Development and application of proteomics in ovarian adenocarcinomas using multi-dimension separation, microarray and mass spectrometry.

by

Evelyn H. Kim

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2009

Doctoral Committee:

Professor David M. Lubman, Chair
Professor Masato Koreeda
Professor K. M. Jairam Menon
Professor Michael D. Morris

Dedicated to my parents Janghee Kim and Guiyeon Lee

My husband Vince Kim

Acknowledgements

First of all, I would like to express my sincere appreciation to my advisor Dr. David M. Lubman, whose encouragement, invaluable advice and support have made this work possible. I am most grateful for his support and understanding of the many situations I have to face as a scientist and a person during the course of my Ph.D. study.

I sincerely thank my doctoral committee members, Dr. Michael D. Morris, and Dr. Jairam K. Menon, and Dr. Masato Koreeda for serving on my dissertation committee and providing helpful advice and suggestions.

I am also deeply grateful to Dr. David E. Misek whose insights into human biology have been crucial to the development of this work and for providing great encouragement, helpful suggestions with my research and guidance in manuscript preparation.

My sincere thanks also for the collaboration of Dr. Kathleen R. Cho and Dr. Rong Wu for providing invaluable samples. I sincerely thank Dr. Yangsun Kim to providing invaluable materials to improve my research.

I would like to thank all past and present group-members for all of their help. Special thanks to Dr. Nathan S. Buchanan, and Dr. Rick Hamler for their valuable

mentoring during the early part of my graduate school study and Dr.Yanfei Wang, Dr. Jia Zhao, Dr. Xiaoping Ao, Dr. Tasneem H. Patwa for wonderful conversations in both chemistry and other topics and their friendship. I would also like to thank Dr. Shun Feng, Chen Li, Huy Vuong for their support and collaborations.

I would like to express my love and appreciation for my parents, Janghee Kim, Guiyeon Lee for encouraging me in my academic pursuits and their prayers over these years and to my mother-in-law Youngmin Kim for her support and understanding over the years.

Lastly, I would like to thank my husband, Vince. Words cannot express my gratitude for the encouragement and unconditional support I have from him. Thank you for your love and your patience through all of this.

The completion of my graduate career would not have been possible without the love and support I have received from my parents, mother in-law and my husband.

Table of Contents

List of Figures

Figure

List of Tables

Table

List of Abbreviations

2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
α-CHCA	α-Cyano-4- hydroxycinnamic acid
ACN	Acetonitrile
ACTH	Adrenocorticotropin
CF	Chromatofocusing
CID	Collision Induced Dissociation
DE	Delayed Extraction
ESI	Electrospary ionization
FA	Formic Acid
HPLC	High performance liquid chromatography
IEX	Ion exchange
LCT	Liquid chromatography time of flight (Micromass mass spectrometer)
IPA	Ingenuity Pathway Analysis
MALDI	Matrix-assisted laser desorption ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular Weight
NP	Non-porous

OG	n-Octyl-β-D-glycopyranoside
PBS	Phosphate buffered saline
pI	isoelectric point
PTM	Post-translational modification
QIT	Quadropole ion trap
RP	Reversed phase
TFA	Trifluoroacetic acid
TOF	Time of flight

Chapter 1

Introduction

1.1.Proteomics in Cancer Research

Cancer is one of the leading causes of death in the United State and is believed to develop through a multi faceted process, involving genetic mutation and functional protein expression changes [1]. Most genetic studies in cancer research have revealed low correlations between mRNA and proteins [2-5] although it is too premature to completely disregard mRNA studies and prefer protein profiling. However, this result has prompted further studies of proteins in cancer research not only to understand carcinogenesis, tumor progression and metastasis but also to develop methods for early diagnosis. The study of the proteome has great challenges due to its complexity and dynamic range of biological samples to be analyzed. The proteomes are heterogeneous and the dynamic range of their expression levels in response to the change in cell/tissue and molecular environments exceeds an order of 10^{10} magnitude[6,7] Analytical technologies for proteomics has improved to overcome the problems stated above and posses the ability to detect low concentration of proteins or peptides with biological significance[8,9]. One approach at the molecular level involves comparing protein expression in normal and cancerous samples [10]. The motivation of this research is that the profiling of protein expression-changes in cancerous cells helps define them in nature, predicting their pathological behavior particularly in cancer [11], indicating the responsiveness to

treatment of a certain tumor and ultimately providing the opportunity for discovering new biomarkers for early detection of cancer. Early cancer detection is not simple since most cancers are asymptomatic until the late stages which are associated with high mortality rates. Thus, the discovery of biomarkers that are highly effective in detecting cancer has become an important issue for early detection as well as treatment.

Also the study of proteins involved in signaling pathway networks in cancer has also played an important role in proteomics in disease research [12]. The interactions of proteins in complicated biological networks determines the function of the organism and are indicative of biological complexity downstream from the alterations within the genes of a neoplastic cell[13,14]. Using noninvasive techniques in proteomics, it is important to find proteins that are highly expressed in cancer than normal, and are involved in signal pathways related to tumor progressions.

1.2. Liquid Separation.

Current proteomics studies of cancer can be divided into a few main areas: sample preparation, separation, and mass spectrometry. Since most biological samples are very complex mixtures, attempting to yield comprehensive outcome requires the separation of the samples before using mass spectrometry. The most popular method applied for fractionation of protein mixtures has been 2D-PAGE [15], enabling the separation of proteins in spot on gels where spot patterns can be compared between samples. However, the 2D-PAGE method has several limitations such as reproducibility, throughput and interfacing with mass spectrometry [16, 17]. Alternative multidimensional liquid separations have been suggested to provide an ideal means to fractionate and purify

proteins. The major advantage of the liquid-based technique over gel-based separations is that liquid-based separations are amenable to large-scale automation and are also easily coupled with mass spectrometry. Liquid-based methods in proteomics are performed with either bottom-up or top-down approaches. Bottom-up methods are often called shotgun approaches [18, 19], in which the entire samples are digested and then separated using strong cation exchange (SCX) followed by a Reversed Phase High Perforamance Liquid Chromatography (RP-HPLC) separation. This shotgun method is rapid in the identification of proteins, resulting in a high possibility of obtaining false positive as well as losing intact protein information such as protein truncation and protein isoforms. In contrast to the bottom up method, top-down approaches preserve the intact protein, enabling the study of the structural analysis of proteins.

Among liquid separation techniques for the top-down approach, the combination of two methods have been popular, namely chromatofocusing (CF) and non-porous reverse phase high performance liquid chromatography (NP-RP-HPLC) [20,21]. Chromatofocusing (CF) [22] is a technique that separates proteins using a pH gradient with a weak anion exchange column. In this process a biological sample is loaded into a column equilibrated with a starting buffer with higher pH (about pH 7.4). Fractionated proteins were eluted gradually based on their pI, using a low pH buffer (elution buffer, about pH 4.0) when a pH gradient was formed inside the column between two buffer systems. Each fractionated sample from CF is further fractionated using non porous-reverse phase-chromatography techniques based on hydrophobilicy of proteins. This NP-RP-HPLC method provides fast separation, sufficient peak capacity and high recovery. Furthermore, using on-line NP-RP-HPLC with ESI (Electrospary ionization)-TOF (Time

of flight)-MS, the molecular weight (MW) of proteins can be simultaneously obtained. This 2-D liquid separation combination (CF, NP-RP-HPLC) can provide information on intact proteins such as pI, hydrophobicity and expression levels to be compared between cancer and disease-free samples. Futhermore, obtained information can be visualized using software creating a 2-D mass map. A 2-D mass map analogous to that from 2D-PAGE can display differentially expressed proteins between the cancerous and disease-free samples. This approach can be used to suggest potential biomarkers along with information obtained from the mass spectrometer.

1.3. Mass spectrometry.

Mass spectrometry has played a great role in proteomics by identifying proteins and providing useful information such as post translational modification in biological samples. Although great improvements in separation technologies have been made, the usage of mass spectrometry in the proteomics field has been inevitable due to the detection limits when biological-sample availability is diminishingly low. The application of mass spectrometry in proteomics study has been based on two fundamental instruments, electrospray ionization mass spectrometry (ESI-MS) [23, 24] and Matrix-assisted laser-desorption ionization (MALDI) [25, 26]. Both techniques employ a common soft ionization method for biomolecules. Electrospray ionization (ESI) generates multiply-charged analyte ions in the aqueous state under an electric field, which allows the calculation of the molecular weight of the analytes, converting ion m/z to mass. The advantage of using the ESI technique is the capability of interfacing with on-line chromatography. On-line LC-ESI-MS can decrease the analysis time as well as increase

improved peak capacity by well-performed separation. Since sample availability is limited, nano-LC-ESI-MS has been performed more frequently in proteomics [27]. In nano-LC-ESI-MS, only a nano- to micro-liter of samples is needed for spraying using a high voltage at the end of the spray tip. The flow rate of nano-LC-ESI-MS is 20-300nl/min which is 1000 times smaller than that of standard LC-ESI-MS. Due to the small amount of usage of solvent in nano-LC-ESI-MS, droplet size by spraying at the end of capillary in ESI is decreased, generating a more stable spray compared to that from regular LC-ESI-MS.

Another soft ionization method is matrix-assisted laser-desorption ionization (MALDI). MALDI has been applied to samples which have been separated by liquid chromatography, dried and digested using an enzyme. The digested sample and matrix are spotted onto a plate for MALDI and crystallized. Common matrixes used in MALDI are alpha-cyano-4-hydroxycinnamic acid (α-CHCA) and 2, 5-dihydroxybenzoic acid (DHB). The mechanism of generating ions in MALDI has not yet been firmly established, though this technique has been available for the past three decades. Generally, however, matrices facilitate the absorption and transference of the pulsed laser energy to the sample (analyte) on the spot. Those analyte molecules are simply ionized by protonation, generating an $[M+H]^+$ ion (where M is the mass of the analyte molecules). The mass of the analyte molecules can be analyzed by different mass analyzers. The most popular mass analyzer for MALDI is time-of-flight (TOF). MALDI-TOF has a great capability to analyze a complex biological sample (mainly peptide mixtures), which is equipped with a reflectron and an ion lens, improving resolution. Also the TOF instrument has unlimited

mass range for analyzing biomolecules. Therefore, MALDI-TOF has been an ideal technique to apply in proteomics.

1.4. Microarray technology.

Microarrray technology has been initiated to study DNA in the fields of genomics in order to monitor changes in gene expression [28-30]. The principle of microarray is that biomolecules are placed into 10-12 blocks of formatted (12 x 12 per block) slides and each biomolecule can be reacted with antibody or fluorescence dyes so that those reactions can be scanned. Although each spot contains a low concentration of biomolecules, the ratio of surface-to-volume ratio is relatively high, allowing lower concentrations of the target biomolecule to be monitored [31]. As a result of monitoring, the reaction of a femto-molar level of biomolecules can be easily detected [32]. Given the success of microarray techniques with DNA analysis, this technique has been applied to proteins. Initially, there were concerns about applying microarray techniques to proteins due to the tendency of protein denaturation or non-specific binding of proteins with dye or antibody. However, advances in microarray technology have enabled protein microarray technology to perform immunoassay, autoantibodies detection [33, 34], and protein-protein interactions.

Protein microarray formats can be divided into two major classes. Forward-phase arrays (FPA) and reverse-phase arrays (RPA). In FPA, each spot contains one immobilized antibody or bait protein which is able to react with the proteins in the cell or tissue lysate solution. Proteins in the sample are then loaded directly into each spot on the array and incubated to allow for the reaction. In contrast, in RPA, proteins in samples are

immobilized first, and each spot on the array undergoes reaction with specific antibody or other proteins. RPA provides high linearity, sensitivity, and the possibility of combining multi-dimensional separation methods [35]. The utility of reverse-phase protein microarrays lies in their ability to provide a map of known cell signaling proteins which could be the starting point for drug development. RPA could also be used to perform autoimmune humoral response assays.

1.5. Software

The utility of software in proteomics has tremendously increased with the increased amount of data generated by improved mass-spectrometry technology. Mass spectrometry provides a mass spectrum for digested peptides from proteins using enzymes and mass spectrum providing sequence information. This could be introduced into software, which then can be used to search and compare experimental data and theoretical data in the data bases. Extensive software has been used for protein identification and/or protein post translational modification (phosphorylation, glycosylation) [36] studies.

The softwares that have been generally used are MOWSE, MASOT (Matrix Science Inc., Boson, MA), which are available on the web and Sequest (Thermo Fisher Scientific) for LC-ESI/MS/MS. Each program has scoring methods which count the number of measured peptide masses that match calculated theoretical peptide masses in the database. Sequest uses data from peptide fragmentation mass spectra. A cross correlation function in Sequest is calculated between the measured fragment mass spectrum and the protein sequences and scores the proteins in the database. Other

softwares could be in-use to study protein-protein interaction such as molecular concept map (University of Michigan, Ann Arbor, MI), GeneGo (GeneGo, Inc., St. Joseph, MI), and Ingenuity Pathway Analysis (Ingenuity Systems, Inc., Redwood City, CA). These programs are able to explore the network of relationships among molecules based on biologically related gene sets. They can provide opportunities to study disease classification, and diagnosis as well as study the signaling pathway in cancer progression.

This dissertation is intended to present a cancer study, mainly ovarian cancer, utilizing a high throughput proteomic approach as well as scaled-down proteomics to detect potential biomarkers in the disease as well as the pathological behavior of a cancer cell.

Chapter 2 discusses the utilization of 2-D liquid-based separation/mass mapping techniques to elucidate molecular weight and pI measurements of the differentially expressed intact proteins. 2-D protein mass maps were generated to facilitate the analysis of protein expression between both the low stage and high stage tumors. These mass maps (over a pI range of 5.6–4.6) revealed that the low stage ovarian endometrioid adenocarcinomas (OEAs) demonstrated protein over-expression at the lower pI ranges (pI 4.8–4.6) in comparison to the high stage tumors, which demonstrated protein over-expression in the higher pI ranges (pI 5.4–5.2). These data suggest that both low and high stage OEAs have characteristic pI signatures of abundant protein expression probably reflecting, at least in part, the different signaling pathway defects that characterize each group. In this study, the low stage OEAs were distinguishable from high stage tumors based upon their proteomic profiles. Interestingly, when only high-grade (grade 2 or 3)

OEAs were included in the analysis, the tumors still tended to cluster according to stage, suggesting that the altered protein expression was not solely dependent upon tumor cell differentiation. Further, these protein profiles clearly distinguish OEA from other types of ovarian cancer at the protein level.

Chapter 3 presents development of a micro separation technique, micro-CF, for analyzing two ovarian cancer cell lines MDAH 2774, and TOV 112D. Due to the limited amount of sample availability, micro separation techniques have been important in proteomics research. Especially there have been very limited improvements for separating intact proteins based on pI. Integrated microCF-nanoLC-MS/MS methods detected 700-800 proteins from ovarian cancer cell line lysates using 10μg of samples and compared their expression between two samples using label free semi-quantitation method.

Chapter 4 describes that Panc-1 cell line was used as bait and the method including the combination of protein microarrays with two dimensional separations was used to explore specific autoantibodies driven from pancreatic ductal adenocarcinoma (PDAC). Fractionated protein solutions were spotted on nitrocellulose slides and probed with 38 pancreatic cancer sera, 23 pancreatitis sera, and 25 normal sera. The data obtained from protein microarrays was analyzed by 4 different statistical analyses (COPA, OS, Wilcoxon, Pam). Each statistics generated the list of differential response proteins in comparison between cancer vs normal, cancer vs pancreatitis, pancreatitis vs normal. Among identified proteins using LC-MS/MS, annxein 2, Malate dehydrogenase,

cytoplasmic (MDH1) showed higher response in cancer sera against normal. This work suggests that using a protein microarray approach with several statistical methods to study the humoral response against pancreatic cancer may be an effective technique to identify potential biomarkers for disease.

Chapter 5 presents the study on mouse of model ovarian endometrioid adenocarcinomas using cIEF followed by tandem mass spectrometry. The obtained mice with cancer have genetic defects which are PTEN, Apc defects and p53 mutations. Mice were grouped into two such as mice with PTEN, Apc defect, which is considered as low grade cancer, and mice with PTEN, Apc defect with p53 mutation, which is considered as high grade cancer. Mice tumor samples were lysed, digested and separated using cIEF and analyzed using mainly LC-MS/MS. To understand genetic defects effect on a molecular level in tumor progression, identified proteins were analyzed further using Ingenuity Pathway Analysis software (IPA). Spectral count numbers for listed proteins in IPA were also uploaded to compare expression level changes. Any change in canonical pathway was carefully observed to study ovarian cancer progression based on genetic defects. The Right-tailed Fisher's exact was used for calculating the significance values for analyses of network and pathway generation.

Lastly, Chapter 6 described an on-plate digestion method where intact proteins were digested on MALDI plate and digested proteins on plate were identified. This work presents a method retaining the advantages of top-down proteomics.

1.6 References

[1] Hannash, S. *Nature* **2003**, *422*, 226-232.

[2] Anderson L.; Seilhamer J. *electrophoresis* **1997**, *18*, 533-537.

[3] Gygi S. P.; Rochon Y.; Franza B. R.; Aebersold R. *Mol. Cell. Biol.* **1999**, *19*, 1720-1730.

[4] Ideker T.; Thorsson, V.; Ranish, J. A.; Christmas, R.; Buhler, J.; Eng. J. K.; Bumgarner, R.; Goodlett, D. R.; Aebersold, R.; Hood, L. *Science* **2001**, *29*, 929-934.

[5] Ghaemmaghami S.; Huh W. K.; Bower K.; Houson R. W.; Belle A.; Dephoure N.; O'Shea E. K.; Weissman J. S. *Nature*, **2003**, 425, 737-741.

[6] Jacobs, J. M.; Adkins J. N.; Qian, W.-J.; Liu, T.; Shen, Y.; David, I.; Camp, D. G., II; Smith R. D. *J. Proteome Res*. **2005**, *4*, 1073-1085.

[7] Issaq, H. J.; Chen, K. C.; Janini, G. M.; Conrads, T. P.; Veenstra, T. D. *J. Chromatogr. B* **2005**, *817*, 35-47.

[8] Stevenes, E. V.; Posadas, E. M.; Davison, B.; Kohn, E. C. *Annals of Oncology* **2004**, *15*, 167-171.

[9] Pastwa, E.; Somiari, S. B.; Czyz, M.; Somiari R. I. *Proteomics Clin. App*. **2007**, *1*, 4-17.

[10] Jansen R. C.; Nap J. P.; Mlynarova, L. *Nat. Biotechnol*. **2002**, *20(1)*, 19.

[11] Wu, W.; Hu W.; Kavanagh J. J. *Int J. Gyenocol. Cancer* **2002**, *12*, 409-423.

[12] Gschwind A.; Fischer O. M.; Ullrich A. *Nat. Rev*. **2004**, *4*, 361-370.

[13] Nakal K. J. *Str. Biol*. **2001**, *134*, 103-116.

[14] Harrison P. M.; Kumar, A.; Lang N.; Snyder M.; Gerstein M. *Nucleic Acids Res.* **2002**, *30*, 1083-1090.

[15] O'Farrell P. H. *J. Biol. Chem*. **1975**, *250*, 4007-4021.

[16] Shevchenko A.; Wilm M.; Vorm O.; Mann M. *Anal. Chem*, **1996**, *68*, 850-858.

[17] Gygi S. P.; Corthals G. L.; Zhang Y.; Rochon Y.; Aebersold R. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9390-9395.

[18] Link A.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat. Biotechnol*. **1999**, *17*, 676-682.

[19] Wolters, D. A.; Washburn, M. P.; Yates, J. R. *Anal. Chem*. **2001**, *72*, 5683-5690.

[20] Hamler, R. L.; Zhu, K.; Buchanan N. S.; Kreunin P.; Kachman M. T.; Miller, F. R.; Lubman, D. M. *Proteomics* **2004**, *4*, 562-577.

[21] Wang H.; Kachman, M. T.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Proteomics* **2004**, *4*, 2476-2495.

[22] Liu Y. S.; Anderson D. J. *J chromatogr. A.* **1997**, *762*, 47-54.

[23] Fenn J. B.; Mann M.; Meng C. K.; Wong S. F.; Whitehouse C. M. *Science* **1989**, *246*, 64-71.

[24] Gaskell S. J. *J. Mass Spectrom.*, **1997**, *32*, 677-688.

[25] Tanaka K.; Waki H.; Ido Y.; Akita S.; Yoshida Y.; Yoshida T. *Rapid Commun. Mass Spectrom*. **1998**, *2*, 151-153.

[26] Karas M.; Hillenkamp F. *Anal. Chem*. **1998**, *60*, 2299-2301.

[27] Shen, Y.; Zhao, R.; Bergers, S. J.; Anderson, G. A.; Rodrigues, N.; Smith, R. D. *Anal Chem* **2002**, *74*, 4235-4249.

[28] DeRisi J. L.; Iyer V. R.; Brown P. O. *Science* **1997**, *278*, 680-686.

[29] The Chipping Forecast *Nat. Genet*. **1999**, *21*, Suppl: 1-60.

[30] The Chipping Forecase *Nat. Genet*. **2002**, *32*, Suppl: 452-461.

[31] Ekins. R. P. *J. Pharm. Biomed. Anal*. **1989**, *7*, 155-168.

[32] Finckh P.; Berger H.; Karl J.; Eichenlaub U.; Weindel K.; Hornauer H.; Lenz H.; Sluka P.; Weinreich G. E.; Chu, F.; Ekins R. *Proc. U. K. Natl. Ext. Qual. Assess. Serv. Meeting* **1998**, *3*, 155-165.

[33] Joos T. O.; Schrenk M.; Hopfl P.; Kroger K.; Chowdhury U.; Stoll D.; Schorner D.; Durr M.; Herick K.; Rupp S.; Sohn K.; Hammerle H. *Electrophoresis* **2000**, *21*, 2641-2650.

[34] Robinson W. H.; Digennaro C.; Hueber W.; Haab B. B.; Kamachi M.; Dean E. J.; Fournel S.; Fong D.; Genovese, M. C.; Neuman de Vegvar, H. E.; Skriner, K.; Hirschberg, D. L.; Morris, R. I.; Muller, S.; Pruijn, G. J.; Van Venrooij, W. J.; Smolen, J. S.; Brown, P. O.; Steinman, L.; Utz1et, P. J. *Nat. Med*. **2002**, *8*, 295-301.

[35] Yan F.; Sreekumar A.; Laxman B.; Chinnaiyan A. M.; Lubman D. M.; Barder T. J. *Proteomics* **2003**, *3*, 1228-1235.

[36] Patwa T. H.; Zhao J.; Anderson M. A.; Simeone D. M.; Lubman D. M. *Anal. Chem*. **2006**, *78*, 6411-6421.

Chapter 2

Comparative proteomic analysis of low stage and high stage endometrioid ovarian adenocarcinomas

2.1 Introduction

Ovarian cancer, the second most common gynecological malignancy, accounts for 3% of all cancers among women in the United States [1]. Five-year survival rates can be as high as 94% with early detection of the malignancy. However, as ovarian cancer is insidious in onset, less than 20% of ovarian tumors present at early stages of tumor development. The 5-year survival of women with advanced stage ovarian cancer is approximately 28%, largely because existing therapies for widespread disease are rarely curative [1].

Approximately 90% of ovarian cancers are thought to arise from the ovarian surface epithelium or from surface epithelial inclusion glands and/or cysts. Epithelial ovarian cancers (carcinomas) occur as several distinct morphological subtypes, including serous, endometrioid, clear cell, and mucinous tumors, based on their appearance under

the light microscope. Combined, serous and endometrioid carcinomas account for ~75-80% of all ovarian cancer. Substantial evidence suggests that the different morphological subtypes of ovarian carcinoma likely represent biologically and genetically distinct disease entities [2]. For example, about 85% of mucinous ovarian adenocarcinomas harbor *K-RAS* gene mutations, while *K-RAS* mutations are much less frequently observed in clear cell, endometrioid, and typical (high grade) serous carcinomas [3,4,5]. Likewise, mutations of the *CTNNB1* gene (encoding β-catenin) are observed in 16-38% of ovarian endometrioid adenocarcinomas, but are uncommon in the other types of ovarian carcinomas [6,7,8,9,10,11]. As in other cancers, the genes mutated in ovarian cancer typically encode proteins that function in conserved signaling pathways [12].

Recent investigations of ovarian endometrioid cancers have provided insights into the signaling pathways dysregulated during ovarian cancer pathogenesis. For example, the canonical Wnt signaling pathway is frequently deregulated in ovarian endometrioid adenocarcinomas (OEAs), and usually occurs as a consequence of activating mutations in the gene (*CTNNB1*) that encodes β-catenin [13,7,9,14,15]. Similarly, mutations predicted to deregulate the PI3K/Pten/Akt signaling pathway, such as inactivating mutations of the *PTEN* or activating mutations of the PI3K catalytic subunit, *PIK3CA*, are also frequently observed in OEAs [16, 17, 18]. More recently, Wu

and colleagues [19] have demonstrated that OEAs with defects in Wnt/β-catenin and/or PI3K/Pten/Akt signaling are usually low grade, low stage tumors, whereas high grade, high stage (stages 3/4) OEAs typically harbor p53 mutations and lack Wnt/β-catenin and PI3K/Pten/Akt signaling pathway defects.

Though comprehensive studies of tumor RNA and DNA have provided a number of insights into ovarian cancer pathogenesis, proteins are the major effector molecules in tumor cells. Protein levels may be discordant with corresponding transcript levels, and post-translational modifications can have biologically critical effects on protein function. To date, a number of proteomics-based studies have been conducted on ovarian tumors or on sera from patients with ovarian tumors using several methods, including 2-D PAGE [20,21], Surface Enhanced Laser Desorption Ionization (SELDI) [22], and 2-D liquid-based separation methodologies [23,24]. The latter methodology not only has greater reproducibility, but also has the ability to identify and quantify proteins, and the capability to compare results amongst different sets of experiments and samples [25]. Cancer-specific protein expression patterns in specific histologic subtypes of ovarian carcinomas have not been defined.

In this study, we compared protein expression profiles for the low stage (FIGO stage 1 or 2) and high stage (FIGO stage 3 or 4) OEAs. In general, the low stage tumors

lacked p53 mutations and had frequent *CTNNB1, PTEN,* and/or *PIK3CA* mutations, while the high stage tumors were high grade, had mutant p53 and lacked mutations predicted to deregulate Wnt/β-catenin and PI3K/Pten/Akt signaling. We employed 2-D liquid-based separation coupled with ESI-TOF MS to accurately obtain intact protein molecular weights, and both MALDI-MS, and MALDI-QIT-MS for protein identification. Using these methods, protein expression profiles of both low and high stage OEAs were obtained from consecutive pH ranges (pH 5.6-4.6), and were visualized by a 2-D mass map. We have identified proteins distinguishing low versus high stage OEA by mass spectrometry. Interestingly, when only high-grade (grade 2 or 3) OEAs were included in the analysis, the tumors still tended to cluster according to stage, suggesting that the altered protein expression was not solely dependent upon tumor cell differentiation. Additionally, we have demonstrated that the proteomic profiles of OEA are distinguishable from those of other types of ovarian cancer.

2.2. Materials and Methods

2. 2. 1. Sample preparation

We have analyzed twelve OEA tumor samples: six were low stage (stage 1) and six were high stage (stages 3/4). Further, two tumors were analyzed twice, using independent

tumor lysates. We also analyzed two OEAs whose staging and mutational status remained hidden to the investigators until after completion of experiments within this study. The OEA proteomic data obtained herein was also compared to proteomic data obtained from a previously published analysis of 16 ovarian clear cell carcinomas and 8 ovarian serous carcinomas [26], using principal components analysis to view tumor interrelationships. The University of Michigan Institutional Review Board (IRB) approved the experimental protocol. The OEAs were histologically analyzed and staged by a Board-Certified Pathologist (K.R.C.) prior to utilization in this study. The tumor tissues were stored at -80°C until use. The tissue samples were solubilized in lysis buffer using a mini-bead beater (Biospec, Inc., Bartlesville, OK). The lysis buffer was composed of 7M urea (Sigma-Aldrich, St. Louis, MO), 2M thiourea (Sigma-Aldrich), 100mM DTT (Sigma-Aldrich), 0.2% n-octyl-β-D-glucopyranoside (OG; Sigma-Aldrich), 0.5% Biolyte ampholytes, pH 3–10 (Bio-Rad, Hercules, CA), 10% glycerol (Sigma-Aldrich), and 2% (v/v) of 50X diluted protease inhibitor cocktail solution. The homogenized tissue samples were incubated at room temp for 30 min, then centrifuged (35,000 x g, 1 h, 4°C) to pellet insoluble material. The lysed samples were exchanged from the lysis buffer solution to the buffer solution for further chromatofocusing experiments, using a PD–10G column (Amersham Bioscience, Piscataway, N.J.). The

Bradford protein assay (Bio-Rad) was used for protein quantitation.

2. 2. 2. Chromatofocusing

Chromatofocusing (CF) was performed on a Beckman System Gold model 127s binary HPLC pump with high-pressure mixer (Fullerton, CA, USA), and a HPCF 1D column (2.1 x 250 mm) (Eprogen, Inc., Darien, IL). Two buffer solutions (a start buffer (SB) and an elution buffer (EB)) were utilized for the generation of a pH gradient on the CF column. The SB solution was composed of 6M urea and 25mM Bis-Tris (pH 7.4). The EB solution was composed of 6M urea and 10% polybuffer74 (pH 3.8). Both buffer solutions were brought to pH by addition of a saturated solution of iminodiacetic acid. The CF column was pre-equilibrated with SB, after which the mobile phase was switched to EB solution for pH gradient initiation. The pH gradient was monitored on-line with a post detector pH flow cell (Lazar Research Laboratories, Inc., Los Angeles, CA). The UV absorbance of the eluent was monitored on-line at 280nm. The flow rate was 0.2ml/min, with fractions collected in 0.2 pH units. Each fraction was stored at -80°C until use.

2. 2. 3. On-line mass mapping by NPS-RP-HPLC with LC–ESI–TOF MS

The second dimensional separation and MW determination of proteins were performed simultaneously using ESI-TOF-MS (LCT system, Waters-Micromass, Manchester, UK) coupled with non-porous silica (NPS) reverse phase HPLC (RP-HPLC). The separation was performed with a ODS-I column (4.6 x 33mm), which is a packed column with polymeric C18, 1.5μm silica (Eprogen, Inc., Darien, IL). The UV absorbance was monitored on-line at 214nm on a Beckman 166 System Gold HPLC system. A water/acetonitrile gradient, containing 0.1% trifluoroacetic acid (TFA), was used, with 0.3% of formic acid (v/v) in 50% acetonitrile solution added separately with a post column splitter before ESI-TOF-MS. The gradient was as follows: from 5% to 10% in 1 min, 10% to 25% in 1 min, 25% to 35% in 3 min, 35% to 45% in 5 min, 45% to 75% in 10 min, 75% to 100% in 1 min, 100% in 1 min, and 100% to 5% in 1 min. The flow rate was 0.4ml/min for the separation. Approximately 30% of the eluent from HPLC was delivered on-line to the ESI-TOF MS, with the remainder collected using an automated fraction collector. The ESI-TOF MS was set as follows: capillary voltage 3000V, sample cone 37V, extraction cone 5V, desolvation temperature 200°C, source temperature 110°C, desolvation gas flow 600L/h, and maximum nebulizer gas flow. Prior to each experiment, the instrument was externally calibrated by direct injection of the NaI/CsI solution. Five pM bovine insulin (Sigma-Aldrich) was added to the sample for internal

standard and normalization. The intact MW value was determined by deconvoluting the combined ESI spectra of the protein utilizing Protein Trawler™ (BioAnalyte, Inc., Portland, ME). The Protein Trawler™ program allows fast deconvolution processing based on MaxEnt 1 algorithm in MassLynx V 4.0 program.

