the precipitate before mRPLC. 92.4% of peptides and 98.1% of identifiable proteins were found in the eluate from mRPLC. 135

Figure 6.1 Workflow of the SCX-RPLC methods. 140

Figure 6.2 Workflow of the mRP-RPLC methods. 141

Figure 6.3 (A) Venn diagram of the phosphopeptides identified by the replicate experiments of the SCX-RPLC method. (B) Venn diagram of the phosphoproteins identified by the replicate experiments of the SCX-RPLC method. A total of 4297 different phosphopeptides and 1585 unique phosphoproteins were identified. 147

Figure 6.4 (A) Venn diagram of the phosphopeptides identified by the replicate experiments of the mRP-RPLC method. (B) Venn diagram of the phosphoproteins identified by the replicate experiments of the mRP-RPLC method. A total of 4519 different phosphopeptides and 1585 unique phosphoproteins were identified. 148

Figure 6.5 (A) Venn diagram of the phosphopeptides identified by the SCX-RPLC and mRP-RPLC methods. (B) Venn diagram of the phosphoproteins identified by the SCX-RPLC and mRP-RPLC methods. A total of 6278 different phosphopeptides and 1947 unique phosphoproteins were identified. 149

Figure 6.6 Distribution of phosphopeptides from SCX and mRPLC. 151

Figure 6.7 (A) Molecular weight distribution of unique protein of the SCX-RPLC and mRP-RPLC methods. (B) Peptide GRAVY distribution of the SCX-RPLC and mRP-RPLC methods. 153

Figure 6.8 Functional pathway analysis of the MDA-MB-231 phosphoproteins. 154

Figure 6.9 New phosphoproteins identified by the SCX-RPLC and mRP-RPLC method. 155

Figure 7.1 Workflow of analyzing breast cancer tissue samples. 200

Figure 7.2 Functional pathway analysis of the tumor genesis related phosphoproteins. 202

Figure 7.3 Functional pathway analysis of the tumor invasiveness related phosphoproteins. 203

Figure 7.4 Electrospray ionization tandem mass spectrometry (ESI MS/MS) spectrum of a phosphoserine peptide generated from phosphorylated 14-3-3 protein epsilon (phospho-Serine^{210})..... 205

List of Abbreviations

µ	micro- (10^{-6})
2D-LC	Two dimensional liquid chromatography
Å	Angstrom- (10^{-10})
ACN	Acetonitrile
BSA	Bovin serum albumin
CAD	Collision-activated dissociation
CE	Capillary electrophoresis
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CID	Collision-induced dissociation
CNBr	Cyanogen bromide
DC	Direct current
DHB	2,5-dihydroxyenzoic acid
DMEM	Dulbecco's modified eagle's medium
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ECD	Elelctron capture dissociation
ERLIC	Electrostatic repulsion-hydrophilic interaction chromatography
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FBS	Fetal bovine serum
FTICR	Fourier-transform ion cyclotron resonance
GO	Gene ontology
GRAVY	Grand average of hydrophathy
HCD	High energy C-trap dissociation
HILIC	Hydrophilic interaction liquid chromatography

HPLC	High performance liquid chromatography
IAA	Iodoacetamide
IE	Ion exchange
IEF	Isoelectric focusing
IMAC	Immobilized metal ion affinity chromatography
LIT	Linear ion trap
LTQ	Linear trap quadrupole
m	Milli- (10^{-3})
m/z	Mass to charge
MAAH	Microwave-assisted acid hydrolysis
MALDI	Matrix-assisted laser desorption/ionization
MAPS	Microwave-assisted protein solubilization
MeOH	Methanol
MOAC	Metal oxide affinity chromatography
mRP-C18	Macro-porous C18 reversed-phase
MS	Mass spectrometry
MW	Molecular weight
n	Nano- (10^{-9})
NP-40	Nonidet P40-substitute
NTA	Nitrilotriacetic acid
PBS	Phosphate buffered saline
pI	Isoelectric point
PMSF	Phenylmethyl sulfonyl fluoride
PPS	Sodium 3-(4-(1,1-bis(hexyloxy)ethyl)pyridinium-1-yl)propane-1-sulfonate
PTM	Post-translational modification
QTOF	Quadrupole time-of-flight

RCF	Relative centrifugal force (×g)
RF	Radio frequency
RP	Reversed-phase
SA	Sinapinic acid
SAX	Strong anion exchange
SCX	Strong cation exchange
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TFA	Trifluoroacetic acid
TMD	Transmembrane domain
TOF	Time-of-flight
Tris	Tris(hydroxymethyl) aminomethane
Triton X-100	t-Octylphenoxypolyethoxyethanol
UV	Ultraviolet
VAPS	Vortex-assisted protein solubilization
WAX	Weak anion exchange
WCX	Weak cation exchange

Chapter 1

Introduction to Proteome Analysis by Mass Spectrometry

1.1 Research objective and scope of the thesis work

As a relatively new research field, proteomics attracts increasing interests in many research areas including medicine and biology. Unlike traditional techniques focusing on the study of functions and properties of individual proteins, proteomics allows characterization of the comprehensive proteome with high throughput. Recently, rapid progresses have been made in instrumental technologies employed for proteomics.[1-6] However, development of sample preparation and fractionation methods for handling a complex proteome sample is still lagging behind.

One challenge is related to protein solubilization. During sample preparation, interfering compounds in the cell lysate are removed by protein precipitation using an organic solvent, such as acetone. After this, protein solubilization is needed to disrupt the interactions between proteins. Such solubilization is usually done by salt buffers (e.g., ammonium bicarbonate), chaotropes (e.g., urea) or detergents (e.g., sodium dodecyl sulfate or SDS) with the assistance of vortexing. Protein solubilization can also be facilitated by sonication[7],or occasionally, heating.[8] Considering the compatibility with the downstream work, usually involved in mass spectrometry (MS) analysis, salt buffers and chaotropes are often used since they can be easily removed from the sample prior to MS analysis. However, the solubilization process is time-consuming, usually takes several hours, and some hydrophobic proteins can not be solubilized completely, resulting in sample loss.

Another challenge is related to protein and peptide separation. A solubilized protein sample generally comes in the form of a complex mixture. The relative abundance of proteins could be dramatically different. The presence of high abundant proteins suppresses the detection of low abundant proteins. Therefore, sample fractionation on either the protein or peptide level or both is needed. This step can be performed by two-dimensional (2D) gel electrophoresis (2D-gel or a variety of liquid chromatography (LC) techniques.[9-14] As a traditional fractionation technique for proteome analysis by mass spectrometry, 2D-gel has several fundamental disadvantages in the analysis of biological samples. It is challenging for 2D-gel to detect low abundant proteins,[15] membrane proteins[16] and proteins with very basic isoelectric point (pI) and extreme molecular weight (MW).[17] Moreover, extracting peptides from the sample spots is labor intensive and may introduce sample loss. An alternative fractionation technique suitable

to be coupled with mass spectrometry is high-performance liquid chromatography (HPLC). HPLC is normally performed in a two-dimensional (2D) manner.[14, 18, 19] Combining two separation mechanisms, complex samples can be simplified based on the orthogonality and peak capacities of the two individual dimensions. Nevertheless, in mass spectrometry analysis, it is never easy to avoid ion suppression problem, which means complex samples to be analyzed are always underrepresented in terms of the number of proteins identified. Therefore, in some cases, three-dimensional (3D) fractionation is necessary for the analysis of complex samples.[20] On the other hand, since different LC methods vary in their separation mechanisms, even for the same starting material, after being fractionated by different combinations of 2D-LC methods, the composition of proteins/peptides in individual fractions can be dramatically different. As a result, proteins/peptides identified by different 2D-LC methods could be complementary to each other. Thus, it should be useful to investigate the effect of novel combinations of LC methods. on proteome coverage, i.e., the total number of proteins identifiable by MS.

To overcome some of the limitations of traditional methods or simply evaluate a novel protocol alternative to the conventional techniques, my thesis work mainly focuses on the development of protein solubilization and fractionation methods in shotgun proteomic analysis and the application of these techniques to breast cancer research. The major goal is to detect as many proteins as possible from a biological sample (e.g., cells and tissues) with high sample handling throughput. This work should form a strong foundation for further technique development in large-scale quantitative proteome analysis, protein-protein interaction investigation and the analysis of post-translational modifications (PMTs). In Chapter 2 of my thesis, I will describe a protein fractionation method using sequential solubilization and precipitation for analyzing the breast cancer cell line, MCF-7. Chapter 3 describes the results from the comprehensive mapping of the *E. coli* proteome. Chapter 4 reports the development of the microwave-assisted protein solubilization (MAPS) method for comprehensive analysis of the *E. coli* integral membrane extract. In Chapter 5, the MAPS method is combined with the protein-level separation of the proteome using a macro-porous C18 reversed-phase LC (mRP-C18) This new protocol is applied to profile the *E. coli* proteome. Chapters 6 and 7 describe the applications of the mRP-C18 LC and the traditional strong-cation exchange (SCX) LC method for the comprehensive analysis of the phosphoproteome of the breast cancer cell line, MDA-MB-231, and human breast cancer tissue samples, respectively. In Chapter 8, I will briefly draw some conclusions regarding the above work and comment on the future work related to my thesis.

1.2 Overview of protein identification by mass spectrometry

1.2.1 Protein identification strategies by mass spectrometry

There are two widely use strategies for proteome analysis by mass spectrometry: bottom-up (or shotgun) and top-down. The shotgun strategy is the most popular approach for the analysis of highly complex samples because of its high throughput. In the shotgun strategy, proteins are first enzymatically or chemically digested into peptides.[21] Before introduced into mass spectrometry, proteins or peptides are usually subjected to fractionation, either being done by gel separation[22] or by liquid chromatography.[19] The simplified peptide mixtures are then subjected to tandem mass spectrometric analysis. The fragment ion spectra showing different patterns of fragment ion peaks are collected, and searched against a proteome database for peptide and protein identification. The major limitation of the shotgun approach is that not all peptides digested from proteins can be detected by mass spectrometry, especially those of low abundance or with modifications.

As a recently developed approach, "top-down" allows the gas phase ionization of intact proteins without digestion and the fragment ion spectra of the intact protein ions are collected in a high resolution mass spectrometer.[23] Through the accurate measurement of the m/z values of the fragment ions, the amino acid sequence of a protein is determined. One of the advantages of the top-down strategy is that it can examine the intact protein sequence, which provides a more reliable way of protein identification, compared to sequencing a short peptide in the shotgun method.[24] Another advantage is that, because the method examines the entire sequence, it generates better results in the characterization of protein isoforms and post-translational modifications. However, this approach is still not widely used in proteome analysis due to the following limitations. Firstly, the analysis of large proteins is not always possible because of the low fragmentation efficiency for large species. And large proteins are not always soluble in MS compatible solutions. Secondly, the top-down strategy is only suitable to work with a single protein or simple protein mixture. Separation with superior resolution is required for the analysis of complex samples. However, high efficiency separation of proteins is much more difficult than peptides due to protein degradation, protein-protein interactions, different comformers and post translational modifications of proteins. Therefore, this method is still under development for efficient proteome analysis with high throughput and large proteome coverage.

In my thesis work, the shotgun approach was used for protein identification.

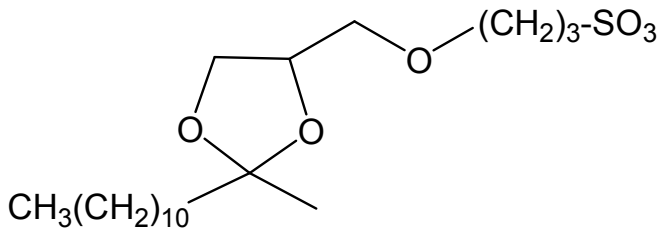
1.2.2 Overview of sample preparation methods for shotgun proteome analysis

In shotgun proteome analysis, several steps are involved in sample preparation. There are many different types of proteomic samples, such as cultured cell lines, tissue samples, body fluids, etc. Prior to further analysis, proteins need to be extracted and purified from the crude samples. The most commonly adopted method for protein purification is protein precipitation by acetone or trichloroacetic acid (TCA). By doing precipitation, impurities, such as salts and detergents can be removed from the protein sample.

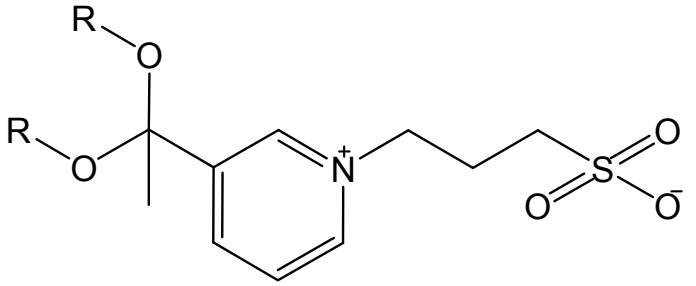
Next, the protein precipitates are solubilized in solvents. Complete solubilization of proteins is a key step in shotgun proteome analysis. Many research efforts have been devoted to the development of efficient and complete protein solubilization methods by MS-compatible solvents, especially for hydrophobic proteins.[25] Buffer solutions, such as NH_4HCO_3, organic solvents, such as chloroform and methanol,[26] and chaotropic agents, such as urea, have all been employed to dissolve proteins. However, protein solubilization by these reagents may take a long time and not all the hydrophobic proteins can be dissolved. Sodium dodecyl sulfate (SDS) is one of the most powerful detergents utilized to solubilize hydrophobic proteins. However, it severely affects enzyme activity, therefore decreasing digestion efficiency. More importantly, it is not compatible with downstream mass spectrometric analysis. Nowadays, some commercially available detergents have been developed to solubilize proteins with high efficiency, such as RapiGest[27] and PPS[28] (Figure 1.1). Under the acidic conditions, they are degraded into products without surfactant properties. Therefore, they are more suitable for the analysis of hydrophobic proteins. Protein solubilization can also be improved by other physical assistances, such as sonication and heating. However, in order to prevent possible protein degradation, it is not desirable to leave protein samples at an elevated temperature for a long time.

After solubilization, proteins are subjected to digestion. This step can be achieved by enzymatic (e.g., trypsin) or chemical reaction (e.g., CNBr). During digestion, it is desirable to keep solubilized proteins in their denatured forms so that many potential cleavage sites can contact with the enzyme to facilitate the digestion process and eliminate the possibility of miss-cleavage. As a result of complete digestion, more comprehensive and reliable mass spectra can be obtained in the downstream workflow. Methanol[29] and detergents[30] can be used to increase protein digestion efficiency.

In Chapter 4, a microwave-assisted protein solubilization (MAPS) method is described to increase the protein solubility in MS- compatible reagents (i.e., NH_4HCO_3 and urea).



RapiGest
Sodium 3-[(2-Methyl-2-undecyl-1,3-dioxolan-4 yl)-Methoxyl]-1-Propanesulfonate



1a: R= -$(CH_2)_5CH_3$
1b: R= -$(CH_2)_7CH_3$

PPS
sodium 3-(4-(1,1-bis(hexyloxy)ethyl)pyridinium-1-yl)propane-1-sulfonate

Figure 1.1 Structures of RapiGest and PPS.

Under microwave irradiation, proteins are more likely to be irreversibly denatured during the solubilization process. As a result, compared to the conventional vortexing-assisted solubilization, protein solubility can be done within a much shorter processing time by using MAPS. The efficiencies of protein digestion and proteome identification are also enhanced.

1.2.3 Protein and peptide fractionation methods

Complex proteome samples usually contain thousands of proteins. Since mass spectrometry itself does not have sufficient resolving power to analyze proteins and peptides with similar masses and various concentrations at once, separation of a complex sample prior to mass spectrometric analysis is required.

There are many methods available for the separation of proteins and peptides. For protein separation, the most conventional technique is two-dimensional gel electrophoresis (2D-gel).[2] Proteins are first separated by isoelectric focusing (IEF) according to their isoelectric points (IP), and then by SDS-polyacrylamide gel electrophoresis (PAGE) based on their molecular weight or size. This technique provides two-dimensional protein level separations with high orthogonality. However, it is difficult for 2D-gel to detect hydrophobic proteins as proteins tend to precipitate when pH is close to their pI during IEF separation. Besides 2D-gel, several other techniques are developed for protein separation, such as size exclusion chromatography (SEC) and affinity chromatography. However, size exclusion suffers from low protein recovery and separation resolution. Non-specific binding, limited lifetime, low productivity and high costs limit the employment of affinity chromatography for large-scale proteome analysis.

In the case of peptides, they are relatively small in sizes and have less interaction among themselves. Therefore, more varieties of separation techniques based on liquid chromatography (LC) are available for peptides. LC separations are based on the interactions between stationary phase and the species to be separated. There are different kinds of interactions, such as ionic interactions, hydrophobic interactions, hydrophilic interactions, or the mix of two modes, determined by the functional groups on the stationary phase and the analytes. The most widely used separation method for peptides is reversed-phase (RP), where the stationary phase contains non-polar silane chains (e.g., C_8, C_{18}) as the functional groups, and peptides are retained by hydrophobic forces. Increasing percentage of an organic solvent disrupts the interaction between the peptides and the stationary phase. Thus peptides are eluted out in the order of increasing hydrophobicity. Strong cation exchange (SCX) is also a popular separation technique for peptides. In SCX, the functional group of stationary phase is anionic, such as the $-SO_3^-$ group. At an acidic

condition (i.e., pH 2.7-3.0), most peptides are positively charged and thus they are retained on the column by ionic forces. As a salt (e.g., KCl) with a gradually increasing concentration is introduced onto the column, peptides are eluted out in order of their increasing charge states. Sometimes organic solvents (e.g., acetonitrile) are added into the mobile phase to minimize hydrophobic interactions between the peptides and the stationary phase. Other less commonly used separation methods include hydrophilic interaction chromatography (HILIC),[31] strong anion exchange (SAX),[32] and electrostatic repulsion-hydrophilic interaction chromatography (ERLIC).[33] Table 1.1 lists the major separation methods used in proteome analysis for separating proteins and peptides.

Table 1.1 Summary of the properties of LC methods commonly used for protein and peptide separation.

Separation Method	Separation Mechanism	References
Reversed-phase (RP)	Hydrophobic interaction	34
Hydrophilic interaction chromatography (HILIC)	Hydrophilic interaction	35
Ion exchange (SCX, SAX, WCX and WAX)	Ionic interaction	32, 36-38
Electrostatic repulsion-hydrophilic interaction chromatography (ERLIC)	Mixed mode	33
Isoelectric focusing (IEF)	Isoelectric point	39
Capillary electrochromatography (CEC)	Charge to size ratio and hydrophobic interaction	40
Affinity chromatography	Affinity adsorption	41

The separation technologies listed in Table 1.1 are not limited to separation of proteins or peptides alone. For instance, there are gels specifically designed for peptide separation, and also most of these peptide separation methods have columns with large pore sizes of stationary phase also available for the separation of proteins. However, since the separation of protein samples suffer from all kinds of problems, e.g., inter-protein interactions, protein degradation and different conformers and post-translational modifications (PTMs) of proteins, generally it is more difficult to obtain good resolution for protein separations than peptides.

In the case of complex samples, such as whole cell lysates, the capacity of a single separation method is usually not enough to perform comprehensive proteome analysis, due to the under-sampling problem associated with MS detection (i.e., only a few peptides or proteins are detected when many species are eluted out together). Therefore, multidimensional separation, which combines different or orthogonal separation techniques, is required. There are two criteria for setting up a multidimensional separation scheme. First, in different dimensions, samples must be separated based on different

physical or chemical properties. In other words, the separation methods should be orthogonal. Second, effective interfacing between the two dimensions should be in place such that the separated sample from the earlier dimension will not be remixed when transferred into the subsequent dimension. The 2D-gel technique described previously is a good example of two-dimensional separations of proteins. Other commonly used 2D-LC configurations are, but not limited to, SCX-RPLC, high pH RPLC-low pH RPLC,[42] HILIC-RPLC,[35] and IEF-RPLC.[43] After separation, peptide fractions are introduced into mass spectrometry for further analysis. Normally, prior to MS, the separations of proteins and peptides do not reach more than three dimensions. There are several reasons: first, as the number of separation dimensions increases, the analysis time increases significantly; second, an increasing number of dimensions increases the chance of sample loss during separation; finally, for same amount of sample, an increasing number of dimensions dilutes the concentration of each fraction, which results in the difficulty of detecting low abundance analytes.

For 2D-LC MS analysis, there are two instrumental configurations. One is on-line 2D-LC MS and the other one is off-line 2D-LC MS. A well-known on-line 2D-LC setup is the MudPIT (multidimesional protein identification technology)[36] (Figure 1.2 A). It combines SCX and RPLC separations before the MS analysis. A fused-silica microcapillary column is packed with C_{18} reversed-phase material followed by SCX material. The analytes introduced onto the column are first separated by SCX, followed by reversed-phase separation, during which salts are removed from the sample so that it will not interfere with downstream MS analysis. The separation is performed in cycles. Each cycle starts with a step increase of salt concentration in the SCX to elute a portion of peptides. A gradient of increasing percentage of organic buffer solution is performed to gradually elute the peptides into the MS. The whole process is controlled by a fully automated software system. On-line 2D-LC MS systems simplify the separation process and minimize sample loss. However, it is not a good choice when the sample loading amount needs to be optimized before the second dimension of separation.

In the off-line 2D-LC MS configuration, peptide fractions eluted from the first dimension are individually collected, and then loaded onto the second dimension. The separated analytes are then analyzed by mass spectrometry. Compared to on-line 2D-LC MS systems, off-line systems involve fraction collection into vials, and thus suffer from more chances of potential sample loss.[9, 19] However, off-line systems have several advantages over on-line systems. First, because two separation dimensions are isolated from each other, the gradient conditions in the 1st dimension can be adjusted without any restriction so that optimized separation results can be achieved. Second, after the 1st dimension separation, the sample amount loaded on to the second dimension can also be optimized so that the situation of under- or over-loading can be avoided. For example, in the case of SCX-RPLC 2D-LC MS system, a strategy of maximized sample injection into the RPLC

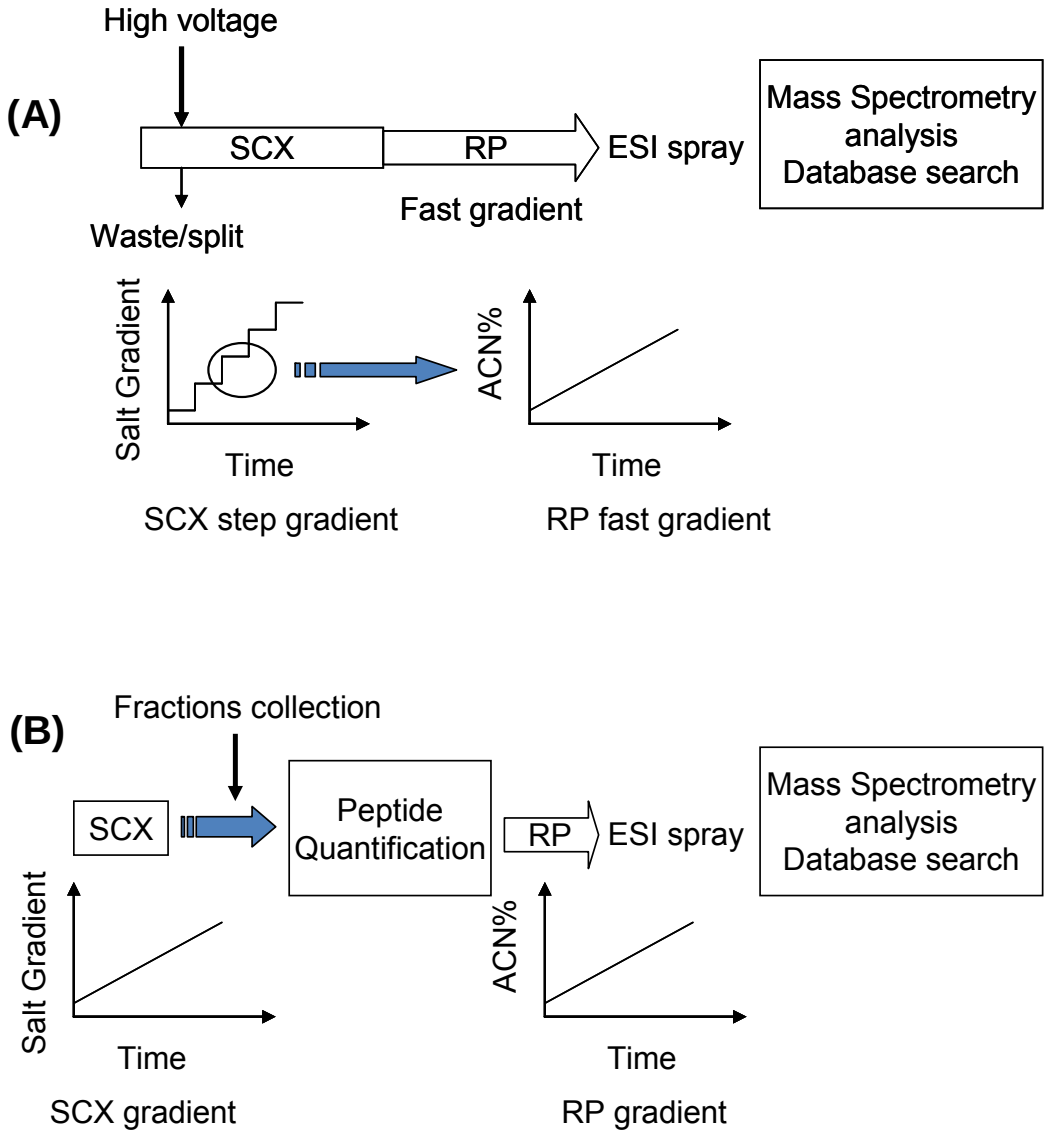


Figure 1.2 Schematic diagram of the (A) on-line and (B) off-line LC strategy.

MS/MS based on the UV-absorbance of SCX fractions was developed recently by N. Wang and coworkers.[34] By using this method, the sample amount injected into the RPLC MS/MS can be fully optimized to enhance the proteome analysis efficiency.

In my thesis, an off-line SCX-RPLC MS configuration is used for most of the work. In Chapter 6, I also applied a new 2D-LC fractionation method that combines reversed-phase separation at the protein and peptide levels for the analysis of the phosphoproteome of MDA-MB-231 cells. Compared to the traditional SCX-RPLC 2D separation of phosphopeptides, the new method is demonstrated to provide similar capability of phosphoprotein identification, and it can generate complementary information of the phosphoproteome to that from the SCX-RPLC method.

1.2.4 Tandem mass spectrometry and instrumentation

1.2.4.1 General introduction of mass spectrometry

The analysis of proteins and peptides by mass spectrometry involves three major steps: ionization, ion transportation, and mass analysis with ion detection.

First, proteins or peptides are ionized and remain in the gas phase. Since proteins and peptides are difficult to ionize and are thermally unstable, it is desirable to have a soft ionization technique for ionization. The two primary methods for ionization of proteins and peptides are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). Both of them allow the ionization of intact proteins and peptides without thermal degradation. ESI was originally developed in 1968,[44] and Dr. M. Dole first introduced this concept into mass spectrometry. After Dr. J Fenn's further development,[45] ESI became the most popular ionization technique in MS analysis of biological molecules. The ESI mechanism is described in Figure 1.3. A mixture of analytes with solvent is ejected from a metal tip to form a Taylor cone, at the end of which, charged droplets with analytes are emitted. As the solvent evaporates, the sizes of the droplets decrease and they become unstable. Once they reach their Rayleigh limit, droplets will emit smaller charged droplets with less analytes and solvent. This process repeats until analyte molecules with single or multiple charges are produced and transported into the MS. The advantages for the ESI technique are quite clear. As a soft ionization method, it can analyze fragile biomolecules without fragmentation. Moreover, because ions produced by ESI may be highly charged, the m/z ratio can be small for high mass molecules. Thus, a mass analyzer with a limited range of mass-to-charge ratio (m/z) (e.g., a quadrupole MS) can be used for the analysis of large molecules. However, the limitations of the ESI technique are as follows. First, during ionization, because only molecules on the surface of the droplets can be ionized, ion suppression can happen. In

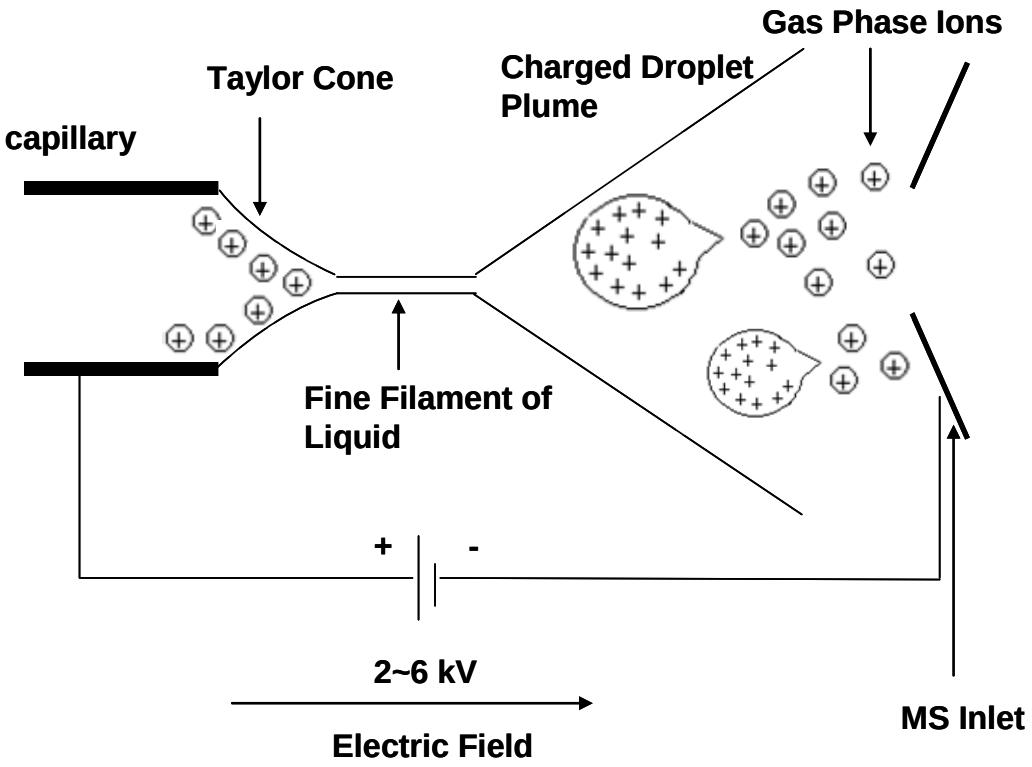


Figure 1.3 Schematic diagram of electrospray ionization (ESI).

the case of analyzing complex mixtures or samples with highly ionizable contaminates, such as salts and detergents, the results are unreliable. Second, not all analytes are ionized with high efficiency in ESI.[46, 47] On the other hand, MALDI was invented by Franz Hillenkamp and Michael Karas in 1985[48] and then applied to the ionization of proteins by Koichi Tanaka. The MALDI mechanism is described in Figure 1.4. In the MADLI technique, analytes are mixed with a matrix at a small ratio (e.g., 1:1000) and the mixture is deposited onto a metal plate which is inserted into a mass spectrometer. The desorption of matrix and analytes is triggered by a pulsed UV laser beam. Analyte molecules are then lifted by the matrix material into the gas phase and ionized. Since the ion suppression effect in the MALDI is not as severe as in the ESI, a limited amount of detergents and salts can be tolerated in MALDI analysis. However, this technique has some disadvantages, such as the difficulty of interfacing with LC separation for online analysis and the matrix interferences that may prevent from the analysis of low mass peptides ($m/z < 600$).

After ionization, ions are transported into a mass analyzer and are separated according to their m/z values. For proteomic research, the most commonly used mass analyzers are quadrupole (Q), time-of-flight (TOF), Fourier-transform ion cyclotron resonance (FTICR) and different types of ion traps (e.g., quadrupole ion trap, linear ion trap (LIT or LTQ) and Orbitrap). In order to combine the varying capacities of different mass analyzers, several types of hybrid instruments have been designed, such as Q–q–Q, Q–q–LIT, Q–TOF, TOF–TOF, and LTQ–FTICR. In my thesis, most work is done by using the Waters Q-TOF premier system, which combines the quadrupole and the TOF mass analyzers to form a tandem mass spectrometer to generate MS/MS spectra. This system will be discussed in detail in the next section.

After separation and/or selection, ions are transferred into the final part of the mass spectrometer: the detector. The type of detector used is dependent on the type of mass analayzer. A typical detector used in TOF instruments is microchannel plate detector (MCP). In a MCP, each microchannel is an electron multiplier. An ion leaving the mass analyzer enters one of the channels and hits the surface of a semi-conducting material, inducing the emission of 1-3 electrons. Electrons propagate through the channel by such bombardments, which amplify the original signal by 15 to 16 orders of magnitude, and finally the amplified electron signal reaches the anode and is detected. Many such processes occur in each micro channel of the MCP detector, allowing the detection of a large number of ions at the same time.

1.2.4.2 Quadrupole time-of-flight (Q-TOF) MS

Figure 1.5 displays the schematic of the Q-TOF premier systems from Waters. The system consists of an ESI source, followed by a quadrupole unit, which is connected to a collision cell. The ions are detected by the orthogonal acceleration TOF mass analyzer. The quadrupole is constructed of four parallel cylindrical rods. DC and RF voltages are

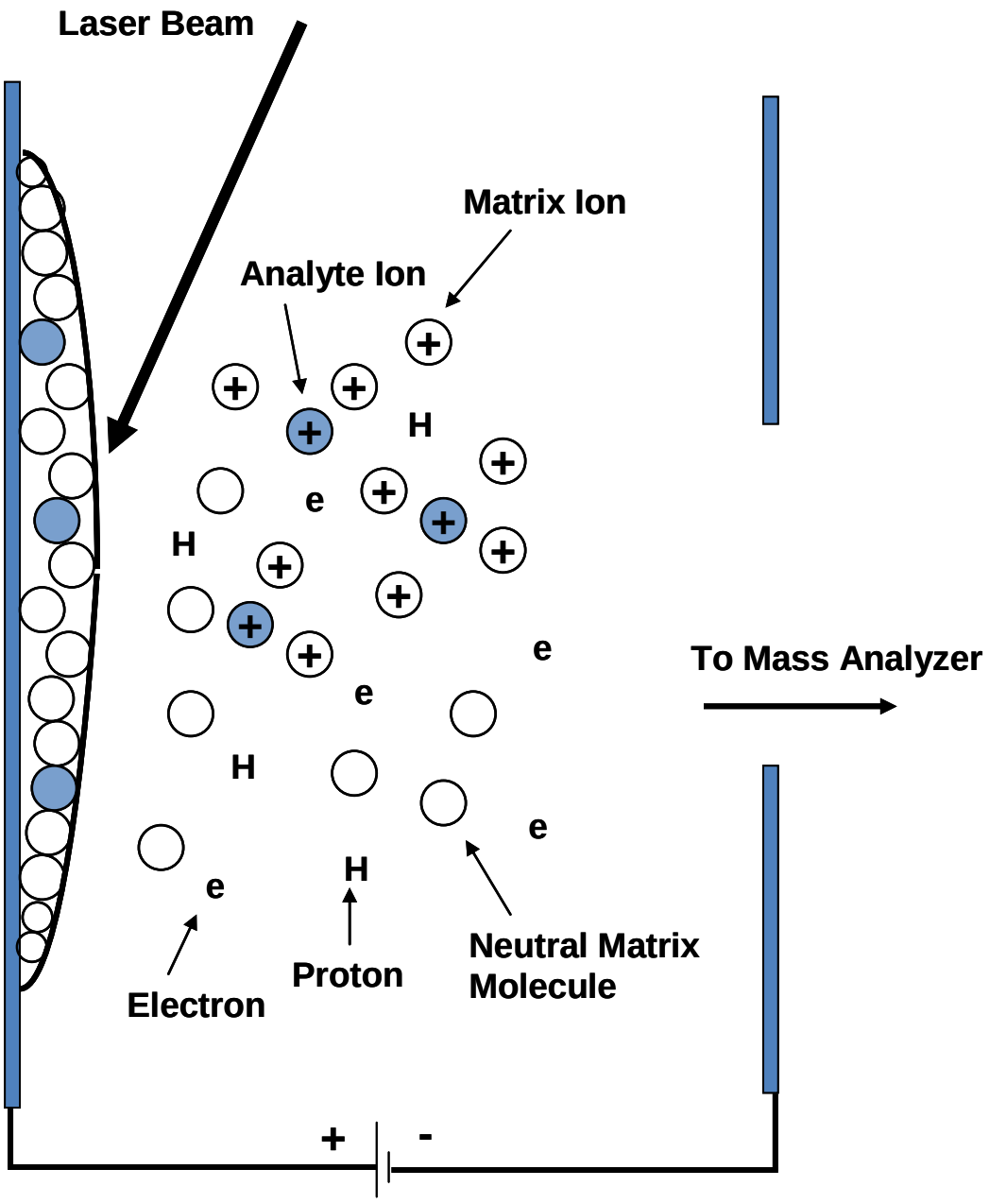


Figure 1.4 Schematic diagram of matrix-assisted laser desorption ionization (MALDI).

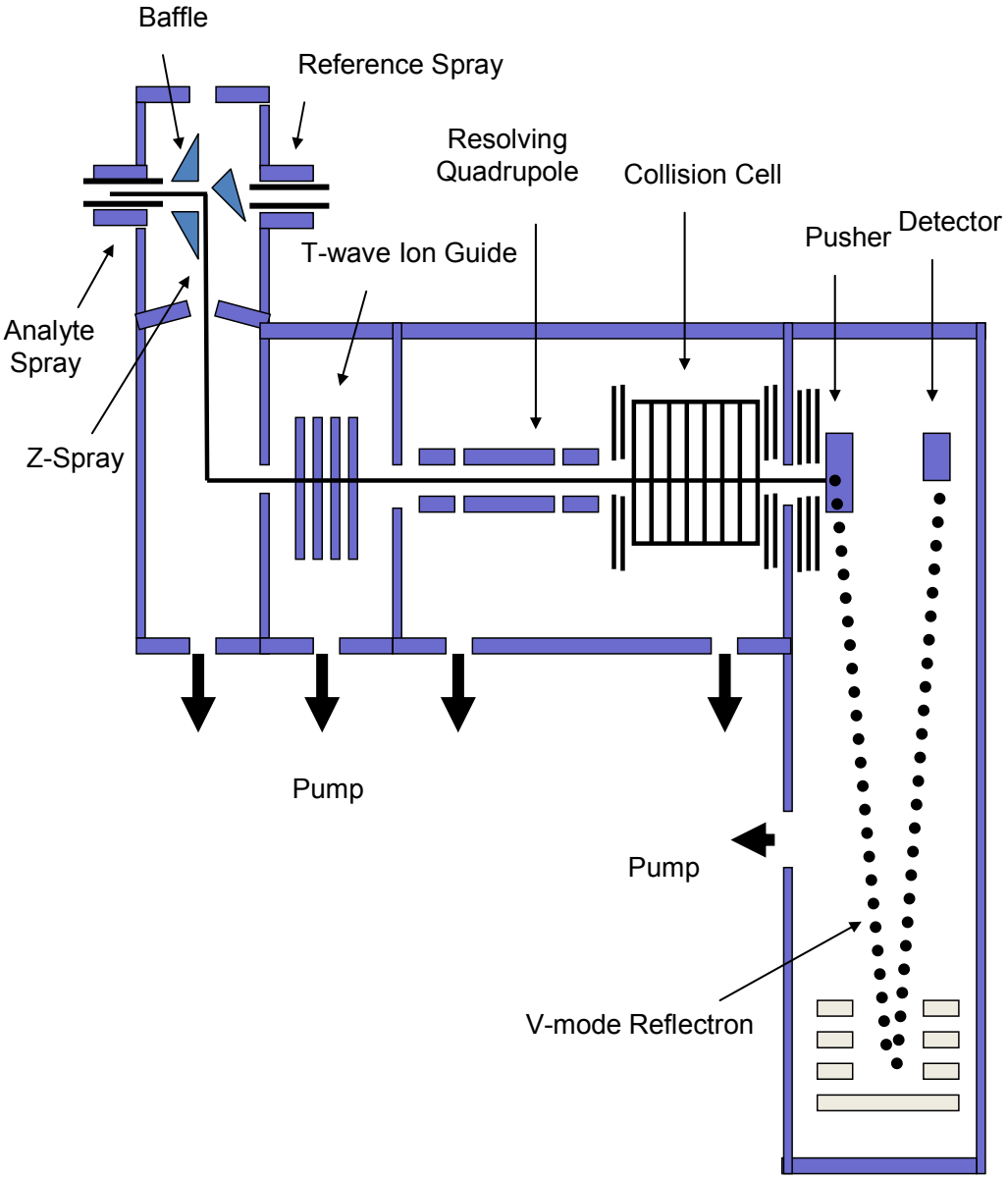


Figure 1.5 Schematic diagram of the Waters Quadrupole-time-of-flight mass spectrometric instrument (Q-TOF).

applied to the rods to form an oscillating electric field. At a given ratio of DC and RF voltages, only ions of a certain m/z ratio have stable traveling trajectories, allowing them to reach the next compartment. Other ions will collide with the rods. The transmitted ions enter the TOF tube. In the TOF analyzer, ions are accelerated by an electric field. The velocity of an accelerated ion is:

$$v = \left(\frac{2qU}{m} \right)^{\frac{1}{2}}$$

where U is the voltage, q is the charge of the ion, and m is the mass of the ion. And its time of flight is:

$$t = \frac{d}{v}$$

where d represents the length of the ion path. Since U and d are constant at a given flight tube with an electric field of known strength, the ion velocity and in turn, the flight time of the ion is determined by its mass-to-charge ratio only. Thus, the accurate m/z value can be measured by the TOF mass analyzer to generate a MS spectrum. On the other hand, ions transmitted through the quadrupole (precursor ions) can also enter the collision cell where they are fragmented into product ions. The generated product or fragment ions are then analyzed by the TOF analyzer and their accurate m/z values are recorded in the MS/MS spectrum. Normally, a V-mode reflectron TOF analyzer is used in the Q-TOF premier system (Figure 1.5). By using an ion reflector or ion mirror, the reflectron TOF analyzer compensates the initial energy distribution and focus ions having the same m/z value to the detector. Thus, compared to linear TOF analyzers, the reflectron TOF increases mass resolution from 300 to higher than 5000.

The Q-TOF premier system can be operated in three modes: 1). In MS mode, when the quadrupole is set to be RF only, such that all the ions pass through the quadrupole and be transported to the TOF analyzer. 2). In MS/MS mode, when the resolving DC of the quadrupole is turned on, the quadrupole operates to select the candidate ions for fragmentation. 3). In data directed analysis (DDA), the instrument is set to switch between MS and MS/MS modes depending on the ions detected during the MS mode.

There are two special designs in the Waters Q-TOF premier system. First, in the ion source part, the Z spray source technology is used to avoid direct introduction of analytes into the ion source. Thus, contamination from the sample itself and noise from unionized samples and other neutral molecules can be reduced. Second, a real time mass calibration is necessary in the Q-TOF instrument since the Q-TOF analyzer is sensitive to the changes in environment or experimental conditions, such as temperature. Traditionally, this was achieved by teeing-in a reference solution into the sample flow. However, it may suffer from ion suppression effect. The Waters Q-TOF premier system introduces the design of the lockspray source, where a reference spray and the sample electrospray are introduced

into the ion source separately. An oscillating baffle with programmed frequency allows the two electrospray streams to be sampled separately. As a result, the limitation of the ion suppression effect can be overcome.

1.2.5 Peptide ion fragmentation

Sequencing a peptide by tandem mass spectrometry (MS/MS) involves two steps. First, the m/z value of the peptide is measured so that the molecular weight of the peptide can be determined. Second, the peptide ion is fragmented by various techniques and the m/z values of the product ions are recorded. In the past several decades, many fragmentation techniques for proteins and peptides have been developed. In shotgun proteome analysis, the most popular MS/MS technique is the collision-induced dissociation (CID),[51] also referred as collision-activated dissociation (CAD). In CID, precursor ions are kinetically excited in the vacuum and undergo collision with neutral gas species (e.g., N_2, Ar, He) in the collision cell. Part of the kinetic energy is converted into vibrational energy which is then rapidly distributed into all covalent bonds of the ion. This process results in decomposition of the ions when the internal energy of the ion exceeds the activation barrier required for bond cleavage. These fragment ions can then be analyzed by a mass analyzer (e.g., TOF). CID can be performed in two modes: high energy and low energy.[52-54] The low energy CID mode, which is usually performed around 100 eV, is commonly used in most MS instruments for proteome analysis. Peptide ions fragmented under low energy CID usually dissociate at the peptide amide bonds, mainly producing y and b ions, and occasionally, a ions. In contrast, high energy CID is performed at the keV range, generating mostly v and w ions which refer to the fragmentation at the peptide backbone. Different patterns of product ions are illustrated in Figure 1.6. Compared to the low energy CID, more information is generated by the high energy CID, but it also increases the complexity of the spectrum, resulting in difficulty to interpret the spectral data.

Besides CID, electron-capture dissociation (ECD),[55, 56] electron-transfer dissociation (ETD)[57, 58] and high energy C-trap dissociation (HCD)[59] have been developed to be the alternative methods for dissociating the peptide or protein ions. In ECD, a thermal electron is captured by the peptide/protein ions in a FTICR MS instrument, resulting in backbone cleavage of the N-C bond to produce c and z ions. In ETD, radical anions transfer electrons to peptide/protein ions in an RF quadrupole ion trapping (QIT) instrument, inducing the backbone fragmentation to produce c and z ions. In HCD, fragmentation takes place in an octupole collision cell located at the end of the C-trap in a LTQ Orbitrap mass spectrometer. An RF voltage of 1.5-2.5 kV is applied to the collision cell and a neutral gas is supplied to facilitate the fragmentation process. Compared with CID, these new techniques provide better or

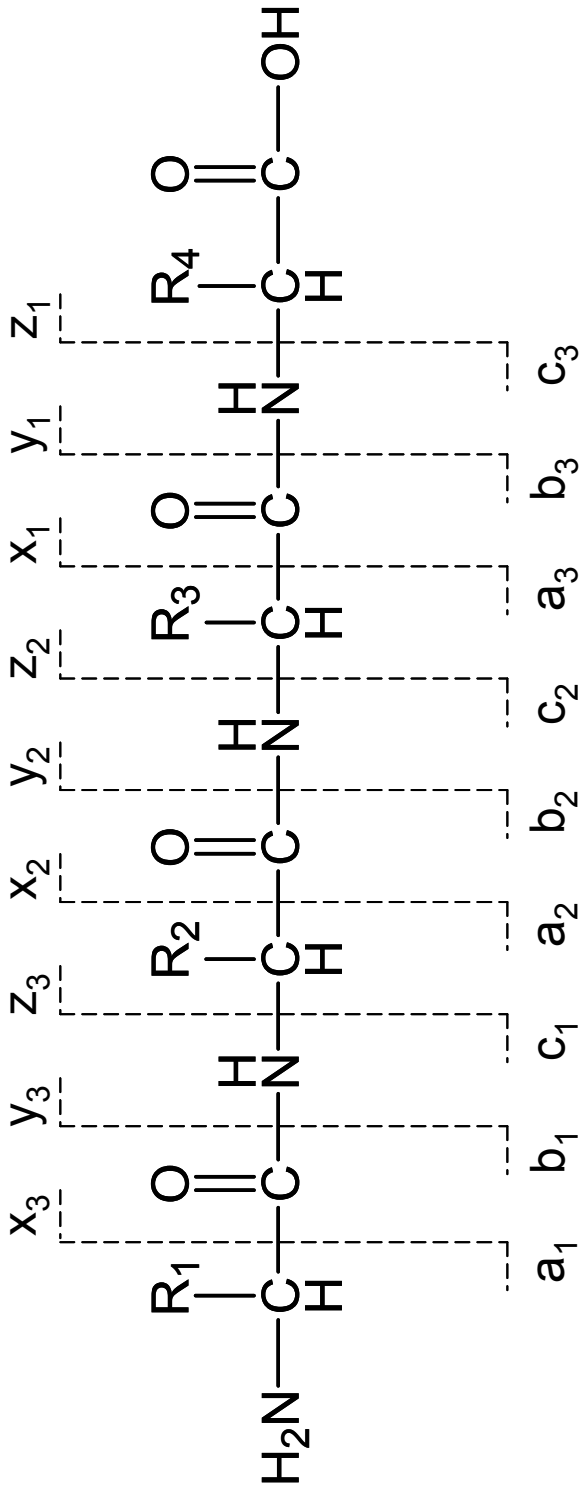


Figure 1.6 Cleavage patterns of the precursor ion.

complementary sequence coverage and preservation of labile post-translational modification groups (e.g., phosphorylation and glycosylation). However, they are still not widely used since all of them require specific expensive instrumentations.

In my thesis, low energy CID fragmentation is used for all of the work.

1.2.6 Protein identification by database searching based on MS and MS/MS

After MS and MS/MS spectra are obtained, proteins and peptides are identified by a database search. This is usually conducted in two ways, peptide mass fingerprinting (PMF)[60] and MS/MS search.[61]

In PMF, unknown proteins are usually digested by trypsin, generating peptides with specific cleavage sites. After mass spectrometric analysis, the mass list of peptides is compared to the proteome database, in which proteins are theoretically cut into specific patterns of peptides, depending on the protease used. The amino acid sequences of proteins in the proteome database are generally obtained by translating a genome sequence For each protein, a list of peptides with their calculated masses predicted from the theoretical digestion using an enzyme are entered into the database. The best match is statistically determined. Search engines available online are listed in Table 1.2. The major disadvantage of the PMF method is that the PMF algorithms assume the peptides come from a single protein,[62] which limits its application to complex protein samples.

Table 1.2 List of PMF search engines.

Search Engines	Website
Mascot	http://www.matrixscience.com/
PeptIdent	http://ca.expasy.org/tools/peptident.html/
MOWSE	http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse/
MS-FIT	http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm/
PeptideSearch	http://www.mann.embl-eidelberg.de/GroupPages/Pagelink/peptidesearchpage.html/

In the MS/MS search, tandem mass spectra are collected, which contain the information of peptide masses as well as the amino acid sequence information of the peptides. The experimental mass values are compared with the theoretical peptide mass and fragment ion mass values that are obtained by applying cleavage rules from a specific peptide fragmentation method (e.g., CID) to the protein sequences in the proteome database.

Search engines available online for MS/MS search are listed in Table 1.3. In my thesis, the Mascot search engine is used for all of the work. Mascot is based on the algorithm of MOWSE[63] with the addition of a probability-based scoring system. The scores are calculated according to the probability that the match of the experimental data to a protein sequence is a random match. A more reliable match result can be achieved by picking various search filters, such as taxonomy, enzyme, modifications and specific instrumentations. Because the fragmentation patterns of peptide ions are unique, the MS/MS search ensures more confident identification of proteins and peptides and is applicable to protein mixtures. Therefore, the MS/MS search method is widely used in analyzing complex mixtures.

Table 1.3 List of MS/MS search engines.

Search Engines	Website
Mascot	http://www.matrixscience.com/
SEQUEST	http://field.scripp.ed/sequest/
X!tandem	http://www.thegpm.org/TANDEM/index.html/
probID	http://projects.sysyemsbiology.net/probid/
Pepsea	http://pepsea.protana.com/PA_peptidePatternForm.html/
Pepfrag	http://www.proteomemetrics.com/prowl/pepfragch.html/

1.3 Analysis of phosphoproteome

1.3.1 Enrichment strategies of phosphoproteome

Analysis of the phosphoproteome by mass spectrometry is challenging as phosphoproteins are usually presented at a very small percentage, compared to their unmodified counterparts.[64] Moreover, in positive mode MS, which is commonly adopted in proteome analysis, the ionization efficiency of phosphopeptides are very low due to their negative phosphate group and ion suppression effects.[65] Therefore, enrichment of phosphoproteins/peptides is necessary to reduce the complexity of a protein sample and increase the detection sensitivity of phosphoproteins. Theoretically, phosphoproteome enrichment can be performed at both the protein and peptide levels. However, for the protein level enrichment, inter-protein interactions between binding partners and protein aggregation problems may cause low specificity. In addition, the phosphate groups may also be wrapped by other residues when proteins are not denatured well, causing low recovery of the phosphoproteins, if the enrichment is based on selective binding to the phosphate groups of the proteins. As a result, enrichment of the phosphoproteome is often performed at the peptide level.

1.3.1.1 Immobilized metal affinity chromatography (IMAC)

The IMAC technique was introduced in 1975 by Porath et al.[66] It utilizes the affinity interaction between metal cations (e.g., Fe (III), Ga (III)) and negatively charged phosphate groups. It generally enriches phosphorylated serine (pSer), threonine (pThr) and tyrosine (pTyr) peptides (Figure 1.7). The IMAC stationary phase is made by chelating the carboxylic groups on iminodiacetic acid (IDA) or the amino groups on nitrilotriacetic acid (NTA)[67] with metal cations. Figure 1.8 shows the chemical structure of the metal ion chelating molecules.

Negatively charged chemical groups other than phosphate group, such as carboxylic group on aspartic and glutamic acids, can also bind to the metal ions. Achieving a high enrichment specificity has always been a challenge in IMAC. In order to reduce the non-specific binding onto the IMAC resin, acidic residues are methyl-esterificated by chemical derivatization before IMAC enrichment.[68] However, products from side reactions may complicate the mass spectra, making interpretation difficult. Also the derivatization can not be finished completely, which sacrifices sensitivity. The other shortcoming of the conventional IMAC method is that multi-phosphorylated peptides are more easily enriched, which causes the undersampling problem for the detection of mono-phosphorylated peptides. Recently, a new protocol of sequential elution from IMAC (SIMAC) was developed by the Larsen group,[69] which sequentially eluted mono- and multi-phosphorylated peptides from the IMAC resin, allowing individual analysis of the two pools of phosphopeptides. In Chapter 6, a sequential enrichment of phosphopeptides modified from the Larsen protocol is used. In addition to SIMAC, there are several other new protocols developed and reported based on the IMAC method for the enrichment of phosphopeptides; for example, by modifying the stationary phase and supports,[70-74] specificity and efficiency can be improved. They will be discussed in detail in Chapter 6.

1.3.1.2 Metal oxide affinity chromatography (MOAC)

Phosphopeptides enrichment by MOAC benefits from the affinity between metal oxides and phosphate groups, where metal oxides act as Lewis acids and phosphate groups act as Lewis bases. Thanks to its high recovery, specificity and thermal stability, MOAC have been increasingly used in recent years. TiO_2, ZrO_2, Al_2O_3, and Nb_2O_5 have all been applied to the enrichment of phosphopeptides, among which, the most commonly used is TiO_2.[75-78] In order to improve phosphopeptide binding specificity for TiO_2, many efforts have been devoted to the optimization of the sample loading process. For instance, the binding strength to TiO_2 of 2,5-dihydroxybenzoic acid (DHB) is found to be stronger than that of non-phosphorylated acidic peptides, but weaker than phosphopeptides. Therefore, DHB was introduced into the loading buffer as a competitor to prevent non-specific binding.[79] Similar functions are also found with ammonium glutamate,[80] glutamic acid[76]

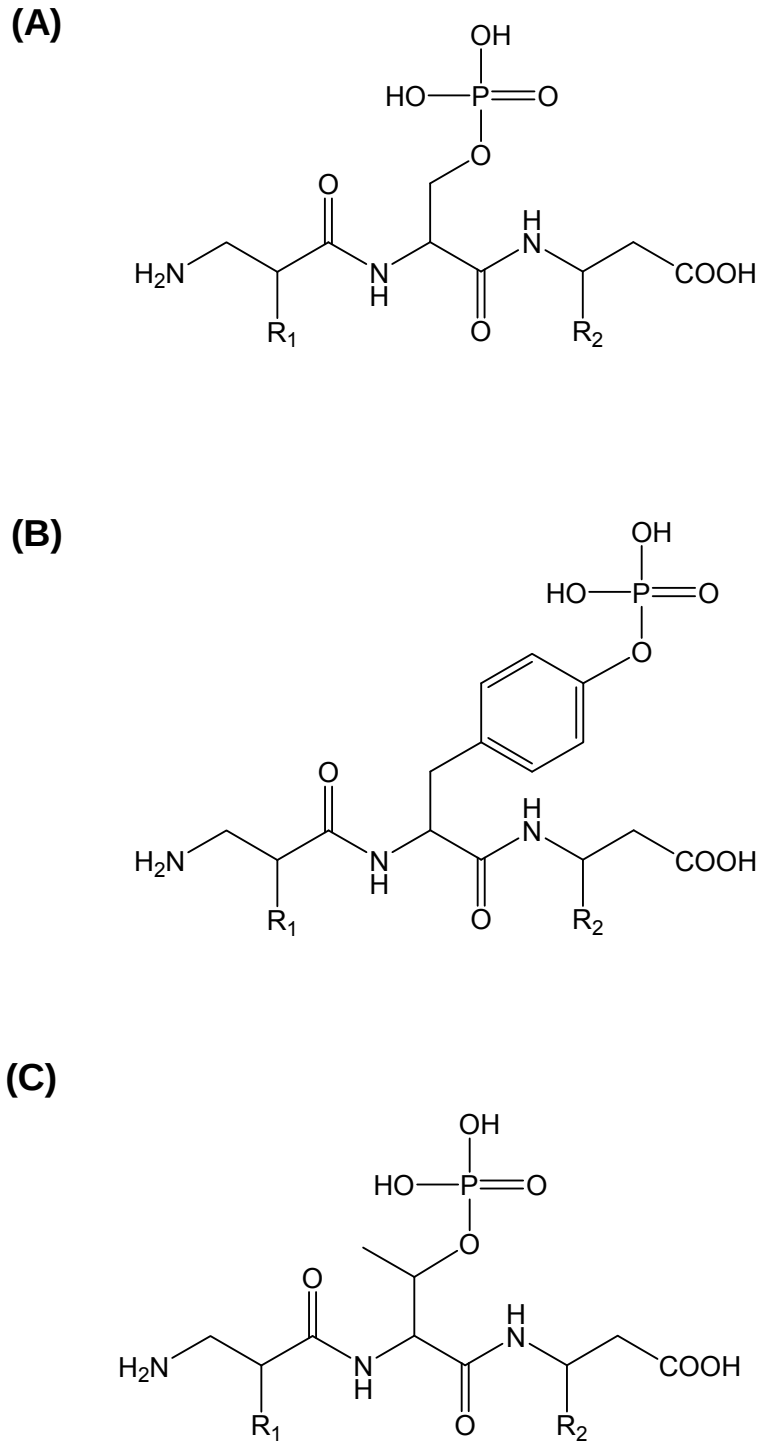


Figure 1.7 Structures of (A) phosphoserine. (B) phosphothreonine. (C) phosphotyrosine.

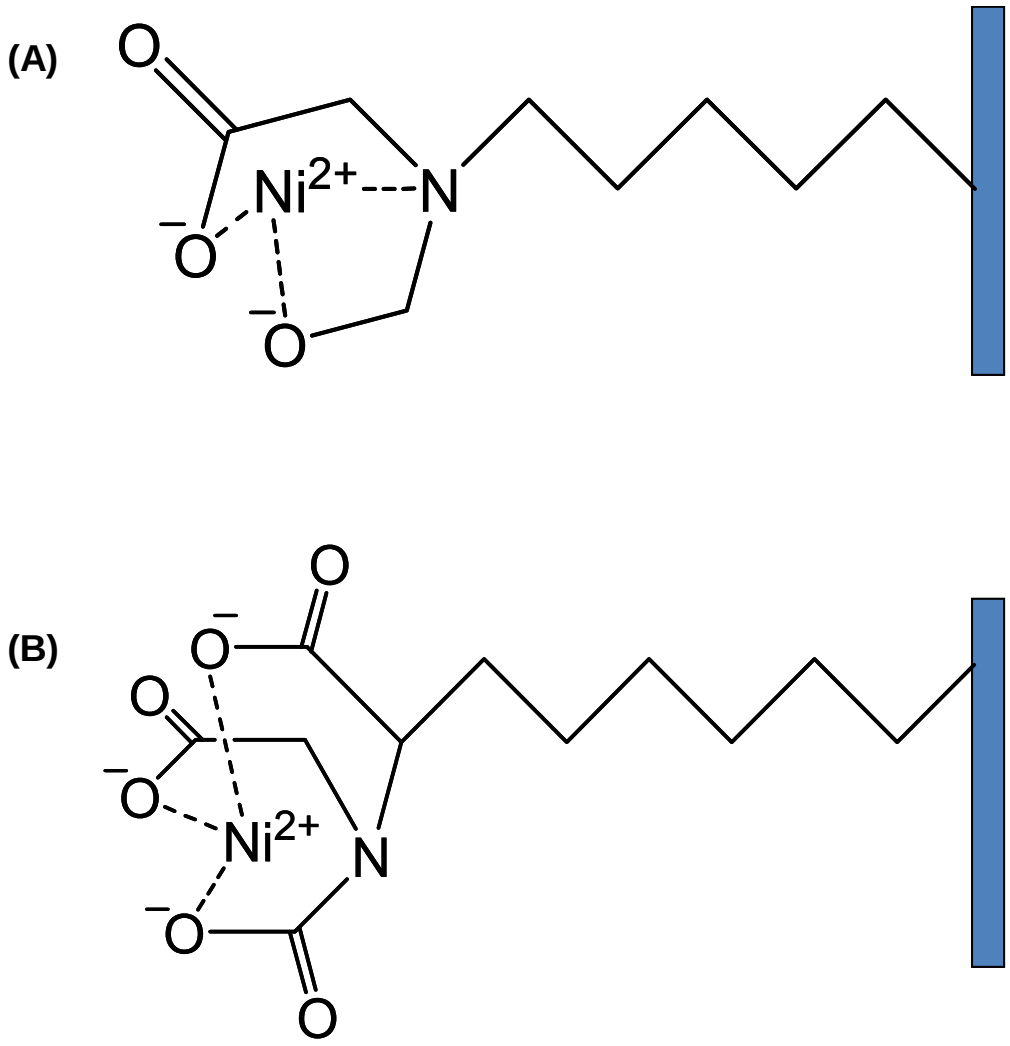


Figure 1.8 Structures of metal ion chelating stationary phase in IMAC (A) iminodiacetic acid (IDA). (B) nitrilotriacetic acid (NTA).

and phithalic acid.[75] On the other hand, the ratio of MOAC resin to the peptide sample amount is also found to be important.[81] Both insufficient and excessive amounts of TiO_2 could decrease the selectivity. Since the optimal peptide-to-TiO_2 ratio is largely sample dependent, a more general rule needs to be developed.

1.3.1.3 Strong cation exchange chromatography (SCX)

SCX, which separates peptides according to their charge states, was introduced as a phosphopeptide enrichment method in 2004.[82] More than 97% of tryptic peptides carry two positive charges under the acidic conditions (pH 2.7-3.0).[83] The charge states of phosphopeptides, which are also generated by trypsin digestion, are generally lower because of the negative charge of the phosphate group. Therefore, phosphopeptides are eluted in the front part of the salt gradient in SCX before the elution of other tryptic peptides. There are several shortcomings in using this method. First, peptides containing aspartic and gultamic acids may be co-eluted with the phosphopeptides due to their carboxylic groups. Second, phosphopeptides containing basic residues (e.g., histidine) carry an additional positive charge at low pH, which balances out the charge effect of the phosphate group, causing them to co-elute with other peptides. Finally, SCX enrichment is only applicable to the tryptic peptides while the use of other enzymes for digestion may not produce large differences in charge states between the phosphopeptides and non-phosphorylated peptides. As more efficient technologies such as optimized IMAC and MOAC have been developed for the enrichment of phosphopeptides, SCX is no longer used alone for enriching phosphopeptides. It has now become a separation method coupled with other enrichment techniques for the analysis of phosphoproteome.[84, 85]

1.3.1.4 Immunoaffinity chromatography

Using antibodies targeting phosphorylated residues to capture phosphoproteins in a generic fashion is probably the most traditional and simplest way to enrich phosphoproteome.[86, 87] Antibodies targeting phosphorylated serine, threonine, and tyrosine residues can be used for the overall enrichment of phosphoproteins, and they can also be used for enrichment at the peptide level. Non-specific binding and high cost of the antibodies are the major limitations of this method. Particular attention is given to the analysis of tyrosine phosphorylated proteins due to their extremely low abundance and the importance in many biological processes. There are highly specific antibodies against pTyr only, allowing the enrichment of tyrosine phosphorylated proteins prior to MS analysis. In the case of complex biological samples, it is necessary to combine the enrichment capacity of pTyr antibody with other phosphoproteome enrichment techniques. Recently, Zheng et al. developed a double enrichment protocol for the capture of tyrosine phosphorylated peptides.[88] To enhance the specificity, they performed phosphotyrosine enrichment by pTyr antibody followed by IMAC for the characterization of interferon α-induced pTyr proteomic changes in Jurkat cells. In contrast, anti-phosphoserine and anti-threonine

antibodies have not been widely used, mainly due to their low specificity. Only a limited number of papers reported the application of anti-phosphoserine and anti-threonine antibodies for the enrichment of phosphoproteins/peptides.[89, 90]

1.3.1.5 Chemical transformation methods

Under alkaline conditions, the phosphate group on phosphorylated residues (pSer and pThr) undergoes β-elimination to generate a reactive dehydroalanine.[91] This residue is a Michael acceptor and can readily react with a nucleophile (e.g. dithiothreitol (DTT) and 1,2-ethanedithiol (EDT)), that can be linked to a immobilizing agent, e.g. biotin affinity tag.[92] The enrichment of phosphopeptides can then be performed based on activated thiol and biotin–avidin resins. One major advantage of this technique is that in MS, thiol based derivatives are more stable that phosphate groups, resulting in enhanced MS/MS spectra of phosphopeptides and their detection. However, there are two potential disadvantages of this method. First, β-elimination of O-linked glycoproteins may cause false identification of phosphopeptides. Therefore extra procedures needs to be taken to eliminate these proteins by affinity chromatography, or remove glycosylations by a glycosidase.[93] Second, during the β-elimination process, oxidation of cysteines is necessary to prevent their reactivity,[94] which also results in the oxidation of tryptohan residues,[95] complicating the MS/MS spectra of certain peptides.

1.3.2 Analysis of phosphopeptides by mass spectrometry

As the most established fragmentation method, low energy collision-induced fragmentation (CID) has been widely used for peptide sequencing. As CID is a "slow heating" process,[96] the most labile bonds are fragmented preferentially. Therefore, the phosphorylation groups often selectively undergo β-elimination of the phosphoester bond in the gas phase, with the neutral loss of 98 Da (H_3PO_4) or 80 Da (HPO_3). This process happens in both positive and negative ion detection modes. Due to inadequate fragmentation of the peptide backbone in CID, the identification of peptide phosphorylation is difficult, especially for the determination of site location. For phosphotyrosine-containing peptides, the positive mode CID can often help generate phosphotyrosine specific immonium ion at m/z 216.043 Da,[97, 98] which may be formed by the cleavage of the N-terminal residue or by two backbone cleavages surrounding the phosphotyrosine residue.

Other than CID, electron-capture dissociation (ECD), electron-transfer dissociation (ETD) and high energy C-trap dissociation (HCD) are also used for phosphopeptide analysis. In these methods, the phosphate groups are largely intact upon fragmentation of the precursor ions and more comprehensive sequence information of product ions can be obtained from

the peptide backbone cleavage, overcoming the problems associated with CID.[57] However, ECD and ETD involve the addition or loss of electrons, inducing charge state reductions. Therefore they are limited to the analysis of multiple charged precursor ions. On the other hand, HCD is still not routinely used because it requires special expensive instrumentation.

1.4 References

1. Chalkley, R. J.; Hansen, K. C.; Baldwin, M. A., Bioinformatic methods to exploit mass spectrometric data for proteomic applications. In *Biological Mass Spectrometry*, Elsevier Academic Press Inc: San Diego, 2005; Vol. 402, pp 289-312.
2. Gorg, A.; Weiss, W.; Dunn, M. J., Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **2004,** 4, (12), 3665-3685.
3. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Science* **1989,** 246, (4926), 64-71.
4. Karas, M.; Hillenkamp, F., Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons. *Analytical Chemistry* **1988,** 60, (20), 2299-2301.
5. Henzel, W. J.; Watanabe, C.; Stults, J. T., Protein identification: The origins of peptide mass fingerprinting. *Journal of the American Society for Mass Spectrometry* **2003,** 14, (9), 931-942.
6. Canas, B.; Lopez-Ferrer, D.; Ramos-Fernandez, A.; Camofeita, E.; Calvo, E., Mass spectrometry technologies for proteomics. *Briefings in Functional Genomics & Proteomics* **2006,** 4, (4), 295-320.
7. McArthur, H. A. I.; Reynolds, P. E., Solubilization of the Membrane-Bound D-Alanyl-D-Alanine Carboxypeptidase of Bacillus-Coagulans Ncib-9365. *Biochimica Et Biophysica Acta* **1979,** 568, (2), 395-407.
8. Fischer, M., Limiting factors for the enzymatic accessibility of soybean protein. In *Limiting factors for the enzymatic accessibility of soybean protein*, 2006; p 145.
9. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology* **1999,** 17, (7), 676-682.
10. Peng, J. M.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P., Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *Journal of Proteome Research* **2003,** 2, (1), 43-50.
11. Bundy, J. L.; Cargile, B. J.; Sevinsky, J. R.; Essader, A. S.; Bjellquvist, B.; Freeman, T. W.; Stephenson, J. L., Jr., Isoelectric focusing of peptides: A ideal first dimension separation in shotgun proteomics. *Molecular & Cellular Proteomics* **2004,** 3, (10), S279-S279.

12. Figeys, D.; Ducret, A.; Yates, J. R.; Aebersold, R., Protein identification by solid phase microextraction-capillary zone electrophoresis-microelectrospray-tandem mass spectrometry. *Nature Biotechnology* **1996,** 14, (11), 1579-1583.
13. Badock, V.; Steinhusen, U.; Bommert, K.; Otto, A., Prefractionation of protein samples for proteome analysis using reversed-phase high-performance liquid chromatography. *Electrophoresis* **2001,** 22, (14), 2856-2864.
14. Jacobs, I. M.; Mottaz, H. M.; Yu, L.-R.; Anderson, D. J.; Moore, R. J.; Chen, W.-N. U.; Auberry, K. J.; Strittmatter, E. F.; Monroe, M. E.; Thrall, B. D.; Camp, D. G., II; Smith, R. D., Multidimensional proteome analysis of human mammary epithelial cells. *Journal of Proteome Research* **2004,** 3, (1), 68-75.
15. Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R., Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (17), 9390-9395.
16. Santoni, V.; Molloy, M.; Rabilloud, T., Membrane proteins and proteomics: Un amour impossible? *Electrophoresis* **2000,** 21, (6), 1054-1070.
17. Oh-Ishi, M.; Satoh, M.; Maeda, T., Preparative two-dimensional gel electrophoresis with agarose gels in the first dimension for high molecular mass proteins. *Electrophoresis* **2000,** 21, (9), 1653-1669.
18. Hood, B. L.; Zhou, M.; Chan, K. C.; Lucas, D. A.; Kim, G. J.; Issaq, H. J.; Veenstra, T. D.; Conrads, T. P., Investigation of the mouse serum proteome. *Journal of Proteome Research* **2005,** 4, (5), 1561-1568.
19. Yates, J. R.; Wolters, D. A.; Washburn, M. P., An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry* **2001,** 73, (23), 5683-5690.
20. Zhang, J.; Xu, X.; Gao, M.; Yang, P.; Zhang, X., Comparison of 2-D LC and 3-D LC with post- and pre-tryptic-digestion SEC fractionation for proteome analysis of normal human liver tissue. *Proteomics* **2007,** 7, (4), 500-512.
21. McDonald, W. H.; Yates, J. R., Shotgun proteomics and biomarker discovery. *Disease Markers* **2002,** 18, (2), 99-105.
22. Yates, J. R.; Lin, D.; Tabb, D. L., Large-scale protein identification using mass spectrometry. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2003,** 1646, (1-2), 1-10.
23. McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X. M.; Infusini, G.; Jiang, H.; Kong, X. L.; Begley, T. P., Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *Febs Journal* **2007,** 274, (24), 6256-6268.
24. McLafferty, F. W.; Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K., Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *Journal of the American Chemical Society* **1999,** 121, (4), 806-812.
25. Wu, C. C.; Yates, J. R., The application of mass spectrometry to membrane proteomics.

Nature Biotechnology **2003,** 21, (3), 262-267.
26. Molloy, M. P.; Herbert, B. R.; Williams, K. L.; Gooley, A. A., Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. *Electrophoresis* **1999,** 20, (4-5), 701-704.
27. Yu, Y. Q.; Gilar, M.; Gebler, J. C., A complete peptide mapping of membrane proteins: a novel surfactant aiding the enzymatic digestion of bacteriorhodopsin. *Rapid Communications in Mass Spectrometry* **2004,** 18, (6), 711-715.
28. Norris, J. L.; Porter, N. A.; Caprioli, R. M., Mass spectrometry of intracellular and membrane proteins using cleavable detergents. *Analytical Chemistry* **2003,** 75, (23), 6642-6647.
29. Li, L.; Zhang, N.; Chen, R.; Young, N.; Wishart, D.; Winter, P.; Weiner, J. H., Comparison of SDS- and methanol-assisted protein solubilization and digestion methods for Escherichia coli membrane proteome analysis by 2-D LC-MS/MS. *Proteomics* **2007,** 7, (4), 484-493.
30. Yates, J. R.; Chen, E. I.; Cociorva, D.; Norris, J. L., Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *Journal of Proteome Research* **2007,** 6, (7), 2529-2538.
31. Mitchell, C. R.; Bao, Y.; Benz, N. J.; Zhang, S., Comparison of the sensitivity of evaporative universal detectors and LC/MS in the HILIC and the reversed-phase HPLC modes. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2009,** 877, (32), 4133-4139.
32. Takahashi, N.; Takahashi, Y.; Heiny, M. E.; Putnam, F. W., Purification of Hemopexin and Its Domain Fragments by Affinity-Chromatography and High-Performance Liquid-Chromatography. *Journal of Chromatography* **1985,** 326, (JUN), 373-385.
33. Alpert, A. J., Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Analytical Chemistry* **2008,** 80, (1), 62-76.
34. Li, L.; Wang, N.; Xie, C. H.; Young, J. B., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
35. Heck, A. J. R.; Boersema, P. J.; Divecha, N.; Mohammed, S., Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *Journal of Proteome Research* **2007,** 6, (3), 937-946.
36. Yates, J. R.; Washburn, M. P.; Wolters, D., Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **2001,** 19, (3), 242-247.
37. Ludewig, R.; Nietzsche, S.; Scriba, G. K. E., A weak cation-exchange monolith as stationary phase for the separation of peptide diastereomers by CEC. *Journal of Separation Science* 34, (1), 64-69.
38. Dizdaroglu, M.; Krutzsch, H. C., Comparison of Reversed-Phase and Weak

Anion-Exchange High-Performance Liquid-Chromatographic Methods for Peptide Separations. *Journal of Chromatography* **1983,** 264, (2), 223-229.
39. Radola, B. J., Preparative Isoelectric Focusing - a New Effective Method for Isolation and Purification of Proteins. *Chemiker-Zeitung* **1974,** 98, (11), 549-554.
40. Behnke, B.; Metzger, J. W., Tryptic digest mapping by gradient capillary electrochromatography. *Electrophoresis* **1999,** 20, (1), 80-83.
41. Urbas, L.; Brne, P.; Gabor, B.; Barut, M.; Strlic, M.; Petric, T. C.; Strancar, A., Depletion of high-abundance proteins from human plasma using a combination of an affinity and pseudo-affinity column. *Journal of Chromatography A* **2009,** 1216, (13), 2689-2694.
42. Gilar, M.; Olivova, P.; Chakraborty, A. B.; Jaworski, A.; Geromanos, S. J.; Gebler, J. C., Comparison of 1-D and 2-D LC MS/MS methods for proteomic analysis of human serum. *Electrophoresis* **2009,** 30, (7), 1157-1167.
43. Martins-de-Souza, D.; Maccarrone, G.; Reckow, S.; Falkai, P.; Schmitt, A.; Turck, C. W., Shotgun mass spectrometry analysis of the human thalamus proteome. *Journal of Separation Science* **2009,** 32, (8), 1231-1236.
44. Dole, M.; Mack, L. L.; Hines, R. L., Molecular Beams of Macroions. *Journal of Chemical Physics* **1968,** 49, (5), 2240-2245.
45. Yamashita, M.; Fenn, J. B., Electrospray Ion-Source - Another Variation on the Free-Jet Theme. *Journal of Physical Chemistry* **1984,** 88, (20), 4451-4459.
46. Cohen, S. L.; Chait, B. T., Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Analytical Chemistry* **1996,** 68, (1), 31-37.
47. Zhu, Y. F.; Lee, K. L.; Tang, K.; Allman, S. L.; Taranencko, N. I.; Chen, C. H., Revisit of Maldi for Small Proteins. *Rapid Communications in Mass Spectrometry* **1995,** 9, (13), 1315-1320.
48. Karas, M.; Bachmann, D.; Hillenkamp, F., Influence of the Wavelength in High-Irradiance Ultraviolet-Laser Desorption Mass-Spectrometry of Organic-Molecules. *Analytical Chemistry* **1985,** 57, (14), 2935-2939.
49. Cech, N. B.; Enke, C. G., Relating electrospray ionization response to nonpolar character of small peptides. *Analytical Chemistry* **2000,** 72, (13), 2717-2723.
50. Hop, C.; Bakhtiar, R., Homocysteine thiolactone and protein homocysteinylation: mechanistic studies with model peptides and proteins. *Rapid Communications in Mass Spectrometry* **2002,** 16, (11), 1049-1053.
51. Shukla, A. K.; Futrell, J. H., Tandem mass spectrometry: dissociation of ions by collisional activation. *Journal of Mass Spectrometry* **2000,** 35, (9), 1069-1090.
52. Biemann, K.; Martin, S. A., Mass-Spectrometric Determination of the Amino-Acid-Sequence of Peptides and Proteins. *Mass Spectrometry Reviews* **1987,** 6, (1), 1-75.
53. Yost, R. A.; Boyd, R. K., Tandem Mass-Spectrometry - Quadrupole and Hybrid Instruments. *Methods in Enzymology* **1990,** 193, 154-200.
54. McLuckey, S. A., Principles of Collisional Activation in Analytical

Mass-Spectrometry. *Journal of the American Society for Mass Spectrometry* **1992,** 3, (6), 599-614.
55. McLafferty, F. W.; Zubarev, R. A.; Kelleher, N. L., Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society* **1998,** 120, (13), 3265-3266.
56. Zubarev, R., Protein primary structure using orthogonal fragmentation techniques in Fourier transform mass spectrometry. *Expert Review of Proteomics* **2006,** 3, (2), 251-261.
57. Hunt, D. F.; Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** 101, (26), 9528-9533.
58. McLuckey, S. A.; Pitteri, S. J.; Chrisman, P. A., Electron-transfer ion/ion reactions of doubly protonated peptides: Effect of elevated bath gas temperature. *Analytical Chemistry* **2005,** 77, (17), 5662-5669.
59. Horning, S.; Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Mann, M., Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* **2007,** 4, (9), 709-712.
60. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C., Identifying Proteins from 2-Dimensional Gels by Molecular Mass Searching of Peptide-Fragments in Protein-Sequence Databases. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** 90, (11), 5011-5015.
61. Loo, R. R. O.; Hayes, R.; Yang, Y. N.; Hung, F.; Ramachandran, P.; Kim, N.; Gunsalus, R.; Loo, J. A., Top-down, bottom-up, and side-to-side proteomics with virtual 2-D gels. *International Journal of Mass Spectrometry* **2005,** 240, (3), 317-325.
62. Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Shevchenko, A.; Boucherie, H.; Mann, M., Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (25), 14440-14445.
63. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J., Rapid Identification of Proteins by Peptide-Mass Fingerprinting. *Current Biology* **1993,** 3, (6), 327-332.
64. Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A., Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends in Biotechnology* **2002,** 20, (6), 261-268.
65. Bennett, K. L.; Stensballe, A.; Podtelejnikov, A. V.; Moniatte, M.; Jensen, O. N., Phosphopeptide detection and sequencing by matrix-assisted laser desorption/ionization quadrupole time-of-flight tandem mass spectrometry. *Journal of Mass Spectrometry* **2002,** 37, (2), 179-190.
66. Porath, J.; Carlsson, J.; Olsson, I.; Belfrage, G., Metal Chelate Affinity Chromatography, A New Approach to Protein Fractionation. *Nature* **1975,** 258, (5536), 598-599.

67. Hochuli, E.; Dobeli, H.; Schacher, A., New Metal Chelate Adsorbent Selective for Proteins and Peptides Ccontaining Neihboring Histidine-redidues. *Journal of Chromatography* **1987,** 411, 177-184.
68. Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M., Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nature Biotechnology* **2002,** 20, (3), 301-305.
69. Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R., SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Molecular & Cellular Proteomics* **2008,** 7, (4), 661-671.
70. Iliuk, A. B.; Martin, V. A.; Alicie, B. M.; Geahlen, R. L.; Tao, W. A., In-depth Analyses of Kinase-dependent Tyrosine Phosphoproteomes Based on Metal Ion-functionalized Soluble Nanopolymers. *Molecular & Cellular Proteomics* 9, (10), 2162-2172.
71. Feng, S.; Ye, M. L.; Zhou, H. J.; Jiang, X. G.; Jiang, X. N.; Zou, H. F.; Gong, B. L., Immobilized zirconium ion affinity chromatography for specific enrichment of phosphopeptides in phosphoproteome analysis. *Molecular & Cellular Proteomics* **2007,** 6, (9), 1656-1665.
72. Zhou, H. J.; Ye, M. L.; Dong, J.; Han, G. H.; Jiang, X. N.; Wu, R. N.; Zou, H. F., Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *Journal of Proteome Research* **2008,** 7, (9), 3957-3967.
73. Yu, L. R.; Issaq, H. J.; Veenstra, T. D., Phosphoproteornics for the discovery of kinases as cancer biornarkers and drug targets. *Proteomics Clinical Applications* **2007,** 1, (9), 1042-1057.
74. Wei, J. Y.; Zhang, Y. J.; Wang, J. L.; Tan, F.; Liu, J. F.; Cai, Y.; Qian, X. H., Highly efficient enrichment of phosphopeptides by magnetic nanoparticles coated with zirconium phosphonate for phosphoproteome analysis. *Rapid Communications in Mass Spectrometry* **2008,** 22, (7), 1069-1080.
75. Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R., Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* **2007,** 4, (3), 231-237.
76. Wu, J.; Shakey, Q.; Liu, W.; Schuller, A.; Follettie, M. T., Global profiling of phosphopeptides by titania affinity enrichment. *Journal of Proteome Research* **2007,** 6, (12), 4684-4689.
77. Sugiyama, N.; Masuda, T.; Shinoda, K.; Nakamura, A.; Tomita, M.; Ishihama, Y., Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Molecular & Cellular Proteomics* **2007,** 6, (6), 1103-1109.
78. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M.,

Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006,** 127, (3), 635-648.
79. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & Cellular Proteomics* **2005,** 4, (7), 873-886.
80. Yu, L. R.; Zhu, Z. Y.; Chan, K. C.; Issaq, H. J.; Dimitrov, D. S.; Veenstra, T. D., Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra. *Journal of Proteome Research* **2007,** 6, (11), 4150-4162.
81. Li, Q. R.; Ning, Z. B.; Tang, J. S.; Nie, S.; Zeng, R., Effect of Peptide-to-TiO(2) Beads Ratio on Phosphopeptide Enrichment Selectivity. *Journal of Proteome Research* **2009,** 8, (11), 5375-5381.
82. Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P., Phosphoproteomic analysis of the developing mouse brain. *Molecular & Cellular Proteomics* **2004,** 3, (11), 1093-1101.
83. Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J. X.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P., Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** 101, (33), 12130-12135.
84. Jensen, O. N.; Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M., Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Molecular & Cellular Proteomics* **2005,** 4, (3), 310-327.
85. Heck, A. J. R.; Pinkse, M. W. H.; Mohammed, S.; Gouw, L. W.; van Breukelen, B.; Vos, H. R., Highly robust, automated, and sensitive on line TiO_2-based phosphoproteomics applied to study endogenous phosphorylation in Drosophila melanogaster. *Journal of Proteome Research* **2008,** 7, (2), 687-697.
86. Pandey, A.; Fernandez, M. M.; Steen, H.; Blagoev, B.; Nielsen, M. M.; Roche, S.; Mann, M.; Lodish, H. F., Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. *Journal of Biological Chemistry* **2000,** 275, (49), 38633-38639.
87. Gronborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara, O.; Mann, M.; Jensen, O. N.; Pandey, A., A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies - Identification of a novel protein, Frigg, as a protein kinase A substrate. *Molecular & Cellular Proteomics* **2002,** 1, (7), 517-527.
88. Wang, Y. K.; Zheng, H. Y.; Hu, P.; Quinn, D. F., Phosphotyrosine proteomic study of interferon alpha signaling pathway using a combination of immunoprecipitation and immobilized metal affinity chromatography. *Molecular & Cellular Proteomics* **2005,** 4, (6), 721-730.
89. Pandey, A.; Gronborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara,

O.; Mann, M.; Jensen, O. N., A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies - Identification of a novel protein, Frigg, as a protein kinase A substrate. *Molecular & Cellular Proteomics* **2002,** 1, (7), 517-527.

90. Lienhard, G. E.; Kane, S.; Sano, H.; Liu, S. C. H.; Asara, J. M.; Lane, W. S.; Garner, C. C., A method to identify serine kinase substrates - Akt phosphorylates a novel adipocyte protein with a Rab GTPase-activating protein (GAP) domain. *Journal of Biological Chemistry* **2002,** 277, (25), 22115-22118.

91. Mega, T.; Nakamura, N.; Ikenaka, T., MODIFICATIONS OF SUBSTITUTED SERYL AND THREONYL RESIDUES IN PHOSPHOPEPTIDES AND A POLYSIALOGLYCOPROTEIN BY BETA-ELIMINATION AND NUCLEOPHILE ADDITIONS. *Journal of Biochemistry* **1990,** 107, (1), 68-72.

92. McLachlin, D. T.; Chait, B. T., Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Anal Chem* **2003,** 75, (24), 6826-6836.

93. Knight, Z. A.; Schilling, B.; Row, R. H.; Kenski, D. M.; Gibson, B. W.; Shokat, K. M., Phosphospecific proteolysis for mapping sites of protein phosphorylation. *Nat. Biotechnol.* **2003,** 21, (9), 1047-1054.

94. Meyer, H. E.; Hoffmann-Posorske, E.; Heilmeyer, L. M. G. J., DETERMINATION AND LOCATION OF PHOSPHOSERINE IN PROTEINS AND PEPTIDES BY CONVERSION TO S ETHYLCYSTEINE. In *Hunter, T. And B. M. Sefton*, 1991; pp 169-185.

95. Oda, Y.; Nagasu, T.; Chait, B. T., Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **2001,** 19, (4), 379-382.

96. McLuckey, S. A.; Goeringer, D. E., Slow heating methods in tandem mass spectrometry. *J Mass Spectrom* **1997,** 32, (5), 461-474.

97. Mann, M.; Steen, H.; Kuster, B., Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning. *J Mass Spectrom* **2001,** 36, (7), 782-790.

98. Mann, M.; Steen, H.; Kuster, B.; Fernandez, M.; Pandey, A., Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal Chem* **2001,** 73, (7), 1440-1448.

Chapter 2

Comprehensive Analysis of MCF-7 Membrane Proteome by Sequential Protein Precipitation and Solubilization Combined with 2D-LC MS/MS

2.1 Introduction

Membrane proteins can be defined as protein molecules that are attached to, or associated with the membrane of a cell or an organelle.[1] They play crucial roles in many important biological processes, such as vesicle and nutrient transport, intercellular communication, cell movement, and energy production.[2] Comprehensive studies of membrane proteins can help to further understand the principle of biological processes and possibly find new biomarkers and drug targets. However, as membrane proteins are usually highly hydrophobic, characterization of membrane proteome is challenging.[3] In shotgun analysis of membrane proteome, it is difficult to maintain high solubility of the hydrophobic proteins throughout the entire analysis process without introducing reagents which may interfere with downstream LC MS/MS, such as sodium dodecyl sulfate (SDS). Therefore, for the analysis of membrane proteome, it is desirable to solubilize relatively hydrophilic proteins in MS compatible solvents, or solvents that can be removed easily prior to the MS analysis (e.g., ammonium bicarbonate) in the first step, followed by the solubilization of more hydrophobic proteins in a detergent with high protein solubility (e.g., SDS). On the other hand, since most integral membrane proteins are of low abundance and high hydrophobicity, they can be easily suppressed during mass spectrometric analysis and thus not detected. Therefore, it is necessary to perform fractionations at the protein or peptide level to increase the likelihood of identification of low abundant membrane proteins.

By using the sequential protein solubilization method developed in our lab,[4] fractionation at the protein level can be achieved by sequentially solubilizing proteins in solvents of different solubilization strength (i.e. NH_4HCO_3, methanol, urea and SDS). However, different solvents possess dramatically different protein solubilities, which means that protein amount in each solubilized fraction varies greatly. For some fractions containing a high amount of proteins, the sample is still not fully simplified. Therefore, in the case of complex proteome analysis, improved fractionation at the protein level is needed. In this study, a new protocol of sequential protein precipitation and solubilization combined with 2D-LC MS/MS was developed and applied to the comprehensive analysis of the MCF-7

membrane proteome. Two steps of separation at the protein level by sequential protein precipitation and solubilization were performed prior to 2-D LC-ESI-MS/MS to reduce the complexity of the membrane protein extracts. The results generated from these experiments will be described in detail.

2.2 Experimental

2.2.1 Chemicals and reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC/MS-grade water, acetone, formic acid, methanol (MeOH), and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada). Human breast cancer cell line, MCF-7 cells (ATCC HTB-22), was purchased from the American Type Culture Collection (Manassas, VA).

2.2.2 Membrane protein extraction

MCF-7 membrane proteins were extracted as described previously.[5] MCF-7 cells were grown in 15 cm diameter plates in ATCC medium at 37 °C for 2 weeks. The growth medium was aspirated to leave a monolayer of cells on the plates, which were then placed on an ice-cold metal tray. The plates were washed three times with ice-cold 10 ml phosphate buffered saline (PBS)-buffer (0.9 mM $CaCl_2$, 0.5 mM $MgCl_2$, 0.7 mM KH_2PO_4, 8 mM $NaHPO_4$, 1 mM KCl, and 0.1 M NaCl) and then 2.5 ml saponin lysis buffer [0.2% saponin in 50 mM Tris-HCl pH 7.5 with 1 mM phenylmethylsulfonyl fluoride (PMSF)] was added to each dish. Following incubation on ice for 5 min with constant rocking, the cells were scraped from the plates and collected with the buffer solution. The suspension was centrifuged at 17,400 g for 15 min at 4 °C. The supernatant contained cytosolic proteins. The cell pellets were washed by resuspending in a total volume of 10 ml wash buffer (50 mM Tris-HCl pH 7.5 with 1 mM PMSF). The pellets were combined into one centrifuge tube and centrifuged at 17,400 g for 15 min at 4 °C. The pellet was resuspended in 12 ml Triton X-100 buffer (1% Triton X-100 in 50 mM Tris-HCl pH 7.5 with 1 mM PMSF) and incubated on ice for 15 min with vortexing at 5 min intervals to release membrane proteins. The preparation was then centrifuged (as above) to remove

insoluble materials. The supernatant was stored as aliquots at -80 °C for future use.

2.2.3 Sequential protein precipitation

Sequential precipitation techniques were employed to reduce the protein complexity of the fractions to be analyzed. 450 mM dithiothreitol (DTT) was added to the supernatant to a final concentration of 30 mM. The pH of the solution was adjusted to 8.0 by using 1 M NH_4HCO_3. The solution was then incubated for 1 h at 37 °C to reduce disulfide bonds. Iodoacetamide (IAA) (450 mM) was added to the solution with the DTT/IAA ratio of 1:2 (mole/mole) and the solution was left to stand for 1 h at the room temperature in the dark so that all reduced disulfide bonds were blocked by the carbamidomethylation reaction. The cloudy solution was centrifuged at the setting of 20,817 g for 10 min at 4°C. The pellet was stored at -80 °C as the reduction/alkylation fraction, whereas the supernatant was then transferred to a new vial and gradually mixed with four volumes of cold acetone (-80 °C) (with intermittent vortexing) to precipitate the proteins and remove detergents and other chemicals. The mixture was kept at -20 °C overnight and then centrifuged as described above. The supernatant was decanted and properly disposed. The remaining acetone was evaporated at room temperature and the protein precipitate was stored at -80 °C for future use.

2.2.4 Sequential protein solubilization

Sequential solubilization techniques were employed to maximize the dissolution efficiency of the two protein precipitates produced from the previous step (Figure 2.1). They were performed as previously described with minor changes.[4] The sequential solubilization steps were: (1) ammonium bicarbonate-assisted solubilization, (2) methanol-assisted solubilization,[6, 7] and (3) SDS-assisted solubilization.[6-8] For each of the precipitated protein fractions, ammonium bicarbonate (25 mM, pH 8.0) was first added to the pellets. Intermittent vortexing was applied for 2 h, and then centrifuged at 20,817 g for 10 min at 4°C (the same vortexing and centrifugation steps were applied in the subsequent solubilization processes). The remaining pellet was re-suspended in 60% methanol with sufficient vortexing for 2 h. The solubilized proteins were transferred into a different vial, and MeOH in the solution was evaporated using a SpeedVac (Thermo Savant, Milford, MA). Finally, 2% SDS was employed to dissolve the remaining pellet. The same sequential solubilization procedure was employed for both protein precipitates. However, for the reduction/alkylation fraction, since the pellet amount was small，the methanol-assisted solubilization step for this fraction was skipped.

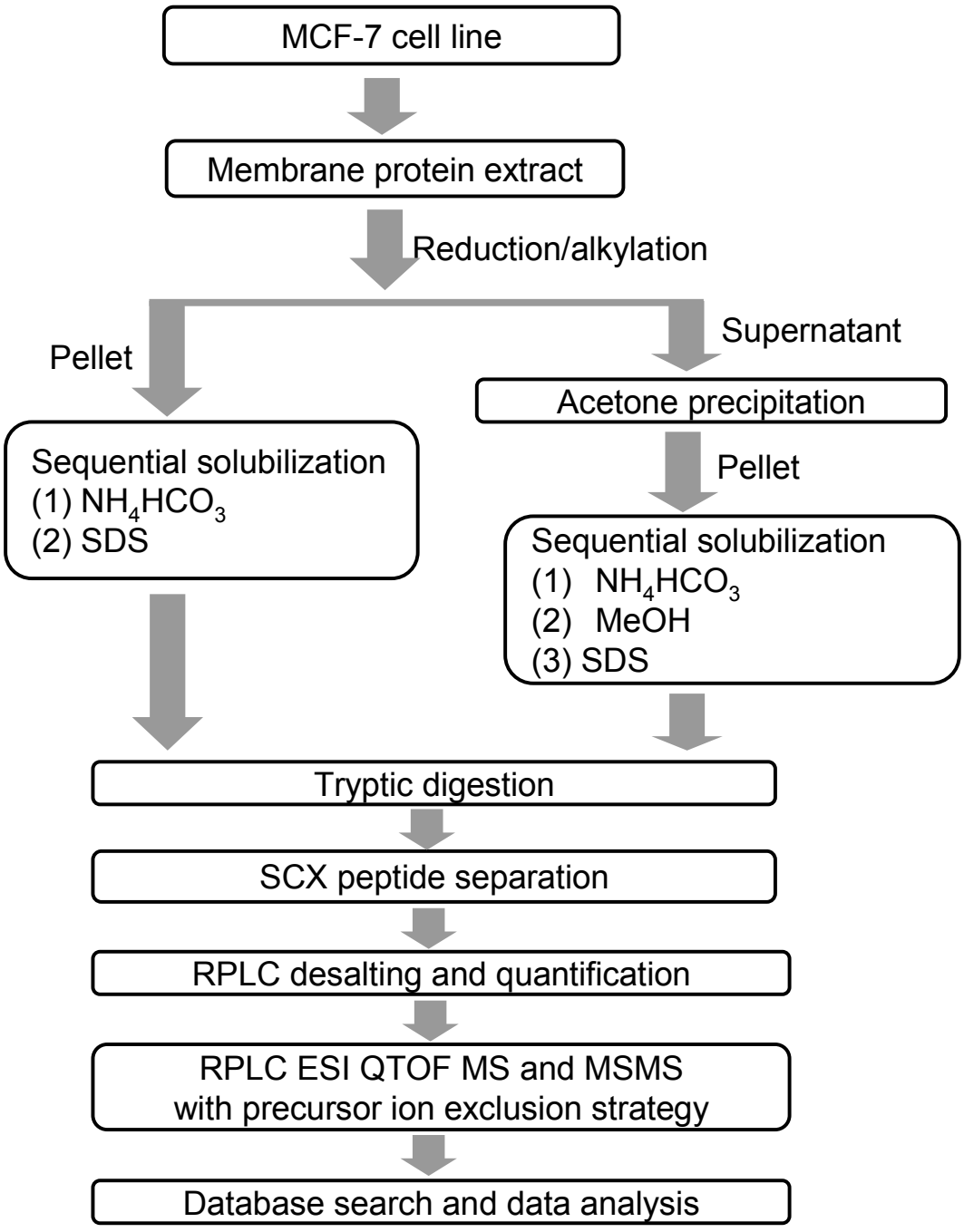


Figure 2.1 Work flow for sample preparation, separation and analysis.