2. 2. 4. Trypsin Digestion

All of the NPS RP-HPLC sample fractions were lyophilized to 20 μl, with any remaining TFA neutralized by addition of 10% (v/v) NH_4HCO_3 into samples. After addition of 10% (v/v) DTT and 0.25μg of TPCK-treated trypsin (Promega, Madison, WI) to each sample, the samples were vortexed, and then incubated at 37°C for 18h. The tryptic digestion was terminated by addition of 2.5 % (v/v) of TFA.

2. 2. 5. MALDI-MS and MALDI-QIT-MS Analysis

The digested samples were desalted on 2 μm C18 ZipTips (Millipore) before spotting onto a MALDI/QIT plate. MALDI/QIT targets were prepared by depositing 0.5μl of the sample on the plate with 0.5μl of matrix solution (1:4 dilution of saturated α-cyano-4-hydroxycinnamic acid (α-CHCA, in 50% ACN, 1% TFA), containing 50 fmol of angiotensin I, adrenocorticotropin (ACTH, amino acid 1-17), and adrenocorticotropin

(ACTH, 18-39) for the internal calibration. The mixture was dried at ambient conditions. The delayed extraction reflectron mode was utilized during MALDI–TOF MS analysis. The operation condition was as follows: operating voltage 20kV, extraction voltage 19950V, focusing voltage 16kV, reflectron 24.5kV, and nitrogen laser operating 337nm. The pulse voltage was 2300V, and sensitivity was 100mV. Peptide mass spectra, with a mass range from 800–5000 Da, were obtained during MALDI–TOF MS analysis. The matrix solution for MALDI-QIT was prepared from DHB (2, 5,-dihydroxybenzoic acid) in 50% ACN, 0.1% TFA. Data acquisition and processing during MALDI-QIT-TOF analysis were controlled by Kratos "Launchpad" software, with the standard instrument setting for optimum transmission at medium mass. The MALDI-QIT-TOF was externally calibrated using fullerite deposited directly onto the plate. The data acquired from MALDI-TOF MS analysis was searched in MS-fit (http://prospector.ucsf.edu/ucsfhtml13.4/msfit.htm) for protein identification. For the MS-fit, the SWISS-PROT (http://us.expasy.org) databases were searched using the following parameters: (1) species: homo sapiens; (2) maximum number of missed cleavages: 1; (3) possible modifications: peptide N-terminal glutamine to pyroglutamic acid, oxidation of methionine, protein N-terminus acetylated, and phosphorylation of serine, threonine and tyrosine; (4) peptide mass tolerance: 50 ppm; (5) MW ranged from

1000 to 90000 Da; (6) pI range of proteins 4–8. The MALDI-QIT-TOF confirmed the identified proteins by MALDI-MS. The data acquired from MALDI-QIT-TOF were searched in Mascot using SWISS-PROT with the following parameters; (1) species: homo sapiens (2) possible modifications: peptide N-terminal glutamine to pyroglutamic acid, oxidation of methionine, protein N-terminus acetylated, and phosphorylation of serine, threonine and tyrosine (3) peptide tolerance: 1.2Da (4) MS/MS tolerance: 0.8Da. (5) Peptide charge: +1. The following criteria were used to obtain a confident search result: (1) MOWSE score was at least 10^3; (b) four minimum matched peptides (c) two or more consecutive y-or b-series ions for MS/MS; (d) total six or more y-or b ions for MS/MS.

2. 2. 6. Data analysis

To compare protein expression between low stage (stage 1) versus high stage (stages 3/4) OEAs, MW and intensity of proteins obtained by deconvoluting the ESI spectra were converted to a single text file, and were then transformed by logarithm10 (max[X+100,0]+100). Higher abundance proteins were selected and colored in the mass map if the t test p-value of the log-transformed data between the two groups (low and high stage OEAs) was <0.05, with $\geq$2-fold increase in the protein expression ratio. Both

hierarchical clustering (HC) and principal component analysis (PCA) techniques were utilized to view sample relationships, based upon protein expression similarities. The HC produced dendrograms in which points are close if they have greater correlation, found by calculating the Pearson correlation coefficient. PCA facilitated a two-dimensional view of protein expression of different types of ovarian cancer (serous, clear, endometrioid) at pH 5.4-5.2, thus transforming the dataset to a new coordinate system such that the greatest variance from expression dataset turn into the first principal component, the second greatest variance becomes the second component, and so on. Two components displayed the relationship of the dataset in which similar data are more closely located. Two sets of protein expression data were compared by analysis with DeltaVue software (Beckman-Coulter), with each protein map having a different color (red, green) and differences between the two protein maps shown in a third map (middle).

2. 2. 7. Ovarian tumor tissue microarrays and immunohistochemistry

Two ovarian tumor tissue microarrays were constructed as described [27], with the tissues randomly selected from the University of Michigan Pathology archives. These two tissue microarrays (combined) contained triplicates of ovarian tumor samples, including 5 clear cell carcinomas, 26 ovarian papillary serous carcinomas, 6 ovarian

endometrioid carcinomas, 15 metastatic ovarian carcinomas, 8 mixed ovarian serous/endometrioid carcinomas, 4 ovarian serous/clear cell carcinomas, 1 papillary serous cystadenocarcinoma and 1 mixed anaplastic/endometrioid carcinoma. Immunohistochemistry for Lamin A/C was performed using a mouse monoclonal antibody (clone JOL2, Chemicon) at 1:50 dilution; immunohistochemistry for S100A9 (Calgranulin B) was performed using a mouse monoclonal antibody (Clone 1C10, Novus Biologicals) at 1:600 dilution. Anti-Lamin A/C and anti-S100A9 primary antibodies were hybridized to both tissue microarrays for 30 min at RT using citrate buffer (pH 6.0) and microwave antigen retrieval (10 min) and the Dako automated instrument (Dako Cytomation, Carpinteria, CA). Detection of the primary antibody was performed using the Dako Envision kit.

2.3. Results and Discussion

In this work, 12 ovarian endometrioid carcinoma tumor samples (six low stage (stage 1) and six high stage (stages 3/4) were analyzed in the initial training set by 2-D liquid-based protein separation methodology. Differentially expressed proteins in the low stage versus high stage tumors were subsequently identified using mass spectrometry. Extracted tumor proteins were resolved by chromatofocusing in the first dimensional

separation (based on pI), followed by NPS RP-HPLC separation in the second dimension (based on hydrophobicity). Intact protein MW values obtained by ESI-TOF MS were deconvoluted by Protein Trawler^{TM} to generate a list of protein MW and intensities. Proteins were identified using MALDI-TOF-MS, and confirmed by MALDI-TOF-QIT MS/MS. 2-D mass maps (based upon the intact protein pI and MW) facilitated comparison of protein expression between different samples. OEA tumor samples were categorized as either low stage (stage 1) or high stage (stages 3/4) tumors. Five of six low stage OEAs had *CTNNB1*, *PTEN*, and/or *PIK3CA* mutations and lacked p53 mutations. In contrast, five of six high stage (stages 3/4) OEAs had mutant p53 and lacked mutations predicted to deregulate Wnt/β-catenin and PI3K/Pten/Akt signaling (Table 2.1).

We used chromatofocusing (CF) to resolve extracted proteins (based upon protein pI) to compare protein expression profiles associated with the clinical stage of the tumors. We utilized a pH range of 5.6-4.6, as the greatest number of the proteins from the CF column eluted in this range according to the UV absorption profile. Within the selected pH range, the second dimensional separation and the ESI-TOF-MS were simultaneously performed to generate 2D-mass maps displaying both protein MW and expression level. Moreover, the 2-D mass maps facilitated comparison of differential protein expression

between the low stage and high stage tumors. A comparison of the mass maps for low stage and high stage OEA tumors from pH 5.6-4.6, at 0.2 pI intervals, is shown in Figure 2.1. The blue bands represent over-expression of proteins in high stage tumors; the purple bands indicate proteins with over-expression in low stage tumors. Overall, more proteins were over-expressed in high stage OEAs, than were over-expressed in low stage OEAs. As shown in Figure 2.1, proteins over-expressed in the high stage OEAs were more frequently present in the pI range of 5.4-5.2, whereas proteins over-expressed in low stage OEAs were more frequently present in a lower pI range (e.g., pH 4.8-4.6). This subtle pH change in terms of protein expression based on OEA stages may indicate post-translational changes in protein expression during tumor progression.

Differences in protein expression between low stage and high stage OEAs might be related to the mutational status of the OEAs. *CTNNB1*, *PTEN*, *PIK3CA* and/or p53 mutations may all have profound effects on global protein expression in the tumor cells. Table 2.2 indicates the proteins over-expressed in low stage OEAs. Some of these proteins were previously identified as candidate markers of ovarian cancer, including Calgranulin B, and protein phosphatase 2, catalytic subunit, alpha isoform. Calgranulin B shows the most elevated level of expression difference between low stage and high stage OEAs in this study. It is a member of S100 family, and is a Ca^{2+}-binding protein

that plays a role in carcinogenesis through the inhibition of p53, and forms a complex with Calgranulin A to prevent phosphorylation of various molecules necessary for normal transcription and translation [24]. Additionally, Calgranulin B serves as a ligand that can bind to the receptor for advanced glycation end products (RAGE). This receptor-ligand complex may activate a cellular signaling cascade that influences cell division [28]. Calgranulin B was previously identified in both cystic ovarian carcinomas and in the serum of corresponding patients. This protein may have utility as a diagnostic biomarker to distinguish between malignant and benign ovarian tumors [29], and may have utility for the diagnosis of OEA and other forms of ovarian cancer.

Protein phosphatase 2, catalytic subunit, alpha isoform (PP2A) has been intensely studied in recent years [30, 31]. We have demonstrated an over-expression of PP2A in the low stage OEAs. PP2A may regulate the Wnt signaling pathway at multiple levels [32]. It is plausible that over-expression of PP2A results as part of futile negative feedback in tumors with downstream mutations in the Wnt signaling pathway. Thus, the over-expression of PP2A proteins in low stage OEAs is not surprising, as five of six low stage tumors used in the initial training set have a mutated *CTNNB1* gene encoding β-catenin. Other proteins of interest in OEA include Ras-related protein Rab1A (RAB1A), RUN and FYVE domain containing protein 1 (RUFY1), and inorganic pyrophosphatase

2, (PPase 2, PPA2; Figure 2.2). RAB1A is a member of an oncogene family, and plays a role in the regulation of vesicular transport from the endoplasmic reticulum to the Golgi complex [33]. In a study of alternative mRNA splicing in cancer, the Rab1A transcripts are only alternatively spliced variants [34], not with the regularly spliced variants found in normal tissues. The absence of regularly spliced isoforms may be associated with tumorigenesis, although that study did not include analysis of ovarian cancer. Rab1A protein may be associated with tumor development. Run and FYVE domain containing protein1 (RUFY1) contains C-terminal RUN domain and N-terminal FYVE domain with two coiled-coil domains in between. RUFY1 protein is tyrosine-phosphorylated and interacts with Etk (epithelial and endothelial tyrosine kinase) through the SH3 and SH2 domains [35]. Although the role of RUFY1 proteins is unclear, interaction with Etk appears to be involved in the regulation of endocytosis. Inorganic pyrophosphatase 2 is essential for mitochondrial function. According to previous studies [36, 37], inorganic pyrophosphatase 2 participates in regulation of cellular pH in the tumor, and is related to alkaline phosphatase, which regulates intracellular to extracellular movement of inorganic pyrophosphate.

Table 2.3 indicates the over-expressed proteins in the high stage OEAs, some of which were previously described in other types of malignancies. For example, ATP

synthase D (Figure 2.3) is over-expressed in high stage tumors in the pH 4.8-4.6 range, and has also been shown to be over-expressed in lung cancer [38]. It may play an important role in H^+ transportation, which is an essential mechanism for tumor adaptation in a microenvironment. High-affinity cGMP-specific 3', 5'-cyclic phosphodiesterase 9A (cIAP1) has been studied in both lung [39] and cervical cancer [40]. cIAP1 is over-expressed in lung cancer, and functions as a mediator of tumor necrosis factor (TNF) receptor signaling. This protein has anti-apoptotic activity [41] and protein expression is induced by hypoxia [42]. Tumor protein D53 is a member of tumor protein D52 family and its expression is elevated in breast cancer cells [43]. This protein shows over-expression at the G2-M transition and is a regulator of cell cycle progression.

Not all of the proteins identified in our study have obvious biological relevance in terms of OEA stage. Lamin A/C is involved in nuclear stability and chromatin structure, and has been primarily studied in lung cancer [44]. In lung cancer, the expression of Lamin A/C decreases along with the expression of other cytoskeletal proteins, including Vimentin. However, a previous analysis of ovarian cancer cell lines demonstrated Lamin A/C over-expression. In the present study, we have demonstrated a 1.8-fold increase in Lamin A/C expression in the stage 3/4 tumors that were analyzed.

Interestingly, four of the eight over-expressed proteins in low stage OEAs are

metal binding proteins. In high stage OEAs, three of nine over-expressed proteins are metal binding proteins. Most of the metal binding proteins are associated with protein structural changes, which may perturb protein functionality. Further, several ATP and GTP-binding proteins were over-expressed in the high stage OEAs. The increased expression of ATP and GTP-binding proteins in high stage OEAs may be associated with cancer progression.

In order to confirm the reproducibility of our findings, we repeated the assays for samples OE-24T and OE-28T. Thus, tumor OE-24T(R) was actually a second, independent sample of tumor OE-24T; tumor OE-28T(R) was actually a second, independent sample of tumor OE-28T. Both of the repeated tumor samples were processed in an identical manner to all the other tumor samples in our study, using the two-dimensional liquid-based separation/mass mapping methodology. Figure 2.4 displays the 2-D mass map comparing OE-24T(R) versus OE-24T at pH 5.0-4.8. Figure 2.5 displays the 2-D mass map comparing OE-28T(R) versus OE-28T at pH 5.4-5.2. In both cases, the protein bands of the tumor tissue from the second (independent) analysis matched well with the protein bands of the tumor tissue from the initial analysis. As expected, as both of the repeated tumors were high stage tumors, both tumors partitioned into the high stage tumors in the dendrogram (Figure 2.6).

To determine whether the protein signatures associated with high stage versus low stage tumors could correctly classify OEAs of unknown stage, we further analyzed two OEA (OE-37T and OE-104T) tumors whose staging and mutational status remained hidden to the investigators until after completion of experiments contained in this study, using the two-dimensional liquid-based separation/mass mapping methodology. Based upon protein identification and the mass mapping data described in this study, we have correctly assigned a clinical stage to both tumor samples. OE-37T partitioned into the other stage I OEAs in the hierarchical clustering dendrogram (Figure 2.6). At pH 5.4-5.2, the protein bands from OE-37T matched well with a stage 1 sample, OE-47T (Figure 2.7). OE-37T is a well-differentiated stage 1 OEA with mutant β-catenin. Figure 2.7 displays the 2-D mass map comparing OE-37T versus OE-47T at pH 5.4-5.2. The protein bands from the OE-37T tissue matched well with the protein bands in the OE-47T tumor tissue. OE-37T tissue partitioned into the other stage 1 tumors (Figure 2.5), as OE-37T has features similar to those of other stage 1 cancer tissue samples having mutated *CTNNB1* and nuclear β-catenin accumulation, indicating that tumors with similar mutations may show close relationships in the dendrogram. We also analyzed a second tumor sample, OE-104T, whose staging and mutational status remained hidden from the investigators until after analysis was completed. It clustered into the high stage

OEA group in Figure 6 and the protein bands at pH 5.2-5.0 matched well with the bands from OE-39T, which is a stage 3 OEA (Figure 2.8). Indeed, OE-104T is a stage 3 OEA that has a p53 mutation. Figure 8 displays the 2-D mass map comparing OE-104T versus OE-39T at pH 5.2-5.0. These results suggest that clustering using proteomic data alone may help to reveal staging information about the tumors. Our study shows that it is possible to resolve low stage OEA from high stage OEA on a proteomic level.

Further, we explored whether we would be able to distinguish OEA tumor staging independent of tumor grade (degree of tumor cell differentiation). To this end, we reanalyzed and re-clustered our data, using only the high grade OEAs (either low or high stage) previously analyzed. We compared the three (out of six) independent high grade OEAs that were diagnosed at low stage with the six (out of 7) independent high grade OEAs that were diagnosed high stage. Interestingly, although we did not find perfect segregation of tumor stage in our clustering dendrogram, we did find that all three low stage OEAs tended to cluster together (Figure 2.9). These results suggest that the altered protein expression was not solely dependent upon tumor cell differentiation.

The proteomic maps produced in this work may have utility to classify OEA from other ovarian carcinomas. Gene expression data suggests that OEA is similar to clear ovarian cancer in advanced stages and to serous ovarian cancer in the early stages

[45]. We have compared our proteomic data for endometrioid ovarian carcinomas obtained within to that of serous and clear cell proteomic data obtained previously [26]. Figure 2.10 shows the PCA of protein mass map data of three different sub-types (serous, endometrioid, clear) of ovarian cancer from the pI range 5.4-5.2, which is the pI range from the first dimension separation with the highest abundance of proteins. It is possible to classify the different sub-types of ovarian cancer by analysis of proteomic data, thus suggesting that different sub-types of ovarian cancer may be discernable at the protein level (Figure 2.10).

Immunohistochemical analysis of Lamin A/C and S100A9.

To further confirm the stage-dependent protein identifications, Lamin A/C and S100A9 (calgranulin B) expression in ovarian endometrioid tumors was assessed by immunohistochemistry (Fig. 2.11), using mouse monoclonal anti-Lamin A/C and anti-S100A9 antibodies and the ovarian tumor tissue microarrays. Strong nuclear membrane immunoreactivity for Lamin A/C was documented (Fig. 2. 11A, Fig. 2. 11B) in the majority of the ovarian endometrioid tumors. There were no significant staining differences observed with regard to tumor stage, in apparent discordance with the findings from the mass mapping experiments. However, using immunohistochemistry to

discern small differences in protein expression is difficult. We documented only a 1.8-fold difference in expression between early and late stage ovarian endometrioid tumors in the mass mapping experiments. This fold-change in expression levels may not be apparent by immunohistochemistry. Diffuse and consistent cytosolic immunoreactivity for S100A9 was documented in early stage ovarian endometrioid tumors, whereas late stage tumors exhibited low levels of immunoreactivity (Fig. 2.11C, Fig. 2.11D). These findings were concordant with the findings from the mass mapping experiments.

2.4. Conclusion

Gene expression studies have yielded limited results in distinguishing endometrioid ovarian cancer from other types of ovarian cancer. We have demonstrated that the profiling of proteins in tumor samples, based on pI and molecular weight using on-line mass mapping, has facilitated the correct staging of OEA. Protein expression levels for two tumors whose staging and mutational status was not revealed to the investigators until after analysis was completed were predictive of correct tumor stage in mass mapping dendrograms. Further, we have demonstrated that endometrioid ovarian cancer has a unique protein signature, as compared to other subtypes of ovarian cancer. Protein profiling can facilitate stage-dependent ovarian tumor classification.

Table 2.1. The Summary of OEAs utilized in this study

Tumor	Clinical data			*CTNNB1* mutation (exon 3)		*PTEN* mutation (exons 1-9)		*PIK3CA* mutation (exons 9 & 20)		*TP53* mutation (exons 5-8)		IHC
ID	Age	Stage	Grade	AA	Codon	AA	Codon (exon)	AA	Codon (exon)	AA	Codon (exon)	p53
OE-15T	61	1C	3									-
OE-17T	67	1C	2	Ser-Cys	33							-
OE-20T	69	4	2									+
OE-24T	66	3C	1							Tyr-Asn	236 (Exon 7)	+
OE-25T	58	3	3							Arg-His	273 (Exon 8)	+
OE-28T	73	4	3									+
OE-35T	70	4	3							Stop	196 (Exon 6)	-
OE-37T	40	1A	1	Ser-Cys	37							-
OE-39T	73	3C	2							Val-Leu	216 (Exon 6)	+
OE-44T	69	1	1	Ser-Cys	33							-
OE-47T	72	1A	1	Ser-Ala	37							-
OE-48T	37	1C	1	Gly-Glu	34	Thr-Pro	131 (Exon 5)					-
OE-71T	49	1C	2	Gly-Val	34			Glu-Lys	542 (Exon 9)			-
OE-104T	61	3C	3							Pro-Ser	151 (Exon 5)	+

Table 2.2 Over-expressed proteins in low stage (stage I) ovarian endometrioid tumors.

Protein Name	Access no.	Protein MW/pI	Observed MW	Observed pH range	Protein fold-changes	Mowse score	Sequence Coverage (%)
Calgranulin B	P06702	13242/5.7	13524	5.4-5.2	7.9	1.04E+05	54
Ras-related protein Rab-1A (YPT1-related protein)	P62820-3	13744/5.4	13795	5.2-5.0	2	7361	53
E2F transcription factor 7	Q96AV8	14660/5.10	14791	5.4-5.2	2.1	5642	36
Protein phosphatase 2, catalytic subunit, alpha isoform	P67775	35571/ 5.30	35504	5.6-5.4	2.7	2.11E+06	33
Inorganic pyrophosphatase 2, mitochondrial precursor (PPase 2)	Q9H2U2	37963/7.1	38065	5.4-5.2	2.56	2.29E+04	41
Ubiquitin ligase protein RNF8	O76064	55518/7.1	55557	5.0-4.8	4.8	8.84E+04	41
RUN and FYVE domain containing protein 1	Q96T51	69079/5.6	69153	4.8-4.6	5.5	2.10E+06	37
Myotubularin-related protein 6	Q9Y217	71955/7.6	71961	5.2-5.0	2	1.28E+08	42

Table 2.3 Over-expressed proteins in high stage (stage III/IV) ovarian endometrioid tumors.

Protein Name	Access no.	Protein MW/pI	Observed MW	Observed pH range	Protein fold-changes	Mowse score	Sequence Coverage (%)
ATP synthase D chain, mitochondrial	O75947	18479/5.21	18594	5.4-5.2	2.39	6689	72
Tumor protein D53 (hD53) (D52-like 1)	Q16890	22449/5.5	22464	5.4-5.2	2	2.13E+05	47
Ras-related protein Rab-14	P61106	23766/5.8	24085	5.4-5.2	2.8	1.75E+04	44
Potassium channel tetramerisation domain containing protein 4	Q8WVF5	29996/7.0	29979	5.4-5.2	2.9	7037	40
Alkaline phosphatase, tissue-nonspecific isozyme precursor	P05186	57279/6.2	57240	5.4-5.2	1.3	6.62E+06	33
T-complex protein 1, theta subunit (TCP-1-theta)	P50990	59621/5.4	59565	5.6-5.4	2	2.90E+08	36
Lamin A/C (70 kDa lamin)	P02545-2	65134.8/6.40	65145	4.8-4.6	1.8	2.93E+11	48
High-affinity cGMP-specific 3',5'-cyclic phosphodiesterase 9A	O76083	68493/5.9	68468	5.4-5.2	7.4	2.58E+06	33
Baculoviral IAP repeat-containing protein 2	Q13490	69899.67/6.27	69844	5.4-5.2	2.3	2.10E+06	37
Vesicle-fusing ATPase	P46459	82655/6.4	82717	5.2-5.0	2.1	9.86E+05	29

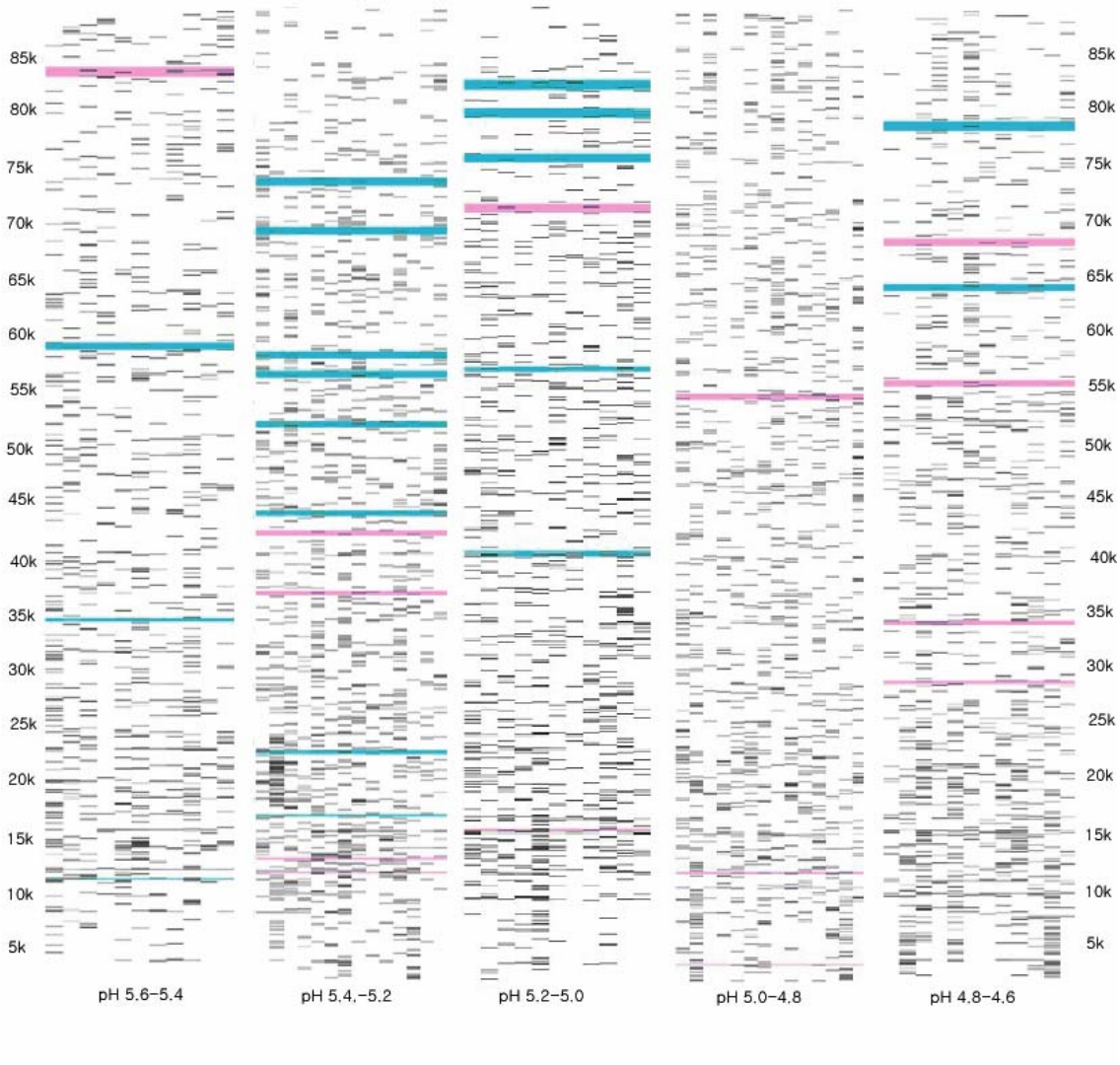


Figure 2.1 Comparison of mass maps showing protein expression differences between low stage (stage 1) and high stage (stage 3/4) ovarian endometrioid tumors across several pH ranges. The blue bands represent over-expression of proteins in the high stage tumors. The purple bands represent over-expression of proteins in the low stage tumors.

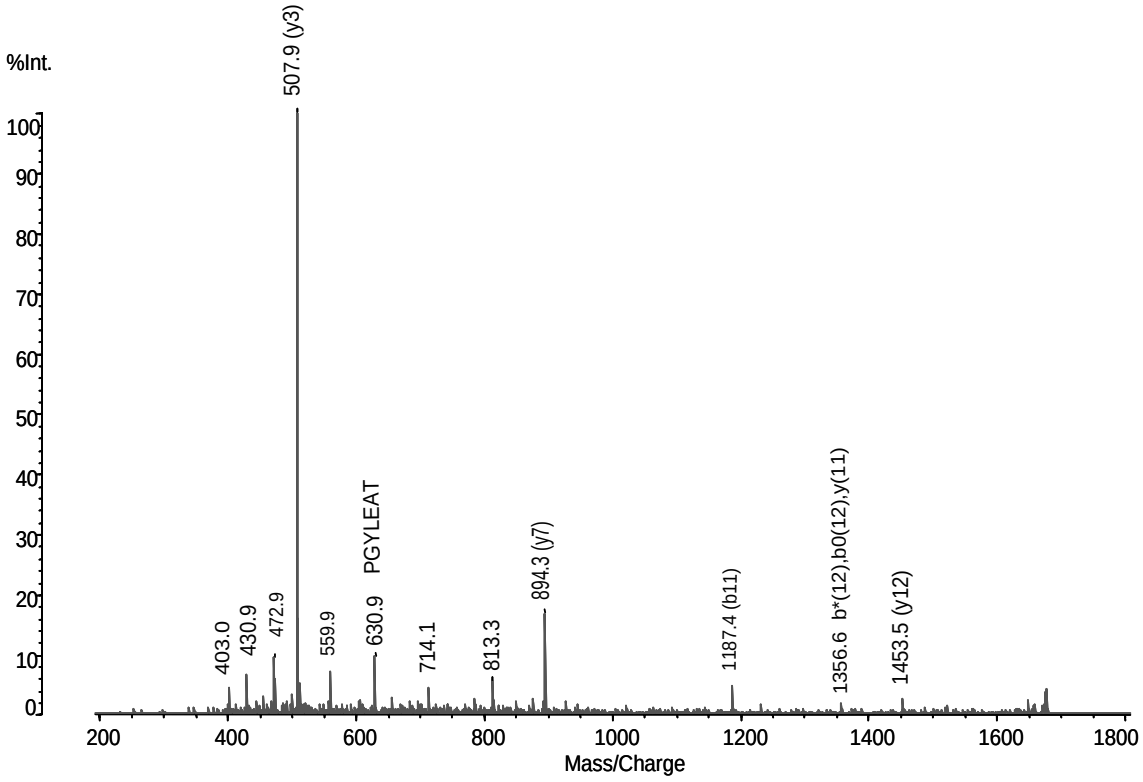


Figure 2.2 Identification of Inorganic pyrophosphatase. QIT-TOF-MS spectra of the peak for m/z 1694.8

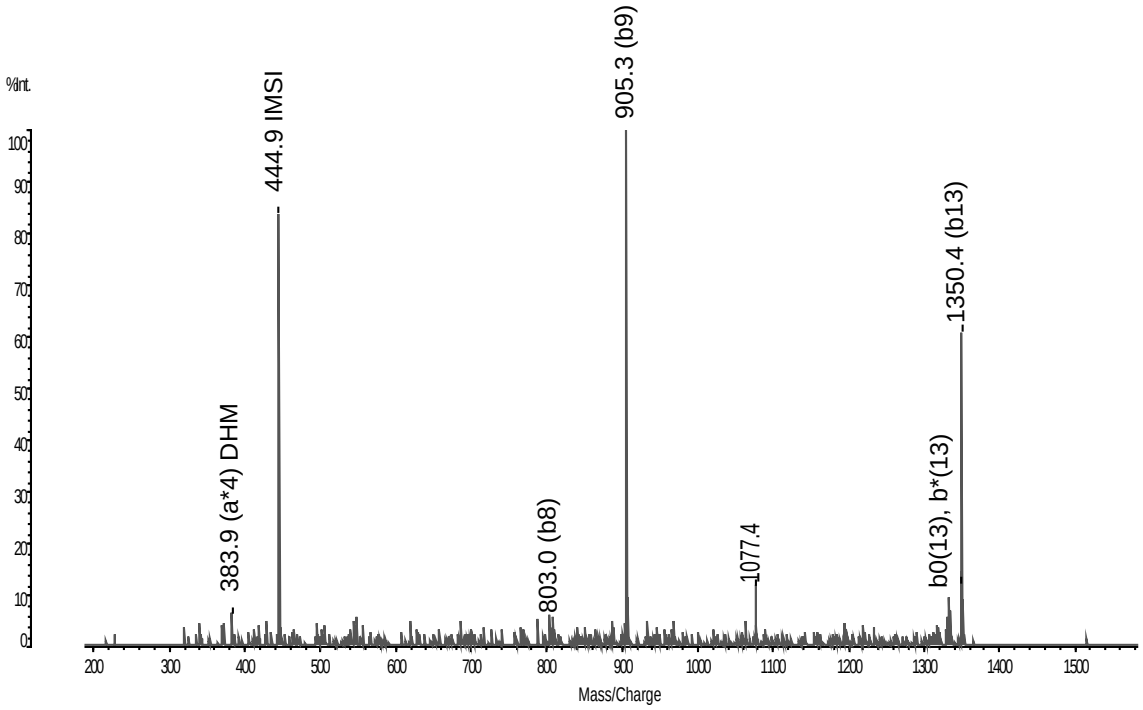


Figure 2.3 Identification of ATP synthase, synthase D chain, mitochondrial. QIT-TOF-MS spectra of peak m/z 1351.6

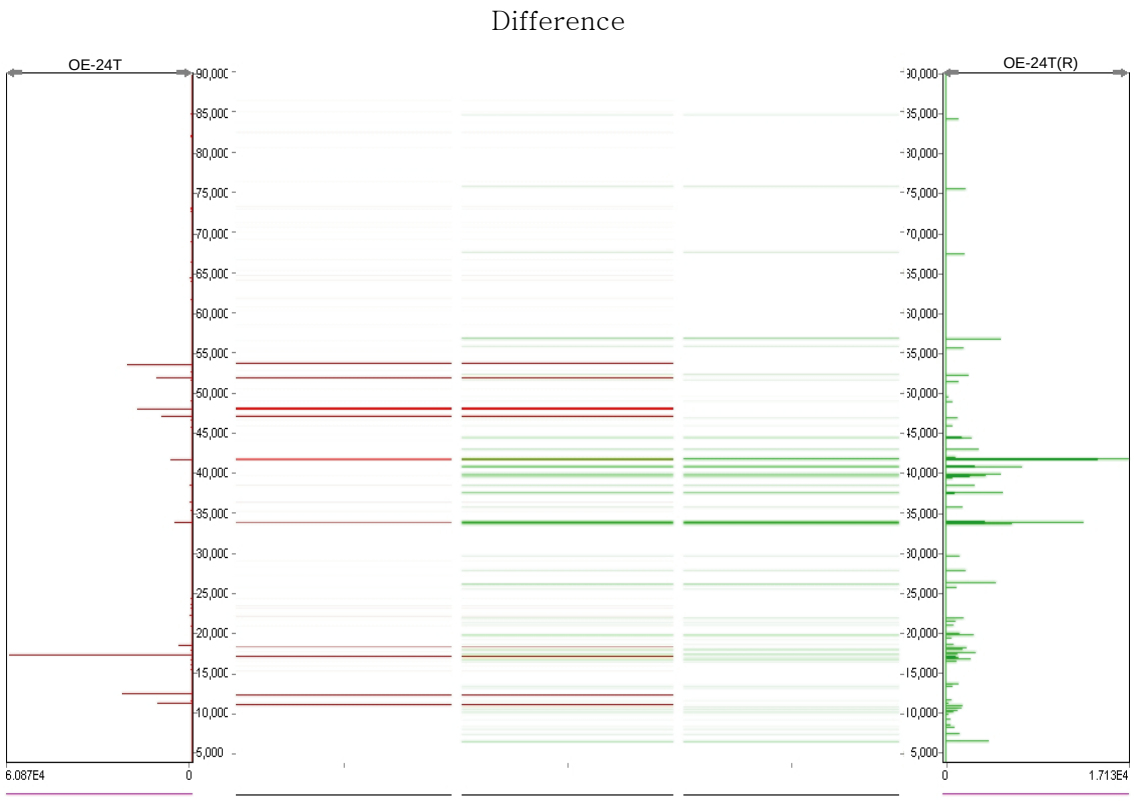


Figure 2.4 A mass map showing comparison of protein expression between two individual assays from the ovarian endometrioid tumor. Shown is OE-24T (left) and on the right is a repeat analysis, using a lysate prepared from a different region of the same tumor).