2.2.5 In-solution digestion

The five protein fractions generated by the above solubilization steps (two from the reduction/alkylation precipitate and three from the acetone precipitate) were individually treated by tryptic digestion. The SDS-assisted fractions were first diluted 40 fold to ensure high digestion efficiency. Trypsin was added into the supernatant at an enzyme/protein ratio of 1:40, and the digestion was conducted at 37 °C for two days with 10% (of the amount added for the first day) additional fresh trypsin added before the second day digestion. Reactions were stopped by adding 10% TFA to adjust the pH to 2.5. The digested peptide solutions were stored at -80 °C until further analysis

2.2.6 Strong cation exchange chromatography

SCX separation of the tryptic peptides was performed as previously described with minor changes.[6] A 2.1×250 mm polySULFOETHYL^{TM} A column (particle size of 5 μm diameter and 300 Å pore, PolyLC, Columbia, MD) was connected to an Agilent 1100 HPLC system (Palo Alto, CA). Gradient elution was performed with Buffer A (10 mM KH_2PO_4, pH 2.7) and Buffer B (10 mM KH_2PO_4, 500 mM KCl, pH 2.7). Protein digests from each fractions were loaded separately onto the SCX column, peptides were eluted using linear gradients (100% Buffer A for 15 min, 0-4% Buffer B for 1 min, 4-20% Buffer B for 16 min, 20-60% Buffer B for 22 min, 60-100% Buffer B for 16 min, held for another 5 min, and then ramp back to 0.0 Buffer B for 2 min) at 0.20 mL/min with auto collection of 1-min fractions. Adjacent fractions were pooled together according to their amount based on the chromatography UV absorption signal recorded at 214 nm.

2.2.7 Peptide desalting and quantification by RPLC

Desalting and quantification were carried out using an Agilent 1100 HPLC system (Palo Alto, CA) using a method developed by Wang et al.[9] Desalting of tryptic peptides was performed on a 4.6×50 mm Polaris C18 A column (3 μm particles and 300 Å pore size Varian, USA). After loading of the peptide sample, the column was flushed with Buffer A (0.1% TFA in water) and the salts were effectively removed. Subsequently, the concentration of Buffer B (0.1% TFA in ACN) in the mobile phase was step-wise increased to 85% to ensure complete elution of the peptide fractions from the column. During the peptide elution process, a chromatographic peak was produced and based on the peak area of UV absorbance at 214 nm, the amount of peptides was determined. BSA digests of various amounts were used as standards for the generation of a linear calibration between

the peak area and the injected peptide amount. The calibration curve was generated as y=410.7x+1600, where y refers to the peak area of the peptide sample, and 1600 refers to the peak area of the blank run and x refers to the peptide amount of the sample in μ g. The linear range of the calibration curve was from 0.5 μ g to 5.0 μ g of peptides (R^2=0.998).

2.2.8 LC-ESI QTOF MS and MS/MS analysis

This step was performed as previously described with minor changes.[6] 1-μg portions of peptides of each desalted SCX fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA) equipped with a nanoACQUITY Ultra Performance LC system (Waters). Peptide solution from each SCX fraction was injected onto a 75 μm × 100 mm Atlantis dC18 column (Waters) (3 μm particles and 100 Å pore size). Buffer A consisted of 0.1% formic acid in water, and Buffer B consisted of 0.1% formic acid in ACN. Peptides were first separated using 120 min gradients (2-7% Buffer B for 2 min, 7-20% Buffer B for 83 min, 20-30% Buffer B for 25 min, 30-45% Buffer B for 5 min, 45-90% Buffer B for 5 min). All samples were electrosprayed into the mass spectrometer (fitted with a nanoLockSpray source) at a flow rate of 0.35 μL/min. Mass spectra were acquired from m/z 350-1600 for 0.8 s, followed by 4 data-dependent MS/MS scans from m/z 50-1990 for 0.8 s each. Leucine Enkephalin and (Glu1)-Firinopeptide B, a mixed mass calibrant (i.e., lock-mass), was infused at a flow rate of 0.25 μL/min, and an MS scan was acquired for 1 s every 1 min throughout the run. The collision energy used to perform MS/MS was automatically varied according to the mass and charge state of the eluting peptide. Peptide precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides identified from the adjacent two SCX fractions to enable additional and less abundance peptides to be analyzed and identified.[10] An exclusion list was generated based on Mascot (Matrix Science, London, U.K.) searching results of peptides with a score 10 points above the identity threshold.

2.2.9 Protein database search

Database searches were performed as previously described with minor changes.[6] Raw search data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.2.5 (Waters). Peptide sequences were identified by automated database searching of peak list files using the Mascot search program. Database searching was restricted to Homo sapiens (16054 sequences) in the Swiss-Prot database. The following search parameters were selected for all database searching: enzyme, trypsin; missed cleavages, 1; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+, and 3+; fixed modification, carbamidomethyl (C); variable

modifications, acetyl (protein), oxidation (M), Pyro_glu (N-term Q), Pyro_glu (N-term E). The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at the confidence level of 95% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but different access ID numbers were also removed from the list. To gauge the false positive peptide matching rate, target-decoy search strategy was applied by searching the MS/MS spectra against the forward and reverse human proteome sequences.[11, 12]

2.2.10 Hydropathy calculation and annotation of localization

All peptides and proteins identified were examined using the ProtParam program available at the EXPASY Web site (http:// us.expasy.org/tools/protparam.html), which allows for calculation of the grand average of hydropathy (GRAVY).[13] The Gene Ontology database (http://geneontology.org/) was used to classify proteins on the basis of cellular location. Proteins not described in the Gene Ontology database and proteins with unspecified cellular location were categorized as 'unknown'. Transmembrane domain (TMD) of identified proteins were predicted by TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) according to protein primary sequences.

2.3 Results and discussion

2.3.1 Protein and peptide identification

To generate a comprehensive view of the MCF-7 membrane proteome, separations at both the protein and peptide levels were performed. Figure 2.1 illustrates the workflow of sample preparation, fractionation and analysis. Membrane proteins were first enriched from the MCF-7 cell line as described previously.[5] The membrane protein extract was subjected to reduction by DTT and alkylation by IAA sequentially. The precipitate generated after this step was kept as the reduction/alkylation fraction, and the supernatant was then treated by cold acetone to obtain the other protein precipitate fraction or acetone fraction. Both of the protein precipitate fractions were sequentially solubilized by NH_4HCO_3, methanol and SDS, respectively. All solubilized fractions were individually subjected to tryptic digestion and the resultant peptides were separated and analyzed by

2-D (strong cation exchange and RPLC) LC QTOF MS/MS. Table 2.1 summaries the protein amounts, the numbers of SCX fractions and the numbers of proteins and peptides identified from the five protein fractions generated by sequential protein precipitation and solubilization. Figures 2.2 and 2.3 show the number distributions of the identified proteins and peptides in various fractions.

Table 2.1 Summary of the protein amounts, the numbers of SCX fractions and the numbers of proteins and peptides identified from the five protein fractions generated by sequential protein precipitation and solubilization.

Protein fraction	Reduction/alkylation		Acetone			Total
	NH_4HCO_3	SDS	NH_4HCO_3	MeOH	SDS	
Protein amount (μg)	406.45	35.71	446.08	17.38	41.58	947.20
Number of fractions after SCX	22	11	23	4	13	73
Number of Proteins identified	3153	2715	2943	1041	2391	5011
Number of Peptides identified	11369	7820	8437	2677	8175	24261

As Figure 2.2 shows, in the reduction/alkylation fraction, 3153 unique proteins with 11369 different peptides were identified from the NH_4HCO_3-assisted solubilization fraction, and 2715 unique proteins with 7820 different peptides from the SDS-assisted solubilization fraction, with a total of 3812 unique proteins and 16476 different peptides. In the acetone fraction (Figure 2.3), 2943 unique proteins and 8437 different peptides were identified from the NH_4HCO_3-assisted solubilization fraction, 1041 proteins and 2677 peptides from the methanol-assisted solubilization fraction, and 2391 proteins with 8175 peptides from the SDS-assisted solubilization fraction, with a total of 3784 unique proteins and 14674 different peptides. In total, 5011 unique proteins and 24261 different peptides were identified from the membrane protein extract. Among these proteins, 64% proteins were identified based on two or more peptide matches. In terms of false positive rate, reversed sequence search resulted in 388 matches, whereas the forward and reversed sequences search resulted in 73484 matches. Therefore, the false positive peptide matching rate was estimated to be 0.53%.

2.3.2 Sequential precipitation and solubilization validation

Figure 2.4 shows the comparison of the proteins and peptides identified from two sequentially precipitated protein fractions. There were 3811 unique proteins and 16476 peptides identified from the reduction/alkylation fraction and 3784 unique proteins with

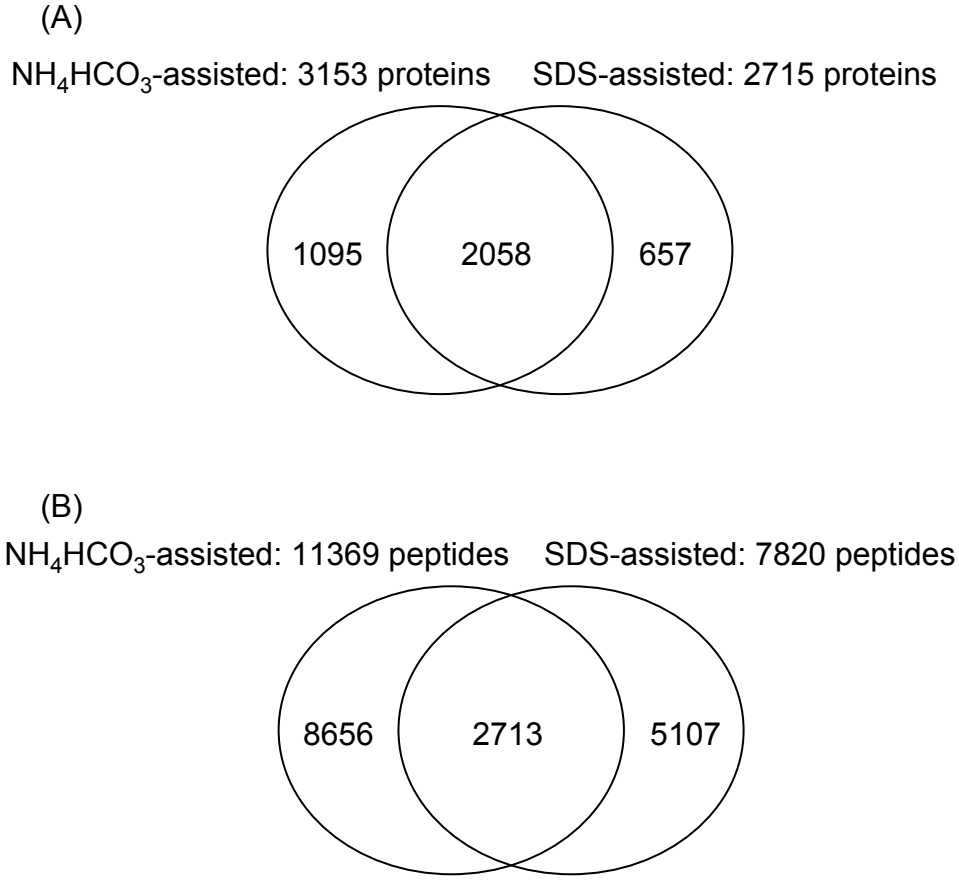


Figure 2.2 (A) Venn diagram of the proteins identified from reduction/alkylation fraction. A total number of 3812 proteins were identified: 3153 proteins from the NH_4HCO_3-assisted solubilization fraction, 2717 proteins from the SDS-assisted solubilization fraction. (B) Venn diagram of the peptides identified from reduction/alkylation fraction. A total number of 16476 peptides were identified: 11369 peptides from the NH_4HCO_3-assisted solubilization fraction, 7820 peptides from the SDS-assisted solubilization fraction.

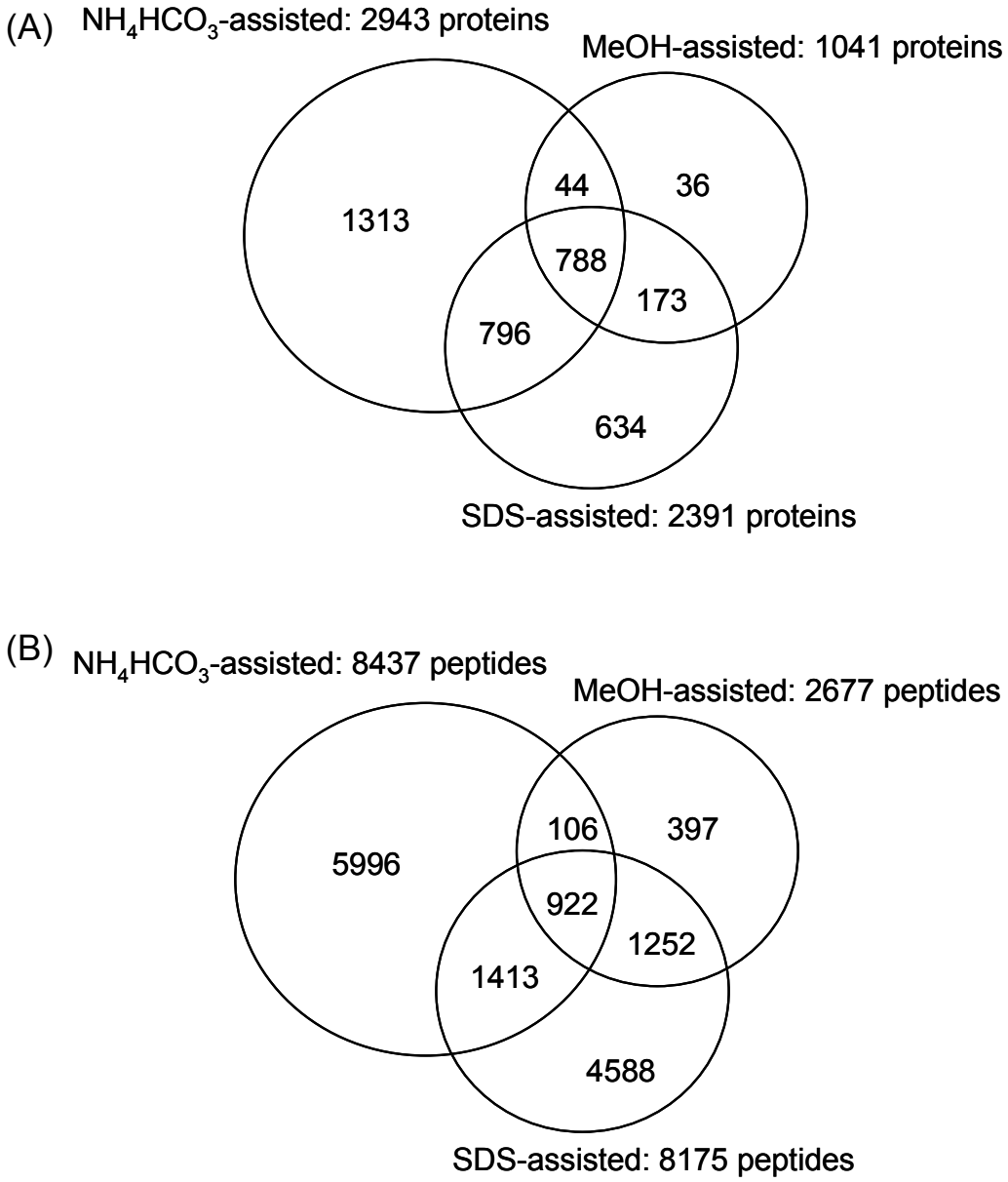


Figure 2.3 (A) Venn diagram of the proteins identified from acetone fraction. A total number of 3784 proteins were identified: 2943 proteins from the NH_4HCO_3-assisted solubilization fraction, 1041 proteins from the methanol-assisted fraction, 2391 proteins from the SDS-assisted solubilization fraction. (B) Venn diagram of the peptides identified from acetone fraction. A total number of 14674 peptides were identified: 8437 peptides from the NH_4HCO_3-assisted solubilization fraction, 2677 peptides from the methanol-assisted fraction, 8175 peptides from the SDS-assisted solubilization fraction.

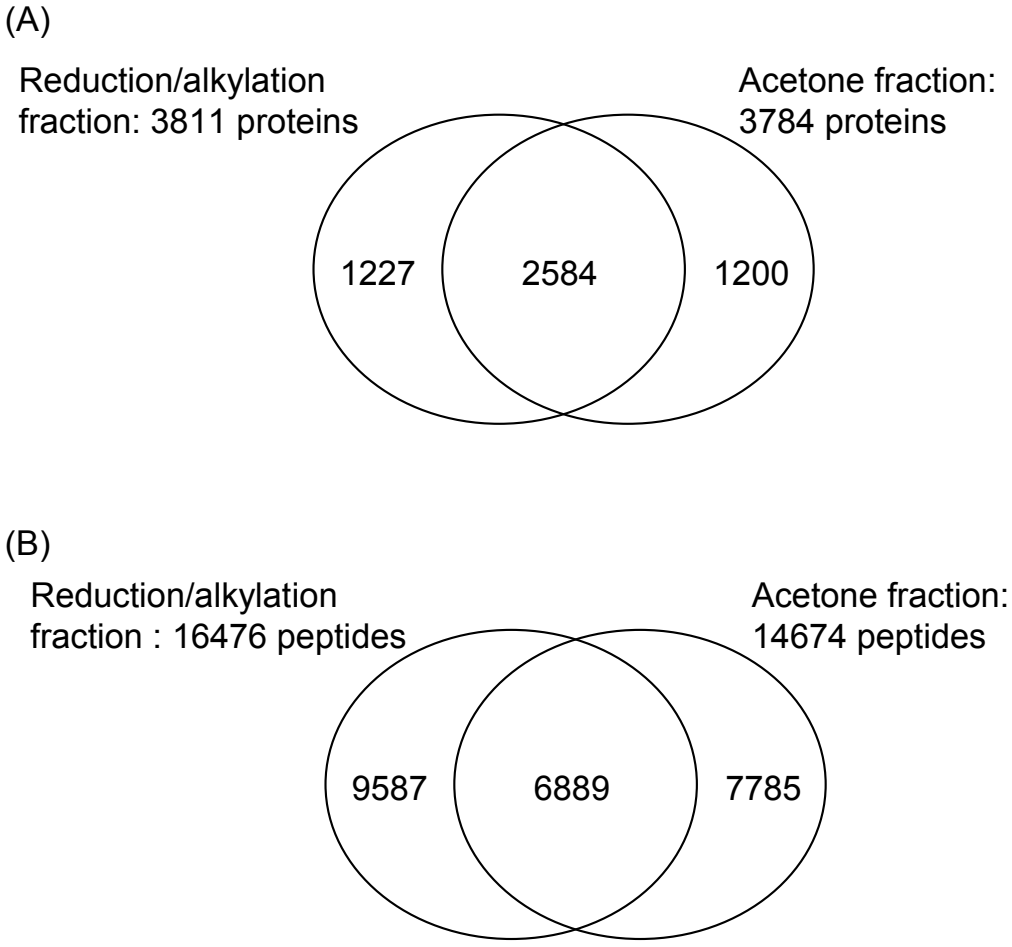


Figure 2.4 (A) Venn diagram of the proteins identified from two sequential precipitation techniques. A total number of 5011 proteins were identified, with the overlap of 51% between two seuqential precipitation fractions. (B) Venn diagram of the peptides identified from two sequential precipitation techniques. A total number of 24261 peptides were identified, with the overlap of 28% between two sequential precipitation fractions.

14674 peptides identified from the acetone fraction, with an overlap of 51% at the protein level and 28% at the peptide level. Figure 2.5 shows the comparison of the proteins identified from different solubilization solvents. Proteins detected from different precipitation fractions, but from the same solubilization solvent (e.g., NH_4HCO_3) were combined together. It should be noted that, since only 36 unique proteins were identified from the methanol-assisted solubilization fraction, it was combined with the result of SDS-assisted solubilization fractions, and will be discussed later as one fraction. As Figure 2.5 shows, there were 4064 unique proteins and 15444 different peptides found from the NH_4HCO_3-assisted fraction and 3698 unique proteins with 13813 different peptides identified from the methanol- and SDS-assisted fraction, with an overlap of 55% at the protein level and 21% at the peptide level. It has been well accepted that sequential solubilization of proteins by solvents with different solubilization strengths (e.g., NH_4HCO_3 and SDS) can achieve some level of proteome fractionation.[14] Since the overlap between the two sequential precipitation fractions (51% at the protein level and 28% at the peptide level) is similar with that of the two sequential solubilization fractions, it is clear that sequential precipitation can also simplify the proteome sample at the protein level. On the other hand, as shown in Figures 2.2 and 2.3, in the reduction/alkylation and acetone fractions, the overlaps between sequential solubilization fractions (i.e., NH_4HCO_3- and MeOH/SDS-solubilized fractions) were as low as 54% and 43%, respectively. This result indicates that fractionation by sequential precipitation is, to some extent, orthogonal to the sequential solubilization method, and thus, by combining the two methods, we can further reduce the complexity of the sample.

In this work, replicate experiments were not performed, since from our experience, the shotgun proteome profiling protocols developed in our lab hold very high reproducibility which will be demonstrated in later chapters. For example, in Chapter 4, the overlaps between the two replicate experiments of the vortex-assisted sequential protein solubilization (VAPS) method for the profiling of the *E. coli* integral membrane extract were 77% at the protein level and 55% at the peptide level (see Chapter 4 for detailed information). In Chapter 6, the overlaps between the two replicate experiments of the SCX-RPLC method for the phosphoproteome profiling of the MDA-MB-231 cell line were 77% at the protein level and 58% on the peptide level (see Chapter 6 for detailed information). Although different proteomes were profiled in Chapters 4 and 6, the overlaps between replicate experiments in these works are very similar. Since the similar protocols and the same instrument were used in this current work and those in Chapters 4 and 6, we would expect that comparable overlaps could also be obtained if replicate experiments were preformed by the sequential protein precipitation and solubilization method. Because the overlaps between the replicate experiments in Chapters 4 and 6 are all significantly higher than the overlaps between the sequential precipitation and

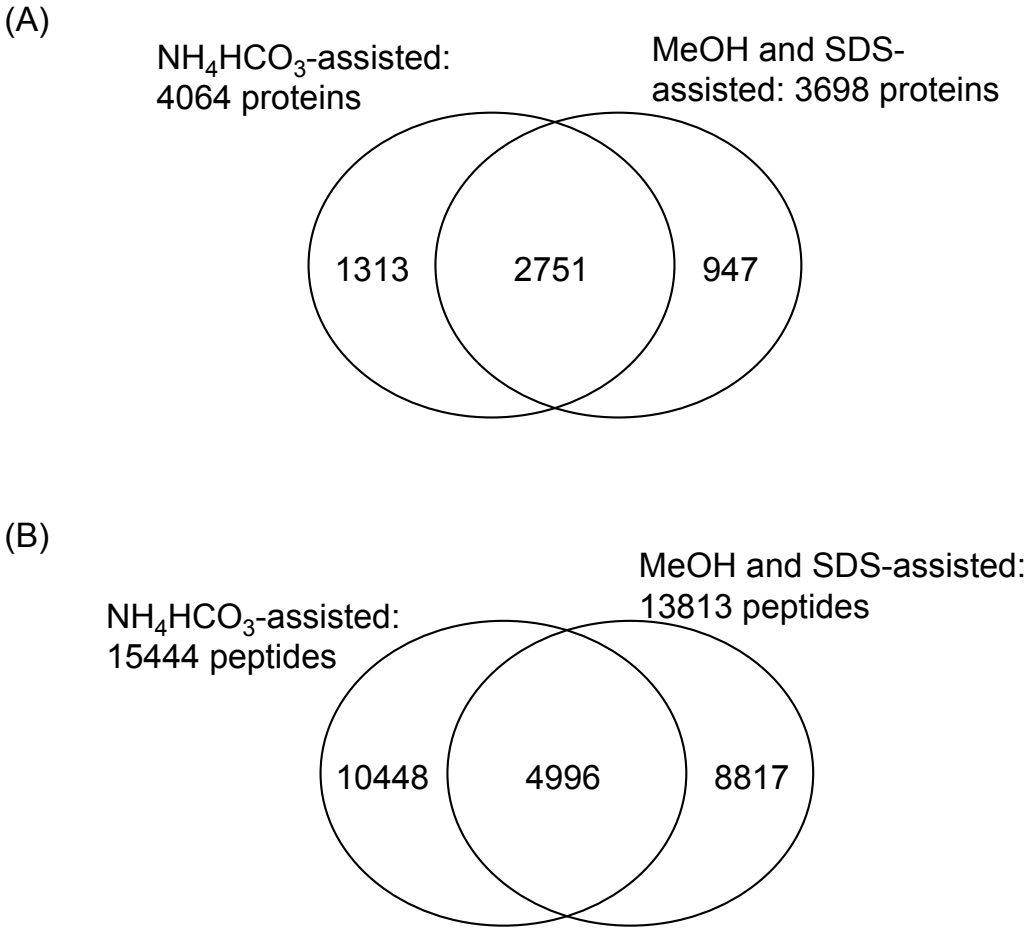


Figure 2.5 (A) Venn diagram of the proteins identified from sequential solubilization techniques. MeOH and SDS-assisted solubilization fractions were combined together. The overlap between two sequential solubilization fractions was 55%. (B) Venn diagram of the peptides identified from sequential solubilization techniques. MeOH and SDS-assisted solubilization fractions were combined together. The overlap between two sequential solubilization fractions was 21%.

solubilization fractions in this work, we can conclude that the technique of sequential precipitation and solubilization did achieve some orthogonality for protein fractionation.

2.3.3 Protein level separation characterization

We have annotated some of the physicochemical and biological characteristics of the obtained proteome data to rationalize the separation effect of sequential precipitation and provide a better understanding of the proteome itself. To better characterize the sequential precipitation and solubilization technique, only unique proteins identified from the individual fractions were compared.

The number of cysteines of each identified protein from the reduction/alkylation and acetone fractions was first investigated. Figure 2.6 shows the distribution of numbers of cysteine sites in proteins detected from the two sequential precipitation fractions. Proteins identified from the reduction/alkylation fraction appear to have more cysteine sites than that of the acetone fraction. The coupling of thiol groups from two cysteine residues in a protein results in the formation of a disulfide bond in the absence of reducing reagent. The disulfide bond facilitates the folding of a protein in two ways.[15] First, it holds two portions of the protein together, converting the protein molecule towards the folded conformer. Second, it can act as the nucleus of a hydrophobic core of the folded protein, which means the local hydrophobic resides may condense around the disulfide bond and onto each other through hydrophobic interactions. Once treated by DTT, disulfide bonds are reduced and the proteins are partially unfolded or denatured. Therefore, the hydrophobic sites of a protein, if present, originally conserved by the disulfide bonds are exposed into the aqueous solution. As a result, these proteins with more disulfide bonds would have a higher probability to be precipitated out during the reduction/alkylation process.

The detectability of proteins between the two sequential precipitation fractions was also found to be dependent on molecular mass. As Figure 2.7 shows, the reduction/alkylation fraction contains 73.9% proteins with molecular mass of higher than 40 kDa and 26.1% proteins with lower than 40 kDa. In contrast, among the proteins recovered from the acetone fraction, 57.5% proteins have molecular mass of higher than 40 kDa, whereas molecular masses of 42.5% proteins are lower than 40 kDa. Thus, reduction/alkylation tends to recover more proteins with high molecular mass, whereas acetone tends to recover more proteins with low molecular mass. This may be attributed to the fact that proteins with relatively large sizes are more readily precipitated out first and removed from the solution by the reduction/alkylation step prior to the acetone precipitation. Furthermore, this result is consistent with the previous statement that proteins with more cysteine sites would have more chance to be precipitated out during the reduction/alkylation step,

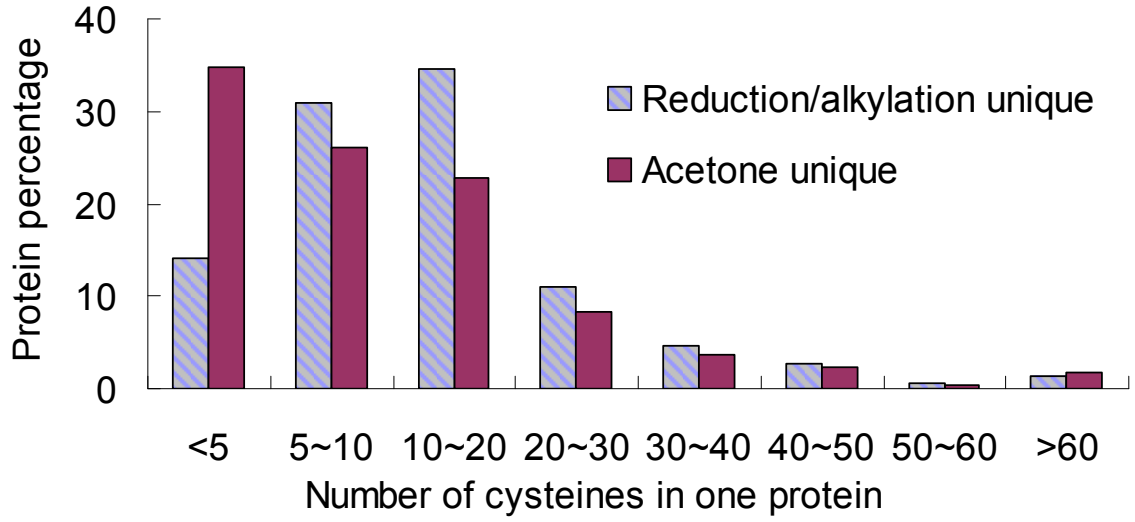


Figure 2.6 Comparison of cystein residues in one protein from two sequential precipitation techniques.

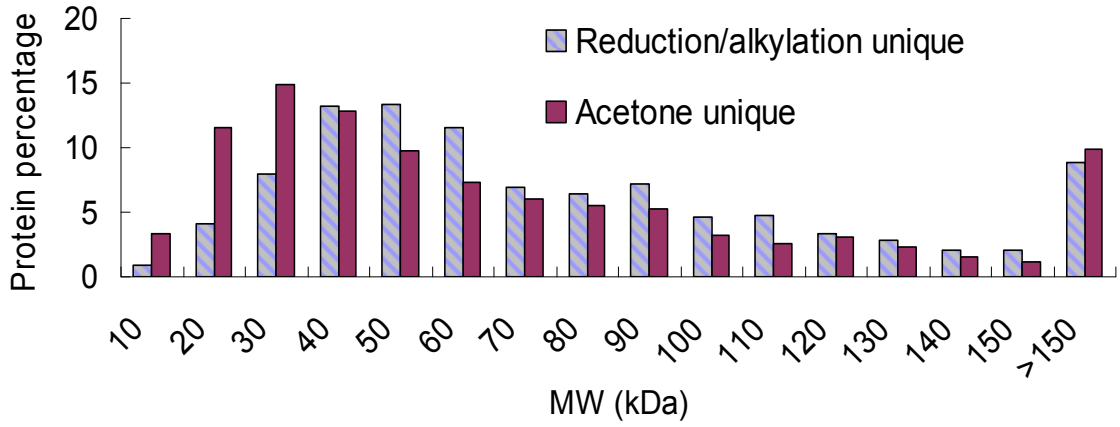


Figure 2.7 Comparison of molecular mass (MW) of identified proteins from two sequential precipitation techniques.

because large proteins tend to have more cysteine residues.

GRAVY is a commonly used parameter to gauge the hydropathicity of proteins or peptides. In this work, proteins were categorized into four groups according to their GRAVY indices: proteins with GRAVY indices lower than -0.5 were considered hydrophilic, higher than -0.5 but lower than 0 were considered mildly hydrophilic, higher than 0 but lower than 0.5 were considered mildly hydrophobic, and higher than 0.5 were considered hydrophobic. The distribution of GRAVY indices for proteins found in each fraction is summarized in Figure 2.8.

Hydrophilic and mild hydrophilic proteins were found to be the majority of the proteins found in each fraction (83.9% out of 5011 proteins in the whole membrane protein extract, 87.9% out of 1340 unique proteins in the NH_4HCO_3-assisted solubilization fraction, and 82.4% out of 1332 unique proteins in the methanol and SDS-assisted solubilization fractions). This is consistent with the fact that most of the proteins present in the protein extract are actually not the membrane proteins (see below). The methanol and SDS-assisted solubilization method appears to be better at recovering the hydrophobic proteins, since the methanol and SDS-assisted solubilization fraction contains a higher percentage of mildly hydrophobic and hydrophobic proteins (17.6% out of 1332 unique proteins) compare to the NH_4HCO_3-assisted solubilization fraction (12.1% out of 1340 unique proteins) (Figure 2.6). These results indicate that methanol and SDS-assisted solubilization show better performance in solubilizing higher hydrophobicity proteins. This result is consistent with previous studies[6] and gives better differentiation. However, for the two protein fractions generated from sequential precipitation by reduction/alkylation and acetone, no significant difference could be observed in terms of GRAVY (data not shown). This is because the GRAVY value of a protein is calculated by adding the hydropathy value for each amino acid residue together and dividing by the length of the protein sequence,[13] and does not take into account local deviations from the average, which may play an important role in the solubility properties of proteins. In that case, the hydrophobic domains in a protein, which was conserved by disulfide bonds between cysteines can not be necessarily reflected in the protein GRAVY value.

2.3.4 Subcellular location

Subcellular location is a key functional characteristic of proteins. Proteins can only function optimally in a specific subcellular localization; hence, the determination of subcellular location of each protein is an important step for large-scale proteomic analysis to provide reliable annotations regarding the biological functions of proteins. Using the ExPASY and Gene Ontology (GO) (http://www.geneontology.org/) database, identified

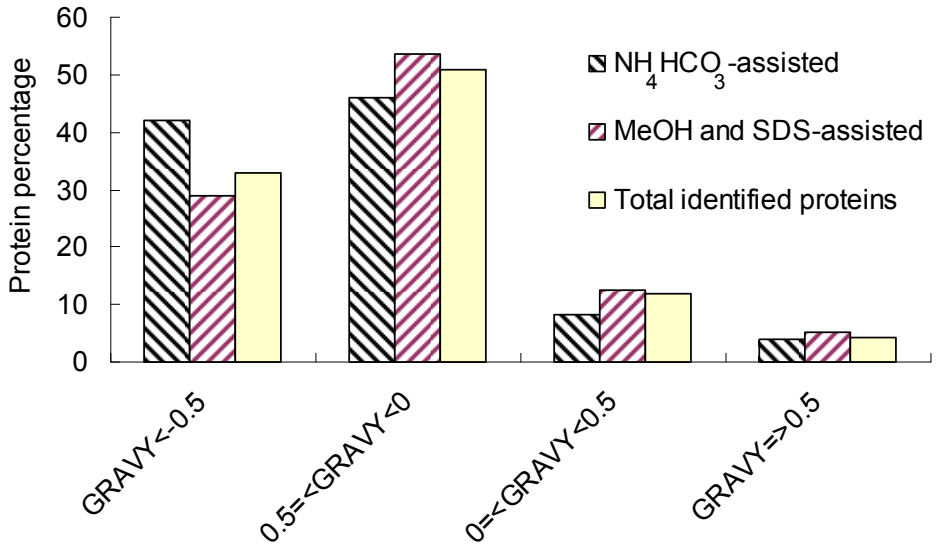


Figure 2.8 Distribution of GRAVY (Grand Average of Hydropathy) indices of identified proteins.

proteins were sorted according to their subcellular location in Figure 2.9. Then they were further examined by TMHMM Server, v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), a transmembrane domain (TMD) predicting program which can perform transmembrane helix prediction within protein primary sequences. Overall, three groups of identified proteins were examined by TMHMM. They were described by ExPASY or GO as follows: 1) no-membrane proteins, 2) membrane proteins, and 3) integral membrane proteins. The TMD distribution of the identified membrane proteins and integral membrane proteins are illustrated in Figure 2.10.

As Figure 2.9 shows, the largest proportion of proteins had a subcellular localization in the membrane catalog (44.4%), and a large portion had an unknown localization (14.5%). Furthermore, among the identified proteins not described by ExPASY or GO as being membrane proteins, it was predicted by TMHMM that 92 non-membrane proteins (1.8% out of 5011 proteins) contain one or more TMDs. Thus, overall 46.2% proteins recovered from the protein extract were predicted as membrane-associated or membrane-bound proteins. Among these, 1077 proteins (21.5% out of 5011 proteins) were assigned to be integral membrane proteins by ExPASY or GO. Since normally it has been estimated that 20-30% of the human genome encodes membrane proteins,[16] and in an un-fractionated proteome, integral membrane proteins only make up 5-15% of identified proteins,[17] it is clear that we did achieve enrichment of the membrane proteins from the MCF-7 cell line. However, the membrane protein purification procedure did not give 100% pure membrane proteins. Moreover, even the membrane proteins may not have the overall high hydrophobicity, as some portions of the proteins may be very hydrophilic. Thus, in terms of protein GRAVY, the majority of identified proteins (83.9%) are considered to be hydrophilic or mildly hydrophilic. The relatively low efficiency of membrane protein enrichment in our procedure may be due to several reasons. First, the cytosolic proteins were removed after hypo-osmotic lysis under non-denaturing conditions. Therefore, proteins that have stable associations with membrane proteins (e.g., membrane bound organellar associated proteins, cytoskeletal elements) may not be separated from the membrane protein extract. Thus, simple physical entrapment can bring soluble proteins into the pellet.[2] Second, during the cell lysis, the membrane proteins were released and re-suspended by using Triton X-100 buffer, and the insoluble materials were centrifuged down and removed. Therefore, portions of hydrophobic membrane proteins may be discarded with the pellet, rather than being solubilized into the Triton X-100 supernatant, which means not all membrane proteins could be extracted with high efficiency. Furthermore, after tryptic digestion of SDS-assisted solubilization fractions, trace amount of precipitation was observed. This precipitate should be mainly composed of hydrophobic proteins, which could not be recovered by our current technique. All the above issues need to be addressed in future work.

In terms of the TMD distribution, as Figure 2.10 shows, 1336 (60.1%) out of 2224

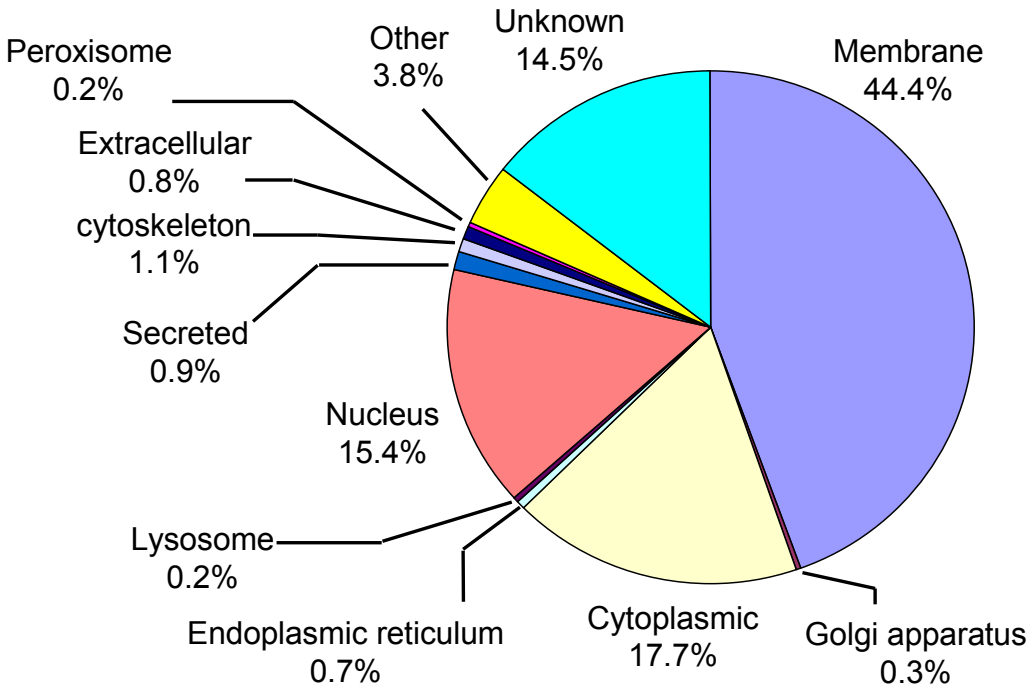


Figure 2.9 Subcellular localizations of the MCF-7 membrane extract.

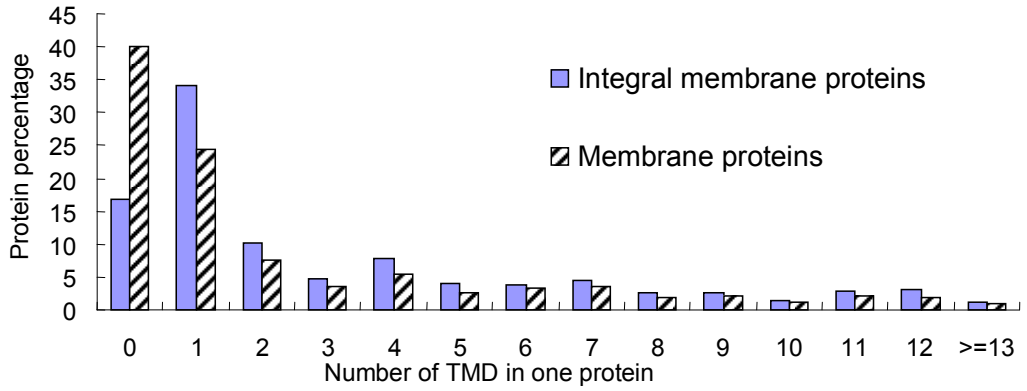


Figure 2.10 Transmembrane domain (TMD) distribution of identified integral membrane proteins and membrane proteins.

identified membrane proteins contain one or more TMD, and 896 (83.2%) out of 1077 identified integral membrane proteins contain one or more TMD. Among these, 629 highly hydrophobic membrane proteins with three or more predicted TMDs were identified.

In addition, the NH_4CO_3-assisted solubilization method recovered 437 membrane proteins (32.6 % out of 1340 unique proteins) according to ExPASY or GO, whereas the MeOH and SDS-assisted solubilization method recovered 588 membrane proteins (44.1% out of 1332 unique proteins) (data now shown). This result is consistent with our previous statement that the MeOH and SDS-assisted solubilization method recovers more hydrophobic proteins.

2.3.5 Comparison to other work

The MCF-7 cell line has been widely used for cancer biology studies and thus there is a great interest in generating a comprehensive proteome profile of this cell type. There are several reports on proteome profiling of MCF-7 cells (Table 2.2). However, most of the previous work was not focused on the membrane subproteome profiling. For example, Sarvaiya et al. identified a total of 3873 proteins from the whole cell lysate of MCF-7 cells using 2D-LC MS/MS,[18] which was, to our knowledge, the largest number of proteins identified in MCF-7 cells reported in the literature. The protein identification number is greater in our work, even though only the membrane fraction of MCF-7 cells was investigated. It can be expected that, using the protocols described in this chapter, we could achieve a much greater proteome coverage if the whole cell lysate of the MCF-7 cell line was studied.

Table 2.2 List of previously published MCF-7 profiling works.

Cellular compartments	Separation methods	Instruments	Number of proteins identified	Number of peptides identified	Search engine and confidence level	References
Whole cell lysate	RPLC	ESI-QTOF	726	2659	Mascot (95%)	10
Whole cell lysate	RPLC	MALDI-MS ESI-QTOF	N/A	2300	Mascot (N/A)	19
Membrane fraction	2-DE	MALDI-TOF	1000	N/A	Aldente software (N/A)	20
Whole cell	2-DE	MALDI-MS	764	N/A	MS-Fit	21

lysate					(N/A)	
Whole cell lysate	SCX-RPLC	ESI-QTOF	2911	12417	Mascot (95%)	22
Whole cell lysate	RPLC	ESI-QTOF	665	2161	Mascot (95%)	23
Whole cell lysate	SCX-RPLC	Linear ion trap	2286	N/A	Bioworks 3.3 software (99.9%)	24
Whole cell lysate	2-DE	MALDI -TOF	947	N/A	Mascot (N/A)	25
Whole cell lysate	RPLC	ESI-LTQ	1859	6597	Bioworks 3.3 software (99.9%)	26
Whole cell lysate	SCX-RPLC	ESI-LTQ ion trap	3873	14217	Bioworks 3.2 software (95%)	18
Whole cell lysate	SDS-PAGE RPLC	ESI-QSTAR -XL QTOF	488 (with MDA-MB -231 cell line)	3304 (with MDA-MB -231 cell line)	Mascot (95%) X! Tandem (95%)	27
Membrane fraction	SCX-RPLC	ESI-ion trap	430	N/A	Bioworks 3.1 software (N/A)	28
Whole cell lysate	Sucrose gradient density centrifugation SDS-PAGE RPLC	LTQ -Orbitrap	2184	15527	SEQUEST (95%)	29
Whole cell lysate	Microfluidic LC chip	ESI-LTQ ion trap	40-50	N/A	Bioworks 3.3 software (99.9%)	30

2.4 Conclusions

We have developed and applied a shotgun technique for the analysis of the MCF-7 human breast cancer membrane proteome using a combination of protein and peptide separations to simply the proteome sample, followed by tandem MS analysis. A sequential protein precipitation and solubilization technique was employed to reduce the protein complexity

of the proteome sample. The sequential solubilization technique appears to separate proteins mainly according to their hydrophobicity, because it was observed that there were significant differences in the GRAVY of identified proteins from various solvent-assisted solubilization fractions. Overall, a total of 5011 proteins and 24261 peptides were identified with high confidence and a low false positive rate (0.53%). Among these, 46.2% of the identified proteins were predicted as membrane-associated or membrane-bound proteins, and 21.5% of the membrane proteins were predicted as integral membrane proteins. Compared to other recently reported MCF-7 proteome profiling work,[10, 18-32] these numbers can be considered as the state-of-the-art in terms of the coverage of the membrane proteome, suggesting that our technique is relatively efficient to recover and identify hydrophobic proteins. At this stage, the mechanisms of protein precipitation by reduction/alkylation have not been fully studied. To optimize the performance of the protocol, further experiments are needed to investigate the effects of various parameters during reduction and alkylation (e.g., pH, salt concentration) on the amount of protein precipitate and the separation efficiency. In addition, since the enrichment of membrane proteins from MCF-7 cells was not ideally efficient, more research efforts should be devoted to the development of improving membrane protein extraction technique.

2.5 References

1. Weinglass, A. B.; Whitelegge, J. P.; Kaback, H. R., Integrating mass spectrometry into membrane protein drug discovery. *Current Opinion in Drug Discovery & Development* **2004,** 7, (5), 589-599.
2. Santoni, V.; Molloy, M.; Rabilloud, T., Membrane proteins and proteomics: Un amour impossible? *Electrophoresis* **2000,** 21, (6), 1054-1070.
3. Fischer, F.; Wolters, D.; Rogner, M.; Poetsch, A., Toward the complete membrane proteome - High coverage of integral membrane proteins through transmembrane peptide detection. *Molecular & Cellular Proteomics* **2006,** 5, (3), 444-453.
4. Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R.; Li, L., Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *Journal of Proteome Research* **2008,** 7, (8), 3583-3590.
5. Zhong, H. Y.; Marcus, S. L.; Li, L., Microwave-assisted acid hydrolysis of proteins combined with liquid chromatography MALDI MS/MS for protein identification. *Journal of the American Society for Mass Spectrometry* **2005,** 16, (4), 471-481.
6. Wang, N.; MacKenzie, L.; De Souza, A. G.; Zhong, H. Y.; Goss, G.; Li, L., Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *Journal of Proteome*

Research **2007,** 6, (1), 263-272.
7. Zhang, N.; Chen, R.; Young, N.; Wishart, D.; Winter, P.; Weiner, J. H.; Li, L., Comparison of SDS- and methanol-assisted protein solubilization and digestion methods for Escherichia coli membrane proteome analysis by 2-D LC-MS/MS. *Proteomics* **2007,** 7, (4), 484-493.
8. Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D., Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *Journal of Proteome Research* **2002,** 1, (4), 351-360.
9. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
10. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical Chemistry* **2008,** 80, (12), 4696-4710.
11. Higdon, R.; Hogan, J.; Van Belle, G.; Kolker, E., Randomized sequence databases for tandem mass spectrometry peptide and protein identification. *Omics-A Journal of Integrative Biology* **2005,** 9, (4), 364-379.
12. Qian, W.; Liu, T.; Monroe, M.; Strittmatter, E.; Jacobs, J.; Kangas, L.; Petritis, K.; Camp, D.; Smith, R., Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: The human proteome. *Journal of Proteome Research* **2005,** 4, (1), 53-62.
13. Kyte, J.; Doolittle, R. F., A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology* **1982,** 157, (1), 105-132.
14. Li, L.; Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R., Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *Journal of Proteome Research* **2008,** 7, (8), 3583-3590.
15. Zhou, N. E.; Kay, C. M.; Hodges, R. S., Disulfide Bond Contribution to Protein Stability - Positional Effects of Substitution in the Hydrophobic Core of the 2-Stranded Alpha-Helical Coiled-Coil. *Biochemistry* **1993,** 32, (12), 3178-3187.
16. Stevens, T. J.; Arkin, I. T., Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins-Structure Function and Genetics* **2000,** 39, (4), 417-420.
17. Speers, A. E.; Wu, C. C., Proteomics of integral membrane proteins-theory and application. *Chemical Reviews* **2007,** 107, (8), 3687-3714.
18. Sarvaiya, H. A.; Yoon, J. H.; Lazar, I. M., Proteome profile of the MCF7 cancer cell line: a mass spectrometric evaluation. *Rapid Communications in Mass Spectrometry* **2006,** 20, (20), 3039-3055.

19. Tammen, H.; Kreipe, H.; Hess, R.; Kellmann, M.; Lehmann, U.; Pich, A.; Lamping, N.; Schulz-Knappe, P.; Zucht, H. D.; Lilischkis, R., Expression profiling of breast cancer cells by differential peptide display. *Breast Cancer Research and Treatment* **2003,** 79, (1), 83-93.
20. Ruan, Y. S.; Wan, M. X., An optimized procedure for solubilization, reduction, and transfer of human breast cancer membrane-enriched fraction by 2-DE. *Electrophoresis* **2007,** 28, (18), 3333-3340.
21. Chen, S. T.; Pan, L. T.; Tsai, Y. C.; Huang, C. M., Proteomics reveals protein profile changes in doxorubicin - treated MCF-7 human breast cancer cells. *Cancer Letters* **2002,** 181, (1), 95-107.
22. Wang, N.; Xie, C. H.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
23. Wang, N.; Xu, M. G.; Wang, P.; Li, L., Development of Mass Spectrometry-Based Shotgun Method for Proteome Analysis of 500 to 5000 Cancer Cells. *Analytical Chemistry* 82, (6), 2262-2271.
24. Yang, X.; Lazar, I. M., MRM screening/biomarker discovery with linear ion trap MS: a library of human cancer-specific peptides. *BMC Cancer* **2009,** 9, 11.
25. Ou, K.; Yu, K.; Kesuma, D.; Hooi, M.; Huang, N.; Chen, W.; Lee, S. Y.; Goh, X. P.; Tan, L. K.; Liu, J.; Soon, S. Y.; Rashid, S. B. A.; Putti, T. C.; Jikuya, H.; Ichikawa, T.; Nishimura, O.; Salto-Tellez, M.; Tan, P., Novel breast cancer biomarkers identified by integrative proteomic and gene expression mapping. *Journal of Proteome Research* **2008,** 7, (4), 1518-1528.
26. Armenta, J. M.; Perez, M.; Yang, X.; Shapiro, D.; Reed, D.; Tuli, L.; Finkielstein, C. V.; Lazar, I. M., Fast proteomic protocol for biomarker fingerprinting in cancerous cells. *Journal of Chromatography A* 1217, (17), 2862-2870.
27. Lawlor, K.; Nazarlan, A.; Lacomis, L.; Tempst, P.; Villanueva, J., Pathway-Based Biomarker Search by High-Throughput Proteomics Profiling of Secretomes. *Journal of Proteome Research* **2009,** 8, (3), 1489-1503.
28. Xiang, R.; Shi, Y.; Dillon, D. A.; Negin, B.; Horvath, C.; Wilkins, J. A., 2D LC/MS analysis of membrane proteins from breast cancer cell lines MCF7 and BT474. *Journal of Proteome Research* **2004,** 3, (6), 1278-1283.
29. Qattan, A. T.; Mulvey, C.; Crawford, M.; Natale, D. A.; Godovac-Zimmermann, J., Quantitative Organelle Proteomics of MCF-7 Breast Cancer Cells Reveals Multiple Subcellular Locations for Proteins in Cellular Functional Processes. *Journal of Proteome Research* 9, (1), 495-508.
30. Armenta, J. M.; Dawoud, A. A.; Lazar, I. M., Microfluidic chips for protein differential expression profiling. *Electrophoresis* **2009,** 30, (7), 1145-1156.
31. Goncalves, A.; Charafe-Jauffret, E.; Bertucci, F.; Audebert, S.; Toiron, Y.; Esterni, B.; Monville, F.; Tarpin, C.; Jacquemier, J.; Houvenaeghel, G.; Chabannon, C.; Extra, J. M.;

Viens, P.; Borg, J. P.; Birnbaum, D., Protein profiling of human breast tumor cells identifies novel biomarkers associated with molecular subtypes. *Molecular & Cellular Proteomics* **2008,** 7, (8), 1420-1433.
32. You, J.; Wang, L. W.; Saji, M.; Olesik, S. V.; Ringel, M. D.; Lucas, D. M.; Byrd, J. C.; Freitas, M. A., High-sensitivity TFA-free LC-MS for profiling histones. *Proteomics* 11, (16), 3326-3334.

Chapter 3

Comprehensive Proteome Profiling of *E. coli* K12 Cell Line by Shotgun Proteomic Strategy*

3.1 Introduction

The bacterium *Escherichia coli* (*E. coli*) is one of the most popular model organisms for microbiology studies, and also plays important role in biological engineering (e.g., acting as the host for the production of recombinant proteins).[1] Although *E. coli* was one of the first organisms having its genome fully sequenced,[2] its proteome has not been fully investigated. This is mainly caused by two reasons. First, the relative abundance of proteins varies in tremendous extent due to the expression level difference and the regulation of proteins. The low abundant proteins are prone to be suppressed, during the analytical detection process, by the proteins expressed in high amounts. Second, the post-translational modifications (PTMs) further complicate the proteome, making it difficult to characterize proteins, compared to the analysis of the genome or DNA. There have been a number of publications related to the analysis of the *E. coli* proteome, utilizing gel-base and/or solution-based techniques.[3-5] For example, Hunt et al. identified 1147 unique proteins from the membrane fraction of *E. coli* using 2D-LC MS/MS,[5] which was the largest number of proteins identified in *E. coli* in the literature. However, none of these results reached the same scale as the genome analysis. Therefore, the first step of our goal is to develop a protocol that is able to identify all the proteins present in a proteome sample at a certain time point. This will pave a solid foundation from which we carry out the future work in the characterization of PTMs, quantification of the whole proteome and investigating the protein-protein interactions. The ultimate goal of this research is to develop and apply powerful techniques to examine the entire proteome of a biological sample.

As described in Chapter 2, sequential protein precipitation and solubilization combined with 2D-LC MS/MS was developed to improve the analysis of MCF-7 membrane proteome. The MCF-7 proteome can potentially contain over 20,000 proteins based on the estimation of the number of genes present in human cells. Analysis of the whole proteome of human cells is a major challenge. In this work, the *E. coli* K12 was chosen as

---

* Dr. Nan Wang contributed partially to sample preparation, data acquisition and data processing of this work. Dr. Joe Weiner performed bioinformatic characterization of *E. coli* proteins.

there are only about 4300 genes predicted from the genome of the *E. coli* K12 strain.[6-10] Since the proteins expressed in the cell line and their relative abundances can vary according to different culture environments, it is difficult to predict the size of the total proteome under a certain growth condition. However, the predicted gene number in *E. coli* K12 is about 4300[6-10] and thus the upper limit of 4300 possibly expressed proteins isexpected in the *E. coli* K12 cells, which should provide a useful metrics to gauge the detectability of a proteome analysis technique.

In this work, the *E. coli* K12 cell line was cultured in a rich media to the early stationary growth phase. Three cellular compartments, namely cytoplasm, peripheral membrane, and integral membrane, were separated for individual sub-proteome analysis to maximize the overall proteome coverage. Similar technique was also reported for the fractionation of mammalian cells.[11] The cytoplasm and peripheral membrane fractions were sequentially precipitated as described in Chapter 2. All the fractions were subjected to sequential solubilization by ammonium bicarbonate, methanol and SDS, followed by 2D-LC MS/MS analysis. Low molecular weight proteins were also enriched by a molecular weight cutoff filter to further improve the proteome coverage. The same identification protocol was applied to the analysis of the low molecular weight enriched samples. The results generated from these experiments will be described in detail.

3.2 Experimental

3.2.1 Chemicals and reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), sodium carbonate (Na_2CO_3) and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC/MS-grade water, acetone, formic acid, methanol (MeOH), and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada). The BCA assay kit was from Pierce (Rockford, IL).

3.2.2 Cell culture and protein extraction

Escherichia coli K-12 (*E. coli,* ATCC 47076) was from the American Type Culture Collection. A single *E. coli* K12 colony was used to inoculate 50 mL of LB broth (BBL, Becton Dickinson). The culture was incubated overnight with shaking at 37 ˚C. This saturated culture (12.5 mL) was added to 500 mL of growth medium in a 2-L baffled Erlenmeyer flask. Cells were harvested in the early stationary phase by centrifugation at 11 300 *g* for 15 min at 4 ˚C, resuspended, washed in 150 ml PBS buffer, and collected by centrifugation at 9,820 *g* for 15 min at 4 ˚C.

The *E. coli* cells were suspended in 30 mL PBS buffer and three tablets of Roche mini protease inhibitor cocktail were added. The suspension was passed twice through a French press (Aminco Rochester, NY) using a rapid fill kit at 20,000 psi. Unbroken cells and debris were removed by centrifugation at 7,740 *g* for 15 min at 4 ˚C. The supernatant was collected for the next step fractionation.

Three cellular compartments, namely cytoplasm, peripheral membrane and integral membrane, were isolated using a centrifugation and carbonate fractionation procedure[12] with some modifications. The lysate was centrifuged in a Beckman Type 55.2Ti rotor for 55 min at 118,000 *g*, and the pellet was resuspended, washed in 20 mL of 50 mM NH_4HCO_3, and collected by centrifugation in a Beckman Type 55.2Ti rotor for 40 min at 116,811 *g*. The supernatants generated from the above centrifugations were collected as the cytoplasm fraction. The pellet containing membrane proteins was suspended again in 11 mL of 50 mM ammonium NH_4HCO_3, and was then diluted with ice-cold 100 mL of 100 mM Na_2CO_3 (pH 11.0). The solution was stirred slowly in an ice bath for 1 h to extract peripheral membranes. The extract was divided equally into two tubes, and centrifuged in a Beckman Type 45Ti rotor for 60 min at 115,000 *g*. The supernatant was collected as the peripheral membrane fraction. Each pellet was gently rinsed with 5 mL of water, suspended in 7 mL of 50 mM MOPS buffer (pH 7.3) and transferred to an 8 mL tube. The tubes were centrifuged in a Beckman Type 70.1Ti rotor at 115 000 *g* for 25 min. The pellet was collected as the integral membrane proteins. All the fractions were then stored at -80 ˚C for future use.

3.2.3 Sequential protein precipitation

Sequential precipitation techniques were employed as described in Chapter 2 to the cytoplasm and peripheral membrane fractions, respectively. 450 mM dithiothreitol (DTT) was added to the supernatant with the final concentration of 30 mM. The pH of the solution was adjusted to 8.0 by adding 1 M NH_4HCO_3. The solution was incubated for 1 h

at 37°C to reduce disulfide bonds. Iodoacetamide (IAA) at 450 mM was added to the DTT/IAA ratio of 1:2 (mole/mole) and the solution was left to stand for 1 h at room temperature in the dark so that all reduced disulfide bonds could be blocked by carbamidomethylation reaction. The cloudy solution was centrifuged at the setting of 20,817 g for 10 min at 4 °C. The pellet was stored at -80 °C as the reduction/alkylation fraction, whereas the supernatant was then transferred to a new vial and gradually mixed with four volumes of cold acetone (-80 °C) (with intermittent vortexing) to precipitate the proteins and remove detergents and other chemicals. The mixture was kept at -20 °C overnight and centrifuged as described above and the supernatant was decanted and properly disposed. The remaining acetone was evaporated at room temperature and the protein precipitate was stored at -80 °C for future use.

3.2.4 Sequential protein solubilization

Before sequential protein solubilization, the integral membrane fraction was carefully suspended in four volumes of cold acetone (-80 °C) to remove detergents and other chemicals. The mixture was kept at -20 °C overnight and centrifuged at the setting of 20,817 g for 10 min at 4 °C. The supernatant was decanted and remaining acetone was evaporated at room temperature. The pellet, together with four fractions of protein precipitates generated from the cytoplasm and peripheral fractions, was subjected to sequential protein solubilization. They were performed as previously described with minor changes.[13] The sequential solubilization steps were: (1) ammonium bicarbonate-assisted solubilization, (2) methanol-assisted solubilization,[14, 15] and (3) SDS-assisted solubilization.[14-16] For each of the precipitated protein fractions, ammonium bicarbonate (25 mM, pH 8.0) was first added to the pellets. Intermittent vortexing was applied for 2 h, and then centrifuged at 20,817 g for 10 min at 4 °C (the same vortexing and centrifugation steps were applied in the subsequent solubilization processes). The remaining pellet was resuspended in 60% methanol, with sufficient vortexing for 2 h. The solubilized proteins were transferred into a different vial, and MeOH in the solution was evaporated using a SpeedVac (Thermo Savant, Milford, MA). Finally, 2% SDS was employed to dissolve the remaining pellet. The same sequential solubilization procedure was applied to all of the sequential protein precipitate fractions. The concentration of each protein fraction was determined by BCA assay. For some of the fractions, since the pellet amount was small，methanol-assisted solubilization step was skipped. Table 3.1 summarizes the 13 proteins fractions generated from the three cellular compartments and subsequent sequential protein precipitation and solubilization.

Table 3.1 Summary of the 13 protein fractions generated from the three cellular compartments and subsequent sequential protein precipitation and solubilization.

	Cytoplasm fraction	Peripheral membrane fraction	Integral membrane fraction
Reduction/alkylation fraction	NH_4HCO_3-solubilized	NH_4HCO_3-solubilized	N/A
	SDS-solubilized	MeOH-solubilized	
		SDS-solubilized	
Acetone fraction	NH_4HCO_3-solubilized	NH_4HCO_3-solubilized	NH_4HCO_3-solubilized
	MeOH-solubilized	SDS-solubilized	MeOH-solubilized
	SDS-solubilized		SDS-solubilized

3.2.5 In-solution digestion

Before tryptic digestion, MeOH in the MeOH-solubilized protein samples was evaporated using SpeedVac (Thermo Savant, Milford, MA). The SDS-solubilized protein samples were subjected to a 20-fold dilution. The three protein fractions generated from the integral membrane extract were subjected to reduction by DTT and alkylation by IAA as described above. Trypsin solution was added into the supernatant at an enzyme/protein ratio of 1:40, and the digestion was conducted at 37 °C for two days with 10% (of the amount added for the first day) more fresh trypsin added before the second day digestion. Reactions were stopped by adding 10% TFA to adjust the pH to 2.5. The digestion solutions were stored at -80 °C until further analysis.

3.2.6 Strong cation exchange (SCX) liquid chromatography

A 2.1×250 mm highly hydrophilic polysulfoethyl A column (5 μm diameter, 300 Å pore, Poly LC, Columbia, MD) was used for the strong-cation exchange separation of the tryptic peptides. Solvent A (5 mM KH_2PO_4, 20% ACN, pH 2.7) and solvent B (solvent A with 500 mM KCl) were used to develop a salt gradient (0 to 4% B for 1 min, 4 to 20% B for 16 min, 20 to 60% B for 22 min, ramped to 100% B in 6 min, held for another 10 min). One min or 2 min fractions were collected based on sample loading amount, and then adjacent fractions were pooled according to the UV absorbance values in the chromatograms.

3.2.7 Peptide desalting and quantification by RPLC

Desalting and quantification were carried out using an Agilent 1100 HPLC system (Palo Alto, CA) with a 4.6 mm × 50 mm Polaris C18 A column with 3 μm particles and 300 Å pore size (Varian, CA). The eluted peptides were monitored and quantified using a UV detector operated at 214 nm using a method previously described.[17] After loading of the peptide sample, the column was flushed with mobile phase A (0.1% TFA in water) and the salts were effectively removed. Subsequently, the concentration of mobile phase B (0.1% TFA in ACN) in the mobile phase was step-wise increased to 85% to ensure complete elution of the peptide fractions from the column. During the peptide elution process a chromatographic peak was produced and, based on the peak area, the amount of peptides was determined. BSA digests of various amounts were used as standards for the generation of a linear calibration between the peak area and the injected peptide amount. The calibration curve was generated as $y = 420x + 2056$, where y refers to the peak area of the peptide sample, 2056 refers to the peak area of the blank wash, and x refers to the peptide amount analyzed in μg. The linear range of the calibration curve was from 0.5 μg to 5.0 μg of peptides (R^2=0.998).

3.2.8 Mass spectrometric analysis

This step was performed as previously described with minor changes.[17, 18] 1 μg portions of peptides of each desalted SCX fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA) equipped with a nanoACQUITY Ultra Performance LC system (Waters). Peptide solution from each SCX fraction was injected onto a 75 μm × 100 mm Atlantis dC18 column with 3 μm particles and 100 Å pore size (Waters). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Mass spectra were acquired from m/z 300-1600 for 0.8 s, followed by 4 data-dependent MS/MS scans from m/z 50-1900 for 0.8 s each. The collision energy used to perform MS/MS was automatically varied according to the mass and charge state of the eluting peptide. A mass calibrant (i.e., lock-mass) was infused at a rate of 350 nL/min, and an MS scan was acquired for 1 s every 1 min throughout the run. Peptide precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides identified from the adjacent two SCX fractions to enable additional and less abundant peptides to be analyzed and identified.[14] An exclusion list was generated based on Mascot (Matrix Science, London, U.K.) searching results of peptides with a score 10 points above the identity threshold.

3.2.9 Protein database search

Database searches were performed as previously described with minor changes.[17, 18] Raw MS/MS data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.5 (Waters). Peptide sequences were identified by automated database searching of peak list files using the Mascot search program against the *E. coli* K12 database (4337 protein sequences). The following search parameters were selected for all database searching: enzyme, trypsin; missed cleavages, 1; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+, and 3+; fixed modification, carbamidomethyl (C); variable modifications, N-acytyl (Protein), oxidation (M), Pyro-Glu (N-term Q), Pyro-Glu (N-term E). The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, and Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses, were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at the confidence level of 99% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but different access ID numbers were also removed from the list. To gauge the false positive peptide matching rate, target-decoy search strategy was applied by searching the MS/MS spectra against the forward and reverse *E. coli* K12 proteome sequences.[19]

3.2.10 Hydropathy calculation and annotation of localization

All peptides and proteins identified were examined using the ProtParam program, available at the EXPASY Web site (http://us.expasy.org/tools/protparam.html), which allows calculation of the grand average of hydropathy (GRAVY).[20] The Gene Ontology database (http://geneontology.org/) was used to classify proteins on the basis of cellular location. Transmembrane domains (TMD) of identified proteins were predicted by TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) according to the protein primary sequences.

3.3 Results and discussion

The proteins expressed in *E. coli* cells have a wide dynamic range in their concentrations. To increase the likelihood of detecting the low abundance proteins, the cell lysates were fractionated according to their cellular properties and further separated by a

solubility-based protein fractionation method. The workflow for comprehensive profiling of the *E. coli* K12 proteome is illustrated in Figures 3.1 and 3.2. As shown in Figure 3.1, the *E. coli* K12 cell line was cultured in a rich media to the early stationary growth phase. Cells were lysed by French press and isolated into three cellular compartments: cytoplasm, peripheral membrane and integral membrane. As shown in Figure 3.2, the cytoplasm and peripheral membrane fractions were subjected to sequential protein precipitation by reduction/alkylation and cold acetone. The integral membrane fraction was washed by cold acetone thoroughly. All the protein precipitates were subjected to sequential protein solubilization by NH_4HCO_3, methanol and SDS, generating 13 protein fractions. These protein fractions were then digested by trypsin and analyzed by off-line 2D-LC MS/MS.

3.3.1 Protein identification results

Figures 3.3-3.6 summarize the numbers of unique proteins identified from the 13 protein fractions. As Figures 3.3A and B show, shotgun proteome analysis of the cytoplasm fraction resulted in the identification of 2913 proteins, representing 68% of the 4300 proteins predicted to be expressed in the *E. coli* K12 cell line. In the reduction/alkylation fraction shown in Figure 3.3A, we identified 1870 unique proteins from the NH_4HCO_3-solubilized fraction and 2254 unique proteins from the SDS-solubilized fraction, resulting in 2595 unique proteins identified from the reduction/alkylation fraction. In the acetone fraction shown in Figure 3.3B, we detected 836 unique proteins from the NH_4HCO_3-solubilized fraction, 2057 unique proteins from the MeOH-solubilized fraction, and 911 unique proteins from the SDS-solubilized fraction. In total, 2276 unique proteins were identified from the acetone fraction. As Figures 3.4A and B show, analysis of the peripheral membrane fraction resulted in the identification of 2579 unique proteins. In the reduction/alkylation fraction shown in Figure 3.4A, we identified 1453 unique proteins from the NH_4HCO_3-solubilized fraction, 1215 unique proteins from the MeOH-solubilized fraction and 1770 unique proteins from the SDS-solubilized fraction. In total, 2069 unique proteins were identified from the reduction/alkylation fraction. In the acetone fraction shown in Figure 3.4B, we detected 1355 unique proteins from the NH_4HCO_3-solubilized fraction, and 1634 unique proteins from the SDS-solubilized fraction, resulting in 1961 unique proteins identified from the acetone fraction. As Figure 3.5 shows, in the integral membrane fraction, we identified 106 unique proteins from the NH_4HCO_3-solubilized fraction, 101 unique proteins from the MeOH-solubilized fraction and 1650 unique proteins from the SDS-solubilized fraction, resulting in the identification of 1684 unique proteins. The overlaps between the two sequential protein precipitation fractions, namely the reduction/alkylation and acetone fractions, were 67% for the cytoplasm and 56% for the peripheral membrane extract. These numbers are slightly higher than the overlap between the two sequential protein precipitation fractions in

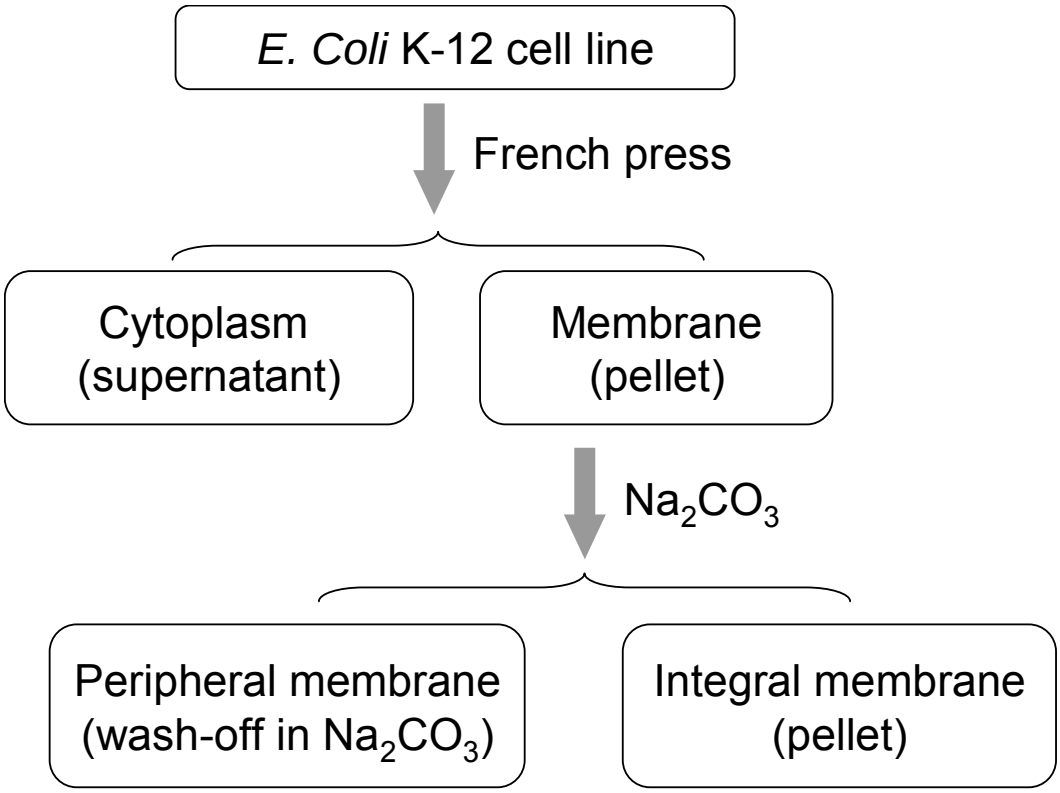


Figure 3.1 Workflow of cellular compartments fractionation for *E. coli* K12 cell line.

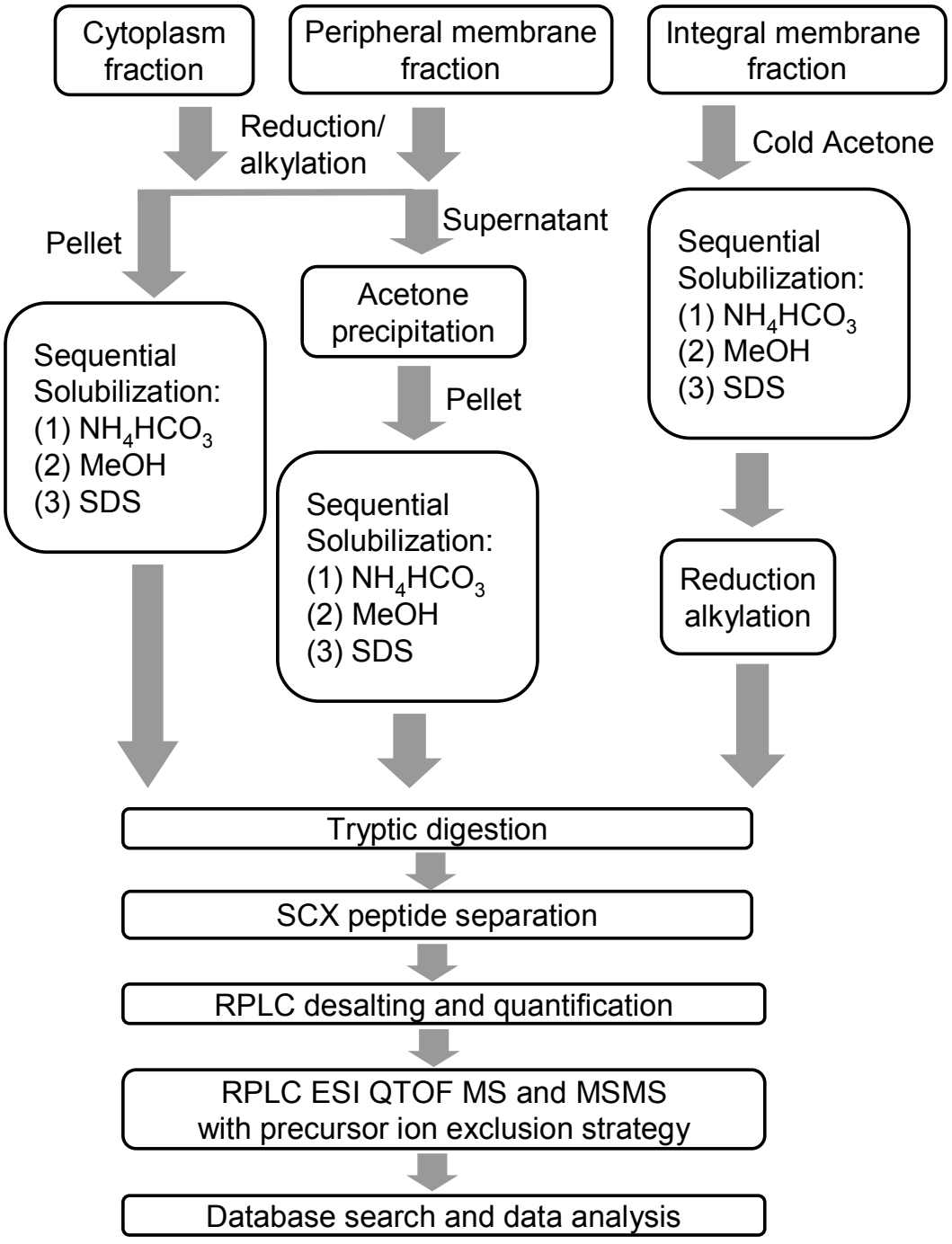
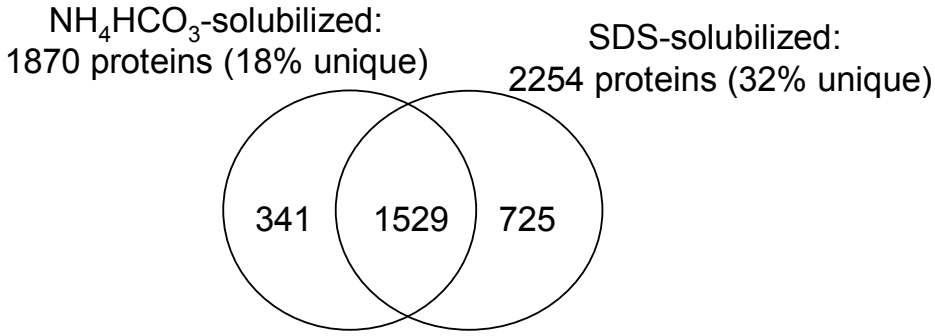


Figure 3.2 Workflow of sequential protein precipitation and solubilization followed by 2D-LC MS/MS analysis for *E. coli* K12 cell line.

(A).



(B).

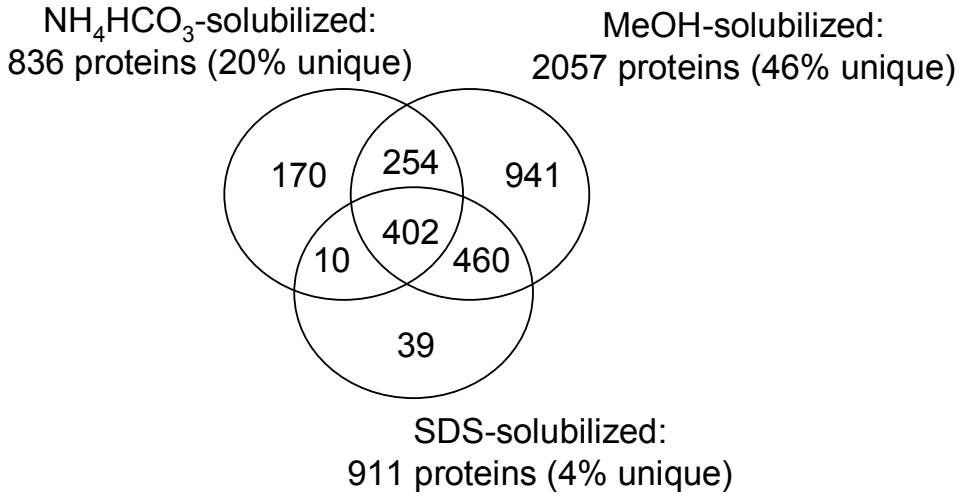
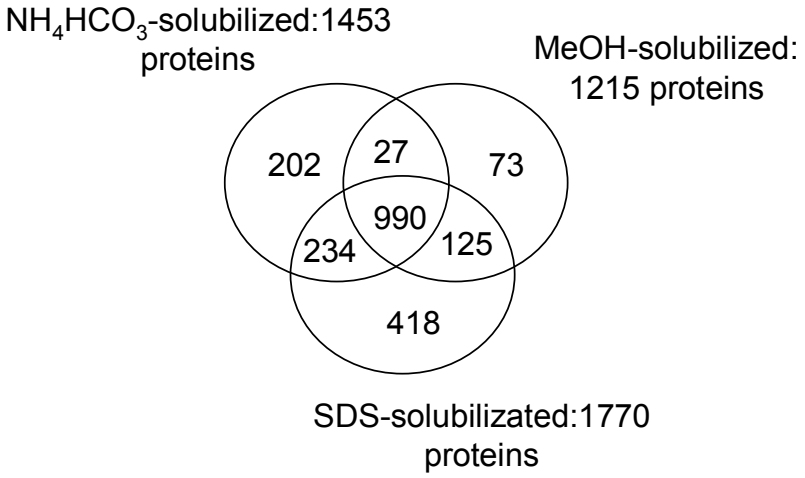


Figure 3.3 (A) Venn diagram of the 2595 unique proteins identified from the cytoplasm_reduction/alkylation fraction. (B) Venn diagram of the 2276 unique proteins identified from the cytoplasm_acetone fraction, with a total number of 2913 protein identified from the cytoplasm.

(A).



(B).

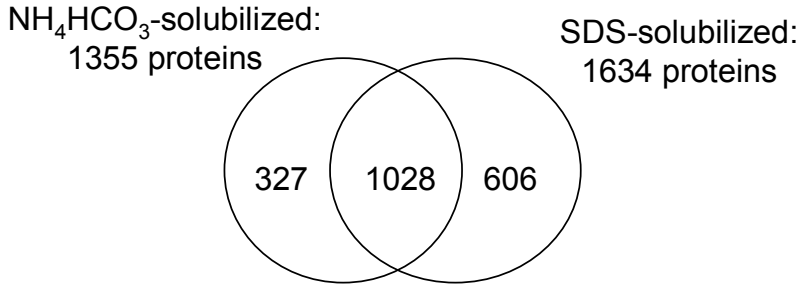


Figure 3.4 (A) Venn diagram of the 2069 unique proteins identified from the peripheral membrane extract_reduction/alkylation fraction. (B) Venn diagram of the 1961 unique proteins identified from the peripheral membrane extract_acetone fraction, with a total number of 2579 protein identified from the cytoplasm.

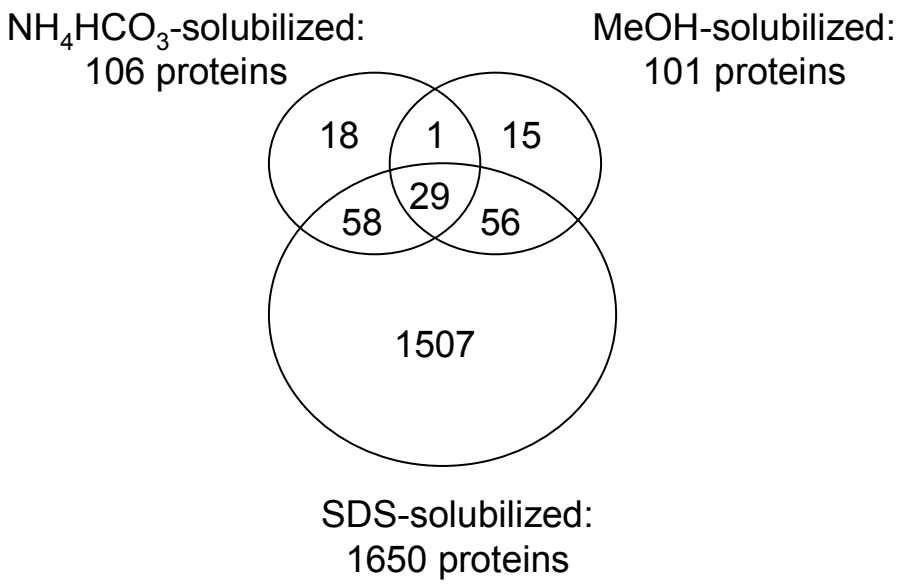


Figure 3.5 Venn diagram of the 1684 unique proteins identified from the integral membrane extract.

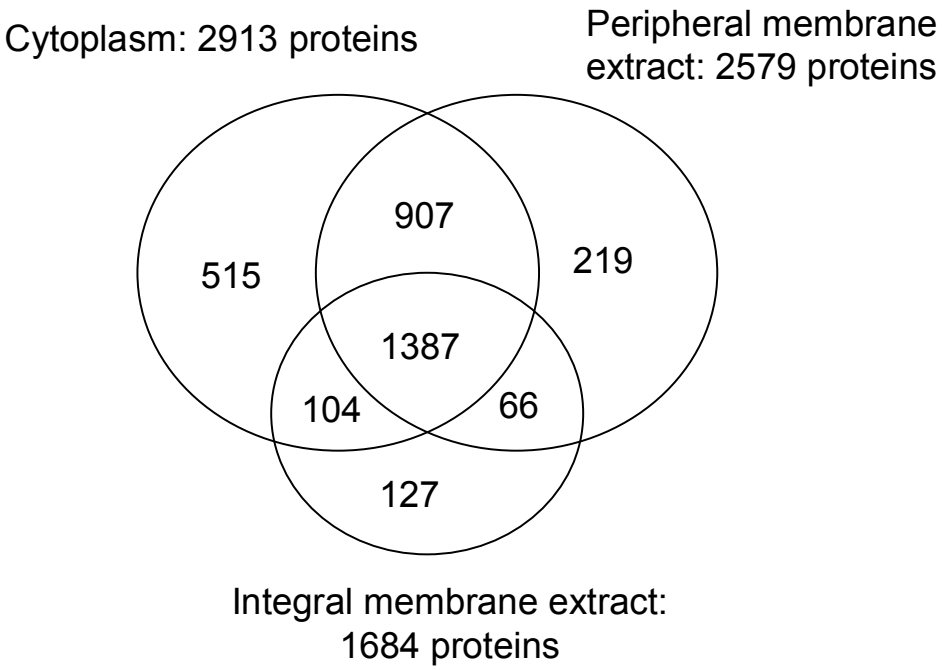


Figure 3.6 Venn diagram of the 3325 unique proteins identified from the *E. coli* K12 cell line.

Chapter 2 (51%), which may be due to the much smaller size of the *E. coli* proteome, compared to the MCF-7 proteome.

As Figure 3.6 illustrates, 3325 unique proteins were identified from the three cellular compartments of the *E. coli* K12 cell line. There were 1387 common proteins detected from all of the three compartments. The overlap of the identified proteins between the cytoplasm and the peripheral membrane fractions was (1387+907). The overlap between the peripheral membrane and the integral membrane fractions was (1387+66). And the overlap between the cytoplasm and the integral membrane fractions was (1387+104). These results indicate that the proteins in the cytoplasm and the peripheral membrane fractions share more similar physiochemical properties compared to the proteins in the integral membrane fraction. It should be noted that the total number of unique proteins identified from the integral membrane fraction was much smaller than the other two fractions, which may be due to two reasons. First, the integral membrane fraction contains a smaller number of unique proteins than the other fractions. Second, there are more hydrophobic proteins in the integral membrane fraction, which are more challenging to be detected than the more hydrophilic proteins because of their low solubility.

3.3.2 Properties of identified proteins

The 3325 unique proteins identified from the cell extract represent 77% of the predicted proteins in the *E. coli* K12 cell line. There are a number of reports on the analysis of the *E. coli* proteome[3, 4] and, to our knowledge, this is the most comprehensive proteome profile of *E. coli* K12. However, 975 predicted proteins were still not found. The incomplete coverage of the proteome may be due to two reasons. First, it may reflect the true situation of the proteome in the *E. coli* cells cultured under this specific condition. Second, the technique we applied may have bias towards proteins with certain properties. Therefore, to gauge the possible technical bias and better understand the physical properties of identified proteins, the distributions of hydrophobicity, transmembrane domains (TMD), subellular location and molecular weight (MW) of the identified proteins were investigated and compared with that of the 4300 predict proteins.

Figure 3.7A shows the protein distribution as a function of protein hydrophobicity, gauged by the GRAVY indices grouped into six bins. Figure 3.7B displays the distribution of the number of transmembrane domains (TMDs) of the identified proteins and the 4300 predict *E. coli* K12 proteins. There is no apparent difference in the GRAVY and TMD distribution. It should be noted that proteins identified in the integral membrane fraction appear to have higher GRAVY value and higher number of TMDs while the proteins identified in the other two cellular compartments are similar. This result is consistent with

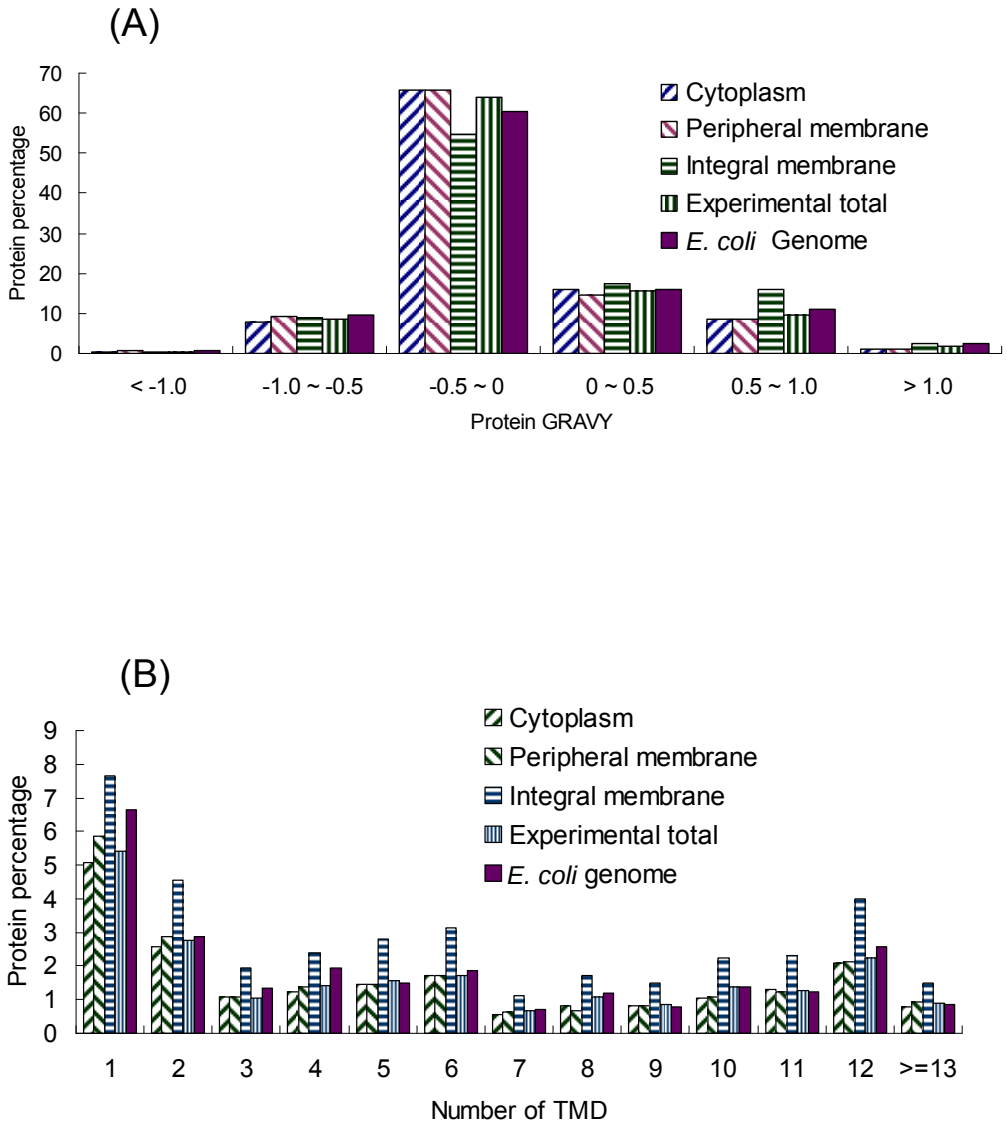


Figure 3.7 (A) GRAVY and (B) TMD distribution of identified proteins.

our expectation that the more hydrophobic proteins are present in the integral membrane fraction.

Figure 3.8A displays the distribution of proteins as a function of molecular weight (MW) or mass. Compared to the predicted proteome, proteins with MW higher than 60 kDa were well represented in the identified proteome. However, the low mass proteins are severely under-represented, especially for the MW of lower than 20 kDa (Figure 3.8B). There are 833 proteins predicted to be expressed in the MW range of 10-20 kDa and 333 proteins with MW of lower than 10 kDa. About 30% of predicted 10-20 kDa-proteins and 50% of predicted <10 kDa-proteins were not found in our list. Under-representation of the low mass proteins may be attributed to a bias introduced during the protein sample processing step. For example, the low mass proteins were not precipitated as readily as the high mass ones using acetone precipitation and thus some of the low mass proteins remain in the solution or were washed away. It is also possible that there is a bias towards the detection of high MW proteins by the shotgun proteome analysis approach in general.

To investigate if the process of acetone precipitation caused this bias on MW, the supernatant from the cytoplasm fraction after acetone precipitation was analyzed by the shotgun method as described above, except the sample was not subjected to SCX separation, and only one LC-MS/MS run was performed. Only 61 unique proteins were identified. First, compared with an average of more than 300 unique proteins routinely identified from one SCX fraction from the *E. coli* proteome extract, this number of protein identification is very low. Second, among these 61 proteins, only 3 low MW proteins were detected uniquely to the supernatant of acetone precipitation (data not shown). These observations indicate acetone precipitation is not the major cause for such bias on protein MW.

To address if the current strategy is biased towards the detection of high MW proteins, the cytoplasm and peripheral membrane fractions were combined and fractionated by enriching the low MW proteins using molecular weight cutoff filters. As shown in Figure 3.9, the proteome sample was fractionated using a combination of two MW cutoff filters into three fractions, namely MW lower than 10 kDa, MW between 10-30 kDa and MW lower than 30 kDa. The results of identified protein in the three low MW samples are presented in Figure 3.10. A total of 2241 unique protein were identified. However, compared with the 3325 unique proteins identified from the cytoplasm, peripheral and integral membrane fractions previously, only 56 new proteins were identified. Figure 3.11 A shows the distribution of the new 56 proteins as a function of molecular weight. Compared with the distribution shown in Figure 3.8A, a larger proportion of low MW proteins was identified from these low MW protein samples. On the other hand, it is not surprising that some proteins with MW of higher than 30 kDa were still detected. Firstly, because protein fragments generated from the degradation of large proteins can pass

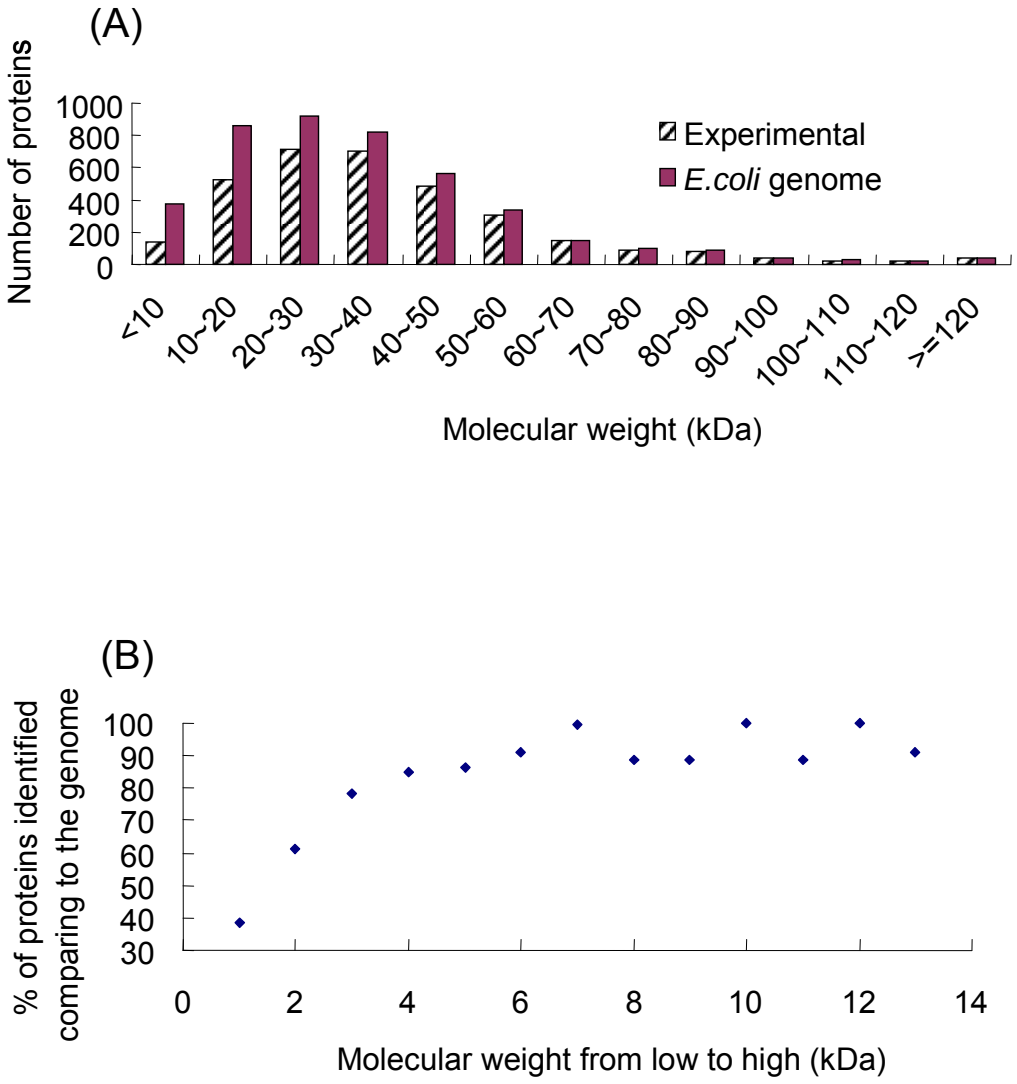


Figure 3.8 (A) molecular weight distribution of identified proteins. (B) the percentage of proteins identified comparing to the genome as a function of molecular weight.

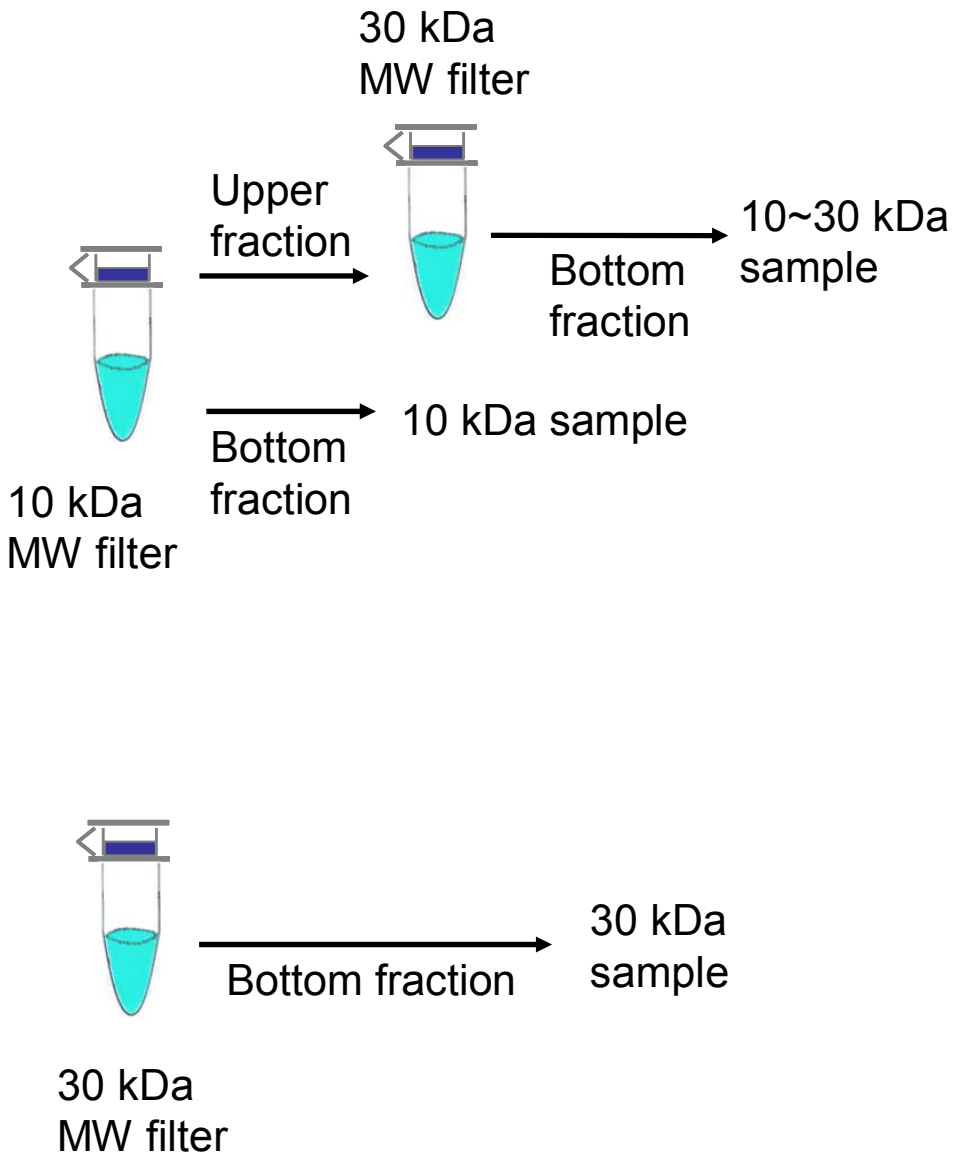


Figure 3.9 Workflow of enrichment of low molecular weight proteins by molecular weight cut off filters.

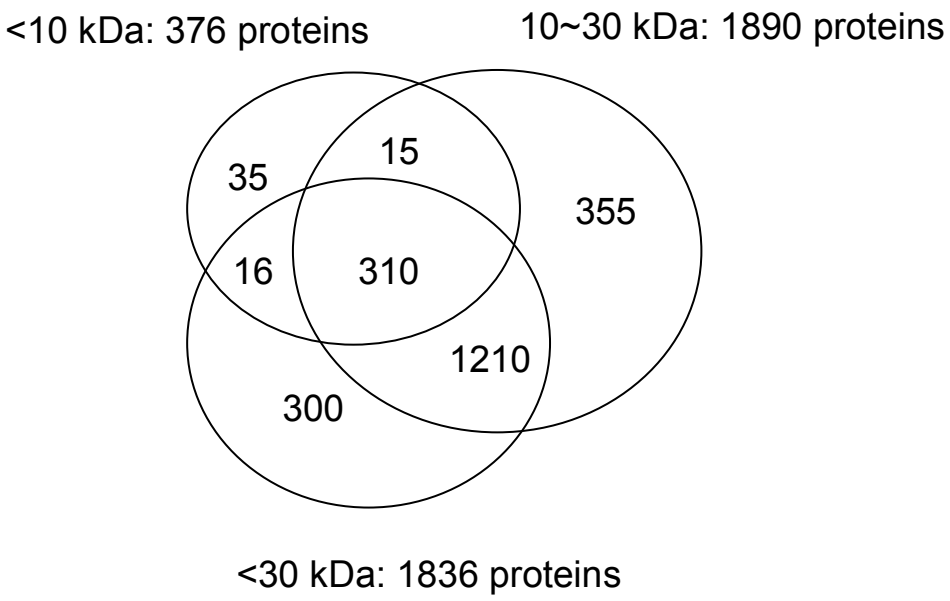


Figure 3.10 Venn diagram of the unique proteins identified from the low molecular weight fractions. 2241 unique proteins were identified.

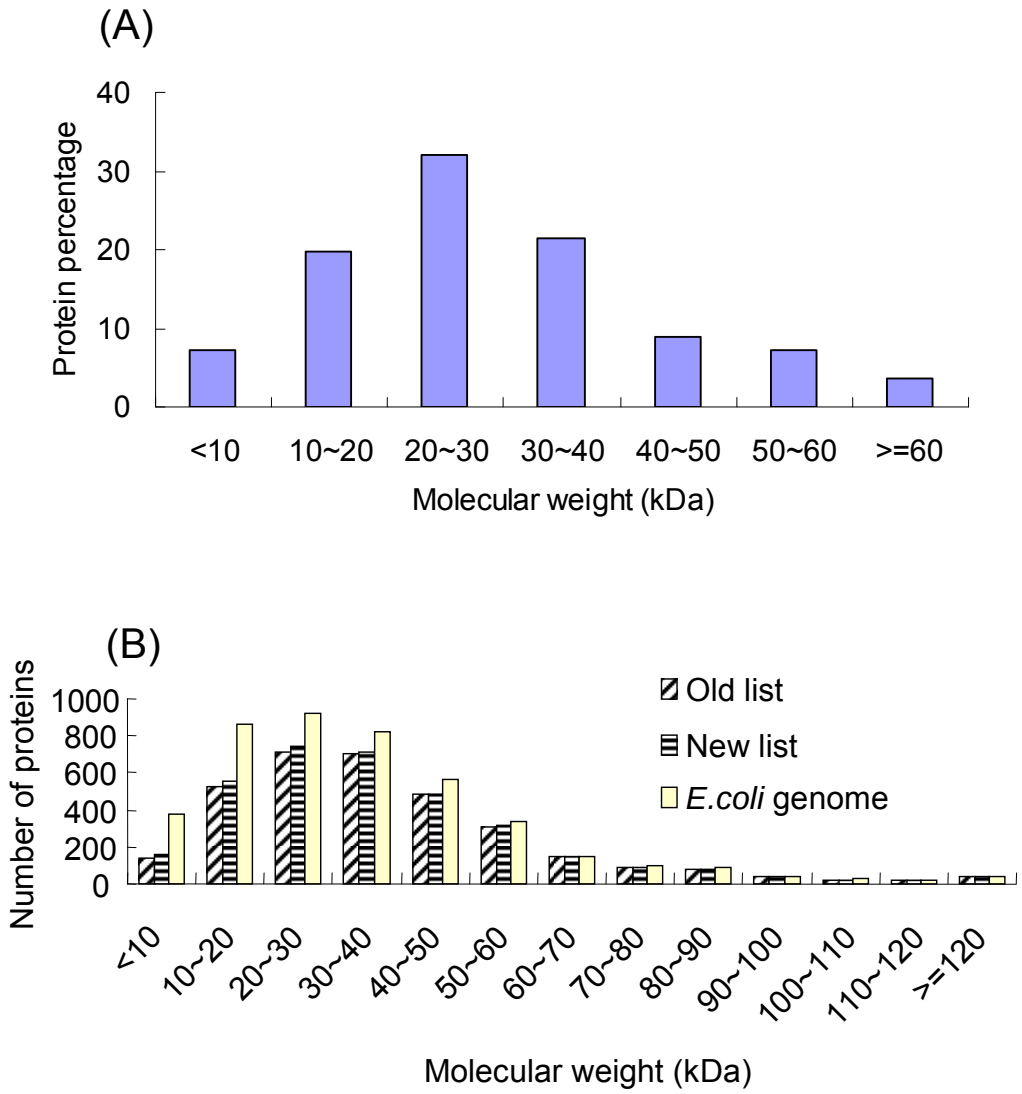


Figure 3.11 (A) Distribution of the percentage of unique proteins identified in the low molecular weight fractions. (B) Molecular weight distribution of all the identified proteins and the genome-predicted proteins.

through the MW cutoff filter. Secondly, the shot proteomic strategy identifies proteins based on sequence match of one or several peptides digested from the protein, which cannot differentiate a protein fragment from the intact protein.

Combining the proteins identified from the low MW protein samples and the 3325 proteins detected from the cytoplasm, peripheral and integral membrane fractions, a total of 3418 unique proteins were identified from the *E. coli* K12 cell line with higher than 99% confidence. This result represents 79% proteome coverage of the 4300 predicted proteins, and is the most comprehensive proteome profile of the *E. coli* K12 cell line. However, from this work, there are still two issues remaining to be addressed. First, Figure 3.11B shows the molecular weight distribution of the 3418 proteins. Compared to the genome-predicted proteins, it is clear that proteins with low MW are still under-represented. We re-examined these 882 missing proteins carefully and found no unique properties that could prevent them from being detected. For example, most of them are not hydrophobic proteins according to the protein GRAVY. In the future, antibody-based techniques such as Western-blot may be applied to investigate whether the 882 missing proteins are present in the *E. coli* K12 cell line cultured in the rich media condition and harvested at the early stationary growth phase. Second, to maximize the proteome coverage, multi-levels of protein and peptide fractionation were performed, which resulted in the consumption of about 10 mg of starting material and more than 300 LC-MS/MS runs. Therefore, in the future, experimental procedures need to be simplified to increase throughput and reduce sample consumption.

3.4 Conclusions

A comprehensive proteome profile of the *E. coli* K12 cell line was generated for the first time. By using a combination of cell lysate pre-fractionation, sequential protein precipitation and solubilization, low molecular weight cutoff filter and 2D-LC MS/MS, a total of 3418 proteins was identified with higher than 99% confidence, representing 79% of the predicted proteome. This is the most comprehensive proteome profile of the *E. coli* cell line. We also found that 882 proteins not detected from this work are mainly of low molecular weight (MW <60 kDa). Whether these missing proteins are present in the cell line under our culture condition still remains to be further investigated. Future work should also focuse on increasing the sample handling throughput and reducing sample consumption by simplifying the experimental procedures.

3.5 References

1. Cornelis, P., Expressing genes in different Escherichia coli compartments. *Curr Opin Biotech* **2000,** 11, (5), 450-454.
2. Blattner, F. R.; Plunkett, G.; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; ColladoVides, J.; Glasner, J. D.; Rode, C. K.; Mayhew, G. F.; Gregor, J.; Davis, N. W.; Kirkpatrick, H. A.; Goeden, M. A.; Rose, D. J.; Mau, B.; Shao, Y., The complete genome sequence of Escherichia coli K-12. *Science* **1997,** 277, (5331), 1453-1458.
3. Molloy, M. P.; Herbert, B. R.; Slade, M. B.; Rabilloud, T.; Nouwens, A. S.; Williams, K. L.; Gooley, A. A., Proteomic analysis of the Escherichia coli outer membrane. *European Journal of Biochemistry* **2000,** 267, (10), 2871-2881.
4. Mee-Jung, H.; Sang Yup, L., The Escherichia coli proteome: past, present, and future prospects. *Microbiology and Molecular Biology Reviews* **2006,** 70, (2), 362-439.
5. Corbin, R. W.; Paliy, O.; Yang, F.; Shabanowitz, J.; Platt, M.; Lyons, C. E.; Root, K.; McAuliffe, J.; Jordan, M. I.; Kustu, S.; Soupene, E.; Hunt, D. F., Toward a protein profile of Escherichia coli: Comparison to its transcription profile. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (16), 9232-9237.
6. Loo, R. R. O.; Cavalcoli, J. D.; VanBogelen, R. A.; Mitchell, C.; Loo, J. A.; Moldover, B.; Andrews, P. C., Virtual 2-D gel electrophoresis: Visualization and analysis of the E-coli proteome by mass spectrometry. *Analytical Chemistry* **2001,** 73, (17), 4063-4070.
7. Vijayendran, C.; Burgerneister, S.; Friehs, K.; Niehaus, K.; Flaschel, E., 2DBase: 2D-PAGE database of Escherichia coli. *Biochemical and Biophysical Research Communications* **2007,** 363, (3), 822-827.
8. Watt, R. M.; Wang, J.; Leong, M.; Kung, H.-f.; Cheah, K. S. E.; Liu, D.; Danchin, A.; Huang, J.-D., Visualizing the proteome of Escherichia coli: an efficient and versatile method for labeling chromosomal coding DNA sequences (CDSs) with fluorescent protein genes. *Nucleic Acids Research* **2007,** 35, (6).
9. Maillet, I.; Berndt, P.; Malo, C.; Rodriguez, S.; Brunisholz, R. A.; Pragai, Z.; Arnold, S.; Langen, H.; Wyss, M., From the genome sequence to the proteome and back: Evaluation of E-coli genome annotation with a 2-D gel-based proteomics approach. *Proteomics* **2007,** 7, (7), 1097-1106.
10. VanBogelen, R. A.; Abshire, K. Z.; Moldover, B.; Olson, E. R.; Neidhardt, F. C., Escherichia coli proteome analysis using the gene-protein database. *Electrophoresis* **1997,** 18, (8), 1243-1251.
11. Holden, P.; Horton, W. A., Crude subcellular fractionation of cultured mammalian cell lines. *BMC research notes* **2009,** 2, 243.
12. Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D., Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2002,** 1, (4), 351-360.
13. Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R.; Li, L., Proteome

profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *Journal of Proteome Research* **2008,** 7, (8), 3583-3590.

14. Wang, N.; MacKenzie, L.; De Souza, A. G.; Zhong, H. Y.; Goss, G.; Li, L., Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *Journal of Proteome Research* **2007,** 6, (1), 263-272.

15. Zhang, N.; Chen, R.; Young, N.; Wishart, D.; Winter, P.; Weiner, J. H.; Li, L., Comparison of SDS- and methanol-assisted protein solubilization and digestion methods for Escherichia coli membrane proteome analysis by 2-D LC-MS/MS. *Proteomics* **2007,** 7, (4), 484-493.

16. Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D., Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *Journal of Proteome Research* **2002,** 1, (4), 351-360.

17. Wang, N.; Xie, C. H.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Anal. Chem.* **2009,** 81, (3), 1049-1060.

18. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem.* **2008,** 80, (12), 4696-4710.

19. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Meth* **2007,** 4, (3), 207-214.

20. Kyte, J.; Doolittle, R. F., A Simple Method for Displaying the Hydropathic Character of A Protein. *J. Mol. Biol.* **1982,** 157, (1), 105-132.

Chapter 4

Microwave-assisted Protein Solubilization for Mass Spectrometry-based Shotgun Proteome Analysis

4.1 Introduction

Protein solubilization is a key step in the shotgun proteome analysis workflow. Various solvents, buffers and surfactants can be used to dissolve proteins and the choice of a solubilization reagent is determined by many factors, including compatibility with the downstream processes. For example, if a liquid chromatography (LC)-based protein separation is performed to fractionate a proteome sample at the protein level, a buffer solution compatible with LC is likely chosen over a surfactant that may interfere with the separation process. However, because of the relatively low solubilizing capability, protein solubilization in a buffer solution may take a long time and, in some cases, such as in analyzing a membrane proteome fraction,[1] not all the proteins can be dissolved. For shotgun proteome analysis, efficient and complete digestion of the solubilized proteins is also important. The method of protein solubilization may affect the structures of the solubilized proteins in the solution used for digestion.[2-6] It is desirable to keep the solubilized proteins in an unfolded form so that their digestion sites are exposed to an enzyme molecule, such as trypsin, for efficient digestion without the occurrence of missed cleavage at many sites.

Aside from the selection of a proper reagent to dissolve the proteins, some physical means, such as vortex mixing, is often used to accelerate the solubilization process. However, vortex-assisted protein solubilization (VAPS) is still a slow process, particularly when a reagent with low solubilizing capability is used. In this chapter, we describe a simple and efficient technique to dissolve proteins for shotgun proteome analysis. It is based on the use of a domestic microwave oven to irradiate the protein sample mixed with a reagent to assist protein solubilization. Microwave irradiation has recently been applied to accelerate chemical or enzymatic reactions for proteomics applications.[7-16] In our laboratory, during the course of the development of microwave-assisted acid hydrolysis (MAAH) as a means of degrading proteins for protein sequencing or identification,[10] Zhong *et al.* observed that hydrophobic proteins merely suspended in 25% trifluoroacetic acid (TFA) could be dissolved, under microwave irradiation, and degraded into peptides. If the same protein suspension was subjected to conventional heating, protein pellets could still be seen in the solution after prolonged heat exposure.[10] Thus, microwave irradiation must play a role in accelerating the protein solubilization and degradation processes,

although the mechanism is not entirely clear.

The objective of this study was to develop a microwave-assisted protein solubilization (MAPS) method to increase the sample handling speed and protein identification efficiency for shotgun proteome analysis. Compared to vortex-assisted protein solubilization, MAPS could reduce the solubilization time significantly and increase the number of peptides and proteins identified. This method was combined with sequential protein solubilization using various reagents to pre-fractionate a complex proteome sample. The performance of this method is illustrated in the proteome analysis of the *E. coli* K-12 integral membrane protein extract.

4.2 Experimental

4.2.1 Chemicals and reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), sodium carbonate (Na_2CO_3) and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC/MS-grade water, acetone, formic acid, methanol (MeOH), and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada). A domestic 900 W (2450 MHz) microwave oven (Sunbeam or Panasonic purchased from a local store) was used to perform microwave-assisted protein solubilization experiments.

4.2.2 Cell culture and membrane preparation

Escherichia coli K-12 (*E. coli,* ATCC 47076) was from the American Type Culture Collection. A single *E. coli* K12 colony was used to inoculate 50 mL of LB broth (BBL, Becton Dickinson). The culture was incubated overnight with shaking at 35 ˚C. This saturated culture (12.5 mL) was added to 500 mL of growth medium in a 2-L baffled Erlenmeyer flask. Cells were harvested in the early stationary phase by centrifugation at 11,300 *g* for 15 min at 4 ˚C, resuspended, washed in 150 mL PBS buffer, and collected by centrifugation at 9,820 *g* for 15 min at 4 ˚C.

The *E. coli* cells were suspended in 30 mL water and three tablets of Roche mini protease

inhibitor cocktail were added. The suspension was passed twice through a French press (Aminco Rochester, NY) using rapid fill kit at 20,000 psi. Unbroken cells and debris were removed by centrifugation at 7,740 *g* for 15 min at 4 ˚C. The supernatant was collected as the cell lysate. The membrane proteins were isolated using a carbonate fractionation procedure[17] with some modifications. The lysate was centrifuged in a Beckman Type 55.2Ti rotor for 55 min at 118,000 *g*, and the pellet was resuspended, washed in 20 mL of 50 mM NH_4HCO_3, and collected by centrifugation in a Beckman Type 55.2Ti rotor for 40 min at 116,811 *g*. The pellet containing membrane proteins was suspended again in 11 mL of 50 mM ammonium NH_4HCO_3, and the protein concentration was determined by BCA assay.

The membranes were diluted with ice-cold 100 mL of 100 mM Na_2CO_3 (pH 11.0). The solution was stirred slowly in an ice bath for 1 h to extract membranes. The extract was divided equally into two tubes, and centrifuged in a Beckman Type 45Ti rotor for 60 min at 115,000 *g*. Each pellet was gently rinsed with 5 mL of water, suspended in 7 mL of 50 mM MOPS buffer (pH 7.3) and transferred to an 8 mL tube. The tubes were centrifuged in a Beckman Type 70.1Ti rotor at 115,000 *g* for 25 min. The pellet was collected as integral membrane proteins and stored at -80 ˚C.

4.2.3 Sequential microwave-assisted protein solubilization (MAPS) and vortex-assisted protein solubilization (VAPS).

E. coli integral membrane proteins were solubilized in 0.5% SDS and equally split. Acetone, precooled to -80 ˚C, was added gradually (with intermittent vortexing) to the protein extract to a final concentration of 80% (v/v). The mixture was kept at -20 ˚C overnight and centrifuged at 20,800 *g* for 10 min at 4 ˚C. The supernatant was decanted and properly disposed. Acetone was evaporated at room temperature. About 650 µg of the pellet was subjected to three levels of sequential microwave-assisted protein solubilization to reduce the protein complexity of the membrane fraction. In the NH_4HCO_3-assisted method, 25 µL of NH_4HCO_3 (25 mM, pH 8.0) was added to the pellet in a 1.5-mL polypropylene vial with a flat top cap (Rose Scientific, Edmonton, Alberta). A sealed sample vial was placed on the rotating tray inside a microwave beside a Rubbermaid 118-mL (1/2 cup) plastic container containing 100 mL of water. Microwave irradiation was applied for a total of 3 min in 30 s cycles with intermittent homogenization applied between cycles by vortex. For each cycle, the initial water temperature was 23.2 ˚C and raised to 68.2 ˚C after 30 s of microwave. The temperature of the sample vial was unknown; but it should be similar to the water temperature. After MAPS, the solution was centrifuged at 20,800 *g* for 12 min at 4 °C. The same microwave irradiation procedure

with homogenization and centrifugation steps was applied in the subsequent solubilization processes using different reagents. The remaining pellet was resuspended in 8 M urea and then subjected to microwave irradiation for 3 min in 30 s cycles. After removing the solubilized proteins, the remaining pellet was suspended in 2% SDS and microwaved for 3 min also in 30 s cycles, then centrifuged at 20,800 *g* for 12 min at 4 °C.

The vortex-assisted protein sequential solubilization was performed in the same manner. The three reagents, 25 mM of NH_4HCO_3, 8 M urea, and 2% SDS, were sequentially employed to dissolve the membrane proteins with the assistance of 2-2.5 h vortexing; the optimized time used is discussed in the Results and Discussion section.

4.2.4 In-solution digestion

The protein fractions generated by the above solubilization steps were individually reduced with 28 μL of 900 mM DTT for 1 h at 37 °C. Free thiol groups were blocked by reaction with a double volume of 900 mM iodoacetamide for 1 h at room temperature in the dark. Before tryptic digestion, the urea and SDS fractions were diluted 8- and 20-fold, respectively. Trypsin solution was added to the supernatant at an enzyme/protein ratio of 1:40, and the digestion was conducted at 37 °C overnight. The digest solutions were stored at -80 °C until further analysis.

4.2.5 Cation exchange chromatography

A 2.1×250 mm highly hydrophilic polysulfoethyl A column (5 μm diameter, 300 Å pore, PolyLC Inc. U.S.) was used for the strong-cation exchange separation of the tryptic peptides. Gradient elution was performed with mobile phases A (10 mM KH_2PO_4 with 20% ACN, pH 2.7) and B (10 mM KH_2PO_4, 500 mM KCl with 20% ACN, pH 2.7). The gradient was from 0 to 4% B for 1 min, 4 to 20% B for 16 min, 20 to 60% B for 22 min, ramped to 100% B in 6 min, and held for another 10 min. One min or 2 min fractions were collected based on the sample loading amount, and then pooled into 3 to 29 fractions according to the UV absorbance values in the chromatograms (see below).

4.2.6 Peptide desalting and quantification by RPLC

Desalting and quantification were carried out using an Agilent 1100 HPLC system (Palo Alto, CA) with 4.6 mm × 50 mm Polaris C18 A column with 3 μm particles and 300 Å pore

size (Varian, CA). The eluted peptides were monitored and quantified using a UV detector operated at 214 nm using a method previously described.[18] After loading of the peptide sample, the column was flushed with mobile phase A (0.1% TFA in water) and the salts were effectively removed. Subsequently, the concentration of mobile phase B (0.1% TFA in ACN) in the mobile phase was step-wise increased to 85% to ensure complete elution of the peptide fractions from the column. During the peptide elution process a chromatographic peak was produced and, based on the peak area, the amount of peptides was determined. BSA digests of various amounts were used as standards for the generation of a linear calibration between the peak area and the injected peptide amount. The calibration curve was generated as $y = 417x + 2139$, where y refers to the peak area of the peptide sample in µg, 2139 refers to the peak area of the blank wash, and x refers to the peptide amount analyzed. The linear range of the calibration curve was from 0.5 µg to 5.0 µg of peptides (R^2=0.998).

4.2.7 LC-ESI QTOF MS and MS/MS analysis

This step was performed as previously described with minor changes.[18, 19] 1 µg portions of peptides of each desalted SCX fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA) equipped with a nanoACQUITY Ultra Performance LC system (Waters). Peptide solution from each SCX fraction was injected onto a 75 µm × 100 mm Atlantis dC18 column with 3 µm particles and 100 Å pore size (Waters). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Mass spectra were acquired from m/z 300-1600 for 0.8 s, followed by 4 data-dependent MS/MS scans from m/z 50-1900 for 0.8 s each. The collision energy used to perform MS/MS was automatically varied according to the mass and charge state of the eluting peptide. A mass calibrant (i.e., lock-mass) was infused at a rate of 350 nL/min, and an MS scan was acquired for 1 s every 1 min throughout the run. Peptide precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides identified from the adjacent two SCX fractions to enable additional and less abundant peptides to be analyzed and identified.[19] An exclusion list was generated based on Mascot (Matrix Science, London, U.K.) searching results of peptides with a score 10 points above the identity threshold.

4.2.8 Protein database search

Database searches were performed as previously described with minor changes.[18, 19] Raw search data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.5 (Waters). Peptide sequences were identified by automated

database searching of peak list files using the Mascot search program. Database searching was in the *E. coli* K12 database (4337 sequences). The following search parameters were selected for all database searching: enzyme, trypsin; missed cleavages, 1; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+, and 3+; fixed modification, carbamidomethyl (C); variable modifications, N-acytyl (Protein), carbamyl (N-term), oxidation (M), Pyro-Glu (N-term Q), Pyro-Glu (N-term E). The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, and Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses, were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at a confidence level of 99% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but different access ID numbers were also removed from the list. To gauge the false positive peptide matching rate, target-decoy search strategy was applied by searching the MS/MS spectra against the forward and reverse *E. coli* K12 proteome sequences.[20]

4.2.9 Hydropathy calculation and annotation of localization

All peptides and proteins identified were examined using the ProtParam program, available at the EXPASY Web site (http://us.expasy.org/tools/protparam.html), which allows calculation of the grand average of hydropathy (GRAVY).[21] The Gene Ontology database (http://geneontology.org/) was used to classify proteins on the basis of cellular location. Transmembrane domains (TMD) of identified proteins were predicted by TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) according to protein primary sequences.

4.3 Results and discussion

4.3.1 Standard test

Membrane proteins are generally more difficult to dissolve in a solvent, than other types of proteins. Bacteriorhodopsin (BR) is a transmembrane protein that has been used as a standard for MS and proteomics method development.[2, 10, 22-24] We used this standard to gauge the solubilization of membrane proteins under different conditions. In our experiment, 500 mg of BR was solubilized in 0.25% SDS and split equally into seven 1.5 mL aliquots. Each aliquot was precipitated by cold acetone (-80 °C), mimicking a common practice where proteins are precipitated from a cell or tissue extract (see below).

The protein pellet was subjected to solubilization in 25 mM NH_4HCO_3, 60% methanol or 8 M urea with the assistance of either 2 h of vortexing or 3 min of microwave irradiation in 6 cycles of 30 s. The last aliquot was completely dissolved in 2% SDS as a control. Figure 4.1A shows the SDS-PAGE results from the solutions of the seven protein-pellet samples dissolved by 25 mM NH_4HCO_3 with vortex (lane 1) or microwave (lane 2), 60% methanol with vortex (lane 3) or microwave (lane 4), and 8 M urea with vortex (lane 5) or microwave (lane 6). No protein band is visible for lanes 1 to 4, indicating that the dissolving power of NH_4HCO_3 or MeOH was not sufficiently high to solubilize any significant amount of the acetone-precipitated protein pellet. In 8 M urea, a weak band is observed from the vortex sample (lane 5) and a more intense band from the microwave sample (lane 6). However, lane 6 is less intense than lane 0 where the pellet was dissolved in 2% SDS. These results indicate that BR can be partially dissolved in 8 M urea and that microwave irradiation improves protein solubilization to a large extent, compared to vortexing. To optimize the MAPS method and examine its performance when dissolving a wide range of proteins, compared to the commonly used vortex method, we used the *E. coli* membrane proteome fraction as the test sample in the subsequent experiments.

4.3.2 Solubilization time and solubility

The time dependence of protein solubilization for both the vortex and microwave methods was studied. In this experiment, 25 μL of 25 mM NH_4HCO_3 or 8 M urea was added to a vial containing 650 μg of the *E. coli* integral membrane protein pellet. After vortex or microwave for a specific time period, the solution was centrifuged for 10 min, and the protein concentration of the supernatant was determined by the BCA assay. Duplicate runs were performed for individual samples. Figure 4.2A shows the concentrations of dissolved proteins in 25 mM NH_4HCO_3 as a function of vortex time. It appears that solubilization reaches a plateau when the vortex time approaches 150 min. Figure 4.2B shows the results using 8 M urea with vortex for protein solubilization. In this case, the concentration levels off when the vortex time is about 120 min. Panels C and D in Figure 4.3 present the concentrations of dissolved proteins as a function of the microwave time in 25 mM NH_4HCO_3 and 8 M urea, respectively. The concentration reaches a plateau when the microwave time is 3 min for NH_4HCO_3 and 2.5 min for urea. It should be noted that, for the MAPS experiment, we applied the microwave irradiation in a discontinuous mode with a 30 s cycle, i.e., after microwave irradiation for 30 s, the sample vial was taken out and placed on ice to cool the sample to about room temperature (~40 s) followed by intermittent homogenization using vortex (~2 min). Then the vial was placed inside the microwave oven for another round of 30 s irradiation. In this way, overheating the sample was avoided. The total experiment time for a 3 min microwave solubilization process was about 20 min.

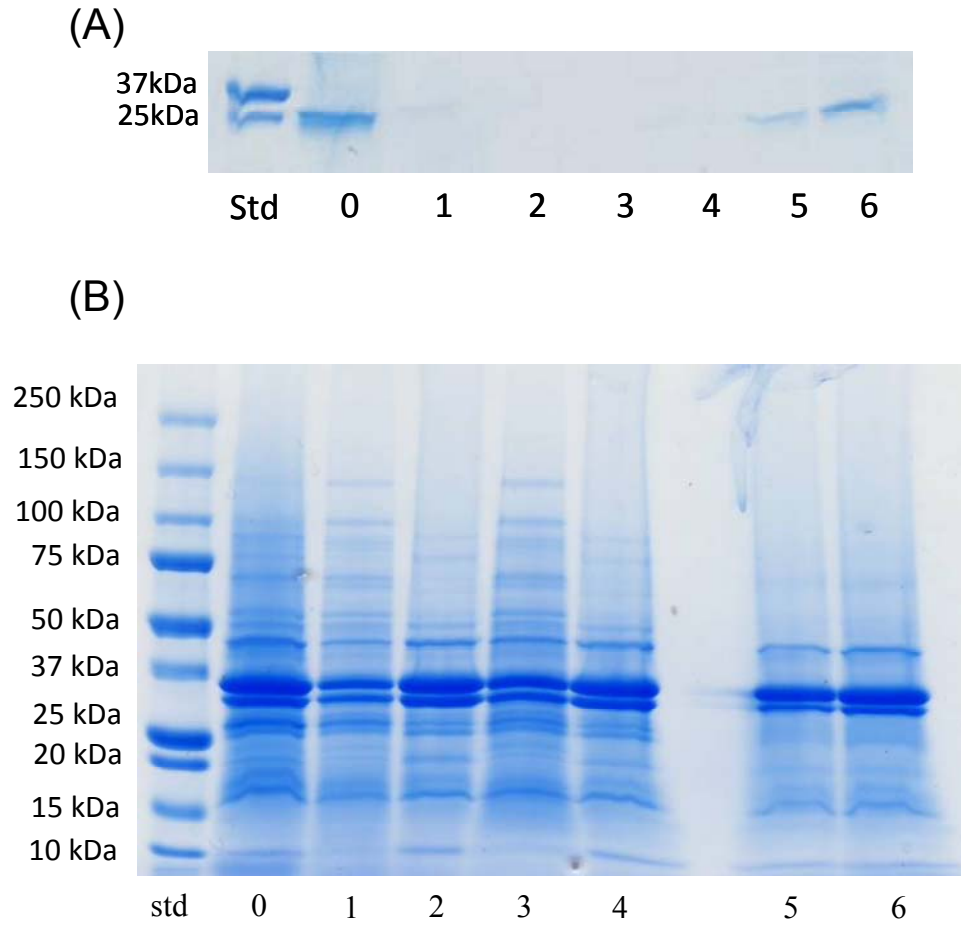


Figure 4.1 (A) SDS-PAGE images of acetone-precipitated bacteriorhodopsin (BR) dissolved in 2% SDS (lane 0), 25 mM NH_4HCO_3 with vortex (lane 1) or microwave (lane 2), 60% methanol with vortex (lane 3) or microwave (lane 4), and 8 M urea with vortex (lane 5) or microwave (lane 6). (B) SDS-PAGE images of the *E. coli* integral membrane protein extract solubilized in 2% SDS (lane 0), 25 mM NH_4HCO_3 with microwave (lane 1), and 25 mM NH_4HCO_3 with vortex (lane 3). The remaining protein pellets from both cases are dissolved in 8 M urea with microwave (lane 2) and vortex (lane 4), respectively. After solubilization by 8 M urea with microwave, the remaining proteins were solubilized completely in 2% SDS (lane 5). After solubilization by 8 M urea with vortex, the remaining proteins were also completely solubilized in 2% SDS (lane 6).

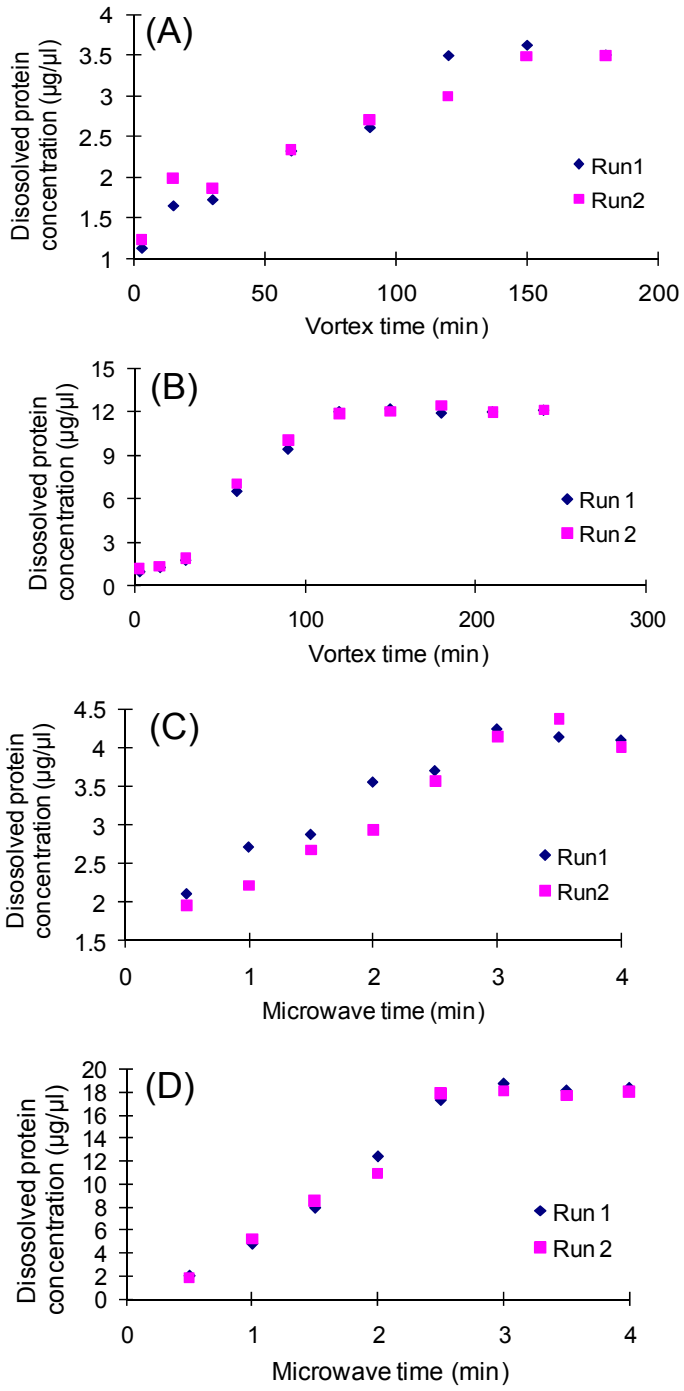


Figure 4.2 The concentration of dissolved proteins from 650 µg of the *E. coli* integral The results shown in Figure 4.2 clearly demonstrate that MAPS is a much more efficient membrane protein pellet in (A) 25 mM NH_4HCO_3 and (B) 8 M urea as a function of vortex time, and (C) 25 mM NH_4HCO_3 and (D) 8 M urea as a function of the total microwave irradiation time in 30 s cycles.

technique than the conventional vortex method. The optimal microwave time is dependent on the type of solvent used, as in the case of the vortex method. Urea has a stronger dissolving power than NH_4HCO_3 and hence both the optimal microwave time and vortex time are shorter when urea is used. A much greater amount of protein was solubilized in urea than in NH_4HCO_3, and in both solvents the MAPS technique dissolved larger amounts of protein than the vortex method.

A more rigorous comparison of the solubility differences between the MAPS and vortex methods is shown in Figure 4.3. In this set of experiments, equal amounts of the *E. coli* integral membrane proteins (650 μg per aliquot) were solubilized in 8 M urea with the assistance of vortex or microwave irradiation under the optimal solubilization time, as described above. In brief, 25 μL of 8 M urea was added to the protein pellet, followed by either vortex for 2 h or microwave irradiation for 3 min. The supernatant was taken out and the remaining pellet was re-dissolved by adding 25 μL of fresh 8 M urea using the same procedure as the first round. This process was repeated for 7 rounds. The concentration of solubilized proteins from each round was determined by the BCA assay. Any residual pellet after 7 rounds of solubilization was dissolved in 2% SDS with 3 min microwave irradiation at a 30 s cycle, and the concentration of SDS-solubilized proteins was also determined by the BCA assay.

Figure 4.3A plots the amount of dissolved proteins from different rounds of solubilization in 8 M urea with vortex, while Figure 4.3B shows the results from MAPS. The data shown from the replicate runs (n=2 for vortex and n=3 for microwave) indicate that both methods are reproducible. In both cases, a large fraction of proteins was dissolved in the first round. The total amount of protein dissolved in urea with vortex after 7 rounds of solubilization was 551 ± 5 μg. The amount of the remaining proteins dissolved in the SDS solution was 102 ± 5 μg. The sum of these two is equal to the amount of the starting material, i.e., 650 μg, indicating there was no sample loss during the solubilization process. In the case of MAPS, 627 ± 8 μg of proteins were dissolved after 7 rounds and the remaining proteins dissolved in SDS were found to be 29 ± 10 μg. These results indicate that about 85% of proteins can be dissolved in 8 M urea by 7 rounds of vortexing, compared to about 96% using MAPS. Moreover, the 7 rounds of the vortex experiments took more than 14 h, while 7 rounds MAPS experiments took about 140 min (7 × 20 min). This clearly indicates that MAPS is much more efficient in dissolving the membrane protein fraction in 8 M urea.

The comparison of the protein amount dissolved in each round with the assistance of vortex or microwave irradiation is shown in Figure 4.3C. A much greater amount of protein was dissolved in 8 M urea using microwave irradiation than vortex (450 ± 20 μg vs. 300 ± 5 μg) from the first round of solubilization. In the second round, a similar amount was dissolved. Only small amounts of proteins were dissolved in the 3rd round and beyond with MAPS,

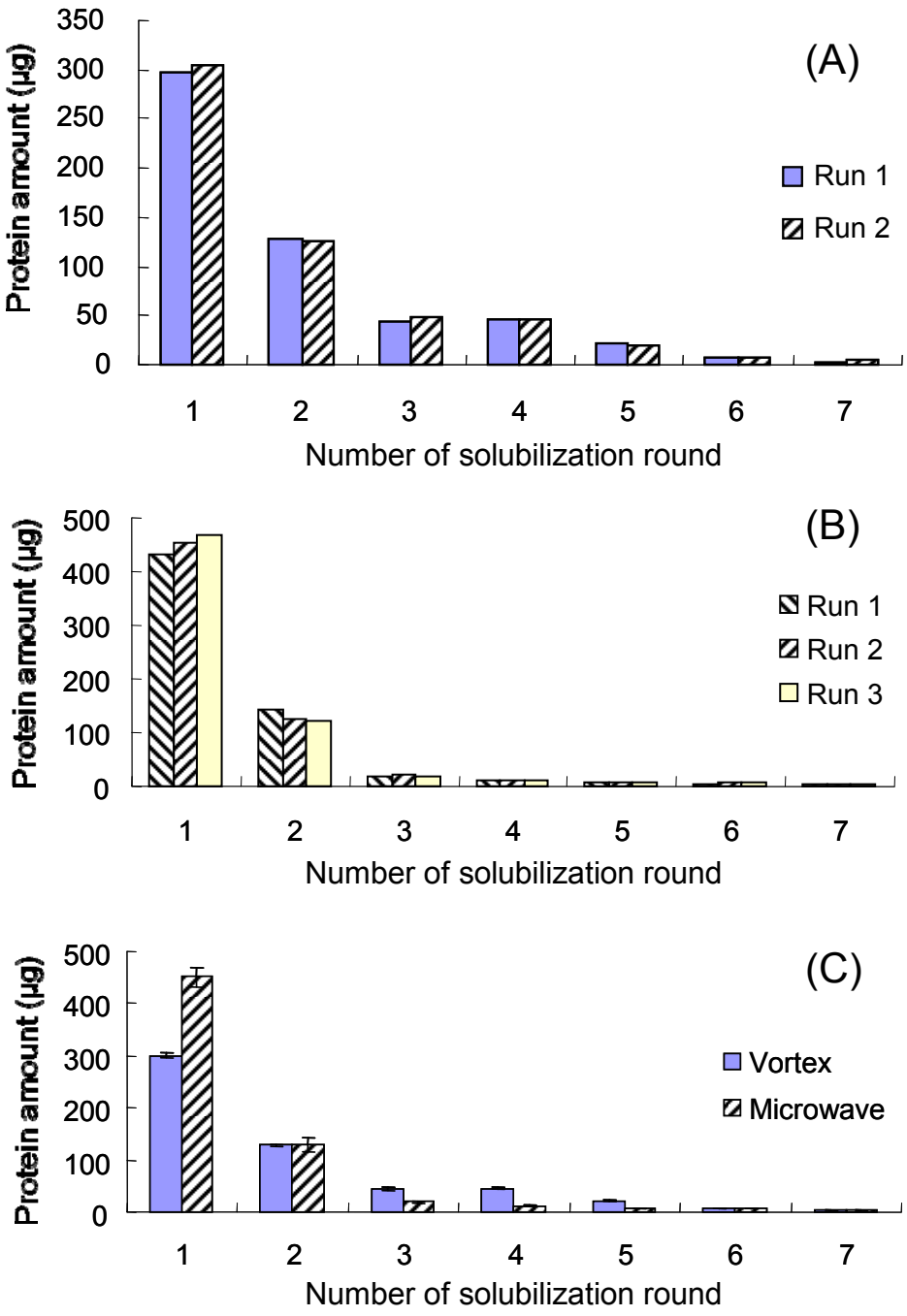


Figure 4.3 The amount of dissolved proteins from different rounds of solubilization in 8 M urea with (A) vortex and (B) MAPS. (C) comparison of the vortex and MAPS results.

while more protein continued to be dissolved in subsequent rounds of the vortex procedure.

The compositions of the dissolved proteins in urea were examined by SDS-PAGE. Figure 4.1B shows the gel images obtained from the *E. coli* samples with the same amount of sample loading for each run. Lane 0 is from the control, i.e., the integral membrane protein extract completely solubilized in 2% SDS. Lanes 1 and 3 are from the proteins dissolved in 25 mM NH_4HCO_3 by using the MAPS method and the vortex method, respectively. These two gel images are similar, suggesting that there is no significant difference in the composition of the proteins. The remaining protein pellets from both cases were further dissolved in 8 M urea by the MAPS method (lane 2) and the vortex method (lane 4), respectively. There is no significant difference between these two images. In both cases, some protein pellets still remained in the vials and they were further dissolved in 2% SDS. Lanes 5 and 6 are from the remaining pellets dissolved in 2% SDS. These two images are also similar. Overall, the SDS-PAGE data do not indicate any major difference in the protein compositions from the protein pellets dissolved in 25 mM NH_4HCO_3 or 8 M urea by the two methods. Therefore, the SDS-PAGE results did not reveal any significant protein degradation during the microwave irradiation process.

Similar results were obtained from the solubility comparison of vortex and MAPS for the proteins dissolved in NH_4HCO_3 (Figure 4.4). The microwave method dissolved a much larger amount of protein after 7 rounds than did the vortex method (331 μg vs. 177 μg dissolved from ~650 μg of starting material). Compared to the urea sample, a higher proportion of the proteins was dissolved after 3 rounds of solubilization (see Figure 4.4). This is mainly due to the fact that NH_4HCO_3 has a lower solubilization power than urea and, thus, proteins tend to stay in both solution and solid forms longer. With urea, the equilibrium shifts to the solution form more rapidly, resulting in a more complete dissolution after 3 rounds. For the SDS samples, one round of microwave or vortex was found to be sufficient to solubilize all the protein pellets.

4.3.3 BSA tryptic digestion

Another important step in the proteome sample preparation is related to trypsin digestion of the solubilized sample. Digestion will be more complete if the protein molecules are full denatured to allow trypsin to access all the digestion sites in the protein sequence. To compare the performance of trypsin digestion of proteins prepared using different solubilization methods, we examined the number of tryptic peptides identified and the sequence coverage of these peptides for a standard protein, BSA. In addition to vortexing, there were other techniques also reported to assist in protein solubilization, e.g. sonication[25, 26] and heating[27].

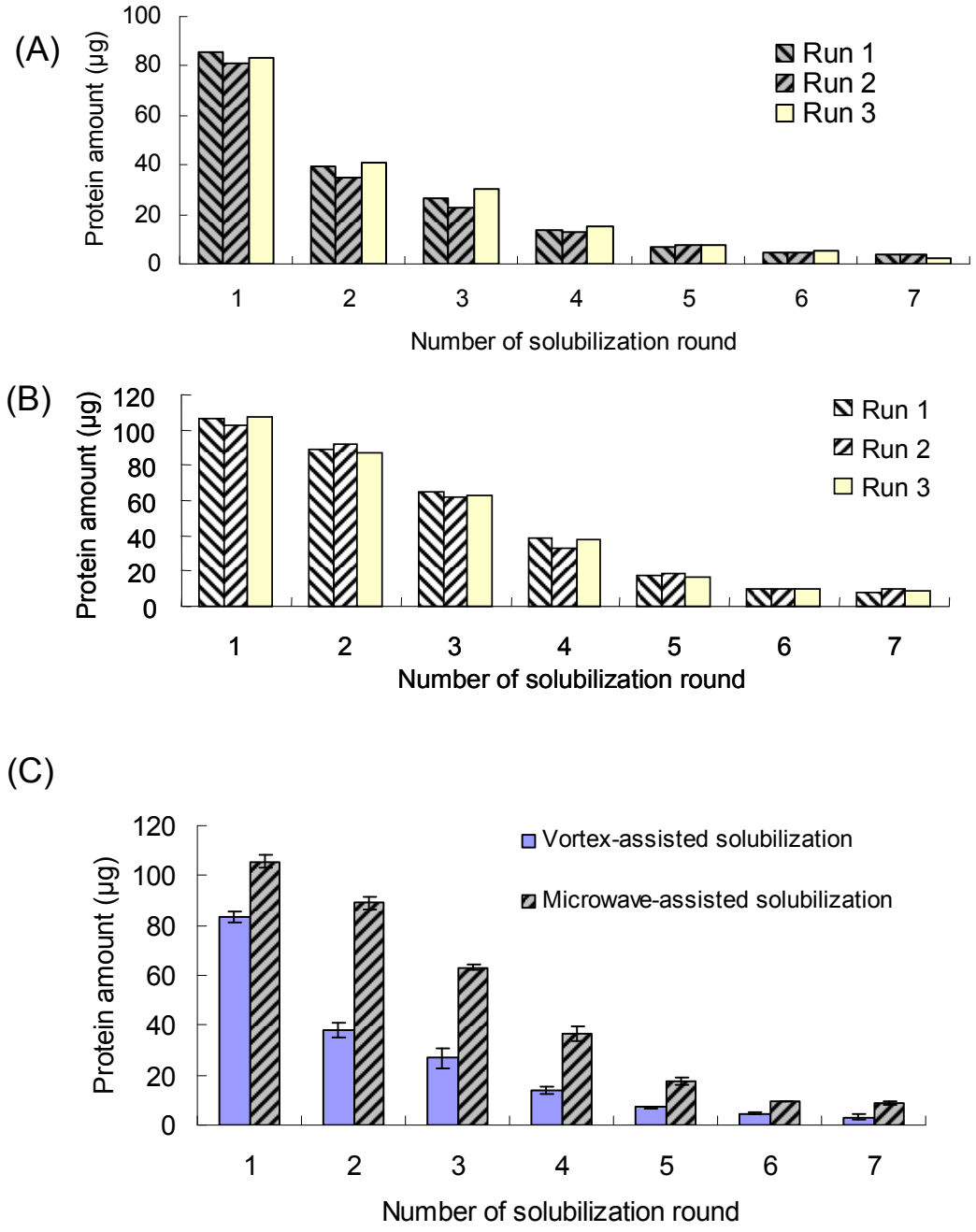


Figure 4.4 The amount of dissolved proteins from different rounds of solubilization in 25 mM NH_4HCO_3 with (A) vortex and (B) MAPS (C) comparison of the vortex and MAPS results.

In our comparison experiments, eight aliquots of BSA (1 mg each) were individually solubilized in 1 mL of 25 mM NH_4HCO_3 or 8 M urea with the assistance of 2 h vortexing, 10 min heating at 65 ° C, 30 min sonication or 6 × 30 s microwave with intermittent cooling on ice, respectively. All the experiments were done in replicates (16 samples in total). All the samples were subjected to overnight tryptic digestion according to the same procedure described previously in section 4.2.4, except that the enzyme/protein ratio was 1:200. After removing salts from the digested samples by a reversed-phase column, a 0.2 μg portion of the BSA digest from each sample was analyzed by LC MS/MS.

Table 4.1 shows the number of peptides identified from each sample and their sequence coverage. Redundant peptides caused by modifications (i.e., oxidation and carbamylation) were removed. As Table 4.1A shows, in the NH_4HCO_3 solubilized samples, vortex-, sonication- and heating-assisted methods identified an average of 43, 46 and 44 unique peptides, respectively. The sequence coverages were all under 74%. The MAPS method identified 52 unique peptides in both replicate experiments and the sequence coverages were higher than 81%. As Table 4.1B shows, in the urea-solubilized samples, the average numbers of identified peptides from replicate experiments using the vortex-, sonication-, and heating-assisted methods were 41, 46 and 49, respectively. The sequence coverages were all lower than 78%. The MAPS method identified an average of 70 unique peptides in the two replicate experiments and both sequence coverages were higher than 85%. These results clearly demonstrate that, compared to the traditional solubilization methods, MAPS increased the BSA sequence coverage and the number of identified peptides, indicating that MAPS can increase the overall tryptic digestion efficiency. This can be attributed to the more efficient denaturing of the protein molecules by the microwave irradiation during the solubilization process, compared to other methods. Thus, the MAPS method not only facilitates protein solubilization, but also denatures the proteins to an extent leading to improved tryptic digestion.

4.3.4 Sequential solubilization

For analyzing a complex proteome sample, protein-level fractionation can be useful to generate a more comprehensive proteome coverage. Sequential protein solubilization, after proteins from cell lysates, tissue extracts, or other types of samples have been precipitated and washed to remove potentially interfering materials, is a simple method of protein-level fractionation.[28-31] We have compared the performance of vortex and MAPS for dissolving proteins in shotgun proteome analysis with sequential protein solubilization. In this experiment, equal amounts of the *E. coli* integral membrane proteins (i.e., 650 μg) were sequentially solubilized by NH_4HCO_3, urea, and then SDS,

Table 4.1 BSA test results: number of identified peptides and sequence coverage by different solubilization methods. (A) BSA solubilized in 25 mM NH_4HCO_3. (B) BSA solubilized in 8 M urea.

(A)

25 mM NH_4HCO_3 solubilized BSA	Identified peptide number	sequence coverage
Vortex for 2 h_replicate 1	43	71
Vortex for 2 h_replicate 2	43	70
Sonication for 30 min_replicate 1	45	73
Sonication for 30 min_replicate 2	46	74
Heating 65 °C for 10 min_replicate 1	43	69
Heating 65 °C for 10 min_replicate 2	44	70
Microwave for 6 × 30 s_replicate 1	52	81
Microwave for 6 × 30 s_replicate 2	52	83

(B)

8 M urea solubilized SA	Identified peptide number	sequence coverage
Vortex for 2 h_replicate 1	45	72
Vortex 2 h_replicate 2	36	68
Sonication for 30 min_replicate 1	45	78
Sonication for 30 min_replicate 2	46	78
Heating 65 °C for 10 min_replicate 1	49	78
Heating 65 °C for 10 min_replicate 2	49	77
Microwave 6 × 30 s_replicate 1	66	85
Microwave 6 × 30 s_replicate 2	76	91

either with the assistance of vortexing or microwave irradiation. About 179.6 μg (7 rounds), 360.5 μg (6 round) and 108.7 μg (1 round) of proteins were dissolved in the three solvents with vortex, respectively, with the number of rounds used as indicated. In the case of MAPS, 325.1 μg (7 rounds), 289.0 μg (5 rounds) and 40.3 μg (1 round) were dissolved in the three solvents, respectively. The dissolved proteins were digested by trypsin and the peptides were analyzed by 2D-LC MS/MS. Biological replicate experiments were performed for both the vortex and MAPS methods.

Figures 4.5, 4.6 and 4.7 present Venn diagrams indicating the results of the numbers of proteins identified from the replicate experiments of the two methods. As Figures 4.5A and 4.6A show, in the two replicate experiments, the vortex method identified 585 and 577 proteins from the NH_4HCO_3 fraction, 889 and 827 proteins from the urea fraction and 288 and 269 proteins from the SDS fraction, respectively. In total, 954 proteins were identified from the first replicate and 919 proteins from the second replicate. There were many common proteins identified from the NH_4HCO_3 and urea fractions, indicating that after NH_4HCO_3 solubilization these common proteins were still present in the pellet that was subsequently dissolved by urea. Much less overlap was observed between the NH_4HCO_3 fraction and the SDS fraction. This is understandable considering that the same protein can be present in both solution and solid forms in a solvent and the equilibrium shifts to the solution form if the protein is very soluble. A protein may exist in different conformers (e.g., native and denatured forms) which can have different solubilities. Thus, many proteins that were only partially dissolved in the NH_4HCO_3 fraction remained in the pellet and were soluble in urea, resulting in the identification of these proteins in both fractions. After urea solubilization, only a small number of unique proteins are detectable in the SDS fraction. Figures 4.5A and 4.6A show that in the replicate experiments, 63 proteins (6.6%) and 84 proteins (9.1%) were identified from the NH_4HCO_3 fraction only, 319 unique proteins (33.4%) and 287 proteins (31,2%) were found in the urea fraction and only 2 proteins (0.2%) and 7 proteins (0.8%) were uniquely detected in the SDS fraction, respectively.

Figures 4.5B and 4.6B show that in the two replicate experiment, there were 893 and 857 proteins identified in NH_4HCO_3, 968 and 912 proteins in urea and 477 and 493 proteins in SDS and the number of proteins uniquely detected was 166 (14.4%) and 152 (13.6%), 203 (17.8%) and 160 (14.3%), and 15 (1.3%) and 41 (3.7%), respectively. These results indicate that, in the MAPS method, the number of unique proteins identified in NH_4HCO_3 and urea was more evenly distributed, suggesting that more proteins were completely dissolved in NH_4HCO_3 with the assistance of microwave irradiation than by the vortex method. These fully dissolved proteins in NH_4HCO_3 were only detected in the NH_4HCO_3 fraction. This is consistent with the observation discussed above that a larger amount of protein was soluble in NH_4HCO_3 using MAPS (i.e., 325.1 μg in the microwave method vs. 179.6 μg in the vortex method). Apparently, the microwave process has assisted in

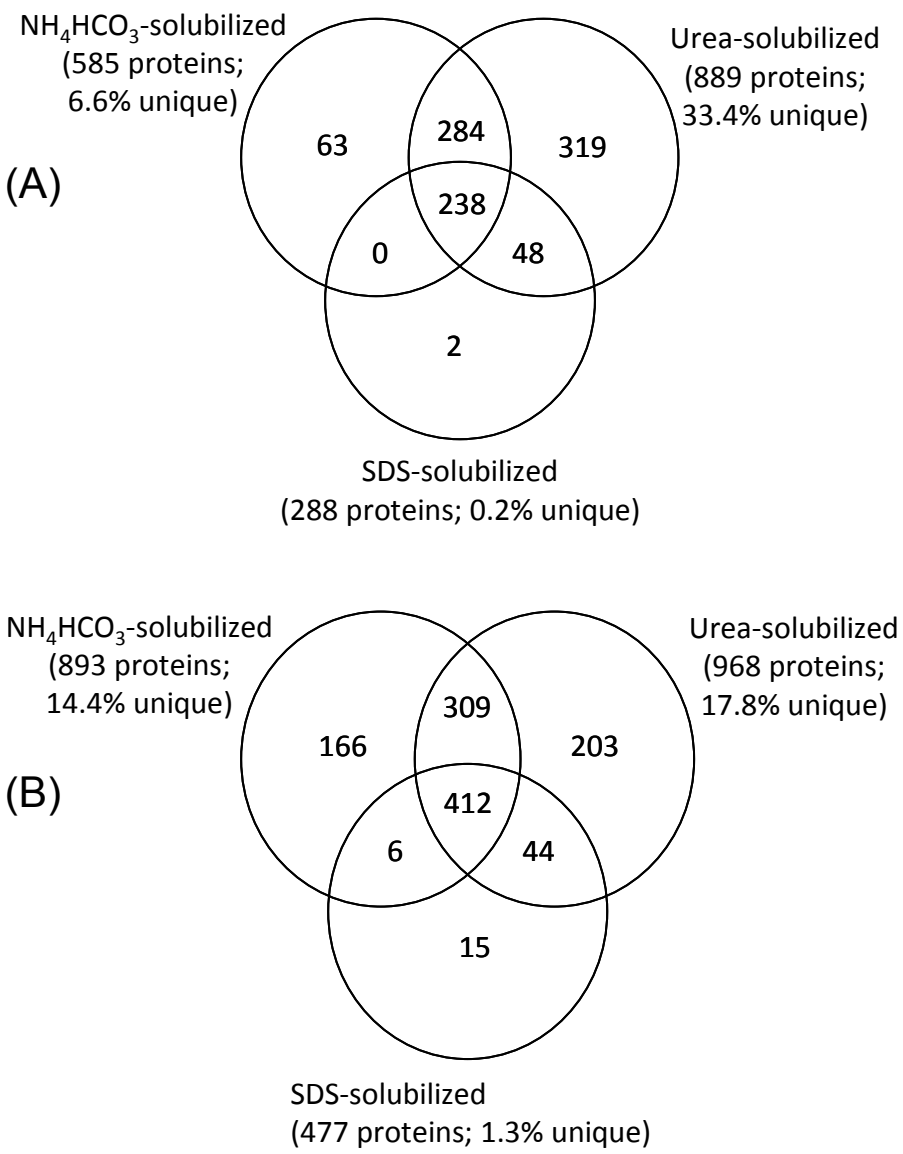


Figure 4.5 Venn diagram of the proteins identified from the two methods (replicate one): (A) vortex and (B) MAPS.

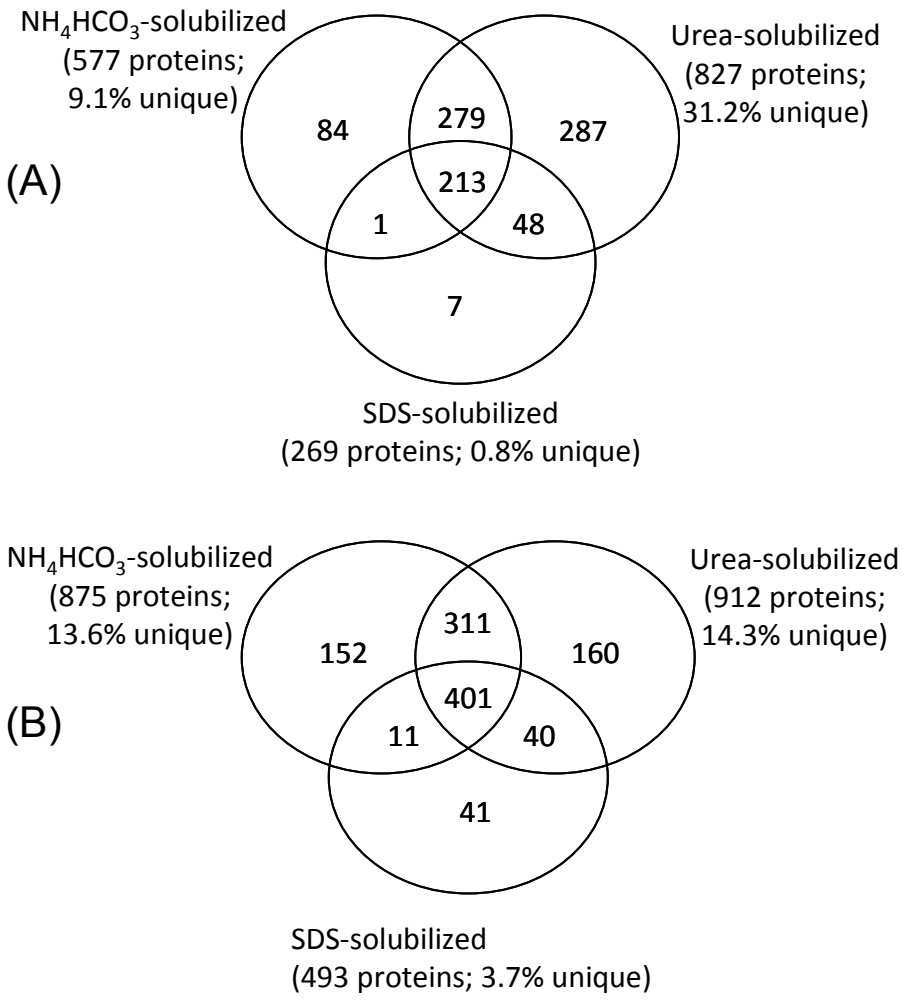


Figure 4.6 Venn diagram of the proteins identified from the two methods (replicate two): (A) vortex and (B) MAPS.

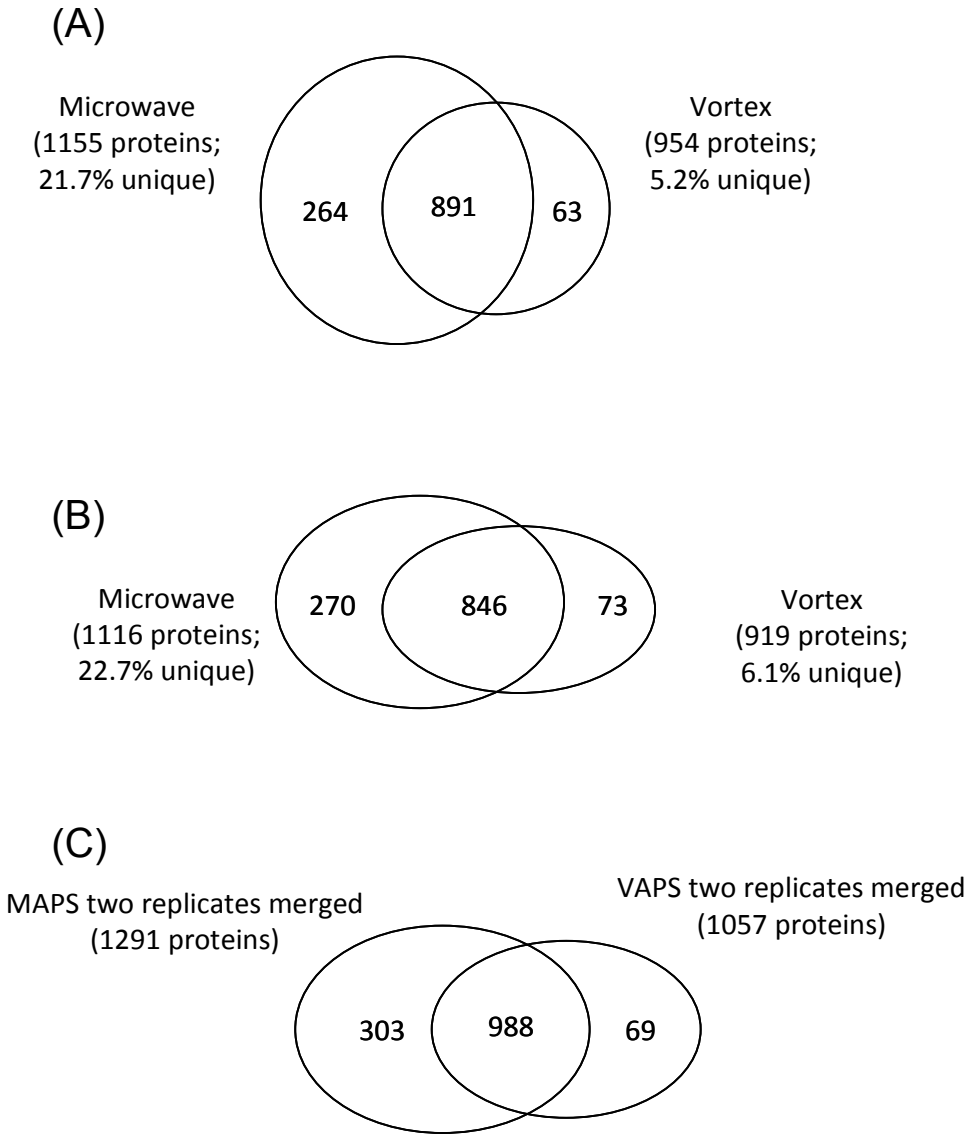


Figure 4.7 Venn diagram of the identified proteins by the MAPS and vortex methods: (A) replicate one. (B) replicate two. (C) merged data from the two replicates.

shifting the equilibrium to the solution form when a protein is partially soluble in a solvent. However, the data shown in Figures 4.5B and 4.6B also indicate that many proteins were still identified in both the NH_4HCO_3 and urea fractions; these proteins were not completely dissolved in NH_4HCO_3. This is the limitation of the sequential solubilization method, i.e., it is not completely orthogonal for protein fractionation. Nevertheless, the MAPS method improves the performance of sequential solubilization by dissolving proteins more completely in a solvent, and thereby resulting in better protein separation between the two solvent fractions. In MAPS, the number of proteins identified was more evenly distributed in the three fractions (884, 940 and 485) than with the vortex method (581, 858 and 279) (average numbers of identified proteins are presented here). In total, 1155 and 1116 proteins were identified from the three fractions by MAPS in the replicate experiments.

Figure 4.7C shows the comparison of the proteins identified using the vortex and microwave methods from combined results of two replicate experiments. The two methods identified a total of 1360 proteins. Combining the identified proteins in the two replicate experiments, the MAPS identified 1291 proteins and the number of proteins detected by the vortex method was 1057. The MAPS method identified most of the proteins found by the vortex method; only 69 out of 1360 proteins (5.1%) were detected by the vortex method only. There were 234 more proteins identified by MAPS in the replicate experiments, representing about a 22% increase over the vortex method.

Figures 4.7 A and B show the number distribution of the proteins identified using the vortex and microwave methods from the two replicate experiments, respectively. The same distributions were observed in both the replicate experiments. Most identified proteins were found by the MAPS method (95% and 94%); only 63 out of 1218 proteins (5.2%) and 73 out of 1189 proteins (6.1%) were detected by the vortex method only.

The comparisons of the two methods at the peptide level are shown in Figures 4.8, 4.9 and 4.10. In the first replicate, from a total of 50 LC-MS/MS runs (21, 26 and 3 runs from the NH_4HCO_3, urea and SDS fractions, respectively), 4677 unique peptides from 954 proteins were identified from the three fractions prepared with the vortex method. In the second replicate, from a total of 58 LC-MS/MS runs (22, 29 and 7 runs from the NH_4HCO_3, urea and SDS fractions, respectively), 5022 unique peptides from 919 proteins were identified. Figures 4.8A and 4.9A show the number distribution of the peptides identified in the three fractions from the replicate experiments. There were 1774 and 1838 peptides identified from the NH_4HCO_3 fraction, 4136 and 4311 peptides from the urea fraction and 560 and 573 peptides from the SDS fraction, respectively. Using the microwave method, 7973 unique peptides (1155 proteins) were identified from a total of 58 LC-MS/MS runs (22, 29 and 7 runs from the NH_4HCO_3, urea and SDS fractions, respectively) in the first replicate experiment, 8010 unique peptides (1116 proteins) were identified from a total of 58

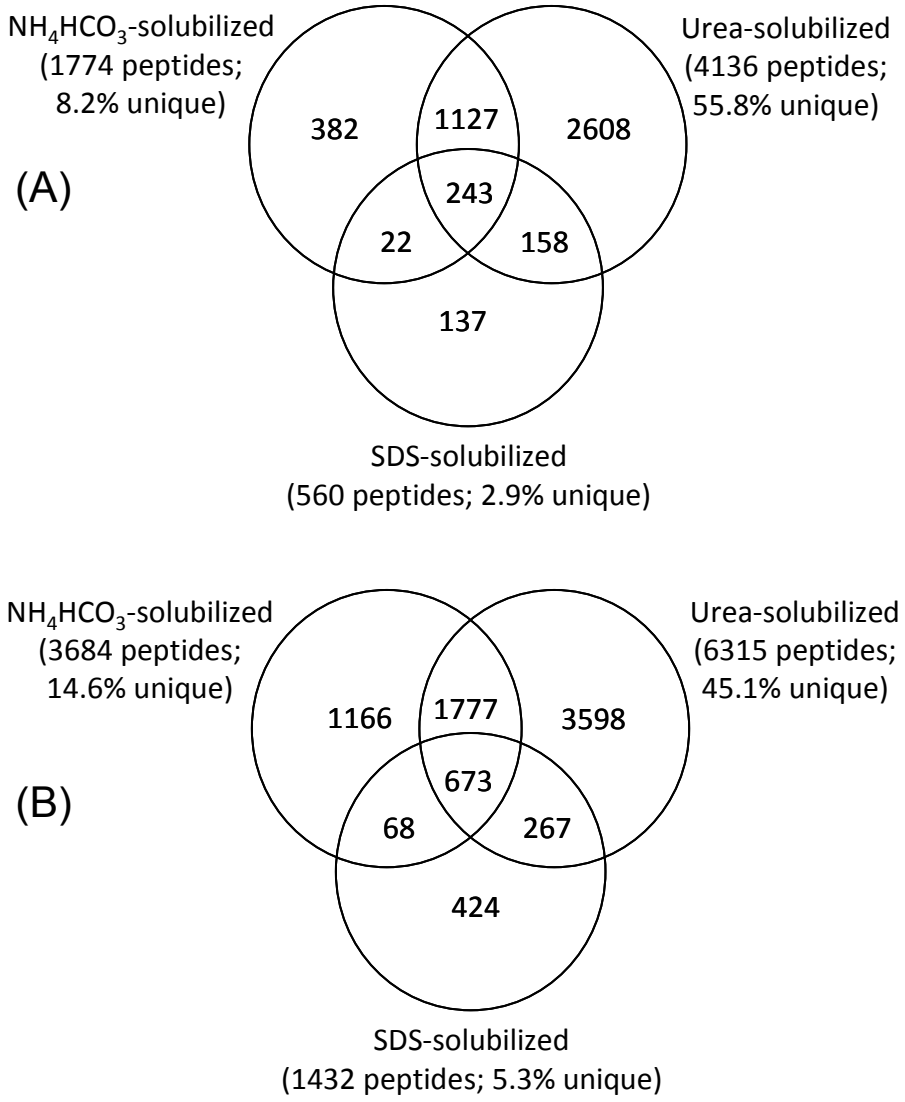
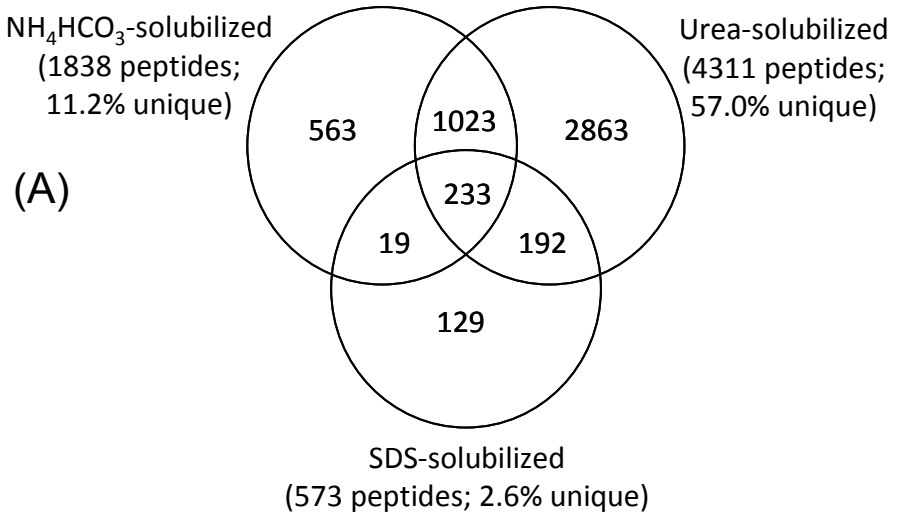


Figure 4.8 Venn diagram of the peptides identified from the two methods (replicate one): (A) vortex and (B) MAPS.



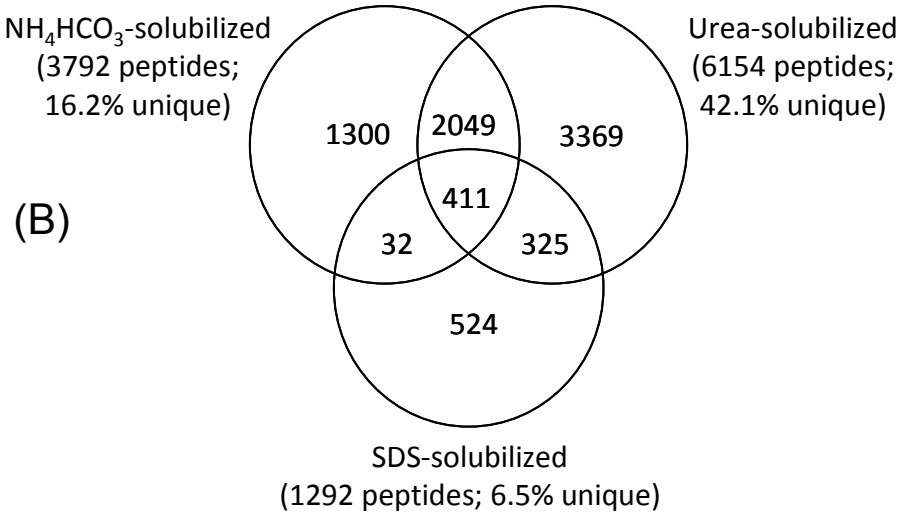


Figure 4.9 Venn diagram of the peptides identified from the two methods (replicate two): (A) vortex and (B) MAPS.

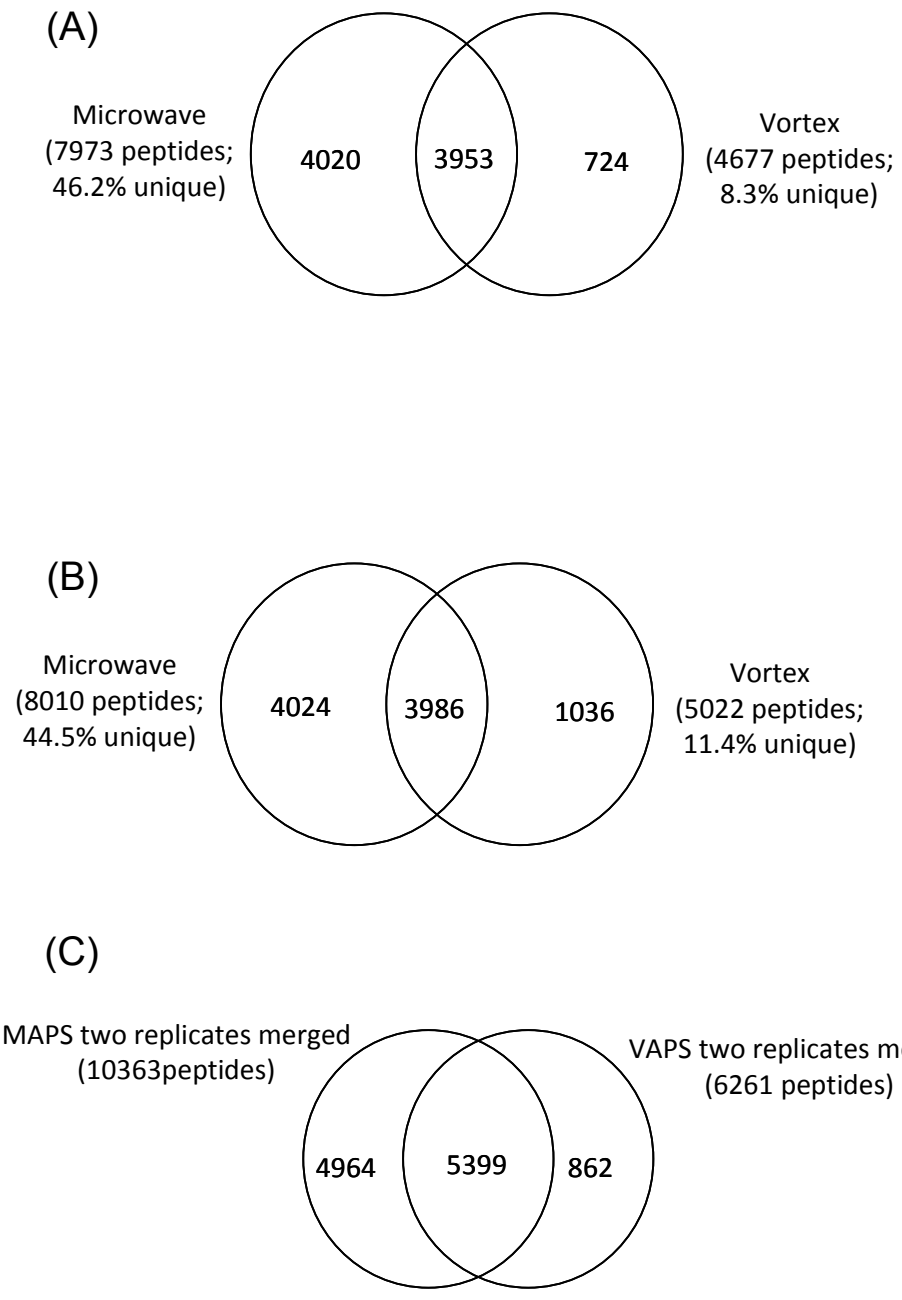


Figure 4.10 Venn diagram of the identified peptides by the MAPS and vortex methods: (A) replicate one. (B) replicate two. (C) merged data from the two replicates.

LC-MS/MS runs (22, 29 and 7) runs from the NH_4HCO_3, urea and SDS fractions, respectively) in the second replicate experiment. In the first replicate, the slight difference in the number of runs was due to the pooling of SCX fractions; the number of SCX fractions collected was the same in both methods. However, in the second replicate, both vortex and MAPS samples were analyzed by the same number of 58 LC-MS/MS runs. Compared to the first replicate, only 7.3% increase in the number of identified peptides was found in the vortex method. This result indicates that the slight difference in the number of LC-MS/MS runs in the first replicate was not the major reason causing the significant difference in the number of identified peptides between the vortex and MAPS methods. Figures 4.8B and 4.9B show the distribution of the peptide numbers found in the three fractions from the two replicate experiments. There were 3684 and 3792 peptides identified from the NH_4HCO_3 fraction, 6315 and 6154 peptides from the urea fraction and 1432 and 1292 peptides from the SDS fraction, respectively. These results are consistent with the protein distributions shown in Figures 4.5 and 4.6. The number of peptides identified in the three fractions is more evenly distributed in MAPS (3738, 6235 and 1362), compared to the vortex method (1806, 4224 and 567) (average data from the two replicates were taken). Figure 4.10C shows the number distribution of the peptides identified by the two methods from the combined results of two replicate experiments. The total number of peptides identified by the two methods combined is 11225. About 92% (10363) of peptides can be identified by the MAPS method, compared to 56% (6261) of peptides detected by the vortex method. Figures 4.10A and 4.10B show the number distribution of the peptides identified using the vortex and microwave methods from the two replicate experiments, respectively. The same distributions were observed in the two replicate experiments as the combined data. Most identified peptides were found by the MAPS method (92% and 89%); only 724 out of 8697 peptides (8.3%) and 1036 out of 9046 peptides (11.4%) were detected by the vortex method only.

Table 4.2 shows the total peptide numbers identified from the two replicates of the microwave and vortex methods, the redundant peptides introduced by modifications, and the peptide numbers after redundant peptides were removed. After removing redundant peptides caused by modifications, 7758 and 7613 peptides were identified the MAPS method, and 4634 and 4955 peptides by the vortex method. The MAPS method still shows a significantly higher number of peptides than the vortex method.

Figures 4.11 and 4.12 show the comparison of the identified peptides from each pooled SCX fraction by the MAPS and vortex methods in two replicate experiments. Compared to the vortex method, all the fractions solubilized by different solutions (i.e., NH_4HCO_3, urea and SDS) in the MAPS method show significantly higher numbers of identified peptides. This result further indicates that the significant increase in the number of peptides identified by the microwave method in the first replicate experiment was not due to the use of extra eight LC-MS/MS runs. It should be noted that, in the first replicate, because the

SCX fractions in the vortex and MAPS samples were not pooled in the same manner, only part of the SCX fractions were taken into comparison. They were 19 SCX fractions in the NH_4HCO_3-solubilized fraction, 22 SCX fractions in the urea-solubilized fraction and 2 SCX fractions in the SDS-solubilized fraction. However, the number of peptides from these fractions account for 100% in the NH_4HCO_3-solubilized fraction, 95% in the urea-solubilized fraction and 83% in the SDS-solubilized fraction of the vortex sample. Therefore, although not all of the SCX fractions were taken into comparison, the results still represented the whole dataset well.

Table 4.2 Modification effects on the identification number of unique peptides.

Replicates	Solubilization method	Total identified peptides	Redundant peptides due to modifications	Total identified peptides after eliminating modification redundant
Replicate 1	MAPS	7973	215	7758
	VAPS	4677	43	4634
Replicate 2	MAPS	8011	398	7613
	VAPS	5022	67	4955

The significant increase in the number of peptides identified by the microwave method is likely related to two major factors. One is the protein concentration difference. In the microwave method, the proteins are more completely dissolved in one solvent, reducing the possibility of dividing the protein amount into two solvent fractions. Higher protein concentration in a solvent fraction would generate higher peptide concentration, which facilitates LC-MS/MS detection. The second factor is the digestion efficiency difference. In the BSA digestion work described earlier, the MAPS method was shown to generate higher BSA sequence coverage than the other three traditional solibilization methods in both the NH_4HCO_3 and urea-solubilized experiments. The current results are consistent with the BSA data. With microwave irradiation, it is possible that the dissolved proteins are unfolded more extensively than those dissolved by vortex. The unfolded proteins are more readily digested by trypsin, resulting in higher concentrations of tryptic peptides, instead of large sizes of peptides produced from the missed cleavage of a number of sites within a protein sequence. Note that, to avoid denaturing trypsin, the urea sample was diluted by 8-fold from 8 M to 1 M and the SDS sample was diluted by 20-fold from 2% to 0.1%. The sample dilution might result in re-folding of some proteins, preventing trypsin from accessing some cleavage sites during the digestion process. Under microwave irradiation, some proteins may be irreversibly denatured during the solubilization process. As a result, the re-folding problem may be less severe in MAPS than that of the vortex method.

It should be noted that the false positive peptide matching rate was estimated to be 0.3% for

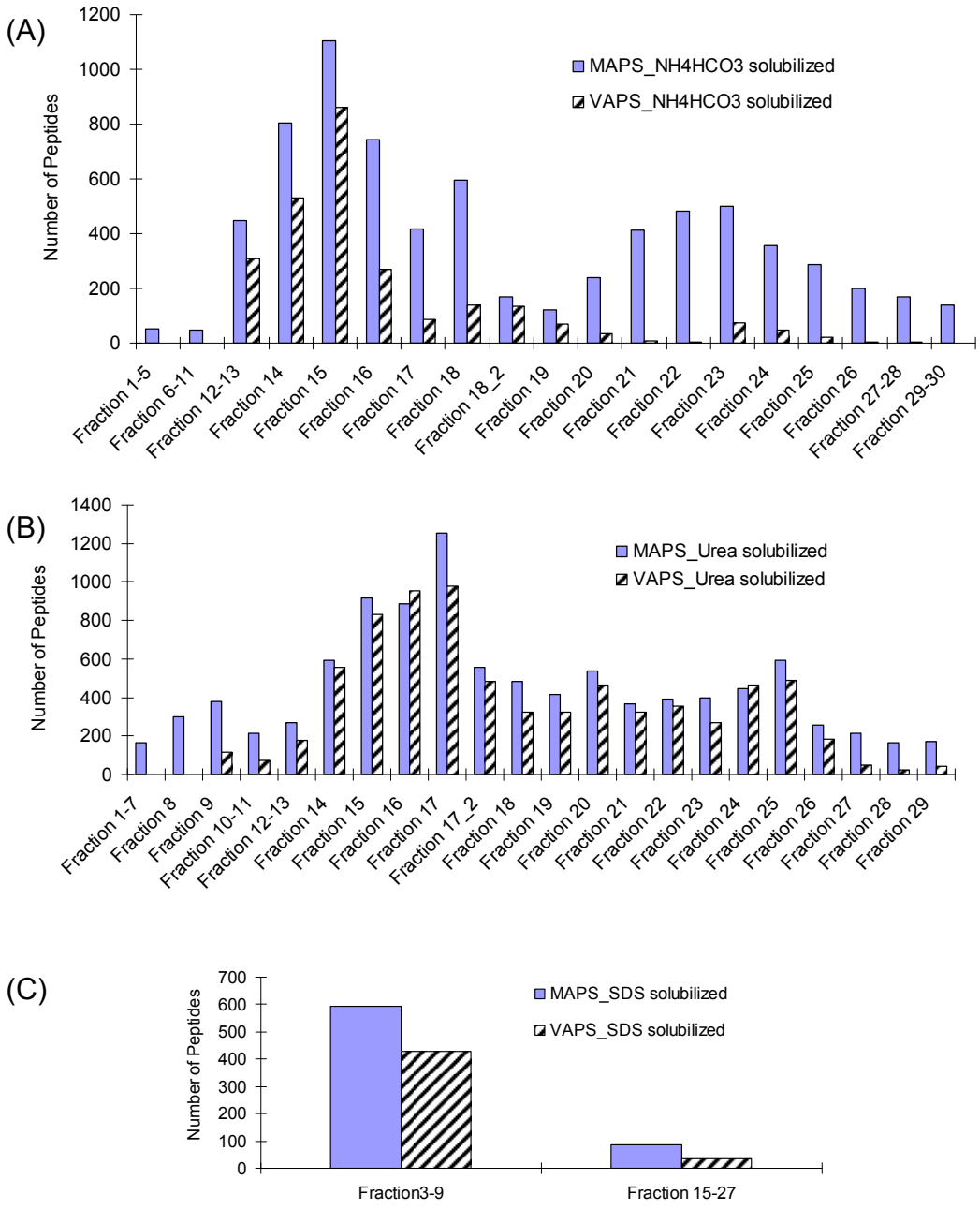


Figure 4.11 Comparison of identified peptides from MAPS and VAPS (replicate one) (A) NH_4HCO_3-solubilized fraction. (B) urea-solubilized fraction. (C) SDS-solubilized fraction.

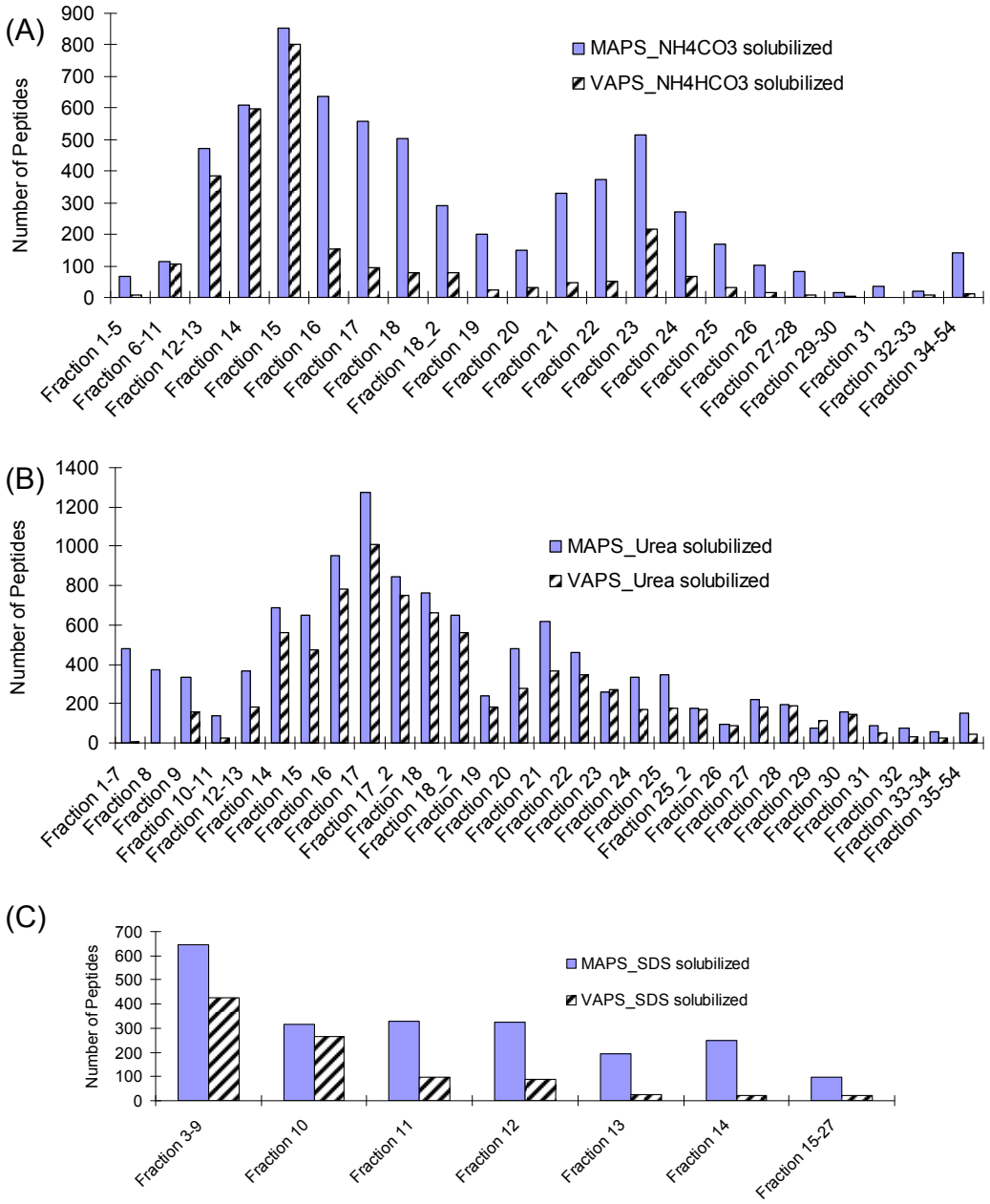


Figure 4.12 Comparison of identified peptides from MAPS and VAPS (replicate two) (A) NH_4HCO_3-solubilized fraction. (B) urea-solubilized fraction. (C) SDS-solubilized fraction.

the vortex method and 0.4% for the microwave method. The rate of carbamylation of peptides was found to be 7.5% for the vortex method and 15.1% for the microwave method. The increase in carbamylation rate was not unexpected as urea was used and the microwave method heated the sample above room temperature. Heating can result in an increase in carbamylation of proteins in urea.[32, 33] The exact temperature of the protein sample during the microwave irradiation was unknown. However, it should not be too high as the sample solution did not boil. The 100 mL of water placed beside the sample vial absorbed most of the radiation during the 30 s microwave irradiation and the final water temperature was 68.2 ˚C. We expect that the protein sample was well below 68.2 ˚C during the most part of the 30 s microwave irradiation.