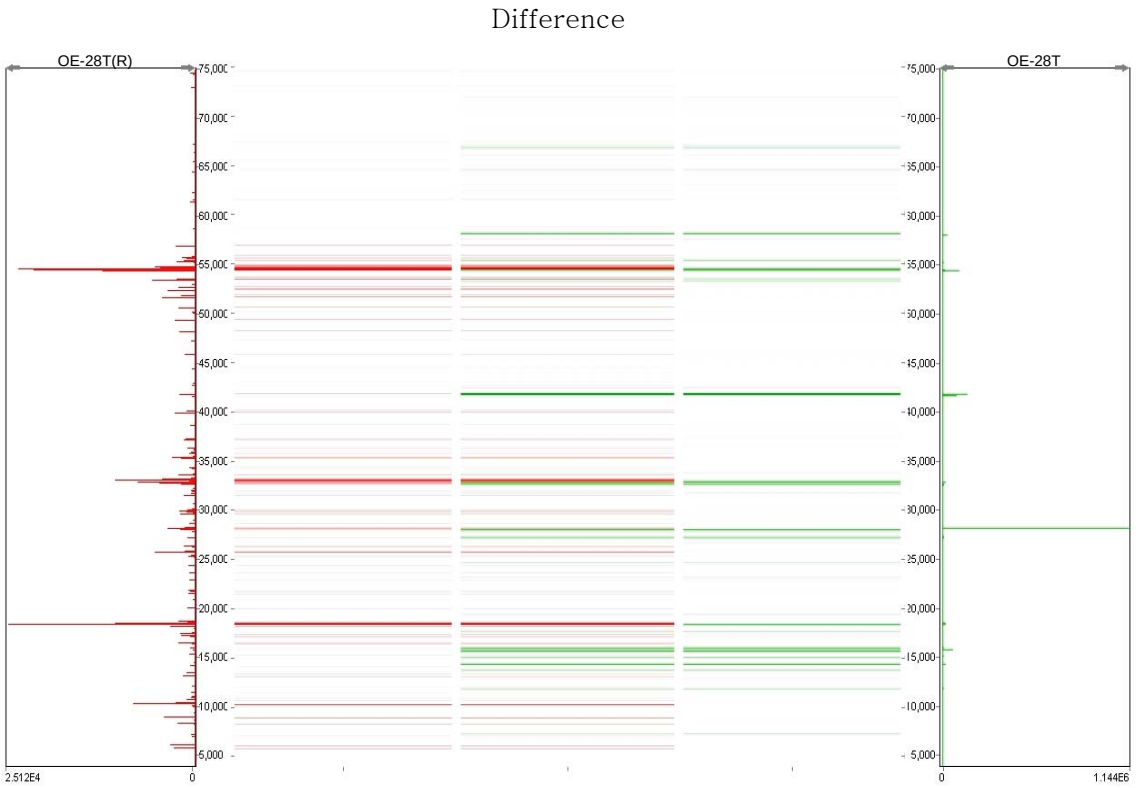


Figure 2.5 A mass map showing comparison of protein expression between two individual assays from the ovarian endometrioid tumor. Shown is OE-28T (right) and on the left is a repeat analysis, using a lysate prepared from a different region of the same tumor).

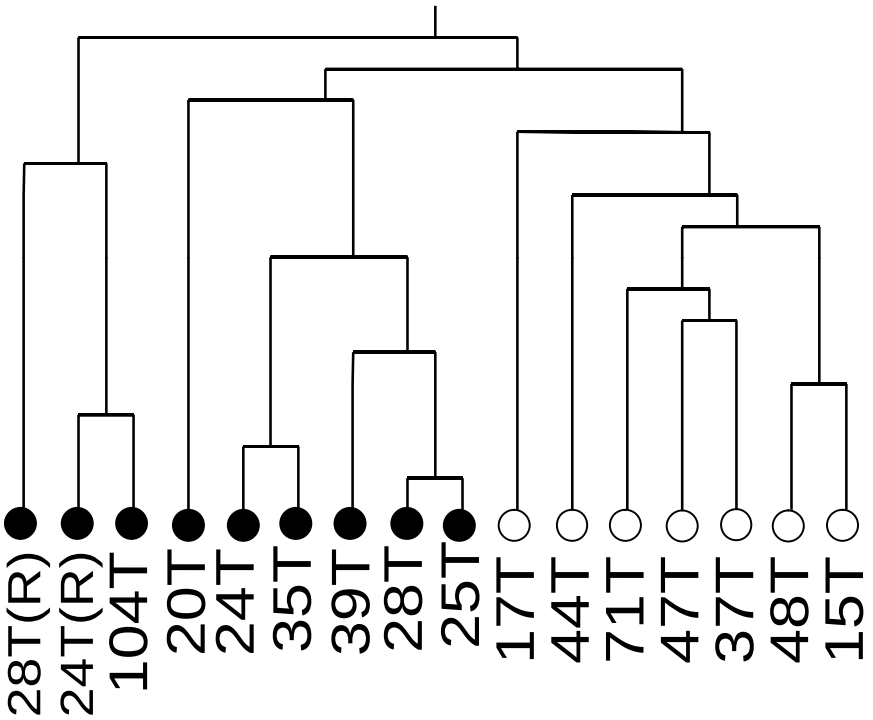


Figure 2.6 A hierarchical clustering dendrogram of the OEAs utilized in this study. Included in the dendrogram are the two repeat sample analyses (OE-28T(R) and OE-24T(R)), corresponding to tumors OE-28T and OE-24T, respectively. Also included are the two tumors (OE-37T and OE-104T) whose stage and mutational status was not revealed to the investigators until after analysis was completed.
White dot: low stage (stage 1).
Black dot: high stage (stage 3/4).

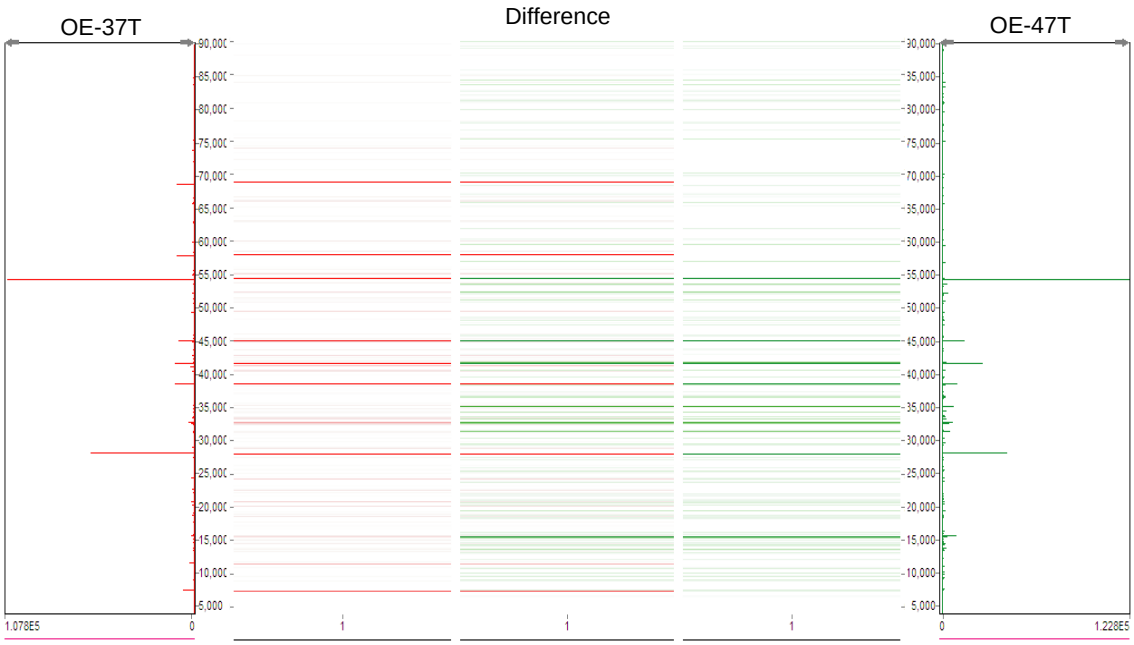


Figure 2.7 A mass map comparison of protein expression between OE-47T, a stage 1 ovarian endometrioid tumor (right) and OE-37T, an OEA whose stage and mutational status was not revealed to the investigators until after analysis was completed.

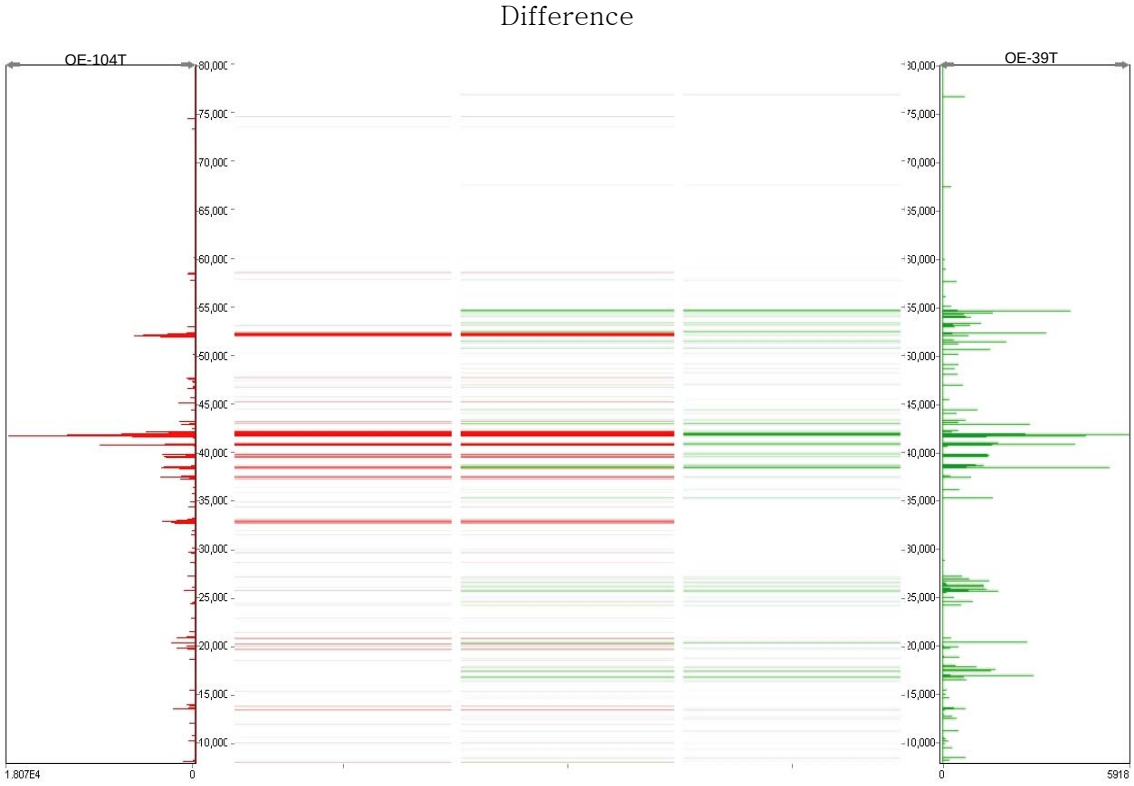
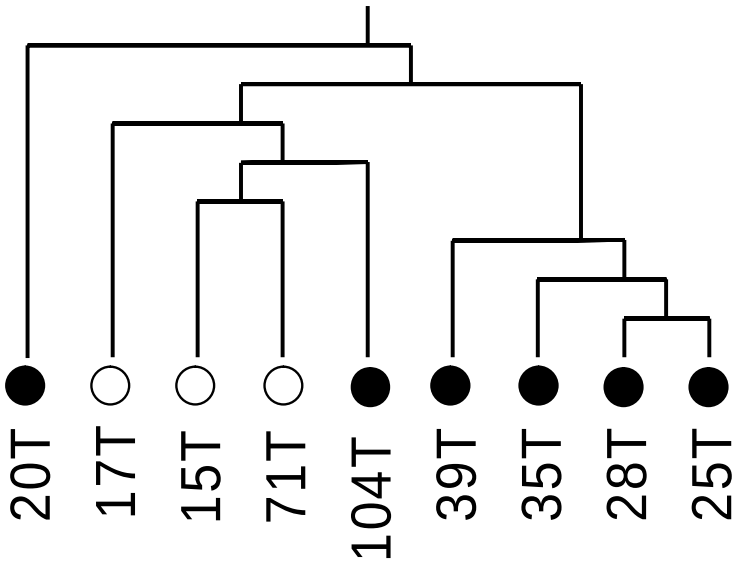


Figure 2.8 A mass map comparison of protein expression between OE-39T, a stage III ovarian endometrioid tumor (right) and OE-104T, an OEA whose stage and mutational status was not revealed to the investigators until after analysis was completed.



All samples are high grade (2 and 3)
Grade 2: 17T, 71T, 20T, 39T
Grade 3: 15T, 25T, 28T, 35T, 104T

Figure 2.9 A hierarchical clustering dendrogram of only the high grade OEAs utilized in this study.
White dot: low stage (stage 1).
Black dot: high stage (stage 3/4)

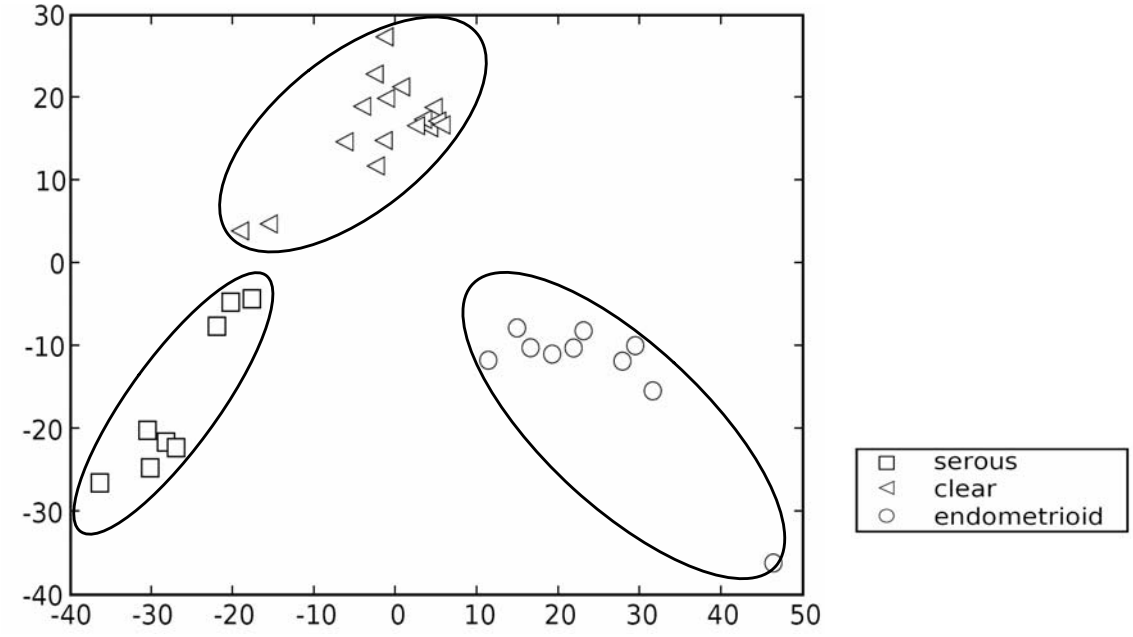


Figure 2.10 PCA analysis of Endometrioid, Serous and Clear cell ovarian cancer.

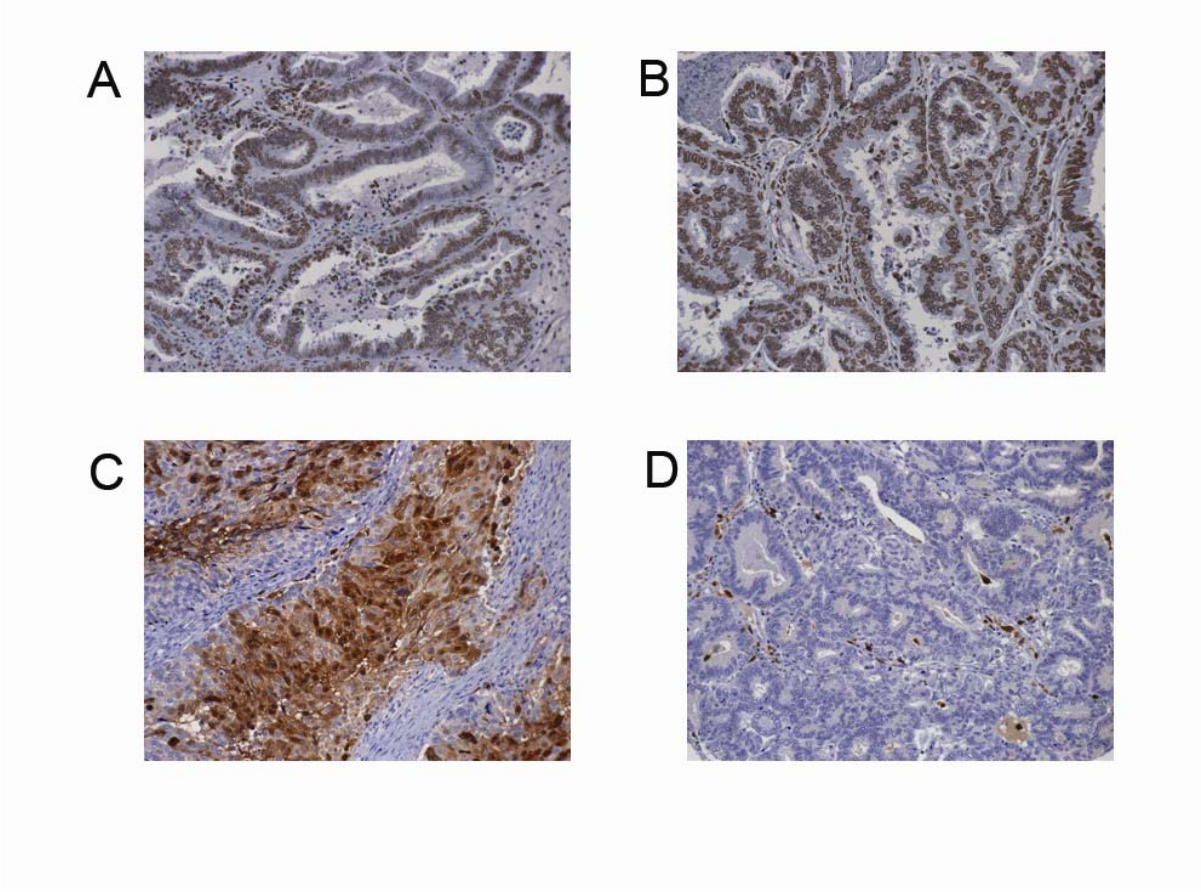


Figure 2.11 Immunohistochemical staining of Lamin A/C and S100A9 (calgranulin B). Both early stage (stage 1; A) and advanced stage (stage 3; B) ovarian endometrioid tumors stained for Lamin A/C immunoreactivity are shown, demonstrating predominantly nuclear membrane immunoreactivity in the neoplastic cells, with slightly higher immunoreactivity in the advanced stage tumor. Both early stage (stage 1; C) and advanced stage (stage 3; D) ovarian endometrioid tumors stained for S100A9 immunoreactivity are shown, demonstrating diffuse and consistent cytosolic immunoreactivity in the neoplastic cells of the early stage tumor (C), with inflammatory cells showing immunoreactivity in the advanced stage tumor (D).

2.5 References

[1] Jemal, A.; Murray, T.; Ward, E.; Samuels, A.; Tiwari, R. C. et. al., *Cancer Statistics*. Ca-A Cancer Journal for Clinicians **2006**, *55*, 10-30.

[2] Bell, D. A. *Mod. Pathol.* **2005**, *18, Suppl 2*, S19-32.

[3] Ichikawa, Y.; Nishida, M.; Suzuki, H.; Yoshida, S.; Tsunoda, H.; Kubo, T.; Uchida, K.; Miwa, M. M *Cancer Research* **1994**, *54*, 33-35.

[4] Cuatrecasas, M.; Villanueva, A.; Matias-Guiu, X.; Prat, J. *Cancer* **1997**, *79*, 1581-1586.

[5] Enomoto, T.; Weghorst, C. M.; Inoue, M.; Tanizawa, O.; Rice, J. M. *Am. J. Pathol.* **1998**, *139*, 777-785.

[6] Gamallo, C.; Palacios, J.; Moreno, G.; Calvo de Mora, J.; Suarez, A.; Armas, A. *Amer. J. Path.* **1999**, *155*, 527-536.

[7] Sagae, S.; Kobayashi, K.; Nishioka, Y.; Sugimura, M.; Ishioka, S.; Nagata, M.; Terasawa, K.; Tokino, T.; Kudo, R. *Jpn. J. Cancer Res.* **1999**, *90*, 510-515.

[8] Wright, K.; Wilson, P.; Morland, S.; Campbell, I.; Walsh, M.; Hurst, T.; Ward, B.; Cummings, M.; Chenevix-Trench, G.. *Int. J. Cancer* **1999**, *82*, 625-629.

[9] Wu, R.; Zhai, Y.; Fearon, E. R.; Cho, K. R. *Cancer Res.* **2001**, *61*, 8247-8255.

[10] Kildal, W.; Risberg, B.; Abeler, V.; Kristensen, G..;, Sudbø, J.; Nesland, J.; Danielsen, H. *Eur. J. Cancer* **2005**, *41*, 1127-1134.

[11] Catasus, L.; Bussaglia, E.; Rodrguez, I.; Gallardo, A.; Pons, C.; Irving, J.; Prat, J. *Hum. Pathol.* **2004**, *35*, 1360-1368.

[12] Vogelstein, B.; Kinzler, K. W. *Nature Medicine* **2004**, *10*, 789-799.

[13] Palacios J.; Gamallo C. *Cancer Research* **1998**, *58*, 1344-1347.

[14] Zhai, Y.; Wu, R.; Schwartz, D. R.; Darrah, D.; Reed, H.; Kolligs, F. T.; Nieman, M. T.; Fearon, E. R.; Cho, K. R. *Am. J. Pathol.* **2002**, *160*, 1229-1238.

[15] Schwartz, D. R; Wu. R.; Kardia S. L; Levin A. M; Huang C. C. *Cancer Res.* **2003**, *63*, 2913-2922.

[16] Tashiro, H.; Blazes, M. S.; Wu, R.; Cho, K. R.; Bose, S.; Wang, S. I.; Li, J.; Parsons, R.; Ellenson, L. H. *Cancer Res.* **1997**, *57*, 3935-3940.

[17] Obata K.; Morland, S. J.; Watson, R. H.; Hitchcock, A.; Chenevix-Trench, G..; Thomas, E. J.; Campbell, I. G. *Cancer Res*.**1998**, *58*, 2095-2097.

[18] Oda, K.; Stokoe, D.; Taketani, Y.; McCormick F. *Cancer Research* **2005**, *65*, 10669-10703.

[19] Wu, R.; Hendrix-Lucas, N.; Kuick, R.; Zhai, Y.; Schwartz, D. R.; Akyol, A.; Hanash, S.; Misek, D. E.; Katabuchi, H.; Williams, B. O.; Fearon, E. R.; Cho, K. R. *Cancer Cell* **2007**, *11*, 321-333.

[20] Ahmed, N.; Oliva, K. T.; Barker, G..; Hoffmann, P.; Reeve, S.; Smith, I. A.; Quinn, M. A.; Rice, G. E. *Proteomics* **2005**, *5*, 4625-4636.

[21] Jones, M.B.; Krutzsh H.; Shu, H.; Zhao, Y.; Liotta, L. A.; Kohn, E. C.,; Petricoin III, E. *Proteomics* **2002**, *2*, 76-84.

[22] Petricoin III, E.; Ardekani, A. M.; Hitt, B. A.; Levine, P. J.; Fusaro, P. J.; Steinberg , S.; Mills, G..; Simone, C.; Fishman, D.; Kohn, E. *The Lancet* **1999**, *359*, 572-577.

[23] Kachman, M. T.; Wang H.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Anal. Chem.* **2002**, *74*, 1779-1791.

[24] Wang, J.; Cai, Y.; Xu, H.; Zhao, J.; Xu, X.; Han, Y. L.; Xu, Z. X.; Chen, B. S.; Hu, H.; Wu, M.; Wang, M. R. *Cell Research* **2004**, *14*, 46-53.

[25] Hamler, R. L.; Zhu, K.; Buchannan, N. S.; Kreunin, P.; Kachman, M. T.; Miller, F. R.; Lubman, D. M. *Proteomics* **2004**, *4,* 562-577.

[26] Zhu, Y.; Wu, R., Sangha, N.; Yoo, C.; Cho, K. R.; Shedden, K. A.; Katabuchi, H.; Lubman, D. M. *Proteomics* **2006**, *6,* 5846-5856.

[27] Kononen, J.; Bubendorf, L.; Kallioniemi, A.; Barlund, M.; Schraml, P.; Leighton, S.; Torhorst, J.; Mihatsch, M. J.; Sauter, G..; Kallionimeni, O.-P. *Nature Med.* **1998**, *4,* 844-847.

[28] Fels, L. M.; Buschmann, T.; Meuer, J.; Reymond, M. A.; Lamer, S.; Christoph Röckenc, C.; Matthias P. A.; Ebertd, M. P. A. *Digestive Disease* **2003**, *21*, 292-298.

[29] Ott, H. W.; Lindner H.; Sarg, B.; Mueller-Holzner, E.; Abendstein, B.; Bergant, A.; Fessler, S.; Schwaerzler, P.; Zeimet, A.; Marth, C.; Illmensee, K. *Cancer Research* **2003**, *63*, 7507-7514.

[30] Chen, W.; Arroyo, J. D.; Timmons, J. C.; Possemato, R.; Hahn, W., C. *Cancer Research* **2005**, *65*, 8183-8192.

[31] Polakis, P. *Gene and Development*, **2000**, *14*, 1837-1851.

[32] Radcliffe, M. J.; Itoh, K.; Sokol, S. Y. *J. Biol. Chem.* **2000**, *275*, 35680-35683.

[33] Tisdale, E. J.; Bourne, J. R; Khosravi-Rar, R.; Der. C. J.; Balch, W. E. *J. Cell. Biol*. **1992**, *119*, 749-761.

[34] Wang, Z.; Lo, H. S.; Yang. H.; Gere, S.; Hu, Y.; Buetow, K. H.; Lee, M. P. *Cancer Research* **2003**, *63*, 655-657.

[35] Yang, J.; Kim O.; Wu, J.; Qiu, Y. *J. Biol. Chem*. **2002**, *277*, 30219-30226.

[36] Jahde, E.; Volk, T.; Atema, A.; Smets, L. A.; Glusenkamp, K. H.; Rajewsky, M. F. *Cancer Research* **1992**, *52*, 6209-6215.

[37] Terkeltaub R. A. *Am. J. Physiol. Cell Physiol.* **2001**, *281*, C1-C11.

[38] Chen, G.; Gharib, T. G..; Huang, C.; Thomas, D. G..; Shedden K. A.; Taylor, J. M. G..; Kardia, S. L. R.; Misek, D. E.; Giordano, T. J.; Iannettoni, M. D.; Orringer, M. B.; Hanash, S. M.; Beer, D. G.. *Clin. Cancer Res*. **2002**, *8*, 2298-2305.

[39] Dai, Z.; Zhu, W. G..; Morrison, C.D.; Brena, R. M.; Smiraglia, D. J.; Raval, A.; Wu, Y. Z.; Rush, L. J.; Ross, P.; Molina, J. R.; Otterson, G. A.; Plass, C. *Exp. Cell Res*. **2004**, *298*, 535-548.

[40] Liu, S. S.; Tsang, B. K.; Cheung, A. N.; Xue, W. C.; Cheng, D. K.; Nga, T. Y.; Wonga, L. C.; Ngan, H. Y. S. *Eur. J. Cancer* **2001**, *37*, 1104-1110.

[41] Clem, R. J.; Sheu, T. T.; Richter, B. W.; He, W. W.; Thornberry, N. A.; Duckett, C. S.; Hardwick, J. M. *J. Biol. Chem*. **2001,** *276*, 7602-7608.

[42] Dong, Z.; Venkatachalam, M. A.; Wang, J.; Patel, Y.; Saikumar, P.; Semenza, G. L.; Force, T.; Nishiyama, J. *J. Biol. Chem.* **2001**, *276*, 18702-18709.

[43] Boutrous, R.; Bryne, J. A. *Exp. Cell Res*. **2005**, *310*, 152-165.