In summary, the microwave method identified 22% more proteins and 66% more peptides than the vortex method. On average, the vortex method resulted in 5.8 unique peptide identifications per one uniquely identified protein, while the microwave method resulted in 8.0 unique peptide identifications per one uniquely identified protein. These results clearly demonstrate that the microwave method can increase the proteome coverage as well as the confidence of protein identification.

Finally, to provide better understanding of the differences in peptide and protein identification results, we have examined the possible correlation between protein or peptide property and detectability of either method. Figure 4.13A shows the molecular weight distribution of the proteins identified by the two methods. The number of identified proteins were grouped into several bins as a function of the molecular masses, and then calculated into percentages. The two distributions shown in Figure 13A appear to be quite similar.

Figure 4.13B shows the protein distribution as a function of protein hydrophobicity, gauged by the GRAVY indices grouped into four bins. Compared to the GRAVY distribution of all the *E. coli* K12 proteins found in the proteome database, the two experimental distributions are biased to the positive GRAVY indices, which is consistent with the fact that the hydrophobic membrane proteins were enriched in the analyzed samples. However, there is no apparent difference between the vortex and microwave methods in terms of the protein GRAVY distribution. Figure 4.14A shows the distribution of identified peptides as a function of the hydrophobicity (GRAVY). 39.3% of peptides detected by the microwave method have positive GRAVY indices, compared to 38.4% for the vortex method. In addition, for the microwave method, the proportion of mildly hydrophilic peptides (GRAVY ranging from 0 to -0.5) is higher than that of the vortex method (36.8% for microwave vs. 34.2% for vortex). Thus, the microwave method appears to detect more hydrophobic peptides than the vortex method.

We have also examined the distribution of the number of membrane proteins identified by

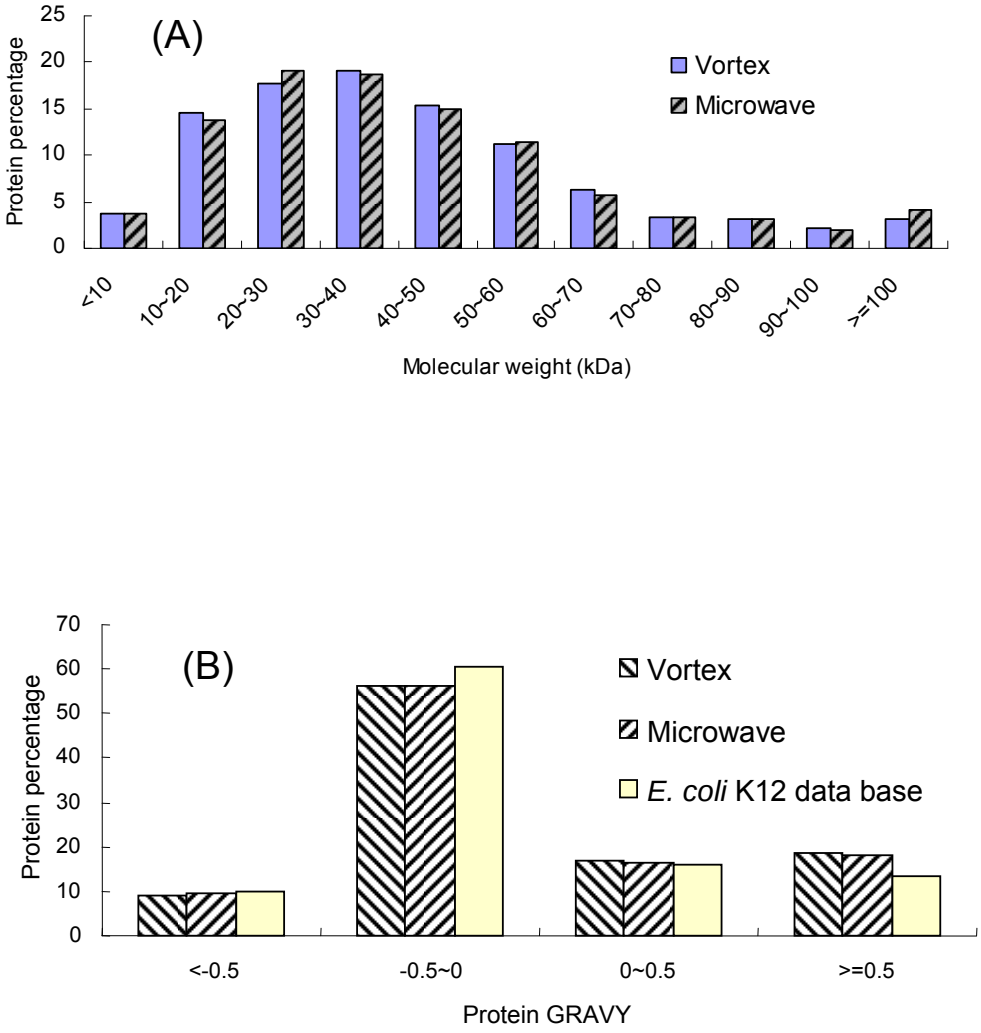


Figure 4.13 (A) Molecular weight distribution of the proteins identified by the vortex and MAPS methods. (B) Protein distribution as a function of protein hydrophobicity, gauged by the GRAVY indices grouped into four bins.

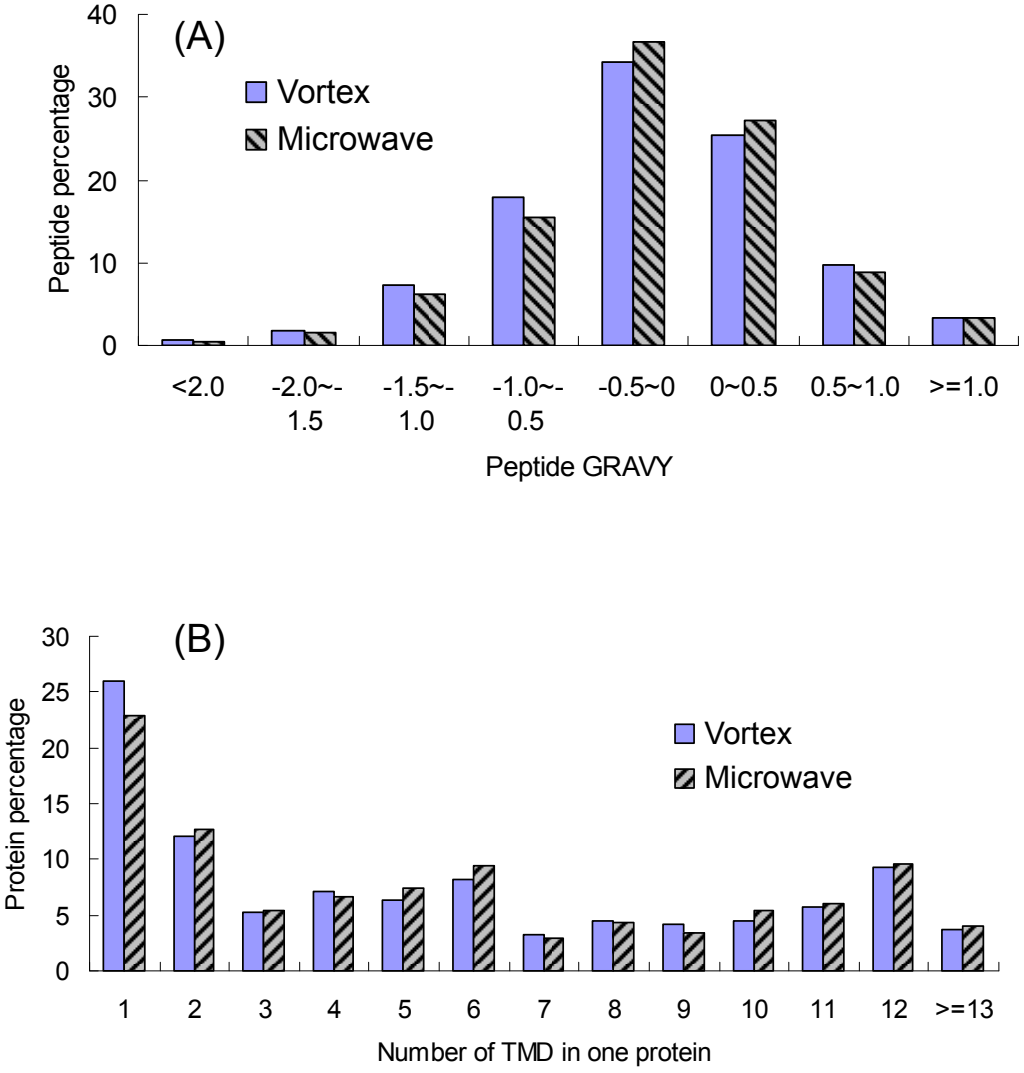


Figure 4.14 (A) Peptide distribution as a function of peptide hydrophobicity, gauged by the GRAVY indices grouped into eight bins. (B) Distribution of the number of transmembrane domains (TMDs) of identified proteins by the vortex and microwave methods.

the two methods. A total of 788 membrane proteins were identified. Among these, 615 proteins were identified by both methods, 126 proteins were identified exclusively by the microwave method and only 45 proteins were identified exclusively by the vortex method. This result indicates that 94% of the membrane proteins could be identified if only the microwave method were used, compared to 84% if the vortex method were exclusively used. Figure 4.14B displays the distribution of the number of transmembrane domains (TMDs) of identified proteins by the vortex and microwave methods. The protein percentage was calculated by the number of proteins in each catalog divided by the total number of proteins with at least one TMD identified. Overall, the distributions are similar. However, the proportions of identified proteins with higher than one TMD are 77.1% for the microwave method and 74.0% for the vortex method. Therefore, the microwave method is slightly more efficient in detecting the very hydrophobic membrane proteins.

4.4 Conclusions

We have developed a microwave-assisted protein solubilization method based on the use of a domestic microwave oven and the application of a series of 30 s of irradiation to dissolve proteins. MAPS offers three major advantages over the commonly used vortex-assisted solubilization method. First of all, MAPS reduces the overall solubilization time significantly. It was shown that, to dissolve a maximum amount of the *E. coli* K12 integral membrane proteins in NH_4HCO_3 or urea, two hours of vortex was needed, compared to 20 min required for carrying out 6 cycles of 30 s microwave irradiation (i.e., 3 min of total microwave time). After the first round of dissolution, more proteins can be dissolved by adding fresh reagent to the remaining protein pellet. To reach the maximum amount of proteins soluble in NH_4HCO_3 or urea, 7 rounds of 2 h vortex are needed, compared to 3 rounds of 20 min MAPS. Second, MAPS increases the total amount of proteins dissolvable in NH_4HCO_3 or urea. For example, starting with about 650 μg of the *E. coli* K12 integral membrane protein fraction, 85% was dissolved in urea by 7 rounds of 2 h vortex, compared to 95% by 3 rounds of 20 min MAPS. In the case of NH_4HCO_3, starting with about 650 μg of the membrane protein fraction, 28% was dissolved in NH_4HCO_3 by 7 rounds of 2 h vortex, compared to 51% by 7 rounds of 20 min MAPS. Finally, MAPS, in combination with sequential protein solubilization and shotgun LC-MS/MS analysis, allows identification of more proteins and peptides from a proteome sample than does the vortex method. By using the vortex method, a total of 1057 proteins and 6261 peptides were identified from the replicate experiments, while the MAPS method identified 1291 proteins and 10363 peptides. The significant increase in the number of proteins and peptides identified (22% and 66%, respectively) may be attributed to the increase in the amount of protein dissolved and higher digestion efficiency of the proteins that were likely more effectively denatured under microwave irradiation. The MAPS

method appears to identify slightly more hydrophobic proteins than the vortex method. In addition to the above listed advantages, MAPS can be carried out using an inexpensive microwave oven and thus can be readily implemented.

4.5 References

1. Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Yates, J. R., A method for the comprehensive proteomic analysis of membrane proteins. *Nature Biotechnology* **2003,** 21, (5), 532-538.
2. Rey, M.; Mrazek, H.; Pompach, P.; Novak, P.; Pelosi, L.; Brandolin, G.; Forest, E.; Havlicek, V.; Man, P., Effective Removal of Nonionic Detergents in Protein Mass Spectrometry, Hydrogen/Deuterium Exchange, and Proteomics. *Analytical Chemistry* **2010,** 82, (12), 5107-5116.
3. Botelho, D.; Wall, M. J.; Vieira, D. B.; Fitzsimmons, S.; Liu, F.; Doucette, A., Top-Down and Bottom-Up Proteomics of SDS-Containing Solutions Following Mass-Based Separation. *Journal of Proteome Research* **2010,** 9, (6), 2863-2870.
4. Chen, E. I.; Cociorva, D.; Norris, J. L.; Yates, J. R., Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *Journal of Proteome Research* **2007,** 6, (7), 2529-2538.
5. Zhang, N.; Chen, R.; Young, N.; Wishart, D.; Winter, P.; Weiner, J. H.; Li, L., Comparison of SDS- and methanol-assisted protein solubilization and digestion methods for Escherichia coli membrane proteome analysis by 2-D LC-MS/MS. *Proteomics* **2007,** 7, (4), 484-493.
6. Zhang, N.; Li, L., Effects of common surfactants on protein digestion and matrix-assisted laser desorption/ionization mass spectrometric analysis of the digested peptides using two-layer sample preparation. *Rapid Communications in Mass Spectrometry* **2004,** 18, (8), 889-896.
7. Lill, J. R.; Ingle, E. S.; Liu, P. S.; Pham, V.; Sandoval, W. N., Microwave-assisted proteomics. *Mass Spectrom Rev* **2007,** 26, (5), 657-671.
8. Sandoval, W. N.; Pharn, V. C.; Lill, J. R., Recent developments in microwave-assisted protein chemistries - can this be integrated into the drug discovery and validation process? *Drug Discovery Today* **2008,** 13, (23-24), 1075-1081.
9. Zhong, H. Y.; Zhang, Y.; Wen, Z. H.; Li, L., Protein sequencing by mass analysis of polypeptide ladders after controlled protein hydrolysis. *Nature Biotechnology* **2004,** 22, (10), 1291-1296.
10. Zhong, H. Y.; Marcus, S. L.; Li, L., Microwave-assisted acid hydrolysis of proteins combined with liquid chromatography MALDI MS/MS for protein identification. *Journal of the American Society for Mass Spectrometry* **2005,** 16, (4), 471-481.
11. Chen, W. Y.; Chen, Y. C., Acceleration of microwave-assisted enzymatic digestion

reactions by magnetite beads. *Analytical Chemistry* **2007,** 79, (6), 2394-2401.
12. Pramanik, B. N.; Mirza, U. A.; Ing, Y. H.; Liu, Y. H.; Bartner, P. L.; Weber, P. C.; Bose, M. K., Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes. *Protein Science* **2002,** 11, (11), 2676-2687.
13. Swatkoski, S.; Gutierrez, P.; Wynne, C.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C., Evaluation of microwave-accelerated residue-specific acid cleavage for proteomic applications. *Journal of Proteome Research* **2008,** 7, (2), 579-586.
14. Wang, N.; Li, L., Reproducible Microwave-Assisted Acid Hydrolysis of Proteins Using a Household Microwave Oven and Its Combination with LC-ESI MS/MS for Mapping Protein Sequences and Modifications. *Journal of the American Society for Mass Spectrometry* **2011,** 21, (9), 1573-1587.
15. Reiz, B.; Li, L., Microwave-Assisted Acid and Base Hydrolysis of Intact Proteins Containing Disulfide Bonds for Protein Sequence Analysis by Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2010,** 21, (9), 1596-1605.
16. Sze, S. K.; Wang, W.; Meng, W.; Yuan, R. D.; Guo, T. N.; Zhu, Y.; Tam, J. P., Elucidating the Structure of Cyclotides by Partial Acid Hydrolysis and LC-MS/MS Analysis. *Analytical Chemistry* **2009,** 81, (3), 1079-1088.
17. Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D., Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2002,** 1, (4), 351-360.
18. Wang, N.; Xie, C. H.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
19. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical Chemistry* **2008,** 80, (12), 4696-4710.
20. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Meth* **2007,** 4, (3), 207-214.
21. Kyte, J.; Doolittle, R. F., A Simple Method for Displaying the Hydropathic Character of A Protein. *Journal of Molecular Biology* **1982,** 157, (1), 105-132.
22. Trimpin, S.; Deinzer, M. L., Solvent-free MALDI-MS for the analysis of a membrane protein via the mini ball mill approach: Case study of bacteriorhodopsin. *Analytical Chemistry* **2007,** 79, (1), 71-78.
23. Zhang, N.; Li, N.; Li, L., Liquid chromatography MALDI MS/MS for membrane proteome analysis. *Journal of Proteome Research* **2004,** 3, (4), 719-727.
24. Quach, T. T. T.; Li, N.; Richards, D. P.; Zheng, J.; Keller, B. O.; Li, L., Development and applications of in-gel CNBr/tryptic digestion combined with mass spectrometry for the analysis of membrane proteins. *Journal of Proteome Research* **2003,** 2, (5), 543-552.
25. Zintl, A.; Pennington, S. R.; Mulcahy, G., Comparison of different methods for the

solubilisation of Neospora caninum (Phylum Apicomplexa) antigen. *Veterinary Parasitology* **2006,** 135, (3-4), 205-213.
26. Manadas, B. J.; Vougas, K.; Fountoulakis, M.; Duarte, C. B., Sample sonication after trichloroacetic acid precipitation increases protein recovery from cultured hippocampal neurons, and improves resolution and reproducibility in two-dimensional gel electrophoresis. *Electrophoresis* **2006,** 27, (9), 1825-1831.
27. Cacioppo, E.; Munson, S.; Pusey, M. L., Protein Solubilities Determined by a Rapid Technique and Modification of That Technique to a Micromethod. *Journal of Crystal Growth* **1991,** 110, (1-2), 66-71.
28. Wang, N.; MacKenzie, L.; De Souza, A. G.; Zhong, H. Y.; Goss, G.; Li, L., Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *Journal of Proteome Research* **2007,** 6, (1), 263-272.
29. Jacobs, D. I.; van Rijssen, M. S.; van der Heijden, R.; Verpoorte, R., Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured Catharanthus roseus cells yields 52% more spots after two-dimensional electrophoresis. *Proteomics* **2001,** 1, (11), 1345-1350.
30. Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R.; Li, L., Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *Journal of Proteome Research* **2008,** 7, (8), 3583-3590.
31. De Souza, A. G.; MacCormack, T. J.; Wang, N.; Li, L.; Goss, G. G., Large-Scale Proteome Profile of the Zebrafish (Danio rerio) Gill for Physiological and Biomarker Discovery Studies. *Zebrafish* **2009,** 6, (3), 229-238.
32. Angel, P. M.; Orlando, R., Quantitative carbamylation as a stable isotopic labeling method for comparative proteomics. *Rapid Communications in Mass Spectrometry* **2007,** 21, (10), 1623-1634.
33. McCarthy, J.; Hopwood, F.; Oxley, D.; Laver, M.; Castagna, A.; Righetti, P. G.; Williams, K.; Herbert, B., Carbamylation of proteins in 2-D electrophoresis - Myth or reality? *Journal of Proteome Research* **2003,** 2, (3), 239-242.

Chapter 5

Microwave-assisted Protein Solubilization Combined with Three-dimensional Liquid Chromatography MS/MS for Improving the Efficiency of Shotgun Proteome Analysis

5.1 Introduction

As one of the most widely used model system for biological studies, the microorganism *E. coli* has a relatively simple proteome. There are only about 4300 unique proteins or protein groups can be potentially expressed in *E. coli* K12 cells.[1] Therefore, during the course of developing improved analytical techniques for shotgun proteome analysis, profiling the *E. coli* proteome can be used to gauge the detectability of a certain technique. In Chapter 3, I have reported a comprehensive proteome profile of the *E. coli* K12 cell line using the sequential protein precipitation and solubilization method. A total of 3330 unique proteins were identified, representing 77% of the predicted proteome. However, there are several issues associated with this work.

First, protein solubilization was performed sequentially by using ammonium bicarbonate, methanol and SDS with the assistance of vortexing and sonication, which took more than a day to finish the whole process. To increase protein solubility in MS-compatible solvents as well as sample handling speed and protein identification efficiency, a new method of microwave-assisted protein solubilization (MAPS) was developed, as described in Chapter 4. Second, solubility-based protein level fractionation suffers from several drawbacks, such as low resolution, labor intensive, sample loss during sample transfer between different vials, and the subjective nature of the switching point for different solubilizing solvents. Therefore, finding an alternative way to perform protein fractionation is desirable. Protein fractionation based on liquid chromatography (LC) holds great potential for high efficiency separation of complex samples.[2, 3] The LC methods normally used for protein separation are based on size or molecular weight (e.g., size exclusion)[4, 5] and proteins functions (e.g., immunoaffinity chromatography).[6, 7] Among these methods, size exclusion suffers from low separation resolution, protein recovery and reproducibility. Immunoaffinity chromatography holds the highest specificity to reduce the complexity of biological samples, especially in the analysis of human plasma/serum. However, relatively low productivity, limited lifetime and high expenses make it not practical to be applied to large-scale proteome analysis. In this work, an LC-based protein separation by a macro porous-C18 column (mRP-C18) is used. The advantages of providing better

resolution, higher protein recovery, faster and more convenient procedure are expected, compared to the solubility-based protein separation method used in the previous chapters. Last but not the least, in the *E. coli* profiling work documented in Chapter 3, the protein sample was separated into more than 300 hundred individual fractions, and was analyzed by hundreds of LC-MS/MS runs, which required a great amount of starting material, long time for sample fractionation and inefficient usage of the MS instrument time.

In this work, we wish to address two issues: 1)can we develop a new protocol for faster analysis of the *E. coli* K12 proteome? and 2) by using this protocol, within a limited number of LC-MS/MS runs, how much proteome coverage can we reach? A new three-dimensional (3D)-LC MS/MS protocol was developed, in which microwave-assisted urea protein solubilization was combined with protein separation by the mRP-C18 column, followed by 2D-LC MS/MS analysis. As a comparison, the proteome coverage of a 2D-LC method based on the combination of the SCX and the capillary RPLC separations was also examined.

5.2 Experimental

5.2.1 Chemicals and Reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), sodium carbonate (Na_2CO_3) and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC-MS-grade water, acetone, formic acid, methanol (MeOH), and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada). A domestic 900 W (2450 MHz) microwave oven (Sunbeam or Panasonic purchased from a local store) was used to perform microwave-assisted protein solubilization experiments.

5.2.2 Cell culture and protein extraction

Escherichia coli K-12 (*E. coli,* ATCC 47076) was from the American Type Culture Collection. A single *E. coli* K12 colony was used to inoculate 50 mL of LB broth (BBL, Becton Dickinson). The culture was incubated overnight with shaking at 35 ˚C. This saturated culture (12.5 mL) was added to 500 mL of growth medium in a 2-L baffled

Erlenmeyer flask. Cells were harvested in the early stationary phase by centrifugation at 11,300 *g* for 15 min at 4 ˚C, resuspended, washed in 150 ml PBS buffer, and collected by centrifugation at 9,820 *g* for 15 min at 4 ˚C.

The *E. coli* cells were suspended in 45 mL PBS and three tablets of Roche mini protease inhibitor cocktail were added. The suspension was passed twice through a French press (Aminco Rochester, NY) using rapid fill kit at 20,000 psi. To extract membrane proteins, 5 mL 10% Triton X-100 (v/v) was added to the cell lysate to reach the final concentration of 1% Triton X-100 (v/v). The mixture was incubated with shaking at 4 ˚C for 20 min and centrifuged at 3,200 g for 15 min at 4 ˚C to remove unbroken cells, cell debris and cell walls. The supernatant was collected as the whole cell lysate. It was then stored at -80 ˚C for future use.

5.2.3 Protein quantification and purification by acetone

Protein concentration of the whole cell lysate was measured by BCA protein assay using bovine serum albumin (BSA) as the standard. Two aliquots of 500-μg protein sample were transferred out for further analysis. To remove the salts and other impurities, acetone, pre-cooled to -80 ˚C, was added gradually (with intermittent vortexing) to the two aliquots of protein extract to a final concentration of 80% (v/v). The mixture was kept at -20 ˚C overnight, transferred into 1.5-mL polypropylene vials with a flat top cap (Rose Scientific, Edmonton, Alberta) and centrifuged at 20,800 *g* for 10 min at 4 ˚C. The supernatant was decanted and properly disposed. Acetone was evaporated at the room temperature.

5.2.4 Protein solubilization by SDS and microwave-assisted protein solubilization (MAPS) in urea

Two aliquots of protein pellets were subjected to different solubilization methods (Figure 5.1). The first aliquot was solubilized in 2% SDS with sufficient vortex. The second aliquot of the protein pellets was dissolved in 8 M urea with the assistance of microwave irradiation as described in Chapter 3. Briefly, 8 M urea was added to the sample first. A sealed sample vial was then placed on the rotating tray inside a microwave. Microwave irradiation was applied for a total of 3 min in 30 s cycles with intermittent homogenization applied by vortex between cycles. Solubilized protein samples were reduced with 10 mM dithiothreitol (DTT) for 1 h at 37 ˚ C, and then alkylated with 20 mM iodoacetamide (IAA) for 1 h at the room temperature in dark.

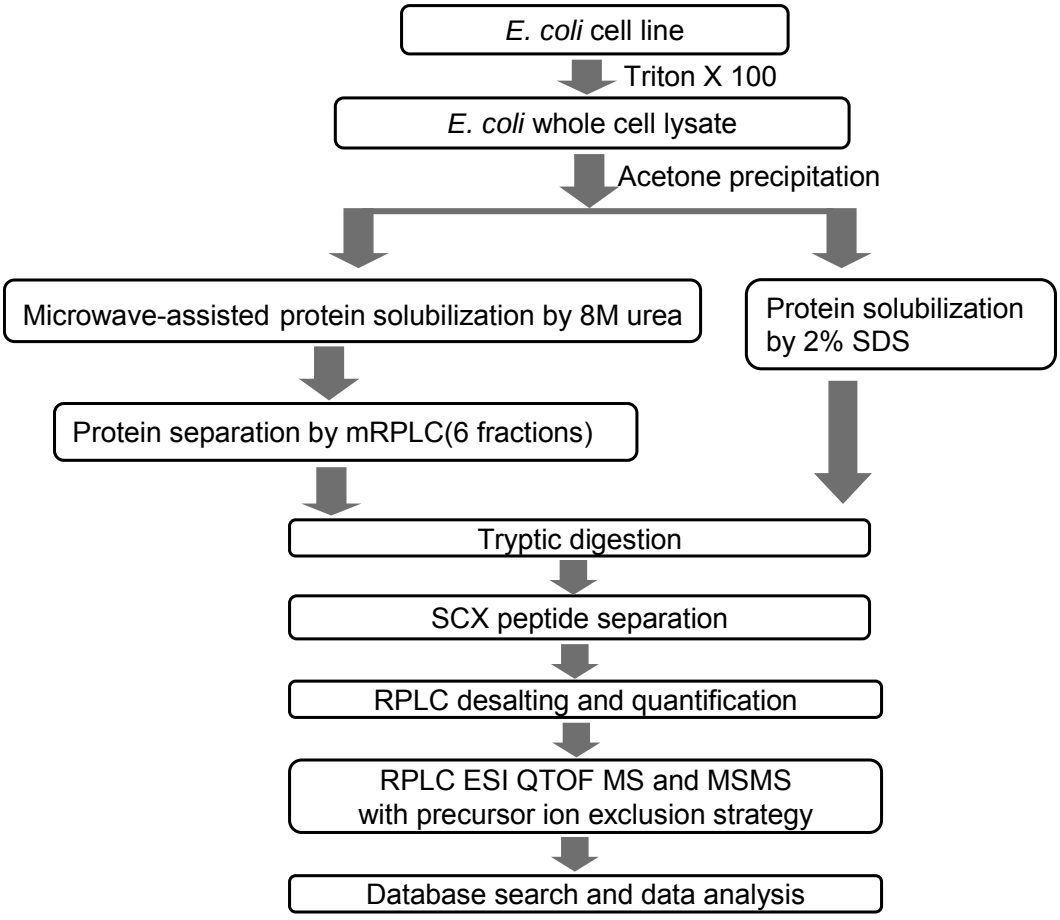


Figure 5.1 Workflow for sample preparation, separation and analysis.

5.2.4 Macro-porous reversed-phase C18 (mRP-C18) separation

The protein sample solubilized in urea by MAPS was separated by a 4.6 x 50 mm Macro-porous Reversed-phase C18 Column (5-μm particle, 5-μm pore size) (Agilent, Canada). Prior to mRP-C18 separation, 10% TFA was added into the protein sample to reach the final TFA concentration of 0.1%. Acetic acid was then added to adjust the sample pH to ~3.0. The protein sample was centrifuged at 20,800 g for 10 min at 4 ˚C to remove precipitates and the supernatant was subjected to the protein level separation. Solvent A consisted of 0.1% TFA in water, and solvent B consisted of 0.1% TFA in ACN. A 50-minute gradient (3 to 30% B for 10 min, 30 to 40% B for 13 min, 40 to 60% B for 12 min, ramped to 100% B in 5 min, held for another 5 min, 100 to 3% B for 5 min) was performed at a flow rate of 0.75 mL/min. Individual 2-min fractions were collected and pooled into 6 fractions containing an approximately equal amount of proteins in each fraction according to the UV absorbance values in the chromatogram recorded at 214 nm. Since urea introduced during the protein solubilization step was all eluted into the first pooled fraction, solvent in that fraction was carefully evaporated using SpeedVac (Thermo Savant, Milford, MA) to reach the final urea concentration of 1 M. Solvents in the other five pooled fractions were evaporated to the final volume of 500 μL. To avoid protein precipitation, 0.5 mmol of urea was added to each pooled fraction before solvent evaporation to reach the final concentration of 1M.

5.2.5 In-solution digestion

Before tryptic digestion, the SDS-solubilized protein sample was subjected to a 20-fold dilution. Trypsin solution was added to the SDS-solubilized sample and the protein sample eluted from the mRP-C18 column at an enzyme/protein ratio of 1:40, and the digestion was conducted at 37 °C for two overnights. Ten percent more fresh trypsin of the amount added for the first overnight was added before the second overnight digestion. Ten percent trifluoroacetic acid (TFA) was added to each peptide mixture to reach the final pH of 2.5. The mixtures were centrifuged at 20,000 g for 10 min and the supernatants transferred into fresh vials and stored at -80 °C until further analysis.

5.2.6 Strong cation exchange (SCX) liquid chromatography

A 2.1×250 mm highly hydrophilic polysulfoethyl A column (5 μm diameter, 300 Å pore, Poly LC, Columbia, MD) was used for the strong-cation exchange separation of the tryptic peptides. Solvent A (5 mM KH_2PO_4, 20% ACN, pH 2.7) and solvent B (solvent

A with 500 mM KCl) were used to develop a salt gradient (0 to 4% B for 1 min, 4 to 20% B for 16 min, 20 to 60% B for 22 min, ramped to 100% B in 6 min, held for another 10 min). One min or 2 min fractions were collected based on sample loading amount, and then adjacent fractions were pooled according to the UV absorbance values in the chromatograms.

5.2.7 Peptide desalting and quantification by RPLC

Desalting and quantification were carried out in an Agilent 1100 HPLC system (Palo Alto, CA) with 4.6 mm × 50 mm Polaris C18 A column with 3 μm particles and 300 Å pore size (Varian, CA). The eluted peptides were monitored and quantified using a UV detector operated at 214 nm using a method previously described.[8] After loading of the peptide sample, the column was flushed with mobile phase A (0.1% TFA in water) and the salts were effectively removed. Subsequently, the concentration of mobile phase B (0.1% TFA in ACN) in the mobile phase was step-wise increased to 85% to ensure complete elution of the peptide fractions from the column. During the peptide elution process, a chromatographic peak was produced and, based on the peak area, the amount of peptides was determined. BSA digests of various amounts were used as standards for the generation of a linear calibration between the peak area and the injected peptide amount. The calibration curve was generated as $y = 420x + 2056$, where y refers to the peak area of the peptide sample, 2056 refers to the peak area of the blank wash, and x refers to the peptide amount analyzed in μg. The linear range of the calibration curve was from 0.5 μg to 5.0 μg of peptides (R^2=0.998).

5.2.8 Mass spectrometric analysis

This step was performed as previously described with minor changes.[8, 9] 1 μg portions of peptides of each desalted SCX fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA) equipped with a nanoACQUITY Ultra Performance LC system (Waters). Peptide solution from each SCX fraction was injected onto a 75 μm × 100 mm Atlantis dC18 column with 3 μm particles and 100 Å pore size (Waters). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Mass spectra were acquired from m/z 300-1600 for 0.8 s, followed by 4 data-dependent MS/MS scans from m/z 50-1900 for 0.8 s each. The collision energy used to perform MS/MS was automatically varied according to the mass and charge state of the eluting peptide. A mass calibrant (i.e., lock-mass) was infused at a rate of 350 nL/min, and an MS scan was acquired for 1 s every 1 min throughout the run. Peptide precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides

identified from the adjacent two SCX fractions to enable additional and less abundant peptides to be analyzed and identified.[9] An exclusion list was generated based on Mascot (Matrix Science, London, U.K.) searching results of peptides with a score of 10 points above the identity threshold.

5.2.9 Protein database search

Database searches were performed as previously described with minor changes.[8, 9] Raw MS/MS data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.5 (Waters). Peptide sequences were identified by automated database searching of peak list files using the Mascot search program against the *E. coli* K12 database (4337 sequences). The following search parameters were selected for all database searching: enzyme, trypsin; missed cleavages, 1; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+, and 3+; fixed modification, carbamidomethyl (C); variable modifications, oxidation (M). The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, and Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses, were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at the confidence level of 99% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but different access ID numbers were also removed from the list. To gauge the false positive peptide matching rate, target-decoy search strategy was applied by searching the MS/MS spectra against the forward and reverse *E. coli* K12 proteome sequences.[10]

5.2.10 Hydropathy Calculation and Annotation of Localization

All peptides and proteins identified were examined using the ProtParam program, available at the EXPASY Web site (http://us.expasy.org/tools/protparam.html), which allows calculation of the grand average of hydropathy (GRAVY).[11] The Gene Ontology database (http://geneontology.org/) was used to classify proteins on the basis of cellular location. Transmembrane domains (TMD) of identified proteins were predicted by TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) according to protein primary sequences.

5.3 Results and discussion

5.3.1 Carbamylation rate

In this work, the microwave-assisted protein solubilization (MAPS) method was done in a different way from that described in Chapter 4. In brief, the sample vial was directly put in an ice bath rather than beside a beaker of water. The ice bath avoided the temperature rise caused by microwave irradiation and thus, reduced the extent or occurrence rate of carbamylation introduced by urea. Two modes of Mascot database search, with and without carbamylation on lysine and N-term as a variable modification parameter, were performed to the peak list files generated from both 2D- and 3D-methods. The results including the number of proteins and peptides identified, the number of carbamylated peptide and the rate of carbamylation are listed in Table 5.1. First, by using the new MAPS protocol, the carbamylation rate in the dataset of the 3D-LC method was found to be as low as 0.25%. Second, surprisingly, although no urea was used in the 2D-LC method, a similar rate of carbamylation (0.15%) was found in the result of the 2D-LC method; these carbamylated peptides were possibly resulted from random matches, i.e., false positives. Since it is difficult to judge whether these carbamylated peptide identifications in the 3D-LC method are real positive matches, carbamylation was not included as the variable modification parameter in the Mascot search for the 3D-LC method. Nevertheless, the carbamylation rate introduced by the MAPS method is very low.

Table 5.1 Comparison of search results by different searching parameters.

	Searching parameters	Protein number	Peptide number	Carbamylated peptide number	Carbamylation rate
2D-LC method	with carbamylation	1741	8432	13	0.15%
	without carbamylation	1901	9221	-	-
3D-LC method	with carbamylation	2072	12821	32	0.25%
	without carbamylation	2136	13712	-	-

5.3.2 Protein and peptide identification

The performance of a 3D-LC MS/MS protocol was evaluated in this work. In brief, 500

μg of the *E. coli* whole cell lysate was purified by cold acetone precipitation and re-solubilized in 8 M urea by MAPS. Solubilized proteins were separated by a macro-porous reversed-phase C18 (mRP-C18) column on the protein level into 6 fractions. Each protein fraction was then subjected to tryptic digestion, SCX fractionation of the peptides, followed by capillary RPLC-MS/MS analysis of the SCX peptide fractions. As a comparison, the same amount of the *E. coli* whole cell lysate sample was analyzed by a traditional 2D-LC MS/MS method where the protein sample purified by cold acetone was subjected to vortex-assisted solubilization by 2% SDS and tryptic digestion followed by SCX and capillary RPLC-MS/MS analysis. Figure 5.2 shows the comparison of proteins and peptides identified by using the 3D- and 2D-LC methods from the *E. coli* whole cell lysate. As Figures 5.2A and 5.3B show, from a total of 37 LC-MS/MS runs, 1901 unique proteins and 9221 different peptides were identified by the 2D-LC method, representing 44% of the possible 4300 proteins in the *E. coli* K12 cell line, and 2136 unique proteins with 13712 peptides were identified by the 3D-LC method by the same number of LC-MS/MS runs, representing 50% of the proteome. In total, 2333 unique proteins and 17158 different peptides were detected by the two methods, representing 54% of the total proteome, and the overlaps between the two methods are 73% at the protein level and 34% at the peptide level. Compared to the 2D-LC method, the 3D-LC method demonstrated 12% increase in the protein identification number and 41% increase in the peptide identification number. At the protein level, the larger overlap and smaller identification number difference between these two methods are understandable, considering the proteome of the *E. coli* K12 is relatively simple. However, the differences are much greater at the peptide level, as more diverse peptides were produced from the proteome digests. It should be noted that there were 36 and 90 sequence-redundant peptides found in the lists of the 2D- and 3D-LC methods, respectively. They were introduced due to the modification of peptides (i.e., the same peptides but with different modification groups, such as the presence or absence of oxidation in an amino acid in the peptide sequence). These modifications can be real from the sample or artifacts introduced during the sample preparation and analysis steps. However, compared to the large dataset of peptide identification (17158 different peptides), the number of sequence-redundant peptides is very small, indicating that the sample handling process including MAPS did not introduce a significant number of modifications.

In terms of the false positive rate, by using the 2D-LC method, reversed sequence search resulted in 436 matches, whereas the forward and reversed sequences search resulted in 27157 matches. Thus, the false positive peptide matching rate for the 2D-LC method was estimated to be 1.6%. By using the 3D-LC method, reversed sequence search resulted in 731 matches, whereas the forward and reversed sequences search resulted in 48815 matches. The false positive peptide matching rate for the 3D-LC method was estimated to be 1.5%.

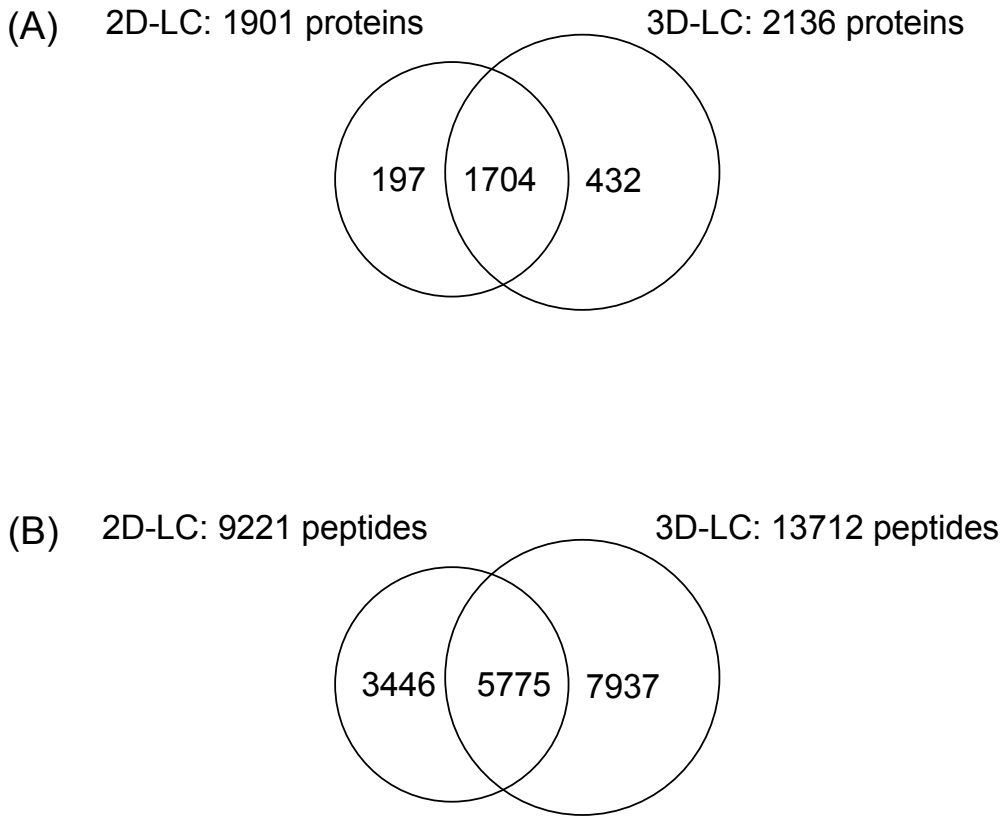


Figure 5.2 (A) Venn diagram of the proteins identified by 2D- and 3D-LC methods. A total number of 2333 unique proteins were identified. (B) Venn diagram of the peptides identified by 2D- and 3D-LC methods. 17158 different peptides were identified in total. (C) Venn diagram of the membrane proteins identified by 2D- and 3D-LC methods. 528 unique membrane proteins were identified in total.

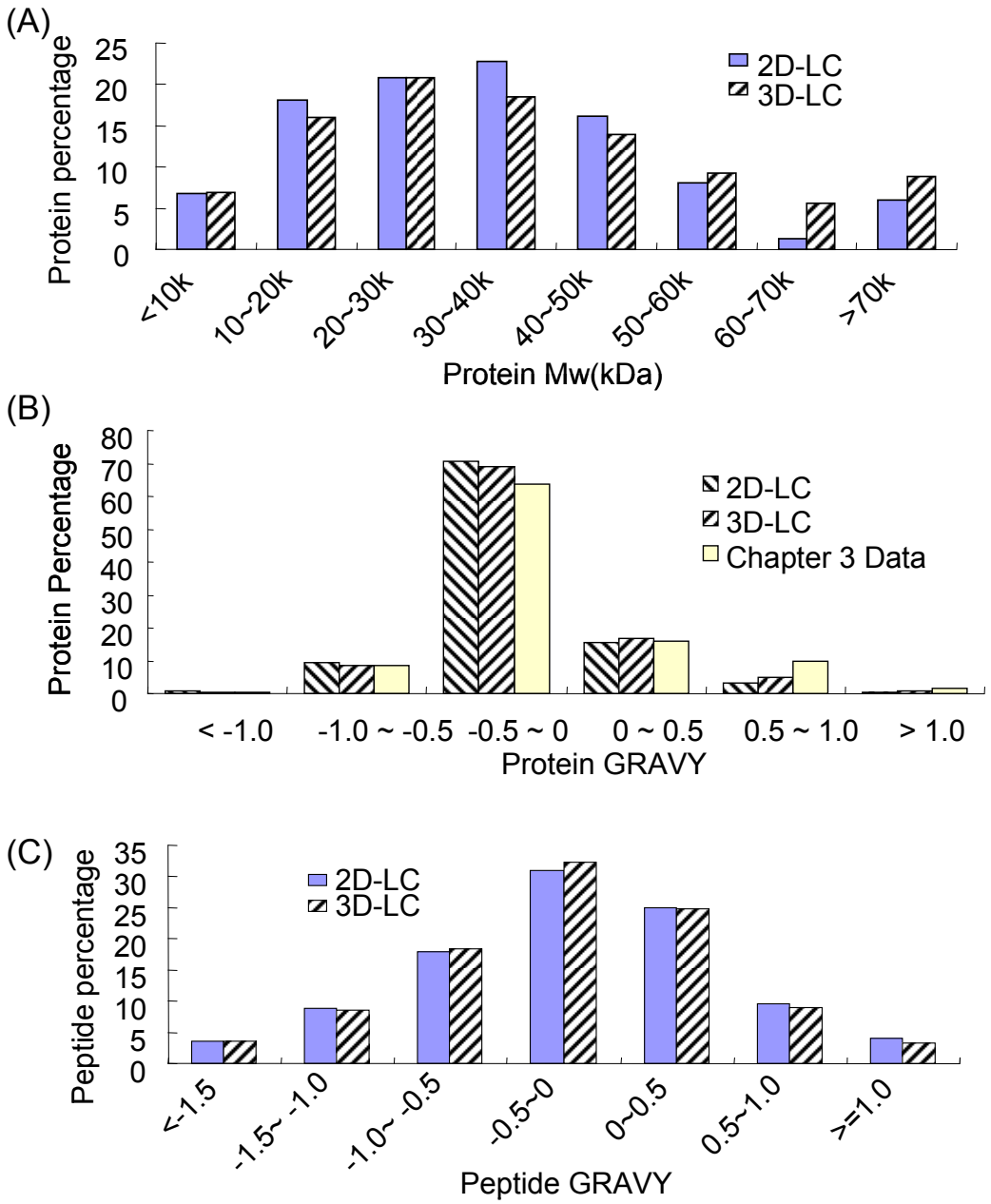


Figure 5.3 (A) The molecular weight distribution of proteins uniquely identified by 2D- and 3D-LC methods. (B) Protein GRAVY distribution of identified proteins by the 2D-and 3D-LC methods and in Chapter 3. (C) Peptide GRAVY distribution of identified peptides by the 2D- and 3D-LC methods.

As it was described in Chapter 4, the variations of two replicate experiments on the protein identification numbers were 4% for the vortex method and 3% for the microwave method. The variations on the peptide identification numbers were 7% for the vortex method and 1% for the microwave method. The overlaps of identified proteins between two replicates were 77% for the vortex method and 76% for the microwave method. It is clear that the reproducibility of both the MAPS and vortex methods are very high. Since the same techniques were used in this work, replicate experiments were not conducted, but expected to be similar to those shown in Chapter 4.

5.3.3 Characterization of identified proteins and peptides

To provide better understanding of the differences in peptide and protein identification results, we have examined the possible correlation between protein or peptide property and detectability of either method. Figure 5.3A shows the molecular weight distribution of the proteins identified by the two methods. The number of identified proteins was grouped into several bins as a function of the molecular masses, and then calculated into percentages. The two distributions shown in Figure 5.3A appear to be quite similar, except that slightly more proteins in the mass range higher than 50 kDa were detected in the 3D-LC method. Figures 5.3B and 5.3C show the protein and peptide distribution as a function of hydrophobicity, gauged by the GRAVY indices grouped into six and seven bins, respectively. There is no apparent difference between the 2D- and 3D-LC methods in terms of the protein and peptide GRAVY distribution. However, it should be noted that the *E. coli* proteins identified in Chapter 3 (Figure 3.7) show higher percentage in the category of highly hydrophobic proteins (i.e. GRAVY > 0.5). This result indicates the protein extraction method employed in this work might not be efficient in extracting highly hydrophobic proteins.

We have also examined the distribution of the number of membrane proteins identified by the two methods. As Figure 5.4A shows, a total of 528 membrane proteins were identified. Among these, 329 proteins were identified by both methods, 145 proteins were identified exclusively by the 3D-LC method and 54 proteins were identified exclusively by the 2D-LC method. This result indicates that 90% of the membrane proteins could be identified if only the 3D-LC method was used, compared to 73% if the 2D-LC method were exclusively used. Figure 5.4B displays the distribution of the number of transmembrane domains (TMDs) of identified proteins by the 2D- and 3D-LC methods and in Chapter 3. The protein percentage was calculated by the number of proteins in each catalog divided by the total number of proteins with at least one TMD identified. Overall, the distributions are similar. However, the proportions of identified proteins with higher than one TMD are 68% for the 3D-LC method and 61% for the 2D-LC method. These

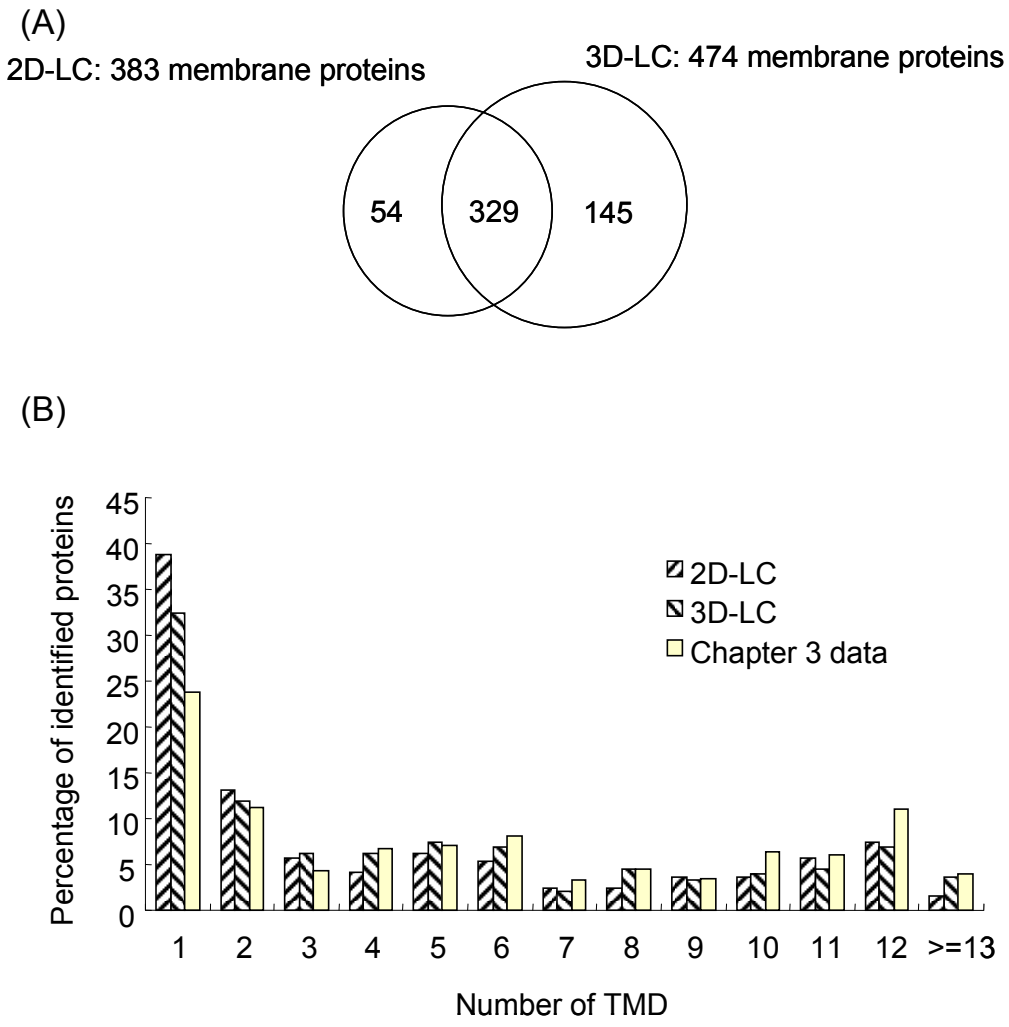


Figure 5.4 (A) Venn diagram of the membrane proteins identified by 2D and 3D-LC methods. A total number of 528 unique membrane proteins were identified. (B) Transmembrane domain (TMD) distribution of identified proteins by 2D and 3D-LC methods. The percentage is calculated based on the number of proteins with at least one TMD.

results indicate that the 3D-LC method is slightly more efficient in detecting the very hydrophobic membrane proteins. In the Chapter 3 data, the proportion of identified proteins with more than one TMD is 76%, which is consistent with the distribution of protein GRAVY.

5.3.4 Resolution of the mRP-C18 separation

Figure 5.5 shows the chromatogram generated from the separation of the *E. coli* whole cell lysate by the mRP-C18 column. The six fractions collected for further analysis were: 1^{st}-12^{th} min, 13^{th}-16^{th} min, 17^{th}-22^{nd} min, 23^{rd}-28^{th} min, 29^{th}-36^{th} min and 37^{th}-60^{th} min. The resolution of fractionation appeared to be fine from the chromatogram. However, the overlap of identified proteins between adjacent fractions was still as high as 54% (data not shown). Figure 5.6 shows the distribution of the number of identified proteins' appearance times in different mRP-C18 fractions by the 3D-LC method. Out of 2136 identified proteins, there were only 565 proteins identified from one unique mRP-C18 fraction, and 261 proteins were found in all the six fractions. Large overlaps may be possibly caused by several reasons. First, intra-protein interactions can cause the co-elution of two proteins at different retention times than those of individual proteins. Second, different conformations and post-translational modifications may cause the same protein eluted at different retention times. Third, protein degradation can also introduce retention time shifts since different sizes of protein molecules may have variable retention behaviors. The first two problems are due to insufficient denaturing of proteins, which may be improved by further optimization of microwave-assisted protein solubilization (e.g., higher irradiation energy, longer time, etc.) or the utilization of MS-compatible detergents (e.g.. acid cleavable detergents such as RapiGest). However, higher irradiation energy and longer microwaving time must be carefully used to avoid protein degradation. Furthermore, based on our current separation resolution, the proteome identification efficiency can also be increased by applying the peptide precursor ion exclusion (PIE) strategy to two adjacent protein fractions. This strategy was applied in this work.

5.3.5 Characterization of the precipitation induced by acidification before mRP-C18 separation

Prior to mRP-C18 separation, 10% TFA was added into the protein sample to reach the final TFA concentration of 0.1%. Acetic acid was then added to adjust to sample pH to ~3.0. A visible amount of protein precipitate was observed after acidification. The precipitate was collected, solubilized again in 8 M urea by MAPS, and digested by trypsin.

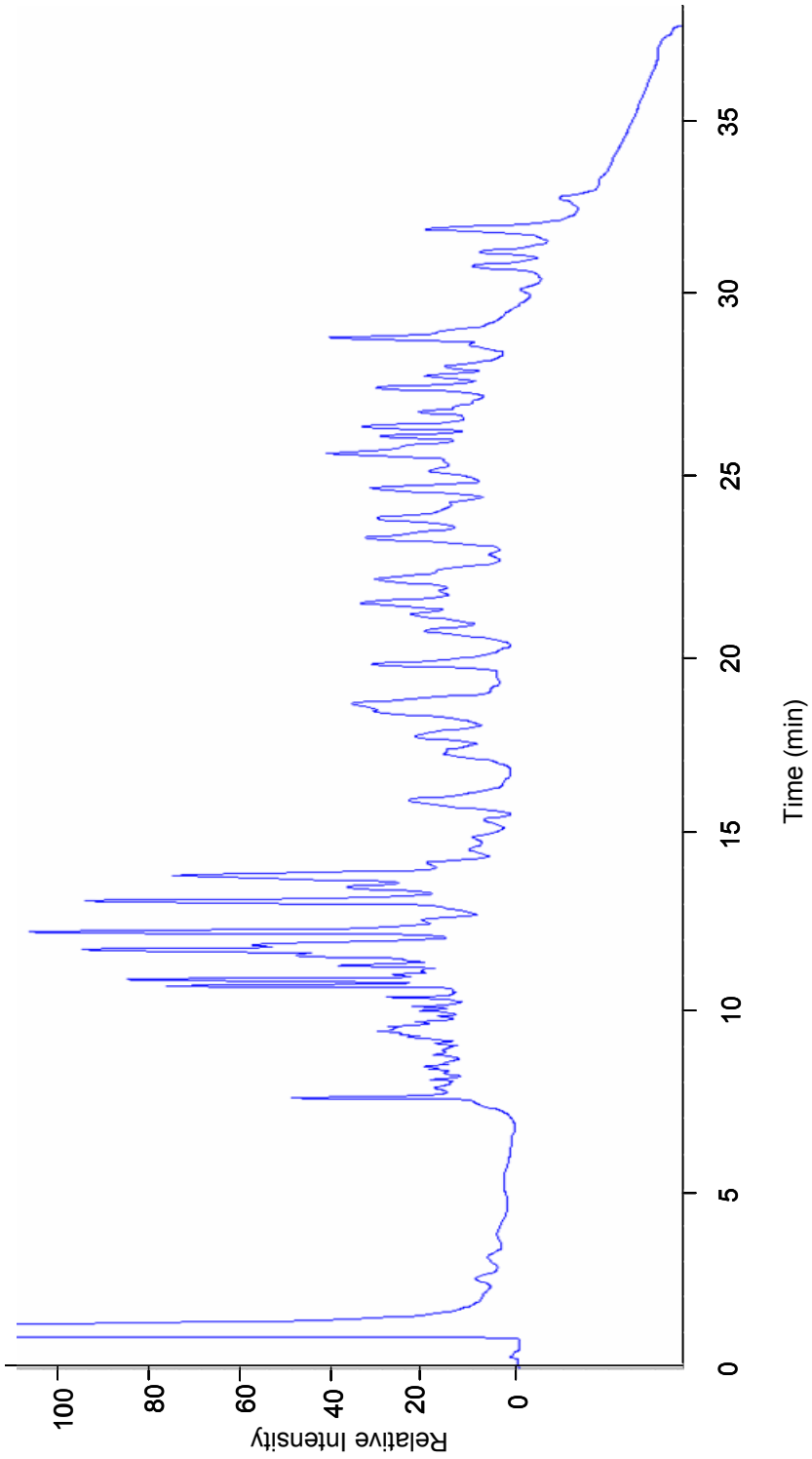


Figure 5.5 Chromatogram generated from the separation of the *E.coli* whole cell lysate by mRP-C18 column. Flow rate: 0.75 mL/min; Detection wavelength: 214 nm.

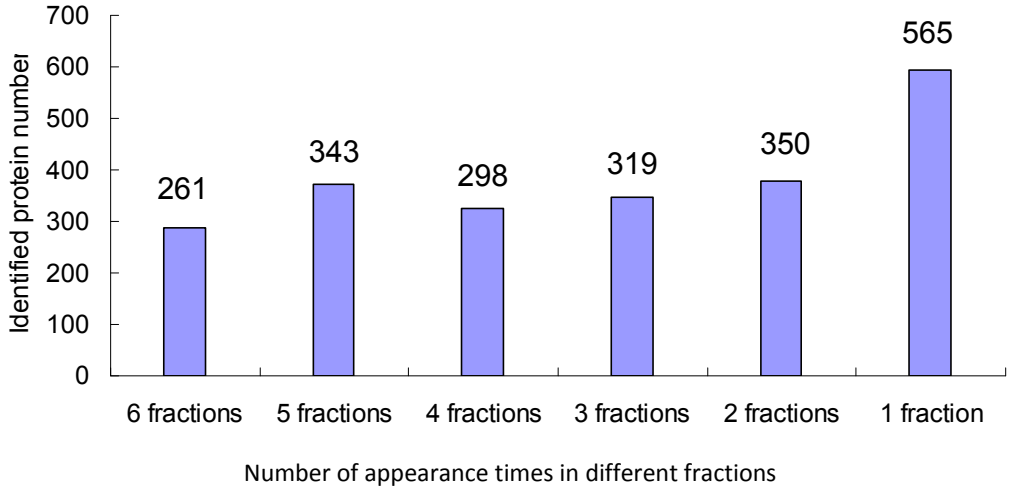


Figure 5.6 Distribution of number of identified proteins' appearance times in different mRP-C18 fractions by the 3D-LC method.

The protein digests were then subjected to SCX fractionation, followed by capillary RPLC MS/MS analysis with the same parameter setting as described above. Four pooled SCX fractions were collected, resulting in the identification of 1035 unique proteins and 4109 different peptides (protein and peptide lists not shown). Figures 5.7A and 5.7B present the comparison results of the proteins and peptides identified from the precipitate and the 3D-LC method. As Figures 5.7A and 5.7B show, compared to the results of the 3D-LC method, there were 1222 extra peptides and 46 extra proteins detected in the precipitate, representing only 8.9% and 2.1% increase in the peptide and protein identification number, respectively. The properties of these extra proteins and peptides found in the precipitate were investigated, including protein molecular weight (MW), protein and peptide GRAVY, protein TMD and the number of membrane proteins. No significant property differences from those of the proteins and peptides detected by the 3D-LC method were found (data not shown). These results indicate that prior to the mRP-C18 separation, the majority detectable proteins and peptides in the *E. coli* whole cell lysate still stay in the solution after acidification.

5.4 Conclusions

The purpose of this study was to evaluate the performance of the new 3D-LC MS/MS protocol for the analysis of the *E. coli* K12 proteome in terms of the proteome coverage that could be achieved with a limited amount of starting material and a limited number of LC MS/MS runs. In this method, microwave-assisted urea protein solubilization was combined with protein separation by the mRP-C18 column, followed by 2D-LC MS/MS analysis. This 3D-LC method identified 2136 unique proteins and 13712 different peptides from 500 μg of the *E. coli* whole cell lysate with 37 LC-MS/MS runs. These proteins represent 50% of the *E. coli* K12 proteome. This coverage is less than what we have achieved in Chapter 3 (i.e.,77%). However, in the work described in Chapter 3, more than 300 hundred of LC-MS/MS runs were performed and more than 100 mg of the *E. coli* protein sample was consumed. In contrast, less than 10% of the instrument time and less than 0.5% of the starting material were used in the 3D-LC method.

In this work, a traditional method where proteins were solubilized in 2% SDS followed by 2D-LC MS/MS was also taken into comparison. With the same amount of the protein sample and the same number of LC-MS/MS runs using the same instrument setting as in the 3D-LC method, the 2D-LC method identified 1901 unique proteins and 9221 different peptides. These proteins represent 44% of the *E. coli* K12 proteome. The overlap of protein identification between the 2D- and 3D-LC methods is as high as 74%. Compared to the 2D-LC method, the proteome coverage of the 3D-LC method is slightly better. The significant increase in the number of peptides identified (49%) may be attributed to two

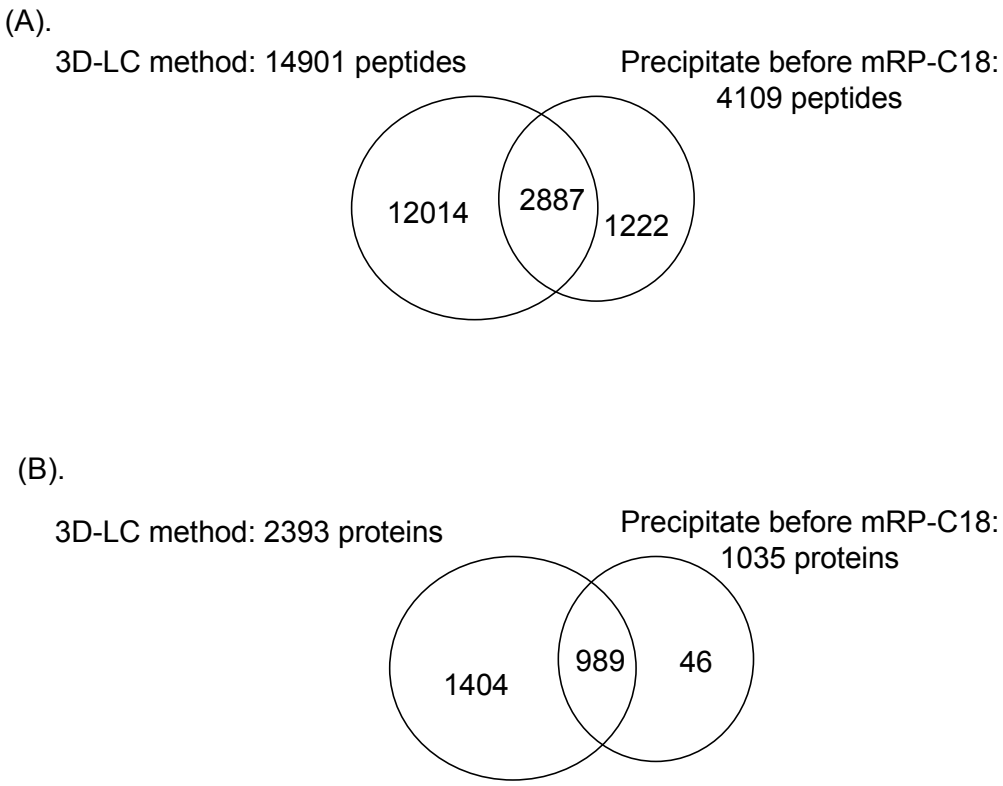


Figure 5.7 Venn diagram of the unique (A) peptides and (B) proteins identified from the eluate of mRPLC and the precipitate before mRPLC. 92.4% of peptides and 98.1% of identifiable proteins were found in the eluate from mRPLC.

factors. First, higher digestion efficiency of the proteins can be achieved in the 3D-LC method, because proteins are more effectively denatured under microwave irradiation. Second, the protein level fractionation by the mRP-C18 column provides an additional dimension of separation to the protein sample, which further simplified the sample to be analyzed by ESI-MS/MS. Although the total number of LC-MS/MS runs was the same, the additional peak capacity introduced by the mRP-C18 separation further reduces the ion suppression effect and thus increases the proteome identification efficiency. Furthermore, the 3D-LC method appears to identify slightly more membrane proteins than the 2D-LC method, which can also be attributed to the reduced ion suppression effect.

Despite the improvement achieved by the 3D-LC method, there are still several issues associated with this method. First, further investigation of the proteins uniquely identified in Chapter 3, but not by the 3D-LC method in this work, revealed that these proteins have higher GRAVY and TMD than the proteins detected by the 3D-LC method. The same situation happens for the 2D-LC method. Since all the protein samples were solubilized in both methods, this result indicates that there were hydrophobic proteins not being successfully extracted out from the *E. coli* cells. Therefore, to obtain better proteome coverage, more efforts should be made to improve the extraction efficiency of hydrophobic proteins in the future. Detergents such as NP40 or SDS are expected to improve the extraction efficiency. Second, the resolution of the mRP-C18 separation still needs some improvement to further enhance the protein identification efficiency. It was observed that the overlap of identified proteins between adjacent fractions was as high as 54%, which was mainly due to insufficient denaturing of proteins or possible protein degradation. These problems may be partially resolved by further optimization of the MAPS process or by using novel detergents such as Rapi-Gest, perhaps in combination with MAPS, for protein solubilization and denaturing.

5.5 References

1. Blattner, F. R.; Plunkett, G.; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; ColladoVides, J.; Glasner, J. D.; Rode, C. K.; Mayhew, G. F.; Gregor, J.; Davis, N. W.; Kirkpatrick, H. A.; Goeden, M. A.; Rose, D. J.; Mau, B.; Shao, Y., The complete genome sequence of Escherichia coli K-12. *Science* **1997,** 277, (5331), 1453-&.
2. Nagele, E.; Vollmer, M.; Horth, P.; Vad, C., 2D-LC/MS techniques for the identification of proteins in highly complex mixtures. *Expert Review of Proteomics* **2004,** 1, (1), 37-46.
3. Horvatovich, P.; Hoekman, B.; Govorukhina, N.; Bischoff, R., Multidimensional chromatography coupled to mass spectrometry in analysing complex proteomics samples. *Journal of Separation Science* 33, (10), 1421-1437.

4. Josic, D.; Horn, H.; Schulz, P.; Schwinn, H.; Britsch, L., Size-exclusion chromatography of plasma proteins with high molecular masses. *Journal of Chromatography A* **1998,** 796, (2), 289-298.
5. Stroink, T.; Wiese, G.; Lingeman, H.; Bult, A.; Underberg, W. J. M., Development of an on-line size exclusion chromatographic - reversed-phase liquid chromatographic two-dimensional system for the quantitative determination of peptides with concentration prior to reversed-phase liquid chromatographic separation. *Analytica Chimica Acta* **2001,** 444, (2), 193-203.
6. Cho, B. Y.; Zou, H. F.; Strong, R.; Fisher, D. H.; Nappier, J.; Krull, I. S., Immunochromatographic analysis of bovine growth hormone releasing factor involving reversed-phase high-performance liquid chromatography immunodetection. *Journal of Chromatography A* **1996,** 743, (1), 181-194.
7. Hoos, J. S.; Sudergat, H.; Hoelck, J. P.; Stahl, M.; de Vlieger, J. S. B.; Niessen, W. M. A.; Lingeman, H.; Irth, H., Selective quantitative bioanalysis of proteins in biological fluids by on-line immunoaffinity chromatography-protein digestion-liquid chromatography-mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2006,** 830, (2), 262-269.
8. Wang, N.; Xie, C. H.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
9. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical Chemistry* **2008,** 80, (12), 4696-4710.
10. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Meth* **2007,** 4, (3), 207-214.
11. Kyte, J.; Doolittle, R. F., A Simple Method for Displaying the Hydrpathic Character of a Protein. *Journal of Molecular Biology* **1982,** 157, (1), 105-132.

Chapter 6

Multidimensional mRP-RPLC Separation of Proteins and Peptides Combined with ESI-MS/MS for Comprehensive Profiling of the Phosphoproteome of MDA-MB-231 cells

6.1 Introduction

Protein phosphorylation is a post-translational modification normally found on serine (Ser), threonine (Thr) and tyrosine (Tyr) residues.[1] The study of phosphoproteome is important as phosphorylation regulates the functions of proteins and thus modulates many biological processes, such as complex formation, regulation of transcription factors, cell proliferation and apoptosis.[2] However, it is a challenging task to study protein phosphorylation using shotgun proteomic strategy, because many phosphorylated proteins are present at low concentrations compared to their native counterparts.[3] To reduce ion suppression in mass spectrometry, enrichment and fractionation of phosphopeptides become crucial for the analysis of the phosphoproteome of biological samples. Despite the huge progress made to increase the enrichment efficiency for phosphopeptides,[4-11] there is still rooms for improvement in the area of protein and peptide fractionation to reduce the sample complexity that would further enhance the detectability of phosphopeptides.

Conventionally, two-dimensional LC-MS/MS combined with phosphopeptide enrichment is employed for phosphoproteome analysis, where tryptic digests are separated by strong cation exchange (SCX) first, followed by phosphopeptide enrichment by immobilized metal ion affinity chromatography (IMAC) or titanium dioxide (TiO_2).[8] The individual fractions are then analyzed by reversed-phase (RP) LC-MS/MS. As the SCX method separates the peptides based on their charge states, phosphopeptides are always eluted out in the low salt region of the gradient due to the negatively charged phosphate groups. SCX is not an ideal fractionation method for phosphopeptides. It is desirable to use additional fractionation or alternative fractionation method to simplify the phosphoproteome.[12] As it has been described in the previous chapters, fractionation of the proteome whole cell lysate at the protein level by the mRP-C18 column is shown to be orthogonal, to some extent, to the separation by capillary RPLC at the peptide level. One would expect that the use of protein separation should add another dimension of simplifying the phosphoproteome or an alternative approach to the traditional peptide level enrichment scheme for phosphoproteome profiling.

In this work, a new mRP-RPLC fractionation protocol was developed for analyzing the phosphoproteome of the MDA-MB-231 cell line which is widely used as a model system for breast cancer research.[13-35] In this protocol, MDA-MB-231 whole cell lysate was first separated by the mRP-C18 column at the protein level into 30 fractions. Each fraction was subjected to tryptic digestion, sequential phosphopeptide enrichment by IMAC and TiO_2 modified from the SIMAC method, [11, 36] followed by capillary RPLC-MS/MS analysis. As a comparison, the same sample was also analyzed by the traditional SCX-RPLC method, where the tryptic digest of the MDA-MB-231 whole cell lysate was subjected to sequential phosphopeptide enrichment by IMAC and TiO_2, followed by SCX fractionation and capillary RPLC-MS/MS analysis. The results generated from the replicate experiments indicate that these two methods have similar ability in the identification of the phosphoproteome, but produce complementary information to each other. Therefore, the coverage of the MDA-MB-231 phosphoproteome was increased by combining the identification results from both methods. Among the 1947 phosphoproteins identified in this work, we identified 57 phosphoproteins previously reported to be functionally related to the breast cancer, and 180 phosphoproteins that have never been reported to be phosphorylated.

6.2 Experimental

6.2.1 Chemicals and reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4) and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC/MS-grade water, acetone, formic acid, and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada).

6.2.2 Cell culture and protein extraction

MDA-MB-231 was cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma, Ontario, Canada), supplemented with 10% heat inactivated fetal bovine serum (FBS) and 150 units/mL penicillin, under an atmosphere of 95% O_2 and 5% CO_2 in 98% humidity at 37 °C. Figures 6.1 and 6.2 show the workflow used to profile the phosphoproteome of

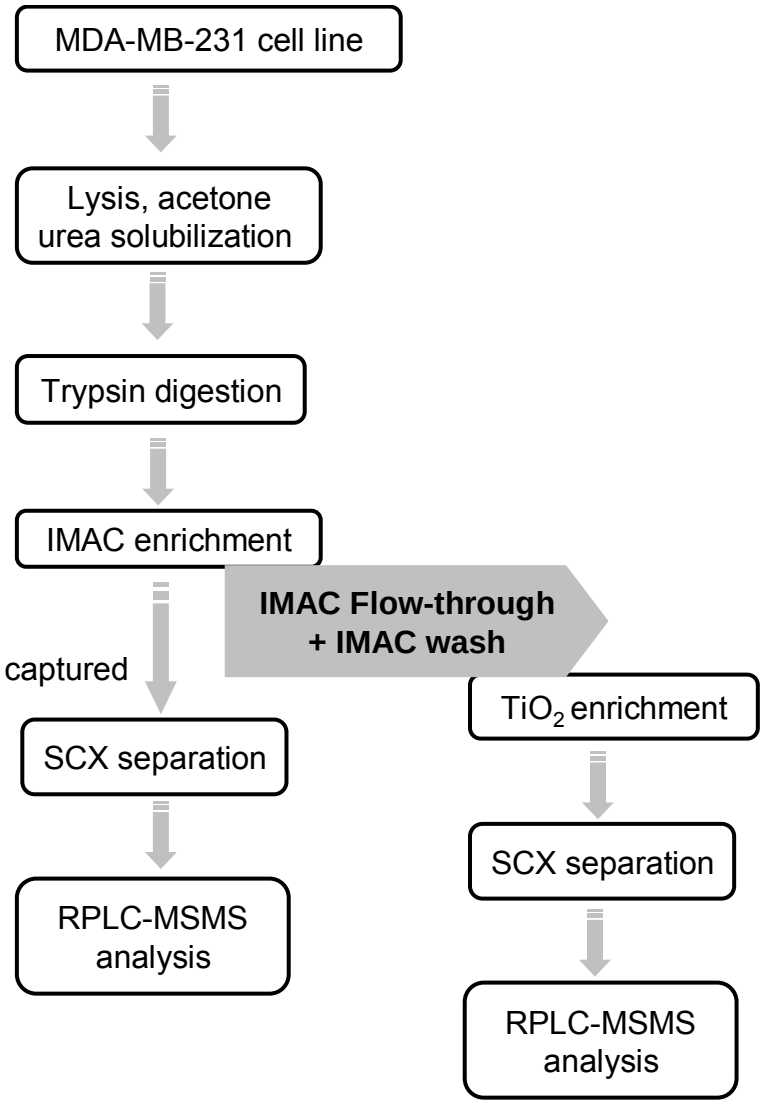


Figure 6.1 Workflow of the SCX-RPLC methods.

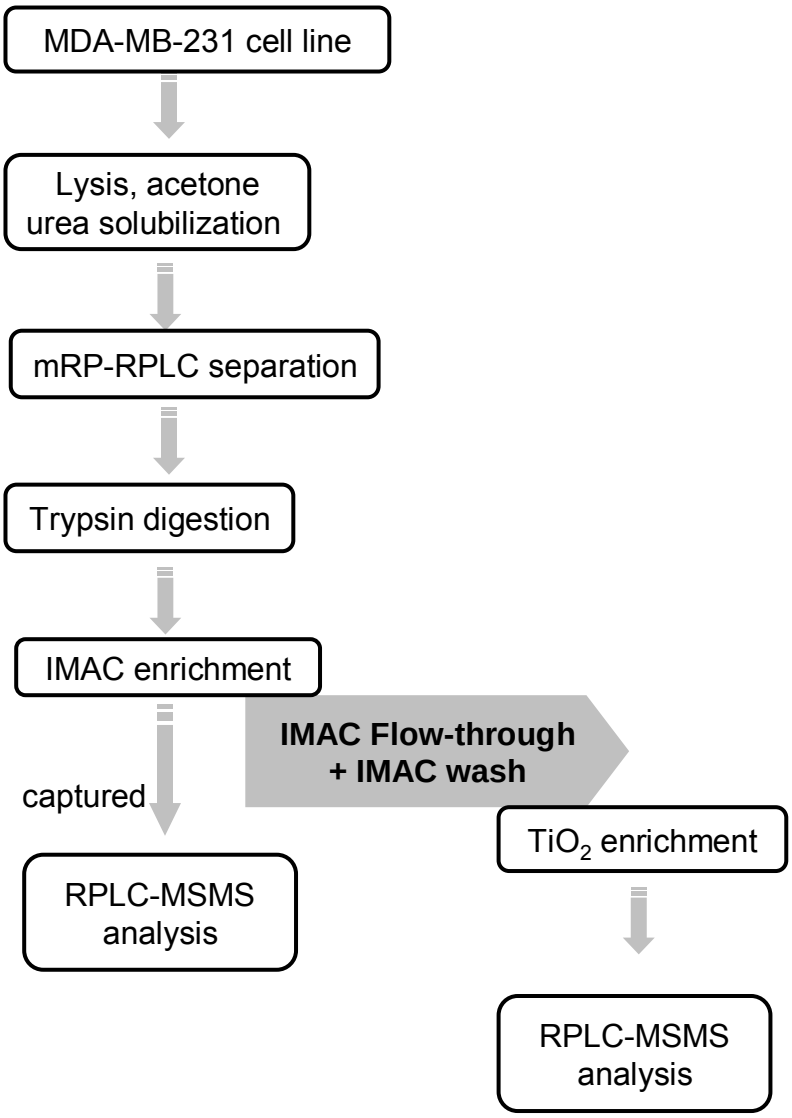


Figure 6.2 Workflow of the mRP-RPLC methods.

MDA-MB-231 cells. Cells were lysed in CelLytic™ M buffer (Sigma, Ontario, Canada), 1 mM phenylmethylsulfonyl fluoride, protease inhibitor cocktail, and phosphotease inhibitor cocktail (Sigma, Ontario, Canada). The lysates were centrifuged at 15,000 g for 15 min to remove insoluble materials. A BCA assay Kit (Bio-Rad, Ontario, Canada) was used to determine the protein concentration of the lysate. About 10 mg of each sample were used to further analysis. To remove the salts and other impurities, acetone was added into cell lysates with vortexing to final acetone concentration of 80%. The mixture was incubated at -20 °C overnight. The mixture was then centrifuged at 20,000 g for 10 min. After decanting the supernatant, the pellets were dissolved with 8 M urea with intermittent vortexing. The protein mixture was equally split, reduced with 10 mM dithiothreitol (DTT) for 1 h at 37 °C, and then alkylated with 20 mM iodoacetamide (IAA) for 1 h at room temperature in dark.

6.2.3 In-solution digestion

Trypsin solution was added into one aliquot of the MDA-MB-231 protein sample at a ratio of 1:100 after 10-fold dilution. The mixture was incubated at 37 °C overnight. Ten percent trifluoroacetic acid (TFA) was added to the peptide mixture to make the final pH ~2.5. The mixture was centrifuged at 20,000 g for 10 min and the supernatant transferred into a fresh vial for further analysis.

6.2.4 Sequential phosphopeptides enrichment by immobilized metal ion affinity chromatography (IMAC) and titanium dioxide

Prior to IMAC, the peptides were desalted by using Agilent 1100 HPLC with 4.6 × 50 mm C18 column (3 μm particles and 300 Å pore size, Varian, Ontario, Canada), and then lyophilized. Sequential phosphopeptide enrichment was performed as descried previously[36] with modifications. After re-solubilization with 30% acetonitrile (ACN) mixed with 250 mM acetic acid, the peptides were loaded onto pre-equilibrated Fe-IMAC resin (Phos-Select iron affinity gel, Sigma, Ontario, Canada). The resin was washed three times with 30% acetonitrile mixed with 250 mM acetic acid, two times with water after overnight incubation at 4 °C. The IMAC flow-through and washing buffer were collected for further analysis by using the TiO_2 beads (see below). The phosphopeptides were then released from the resin with 400 mM ammonium hydroxide. After lyophilization, the pooled flow-through and wash from the IMAC was enriched for phosphopeptides using the TiO_2 beads (Titansphere, 5 μm) (GL-Sciences, Inc. Japan). The enrichment procedure has been described elsewhere[36] with modifications. In brief, the lyophilized sample was

re-solubilized in 200 μL loading buffer (65% ACN/2% TFA/saturated by phithalic acid), and then incubated with pre-equilibrated TiO_2 beads. The incubated beads were then washed with 800 μL wash buffer I (65% ACN/0.5% TFA) and wash buffer II (65% ACN/0.1% TFA). The phosphopeptides were eluted once with 200 μL elution buffer I (300 mM NH_4OH/50% ACN) and twice with 200 μL elution buffer II (500 mM NH_4OH/60% ACN). All the incubation, washing, as well as elution procedures were performed at room temperature for 20 minutes.

6.2.5 Strong cation exchange (SCX) liquid chromatography

A 2.1 x 250 mm poly SULFOETHYL^{TM} A column (5-μm particles, 300-Å pore) (Poly LC, Columbia, MD) was used for the SCX separation of the phosphopeptide-enriched mixture. Solvent A (5 mM KH_2PO_4, 20% ACN, pH 2.7) and solvent B (solvent A with 500 mM KCl) were used to develop a salt gradient (0-7% for 5 min, 7-42% for 35 min and 42-100% for 2 min). All the samples collected from SCX were desalted on an Agilent 1100 HPLC system and quantified by UV absorbance.[37]

6.2.6 mRP-C18 separation，in-solution digestion and sequential phosphopeptide enrichment by IMAC and TiO_2 beads

The second aliquot of the MDA-MB-231 whole cell lysate was separated by a 4.6 x 50 mm Macroporous Reversed-phase C18 Column (5-μm particle) (Agilent, Canada). Solvent A consisted of 0.1% TFA in water, and solvent B consisted of 0.1% TFA in ACN. A 50-min gradient (3 to 30% B for 10 min, 30 to 40% B for 13 min, 40 to 60% B for 12 min, ramped to 100% B in 5 min, held for another 5 min, 100 to 3% B for 5 min) was performed at a flow rate of 0.75 mL/min. Individual 1-min or 2-min fractions were collected based on the sample loading amount, and pooled into 30 fractions with an approximately equal amount of proteins according to the UV absorbance values in the chromatograms recorded at 214 nm. Each fraction was digested by trypsin, and phosphopeptides were sequentially enriched by IMAC and TiO_2 beads as described previously. All the phosphopeptide-enriched samples eluted from IMAC and TiO_2 beads were desalted individually on an Agilent 1100 HPLC system and quantified by the UV absorbance.[37]

6.2.7 Mass spectrometric analysis

Tandem mass spectrometry analysis was performed as previously described.[38] Approximately 1.0 μg of peptides from individual fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA). Peptide solution was injected onto a 75 × 100 mm Atlantic dC18 column (Waters, Milford, MA). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Peptides were separated using a 120-min gradient (2 to 7% B for 2 min, 7 to 8% B for 14 min, 8 to 20% B for 69 min, 20 to 30% B for 25 min, 30 to 45% B for 5 min, and 45 to 90% B for 5 min) and electrosprayed into the mass spectrometer fitted with a nanoSpray source at a flow rate of 350 nL/min. Mass spectra were acquired from m/z 300-1600 for 0.8 second, followed by 6 data-dependent MS/MS scans from m/z 50-1990 for 1.0 second each. The collision energy used to perform MS/MS was automatically varied according to the mass and the charge state of the eluting peptide. A mass calibrant (i.e., lock-mass) was infused at a flow rate of 300 nL/min, and an MS scan was acquired for 1 s every 1 min throughout the run. Precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides identified from the adjacent two SCX/mRP-C18 fractions to enable additional less abundance peptides to be analyzed and identified. An exclusion list was generated as described previously.[38]

6.2.8 Database search and data analysis

Raw MS/MS data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.3 (Waters, Milford, MA). Protein identification was performed by using Mascot 2.2 search engine (http://www.matrixscience.com/) for searching the Swiss-Prot database (http://ca.expasy.org/sprot/) (Swiss-Prot database; version 57.4; 410518 sequences). Searching was restricted *Homo sapiens* (20401 sequences) and preformed using the following parameters: fixed modification, carbamidomethyl (C); variable modifications, oxidation (M), phosphorylation on serine, threonine or tyrosine; enzyme, trypsin; missed cleavages, 2; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+ and 3+.

The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, and Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses, were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at the confidence level of 95% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but

different access ID numbers were also removed from the list. The false positive rate was determined by employing the target-decoy search strategy.[39] The overall false positive rate was 1.5%.

A 79.96 Da mass increase found on a serine, threonine or tyrosine residue indicates phosphorylation of the residue. The spectra of phosphoproteins that were previously found to be functionally involved in the breast cancer and all the newly found phosphorylated proteins were manually checked based on the following rules. First, the y and/or b ions of serine and threonine phosphorylated peptides can be subjected to the neutral losses of H_3PO_4. An 18 Da mass difference between serine and dehydroalanine or threonine and dehydroaminobutyric acid was also considered. Second, as neutral loss is not normally found in the fragmentation of phosphotyrosine, immonium ion at the m/z value of 216 indicates the present of phosphotyrosine. Additionally, the cognate pair of fragment ions (y, y-79.96, and/or b, b-79.96) and y and/or b ions containing a phosphate group were also considered. Finally, continuous fragment ion series (y and/or b ions) around the phosphorylated residue were investigated, which can further confirm the localization of a phosphorylation site.

All peptides and proteins identified were examined using the ProtParam program, available at the EXPASY Web site (http://us.expasy.org/tools/protparam.html), which allows calculation of the grand average of hydropathy (GRAVY).[40]

Functional pathway analysis was generated by the Ingenuity Pathway Analysis (IPA) software (www.ingenuity.com) (Ingenuity Systems, Redwood, CA). The significance of the association between the data set and the canonical pathway was measured in two ways. First, a ratio of the number of proteins from the data set that map to the pathway to the total number of proteins that map to the canonical pathway is calculated. Second, Fisher's exact test was used to calculate a p-value representing the probability of the association between the proteins in the data set and the canonical pathway. A p-value less than 0.05 indicates a protein is not randomly associated with a canonical pathway.

6.3 Results and discussion

6.3.1 Identification of phosphopeptides

In this work, two methods were used to analyze the phosphoproteome of the MDA-MB-231 cells: SCX-RPLC and mRP-RPLC. Figures 6.1 and 6.2 demonstrate the workflow of these two methods. In the SCX-RPLC method, cell lysates were first digested by trypsin, fractionated by SCX into 30 fractions. Individual fractions were

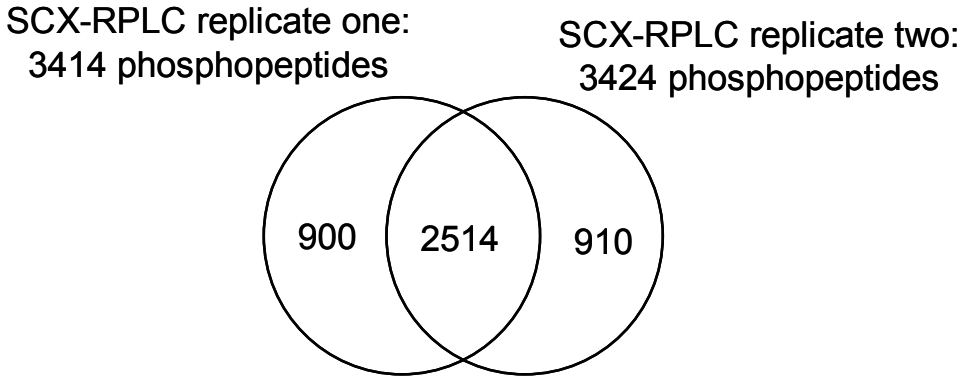
subjected to sequential phosphopeptide enrichment by IMAC and TiO_2 beads and finally analyzed by capillary RPLC-MS/MS. In the mRP-RPLC method, cell lysates were first fractionated by an mRP-C18 column at the protein level into 30 fractions. Individual fractions were subjected to tryptic digestion, followed by sequential phosphopeptides enrichment by IMAC and TiO_2 beads. Each phosphopeptide sample was finally analyzed by capillary RPLC-MS/MS with the same instrument setting as the SCX-RPLC method. Replicate experiments were performed for both methods.

As Figure 6.3 shows, by using the traditional SCX-RPLC method, 1422 unique phosphoproteins with 3414 different phosphopeptides were identified from the first replicate experiment, and 1383 unique proteins with 3424 different phosphopeptides were detected from the second replicate experiment. The overlaps between the two replicates were 77% at the protein level and 59% at the peptide level. As Figure 6.4 shows, by using the mRP-RPLC method, 1455 unique phosphoproteins with 3699 different phosphopeptides were identified from the first replicate experiment, and 1399 unique proteins with 3738 different phosphopeptides were detected from the second replicate experiment. The overlaps between the two replicates were 80% at the protein level and 65% at the peptide level. Combining the results of the two replicate experiments from the SCX-RPLC method, 1585 unique phosphoproteins with 4297 different phosphopeptides were identified. Using the mRP-RPLC method, 1502 unique phosphoproteins with 4029 different phosphopeptides were found. Combining the results of the two methods, a total number of 1947 unique phosphoproteins and 6278 different phosphopeptides were identified. These results clearly show that the SCX-RPLC and the mRP-RPLC methods have similar ability for the identification of the MDA-MB-231 phosphoproteome in terms of the number of identified phosphoproteins and phosphopeptides. On the other hand, as Figures 6.5A and 6.5B show, the overlaps between these two methods were 33% at the peptide level and 59% at the protein level. Considering the overlaps between the replicate experiments of either method were significantly lower than the overlaps between the two methods, we can conclude that the phosphoproteins and phosphopeptides identified by the SCX-RPLC and the mRP-RPLC methods are complementary to each other. Therefore, by combining the results of these two methods, the coverage of the MDA-MB-231 phosphoproteome can be enhanced.

6.3.2 Sequential phosphopeptide enrichment

To increase the extent of phosphopeptide enrichment, sequential phosphopeptide enrichment by IMAC and TiO_2 beads was performed in both SCX-RPLC and mRP-RPLC methods. The phosphopeptides, mainly mono-phosphorylated, in the flow through of the IMAC enrichment was captured by the TiO_2 beads. The enriched phosphopeptides by

(A)



(B)

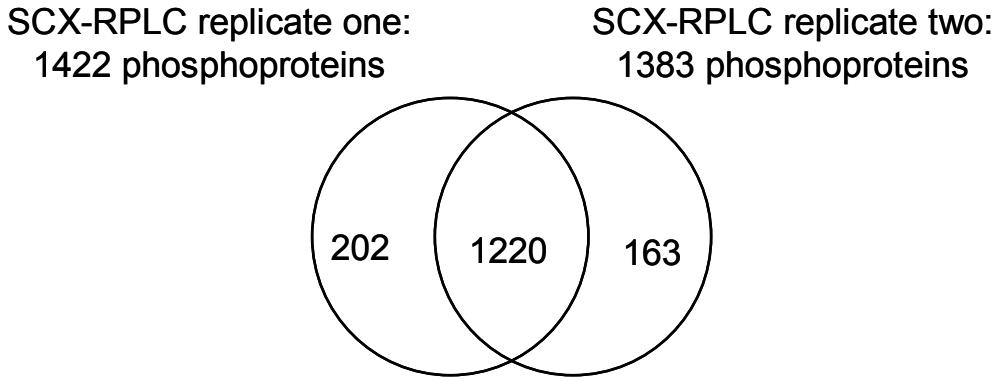
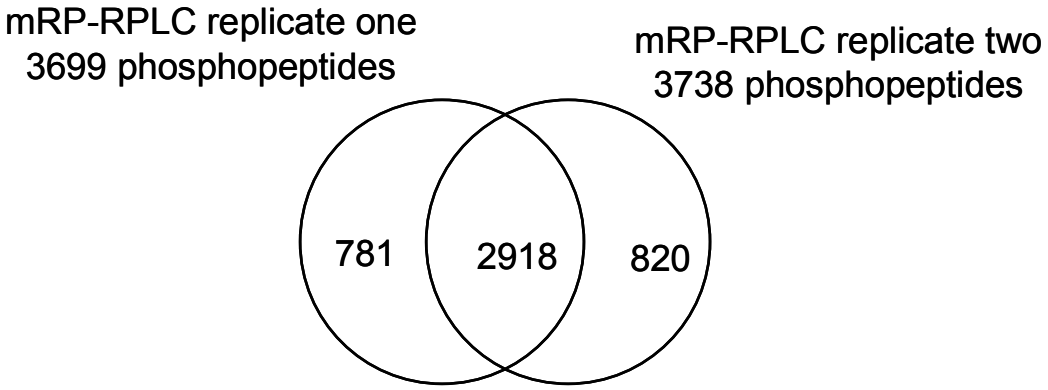


Figure 6.3 (A) Venn diagram of the phosphopeptides identified by the replicate experiments of the SCX-RPLC method. (B) Venn diagram of the phosphoproteins identified by the replicate experiments of the SCX-RPLC method. A total of 4297 different phosphopeptides and 1585 unique phosphoproteins were identified.

(A)



(B)

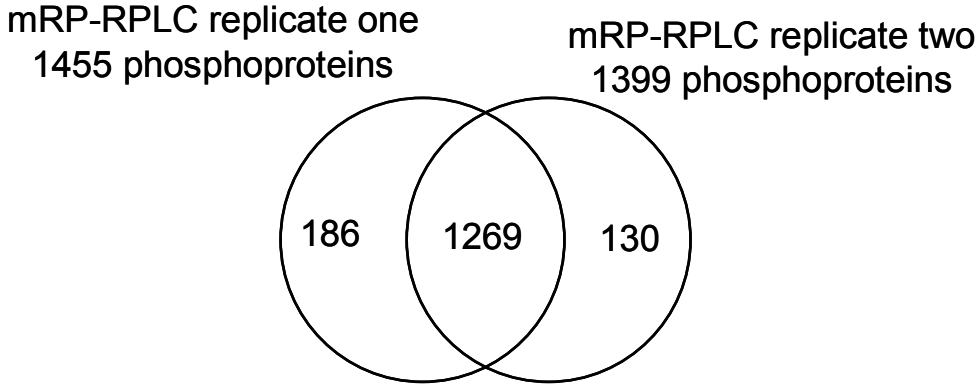


Figure 6.4 (A) Venn diagram of the phosphopeptides identified by the replicate experiments of the mRP-RPLC method. (B) Venn diagram of the phosphoproteins identified by the replicate experiments of the mRP-RPLC method. A total of 4519 different phosphopeptides and 1585 unique phosphoproteins were identified.

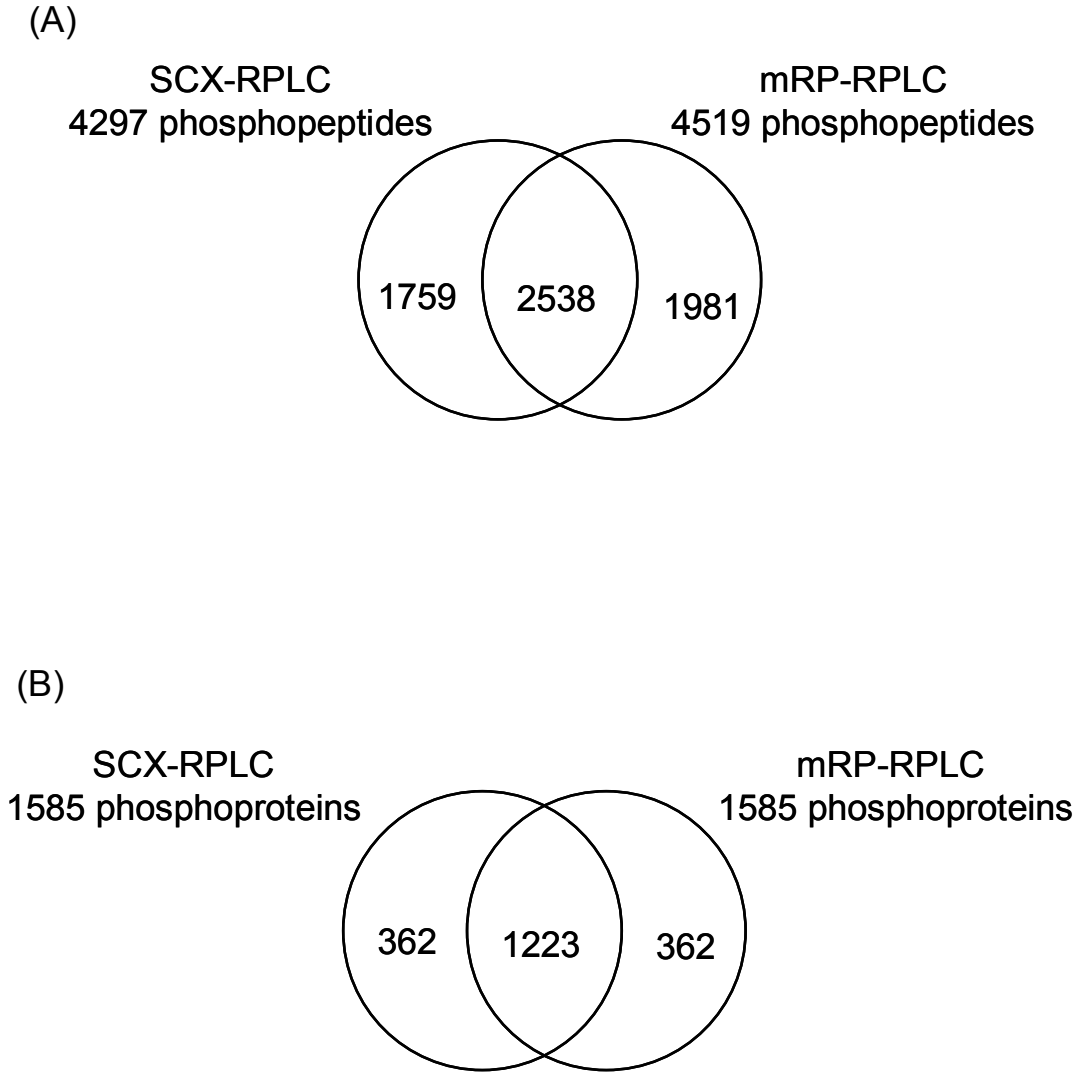


Figure 6.5 (A) Venn diagram of the phosphopeptides identified by the SCX-RPLC and mRP-RPLC methods. (B) Venn diagram of the phosphoproteins identified by the SCX-RPLC and mRP-RPLC methods. A total of 6278 different phosphopeptides and 1947 unique phosphoproteins were identified.

IMAC and TiO_2 were then individually subjected to subsequent sample fractionation and MS analysis. Table 6.1 summaries the properties of the identified phosphopeptides enriched by the IMAC and TiO_2 beads. As Table 6.1 shows, in all replicate experiments of the mRP-RPLC and SCX-RPLC methods, about half of the phosphopeptides identified from the IMAC enriched samples are mono-phosphorylated, and close to 100% of the phosphopeptides identified from the TiO_2 enriched samples are mono-phosphorylated. By using the IMAC enrichment, 276±19 peptides were found to be highly phosphorylated (triple or higher), and almost no multi-phosphorylated peptides were detected by using the TiO_2 enrichment. This result is consistent with the previous finding that IMAC is more prone to enrich multi-phosphorylated peptides.[36]

6.3.3 Characterization of the SCX-RPLC and mRP-RPLC methods

To better understand the differences of the phosphoproteins and phosphopeptides detected by the SCX-RPLC and mRP-RPLC methods, the possible correlation between the physiochemical property of phosphoproteins and phosphopeptides and the detectability of either method is examined.

The SCX technique fractionates phosphopeptides according to their charges. Therefore most phosphopeptides are eluted early in the salt gradient due to their reduced charge state introduced by the phosphate groups in the peptide sequences. Thus, SCX acts more like another enrichment method rather than a fractionation method. In contrast, the mRP-C18 column fractionates phosphoproteins according to their hydrophobic interactions with the stationary phase. Because of diverse variations of hydrophobicity of phosphoproteins, they will be separated and eluted out at the entire chromatographic separation time window. In other words, mRP-C18 does not have the enrichment effect, but provides a means of separating and fractionating the phosphoproteins. The distributions of the identified phosphopeptides from the SCX and mRP-C18 fractions shown in Figure 6.6 supports this statement. In the SCX-RPLC method, most phosphopeptides are detected in the front part of the SCX gradient, whereas phosphopeptides identified by the mRP-RPLC method are more evenly distributed throughout the mRP-C18 gradient. However, although the mRP-C18 column demonstrated better fractionation effect for the phosphoproteome, we did not identify more phosphoproteins and phosphopeptides by using the mRP-RPLC method, compared to the SCX-RPLC method. This result may be due to two reasons. First, as described in Chapter 5, insufficient denaturing and protein degradation may cause low resolution of protein separation. High abundant phosphoproteins and phosphopeptides were repeatedly detected due to ion suppression effect even though the

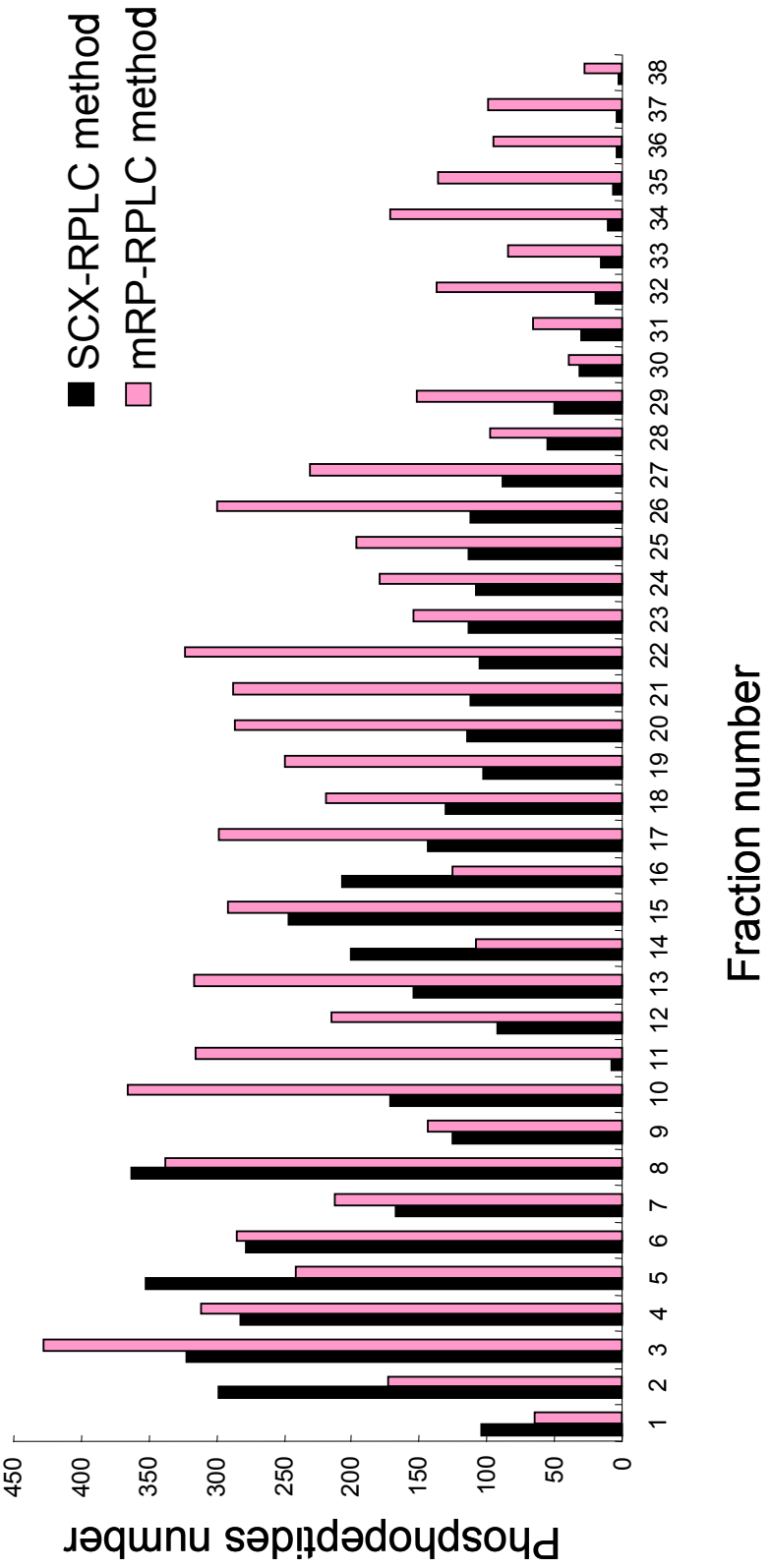


Figure 6.6 Distribution of phosphopeptides from SCX and mRPLC.

precursor ion exclusion (PIE) strategy was used. Second, an elevated temperature was used in the mRP-C18 fractionation to increase sample recovery. Although no evidence of severe de-phosphoryation was observed, it is possible that the phosphate groups might undergo hydrolysis reactions during the sample fractionation process. Therefore, our future work will be focused on the improvement of the mRP-C18 separation, e.g., investigating the stability of phosphoproteins during the mRP-C18 separation under an elevated temperature.

Figure 6.7A shows the phosphopeptide GRAVY distribution of these two methods. It is understandable that most phosphopeptides detected are hydrophilic or mildly hydrophobic due to the phosphate groups. And there is no significant difference found between these two methods. Figure 6.7B shows the molecular weight distribution of the identified phosphoproteins by using the two methods. Similar distributions are found for these two methods except the mRP-RPLC method appears to recover more phosphoproteins with MW of lower than 30 kDa.

6.3.3 Bioinformatic analysis

Using the Ingenuity Pathway Analysis (IPA) software, 1947 phosphoproteins detected by the mRP-RPLC and SCX-RPLC methods were examined to see if they could be grouped according to the canonical functional cellular pathways. As illustrated in Figure 6.8, the pathways to which the highest number of phosphoproteins belong are cellular growth and proliferation (n=507), gene expression (n=438), cellular assembly and organization (n=431), cellular development (n=366), cell cycle (n=325) and cell death or apoptosis (n=325).

As summarized in Table 6.2, 57 phosphoproteins identified by the SCX-RPLC and mRP-RPLC methods were previously reported to be functionally involved in breast cancer. The biological processes they involve in include cell proliferation, cell cycle regulation, cell apoptosis related, tumor genesis, tumor metastasis, and drug sensitivity. Six proteins were exclusively found to be phosphorylated by the mRP-RPLC method and they are BRCA1, CDK7, FGFR1, IRS1, SHC1 and TSC2. Six phosphoproteins were detected exclusively by the SCX-RPLC method: Ets1, p70S6K, LIMK, JNK1, JNK2 and MEK1. In addition, in our experimental dataset, we found 180 proteins that have never been reported to be phosphorylated. These proteins are listed in Table 6.3. Among these phosphorylated proteins, 127 phosphoproteins were detected by the mRP-RPLC method and 120 phosphoproteins were detected by the SCX method (Figure 6.9). This list can serve as the entry point for further biological validation.

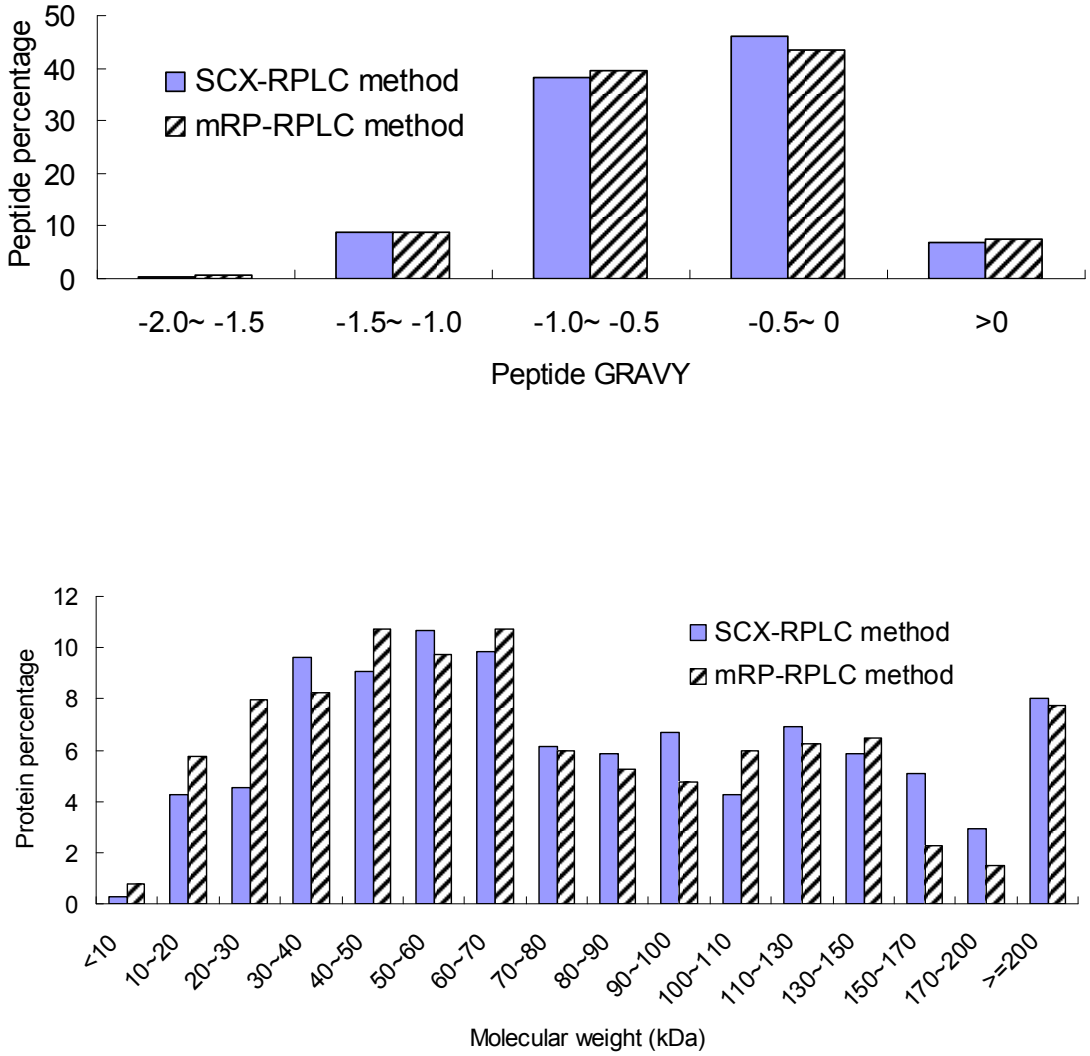


Figure 6.7 (A) Molecular weight distribution of unique protein of the SCX-RPLC and mRP-RPLC methods. (B) Peptide GRAVY distribution of the SCX-RPLC and mRP-RPLC methods.

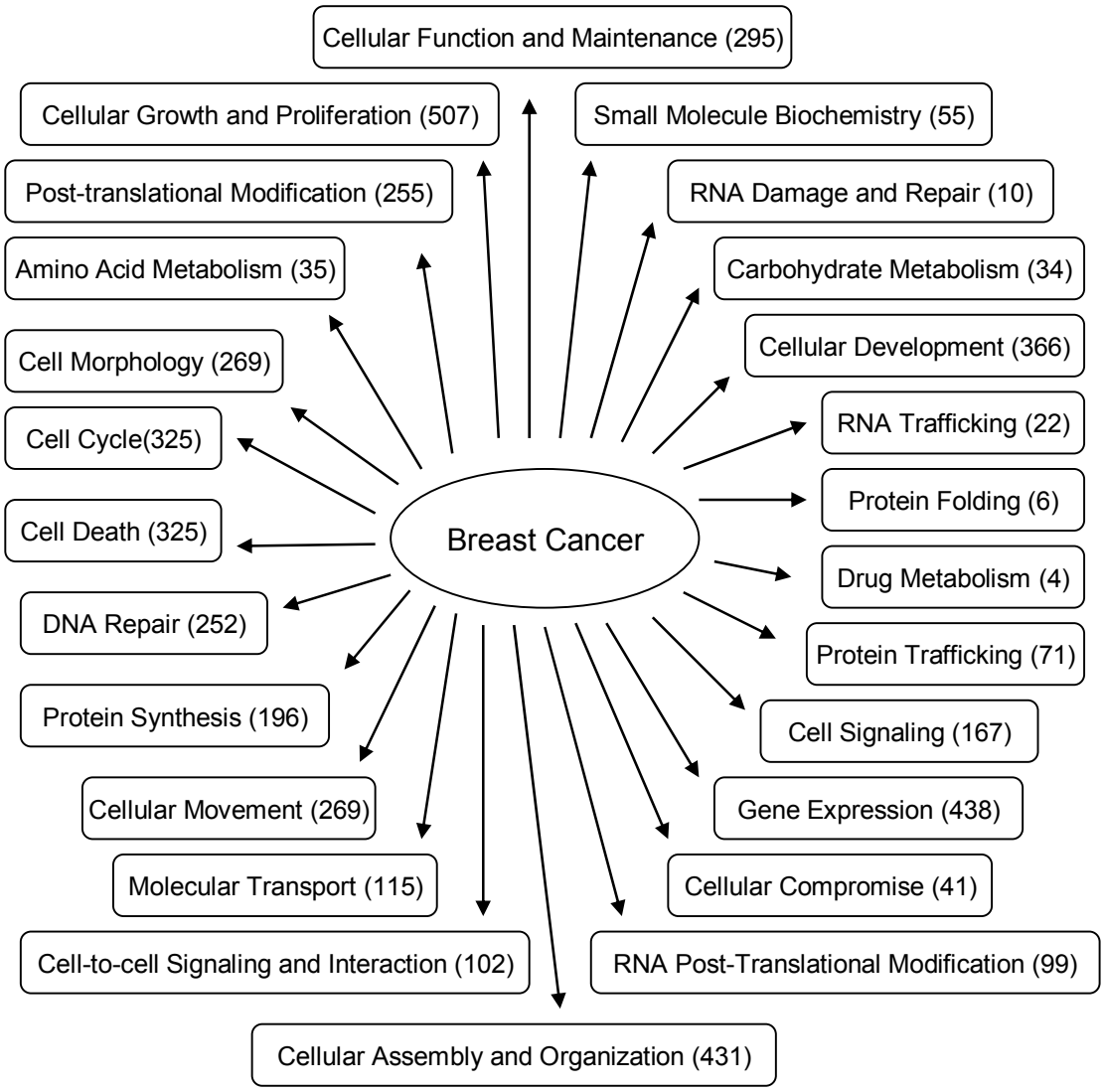


Figure 6.8 Functional pathway analysis of the MDA-MB-231 phosphoproteins.

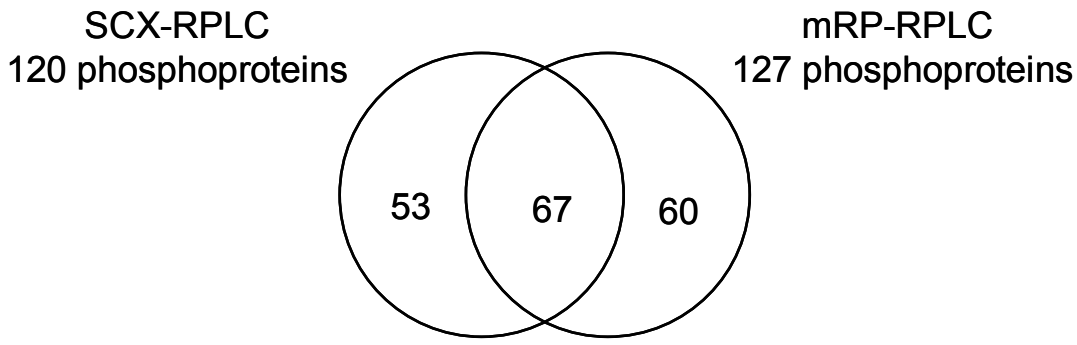


Figure 6.9 New phosphoproteins identified by the SCX-RPLC and mRP-RPLC method.

6.4 Conclusions

In this study, we have developed a new mRP-RPLC protocol for analyzing the phosphoproteome of the mDA-MB-231 cell line. In this method, protein sample is separated by the mRP-C18 column at the protein level first. Each protein fraction was then subjected to tryptic digestion, sequential phosphopeptide enrichment by IMAC and TiO_2, followed by capillary RPLC-MS/MS analysis. As a comparison, a traditional SCX-RPLC method was also used to evaluate the phosphoproteome detectability of the mRP-RPLC method. By using the same sample amount and number of LC-MS/MS runs with the same instrument setting, these two methods showed similar ability in the identification of phosphoproteins and phosphopeptides. There were 1585 unique phosphoproteins with 4519 different phosphopeptides identified by the mRP-RPLC method, and 1585? unique phosphoproteins with 4297 different phosphopeptides identified by the SCX-RPLC method. A total number of 1974 unique phosphoproteins and 6278 phosphopeptides were identified from these two methods. Since the protein and peptide overlaps between the results of the mRP-RPLC and SCX-RPLC methods were much lower than the overlaps between the replicate experiments of the same method, it can be concluded that the mRP-RPLC method provides complementary information to the conventional SCX-RPLC method, and thus, can increase the coverage of the MDA-MB-231 phosphoproteome. By using these two methods, we found 57 phosphoproteins that were previously reported to be functionally related to breast cancer and 180 new phosphoproteins which have never been reported to be phosphorylated. Future work will be focused on the improvement of the mRP-C18 separation and the validation of the newly identified phosphoproteins for investigating their biological functions in breast cancer development.

Table 6.1 Properties of phosphopeptides enriched by the IMAC and TiO_2 beads.

	Replicates	Enrichment methods	Mono-phosphorylated peptides	Multi-phosphorylated peptides	Percentage of mono-phosphorylated peptides	Triple- (or higher) phosphorylated peptides
mRP-RPLC method	Replicate one	IMAC	1829	1400	57%	272
		TiO_2	993	19	98%	0
	Replicate two	IMAC	1881	1362	58%	272
		TiO_2	928	17	98%	0
SCX-RPLC method	Replicate one	IMAC	1661	1391	54%	295
		TiO_2	884	35	96%	1
	Replicate two	IMAC	1821	1243	59%	266
		TiO_2	825	19	98%	0

Table 6.2 List of previously reported phosphoproteins found to be functionally involved in the breast cancer that were identified in this work.

#	Swiss-Prot ID	Protein Name	Peptides	Function	Signalling Pathway	Reference
1	P31749	AKT	SGSPSDNSGAEEMEVSLAKPK + 2 Phospho (ST) SGSPSDNSGAEEMEVSLAKPK + 3 Phospho (ST) SGSPSDNSGAEEMEVSLAKPK + Oxidation (M); 3 Phospho (ST)	cell survival and metabolism regulation	ErbB/HER and PI3K/AKT	41-44

			SGSPSDNSGAEEMEVSLAKPK + Oxidation (M); Phospho (ST) SGSPSDNSGAEEMEVSLAKPK + Phospho (ST)			
2	Q9Y478	AMPK	SHNNFVAILDLPEGEHQYK + Phospho (ST)	regulating cellular and organismal energy balance	AMPK	45
3	O43741	AMPK	DLSSSPPGPYGQEMYAFR + Oxidation (M); Phospho (ST) DLSSSPPGPYGQEMYAFR + Phospho (ST) IMVGSTDDPSVFSLPDSK + Oxidation (M); Phospho (ST) IPLIKSHNDFVAILDLPEGEHQYK + Phospho (ST) SHNDFVAILDLPEGEHQYK + Phospho (ST)	regulating cellular and organismal energy balance	AMPK	45
4	Q92934	BAD	HSSYPAGTEDDEGMGEEPSPFR + Oxidation (M); Phospho (ST) HSSYPAGTEDDEGMGEEPSPFR + Oxidation (M); Phospho (Y) RMSDEFVDSFK + Oxidation (M); Phospho (ST) RMSDEFVDSFK + Phospho (ST) RMSDEFVDSFKK + Oxidation (M); Phospho (ST) RMSDEFVDSFKK + Phospho (ST) SAPPNLWAAQR + Phospho (ST) SRHSSYPAGTEDDEGMGEEPSPFR + Oxidation (M); Phospho (ST) SRSAPPNLWAAQR + Phospho (ST)	Apoptosis	ErbB/HER and PI3K/AKT	46
5	Q9HB09	BCL-2	ACPGPPPPSPEPLAR + Phospho (ST) LSSDSFAR + Phospho (ST) RGEAAGSPVPTPPR + 2 Phospho (ST) RGEAAGSPVPTPPRSPAQEEPTDFLSR + 2 Phospho (ST)	Apoptosis	Apoptosis and PI3K/AKT	47, 48
6	P35222	beta-ca	RTSMGGTQQQFVEGVR + Oxidation (M); Phospho (ST)	Tumorigenesis	Wnt/ß-Cateni	49, 50

		tenin	RTSMGGTQQQFVEGVR + Phospho (ST) TSMGGTQQQFVEGVR + Oxidation (M); Phospho (ST) TSMGGTQQQFVEGVR + Phospho (ST)		n	
7	P16070	CD44	KPSGLNGEASKSQEMVHLVNK + 2 Phospho (ST) KPSGLNGEASKSQEMVHLVNK + Oxidation (M); Phospho (ST) KPSGLNGEASKSQEMVHLVNK + Phospho (ST) SQEMVHLVNK + Phospho (ST)	tumor growth and progression	PI3K/AKT	51
8	P50613	CDK7	AYTHQVVTR + Phospho (ST)	cell cycle progression (regulation)	Cell Cycle: G2/M DNA Damage Checkpoint	52
9	P45983	JNK1	TAGTSFMMTPYVVTR + Phospho (ST); Phospho (Y)	Apoptosis	ErbB/HER Signaling and SAPK/JNK	53
10	P45984	JNK2	TACTNFMMTPYVVTR + Phospho (ST); Phospho (Y)	Apoptosis	ErbB/HER	53
11	P23528	cofilin	LGGSAVISLEGKPL + Phospho (ST)	actin depolymerization	TGF-ß	54
12	P04049	c-Raf	MPVSSQHRYSTPHAFTFNTSSPSSEGSLSQR + Phospho (ST) SHSESASPSALSSSPNNLSPTGWSQPK + 2 Phospho (ST) STSTPNVHMVSTTLPVDSR + Oxidation (M); Phospho (ST) STSTPNVHMVSTTLPVDSR + Phospho (ST)	Apoptosis and transcription factor	ErbB/HER	55, 56
13	P15336	CREB	KMPLDLSPLATPIIR + Oxidation (M); Phospho (ST) MPLDLSPLATPIIR + Oxidation (M); Phospho (ST) MPLDLSPLATPIIR + Phospho (ST) NDSVIVADQTPTPTR + 2 Phospho (ST)	transcription factor	B Cell Receptor	57-59

			NDSVIVADQTPTPTR + Phospho (ST)			
14	O00418	eEF-2	AGHDGDSDGDSDDEEGYFICPITDDPSSNQNVNSK + 2 Phosph o (ST) KYESDEDSLGSSGR + 2 Phospho (ST) KYESDEDSLGSSGR + Phospho (ST) YSSSGSPANSFHFK + 2 Phospho (ST)	linking cellular energy status and the inhibition of protein synthesis	AMPK	60
15	P00533	EGFR	ALMDEEDMDDVVDADEYLIPQQGFFSSPSTSR + 2 Oxidation (M); Phospho (ST) ALMDEEDMDDVVDADEYLIPQQGFFSSPSTSR + Oxidation (M); Phospho (ST) ALMDEEDMDDVVDADEYLIPQQGFFSSPSTSR + P hospho (ST) GSHQISLDNPDYQQDFFPK + Phospho (ST) MHLPSPTDSNFYR + 2 Phospho (ST) MHLPSPTDSNFYR + Oxidation (M); 2 Phospho (ST) TPLLSSLSATSNNSTVACIDR + 2 Phospho (ST) TPLLSSLSATSNNSTVACIDR + 3 Phospho (ST)	apoptosis	ErbB/HER	61
16	Q13541	eIF4E-binding protein 1	DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDK + Phospho (ST) DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR + Oxidation (M); Phospho (ST) DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR+ Phospho (ST) FLMECRNSPVTK + Oxidation (M); Phospho (ST) FLMECRNSPVTK + Phospho (ST) FLMECRNSPVTKTPPR + 2 Phospho (ST) FLMECRNSPVTKTPPR + Oxidation (M); 2 Pho spho (ST) NSPVTKTPPR + 2 Phospho (ST)	cell cycle regulation	Insulin Receptor	32, 62

			NSPVTKTPPRDLPTIPGVTSPSSDEPPMEASQSHLR + 2 Phospho (ST) NSPVTKTPPRDLPTIPGVTSPSSDEPPMEASQSHLR + 3 Phospho (ST) NSPVTKTPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Oxida tion (M); 2 Phospho (ST) NSPVTKTPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST) RVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y) TPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST) TPPRDLPTIPGVTSPSSDEPPMEASQSHLRNSPEDK + 2 Phospho (ST) TPPRDLPTIPGVTSPSSDEPPMEASQSHLRNSPEDK + Oxidati on (M); 2 Phospho (ST) VVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y)			
17	P27361	ERK1	IADPEHDHTGFLTEYVATR + Phospho (ST) IADPEHDHTGFLTEYVATR + Phospho (ST); Phospho (Y) IADPEHDHTGFLTEYVATR + Phospho (Y)	regulation of cell growth and differentiation	ErbB/HER	63, 64
18	P28482	ERK2	VADPDHDHTGFLTEYVATR + Phospho (ST) VADPDHDHTGFLTEYVATR + Phospho (ST); Phospho (Y) VADPDHDHTGFLTEYVATR + Phospho (Y)	regulation of cell growth and differentiation	ErbB/HER	65
19	P14921	Ets1	VPSYDSFDSEDYPAALPNHKPK + Phospho (ST) VPSYDSFDSEDYPAALPNHKPK + Phospho (ST); Phospho (Y)	tumor metastasis	B Cell Receptor	66, 67

20	Q05397	FAK	GSIDREDGSLQGPIGNQHIYQPVGKPDPAAPPK + Phospho (ST) LQPQEISPPPTANLDR + Phospho (ST) LSRGSIDREDGSLQGPIGNQHIYQPVGKPDPAAPPK + 2 Phospho (ST) RQATVSWDSGGSDEAPPKPSRPGYPSPR + Phospho (ST)	apoptosis	PI3K/Akt	67, 68
21	P11362	FGFR1	LAKSIPLR + Phospho (ST)	positive regulation of cell proliferation	ESC Pluripotency and Differentiation	69
22	P21333	FLNa	AFGPGLQGGSAGSPAR + Phospho (ST) IPEISIQDMTAQVTSPSGK + Oxidation (M); Phospho (ST) RAPSVANVGSHCDLSLK + Phospho (ST) RRAPSVANVGSHCDLSLK + Phospho (ST)	embryonic cell migration	MAPK/Erk in Growth and Differentiation	14
23	Q13480	Gab1	SYSHDVLPK + Phospho (ST) VKPAPLEIKPLPEWEELQAPVRSPITR + Phospho (ST)	cell proliferation	PI3K/Akt	70
24	P49840	GSK3	GEPNVSYICSR + Phospho (Y)	apoptosis	PI3K/Akt	16
25	P04626	HER2	SGGGDLTLGLEPSEEEAPR + Phospho (ST)	negative regulation of apoptosis	Wnt signaling pathway and PI3K/Akt	71, 72
26	Q00613	HSF1	GHTDTEGRPPSPPPTSTPEK + 2 Phospho (ST) GHTDTEGRPPSPPPTSTPEK + Phospho (ST) VEEASPGRPSSVDTLLSPTALIDSILR + Phospho (ST) VKEEPPSPPQSPR + 2 Phospho (ST) VKEEPPSPPQSPRVEEASPGRPSSVDTLLSPTALIDSILR + 2	negative regulation of cell proliferation	SAPK/JNK Signaling Cascades	73

			Phospho (ST) VKEEPPSPPQSPRVEEASPGRPSSVDTLLSPTALIDSILR + 3 Phospho (ST)			
27	P04792	Hsp27	GPSWDPFR + Phospho (ST) GPSWDPFRDWYPHSR + Phospho (ST)	apoptosis	Signaling Pathways Activating p38 MAPK	74, 75
28	Q9Y4H2	IRS2	ASSPAESSPEDSGYMR + 3 Phospho (ST) HNSASVENVSLR + Phospho (ST) SNTPESIAETPPAR + 2 Phospho (ST) SPGEYINIDFGEPGAR + Phospho (ST)	cell proliferation	Insulin Receptor	76
29	P51812	p90RSK	TPKDSPGIPPSANAHQLFR + 2 Phospho (ST)	apoptosis	Mitochondrial Control of Apoptosis	59
30	P35568	IRS1	HSSETFSSTPSATR + Phospho (ST) SVTPDSLGHTPPAR + 2 Phospho (ST)	positive regulation of cell proliferation	PI3K/Akt	77
31	P53667	LIMK	SPGAGSLGSPASQR + Phospho (ST)	tumor metastasis	TGF-ß	78
32	Q16539	p38MAPK	HTDDEMTGYVATR + Oxidation (M); Phospho (ST); Phospho (Y) HTDDEMTGYVATR + Oxidation (M); Phospho (Y) HTDDEMTGYVATR + Phospho (ST) HTDDEMTGYVATR + Phospho (ST); Phospho (Y) HTDDEMTGYVATR + Phospho (Y)	cell cycle regulation and drug sensitivity (chemotaxis)	NF-kB	79
33	P29966	MARCKS	AEDGATPSPSNETPK + 2 Phospho (ST) AEDGATPSPSNETPKK + 2 Phospho (ST)	tumor metastasis	FAK/PKC/MARCKS	80

			EAPAEGEAAEPGSPTAAEGEAASAASSTSSPK + 2 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPK + 3 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PK + 2 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PK + 4 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PK + 5 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PK + Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PKK + 2 Phospho (ST) EAPAEGEAAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNET PKK + 3 Phospho (ST) FSFKKSFK + 2 Phospho (ST) GEPAAAAAPEAGASPVEKEAPAEGEAAEPGSPTAAEGEAASAASS TSSPK + Phospho (ST) KRFSFK + Phospho (ST) KSFKLSGFSFK + 2 Phospho (ST) RFSFKK + Phospho (ST) SFKLSGFSFKK + 2 Phospho (ST) SFKLSGFSFKK + 3 Phospho (ST)			
34	Q02750	MEK1	TPGRPLSSYGMDSRPPMAIFELLDYIVNEPPPK + 2 Oxidation (M); Phospho (ST)	cell profliferation	ErbB/HER	81
35	P08581	Met	SVSPTTEMVSNESVDYR + 2 Phospho (ST)	cell proliferation	PI3K/Akt	82

			SVSPTTEMVSNESVDYR + 3 Phospho (ST) SVSPTTEMVSNESVDYR + Oxidation (M); 3 Phospho (ST) SVSPTTEMVSNESVDYR + Oxidation (M); Phospho (ST) VHTPHLDR + Phospho (ST)			
36	P56945	p130Cas	SSQSASSLEVAGPGR + 2 Phospho (ST)	cell migration and apoptosis	AND-34/BCAR3/NSP2	83
37	P23443	p70S6K	TPVSPVKFSPGDFWGR + 2 Phospho (ST)	drug sensitivity	PI3K/Akt	84
38	Q15418	S6 kinase (p90RSK)	GFSFVATGLMEDDGKPR + Phospho (ST) KAYSFCGTVEYMAPEVVNR + Phospho (ST) TPKDSPGIPPSAGAHQLFR + 2 Phospho (ST) TPKDSPGIPPSAGAHQLFR + Phospho (ST)	signal transduction	mTOR	32
39	P49023	paxillin	ELDELMASLSDFK + Oxidation (M); Phospho (ST) GLEDVRPSVESLLDELESSVPSPVPAITVNQGEMSSPQR + Oxidation (M); Phospho (ST) LLLELNAVQHNPPGFPADEANSSPPLPGALSPLYGVPETNSPLGGK + 2 Phospho (ST) QKSAEPSPTVMSTSLGSNLSELDR + Oxidation (M); 2 Phospho (ST) QKSAEPSPTVMSTSLGSNLSELDR + Oxidation (M); 3 Phospho (ST) SAEPSPTVMSTSLGSNLSELDR + 2 Phospho (ST) SAEPSPTVMSTSLGSNLSELDR + Oxidation (M); 2 Phospho (ST) SAEPSPTVMSTSLGSNLSELDR + Oxidation (M); Phospho (ST) SAEPSPTVMSTSLGSNLSELDR + Phospho (ST) SSPGGQDEGGFMAQGK + Phospho (ST) TGSSSPPGGPPKPGSQLDSMLGSLQSDLNK + Oxidation (M);	cell adhesion	PI3K/Akt	85

			Phospho (ST) TGSSSPPGGPPKPGSQLDSMLGSLQSDLNK + Phospho (ST) TSSVSNPQDSVGSPCSR + Phospho (ST) VGEEEHVYSFPNK + Phospho (Y)			
40	P06400	Rb	DREGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + 2 Phospho (ST) DREGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + Oxidation (M); 2 Phospho (ST) DREGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + Oxidation (M); Phospho (ST); Phospho (Y) DREGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + Phospho (ST); Phospho (Y) EGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + 2 Phospho (ST) EGPTDHLESACPLNLPLQNNHTAADMYLSPVRSPK + Oxidation (M); 2 Phospho (ST) FPSSPLR + Phospho (ST) IPGGNIYISPLKSPYK + 2 Phospho (ST) IPGGNIYISPLKSPYK + Phospho (ST) ISEGLPTPTKMTPR + 2 Phospho (ST) ISEGLPTPTKMTPR + Oxidation (M); 2 Phospho (ST) ISEGLPTPTKMTPR + Oxidation (M); Phospho (ST) ISEGLPTPTKMTPR + Phospho (ST) KSNLDEEVNVIPPHTPVR + Phospho (ST) KTAATAAAAAAEPPAPPPPPPPEEDPEQDSGPEDLPLVR +	cell cycle progression (regulation)	Cell Cycle: G1/S Checkpoint	86

			Phospho (ST) LFLDHDKTLQTDSIDSFETQRTPR + Phospho (ST) SPYKFPSSPLR + 2 Phospho (ST) SPYKFPSSPLR + Phospho (ST) TAATAAAAAAEPPAPPPPPPPEEDPEQDSGPEDLPLVR + Phospho (ST) TLQTDSIDSFETQRTPR + Phospho (ST) TNILQYASTRPPTLSPIPHIPRSPYK + Phospho (ST); Phospho (Y) TNILQYASTRPPTLSPIPHIPRSPYKFPSSPLR + 2 Phospho (ST); Phospho (Y) TNILQYASTRPPTLSPIPHIPRSPYKFPSSPLR + 3 Phospho (ST)			
41	P04049	Raf	RDSSDDWEIPDGQITVGQR + Phospho (ST) SASEPSLNR + Phospho (ST) SSSAPNVHINTIEPVNIDDLIR + Phospho (ST)	cell apoptosis related	ErbB/HER	87
42	Q15311	RBP1	RTEGYAAFQEDSSGDEAESPSK + 2 Phospho (ST) TEGYAAFQEDSSGDEAESPSK + 2 Phospho (ST) TGEPSPPHDILHEPPDVVSDDEKDHGK + 2 Phospho (ST) TGEPSPPHDILHEPPDVVSDDEKDHGK + Phospho (ST) TPSSEEISPTK + Phospho (ST) TPSSEEISPTKFPGLYR + 2 Phospho (ST) TPSSEEISPTKFPGLYR + Phospho (ST)	drug sensitivity	cyclin/CDK	88
43	P62753	ribosomal protein S6 (S6)	RLSSLR + 2 Phospho (ST) RLSSLRASTSK + 2 Phospho (ST) RLSSLRASTSK + 3 Phospho (ST)	cell growth and proliferation	PI3K/Akt	89

44	P29353	Shc1	ELFDDPSYVNVQNLDK + Phospho (ST)	positive regulation of cell proliferation	ErbB/HER	90
45	P12931	Src	RRSLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST) SLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST)	membrane organization and biogenesis	ErbB/HER	91
46	P40763	STAT3	FICVTPTTCSNTIDLPMSPR + Oxidation (M); Phospho (ST) FICVTPTTCSNTIDLPMSPR + Phospho (ST) YCRPESQEHPEADPGSAAPYLK + Phospho (Y)	drug sensitivity	ErbB/HER	92
47	P67809	YB1	NEGSESAPEGQAQQR + 2 Phospho (ST) NEGSESAPEGQAQQR + Phospho (ST) NYQQNYQNSESGEK + Phospho (ST) NYQQNYQNSESGEKNEGSESAPEGQAQQR + 2 Phospho (ST) NYQQNYQNSESGEKNEGSESAPEGQAQQR + Phospho (ST)	drug sensitivity	PI3K/Akt	91
48	P38398	BRCA1	GELSRSPSPFTHTHLAQGYR + 3 Phospho (ST)	cell apoptosis related	Cell Cycle: G2/M DNA Damage Checkpoint	93, 94
49	P11274	BCR	ASASRPQPAPADGADPPPAEEPEARPDGEGSPGK + Phospho (ST) HQDGLPYIDDSPSSSPHLSSK + 2 Phospho (ST) HQDGLPYIDDSPSSSPHLSSK + Phospho (ST) RHQDGLPYIDDSPSSSPHLSSK + 2 Phospho (ST) RHQDGLPYIDDSPSSSPHLSSK + Phospho (ST)	tumor genesis	PI3K/Akt	95
50	P30305	CDC25	SKSLCHDEIENLLDSDHR + Phospho (ST)	cell proliferation	MAPK/Erk in	96

					Growth and Differentiation	
51	Q12929	EPS8	RLSTEHSSVSEYHPADGYAFSSNIYTR + Phospho (ST)	tumor metastasis	EGFR	97
52	Q13547	HDAC1	IACEEEFSDSEEEGEGGR + 2 Phospho (ST) IACEEEFSDSEEEGEGGRK + 2 Phospho (ST) MLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKR + 2 Oxidation (M); Phospho (ST) MLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKR + Oxidation (M); Phospho (ST) MLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKR + Phospho (ST)	cell apoptosis related	Notch Signaling	98
53	P17096	HMGA1	EEEEGISQESSEEEQ + 3 Phospho (ST) KLEKEEEEGISQESSEEEQ + 3 Phospho (ST) KQPPVSPGTALVGSQK + Phospho (ST) KQPPVSPGTALVGSQKEPSEVPTPK LEKEEEEGISQESSEEEQ + 3 Phospho (ST) QPPVSPGTALVGSQKEPSEVPTPK + Phospho (ST)	cell apoptosis related	Ras/ERK	99
54	O60934	NBS1	IPNYQLSPTKLPSINK + Phospho (ST)	cell proliferation	Cell Cycle: G2/M DNA Damage Checkpoint	100
55	P55327	TPD52	KLEDVKNSPTFK + Phospho (ST) NSPTFKSFEEK + Phospho (ST) NSPTFKSFEEKVENLK + Phospho (ST)	tumor genesis	Cell Cycle: G2/M DNA Damage Checkpoint	101

56	P49815	TSC2	SSSSPELQTLQDILGDPGDK + Phospho (ST)	cell proliferation	HER-2	102
57	Q9Y6Q9	SRC-3	AVSLDSPVSVGSSPPVK + Phospho (ST)	cell apoptosis related	Notch	103

Table 6.3 List of new phosphoproteins found in this work.

Name	SwissProt ID	Unique Peptide Sequence	Score	Mascot score for identity
ZNHI6_HUMAN	Q9NWK9	SGGGLHSVAEGVRLSPEPGR + Phospho (ST)	56	31
ZNF35_HUMAN	P13682	KSGGKYSLNSGAVK + Phospho (ST)	35	30
ZN593_HUMAN	O00488	RLAVPTEVSTEVPEMDTST + Phospho (ST)	65	32
		RLAVPTEVSTEVPEMDTST + Oxidation (M); Phospho (ST)	40	32
		LAVPTEVSTEVPEMDTST + Oxidation (M); Phospho (ST)	36	31
ZN511_HUMAN	Q8NB15	SPASAEAPGDSGERSEGEAMEICSEPVAASPAPAGER + Oxidation (M); Phospho (ST)	33	33
ZFAN3_HUMAN	Q9H8U3	SCGTDSQSENEASPVKRPR + 3 Phospho (ST)	54	29
		SCGTDSQSENEASPVKRPR + 2 Phospho (ST)	59	30
		SCGTDSQSENEASPVK + 3 Phospho (ST)	28	23
YB041_HUMAN	A6ND21	NIEEVTSLPK + Phospho (ST)	36	29
WWC1_HUMAN	Q8IX03	VDKETNTETPAPSPTVVRPK + 2 Phospho (ST)	70	32
		TDRGSHSDLWSSSSSLESSSFPLPK + Phospho (ST)	56	32
		SKTFSPGPQSQYVCR + Phospho (ST)	33	31
		LGASEAAAFDSDESEAVGATR + Phospho (ST)	61	31
		ETNTETPAPSPTVVRPK + 2 Phospho (ST)	37	31

WD42A_HUMAN	Q5TAQ9	VHDRSEEEEEEEEEEEEQPRR + Phospho (ST)	88	31
		VHDRSEEEEEEEEEEEEQPR + Phospho (ST)	144	30
		SEEEEEEEEEEEEQPR + Phospho (ST)	45	27
		DQDSSDDERALEDWVSSETSALPRPR + 2 Phospho (ST)	66	32
VPS4A_HUMAN	Q9UN37	ENQSEGKGSDSDSEGDNPEK + 2 Phospho (ST)	37	26
UBE3C_HUMAN	Q15386	VLQTFLSQLPVSPASASCHDSASDSEEESEEADKPSSPEDGR + 3 Phospho (ST)	47	32
		VLQTFLSQLPVSPASASCHDSASDSEEESEEADKPSSPEDGR + 2 Phospho (ST)	54	32
TXND1_HUMAN	Q9H3N1	VEEEQEADEEDVSEEEAESKEGTNK + Phospho (ST)	75	31
		SLGPSLATDKS + Phospho (ST)	31	29
TRI11_HUMAN	Q96F44	LHPPSPVPQGVCPAHREPLAAFCGDELR + Phospho (ST)	55	33
TMSL6_HUMAN	A9Z1Y9	TETQEKNPLPSK + Phospho (ST)	34	30
		SDKSDMAEIEKFDK + Oxidation (M); Phospho (ST)	38	31
TISD_HUMAN	P47974	RLSDSPVFDAPPSPPDSLSDR + Phospho (ST)	32	32
		RHSASNLHALAHPAPSPGSCSPK + Phospho (ST)	56	32
		RHSASNLHALAHPAPSPGSCSPK + 2 Phospho (ST)	36	32
TDG_HUMAN	Q13569	KPRTTEPK + Phospho (ST)	31	28
TCAB1_HUMAN	Q9BUR4	VFPEPTESGDEGEELGLPLLSTR + Phospho (ST)	126	32
		TLETQPLAPDCCPSDQDPAPAHPSPHASPMNK + Oxidation (M); 2 Phospho (ST)	63	32
		IEEQELSENTSLPAEEANGSLSEEEANGPELGSGK + 2 Phospho (ST)	38	33
		GDPPRLSPDPVAGSAVSQELR + Phospho (ST)	51	31
		EGDPVSLSTPLETEFGSPSELSPR + Phospho (ST)	65	32
		EGDPVSLSTPLETEFGSPSELSPR + 2 Phospho (ST)	43	32
TBC24_HUMAN	Q9ULP9	HPELTKPPPLMAAEPTAPLSHSASSDPADRLSPFLAAR + Phospho (ST)	71	31
		HPELTKPPPLMAAEPTAPLSHSASSDPADRLSPFLAAR + Oxidation (M); Phospho (ST)	40	31

SYNG_HUMAN	Q9UMZ2	SLSLGDKEISR + Phospho (ST)	37	30
		SGSLDDSFSDFQELPASSK + Phospho (ST)	33	31
		ETSFGSSENITMTSLSK + Phospho (ST)	42	31
		ETSFGSSENITMTSLSK + Oxidation (M); Phospho (ST)	59	32
STUB1_HUMAN	Q9UNE7	LGAGGGSPEKSPSAQELK + 2 Phospho (ST)	71	31
SPEC1_HUMAN	Q5M775	SPLSGIPVR + Phospho (ST)	33	29
		KSPSLESLSRPPSLGFGDTR + Phospho (ST)	45	31
		ELSDLEEENR + Phospho (ST)	36	28
SOX9_HUMAN	P48436	ALQADSPHSSSGMSEVHSPGEHSGQSQGPPTPPTTPK + Oxidation (M); 2 Phospho (ST)	49	32
SOLO_HUMAN	Q8TER5	RIQQHLGEEASPR + Phospho (ST)	47	31
SMG1L_HUMAN	Q6P435	LSNLLR + Phospho (ST)	31	28
SLFN5_HUMAN	Q08AF3	FSGLER + Phospho (ST)	36	26
SLD5_HUMAN	Q9BRT9	TEEVDFLGQDSDGGSEEVVLTPAELIER + 2 Phospho (ST)	36	33
SKP1_HUMAN	P63208	GKTPEEIRK + Phospho (ST)	40	29
SK2L2_HUMAN	P42285	MSPTIGKQLLK + Oxidation (M); Phospho (ST)	34	29
SGEF_HUMAN	Q96DR7	TVHRSPLLLGAQR + Phospho (ST)	45	29
		ALDIDSDEESEPK + Phospho (ST)	36	29
SFRS9_HUMAN	Q13242	GSPHYFSPFRPY + Phospho (ST)	74	30
SFRS7_HUMAN	Q16629	YFQSPSR + Phospho (ST)	30	27
		YFQSPSR + 2 Phospho (ST)	40	24
		SRSISLR + 2 Phospho (ST)	47	27
		SISLRR + 2 Phospho (ST)	43	27
		SISLR + 2 Phospho (ST)	28	24
		RSRSISLR + 3 Phospho (ST)	35	26

SFRS6_HUMAN	Q13247	SVSPPPKR + 2 Phospho (ST)	36	28
		SNSPLPVPPSK + 2 Phospho (ST)	44	28
		ARSVSPPPKR + 2 Phospho (ST)	31	29
SFRS2_HUMAN	Q01130	VDNLTYRTSPDTLR + Phospho (ST)	50	31
		SRSPPPVSK + 2 Phospho (ST)	40	28
		SPPKSPEEEGAVSS + 3 Phospho (ST)	35	25
		SKSPPKSPEEEGAVSS + 4 Phospho (ST)	46	25
		SKSPPKSPEEEGAVSS + 2 Phospho (ST)	51	30
SFRS1_HUMAN	Q07955	VKVDGPRSPSYGR + Phospho (ST); Phospho (Y)	32	30
		VKVDGPRSPSYGR + Phospho (ST)	66	30
		VKVDGPRSPSYGR + 2 Phospho (ST)	40	30
		VDGPRSPSYGRSR + 3 Phospho (ST)	38	28
		VDGPRSPSYGRSR + 2 Phospho (ST); Phospho (Y)	33	27
		VDGPRSPSYGRSR + 2 Phospho (ST)	54	29
		VDGPRSPSYGR + Phospho (ST); Phospho (Y)	39	28
		VDGPRSPSYGR + Phospho (ST)	43	29
		VDGPRSPSYGR + 2 Phospho (ST)	43	27
SFRIP_HUMAN	Q99590	TEELIESPKLESSEGEIIQTVDR + Phospho (ST)	41	32
		FHSPSTTWSPNKDTPQEK + 2 Phospho (ST)	38	31
SFR2B_HUMAN	Q9BRL6	SKRPPKSPEEEGQMSS + 2 Phospho (ST)	47	30
		RPPKSPEEEGQMSS + 2 Phospho (ST)	42	28
SFR11_HUMAN	Q05519	TKECSVEKGTGDSLR + Phospho (ST)	78	31
		LNHVAAGLVSPSLK + Phospho (ST)	48	28
		KPIETGSPK + Phospho (ST)	48	29
		DYDEEEQGYDSEKEKK + Phospho (ST)	113	30

		DYDEEEQGYDSEKEK + Phospho (ST)	113	28
SF04_HUMAN	Q8IWZ8	DVDASPSPLSVQDLK + Phospho (ST)	44	31
		AVQQHQHGYDSDEEVDSELGTWEHQLR + Phospho (ST)	96	33
SDCB2_HUMAN	Q9H190	GKIVSLVKGSSAAR + 2 Phospho (ST)	30	30
SDC10_HUMAN	Q6UX04	QQSKKGTSR + Phospho (ST)	31	29
SCHI1_HUMAN	Q9P0W5	DDRSPAREPGDVSAR + Phospho (ST)	57	30
SAPS3_HUMAN	Q5H9R7	NTVDLVTTCHIHSSSDDEIDFK + 3 Phospho (ST)	53	30
		NTVDLVTTCHIHSSSDDEIDFK + 2 Phospho (ST)	67	31
		FADQDDIGNVSFDR + Phospho (ST)	67	29
SAPS1_HUMAN	Q9UPN7	SGGSTDSEDEEEEDEEEEEDEEGIGCAAR + 3 Phospho (ST)	79	23
		NMVDLVNTHHLHSSSDDEDDRLK + Oxidation (M); Phospho (ST)	36	32
		NMVDLVNTHHLHSSSDDEDDRLK + 3 Phospho (ST)	40	31
		IQQFDDDEEEEDEEEAQGSGESDGEDGAWQGSQLAR + 2 Phospho (ST)	44	29
SAM4A_HUMAN	Q9UPU9	SDSVDYGQTHYYHQR + Phospho (ST)	120	30
		GRSDSVDYGQTHYYHQR + Phospho (ST)	60	31
S41A3_HUMAN	Q96GZ6	RLDSCGKPGELGLPHPLSTGGLPVASEDGALR + Phospho (ST)	59	31
RWD2B_HUMAN	P57060	KFSIFEEK + Phospho (ST)	41	29
RUSD1_HUMAN	Q9UJJ7	GPRPGSPSALLPGPGRPPPPPTKPPETEAQR + Phospho (ST)	42	29
RS16_HUMAN	P62249	TATAVAHCKR + Phospho (ST)	30	29
RNF5_HUMAN	Q99942	LKTPPRPQGQRPAPESR + Phospho (ST)	34	29
RN185_HUMAN	Q96GF1	EKTPPRPQGQRPEPENR + Phospho (ST)	58	31
RLBL1_HUMAN	Q8IUQ0	LMKRSQSVVEAGTLK + 2 Phospho (ST)	34	31
RICS_HUMAN	A7KAX9	SAKSEESLTSLHAVDGDSK + Phospho (ST)	68	32
RGPD4_HUMAN	Q7Z3J3	SALSPSKSPAK + 2 Phospho (ST)	31	30
RGPD3_HUMAN	A6NKT7	EDALDDSVSSSSVHASPLASSPVRK + Phospho (ST)	33	32

RET3_HUMAN	P10745	FDAMAELETVK + Phospho (ST)	31	29
RAMA1_HUMAN	Q8IX90	YGYSPR + Phospho (ST)	35	26
		SPRSPQLSDFGLER + Phospho (ST)	38	31
		SPQLSDFGLER + Phospho (ST)	57	30
RABE2_HUMAN	Q9H5N1	HAPSLHGSTELLPLSR + Phospho (ST)	76	31
RA1L3_HUMAN	P0C7M2	SKSESPKEPEQLR + 2 Phospho (ST)	42	30
		SESPKEPEQLR + 2 Phospho (ST)	38	29
R3GEF_HUMAN	Q8TBN0	HKSTSSTLCPAVCPAAGHTLTPDR + Phospho (ST)	54	32
QSK_HUMAN	Q9Y2K2	TWCGSPPYAAPELFEGK + Phospho (ST)	57	31
		RFSDGAASIQAFK + Phospho (ST)	34	31
		RASDGGANIQLHAQQLLK + Phospho (ST)	46	31
		IQPSSPPPNHPNNHLFR + Phospho (ST)	38	31
		FSDGAASIQAFK + Phospho (ST)	48	29
PTOV1_HUMAN	Q86YD1	SRSWPASPR + Phospho (ST)	48	29
PTAD1_HUMAN	Q9P035	WLDESDAEMELR + Oxidation (M); Phospho (ST)	70	29
PSG1_HUMAN	P11464	SDPVTLNLLPK + Phospho (ST)	37	29
PPIB_HUMAN	P23284	TDSRDKPLKDVIIADCGK + Phospho (ST)	50	31
PLEC1_HUMAN	Q15149	SSSVGSSSSYPISPAVSR + Phospho (ST); Phospho (Y)	75	31
		SSSVGSSSSYPISPAVSR + Phospho (ST)	93	31
		SSSVGSSSSYPISPAVSR + 2 Phospho (ST); Phospho (Y)	50	29
		SSSVGSSSSYPISPAVSR + 2 Phospho (ST)	130	30
		GYYSPYSVSGSGSTAGSR + Phospho (ST)	39	30
		GYYSPYSVSGSGSTAGSR + 2 Phospho (ST)	31	28
PHTF2_HUMAN	Q8N3S3	TITNVSDEVSSEEGPETGYSLR + 3 Phospho (ST)	52	30
PELI3_HUMAN	Q8N2H9	TPAGLLR + Phospho (ST)	35	27

PDE12_HUMAN	Q6L8Q7	EAKPGAAEPEVGVPSSLSPSSPSSSWTETDVEER + 2 Phospho (ST)	76	33
PCTK1_HUMAN	Q00536	RVSLSEIGFGK + Phospho (ST)	39	30
		LGEGTYATVYK + Phospho (ST)	46	29
PAR6B_HUMAN	Q9BYG5	HGAGSGCLGTMEVK + Phospho (ST)	110	29
		HGAGSGCLGTMEVK + Oxidation (M); Phospho (ST)	50	29
P80C_HUMAN	P38432	AFQLEEGEETEPDCK + Phospho (ST)	38	29
P34_HUMAN	Q6PD74	AFWMAIGGDRDEIEGLSSDEEH + Oxidation (M); 2 Phospho (ST)	48	30
		AFWMAIGGDRDEIEGLSSDEEH + 2 Phospho (ST)	35	30
P2R3A_HUMAN	Q06190	KPGTPLPPPATSPSSPRPLSPVPHVNNVVNAPLSINIPR + 2 Phospho (ST)	60	28
OCC1_HUMAN	Q8TAD7	DVTEESVTEDDKRR + Phospho (ST)	43	31
NUPL1_HUMAN	Q9BVL2	TQKTPPGLQHEYAAPADYFR + Phospho (ST)	64	32
NOD2_HUMAN	Q9HC29	RLLDLATVK + Phospho (ST)	32	26
NDUC2_HUMAN	O95298	SLPPPK + Phospho (ST)	43	28
NBR1_HUMAN	Q14596	VPHNTPVDVTPCMSPLPHDSPLIEKPGLGQIEEENEGAGFK + 2 Phospho (ST)	33	33
NAP5_HUMAN	O14513	SPQLLR + Phospho (ST)	34	27
NALP6_HUMAN	P59044	LSEAGLR + Phospho (ST)	37	28
MYO1H_HUMAN	Q8N1T3	INSSLVNKVGQR + Phospho (ST)	32	30
MT2_HUMAN	P02795	MDPNCSCAAGDSCTCAGSCK + Phospho (ST)	39	24
MSH4_HUMAN	O15457	EDFPRTEQVPEK + Phospho (ST)	33	30
MRP9_HUMAN	Q96J65	QLLCVARALLRNSK + Phospho (ST)	30	29
MINA_HUMAN	Q8IUF8	RLSGFLR + Phospho (ST)	34	27
MCLN2_HUMAN	Q8IZK6	KRSDDHLIPIS + Phospho (ST)	39	30
MBOA7_HUMAN	Q96N66	AGGGPTLQCPPPSSPEK + Phospho (ST)	31	30
MAT2B_HUMAN	Q9NZL9	YEMACAIADAFNLPSSHLRPITDSPVLGAQRPR + Oxidation (M); Phospho (ST)	33	32
M6PBP_HUMAN	O60664	VSGAQEMVSSAK + Phospho (ST)	37	29

		LGSLSER + Phospho (ST)	35	29
LRIG1_HUMAN	Q96JA1	ACDASPESTPLTGQLPGK + Phospho (ST)	32	31
LRC53_HUMAN	A6NM62	SLSSLPK + 2 Phospho (ST)	32	27
LFDH_HUMAN	Q008S8	VDFSTVLPR + Phospho (ST)	30	28
LARP5_HUMAN	Q92615	FTSSQTQSPTPPKPPSPSFELGLSSFPPLPGAAGNLK + 2 Phospho (ST)	43	32
		EPSVPASCAVSATYERSPSPAHLPDDPK + 2 Phospho (ST)	40	33
		EPPSSPLQPQK + Phospho (ST)	43	30
LAP4_HUMAN	Q14160	RSEACPCQPDSGSPLPAEEEK + Phospho (ST)	52	30
		NSLESISSIDR + Phospho (ST)	44	29
		MVEPENAVTITPLRPEDDYSPR + Phospho (ST)	48	32
		MVEPENAVTITPLRPEDDYSPR + Oxidation (M); Phospho (ST)	37	32
		MAESPCSPSGQQPPSPPSPDELPANVK + 2 Phospho (ST)	41	32
		LPLLPPESPGPLR + Phospho (ST)	34	29
		LAEAPSPAPTPSPTPVEDLGPQTSTSPGRLSPDFAEELR + Phospho (ST)	53	32
		LAEAPSPAPTPSPTPVEDLGPQTSTSPGRLSPDFAEELR + 2 Phospho (ST)	61	33
		ALSPAELR + Phospho (ST)	51	28
		AFAAVPTSHPPEDAPAQPPTPGPAASPEQLSFR + Phospho (ST)	43	33
		AFAAVPTSHPPEDAPAQPPTPGPAASPEQLSFR + 2 Phospho (ST)	75	33
		AEEEEASTEEEDKEGAVVSAPSVK + 2 Phospho (ST)	65	31
KV106_HUMAN	P01598	ASSLESGVPSR + Phospho (ST)	32	28
KDM1_HUMAN	O60341	KLPPPPPQAPPEEENESEPEEPSGVEGAAFQSR + Phospho (ST)	44	33
		EMDESLANLSEDEYYSEEER + Phospho (ST); Phospho (Y)	59	28
		EMDESLANLSEDEYYSEEER + 2 Phospho (ST)	86	28
K1688_HUMAN	Q9C0H5	AFSEDEALAQQENR + Phospho (ST)	44	30
K0819_HUMAN	Q7RTP6	RPDSPTRPTLR + Phospho (ST)	57	30

		LGSPLAVDEALR + Phospho (ST)	45	29
K0652_HUMAN	O75143	ASPHDVLETIFVR + Phospho (ST)	59	30
K0574_HUMAN	O60320	SLSRLR + 2 Phospho (ST)	43	27
K0329_HUMAN	O15040	FNAISSEDFDQELVVKPIK + 2 Phospho (ST)	40	32
JKIP3_HUMAN	Q5VZ66	KNEDLSHALR + Phospho (ST)	31	30
IQWD1_HUMAN	Q58WW2	DSALQDTDDSDDDPVLIPGAR + 2 Phospho (ST)	80	31
		DGEQSPNVSLMQR + Oxidation (M); Phospho (ST)	32	30
IMPA3_HUMAN	Q9NX62	TREGAEDK + Phospho (ST)	27	27
IL1A_HUMAN	P01583	VPDMFEDLKNCYSENEEDSSSIDHLSLNQK + Phospho (ST)	86	33
		RRLSLSQSITDDDLEAIANDSEEEIIKPR + Phospho (ST)	56	32
		RRLSLSQSITDDDLEAIANDSEEEIIKPR + 2 Phospho (ST)	67	33
		RLSLSQSITDDDLEAIANDSEEEIIKPR + 3 Phospho (ST)	106	33
		RLSLSQSITDDDLEAIANDSEEEIIKPR + 2 Phospho (ST)	143	33
		NCYSENEEDSSSIDHLSLNQK + Phospho (ST)	106	31
		LSLSQSITDDDLEAIANDSEEEIIKPR + 2 Phospho (ST)	72	33
IIP45_HUMAN	Q5JXC2	RKSFDASDTLALPR + Phospho (ST)	37	30
IFRD2_HUMAN	Q12894	SSAQADSGSSDDEAASEAR + 3 Phospho (ST)	141	23
HYLS1_HUMAN	Q96M11	KLPFPLSPS + Phospho (ST)	49	29
HUCE1_HUMAN	O43159	ALEAASLSQHPPSLCISDSEEEEEER + 2 Phospho (ST)	83	32
HTRA2_HUMAN	O43464	LSVGVTEPR + Phospho (ST)	34	28
HRX_HUMAN	Q03164	GPGEPDSPTPLHPPTPPILSTDR + 2 Phospho (ST)	42	32
HMGX4_HUMAN	Q9UGU5	SPPTTMLLPASPAKAPETEPIDVAAHLQLLGESLSLIGHR + Oxidation (M); 2 Phospho (ST)	35	30
		KHSSDDYYYGDISSLESSQK + Phospho (Y)	38	32
HEAT2_HUMAN	Q86Y56	ENEEDLKDKLDFAPPTPPHYPPHER + Phospho (ST)	46	33

GRIPE_HUMAN	Q6GYQ0	SSSTSDILEPFTVER + Phospho (ST)	39	31
GR65_HUMAN	Q9BQQ3	KPPGTPPPSALPLGAPPPDALPPGPTPEDSPSLETGSR + 2 Phospho (ST)	34	33
GP2_HUMAN	P55259	YCTVPR + Phospho (ST)	30	26
GCFC_HUMAN	Q9Y5B6	MADHLEGLSSDDEETSTDITNFNLEKDR + 2 Phospho (ST)	43	32
		MADHLEGLSSDDEETSTDITNFNLEK + Oxidation (M); 2 Phospho (ST)	92	31
		LVREDENDASDDEDDDEKR + Phospho (ST)	54	30
GASP2_HUMAN	Q96D09	TNREDCFESESEDEFYK + 2 Phospho (ST)	56	26
GALR1_HUMAN	P47211	YVAIVHSRR + Phospho (ST); Phospho (Y)	33	30
FUSIP_HUMAN	O75494	SRSRSFDYNYR + 3 Phospho (ST)	33	25
		SRSFDYNYR + 2 Phospho (ST)	61	24
		SFDYNYR + Phospho (ST)	39	25
FTSJ2_HUMAN	Q8N1G2	VAELALSLSSTSDDEPPSSVSHGAK + Phospho (ST)	35	32
		ASTTSLSGSDSETEGKQHSSDSFDDAFK + 3 Phospho (ST)	45	29
		ASTTSLSGSDSETEGK + 3 Phospho (ST)	59	25
		ASTTSLSGSDSETEGK + 2 Phospho (ST)	30	27
FRAP_HUMAN	P42345	TRTDSYSAGQSVEILDGVELGEPAHKK + Phospho (ST)	35	32
FOSL1_HUMAN	P15407	RRPCEQISPEEEER + Phospho (ST)	61	31
FMO4_HUMAN	P31512	TAAQVLLSTR + 3 Phospho (ST)	36	28
FAN_HUMAN	Q92636	EAVTGITVSR + 2 Phospho (ST)	34	29
FA44A_HUMAN	Q8NFC6	YYSDSDDELTVEQR + Phospho (ST); Phospho (Y)	57	27
		YYSDSDDELTVEQR + 2 Phospho (ST)	93	27
		SVSDPVEDKKEQESDEEEEEEEEDEPSGATTR + Phospho (ST)	94	32
		RLSESLHVVDENKNESK + Phospho (ST)	59	32
		KQHYLSSEDEPDDNPDVLDSR + Phospho (ST); Phospho (Y)	55	31
		KQHYLSSEDEPDDNPDVLDSR + 2 Phospho (ST)	74	31

		EQESDEEEEEEEEDEPSGATTR + Phospho (ST)	45	27
		DKDVTLSPVK + Phospho (ST)	31	29
FA38A_HUMAN	Q92508	TASELLLDR + Phospho (ST)	42	29
FA29A_HUMAN	Q7Z4H7	TPENLITEIR + Phospho (ST)	49	30
EYS_HUMAN	Q5T1H1	SIIAPGR + Phospho (ST)	34	28
EXT1_HUMAN	Q16394	IAESYQNILAAIEGSR + 2 Phospho (ST)	40	31
ESAM_HUMAN	Q96AP7	ALRPPHGPPRPGALTPTPSLSSQALPSPR + Phospho (ST)	32	28
ERG19_HUMAN	P53602	RNSRDGDPLPSSLSCK + Phospho (ST)	54	31
ENW1_HUMAN	Q9UQF0	EVISQLTR + Phospho (ST)	34	30
DPCD_HUMAN	Q9BVM2	KDTKMSFQWR + Phospho (ST)	31	30
DNMBP_HUMAN	Q6XZF7	SHSDASVGSHSSTESEHGSSSPR + Phospho (ST)	33	30
		HETSDHEAEEPDCIISEAPTSPLGHLTSEYDTDR + Phospho (ST)	127	33
DDHD1_HUMAN	Q8NEL9	EPTSVSENEGISTIPSPVTSPVLSR + 2 Phospho (ST)	47	33
DCBD1_HUMAN	Q8N8Z6	KGSTFRPMDTDAEEAGVSTDAGGHYDCPQR + Phospho (ST)	74	32
		HSLSSGGFSPVAGVGAQDGDYQRPHSAQPADR + Phospho (ST)	86	33
DBC1_HUMAN	O60477	LPTLLR + Phospho (ST)	35	27
DAOA_HUMAN	P59103	LMGADSLQLFR + Phospho (ST)	40	30
DAAM2_HUMAN	Q86T65	FQTLLNELDRSLGRYR + Phospho (ST); Phospho (Y)	38	32
CTL2_HUMAN	Q8IWA5	HGTPQKYDPTFK + Phospho (ST)	38	30
CSF1_HUMAN	P09603	ADSPLEQPEGSPLTQDDR + Phospho (ST)	84	31
CS061_HUMAN	Q9H0W8	WKEPGSGGPQNLSGPGGR + Phospho (ST)	58	31
		AGPGSSPLFSLLPGYR + Phospho (ST)	49	31
CS044_HUMAN	Q9H6X5	KLFSVPSQLR + Phospho (ST)	30	27
CS034_HUMAN	Q8NCQ2	ATVITTR + Phospho (ST)	29	29
CQ079_HUMAN	Q9NQ92	GTQSIPNDSPAR + Phospho (ST)	45	29

CQ064_HUMAN	Q86WR6	DSLRELSQKPK + 2 Phospho (ST)	30	30
CQ049_HUMAN	Q8IXM2	VYEDSGIPLPAESPK + Phospho (ST)	47	31
		KVASGVLSPPPAAPPPSSSSVPEAGGPPIKK + Phospho (ST)	60	28
CP014_HUMAN	Q9BUT9	RAPSTSPSFEGTQETYTVAHEENVR + Phospho (ST); Phospho (Y)	45	32
		RAPSTSPSFEGTQETYTVAHEENVR + Phospho (ST)	92	33
CN133_HUMAN	Q9H9C1	TRPGSFQSLSDALSDTPAK + Phospho (ST)	40	31
CK063_HUMAN	Q6NUN7	SKPFSELSDSDLEEK + 2 Phospho (ST)	33	30
CK059_HUMAN	Q6IAA8	KLLLDPSSPPTK + Phospho (ST)	32	30
CIR_HUMAN	Q86X95	NLTANDPSQEYVASEGEEDPEVEFLK + Phospho (ST)	47	32
CG029_HUMAN	Q96FA7	EKGLLESMGLLASER + Phospho (ST)	32	31
CG028_HUMAN	O95766	HIEPELAGRDSPIR + Phospho (ST)	42	31
CG027_HUMAN	Q6PJG6	GSPNTASAEATLPR + Phospho (ST)	46	29
CDR2_HUMAN	Q01850	SSSETILSSLAGSDIVK + Phospho (ST)	57	31
		SPGKCDQEKPAPSFACLK + Phospho (ST)	32	31
CDC2_HUMAN	P06493	VYTHEVVTLWYR + Phospho (ST)	74	31
		IGEGTYGVVYK + Phospho (Y)	68	30
		IGEGTYGVVYK + Phospho (ST); Phospho (Y)	85	29
		IGEGTYGVVYK + Phospho (ST)	85	29
CD2L7_HUMAN	Q9NYV4	SGSYSGRSPSPYGR + 2 Phospho (ST)	32	28
		RTPTMPQEEAAACPPHILPPEK + Phospho (ST)	61	32
		RQSVSPPYKEPSAYQSSTR + 2 Phospho (ST)	90	32
		NSSPAPPQPAPGK + Phospho (ST)	58	30
		LYNSEESRPYTNK + Phospho (ST)	69	30
		HSSISPVRLPLNSSLGAELSR + 2 Phospho (ST)	85	31
		HLLTDLPLPPELPGGDLSPPDSPEPK + 2 Phospho (ST)	75	32

		ESKGSPVFLPR + Phospho (ST)	69	30
CD2L5_HUMAN	Q14004	SRKSPSPAGGGSSPYSR + 3 Phospho (ST)	33	29
		SLSPLGGRDDSPVSHR + 3 Phospho (ST)	61	29
		LYSSEESRPYTNK + Phospho (ST)	68	31
		ILELTPEPDRPR + Phospho (ST)	57	30
		HSSISPSTLTLK + 2 Phospho (ST)	56	30
		GGDVSPSPYSSSSWR + Phospho (ST)	33	29
		ASNTSTPTKGNTETSASASQTNHVK + Phospho (ST)	46	32
CD2L2_HUMAN	Q9UQ88	EYGSPLKAYTPVVVTQWYR + Phospho (ST); Phospho (Y)	47	32
		AYTPVVVTQWYR + Phospho (ST)	66	30
CD2L1_HUMAN	P21127	RGTSPRPPEGGLGYSQLGDDDLKETGFHLTTTNQGASAAGPGFSLK + 2 Phospho (ST)	49	32
		RGTSPRPPEGGLGYSQLGDDDLK + 2 Phospho (ST)	76	32
		REDSMEDRGEEDDSLAIKPPQQMSR + Phospho (ST)	37	33
		RDSLEEGELRDHR + Phospho (ST)	43	31
		RDSLEEGELR + Phospho (ST)	42	29
		GTSPRPPEGGLGYSQLGDDDLKETGFHLTTTNQGASAAGPGFSLK + Phospho (ST); Phospho (Y)	35	33
		GTSPRPPEGGLGYSQLGDDDLKETGFHLTTTNQGASAAGPGFSLK + Phospho (ST)	55	32
		GTSPRPPEGGLGYSQLGDDDLK + Phospho (ST)	61	32
		GTSPRPPEGGLGYSQLGDDDLK + 2 Phospho (ST)	40	32
		EYGSPLKAYTPVVVTLWYR + Phospho (ST); Phospho (Y)	70	32
		EYGSPLKAYTPVVVTLWYR + 2 Phospho (ST)	55	32
		DLLSDLQDISDSER + Phospho (ST)	70	31

		AYTPVVVTLWYR + Phospho (ST)	50	30
CCD99_HUMAN	Q96EA4	SVSIHTPVVSLSPHK + Phospho (ST)	49	30
CCD38_HUMAN	Q502W7	KITQVYK + Phospho (ST)	36	29
CC85A_HUMAN	Q96PX6	HALGGSLEHLPR + Phospho (ST)	50	30
CA1L1_HUMAN	Q08AD1	YDGESDKEQFDDDQK + Phospho (ST)	51	28
		SISNEGLTLNNSHVSK + Phospho (ST)	84	31
		SESVEGFLSPSR + Phospho (ST)	32	29
		LNQSSPDNVTDTK + Phospho (ST)	39	30
		GALSPITDNTEVDTGIHVPSEDIPETMDEDSSLR + Oxidation (M); Phospho (ST)	64	33
CA149_HUMAN	Q9HAF1	REPGSGTESDTSPDFHNQENEPSQEDPEDLDGSVQGVKPQK + 3 Phospho (ST)	67	32
		REPGSGTESDTSPDFHNQENEPSQEDPEDLDGSVQGVKPQK + 2 Phospho (ST)	62	32
BRAF1_HUMAN	P15056	SSSAPNVHINTIEPVNIDDLIR + Phospho (ST)	41	32
		SASEPSLNR + Phospho (ST)	42	28
		RDSSDDWEIPDGQITVGQR + Phospho (ST)	55	32
BGAL_HUMAN	P16278	TEAVASSLYDILAR + Phospho (ST); Phospho (Y)	31	31
BAT2L_HUMAN	Q5JSZ5	VRSPDEALPGGLSGCSSGSGHSPYALER + 2 Phospho (ST)	51	33
		LKFSDDEEEEEVVKDGRPK + Phospho (ST)	68	32
		LKFSDDEEEEEVVK + Phospho (ST)	106	31
		EAGSPAQEFK + Phospho (ST)	39	28
BASP_HUMAN	P80723	SDGAPASDSKPGSSEAAPSSK + 2 Phospho (ST)	47	30
		KTEAPAAPAAQETKSDGAPASDSKPGSSEAAPSSK + Phospho (ST)	46	33
		KAEGAATEEEGTPKESEPQAAAEPAEAK + Phospho (ST)	33	33
		AEGAATEEEGTPKESEPQAAAEPAEAK + Phospho (ST)	96	32
BAG4_HUMAN	O95429	SSGNSPTPVSR + Phospho (ST)	66	29
BA2D1_HUMAN	Q9Y520	SVEDVRPHHTDANNQSACFEAPDQK + Phospho (ST)	130	32

		SFSSQRPVDR + Phospho (ST)	50	29
		LPDLSPVENKEHKPGPIGK + Phospho (ST)	52	31
		AFGSGIDIKPGTPPIAGR + Phospho (ST)	38	30
B4GT1_HUMAN	P15291	LPQLVGVSTPLQGGSNSAAAIGQSSGELR + Phospho (ST)	57	30
AS250_HUMAN	Q2PPJ7	RLSNSSLCSIEEEHR + 2 Phospho (ST)	52	30
ARD1A_HUMAN	P41227	GLAAEDSGGDSKDLSEVSETTESTDVK + Phospho (ST)	69	32
		DLSEVSETTESTDVK + 2 Phospho (ST)	55	28
ALEX_HUMAN	P84996	SISDPPAPRSR + 2 Phospho (ST)	35	29
AHTF1_HUMAN	Q8WYP5	EVSPSDVR + Phospho (ST)	37	27
ADX_HUMAN	P10109	SSSEDKITVHFINR + Phospho (ST)	52	31
ABCC9_HUMAN	O60706	FNLDPECKCTDDR + Phospho (ST)	34	28
1A68_HUMAN	P01891	KGGSYSQAASSDSAQGSDVSLTACKV + Phospho (ST)	37	32
		KGGSYSQAASSDSAQGSDVSLTACK + Phospho (ST)	38	32

6.5 References

1. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M., Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006,** 127, (3), 635-648.

2. Bode, A. M.; Dong, Z. G., Post-translational modification of p53 in tumorigenesis. *Nature Reviews Cancer* **2004,** 4, (10), 793-805.

3. Hunter, T.; Sefton, B. M., Transforming Gene-Product of Rous-Sarcoma Virus Phosphorylates Tyrosine. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1980,** 77, (3), 1311-1315.

4. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & Cellular Proteomics* **2005,** 4, (7), 873-886.

5. Imanishi, S. Y.; Kochin, V.; Eriksson, J. E., Optimization of phosphopeptide elution conditions in immobilized Fe(III) affinity chromatography. *Proteomics* **2007,** 7, (2), 174-176.

6. Kweon, H. K.; Hakansson, K., Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis. *Analytical Chemistry* **2006,** 78, (6), 1743-1749.

7. Sugiyama, N.; Nakagami, H.; Mochida, K.; Daudi, A.; Tomita, M.; Shirasu, K.; Ishihama, Y., Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in Arabidopsis. *Molecular Systems Biology* **2008,** 4.

8. Villen, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P., Large-scale phosphorylation analysis of mouse liver. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** 104, (5), 1488-1493.

9. Zhou, H.; Ye, M.; Dong, J.; Han, G.; Jiang, X.; Wu, R.; Zou, H., Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *Journal of Proteome Research* **2008,** 7, (9), 3957-3967.

10. Li, Y.; Leng, T.; Lin, H.; Deng, C.; Xu, X.; Yao, N.; Yang, P.; Zhang, X., Preparation of Fe3O4@ZrO2 core-shell microspheres as affinity probes for selective enrichment and direct determination of phosphopeptides using matrix-assisted laser desorption ionization mass spectrometry. *Journal of Proteome Research* **2007,** 6, (11), 4498-4510.

11. Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R., SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Molecular & Cellular Proteomics* **2008,** 7, (4), 661-671.

12. Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P., Phosphoproteomic analysis of the developing mouse brain. *Molecular & Cellular Proteomics* **2004,** 3, (11),

1093-1101.
13. Zhang, X. T.; Kang, L. G.; Ding, L.; Vranic, S.; Gatalica, Z.; Wang, Z. Y., A positive feedback loop of ER-alpha 36/EGFR promotes malignant growth of ER-negative breast cancer cells. *Oncogene* **2011,** 30, (7), 770-780.
14. Zhong, Z. J.; Yeow, W. S.; Zou, C. H.; Wassell, R.; Wang, C. G.; Pestell, R. G.; Quong, J. N.; Quong, A. A., Cyclin D1/Cyclin-Dependent Kinase 4 Interacts with Filamin A and Affects the Migration and Invasion Potential of Breast Cancer Cells. *Cancer Research* **2010,** 70, (5), 2105-2114.
15. Yuan, T. C.; Wang, Y. P.; Zhao, Z. Z. J.; Gu, H. H., Protein-tyrosine Phosphatase PTPN9 Negatively Regulates ErbB2 and Epidermal Growth Factor Receptor Signaling in Breast Cancer Cells. *Journal of Biological Chemistry* **2010,** 285, (20), 14861-14870.
16. Ye, Y.; Xiao, Y.; Wang, W.; Yearsley, K.; Gao, J. X.; Shetuni, B.; Barsky, S. H., ER alpha signaling through slug regulates E-cadherin and EMT. *Oncogene* **2010,** 29, (10), 1451-1462.
17. Pichot, C. S.; Arvanitis, C.; Hartig, S. M.; Jensen, S. A.; Bechill, J.; Marzouk, S.; Yu, J. D.; Frost, J. A.; Corey, S. J., Cdc42-Interacting Protein 4 Promotes Breast Cancer Cell Invasion and Formation of Invadopodia through Activation of N-WASp. *Cancer Research* **2010,** 70, (21), 8347-8356.
18. Migneco, G.; Whitaker-Menezes, D.; Chiavarina, B.; Castello-Cros, R.; Pavlides, S.; Pestell, R. G.; Fatatis, A.; Flomenberg, N.; Tsirigos, A.; Howell, A.; Martinez-Outschoorn, U. E.; Sotgia, F.; Lisanti, M. P., Glycolytic cancer associated fibroblasts promote breast cancer tumor growth, without a measurable increase in angiogenesis Evidence for stromal-epithelial metabolic coupling. *Cell Cycle* **2010,** 9, (12), 2412-2422.
19. Konson, A.; Pradeep, S.; Seger, R., Phosphomimetic Mutants of Pigment Epithelium-Derived Factor with Enhanced Antiangiogenic Activity as Potent Anticancer Agents. *Cancer Research* **2010,** 70, (15), 6247-6257.
20. Holland, S. J.; Pan, A.; Franci, C.; Hu, Y. M.; Chang, B.; Li, W. Q.; Duan, M.; Torneros, A.; Yu, J. X.; Heckrodt, T. J.; Zhang, J.; Ding, P. Y.; Apatira, A.; Chua, J.; Brandt, R.; Pine, P.; Goff, D.; Singh, R.; Payan, D. G.; Hitoshi, Y., R428, a Selective Small Molecule Inhibitor of Axl Kinase, Blocks Tumor Spread and Prolongs Survival in Models of Metastatic Breast Cancer. *Cancer Research* **2010,** 70, (4), 1544-1554.
21. Chiavarina, B.; Whitaker-Menezes, D.; Migneco, G.; Martinez-Outschoorn, U. E.; Pavlides, S.; Howell, A.; Tanowitz, H. B.; Casimiro, M. C.; Wang, C. G.; Pestell, R. G.; Grieshaber, P.; Caro, J.; Sotgia, F.; Lisanti, M. P., HIF1-alpha functions as a tumor promoter in cancer associated fibroblasts, and as a tumor suppressor in breast cancer cells Autophagy drives compartment-specific oncogenesis. *Cell Cycle* **2010,** 9, (17), 3534-3551.
22. Adams, L. S.; Phung, S.; Yee, N.; Seeram, N. P.; Li, L. Y.; Chen, S. A., Blueberry Phytochemicals Inhibit Growth and Metastatic Potential of MDA-MB-231 Breast Cancer Cells through Modulation of the Phosphatidylinositol 3-Kinase Pathway. *Cancer Research* **2010,** 70, (9), 3594-3605.
23. Wang, S.; Liu, Q.; Zhang, Y.; Liu, K.; Yu, P. F.; Liu, K.; Luan, J. L.; Duan, H. Y.; Lu, Z.

Q.; Wang, F. F.; Wu, E. X.; Yagasaki, K.; Zhang, G. Y., Suppression of growth, migration and invasion of highly-metastatic human breast cancer cells by berbamine and its molecular mechanisms of action. *Molecular Cancer* **2009,** 8, -.
24. Sung, B.; Murakami, A.; Oyajobi, B. O.; Aggarwal, B. B., Zerumbone Abolishes RANKL-Induced NF-kappa B Activation, Inhibits Osteoclastogenesis, and Suppresses Human Breast Cancer-induced Bone Loss in Athymic Nude Mice. *Cancer Research* **2009,** 69, (4), 1477-1484.
25. Gonzalez, M. E.; Li, X.; Toy, K.; DuPrie, M.; Ventura, A. C.; Banerjee, M.; Ljungman, M.; Merajver, S. D.; Kleer, C. G., Downregulation of EZH2 decreases growth of estrogen receptor-negative invasive breast carcinoma and requires BRCA1. *Oncogene* **2009,** 28, (6), 843-853.
26. Ta, H. Q.; Thomas, K. S.; Schrecengost, R. S.; Bouton, A. H., A Novel Association between p130(Cas) and Resistance to the Chemotherapeutic Drug Adriamycin in Human Breast Cancer Cells. *Cancer Research* **2008,** 68, (21), 8796-8804.
27. Jiang, P.; Enomoto, A.; Jijiwa, M.; Kato, T.; Hasegawa, T.; Ishida, M.; Sato, T.; Asai, N.; Murakumo, Y.; Takahashi, M., An actin-binding protein girdin regulates the motility of breast cancer cells. *Cancer Research* **2008,** 68, (5), 1310-1318.
28. Bhaumik, D.; Scott, G. K.; Schokrpur, S.; Patil, C. K.; Campisi, J.; Benz, C. C., Expression of microRNA-146 suppresses NF-kappa B activity with reduction of metastatic potential in breast cancer cells. *Oncogene* **2008,** 27, (42), 5643-5647.
29. Barkan, D.; Kleinman, H.; Simmons, J. L.; Asmussen, H.; Kamaraju, A. K.; Hoenorhoff, M. J.; Liu, Z. Y.; Costes, S. V.; Cho, E. H.; Lockett, S.; Khanna, C.; Chambers, A. F.; Green, J. E., Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton. *Cancer Research* **2008,** 68, (15), 6241-6250.
30. Mertens-Talcott, S. U.; Chintharlapalli, S.; Li, M. R.; Safe, S., The oncogenic microRNA-27a targets genes that regulate specificity protein transcription factors and the G(2)-M checkpoint in MDA-MB-231 breast cancer cells. *Cancer Research* **2007,** 67, (22), 11001-11011.
31. Jallal, H.; Valentino, M. L.; Chen, G. P.; Boschelli, F.; Ali, S.; Rabbani, S. A., A Src/Abl kinase inhibitor, SKI-606, blocks breast cancer invasion, growth, and metastasis in vitro and in vivo. *Cancer Research* **2007,** 67, (4), 1580-1588.
32. Dowling, R. J. O.; Zakikhani, M.; Fantus, I. G.; Pollak, M.; Sonenberg, N., Metformin inhibits mammalian target of rapamycin-dependent translation initiation in breast cancer cells. *Cancer Research* **2007,** 67, (22), 10804-10812.
33. Wang, Y.; Lam, J. B.; Lam, K. S. L.; Liu, J.; Lam, M. C.; Hoo, R. L. C.; Wu, D. H.; Cooper, G. J. S.; Xu, A. M., Adiponectin modulates the glycogen synthase kinase-3 beta/beta-catenin signaling pathway and attenuates mammary tumorigenesis of MDA-MB-231 cells in nude mice. *Cancer Research* **2006,** 66, (23), 11462-11470.
34. Troussard, A. A.; McDonald, P. C.; Wederell, E. D.; Mawji, N. M.; Filipenko, N. R.; Gelmon, K. A.; Kucab, J. E.; Dunn, S. E.; Emerman, J. T.; Bally, M. B.; Dedhar, S., Preferential dependence of breast cancer cells versus normal cells on integrin-linked kinase

for protein kinase B/Akt activation and cell survival. *Cancer Research* **2006,** 66, (1), 393-403.
35. Bigelow, R. L. H.; Cardelli, J. A., The green tea catechins, (-)-epigallocatechin-3-gallate (EGCG) and (-)-epicatechin-3-gallate (ECG), inhibit HGF/Met signaling in immortalized and tumorigenic breast epithelial cells. *Oncogene* **2006,** 25, (13), 1922-1930.
36. Carrascal, M.; Ovefletro, D.; Casas, V.; Gay, M.; Ablan, J., Phosphorylation Analysis of Primary Human T Lymphocytes Using Sequential IMAC and Titanium Oxide Enrichment. *Journal of Proteome Research* **2008,** 7, (12), 5167-5176.
37. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
38. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical Chemistry* **2008,** 80, (12), 4696-4710.
39. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* **2007,** 4, (3), 207-214.
40. Kyte, J.; Doolittle, R. F., A Simple Method for Displaing the Hydropathic Character of a Protein. *Journal of Molecular Biology* **1982,** 157, (1), 105-132.
41. Thangapazham, R. L.; Passi, N.; Maheshwari, R. K., Green tea polyphenol and epigallocatechin gallate induce apoptosis and inhibit invasion in human breast cancer cells. *Cancer Biology & Therapy* **2007,** 6, (12), 1938-1943.
42. Koide, H.; Asai, T.; Furuya, K.; Tsuzuku, T.; Kato, H.; Dewa, T.; Nango, M.; Maeda, N.; Oku, N., Inhibition of Akt (ser473) Phosphorylation and Rapamycin-Resistant Cell Growth by Knockdown of Mammalian Target of Rapamycin with Small Interfering RNA in Vascular Endothelial Growth Factor Receptor-1-Targeting Vector. *Biological & Pharmaceutical Bulletin* **2011,** 34, (5), 602-608.
43. Bigelow, R. L. H.; Williams, B. J.; Carroll, J. L.; Daves, L. K.; Cardelli, J. A., TIMP-1 overexpression promotes tumorigenesis of MDA-MB-231 breast cancer cells and alters expression of a subset of cancer promoting genes in vivo distinct from those observed in vitro. *Breast Cancer Research and Treatment* **2009,** 117, (1), 31-44.
44. Jiang, J. H.; Slivova, V.; Harvey, K.; Valachovicova, T.; Sliva, D., Ganoderma lucidum suppresses growth of breast cancer cells through the inhibition of Akt/NF-kappa B signaling. *Nutrition and Cancer-an International Journal* **2004,** 49, (2), 209-216.
45. Linher-Melville, K.; Zantinge, S.; Sanli, T.; Gerstein, H.; Tsakiridis, T.; Singh, G., Establishing a relationship between prolactin and altered fatty acid beta-Oxidation via carnitine palmitoyl transferase 1 in breast cancer cells. *Bmc Cancer* **2011,** 11, -.
46. Gauci, S.; Helbig, A. O.; Slijper, M.; Krijgsveld, J.; Heck, A. J. R.; Mohammed, S.,

Lys-N and Trypsin Cover Complementary Parts of the Phosphoproteome in a Refined SCX-Based Approach. *Analytical Chemistry* **2009,** 81, (11), 4493-4501.
47. Raobaikady, B.; Reed, M. J.; Leese, M. P.; Potter, B. V. L.; Purohit, A., Inhibition of MDA-MB-231 cell cycle progression and cell proliferation by C-2-substituted oestradiol mono- and bis-3-O-sulphamates. *International Journal of Cancer* **2005,** 117, (1), 150-159.
48. Liao, C. H.; Pan, S. L.; Guh, J. H.; Teng, C. M., Genistein inversely affects tubulin-binding agent-induced apoptosis in human breast cancer cells. *Biochemical Pharmacology* **2004,** 67, (11), 2031-2038.
49. Wang, P. S.; Chou, F. S.; Bloomston, M.; Vonau, M. S.; Saji, M.; Espinosa, A.; Pinzone, J. J., Thiazolidinediones Downregulate Wnt/beta-Catenin Signaling Via Multiple Mechanisms in Breast Cancer Cells. *Journal of Surgical Research* **2009,** 153, (2), 210-216.
50. Zhang, T. F.; Yu, S. Q.; Guan, L. S.; Wang, Z. Y., Inhibition of breast cancer cell growth by the Wilms' tumor suppressor WT1 is associated with a destabilization of beta-catenin. *Anticancer Research* **2003,** 23, (5A), 3575-3584.
51. Bourguignon, L. Y. W.; Singleton, P. A.; Zhu, H. B.; Zhou, B., Hyaluronan promotes signaling interaction between CD44 and the transforming growth factor beta receptor I in metastatic breast tumor cells. *Journal of Biological Chemistry* **2002,** 277, (42), 39703-39712.
52. Wesierska-Gadek, J.; Gritsch, D.; Zulehner, N.; Komina, O.; Maurer, M., Roscovitine, a Selective CDK Inhibitor, Reduces the Basal and Estrogen-Induced Phosphorylation of ER-alpha in Human ER-Positive Breast Cancer Cells. *Journal of Cellular Biochemistry* **2011,** 112, (3), 761-772.
53. Butt, A. J.; Dickson, K. A.; Jambazov, S.; Baxter, R. C., Enhancement of tumor necrosis factor-alpha-induced growth inhibition by insulin-like growth factor-binding protein-5 (IGFBP-5), but not IGFBP-3 in human breast cancer cells. *Endocrinology* **2005,** 146, (7), 3113-3122.
54. Wang, J. N.; Wan, W. Z.; Sun, R. H.; Liu, Y.; Sun, X. J.; Ma, D. L.; Zhang, N., Reduction of Akt2 expression inhibits chemotaxis signal transduction in human breast cancer cells. *Cellular Signalling* **2008,** 20, (6), 1025-1034.
55. Yang, W.; Ju, J. H.; Lee, K. M.; Shin, I., Akt isoform-specific inhibition of MDA-MB-231 cell proliferation. *Cellular Signalling* **2011,** 23, (1), 19-26.
56. Nakamura, M.; Han, B.; Nishishita, T.; Bai, Y.; Kakudo, K., Calcitonin targets extracellular signal-regulated kinase signaling pathway in human cancers. *Journal of Molecular Endocrinology* **2007,** 39, (5-6), 375-384.
57. Shankar, E.; Krishnamurthy, S.; Paranandi, R.; Basu, A., PKC epsilon induces Bcl-2 by activating CREB. *International Journal of Oncology* **2010,** 36, (4), 883-888.
58. Naviglio, S.; Di Gesto, D.; Illiano, F.; Chiosi, E.; Giordano, A.; Illiano, G.; Spina, A., Leptin Potentiates Antiproliferative Action of cAMP Elevation via Protein Kinase A Down-Regulation in Breast Cancer Cells. *Journal of Cellular Physiology* **2010,** 225, (3), 801-809.
59. Amorino, G. P.; Hamilton, V. M.; Valerie, K.; Dent, P.; Lammering, G.;

Schmidt-Ullrich, R. K., Epidermal growth factor receptor dependence of radiation-induced transcription factor activation in human breast carcinoma cells. *Molecular Biology of the Cell* **2002,** 13, (7), 2233-2244.
60. Parmer, T. G.; Ward, M. D.; Yurkow, E. J.; Vyas, V. H.; Kearney, T. J.; Hait, W. N., Activity and regulation by growth factors of calmodulin dependent protein kinase III (elongation factor 2-kinase) in human breast cancer. *British Journal of Cancer* **1999,** 79, (1), 59-64.
61. Boerner, J. L.; Demory, M. L.; Silva, C.; Parsons, S. J., Phosphorylation of Y845 on the epidermal growth factor receptor mediates binding to the mitochondrial protein cytochrome c oxidase subunit II. *Molecular and Cellular Biology* **2004,** 24, (16), 7059-7071.
62. Greenberg, V. L.; Zimmer, S. G., Paclitaxel induces the phosphorylation of the eukaryotic translation initiation factor 4E-binding protein 1 through a Cdk1-dependent mechanism. *Oncogene* **2005,** 24, (30), 4851-4860.
63. Chen, P.; Lu, N.; Ling, Y.; Chen, Y.; Hui, H.; Lu, Z. J.; Song, X. M.; Li, Z. Y.; You, Q. D.; Guo, Q. L., Inhibitory effects of wogonin on the invasion of human breast carcinoma cells by downregulating the expression and activity of matrix metalloproteinase-9. *Toxicology* **2011,** 282, (3), 122-128.
64. Soto-Guzman, A.; Robledo, T.; Lopez-Perez, M.; Salazar, E. P., Oleic acid induces ERK1/2 activation and AP-1 DNA binding activity through a mechanism involving Src kinase and EGFR transactivation in breast cancer cells. *Molecular and Cellular Endocrinology* **2008,** 294, (1-2), 81-91.
65. Bocca, C.; Bozzo, F.; Martinasso, G.; Canuto, R. A.; Miglietta, A., Involvement of PPAR alpha in the growth inhibitory effect of arachidonic acid on breast cancer cells. *British Journal of Nutrition* **2008,** 100, (4), 739-750.
66. Lindemann, R. K.; Braig, M.; Ballschmieter, P.; Guise, T. A.; Nordheim, A.; Dittmer, J., Protein kinase C alpha regulates Ets1 transcriptional activity in invasive breast cancer cells. *International Journal of Oncology* **2003,** 22, (4), 799-805.
67. Villa-Moruzzi, E., Targeting of FAK Ser(910) by ERK5 and PP1 delta in non-stimulated and phorbol ester-stimulated cells. *Biochemical Journal* **2007,** 408, 7-18.
68. Walsh, C.; Tanjoni, I.; Uryu, S.; Tomar, A.; Nam, J. O.; Luo, H.; Phillips, A.; Patel, N.; Kwok, C.; McMahon, G.; Stupack, D. G.; Schlaepfer, D. D., Oral delivery of PND-1186 FAK inhibitor decreases tumor growth and spontaneous breast to lung metastasis in pre-clinical models. *Cancer Biology & Therapy* **2010,** 9, (10), 778-790.
69. Nurcombe, V.; Smart, C. E.; Chipperfield, H.; Cool, S. M.; Boilly, B.; Hondermarck, H., The proliferative and migratory activities of breast cancer cells can be differentially regulated by heparan sulfates. *Journal of Biological Chemistry* **2000,** 275, (39), 30009-30018.
70. Bourguignon, L. Y. W.; Singleton, P. A.; Zhu, H. B.; Diedrich, F., Hyaluronan-mediated CD44 interaction with RhoGEF and Rho kinase promotes Grb2-associated binder-1 phosphorylation and phosphatidylinositol 3-kinase signaling

leading to cytokine (Macrophage-Colony stimulating factor) production and breast tumor progression. *Journal of Biological Chemistry* **2003,** 278, (32), 29420-29434.
71. Gril, B.; Palmieri, D.; Bronder, J. L.; Herring, J. M.; Vega-Valle, E.; Feigenbaum, L.; Liewehr, D. J.; Steinberg, S. M.; Merino, M. J.; Rubin, S. D.; Steeg, P. S., Effect of lapatinib on the outgrowth of metastatic breast cancer cells to the brain. *Journal of the National Cancer Institute* **2008,** 100, (15), 1092-1103.
72. Aguilar, Z.; Slamon, D. J., The transmembrane heregulin precursor is functionally active. *Journal of Biological Chemistry* **2001,** 276, (47), 44099-44107.
73. Kiang, J. G.; Gist, I. D.; Tsokos, G. C., Regulation of heat shock protein 72 kDa and 90 kDa in human breast cancer MDA-MB-231 cells. *Molecular and Cellular Biochemistry* **2000,** 204, (1-2), 169-178.
74. Shin, K. D.; Lee, M. Y.; Shin, D. S.; Lee, S.; Son, K. H.; Koh, S.; Paik, Y. K.; Kwon, B. M.; Han, D. C., Blocking tumor cell migration and invasion with biphenyl isoxazole derivative KRIBB3, a synthetic molecule that inhibits Hsp27 phosphorylation. *Journal of Biological Chemistry* **2005,** 280, (50), 41439-41448.
75. Shi, Y. Y.; Small, G. W.; Orlowski, R. Z., Proteasome inhibitors induce a p38 mitogen-activated protein kinase (MAPK)-dependent anti-apoptotic program involving MAPK phosphatase-1 and Akt in models of breast cancer. *Breast Cancer Research and Treatment* **2006,** 100, (1), 33-47.
76. Jackson, J. G.; Zhang, X. H.; Yoneda, T.; Yee, D., Regulation of breast cancer cell motility by insulin receptor substrate-2 (IRS-2) in metastatic variants of human breast cancer cell lines. *Oncogene* **2001,** 20, (50), 7318-7325.
77. del Rincon, S. V.; Guo, Q.; Morelli, C.; Shiu, H. Y.; Surmacz, E.; Miller, W. H., Retinoic acid mediates degradation of IRS-1 by the ubiquitin-proteasome pathway, via a PKC-dependant mechanism. *Oncogene* **2004,** 23, (57), 9269-9279.
78. Scott, R. W.; Hooper, S.; Crighton, D.; Li, A.; Konig, I.; Munro, J.; Trivier, E.; Wickman, G.; Morin, P.; Croft, D. R.; Dawson, J.; Machesky, L.; Anderson, K. I.; Sahai, E. A.; Olson, M. F., LIM kinases are required for invasive path generation by tumor and tumor-associated stromal cells. *Journal of Cell Biology* **2010,** 191, (1), 169-185.
79. Sarkar, S.; Mazumdar, A.; Dash, R.; Sarkar, D.; Fisher, P. B.; Mandal, M., ZD6474, a dual tyrosine kinase inhibitor of EGFR and VEGFR-2, inhibits MAPK/ERK and AKT/PI3-K and induces apoptosis in breast cancer cells. *Cancer Biology & Therapy* **2010,** 9, (8), 592-603.
80. Cho, S. J.; La, M. H.; Ahn, J. K.; Meadows, G. G.; Joe, C. O., Tob-mediated cross-talk between MARCKS phosphorylation and ErbB-2 activation. *Biochemical and Biophysical Research Communications* **2001,** 283, (2), 273-277.
81. Tiedemann, K.; Hussein, O.; Sadvakassova, G.; Guo, Y. B.; Siegel, P. M.; Komarova, S. V., Breast Cancer-derived Factors Stimulate Osteoclastogenesis through the Ca2+/Protein Kinase C and Transforming Growth Factor-beta/MAPK Signaling Pathways. *Journal of Biological Chemistry* **2009,** 284, (48), 33662-33670.
82. Lee, W. J.; Chen, W. K.; Wang, C. J.; Lin, W. L.; Tseng, T. H., Apigenin inhibits

HGF-promoted invasive growth and metastasis involving blocking PI3K/Akt pathway and beta 4 integrin function in MDA-MB-231 breast cancer cells. *Toxicology and Applied Pharmacology* **2008,** 226, (2), 178-191.