[44] Waldburg, N.; Kähne, T.; Reisenauer, A.; Rocken, C.; Welte, T.; Bühling, F. *Pathol. Res. Pract*. **2004**, *200*, 147-154.

[45] Schwartz, D. R.; Kardia, S. L.; Shedden, K. A.; Kuick, R.; Michailidis, G..; Taylor, J. M. G..; Misek, D. E.; Wu, R.; Zhai, Y.; Darrah, D. M.; Reed, H.; Ellenson, L. H.; Giordano, T. J.; Fearon, E. R.; Hanash, S. M.; Cho, K. R. *Cancer Research* **2002**, *62*, 4722-4729.

Chapter 3

Micro - Proteome Analysis Using Micro-Chromatofocusing in Intact Protein Separations

3.1. Introduction

An important issue in current proteomic analysis is the ability to work with small amounts of samples. This becomes particularly important in situations where only a limited amount of sample may be available such as tissue or fluid samples extracted from aspirates [1], laser capture micro-dissection [2], tumor micro-environment experiments [3] or stem cell research [4] . In these cases, there may be <100,000 cells available or fluids in the amount of 50μl or less. This number may correspond to only several micrograms of total protein where if there are a thousand proteins present, then on average there may only be tens of femtomoles of each protein available. It thus becomes critical to find methods capable of separating and analyzing such small amounts of sample with the ability to identify the presence of large numbers of proteins.

A number of methods have been developed to fractionate the protein content of a cell on a micro-proteomic scale [5, 6, 7, 8]. These methods may use an approach based upon fractionation of intact proteins [6,9] or alternatively fractionation of the protein content of a cell following total digestion of the proteins into peptides [5,10,11,12]. Both of the approaches have been used in previous work for analysis of small amounts of

sample. An advantage of the intact protein method is that it may allow isolation of the original protein for further analysis by protein molecular weight measurement or antibody analysis. The presence of the intact protein information may also ultimately provide information on posttranslational modifications [13].

Ion exchange chromatography (IEC) is a widely used process that separates inorganic ions and proteins based on charge-charge interaction [14, 15, 16]. Protein separation by IEC can be performed using two methods, the salt-gradient method, and the pH- gradient method. Compared with the salt-gradient method, the pH-gradient method has the ability to produce protein separation with a high degree of fractionation based upon an on-column pH-based separation. Despite the capabilities of the pH-gradient method, IEC has been limited because of the difficulty producing linear pH gradients. Several improvements have been made in this area where, for example, Sluyterman and co-workers [17, 18] generated linear-pH gradients with no buffer mixture externally in a process called chromatofocusing (CF). During CF, a quasi linear-pH gradient was produced using simple buffer solutions where proteins elute down the column creating focused bands. Most CF has been conducted using a weak anion exchanger. In other work, Frey and coworkers [19, 20], have developed an advanced theory for CF using simple buffer solutions. Also, Shan and Anderson [21, 22] have generated pH-gradient IEC with amine buffer solutions.

CF has been exploited in proteomics and used as the first dimension in multi-dimensional separation because of its ability to obtain pI information, a fundamental property of proteins. This information is important for identification of proteins and ultimately, posttranslational modifications. For example, Lubman and coworkers [23, 24,

25] studied various cancers, e.g. breast, ovary, and pancreatic using chromatofocusing, for separation followed by mass spectrometry, and showed a difference of protein expression and post-translational modifications with a mass map based on protein pI and molecular weight of proteins. In this work, they used an analytical scale column requiring several hundred micrograms to several milligrams for separation. Alternatively, a method has been developed using a micro-chromatofocusing (micro-CF) procedure with a weak anion exchanger packed in a capillary column. This method is a scaled down version of the CF separation currently used in the Beckman PF2D instrument. However, instead of milligram amounts of starting material 2-30μg of material can be fractionated based upon pH. The intact protein can be collected according to pH value and further analyzed by a second dimension. In this work, we demonstrate digestion of the eluted fraction of intact proteins from micro-CF followed by nano-RPLC-MS/MS. This micro-proteomic procedure is demonstrated for analysis of 700-800 proteins from only 10 μg of two ovarian cell lines (MDAH 2774 and TOV 112D).

3.2. Experimental Section

3.2.1. Chemicals and Materials.

Urea, Thiourea, bis-tris, acetonitrile, n-Octyl-D-glucopyranoside, DTT and formic acid were purchased from Sigma-Aldrich. PMSF was purchased from Bio-Rad. Polybuffer 74 was obtained from Amersham Biosciences (Pittsburgh, PA). The packing material (AX-300) for micro - CF column was obtained from Eprogen, Inc. (Darien, IL). The fused-

silica capillary with 200um i.d./360um o.d. was purchased from Polymicro Technologies, LLC. (Phoenix, AZ).

3.2.2. Micro-Chromatofocusing.

The column used in this experiment was a fused silica capillary column (200μm x 200mm), packed with anion exchange packing materials (AX-300) (Eprogen, Inc., Darien, IL), using a packing bomb built in house at a pressure of 2000psi for 2-3h. The experiments were performed using an ultra-pure II capillary pump with high-pressure mixer (Micro-Tech Scientific, Vista, CA). Two buffer solutions, a starting buffer (SB) and an elution buffer (EB), were used to generate a pH gradient on the CF column. The SB solution was composed of 6M urea and 25mM Bis-tris. The upper limit pH of the SB solution was set at pH 7.8. The EB solution was prepared with 6M urea and polybuffer74, while the lower limit of the EB was set at pH 3.8-4.0. The pH for both buffer solutions was established by addition of a saturated solution of iminodiacetic acid. The CF column was pre-equilibrated with the SB solution. 5-20μg of sample was injected into the micro-CF column and the mobile phase was then switched to the EB solution to initiate the pH gradient. The pH gradient in the column was monitored by an on-line post column detector flow-through pH electrode (Microelectrodes, Inc., Bedford, NH). The flow rate was 5ul/min, with fractions collected in 0.2 pH units. The pH fraction above 7.5 and the elution by salt wash solution (1M NaCl) were collected separately for further analysis. Each sample of the MDAH and TOV cell lines was run by micro-CF three times for reproducibility.

3.2. 3. Analytical Chromatofocusing

Chromatofocusing (CF) was performed on a Beckman System Gold model 127s binary HPLC pump with high-pressure mixer (Fullerton, CA, USA), using a HPCF 1D column (2.1 x 250 mm) (Eprogen, Inc., Darien, IL). Two buffer solutions (a start buffer (SB) and an elution buffer (EB)) were utilized for the generation of a pH gradient on the CF column. The SB solution was composed of 6M urea and 25mM Bis-Tris (pH 7.8). The EB solution was composed of 6M urea and 10% polybuffer74 (pH 3.8-4.0). Both buffer solutions were brought to pH by addition of a saturated solution of iminodiacetic acid. The CF column was pre-equilibrated with SB, after which the mobile phase was switched to the EB solution for pH gradient initiation. The pH gradient was monitored on-line with a post detector pH flow cell (Lazar Research Laboratories, Inc., Los Angeles, CA). The UV absorbance of the eluent was monitored on-line at 280nm. The flow rate was 0.2ml/min, with fractions collected in 0.2 pH units. Each fraction was stored at -80°C.

3.2.4 Trypsin Digestion

The NPS RP-HPLC sample fractions were lyophilized by vacuum centrifugation to remove remaining solvent in the sample tubes. The remaining TFA was neutralized by adding 10% (v/v) NH_4HCO_3 into the samples. 10% (v/v) DTT and 0.01μg of TPCK-treated trypsin (Promega, Madison, WI) were added. The samples were vortexed and then incubated at 37°C for approximately 18h. The tryptic digestion was terminated by addition of 2.5 % (v/v) of TFA.

3.2.5 Protein Identification and quantification

The digested peptide mixture was analyzed by nano-flow reverse-phase LC/MS/MS using the LTQ mass spectrometer with a nano-spray ESI ion source (Thermo, San Jose, CA). The samples were separated using a (0.1 x 150mm) capillary reverse phase column (MichromBioresources, Auburn, CA) with a flow rate of 5ul/min. An acetonitrile:water gradient method was used, starting with 5% acetonitrile which was ramped to 60% in 25min and to 90% in another 5min. Both solvent A (water) and B(ACN) contained 0.1% formic acid. The electrospray voltage was 2.6V, with a capillary temperature of 200°C and a capillary voltage of 4kV. The normalized collision energy was set at 35% for MS/MS. The MS/MS spectra obtained were analyzed using the Sequest feature of Bioworks 3.1 SR1. Peptide ions were assigned with the Xcorr values to consider >3.5 for +3, >2.5 for +2, and >1.7 for +1 ions and a ΔCn of 0.1 or higher. To further validate data obtained from Sequest, Protein prophet/peptide prophet software [26] modified in house were used to provide a confidence level in identification of 95%. In order to quantify proteins, DTA select program has been used as described previously [27]. The DTA program is a form of spectral counting where the program reads individual spectra from the Sequest output files and evaluates the result using a filter set by the user. This then is followed by generating a dtaselect.html file and text file containing information on identified proteins and the corresponding spectral counts.

3.2.6 Description of MDAH and TOV

Two ovarian endometrioid cancer cell lines (MDAH 2774, TOV 112D) were obtained from the American Type Culture Collection. The cell lines were cultured as described previously [28]. Cultured cells were removed by scrapping with a cell scraper (Costar,

Cambridge, MA) and washed three times with PBS solution. Then, cells were lysed with lysis buffer solution composed of 6M urea (Sigma), 2M thiourea (Sigma), 1% n-Octyl-D-glucopyranoside(Sigma), 2mM DTT(Sigma), and 2.5mM PMSF(Bio-Rad). The lysed solution containing cells were centrifuged at 35,000rpm at 4°C for 1h and the supernatant was stored at -80°C for further use.

3.3. Results and Discussion

Micro CF column

Chromatofocusing is an on-column, non-gel fractionation technique that has been used as the first dimension in multi-dimensional separations based on the pI values of proteins in complex mixtures. The dimension of the column used for chromatofocusing (CF) on an analytical scale in our previous work [25] has been 2.1mm (i.d.) x 250mm, while the dimension of the column for micro-CF is 200μm (i.d.) x 200mm. These two columns, which differ in volume by a factor of 100, result in several experimental differences since columns size affects both the amount of protein loaded onto the column and the composition of the elution buffer (EB) solution. The amount of protein loaded onto the column for multidimensional separation is an important factor in proteomic research where it will determine the number of proteins detected, especially those at lower abundance. The amount of protein required for analytical scale CF is ~ 5mg, whereas in micro-CF the total protein loading can be reduced to typically 5-20μg of sample. Indeed, the protein loading in these micro-CF experiments can be as little as 2μg of total protein.

Another difference between CF and micro-CF is the time required to generate a pH gradient. In the analytical scale CF used in prior work, the EB solution contained a 10% polybuffer solution, which generates a 40-70 min pH gradient. However, in micro-CF separation, a 10% polybuffer solution generates a pH gradient separation for only a limited time before reaching a lower pH. This rapid pH gradient was insufficient to generate adequate protein separation based on pI or to collect sufficient volume of the fractionated samples for further experiments. To generate a longer elution time, a 2% polybuffer solution was used in micro-CF for the pH gradient. Highly repeatable pH gradients (pH 7.6-3.8) were obtained for micro-CF using the EB solution as shown for 10μg of sample injected on-column in Figure 3.1. Compared to the gradient change of analytical CF involving a 10% polybuffer solution, the 2% polybuffer system yielded a pH gradient that was twice as long in our experiments.

Separation and identification of proteins using nano ESI-LC-MS

Nano-ESI-LC-MS

The proteins separated by micro-CF were collected into pH fractions and digested into peptides using trypsin. The peptides from each fraction were further analyzed using nano-flow HPLC-ESI-MS/MS. The proteins from the pH fractions and the salt gradient were analyzed in these experiments. Figure 3.2 shows the base peak chromatogram for three pH fractions of a microCF/nano-HPLC-ESI MS/MS of the peptides detected in a human ovarian cancer cell line (TOV112D). Though there are some overlapping peptide peaks from high abundance proteins across adjacent pH ranges, low abundance proteins such as p53 were still identified based upon the reverse phase separation and MS/MS

analysis. The reproducibility of the peptide separation and identification was tested using the combined microCF/nano-HPLC-ESI MS/MS in multiple runs of tryptic digest samples. Two consecutive experiments were performed for two different analyses from repeat micro-CF experiments with the same pH range (pH 5.6-5.4). There is only a small difference in elution times of the peptides from the fractionated solution, usually <0.2min. Figure 3.3 illustrates the base peak chromatogram and tandem mass spectrometry of the sample solution. Mass spectra were generated for the peptide K. MEKETAENYLGHTAK.N from GRP75 (Stress-70 protein, mitochondrial precursor) at 19.47min from Figure 3.3B and 19.33min in Figure 3.3D, respectively. Figure 3.3 demonstrates the reproducibility of the combined micro-CF/nano-HPLC-ESI-MS/MS method.

Comparison of ovarian cancer cell line proteins

In the integrated micro-CF/nano HPLC-ESI-MS/MS approach, 700-800 proteins were detected using 10μg of sample in each of two cell lines. Each of these samples was run three times to increase the number of total proteins detected. However, after the second run the number of proteins detected did not increase substantially. The pI distribution and location of the identified proteins is illustrated in Figure 3.4 and 3.5 respectively. The pI distribution of proteins using micro-CF is similar to that obtained from analytical scale CF. As can be seen, most of the proteins are located in the Cytosol and Nuclear compartments. There was little difference in the location of the proteins in the two ovarian endometrioid cell lines.

Table 3.1 lists 173 proteins common to both the MDAH 2774 and TOV 112D cell lines. This Table also contains quantitative information on protein expression based upon the spectral counting method. For example, structural proteins such as Keratin 9, 19, 22 and the Histones AA, BA are found with a high Xcorr score from the ESI-MS/MS Sequest database search result. It should be noted that there are large differences in expression between the histones and keratins between the cell lines, where the presence of the different keratins is often used as a marker of the type of cancer [29, 30]. Of the common proteins in both cell lines Table 3.1 shows that less than half of the proteins undergo pI shifts. Also, approximately 25% of the proteins in Table 3.1 are shifted to acidic pH compared to the database value. According to the nano HPLC-ESI/MS/MS result, half of the proteins with a pI shift toward the acidic range showed possible phosphorylations, while the remainder showed possible methionine oxidation or no PTM present. Many of these proteins have been studied in previous work using analytical scale chromatofocusing in breast cancer cell lines and showed distinctive phosphorylations in aggressive cancers [13]. Some of these proteins include the four histones detected, which have a strong shift towards lower pI compared to that expected. In addition, other proteins such as elongation factor1-alpha1 and arginine splice factor have been studied by these methods and shown to be multiply-phosphorylated in cancer cells. Interestingly, proteins with a pI shift toward the acidic range appear to be involved in protein binding at the molecular function level such as ATP binding. This may indicate that these proteins play a role in signal pathways involved in cancer.

Ovarian cancer cell lines-Pathway proteins detected

Two ovarian endometrioid adenocarcinoma (OEA) cell lines have been studied to gain insight into the progression of disease through potential proteomic studies of cell signaling pathways. Among the signal pathways, the Wnt signaling pathway has been known to play a critical role in cell proliferation, regulating morphology, and cell fate at the cellular level [31] and is known to be deregulated in endometrioid ovarian cancer [32]. Moreover, mutations such as the CTNNB1 gene mutation and the PTEN tumor suppressor gene have also been shown to be involved in this pathway [33]. The CTNBB1 gene encodes beta-catenin which is regulated by a multi-protein complex (Beta-catenin - Axin-adenomatous polyposis coli (APC)-glycogen synthase kinase (GSK-3β)). In this protein complex, GSK-3β phosphorylates NH_2-terminal β-catenin, producing beta-catenin for ubiquitination and proteasome degradation. With the Wnt signal, beta-catenin degradation is inhibited allowing β-catenin to enter the nucleus, activating the TCF/LEF complex including Bcl-9, PYCO, CBP and turning on genes such as c-Myc and cyclin-D.

In this work, two ovarian endometrioid cell lines (MDAH2774, TOV112D) were studied. MDAH2774 and TOV112D are cultured cell lines derived from human ovarian endometrioid adenocarcinomas. TOV112D harbors the β-catenin mutation, Wnt pathway defect and MDAH2774 has an Axin1 mutation, also a Wnt pathway defect. The effect of these mutations is an increase in the cellular level of beta-catenin and subsequent transcription of Wnt target genes. In considering the Wnt pathway, no Wnt antagonists e.g. secreted frizzled-related proteins (SFRPs) or Wnt-inhibit-factor-1 (WIF-1) were observed in either cell line at detectable levels with the capability of our current experiment. However, the Receptor (Frizzled-1[precursor]) and low density lipoprotein

receptor for Wnt ligand were detected in each cell line, which suggests the activation of the cytoplasmic protein disheveled (Dsh, Dvl). The mechanism for Dsh/Dvl activation is not clear yet, but CSK21 (casein kinase II subunit alpha) found in TOV112D has been suggested as being involved in this process [34]. Dishevelled proteins have been shown to recruit GSK-3 binding protein (GBP), detected in both cell lines herein to the multiprotein complex. The resulting complex in both cell lines (TOV112D, MDAH2774) was expected to lead to activation of the TCF/LEF target since the mutations of CTNNB1 (TOV112D) and Axin1(MDAH2774) have been shown in prior work [32]. Bcl-9, and CBP (CREB-binding protein), which were detected in TOV112D, are a part of the β-catenin complex (LEF/TCF- β-catenin-Bcl9-CBP) in the nucleus. This finding may support the higher activation of TCF-dependent transcription in TOV112D obtained compared to MDAH2774. Also p53 was found in MDAH2774, which could provide an alternate pathway for tumor activation in this cell line.

Since both cell lines have different mutations, different cellular behavior may result according to protein expression. Compared with MDAH2774, the TOV112D cell line expresses a higher number of proteins involved with cell activation, for example, T-cell surface antigens CD2, CD4, and the Wiskott-Aldrich syndrome protein family member 3(WASF3). Also detected was p85A (phosphatidylinositol 3-kinase regulatory subunit alpha) which is necessary for the insulin-stimulated increase in glucose uptake and glycogen synthesis in insulin-sensitive tissues. In terms of glucose transport, GTR4 (Solute carrier family 2, facilitated glucose transporter member 4) was found in TOV112D. Raf1 protein was detected which is involved in B-cell receptor signaling and

IL-6 signaling where IL6 (Interleukin-6 [Precursor]) was detected. Several proteins (M3K9, M3K10 and M3K11) involved in the MAPKinase signal pathway were also more often observed in TOV112D than in MDAH2774. Although these cell lines are both endometrioid cancers, there may be different pathways at work that are responsible for their aggressiveness and pathologies.

3.4. Conclusion

It is shown that cell lysates can be analyzed using the combined method, micro-CF/nano-HPLC-ESI-MS/MS, where only limited amounts of sample are available. The micro-CF method involves separation of intact proteins where the pI information can be used for identifying the potential presence of PTMs and also the change in PTM content between cell lines. The method can identify large numbers of proteins (700-800) with only 10-20μg total protein where the proteins observed can be associated with protein pathways. The proteins identified can be associated with known genetic defects in these cancer cell lines and these pathways can potentially be associated with the aggressiveness or the progression of disease.

Table 3.1: Proteins identified from both cell lines, TOV112D, and MDAH2774.

Access No.	Protein Name	Expression change(fold)		Theritical	Experimental pI	
		TOV112D	MDAH2774	pI	TOV 112D	MDAH 2774
P31947	14-3-3 protein sigma		98.4	4.68	pH4.6-4.4	pH4.4-4.2
P62736	Actin, aortic smooth muscle		6.3	5.24	pH5.6-5.4	pH5.2-5.0
P60709	Actin, cytoplasmic 1		3.6	5.29	pH5.6-5.4	pH5.2-5.0
P63261	Actin, cytoplasmic 2 (Gamma-actin)	—	—	5.31	abovepH7.5	abovepH7.5
P12814	Alpha-actinin-1 (Alpha-actinin cytoskeletal isoform)		1.1	5.25	pH5.2-5.0	pH4.2-4.0
O43707	Alpha-actinin-4 (Non-muscle alpha-actinin 4)	—	—	5.27	abovepH7.5	above pH 7.5
Q9H2P0	Activity-dependent neuroprotector	3.1		6.97	pH6.2-6.0	pH6.8-6.6
P51825	AF4/FMR2 family member 1,Proto-oncogene AF4	2.7		9.26	abovepH7.5	abovepH7.5
Q16352	Alpha-internexin	—	—	5.34	abovepH7.5	salt washing
P02768	Serum albumin [Precursor]		8.9	5.67	pH5.6-5.4	pH5.6-5.4
Q01484	Ankyrin-2	—	—	5.03	pH5.0-4.8	pH5.2-5.0
Q86YR6	Ankyrin repeat domain-containing protein 21		2.6	6.32	pH6.6-6.4	pH7.0-6.8
P50995	Annexin A11 (Annexin XI)	—	—	7.53	salt washing	salt washing
P07355	Annexin A2		7.2	7.56	pH5.8-5.6	pH6.0-5.8
P12429	Annexin A3		8.8	5.63	pH5.6-5.4	pH5.6-5.4
Q9NR81	Rho guanine nucleotide exchange factor 3	4.4		6.03	pH4.8-4.6	pH4.4-4.2
Q68CP9	AT-rich interactive domain-containing protein 2	3.2		7.08	pH7.2-7.0	pH7.2-7.0
P06576	ATP synthase subunit beta, mitochondrial precursor		15.5	5	pH5.6-5.4	pH5.6-5.4
O75173	ADAMTS-4 [Precursor]		3.2	8.23	pH4.6-4.4	abovepH7.5
Q8WWM7	Ataxin-2-like protein	2.7		8.7	above pH7.5	abovepH7.5
Q8IY92	BTB/POZ domain-containing protein 12		1.5	5.32	pH6.2-6.0	abovepH7.5
Q05682	Caldesmon (CDM)	1.8		5.63	abovepH7.5	abovepH7.5
P62158	Calmodulin (CaM)	6.7		4.09	pH4.4-4.2	pH4.4-4.2
P83916	Chromobox protein homolog 1	—	—	4.85	above pH7.5	abovepH7.5
P47902	Homeobox protein CDX-1	3.3		9.58	pH6.2-6.0	pH6.0-5.8
P49454	Centromere protein F		1.7	5.03	pH4.8-4.6	pH4.6-4.4
Q86VQ0	Uncharacterized protein C6orf152		1.7	7.31	pH6.6-6.4	pH6.6-6.4
Q9HAC7	Uncharacterized protein C7orf10		2.6	8.54	salt washing	salt washing
Q9P2M7	Cingulin		2.2	5.46	pH7.2-7.0	pH7.2-7.0
P53618	Coatomer subunit beta		4.3	5.72	pH4.8-4.6	pH4.6-4.4
Q92828	WD repeat-containing protein 2		2	8.24	pH5.6-5.4	pH6.0-5.8
Q8IVV8	Protein FAM77A		3.2	5.43	above pH7.5	abovepH7.5
Q14118	Dystrophin-associated glycoprotein 1	2.4		9.31	above pH7.5	above pH7.5
Q9UIK4	Death-associated protein kinase 2	28.5		6.45	pH6.8-6.6	abovepH7.5
P16989	DNA-binding protein A	27		9.77	above pH7.5	pH7.2-7.0
Q13561	Dynactin subunit 2 (Dynactin complex 50 kDa subunit)	11.9		5.1	pH5.2-5.0	pH4.6-4.4
P17661	Desmin	2.6		5.21	pH5.4-5.2	pH5.4-5.2
O00273	DNA fragmentation factor subunit alpha		1.4	4.68	pH4.8-4.6	PH4.8-4.6
O60469	Down syndrome cell adhesion molecule [Precursor]	1.7		7.81	pH5.2-5.0	pH5.2-5.0
Q8TE73	Ciliary dynein heavy chain 5		2.4	5.79	pH4.8-4.6	pH4.8-4.6
Q8WXU2	Dyslexia susceptibility 1 candidate gene 1 protein	1.2		8.95	pH5.2-5.0	pH5.2-5.0
P68104	Elongation factor 1-alpha 1		35.8	9.1	pH7.6-7.4	pH7.6-7.4
P26641	Elongation factor 1-gamma (EF-1-gamma)	1		6.27	pH5.2-5.0	pH4.6-4.4
P13639	Elongation factor 2		3.2	6.42	6.2-6.0	pH6.0-5.8
Q14677	Clathrin interactor 1 (Epsin-4) (Epsin-related protein)	1.7		6.01	pH6.2-6.0	pH6.2-6.0
P15311	Ezrin (p81) (Cytovillin) (Villin-2)		4.7	5.95	pH6.0-5.8	pH6.0-5.8
Q5JRC9	Protein FAM47A		2.6	9.15	above pH7.5	above pH7.5
Q96AY3	FK506-binding protein 10 [Precursor]	3.8		5.26	pH5.4-5.2	pH5.4-5.2

Access No.	Protein Name	Expression change(fold)		Theritical	Experimental pI	
		TOV112D	MDAH2774	pI	TOV 112D	MDAH 2774
Q16595	Frataxin, mitochondrial [Precursor]		6.8	5.94	pH4.8-4.6	pH4.6-4.4
Q5H8C1	FRAS1-related extracellular matrix protein 1 [Precursor]	2.7		5.5	pH4.6-4.4	salt washing
P32455	Interferon-induced guanylate-binding protein 1		1.2	5.97	pH7.6-7.4	pH7.2-7.0
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2	1.3		5.08	above pH7.5	abovepH7.5
Q02153	Guanylate cyclase soluble subunit beta-1 (GCS-beta-1)	3.7		5.2	salt washing	pH5.4-5.2
Q96N19	Integral membrane protein GPR137		2.5	8.78	salt washing	pH4.8-4.6
P38646	Stress-70 protein, mitochondrial [Precursor]		150.7	5.44	pH5.6-5.4	pH5.6-5.4
P11021	78 kDa glucose-regulated protein precursor		2.7	5.01	pH5.2-5.0	pH5.2-5.0
P28001	Histone H2A type 1-E		244	11.05	pH6.8-6.6	pH7.6-7.4
P0C0S8	Histone H2A type 1		531.9	10.9	above 7.5	pH7.2-7.0
P62807	Histone H2B type 1-C/E/F/G/I		100	10.32	pH7.6-7.4	pH7.6-7.4
P62805	Histone H4		46.2	11.36	pH7.6-7.4	pH7.6-7.4
P82979	Nuclear protein Hcc-1	2.6		6.12	pH4.6-4.4	pH4.6-4.4
Q16836	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	1.1		8.38	pH7.4-7.2	pH7.0-6.8
Q8TBE9	N-acylneuraminate-9-phosphatase(NANP_human)		9.7	6.01	pH4.6-4.4	pH6.2-6.0
P31943	Heterogeneous nuclear ribonucleoprotein H (hnRNP H)		58.6	5.89	pH5.8-5.6	pH5.6-5.4
P52597	Heterogeneous nuclear ribonucleoprotein F (hnRNP F)	2		5.38	pH5.6-5.4	pH5.6-5.4
P61978	Heterogeneous nuclear ribonucleoprotein K (hnRNP K)		1.8	5.39	pH5.2-5.0	pH4.8-4.6
O43390	Heterogeneous nuclear ribonucleoprotein R (hnRNP R)	1.1		8.23	salt washing	salt washing
P34931	Heat shock 70 kDa protein 1L		15.3	5.76	pH5.8-5.6	pH5.6-5.4
P07900	Heat shock protein HSP 90-alpha		13.8	4.94	pH6.6-6.4	pH4.4-4.2
P08107	Heat shock 70 kDa protein 1		7.7	5.48	pH5.4-5.2	pH5.2-5.0
P54652	Heat shock-related 70 kDa protein 2		6	5.56	pH5.6-5.4	pH5.6-5.4
P34932	Heat shock 70 kDa protein 4	1.2		5.18	pH5.2-5.0	pH4.8-4.6
P11142	Heat shock cognate 71 kDa protein		7.8	5.37	pH5.6-5.4	pH5.6-5.4
Q9Y3Y2	Uncharacterized protein C1orf77		1.6	12.24	above pH7.5	7.6-7.4
Q7Z6Z7	E3 ubiquitin-protein ligase HUWE1		2.2	5.1	above pH7.5	above pH7.5
P20810	Calpastatin	4.5		4.98	above pH7.5	above pH7.5
P63241	Eukaryotic translation initiation factor 5A-1	1.4		5.08	pH5.0-4.8	pH5.0-4.8
P01563	Interferon alpha-2 [Precursor]	1.3		5.99	pH6.0-5.8	pH6.0-5.8
Q12906	Interleukin enhancer-binding factor 3	1.3		8.86	salt washing	pH5.2-5.0
Q14643	Inositol 1,4,5-trisphosphate receptor type 1		2.3	5.71	pH5.8-5.6	pH5.8-5.6
P29375	Histone demethylase JARID1A	1.1		6.42	above pH7.5	pH7.0-6.8
O94953	JmjC domain-containing histone demethylation protein 3B	34.1		6.72	pH6.8-6.6	above pH7.5
Q9HCJ3	Ribonucleoprotein PTB-binding 2	1.4		7.12	pH7.0-6.8	above pH7.5
P13645	Keratin, type I cytoskeletal 10	34		5.13	pH5.8-5.6	pH6.0-5.8
P13646	Keratin, type I cytoskeletal 13	3.6		4.91	pH7.6-7.4	pH7.6-7.4
P08779	Keratin, type I cytoskeletal 16	3.7		4.98	pH7.6-7.4	pH5.2-5.0
P12035	Keratin, type II cytoskeletal 3	2.1		8.16	pH5.6-5.4	pH5.6-5.4
P05787	Keratin, type II cytoskeletal 8		30	5.52	pH5.6-5.4	pH5.6-5.4
P00568	Adenylate kinase isoenzyme 1	11.1		8.73	above 7.5	pH7.2-7.0
O95069	Potassium channel subfamily K member 2	1.4		8.46	above pH7.5	pH7.2-7.0
Q9NQT8	Kinesin-like protein KIF13B		3.4	5.56	pH7.2-7.0	above pH7.5
P46013	Antigen KI-67	1.1		9.49	above pH7.5	pH6.6-6.4
O00139	Kinesin-like protein KIF2A	6.8		6.04	pH6.2-6.0	above pH7.5
Q9BVG8	Kinesin-like protein KIFC3	4.3		7.62	pH7.4-7.2	pH7.2-7.0
P05455	Lupus La protein		2.3	6.68	pH6.0-5.8	pH5.6-5.4
P20700	Lamin-B1		5.4	5.11	pH5.2-5.0	pH5.0-4.8