83. Liu, Z. M.; Adams, H. C.; Whitehead, I. P., The Rho-specific Guanine Nucleotide Exchange Factor Dbs Regulates Breast Cancer Cell Migration. *Journal of Biological Chemistry* **2009,** 284, (23), 15771-15780.

84. Stuart, E. C.; Jarvis, R. M.; Rosengren, R. J., In vitro mechanism of action for the cytotoxicity elicited by the combination of epigallocatechin gallate and raloxifene in MDA-MB-231 cells. *Oncology Reports* **2010,** 24, (3), 779-785.

85. Su, S.; Li, Y.; Luo, Y.; Sheng, Y.; Su, Y.; Padia, R. N.; Pan, Z. K.; Dong, Z.; Huang, S., Proteinase-activated receptor 2 expression in breast cancer and its role in breast cancer cell migration. *Oncogene* **2009,** 28, (34), 3047-3057.

86. Park, E. J.; Min, H. Y.; Chung, H. J.; Hong, J. Y.; Kang, Y. J.; Hung, T. M.; Youn, U. J.; Kim, Y. S.; Bae, K.; Kang, S. S.; Lee, S. K., Down-regulation of c-Src/EGFR-mediated signaling activation is involved in the honokiol-induced cell cycle arrest and apoptosis in MDA-MB-231 human breast cancer cells. *Cancer Letters* **2009,** 277, (2), 133-140.

87. Suy, S.; Mitchell, J. B.; Ehleiter, D.; Haimovitz-Friedman, A.; Kasid, U., Nitroxides tempol and tempo induce divergent signal transduction pathways in MDA-MB 231 breast cancer cells. *Journal of Biological Chemistry* **1998,** 273, (28), 17871-17878.

88. Sarcevic, B.; Suryadinata, R.; Sadowski, M.; Steel, R., Cyclin-dependent Kinase-mediated Phosphorylation of RBP1 and pRb Promotes Their Dissociation to Mediate Release of the SAP30 center dot mSin3 center dot HDAC Transcriptional Repressor Complex. *Journal of Biological Chemistry* **2011,** 286, (7), 5108-5118.

89. Oh, A. S.; Lahusen, J. T.; Chien, C. D.; Fereshteh, M. P.; Zhang, X. L.; Dakshanamurthy, S.; Xu, J. M.; Kagan, B. L.; Wellstein, A.; Riegel, A. T., Tyrosine Phosphorylation of the Nuclear Receptor Coactivator AIB1/SRC-3 Is Enhanced by Abl Kinase and Is Required for Its Activity in Cancer Cells. *Molecular and Cellular Biology* **2008,** 28, (21), 6580-6593.

90. Filardo, E. J.; Quinn, J. A.; Bland, K. I.; Frackelton, A. R., Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF. *Molecular Endocrinology* **2000,** 14, (10), 1649-1660.

91. Randi, A. S.; Pontillo, C. A.; Garcia, M. A.; Pena, D.; Cocca, C.; Chiappini, F.; Alvarez, L.; de Pisarev, D. K., Activation of c-Src/HER1/STAT5b and HER1/ERK1/2 Signaling Pathways and Cell Migration by Hexachlorobenzene in MDA-MB-231 Human Breast Cancer Cell Line. *Toxicological Sciences* **2011,** 120, (2), 284-296.

92. Yeh, Y. T.; Fu, O. Y.; Chen, I. F.; Yang, S. F.; Wang, Y. Y.; Chuang, H. Y.; Su, J. H.; Hou, M. F.; Yuan, S. S. F., STAT3 ser727 phosphorylation and its association with negative estrogen receptor status in breast infiltrating ductal carcinoma. *International Journal of Cancer* **2006,** 118, (12), 2943-2947.

93. Nelson, A. C.; Lyons, T. R.; Young, C. D.; Hansen, K. C.; Anderson, S. M.; Holt, J. T.,

AKT regulates BRCA1 stability in response to hormone signaling. *Molecular and Cellular Endocrinology* 319, (1-2), 129-142.
94. Miki, Y.; Swensen, J.; Shattuckeidens, D.; Futreal, P. A.; Harshman, K.; Tavtigian, S.; Liu, Q. Y.; Cochran, C.; Bennett, L. M.; Ding, W.; Bell, R.; Rosenthal, J.; Hussey, C.; Tran, T.; McClure, M.; Frye, C.; Hattier, T.; Phelps, R.; Haugenstrano, A.; Katcher, H.; Yakumo, K.; Gholami, Z.; Shaffer, D.; Stone, S.; Bayer, S.; Wray, C.; Bogden, R.; Dayananth, P.; Ward, J.; Tonin, P.; Narod, S.; Bristow, P. K.; Norris, F. H.; Helvering, L.; Morrison, P.; Rosteck, P.; Lai, M.; Barrett, J. C.; Lewis, C.; Neuhausen, S.; Cannonalbright, L.; Goldgar, D.; Wiseman, R.; Kamb, A.; Skolnick, M. H., A Strong Candidate for the Breast and Ovarian-cancer Susceptibility Gene BRCA1. *Science* **1994,** 266, (5182), 66-71.
95. Srinivasan, D.; Sims, J. T.; Plattner, R., Aggressive breast cancer cells are dependent on activated Abl kinases for proliferation, anchorage-independent growth and survival. *Oncogene* **2008,** 27, (8), 1095-1105.
96. Kramer, A.; Mailand, N.; Lukas, C.; Syljuasen, R. G.; Wilkinson, C. J.; Nigg, E. A.; Bartek, J.; Lukas, J., Centrosome-associated Chk1 prevents premature activation of cyclin-B-Cdk1 kinase. *Nature Cell Biology* **2004,** 6, (9), 884-U71.
97. Sachdev, S.; Bu, Y.; Gelman, I. H., Paxillin-Y118 phosphorylation contributes to the control of Src-induced anchorage-independent growth by FAK and adhesion. *Bmc Cancer* **2009,** 9, 14.
98. De Amicis, F.; Giordano, F.; Vivacqua, A.; Pellegrino, M.; Panno, M. L.; Tramontano, D.; Fuqua, S. A. W.; Ando, S., Resveratrol, through NF-Y/p53/Sin3/HDAC1 complex phosphorylation, inhibits estrogen receptor {alpha} gene expression via p38MAPK/CK2 signaling in human breast cancer cells. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 25, (10), 3695-707.
99. Treff, N. R.; Pouchnik, D.; Dement, G. A.; Britt, R. L.; Reeves, R., High-mobility group A1a protein regulates Ras/ERK signaling in MCF-7 human breast cancer cells. *Oncogene* **2004,** 23, (3), 777-785.
100. Bogdanova, N.; Feshchenko, S.; Schurmann, P.; Waltes, R.; Wieland, B.; Hillemanns, P.; Rogov, Y. I.; Dammann, O.; Bremer, M.; Karstens, J. H.; Sohn, C.; Varon, R.; Dork, T., Nijmegen BREAKAGE SYNDROME mutations and risk of breast cancer. *International Journal of Cancer* **2008,** 122, (4), 802-806.
101. Hwang, S.-K.; Kim, H.-H., The functions of mTOR in ischemic diseases. *Bmb Reports* 44, (8), 506-511.
102. Adelaide, J.; Finetti, P.; Bekhouche, I.; Repellini, L.; Geneix, J.; Sircoulomb, F.; Jauffret, E. C.; Cervera, N.; Desplans, J.; Parzy, D.; Schoenmakers, E.; Viens, P.; Jacquemier, J.; Birnbaum, D.; Bertucci, F.; Chaffanet, M., Integrated profiling of basal and luminal breast cancers. *Cancer Research* **2007,** 67, (24), 11565-11575.
103. Amazit, L.; Pasini, L.; Szafran, A. T.; Berno, V.; Wu, R. C.; Mielke, M.; Jones, E. D.; Mancini, M. G.; Hinojos, C. A.; O'Malley, B. W.; Mancini, M. A., Regulation of SRC-3 intercompartmental dynamics by estrogen receptor and phosphorylation. *Molecular and Cellular Biology* **2007,** 27, (19), 6913-6932.

Chapter 7

Large-scale Phosphoproteome Profiling of Human Breast Cancer Tissues

7.1 Introduction

Breast cancer is one of the most common public health problems in the western world, causing more than 40,000 death in the north American every year.[1] As a high-throughput technology, gene expression microarray has identified many potential breast cancer biomarkers. However, gene expression profiling cannot predict protein dynamics and the regulation of post-translational modifications (e.g., phosphorylation, glycosylation) which eventually determine the behavior of a cell. Therefore, despite the achievements made in genetic study of breast cancer,[2-7] information on the proteome is still believed to be the missing link to translate molecular knowledge into clinical screening and therapies.[8]

Discovering and investigating breast cancer biomarkers can be done from various sample sources, such as breast cancer cell lines, serum/plasma, biopsies, and tumor tissues.[9-14] As the most accessible sample source, breast cancer cell lines can provide proteomic information potentially related to the breast cancer. However, since each cell line was developed from individual patient, and mutations may be introduced after generations of culture process, cell lines can not accurately reflect the microenvironment of tumor development. Therefore, to obtain direct knowledge of breast cancer related proteins, direct investigation of tumor tissue samples becomes important.

In the tissue samples, the behavior of cells and proteins can be influenced extensively by a variety of modifications that take place on the expressed proteins. As one of the most important post-translational modifications, phosphorylation plays critical roles in cell signaling, protein-protein interactions, protein function activation, cell proliferation and apoptosis. In the past few decades, numerous efforts have been devoted into the phosphoproteome study of tissue samples.[15-23] Among many technologies and strategies, mass spectrometry (MS)-based phosphoproteome study has evolved into a powerful tool to characterize the global phosphoproteomic dynamics of tissue samples. For example, recently, Han et al.[24] applied strong anion exchange (SAX) chromatography for phosphopeptide fractionation and enrichment. Combined with the Fe^{3+} immobilized metal affinity chromatography (Fe^{3+}-IMAC), 274 phosphorylation sites from 305 unique

phosphopeptides corresponding to 168 proteins were identified from the human liver tissue sample. More recently, to study insulin signaling, Monetti et al.[25] reported a 'spike-in' stable-isotope labeling with amino acids in cell culture (SILAC) methodology using labeled mouse liver cell lines as the internal standards. Using this approach they identified 15,000 phospho-sites and quantitatively compared 10,000 sites in response to insulin treatment. Despite the achievements on the phosphoproteome study of human or mouse tissue samples, unfortunately, there have been very few publications on the large-scale analysis of the human breast cancer tissue phosphoproteome.

In this work, the phospoproteome of human ductal breast cancer tissue samples were analyzed. Proteins were extracted from the Optimal Cutting Temperature (O.C.T.)-medium embedded fresh frozen tissue, followed by tryptic digestion. Sequential enrichment of the phosphopeptides by IMAC and TiO_2 beads was utilized to increase the phosphoproteome coverage. Enriched phosphopeptides were subjected to strong cation exchange (SCX) fractionation, followed by capillary reversed-phase LC-MS/MS analysis. Seven cases of human breast tissues have been analyzed, including one case of stage III B, one case of stage II A, one case of stage I A and four cases of normal tissues obtained from reductive surgery. A total number of 1624 phosphoproteins with 3606 phosphopeptides corresponding to 3135 phospho-sites were detected from the seven cases of tissue samples. Among these phosphoproteins, 875 were identified as the potential biomarkers related to tumor genesis, and 297 related to tumor invasiveness (metastasis). The majority of these phosphoproteins have never been previously reported. They can serve as the candidates for further biological study, providing novel insights into the human breast cancer.

7.2 Experimental

7.2.1 Chemicals and reagents

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) and were of analytical grade. Phosphoric acid (H_3PO_4), potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4) and ammonium bicarbonate (NH_4HCO_3) were purchased from EMD Chemical Inc. (Mississauga, ON, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA). Sequencing grade modified trypsin, LC/MS-grade water, acetone, formic acid, and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, Canada).

7.2.2 Protein extraction from uuman breast cancer tissue samples

Seven cases of O.C.T-embedded frozen human breast tumor tissue and human normal breast tissue samples were obtained from the Alberta Cancer Research Tumor Bank (Cross Cancer Institute, Edmonton, AB, Canada). Informed consents were obtained from the patients for banking and use of tissues for research. This study was approved by the institutional ethics board of the University of Alberta. The clinical information about the patients is summarized in Table 7.1

Table 7.1 Clinical information of the patient samples.

Patient ID	Days since last update	Age	Gender	Deceased	Invasive type	HER2 status	tumor size (cm)
GT388	373	50	F	n	Ductal	Neg	2
MT1355	370	78	F	n	Ductal	Neg	3
MT324	0	72	F	y	Ductal	Pos	10

Fresh frozen tissue samples were slowly thawed on ice. O.C.T. and blood were removed by a simple rinse of the frozen tissue in cold 70% ethanol solution followed by rinse with phosphoate buffered saline (PBS) five times. The tissue samples were then homogenized by a glass pestle in 6 volumes of iced cold radioimmunoprecipitation assay (RIPA) buffer followed by 20-min incubation on ice. The lysed samples were centrifuged for 20 min at 15,000 g to pellet the tissue debris. The supernatants were then transferred to new tubes. To remove the salts and other impurities, acetone was added into the cell lysates with vortexing to final acetone concentration of 80%. The mixtures were incubated at -20 °C overnight, and were then centrifuged at 20,000 g for 10 minutes. After decanting the supernatant, the pellets were dissolved with 8 M urea with intermittent vortexing.

7.2.3 In-solution digestion

The protein mixtures were individually reduced with 10 mM dithiothreitol (DTT) for 1 h at 37 °C, and then alkylated with 20 mM iodoacetamide (IAA) for 1 h at room temperature in dark. Trypsin solution was added into the protein samples at a ratio of 1:100 after 10-fold dilution. The mixtures were incubated at 37 °C overnight. Ten percent trifluoroacetic acid (TFA) was added to the peptide mixtures to make the final pH ~2.5. The mixtures were

then centrifuged at 20,000 g for 10 min and the supernatant of each sample was transferred into a fresh vial for further analysis.

7.2.4 Sequential phosphopeptides enrichment by immobilized metal ion affinity chromatography (IMAC) and titanium dioxide

Prior to IMAC, the peptide samples were desalted by using Agilent 1100 HPLC with 4.6 × 50 mm C 18 column (Varian, Ontario, Canada), and then lyophilized. Sequential phosphopeptide enrichment was performed as descried previously[26] with modifications. After re-solubilization with 30% acetonitrile (ACN) mixed with 250 mM acetic acid, the peptides were loaded onto pre-equilibrated Fe-IMAC resin (Phos-Select iron affinity gel; Sigma, Ontario, Canada). The resin was washed three times with 30% acetonitrile (ACN) mixed with 250 mM acetic acid, two times with water after overnight incubation at 4 °C. The IMAC flow-through and washing buffer were collected for further analysis by TiO_2 beads (see below). The phosphopeptides were then released from the resin with 400 mM ammonium hydroxide. After lyophilization, the pooled flow-through and wash from the IMAC was enriched for phosphopeptides using TiO_2 beads (Titansphere, 5 μm) (GL-Sciences, Inc. Japan). The enrichment procedure has been described elsewhere[26] with some modifications. In brief, the lyophilized sample was re-solubilized in 200 μL loading buffer (65% ACN/2% TFA/saturated by phithalic acid), and then incubated with pre-equilibrated TiO_2 beads. The incubated beads were then washed with 800 μL wash buffer I (65% ACN/0.5% TFA) and wash buffer II (65% ACN/0.1% TFA). The phosphopeptides were eluted once with 200 μL elution buffer I (300 mM NH_4OH/50% ACN) and twice with 200 μL elution buffer II (500 mM NH_4OH/60% ACN). All the incubation, washing, as well as elution procedures were processed at room temperature for 20 minutes.

7.2.5 Strong cation exchange (SCX) liquid chromatography

A 2.1 x 250 mm poly SULFOETHYLTM A column (5-μdiameter, 300-Å pore) (Poly LC, Columbia, MD) was used for the SCX separation of the phosphopeptide-enriched mixture. Solvent A (5 mM KH_2PO_4, 20% ACN, pH 2.7) and solvent B (solvent A with 500 mM KCl) were used to develop a salt gradient (0-7% for 5 minutes, 7-42% for 35 minutes and 42-100% for 2 minutes). All the samples collected from SCX were desalted on an Agilent 1100 HPLC system and quantified by the UV absorbance.[27]

7.2.6 Mass spectrometric analysis

Tandem mass spectrometry analysis was performed as previously described.[28] Approximately 1.0 μg of peptides from individual fraction were analyzed using a QTOF Premier mass spectrometer (Waters, Milford, MA). Peptide solution was injected onto a 75 × 100 mm Atlantic dC18 column (Waters, Milford, MA). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Peptides were separated using a 120-minute gradient (2 to 7% B for 2 min, 7 to 8% B for 14 min, 8 to 20% B for 69 min, 20 to 30% B for 25 min, 30 to 45% B for 5 min, and 45 to 90% B for 5 min) and electrosprayed into the mass spectrometer fitted with a nanoSpray source at a flow rate of 350 nL/min. Mass spectra were acquired from m/z 300-1600 for 0.8 second, followed by 6 data-dependent MS/MS scans from m/z 50-1990 for 1.0 second each. The collision energy used to perform MS/MS was automatically varied according to the mass the charge state of the eluting peptide. A mass calibrant (i.e., lock-mass) was infused at a flow rate of 300 nL/min, and an MS scan was acquired for 1 second every 1 minute throughout the run. Precursor ion exclusion (PIE) strategy was applied to exclude relatively high-abundance peptides identified from the adjacent two SCX/mRP-C18 fractions to enable additional less abundance peptides to be analyzed and identified. An exclusion list was generated as described previously.[28]

7.2.7 Database search and data analysis

Raw MS/MS data were lock-mass-corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.3 (Waters, Milford, MA). Protein identification was performed by using Mascot 2.2 search engine (http://www.matrixscience.com/) for searching the Swiss-Prot database (http://ca.expasy.org/sprot/) (Swiss-Prot database; version 57.4; 410518 sequences). Searching was restricted *Homo sapiens* (20401 sequences) and preformed using the following parameters: fixed modification, carbamidomethyl (C); variable modifications, oxidation (M), phosphorylation on serine, threonine or tyrosine; enzyme, trypsin; missed cleavages, 2; peptide tolerance, 30 ppm; MS/MS tolerance, 0.2 Da; peptide charge, 1+, 2+ and 3+.

The search results, including protein names, access IDs, molecular mass, unique peptide sequences, ion score, and Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses, were extracted to Excel files using in-house software. All the identified peptides with scores lower than the Mascot threshold score for identity at the confidence level of 95% were then removed from the protein list. The redundant peptides for different protein identities were deleted, and the redundant proteins identified under the same gene name but

different access ID numbers were also removed from the list. The false positive rate was determined by employing the target-decoy search strategy.[29] The overall false positive rate was 1.5%.

A 79.96 Da mass increase found on a serine, threonine or tyrosine residue indicates phosphorylation of the residue. The spectra of all phosphoproteins that were found to be potentially related to the genesis and metastasis of breast cancer were manually checked based on the following rules. First, the y and/or b ions of serine and threonine phosphorylated peptides can be subjected to the neutral losses of H_3PO_4. An 18 Da mass difference between serine and dehydroalanine or threonine and dehydroaminobutyric acid was also considered. Second, as neutral loss is not normally found in the fragmentation of phosphotyrosine, immonium ion at the m/z value of 216 indicates the present of phosphotyrosine. Additionally, the cognate pair of fragment ions (y, y-79.96, and/or b, b-79.96) and y and/or b ions contaning a phosphate group were also considered. Finally, continuous fragment ion series (y and/or b ions) around the phosphorylated residue were investigated, which can further confirm the localization of a phosphorylation site.

All peptides and proteins identified were examined using the ProtParam program, available at the EXPASY Web site (http://us.expasy.org/tools/protparam.html), which allows calculation of the grand average of hydropathy (GRAVY).[30]

Functional pathway analysis was generated by the Ingenuity Pathway Analysis (IPA) software (www.ingenuity.com) (Ingenuity Systems, Redwood, CA). The significance of the association between the data set and the canonical pathway was measured in two ways. First, a ratio of the number of proteins from the data set that map to the pathway to the total number of proteins that map to the canonical pathway is calculated. Second, Fisher's exact test was used to calculate a p value representing the probability of the association between the proteins in the data set and the canonical pathway. A p-value less than 0.05 indicates a protein is not randomly associated with a canonical pathway.

7.3 Results and discussion

Proteins were extracted by RIPA buffer from three cases of human breast cancer tissue samples and four cases of normal human breast tissue samples obtained from reductive surgery. After purified by acetone, the protein samples were digested by trypsin, followed by sequential phosphopeptide enrichment by IMAC and TiO_2 beads. The enriched phosphopeptides were then subjected to 2D-LC MS/MS analysis by SCX and capillary RPLC. Each tissue sample, including normal tissue samples, was analyzed individually (Figure 7.1). The numbers of phosphoproteins and peptides identified from

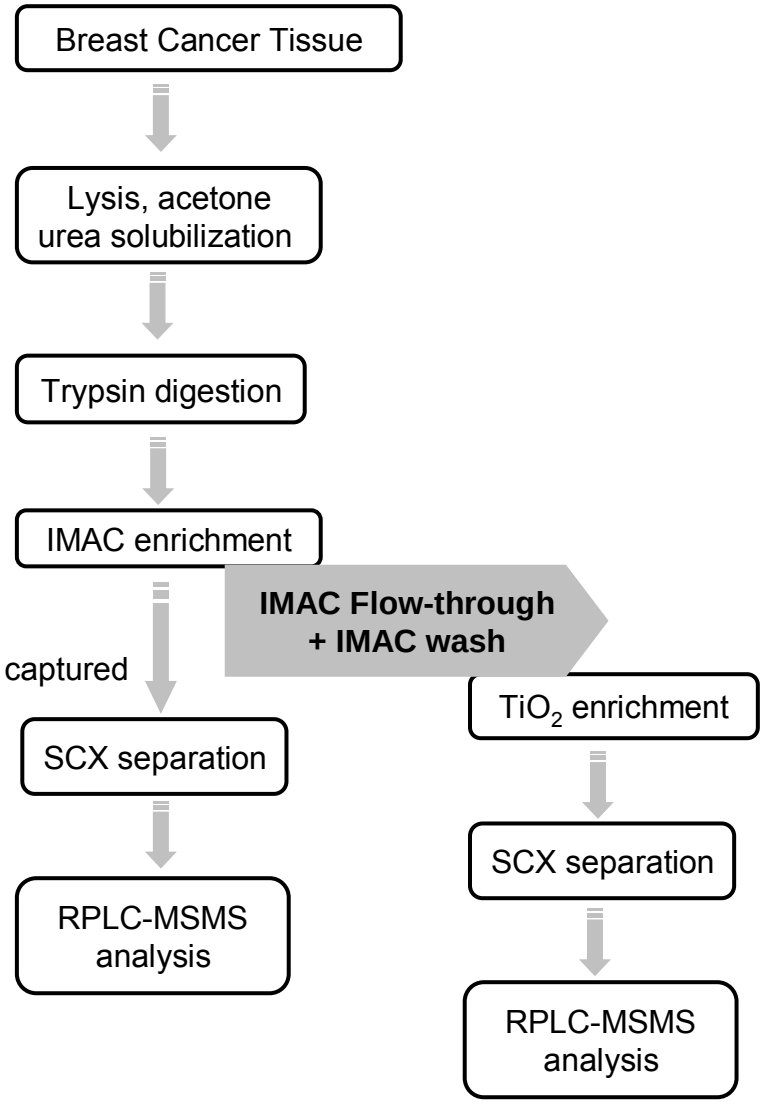


Figure 7.1 Workflow of analyzing breast cancer tissue samples.

each tissue sample are summarized in Table 7.2. As shown in Table 7.2, there were 1154 unique phosphoproteins with 2409 different phosphopeptides detected from the stage III B tissue, 1259 unique phosphoproteins with 2398 different phosphopeptides identified from the stage II A tissue, 149 unique phosphoproteins and 201 phosphopeptides detected from the stage I A sample. Combining the identification results of four normal tissue samples, a total number of 378 phosphoproteins and 614 phosphopeptides were identified from the normal tissue samples.

Table 7.2 Summary of phosphoproteins and phosphopeptides identified from the human breast tissue samples.

	Number of phosphoproteins	Number of phosphopeptides
MT324 (Stage IIIB)	1154	2409
MT1355 (Stage IIA)	1259	2398
GT388 (Stage IA)	149	201
MT1395 (Normal)	164	237
MT1482 (Normal)	175	268
MT1490 (Normal)	287	421
MT1543 (Normal)	225	162
Normal_combined	378	614
total	1624	3606

To generate the list of phosphoproteins possibly related to the genesis and invasiveness of breast cancer tumor, two kinds of comparisons were performed. First, the phosphoproteins detected in the stage III B tissue that are unique from the normal tissue phosphoproteins were treated as the candidates related to tumor genesis. Second, the phosphoproteins identified in the stage III B tissue that are unique from the stage II A and stage I A were treated as the candidates related to tumor invasiveness. Tables 7.3 and 7.4 summarize the 875 and 297 phosphoproteins related to the genesis and invasiveness of breast cancer, respectively. The mass spectra of these phosphoproteins have all been checked manually to ensure the accurate assignments of the phosphorylation sites.

Using the Ingenuity Pathway Analysis (IPA) software, 875 and 297 phosphoproteins related to the genesis and invasiveness of breast cancer were examined to see if they could be grouped according to canonical functional cellular pathways. As illustrated in Figure 7.2, the pathways to which the highest number of phosphoproteins belong are cellular growth and proliferation (n=227), gene expression (n=209), cellular assembly and organization (n=167), cell death or apoptosis (n=140), cellular development (n=131) and cellular function and maintenance (n=125) for the tumor genesis related phosphoproteins. Figure 7.3 shows the pathways to which the highest number of

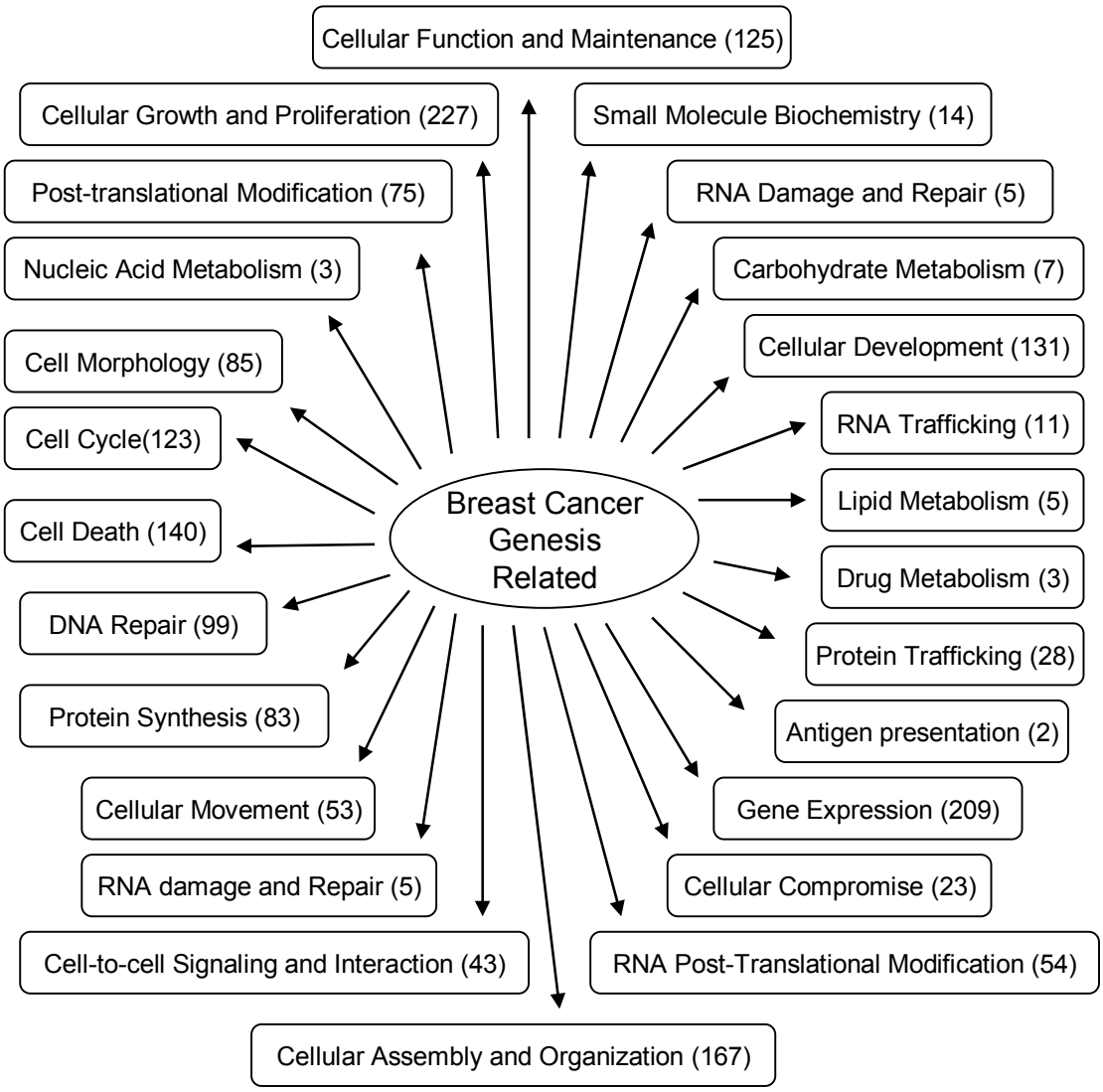


Figure 7.2 Functional pathway analysis of the tumor genesis related phosphoproteins.

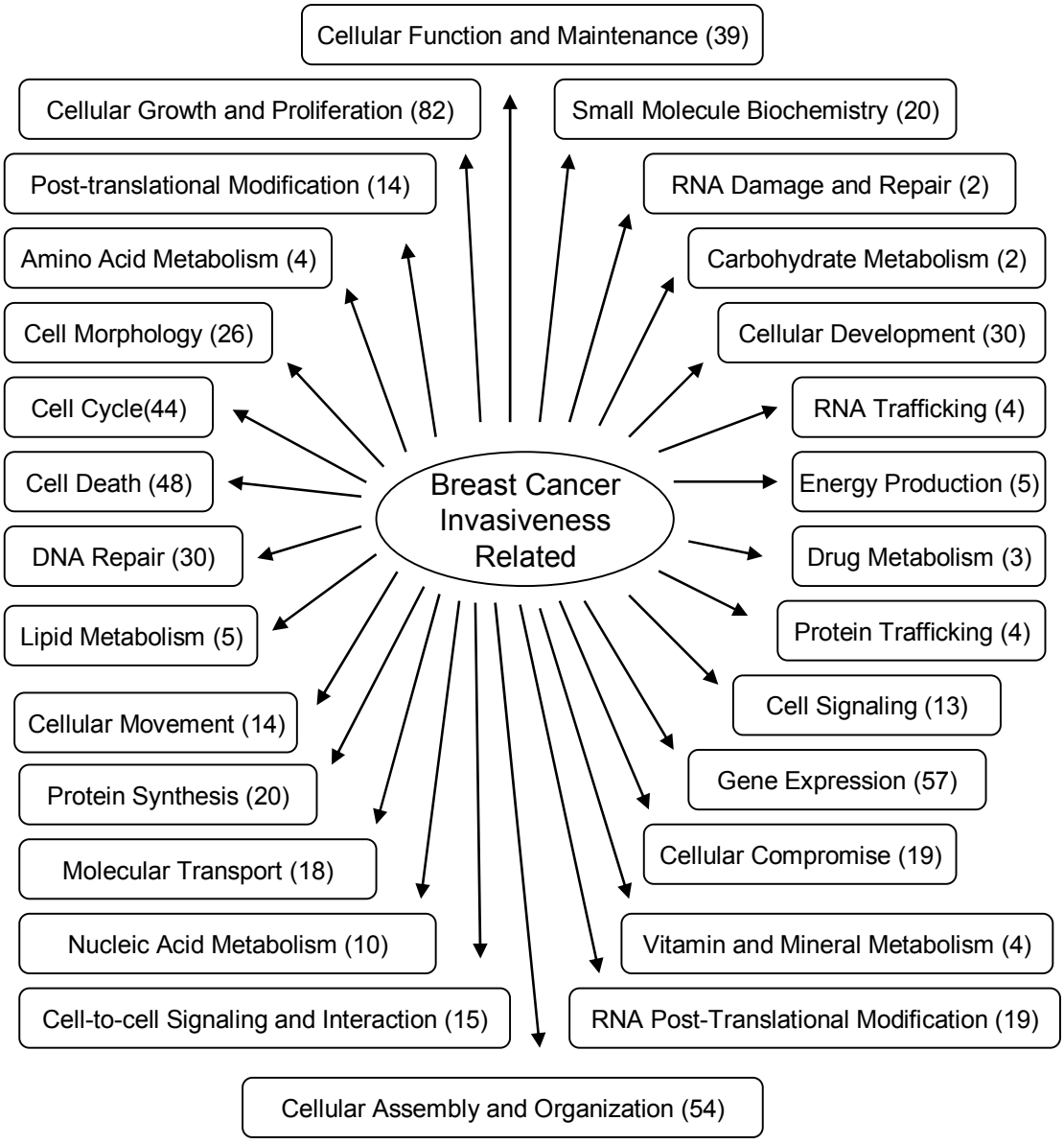


Figure 7.3 Functional pathway analysis of the tumor invasiveness related phosphoproteins.

phosphoproteins belong are cellular growth and proliferation (n=82), gene expression (n=57), cellular assembly and organization (n=54), cell death or apoptosis (n=48), cell cycle (n=44) and cellular Function and maintenance (n=39) for the tumor invasiveness related phosphoproteins.

As summarized in Tables 7.5 and 7.6, among the 875 and 297 candidates of tumor genesis and invasiveness biomarkers, a total of 19 tumor genesis related and 7 tumor invasiveness related phosphoproteins identified in this work were previously reported to be functionally involved in breast cancer. The biological processes they involve in include cell proliferation, cell cycle regulation, cell apoptosis related, tumor genesis, tumor metastasis, and drug sensitivity. In addition, there are 855 and 290 tumor genesis and invasiveness related phosphoproteins that have never been documented in the literature. For example, Figure 7.4 shows the MS/MS spectrum of a phosphoserine peptide generated from phosphorylated 14-3-3 protein epsilon. 14-3-3 protein epsilon is a protein of the 14-3-3 family of proteins which mediates signal transduction by binding to phosphoserine-containing proteins. It is involved in a large number of important cellular processes that include cell cycle progression, growth, differentiation, and apoptosis. It can bind to tuberin leading to increased cell growth.[31] Other than tuberin, more than 300 proteins have been reported to interact with 14-3-3 protein epsilon.[32] However, phosphorylated 14-3-3 protein epsilon has never been found in highly invasive human breast cancer tissues. The regulation mechanism of its phosphorylation is unclear. Therefore, further efforts are needed to study the biological importance and functional significance of these phosphoproteins.

7.4 Conclusions

In this study, we profiled the phosphoproteomes of seven cases of human breast tissues, including high and low metastasis breast cancer, early stage breast cancer, and normal tissues. By using sequential phosphopeptide enrichment by IMAC and TiO_2 beads combined with 2D-LC MS/MS, a total number of 1624 unique phosphoproteins with 3606 different phosphopeptides were identified from these tissue samples. Among these phosphoproteins, 875 and 297 phosphorylated proteins were found to be possibly related to the genesis and metastasis of human breast cancer, respectively. The majority of these candidates have never been reported in the breast cancer. Whether they are functionally involved in the genesis and development human breast cancer still need to be further explored. Therefore, future work will be focused on validation of the new phosphoprotein candidates to study their biological functions in breast cancer.

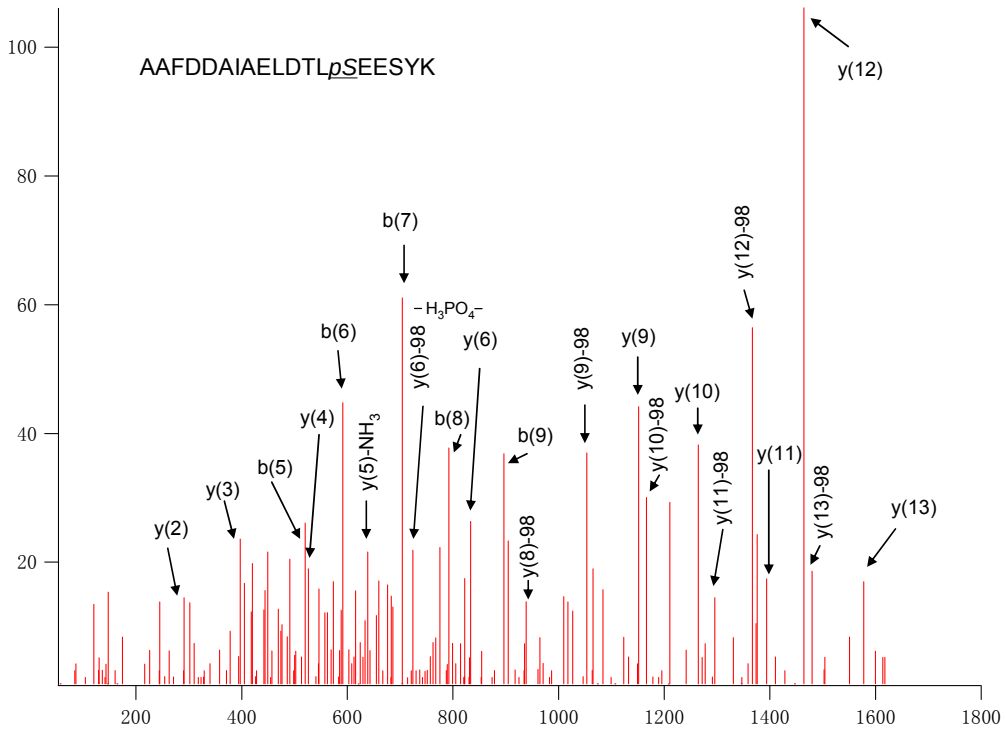


Figure 7.4 Electrospray ionization tandem mass spectrometry (ESI MS/MS) spectrum of a phosphoserine peptide generated from phosphorylated 14-3-3 protein epsilon (phospho-Serine[210]).

Table 7.3 Summary of the 875 phosphoproteins possibly related to the genesis of breast cancer identified in this work.

#	Name	Swiss-Prot ID	Unique Peptide Sequence	Score	Mascot score for identity
1	ZO3_HUMAN	O95049	SPEASQTDSPVESPR + 2 Phospho (ST)	40	28
2	ZO1_HUMAN	Q07157	VQIPVSRPDPEPVSDNEEDSYDEEIHDPR + Phospho (ST); Phospho (Y)	37	33
3	ZNF48_HUMAN	Q96MX3	SPSHSGEPFGLPGLEPEPGGPQAGEPPPPLAGDKPHK + Phospho (ST)	73	32
4	ZN828_HUMAN	Q96JM3	GQESSSDQEQVDVESIDFSK + 2 Phospho (ST)	44	29
			KPGPPLSPEIR + Phospho (ST)	31	28
			KPSPSESPEPWKPFPAVSPEPR + 2 Phospho (ST)	58	33
			KTSPASLDFPESQK + Phospho (ST)	70	31
			LAPVPSPEPQKPAPVSPESVK + 2 Phospho (ST)	38	31
			LAPVPSPEPQKPAPVSPESVK + Phospho (ST)	41	30
5	ZN687_HUMAN	Q8N1G0	ATDIPASASPPPVAGVPFFK + Phospho (ST)	48	31
			HGLQLGAQSPGR + Phospho (ST)	34	29
			TPLDLFAHFGPEPGDHSDPLPPSAPSPTR + Phospho (ST)	37	33
6	ZN652_HUMAN	Q9Y2D9	ESGSPYSVLVDTK + Phospho (ST)	34	30
7	ZN609_HUMAN	O15014	FCDSPTSDLEMR + Phospho (ST)	33	28
8	ZN446_HUMAN	Q9NWS9	TEEPLGSPHPSGTVESPGEGPQDTR + Phospho (ST)	41	32
9	ZN318_HUMAN	Q5VUA4	DISPEKSELDLGEPGPPGVEPPPQLLDIQCK + Phospho (ST)	34	32
10	ZN295_HUMAN	Q9ULJ3	ESEVCPVPTNSPSPPPLPPPPPLPK + 2 Phospho (ST)	44	33
			IQPLEPDSPTGLSENPTPATEK + Phospho (ST)	35	32
			SLSMDSQVPVYSPSIDLK + Phospho (ST)	59	32
			TEPSSPLSDPSDIIR + Phospho (ST)	32	31
11	ZN260_HUMAN	Q3ZCT1	TFSLKQNLVEHKK + Phospho (ST)	30	30

12	ZKSC1_HUMAN	P17029	EATGLSPQAAQEK + Phospho (ST)	57	30
13	ZHX3_HUMAN	Q9H4I2	VPEASSEPFDTSSPQAGR + Phospho (ST)	35	31
14	ZFY19_HUMAN	Q96K21	LPDSDDDEDEETAIQR + Phospho (ST)	61	30
15	ZFR_HUMAN	Q96KR1	RRDSDGVDGFEAEGK + Phospho (ST)	73	31
16	ZFP91_HUMAN	Q96JP5	RSSPSARPPDVPGQQPQAAK + Phospho (ST)	55	31
			SSPSARPPDVPGQQPQAAK + Phospho (ST)	56	31
17	ZF106_HUMAN	Q9H2Y7	AAHVPENSDTEQDVLTVKPVR + Phospho (ST)	66	31
18	ZEP2_HUMAN	P31629	MLKGISSSSLK + Oxidation (M); Phospho (ST)	30	29
19	ZDHC5_HUMAN	Q9C0B5	GDSLKEPTSIAESSR + Phospho (ST)	61	31
			LLRQSPPLPGREEEPGLGDSGIQSTPGSGHAPR + Phospho (ST)	57	32
20	ZCCHV_HUMAN	Q7Z2W4	SCTPSPDQISHR + 2 Phospho (ST)	46	27
			SCTPSPDQISHR + 3 Phospho (ST)	31	23
			SSLGSLQTPEAVTTR + Phospho (ST)	41	30
21	ZC3HE_HUMAN	Q6PJT7	ISPPIKEEETKGDSVEK + Phospho (ST)	53	31
22	ZC3HD_HUMAN	Q5T200	EVSPEVVR + Phospho (ST)	31	28
			SKGDSDISDEEAAQQSK + 2 Phospho (ST)	43	29
			SKLSPSPSLR + 2 Phospho (ST)	30	29
			TPSPPPPIPEDIALGK + 2 Phospho (ST)	65	31
			TSAVSSPLLDQQR + Phospho (ST)	33	30
23	ZC3H4_HUMAN	Q9UPT8	AAKPGPAEAPSPTASPSGDASPPATAPYDPR + 2 Phospho (ST)	40	33
			EYSPPYAPSHQQYPPSHATPLPK + Phospho (ST)	66	32
			TGSGSPFAGNSPAR + Phospho (ST)	50	29
			YREYSPPYAPSHQQYPPSHATPLPK + Phospho (ST)	73	33
24	ZC3H1_HUMAN	O60293	KPISDNSFSSDEEQSTGPIK + 2 Phospho (ST)	67	31
25	ZC11A_HUMAN	O75152	RLSSASTGKPPLSVEDDFEK + Phospho (ST)	42	32

26	ZBTB1_HUMAN	Q9Y2K1	VTDKDCNESTDNDELEDEPEEPFYR + 2 Phospho (ST)	32	29
27	ZBT7A_HUMAN	O95365	GGAPDPSPGATATPGAPAQPSSPDAR + Phospho (ST)	47	32
			HFKDEDEDEDVASPDGLGR + Phospho (ST)	82	30
28	YYAP1_HUMAN	Q9H869	LEPQELSPLSATVFPK + Phospho (ST)	31	30
29	YTDC2_HUMAN	Q9H6S0	VDGIPNDSSDSEMEDK + Phospho (ST)	39	29
30	YTDC1_HUMAN	Q96MU7	GISPIVFDR + Phospho (ST)	54	29
			LSSESHHGGSPIHWVLPAGMSAK + Phospho (ST)	43	32
31	YJ005_HUMAN	Q6ZSR9	KKSMEELTVIQCTSQELPAQTGLLSQTGDVPLPAGR + Phospho (ST)	43	31
32	YETS2_HUMAN	Q9ULM3	IVSGSPISTPSPSPLPR + Phospho (ST)	36	30
33	YD003_HUMAN	Q9NXL2	RKTDTVVESSVSGDHSGTLR + Phospho (ST)	107	31
34	YBOX1_HUMAN	P67809	AADPPAENSSAPEAEQGGAE + Phospho (ST)	57	29
			NEGSESAPEGQAQQR + Phospho (ST)	33	29
35	XRCC1_HUMAN	P18887	TKPTQAAGPSSPQKPPTPEETK + 2 Phospho (ST)	35	32
36	XPO6_HUMAN	Q96QU8	HSVTAATPPPSPTSGESGDLLSNLLQSPSSAK + 2 Phospho (ST)	53	33
37	XIRP2_HUMAN	A4UGR9	DGSGQMLEIK + Phospho (ST)	30	27
38	WWTR1_HUMAN	Q9GZV5	SHSSPASLQLGTGAGAAGSPAQQHAHLR + Phospho (ST)	95	32
39	WRN_HUMAN	Q14191	ICALTK + Phospho (ST)	30	28
40	WNK1_HUMAN	Q9H4A3	DVDDGSGSPHSPHQLSSK + 3 Phospho (ST)	54	27
			DVDDGSGSPHSPHQLSSK + Phospho (ST)	54	30
			KEKPELSEPSHLNGPSSDPEAAFLSR + Phospho (ST)	35	32
41	WIPF1_HUMAN	O43516	NLSLSSSTPPLPSPGR + Phospho (ST)	35	31
42	WDR75_HUMAN	Q8IWA0	EIPEDVDMEEEKESEDSDEENDFTEK + 2 Phospho (ST)	30	29
43	WDR55_HUMAN	Q9H6Y2	TCEERPAEDGSDEEDPDSMEAPTR + Phospho (ST)	67	29
44	WDR44_HUMAN	Q5JSH3	EYVSNDAAQSDDEEKLQSQPTDTDGGR + Phospho (ST)	82	32

			LTQTSSTEQLNVLETETEVLNK + Phospho (ST)	86	32
			RKSELEFETLK + Phospho (ST)	38	30
45	WD42A_HUMAN	Q5TAQ9	VHDRSEEEEEEEEEEEEQPR + Phospho (ST)	42	30
46	WBP11_HUMAN	Q9Y2W2	RRDEDMLYSPELAQR + Phospho (ST)	49	31
47	VPRBP_HUMAN	Q9Y4B6	HLPSPPTLDSIITEYLR + Phospho (ST)	44	30
48	VP13D_HUMAN	Q5THJ4	EVQDKDYPLTPPPSPTVDEPK + 2 Phospho (ST)	33	32
			EYLSQSCPSVSNVEYPDMPR + Phospho (ST)	85	32
			VHTSGFGYQSELELR + Phospho (ST)	35	31
49	VP13B_HUMAN	Q7Z7G8	KEDEVSIGSAPLAK + 2 Phospho (ST)	30	30
50	VIR_HUMAN	Q69YN4	SFLSEPSSPGR + Phospho (ST)	44	29
			VISHDRDSPPPPPPPPPPQPQPSLK + Phospho (ST)	61	31
51	VIP2_HUMAN	O43314	HFFHHADEDDEEEDDSPPER + Phospho (ST)	117	29
52	VDAC2_HUMAN	P45880	LTFDTTFSPNTGK + Phospho (ST)	43	30
53	VDAC1_HUMAN	P21796	LTFDSSFSPNTGK + Phospho (ST)	70	30
			LTFDSSFSPNTGKK + Phospho (ST)	44	31
54	VASP_HUMAN	P50552	MKSSSSVTTSETQPCTPSSSDYSDLQR + Phospho (ST)	55	32
55	VAMP4_HUMAN	O75379	NLLEDDSDEEEDFFLR + Phospho (ST)	118	30
56	UTP18_HUMAN	Q9Y5J1	TSSDDESEEDEDDLLQR + 2 Phospho (ST)	51	26
			VQEHEDSGDSEVENEAK + 2 Phospho (ST)	31	27
57	UT14A_HUMAN	Q9BVJ6	DYLLSESEDEGDNDGER + 2 Phospho (ST)	40	26
58	USO1_HUMAN	O60763	DLGHPVEEEDELESGDQEDEDDESEDPGK + Phospho (ST)	70	29
			DLGHPVEEEDELESGDQEDEDDESEDPGKDLDHI + Phospho (ST)	82	31
			LKDLGHPVEEEDELESGDQEDEDDESEDPGKDLDHI + Phospho (ST)	78	32
59	URFB1_HUMAN	Q6BDS2	TVSQQSFDGVSLDSSGPEDR + 2 Phospho (ST)	49	29
60	UNK_HUMAN	Q9C0B0	NSSLGSPSNLCGSPPGSIR + Phospho (ST)	64	31

61	UNG_HUMAN	P13051	HAPSPEPAVQGTGVAGVPEESGDAAAIPAK + Phospho (ST)	83	32
			KAPAGQEEPGTPPSSPLSAEQLDR + 2 Phospho (ST)	40	32
62	UNC5B_HUMAN	Q8IZJ1	SASLGSQQLLGLPR + Phospho (ST)	52	30
63	ULK1_HUMAN	O75385	LHSAPNLSDLHVVRPK + Phospho (ST)	55	30
			TPSSQNLLALLAR + Phospho (ST)	49	29
64	UGDH_HUMAN	O60701	IPYAPSGEIPK + Phospho (ST)	33	30
65	UFO_HUMAN	P30530	ESFPAPVVILPFMK + Phospho (ST)	38	30
66	UBXN7_HUMAN	O94888	SESLIDASEDSQLEAAIR + Phospho (ST)	65	31
67	UBR5_HUMAN	O95071	WLDGASFDNER + Phospho (ST)	41	28
68	UBR4_HUMAN	Q5T4S7	AAPPPPPPPPPLESSPR + Phospho (ST)	49	30
			HVTLPSSPR + Phospho (ST)	40	29
			SNTPMGDKDDDDDDDADEK + Phospho (ST)	126	26
			TGSTSSKEDDYESDAATIVQK + Phospho (ST)	59	32
			TSPADHGGSVGSESGGSAVDSVAGEHSVSGR + Phospho (ST)	82	32
69	UBP8_HUMAN	P40818	SYSSPDITQAIQEEEK + Phospho (ST)	58	31
			SYSSPDITQAIQEEEKR + Phospho (ST)	139	31
70	UBP40_HUMAN	Q9NVE5	RKSQEALHEQSSYILSSAETPARPR + Phospho (ST)	65	32
71	UBP2L_HUMAN	Q14157	RYPSSISSSPQK + Phospho (ST)	53	30
			SPAVATSTAAPPPPSSPLPSK + Phospho (ST)	74	31
			STSAPQMSPGSSDNQSSSPQPAQQK + Oxidation (M); 2 Phospho (ST)	33	30
			YPSSISSSPQK + Phospho (ST)	33	29
72	UBP20_HUMAN	Q9Y2K6	AVPIAVADEGESESEDDDLKPR + 2 Phospho (ST)	72	32
73	UBP16_HUMAN	Q9Y5T5	NINMDNDLEVLTSSPTR + Oxidation (M); Phospho (ST)	61	31
74	UBP10_HUMAN	Q14694	TCNSPQNSTDSVSDIVPDSPFPGALGSDTR + Phospho (ST)	39	33
75	UBE2O_HUMAN	Q9C0C9	LIHGEDSDSEGEEEGR + 2 Phospho (ST)	87	27

			NMTVEQLLTGSPTSPTVEPEKPTR + 2 Phospho (ST)	33	33
76	UB2J1_HUMAN	Q9Y385	RLSTSPDVIQGHQPR + Phospho (ST)	104	31
77	U3IP2_HUMAN	O43818	MNEEISSDSESESLAPR + 3 Phospho (ST)	30	26
78	TYDP1_HUMAN	Q9NUW8	WTISSSDESEEEKPKPDKPSTSSLLCAR + Phospho (ST)	64	33
79	TXND1_HUMAN	Q9H3N1	KVEEEQEADEEDVSEEEAESK + Phospho (ST)	117	31
			VEEEQEADEEDVSEEEAESKEGTNKDFPQNAIR + Phospho (ST)	33	33
80	TXLNA_HUMAN	P40222	SSPGQPEAGPEGAQERPSQAAPAVEAEGPGSSQAPR + Phospho (ST)	34	33
81	TUSC2_HUMAN	O75896	RGSMFYDEDGDLAHEFYEETIVTK + Phospho (ST)	122	32
82	TTC7B_HUMAN	Q86TV6	LPISSSTSNLHVDR + Phospho (ST)	59	30
83	TTC15_HUMAN	Q8WVT3	VRDEAEPGGEGDPGPEPAGTPSPSGEADGDCAPEDAAPSSGGAPR + 2 Phospho (ST)	35	31
84	TSYL1_HUMAN	Q9H0U9	YITNLEVKELR + Phospho (ST)	45	29
85	TSC2_HUMAN	P49815	SSSSPELQTLQDILGDPGDK + Phospho (ST)	120	32
86	TRUA_HUMAN	Q9Y606	VPSPLEGSEGDGDTD + Phospho (ST)	49	27
87	TRPS1_HUMAN	Q9UHF7	ATEETGQAQSGQANCQGLSPVSVASK + Phospho (ST)	44	32
			HPNYSPPGSPIEK + 2 Phospho (ST)	56	29
			SKTDLLVNDNPDPAPLSPELQDFK + Phospho (ST)	55	32
88	TRIPC_HUMAN	Q14669	SESPPAELPSLR + Phospho (ST)	44	30
			VREDDEDSDDDGSDEEIDESLAAQFLNSGNVR + 2 Phospho (ST)	72	31
			YSPPRDDDKVDNQAK + Phospho (ST)	45	31
89	TREF1_HUMAN	Q96PN7	SSPSHSTTSGETDPTTIFPCK + Phospho (ST)	32	31
90	TPC10_HUMAN	P48553	RQESSSSLEMPSGVALEEGAHVLR + Oxidation (M); 2 Phospho (ST)	63	33
			RQESSSSLEMPSGVALEEGAHVLR + Phospho (ST)	65	32
91	TP53B_HUMAN	Q12888	ECSEAMEVETSVISIDSPQK + Oxidation (M); Phospho (ST)	68	32
			ETAVPGPLGIEDISPNLSPDDK + Phospho (ST)	52	32

			GNLLHFPSSQGEEEKEK + Phospho (ST)	40	31
			GSGEKPVSAPGDDTESLHSQGEEEFDMPQPPHGHVLHR + Oxidation (M); 2 Phospho (ST)	64	32
			IDEDGENTQIEDTEPMSPVLNSK + Phospho (ST)	76	32
			LMLSTSEYSQSPK + Oxidation (M); Phospho (ST); Phospho (Y)	34	29
			LMLSTSEYSQSPK + Phospho (ST); Phospho (Y)	39	29
			MVIQGPSSPQGEAMVTDVLEDQK + 2 Oxidation (M); Phospho (ST)	87	32
			MVIQGPSSPQGEAMVTDVLEDQK + Oxidation (M); Phospho (ST)	43	32
			MVIQGPSSPQGEAMVTDVLEDQK + Phospho (ST)	104	32
			NSPEDLGLSLTGDSCK + Phospho (ST)	104	30
			SDSPEIPFQAAAGPSDGLDASSPGNSFVGLR + Phospho (ST)	125	33
			SEALSSVLDQEEAMEIK + 2 Phospho (ST)	55	31
			SEALSSVLDQEEAMEIK + Oxidation (M); 2 Phospho (ST)	47	30
			SGTAETEPVEQDSSQPSLPLVR + Phospho (ST)	36	32
			SKLPDGPTGSSEEEEEFLEIPPFNK + 2 Phospho (ST)	52	33
			SPEPEVLSTQEDLFDQSNK + Phospho (ST)	117	32
			STPFIVPSSPTEQEGR + Phospho (ST)	36	31
92	TOX4_HUMAN	O94842	LSTTPSPTSSLHEDGVEDFR + 2 Phospho (ST)	33	31
			LSTTPSPTSSLHEDGVEDFRR + 2 Phospho (ST)	36	32
93	TOP2B_HUMAN	Q02880	FDSNEEDSASVFSPSFGLK + Phospho (ST)	42	31
			KTSFDQDSDVDIFPSDFPTEPPSLPR + Phospho (ST)	79	33
			TSFDQDSDVDIFPSDFPTEPPSLPR + Phospho (ST)	35	33
			VVEAVNSDSDSEFGIPK + 2 Phospho (ST)	35	31
94	TOP2A_HUMAN	P11388	GSVPLSSSPPATHFPDETEITNPVPK + Phospho (ST)	97	32
			VPDEEENEESDNEKETEK + Phospho (ST)	81	29

95	TOM20_HUMAN	Q15388	IVSAQSLAEDDVE + Phospho (ST)	31	30
96	TOIP1_HUMAN	Q5JTV8	DSHSSEEDEASSQTDLSQTISK + 2 Phospho (ST)	34	30
			GLRDSHSSEEDEASSQTDLSQTISK + 3 Phospho (ST)	35	30
97	TOE1_HUMAN	Q96GM8	AADSDDGAVSAPAASDGGVSK + Phospho (ST)	37	31
98	TOB1_HUMAN	P50616	TSPINLGLNVNDLLK + Phospho (ST)	42	30
99	TMF1_HUMAN	P82094	SATPVNCEQPDILVSSTPINEGQTVLDK + Phospho (ST)	46	32
			SVSEINSDDELSGK + Phospho (ST)	31	30
100	TMCC1_HUMAN	O94876	ALGVISNFQSSPK + Phospho (ST)	37	30
			FGSADNIPNLK + Phospho (ST)	38	29
101	TM87B_HUMAN	Q96K49	WVEENIPSSFTDVALPVLVDSDEEIMTR + Oxidation (M); Phospho (ST)	53	33
			WVEENIPSSFTDVALPVLVDSDEEIMTR + Phospho (ST)	45	33
102	TM87A_HUMAN	Q8NBN3	WVEENVPSSVTDVALPALLDSDEER + Phospho (ST)	96	32
103	TM63B_HUMAN	Q5T3F8	LTSVSSSVDFDQR + Phospho (ST)	45	30
104	TM1L1_HUMAN	O75674	EATNTTSEPSAPSQDLLDLSPSPR + Phospho (ST)	48	33
105	TLE3_HUMAN	Q04726	DAPTSPASVASSSSTPSSK + Phospho (ST)	69	31
			ESSANNSVSPSESLR + Phospho (ST)	45	30
			YDSDGDKSDDLVVDVSNEDPATPR + 2 Phospho (ST)	60	31
			YDSDGDKSDDLVVDVSNEDPATPR + Phospho (ST); Phospho (Y)	58	30
			YDSDGDKSDDLVVDVSNEDPATPR + Phospho (Y)	45	32
106	TISD_HUMAN	P47974	LHHSLSFSGFPSGHHQPPGGLESPLLLDSPTSR + Phospho (ST)	55	32
107	TIM8A_HUMAN	O60220	SKPVFSESLSD + Phospho (ST)	41	29
108	TIF1A_HUMAN	O15164	NESEDNKFSDDSDDDFVQPR + 2 Phospho (ST)	38	28
			SEWLDPSQKSPLHVGETR + Phospho (ST)	32	32
109	TICN2_HUMAN	Q92563	QNGSASSVAGPASGLDK + Phospho (ST)	30	30
110	THOC5_HUMAN	Q13769	ALFKPPEDSQDDESDSDAEEEQTTK + 2 Phospho (ST)	91	30

111	THOC1_HUMAN	Q96FV9	TGEDEDEEDNDALLKENESPDVR + Phospho (ST)	66	31
112	TGS1_HUMAN	Q96RS0	DRPHASGTDGDESEEDPPEHKPSK + 2 Phospho (ST)	46	30
113	TFP11_HUMAN	Q9UBB9	EEATYGVWAERDSDDERPSFGGK + Phospho (ST)	41	32
			GAAEEAELEDSDDEEKPVKQDDFPK + Phospho (ST)	48	32
			KGAAEEAELEDSDDEEKPVKQDDFPK + Phospho (ST)	54	33
114	TFG_HUMAN	Q92734	NVMSAFGLTDDQVSGPPSAPAEDR + Oxidation (M); Phospho (ST)	93	32
			NVMSAFGLTDDQVSGPPSAPAEDR + Phospho (ST)	41	32
115	TF3C2_HUMAN	Q8WUA4	AHFNAMFQPSSPTR + Oxidation (M); Phospho (ST)	71	30
			AHFNAMFQPSSPTR + Phospho (ST)	56	30
116	TF3C1_HUMAN	Q12789	KNSSTDQGSDEEGSLQK + 2 Phospho (ST)	60	29
			NSSTDQGSDEEGSLQK + Phospho (ST)	58	29
			TSQPPVPQGEAEEDSQGKEGPSGSGDSQLSASSR + Phospho (ST)	34	33
117	TENS3_HUMAN	Q68CZ2	WDSYENLSADGEVLHTQGPVDGSLYAK + Phospho (Y)	79	33
118	TEBP_HUMAN	Q15185	DWEDDSDEDMSNFDR + Oxidation (M); Phospho (ST)	92	24
			DWEDDSDEDMSNFDR + Phospho (ST)	120	22
			FSEMMNNMGGDEDVDLPEVDGADDDSQDSDDEKMPDLE + 2 Phospho (ST)	39	28
			FSEMMNNMGGDEDVDLPEVDGADDDSQDSDDEKMPDLE + Oxidation (M); 2 Phospho (ST)	36	28
119	TE2IP_HUMAN	Q9NYB0	YLLGDAPVSPSSQK + Phospho (ST)	54	31
120	TCP4_HUMAN	P53999	EQISDIDDAVR + Phospho (ST)	31	30
121	TCOF_HUMAN	Q13428	LDSSPSVSSTLAAK + Phospho (ST)	82	30
			LGAGEGGEASVSPEK + Phospho (ST)	57	30
			TSQVGAASAPAKESPR + Phospho (ST)	70	31
			TVANLLSGKSPR + Phospho (ST)	43	29
122	TCEA3_HUMAN	O75764	GLECSDWKPEAGLSPPR + Phospho (ST)	37	31

123	TCEA1_HUMAN	P23193	EPAITSQNSPEAR + Phospho (ST)	63	30
124	TBD2B_HUMAN	Q9UPU7	DTSPDKGELVSDEEEDT + Phospho (ST)	38	29
125	TBCD5_HUMAN	Q92609	NISSSPSVESLPGGR + 2 Phospho (ST)	46	29
			NISSSPSVESLPGGR + Phospho (ST)	54	31
			SESMPVQLNK + Oxidation (M); Phospho (ST)	34	28
			SESMPVQLNK + Phospho (ST)	31	29
126	TBCD4_HUMAN	O60343	LGSVDSFER + 2 Phospho (ST)	36	26
			LGSVDSFER + Phospho (ST)	41	28
			SLTSSLENIFSR + Phospho (ST)	47	30
			TSSTCSNESLSVGGTSVTPR + 2 Phospho (ST)	97	30
			TSSTCSNESLSVGGTSVTPR + Phospho (ST)	83	31
127	TBC9B_HUMAN	Q66K14	KASVVDPSTESSPAPQEGSEQPASPASPLSSR + 2 Phospho (ST)	45	33
			KASVVDPSTESSPAPQEGSEQPASPASPLSSR + Phospho (ST)	161	33
128	TBC30_HUMAN	Q9Y2I9	RRDSLDSSTEASGSDVVLGGR + Phospho (ST)	66	32
129	TBC15_HUMAN	Q8TC07	NDSPTQIPVSSDVCR + Phospho (ST)	77	30
130	TBA1A_HUMAN	Q71U36	TIGGGDDSFNTFFSETGAGK + Phospho (ST)	75	31
131	TB22A_HUMAN	Q8WUA7	SQSLPHSATVTLGGTSDPSTLSSSALSER + Phospho (ST)	154	33
			SVSESHTSCPAESASDAAPLQR + Phospho (ST)	59	31
132	TANC2_HUMAN	Q9HCD6	SSSQLGSPDVSHLIR + Phospho (ST)	39	31
			YQQETSVSQLPGRPKSPLSK + Phospho (ST)	43	31
133	TAGL2_HUMAN	P37802	NFSDNQLQEGK + Phospho (ST)	66	28
134	TAF12_HUMAN	Q16514	LSPENNQVLTK + Phospho (ST)	34	30
135	TACC2_HUMAN	O95359	SSDSEEAFETPESTTPVK + Phospho (ST)	60	30
136	TAC2N_HUMAN	Q8N9U0	KVELSSSSQHGPSYDVYNPFYMYQHISPDLSR + Phospho (ST)	103	33
			LYGSVCDLR + Phospho (ST)	51	28

			SLDTITLSGDER + Phospho (ST)	70	30
137	T2FA_HUMAN	P35269	IHDLEDDLEMSSDASDASGEEGGR + Oxidation (M); 2 Phospho (ST)	58	28
138	SYNJ1_HUMAN	O43426	SSPNPFITGLTR + Phospho (ST)	41	30
			TSPCQSPTISEGPVPSLPIRPSR + 2 Phospho (ST)	33	33
139	SYNG_HUMAN	Q9UMZ2	ETSFGSSENITMTSLSK + Oxidation (M); Phospho (ST)	80	31
140	SYNE2_HUMAN	Q8WXH0	RGSMSYLAAVEEEVEESSVK + Phospho (ST)	50	32
141	SYMC_HUMAN	P56192	TSPKPAVVETVTTAKPQQIQALMDEVTK + Phospho (ST)	93	30
142	SYIC_HUMAN	P41252	APLKPYPVSPSDK + Phospho (ST)	31	30
143	SYEP_HUMAN	P07814	EYIPGQPPLSQSSDSSPTR + Phospho (ST)	48	32
144	SVIL_HUMAN	O95425	EMEKSFDEQNVPK + Phospho (ST)	31	30
			YGSFEEAEASYPILNR + Phospho (ST)	89	31
145	STXB5_HUMAN	Q5T5C0	KLSLPTDLKPDLDVK + Phospho (ST)	42	29
146	STX16_HUMAN	O14662	SIAAELDELADDR + Phospho (ST)	38	30
147	STUB1_HUMAN	Q9UNE7	LGAGGGSPEKSPSAQELK + Phospho (ST)	42	31
148	STT3B_HUMAN	Q8TCJ2	ENPPVEDSSDEDDKR + 2 Phospho (ST)	44	27
			ENPPVEDSSDEDDKR + Phospho (ST)	77	29
149	STMN1_HUMAN	P16949	ASGQAFELILSPR + Phospho (ST)	65	30
			RASGQAFELILSPR + Phospho (ST)	31	29
150	STK39_HUMAN	Q9UEW8	TEDGDWEWSDDEMDEK + Phospho (ST)	66	26
151	STK11_HUMAN	Q15831	IDSTEVIYQPR + Phospho (ST)	37	30
152	STIM1_HUMAN	Q13586	AEQSLHDLQER + Phospho (ST)	62	30
			DLTHSDSESSLHMSDR + Phospho (ST)	70	30
153	STAT3_HUMAN	P40763	FICVTPTTCSNTIDLPMSPR + Phospho (ST)	38	32
154	STAT1_HUMAN	P42224	LQTTDNLLPMSPEEFDEVSR + Phospho (ST)	40	32
155	STAR3_HUMAN	Q14849	GPLLFSGALSEGQFYSPPESFAGSDNESDEEVAGKK + 2 Phospho (ST)	50	33

156	ST65G_HUMAN	O94864	TEESEPLPSCPGSPPLPDDLLPLDCK + Phospho (ST)	53	33
157	SSRP1_HUMAN	Q08945	EGMNPSYDEYADSDEDQHDAYLER + Phospho (ST)	103	30
158	SSRA_HUMAN	P43307	VEMGTSSQNDVDMSWIPQETLNQINKASPR + Phospho (ST)	37	33
159	SSH3_HUMAN	Q8TE77	RQSFAVLR + Phospho (ST)	36	28
			SPPGSGASTPVGPWDQAVQR + Phospho (ST)	160	32
160	SSFA2_HUMAN	P28290	SQSLPTTLLSPVR + Phospho (ST)	53	29
161	SSF1_HUMAN	Q9NQ55	VGGSDEEASGIPSR + Phospho (ST)	71	29
162	SSBP4_HUMAN	Q9BWG4	SSPGAVAGLSNAPGTPR + Phospho (ST)	68	31
163	SSBP3_HUMAN	Q9BWW4	NSPNNISGISNPPGTPR + Phospho (ST)	66	31
164	SRRT_HUMAN	Q9BXP5	TQLWASEPGTPPLPTSLPSQNPILK + Phospho (ST)	64	30
165	SRPR_HUMAN	P08240	GTGSGGQLQDLDCSSSDDEGAAQNSTKPSATK + 3 Phospho (ST)	50	30
166	SRPK2_HUMAN	P78362	TVSASSTGDLPK + 2 Phospho (ST)	31	28
167	SRP72_HUMAN	O76094	TVSSPPTSPRPGSAATVSASTSNIIPPR + 2 Phospho (ST)	86	32
168	SRF_HUMAN	P11831	ALIQTCLNSPDSPPR + Phospho (ST)	49	31
169	SRC_HUMAN	P12931	RRSLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST)	88	33
			SLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST)	120	33
170	SR140_HUMAN	O15042	SEEHHLYSNPIKEEMTESK + Phospho (ST)	34	31
171	SQSTM_HUMAN	Q13501	EVDPSTGELQSLQMPESEGPSSLDPSQEGPTGLK + Phospho (ST)	53	33
			RFSFCCSPEPEAEAEAAAGPGPCER + Phospho (ST)	86	31
			SRLTPVSPESSSTEEK + 2 Phospho (ST)	54	31
			SSSQPSSCCSDPSKPGGNVEGATQSLAEQMR + Oxidation (M); Phospho (ST)	47	32
172	SQRD_HUMAN	Q9Y6N5	RYPNVFGIGDCTNLPTSK + Phospho (ST)	55	32
			YPNVFGIGDCTNLPTSK + Phospho (ST)	40	31
173	SPT5H_HUMAN	O00267	SAAGSEKEEEPEDEEEEEEEEEYDEEEEEEDDDRPPK + 2 Phospho (ST)	102	28
174	SPN90_HUMAN	Q9NZQ3	RGPSASSVAVMTSSTSDHHLDAAAAR + Phospho (ST)	42	33

175	SPF45_HUMAN	Q96I25	SPTGPSNSFLANMGGTVAHK + Oxidation (M); Phospho (ST)	120	32
			SPTGPSNSFLANMGGTVAHK + Phospho (ST)	59	31
176	SPD2A_HUMAN	Q5TCZ1	AASQGSDSPLLPAQR + Phospho (ST)	35	31
177	SP30L_HUMAN	Q9HAJ7	TSDDGGDSPEHDTDIPEVDLFQLQVNTLR + Phospho (ST)	62	33
178	SP3_HUMAN	Q02447	IGPPSPGDDEEEAAAAAGAPAAAGATGDLASAQLGGAPNR + Phospho (ST)	111	33
179	SP110_HUMAN	Q9HB58	DKEDPQEMPHSPLGSMPEIR + Phospho (ST)	37	32
180	SP100_HUMAN	P23497	LNECISPVANEMNHLPAHSHDLQR + Oxidation (M); Phospho (ST)	62	33
			LNECISPVANEMNHLPAHSHDLQR + Phospho (ST)	62	33
			VIGQDHDFSESSEEEAPAEASSGALR + 2 Phospho (ST)	44	31
181	SNX15_HUMAN	Q9NRS6	EEGAAPSPTHVAELATMEVESAR + Phospho (ST)	52	32
182	SNX1_HUMAN	Q13596	LPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSK + 2 Phospho (ST)	89	33
			RFSDFLGLYEK + Phospho (ST)	53	30
183	SNW1_HUMAN	Q13573	GPPSPPAPVMHSPSR + 2 Phospho (ST)	48	30
184	SNUT1_HUMAN	O43290	RVSEVEEEKEPVPQPLPSDDTR + Phospho (ST)	48	32
185	SNPC4_HUMAN	Q5SXM2	VGSESEDEDLLSELELADR + 2 Phospho (ST)	35	31
186	SNP23_HUMAN	O00161	TTWGDGGENSPCNVVSK + Phospho (ST)	33	31
187	SNIP1_HUMAN	Q8TAD8	RPDHSGGSPSPPTSEPAR + 2 Phospho (ST)	40	30
188	SMRCD_HUMAN	Q9H4L7	ANTPDSDITEK + 2 Phospho (ST)	34	25
			KLSSSSEPYEEDEFNDDQSIKK + 2 Phospho (ST)	50	32
189	SMRC2_HUMAN	Q8TAQ2	DMDEPSPVPNVEEVTLPK + Oxidation (M); Phospho (ST)	53	31
			DMDEPSPVPNVEEVTLPK + Phospho (ST)	76	32
			SPSPSPTPEAK + 2 Phospho (ST)	38	27
			TLTDEVNSPDSDRR + Phospho (ST)	49	31
190	SMRC1_HUMAN	Q92922	KHSPSPPPPTPTESR + 2 Phospho (ST)	60	30

191	SMN_HUMAN	Q16637	GTGQSDDSDIWDDTALIK + 2 Phospho (ST)	87	30
			GTGQSDDSDIWDDTALIK + Phospho (ST)	59	31
			RGTGQSDDSDIWDDTALIK + 2 Phospho (ST)	39	31
192	SMCA5_HUMAN	O60264	GGPEGVAAQAVASAASAGPADAEMEEIFDDASPGK + Phospho (ST)	132	33
193	SMCA4_HUMAN	P51532	GRPPAEKLSPNPPNLTK + Phospho (ST)	38	30
			KAENAEGQTPAIGPDGEPLDETSQMSDLPVK + 2 Phospho (ST)	53	33
			KAENAEGQTPAIGPDGEPLDETSQMSDLPVK + Oxidation (M); 2 Phospho (ST)	47	33
194	SMCA2_HUMAN	P51531	AKPVVSDFDSDEEQDER + 2 Phospho (ST)	59	29
195	SMC4_HUMAN	Q9NTJ3	EEGPPPPSPDGASSDAEPEPPSGR + Phospho (ST)	33	32
			TESPATAAETASEELDNR + Phospho (ST)	63	31
196	SMAP_HUMAN	O00193	SASPDDDLGSSNWEAADLGNEER + Phospho (ST)	162	31
197	SLTM_HUMAN	Q9NWH9	DGQDAIAQSPEK + Phospho (ST)	42	28
			DGQDAIAQSPEKESK + Phospho (ST)	68	31
			SPGHMVILDQTK + Phospho (ST)	70	30
198	SLK_HUMAN	Q9H2G2	ASSDLSIASSEEDKLSQNACILESVSEK + 2 Phospho (ST)	56	33
			RDSFIGTPYWMAPEVVMCETSK + Phospho (ST)	45	32
199	SIN3A_HUMAN	Q96ST3	GDLSDVEEEEEEEMDVDEATGAVK + Phospho (ST)	38	30
			SPPVQPHTPVTISLGTAPSLQNNQPVEFNHAINYVNK + Phospho (ST)	68	31
200	SIGIR_HUMAN	Q6IA17	SSEVDVSDLGSR + Phospho (ST)	35	29
201	SI1L1_HUMAN	O43166	SSPKEELHPAAPSQLAPSFSSSSSSSSGPR + Phospho (ST)	56	33
			TLSDESIYNSQR + Phospho (ST)	42	30
202	SHOT1_HUMAN	A0MZ66	SMPVLGSVSSVTK + Oxidation (M); Phospho (ST)	34	30
			SMPVLGSVSSVTK + Phospho (ST)	40	30
			VTAEADSSSPTGILATSESK + Phospho (ST)	118	31
203	SH3B4_HUMAN	Q9P0V3	SYSLSELSVLQAK + Phospho (ST)	65	30

204	SH2B1_HUMAN	Q9NRF2	SSEDLAGPLPSSVSSSSTTSSKPK + Phospho (ST)	64	32
205	SH23A_HUMAN	Q9BRG2	SFSEDTLMDGPAR + Phospho (ST)	35	29
206	SGTA_HUMAN	O43765	SPARTPPSEEDSAEAER + 2 Phospho (ST)	55	29
			TPPSEEDSAEAER + Phospho (ST)	39	28
207	SGPP1_HUMAN	Q9BX95	RNSLTGEEGQLAR + Phospho (ST)	56	30
208	SFRS9_HUMAN	Q13242	GSPHYFSPFRPY + 2 Phospho (ST)	42	28
			GSPHYFSPFRPY + Phospho (ST)	51	30
209	SFRS8_HUMAN	Q12872	SGVSSDNEDDDDEEDGNYLHPSLFASK + Phospho (ST)	121	31
210	SFRS7_HUMAN	Q16629	SISLRR + 2 Phospho (ST)	35	27
			YFQSPSR + Phospho (ST)	35	28
211	SFRS2_HUMAN	Q01130	SKSPPKSPEEEGAVSS + 2 Phospho (ST)	34	30
			SRSPPPVSK + 2 Phospho (ST)	31	28
212	SFRIP_HUMAN	Q99590	FHSPSTTWSPNK + Phospho (ST)	39	30
			SSSNDSVDEETAESDTSPVLEK + Phospho (ST)	35	31
			TEELIESPKLESSEGEIIQTVDR + Phospho (ST)	91	32
			VETVSQPSESPKDTIDK + Phospho (ST)	45	31
213	SFR15_HUMAN	O95104	IEIIQPLLDMAAGTSNAAPVAENVTNNEGSPPPPVK + Phospho (ST)	36	31
214	SFR14_HUMAN	Q8IX01	KMSFDIIDK + Phospho (ST)	41	29
215	SFR11_HUMAN	Q05519	KPIETGSPK + Phospho (ST)	31	28
216	SF3B2_HUMAN	Q13435	GFEEEHKDSDDDSSDDEQEKKPEAPK + 2 Phospho (ST)	45	30
			GFEEEHKDSDDDSSDDEQEKKPEAPK + 3 Phospho (ST)	56	30
			SSLGQSASETEEDTVSVSK + 2 Phospho (ST)	31	30
217	SF3A1_HUMAN	Q15459	FGESEEVEMEVESDEEDDKQEK + Oxidation (M); Phospho (ST)	80	30
218	SF04_HUMAN	Q8IWZ8	AVQQHQHGYDSDEEVDSELGTWEHQLR + Phospho (ST)	46	33
219	SF01_HUMAN	Q15637	SPSPEPIYNSEGK + 2 Phospho (ST)	59	28

			TGDLGIPPNPEDRSPSPEPIYNSEGK + Phospho (ST); Phospho (Y)	41	33
220	SEPT1_HUMAN	Q8WYJ6	EEEIHIYQFPECDSDEDEDFKR + Phospho (ST)	54	31
221	SEC62_HUMAN	Q99442	EELEQQTDGDCEEDEEEENDGETPK + Phospho (ST)	45	28
222	SDS3_HUMAN	Q9H7L9	RPASPSSPEHLPATPAESPAQR + 2 Phospho (ST)	45	32
			RPASPSSPEHLPATPAESPAQR + Phospho (ST)	58	32
223	SDCG1_HUMAN	O60524	DELNEELIQEESSEDEGEYEEVRK + 2 Phospho (ST)	80	31
224	SDA1_HUMAN	Q9NVU7	KYIEIDSDEEPR + Phospho (ST)	35	30
			YIEIDSDEEPR + Phospho (ST)	49	29
			YIEIDSDEEPRGELLSLR + Phospho (ST)	42	32
225	SCAM3_HUMAN	O14828	DGGNPFAEPSELDNPFQDPAVIQHRPSR + Phospho (ST)	37	33
226	SC24B_HUMAN	O95487	TPPTANHPVEPVTSVTQPSELLQQK + Phospho (ST)	50	31
227	SC16A_HUMAN	O15027	FTGSFDDDPDPHRDPYGEEVDR + Phospho (ST)	96	31
			FTGSFDDDPDPHRDPYGEEVDRR + Phospho (ST)	69	31
			GSVSQPSTPSPPKPTGIFQTSANSSFEPVK + 2 Phospho (ST)	41	33
			LLPSAPQTLPDGPLASPAR + Phospho (ST)	57	30
228	SATT_HUMAN	P43007	SEEETSPLVTHQNPAGPVASAPELESK + Phospho (ST)	62	33
229	SAP30_HUMAN	O75446	KGSDDDGGDSPVQDIDTPEVDLYQLQVNTLR + 2 Phospho (ST)	34	33
230	SAMH1_HUMAN	Q9Y3Z3	NFTKPQDGDVIAPLITPQKK + Phospho (ST)	37	29
231	SAMD1_HUMAN	Q6SPF0	AAAAAATAPPSPGPAQPGPR + Phospho (ST)	33	31
232	SAM14_HUMAN	Q8IZD0	VTDGCGSPLHRLR + Phospho (ST)	36	30
233	SAFB1_HUMAN	Q15424	ESSTSEGADQKMSSPEDDSDTKR + 2 Phospho (ST)	34	29
			TDCEPVGLEPAVEQSSAASELAEASSEELAEAPTEAPSPEAR + Phospho (ST)	52	33
234	S39A3_HUMAN	Q9BRY0	EKPSFIDLETFNAGSDVGSDSEYESPFMGGAR + Oxidation (M); 2 Phospho (ST)	37	32
235	S2546_HUMAN	Q96AG3	SFSTGSDLGHWVTTPPDIPGSR + 2 Phospho (ST)	35	32
			SFSTGSDLGHWVTTPPDIPGSR + Phospho (ST)	48	32

236	S12A2_HUMAN	P55011	DEGPAAAGDGLGRPLGPTPSQSR + Phospho (ST)	36	32
237	RU17_HUMAN	P08621	YDERPGPSPLPHR + Phospho (ST)	65	31
238	RSRC2_HUMAN	Q7L4I2	EQSEVSVSPR + Phospho (ST)	39	29
239	RS3_HUMAN	P23396	DEILPTTPISEQK + Phospho (ST)	40	31
240	RS17_HUMAN	P08708	LLDFGSLSNLQVTQPTVGMNFK + Phospho (ST)	64	32
241	RS10L_HUMAN	Q9NQ39	AEAGAGSATEFQFR + Phospho (ST)	30	30
242	RRP12_HUMAN	Q5JTH9	GDSIEEILADSEDEEDNEEEER + Phospho (ST)	37	29
243	RRN3_HUMAN	Q9NYV6	EGDVDVSDSDDEDDNLPANFDTCHR + 2 Phospho (ST)	37	28
244	RRMJ3_HUMAN	Q8IY81	ALDISLSSGEEDEGDEEDSTAGTTK + 2 Phospho (ST)	91	30
245	RRAGC_HUMAN	Q9HB90	MSPNETLFLESTNK + Oxidation (M); Phospho (ST)	74	30
			MSPNETLFLESTNK + Phospho (ST)	48	31
246	RPTOR_HUMAN	Q8N122	VLDTSSLTQSAPASPTNK + Phospho (ST)	67	31
247	RPR1B_HUMAN	Q9NQG5	TFQQIQEEEDDDYPGSYSPQDPSAGPLLTEELIK + Phospho (ST)	137	33
248	RPGF6_HUMAN	Q8TEU7	LPEGPVDSEDDEEEDEEIDRTDPLQGR + Phospho (ST)	46	32
249	RPGF1_HUMAN	Q13905	ATSGSSLPVGINR + Phospho (ST)	43	30
250	RPB1_HUMAN	P24928	YSPTSPTYSPTSPK + Phospho (ST); Phospho (Y)	33	29
			YSPTSPTYSPTSPVYTPTSPK + Phospho (ST); Phospho (Y)	32	32
251	RNF34_HUMAN	Q969K3	ASLSDLSSLDDVEGMSVR + Phospho (ST)	49	31
252	RN216_HUMAN	Q9NWF9	VLPQTILYKYYER + 2 Phospho (Y)	35	31
253	RMP_HUMAN	O94763	KNSTGSGHSAQELPTIR + Phospho (ST)	31	31
254	RLA2_HUMAN	P05387	KEESEESDDDMGFGLFD + 2 Phospho (ST)	45	26
			KEESEESDDDMGFGLFD + Oxidation (M); 2 Phospho (ST)	74	25
			KEESEESDDDMGFGLFD + Oxidation (M); Phospho (ST)	124	27
			YVASYLLAALGGNSSPSAK + Phospho (ST)	45	31

255	RLA0L_HUMAN	Q8NHW5	EESEESDEDMGFGLFD + 2 Phospho (ST)	52	23
			EESEESDEDMGFGLFD + Oxidation (M); 2 Phospho (ST)	49	22
			EESEESDEDMGFGLFD + Oxidation (M); Phospho (ST)	29	25
			EESEESDEDMGFGLFD + Phospho (ST)	31	26
			VEAKEESEESDEDMGFGLFD + 2 Phospho (ST)	65	28
			VEAKEESEESDEDMGFGLFD + Oxidation (M); 2 Phospho (ST)	39	27
256	RIR2_HUMAN	P31350	VPLAPITDPQQLQLSPLK + Phospho (ST)	53	26
			VPLAPITDPQQLQLSPLKGLSLVDKENTPPALSGTR + Phospho (ST)	47	26
257	RIPK2_HUMAN	O43353	SLPAPQDNDFLSR + Phospho (ST)	38	30
258	RIOK2_HUMAN	Q9BVS4	EGSEFSFSDGEVAEK + 3 Phospho (ST)	44	23
259	RICTR_HUMAN	Q6R327	NDSGEENVPLDLTR + Phospho (ST)	39	30
260	RICS_HUMAN	A7KAX9	LSPFFTLDLSPTEDK + Phospho (ST)	54	31
			SAKSEESLTSLHAVDGDSK + Phospho (ST)	47	32
			SEESLTSLHAVDGDSK + Phospho (ST)	80	30
261	RIC8A_HUMAN	Q9NPQ8	GLMAGGRPEGQYSEDEDTDTDEYK + Oxidation (M); 2 Phospho (ST)	54	29
			GLMAGGRPEGQYSEDEDTDTDEYK + Oxidation (M); Phospho (ST); Phospho (Y)	37	28
262	RHG12_HUMAN	Q8IWW6	ATTPPNQGRPDSPVYANLQELK + 2 Phospho (ST)	45	32
263	RFX7_HUMAN	Q2KHR2	NLSGSTLYPVSNIPR + Phospho (ST)	48	31
264	RFIP1_HUMAN	Q6WKZ4	DSGSDTASAIIPSTTPSVDSDDESVVKDK + Phospho (ST)	48	33
			HLFSSTENLAAGSWK + Phospho (ST)	45	31
			NNMTASMFDLSMK + Phospho (ST)	33	28
265	RFFL_HUMAN	Q8WZ73	RASLSDLTDLEDIEGLTVR + Phospho (ST)	58	31
			VPAEDETQSIDSEDSFVPGR + 2 Phospho (ST)	39	31
266	RFC1_HUMAN	P35251	IIYDSDSESEETLQVK + 2 Phospho (ST)	36	30
			IIYDSDSESEETLQVK + Phospho (ST); Phospho (Y)	35	31

267	RERE_HUMAN	Q9P2R6	EEALDDAEEPESPPPPPRSPSPEPTVVDTPSHASQSAR + 3 Phospho (ST)	80	32
			TSPINEDIR + Phospho (ST)	38	29
268	RER1_HUMAN	O15258	VDPSLMEDSDDGPSLPTK + Oxidation (M); Phospho (ST)	54	31
269	REQU_HUMAN	Q92785	NRPGLSYHYAHSHLAEEEGEDKEDSQPPTPVSQR + 2 Phospho (ST)	54	33
			NRPGLSYHYAHSHLAEEEGEDKEDSQPPTPVSQR + Phospho (ST)	66	33
			VDDDSLGEFPVTNSR + Phospho (ST)	47	30
270	REPS1_HUMAN	Q96D71	LKSEDELRPEVDEHTQK + Phospho (ST)	77	32
			RLKSEDELRPEVDEHTQK + Phospho (ST)	76	32
			SSSSQTLTQFDSNIAPADPDTAIVHPVPIR + 2 Phospho (ST)	33	33
			TSADAQEPASPVVSPQQSPPTSPHTWR + 2 Phospho (ST)	40	33
271	REM2_HUMAN	Q8IYK8	RASPPGTPTPEADATLLK + 2 Phospho (ST)	41	31
272	REEP4_HUMAN	Q9H6H4	SFSMQDLR + Oxidation (M); Phospho (ST)	40	27
			SFSMQDLR + Phospho (ST)	36	28
273	REEP3_HUMAN	Q6NUK4	SFSMHDLTTIQGDEPVGQRPYQPLPEAK + Oxidation (M); Phospho (ST)	86	33
			SFSMHDLTTIQGDEPVGQRPYQPLPEAK + Phospho (ST)	118	32
274	REEP1_HUMAN	Q9H902	SFSMQDLTTIR + Oxidation (M); Phospho (ST)	43	29
			SFSMQDLTTIR + Phospho (ST)	52	30
275	RCL_HUMAN	O43598	YFEADPPGQVAASPDPTT + Phospho (ST)	32	31
276	RCC1_HUMAN	P18754	SPPADAIPK + Phospho (ST)	30	29
277	RBP2_HUMAN	P49792	DSLITPHVSR + 2 Phospho (ST)	35	29
			EDALDDSVSSSSVHASPLASSPVR + Phospho (ST)	103	32
			EDALDDSVSSSSVHASPLASSPVRK + Phospho (ST)	96	32
			FESPATGILSPR + Phospho (ST)	49	30
			KQSLPATSIPTPASFK + Phospho (ST)	45	30
			NHETDGGSAHGDDDDDGPHFEPVVPLPDK + Phospho (ST)	93	32