Access No.	Protein Name	Expression change(fold)		Theritical	Experimental pI	
		TOV112D	MDAH2774	pI	TOV 112D	MDAH 2774
P02545	Lamin-A/C		16.3	6.57	pH6.2-6.0	pH5.6-5.4
Q14847	LIM and SH3 domain protein 1		1.6	6.61	pH6.8-6.6	pH4.6-4.4
Q6JVE6	Epididymal-specific lipocalin-10 [Precursor]	1.5		10.25	above pH7.5	above pH7.5
P07195	L-lactate dehydrogenase B chain		4.3	5.72	pH6.2-6.0	pH5.6-5.4
Q7L1W4	Leucine-rich repeat-containing protein 8D		1.5	7.76	7.6-7.4	pH5.2-5.0
Q9NX58	Cell growth-regulating nucleolar protein	1.3		9.57	pH4.8-4.6	pH6.0-5.8
Q9UPN3	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5	1.6		5.27	pH7.0-6.8	above pH7.5
P27816	Microtubule-associated protein 4	1		5.32	pH5.0-4.8	pH4.6-4.4
P43243	Matrin-3		2.8	5.87	pH4.8-4.6	pH4.8-4.6
P41594	Metabotropic glutamate receptor 5 [Precursor]		2.9	8.09	pH5.8-5.6	pH5.6-5.4
P19105	Myosin regulatory light chain 2, nonsarcomeric		2.3	4.67	pH4.8-4.6	pH4.8-4.6
P49006	MARCKS-related protein		4	4.68	pH4.6-4.4	pH4.6-4.4
P60660	Myosin light polypeptide 6		22.2	4.56	above pH7.5	pH4.8-4.6
Q15746	Myosin light chain kinase, smooth muscle	—	—	5.85	pH6.20-6.0	above pH7.5
O95251	Histone acetyltransferase MYST2		1.4	9.01	pH4.8-4.6	pH7.0-6.8
P15531	Nucleoside diphosphate kinase A	3		5.83	pH5.6-5.4	pH5.2-5.0
P20929	Nebulin	1.4		9.1	salt washing	pH5.2-5.0
P82970	Nucleosome-binding protein 1		1.8	4.5	pH4.4-4.2	pH4.4-4.2
Q9UNZ2	NSFL1 cofactor p47		1.2	4.99	pH4.6-4.4	pH4.6-4.4
Q02818	Nucleobindin-1 [Precursor]	6		5.09	pH5.2-5.0	pH5.2-5.0
Q96KK4	Olfactory receptor 10C1	—	—	8.64	above pH7.5	above pH7.5
Q9Y4L1	Hypoxia up-regulated protein 1 [Precursor]	1.2		5.07	pH5.0-4.8	pH4.8-4.6
Q9Y6V0	Protein piccolo		1.4	6.12	pH6.2-6.0	pH6.2-6.0
O95831	Apoptosis-inducing factor 1, mitochondrial [Precursor]	5.5		6.86	pH6.2-6.0	pH5.2-5.0
P30101	Protein disulfide-isomerase A3 [Precursor]	6.9		5.61	pH5.6-5.4	pH5.6-5.4
P41219	Peripherin		1.6	5.37	above pH7.5	above pH7.5
O43933	Peroxisome biogenesis factor 1	—	—	5.91	pH4.6-4.4	pH7.2-7.0
Q9NQP4	Prefoldin subunit 4	3		4.42	pH4.4-4.2	pH4.4-4.2
P78562	Phosphate-regulating neutral endopeptidase	1.8		8.91	salt washing	pH5.2-5.0
Q9ULU4	Protein kinase C-binding protein 1	1.8		6.83	7.4-7.2	pH7.2-7.0
Q96MT3	Prickle-like protein 1	4.1		5.84	pH5.2-5.0	pH5.2-5.0
P17980	26S protease regulatory subunit 6A		13.3	5.13	pH5.2-5.0	pH7.6-7.4
Q15678	Tyrosine-protein phosphatase non-receptor type 14	—	—	8.45	above pH7.5	above pH7.5
P11498	Pyruvate carboxylase, mitochondrial [Precursor]	1.1		6.14	pH6.8-6.6	pH7.0-6.8
Q7Z5J4	Retinoic acid-induced protein 1	16.5		9.03	pH5.6-5.4	pH7.2-7.0
P43487	Ran-specific GTPase-activating protein		4.6	5.19	pH5.2-5.0	pH5.0-4.8
O76081	Regulator of G-protein signaling 20	5.4		6.48	pH6.6-6.4	pH6.6-6.4
Q99729	Heterogeneous nuclear ribonucleoprotein A/B	1.5		9.04	above pH7.5	ph4.8-4.6
P62263	40S ribosomal protein S14	1.3		10.08	pH6.2-6.0	pH7.6-7.4
Q92545	Transmembrane protein 131 [Fragment]	1.8		no info	pH7.2-7.0	abovepH7.5
Q15413	Ryanodine receptor 3	—	—	5.45	pH7.2-7.0	abovepH7.5
Q92562	SAC domain-containing protein 3		1.5	6.46	pH7.2-7.0	above pH7.5
Q9Y5Y9	Sodium channel protein type 10 subunit alpha	2.8		5.67	pH6.4-6.2	pH6.4-6.2
Q9H190	Syntenin-2	17.8		9.15	pH7.6-7.4	pH4.0-3.8
O43175	D-3-phosphoglycerate dehydrogenase	159.5		6.31	pH6.6-6.4	pH5.2-5.0
O15047	Histone-lysine N-methyltransferase	1.5		5.07	pH5.2-5.0	pH5.2-5.0
P23246	Splicing factor, proline- and glutamine-rich		1.3	9.45	pH6.4-6.2	pH6.2-6.0
P84103	Splicing factor, arginine/serine-rich 3	1.1		11.64	above pH7.5	pH5.2-5.0

Access No.	Protein Name	Expression change(fold)		Theritical	Experimental pI	
		TOV112D	MDAH2774	pI	TOV 112D	MDAH 2774
Q13813	Spectrin alpha chain, brain	7		5.22	pH5.0-4.8	pH5.4-5.2
Q01082	Spectrin beta chain, brain 1	4		5.39	pH5.0-4.8	pH6.2-6.0
Q13748	Tubulin alpha-2 chain		2	4.98	pH5.2-5.0	pH5.0-4.8
Q71U36	Tubulin alpha-1A chain		2.7	4.94	pH5.0-4.8	pH5.0-4.8
O75764	Transcription elongation factor A protein 3	—	—	9.32	abovepH7.5	pH5.6-5.4
P48643	T-complex protein 1 subunit epsilon	1.7		5.45	pH5.6-5.4	pH5.2-5.0
P49368	T-complex protein 1 subunit gamma	1.2		6.1	pH4.6-4.4	pH5.6-5.4
Q9UNS1	Timeless homolog	2.3		5.26	pH5.6-5.4	pH5.6-5.4
P11388	DNA topoisomerase 2-alpha	1.4		8.82	above pH7.5	pH7.2-7.0
P55327	Tumor protein D52	1.1		4.94	pH5.0-4.8	pH4.8-4.6
O43399	Tumor protein D54		3	5.26	pH5.0-4.8	pH5.0-4.8
P09493	Tropomyosin-1 alpha chain		1.6	4.69	pH4.4-4.2	pH4.4-4.2
P07951	Tropomyosin beta chain	21.7		4.66	pH4.8-4.6	pH4.8-4.6
P06753	Tropomyosin alpha-3 chain		1.4	4.68	pH4.8-4.6	pH4.8-4.6
Q9Y4A5	Transformation/transcription domain-associated protein		1.4	8.49	above pH7.5	above pH7.5
P26368	Splicing factor U2AF 65 kDa subunit	1.5		9.19	salt washing	pH4.8-4.6
Q00341	Vigilin	3.2		6.43	pH6.6-6.4	pH 6.6-6.4
P08670	Vimentin	1.3		5.06	pH4.8-4.6	pH4.8-4.6
P18206	Vinculin		8.3	5.51	pH5.2-5.0	pH5.6-5.4
Q9P1Q0	Vacuolar protein sorting-associated protein 54		2.7	6.1	pH5.6-5.4	pH5.2-5.0
P04275	Von Willebrand factor [Precursor]		2.6	5.06	pH5.2-5.0	pH5.2-5.0
Q9P2L0	WD repeat protein 35	1.1		5.98	pH5.2-5.0	pH4.2-4.0
Q8NI36	WD repeat protein 36	21.6		7.33	pH7.4-7.2	abovepH7.5
O15213	WD repeat protein 46	23.7		9.69	pH7.6-7.4	abovepH7.5
Q8IZP6	RING finger protein 113B		1.1	7.54	pH5.6-5.2	pH7.6-7.4
Q5TAX3	Zinc finger CCHC domain-containing protein 11		1.2	8.3	pH7.0-6.8	pH7.6-7.4
Q9H4I2	Zinc fingers and homeoboxes protein 3	2.5		5.73	pH7.6-7.4	pH7.0-6.8
Q9UL59	Zinc finger protein 214	3.7		8.89	pH7.6-7.4	pH7.2-7.0
O15015	Zinc finger protein 646	1.1		6.97	pH7.0-6.8	above pH7.5

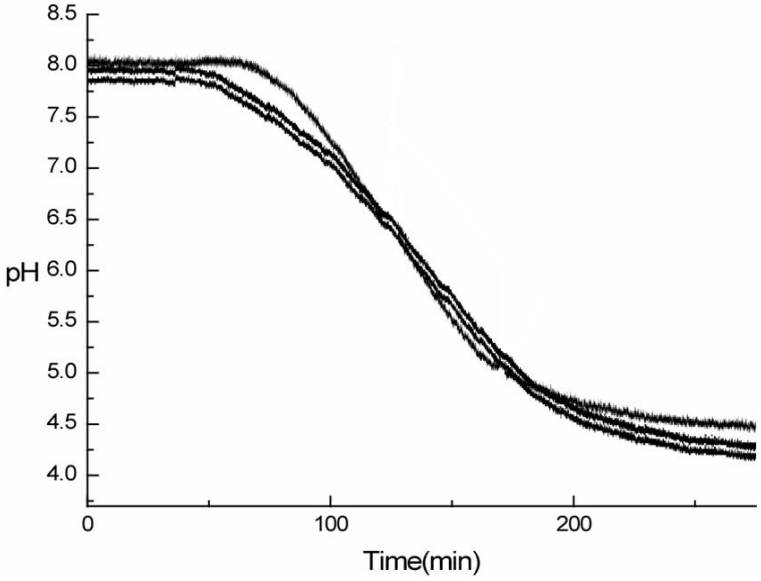


Figure 3.1. pH gradient obtained by micro-CF

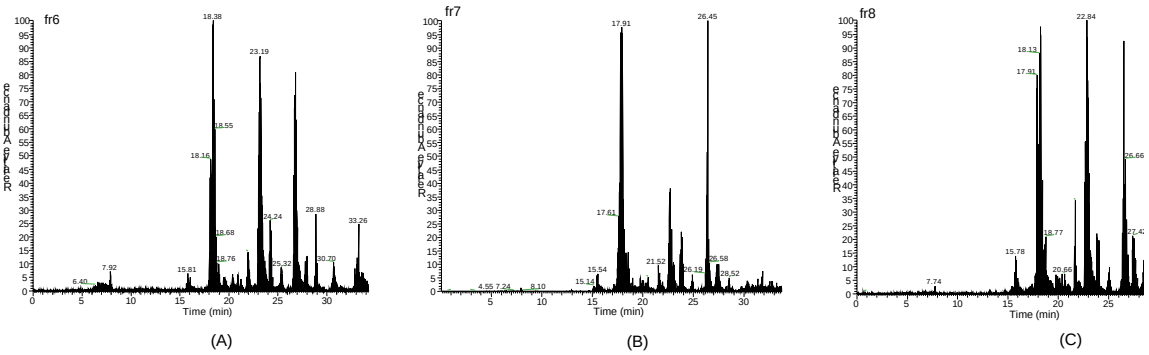


Figure 3.2. Base Peak Chromatogram in TOV112D, (A) fr6, (B) fr7, (C) fr8.

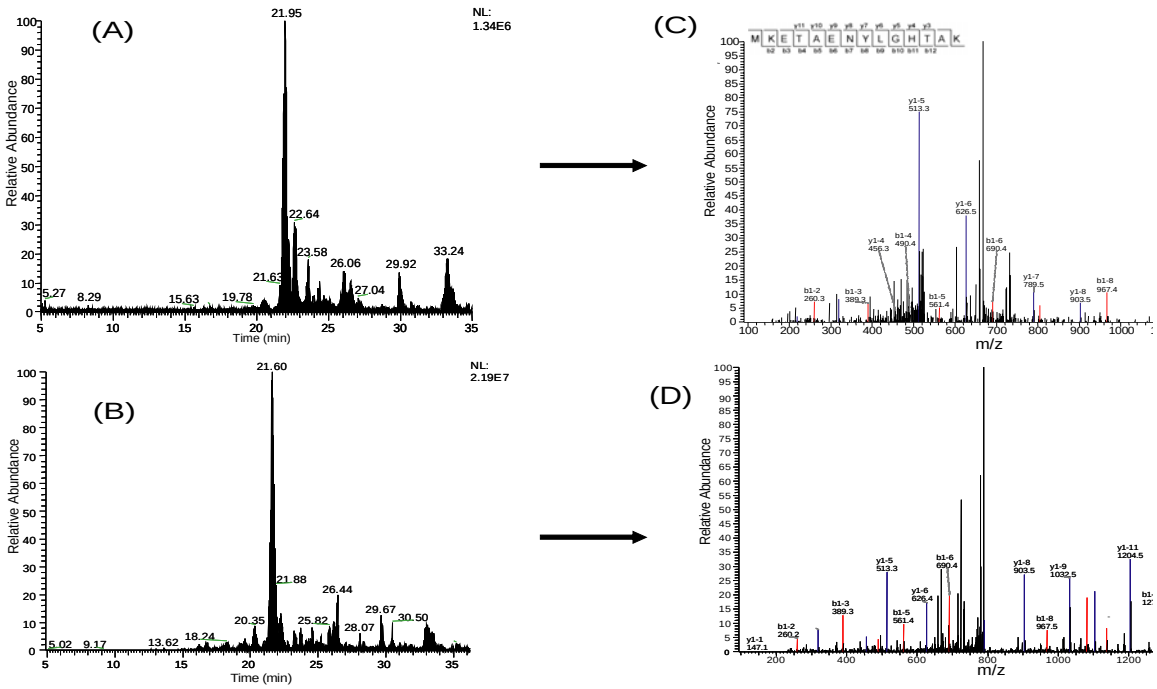


Figure 3.3. Reproducibility analysis of micro-CF/nano-HPLC-ESI-MS/MS separation. Equal amounts of the sample were analyzed. The tandem mass spectrum obtained at 19.43 min in (A) is shown in (C). The Tandem mass spectrum obtained at 19.33 min in (B) is shown in (D).

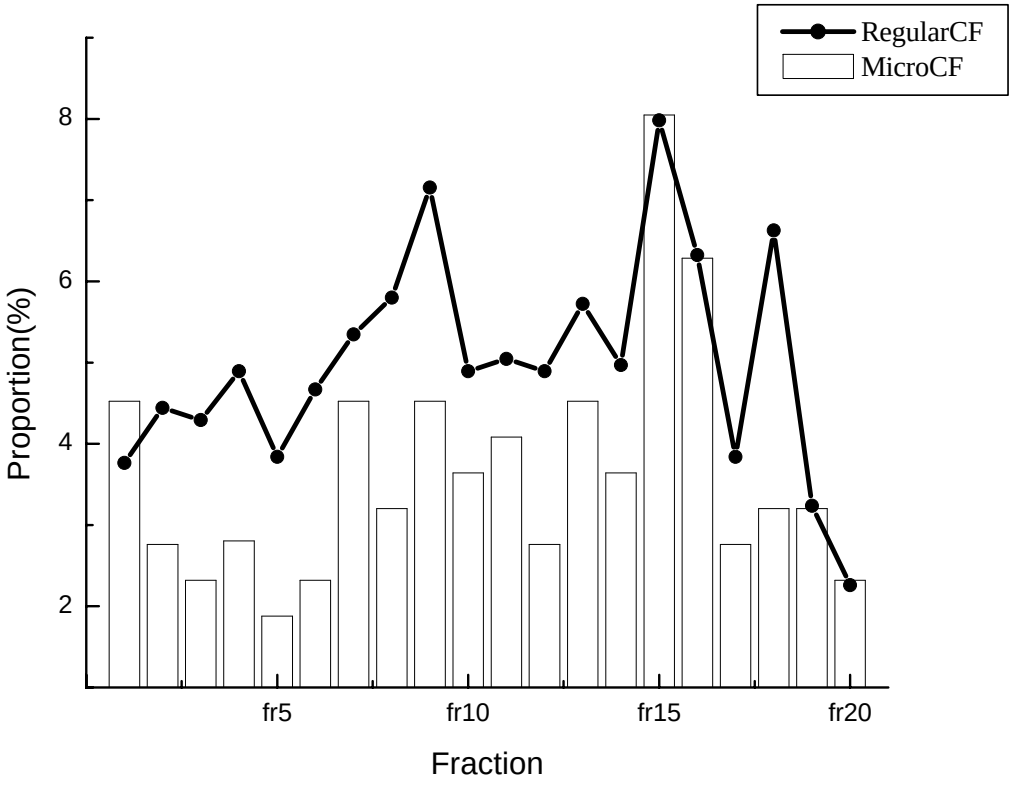


Figure 3.4. pI distribution in both Regular CF and Micro-CF using MDAH2774 cell line.

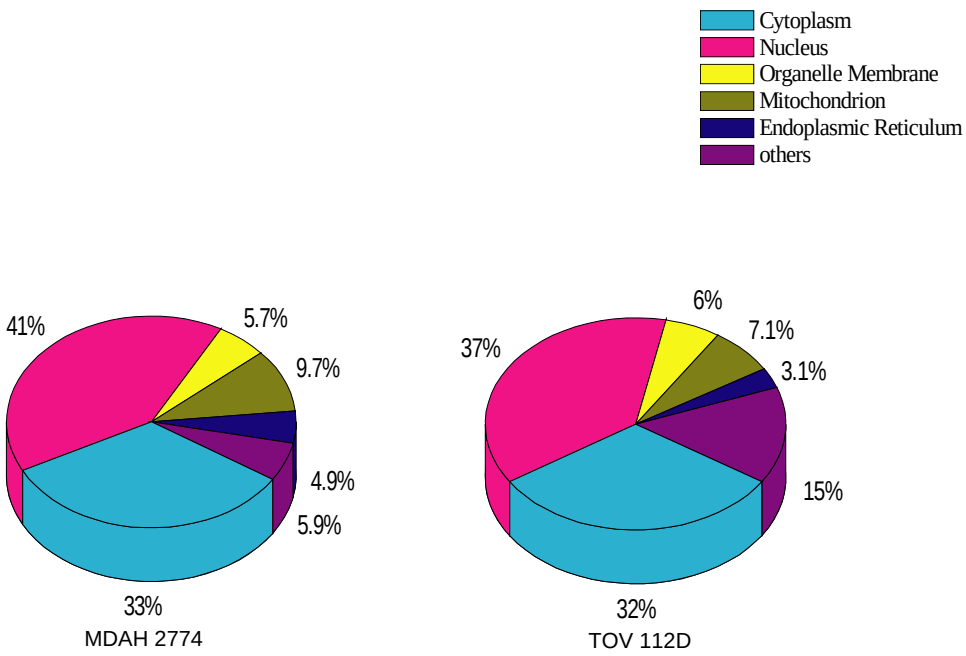


Figure 3.5. Cellular composition for both cell lines

3.5 References

[1] Alexander, H.; Stegner, A. L.; Wagner-Mann, C.; Du Bois, G. C.; Alexander, S.; Sauter, E. R. *Clinical Cancer Research*, **2004**, *10*, 7500-7510.

[2 Emmert-Buck, M. R.; Bonner, R. F.; Smith, P. D.; Chuaqui, R. F.; Zhuang, Z.; Goldstein, S. R.; Weiss, R. A.; Liotta, L. A. *Science*, **1996**, *274*, 998-1001.

[3] Gajewski, T. F. *Clinical Cancer Research*, **2006**, *12*, 2326s-2330s.

[4] Doss, M. X.; Winkler, J.; Chen, S.; Hippler-Altenburg, R.; Sotiriadou, I.; Halbach, M.; Pfannkuche, K.; Liang, H.; Schulz, H.; Hummel, O.; Hübner, N.; Rottscheidt, R.; Hescheler, J.; Sachinidis, A. *Genome Biology*, **2007**, *8*, R56.

[5] Chen, J.; Balgley, B. M.; Devoe, D. L.; Lee, C. S. *Anal. Chem.*, **2003**, *75*, 3145-3152.

[6] Pepaj, M.; Wilson, S. R.; Novotna, K.; Lundanes, E.; Greibrokk, T. J. *Chromatogr. A.*, **2006**, *1120*, 132-141.

[7] Chen, E. I.; Hewel, J.; Felding-Habermann, B.; Yates III, J. R. *Molecular and Cellular Proteomics*, **2006**, *5*, 53-56.

[8] Wu, R.; Hu, L.; Wang, F.; Ye, M.; Zou, H. *J. Chromatogr. A.*, **2008**, *1184*, 369-392.

[9] Nesbitt, C. A.; Lo, J. T. M.; Yeung, K. K. C. *J. Chromatogr. A.*, **2005**, *1073*, 175-180.

[10] Zhou, F.; Johnston, M. V. *Electrophoresis*, **2005**, *26*, 1383-1388.

[11] Metz, T. O.; Jacobs,J. M.; Gritsenko, M. A.; Fonte`s, G.; Qian, W. J.; Camp II, D. G.; Poitout, V.; Smith, R. D. *Journal of Proteome Research*, **2006**, *5*, 3345-3354.

[12] Motoyama, A.; Venable, J. D.; Ruse, C. I.; Yates, III, J. R. *Anal. Chem.*, **2006**, *78* 5109-5118.

[13] Zhu, K.; Zhao, J.; Lubman, D. M.; Miller, F. R.; Barder, T. J. *Anal. Chem.*, **2005**, *77*, 2745-2755.

[14] Xie, R.; Johnson, W.; Spayd,S.; Hall, G. S.; Buckley, B. *Anal. Chim. Acta*, **2006**, *78*, 186-194.

[15] Karlsson,E.; Ryden, L.; Brewer, J.; Janson, J. C. (Editor); Ryden, L. (Editor), Protein Purification, Wiley-VCH., New York, 1998, p145.

[16] Yamamoto, S.; Ishihara, T., *J. Chromatogr. A.*, **1999**, *A852*, 31-38.

[17] Sluyterman, L. A. AE.; Elgersma, O. *J. Chromatogr*., **1978**, *150*, 17-30.

[18] Sluyterman, L. A. AE.; Wijdenes, J. *J. Chromatogr*., **1978**, *150*, 31-44.

[19] Strong, J. C.; Frey, D. D. *J. Chromatogr*., **1997**, *A 769*, 129-143.

[20] Kang, X.; Frey,D. D. *Anal. Chem*., **2002**, *74*, 1038-1045.

[21] Shan, L.; Anderson, D. J. *J. Chromatogr A.*, **1998**, *814*, 43-54.

[22] Shan, L.; Anderson, D. J. *Anal. Chem*., **2002**, *74*, 5641-5649.

[23] Zhao, J.; Zhu, K.; Lubman,D. M.; Miller, F. R.; Barder, T. J. *Proteomics*, **2006**, *6*, 3847-3861.

[24] Wang, Y.; Wu, R.; Cho, K. R.; Shedden, K. A.; Barder, T. J.; Lubman, D. M. *Molecular and Cellular Proteomics*, **2006**, *5*, 3345-3354.

[25] Yang, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal. Chem*., **2003**, *75*, 2299-2308.

[26] Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. *Anal. Chem*., **2003**, *75*, 4646-4658.

[27] Tabb, D. L.; McDonald, W. H.; Yates III, J. R. *Journal of Proteome Research*, **2002**, 1, 21-26.

[28] Wang, H.; Kachman, M. T.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Proteomics*, **2004**, *4*, 2476-2495.

[29] Casanova, M. L.; Bravo, A.; Martínez-Palacio, J.; Fernández-Aceñero, M. J.; Villanueva, C.; Larcher, F.; Conti, C. J.; Jorcano, J. L. *FASEB J*., **2004**, *18*, 1556-1558.

[30] Trask, D. K.; Band, V.; Zajchowski, D. A.; Yaswen, P.; Suh, T.; Sager, R. *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*, 2319-2323.

[31] Dorsky, R. I.; Moon, R. T.; Raible, D. W. *Nature*, **1998**, *396*, 370-373.

[32] Wu, R.; Zhai,Y.; Fearon, E. R.; Cho, K. R. *Cancer Research*, **2001**, *61*, 8247-8255.

[33] Polakis, P. *Genes and Development*, **2000**, *14*, 1837-1851.

[34] Sakanaka, C.; Leong, P.; Xu, L.; Harrison, S. D.; Williams, L. T. *Proc. Natl. Acad. Sci. U.S.A.*, **1999**, *96*, 12548-12552.

Chapter 4

The Identification of Auto-Antibodies in Pancreatic Cancer Patient Sera Using a Naturally Fractionated Panc-1 Cell Line

4.1 Introduction

New methods for detection of cancer are needed to improve diagnosis and treatment to improve the survival rate. In particular, pancreatic ductal adenocarcinoma (PDAC) has one of the poorest survival rates of any cancers, where according to the American Cancer Society, for all stages of pancreatic cancer combined, the one-year relative survival rate is 20%, and the five-year rate is 4% [1]. These low survival rates result from the fact that fewer than 10% of patients' tumors are confined to the pancreas when in most cases, diagnosis of 80% to 90% of PDAC cases are too late for surgical procedures to have a positive outcome. Current markers for PDAC have shown insufficient sensitivity and specificity for early diagnosis including the commonly used CA 19-9 test [2, 3] where serum levels are significantly increased in pancreatitis in addition to pancreatic cancer. Therefore there is a real need for developing new screening tools and biomarkers for PDAC [4]. The idea that there is an immune response to cancer in humans has been demonstrated by the identification of autoantibodies against a number of intracellular antigens in patients with various tumor types [4, 5, 6, 7]. This phenomenon is known as the humoral response and the detection of such autoantibodies has been shown to be of great potential diagnostic and prognostic value in the detection of

cancer and the ability to predict the course of the disease [1, 2]. The basis of this humoral response is that all cells secrete proteins or fragments of proteins into the bloodstream. Cancer cells secrete some proteins that are different than normal cells and these proteins that are different may result in the differential autoantibody response that is observed. Many of these proteins are secreted in relatively low amounts and would be difficult to detect directly. A major advantage of the autoantibody response is that in effect it provides a means to amplify these proteins for detection.

The humoral response has been demonstrated in a number of cases where for example, it has been shown that somatic alterations in the p53 gene elicit a humoral response in 30-40% of affected patients [6]. The detection of anti-p53 antibodies has been shown to sometimes predate the diagnosis as in the case of lung cancer [7,8] and has also been shown to be a potential marker for breast [8] and colorectal [9] cancers. In other work, 60% of patients with lung adenocarcinoma exhibited a humoral response to glycosylated annexins I and/or II whereas none of the noncancerous standards exhibited such a response [4]. There are now a substantial number of examples demonstrating this humoral response to patient sera. In addition, it has been shown that the majority of antigens from tumor cells that elicit this response are not just products of mutated genes. These proteins are often differentiation antigens or other proteins over-expressed in cancer [6].

One of the issues in probing the humoral response is that a platform is required where the response can be tested against large numbers of proteins. Most recently microarrays have been used to provide such a platform where relatively small amounts of often valuable serum can be probed against thousands of protein substrates. There have

been a number of approaches used to generate such protein-antigen arrays. Many of these approaches use recombinant proteins obtained from cDNA libraries [10, 11, 12] or phage display libraries [13]. These methods have been shown to have excellent sensitivity, which is sufficient for measurement of many clinically important proteins in patient blood samples and sera and arrays with large numbers of proteins or protein components have been produced. More recently several groups have developed methods for printing cDNAs on arrays where the proteins are transcribed and translated in situ as needed [14, 15]. A limitation of these technologies though has been that in cellular systems proteins may undergo numerous posttranslational modifications which often play an important role during the cancer process [21-34]. This makes it important to use fractionated cellular proteins as baits to study the autoantibody response. Such methods have included the use of tissue microarrays [35] and the extraction of proteins from cells using either 2-D gel electrophoresis [9] or liquid separation methods [36, 37].

In current work, we have used liquid fractionation methods to produce microarrays for the humoral response experiment against a Panc-1 pancreatic cell line. The methods that are used involve separating intact proteins from cell lysates using two dimensions. A total cell fractionation can be performed using chromatofocusing separation in the first dimension where the proteins are fractionated according to pI [38-42]. Each fraction is then separated in a second dimension by nonporous silica RP HPLC [42-46]. The result is a 2-D separation of the proteins from the cell lysate where relatively pure proteins in the liquid phase are obtained. Using this method isolated proteins in the liquid phase can be collected for spotting on coated glass slides [37]. The protein spots are probed for their humoral response by exposing them to sera from cancer and chronic

pancreatitis patients, type 2 diabetes patients and normal individuals. This method offers a means for comprehensive proteomic analysis of proteins from large numbers of purified proteins as expressed in cancer cells while maintaining their PTMs that are often critical to the humoral response [9]. The method can produce arrays with over a thousand spots and can produce large numbers of slides for testing the response against a large number of patients.

The majority of results to date show that in humoral response trials where there are a large number of patients tested that only a subset of patients with a specific tumor type will develop a response to a specific antigen [5-11, 18, 47], typically 10-20%. The reason for this variability over a patient population is not yet clear, however, the results generated by the humoral response require statistical analysis to identify the potential marker candidates. In the current work, we use several statistical methods to interpret our results including Outlier Sum, COPA, Wilcoxon and Pamr analysis. A comparison of the results from these methods provides candidates for further testing. Based on our results we perform a prevalidation of 4 potential markers for pancreatic cancer against recombinant proteins on a microarray based format against samples from pancreatic cancer, pancreatitis, diabetes type 2 and normal controls. Several potential markers are identified for further work.

4. 2 Materials and Methods.

4. 2. 1 Chemicals

Methanol, acetonitrile, urea, thiourea, iminodiacetic acid, dithiothreitol (DTT), n-octyl-D-glucopyranoside (OG), glycerol, bis-tris, Trifluoroacetic acid (TFA), and PMSF

(Phenylmethanesulfonyl fluoride) were purchased from Sigma (St. Louis, MO). Water was purified using a Milli-Q water filtration system (Millipore Inc., Bedford, MA) and all solvents were HPLC grade unless otherwise specified. Reagents used were in the purest form commercially available. Polybuffer 74 and polybuffer 96 were purchased from GE Healthcare Bio-Sciences Corp. (Piscataway, NJ). 1x PBS and ultra-pure DNase/RNase free distilled water were obtained from Invitrogen (Carlsbad, CA).

4. 2. 2 Serum samples

Eighty six serum samples were obtained at the time of diagnosis following informed consent using IRB-approved guidelines. Sera were obtained from patients with a confirmed diagnosis of pancreatic adenocarcinoma in the Multidisciplinary Pancreatic Tumor clinic at the University of Michigan Hospital. Inclusion criteria for the study included patients with a confirmed diagnosis of pancreatic cancer, the ability to provide written, informed consent, and the ability to provide 40 ml bloods. Exclusion criteria included inability to provide informed consent, patient's actively undergoing chemotherapy or radiation therapy for pancreatic cancer, and patients with other malignancies diagnosed or treated within the last 5 years. Sera were also obtained from patients with chronic pancreatitis who were seen in the gastroenterology Clinic at University of Michigan Medical Center and from control healthy individuals collected at the University of Michigan under the auspices of the Early Detection Research Network (EDRN). The mean age of the tumor group was 65.4 years (range 54-74years) and the chronic pancreatitis group was 54 years (range 45-65). The sera from the normal subject group was age and sex-matched to the tumor group. The chronic pancreatitis group was

sampled when there were no symptoms of acute flare of their disease. All sera were processed using identical procedures. The samples were permitted to sit at room temperature for a minimum of 30 minutes (and a maximum of 60 minutes) to allow the clot to form in the red top tubes, and then centrifuged at 1,300 x g at 4^{o}C for 20 minutes. The serum was removed, transferred to polypropylene, capped tubes in 1 ml aliquots, and frozen. The frozen samples were stored at -70^{o}C until assayed. All serum samples were labeled with a unique identifier to protect the confidentiality of the patient. None of the samples were thawed more than twice before analysis. This study was approved by the Institutional Review Board for the University of Michigan Medical School.

4. 2. 3. Cell culture

The Panc-1 PDAC cell line was cultured in Dulbecco's modified Eagle medium supplemented with 10% fatal bovine serum, 100 units/mL penicillin and 100 units /mL streptomycin (Invitrogen, Carlsbad, CA). Upon reaching 80% confidence, the cells were washed twice in 10mL 1X PBS containing 4mM Na_3VO_4, 10mM NaF and one half of a protease inhibitor cocktail tablet. The sample was then solubilized in 300ul lysis buffer consisting of 7M urea, 2M thiourea, 100mM DTT, 0.5% biolyte ampholyte 3-10, 2% OG, 4mM Na_3VO_4, 10mM NaF and 1mM PMSF at room temperature for 30min, followed by centrifugation at 35000 rpm at 4^{o}C for 1hr. The supernatant was stored at -80°C until further use.