			NHETDGGSAHGDDDDDGPHFEPVVPLPDKIEVK + Phospho (ST)	87	33
			NRPDYVSEEEEDDEDFETAVK + Phospho (ST)	65	31
			SISSPSVSSETMDKPVDLSTR + Oxidation (M); Phospho (ST)	55	32
			SISSPSVSSETMDKPVDLSTR + Phospho (ST)	94	32
			VGEDEDGSDEEVVHNEDIHFEPIVSLPEVEVK + Phospho (ST)	86	33
278	RBP1_HUMAN	Q15311	TEGYAAFQEDSSGDEAESPSK + 2 Phospho (ST)	32	28
			TGEPSPPHDILHEPPDVVSDDEKDHGK + 2 Phospho (ST)	40	32
			TGEPSPPHDILHEPPDVVSDDEKDHGK + Phospho (ST)	78	33
			TPSSEEISPTK + Phospho (ST)	46	29
279	RBM5_HUMAN	P52756	GLVAAYSGDSDNEEELVER + Phospho (ST)	61	31
280	RBM4_HUMAN	Q9BWF3	LHVGNISPTCTNK + Phospho (ST)	44	31
281	RBM23_HUMAN	Q86U06	EKSPVREPVDNLSPEER + Phospho (ST)	67	31
282	RBM12_HUMAN	Q9NTZ6	SRSPHEAGFCVYLK + Phospho (ST)	37	31
283	RBCC1_HUMAN	Q8TDY2	ASVSQTSPQSASSPR + Phospho (ST)	44	30
284	RBBP6_HUMAN	Q7Z6E9	DDATPVRDEPMDAESITFK + Phospho (ST)	57	32
			WDKDDFESEEEDVK + Phospho (ST)	100	28
285	RB3GP_HUMAN	Q15042	KTSASDVTNIYPGDAGK + Phospho (ST)	88	31
286	RB15B_HUMAN	Q8NDT2	DRTPPHLLYSDR + Phospho (ST)	65	31
			DRTPPHLLYSDRDR + Phospho (ST)	50	32
			SLSPVAAPPLR + Phospho (ST)	32	29
			SLSPVAAPPLREPR + Phospho (ST)	38	29
287	RB12B_HUMAN	Q8IXT5	FPPEDFRHSPEDFR + Phospho (ST)	36	31
			RPPEEDFRHSPEEDFR + Phospho (ST)	54	31
			RPPEEDFRHSPEEDFRQSPQEHFR + 2 Phospho (ST)	42	32
288	RB_HUMAN	P06400	TLQTDSIDSFETQRTPR + Phospho (ST)	31	31

289	RANG_HUMAN	P43487	DTHEDHDTSTENTDESNHDPQFEPIVSLPEQEIK + Phospho (ST)	35	33
290	RANB3_HUMAN	Q9H6Z4	VLSPPKLNEVSSDANR + Phospho (ST)	49	31
291	RAI3_HUMAN	Q8NFJ5	AYSQEEITQGFEETGDTLYAPYSTHFQLQNQPPQK + Phospho (ST)	77	33
292	RAGP1_HUMAN	P46060	ILDPNTGEPAPVLSSPPPADVSTFLAFPSPEK + 2 Phospho (ST)	48	33
			KILDPNTGEPAPVLSSPPPADVSTFLAFPSPEK + 2 Phospho (ST)	38	32
293	RAD9A_HUMAN	Q99638	VLPSISLSPGPQPPK + Phospho (ST)	32	29
294	RAD50_HUMAN	Q92878	LFDVCGSQDFESDLDR + Phospho (ST)	52	31
295	RABE1_HUMAN	Q15276	AQSTDSLGTSGSLQSK + 2 Phospho (ST)	45	29
			AQSTDSLGTSGSLQSK + Phospho (ST)	85	31
			GSVHSLDAGLLLPSGDPFSK + 2 Phospho (ST)	42	32
296	RAB12_HUMAN	Q6IQ22	AGGGGGLGAGSPALSGGQGR + Phospho (ST)	104	31
297	R3HD2_HUMAN	Q9Y2K5	ASSFSGISILTR + Phospho (ST)	60	30
298	R3HD1_HUMAN	Q15032	ASSFSGISVLTR + Phospho (ST)	85	30
299	R113A_HUMAN	O15541	AAYGDLSSEEEEENEPESLGVVYK + 2 Phospho (ST)	46	31
			AAYGDLSSEEEEENEPESLGVVYK + Phospho (ST); Phospho (Y)	64	31
			YGVYEDENYEVGSDDEEIPFK + Phospho (ST)	79	31
300	PYRG1_HUMAN	P17812	SGSSSPDSEITELK + 2 Phospho (ST)	57	28
			SGSSSPDSEITELKFPSINHD + 2 Phospho (ST)	56	31
			SGSSSPDSEITELKFPSINHD + Phospho (ST)	53	32
301	PYR1_HUMAN	P27708	ASDPGLPAEEPK + Phospho (ST)	37	29
			IHRASDPGLPAEEPK + Phospho (ST)	83	31
302	PYGO2_HUMAN	Q9BRQ0	GGGTPDANSLAPPGK + Phospho (ST)	41	30
303	PWP2A_HUMAN	Q96N64	SPEAVGPELEAEEK + Phospho (ST)	31	30
304	PWP1_HUMAN	Q13610	LQEEGGGSDEEETGSPSEDGMQSAR + Phospho (ST)	34	29
			NSSISGPFGSR + Phospho (ST)	38	28

305	PVRL2_HUMAN	Q92692	EQTLQGAEEDEDLEGPPSYKPPTPK + Phospho (ST)	45	33
306	PVRL1_HUMAN	Q15223	AGIPQHHPPMAQNLQYPDDSDDEKK + Oxidation (M); Phospho (ST)	34	33
			AGIPQHHPPMAQNLQYPDDSDDEKK + Phospho (ST)	61	33
			AGPLGGSSYEEEEEEEEGGGGGER + 2 Phospho (ST)	58	27
			AGPLGGSSYEEEEEEEEGGGGGER + Phospho (ST); Phospho (Y)	33	28
307	PUR6_HUMAN	P22234	EVYELLDSPGK + Phospho (ST)	34	30
			TKEVYELLDSPGK + Phospho (ST)	47	31
308	PUM1_HUMAN	Q14671	RDSLTGSSDLYK + Phospho (ST)	58	30
309	PTSS1_HUMAN	P48651	TYSECEDGTYSPEISWHHR + Phospho (ST)	106	30
310	PTN23_HUMAN	Q9H3S7	GAAAADLLSSSPESQHGGTQSPGGGQPLLQPTK + Phospho (ST)	69	32
311	PTN12_HUMAN	Q05209	DVDVSEDSPPPLPER + Phospho (ST)	71	31
			TVSLTPSPTTQVETPDLVDHDNTSPLFR + Phospho (ST)	41	32
312	PTAD1_HUMAN	Q9P035	WLDESDAEMELR + Phospho (ST)	62	29
313	PSN1_HUMAN	P49768	AAVQELSSSILAGEDPEER + 2 Phospho (ST)	52	31
314	PSMD2_HUMAN	Q13200	APVQPQQSPAAAPGGTDEKPSGK + Phospho (ST)	64	32
			DKAPVQPQQSPAAAPGGTDEKPSGK + Phospho (ST)	85	32
315	PSMD1_HUMAN	Q99460	TSSAFVGKTPEASPEPK + 2 Phospho (ST)	41	31
316	PRR12_HUMAN	Q9ULL5	KQETAAVCGETDEEAGESGGEGIFR + 2 Phospho (ST)	64	31
317	PRPK_HUMAN	Q96S44	ATTPADGEEPAPEAEALAAAR + Phospho (ST)	85	31
318	PRPF3_HUMAN	O43395	WDEQTSNTKGDDDEESDEEAVKK + Phospho (ST)	75	31
319	PRP31_HUMAN	Q8WWY3	SSGTASSVAFTPLQGLEIVNPQAAEK + 2 Phospho (ST)	40	32
320	PROM2_HUMAN	Q8N271	LSSTSSEETQLFHIPR + 2 Phospho (ST)	65	31
321	PRKDC_HUMAN	P78527	LTPLPEDNSMNVDQDGDPSDR + Phospho (ST)	62	31
322	PRDBP_HUMAN	Q969G5	APEPLGPADQSELGPEQPEAEVGESSDEEPVESR + 2 Phospho (ST)	80	33
323	PRCC_HUMAN	Q92733	IAAPELHKGDSDSEEDEPTK + 2 Phospho (ST)	84	31

			IAAPELHKGDSDSEEDEPTK + Phospho (ST)	105	32
324	PR40B_HUMAN	Q6NWY9	GSPSSHLLGADHGLR + Phospho (ST)	95	31
325	PR38B_HUMAN	Q5VTL8	RSLSPR + Phospho (ST)	30	28
			SQSIEQESQEK + 2 Phospho (ST)	36	26
326	PR38A_HUMAN	Q8NAV1	VSALEEDMDDVESSEEEEEEDEKLER + 2 Phospho (ST)	37	30
			VSALEEDMDDVESSEEEEEEDEKLER + Oxidation (M); 2 Phospho (ST)	44	29
327	PPR3D_HUMAN	O95685	SLSCLSDLDGGVALEPR + Phospho (ST)	62	31
328	PPR1B_HUMAN	Q9UD71	IAESHLQSISNLNENQASEEEDELGELR + Phospho (ST)	175	33
			LSEHSSPEEEASPHQR + 2 Phospho (ST)	67	28
			LSEHSSPEEEASPHQR + Phospho (ST)	104	30
329	PPIL4_HUMAN	Q8WUA2	INHTVILDDPFDDPPDLLIPDRSPEPTR + Phospho (ST)	92	32
330	PPHLN_HUMAN	Q8NEY8	DNTFFRESPVGR + Phospho (ST)	51	30
331	PP1RA_HUMAN	Q96QC0	LIPLDEECSMDETPYVETLEPGGSGGSPDGAGGSK + Oxidation (M); Phospho (ST)	39	33
			VLSPTAAKPSPFEGK + Phospho (ST)	60	30
332	PP1R7_HUMAN	Q15435	GAGQQQSQEMMEVDR + 2 Oxidation (M); Phospho (ST)	91	28
			GAGQQQSQEMMEVDR + Phospho (ST)	50	30
333	PP12C_HUMAN	Q9BZL4	IPEPESPAKPNVPTASTAPPADSR + Phospho (ST)	38	32
334	POGZ_HUMAN	Q7Z3K3	SLDSEPSVPSAAKPPSPEK + Phospho (ST)	36	31
335	PNMT_HUMAN	P11086	SPNAGAAPDSAPGQAAVASAYQR + Phospho (ST)	48	32
336	PNKP_HUMAN	Q96T60	TPESQPDTPPGTPLVSQDEK + 2 Phospho (ST)	36	31
			TPESQPDTPPGTPLVSQDEK + Phospho (ST)	67	32
337	PML_HUMAN	P29590	AVSPPHLDGPPSPR + 2 Phospho (ST)	50	30
			SPRPQQDPARPQEPTMPPPETPSEGR + Phospho (ST)	45	33
			SPVIGSEVFLPNSNHVASGAGEAEER + Phospho (ST)	33	32

338	PLXC1_HUMAN	O60486	KQSQQLELLESELR + Phospho (ST)	55	30
339	PLPL2_HUMAN	Q96AD5	NNLSLGDALAK + Phospho (ST)	36	29
			VQSLPSVPLSCAAYR + Phospho (ST)	40	31
340	PLEC1_HUMAN	Q15149	GYYSPYSVSGSGSTAGSR + Phospho (ST)	40	30
			SSSVGSSSSYPISPAVSR + Phospho (ST)	112	31
341	PLAK_HUMAN	P14923	ALMGSPQLVAAVVR + Oxidation (M); Phospho (ST)	81	30
			ALMGSPQLVAAVVR + Phospho (ST)	67	29
342	PKP4_HUMAN	Q99569	VASPSQGQVGSSSPK + Phospho (ST)	34	30
343	PKP2_HUMAN	Q99959	RLEISPDSSPER + 2 Phospho (ST)	35	28
344	PKN1_HUMAN	Q16512	LNLGTDSDSSPQK + Phospho (ST)	38	30
			TSTFCGTPEFLAPEVLTDTSYTR + Phospho (ST)	70	33
345	PKHF2_HUMAN	Q9H8W4	SPLNDMSDDDDDDDSSD + Oxidation (M); 2 Phospho (ST)	47	18
346	PKCB1_HUMAN	Q9ULU4	RISLSDMPR + Phospho (ST)	38	30
			STSPASEKADPGAVK + 2 Phospho (ST)	32	29
			TGQAGSLSGSPKPFSPQLSAPITTK + 2 Phospho (ST)	50	32
347	PININ_HUMAN	Q9H307	EIAIVHSDAEKEQEEEEQKQEMEVK + Oxidation (M); Phospho (ST)	54	33
			EIAIVHSDAEKEQEEEEQKQEMEVK + Phospho (ST)	68	33
			SLSPGKENVSALDMEK + Oxidation (M); Phospho (ST)	32	31
			SLSPGKENVSALDMEK + Phospho (ST)	113	31
348	PI4KB_HUMAN	Q9UBF8	RLSEQLAHTPTAFK + Phospho (ST)	63	30
			SVENLPECGITHEQR + Phospho (ST)	108	31
349	PHLB2_HUMAN	Q86SQ0	TSASEGNPYVSSTLSVPASPR + Phospho (ST)	45	32
350	PHLB1_HUMAN	Q86UU1	KNSITEISDNEDDLLEYHR + Phospho (ST)	52	32
			TRSPSPTLGESLAPHK + 2 Phospho (ST)	35	31
351	PHLA2_HUMAN	Q53GA4	TAPAAPAEDAVAAAAAAPSEPSEPSRPSPQPKPR + Phospho (ST)	35	33

352	PHIP_HUMAN	Q8WWQ0	VLSDSEDEEKDADVPGTSTR + 2 Phospho (ST)	44	30
353	PHF8_HUMAN	Q9UPP1	DAEYIYPSLESDDDDPALK + Phospho (ST)	48	31
354	PHF3_HUMAN	Q92576	KHSDNEAESIADALSSTSNILASEFFEEEKQESPK + Phospho (ST)	52	33
355	PHF2_HUMAN	O75151	DSDYVYPSLESDEDNPIFK + Phospho (ST)	56	31
			EDKPKPVRDEYEYVSDDGELKIDEFPIR + Phospho (ST)	36	33
356	PHC3_HUMAN	Q8NDX5	MDRTPPPPTLSPAAITVGR + 2 Phospho (ST)	58	32
357	PHAR4_HUMAN	Q8IZ21	SLPITIEMLK + Phospho (ST)	33	29
			SSSPVQVEEEPVR + Phospho (ST)	50	30
358	PHAR2_HUMAN	O75167	ASIANSDGPTAGSQTPPFK + Phospho (ST)	52	31
359	PGK1_HUMAN	P00558	ALESPERPFLAILGGAK + Phospho (ST)	111	30
360	PGAM1_HUMAN	P18669	HGESAWNLENR + Phospho (ST)	30	29
361	PEX14_HUMAN	O75381	EGHSPEGSTVTYHLLGPQEEGEGVVDVK + Phospho (ST)	60	33
			REDKEDEEDEEDDDVSHVDEEDCLGVQR + Phospho (ST)	56	31
362	PERQ2_HUMAN	Q6Y7W6	ALSSGGSITSPPLSPALPK + Phospho (ST)	62	29
			VGVEASEETPQTSSSSARPGTPSDHQSQEASQFER + Phospho (ST)	83	33
363	PELP1_HUMAN	Q8IZL8	AGSNEDPILAPSGTPPPTIPPDETFGGR + Phospho (ST)	43	33
			EGESPAAGPPPQELVEEEPSAPPTLLEEETEDGSDK + Phospho (ST)	48	33
			GSPDGSLQTGKPSAPK + Phospho (ST)	127	31
364	PDCD5_HUMAN	O14737	KVMDSDEDDDY + Phospho (ST)	49	24
365	PCY1A_HUMAN	P49585	MLQAISPK + Oxidation (M); Phospho (ST)	39	29
			MLQAISPK + Phospho (ST)	49	29
366	PCM1_HUMAN	Q15154	VTNDISPESSPGVGR + Phospho (ST)	58	30
367	PCDH1_HUMAN	Q08174	SNSPLPSIQLQPQSPSASK + Phospho (ST)	54	31
368	PCBP2_HUMAN	Q15366	GVTIPYRPKPSSSPVIFAGGQDR + Phospho (ST)	81	30
369	PBIP1_HUMAN	Q96AQ6	EEGRCSSSDDDTDVDMEGLR + Oxidation (M); 3 Phospho (ST)	33	22

370	PB1_HUMAN	Q86U86	ATSPSSSVSGDFDDGHHSVSTPGPSR + Phospho (ST)	61	32
371	PATL1_HUMAN	Q86TB9	STSPIIGSPPVR + Phospho (ST)	44	29
372	PARN_HUMAN	O95453	NNSFTAPSTVGK + Phospho (ST)	56	30
373	PARF_HUMAN	Q3YEC7	ADDFPVRDDPSDVTDEDEGPAEPPPPPK + 2 Phospho (ST)	78	32
			GSPPLPAGPVPSQDITLSSEEEAEVAAPTK + 2 Phospho (ST)	40	33
			GSPPLPAGPVPSQDITLSSEEEAEVAAPTK + 3 Phospho (ST)	37	33
			RADDFPVRDDPSDVTDEDEGPAEPPPPPK + 2 Phospho (ST)	44	32
			RADDFPVRDDPSDVTDEDEGPAEPPPPPK + Phospho (ST)	33	33
374	PAK4_HUMAN	O96013	DKRPLSGPDVGTPQPAGLASGAK + Phospho (ST)	67	31
			RKSLVGTPYWMAPELISR + Phospho (ST)	45	30
			SLVGTPYWMAPELISR + Oxidation (M); Phospho (ST)	46	31
			SLVGTPYWMAPELISR + Phospho (ST)	90	31
375	PACE1_HUMAN	Q8IZE3	KDDVSPVMQFSSK + Phospho (ST)	36	30
376	PA1_HUMAN	Q9BTK6	DLFSLDSEDPSPASPPLR + Phospho (ST)	63	32
377	P80C_HUMAN	P38432	AFQLEEGEETEPDCK + Phospho (ST)	48	29
378	P66B_HUMAN	Q8WXI9	GRLTPSPDIIVLSDNEASSPR + 2 Phospho (ST)	49	32
			GRLTPSPDIIVLSDNEASSPR + 3 Phospho (ST)	43	32
			GRLTPSPDIIVLSDNEASSPR + 4 Phospho (ST)	34	31
			LQQQAALSPTTAPAVSSVSK + Phospho (ST)	55	30
			LTPSPDIIVLSDNEASSPR + Phospho (ST)	54	31
379	P66A_HUMAN	Q86YP4	RPPSPDVIVLSDNEQPSSPR + 2 Phospho (ST)	33	32
380	OXSR1_HUMAN	O95747	TEDGGWEWSDDEFDEESEEGK + Phospho (ST)	88	28
381	OXR1_HUMAN	Q8N573	VVSSTSEEEEAFTEK + Phospho (ST)	39	30
382	OTUD4_HUMAN	Q01804	EESSEDENEVSNILR + 2 Phospho (ST)	84	28
			RPEPSTLENITDDKYATVSSPSK + Phospho (ST)	82	32

383	OSTP_HUMAN	P10451	AIPVAQDLNAPSDWDSR + Phospho (ST)	59	32
			DSYETSQLDDQSAETHSHK + Phospho (ST)	39	31
			EFHSHEFHSHEDMLVVDPK + Oxidation (M); 2 Phospho (ST)	34	30
			EFHSHEFHSHEDMLVVDPK + Oxidation (M); Phospho (ST)	32	31
			FRISHELDSASSEVN + 2 Phospho (ST)	70	30
			FRISHELDSASSEVN + Phospho (ST)	71	31
			GKDSYETSQLDDQSAETHSHK + Phospho (ST)	64	31
			KANDESNEHSDVIDSQELSKVSR + 3 Phospho (ST)	55	31
384	OSTF1_HUMAN	Q92882	TLSNAEDYLDDEDSD + 2 Phospho (ST)	47	24
			TLSNAEDYLDDEDSD + Phospho (ST)	61	26
385	OSBP1_HUMAN	P22059	GDMSDEDDENEFFDAPEIITMPENLGHK + 2 Oxidation (M); Phospho (ST)	85	31
			GDMSDEDDENEFFDAPEIITMPENLGHK + Phospho (ST)	165	32
			MLAESDESGDEESVSQTDK + 2 Phospho (ST)	76	27
			MLAESDESGDEESVSQTDK + Oxidation (M); 2 Phospho (ST)	63	26
			MLAESDESGDEESVSQTDKTELQNTLR + 2 Phospho (ST)	115	32
			TGSNISGASSDISLDEQYK + 2 Phospho (ST)	88	30
386	OSBL8_HUMAN	Q9BZF1	GYSSPEPDIQDSSGSEAQSVKPSTR + 3 Phospho (ST)	35	31
387	OSBL3_HUMAN	Q9H4L5	LHSSNPNLSTLDFGEEK + Phospho (ST)	69	31
388	OSB11_HUMAN	Q9BXB4	SFSLASSSNSPISQR + Phospho (ST)	37	31
389	ODBA_HUMAN	P12694	IGHHSTSDDSSAYR + Phospho (ST)	61	28
			SVDEVNYWDK + Phospho (ST)	43	28
			SVDEVNYWDKQDHPISR + Phospho (ST)	48	31
390	OCLN_HUMAN	Q16625	TEQDHYETDYTTGGESCDELEEDWIR + Phospho (ST)	110	31
391	OCAD1_HUMAN	Q9NX40	KLENSPLGEALR + Phospho (ST)	54	30
			LENSPLGEALR + Phospho (ST)	31	30

392	NUP98_HUMAN	P52948	DSENLASPSEYPENGER + Phospho (ST)	31	29
			NLNNSNLFSPVNR + Phospho (ST)	62	30
			YGLQDSDEEEEEHPSK + Phospho (ST)	90	29
393	NUFP2_HUMAN	Q7Z417	DYEIESQNPLASPTNTLLGSAK + Phospho (ST)	32	32
			NDSWGSFDLR + Phospho (ST)	44	28
394	NUCL_HUMAN	P19338	AAAAAPASEDEDDEDDEDDEDDDDDEEDDSEEEAMETTPAK + 2 Phospho (ST)	47	27
			AAAAAPASEDEDDEDDEDDEDDDDDEEDDSEEEAMETTPAK + Oxidation (M); 2 Phospho (ST)	50	27
			AAAAAPASEDEDDEDDEDDEDDDDDEEDDSEEEAMETTPAK + Phospho (ST)	47	27
			EDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMK + 2 Phospho (ST)	52	25
			EVEEDSEDEEMSEDEEDDSSGEEVVIPQK + Oxidation (M); 3 Phospho (ST)	31	27
			EVEEDSEDEEMSEDEEDDSSGEEVVIPQK + Oxidation (M); 4 Phospho (ST)	35	24
			EVEEDSEDEEMSEDEEDDSSGEEVVIPQKK + 3 Phospho (ST)	29	29
			EVEEDSEDEEMSEDEEDDSSGEEVVIPQKK + Oxidation (M); 2 Phospho (ST)	43	30
			KEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMK + 2 Phospho (ST)	87	26
			KEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMK + Oxidation (M); 2 Phospho (ST)	70	26
			KEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMK + Oxidation (M); Phospho (ST)	110	26
			KEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMK + Phospho (ST)	50	27
			KVVVSPTKK + Phospho (ST)	59	26
395	NU160_HUMAN	Q12769	LIRPEYAWIVQPVSGAVYDRPGASPK + Phospho (ST)	81	30
396	NRDC_HUMAN	O43847	RGSLSNAGDPEIVK + Phospho (ST)	43	30
397	NP1L4_HUMAN	Q99733	EFITGDVEPTDAESEWHSENEEEEK + Phospho (ST)	88	32
			EFITGDVEPTDAESEWHSENEEEEKLAGDMK + Phospho (ST)	43	33

			REFITGDVEPTDAESEWHSENEEEEK + Phospho (ST)	47	32
398	NP1L1_HUMAN	P55209	LDGLVETPTGYIESLPR + Phospho (ST)	67	31
399	NOP58_HUMAN	Q9Y2X3	EEPLSEEEPCTSTAIASPEK + 2 Phospho (ST)	47	30
			EEPLSEEEPCTSTAIASPEK + Phospho (ST)	98	31
			EEPLSEEEPCTSTAIASPEKK + Phospho (ST)	33	32
			HIKEEPLSEEEPCTSTAIASPEK + 2 Phospho (ST)	37	32
			HIKEEPLSEEEPCTSTAIASPEK + Phospho (ST)	138	33
			HIKEEPLSEEEPCTSTAIASPEKK + 2 Phospho (ST)	57	33
400	NOP56_HUMAN	O00567	EELMSSDLEETAGSTSIPK + Oxidation (M); Phospho (ST)	96	31
			EELMSSDLEETAGSTSIPK + Phospho (ST)	107	31
401	NOP14_HUMAN	P78316	HNDIVDSDSDAEDR + 2 Phospho (ST)	57	25
402	NONO_HUMAN	Q15233	FGQAATMEGIGAIGGTPPAFNR + Oxidation (M); Phospho (ST)	46	32
403	NOL9_HUMAN	Q5SY16	LAAFADALEFADEEKESPVEFTGHK + Phospho (ST)	66	32
404	NOL8_HUMAN	Q76FK4	EYDSGDTDEIIAMK + 2 Phospho (ST)	32	26
			FLETDSEEEQEEVNEK + 2 Phospho (ST)	36	28
			NDREYDSGDTDEIIAMK + Oxidation (M); 2 Phospho (ST)	33	28
405	NOC2L_HUMAN	Q9Y3T9	DLFDLNSSEEDDTEGFSER + 2 Phospho (ST)	135	28
			DLFDLNSSEEDDTEGFSER + Phospho (ST)	67	31
406	NOB1_HUMAN	Q9ULX3	KDDSDDDGGGWITPSNIK + Phospho (ST)	34	31
407	NMNA1_HUMAN	Q9HAN9	LEASDCDHQQNSPTLERPGR + Phospho (ST)	41	32
408	NMD3_HUMAN	Q96D46	DSAIPVESDTDDEGAPR + 2 Phospho (ST)	37	28
			DSAIPVESDTDDEGAPR + Phospho (ST)	32	30
409	NKAP_HUMAN	Q8N5F7	IGELGAPEVWGLSPK + Phospho (ST)	68	30
410	NIBAN_HUMAN	Q9BZQ8	ASAILPGVLGSETLSNEVFQESEEEKQPEVPSSLAK + Phospho (ST)	37	31
			RASAILPGVLGSETLSNEVFQESEEEKQPEVPSSLAK + Phospho (ST)	116	31

411	NGDN_HUMAN	Q8NEJ9	LSSEDEEEDEAEDDQSEASGK + 2 Phospho (ST)	30	26
412	NFIX_HUMAN	Q14938	MAFTHHPLPVLAGVRPGSPR + Phospho (ST)	70	30
			SIDDSEMESPVDDVFYPGTGR + Oxidation (M); Phospho (ST)	36	31
			SIDDSEMESPVDDVFYPGTGR + Phospho (ST)	83	32
			SPAAGSSQSSGWPNDVDAGPASLK + Phospho (ST)	40	32
413	NFIC_HUMAN	P08651	SPFNSPSPQDSPR + Phospho (ST)	38	29
414	NFIA_HUMAN	Q12857	ASPHATPSTLHFPTSPIIQQPGPYFSHPAIR + Phospho (ST)	59	31
			SPGSGSQSSGWHEVEPGMPSPTTLK + Phospho (ST)	86	32
			SVEDEMDSPGEEPFYTGQGR + Oxidation (M); Phospho (ST)	79	29
			SVEDEMDSPGEEPFYTGQGR + Phospho (ST)	110	30
415	NF2IP_HUMAN	Q8NCF5	TEFLDLDNSPLSPPSPR + Phospho (ST)	68	31
416	NEK9_HUMAN	Q8TD19	VASEAPLEHKPQVEASSPR + Phospho (ST)	41	31
417	NED4L_HUMAN	Q96PU5	SLSSPTVTLSAPLEGAK + Phospho (ST)	43	30
418	NCOR2_HUMAN	Q9Y618	EGTPPPPPPSR + Phospho (ST)	40	29
			EIAKSPHSTVPEHHPHPISPYEHLLR + Phospho (ST)	78	32
			LEPVSPPSPPHTDPELELVPPR + 2 Phospho (ST)	52	32
			SLGYHGSSYSPEGVEPVSPVSSPSLTHDK + 2 Phospho (ST)	48	33
			SPGNTSQPPAFFSK + Phospho (ST)	67	30
			SPHSTVPEHHPHPISPYEHLLR + Phospho (ST)	69	32
419	NCOR1_HUMAN	O75376	SKSPIPGQGYLGTERPSSVSSVHSEGDYHR + Phospho (ST)	46	33
			SPGSISYLPSFFTK + Phospho (ST)	80	31
			YETPSDAIEVISPASSPAPPQEK + 2 Phospho (ST)	35	31
420	NCF2_HUMAN	P19878	TPEIFR + Phospho (ST)	41	28
421	NCBP1_HUMAN	Q09161	KTSDANETEDHLESLICK + Phospho (ST)	122	31
422	NAP5_HUMAN	O14513	MDISKTKVEK + Oxidation (M); Phospho (ST)	32	29

423	NADAP_HUMAN	Q9BWU0	NWEDEDFYDSDDDTFLDR + Phospho (ST)	140	28
424	NACA_HUMAN	Q13765	VQGEAVSNIQENTQTPTVQEESEEEEVDETGVEVK + Phospho (ST)	158	33
425	MYO9B_HUMAN	Q13459	VQEKPDSPGGSTQIQR + Phospho (ST)	57	31
426	MYO5C_HUMAN	Q9NQX4	MLLEKSFELK + Phospho (ST)	51	30
427	MYL9_HUMAN	P24844	ATSNVFAMFDQSQIQEFK + Phospho (ST)	49	32
428	MYH10_HUMAN	P35580	RQLHLEGASLELSDDDTESK + Phospho (ST)	48	32
429	MYEF2_HUMAN	Q9P2K5	AEVPGATGGDSPHLQPAEPPGEPR + Phospho (ST)	114	32
430	MTMRA_HUMAN	Q9NXD2	RNSLILKPKPDPAQQTDSQNSDTEQYFR + Phospho (ST)	50	33
431	MTMR4_HUMAN	Q9NYA4	SMDDLLSACDTSSPLTR + Phospho (ST)	77	31
432	MTMR3_HUMAN	Q13615	SYDNLTTACDNTVPLASR + Phospho (ST)	49	31
433	MTA2_HUMAN	O94776	GHLSRPEAQSLSPYTTSANR + Phospho (ST)	53	32
434	MSH6_HUMAN	P52701	SEEDNEIESEEEVQPK + Phospho (ST)	62	30
			VHVQFFDDSPTR + Phospho (ST)	38	30
			VISDSESDIGGSDVEFKPDTK + 2 Phospho (ST)	52	31
435	MRVI1_HUMAN	Q9Y6F6	GLSWDSGPEEPGPR + Phospho (ST)	53	29
436	MRP_HUMAN	P49006	GDVTAEEAAGASPAK + Phospho (ST)	74	30
			LSGLSFKR + Phospho (ST)	31	29
437	MREG_HUMAN	Q8N565	ELHYLPFPSP + Phospho (ST)	35	29
438	MRE11_HUMAN	P49959	GVDFESSEDDDDDPFMNTSSLR + 2 Phospho (ST)	79	28
			GVDFESSEDDDDDPFMNTSSLR + Oxidation (M); 2 Phospho (ST)	33	27
			GVDFESSEDDDDDPFMNTSSLR + Phospho (ST)	55	30
			GVDFESSEDDDDDPFMNTSSLRR + 2 Phospho (ST)	47	29
439	MPP10_HUMAN	O00566	KSPVFSDEDSDLDFDISK + 2 Phospho (ST)	31	30
			SPVFSDEDSDLDFDISK + 2 Phospho (ST)	36	29

			SPVFSDEDSDLDFDISK + 3 Phospho (ST)	43	27
440	MOT1_HUMAN	P53985	KESKEEETSIDVAGKPNEVTK + Phospho (ST)	105	31
441	MORC2_HUMAN	Q9Y6X9	SPPLPAVIR + Phospho (ST)	46	28
			SVAVSDEEEVEEEAER + Phospho (ST)	53	31
442	MOL1A_HUMAN	Q7L9L4	HAEATLGSGNLR + Phospho (ST)	32	30
443	MMTA2_HUMAN	Q9BU76	RPAEATSSPTSPERPR + 2 Phospho (ST)	45	31
444	MLL4_HUMAN	Q9UMN6	VEVSPVLRPPITTSPPVPQEPAPVPSPPRAPTPPSTPVPLPEK + 3 Phospho (ST)	36	30
445	MLL2_HUMAN	O14686	ALSPVIPLIPR + Phospho (ST)	41	26
			VSPAAAQLADTLFSK + Phospho (ST)	69	30
446	MINT_HUMAN	Q96T58	ESGVVAVSPEKSESPQKEDGLSSQLK + 2 Phospho (ST)	36	33
			HGSFHEDEDPIGSPR + Phospho (ST)	54	29
447	MINK1_HUMAN	Q8N4C8	SDSVLPASHGHLPQAGSLER + Phospho (ST)	36	32
448	MIB2_HUMAN	Q96AX9	CLLDTDVLR + Phospho (ST)	31	29
449	MIA3_HUMAN	Q5JRA6	DTMDLESSSSEEEKEDDDDALVPDSK + Oxidation (M); Phospho (ST)	39	30
			DTMDLESSSSEEEKEDDDDALVPDSK + Phospho (ST)	33	31
			SEFGSVDGPLPHPR + Phospho (ST)	38	30
450	MFF_HUMAN	Q9GZY8	IVVAGNNEDVSFSRPADLDLIQSTPFKPLALKTPPR + Phospho (ST)	39	28
451	MEPCE_HUMAN	Q7L2J0	DEVVSPLPSALQGPSGSLSAPPAASVISAPPSSSSR + Phospho (ST)	49	31
			DITDPLSLNTCTDEGHVVLASPLK + Phospho (ST)	120	32
			ESPGAAATSSSGPQAQQHR + Phospho (ST)	128	31
452	MEP50_HUMAN	Q9BQA1	KETPPPLVPPAAR + Phospho (ST)	42	28
453	MED24_HUMAN	O75448	LLSSNEDDANILSSPTDR + Phospho (ST)	49	31
454	MED12_HUMAN	Q93074	GDLAFGAPGPRPPSPFDDPADDPEHK + Phospho (ST)	38	32
455	MED1_HUMAN	Q15648	SQTPPGVATPPIPK + Phospho (ST)	55	30

456	MDC1_HUMAN	Q14676	AQPFGFIDSDTDAEEER + Phospho (ST)	35	30
			GPGAPGLAHLQESQAGSDTDVEEGKAPQAVPLEK + Phospho (ST)	47	33
			LLLAEDSEEEVDFLSER + Phospho (ST)	88	32
			SQASMVINSDTDDEEEVSAALTLAHLK + 2 Phospho (ST)	100	33
			SQASMVINSDTDDEEEVSAALTLAHLK + Oxidation (M); 2 Phospho (ST)	91	33
			SQPPGEDSDTDVDDDSRPPGRPAEVHLER + Phospho (ST)	41	33
			SQTTTERDSDTDVEEEELPVENR + 2 Phospho (ST)	59	31
			SQTTTERDSDTDVEEEELPVENR + Phospho (ST)	52	32
			SSVKTPEPVVPTAPELQPSTSTDQPVTSEPTSQVTR + Phospho (ST)	32	31
			SSVKTPEPVVPTAPELQPSTSTDQPVTSEPTYQATR + Phospho (ST)	34	32
			SSVKTPETVVPTAPELQASASTDQPVTSEPTSR + Phospho (ST)	49	32
457	MCRS1_HUMAN	Q96EZ8	GDQVLNFSDAEDLIDDSK + Phospho (ST)	51	31
458	MCM6_HUMAN	Q14566	EIESEIDSEEELINK + Phospho (ST)	50	31
459	MCM3_HUMAN	P25205	DGDSYDPYDFSDTEEEMPQVHTPK + 2 Phospho (ST)	58	29
			DGDSYDPYDFSDTEEEMPQVHTPK + Oxidation (M); 2 Phospho (ST)	39	28
			DGDSYDPYDFSDTEEEMPQVHTPK + Oxidation (M); Phospho (ST)	56	30
			DGDSYDPYDFSDTEEEMPQVHTPK + Phospho (ST)	66	31
			SEDESETEDEEEKSQEDQEQK + 2 Phospho (ST)	45	27
460	MCM2_HUMAN	P49736	GLLYDSDEEDEERPAR + Phospho (ST)	114	31
			GNDPLTSSPGR + Phospho (ST)	48	28
			RGLLYDSDEEDEERPAR + Phospho (ST)	70	31
461	MCAF1_HUMAN	Q6VMQ6	NKQDDDLNCEPLSPHNITPEPVSK + Phospho (ST)	83	33
462	MBD3_HUMAN	O95983	YLGGSMDLSTFDFR + Oxidation (M); Phospho (ST)	54	30
			YLGGSMDLSTFDFR + Phospho (ST)	61	30
463	MBB1A_HUMAN	Q9BQG0	ALGGEDSENEEELGDEAMMALDQSLASLFAEQK + 2 Oxidation (M); Phospho	133	33

			(ST)		
464	MAT2B_HUMAN	Q9NZL9	YEMACAIADAFNLPSSHLRPITDSPVLGAQRPR + Phospho (ST)	50	32
465	MAST4_HUMAN	O15021	SQALGQSAPSLTASLK + Phospho (ST)	48	30
466	MAST2_HUMAN	Q6P0Q8	NQSLGQSAPSLTAGLK + Phospho (ST)	68	31
467	MARK3_HUMAN	P27448	GIAPASPMLGNASNPNK + Oxidation (M); Phospho (ST)	50	31
468	MARK2_HUMAN	Q7KZI7	LDTFCGSPPYAAPELFQGK + Phospho (ST)	80	32
			VPASPLPGLER + Phospho (ST)	36	29
469	MAP7_HUMAN	Q14244	AAPAQVRPPSPGNIRPVK + Phospho (ST)	33	28
			LSSSSATLLNSPDR + Phospho (ST)	65	30
470	MAP1A_HUMAN	P78559	AELEEMEEVHPSDEEEEDATK + Oxidation (M); Phospho (ST)	51	29
			ELVLSSPEDLTQDFEEMKR + 2 Phospho (ST)	40	32
			ELVLSSPEDLTQDFEEMKR + Oxidation (M); 2 Phospho (ST)	34	32
471	MAN1_HUMAN	Q9Y2U8	ENYSDSEEEDDDDVASSR + Phospho (ST)	30	27
472	MAGD2_HUMAN	Q9UNF1	HLDGEEDGSSDQSQASGTTGGR + Phospho (ST)	71	30
473	MACF1_HUMAN	Q9UPN3	AFLAELEQNSPK + Phospho (ST)	41	31
474	MA7D1_HUMAN	Q3KQU3	ESAAPASPAPSPAPSPTPAPPQK + 3 Phospho (ST)	38	31
475	M3K2_HUMAN	Q9Y2U5	DRSSPPPGYIPDELHQVAR + Phospho (ST)	34	32
476	LYRIC_HUMAN	Q86UE4	LSSQISAGEEK + Phospho (ST)	85	29
			SQEPIPDDQKVSDDDKEK + Phospho (ST)	94	31
477	LSR_HUMAN	Q86X29	ARSVDALDDLTPPSTAESGSR + Phospho (ST)	55	32
			GPALTPIRDEEWGGHSPR + Phospho (ST)	33	32
			SRDDLYDQDDSR + Phospho (ST)	50	28
			SRDDLYDQDDSRDFPR + Phospho (ST)	86	31
			SVDALDDLTPPSTAESGSR + 2 Phospho (ST)	41	30
			SVDALDDLTPPSTAESGSR + Phospho (ST)	65	31

478	LSM12_HUMAN	Q3MHD2	TETPPPLASLNVSK + Phospho (ST)	38	31
479	LS14A_HUMAN	Q8ND56	KSPTMEQAVQTASAHLPAPAAVGR + Phospho (ST)	69	32
			SPTMEQAVQTASAHLPAPAAVGR + Oxidation (M); Phospho (ST)	90	32
			SPTMEQAVQTASAHLPAPAAVGR + Phospho (ST)	188	32
			SPVSTRPLPSASQK + Phospho (ST)	94	30
			SSPQLDPLR + Phospho (ST)	46	29
			SSPQLDPLRK + Phospho (ST)	38	30
			SSPQLDPLRKSPTMEQAVQTASAHLPAPAAVGR + 2 Phospho (ST)	63	32
480	LRRF2_HUMAN	Q9Y608	RGSGDTSSLIDPDTSLSELR + Phospho (ST)	58	32
481	LRRF1_HUMAN	Q32MZ4	ALDSNSLENDDLSAPGR + Phospho (ST)	80	31
			IDGATQSSPAEPK + Phospho (ST)	45	29
			NMPGLSAATLASLGGTSSR + Oxidation (M); Phospho (ST)	49	31
482	LRC47_HUMAN	Q8N1G4	EEGSLSDTEADAVSGQLPDPTTNPSAGK + Phospho (ST)	72	33
483	LRBA_HUMAN	P50851	VEGSPTEEANLPTELQDNSLSPAASEAGEK + Phospho (ST)	54	33
484	LPP2_HUMAN	O43688	KPSLSLTLTLGEADHNHYGYPHSSS + Phospho (ST)	40	32
485	LMO7_HUMAN	Q8WWI1	EGFESDTDSEFTFK + Phospho (ST)	50	28
			EVAATEEDVTRLPSPTSPFSSLSQDQAATSK + Phospho (ST)	72	33
			LPSPTSPFSSLSQDQAATSK + Phospho (ST)	36	32
			MYSFDDVLEEGK + Phospho (ST)	60	28
			RGESLDNLDSPR + Phospho (ST)	46	30
			SMSDVSAEDVQNLR + Oxidation (M); Phospho (ST)	44	29
486	LMNB1_HUMAN	P20700	LKLSPSPSSR + Phospho (ST)	31	29
487	LMBL3_HUMAN	Q96JM7	LSGEMPPASPSFPR + Oxidation (M); Phospho (ST)	34	30
488	LMBL2_HUMAN	Q969R5	VKEEHLDVASPDKASSPELPVSVENIK + 2 Phospho (ST)	40	32
489	LIPB2_HUMAN	Q8ND30	LSCSLEDLR + Phospho (ST)	62	28

			TLSINEEEPEGGFSK + Phospho (ST)	39	31
			TQSGNFYTDTLGMAEFR + Phospho (ST)	56	31
490	LIPB1_HUMAN	Q86W92	SQSTTFNPDDMSEPEFK + Phospho (ST)	39	30
491	LIMD1_HUMAN	Q9UGP4	LSPTSLVHPVMSTLPELSCK + Phospho (ST)	41	31
492	LEMD2_HUMAN	Q8NC56	ASVRGSSEEDEDAR + 2 Phospho (ST)	60	26
			WTKPSSFSDSER + Phospho (ST)	32	30
493	LATS1_HUMAN	O95835	SNSFNNPLGNR + Phospho (ST)	54	29
			SVTPPPPPR + Phospho (ST)	46	28
494	LASS2_HUMAN	Q96G23	EETESSEGEEAAAGGGAK + 3 Phospho (ST)	42	23
			LVEDERSDREETESSEGEEAAAGGGAK + 4 Phospho (ST)	75	28
			SDREETESSEGEEAAAGGGAK + 4 Phospho (ST)	31	23
495	LAS1L_HUMAN	Q9Y4W2	MEVGPFSTGQESPTAENAR + Oxidation (M); Phospho (ST)	77	31
			MEVGPFSTGQESPTAENAR + Phospho (ST)	102	31
496	LARP5_HUMAN	Q92615	SPSPAHLPDDPK + Phospho (ST)	36	30
497	LARP4_HUMAN	Q71RC2	ASTASPCNNNINAATAVALQEPR + Phospho (ST)	48	32
498	LARP1_HUMAN	Q6PKG0	ESPRPLQLPGAEGPAISDGEEGGGEPGAGGGAAGAAGAGR + 2 Phospho (ST)	87	33
			ESPRPLQLPGAEGPAISDGEEGGGEPGAGGGAAGAAGAGR + Phospho (ST)	175	33
			GLSASLPDLDSENWIEVK + Phospho (ST)	152	31
			KNTFTAWSDEESDYEIDDR + 2 Phospho (ST)	48	28
			KNTFTAWSDEESDYEIDDRDVNK + 2 Phospho (ST)	44	31
			NTFTAWSDEESDYEIDDR + 2 Phospho (ST)	38	27
			NTFTAWSDEESDYEIDDR + Phospho (ST)	83	29
			NTFTAWSDEESDYEIDDR + Phospho (ST); Phospho (Y)	48	27
			NTFTAWSDEESDYEIDDRDVNK + Phospho (ST); Phospho (Y)	43	30
			SLPTTVPESPNYR + Phospho (ST)	32	31

499	LAP4_HUMAN	Q14160	LAEAPSPAPTPSPTPVEDLGPQTSTSPGRLSPDFAEELR + Phospho (ST)	83	32
			NSLESISSIDR + Phospho (ST)	57	29
500	LAP2B_HUMAN	P42167	HASPILPITEFSDIPR + Phospho (ST)	39	31
501	LAC_HUMAN	0	ASPTSDIESVER + Phospho (ST)	30	29
			EMDESLANLSEDEYYSEEER + Phospho (ST); Phospho (Y)	36	28
			ETVQTTQSPTPVEK + Phospho (ST)	31	31
			FTSSQTQSPTPPKPPSPSFELGLSSFPPLPGAAGNLK + 2 Phospho (ST)	39	32
			IGEGTYGVVYK + Phospho (ST)	54	30
			IGEGTYGVVYK + Phospho (ST); Phospho (Y)	39	28
			IQQFDDGGSDEEDIWEEK + Phospho (ST)	64	30
			LPDLSPVENK + Phospho (ST)	34	29
			NAEEESESEAEEGD + 2 Phospho (ST)	33	20
			NMVDLVNTHHLHSSSDDEDDRLK + Oxidation (M); Phospho (ST)	76	32
			NSSSPVSPASVPGQR + 2 Phospho (ST)	42	29
			RGSDASDFDISEIQSVCSDVETVPQTHRPTPR + Phospho (ST)	85	33
			SEGSEYEEIPK + 2 Phospho (ST)	37	26
			SESPKEPEQLR + Phospho (ST)	52	30
			SNSPLPVPPSK + 2 Phospho (ST)	35	28
			SNSPLPVPPSK + Phospho (ST)	54	29
			SSSVGSSSSYPISPAVSR + 2 Phospho (ST)	58	30
			TWTFCGTPEYVAPEIILNK + Phospho (ST)	47	32
			VEEEQEADEEDVSEEEAESK + Phospho (ST)	72	29
			VFPEPTESGDEGEELGLPLLSTR + Phospho (ST)	102	32
			WLDESDAEMELR + Oxidation (M); Phospho (ST)	51	30
			YFQSPSR + 2 Phospho (ST)	33	24

			ADEASELACPTPK + Phospho (ST)	76	29
			DLAAVNLDSSLADLGLDSLMSVEVR + Oxidation (M); Phospho (ST)	93	32
			DPSQQELPR + Phospho (ST)	55	29
			LFDHPESPTPNPTEPLFLAQAEVYK + Phospho (ST)	87	32
			LGMLSPEGTCK + Oxidation (M); Phospho (ST)	44	28
			LGMLSPEGTCK + Phospho (ST)	51	29
			LNSVQSSERPLFLVHPIEGSTTVFHSLASR + Phospho (ST)	163	30
			RPTPQDSPIFLPVDDTSFR + Phospho (ST)	38	32
502	LA_HUMAN	P05455	FASDDEHDEHDENGATGPVK + Phospho (ST)	97	30
			FASDDEHDEHDENGATGPVKR + Phospho (ST)	128	29
			SPSKPLPEVTDEYKNDVK + Phospho (ST)	35	32
			TKFASDDEHDEHDENGATGPVKR + Phospho (ST)	38	32
503	KS6C1_HUMAN	Q96S38	FLNRSPEESFDIK + 2 Phospho (ST)	43	30
504	KS6A1_HUMAN	Q15418	AYSFCGTVEYMAPEVVNR + Phospho (ST)	124	31
505	KRI1_HUMAN	Q8N9T8	AFVEDSEDEDGAGEGGSSLLQK + Phospho (ST)	122	31
506	KPSH1_HUMAN	P11801	LLTVDPGAR + Phospho (ST)	31	29
507	KPCI_HUMAN	P41743	EGLRPGDTTSTFCGTPNYIAPEILR + Phospho (ST)	71	33
508	KPCA_HUMAN	P17252	TFCGTPDYIAPEIIAYQPYGK + Phospho (ST)	104	32
509	KLDC4_HUMAN	Q8TBB5	SEDEDSLEEAGSPAPGPCPR + 2 Phospho (ST)	55	28
			SEDEDSLEEAGSPAPGPCPR + Phospho (ST)	37	30
510	KLC3_HUMAN	Q6P597	AMSLNTLNVDAPR + Phospho (ST)	85	30
			TLSASTQDLSPH + Phospho (ST)	44	29
511	KLC2_HUMAN	Q9H0B6	ASSLNFLNK + Phospho (ST)	30	29
			TLSSSSMDLSR + Phospho (ST)	35	29
512	KIF1C_HUMAN	O43896	LYADSDSGDDSDKR + 2 Phospho (ST)	58	26

513	KIF1B_HUMAN	O60333	SGLSLEELR + Phospho (ST)	49	29
514	KI67_HUMAN	P46013	AQSLVISPPAPSPR + Phospho (ST)	35	29
			MPCESSPPESADTPTSTR + Phospho (ST)	37	30
515	KHDR1_HUMAN	Q07666	ASPATQPPPLLPPSATGPDATVGGPAPTPLLPPSATASVK + Phospho (ST)	32	28
516	KDM3A_HUMAN	Q9Y4C1	HLEHAPSPSDVSNAPEVK + Phospho (ST)	52	31
517	KDM1_HUMAN	O60341	EPPRASPPGGLAEPPGSAGPQAGPTVVPGSATPMETGIAETPEGR + Oxidation (M); Phospho (ST)	34	32
518	KCRB_HUMAN	P12277	VLTPELYAELR + Phospho (ST)	39	30
519	KCD15_HUMAN	Q96SI1	SPVSPLAAQGIPLPAQLTK + 2 Phospho (ST)	62	30
			SPVSPLAAQGIPLPAQLTK + Phospho (ST)	69	27
520	KCC2D_HUMAN	Q13557	KPDGVKESTESSNTTIEDEDVK + 3 Phospho (ST)	40	30
521	KAPCA_HUMAN	P17612	TWTLCGTPEYLAPEIILSK + Phospho (ST)	117	31
522	K6PP_HUMAN	Q01813	GRSFAGNLNTYK + Phospho (ST)	38	30
			SFAGNLNTYK + Phospho (ST)	52	28
523	K2C8_HUMAN	P05787	SLDMDSIIAEVK + Phospho (ST)	36	30
524	K1671_HUMAN	Q9BY89	STSVDHSSTDLESTDGMEGPPPPDACPEK + Oxidation (M); Phospho (ST)	42	31
525	K1543_HUMAN	Q9P1Y5	APSPSGLMSPSR + Oxidation (M); Phospho (ST)	34	29
			VLTPPHDVDSLPHLR + Phospho (ST)	61	30
526	K1468_HUMAN	Q9P260	AGSISTLDSLDFAR + Phospho (ST)	66	30
527	K1467_HUMAN	A2RU67	SPLGEAPEPDSDAEVAEAAKPHLSEVTTEGYPSEPLGGLEQK + Phospho (ST)	39	33
528	K1143_HUMAN	Q96AT1	IQPQPPDEDGDHSDKEDEQPQVVVLK + Phospho (ST)	35	33
529	K0574_HUMAN	O60320	SLSRLR + 2 Phospho (ST)	32	27
530	K0528_HUMAN	Q86YS7	SQSESSDEVTELDLSHGK + 2 Phospho (ST)	44	30
531	K0284_HUMAN	Q9Y4F5	MIDQVFGVLESPELSR + Phospho (ST)	35	31
			SPEPDGPAPAFLR + Phospho (ST)	40	30

			WASLADSYSDPGLTEDGLGR + Phospho (ST)	42	31
532	JUND_HUMAN	P17535	LAALKDEPQTVPDVPSFGESPPLSPIDMDTQER + Oxidation (M); 2 Phospho (ST)	52	33
533	JIP4_HUMAN	O60271	SASQSSLDKLDQELK + 2 Phospho (ST)	72	31
534	ITSN1_HUMAN	Q15811	SAFTPATATGSSPSPVLGQGEK + 2 Phospho (ST)	82	32
			SAFTPATATGSSPSPVLGQGEK + Phospho (ST)	75	32
535	ITPR3_HUMAN	Q14573	LGFVDVQNCISR + Phospho (ST)	32	30
			VASFSIPGSSSR + Phospho (ST)	32	29
536	ITB4_HUMAN	P16144	MTTTSAAAYGTHLSPHVPHR + Oxidation (M); Phospho (ST)	54	32
			MTTTSAAAYGTHLSPHVPHR + Phospho (ST)	48	32
537	IQWD1_HUMAN	Q58WW2	DSALQDTDDSDDDPVLIPGAR + 2 Phospho (ST)	49	30
538	IQGA2_HUMAN	Q13576	TPEEMKHSQSMIEDAQLPLEQK + Phospho (ST)	36	33
			YGSIVDDER + Phospho (ST)	47	28
539	IQEC1_HUMAN	Q6DN90	MQFSFEGPEK + Phospho (ST)	34	28
540	IPYR2_HUMAN	Q9H2U2	SLVESVSSSPNK + Phospho (ST)	35	30
			SLVESVSSSPNKESNEEEQVWHFLGK + Phospho (ST)	33	32
541	IPP2_HUMAN	P41236	IDEPSTPYHSMMGDDEDACSDTEATEAMAPDILAR + Phospho (ST)	93	32
542	INF2_HUMAN	Q27J81	DPTSLLGVLQAEADSTSEGLEDAVHSR + 2 Phospho (ST)	40	33
			EHNSMWASLSSPDAEAVEPDFSSIER + Phospho (ST)	32	32
			GARPPAAGPGGDEDEDEEDTAPESALDTSLDK + 2 Phospho (ST)	44	32
543	IMA3_HUMAN	O00505	NVPQEESLEDSDVDADFK + Phospho (ST)	72	31
544	IMA2_HUMAN	P52292	NVSSFPDDATSPLQENR + Phospho (ST)	55	31
545	ILF3_HUMAN	Q12906	DSSKGEDSAEETEAKPAVVAPAPVVEAVSTPSAAFPSDATAEQGPILTK + 2 Phospho (ST)	63	32
			DSSKGEDSAEETEAKPAVVAPAPVVEAVSTPSAAFPSDATAEQGPILTK + Phospho (ST)	78	31

			GEDSAEETEAKPAVVAPAPVVEAVSTPSAAFPSDATAEQGPILTK + Phospho (ST)	110	32
			LFPDTPLALDANK + Phospho (ST)	41	30
			RPMEEDGEEKSPSK + Phospho (ST)	52	29
546	IL16_HUMAN	Q14005	DPGVSESPPPGR + Phospho (ST)	49	29
547	IKZF3_HUMAN	Q9UKT9	GLSPNNSGHDSTDTDSNHEER + Phospho (ST)	113	30
548	IF5_HUMAN	P55010	EAEEESSGGEEEDEDENIEVVYSK + 2 Phospho (ST)	53	28
549	IF4G3_HUMAN	O43432	LDFIESDSPCSSEALSK + Phospho (ST)	47	31
			SPVPAQIAITVPK + Phospho (ST)	30	28
550	IF2BL_HUMAN	A6NK07	SGDEMIFDPTMSK + Phospho (ST)	47	29
551	ICLN_HUMAN	P54105	EPVADEEEEDSDDDVEPITEFR + Phospho (ST)	89	31
			FEEESKEPVADEEEEDSDDDVEPITEFR + Phospho (ST)	131	32
552	ICAL_HUMAN	P20810	EGITGPPADSSKPIGPDDAIDALSSDFTCGSPTAAGK + Phospho (ST)	48	33
			SESELIDELSEDFDR + Phospho (ST)	62	30
			SLTPAVPVESKPDKPSGK + Phospho (ST)	41	31
553	IBP5_HUMAN	P24593	IERDSREHEEPTTSEMAEETYSPK + Phospho (ST)	36	32
554	I2BP2_HUMAN	Q7Z5L9	LEEPPELNRQSPNPR + Phospho (ST)	41	31
			RKPSPEPEGEVGPPK + Phospho (ST)	47	30
555	HUCE1_HUMAN	O43159	ALEAASLSQHPPSLCISDSEEEEEER + 2 Phospho (ST)	63	32
556	HSF1_HUMAN	Q00613	VEEASPGRPSSVDTLLSPTALIDSILR + Phospho (ST)	37	30
			VKEEPPSPPQSPR + 2 Phospho (ST)	36	30
557	HS902_HUMAN	Q14568	ESKDKPEIEDVGSDEEEEKK + Phospho (ST)	56	31
558	HOMEZ_HUMAN	Q8IX15	AETPPLPIPPPPPDIQPLER + Phospho (ST)	41	29
559	HOME3_HUMAN	Q9NSC5	EGLGQGQSLEQLEALVQTK + Phospho (ST)	68	31

560	HNRPM_HUMAN	P52272	MGLAMGGGGGASFDR + 2 Oxidation (M); Phospho (ST)	47	28
			MGLSMER + Phospho (ST)	32	24
561	HNRPK_HUMAN	P61978	DYDDMSPR + Phospho (ST)	31	23
			GSYGDLGGPIITTQVTIPK + Phospho (ST)	128	30
			IILDLISESPIK + Phospho (ST)	59	28
			IIPTLEEGLQLPSPTATSQLPLESDAVECLNYQHYK + Phospho (ST)	152	31
			KIIPTLEEGLQLPSPTATSQLPLESDAVECLNYQHYK + Phospho (ST)	37	31
			RDYDDMSPR + Phospho (ST)	71	27
562	HNRPF_HUMAN	P52597	ATENDIYNFFSPLNPVR + Phospho (ST)	32	31
563	HNRPD_HUMAN	Q14103	IFVGGLSPDTPEEK + Phospho (ST)	39	30
564	HNRPC_HUMAN	P07910	DDEKEAEEGEDDRDSANGEDDS + Phospho (ST)	62	26
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIK + 2 Phospho (ST)	81	31
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIK + Oxidation (M); Phospho (ST)	207	32
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIK + Phospho (ST)	93	32
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIKDDEK + 2 Phospho (ST)	123	30
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIKDDEK + Oxidation (M); 2 Phospho (ST)	89	30
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIKDDEK + Oxidation (M); Phospho (ST)	169	31
			MESEGGADDSAEEGDLLDDDDNEDRGDDQLELIKDDEK + Phospho (ST)	172	31
			QAVEMKNDKSEEEQSSSSVK + Phospho (ST)	72	31
565	HNRL1_HUMAN	Q9BUJ2	APQQQPPPQQPPPPQPPPQQPPPPPSYSPAR + Phospho (ST)	34	32
			GRSPQPPAEEDEDDFDDTLVAIDTYNCDLHFK + Phospho (ST)	119	33

			RGRSPQPPAEEDEDDFDDTLVAIDTYNCDLHFK + Phospho (ST)	60	33
			SPQPPAEEDEDDFDDTLVAIDTYNCDLHFK + Phospho (ST)	96	33
566	HN1L_HUMAN	Q9H910	GSGIFDESTPVQTR + Phospho (ST)	33	30
567	HMHA1_HUMAN	Q92619	KSSFNVSDVARPEAAGSPPEEGGCTEGTPAK + Phospho (ST)	51	33
568	HMGA1_HUMAN	P17096	EEEEGISQESSEEEQ + 2 Phospho (ST)	50	24
569	HMCS1_HUMAN	Q01581	RPTPNDDTLDEGVGLVHSNIATEHIPSPAK + Phospho (ST)	102	32
			RPTPNDDTLDEGVGLVHSNIATEHIPSPAKK + Phospho (ST)	115	32
570	HIRP3_HUMAN	Q9BW71	EESEESEAEPVQR + 2 Phospho (ST)	61	26
			EESEESEAEPVQR + Phospho (ST)	47	29
			ESEQESEEEILAQK + Phospho (ST)	37	30
			EVSDSEAGGGPQGER + Phospho (ST)	75	28
			GEESSEEEEKGYK + 2 Phospho (ST)	30	25
			RPPTPCSDPER + 2 Phospho (ST)	35	27
			TLDSDEERPRPAPPDWSHMR + Phospho (ST)	51	32
571	HEBP2_HUMAN	Q9Y5Z4	VYYTAGYNSPVK + Phospho (ST)	37	31
572	HDGR2_HUMAN	Q7Z4V5	GVMAVTAVTATAASDR + Oxidation (M); Phospho (ST)	40	30
573	HDAC7_HUMAN	Q8WUI4	AQSSPAAPASLSAPEPASQAR + Phospho (ST)	43	32
574	HCLS1_HUMAN	P14317	RSPEAPQPVIAMEEPAVPAPLPK + Phospho (ST)	47	30
			SPEAPQPVIAMEEPAVPAPLPK + Phospho (ST)	56	30
575	HCFC1_HUMAN	P51610	SPISVPGGSALISNLGK + Phospho (ST)	32	30
576	HAP28_HUMAN	Q13442	KSLDSDESEDEEDDYQQK + 2 Phospho (ST)	67	27
			SLDSDESEDEEDDYQQK + 2 Phospho (ST)	64	25
			SLDSDESEDEEDDYQQK + Phospho (ST)	105	28
577	H1BP3_HUMAN	Q53T59	GEDAEESLEEEEALDPLGIMR + Phospho (ST)	109	32
			KLSPQDPSEDVSSVDPLK + Phospho (ST)	73	32

578	H15_HUMAN	P16401	SETAPAETATPAPVEKSPAK + Phospho (ST)	84	31
579	H14_HUMAN	P10412	SETAPAAPAAPAPAEKTPVK + Phospho (ST)	39	31
580	GTPB1_HUMAN	O00178	SAMDSPVPASMFAPEPSSPGAAR + 2 Oxidation (M); Phospho (ST)	31	31
			SAMDSPVPASMFAPEPSSPGAAR + Phospho (ST)	68	32
581	GSE1_HUMAN	Q14687	ALSAAVADSLTNSPR + Phospho (ST)	47	30
582	GSCR1_HUMAN	Q9NZM4	SPTPPPTLHLVPEPAAPPPPPPR + Phospho (ST)	49	30
583	GRLF1_HUMAN	Q9NRY4	KVSIVSKPVLYR + Phospho (ST)	38	26
			TSFSVGSDDELGPIR + Phospho (ST)	64	30
584	GRIPE_HUMAN	Q6GYQ0	HFSQSEETGNEVFGALNEEQPLPR + Phospho (ST)	70	33
			RGSSPGSLEIPK + 2 Phospho (ST)	35	29
			SSSTSDILEPFTVER + Phospho (ST)	67	31
585	GRB7_HUMAN	Q14451	ATSLPSIPNPFPELCSPPSQSPILGGPSSAR + Phospho (ST)	133	32
			HLHPSCLGSPPLR + Phospho (ST)	57	31
			RATSLPSIPNPFPELCSPPSQSPILGGPSSAR + 2 Phospho (ST)	62	33
			RATSLPSIPNPFPELCSPPSQSPILGGPSSAR + Phospho (ST)	151	31
			SASDNTLVAMDFSGHAGR + Oxidation (M); Phospho (ST)	144	31
			SASDNTLVAMDFSGHAGR + Phospho (ST)	149	32
			SQPLLIPTTGR + Phospho (ST)	31	29
			TNHRLSLPMPASGTSLSAAIHR + Oxidation (M); Phospho (ST)	53	31
			TNHRLSLPMPASGTSLSAAIHR + Phospho (ST)	70	31
586	GR65_HUMAN	Q9BQQ3	KPPGTPPPSALPLGAPPPDALPPGPTPEDSPSLETGSR + Phospho (ST)	51	31
587	GPTC4_HUMAN	Q5T3I0	DLESCSDDDNQGSK + Phospho (ST)	41	25
588	GPSM3_HUMAN	Q9Y4H4	SAPPSPPPPGTR + 2 Phospho (ST)	31	28
589	GPSM1_HUMAN	Q86YR5	APSSDEECFFDLLTK + Phospho (ST)	50	30
590	GPN1_HUMAN	Q9HCN4	GTLDEEDEEADSDTDDIDHR + Phospho (ST)	98	28

591	GPC5C_HUMAN	Q9NQ84	VPSEGAYDIILPR + Phospho (ST)	76	30
592	GOGA5_HUMAN	Q8TBA6	KKSEPDDELLFDFLNSSQK + Phospho (ST)	93	32
			SEPDDELLFDFLNSSQK + Phospho (ST)	52	31
593	GOGA4_HUMAN	Q13439	EENPESDGEPVVEDGTSVK + Phospho (ST)	104	31
			TSSFTEQLDEGTPNR + Phospho (ST)	95	30
			VPSVESLFR + Phospho (ST)	49	28
594	GNL1_HUMAN	P36915	EEQTDTSDGESVTHHIR + Phospho (ST)	80	31
595	GLCI1_HUMAN	Q86VQ1	EKDRQSPLHGNHITISHTQATGSR + Phospho (ST)	92	32
			TSSLDTITGPYLTGQWPR + Phospho (ST)	53	31
596	GIT2_HUMAN	Q14161	TINNQHSVESQDNDQPDYDSVASDEDTDLETTASK + 2 Phospho (ST)	32	31
597	GIMA5_HUMAN	Q96F15	SEDNLSATPPALR + Phospho (ST)	31	30
598	GCSP_HUMAN	P23378	DATGKEVYRLALQTR + Phospho (Y)	31	29
599	GCP60_HUMAN	Q9H3P7	LEVSVDGLTLSPDPEERPGAEGAPLLPPPLPPPSPPGSGR + Phospho (ST)	34	30
600	GCP6_HUMAN	Q96RT7	ARLATVGDLEEIQR + Phospho (ST)	29	29
601	GCFC_HUMAN	Q9Y5B6	MADHLEGLSSDDEETSTDITNFNLEK + 2 Phospho (ST)	69	31
			MADHLEGLSSDDEETSTDITNFNLEK + Oxidation (M); 2 Phospho (ST)	81	31
602	GBF1_HUMAN	Q92538	ADAPDAGAQSDSELPSYHQNDVSLDR + Phospho (ST)	132	32
			GYTSDSEVYTDHGRPGK + Phospho (ST)	38	31
			SATDADVVNSGWLVVGK + Phospho (ST)	40	31
603	GATA3_HUMAN	P23771	DVSPDPSLSTPGSAGSAR + Phospho (ST)	55	31
604	GAK_HUMAN	O14976	DESEVSDEGGSPISSEGQEPR + 2 Phospho (ST)	47	28
			ESESALMEDRDESEVSDEGGSPISSEGQEPR + 2 Phospho (ST)	50	31
605	GAB1_HUMAN	Q13480	SSSLEGFHNHFK + Phospho (ST)	50	29
606	G3BP2_HUMAN	Q9UN86	STTPPPAEPVSLPQEPPK + Phospho (ST)	45	31
			YEDEVFGDSEPELDEESEDEVEEEQEER + 2 Phospho (ST)	95	28

607	FXR2_HUMAN	P51116	TDGSISGDRQPVTVADYISR + 2 Phospho (ST)	102	32
608	FXR1_HUMAN	P51114	RGPNYTSGYGTNSELSNPSETESER + 2 Phospho (ST)	44	31
609	FXL19_HUMAN	Q6PCT2	HVVRPPPRSPEPDTLPLAAGSDHPLPR + Phospho (ST)	72	30
610	FUSIP_HUMAN	O75494	SFDYNYR + Phospho (ST)	37	25
			SRSFDYNYR + 2 Phospho (ST)	55	26
611	FUND1_HUMAN	Q8IVP5	NPPPQDYESDDDSYEVLDLTEYAR + Phospho (ST)	46	32
612	FRMD8_HUMAN	Q9BZ67	TTSFFSR + Phospho (ST)	33	27
613	FRMD6_HUMAN	Q96NE9	GQSTDSLPQTICR + 2 Phospho (ST)	52	28
614	FR1OP_HUMAN	O95684	EKGPTTGEGALDLSDVHSPPKSPEGK + Phospho (ST)	38	32
615	FOXK2_HUMAN	Q01167	EASGGDSPKDDSKPPYSYAQLIVQAITMAPDK + Oxidation (M); Phospho (ST)	67	33
			EASGGDSPKDDSKPPYSYAQLIVQAITMAPDK + Phospho (ST)	48	33
			EGSPAPLEPEPGAAQPK + Phospho (ST)	66	31
			FAQSAPGSPLSSQPVLITVQR + Phospho (ST)	97	29
			SAPASPNHAGVLSAHSSGAQTPESLSR + Phospho (ST)	83	32
616	FOXK1_HUMAN	P85037	EEAPASPLRPLYPQISPLK + 2 Phospho (ST)	32	31
			SAPASPTHPGLMSPR + Oxidation (M); 2 Phospho (ST)	40	29
			SGGLQTPECLSR + 2 Phospho (ST)	34	28
617	FOXA1_HUMAN	P55317	KDPSGASNPSADSPLHR + Phospho (ST)	56	31
			TGQLEGAPAPGPAASPQTLDHSGATATGGASELK + Phospho (ST)	102	32
618	FNBP4_HUMAN	Q8N3X1	IDENSDKEMEVEESPEK + Phospho (ST)	55	30
619	FNBP1_HUMAN	Q96RU3	TVSDNSLSNSR + Phospho (ST)	35	29
620	FLII_HUMAN	Q13045	NAEAVLQSPGLSGK + Phospho (ST)	55	30
621	FKBP4_HUMAN	Q02790	SNTAGSQSQVETEA + Phospho (ST)	40	28
622	FIP1_HUMAN	Q6UN15	DHSPTPSVFNSDEER + 2 Phospho (ST)	57	28
			DHSPTPSVFNSDEER + Phospho (ST)	93	30

			ERDHSPTPSVFNSDEER + 2 Phospho (ST)	31	29
			ERDHSPTPSVFNSDEER + Phospho (ST)	76	31
623	FCHO2_HUMAN	Q0JRZ9	LSGINEIPRPFSPPVTSNTSPPPAAPLAR + 2 Phospho (ST)	58	32
			NLSNEELTK + Phospho (ST)	52	29
624	FBXW9_HUMAN	Q5XUX1	TWDDDSDPESETDPDAQAK + Phospho (ST)	31	29
625	FA84B_HUMAN	Q96KN1	TTPPPGRPPAPSSEEEDGEAVAH + Phospho (ST)	116	32
626	FA76B_HUMAN	Q5HYJ3	ISNLSPEEEQGLWK + Phospho (ST)	40	31
627	FA44A_HUMAN	Q8NFC6	KQHYLSSEDEPDDNPDVLDSR + Phospho (ST); Phospho (Y)	33	31
628	FA38A_HUMAN	Q92508	TASELLLDR + Phospho (ST)	39	30
629	FA21A_HUMAN	Q641Q2	ERRTPSDDEEDNLFAPPK + 2 Phospho (ST)	50	32
			SHGLESVPVLPGSGEAGVSFDLPAQADTLHSANK + Phospho (ST)	36	33
			TPSDDEEDNLFAPPK + 2 Phospho (ST)	31	28
630	F177A_HUMAN	Q8N128	VIHFVSGETMEEYSTDEDEVDGLEK + Oxidation (M); 2 Phospho (ST)	98	31
			VIHFVSGETMEEYSTDEDEVDGLEK + Oxidation (M); Phospho (ST); Phospho (Y)	89	31
631	F176B_HUMAN	Q9NVM1	SSTLEPEDDDEDEEDTVTR + Phospho (ST)	31	29
632	F125A_HUMAN	Q96EY5	DMQGLSLDAASQPSK + Oxidation (M); Phospho (ST)	54	30
			RNDSIYEASSLYGISAMDGVPFTLHPR + Phospho (ST)	53	33
633	F122B_HUMAN	Q7Z309	RIDFTPVSPAPSPTR + 2 Phospho (ST)	48	31
634	F117A_HUMAN	Q9C073	SLEGLNQELEEVFVK + Phospho (ST)	86	31
635	F10A1_HUMAN	P50502	ADEPSSEESDLEIDK + 2 Phospho (ST)	29	27
			ADEPSSEESDLEIDK + 3 Phospho (ST)	39	25
			KVEEDLKADEPSSEESDLEIDK + 2 Phospho (ST)	84	32
			KVEEDLKADEPSSEESDLEIDK + 3 Phospho (ST)	78	31
			KVEEDLKADEPSSEESDLEIDK + Phospho (ST)	63	32
636	EXOC4_HUMAN	Q96A65	DASVPLIDVTNLPTPR + Phospho (ST)	62	30

637	EXOC1_HUMAN	Q9NV70	LTGSTSSLNK + 2 Phospho (ST)	32	27
			SQSSSLLDMGNMSASDLDVADR + Phospho (ST)	112	32
638	ETV6_HUMAN	P41212	ISYTPPESPVPSYASSTPLHVPVPR + 2 Phospho (ST)	92	33
639	ESYT2_HUMAN	A0FGR8	EPTPSIASDISLPIATQELR + 2 Phospho (ST)	91	32
			EPTPSIASDISLPIATQELR + Phospho (ST)	38	31
640	ESYT1_HUMAN	Q9BSJ8	GSSVDAPPRPCHTTPDSQFGTEHVLR + Phospho (ST)	39	33
641	ESF1_HUMAN	Q9H501	ALAEEASEEELPSDVDLNDPYFAEEVK + 2 Phospho (ST)	52	32
642	ES8L2_HUMAN	Q9H6S3	HSPTSEPTPPGDALPPVSSPHTHR + Phospho (ST)	36	32
643	ERF_HUMAN	P50548	RVSSDLQHATAQLSLEHR + Phospho (ST)	40	31
644	ERCC5_HUMAN	P28715	NAPAAVDEGSISPR + Phospho (ST)	51	30
645	ERBB3_HUMAN	P21860	GESIEPLDPSEK + Phospho (ST)	43	29
646	ERBB2_HUMAN	P04626	FVVIQNEDLGPASPLDSTFYR + Phospho (ST)	45	32
			GAPPSTFKGTPTAENPEYLGLDVPV + Phospho (ST)	57	32
			GTPTAENPEYLGLDVPV + Phospho (ST)	54	31
			SGGGDLTLGLEPSEEEAPR + 2 Phospho (ST)	52	30
			SGGGDLTLGLEPSEEEAPR + Phospho (ST)	123	31
			SGGGDLTLGLEPSEEEAPRSPLAPSEGAGSDVFDGDLGMGAAK + Oxidation (M); Phospho (ST)	48	33
			SPLAPSEGAGSDVFDGDLGMGAAK + 2 Phospho (ST)	107	31
			SPLAPSEGAGSDVFDGDLGMGAAK + Oxidation (M); Phospho (ST)	121	32
			SPLAPSEGAGSDVFDGDLGMGAAK + Phospho (ST)	150	32
			VLGSGAFGTVYK + Phospho (ST)	66	29
			YSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDVRPQPPSPR + 2 Phospho (ST)	41	33
			YSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDVRPQPPSPR + Phospho (ST)	101	32
647	EPN4_HUMAN	Q14677	TIDLGAAAHYTGDKASPDQNASTHTPQSSVK + Phospho (ST)	91	33

648	EPN3_HUMAN	Q9H201	ACRTPESFLGPSASSLVNLDSLVK + 2 Phospho (ST)	63	32
			ACRTPESFLGPSASSLVNLDSLVK + Phospho (ST)	90	32
			EGLEQALPSGKPSSPVELDLFGDPSPSSK + 2 Phospho (ST)	39	33
			EGLEQALPSGKPSSPVELDLFGDPSPSSK + Phospho (ST)	77	32
			TPESFLGPSASSLVNLDSLVK + Phospho (ST)	100	31
649	EPN1_HUMAN	Q9Y6I3	TALPTSGSSAGELELLAGEVPAR + 2 Phospho (ST)	35	32
			TALPTSGSSAGELELLAGEVPAR + Phospho (ST)	113	31
			TPESFLGPNAALVDLDSLVSRPGPTPPGAK + Phospho (ST)	105	31
650	EP15R_HUMAN	Q9UBC2	TVFPGAVPVLPASPPPK + Phospho (ST)	42	28
651	EMAL4_HUMAN	Q9HC35	APVSSTESVIQSNTPTPPPSQPLNETAEEESR + 2 Phospho (ST)	92	33
			ASPSPQPSSQPLQIHR + 2 Phospho (ST)	33	31
			ASPSPQPSSQPLQIHR + Phospho (ST)	78	31
652	EMAL3_HUMAN	Q32P44	AISSANLLVR + Phospho (ST)	38	29
653	ELL_HUMAN	P55199	LGLPLLTDCAQPSRPHGSPSR + Phospho (ST)	33	31
654	EIF3G_HUMAN	O75821	GIPLATGDTSPEPELLPGAPLPPPK + Phospho (ST)	93	30
655	EIF3B_HUMAN	P55884	TEPAAEAEAASGPSESPSPPAAEELPGSHAEPPVPAQGEAPGEQAR + 2 Phospho (ST)	54	33
			TEPAAEAEAASGPSESPSPPAAEELPGSHAEPPVPAQGEAPGEQAR + Phospho (ST)	115	33
656	EI2BE_HUMAN	Q13144	GGSPQMDDIK + Phospho (ST)	65	26
657	EI24_HUMAN	O14681	TVYLQSALSSSTSAEKFPSPHPSPAK + Phospho (ST)	49	32
658	EHMT2_HUMAN	Q96KQ7	SPPSVQSLAMR + Phospho (ST)	41	29
659	EHBP1_HUMAN	Q8NDI1	DLSTSPKPSPIPSPVLGR + 2 Phospho (ST)	32	31
660	EFNB2_HUMAN	P52799	KHSPQHTTTLSLSTLATPK + Phospho (ST)	49	30
661	EF2_HUMAN	P13639	FTDTR + Phospho (ST)	43	23

662	EF1G_HUMAN	P26641	TPEFLR + Phospho (ST)	39	28
			VLSAPPHFHFGQTNRTPEFLR + Phospho (ST)	36	32
663	EDC4_HUMAN	Q6P2E9	DSQDASAEQSDHDDEVASLASASGGFGTK + Phospho (ST)	50	31
			TRSPDVISSASTALSQDIPEIASEALSR + Phospho (ST)	71	31
664	EAP1_HUMAN	Q9H1B7	NSSSPVSPASVPGQR + Phospho (ST)	53	31
665	EAF1_HUMAN	Q96JC9	TSPLKDNPSPEPQLDDIKR + 2 Phospho (ST)	34	32
			TSPLKDNPSPEPQLDDIKR + Phospho (ST)	40	32
666	E41L3_HUMAN	Q9Y2J2	GISQTNLITTVTPEK + Phospho (ST)	55	30
			TTESGSDSESKPDQEAEPQEAAGAQGR + Phospho (ST)	76	32
667	E41L1_HUMAN	Q9H4G0	SLDGAEFSRPASVSENHDAGPDGDKR + Phospho (ST)	45	32
668	E2AK4_HUMAN	Q9P2K8	HERPAGPGTPPPDSGPLAK + Phospho (ST)	46	32
669	DYN2_HUMAN	P50570	EALNIIGDISTSTVSTPVPPPVDDTWLQSASSHSPTPQR + Phospho (ST)	48	32
670	DYHC1_HUMAN	Q14204	TDSTSDGRPAWMR + Phospho (ST)	42	29
671	DPOD3_HUMAN	Q15054	VALSDDETKETENMR + Phospho (ST)	58	30
672	DP13A_HUMAN	Q9UKG1	VNQSALEAVTPSPSFQQR + Phospho (ST)	42	31
673	DOP2_HUMAN	Q9Y3R5	GELSEEELPYYVELPDR + Phospho (ST)	82	31
			SEDSGIGLSASSPELSEHLR + Phospho (ST)	91	32
674	DOK3_HUMAN	Q7L591	ATSLPSLDTPGELR + Phospho (ST)	50	30
675	DOCK7_HUMAN	Q96N67	SLSNSNPDISGTPTSPDDEVR + 2 Phospho (ST)	35	30
			SLSNSNPDISGTPTSPDDEVR + Phospho (ST)	43	31
			SPSGSAFGSQENLR + Phospho (ST)	54	30
676	DOCK6_HUMAN	Q96HP0	SISSSNPDLAVAPGSVDDEVSR + Phospho (ST)	33	32
677	DNMT1_HUMAN	P26358	EADDDEEVDDNIPEMPSPK + Phospho (ST)	59	29
678	DNLI1_HUMAN	P18858	AETPTESVSEPEVATK + Phospho (ST)	64	31
			TIQEVLEEQSEDEDR + Phospho (ST)	32	31

			VLGSEGEEEDEALSPAK + 2 Phospho (ST)	41	29
			VLGSEGEEEDEALSPAK + Phospho (ST)	36	31
679	DNJC2_HUMAN	Q99543	ELSEESEDEELQLEEFPMLK + 2 Phospho (ST)	32	31
			KELSEESEDEELQLEEFPMLK + 2 Phospho (ST)	43	32
			KELSEESEDEELQLEEFPMLK + Oxidation (M); 2 Phospho (ST)	33	32
			NASASFQELEDK + Phospho (ST)	38	30
			NASASFQELEDKK + 2 Phospho (ST)	63	29
			NASASFQELEDKK + Phospho (ST)	48	30
680	DNJC1_HUMAN	Q96KC8	DFDIAEQNESSDEESLRK + 2 Phospho (ST)	50	29
681	DMXL1_HUMAN	Q9Y485	SLALVAHTGYLPHQQDPHHVHR + Phospho (ST)	72	32
682	DJC21_HUMAN	Q5F1R6	DEEDGKDSDEAEDAELYDDLYCPACDK + Phospho (ST)	41	29
683	DIP2B_HUMAN	Q9P265	GTSGSLADVFANTR + Phospho (ST)	31	30
684	DIDO1_HUMAN	Q9BTC0	RNSVERPAEPVAGAATPSLVEQQK + Phospho (ST)	80	31
			SPPEGDTTLFLSR + Phospho (ST)	78	30
			YPLCSADAAVSTTPPGSPPPPPPLPEPPVLK + 2 Phospho (ST)	52	32
			YPLCSADAAVSTTPPGSPPPPPPLPEPPVLK + Phospho (ST); Phospho (Y)	37	32
685	DI3L2_HUMAN	Q8IYB7	RPGTQGHLGPEKEEEESDGEPEDSSTS + Phospho (ST)	100	32
686	DHX9_HUMAN	Q08211	SEEVPAFGVASPPPLTDTPDTTANAEGDLPTTMGGPLPPHLALK + Oxidation (M); Phospho (ST)	61	32
			SEEVPAFGVASPPPLTDTPDTTANAEGDLPTTMGGPLPPHLALK + Phospho (ST)	134	32
687	DHX57_HUMAN	Q6P158	GLSGEEEDDEPDCCNDER + Phospho (ST)	51	26
688	DHX16_HUMAN	O60231	LLEDSEESSEETVSR + 2 Phospho (ST)	51	28
			LLEDSEESSEETVSR + 3 Phospho (ST)	38	26
689	DHB4_HUMAN	P51659	ATSTATSGFAGAIGQK + Phospho (ST)	55	30
690	DESP_HUMAN	P15924	SMSFQGIR + Phospho (ST)	39	28

691	DEOC_HUMAN	Q9Y315	IGASTLLSDIER + Phospho (ST)	36	30
692	DEN4C_HUMAN	Q5VZ89	KSSTGSISNVLFSTQDPVEDAVFGEATNLK + 2 Phospho (ST)	55	33
			RSSLPLDHGSPAQENPESEK + Phospho (ST)	48	32
			THSFENVSCHLPDSR + Phospho (ST)	44	30
			VPSGIFDVNSR + Phospho (ST)	56	29
693	DEK_HUMAN	P35659	EESEEEEDEDDEEEEEEEKEK + Phospho (ST)	72	26
			EESSDDEDKESEEEPPKK + 3 Phospho (ST)	41	25
694	DDX55_HUMAN	Q8NHQ9	KREEGSDIEDEDMEELLNDTR + Phospho (ST)	57	32
695	DDX46_HUMAN	Q7L014	AALGLQDSDDEDAAVDIDEQIESMFNSK + Phospho (ST)	76	32
			KGELMENDQDAMEYSSEEEEVDLQTALTGYQTK + 2 Oxidation (M); 2 Phospho (ST)	35	32
696	DDX41_HUMAN	Q9UJV9	SEAEDEDDEDYVPYVPLR + Phospho (ST)	115	31
			TDEVPAGGSRSEAEDEDDEDYVPYVPLR + 2 Phospho (ST)	32	31
697	DDX24_HUMAN	Q9GZR7	AQAVSEEEEEEEGK + Phospho (ST)	50	28
			SPGKAEAESDALPDDTVIESEALPSDIAAEAR + Phospho (ST)	66	33
698	DDX21_HUMAN	Q9NR30	NEEPSEEEIDAPKPK + Phospho (ST)	55	31
699	DDA1_HUMAN	Q9BW61	TDSPDMHEDT + Phospho (ST)	64	22
700	DCP1A_HUMAN	Q9NPI6	ASSPSPLTIGTPESQR + Phospho (ST)	50	31
701	DC1L2_HUMAN	O43237	DFQDYMEPEEGCQGSPQR + Oxidation (M); Phospho (ST)	36	28
			DFQDYMEPEEGCQGSPQR + Phospho (ST)	52	29
702	DBPA_HUMAN	P16989	SPVGSGAPQAAAPAPAAHVAGNPGGDAAPAATGTAAAASLATAAGSEDAEK + Phospho (ST)	67	33
703	DBNL_HUMAN	Q9UJU6	AMSTTSISSPQPGK + Oxidation (M); Phospho (ST)	69	29
			AMSTTSISSPQPGK + Phospho (ST)	31	30
704	DAXX_HUMAN	Q9UER7	ICTLPSPPSPLASLAPVADSSTR + Phospho (ST)	41	31

			VDSPSHGLVTSSLCIPSPAR + Phospho (ST)	47	32
705	DAP1_HUMAN	P51397	DKDDQEWESPSPPKPTVFISGVIAR + Phospho (ST)	68	32
			EEKDKDDQEWESPSPPKPTVFISGVIAR + Phospho (ST)	33	32
706	CXA1_HUMAN	P17302	KLAAGHELQPLAIVDQRPSSR + Phospho (ST)	34	30
707	CV009_HUMAN	Q6ICG6	KSHSANDSEEFFREDDGGADLHNATNLR + Phospho (ST)	37	32
			SHSANDSEEFFREDDGGADLHNATNLR + Phospho (ST)	62	32
708	CUL4B_HUMAN	Q13620	EDFDSTSSSSSTPPLQPR + Phospho (ST)	35	31
709	CUL4A_HUMAN	Q13619	KGSFSALVGR + Phospho (ST)	54	29
710	CTR9_HUMAN	Q6PD62	GGEFDEFVNDDTDDDLPISK + Phospho (ST)	69	31
			KGGEFDEFVNDDTDDDLPISK + Phospho (ST)	69	32
711	CTGE4_HUMAN	Q8IX94	SNSELEDEILCLEK + Phospho (ST)	63	31
712	CTGE2_HUMAN	Q96RT6	LSGPAELR + Phospho (ST)	37	28
713	CTBP2_HUMAN	P56545	YPPGIVGVAPGGLPAAMEGIIPGGIPVTHNLPTVAHPSQAPSPNQPTK + Phospho (ST)	67	28
714	CT151_HUMAN	Q8NC74	GQDTPKPAGQHGSLSPAAAHTASPEPPTQSGPLTR + 2 Phospho (ST)	59	33
715	CT117_HUMAN	O94964	VYYSPPVAR + Phospho (ST)	30	29
716	CSF1R_HUMAN	P07333	KVMSISIRLK + Phospho (ST)	30	26
717	CRTC3_HUMAN	Q6UUV7	SNPSIQATLNK + Phospho (ST)	36	30
718	CRTC2_HUMAN	Q53ET0	SSHYGGSLPNVNQIGSGLAEFQSPLHSPLDSSR + Phospho (ST)	78	33
			TSSDSALHTSVMNPSPQDTYPGPTPPSILPSR + Oxidation (M); Phospho (ST)	52	33
719	CRIP2_HUMAN	P52943	ASSVTTFTGEPNTCPR + Phospho (ST)	79	30
720	CRIP1_HUMAN	P50238	TLTSGGHAEHEGKPYCNHPCYAAMFGPK + Oxidation (M); Phospho (ST)	47	32
721	CREB1_HUMAN	P16220	ILNDLSSDAPGVPR + Phospho (ST)	30	30
722	CQ085_HUMAN	Q53F19	AEAPAGPALGLPSPEAESGVDR + Phospho (ST)	70	32
723	CQ062_HUMAN	Q9BQA9	LITSFLELHCLESPTELSQSSDSEAGDPASQS + Phospho (ST)	51	33

724	CQ049_HUMAN	Q8IXM2	KVASGVLSPPPAAPPPSSSSVPEAGGPPIKK + Phospho (ST)	57	29
			KVYEDSGIPLPAESPK + Phospho (ST)	82	31
			VYEDSGIPLPAESPK + Phospho (ST)	42	31
725	CQ028_HUMAN	Q8IV36	TGSQEGTSMEGSRPAAPAEPGTLK + Phospho (ST)	95	32
726	CPSF7_HUMAN	Q8N684	DSSDSADGRATPSENLVPSSAR + Phospho (ST)	38	32
727	CPSF2_HUMAN	Q9P2I0	EADIDSSDESDIEEDIDQPSAHK + 3 Phospho (ST)	42	27
728	CPIN1_HUMAN	Q6FI81	KSSPSVKPAVDPAAAK + Phospho (ST)	38	30
729	CP088_HUMAN	Q1ED39	YSVLNNDDYFADVSPLR + Phospho (ST)	34	31
730	COASY_HUMAN	Q13057	LASVLLYSDYGIGEVPVEPLDVPLPSTIRPASPVAGSPK + Phospho (ST)	36	27
731	CO052_HUMAN	Q6ZUT6	SPPTQVAISSDSAR + Phospho (ST)	35	30
732	CNOT4_HUMAN	O95628	SPFEGAVTESQSLFSDNFR + Phospho (ST)	86	32
733	CND3_HUMAN	Q9BPX3	TLHCEGTEINSDDEQESKEVEETATAK + Phospho (ST)	39	32
734	CNBP_HUMAN	P62633	GFQFVSSSLPDICYR + Phospho (ST)	69	31
735	CN133_HUMAN	Q9H9C1	TRPGSFQSLSDALSDTPAK + 2 Phospho (ST)	35	31
			TRPGSFQSLSDALSDTPAK + Phospho (ST)	54	31
736	CLN3_HUMAN	Q13286	FSDSEGEETVPEPR + Phospho (ST)	64	30
			RFSDSEGEETVPEPR + 2 Phospho (ST)	72	29
			RFSDSEGEETVPEPR + Phospho (ST)	63	30
737	CLMN_HUMAN	Q96JQ2	TSHSDSSIYLR + Phospho (ST)	75	30
738	CLK3_HUMAN	P49761	YRSPEPDPYLSYR + Phospho (ST)	32	31
739	CLK2_HUMAN	P49760	DRGDAYYDTDYRHSYEYQR + Phospho (ST)	52	30
740	CLH1_HUMAN	Q00610	TSIDAYDNFDNISLAQR + Phospho (ST)	72	32
741	CLD3_HUMAN	O15551	STGPGASLGTGYDR + 2 Phospho (ST)	81	27
			STGPGASLGTGYDR + Phospho (ST)	82	30
742	CLCN7_HUMAN	P51798	VGHMSSVELDDELLDPDMDPPHPFPK + Phospho (ST)	65	33

743	CL043_HUMAN	Q96C57	EAAVSASDILQESAIHSPGTVEK + Phospho (ST)	38	32
744	CIP4_HUMAN	Q15642	APSDSSLGTPSDGRPELR + 2 Phospho (ST)	32	31
745	CIC_HUMAN	Q96RK0	FAELPEFRPEEVLPSPTLQSLATSPR + 2 Phospho (ST)	64	32
746	CI129_HUMAN	Q5T035	NLTEQNSYSNIPHEGKHTPLYER + Phospho (ST)	44	33
747	CI075_HUMAN	Q4KMQ1	WQRPSSPPPFLPAASEEAEPAEGLR + Phospho (ST)	45	32
748	CHP1_HUMAN	Q99653	ASTLLRDEELEEIKK + Phospho (ST)	41	30
749	CHMP3_HUMAN	Q9Y3E7	VTDALPEPEPPGAMAASEDEEEEEEALEAMQSR + Oxidation (M); Phospho (ST)	58	32
			VTDALPEPEPPGAMAASEDEEEEEEALEAMQSR + Phospho (ST)	109	33
750	CHERP_HUMAN	Q8IWX8	SRSPTPPSSAGLGSNSAPPIPDSR + 3 Phospho (ST)	45	31
751	CHD8_HUMAN	Q9HCK8	HFSTLKDDDLVEFSDLESEDDERPR + Phospho (ST)	67	33
			TASPLPLRPDAPVEK + Phospho (ST)	40	30
			TASPLPLRPDAPVEKSPEETATQVPSLESLTLK + 2 Phospho (ST)	67	31
752	CHD7_HUMAN	Q9P2D1	CEGKEEEEETDGSGKESK + Phospho (ST)	34	29
753	CHD4_HUMAN	Q14839	EDNSEGEEILEEVGGDLEEEDDHHMEFCR + Oxidation (M); Phospho (ST)	85	30
			WGQPPSPTPVPRPPDADPNTPSPKPLEGRPER + 2 Phospho (ST)	39	33
			WGQPPSPTPVPRPPDADPNTPSPKPLEGRPER + Phospho (ST)	33	32
754	CHD3_HUMAN	Q12873	ELQGDGPPSSPTNDPTVK + Phospho (ST)	74	31
			VELSPMQKK + Oxidation (M); Phospho (ST)	30	29
755	CG050_HUMAN	Q9BRJ6	ELDEEGSDPPLPGR + Phospho (ST)	37	30
756	CFDP1_HUMAN	Q9UEE9	LDWESFKEEEGIGEELAIHNR + Phospho (ST)	38	32
757	CF203_HUMAN	Q9P0P8	VDEEDSDEESHHDEMSEQEEELEDDPTVVK + 2 Phospho (ST)	40	30
			VDEEDSDEESHHDEMSEQEEELEDDPTVVK + Oxidation (M); 2 Phospho (ST)	41	29
758	CENPJ_HUMAN	Q9HC77	IERESALEK + Phospho (ST)	32	30
759	CE170_HUMAN	Q5SW79	SESLDPDSSMDTTLILK + Phospho (ST)	52	32
			SIKSDVPVYLK + 2 Phospho (ST)	38	30

760	CDYL1_HUMAN	Q9Y232	TAVDGFQSESPEKLDPVEQGQEDTVAPEVAAEKPVGALLGPGAER + Phospho (ST)	50	31
761	CDS1_HUMAN	Q92903	EGEAAGGDHETESTSDKETDIDDR + Phospho (ST)	108	30
762	CDR2_HUMAN	Q01850	SSSETILSSLAGSDIVK + Phospho (ST)	36	31
763	CDC2_HUMAN	P06493	IGEGTYGVVYK + Phospho (Y)	68	29
			IGEGTYGVVYKGR + Phospho (ST); Phospho (Y)	50	30
			VYTHEVVTLWYR + Phospho (ST)	51	30
764	CD45_HUMAN	P08575	NRNSNVIPYDYNR + Phospho (ST)	59	31
765	CD44_HUMAN	P16070	SQEMVHLVNK + Phospho (ST)	58	29
766	CD2L7_HUMAN	Q9NYV4	AITPPQQPYK + Phospho (ST)	42	30
			DSKPIALKEEIVTPK + Phospho (ST)	38	29
			ESKGSPVFLPR + Phospho (ST)	31	30
			GSPVFLPR + Phospho (ST)	32	28
			HLLTDLPLPPELPGGDLSPPDSPEPK + 2 Phospho (ST)	67	33
			HLLTDLPLPPELPGGDLSPPDSPEPK + Phospho (ST)	57	31
			HSSISPVRLPLNSSLGAELSR + 2 Phospho (ST)	36	31
			LYNSEESRPYTNK + Phospho (ST)	72	30
			NSSPAPPQPAPGK + Phospho (ST)	63	30
			RQSVSPPYKEPSAYQSSTR + 2 Phospho (ST)	47	32
			RQSVSPPYKEPSAYQSSTR + Phospho (ST); Phospho (Y)	46	32
			RTPTMPQEEAAACPPHILPPEK + Oxidation (M); Phospho (ST)	39	32
			RTPTMPQEEAAACPPHILPPEK + Phospho (ST)	42	32
767	CD2L1_HUMAN	P21127	AYTPVVVTLWYR + Phospho (ST)	64	30
			DLLSDLQDISDSER + Phospho (ST)	62	31
			GTSPRPPEGGLGYSQLGDDDLK + Phospho (ST)	46	32