4. 2. 4 Chromatofocusing(CF)

Prior to CF, a PD10 column (Amersham Biosciences) was used to exchange the cell

lysate from the lysis buffer solution to the CF buffer solution according to the manufacturer's protocols. The start buffer consisted of 6M Urea, 0.2% OG, 25mM bis-tris. The elution buffer solution was composed of 6M urea, 0.2% OG, and a 10 fold dilution of polybuffer 96 and polybuffer 74 in a ratio of 3:7. The pH of both buffer solutions (7.9, 4.0) was adjusted with saturated imminodiacetic acid. A chromatofocusing column (weak anion exchange HPCF-1D prep column, 250mmL x 4.6mm ID, Eprogen, Darien, IL) was pre-equilibrated with the start buffer solution and 13mg of the cell lysate was injected into the CF column with multiple injections. Fractionation was started after switching elution buffer and a stable based line achieved. The pH fractions were collected in 0.3 pH intervals and pH was monitored using a flow-through on-line pH probe. UV absorption was recorded at 280 nm wavelength. When a pH of 4.0 was reached, elution buffer solution was switched to a 1M NaCl solution to wash the column followed by Isopropanol to elute out strongly bound proteins from the column. The collected fractions were stored at -80C° until used.

4. 2. 5 Reverse Phase HPLC Separation

An ODSI-1 (8 x 33 mm) column (Eprogen, Inc.) was used to separate the pH fractions of the Panc-1 cell line after CF. Solvent A was 0.1% TFA in water and solvent B was 0.1%TFA in acetonitrile. The gradient was run from 5% to 15% B in 1min, 15% to 25% in 2 min, 25% to 31% in 2min, 31% to 45% in 10min, 41% to 47% in 6min, 47% to 67% in 4min, 67% to 100% B in 3min, and reduced to 5% B in 1min after maintaining 100% B for 1min. The flow rate was 1ml/min and the column temperature 65°C. UV absorption was monitored at 214 nm wavelength. The fractions were collected in 96 well plates and

stored at -80°C.

4. 2. 6. Protein Microarrays

Approximately 30% of the total sample of the fractionated Panc-1 proteins obtained using 2D separation were transferred into 96-well printing plates (Bio-Rad) and lyophilized to dryness. The fractions were reconstituted in printing buffer which was composed of 62.5mM Tris-HCl (pH 6.8), 1% w/v sodium dodecyl sulfate (SDS), 5% w/v dithiothreitol(DTT) and 1% glycerol in 1 X PBS. Reconstituted fractions in printing plate were placed in a shaker overnight at 4°C. The fractions from the printing plate were spotted onto nitrocellulose slides using a non-contact piezoelectric printer (nanoplotter 2 GeSiM). Each spot contained 2.5nL of liquid of ~450μm in diameter, and 600μm in distance between spots. Printed slides were dried on the printer deck overnight and stored in desiccator at 4C if slides were not used immediately.

4. 2. 7. Hybridization of slides

The printed slides were blocked in a solution of 1% BSA in PBS-T overnight. Each serum sample was diluted 1:400 in probe buffer which consisted of 1% BSA, 0.5mM DTT, 5mM magnesium chloride, 0.05% Triton X-100, 5% glycerol in 1 X PBS. The slides were hybridized in diluted serum for 2hrs using a mini-rotator at 4°C. After hybridization, slides were washed five times using probe buffer in 5min, and then re-hybridized with goat-anti-human IgG conjugated with Alexafluor 647 (1μg/mL, Invitrogen, Calsbad, CA) for 1hr at 4°C. The slides were washed five times again with probe buffer for 5min each and dried. All slides were scanned using Axon 4000B microarray scanner (Axon Instruments Inc.,

Foster City, CA).

4. 2. 8. Data acquisition and analysis.

LC-MS/MS

The sample which was not used in microarray experiments in 96 well plates were dried down to approximately 10μL and mixed with 10%(v/v) ammonium bicarbonate, 10% (v/v) DTT, and 1:50 ratio (v/v) TPCK-treated trypsin (Promega, Madison, WI). The solution was incubated at 37°C overnight and the tryptic digestion was terminated by addition of 2.5% (v/v) of TFA. The digested peptide mixture was analyzed by nano-flow reverse-phase LC/MS/MS using the LTQ mass spectrometer with a nano-spray ESI ion source (Thermo, San Jose, CA). The samples were separated using a (0.1 x 150mm) capillary reverse phase column (MichromBioresources, Auburn, CA) with a flow rate of 5ul/min. An acetonitrile:water gradient method was used, starting with 5% acetonitrile which was ramped to 60% in 25min and to 90% in another 5min. Both solvent A (water) and B (ACN) contained 0.3% formic acid. The electrospray voltage was 2.6V, with a capillary temperature of 200°C and a capillary voltage of 4kV. The normalized collision energy was set at 35% for MS/MS. The MS/MS spectra obtained were analyzed using the Sequest feature of Bioworks 3.1 SR1, allowing only one missed cleavage during SwissProt human protein database searching. To further validate data obtained from Sequest, Protein prophet/peptide prophet software [48] modified in house was used to provide a confidence level in identification of 95%. Since there might be more than one protein in a protein spot on the microarray slide, we compared proteins identified in adjacent fractions. If the spot that responded to the humoral response was unique and did

not have an adjacent spot that lit up then the highest scoring protein based on LC-MS/MS analysis and protein prophet/peptide prophet was considered as the likely identification. If more than one protein was identified in the spot, then we also performed mass spec analysis on the adjacent spots. If the proteins were identified in the adjacent spots that did not respond then they were likely not to be the protein with the humoral response in our unique spot. However, if adjacent spots also showed a humoral response then the protein present in all spots was considered as the most likely candidate.

4. 2. 9. Statistical Analysis

GenePix 6.0 software was used to determine single-channel intensities and median local background intensities for each spot. A spot was considered positive for analysis if at least 2X signal to background intensity were measured. To account for variation between arrays, each array was median-centered and scaled by its interquartile range. We perform three pairwise comparisons to assess differences in humoral response between in cancer, normal and pancreatitis groups. . First, we employed the non-parametric Wilcoxon rank-sum test for each pair-wise comparison. We applied the FDR approach to calculate the FDR-adjusted p value and then visualize the results on the p-value plots. We then tried outlier-sum analysis for each pair-wise comparison to determine outlier proteins. In addition, we also used z-score statistics on the foreground data to look for differences between cancer and normal sera and between pancreatitis and normal sera. Finally, a classifier was built from the significantly differential proteins found by these methods.

Outlier Analysis

Two outlier analyses were performed: COPA and OS. COPA (Cancer Outlier Profile Analysis) were performed by using copa package in Bioconductor software and OS (Outlier Sum) by using R (version 2.8). Instead of using the aforementioned quantile normalization, each row of the original foreground data was standardized by subtracting the median and dividing by the MAD (median absolute deviation) of each protein. For each pair-wise comparison, the Outlier Sum statistic for each protein was calculated as the sum of the normalized intensity values in the disease group that beyond the 75th quartile [49] plus the interquartile range of that protein. The limit 75^{th} quartile plus interquartile range is defined to be outlier protein in statistical sense. Proteins with outlier sum statistics ranked top 5% were considered to be differential and used in the later analysis.

Non-parametric method

Three pair-wise Wilcoxon rank sum tests were performed between cancer versus normal, pancreatitis versus normal and cancer versus pancreatitis. Each pH/fraction combination was tested and the p-values were visualized in a grid plot to highlight regions of spots that exhibited differential response between normal and cancer sera. A p-value threshold of 0.05 was used to determine differential proteins for further study.

Prediction Analysis for Microarrays (PAM) classification algorithm

The PAM classification algorithm was used to explore the classificatory power of the proteins found to be differential between two groups using the Wilcoxon rank-sum test + Outlier Sum Statistics or Wilcoxon rank-sum test + the z-score method. From PAM we chose the smallest subset of proteins that gave us the lowest error rate to use as a

classifier.

4. 2. 10. Validations

Recombinant proteins were purchased from Abnova Coporation (Taiwan), and Genway Biotech Inc., (SanDiego, CA). The concentration of each recombinant protein is 10ug/mL. A piezoelectric non-contact printer (Nano Plotter, GeSIM) was used to print all the recombinant protein arrays on ultra-thin nitrocellulose slides (PATH slides, GenTel Bioscience). Each spotting event that resulted in 500 pL of solution being deposited was programmed to occur 5 times/spot to ensure that 2.5 nL was being deposited on each spot. Each recombinant protein was printed in triplicate and 14 identical blocks were printed on each slide. The slides were washed three times with 0.1% Tween in PBS buffer (PBST 0.1) and then blocked with 1% bovine serum albumin (Roche) in PBST 0.1 for one hour. The blocked slides were dried by centrifugation and inserted into a SIMplex (GenTel Bioscience) multi-array device which divides each of the slides by 16 small wells. The wells separate the neighboring blocks and prevent crossing contamination. Serum samples were diluted 10 times with PBST 0.1 containing 0.1% Brij. One hundred microliters of each diluted sample was applied to the recombinant protein array and the hybridization was performed in a humidified chamber for one hour. The slides were then rinsed three times to remove unbound proteins. 1ug/mL goat anti-human IgG conjugated with Alexafluor647 (Invitrogen, Carlsbad, CA) solution was used for detection. After second one-hour hybridization with anti-human IgG, the slides were washed and dried again, then scanned with a microarray scanner (Axon 4000A). The program Genepix Pro 6.0 was used to extract the numerical data. The averaged intensity of the spots in the control block

was considered to be a fixed number as *A*. The intensity for each of the recombinant protein spots was calculated as $S \times A/B$ where *B* is averaged intensity of the spots in the control block on the specific slide

4. 3. Results and Discussion

Panc-1 human pancreatic ductal adenocarcinoma (PDAC) cell line was used as bait to study the humoral response in pancreatic cancer since the panc-1 cell line has been shown to maintain some of the differentiated characteristics of normal mammalian pancreatic ductal epithelial cells [50]. The analytical work flow is illustrated in Figure 4.1. The solubilized protein solution extracted from Panc 1 cell line was fractionated using 2-D liquid separation methods as described consisting of chromatofocusing in the first dimension followed by nonporous reversed phase HPLC where intact proteins were collected as the final product. Fraction collection was performed where liquid eluent from each chromatographic peak was collected into 96 well plates. Each collected protein fraction was separated into several parts for further work. One portion was used for spotting the microarray plates and a second portion was used for protein identification based on LC-MS/MS. There were 1029 protein peaks obtained over a pH range of 8.0-4.0 spotted using the microarray device onto each nitrocellulose coated glass slide. Each slide was hybridized against a patient serum sample where the humoral response was run in this work against 38 cancer serum samples, 23 pancreatitis serum samples and 25 normal controls. Statistical analysis was then performed over this sample set to determine which proteins provided a significant response to patient serum.

Statistical Analysis

The autoantibody response varies over a patient population where statistical methods are required to decide which protein responses are significant. In this work on the humoral response in pancreatic cancer three categories were used to evaluate the differential humoral response toward the disease state using three out of four statistical analysis approaches (COPA, OS, Wilcoxon, and Pam). The categories include pancreatic cancer vs normal, pancreatitis vs normal, and cancer vs pancreatitis. This humoral response study for pancreatic cancer includes evaluating the signal of autoantibodies in pancreatitis against normal/cancer since pancreatitis is often confused for pancreatic cancer in diagnosis [51]. The use of statistical methods in this work is especially important since the variability of the humoral response from patient to patient requires a sufficiently large number of samples for evaluation. There have been two major methods to analyze microarray data, non-parametric, and parametric models [52]. It has been suggested that parametric statistics are less desirable over non-parametric statistics unless there is a strong distributional assumption inherent in the data [53]. However, a number of new approaches to analyze data for microarray experiments including parametric methods have been proposed.

COPA (cancer profile outlier analysis)

The recent statistical method called COPA has been proposed by Tomlins and coworkers [54]. The COPA analysis detects differential expression of proteins from microarray data for cancer studies only in a subset of cancer samples with increased expression. Since the majority of cancers have heterogeneous activation, it appears that

the application of this method using the “subset” idea where some cancers respond to the humoral response and others do not respond may result in an improved performance for microarray data relative to traditional t-statistic analysis. COPA can be categorized as a parametric method to analyze large amounts of microarray data, but it has as yet only found limited applications to microarray work. COPA statistics were used in this work to identify differential expression of signals on spots in each comparison, i.e. cancer vs normal, cancer vs pancreatitis, and pancreatitis vs normal (Figure 4.2). Figure 4.2 illustrates three examples of an ordered bar plot for several spots with the largest mutually exclusive outlier number above the common value (default is 5) as a cutoff for 'outlier status' [49] for each comparison category. The result of the COPA analysis as shown in Table 4.1 indicates 6 proteins as significant in cancer vs normal, 7 proteins as significant in cancer vs pancreatitis and 14 proteins as significant in pancreatitis vs normal with differential response to human sera.

In the case of cancer vs normal, there are 10 protein spots each of which has 8 mutually exclusive outliers where the p-value for having 8 or more outliers by chance (based on permutation) is $\mathbf{P}_{C \text{ vs } N} = 0.006$(data not shown). Among the 10 protein spots, there are 6 unique proteins determined in those 10 spots based on mass spec analysis with an increased humoral response in pancreatic cancer against normal. In addition, there are 20 spots with 6 mutually exclusive outliers for the distinction of cancer vs pancreatitis and a p-value of having 6 or more outliers by chance (based on permutation) is $\mathbf{P}_{c \text{ vs } p} = 0.037$. Seven proteins were identified as unique and 3 out of 7 proteins only showed a differential response in cancer against pancreatitis but not in cancer vs normal or pancreatitis versus normal. Among the 3 proteins, Phosphoglycerate kinase 1 (PGK-1) is

known to be a potential marker for gastric cancer [55] and involved in tumorigenesis of human colorectal cancers [56]. In the autoantibody experiments in the third category between pancreatitis vs normal, 16 protein pairs have 6 mutually exclusive outliers with a p-value having more than or equal to 6 outliers by chance(based on permutation) of $P_{P\ vs\ N} = 0.011$. Fourteen proteins were identified using mass spec analysis and 11 proteins were unique, showing a response in only pancreatitis against normal sera and not in the other categories. In two comparisons using COPA analysis, cancer vs normal and cancer vs pancreatitis only 1 protein appears to be truly unique to each category and does not overlap with results from other categories such as pancreatitis vs normal. The proteins identified are C-myc-binding protein Mm-1 for cancer versus normal and Eukaryotic translation initiation factor 4B for cancer versus pancreatitis as seen in Table 4.1.

OS analysis

Another proposed statistical analysis for microarray data known as outlier sum (OS) has been used to detect differential expression in cancer [57]. OS and COPA use a similar concept to detect increased expression values in a small subset among a class of disease samples, however OS was reported to have improved performance relative to COPA, resulting in a smaller false discovery rate [58, 59]. OS analysis was applied to our humoral response sample set to analyze the signal response with respect to the three categories (Figure 4.3). Unlike COPA which used a common value (cut off 5) for the threshold to characterize the outlier sample expression, the threshold for OS was set as the 75th quartile [49] of the sum over the outlier sample expression. In OS analysis, each comparison generated the highest ranked 9 spots based on the outlier sum statistics for

the humoral response for each category on the microarray slides (Table 4.2). The comparisons between cancer vs normal and between cancer vs pancreatitis show that most of the proteins(7 out of 9) are found in both comparisons for the OS analysis, however in the case of the OS analysis almost all of the proteins identified in pancreatitis vs normal are unique which is not the case for the COPA analysis. Some of the identified proteins have been previously reported as pancreatic cancer markers including sulfide:quinone oxidoreductase, mitochondrial precursor. Phosphoglycerate kinase 1(PGK-1) is one of the proteins identified in both statistical analyses (OS) with an upregulated humoral response in cancer versus normal and cancer versus pancreatitis. PGK1 protein is an enzyme involved in releasing angiostatin and is secreted by tumor cells where secretion can be increased by hypoxia [60]. PGK-1 protein has been flagged in colon cancer [61] and pancreatic cancer [62]. Also a humoral response to PGK-1 was observed in another PDAC cell line, MIAPACA, showing a strong response to pancreatic cancer compared to pancreatitis and normal. Glyceraldehyde-3-phosphate dehydrogenase, liver (GAPDH) protein showed a differential expression in the cancer vs normal comparison, but not in cancer vs pancreatitis based on OS.

Initially it was expected that OS may have improved performance compared to COPA for analyzing microarray data. COPA has been used to analyze gene microarray data based on the assumption that generic translocations occurs in cancers and a given translocation is only likely to occur once per sample Therefore, COPA intended to look for gene pairs that have large number of mutually exclusive outlier in cancer sample but no outlier in normal sample. These candidate genes are likely to involve in recurrent gene translocation with an unknown oncogene. However, there are similar assumptions for the

protein data. For a cancer sample, there may be multiple proteins whose intensities are significantly higher (or lower) than the average intensities of the cancer samples. However, OS analysis shows that most of the identified proteins with higher expression exist in both comparisons, cancer vs normal and cancer vs pancreatitis, showing less specificity. These OS statistics may need a higher threshold based on the intensity of the signal, for example, 90% instead of the 75% which was used in the current protein microarray study.

Wilcoxon anlaysis (Non parametric method).

The application of the Wilcoxon rank sum test to microarray experiments has generally shown improved performance relative to traditional t-test [63] results since humoral response data does not have a perfect Gaussian distribution. In Wilcoxon rank sum tests, three comparisons between cancer vs normal, cancer vs pancreatitis, and pancreatitis vs normal have been performed. The Wilcoxon test generated figures for protein spots with different p values such as $p<0.05$, $0.05<p<0.1$, $0.1<p<0.25$ with different colors for each comparison. The p-value here for <0.05 sets the probability is >95% for having higher response between categories. The generated figures are plotted in a 2D separation mass map format with pI on the x-axis and liquid fraction number corresponding to a spot on the microarray which has been tested against the humoral response on the y-axis. The colored bands in the figure generated using Wilcoxon analysis were identified using mass spec analysis, for spots with only $p<0.05$ for elevated expression in the comparisons between cancer vs normal, cancer vs pancreatitis, and pancreatitis normal as shown in Figure 4.4.

Three proteins were identified for increased response in the comparison of cancer vs normal and pancreatitis vs normal and 2 proteins were also identified for the comparison of cancer vs pancreatitis based on the Wilcoxon statistics. Among these are the Heterogeneous nuclear ribonucleoproteins (hnRNPs) such as hnRNP A/B, hnRNP K, hnRNP Q, and hnRNP D0 (Figure 4.5) which were identified in several statistical analyses thus far. hnPNP proteins appear to show humoral response in multiple sclerosis [64] and some of the hnRNPs showed even increased expression in pancreatic tissue and tumor cells [65]. Elongation factor alpha-1(eEF1A) protein (Figure 4.5) which showed increased response in cancer versus pancreatitis by the Wilcoxon method has been shown as an auto-antigen in Felty's syndrome [66] as well as in a prostate cancer cell line and in breast cancer [67]. It is interesting to note that the Wilcoxon method produces data with both upregulated and downregulated humoral response compared to the statistical outlier methods used which only showed upregulated response based on the outliers observed.

Pam (Prediction Analysis for Microarray) analysis

Pam is neither a non-parametric nor parametric statistics method which can be used to analyze microarray data. Pam is one of a type of clustering method used for classification in this case for cancer, pancreatitis, and normal. It uses a computing shrunken centroid [68] for each class and predicts whether the unknown sample falls into a class based on the nearest centroid. In this study, only the comparison of the differential response between cancer versus normal was applied using fractions determined to be significant from the Wilcoxon test (p value between -0.05 and 0.05) in a total of 60 samples and 93 fractions. The PAM results give 19 signature fractions which

could be best classified between cancer and normal and present an overall classification error rate of 0.297 (Figure 4.6). The same rules were used to determine identification of protein as those used in COPA and OS analysis.

Pam analysis also identified proteins with an increased autoantibody response for pancreatic cancer versus normal sera including 40S, 60S human ribosomal proteins (Table 4.4). 60S Ribosomal proteins L7a, L19 detected in this study have shown increased expression in breast cancer [69] and colorectal cancer [70]. PDZ and LIM domain protein 1 also showed a differential humoral response between cancer vs normal sera. The list of proteins with higher response in pancreatic cancer sera versus normal sera appears to be very different than the other statistical analyses where only 2 of the proteins can be identified in these analyses. However, the Pam test uses a different statistical approach such as clustering so the results may be different than the other methods.

Pre-Validation

For selected proteins that showed a differential humoral response in the discovery experiments a pre-validation experiment was performed using recombinant proteins with a set of independent samples of 16 normal, 24 cancer, 16 type 2 diabetes and 14 pancreatitis sera. Type 2 diabetes serum was tested since some fraction of pancreatic cancer patients will develop this condition. The four proteins chosen for this study were based on our discovery analysis as Annexin 2, Malate dehydrogenase, cytoplasmic, Heterogeneous nuclear ribonucleoprotein D0, and Peroxiredoxin 1 (Figure 4.7). The recombinant proteins were spotted on nitrocellulose slides and probed with

normal, pancreatitis, pancreatic cancer, and diabetic sera.

In order to measure the auto-antibody response that is reactive against the recombinant proteins, the serum must be diluted properly i.e. the amount of available auto-antibody in the serum should be lower than the binding capacity of the specific recombinant protein to avoid saturation while still providing good signal. Therefore, a saturation curve was made using different dilutions of serum to hybridize against identical blocks of the recombinant proteins. The result of the saturation test showed that with ten-time dilution, the recombinant proteins were not saturated and yielded a signal/background ratio of >5. Higher or lower dilution resulted in partial saturation or decreased signal. In the following pre-validation experiment using recombinant proteins, the dilution fold was set at ten times.

In this pre-validation experiment Annexin A2 and MDH1 showed higher differential response in cancer sera against pancreatitis or normal sera. A recent study has revealed increased levels of annexin A2 observed in cancer cells and tissues [71]. The pre-validation experiments illustrated that Annexin A2 showed a higher differential expression in pancreatic cancer sera (8/24 cancer) versus normals(1/16) or pancreatitis(2/14) although there was some response above the baseline shown for the diabetes samples(3/16). Previous studies have shown that Annexin 2 protein can be a marker for pancreatic cancer but could not distinguish it from pancreatitis [72]. This study shows that Annexin2 may have potential as a marker for further validation studies to distinguish cancer from pancreatitis.

Malate dehydrogenase, cytoplasmic (MDH1) is a protein related to delivering NADH across membranes and cell division [73]. An elevated level of expression of

MDH1 was observed in a genetic study with thyroid oncocytic tumors [74]. The prevalidation study on this protein using our independent test set shows that MDH1 clearly responds to pancreatic cancer relative to pancreatitis, diabetes or normal sera samples where 6/24 cancer samples respond, 1/14 normals respond and there is no response for pancreatitis or diabetes. MDH1 may serve as a potential marker for further validation in these humoral response experiments.

Heterogeneous nuclear ribonucleoprotein D0 protein was also tested for its humoral response (Figure 4.7c) where this protein showed a clear differential response between cancer versus normal sera but showed a comparable response for pancreatitis and diabetes. The fourth protein selected, Peroxiredoxin 1, did not show a differential response between pancreatic cancer and normal in the experiment using the recombinant protein. This is unexpected given that the OS analysis showed differential expression of PRDX1 in pancreatic cancer over normal sera and was identified with more than 6 unique peptides in the tandem mass experiment. It could be due to the use of the recombinant proteins, which may lack the required PTMs to induce a humoral response or the protein may not be in a form to induce a humoral response. However, the Peroxiredoxon did show some preferential response towards pancreatic cancer versus pancreatitis which was predicted from the Wilcoxon analysis.

4. 4. Conclusion

A humoral response to tumor proteins may have utility for the detection of pancreatic cancer. We have used 2-D liquid separation and protein microarrays to study the humoral response in pancreatic cancer versus pancreatitis and normal controls.

Several different statistical methods were used to identify proteins that elicited a differential humoral response pattern between the different clinical groups that could be used for further pre-validation. Outlier Sum and COPA statistics based on outlier methods were used to identify proteins that could distinguish pancreatic cancer versus pancreatitis and normals where these two methods generally identified different groups of proteins as significant from each other in each of the groups. Although these methods are both based on outlier techniques they use different criteria for evaluating microarray data. Nevertheless, several proteins were chosen as good prospects for further evaluation including Malate dehydrogenase, cytoplasmic, PGK-1, Heterogeneous nuclear ribonucleoprotein D0 and Peroxiredoxin 1. Wilcoxon and Pamr statistical methods were also used to evaluate the humoral response data where there was some overlap with those proteins identified from outlier analysis but in general a different set of proteins was identified. From these analyses Annexin A2 and Peroxiredoxin were chosen for further evaluation. A pre-validation experiment was run with an independent set of serum samples against recombinant proteins obtained commercially. This experiment showed that Annexin 2 and Malate dehydrogenase, cytoplasmic responded to cancer above the baseline compared to pancreatitis and normal sera. Also PGK-1 from a previous study was shown to respond to cancer versus pancreatitis or normal. Heterogeneous nuclear ribonucleoprotein D0 was shown to respond to cancer versus normal but could not discriminate against pancreatitis or diabetes. Peroxiredoxin 1 showed some discrimination against pancreatitis but not against normal sera or diabetes which was predicted from the Wilcoxon statistics. Further work will require experiments with early stage cancer sera which currently is not available in sufficient numbers for these

experiments. Also, further work on the use of truncated or modified forms of these proteins may provide an improved response. In addition, as larger numbers of samples are collected larger validation sets can be run to test the validity of these potential targets for biomarker response.

Table 4.1 Protein identifications of spots according to COPA results

Cancer vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	Therotical pI	#unique pep match
2f2	P46783	40S ribosomal protein S10.	7.3-7.0	18886	8.31	10.15	2
4d3	Q99471	Prefoldin subunit 5 (C-myc-binding protein Mm-1)	6.7-6.4	17318	13.48	5.94	2
4h3	Q7Z739	YTH domain family protein 3.	6.4-6.1	63822	9.07	6.45	3
7f7	Q14103	Heterogeneous nuclear ribonucleoprotein D0	5.2-4.9	38411	12.54	7.61	3
7g9	P30101	Protein disulfide-isomerase A3 precursor	5.2-4.9	56747	21.09	5.98	9
10a5	P27816	Microtubule-associated protein 4 (MAP 4)	4.3-4.0	120944	22.58	5.32	14

Table 4.1. (Continued)

Cancer vs Pancreatitis

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	Therotical pI	#unique pep match
1f11	P00558	Phosphoglycerate kinase 1	7.9-7.6	44587	18.21	8.3	5
6b11	O75368	SH3 domain-binding glutamic acid-rich-like protein	5.8-5.5	12767	20.57	5.22	2
7f7	Q14103	Heterogeneous nuclear ribonucleoprotein D0	5.2-4.9	38411	12.54	7.61	3
7g9	P30101	Protein disulfide-isomerase A3 precursor	5.2-4.9	56747	21.09	5.98	9
10a5	P27816	Microtubule-associated protein 4 (MAP 4)	4.3-4.0	120944	22.58	5.32	14
10g12	O60506	Heterogeneous nuclear ribonucleoprotein Q	4.3-4.0	69506	13.15	8.68	8
11b8	P23588	Eukaryotic translation initiation factor 4B	salt wash	69184	17.52	5.49	9

Table 4.1. (Continued)

Pancreatitis vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	Therotical pI	#unique pep match
1a7	Q9UNX3	60S ribosomal protein L26-like 1.	7.9-7.6	17246	25.04	10.55	4
2f2	P46783	40S ribosomal protein S10.	7.3-7.0	18886	8.31	10.15	2
2g6	P62937	Peptidyl-prolyl cis-trans isomerase A	7.3-7.0	18001	19.17	7.68	3
4a5	P61353	60S ribosomal protein L27.	6.7-6.4	15788	6.65	10.56	2
4c1	P62249	40S ribosomal protein S16	6.7-6.4	16436	18.37	10.21	3
4c4	Q07020	60S ribosomal protein L18.	6.7-6.4	21622	18.59	11.73	3
4d1	P60174	Triosephosphate isomerase	6.7-6.4	26653	19.17	6.45	4
5g12	Q13572	Inositol-tetrakisphosphate 1-kinase	6.1-5.8	45593	4.64	5.78	1
6f12	P29373	Cellular retinoic acid-binding protein 2	5.5-5.2	15684	25.08	5.42	3
7e6	P40925	Malate dehydrogenase, cytoplasmic	5.2-4.9	36404	3.83	6.91	1
8e9	Q9NQG5	Uncharacterized protein C20orf77	4.9-4.6	36878	9.17	5.73	2
9c12	P12956	ATP-dependent DNA helicase 2 subunit 1	4.6-4.3	69800	6.33	6.23	4
10a12	O60506	Heterogeneous nuclear ribonucleoprotein Q	4.3-4.0	69506	13.15	8.68	8
10g2	Q00765	Receptor expression-enhancing protein 5	Injection peak	21480	10.9	8.25	3

Table 4.2. Protein identifications of spots according to OS results

Cancer vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	% mass cov	Therotical pI	#unique pep match
1b1	P62847	40S ribosomal protein S24.	7.9-7.6	15414	28.59	10.79	3
1f11	P00558	Phosphoglycerate kinase 1	7.9-7.6	44587	18.21	8.3	5
1f6	Q15369	Transcription elongation factor B polypeptide 1	7.9-7.6	12466	17.74	4.74	2
3e5	P04406	Glyceraldehyde-3-phosphate dehydrogenase	7.0-6.7	35900	18.97	8.58	4
4d6	Q9Y6N5	Sulfide:quinone oxidoreductase, mitochondrial precursor	6.7-6.4	49929	11.62	9.18	4
5g2	Q06830	Peroxiredoxin 1 (Thioredoxin peroxidase 2)	6.1-5.8	22097	29.35	8.27	6
8c3	O95881	Thioredoxin domain containing protein 12 precursor	4.9-4.6	19194	37.98	5.25	5
9a9	Q99729	Heterogeneous nuclear ribonucleoprotein A/B	4.6-4.3	36590	9.15	9.04	3
11d5	Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	IPA wash	44939	18.43	8.66	5

Table 4.2. (Continued)

Cancer vs Pancreatitis

Fraction	Access number	Protein Name	Fraction pH	MW	% mass cov	Therotical pI	#unique pep match
1b1	P62847	40S ribosomal protein S24.	7.9-7.6	15414	28.59	10.79	3
1f11	P00558	Phosphoglycerate kinase 1	7.9-7.6	44587	18.21	8.3	5
1f6	Q15369	Transcription elongation factor B polypeptide 1	7.9-7.6	12466	17.74	4.74	2
4d6	Q9Y6N5	Sulfide:quinone oxidoreductase, mitochondrial precursor	6.7-6.4	49929	11.62	9.18	4
5g2	Q06830	Peroxiredoxin 1 (Thioredoxin peroxidase 2)	6.1-5.8	22097	29.35	8.27	6
7a8	Q9NTK5	GTP-binding protein 9(Obg-like ATPase 1)	5.5-5.2	44716	10.75	7.64	3
8c3	O95881	Thioredoxin domain containing protein 12 precursor	4.9-4.6	19194	37.98	5.25	5
9a9	Q99729	Heterogeneous nuclear ribonucleoprotein A/B	4.6-4.3	36590	9.15	9.04	3
11d5	Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	IPA wash	44939	18.43	8.66	5

Table 4.2 (Continued)

Pancreatitis vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	% mass cov	Therotical pI	#unique pep match
1e11	P09012	U1 small nuclear ribonucleoprotein A	7.9-7.6	31129	27.64	9.83	6
2b6	P32969	60S ribosomal protein L9	7.6-7.3	21850	12.55	9.96	3
2d9	Q96JB2	Conserved oligomeric Golgi complex component 3	7.6-7.3	93906	2.17	5.39	2
2h8	P23396	40S ribosomal protein S3.	7.3-7.0	26672	36.03	9.68	8
3a1	P06493	Cell division control protein 2 homolog	7.3-7.0	34074	21.8	8.38	5
3f11	Q9H7Z7	Prostaglandin E synthase 2	7.0-6.7	41917	20.38	9.22	5
4g9	Q13242	Splicing factor, arginine/serine-rich 9	6.7-6.4	25527	14.29	8.74	3
7e6	P40925	Malate dehydrogenase, cytoplasmic	5.2-4.9	36404	3.83	6.91	1
10f8	P06733	Alpha enolase	13	47009	16.73	6.99	5

Table 4.3 Protein identifications of spots according to Wilcoxon results

Cancer vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	#unique pep match
1h3-1h6	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	8.97	37407	11.57	3
3b1,3c1-10	P46776	60S ribosomal protein L27a.	11	16421	6.77	1
4b3-4c1	P23528	Cofilin-1	8.26	18360	26.18	3
10b4-10b8	P61978	Heterogeneous nuclear ribonucleoprotein K	5.39	50945	12.58	5

Pancreatitis vs Normal

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	#unique pep match
1d7-1d11	P62701	40S ribosomal protein S4, X isoform	10.16	29449	29.84	9
7f6-7f2	Q14103	Heterogeneous nuclear ribonucleoprotein D0	7.61	38411	14.63	4
10b5-10b8	P61978	Heterogeneous nuclear ribonucleoprotein K	5.39	50945	12.58	5

Cancer vs Pancreatitis

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	#unique pep match
2b9-2b4	P68104	Elongation factor 1-alpha 1	9.1	50110	11.39	5
5h5-5h12	P07355	Annexin A2	7.56	38449	41.19	13

Table 4.4 Protein identifications of spots according to Pam results.

Fraction	Access number	Protein Name	Fraction pH	MW	%mass cov	#unique pep match
1b12	P63173	60S ribosomal protein L38	10.1	8082	51.3	3
1c2	P62241	40S ribosomal protein S8	10.32	24060	18.71	3
1d1	P22087	Fibrillarin (34 kDa nucleolar scleroderma antigen)	10.18	33764	18.51	4
1d3	Q06830	Peroxiredoxin-1	8.27	22097	27.08	5
1g12	P84098	60S ribosomal protein L19	11.48	23452	24.78	5
1h10-1h11	P62847	40S ribosomal protein S24	10.79	15414	28.59	4
1h2	Q02878	60S ribosomal protein L6	10.59	32577	11.03	3
2a1	Q92945	Far upstream element binding protein 2	8.02	72664	14.6	8
2a11	P62424	60S ribosomal protein L7a	10.61	29847	14.05	5
2a7	P30050	60S ribosomal protein L12	9.48	17808	23.56	4
3d4	O00151	PDZ and LIM domain protein 1	6.55	35919	7.85	2
3d7	Q9Y2S7	Polymerase delta interacting protein 2	8.8	42008	21.19	6
7b3	P53365	Arfaptin-2 (ADP-ribosylation factor-interacting protein 2)	5.72	37833	11.24	3
9f4	P19105	Myosin regulatory light chain 2, nonsarcomeric	4.67	19651	19.72	3
11e5-11e7	P02545	Lamin-A/C	6.57	74095	21.96	12

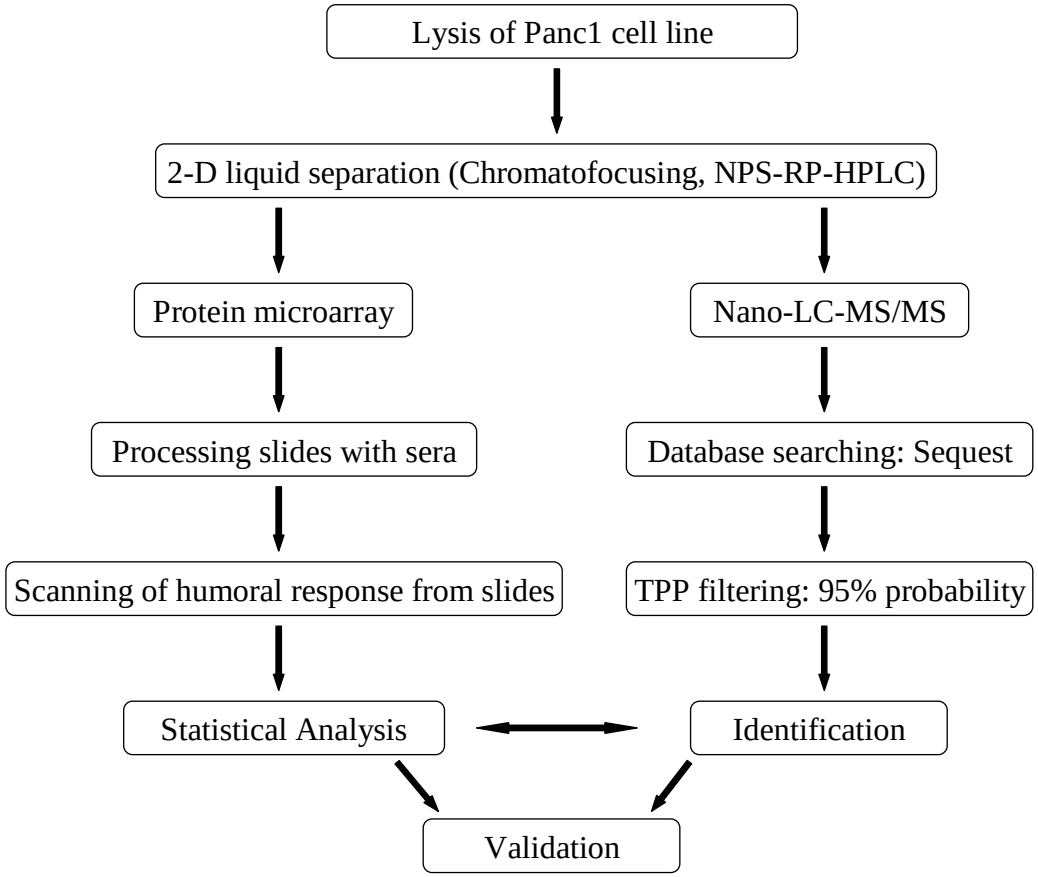


Figure 4.1. Flow Chart

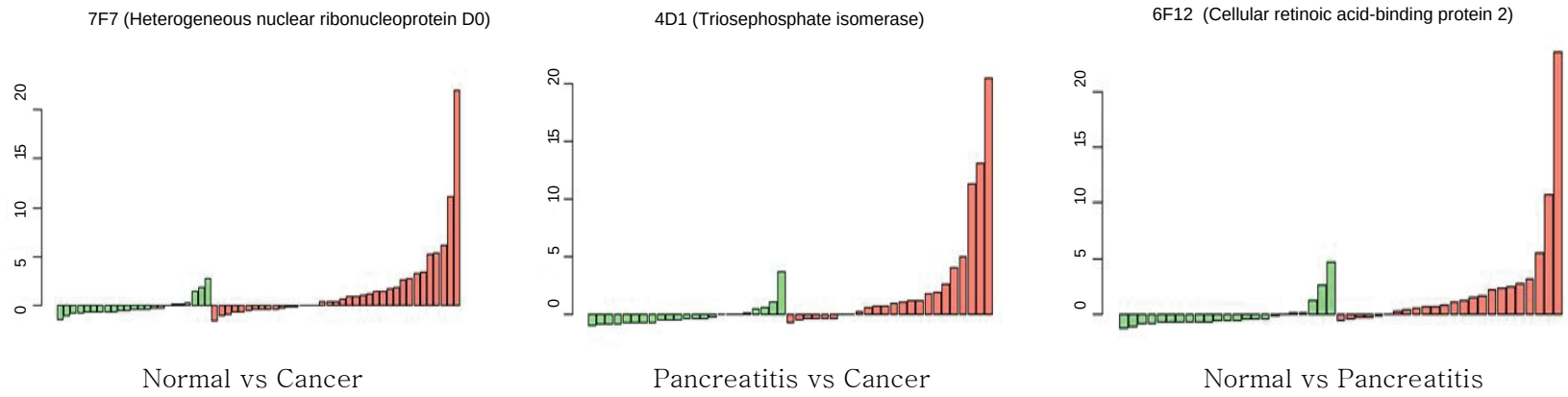


Figure 4.2. Results of COPA analysis for selected spots.

Y axis: COPA outlier number
X axis: samples.

Number on top of the figure is spot location on slides.
Green bar correspond to the normal adjacent samples and Salmon color bar correspond to the tumor samples.

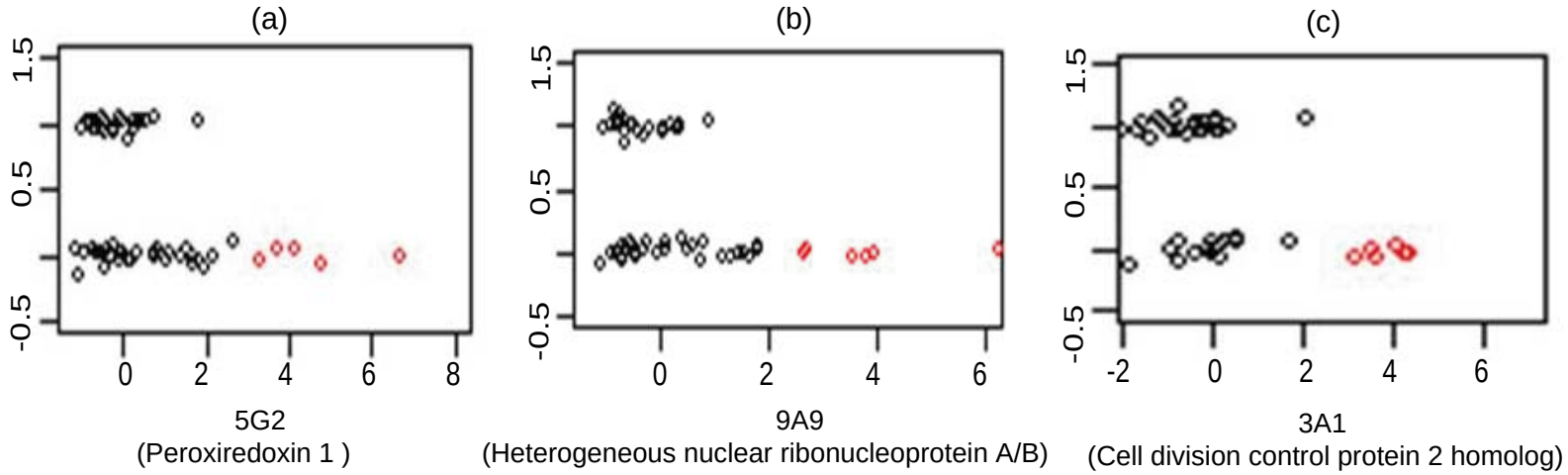
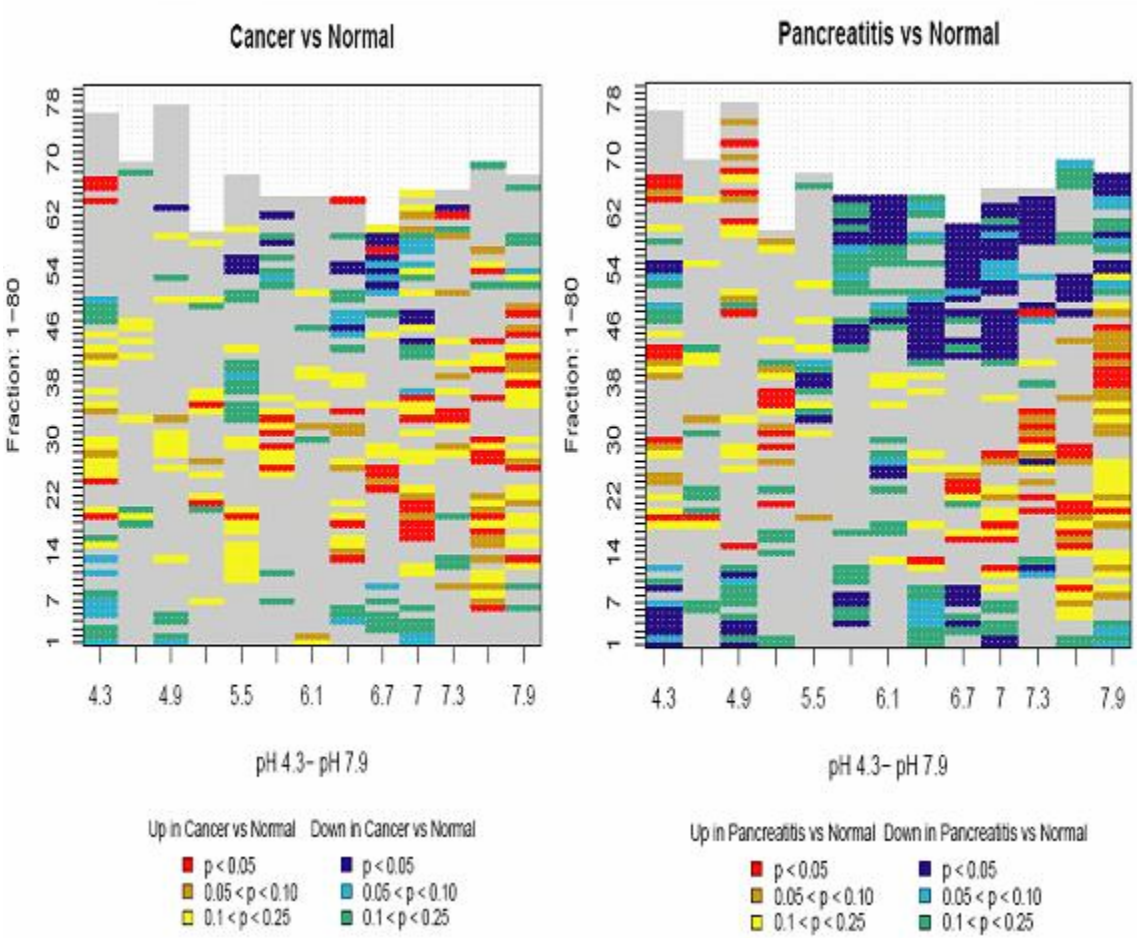


Figure 4.3. OS analysis results for selected spots.

Plots of the expression values in each class for the spot ranked highest by the outliers sum statistics.
The number in brackets is the location of spot on the slides.
The red points are identified as positive outliers; the black points are negative outliers.

X axis: outlier sum number
Y axis: classification, 0: disease state, 1: normal

(a) (b)


Cancer vs Normal
Fraction: 1−80
1 7 14 22 30 38 46 54 62 70 78
4.3 4.9 5.5 6.1 6.7 7 7.3 7.9
pH 4.3− pH 7.9
Up in Cancer vs Normal
p < 0.05
0.05 < p < 0.10
0.1 < p < 0.25
Down in Cancer vs Normal
p < 0.05
0.05 < p < 0.10
0.1 < p < 0.25
Pancreatitis vs Normal
Fraction: 1−80
1 7 14 22 30 38 46 54 62 70 78
4.3 4.9 5.5 6.1 6.7 7 7.3 7.9
pH 4.3− pH 7.9
Up in Pancreatitis vs Normal
p < 0.05
0.05 < p < 0.10
0.1 < p < 0.25
Down in Pancreatitis vs Normal
p < 0.05
0.05 < p < 0.10
0.1 < p < 0.25

(C)

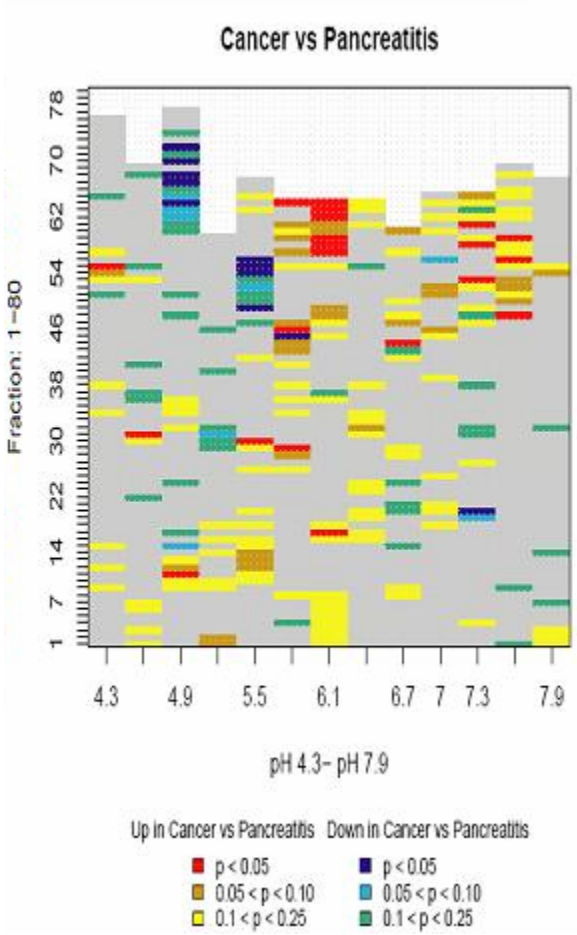


Figure 4.4. The result of Wilcoxon analysis

Following are the p-value plots for the three comparisons
(a)Cancer versus Normal sera
(b)Pancreatitis versus Normal sera
(c)Cancer versus Pancreatitis sera

These p-values are results from three pair-wise Wilcoxon tests.
For those proteins that did not correspond to a plate location or those with a pH value outside the range [4.0, 7.9], have been excluded in the analysis.
There are 866 proteins in the analysis. All the grey grids correspond to p-values >0.25.
pH = 4.3 means "pH is in 4.0-4.3", pH = 7.6 means "pH is in 7.6-7.9", etc.

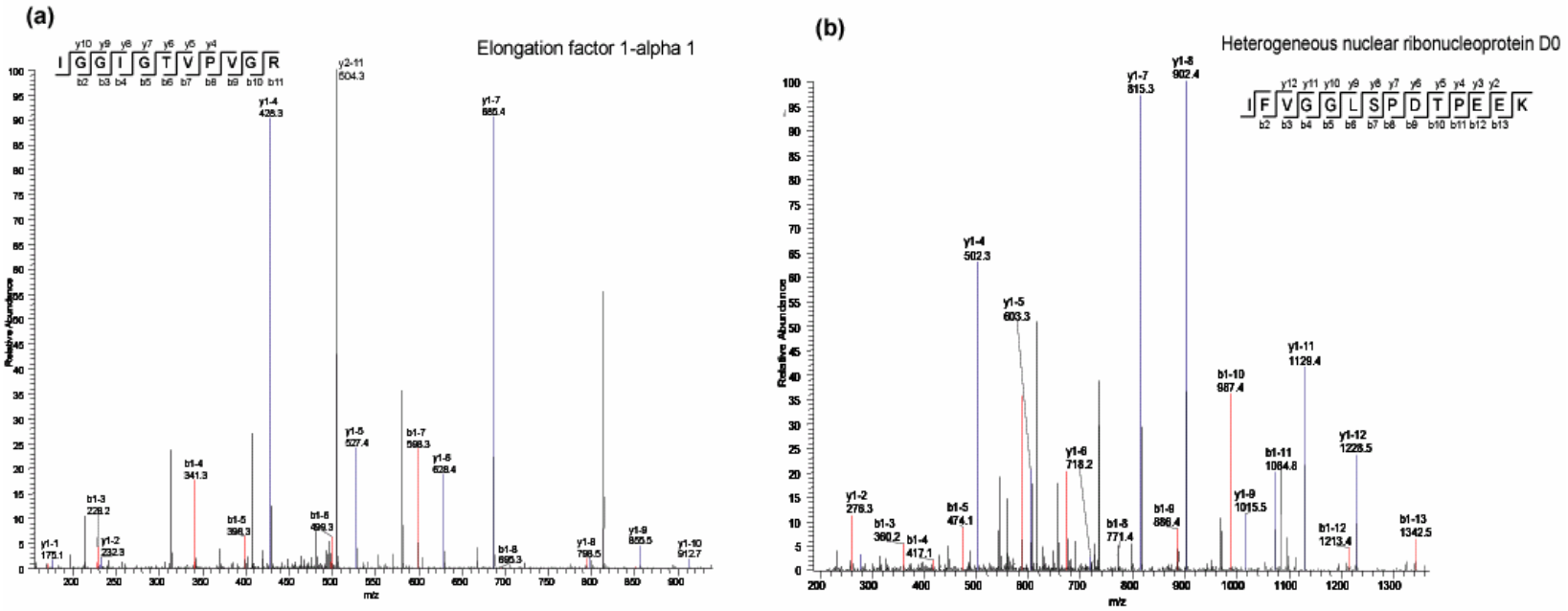


Figure 4.5 Selected microarray shots of differential humoral response as well as selected tandem mass spectrum for sequence confirmation of (a) EF1A1 and (b) hnRNP D0

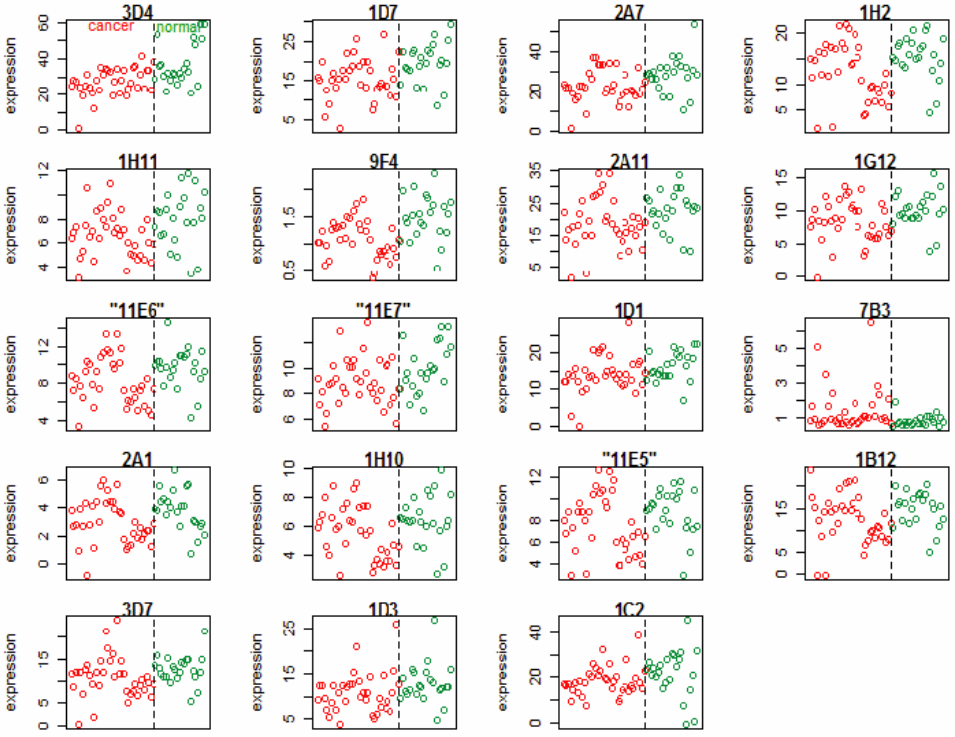


Figure 4.6. Result of Pam for array spots

Only cancer versus normal was compared.
and Only use significant fractions from wilcoxson test
(p value between -0.05 and 0.05)
- Total: 60 samples, 93 fractions

Results
- Pamr gives 19 signatures fractions can be best to classify cancer vs normal
- Overall classification error rate is 0.297

The red dot indicates expression from cancer samples
The green dot indicates expression from normal samples

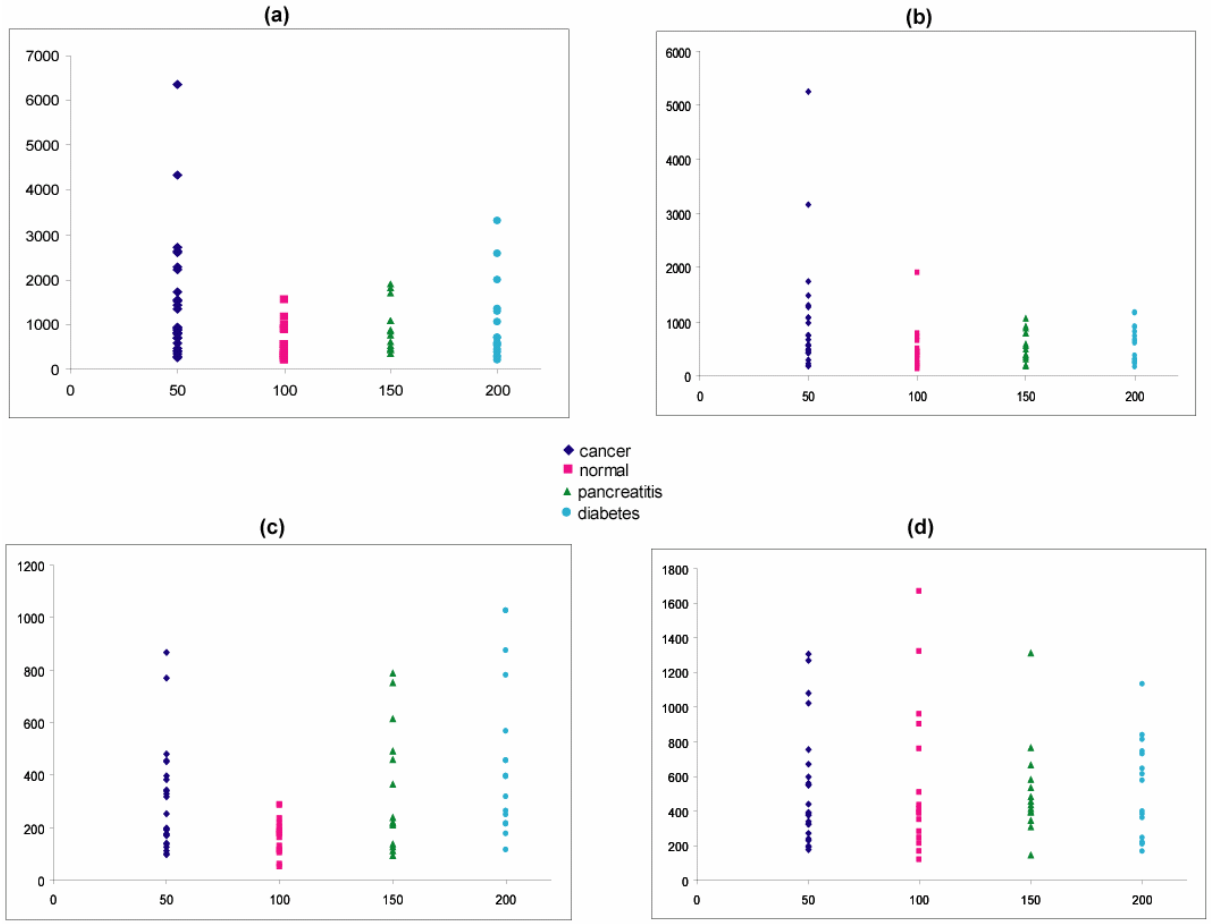


Figure 4.7 Scatterplot illustrating the differential humoral response in 4 diffferent recombinant proteins used for validating initial experimental results. (a) Annexin A2 (b) Malate dehydrogenase, cytoplasmic, (c) Heterogeneous nuclear ribonucleoprotein D0, and (d)Peroxiredoxin 1.

4.5 References

[1] Jemal, A.; Siegel, R.; Ward, E.; Murray, T., et al., Cancer Statistics, 2006. *CA Cancer J Clin* **2006**, *56*, 106-130.

[2] Garcea G.; Neal C. P.; Pattenden C. J.; Steward W. P.; Berry D. P. *Eur. J. Cancer*, **2005**, 41(15), 2213-2236.

[3] Steinberg W., *The American journal of gastroenterology* **1990**, *85(4)*, 350-355.

[4] Tian R.; Wei L. M.; Qin R.-Y.; Li Y.; Du Z.-Y., Xia W., Shi C.-J., Jin H. *Digestive Diseases and Sciences* **2008**, *53(1)*, 65-72.

[5] Haab, B.B.; Dunham, M. J.; Brown. P. O. *Genome Biology* **2001**, *2(2)*, 1-13.

[6] Macbeath, G..; Schreiber, S. L. *Science* **2000**, *289*, 1760-1763.

[7] Cahill, D. J. *Journal of Immunological Methods* **2001**, *250*, 81-91.

[8] Emili, A.Q.; Cagney, G.. *Nature Biotechnology* **2000**, *18*, 393-397.

[9] Walter, G..; Bussow, K.; Cahill, D.; Lueking, A.; Lehrach, H. *Current opinion in Microbiology* **2000**, *3*, 298-302.

[10] Hudson, M.E.; Pozdnyakova, I.; Haines, K.; Mor, G..; Snyder, M. *Proc. Natl. Acad. Sci.* **2007**, *104(44)*, 17494-17499.

[11] Bussow, K.; Cahill, D.; Nietfield, W.; Bancroft, D.; Scherzinger, E., Lehrach, H., Walter, G. *Nucleic Acids Research* **1998**, *26(21)*, 5007-5008.

[12] Trivers, G. E.; De Benedetti, V. M.; Cawley, H. L.; Caron, G.., Harrington, A. M.; Bennett, W. P.; Jett, J. R.; Colby, T. V.; Tazelaar, H.; Pairolero, P.; Miller, R. D.; Harris, C. C. *Clin. Cancer Res.* **1996**, *2*, 1767-1775.

[13] Regele, S.; Vogl, F. D.; Kohler, T.; Kreienberg, R.; Runnebaum, I. E. *Anticancer Res.* **2003**, *23*, 761-764.

[14] Takeda, A.; Shimada, H.; Nakajima, K.; Imaseki, H.; Suzuki, T.; Asano, T.; Ochiai, T.; Isono, K. *Eur. J. Surg*. **2001**, *167*, 50-53.

[15] Zhu, H.; Klemic, J. F.; Chang, S.; Bertone, P.; Casamayor, A.; Klemec, K.G.; Smith, D.; Gerstein, M.; Reed, M. A.; Snyder, M. *Nature Genetics* **2000**, *26*, 283-289.

[16] Bussow, K.; Cahill, D.; Nietfield, W.; Bancroft, D.; Scherzinger, E.; Lehrach, H.; Walter, G. *Nucleic Acids Research* **1998**, *26(21)*, 5007-5008.

[17] Lueking, A.; Horn, M.; Eickhoff, H.; Bussow, K.; Lehrach, H.; Walter, G. *Anal. Biochem.* **1999**, *270*, 103-111.

[18] Kilic, A.; Schuchert, M. J.; Luketich, J. D.; Landreneau, R. J.; Lokshin, A. E.; Bigbee, W. L.; El-Hefnaey, T. *The Journal of Thoracic and Cardiovascular Surgery* **2008**, *136(1)*, 199-204.

[19] Chatterjee, D. K.; Sitaraman, K.; Baptista, C.; Hartley, J.; Hill, T. M.; Munroe, D.J. *PLoS ONE* **2008**, *3(9)*, e3265.

[20] Ramachandran, N.; Hainsworth, E.; Bhullar, B.; Eisenstein, S.; Rosen, B.; Lau, A.Y.; Walter, J. C.; LaBaer, J. *Science* **2004**, *305*, 86-90.

[21] Wang, X.; Yu, J.; Sreekumar, A.; Vrambally, S.; Shen, R.; Giacherio, D.; Mehra, R.; Montie, J. E.; Pienta, K. J.; Sanda, M. G.; Kantoff, P. W.; Rubin, M. A.; Wei, J. T.; Ghosh, D.; Chinnaiyan, A. M. *New Eng. J. Med*. **2005**, *353*, 1224-1235.

[22] Sioud, M.; Hansen, M. H. *Eur. J. Immunol*. **2001**, *31*, 716-725.

[23] Melhem, R.; Hailat, N.; Kuick, R.; Hanash, S. M., *Leukemia* **1997**, *11*, 1690-1695,

[24] Chong, B. E.; Hamler, R.L.; Lubman, D. M.; Ethier, S. P.; Rosenspire, A. J.; Miller, F. R. *Anal. Chem.* **2001**, *73(6)*, 1219-1227.