768	CD2AP_HUMAN	Q9Y5K6	TSSSETEEKKPEKPLILQSLGPK + Phospho (ST)	53	31
769	CCNL1_HUMAN	Q9UK58	GLNPDGTPALSTLGGFSPASKPSSPR + 2 Phospho (ST)	41	32
770	CCDC9_HUMAN	Q9Y3X0	EGAASPAPETPQPTSPETSPK + 2 Phospho (ST)	37	31
771	CCD94_HUMAN	Q9BW85	LLEDSDSEDEAAPSPLQPALRPNPTAILDEAPKPK + 2 Phospho (ST)	72	32
772	CCD86_HUMAN	Q9H6F5	ALVEFESNPEETREPGSPPSVQR + Phospho (ST)	37	32
			LGGLRPESPESLTSVSR + Phospho (ST)	36	31
			LQQGAGLESPQGQPEPGAASPQR + Phospho (ST)	77	32
773	CCD55_HUMAN	Q9H0G5	VEENPDADSDFDAKSSADDEIEETR + 3 Phospho (ST)	37	28
774	CCD43_HUMAN	Q96MW1	AALLAQYADVTDEEDEADEKDDSGATTMNIGSDK + Phospho (ST)	66	33
775	CC85C_HUMAN	A6NKD9	DVGDGSSTSSAGSGGSPDHHHHVPPPLLPPGPHK + Phospho (ST)	34	33
776	CC132_HUMAN	Q96JG6	SAYQEYDSDSDVPEELKR + 2 Phospho (ST)	71	29
			SAYQEYDSDSDVPEELKR + Phospho (ST)	46	31
			SAYQEYDSDSDVPEELKR + Phospho (ST); Phospho (Y)	40	30
777	CBX1_HUMAN	P83916	KADSDSEDKGEESKPK + Phospho (ST)	38	31
778	CBPD_HUMAN	O75976	SLLSHEFQDETDTEEETLYSSK + 2 Phospho (ST)	97	31
779	CB055_HUMAN	Q6NV74	ALSHDSIFIPESGQDATRPVR + Phospho (ST)	76	31
780	CASZ1_HUMAN	Q86V15	TPALAALAALGAPGPAPTAASSP + Phospho (ST)	61	30
781	CASC3_HUMAN	O15234	DPSPEADAPVLGSPEK + Phospho (ST)	38	31
			DPSPEADAPVLGSPEKEEAASEPPAAAPDAAPPPPDRPIEK + 2 Phospho (ST)	38	33
782	CALU_HUMAN	O43852	VHNDAQSFDYDHDAFLGAEEAK + Phospho (ST)	86	32
783	CABL1_HUMAN	Q8TDN4	SSLETLEDIEENAPLR + Phospho (ST)	51	31
784	CA226_HUMAN	A1L170	ASSPSLIER + Phospho (ST)	39	29
			RASSPSLIER + 2 Phospho (ST)	33	29
			SFPHLSKPVAPGSAPLGSGEPGGPGLWVGSSQHLK + Phospho (ST)	39	31
785	BYST_HUMAN	Q13895	MPQDGSDDEDEEWPTLEK + Phospho (ST)	36	30

786	BUD13_HUMAN	Q9BRD0	ARHDSPDLAPNVTYSLPR + Phospho (ST)	40	31
			HDSPDLAPNVTYSLPR + Phospho (ST)	35	31
787	BRE1A_HUMAN	Q5VTR2	ALVVPEPEPDSDSNQER + Phospho (ST)	41	31
788	BRD3_HUMAN	Q15059	SESPPPLSDPK + Phospho (ST)	31	30
789	BRD2_HUMAN	P25440	KADTTTPTPTAILAPGSPASPPGSLEPK + 2 Phospho (ST)	49	32
790	BRAF1_HUMAN	P15056	SSSAPNVHINTIEPVNIDDLIR + Phospho (ST)	58	32
791	BORG2_HUMAN	Q9UKI2	ANSTSDSVFTETPSPVLK + Phospho (ST)	43	31
792	BOP1_HUMAN	Q14137	IGDEYAEDSSDEEDIR + 2 Phospho (ST)	66	26
793	BNIPL_HUMAN	Q7Z465	RLSAPELR + Phospho (ST)	31	29
794	BNIP2_HUMAN	Q12982	KGSITEYTAAEEK + Phospho (ST)	51	30
795	BNI3L_HUMAN	O60238	DHSSQSEEEVVEGEKEVEALKK + Phospho (ST)	42	32
796	BMS1_HUMAN	Q14692	LGPQNFIDETSDIENLLK + Phospho (ST)	73	32
797	BL1S3_HUMAN	Q6QNY0	VAGEAAETDSEPEPEPEPTAAPR + 2 Phospho (ST)	55	31
			VAGEAAETDSEPEPEPEPTAAPR + Phospho (ST)	47	32
798	BIG3_HUMAN	Q5TH69	AGGGDLLLPPSPK + Phospho (ST)	47	29
			ATGSAGLLGDPECEGSPPEHSPEQGR + Phospho (ST)	51	32
			GQDSPLLQRPQHLMDQGQMR + Phospho (ST)	64	32
			HSFSAGPELLR + Phospho (ST)	34	30
			SDVSDIGSDNCSLADEEQTPR + 2 Phospho (ST)	48	29
			SDVSDIGSDNCSLADEEQTPR + Phospho (ST)	56	30
			YSESNFSVDDQDLSR + Phospho (ST)	59	29
799	BI2L1_HUMAN	Q9UHR4	TPASTPVSGTPQASPMIER + 2 Phospho (ST)	33	31
			TPASTPVSGTPQASPMIER + Oxidation (M); Phospho (ST)	33	32
			TPASTPVSGTPQASPMIER + Phospho (ST)	40	31
800	BET1L_HUMAN	Q9NYM9	AQSPGAVEEILDR + Phospho (ST)	56	30

801	BCR_HUMAN	P11274	HQDGLPYIDDSPSSSPHLSSK + Phospho (ST)	69	32
802	BCL9_HUMAN	O00512	SSTPSHGQTTATEPTPAQK + Phospho (ST)	46	31
803	BCL7C_HUMAN	Q8WUZ0	GTEPSPGGTPQPSRPVSPAGPPEGVPEEAQPPR + 2 Phospho (ST)	54	33
			GTEPSPGGTPQPSRPVSPAGPPEGVPEEAQPPR + Phospho (ST)	41	33
804	BBX_HUMAN	Q8WY36	TADGRVSPAGGTLDDKPK + Phospho (ST)	60	31
805	BAT3_HUMAN	P46379	APPQTHLPSGASSGTGSASATHGGGSPPGTR + Phospho (ST)	50	33
			ENASPAPGTTAEEAMSR + Oxidation (M); Phospho (ST)	82	30
806	BAT2L_HUMAN	Q5JSZ5	HIISATSLSTSPTELGSR + Phospho (ST)	45	31
			LKFSDDEEEEEVVK + Phospho (ST)	74	31
807	BAT2_HUMAN	P48634	SEGSEYEEIPK + Phospho (ST)	45	29
			SEGSEYEEIPKR + 2 Phospho (ST)	39	28
808	BAIP2_HUMAN	Q9UQB8	SSSTGNLLDKDDLAIPPPDYGAASR + Phospho (ST)	33	32
809	BAD_HUMAN	Q92934	RMSDEFVDSFK + Phospho (ST)	57	29
810	BA2D1_HUMAN	Q9Y520	SVEDVRPHHTDANNQSACFEAPDQK + Phospho (ST)	135	32
811	B3A2_HUMAN	P04920	NISAGSLGSLLGHHHGQGAESDPHVTEPLMGGVPETR + Oxidation (M); Phospho (ST)	59	33
812	ATX2L_HUMAN	Q8WWM7	EVDGLLTSEPMGSPVSSK + Phospho (ST)	57	31
			LQPSSSPENSLDPFPPR + Phospho (ST)	52	31
			TESVSDKEDKPPLAPSGGTEGPEQPPPPCPSQTGSPPVGLIK + Phospho (ST)	78	32
813	ATX2_HUMAN	Q99700	TSPSGGTWSSVVSGVPR + Phospho (ST)	36	31
814	ATRX_HUMAN	P46100	LTVSDGESGEEKK + 2 Phospho (ST)	41	28
			RPTETNPVTSNSDEECNETVK + 2 Phospho (ST)	50	30
815	ATLA2_HUMAN	Q8NHH9	TSDPSAAVNHVSSTTSLGENYEDDDLVNSDEVMK + Phospho (ST)	62	33
816	ATF2_HUMAN	P15336	MPLDLSPLATPIIR + Phospho (ST)	56	29
817	ATAD1_HUMAN	Q8NBU5	EYVNSTSEESHDEDEIRPVQQQDLHR + Phospho (ST)	37	33

818	AT2B1_HUMAN	P20020	IEDSEPHIPLIDDTDAEDDAPTK + Phospho (ST)	67	32
			IEDSEPHIPLIDDTDAEDDAPTKR + Phospho (ST)	71	33
			SSIHNFMTHPEFR + Phospho (ST)	38	30
819	AT133_HUMAN	Q9H7F0	LVHDSLEDLQMTR + Phospho (ST)	71	31
820	AT131_HUMAN	Q9HD20	DSPTLSNSGIR + Phospho (ST)	48	29
821	ASPC1_HUMAN	Q9BZE9	AAGSPSPLPAPDPAPK + Phospho (ST)	63	30
822	ASML_HUMAN	O95671	HDSIPAADTFEDLSDVEGGGSEPTQR + Phospho (ST)	94	32
823	ARVC_HUMAN	O00192	SLAADDEGGPELEPDYGTATR + Phospho (ST)	47	31
824	ARMX3_HUMAN	Q9UH62	YNDWSDDDDDSNESK + Phospho (ST)	43	24
825	ARM10_HUMAN	Q8N2F6	SAGALEEGTSEGQLCGR + Phospho (ST)	50	30
826	ARIP4_HUMAN	Q9Y4B4	APDPEGLARPVSPDSPEIISELQQYADVAAAR + 2 Phospho (ST)	55	33
827	ARI4B_HUMAN	Q4LE39	DIEVLSEDTDYEEDEVTK + 2 Phospho (ST)	39	29
828	ARI1A_HUMAN	O14497	GPSPSPVGSPASVAQSR + Phospho (ST)	66	31
			VSSPAPMEGGEEEEELLGPK + Phospho (ST)	72	32
829	ARHG2_HUMAN	Q92974	ILSQSTDSLNMR + Oxidation (M); Phospho (ST)	42	30
830	ARFP1_HUMAN	P53367	DLKHSLPSGLGLSETQITSHGFDNTK + Phospho (ST)	72	32
			HSLPSGLGLSETQITSHGFDNTK + Phospho (ST)	40	32
			LAQQGSDLIVPAGGQR + Phospho (ST)	45	30
831	ARFG2_HUMAN	Q8N6H7	AISSDMFFGR + Oxidation (M); Phospho (ST)	42	27
			AISSDMFFGR + Phospho (ST)	54	28
832	ARAP1_HUMAN	Q96P48	LFPEFDDSDYDEVPEEGPGAPAR + Phospho (ST)	53	31
			LFPEFDDSDYDEVPEEGPGAPAR + Phospho (Y)	32	31
833	ARAF_HUMAN	P10398	STSTPNVHMVSTTAPMDSNLIQLTGQSFSTDAAGSR + 2 Oxidation (M); Phospho (ST)	57	33
834	AR6P4_HUMAN	Q66PJ3	SAGEEEDGPVLTDEQK + Phospho (ST)	71	30

835	APT_HUMAN	P07741	IDYIAGLDSR + Phospho (ST)	38	30
836	AP3D1_HUMAN	O14617	HRPSEADEEELAR + Phospho (ST)	47	30
			NTETSKSPEKDVPMVEK + Phospho (ST)	65	31
837	AP3B1_HUMAN	O00203	NFYESDDDQKEK + Phospho (ST)	48	28
838	ANXA2_HUMAN	P07355	GLGTDEDSLIEIICSR + Phospho (ST)	44	31
839	ANS1A_HUMAN	Q92625	SDSDLLTCSPTEDATMGSR + Phospho (ST)	66	30
			SPSFASEWDEIEK + Phospho (ST)	36	29
840	ANR57_HUMAN	Q53LP3	DLVMGSSPQLK + Oxidation (M); Phospho (ST)	38	29
841	ANR17_HUMAN	O75179	EHYPVSSPSSPSPPAQPGGVSR + 2 Phospho (ST)	57	31
			NSPLDCGSASPNK + Phospho (ST)	47	28
842	ANM3_HUMAN	O60678	GAVENEEDLPELSDSGDEAAWEDEDDADLPHGK + 2 Phospho (ST)	75	31
843	ANKL2_HUMAN	Q86XL3	NNSPPTVGAFGHTR + Phospho (ST)	57	30
844	AMPD2_HUMAN	Q01433	TDSDSDLQLYK + Phospho (ST)	57	29
845	ALS2_HUMAN	Q96Q42	RLSLPGLLSQVSPR + 2 Phospho (ST)	39	30
			SESPEPGYVVTSSGLLLPVLLPR + Phospho (ST)	113	29
846	ALG3_HUMAN	Q92685	SGSAAQAEGLCK + Phospho (ST)	32	28
847	ALBU_HUMAN	P02768	KVPQVSTPTLVEVSR + Phospho (ST)	44	29
			TCVADESAENCDK + Phospho (ST)	76	26
848	AKT2_HUMAN	P31751	YFDDEFTAQSITITPPDR + Phospho (ST)	52	32
849	AKA11_HUMAN	Q9UKA4	SVSPTFLNPSDENLK + Phospho (ST)	85	31
850	AIM1_HUMAN	Q9Y4K1	DTCVQSPISSFPCTDLK + Phospho (ST)	48	31
851	AHTF1_HUMAN	Q8WYP5	NLSFNELYPSGTLK + Phospho (ST)	34	31
852	AG2_HUMAN	Q7Z7L8	SFLQSLECLR + Phospho (ST)	36	30
853	AFF4_HUMAN	Q9UHB7	ELLSPLSEPDDRYPLIVK + 2 Phospho (ST)	49	32
			SSSPGKPQAVSSLNSSHSR + Phospho (ST)	82	32

			VNPHKVSPASSVDSNIPSSQGYK + Phospho (ST)	62	32
854	ADNP_HUMAN	Q9H2P0	LMHNASDSEVDQDDVVEWK + 2 Phospho (ST)	30	29
855	ADA17_HUMAN	P78536	SFEDLTDHPVTR + Phospho (ST)	53	30
856	ACTB_HUMAN	P60709	TTGIVMDSGDGVTHTVPIYEGYALPHAILR + Oxidation (M); Phospho (ST)	66	31
857	ACOD_HUMAN	O00767	EKGSTLDLSDLEAEK + Phospho (ST)	61	31
858	ACLY_HUMAN	P53396	AKPAMPQDSVPSPR + Oxidation (M); Phospho (ST)	57	30
			AKPAMPQDSVPSPR + Phospho (ST)	71	31
859	ACINU_HUMAN	Q9UKV3	AESPAEKVPEESVLPLVQK + Phospho (ST)	84	30
			EASSPPPHPQLHSEEEIEPMEGPAPPVLIQLSPPNTDADTR + 2 Phospho (ST)	73	33
			EASSPPPHPQLHSEEEIEPMEGPAPPVLIQLSPPNTDADTR + 3 Phospho (ST)	41	33
			LSEGSQPAEEEEDQETPSR + 2 Phospho (ST)	38	28
			LSEGSQPAEEEEDQETPSR + Phospho (ST)	50	30
			SLIPDIKPLAGQEAVVDLHADDSRISEDETER + Phospho (ST)	67	32
			SLSPGVSR + 2 Phospho (ST)	34	25
			TTSPLEEEER + Phospho (ST)	41	29
860	ACBD5_HUMAN	Q5T8D3	AESSDSGAESEEEEAQEEVK + 2 Phospho (ST)	59	26
861	ACAP2_HUMAN	Q15057	YSISLSPPEQQK + Phospho (ST)	35	30
862	ACACA_HUMAN	Q13085	FIIGSVSEDNSEDEISNLVK + Phospho (ST)	126	32
			IFDEVMGCFSDSPPQSPTFPEAGHTSLYDEDKVPR + 2 Phospho (ST)	61	32
863	ABL2_HUMAN	P42684	GAQASSGSPALPR + Phospho (ST)	65	29
864	AB1IP_HUMAN	Q7Z5R6	SSDTSGSPATPLK + Phospho (ST)	40	29
865	AATF_HUMAN	Q9NY61	YLVDGTKPNAGSEEISSEDDELVEEK + 2 Phospho (ST)	34	32
			YLVDGTKPNAGSEEISSEDDELVEEK + 3 Phospho (ST)	34	31
866	AAPK1_HUMAN	Q13131	DFYLATSPPDSFLDDHHLTRPHPER + Phospho (ST)	63	33
			TSCGSPNYAAPEVISGR + Phospho (ST)	58	31

867	AAKB1_HUMAN	Q9Y478	SHNNFVAILDLPEGEHQYK + Phospho (ST)	84	32
868	AAK1_HUMAN	Q2M2I8	STQLLQAAAAEASLNK + Phospho (ST)	58	30
			VGSLTPPSSPK + 2 Phospho (ST)	30	29
			VQTTPPPAVQGQK + Phospho (ST)	31	30
869	AAAS_HUMAN	Q9NRG9	FSPVLGR + Phospho (ST)	31	29
870	4ET_HUMAN	Q9NRA8	APSPPLSQVFQTR + Phospho (ST)	41	30
871	4EBP1_HUMAN	Q13541	DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR + Phospho (ST)	38	33
			RVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y)	37	33
			TPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST)	106	33
872	2A5D_HUMAN	Q14738	RKSELPQDVYTIK + Phospho (ST)	40	30
873	1A02_HUMAN	P01892	KGGSYSQAASSDSAQGSDVSLTACK + 2 Phospho (ST)	57	31
			KGGSYSQAASSDSAQGSDVSLTACKV + 2 Phospho (ST)	58	31
874	1433Z_HUMAN	P63104	TAFDEAIAELDTLSEESYK + Phospho (ST)	105	32
875	1433E_HUMAN	P62258	AAFDDAIAELDTLSEESYK + Phospho (ST)	74	31

Table 7.4 Summary of the 297 phosphoproteins possibly related to the genesis of breast cancer identified in this work.

#	Name	Swiss-Prot ID	Unique Peptide Sequence	Score	Mascot score for identity
1	ZN652_HUMAN	Q9Y2D9	ESGSPYSVLVDTK + Phospho (ST)	34	30
2	ZN446_HUMAN	Q9NWS9	TEEPLGSPHPSGTVESPGEGPQDTR + Phospho (ST)	41	32
3	ZN318_HUMAN	Q5VUA4	DISPEKSELDLGEPGPPGVEPPPQLLDIQCK + Phospho (ST)	34	32
4	ZN260_HUMAN	Q3ZCT1	TFSLKQNLVEHKK + Phospho (ST)	30	30
5	ZF106_HUMAN	Q9H2Y7	AAHVPENSDTEQDVLTVKPVR + Phospho (ST)	66	31

6	ZEP2_HUMAN	P31629	MLKGISSSSLK + Oxidation (M); Phospho (ST)	30	29
7	ZCCHV_HUMAN	Q7Z2W4	SCTPSPDQISHR + 2 Phospho (ST)	46	27
			SCTPSPDQISHR + 3 Phospho (ST)	31	23
			SSLGSLQTPEAVTTR + Phospho (ST)	41	30
8	ZC3HE_HUMAN	Q6PJT7	ISPPIKEEETKGDSVEK + Phospho (ST)	53	31
9	ZC3H1_HUMAN	O60293	KPISDNSFSSDEEQSTGPIK + 2 Phospho (ST)	67	31
10	YYAP1_HUMAN	Q9H869	LEPQELSPLSATVFPK + Phospho (ST)	31	30
11	YTDC2_HUMAN	Q9H6S0	VDGIPNDSSDSEMEDK + Phospho (ST)	39	29
12	YJ005_HUMAN	Q6ZSR9	KKSMEELTVIQCTSQELPAQTGLLSQTGDVPLPAGR + Phospho (ST)	43	31
13	YD003_HUMAN	Q9NXL2	RKTDTVVESSVSGDHSGTLR + Phospho (ST)	107	31
14	XIRP2_HUMAN	A4UGR9	DGSGQMLEIK + Phospho (ST)	30	27
15	WWTR1_HUMAN	Q9GZV5	SHSSPASLQLGTGAGAAGSPAQQHAHLR + Phospho (ST)	95	32
16	WRN_HUMAN	Q14191	ICALTK + Phospho (ST)	30	28
17	WDR75_HUMAN	Q8IWA0	EIPEDVDMEEEKESEDSDEENDFTEK + 2 Phospho (ST)	30	29
18	WDR55_HUMAN	Q9H6Y2	TCEERPAEDGSDEEDPDSMEAPTR + Phospho (ST)	67	29
19	VPRBP_HUMAN	Q9Y4B6	HLPSPPTLDSIITEYLR + Phospho (ST)	44	30
20	UTP18_HUMAN	Q9Y5J1	TSSDDESEEDEDDLLQR + 2 Phospho (ST)	51	26
			VQEHEDSGDSEVENEAK + 2 Phospho (ST)	31	27
21	UT14A_HUMAN	Q9BVJ6	DYLLSESEDEGDNDGER + 2 Phospho (ST)	40	26
22	URFB1_HUMAN	Q6BDS2	TVSQQSFDGVSLDSSGPEDR + 2 Phospho (ST)	49	29
23	UNK_HUMAN	Q9C0B0	NSSLGSPSNLCGSPPGSIR + Phospho (ST)	64	31
24	UNG_HUMAN	P13051	HAPSPEPAVQGTGVAGVPEESGDAAAIPAK + Phospho (ST)	83	32
			KAPAGQEEPGTPPSSPLSAEQLDR + 2 Phospho (ST)	40	32
25	UFO_HUMAN	P30530	ESFPAPVVILPFMK + Phospho (ST)	38	30

26	UBXN7_HUMAN	O94888	SESLIDASEDSQLEAAIR + Phospho (ST)	65	31
27	UBR5_HUMAN	O95071	WLDGASFDNER + Phospho (ST)	41	28
28	UBP40_HUMAN	Q9NVE5	RKSQEALHEQSSYILSSAETPARPR + Phospho (ST)	65	32
29	UBP16_HUMAN	Q9Y5T5	NINMDNDLEVLTSSPTR + Oxidation (M); Phospho (ST)	61	31
30	U3IP2_HUMAN	O43818	MNEEISSDSESESLAPR + 3 Phospho (ST)	30	26
31	TYDP1_HUMAN	Q9NUW8	WTISSSDESEEEKPKPDKPSTSSLLCAR + Phospho (ST)	64	33
32	TXLNA_HUMAN	P40222	SSPGQPEAGPEGAQERPSQAAPAVEAEGPGSSQAPR + Phospho (ST)	34	33
33	TUSC2_HUMAN	O75896	RGSMFYDEDGDLAHEFYEETIVTK + Phospho (ST)	122	32
34	TSYL1_HUMAN	Q9H0U9	YITNLEVKELR + Phospho (ST)	45	29
35	TRUA_HUMAN	Q9Y606	VPSPLEGSEGDGDTD + Phospho (ST)	49	27
36	TREF1_HUMAN	Q96PN7	SSPSHSTTSGETDPTTIFPCK + Phospho (ST)	32	31
37	TRA2B_HUMAN	P62995	RPHTPTPGIYMGRPTYGSSR + Oxidation (M); Phospho (ST)	50	32
			RSPSPYYSR + Phospho (ST); Phospho (Y)	34	27
38	TOX4_HUMAN	O94842	LSTTPSPTSSLHEDGVEDFR + 2 Phospho (ST)	33	31
			LSTTPSPTSSLHEDGVEDFRR + 2 Phospho (ST)	36	32
39	TOE1_HUMAN	Q96GM8	AADSDDGAVSAPAASDGGVSK + Phospho (ST)	37	31
40	TOB1_HUMAN	P50616	TSPINLGLNVNDLLK + Phospho (ST)	42	30
41	TM63B_HUMAN	Q5T3F8	LTSVSSSVDFDQR + Phospho (ST)	45	30
42	TISD_HUMAN	P47974	LHHSLSFSGFPSGHHQPPGGLESPLLLDSPTSR + Phospho (ST)	55	32
43	TIM8A_HUMAN	O60220	SKPVFSESLSD + Phospho (ST)	41	29
44	TICN2_HUMAN	Q92563	QNGSASSVAGPASGLDK + Phospho (ST)	30	30
45	THOC1_HUMAN	Q96FV9	TGEDEDEEDNDALLKENESPDVR + Phospho (ST)	66	31
46	TENS3_HUMAN	Q68CZ2	WDSYENLSADGEVLHTQGPVDGSLYAK + Phospho (Y)	79	33
47	TCEA3_HUMAN	O75764	GLECSDWKPEAGLSPPR + Phospho (ST)	37	31
48	TBC30_HUMAN	Q9Y2I9	RRDSLDSSTEASGSDVVLGGR + Phospho (ST)	66	32

49	TBC15_HUMAN	Q8TC07	NDSPTQIPVSSDVCR + Phospho (ST)	77	30
50	TBA1A_HUMAN	Q71U36	TIGGGDDSFNTFFSETGAGK + Phospho (ST)	75	31
51	SYNJ1_HUMAN	O43426	SSPNPFITGLTR + Phospho (ST)	41	30
			TSPCQSPTISEGPVPSLPIRPSR + 2 Phospho (ST)	33	33
52	SYNE2_HUMAN	Q8WXH0	RGSMSYLAAVEEEVEESSVK + Phospho (ST)	50	32
53	SYMC_HUMAN	P56192	TSPKPAVVETVTTAKPQQIQALMDEVTK + Phospho (ST)	93	30
54	SYIC_HUMAN	P41252	APLKPYPVSPSDK + Phospho (ST)	31	30
55	STUB1_HUMAN	Q9UNE7	LGAGGGSPEKSPSAQELK + Phospho (ST)	42	31
56	STAT3_HUMAN	P40763	FICVTPTTCSNTIDLPMSPR + Phospho (ST)	38	32
57	STAR3_HUMAN	Q14849	GPLLFSGALSEGQFYSPPESFAGSDNESDEEVAGKK + 2 Phospho (ST)	50	33
58	SSF1_HUMAN	Q9NQ55	VGGSDEEASGIPSR + Phospho (ST)	71	29
59	SRPR_HUMAN	P08240	GTGSGGQLQDLDCSSSDDEGAAQNSTKPSATK + 3 Phospho (ST)	50	30
60	SRPK2_HUMAN	P78362	TVSASSTGDLPK + 2 Phospho (ST)	31	28
61	SRBS2_HUMAN	O94875	TSPGRVDLPGSSTTLTK + Phospho (ST)	36	30
62	SR140_HUMAN	O15042	SEEHHLYSNPIKEEMTESK + Phospho (ST)	34	31
63	SPN90_HUMAN	Q9NZQ3	RGPSASSVAVMTSSTSDHHLDAAAAR + Phospho (ST)	42	33
64	SP30L_HUMAN	Q9HAJ7	TSDDGGDSPEHDTDIPEVDLFQLQVNTLR + Phospho (ST)	62	33
65	SNX15_HUMAN	Q9NRS6	EEGAAPSPTHVAELATMEVESAR + Phospho (ST)	52	32
66	SNPC4_HUMAN	Q5SXM2	VGSESEDEDLLSELELADR + 2 Phospho (ST)	35	31
67	SNIP1_HUMAN	Q8TAD8	RPDHSGGSPSPPTSEPAR + 2 Phospho (ST)	40	30
68	SMRCD_HUMAN	Q9H4L7	ANTPDSDITEK + 2 Phospho (ST)	34	25
			KLSSSSEPYEEDEFNDDQSIKK + 2 Phospho (ST)	50	32
69	SMRC1_HUMAN	Q92922	KHSPSPPPPTPTESR + 2 Phospho (ST)	60	30
70	SMCA5_HUMAN	O60264	GGPEGVAAQAVASAASAGPADAEMEEIFDDASPGK + Phospho (ST)	132	33

71	SMC4_HUMAN	Q9NTJ3	EEGPPPPSPDGASSDAEPEPPSGR + Phospho (ST)	33	32
			TESPATAAETASEELDNR + Phospho (ST)	63	31
72	SLK_HUMAN	Q9H2G2	ASSDLSIASSEEDKLSQNACILESVSEK + 2 Phospho (ST)	56	33
			RDSFIGTPYWMAPEVVMCETSK + Phospho (ST)	45	32
73	SH2B1_HUMAN	Q9NRF2	SSEDLAGPLPSSVSSSSTTSSKPK + Phospho (ST)	64	32
74	SH23A_HUMAN	Q9BRG2	SFSEDTLMDGPAR + Phospho (ST)	35	29
75	SGTA_HUMAN	O43765	SPARTPPSEEDSAEAER + 2 Phospho (ST)	55	29
			TPPSEEDSAEAER + Phospho (ST)	39	28
76	SFRS8_HUMAN	Q12872	SGVSSDNEDDDDEEDGNYLHPSLFASK + Phospho (ST)	121	31
77	SFR14_HUMAN	Q8IX01	KMSFDIIDK + Phospho (ST)	41	29
78	SF04_HUMAN	Q8IWZ8	AVQQHQHGYDSDEEVDSELGTWEHQLR + Phospho (ST)	46	33
79	SAM14_HUMAN	Q8IZD0	VTDGCGSPLHRLR + Phospho (ST)	36	30
80	S39A3_HUMAN	Q9BRY0	EKPSFIDLETFNAGSDVGSDSEYESPFMGGAR + Oxidation (M); 2 Phospho (ST)	37	32
81	S2546_HUMAN	Q96AG3	SFSTGSDLGHWVTTPPDIPGSR + 2 Phospho (ST)	35	32
			SFSTGSDLGHWVTTPPDIPGSR + Phospho (ST)	48	32
82	S12A2_HUMAN	P55011	DEGPAAAGDGLGRPLGPTPSQSR + Phospho (ST)	36	32
83	RS10L_HUMAN	Q9NQ39	AEAGAGSATEFQFR + Phospho (ST)	30	30
84	RRP12_HUMAN	Q5JTH9	GDSIEEILADSEDEEDNEEEER + Phospho (ST)	37	29
85	RRN3_HUMAN	Q9NYV6	EGDVDVSDSDDEDDNLPANFDTCHR + 2 Phospho (ST)	37	28
86	RRMJ3_HUMAN	Q8IY81	ALDISLSSGEEDEGDEEDSTAGTTK + 2 Phospho (ST)	91	30
87	RNF34_HUMAN	Q969K3	ASLSDLSSLDDVEGMSVR + Phospho (ST)	49	31
88	RN216_HUMAN	Q9NWF9	VLPQTILYKYYER + 2 Phospho (Y)	35	31
89	RMP_HUMAN	O94763	KNSTGSGHSAQELPTIR + Phospho (ST)	31	31
90	RIR2_HUMAN	P31350	VPLAPITDPQQLQLSPLK + Phospho (ST)	53	26
			VPLAPITDPQQLQLSPLKGLSLVDKENTPPALSGTR + Phospho (ST)	47	26

91	RIOK2_HUMAN	Q9BVS4	EGSEFSFSDGEVAEK + 3 Phospho (ST)	44	23
92	RICTR_HUMAN	Q6R327	NDSGEENVPLDLTR + Phospho (ST)	39	30
93	RICS_HUMAN	A7KAX9	LSPFFTLDLSPTEDK + Phospho (ST)	54	31
			SAKSEESLTSLHAVDGDSK + Phospho (ST)	47	32
			SEESLTSLHAVDGDSK + Phospho (ST)	80	30
94	REM2_HUMAN	Q8IYK8	RASPPGTPTPEADATLLK + 2 Phospho (ST)	41	31
95	REEP1_HUMAN	Q9H902	SFSMQDLTTIR + Oxidation (M); Phospho (ST)	43	29
			SFSMQDLTTIR + Phospho (ST)	52	30
96	RCL_HUMAN	O43598	YFEADPPGQVAASPDPTT + Phospho (ST)	32	31
97	RCC1_HUMAN	P18754	SPPADAIPK + Phospho (ST)	30	29
98	RBM5_HUMAN	P52756	GLVAAYSGDSDNEEELVER + Phospho (ST)	61	31
99	RBM4_HUMAN	Q9BWF3	LHVGNISPTCTNK + Phospho (ST)	44	31
100	RBM12_HUMAN	Q9NTZ6	SRSPHEAGFCVYLK + Phospho (ST)	37	31
101	RB_HUMAN	P06400	TLQTDSIDSFETQRTPR + Phospho (ST)	31	31
102	RAI3_HUMAN	Q8NFJ5	AYSQEEITQGFEETGDTLYAPYSTHFQLQNQPPQK + Phospho (ST)	77	33
103	RAGP1_HUMAN	P46060	ILDPNTGEPAPVLSSPPPADVSTFLAFPSPEK + 2 Phospho (ST)	48	33
			KILDPNTGEPAPVLSSPPPADVSTFLAFPSPEK + 2 Phospho (ST)	38	32
104	RAD50_HUMAN	Q92878	LFDVCGSQDFESDLDR + Phospho (ST)	52	31
105	RAB12_HUMAN	Q6IQ22	AGGGGGLGAGSPALSGGQGR + Phospho (ST)	104	31
106	PUR6_HUMAN	P22234	EVYELLDSPGK + Phospho (ST)	34	30
			TKEVYELLDSPGK + Phospho (ST)	47	31
107	PSMD2_HUMAN	Q13200	APVQPQQSPAAAPGGTDEKPSGK + Phospho (ST)	64	32
			DKAPVQPQQSPAAAPGGTDEKPSGK + Phospho (ST)	85	32
108	PRKDC_HUMAN	P78527	LTPLPEDNSMNVDQDGDPSDR + Phospho (ST)	62	31
109	PRDBP_HUMAN	Q969G5	APEPLGPADQSELGPEQPEAEVGESSDEEPVESR + 2 Phospho (ST)	80	33

110	PR38B_HUMAN	Q5VTL8	RSLSPR + Phospho (ST)	30	28
			SQSIEQESQEK + 2 Phospho (ST)	36	26
111	PNMT_HUMAN	P11086	SPNAGAAPDSAPGQAAVASAYQR + Phospho (ST)	48	32
112	PKP2_HUMAN	Q99959	RLEISPDSSPER + 2 Phospho (ST)	35	28
113	PKHF2_HUMAN	Q9H8W4	SPLNDMSDDDDDDDSSD + Oxidation (M); 2 Phospho (ST)	47	18
114	PHLA2_HUMAN	Q53GA4	TAPAAPAEDAVAAAAAAPSEPSEPSRPSPQPKPR + Phospho (ST)	35	33
115	PHF8_HUMAN	Q9UPP1	DAEYIYPSLESDDDDPALK + Phospho (ST)	48	31
116	PHF3_HUMAN	Q92576	KHSDNEAESIADALSSTSNILASEFFEEEKQESPK + Phospho (ST)	52	33
117	PGAM1_HUMAN	P18669	HGESAWNLENR + Phospho (ST)	30	29
118	PDCD5_HUMAN	O14737	KVMDSDEDDDY + Phospho (ST)	49	24
119	PCBP2_HUMAN	Q15366	GVTIPYRPKPSSSPVIFAGGQDR + Phospho (ST)	81	30
120	PBIP1_HUMAN	Q96AQ6	EEGRCSSSDDDTDVDMEGLR + Oxidation (M); 3 Phospho (ST)	33	22
121	PARN_HUMAN	O95453	NNSFTAPSTVGK + Phospho (ST)	56	30
122	P80C_HUMAN	P38432	AFQLEEGEETEPDCK + Phospho (ST)	48	29
123	OSBL8_HUMAN	Q9BZF1	GYSSPEPDIQDSSGSEAQSVKPSTR + 3 Phospho (ST)	35	31
124	OSB11_HUMAN	Q9BXB4	SFSLASSSNSPISQR + Phospho (ST)	37	31
125	OCLN_HUMAN	Q16625	TEQDHYETDYTTGGESCDELEEDWIR + Phospho (ST)	110	31
126	OCAD1_HUMAN	Q9NX40	KLENSPLGEALR + Phospho (ST)	54	30
			LENSPLGEALR + Phospho (ST)	31	30
127	NONO_HUMAN	Q15233	FGQAATMEGIGAIGGTPPAFNR + Oxidation (M); Phospho (ST)	46	32
128	NOL9_HUMAN	Q5SY16	LAAFADALEFADEEKESPVEFTGHK + Phospho (ST)	66	32
129	NOL8_HUMAN	Q76FK4	EYDSGDTDEIIAMK + 2 Phospho (ST)	32	26
			FLETDSEEEQEEVNEK + 2 Phospho (ST)	36	28
			NDREYDSGDTDEIIAMK + Oxidation (M); 2 Phospho (ST)	33	28
130	NOB1_HUMAN	Q9ULX3	KDDSDDDGGGWITPSNIK + Phospho (ST)	34	31

131	NMNA1_HUMAN	Q9HAN9	LEASDCDHQQNSPTLERPGR + Phospho (ST)	41	32
132	NMD3_HUMAN	Q96D46	DSAIPVESDTDDEGAPR + 2 Phospho (ST)	37	28
			DSAIPVESDTDDEGAPR + Phospho (ST)	32	30
133	NGDN_HUMAN	Q8NEJ9	LSSEDEEEDEAEDDQSEASGK + 2 Phospho (ST)	30	26
134	NFIA_HUMAN	Q12857	ASPHATPSTLHFPTSPIIQQPGPYFSHPAIR + Phospho (ST)	59	31
			SPGSGSQSSGWHEVEPGMPSPTTLK + Phospho (ST)	86	32
			SVEDEMDSPGEEPFYTGQGR + Oxidation (M); Phospho (ST)	79	29
			SVEDEMDSPGEEPFYTGQGR + Phospho (ST)	110	30
135	NAP5_HUMAN	O14513	MDISKTKVEK + Oxidation (M); Phospho (ST)	32	29
136	MTMR4_HUMAN	Q9NYA4	SMDDLLSACDTSSPLTR + Phospho (ST)	77	31
137	MREG_HUMAN	Q8N565	ELHYLPFPSP + Phospho (ST)	35	29
138	MPP10_HUMAN	O00566	KSPVFSDEDSDLDFDISK + 2 Phospho (ST)	31	30
			SPVFSDEDSDLDFDISK + 2 Phospho (ST)	36	29
			SPVFSDEDSDLDFDISK + 3 Phospho (ST)	43	27
139	MMTA2_HUMAN	Q9BU76	RPAEATSSPTSPERPR + 2 Phospho (ST)	45	31
140	MIB2_HUMAN	Q96AX9	CLLDTDVLR + Phospho (ST)	31	29
141	MFF_HUMAN	Q9GZY8	IVVAGNNEDVSFSRPADLDLIQSTPFKPLALKTPPR + Phospho (ST)	39	28
142	MED24_HUMAN	O75448	LLSSNEDDANILSSPTDR + Phospho (ST)	49	31
143	MCM6_HUMAN	Q14566	EIESEIDSEEELINK + Phospho (ST)	50	31
144	MBD3_HUMAN	O95983	YLGGSMDLSTFDFR + Oxidation (M); Phospho (ST)	54	30
			YLGGSMDLSTFDFR + Phospho (ST)	61	30
145	MBB1A_HUMAN	Q9BQG0	ALGGEDSENEEELGDEAMMALDQSLASLFAEQK + 2 Oxidation (M); Phospho (ST)	133	33
146	MAT2B_HUMAN	Q9NZL9	YEMACAIADAFNLPSSHLRPITDSPVLGAQRPR + Phospho (ST)	50	32
147	MAST4_HUMAN	O15021	SQALGQSAPSLTASLK + Phospho (ST)	48	30
148	MARK2_HUMAN	Q7KZI7	LDTFCGSPPYAAPELFQGK + Phospho (ST)	80	32

			VPASPLPGLER + Phospho (ST)	36	29
149	MAP7_HUMAN	Q14244	AAPAQVRPPSPGNIRPVK + Phospho (ST)	33	28
			LSSSSATLLNSPDR + Phospho (ST)	65	30
150	LRRF2_HUMAN	Q9Y608	RGSGDTSSLIDPDTSLSELR + Phospho (ST)	58	32
151	LPP2_HUMAN	O43688	KPSLSLTLTLGEADHNHYGYPHSSS + Phospho (ST)	40	32
152	LMNB1_HUMAN	P20700	LKLSPSPSSR + Phospho (ST)	31	29
153	LMBL3_HUMAN	Q96JM7	LSGEMPPASPSFPR + Oxidation (M); Phospho (ST)	34	30
154	LMBL2_HUMAN	Q969R5	VKEEHLDVASPDKASSPELPVSVENIK + 2 Phospho (ST)	40	32
155	LATS1_HUMAN	O95835	SNSFNNPLGNR + Phospho (ST)	54	29
			SVTPPPPPR + Phospho (ST)	46	28
156	LAS1L_HUMAN	Q9Y4W2	MEVGPFSTGQESPTAENAR + Oxidation (M); Phospho (ST)	77	31
			MEVGPFSTGQESPTAENAR + Phospho (ST)	102	31
157	LARP5_HUMAN	Q92615	SPSPAHLPDDPK + Phospho (ST)	36	30
158	LARP4_HUMAN	Q71RC2	ASTASPCNNNINAATAVALQEPR + Phospho (ST)	48	32
159	LAP2B_HUMAN	P42167	HASPILPITEFSDIPR + Phospho (ST)	39	31
160	KS6C1_HUMAN	Q96S38	FLNRSPEESFDIK + 2 Phospho (ST)	43	30
161	KS6A1_HUMAN	Q15418	AYSFCGTVEYMAPEVVNR + Phospho (ST)	124	31
162	KPSH1_HUMAN	P11801	LLTVDPGAR + Phospho (ST)	31	29
163	KLDC4_HUMAN	Q8TBB5	SEDEDSLEEAGSPAPGPCPR + 2 Phospho (ST)	55	28
			SEDEDSLEEAGSPAPGPCPR + Phospho (ST)	37	30
164	KIF1C_HUMAN	O43896	LYADSDSGDDSDKR + 2 Phospho (ST)	58	26
165	KDM3A_HUMAN	Q9Y4C1	HLEHAPSPSDVSNAPEVK + Phospho (ST)	52	31
166	KCRB_HUMAN	P12277	VLTPELYAELR + Phospho (ST)	39	30
167	K6PP_HUMAN	Q01813	GRSFAGNLNTYK + Phospho (ST)	38	30
			SFAGNLNTYK + Phospho (ST)	52	28

168	K1467_HUMAN	A2RU67	SPLGEAPEPDSDAEVAEAAKPHLSEVTTEGYPSEPLGGLEQK + Phospho (ST)	39	33
169	K1143_HUMAN	Q96AT1	IQPQPPDEDGDHSDKEDEQPQVVVLK + Phospho (ST)	35	33
170	JUND_HUMAN	P17535	LAALKDEPQTVPDVPSFGESPPLSPIDMDTQER + Oxidation (M); 2 Phospho (ST)	52	33
171	ITPR3_HUMAN	Q14573	LGFVDVQNCISR + Phospho (ST)	32	30
			VASFSIPGSSSR + Phospho (ST)	32	29
172	ITB4_HUMAN	P16144	MTTTSAAAYGTHLSPHVPHR + Oxidation (M); Phospho (ST)	54	32
			MTTTSAAAYGTHLSPHVPHR + Phospho (ST)	48	32
173	IPP2_HUMAN	P41236	IDEPSTPYHSMMGDDEDACSDTEATEAMAPDILAR + Phospho (ST)	93	32
174	INF2_HUMAN	Q27J81	DPTSLLGVLQAEADSTSEGLEDAVHSR + 2 Phospho (ST)	40	33
			EHNSMWASLSSPDAEAVEPDFSSIER + Phospho (ST)	32	32
			GARPPAAGPGGDEDEDEEDTAPESALDTSLDK + 2 Phospho (ST)	44	32
175	IMA3_HUMAN	O00505	NVPQEESLEDSDVDADFK + Phospho (ST)	72	31
176	IMA2_HUMAN	P52292	NVSSFPDDATSPLQENR + Phospho (ST)	55	31
177	IKZF3_HUMAN	Q9UKT9	GLSPNNSGHDSTDTDSNHEER + Phospho (ST)	113	30
178	IF4G1_HUMAN	Q04637	ITKPGSIDSNNQLFAPGGR + Phospho (ST)	45	31
			SFSKEVEER + Phospho (ST)	49	29
179	IF2BL_HUMAN	A6NK07	SGDEMIFDPTMSK + Phospho (ST)	47	29
180	IBP5_HUMAN	P24593	IERDSREHEEPTTSEMAEETYSPK + Phospho (ST)	36	32
181	HOMEZ_HUMAN	Q8IX15	AETPPLPIPPPPPDIQPLER + Phospho (ST)	41	29
182	HNRPM_HUMAN	P52272	MGLAMGGGGGASFDR + 2 Oxidation (M); Phospho (ST)	47	28
			MGLSMER + Phospho (ST)	32	24
183	HNRPG_HUMAN	P38159	RDVYLSPR + Phospho (ST)	34	29

184	HN1L_HUMAN	Q9H910	GSGIFDESTPVQTR + Phospho (ST)	33	30
185	HMCS1_HUMAN	Q01581	RPTPNDDTLDEGVGLVHSNIATEHIPSPAK + Phospho (ST)	102	32
			RPTPNDDTLDEGVGLVHSNIATEHIPSPAKK + Phospho (ST)	115	32
186	HEBP2_HUMAN	Q9Y5Z4	VYYTAGYNSPVK + Phospho (ST)	37	31
187	GRB7_HUMAN	Q14451	ATSLPSIPNPFPELCSPPSQSPILGGPSSAR + Phospho (ST)	133	32
			HLHPSCLGSPPLR + Phospho (ST)	57	31
			RATSLPSIPNPFPELCSPPSQSPILGGPSSAR + 2 Phospho (ST)	62	33
			RATSLPSIPNPFPELCSPPSQSPILGGPSSAR + Phospho (ST)	151	31
			SASDNTLVAMDFSGHAGR + Oxidation (M); Phospho (ST)	144	31
			SASDNTLVAMDFSGHAGR + Phospho (ST)	149	32
			SQPLLIPTTGR + Phospho (ST)	31	29
			TNHRLSLPMPASGTSLSAAIHR + Oxidation (M); Phospho (ST)	53	31
			TNHRLSLPMPASGTSLSAAIHR + Phospho (ST)	70	31
188	GPTC4_HUMAN	Q5T3I0	DLESCSDDDNQGSK + Phospho (ST)	41	25
189	GPSM1_HUMAN	Q86YR5	APSSDEECFFDLLTK + Phospho (ST)	50	30
190	GIT2_HUMAN	Q14161	TINNQHSVESQDNDQPDYDSVASDEDTDLETTASK + 2 Phospho (ST)	32	31
191	GIMA5_HUMAN	Q96F15	SEDNLSATPPALR + Phospho (ST)	31	30
192	GCSP_HUMAN	P23378	DATGKEVYRLALQTR + Phospho (Y)	31	29
193	GCP60_HUMAN	Q9H3P7	LEVSVDGLTLSPDPEERPGAEGAPLLPPPLPPPSPPGSGR + Phospho (ST)	34	30
194	GCP6_HUMAN	Q96RT7	ARLATVGDLEEIQR + Phospho (ST)	29	29
195	GAB1_HUMAN	Q13480	SSSLEGFHNHFK + Phospho (ST)	50	29
196	FXR1_HUMAN	P51114	RGPNYTSGYGTNSELSNPSETESER + 2 Phospho (ST)	44	31
197	FXL19_HUMAN	Q6PCT2	HVVRPPPRSPEPDTLPLAAGSDHPLPR + Phospho (ST)	72	30
198	FUND1_HUMAN	Q8IVP5	NPPPQDYESDDDSYEVLDLTEYAR + Phospho (ST)	46	32
199	FRMD8_HUMAN	Q9BZ67	TTSFFSR + Phospho (ST)	33	27

200	FR1OP_HUMAN	O95684	EKGPTTGEGALDLSDVHSPPKSPEGK + Phospho (ST)	38	32
201	FNBP4_HUMAN	Q8N3X1	IDENSDKEMEVEESPEK + Phospho (ST)	55	30
202	FLNC_HUMAN	Q14315	LGSFGSITR + Phospho (ST)	45	28
203	FKBP4_HUMAN	Q02790	SNTAGSQSQVETEA + Phospho (ST)	40	28
204	FBXW9_HUMAN	Q5XUX1	TWDDDSDPESETDPDAQAK + Phospho (ST)	31	29
205	FA76B_HUMAN	Q5HYJ3	ISNLSPEEEQGLWK + Phospho (ST)	40	31
206	FA38A_HUMAN	Q92508	TASELLLDR + Phospho (ST)	39	30
207	F125A_HUMAN	Q96EY5	DMQGLSLDAASQPSK + Oxidation (M); Phospho (ST)	54	30
			RNDSIYEASSLYGISAMDGVPFTLHPR + Phospho (ST)	53	33
208	ETV6_HUMAN	P41212	ISYTPPESPVPSYASSTPLHVPVPR + 2 Phospho (ST)	92	33
209	ESF1_HUMAN	Q9H501	ALAEEASEEELPSDVDLNDPYFAEEVK + 2 Phospho (ST)	52	32
210	ERF_HUMAN	P50548	RVSSDLQHATAQLSLEHR + Phospho (ST)	40	31
211	ELL_HUMAN	P55199	LGLPLLTDCAQPSRPHGSPSR + Phospho (ST)	33	31
212	EHMT2_HUMAN	Q96KQ7	SPPSVQSLAMR + Phospho (ST)	41	29
213	EHBP1_HUMAN	Q8NDI1	DLSTSPKPSPIPSPVLGR + 2 Phospho (ST)	32	31
214	EGFR_HUMAN	P00533	ELVEPLTPSGEAPNQALLR + Phospho (ST)	35	30
215	EFNB2_HUMAN	P52799	KHSPQHTTTLSLSTLATPK + Phospho (ST)	49	30
216	EF2_HUMAN	P13639	FTDTR + Phospho (ST)	43	23
217	EF1G_HUMAN	P26641	TPEFLR + Phospho (ST)	39	28
			VLSAPPHFHFGQTNRTPEFLR + Phospho (ST)	36	32
218	EAF1_HUMAN	Q96JC9	TSPLKDNPSPEPQLDDIKR + 2 Phospho (ST)	34	32
			TSPLKDNPSPEPQLDDIKR + Phospho (ST)	40	32
219	E41L1_HUMAN	Q9H4G0	SLDGAEFSRPASVSENHDAGPDGDKR + Phospho (ST)	45	32
220	E2AK4_HUMAN	Q9P2K8	HERPAGPGTPPPDSGPLAK + Phospho (ST)	46	32
221	DYN2_HUMAN	P50570	EALNIIGDISTSTVSTPVPPPVDDTWLQSASSHSPTPQR + Phospho (ST)	48	32

222	DYHC1_HUMAN	Q14204	TDSTSDGRPAWMR + Phospho (ST)	42	29
223	DPOD3_HUMAN	Q15054	VALSDDETKETENMR + Phospho (ST)	58	30
224	DP13A_HUMAN	Q9UKG1	VNQSALEAVTPSPSFQQR + Phospho (ST)	42	31
225	DOCK6_HUMAN	Q96HP0	SISSSNPDLAVAPGSVDDEVSR + Phospho (ST)	33	32
226	DNMT1_HUMAN	P26358	EADDDEEVDDNIPEMPSPK + Phospho (ST)	59	29
227	DNLI1_HUMAN	P18858	AETPTESVSEPEVATK + Phospho (ST)	64	31
			TIQEVLEEQSEDEDR + Phospho (ST)	32	31
			VLGSEGEEEDEALSPAK + 2 Phospho (ST)	41	29
			VLGSEGEEEDEALSPAK + Phospho (ST)	36	31
228	DNJC1_HUMAN	Q96KC8	DFDIAEQNESSDEESLRK + 2 Phospho (ST)	50	29
229	DMXL1_HUMAN	Q9Y485	SLALVAHTGYLPHQQDPHHVHR + Phospho (ST)	72	32
230	DJC21_HUMAN	Q5F1R6	DEEDGKDSDEAEDAELYDDLYCPACDK + Phospho (ST)	41	29
231	DIP2B_HUMAN	Q9P265	GTSGSLADVFANTR + Phospho (ST)	31	30
232	DDX55_HUMAN	Q8NHQ9	KREEGSDIEDEDMEELLNDTR + Phospho (ST)	57	32
233	DDX46_HUMAN	Q7L014	AALGLQDSDDEDAAVDIDEQIESMFNSK + Phospho (ST)	76	32
			KGELMENDQDAMEYSSEEEEVDLQTALTGYQTK + 2 Oxidation (M); 2 Phospho (ST)	35	32
234	DDX41_HUMAN	Q9UJV9	SEAEDEDDEDYVPYVPLR + Phospho (ST)	115	31
			TDEVPAGGSRSEAEDEDDEDYVPYVPLR + 2 Phospho (ST)	32	31
235	DDX21_HUMAN	Q9NR30	NEEPSEEEIDAPKPK + Phospho (ST)	55	31
236	DAP1_HUMAN	P51397	DKDDQEWESPSPPKPTVFISGVIAR + Phospho (ST)	68	32
			EEKDKDDQEWESPSPPKPTVFISGVIAR + Phospho (ST)	33	32
237	CXA1_HUMAN	P17302	KLAAGHELQPLAIVDQRPSSR + Phospho (ST)	34	30
238	CV009_HUMAN	Q6ICG6	KSHSANDSEEFFREDDGGADLHNATNLR + Phospho (ST)	37	32
			SHSANDSEEFFREDDGGADLHNATNLR + Phospho (ST)	62	32

239	CUL4A_HUMAN	Q13619	KGSFSALVGR + Phospho (ST)	54	29
240	CTGE2_HUMAN	Q96RT6	LSGPAELR + Phospho (ST)	37	28
241	CT151_HUMAN	Q8NC74	GQDTPKPAGQHGSLSPAAAHTASPEPPTQSGPLTR + 2 Phospho (ST)	59	33
242	CT117_HUMAN	O94964	VYYSPPVAR + Phospho (ST)	30	29
243	CSF1R_HUMAN	P07333	KVMSISIRLK + Phospho (ST)	30	26
244	CRIP1_HUMAN	P50238	TLTSGGHAEHEGKPYCNHPCYAAMFGPK + Oxidation (M); Phospho (ST)	47	32
245	CQ062_HUMAN	Q9BQA9	LITSFLELHCLESPTELSQSSDSEAGDPASQS + Phospho (ST)	51	33
246	CQ049_HUMAN	Q8IXM2	KVASGVLSPPPAAPPPSSSSVPEAGGPPIKK + Phospho (ST)	57	29
			KVYEDSGIPLPAESPK + Phospho (ST)	82	31
			VYEDSGIPLPAESPK + Phospho (ST)	42	31
247	CPIN1_HUMAN	Q6FI81	KSSPSVKPAVDPAAAK + Phospho (ST)	38	30
248	CP088_HUMAN	Q1ED39	YSVLNNDDYFADVSPLR + Phospho (ST)	34	31
249	COASY_HUMAN	Q13057	LASVLLYSDYGIGEVPVEPLDVPLPSTIRPASPVAGSPK + Phospho (ST)	36	27
250	CO052_HUMAN	Q6ZUT6	SPPTQVAISSDSAR + Phospho (ST)	35	30
251	CND3_HUMAN	Q9BPX3	TLHCEGTEINSDDEQESKEVEETATAK + Phospho (ST)	39	32
252	CLK2_HUMAN	P49760	DRGDAYYDTDYRHSYEYQR + Phospho (ST)	52	30
253	CLD3_HUMAN	O15551	STGPGASLGTGYDR + 2 Phospho (ST)	81	27
			STGPGASLGTGYDR + Phospho (ST)	82	30
254	CLCN7_HUMAN	P51798	VGHMSSVELDDELLDPDMDPPHPFPK + Phospho (ST)	65	33
255	CL043_HUMAN	Q96C57	EAAVSASDILQESAIHSPGTVEK + Phospho (ST)	38	32
256	CI129_HUMAN	Q5T035	NLTEQNSYSNIPHEGKHTPLYER + Phospho (ST)	44	33
257	CHD8_HUMAN	Q9HCK8	HFSTLKDDDLVEFSDLESEDDERPR + Phospho (ST)	67	33
			TASPLPLRPDAPVEK + Phospho (ST)	40	30
			TASPLPLRPDAPVEKSPEETATQVPSLESLTLK + 2 Phospho (ST)	67	31
258	CFDP1_HUMAN	Q9UEE9	LDWESFKEEEGIGEELAIHNR + Phospho (ST)	38	32

259	CENPJ_HUMAN	Q9HC77	IERESALEK + Phospho (ST)	32	30
260	CDYL1_HUMAN	Q9Y232	TAVDGFQSESPEKLDPVEQGQEDTVAPEVAAEKPVGALLGPGAER + Phospho (ST)	50	31
261	CCD94_HUMAN	Q9BW85	LLEDSDSEDEAAPSPLQPALRPNPTAILDEAPKPK + 2 Phospho (ST)	72	32
262	CCD55_HUMAN	Q9H0G5	VEENPDADSDFDAKSSADDEIEETR + 3 Phospho (ST)	37	28
263	CCD43_HUMAN	Q96MW1	AALLAQYADVTDEEDEADEKDDSGATTMNIGSDK + Phospho (ST)	66	33
264	CC85C_HUMAN	A6NKD9	DVGDGSSTSSAGSGGSPDHHHHVPPPLLPPGPHK + Phospho (ST)	34	33
265	CC132_HUMAN	Q96JG6	SAYQEYDSDSDVPEELKR + 2 Phospho (ST)	71	29
			SAYQEYDSDSDVPEELKR + Phospho (ST)	46	31
			SAYQEYDSDSDVPEELKR + Phospho (ST); Phospho (Y)	40	30
266	CBX1_HUMAN	P83916	KADSDSEDKGEESKPK + Phospho (ST)	38	31
267	CASC3_HUMAN	O15234	DPSPEADAPVLGSPEK + Phospho (ST)	38	31
			DPSPEADAPVLGSPEKEEAASEPPAAAPDAAPPPPDRPIEK + 2 Phospho (ST)	38	33
268	CAP1_HUMAN	Q01518	SGPKPFSAPKPQTSPSPK + Phospho (ST)	45	30
269	CA226_HUMAN	A1L170	ASSPSLIER + Phospho (ST)	39	29
			RASSPSLIER + 2 Phospho (ST)	33	29
			SFPHLSKPVAPGSAPLGSGEPGGPGLWVGSSQHLK + Phospho (ST)	39	31
270	CA198_HUMAN	Q9H425	LPSPDVR + Phospho (ST)	32	28
			SSSLDALGPTR + Phospho (ST)	43	29
271	BUD13_HUMAN	Q9BRD0	ARHDSPDLAPNVTYSLPR + Phospho (ST)	40	31
			HDSPDLAPNVTYSLPR + Phospho (ST)	35	31
272	BNIPL_HUMAN	Q7Z465	RLSAPELR + Phospho (ST)	31	29
273	BMS1_HUMAN	Q14692	LGPQNFIDETSDIENLLK + Phospho (ST)	73	32
274	BIG3_HUMAN	Q5TH69	AGGGDLLLPPSPK + Phospho (ST)	47	29
			ATGSAGLLGDPECEGSPPEHSPEQGR + Phospho (ST)	51	32

			GQDSPLLQRPQHLMDQGQMR + Phospho (ST)	64	32
			HSFSAGPELLR + Phospho (ST)	34	30
			SDVSDIGSDNCSLADEEQTPR + 2 Phospho (ST)	48	29
			SDVSDIGSDNCSLADEEQTPR + Phospho (ST)	56	30
			YSESNFSVDDQDLSR + Phospho (ST)	59	29
275	BI2L1_HUMAN	Q9UHR4	TPASTPVSGTPQASPMIER + 2 Phospho (ST)	33	31
			TPASTPVSGTPQASPMIER + Oxidation (M); Phospho (ST)	33	32
			TPASTPVSGTPQASPMIER + Phospho (ST)	40	31
276	BCL9_HUMAN	O00512	SSTPSHGQTTATEPTPAQK + Phospho (ST)	46	31
277	BAIP2_HUMAN	Q9UQB8	SSSTGNLLDKDDLAIPPPDYGAASR + Phospho (ST)	33	32
278	ATX2_HUMAN	Q99700	TSPSGGTWSSVVSGVPR + Phospho (ST)	36	31
279	ATLA2_HUMAN	Q8NHH9	TSDPSAAVNHVSSTTSLGENYEDDDLVNSDEVMK + Phospho (ST)	62	33
280	AT2A2_HUMAN	P16615	EFDELNPSAQR + Phospho (ST)	33	29
281	AT131_HUMAN	Q9HD20	DSPTLSNSGIR + Phospho (ST)	48	29
282	ARVC_HUMAN	O00192	SLAADDEGGPELEPDYGTATR + Phospho (ST)	47	31
283	ARIP4_HUMAN	Q9Y4B4	APDPEGLARPVSPDSPEIISELQQYADVAAAR + 2 Phospho (ST)	55	33
284	ARI4B_HUMAN	Q4LE39	DIEVLSEDTDYEEDEVTK + 2 Phospho (ST)	39	29
285	ARAP1_HUMAN	Q96P48	LFPEFDDSDYDEVPEEGPGAPAR + Phospho (ST)	53	31
			LFPEFDDSDYDEVPEEGPGAPAR + Phospho (Y)	32	31
286	ARAF_HUMAN	P10398	STSTPNVHMVSTTAPMDSNLIQLTGQSFSTDAAGSR + 2 Oxidation (M); Phospho (ST)	57	33
287	APT_HUMAN	P07741	IDYIAGLDSR + Phospho (ST)	38	30
288	ANR57_HUMAN	Q53LP3	DLVMGSSPQLK + Oxidation (M); Phospho (ST)	38	29
289	ANKL2_HUMAN	Q86XL3	NNSPPTVGAFGHTR + Phospho (ST)	57	30
290	ALS2_HUMAN	Q96Q42	RLSLPGLLSQVSPR + 2 Phospho (ST)	39	30

			SESPEPGYVVTSSGLLLPVLLPR + Phospho (ST)	113	29
291	ALBU_HUMAN	P02768	KVPQVSTPTLVEVSR + Phospho (ST)	44	29
			TCVADESAENCDK + Phospho (ST)	76	26
292	AHTF1_HUMAN	Q8WYP5	NLSFNELYPSGTLK + Phospho (ST)	34	31
293	ADA17_HUMAN	P78536	SFEDLTDHPVTR + Phospho (ST)	53	30
294	ACOD_HUMAN	O00767	EKGSTLDLSDLEAEK + Phospho (ST)	61	31
295	ABL2_HUMAN	P42684	GAQASSGSPALPR + Phospho (ST)	65	29
296	4EBP1_HUMAN	Q13541	DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR + Phospho (ST)	38	33
			RVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y)	37	33
			TPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST)	106	33
297	1433E_HUMAN	P62258	AAFDDAIAELDTLSEESYK + Phospho (ST)	74	31

Table 7.5 List of phosphoproteins previously reported to be associated with tumor genesis.

#	Swiss-Prot ID	Protein Name	Identified phosphopeptides	Function	Signaling Pathway	Reference
1	P67809	YB-1	AADPPAENSSAPEAEQGGAE + Phospho (ST) NEGSESAPEGQAQQR + Phospho (ST)	drug sensitivity	PI3K/Akt	33
2	Q9H4A3	p65	DVDDGSGSPHSPHQLSSK + 3 Phospho (ST) DVDDGSGSPHSPHQLSSK + Phospho (ST)	regulation of cellular process	JNK/MAPK	34, 35

			KEKPELSEPSHLNGPSSDPEAAFLSR + Phospho (ST)			
3	P40763	STAT3	FICVTPTTCSNTIDLPMSPR + Phospho (ST)	drug sensitivity	ErbB/HER	18, 36
4	P42224	STAT1	LQTTDNLLPMSPEEFDEVSR + Phospho (ST)	cellular response to insulin stimulus	ErbB/HER	37, 38
5	P12931	SRC	RRSLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST) SLEPAENVHGAGGGAFPASQTPSKPASADGHR + Phospho (ST)	membrane organization and biogenesis	ErbB/HER	39, 40
6	Q15311	RBP1	TEGYAAFQEDSSGDEAESPSK + 2 Phospho (ST) TGEPSPPHDILHEPPDVVSDDEKDHGK + 2 Phospho (ST) TGEPSPPHDILHEPPDVVSDDEKDHGK + Phospho (ST) TPSSEEISPTK + Phospho (ST)	drug sensitivity	cyclin/CDK	41
7	P06400	Rb	SKSEEAHAEDSVMDHHFR + Oxidation (M); Phospho (ST)	cell cycle progression (regulation)	Cell Cycle: G1/S Checkpoint	42
8	Q15418	S6 kinase	AYSFCGTVEYMAPEVVNR + Phospho (ST)	signal transduction	mTOR	43, 44
9	P17612	PKA	TWTLCGTPEYLAPEIILSK + Phospho (ST)	cell cycle regulation	AMPK	45
10	Q00613	HSF1	VEEASPGRPSSVDTLLSPTALIDSILR + Phospho (ST) VKEEPPSPPQSPR + 2 Phospho (ST)	negative regulation of cell proliferation	SAPK/JNK Signaling Cascades	46
11	Q13480	Gab-1	SSSLEGFHNHFK + Phospho (ST)	cell proliferation	PI3K/Akt	47
12	P04626	HER2	FVVIQNEDLGPASPLDSTFYR + Phospho (ST)	negative	Wnt	48

			GAPPSTFKGTPTAENPEYLGLDVPV + Phospho (ST) GTPTAENPEYLGLDVPV + Phospho (ST) SGGGDLTLGLEPSEEEAPR + 2 Phospho (ST) SGGGDLTLGLEPSEEEAPR + Phospho (ST) SGGGDLTLGLEPSEEEAPRSPLAPSEGAGSDVFDGDLGMG AAK + Oxidation (M); Phospho (ST) SPLAPSEGAGSDVFDGDLGMGAAK + 2 Phospho (ST) SPLAPSEGAGSDVFDGDLGMGAAK + Oxidation (M); Phospho (ST) SPLAPSEGAGSDVFDGDLGMGAAK + Phospho (ST) VLGSGAFGTVYK + Phospho (ST) YSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDVRPQPPS PR + 2 Phospho (ST) YSEDPTVPLPSETDGYVAPLTCSPQPEYVNQPDVRPQPPS PR + Phospho (ST)	regulation of apoptosis	signaling pathway and PI3K/Akt	
13	P16070	CD44	SQEMVHLVNK + Phospho (ST)	tumor growth and progression	PI3K/AKT	49
14	Q92934	BAD	RMSDEFVDSFK + Phospho (ST)	Apoptosis	ErbB/HER and PI3K/AKT	50, 51
15	P15336	CREB	MPLDLSPLATPIIR + Phospho (ST)	:ranscription factor	B Cell Receptor	51, 52
16	P31751	protein kinase B (AKT2)	YFDDEFTAQSITITPPDR + Phospho (ST)	Apoptosis	PI3K/AKT	53

17	Q9Y478	AMPK	SHNNFVAILDLPEGEHQYK + Phospho (ST)	regulating cellular and organismal energy balance	AMPK	54
18	P30530	AXL	ESFPAPVVILPFMK + Phospho (ST)	tumor metastasis	Gas6/axl	55
19	Q13541	eIF4E-binding protein 1	DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR + Phospho (ST) RVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y) TPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST)	cell cycle regulation	Insulin Receptor	44

Table 7.6 List of phosphoproteins previously reported to be associated with tumor invasiveness.

#	Swiss-Prot ID	Protein Name	Identified phosphopeptides	Function	Signalling Pathway	Reference
1	P40763	STAT3	FICVTPTTCSNTIDLPMSPR + Phospho (ST)	drug sensitivity	ErbB/HER	18, 33
2	P06400	Rb	SKSEEAHAEDSVMDHHFR + Oxidation (M); Phospho (ST)	cell cycle progression (regulation)	Cell Cycle: G1/S Checkpoint	40
3	Q15418	S6 kinase	AYSFCGTVEYMAPEVVNR + Phospho (ST)	signal transduction	mTOR	41, 42
4	Q13480	Gab-1	SSSLEGFHNHFK + Phospho (ST)	cell proliferation	PI3K/Akt	45
5	P00533	EGFR	ELVEPLTPSGEAPNQALLR + Phospho (ST)	apoptosis	ErbB/HER	56, 57
6	P30530	AXL	ESFPAPVVILPFMK + Phospho (ST)	tumor metastasis	Gas6/axl	55
7	Q13541	eIF4E-binding	DLPTIPGVTSPSSDEPPMEASQSHLRNSPEDKR	cell cycle regulation	Insulin	42

protein 1	+ Phospho (ST) RVVLGDGVQLPPGDYSTTPGGTLFSTTPGGTR + Phospho (ST); Phospho (Y) TPPRDLPTIPGVTSPSSDEPPMEASQSHLR + Phospho (ST)	Receptor

7.5 References

1. Jemal, A.; Siegel, R.; Xu, J. Q.; Ward, E., Cancer Statistics, 2010. *Ca-a Cancer Journal for Clinicians* 60, (5), 277-300.

2. Perou, C. M.; Sorlie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Rees, C. A.; Pollack, J. R.; Ross, D. T.; Johnsen, H.; Akslen, L. A.; Fluge, O.; Pergamenschikov, A.; Williams, C.; Zhu, S. X.; Lonning, P. E.; Borresen-Dale, A. L.; Brown, P. O.; Botstein, D., Molecular portraits of human breast tumours. *Nature* **2000,** 406, (6797), 747-752.

3. Sorlie, T.; Perou, C. M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Thorsen, T.; Quist, H.; Matese, J. C.; Brown, P. O.; Botstein, D.; Lonning, P. E.; Borresen-Dale, A. L., Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** 98, (19), 10869-10874.

4. Reis-Filho, J. S.; Tutt, A. N. J., Triple negative tumours: a critical review. *Histopathology* **2008,** 52, (1), 108-118.

5. Peppercorn, J.; Perou, C. M.; Carey, L. A., Molecular subtypes in breast cancer evaluation and management: Divide and conquer. *Cancer Investigation* **2008,** 26, (1), 1-10.

6. Reis, J. S.; Lakhani, S. R., Breast cancer special types: why bother? *Journal of Pathology* **2008,** 216, (4), 394-398.

7. Weigelt, B.; Baehner, F. L.; Reis, J. S., The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *Journal of Pathology* 220, (2), 263-280.

8. Umar, A.; Luider, T. M.; Foekens, J. A.; Pasa-Tolic, L., NanoLC-FT-ICR MS improves proteome coverage attainable for similar to 3000 laser-microdissected breast carcinoma cells. *Proteomics* **2007,** 7, (2), 323-329.

9. Celis, J. E.; Gromov, P.; Gromova, I.; Moreira, J. M. A.; Cabezon, T.; Ambartsumian, N.; Grigorian, M.; Lukanidin, E.; Straten, P. T.; Guldberg, P.; Bartkova, J.; Bartek, J.; Lukas, J.; Lukas, C.; Lykkesfeldt, A.; Jaattela, M.; Roepstorff, P.; Bolund, L.; Orntoft, T.; Brunner, N.; Overgaard, J.; Sandelin, K.; Blichert-Toft, M.; Mouridsen, H.; Rank, F. E., Integrating proteomic and functional genomic technologies in discovery-driven translational breast cancer research. *Molecular & Cellular Proteomics* **2003,** 2, (6), 369-377.

10. Deng, S.-S.; Xing, T.-Y.; Zhou, H.-Y.; Xiong, R.-H.; Lu, Y.-G.; Wen, B.; Liu, S.-Q.; Yang, H.-J., Comparative proteome analysis of breast cancer and adjacent normal breast tissues in human. *Genomics Proteomics & Bioinformatics* **2006,** 4, (3), 165-172.

11. Mbeunkui, F.; Metge, B. J.; Shevde, L. A.; Pannell, L. K., Identification of differentially secreted biomarkers using LC-MS/MS in isogenic cell lines representing a

progression of breast cancer. *Journal of Proteome Research* **2007,** 6, (8), 2993-3002.
12. Ru, Q. H. C.; Zhu, L. W. A.; Silberman, J.; Shriver, C. D., Label-free semiquantitative peptide feature profiling of human breast cancer and breast disease sera via two-dimensional liquid chromatography-mass spectrometry. *Molecular & Cellular Proteomics* **2006,** 5, (6), 1095-1104.
13. An, Y. M.; Fu, Z. M.; Gutierrez, P.; Fenselau, C., Solution isoelectric focusing for peptide analysis: Comparative investigation of an insoluble nuclear protein fraction. *Journal of Proteome Research* **2005,** 4, (6), 2126-2132.
14. Strong, R.; Nakanishi, T.; Ross, D.; Fenselau, C., Alterations in the mitochondrial Proteome of adriamycin resistant MCF-7 breast cancer cells. *Journal of Proteome Research* **2006,** 5, (9), 2389-2395.
15. Ballif, B. A.; Carey, G. R.; Sunyaev, S. R.; Gygi, S. P., Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. *Journal of Proteome Research* **2008,** 7, (1), 311-318.
16. Jin, L. L.; Tong, J. F.; Prakash, A.; Peterman, S. M.; St-Germain, J. R.; Taylor, P.; Trudel, S.; Moran, M. F., Measurement of Protein Phosphorylation Stoichiometry by Selected Reaction Monitoring Mass Spectrometry. *Journal of Proteome Research* 9, (5), 2752-2761.
17. Lin, H. J.; Hsieh, F. C.; Song, H.; Lin, J., Elevated phosphorylation and activation of PDK-I/AKT pathway in human breast cancer. *British Journal of Cancer* **2005,** 93, (12), 1372-1381.
18. Yeh, Y. T.; Fu, O. Y.; Chen, I. F.; Yang, S. F.; Wang, Y. Y.; Chuang, H. Y.; Su, J. H.; Hou, M. F.; Yuan, S. S. F., STAT3 ser727 phosphorylation and its association with negative estrogen receptor status in breast infiltrating ductal carcinoma. *International Journal of Cancer* **2006,** 118, (12), 2943-2947.
19. You, B. C.; Park, S. Y.; Lee, Y. D.; Lee, J. N.; Hwang, Y. J.; Park, H. K., The Expression of Heat Shock Protein 60 kDa in Tissues and Cell Lines of Breast Cancer. *Journal of Breast Cancer* **2008,** 11, (4), 161-171.
20. Su, S.; Li, Y.; Luo, Y.; Sheng, Y.; Su, Y.; Padia, R. N.; Pan, Z. K.; Dong, Z.; Huang, S., Proteinase-activated receptor 2 expression in breast cancer and its role in breast cancer cell migration. *Oncogene* **2009,** 28, (34), 3047-3057.
21. Li, Q. R.; Xing, X. B.; Chen, T. T.; Li, R. X.; Dai, J.; Sheng, Q. H.; Xin, S. M.; Zhu, L. L.; Jin, Y.; Pei, G.; Kang, J. H.; Li, Y. X.; Zeng, R., Large Scale Phosphoproteome Profiles Comprehensive Features of Mouse Embryonic Stem Cells. *Molecular & Cellular Proteomics* 10, (4), 14.
22. MacCoss, M. J.; McDonald, W. H.; Saraf, A.; Sadygov, R.; Clark, J. M.; Tasto, J. J.; Gould, K. L.; Wolters, D.; Washburn, M.; Weiss, A.; Clark, J. I.; Yates, J. R., Shotgun identification of protein modifications from protein complexes and lens tissue. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (12), 7900-7905.
23. Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P.,

Phosphoproteomic analysis of the developing mouse brain. *Molecular & Cellular Proteomics* **2004,** 3, (11), 1093-1101.
24. Han, G. H.; Ye, M. L.; Zhou, H. J.; Jiang, X. N.; Feng, S.; Jiang, X. G.; Tian, R. J.; Wan, D. F.; Zou, H. F.; Gu, J. R., Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. *Proteomics* **2008,** 8, (7), 1346-1361.
25. Monetti, M.; Nagaraj, N.; Sharma, K.; Mann, M., Large-scale phosphosite quantification in tissues by a spike-in SILAC method. *Nature Methods* 8, (8), 655-U74.
26. Carrascal, M.; Ovefletro, D.; Casas, V.; Gay, M.; Ablan, J., Phosphorylation Analysis of Primary Human T Lymphocytes Using Sequential IMAC and Titanium Oxide Enrichment. *Journal of Proteome Research* **2008,** 7, (12), 5167-5176.
27. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-Line Two-Dimensional liquid Chromatography with Maximized Sample Loading to Reversed-Phase Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry for Shotgun Proteome Analysis. *Analytical Chemistry* **2009,** 81, (3), 1049-1060.
28. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical Chemistry* **2008,** 80, (12), 4696-4710.
29. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* **2007,** 4, (3), 207-214.
30. Kyte, J.; Doolittle, R. F., A Simple Method for Displaing the Hydropathic Character of a Protein. *Journal of Molecular Biology* **1982,** 157, (1), 105-132.
31. Liu, M. Y.; Cai, S. L.; Espejo, A.; Bedford, M. T.; Walker, C. L., 14-3-3 interacts with the tumor suppressor tuberin at Akt phosphorylation site(s). *Cancer Research* **2002,** 62, (22), 6475-6480.
32. Moreira, J. M. A.; Shen, T.; Ohlsson, G.; Gromov, P.; Gromova, I.; Celis, J. E., A combined proteome and ultrastructural localization analysis of 14-3-3 proteins in transformed human amnion (AMA) cells. *Molecular & Cellular Proteomics* **2008,** 7, (7), 1225-1240.
33. Randi, A. S.; Pontillo, C. A.; Garcia, M. A.; Pena, D.; Cocca, C.; Chiappini, F.; Alvarez, L.; de Pisarev, D. K., Activation of c-Src/HER1/STAT5b and HER1/ERK1/2 Signaling Pathways and Cell Migration by Hexachlorobenzene in MDA-MB-231 Human Breast Cancer Cell Line. *Toxicological Sciences* **2011,** 120, (2), 284-296.
34. Ghosh-Choudhury, T.; Mandal, C. C.; Woodruff, K.; St Clair, P.; Fernandes, G.; Choudhury, G. G.; Ghosh-Choudhury, N., Fish oil targets PTEN to regulate NF kappa B for downregulation of anti-apoptotic genes in breast tumor growth. *Breast Cancer Research and Treatment* **2009,** 118, (1), 213-228.
35. Wei, Y. Y.; Chen, Y. J.; Hsiao, Y. C.; Huang, Y. C.; Lai, T. H.; Tang, C. H., Osteoblasts-derived TGF-beta 1 enhance motility and integrin upregulation through Akt,

ERK, and NF-kappa B-dependent pathway in human breast cancer cells. *Molecular Carcinogenesis* **2008,** 47, (7), 526-537.
36. Sharma, D.; Kim, S. H.; Nagalingam, A.; Saxena, N. K.; Singh, S. V., Benzyl isothiocyanate inhibits oncogenic actions of leptin in human breast cancer cells by suppressing activation of signal transducer and activator of transcription 3. *Carcinogenesis* **2011,** 32, (3), 359-367.
37. Mariotto, S.; Ciampa, A. R.; de Prati, A. C.; Darra, E.; Vincenzi, S.; Sega, M.; Cavalieri, E.; Shoji, K.; Suzuki, H., Aqueous extract of Arbutus unedo inhibits STAT1 activation in human breast cancer cell line MDA-MB-231 and human fibroblasts through SHP2 activation. *Medicinal Chemistry* **2008,** 4, (3), 219-228.
38. Shang, Y. F.; Baumrucker, C. R.; Green, M. H., The induction and activation of STAT1 by all-trans-retinoic acid are mediated by RAR beta signaling pathways in breast cancer cells. *Oncogene* **1999,** 18, (48), 6725-6732.
39. Yoshida, T.; Nagaharu, K.; Zhang, X. H.; Katoh, D.; Hanamura, N.; Kozuka, Y.; Ogawa, T.; Shiraishi, T.; Imanaka-Yoshida, K., Tenascin C Induces Epithelial-Mesenchymal Transition-Like Change Accompanied by SRC Activation and Focal Adhesion Kinase Phosphorylation in Human Breast Cancer Cells. *American Journal of Pathology* **2011,** 178, (2), 754-763.
40. Wang, P. S.; Chou, F. S.; Porchia, L.; Saji, M.; Pinzone, J. J., Troglitazone Inhibits Cell Migration, Adhesion, and Spreading by Modulating Cytoskeletal Rearrangement in Human Breast Cancer Cells. *Molecular Carcinogenesis* **2008,** 47, (12), 905-915.
41. Sarcevic, B.; Suryadinata, R.; Sadowski, M.; Steel, R., Cyclin-dependent Kinase-mediated Phosphorylation of RBP1 and pRb Promotes Their Dissociation to Mediate Release of the SAP30 center dot mSin3 center dot HDAC Transcriptional Repressor Complex. *Journal of Biological Chemistry* **2011,** 286, (7), 5108-5118.
42. Elangovan, S.; Hsieh, T. C.; Wu, J. M., Growth Inhibition of Human MDA-MB-231 Breast Cancer Cells by delta-Tocotrienol Is Associated with Loss 9 of Cyclin D1/CDK4 Expression and Accompanying Changes in the State of Phosphorylation of the Retinoblastoma Tumor Suppressor Gene Product. *Anticancer Research* **2008,** 28, (5A), 2641-2647.
43. Lu, J. M.; Zhang, K. Q.; Chen, S. A.; Wen, W., Grape seed extract inhibits VEGF expression via reducing HIF-1 alpha protein expression. *Carcinogenesis* **2009,** 30, (4), 636-644.
44. Dowling, R. J. O.; Zakikhani, M.; Fantus, I. G.; Pollak, M.; Sonenberg, N., Metformin inhibits mammalian target of rapamycin-dependent translation initiation in breast cancer cells. *Cancer Research* **2007,** 67, (22), 10804-10812.
45. Han, E. H.; Kim, H. G.; Hwang, Y. P.; Song, G. Y.; Jeong, H. G., Prostaglandin E-2 Induces CYP1B1 Expression via Ligand-Independent Activation of the ER alpha Pathway in Human Breast Cancer Cells. *Toxicological Sciences* **2010,** 114, (2), 204-216.
46. Kiang, J. G.; Gist, I. D.; Tsokos, G. C., Regulation of heat shock protein 72 kDa and 90 kDa in human breast cancer MDA-MB-231 cells. *Molecular and Cellular*

Biochemistry **2000,** 204, (1-2), 169-178.
47. Bourguignon, L. Y. W.; Singleton, P. A.; Zhu, H. B.; Diedrich, F., Hyaluronan-mediated CD44 interaction with RhoGEF and Rho kinase promotes Grb2-associated binder-1 phosphorylation and phosphatidylinositol 3-kinase signaling leading to cytokine (Macrophage-Colony stimulating factor) production and breast tumor progression. *Journal of Biological Chemistry* **2003,** 278, (32), 29420-29434.
48. Gril, B.; Palmieri, D.; Bronder, J. L.; Herring, J. M.; Vega-Valle, E.; Feigenbaum, L.; Liewehr, D. J.; Steinberg, S. M.; Merino, M. J.; Rubin, S. D.; Steeg, P. S., Effect of lapatinib on the outgrowth of metastatic breast cancer cells to the brain. *Journal of the National Cancer Institute* **2008,** 100, (15), 1092-1103.
49. Bourguignon, L. Y. W.; Singleton, P. A.; Zhu, H. B.; Zhou, B., Hyaluronan promotes signaling interaction between CD44 and the transforming growth factor beta receptor I in metastatic breast tumor cells. *Journal of Biological Chemistry* **2002,** 277, (42), 39703-39712.
50. Berndtsson, M.; Konishi, Y.; Bonni, A.; Hagg, M.; Shoshan, M.; Linder, S.; Havelka, A. M., Phosphorylation of BAD at Ser-128 during mitosis and paclitaxel-induced apoptosis. *Febs Letters* **2005,** 579, (14), 3090-3094.
51. Shankar, E.; Krishnamurthy, S.; Paranandi, R.; Basu, A., PKC epsilon induces Bcl-2 by activating CREB. *International Journal of Oncology* **2010,** 36, (4), 883-888.
52. Naviglio, S.; Di Gesto, D.; Illiano, F.; Chiosi, E.; Giordano, A.; Illiano, G.; Spina, A., Leptin Potentiates Antiproliferative Action of cAMP Elevation via Protein Kinase A Down-Regulation in Breast Cancer Cells. *Journal of Cellular Physiology* **2010,** 225, (3), 801-809.
53. Hardy, S.; St-Onge, G. G.; Joly, E.; Langelier, Y.; Prentki, M., Oleate promotes the proliferation of breast cancer cells via the G protein-coupled receptor GPR40. *Journal of Biological Chemistry* **2005,** 280, (14), 13285-13291.
54. Singh, G.; Linher-Melville, K.; Zantinge, S.; Sanli, T.; Gerstein, H.; Tsakiridis, T., Establishing a relationship between prolactin and altered fatty acid beta-Oxidation via carnitine palmitoyl transferase 1 in breast cancer cells. *Bmc Cancer* **2011,** 11.
55. Zhang, Y. X.; Knyazev, P. G.; Cheburkin, Y. V.; Sharma, K.; Knyazev, Y. P.; Orfi, L.; Szabadkai, I.; Daub, H.; Keri, G.; Ullrich, A., AXL is a potential target for therapeutic intervention in breast cancer progression. *Cancer Research* **2008,** 68, (6), 1905-1915.
56. Boerner, J. L.; Demory, M. L.; Silva, C.; Parsons, S. J., Phosphorylation of Y845 on the epidermal growth factor receptor mediates binding to the mitochondrial protein cytochrome c oxidase subunit II. *Molecular and Cellular Biology* **2004,** 24, (16), 7059-7071.
57. Rogers, K. R.; Kikawa, K. D.; Mouradian, M.; Hernandez, K.; McKinnon, K. M.; Ahwah, S. M.; Pardini, R. S., Docosahexaenoic acid alters epidermal growth factor receptor-related signaling by disrupting its lipid raft association. *Carcinogenesis* **2010,** 31, (9), 1523-1530

Chapter 8

Conclusions and Future Work

In the field of proteomics research, it is desirable to detect the entire proteome or all the proteins present in a sample. However, due to the complexity of most samples, this task is challenging. First of all, proteins in a biological sample usually present in a large dynamic range of concentrations. Identification of low abundance proteins is difficult as it is prone to be suppressed by the signal from proteins present in high abundance. Sample fractionation, either at the protein or peptide level or both, can increase the probability of identifying low abundant proteins. Second, the present of hydrophobic proteins in a proteome sample makes the solubilization process time-consuming and low efficient. It is necessary to develop improved protein solubilization techniques with high efficiency and compatible with the downstream mass spectrometry work. Therefore, the goal of my thesis work is to develop new protein solubilization and fractionation techniques to increase the identification efficiency in shotgun proteomic study, namely to detect as many proteins as possible at high sample handling throughout. Several methods have been described in this thesis along with the illustration of their performance in the proteome profiling of complex samples.

In Chapter 2, a sequential protein precipitation and solubilization combined with 2D-LC MS/MS was developed for the fractionation of the MCF-7 membrane proteome at the protein level. This work was an improvement based on the sequential protein solubilization method developed by Gong et al.[1] Sequential precipitation was achieved by reduction/alkylation and acetone precipitation of the protein sample, and was proven to be orthogonal, to some extent, to the sequential solubilization method. It further reduced the complexity of the MCF-7 membrane proteome, thereby increasing the possibility of identification of low abundant proteins. By using the sequential protein precipitation and solubilization method, a total number of 5011 unique proteins with 24261 different peptides were identified from the MCF-7 membrane extracts. Among these, 2536 proteins were predicted as membrane-associated or membrane-bounded proteins, and 1077 proteins were predicted as integral membrane proteins. In Chapter 3, the sequential protein precipitation and solubilization method described in Chapter 2 was applied to the comprehensive proteome mapping of the *E. coli* K12 cell line. In total, 3418 proteins were identified with 99% confidence level, representing 79% of the 4300 proteins predicted from the *E. coli* genome. To our knowledge, this is the most comprehensive profile of the *E. coli* K12 proteome. However, vortex-assisted protein solubilization used in this work is time-consuming and with low efficiency, especially in solubilizing hydrophobic proteins. The usage of SDS may affect tryptic digestion

efficiency as well as interfere with downstream MS analysis. Therefore it is ideal to solubilize as much sample as possible in MS- compatible solvents with high speed.

To address this issue, in Chapter 4, a new method combining microwave-assisted sequential protein solubilization (MAPS) and 2D-LC MS/MS analysis was developed. Compared to the conventional vortex-assisted solubilization method, MAPS speeded up the solubilization process and increased protein solubility. More importantly, by using MAPS, better protein digestion efficiency was achieved. As a result, compared with 1057 proteins and 6261 peptides identified by the vortex method, a total of 1291 distinct proteins and 10363 peptides were identified by the MAPS method in the integral membrane extract of *E. coli* K12 cells.

Despite the progress made in Chapter 4 by using the MAPS method, as a solubility-based protein fractionation method, sequential protein solubilization does not fractionate the proteome with high resolution and throughput. Moreover, the sample handling process is done manually, making this method less robust. Therefore, it is important to develop an alternative method, such as a LC-based technique, for protein fractionation with high resolution and throughput. Although a comprehensive profile of the *E. coli* proteome was generated and described in Chapter 3, it took more than 300 LC-MS/MS runs and consumed a large amount of protein sample to complete the work. Therefore, in Chapter 5, the performance of a new protocol that combines MAPS with 3D-LC MS/MS was evaluated with an objective of increasing the proteome identification efficiency by reducing the number of LC-MS/MS runs and the amount of sample used for analysis. An Agilent macro-porous C18 reversed-phase liquid chromatography column (mRP-C18) specially designed for the separation of proteins was used to achieve a faster and better resolution in protein fractionation. In this work, the *E. coli* K12 whole cell lysate was solubilized in urea by MAPS. Solubilized proteins were were first separated by the mRP-C18 column into six fractions. Each fraction was then subjected to tryptic digestion and 2D-LC [Strong-cation exchange (SCX)-capillary RPLC] MS/MS analysis. In total, 2136 unique proteins from 13712 unique peptides were identified by 37 LC-MS/MS runs. With the same number of LC-MS/MS runs, traditional 2D-LC separation identified 1901 unique proteins and 9221 unique peptides from the same amount of sample. Compared to the 2D-LC method, the 3D-LC method increased the proteome identification efficiency; at the peptide level, more peptides were detected using the same instrument time by the 3D-LC method. Compared with the results generated in Chapter 3, the 3D-LC method detected 64% of the proteins detected in Chapter 3, but only consumed less than 10% of the instrument time and less than 10% of the protein sample amount, thus improving the ratio of the protein identification number to the instrument time.

In Chapter 6, the mRP-C18 fractionation method was applied into the field of

phosphoproteome analysis. Briefly, MDA-MB-231 cell lysates were first separated by the mRP-C18 column at the protein level. Each individual fraction was digested by trypsin, followed by sequential phosphopeptide enrichment by IMAC and TiO_2 modified from a published protocol.[2] Finally the phosphopeptide enriched fractions were individually analyzed by RPLC-MS/MS. In total, 1537 phosphoproteins with 3942 phosphopeptides were identified by duplicate experiments. Compared with 1520 phosphoproteins and 3898 phosphopeptides identified by traditional SCX-RPLC method, the new mRP-RPLC method had a similar identification cability. However, an overlap of 59% in phosphoproteins and 33% in phosphopeptides were found between these two methods, and combining the two datasets, 5384 unique phosphopeptides and 1837 different phosphoproteins were identified. Considering that the overlaps in phosphoproteins and phosphopeptides between two replicate experiments were 77% and 59% for the SCX-RPLC method and 80% and 65% for the mRP-RPLC method, the mRP-RPLC method provided complementary phosphoproteome information to the traditional SCX-RPLC method, resulting in an enhancement of the overall coverage of the MDA-MB-231 cell phosphoproteome. In this dataset, 57 phosphoproteins were previously reported to be functionally related to the beast cancer, and, to our best knowledge, 181 proteins have never been reported to be phosphorylated. Biological validations will need to be performed in the future to further confirm our findings and illustrate the biological significances of these findings.

Finally, in Chapter 7, sequential phosphopeptide enrichment by IMAC and TiO_2 combined with SCX-RPLC MS/MS method was applied to the analysis of human breast cancer tissues. The mRP-RPLC method was not used this chapter due to limited sample amount. The phosphoproteins identified from the stage III B tissue was compared with the stage II A and stage I A tissue samples, generating 297 phosphoproteins possibly related to the metastasis of breast tumor. The phosphoproteins identified from the stage III B tissue was also compared with the results of four cases of normal tissues, generating 875 phosphoproteins possibly related to the genesis of breast tumor. The majority of these candidates have never been reported in the breast cancer. Whether they are functionally involved in the genesis and development of human breast cancer still needs to be further explored.

Looking into future, from the technical development point of view, there are still several aspects that need to be further improved.

In Chapter 3, there are still about 970 *E. coli* proteins that have not been identified in this work. We need to investigate the potential cause for the missing proteins. Antibody-based protein identification methods, such as Western blot, may be used to confirm if they are present in the cells under the culturing condition used to grow the cells. If these proteins are indeed present in the cell, it is an indication that our strategy has

some limitations in detecting proteins with certain properties. And it is necessary to further develop techniques to address this problem. Alternatively, these missing proteins may not be expressed at all in the cells grown under the culturing condition used.

In Chapters 5 and 6, there are several limitations for the mRP-C18 separation utilized for protein fractionation. First of all, the resolution of protein fractionation was not as good as we expected. This may be due to protein degradation and insufficient denaturation of the proteins, introducing migration of protein retention time. Second, to obtain high sample recovery, separation by the mRP-C18 column was operated under an elevated temperature (i.e., 80 ˚C). Although there was no evidence of proteins degradation during mRP-C18 fractionation, it was previously reported that high temperature facilitated the protein degradation process.[3] This can lead to missing or mis-identification of proteins and phosphoproteins. Therefore, we should make more efforts to improve the resolution of mRP-C18 separation, and further investigate the stability of proteins/phosphoproteins during mRP-C18 separation under an elevated temperature.

In Chapters 6 and 7, some important phosphoproteins were identified, including phosphoproteins previously reported to be functionally related to the breast cancer, as well as phosphoproteins never found in breast cancer and phosphoproteins that are potentially related to the genesis and metastasis of breast cancer tumor. There is some important follow up work needs to be performed. For example, the presence of these phosphoproteins needs to be further validated by antibody-based techniques, such as Western blot. For the new phosphoproteins found in Chapter 6 and the phosphoproteins potentially related to the genesis and metastasis of breast cancer tumor found in Chapter 7, we need to further explore whether they are functionally related to breast cancer development using biological tools.

Finally in Chapter 7, the list of phosphoproteins were generated by profiling the phosphoproteomes for 7 cases of human breast tissue samples. This list only tells the "on/off" information on protein phosphorylation. However, it is highly possible that the level of a certain phosphorylation is regulated during breast cancer development, which could only be revealed by quantitative analysis. In the future, a better quantification method needs to be developed to comprehensively monitor the phosphorylation level changes during breast cancer development.

In summary, although there are still many issues needed to be addressed, the techniques developed or optimized in my thesis work improved sample preparation and fractionation efficiency, and therefore enhanced the overall efficiency of proteome identification. In particular, the phosphoproteome results generated from this thesis work provide complementary information to those obtained from the traditional method. Taken together the phosphoproteome profiles generated by the combined methods appear to

provide some future directions where the study of the newly identified phosphoproteins may generate novel insight into breast cancer biology. The techniques developed in this thesis work should also hold a great potential for profiling a wider range of proteomes in a more comprehensive and efficient way.

8.1 References

1. Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R.; Li, L., Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *Journal of Proteome Research* **2008,** 7, (8), 3583-3590.
2. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **2005,** 4, (7), 873-886.
3. Li, J. J.; Dewey, W. C., RELATIONSHIP BETWEEN THERMAL TOLERANCE AND PROTEIN-DEGRADATION IN TEMPERATURE-SENSITIVE MOUSE CELLS. *Journal of Cellular Physiology* **1992,** 151, (2), 310-317.