[25] Minamoto, T.; Buschmann, T.; Habelhah, H.; Matusevich, E.; Tahara, H.; Boerresen-

Dale, A.; Harris, C.; Sidransky, D.; Ronai, Z. *Oncogene* **2001**, *20*, 3341-3347.

[26] Winter, S. F.; Minna, J. D.; Johnson, B. E.; Takahashi, T.; Gazdar, A. F.; Carbone, D. P. *Cancer Research.* **1992**, *52*, 4168-4174.

[27] Gottlieb, T. M.; Oren, M. *Biochimica et Biophysica Acta* **1996**, *1287*, 77-102.

[28] Krueger, K. E.; Srivastava, S. *Molecular & Cellular Proteomics* **2006**, *5(10)*, 1799-1810.

[29] Orntoft, T. F.; Vestergaard, E. M. *Electrophoresis* **1999**, *20*, 362-371.

[30] Hakomori, S.-I. *Cancer Research* **1996**, *56*, 5309-5318.

[31] Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477–488.

[32] Hakomori, S. *Proc. Natl. Acad. Sci.* **2002**, *99(16)*, 10231-10233.

[33] Kurien, B. T.; Hensley, K.; Bachmann, M.; Scolfield, R. H. *Free Radical Biology & Medicine* **2006**, *41*, 549-556.

[34] Hong, S. H.; Misek, D. E.; Wang, H.; Puravs, E.; Giordano, T. J.; Greenson, J. K.; Brenner, D. E.; Simeone, D. M.; Logsdon, C. D.; Hanash, S. M. *Cancer Research* **2004**, *64*, 5504-5510.

[35] Kallioniemi, O. P.; Wagner, U.; Kononen, J.; Sauter, G. *Human Molecular Genetics* **2001**, *10(7)*, 657-662.

[36] Lu, H.; Goodell, V.; Disis, M. L. *J. Proteome Res.* **2008**, *7(4)*, 1388-1394.

[37] Yan, F.; Sreekumar, A.; Laxman, B.; Chinnaiyan, A. M.; Lubman, D. M.; Barder, T. J. *Proteomics* **2003**, *3*, 1228-1235.

[38] Hutchens, T. W.; Janson, J.C. ; Ryden, L.(Ed.) *Protein Purification :Principles, High-Resolution Methods, and Application. 2nd Ed.,* John Wiley & Sons, New York, **1998**, 149-174.

[39] Fountoulakis, M.; Langen, H.; Gray, C.; Takacs, B. *J. of Chromatography A* **1998**, *806(2)*, 279-291.

[40] Strong, J. C.; Frey, D. D. *J. Chromatography A* **1997**, *769(2)*, 129-143.

[41] Liu, Y.; Anderson, D. J. *J. Chromatography A* **1997**, *762*, 207-217.

[42] Chong, B.E.; Yan, F.; Lubman, D. M.; Miller, F. R. *Rapid Commun. Mass Spectrometry* **2001**, *15(4)*, 291-296.

[43] Itoh, H.; Nimura, N.; Kiroshita, T.; Nagae, N.; Nomura, M. *Anal. Biochem.* **1991**, *199*, 7-10.

[44] Banks, J. F.; Gulcicek, E. E. *Anal. Chem.* **1997**, *69*, 3973-3978.

[45] Barder, T. J.; Wohlman, P. J.; Thrall, C.; DuBois, P. D. *LC-GC* **1997**, *15*, 918-926.

[46] Hanson, M.; Unger, K. K.; Mant, C. T.; Hodges, R. S. *Trends in Analytical Chemistry* **1996**, *15*, 102-110.

[47] MacBeath, G.. *Nature Genetics Supplement*, **2002**, *32*, 526-532.

[48] Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. *Anal. Chem*. **2003**, *75*, 4646-4658.

[49] MacDonald, J. W.; Ghosh, D. *Bioinformatics* **2006**, *22(23)*, 2950-2951.

[50] Madden M. E.; Sarras M. P. Jr. *Pancreas*, **1988**, *3(5)*, 512-528.

[51] Chen R.; Pan S.; Cooke K.; Moyes K. W.; Bronner M. P.; Goodlett D. R.; Aebersold R.; Brentnall T. A. *Pancreas* **2007**, *34(1)*, 70–79.

[52] Khondoker M. R.; Glasbey C. A.; Worton B. J. *Biometrical Journal*, **2007**, *49*, 815 – 823.

[53] Giles, P. J.; Kipling, D. *Bioinformatics*, **2003**, *19*, 2254–2262.

[54] Tomlines, S. A.; Rhodes, D. R.; Perner, S.; Dhanasekaran, S. M.; Mehra, R.; Sun, X.

W.; Varambally, S.; Cao, X.; Tchinda, J.; Kuefer, R. *Science* **2005**, *310*, 644-648.

[55] Zieker D.; Konigsrainer I.; Traub F.; Nieselt K.; Knapp B.; Schillinger C.; Stirnkorb C.; Fend F.; Northoff H.; Kupka S.; Brucher B. L. D. M.; Konigsrainer A. C*ellular Physiology and Biochemistry* **2008**, *21*, 429-436.

[56] Yeh C. S.; Wang J. Y.; Chung F. Y.; Lee S. C.; Huang M. Y.; Kuo C. W.; Yang M. J.; Lin S. R. *ONCOLOGY REPORTS* **2008**, *19*, 81-91.

[57] Tibishirani R.; Hastie T. *Biostatistics* **2007,** *8(1)*, 2-8.

[58] Wu B. *Biostatistics* **2007**, *8(3)*, 566-575.

[59] Benjamini Y.; Hochberg Y. *Journal of the Royal statistical society, series B (Methodological)* **1995**, *57*, 289-300.

[60] Daly E.B.; Wind, T.; Jiang, X. M.; Sun, L.; Hogg, P. J. *Biochim. Biophys. Acta.* **2004**, *1691(1)*, 17-22.

[61] Shichijo, S.; Azuma, K.; Komatsu, N.; Ito, M.; Maeda Y.; Ishihara, Y.; Itoh, K. *Clinical Cancer Rresearch* **2006**, *l10*, 5828-5836.

[62] Hwang T.-L.; Ying Liang Y.; Chien K.-Y.; Yu, J.-S. *Proteomics* **2006**, *6*, 2259 – 2272.

[63] Sawilowsky, S. S. *Journal of Modern Applied Statistical Methods* **2005**, *4*, 598-600.

[64] Yukitake M.; Sueoka E.; Sueoka-Aragane N.; Sato A.; Ohashi H.; Yakushiji Y.; Saito M.; Osame M.; Izumo S.; Kuroda Y. *Journal of Neurovirology* **2008**, *14*, 130 – 135.

[65] Yan-Sanders Y.; Hammons G. J.; Lyn-Cook B. D. *Cancer Lett*. **2002**, *183(2)*, 215-220.

[66] Ditzel H. J.; Masaki Y.; Nielsen H.; Farnaes L.; Burton D.R *Proc. Natl. Acad. Sci.*

USA **2000**, *97*, 9234-9239.

[67] Porkka K.; Saramaki O.; Tanner M.; Visakorpi T. *Laboratory Investigation* **2002**, *82*, 629-637.

[68] Tibshirani R.; Hastie T.; Narasimhan B.; Chu G. *Proc. Natl. Acad. Sci.* **2002,** *99*, 6567-6572.

[69] Zhu Y.; Lin H.; Li Z.; Wang M.; Luo J. *Breast Cancer Res. Treat.* **2001**, *69(1)*, 29-38.

[70] Huanga C. J.; Chien C. C.; Yang S. H.; Chang C. C.; Sun H. L.; Cheng Y. C.; Liu C. C.; Lin S. C.; Lin C. M. *J. Cell. Mol. Med*. **2008,** *12(5B)*, 1936-1943.

[71] Vishwanatha J. K.; Chiang Y.; Kumble K. D.; Hollingsworth M. A.; Pour P. M. *Carcinogenesis* **1993**, *14(12)*, 2575-2579.

[72] Chen R.; Brentnall T. A.; Pan S.; Cooke K.; Moyes K. W.; Lane Z.; Crispin D. A.; Goodlett D. R.; Aebersold R.; Bronner M. P. *Molecular & Cellular Proteomics* **2007**, *6(8)*, 1331-1342.

[73] Lo A.S.; Liew C. T.; Ngai S. M.; Tsui S. K.; Fung K. P.; Lee C. Y.; Waye M. M. *J. Cell. Biochem.* **2005**, *94*, 763-773.

[74] Baris O.; Savagner F.; Nasser V.; Loriod B.; Granjeaud S.; Guyetant S.; Franc B.; Rodien P.; Rohmer V.; Bertucci F.; Birnbaum D.; Malthièry Y.; Reynier P.; Houlgatte R. *J. Clin. Endocrinol. Metab.* **2004**, *89(2)*, 994-1005.

Chapter 5

Analysis of protein interaction changes based on genetic defects

5.1. Introduction

Ovarian cancer is a leading cause of cancer-related death among women in the United States [1]. More than 90% of ovarian cancers are thought to arise from the ovarian surface epithelium. Ovarian tumors are classified, based upon histology, as 4 different subtypes: serous, clear, endometrioid, and mucinous [2]. Of these subtypes of ovarian tumors, 10-30% of the tumors are classified as endometrioid ovarian cancer (OEA). OEA often presents similarly to serous ovarian cancer in early stage tumors. However, genetic studies have shown that K-RAS and CTNNB1 (encoding β-catenin) gene mutations occur more frequently in endometrioid and mucinous ovarian cancers [3-6]. Mutation of genes can impact the many signaling pathways involved in cell death, survival, and tumor progression [7].

One obstacle in studying low stage ovarian cancer is the difficulty of obtaining

sufficient amounts of tissue for research. As ovarian cancer is insidious in onset, less than 30% of ovarian cancers are detected at a low stage. Thus, most ovarian cancer presents as high stage tumors with correspondingly high mortality rates. Improvements in the early detection of tumors while they are at low stage, and the development of novel therapies for ovarian carcinomas are crucial to understanding this disease. As such, an alternative method has been developed to facilitate low stage cancer. Genetically engineered mouse (GEM) models of each subtype of ovarian cancer provide access to low stage tumors which faithfully recapitulate human tumors of the same sub-type. Another advantage of using GEM models is increased sample availability. Some studies have used GEM model of serous ovarian tumors [8, 9] and also a GEM model with depletion /activation of K-RAS and/or Pten genes in ovarian cancer [10]. Wu et al have studied gene expression in endometrioid ovarian cancer using high density oligonucleotide arrays [11] and have developed a GEM model of OEA that faithfully recapitulates human OEA [12].

Results from these studies demonstrated that 16-38% of endometrioid ovarian cancer have mutation of the CTNNB gene, rare in other types of ovarian cancer, and that deregulation of PI3K/Pten signaling based on Pten inactivation and/or PIK3CA mutation is significantly associated with Wnt (Wnt/β-catenin) signaling defects. This study also

demonstrated that deregulation of two signaling pathways in murine ovarian surface epithelium results in similar morphology and gene expression to human endometrioid ovarian cancer, especially in early stages [12]. Although gene studies have been very successful, and comprehensive studies of tumor RNA and DNA have provided a number of insights into ovarian cancer pathogenesis, it is known that proteins are the major effector molecules in tumor cells. Protein levels may be discordant with corresponding transcript levels, and posttranslational modifications (PTMs) can have biologically critical effects on protein function.

Within this study, we have performed proteomics analysis of GEM models of OEAs. Two groups of mouse tumor samples were utilized: Group 1 tumors had defects in the PI3K/Pten and Wnt/β-catenin pathway due to the conditional inactivation of tumor suppressor genes Pten and Apc. Group 2 tumors had, in addition to the conditional inactivation of tumor suppressor genes Pten and Apc, a p53 mutation. Four ovarian tumors from the Group 1 mice and three ovarian tumors from the Group 2 mice were analyzed for differences in protein expression, using a combined method of cIEF followed by nano/LC-MS/MS. The results obtained from nano/LC-MS/MS were semi-quantitated by a label free spectral counting method, and the characterization of proteins detected from each mice group was performed using the Ingenuity Pathway Analysis

(IPA) program. In this work, proteins in mouse tumors with signaling pathway defects and/or mutations of tumor suppressor genes were identified and the networks associated with these signaling pathways provided biological insight into cellular perturbations resulting from the genetic defects.

5.2 Experimental section

5.2.1 Sample preparation

We have analyzed seven GEM OEA tumor tissue samples: four mouse tumor samples have both a Pten and Apc genetic defect and three mouse tumors have the Pten and Apc defects, as well as a p53 mutation [12]. The tumors were histologically analyzed by a board-certified pathologist (K. R. C.) prior to utilization in this study. The mouse tumor tissues were solubilized in lysis buffer using a mini-bead beater (Biospec, Bartlesville, OK). The lysis buffer was composed of 7 M urea (Sigma–Aldrich, St. Louis, MO), 2 M thiourea (Sigma–Aldrich), 100 mM DTT (Sigma–Aldrich), 0.2% *n*-octyl-β-D-glucopyranoside (OG; Sigma–Aldrich), pH 3–10 (Bio-Rad, Hercules, CA), 10% glycerol (Sigma–Aldrich), and 2% v/v of 50 x diluted protease inhibitor cocktail solution. The homogenized tissue samples were incubated at room temperature for 30 min, and then centrifuged (35,000 x *g*, 1 h, 4°C) to pellet the insoluble material. The proteins were

quantified with the micro-BCA assay kit from Pierce (Rockford, IL).

5.2.2 Trypsin Digestion

5mM dithiothreitol (DTT) was added and the mixture was incubated at 60°C for 30 min. After cooling, 5mM iodoacetamide (IAA) was added and the mixture was placed in the dark at room temperature for 30 min in order to carboxamidomethylate the cysteine residues. Then 1:50 w/v TPCK-treated trypsin (Promega, Madison, WI) was added to each sample, the samples were vortexed, and then incubated at 37°C for 18 h. The tryptic digestion was terminated by the addition of 2.5% v/v TFA.

5.2.3 cIEF separation.

cIEF was performed using a ProteomeLab PA 800 (Beckman, CA). A 70 cm cIEF (100 μm i.d. 365 μm o.d.) capillary was coated with hydropropyl cellulose for eliminating electroosmotic flow and absorption of peptides onto the capillary wall. Sodium hydroxide solution at pH 10.8 and 0.1 M phosphate acid solution were employed as catholyte and anolyte. One end of the capillary was emerged in the anolyte, while the other end was kept in coaxial metal tubing with a sheath flow composed of catholyte eluting flush with the exit of the capillary. The capillary was initially filled with sample gel buffer

containing 2% ampholyte 3-10 and 10 μg tryptic peptides. The flow rate was controlled by a syringe pump at 2 μl/min, and was adjusted to ensure that a proper droplet formed at the exit to carry the peptides fractionated into individual wells in the sample plate. Isoelectric focusing was performed at 21 kV (300 V/cm) over the entire capillary. The current decreased continuously as the peptides were focused and the process was considered complete when the current stabilized. The focused bands of peptides were sequentially mobilized slowly under pressure towards the cathode and delivered as droplets with catholyte sheath flow into individual wells on a sample plate, where the fractions were collected with a modified Beckman HPLC sample collector. Each cIEF separation was run at least twice for reproducibility and takes approximately 90 min for each run.

5.2.4 Nano-RPLC-ESI-MS/MS

Fractionated peptide solutions from cIEF separation were analyzed by nano-flow reverse phase LC/MS/MS using an LTQ mass spectrometer equipped with a nano spray ESI source (Thermo, San Jose, CA). The samples were injected into a Paradigm AS1 micropump (Michrom Biosciences, Auburn, CA) via Paradigm autosampler (Michrom Biosciences), connected to capillary reverse phase column (0.1 mm×150 mm, Michrom

Biosciences). The samples were separated using a (0.1mm×150mm) capillary reverse phase column (Michrom Bioresources, Auburn, CA) with a flow rate of 5 µl/min. An acetonitrile:water gradient method was used. Both solvent A (water) and B (acetonitrile (ACN)) contained 0.1% formic acid and 2% ACN/water in water (solvent A), ACN (solvent B) respectively. The linear gradient for separation was as follows: from 3% ACN to 12% ACN in 5 min, from 12% ACN to 40% ACN in 30 min, from 40% ACN to 80% ACN in 15 min, and decreased from 80% ACN to 3% ACN in 10 min. The electrospray voltage was 2.4 V, with a capillary temperature of 200° C and a capillary voltage of 4 kV. The normalized collision energy was set at 35% for MS/MS.

5.2.5 Identification, Quantification, and Bioinformatic Analysis of Proteins.

MS/MS spectra obtained were analyzed using the Sequest feature of Bioworks 3.1 SR1 (Thermo Finnigan). Peptide ions were searched in mouse UniProt FASTA database using the following parameters: (1) Enzyme: trypsin; (2) two missed cleavage allowed; (3) peptide ion mass tolerance: 1.5 Da; (3) fragment ion mass tolerance 0.0 Da; (4) mass tolerance for precursor ions 1.4 Da. The identified peptides were further validated through the Trans Proteomic Pipeline (TPP) that was modified in house [13]. In the TPP, the search results were first evaluated by Peptide Prophet, ascertaining that peptides

corresponded to the correct spectra using Bayesian statistics. Protein Prophet utilized the Peptides Prophet results to assign corresponding proteins based on probabilities. In this study, a protein probability score of ≥0.90 was used as the threshold for protein identification to ensure that the minimized overall error rate is below 0.1. After identifying proteins using Sequest followed by TPP, a spectral counting method was applied to compare protein expression level differences between the two groups of mouse tumor tissues. The relative protein abundance fold change was calculated by the ratio of the spectral count of proteins in two sample groups. In order to normalize the data, we first calculated the ratio of the total spectral count of 3 runs of each sample and then multiplied the spectral count of each protein in the numerator by this ratio. If a protein was found in Group 1 tumors but not in Group 2 tumors, then the spectral counting number for that protein in group 2 was assigned as zero.

5.2.6 Ingenuity Pathway Analysis

To understand the genetic defects in the GEM OEA tumors, and their effects on proteins involved in tumor progression, the identified tumor proteins were analyzed using Ingenuity Pathway Analysis software (IPA). The IPA program utilizes a knowledge database that was derived from the scientific literature that contains information on

interactions between genes, proteins and other biological molecules, producing networks / comparison of pathways / metabolism with input molecule lists [14]. Two sets of protein lists, one for the Group 1 tumors (with the Pten and Apc defects) and the other for the Group 2 tumors (with Pten and Apc defects and p53 mutations), were uploaded into IPA analysis software. Spectral count numbers for listed proteins were also uploaded to compare expression level changes. Any change in a canonical pathway was carefully analyzed as an example of OEA progression based upon a genetic defect. The Fisher's exact test p value was used for calculating the significance values for analyses of network and pathway generation.

5.3 Results and Discussion.

Our work presents the study of proteomics in the mouse model of ovarian endometrioid adenocarcinomas (OEAs) with defects in either the Pten and Apc genes or Pten, Apc and p53 genes using cIEF-LC/MS/MS (Figure 5.1, and 5.2). Seven ovarian endometrioid carcinomas (OEAs) mouse tumor samples - four tumors with defects in the Wnt/β catenin and PI3K/Pten signaling pathway (Pten and Apc genes; Group 1 tumors) and three mice tumors with defects in the Wnt/β-catenin, PI3K/Pten signaling pathway and p53 mutation (Pten, Apc *and* p53 genes; designated Group 2 tumors) - were analyzed

in the initial training set using a shotgun method, followed by bioinformatics analysis using the Ingenuity Pathway Analysis (IPA). The mice used in this study were genetically engineered to possess defects in the signaling pathways in order to study morphology and biological behavior of OEA. These tumors faithfully recapitulated human ovarian endometrioid carcinomas with the same signaling pathway defects.

Mice with genetic defects on Wnt/β catenin and PI3K/Pten signaling pathway.

Four mouse tumor samples (Group 1) with genetic defects in Wnt/β-catenin and PI3K/Pten signaling pathways were analyzed. A total of 1025 proteins were detected from the Group 1 mice and were categorically assigned into subcellular compartments. The majority of the identified proteins are normally localized in the cytoplasm and nucleus (Figure 5.3). The proteins normally localized in the plasma membranes and extracellular spaces occupy 8-10% of each total proteome. The location of 20% of the detected proteins was unknown. In some cases proteins may have been located in multiple subcellular compartments. The top 5 associated network functions generated from the IPA for both tumor groups are presented in Table 5.1.

The defects in the Wnt/β-catenin signaling pathway in Group 1 mice were caused by conditional inactivation of the Adenomatous Polyposis Coli (Apc) tumor

suppressor gene. Apc gene is a key tumor suppressor gene, and the mutations in this gene are often found in other types of tumors [15,16]. Apc is considered to be involved in regulation of biological processes such as cell-cell adhesion, proliferation, and differentiation. The Apc protein is known to be a part of a multi-protein complex that negatively regulates the Wnt signaling pathway by promoting the proteasomal degradation of beta-catenin [17]. The Apc protein also interacts with microtubules, a part of the cytoskeleton network, and associates with the plasma membrane in an actin-dependent manner [18].

In this study, the actin cytoskeleton was shown to be one of the most important compartments in the Group 1 mouse tumors. The Group 1 tumors possess increased expression of proteins related to this compartment, as shown by IPA (Ingenuity Pathway Analysis). Increased actin and myosin were found, as well as IQGAP1 (IQ motif containing GTPase-activating-like protein 1), which binds to activated CDC42, and which plays a role in promoting polarized reorganization of the microtubule cytoskeleton [19]. The IQGAP1 may act as a connector between Apc and F-actin [20], playing an important role in cell migration. It shows a twelve fold increased expression in Group 1 mouse tumor samples compared to that in Group 2 tumor samples.

Detection of increased expression of proteins involved in cell migration could help elucidate mechanisms of tumor metastasis. Some studies suggested that depletion of Apc partly affects the distribution of F-actin and ZO-1 (tight junction protein 1) in cancer [21]. ZO-1 protein may bind to tropomyosin and the actins (such as G-actin, and F-actin), and interacting with Myosin, in Group 1 tumors, though distribution of the proteins, ZO-1 and F-actin could not be tested further. Also, ZO-1 was shown to interact with spectrin, alpha, non-erythrocytic 1 (SPTAN1). The SPTAN1 protein was one of the proteins detected in apoptotic signaling in Group 1 mice, and is known as a protein that undergoes cleavage during cell death, and may play an important role in membrane stability and cytoskeletal integrity [21].

In Group 1 tumor samples, there is another tumor suppressor gene (Pten), the inactivation of which causes deregulation of the PI3K/Pten signaling pathway. Pten (Phosphatase and Tensin Homolog deleted on chromosome Ten) was discovered relatively recently, yet is known to play important roles in cellular migration, apoptosis and embryonic development [22-24]. Pten mutations have been found in several types of tumors [25, 26] including OEA's [27]. Some studies suggested that Pten may associate with focal adhesion kinase (FAK) activity [28], and also be involved in activation of PI3K [29] and Akt [30]. In the Group 1 tumors, proteins involved in FAK signaling were

not detected, however PI3K was detected (Figure 5.5), suggesting involvement of calmodulin, CaMKII (calmodulin-dependent protein kinase II), LDL (beta lipoprotein), Rho (Ras homolog) and LASP1 (LIM and SH3 protein 1). Interestingly, most of the above mentioned proteins act indirectly on the PI3K protein, resulting in modulation of the LASP1 protein. This finding is concordant with the results of a study in a breast cancer cell line [31].

Akt proteins in the Group 1 tumor group showed interaction with Hsp90 (heat shock protein 90KDa), and CDC 37 (cell division cycle 37 homolog) using IPA analysis (Figure 5.5). The results from a recent study illustrated that Akt binds with the Hsp90 protein and CDC37, forming a complex, which is destabilized by the Hsp90 inhibitor function [32]. Akt failed to bind to the Cdc37 chaperone in cells expressing NPM-ALK, which also correlates with increased Akt stability [33]. However, the result from our IPA suggested that the Akt protein may be inhibited by CRBP1 (RBP1, retinol binding protein 1, cellular) which interacts with transthyretin (TTR, pre-albumin, amyloidosis type I), but is only detected in Group 1 tumors (data not shown). The expression level of RBP1 in the group 1 mice is almost 25-fold greater than that in the Group 2 tumors. This result may result in lower expression of Akt in Group 1 tumors compared to that in Group 2 tumors (Figure 5.4), although the inhibition of Akt by RBP1 was also found in

the Group 2 tumors. The potential synergistic effects of p53 mutation in Group 2 tumor samples are not clear.

Mice with genetic defects on the Wnt/β catenin and PI3K/Pten signaling pathways and p53 mutations.

Three tumor samples (Group 2 tumors) with inactivation of the Pten and Apc tumor suppressor genes and the p53 mutation were analyzed. 1394 proteins were detected. The distribution of subcellular compartmentalization in Group 2 tumor samples was similar to that found in Group 1 tumors (Figure 5.3). There are increased numbers of proteins assigned to the nucleus and plasma membrane compartments, although the amount of increase is not significant.

The results from a recent study suggested that the murine OEA tumors that developed in the setting of Apc and Pten inactivation presented nuclear accumulation of the beta-catenin protein [12]**.** It has been reported that mutations in Apc caused accelerated Wnt signaling by stabilizing beta-catenin [34]. The amount of accumulated beta-catenin protein in the nucleus in both tumor samples could not be measured in this study. However, in Group 2 tumor samples, the Wnt protein was found, showing indirect interaction with MYH4 (myosin, heavy chain4), CK1 (Casein Kinase 1) and CTNNB1

(beta catenin). Notably, there was an association found between MYH4 and IQGAP1 (IQmotif containing GTPase activating protein1) which interacts with CTNNB1 (Catenin:cadherin-associated protein, beta, 88 KDa). This association was not detected in Group 1 tumors, although the IQGAP1 showed lower expression in Group 2 tumors than was found in the Group 1 tumors.

The genetic difference between the two tumor groups is the existence of the p53 mutation. All of the Group 2 mice have a mutated p53 gene, in addition to inactivation of the Pten and Apc genes. TP53 is a well known tumor suppressor gene, involved in the DNA repair process, in the cell cycle, and in apoptosis. At the protein level, the results from the current study suggested that p53 has an association with DAPK1 (death associated protein kinase 1), with MAPK 3 and MAPK 9 proteins, and with CSNK1A1 (casein kinase I, alpha1) using IPA. The DAPK1 protein has been shown to play a role in stabilizing p53 and apoptosis, while the MAPK family proteins are known to act in response to membrane damage and oxidative stress [35]. With respect to apoptosis in the p53 mutated tumors [36], the well-known mdm2 protein was not found in Group 2 mice, suggesting that mutated p53 may be altered in conformation, thereby inducing anti-apoptosis [37], although CSNK1A1 (casein kinase I, alpha1), and creatine kinase were shown to bind p53, suggesting involvement of the Wnt/beta-catenin pathway. Also,

oncostatin M (OSM) protein is known to induce growth arrest and differentiation of cells. OSM was found in Group 2 tumors, suggesting direct involvement with p53 and CTNNB1 in three mouse tumor samples. Another protein, LGALS3 (lectin, galactoside-biding, soluble, 3) also showed a similar relationship to that observed between p53 and CTNNB1 except that p53 seemed to repress LGALS3 expression (Figure 5.6).

Akt (serine-threonine kinase) protein promotes cell survival by protein phosphorylation, in opposition to p53 which induces cell death when damage of DNA is not repairable. Akt and p53 are known to link to the positive feedback loop between them through Pten and Mdm2 [38, 39]. The feedback between p53 and Akt involves PIP3 (phosphatidylinositol-3,4,5-trisphosphate) and Pten (phosphatase and tensin homolog). There are two theories about this feedback. One is that Pten expression causes p53 indirectly inhibiting production of PIP3; the other is that activity of a subunit of PI3K (phosphatidylinositol 3-kinase) catalyzes the formation of PIP3 [40]. In the Group 2 tumor samples, the cell death-survival association between p53 and Akt was not shown clearly at the protein level. Since crosstalk between these two proteins coexists in two stable pathways and also involves other pathways, it is more complex than what can be ascertained from our study. However, as shown in Figure 5.6, p53 can act independently of PI3K proteins, but not the other way around, in a reaction with mdm2, which is

consistent with the well known regulatory network. Other proteins like MAPK9 (mitogen-activated protein kinase 9) directly interacts on TP53 through several protein interactions from PI3K to MAPK9. For example, p53 exerts effects upon Akt indirectly and also by direct action on PRKDC (protein kinase, DNA-activated, catalytic polypeptide) from TP53 to Akt; TP53 interacts with DAPK1 which, in turn, interacts with CDC37 (cell division cycle 37 homolog); CDC37 interacts with PIK3CA, which directly interacts with Akt. The expression level of CDC37 protein showed no difference between both tumor groups; however the protein-protein interactions were quite different, presumably resulting in changes in biological behavior based on tumor progression.

The comparison of profiling of proteins in two mice groups

The difference between the two mouse tumor groups is that the Group 2 tumors have a p53 mutation in addition to defects in two signaling pathways. Table 5.2 illustrates the list of proteins detected in tumor samples with expression change between the two tumor samples. In general, the Group 1 tumors without p53 mutation presented up-regulated expression of proteins in apoptotic signaling and in the actin cytoskeleton, while showing fewer proteins involved in cellular signaling and metabolism. The Group 2 tumors showed increases in the number of proteins and/or increased expression of

proteins involved in signaling pathways and metabolism, such as estrogen receptor signaling, and glutathione metabolism.

Inflammation-mediated protein expression changes involved proteins from both tumor Groups, a critical part of tumor progression [41]. We have identified seven positive and five negative acute-phase response proteins. Among the negative acute phase response proteins, transferrin showed 5.3 fold increased expression in the Group 2 tumors and alpha-2-HS-glycoprotein demonstrated approximately a 2-fold increase. Interestingly, for apolipoprotein A1, 2 proteins show similar expression levels in both tumor groups, although the expression level of these proteins were not compared with those in regular mice without tumors or in OEA tumors with other genetic defects. However, for positive acute–phase response, three proteins presented higher expression in Group 1 tumors and four proteins showed higher expression in Group 2 tumors. These results may suggest that 3 isoforms of Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1, could be new target molecules to measure as early detection biomarkers for OEA, as these proteins showed 3-fold increases in expression in Group 1 tumors, which are considered to be low stage tumors [12].

Among other identified proteins in the mouse OEA tumors, heat shock protein 27 kDa, heat shock 70K protein 5, 8, 9, and heat shock 90 kDa alpha protein showed

high expression in Group 1 tumors. Heat shock 27 kDa protein 1 and heat shock protein 90 kDa alpha (cytosolic), class A member 1 showed approximately 13-fold and 27-fold increased expression, respectively, over that in Group 2 samples. Heat shock proteins have been shown to be over expressed in various human cancers, implying cellular stress, tumor cell proliferation and cell death [42]. Heat shock proteins 27k, 70k, 90kDa are associated with several oncogenes and may have utility in determining prognosis of cancer [43-45].

When heat shock proteins were involved with increased aryl hydrocarbon receptor signaling in Group 1 tumors, FAS (fatty acid synthase) protein was also identified, suggesting involvement in aryl hydrocarbon receptor signaling as well as fatty acid biosynthesis in Group 1 tumors. FAS protein is known to be involved with functioning estrogen and progesterone receptors, but the mechanism of this protein's function is not clear. The expression of FAS protein increased 7-fold in Group 1 tumors, as compared to that found in Group 2 tumors, and has been demonstrated to be up-regulated in breast and prostate cancer cell lines [46, 47]. Interestingly, FAS protein was reported to have an inverse relationship with Pten expression in prostate cancer[48]. Additionally, most of the proteins involved in PI3K/Pten signaling pathways showed

lower expression in the Group 1 tumors in this study, suggesting that mutated p53 may play a role in their expression in endometrioid ovarian cancer.

5.4 Conclusion

Gene expression studies using GEM tumors have yielded limited results at the molecular level in endometrioid ovarian cancer. We have demonstrated that the profiling of proteins based on defects in signaling pathways and gene mutation in two groups of tumor samples has facilitated the study of endometrioid ovarian cancer. Mouse tumors with the p53 mutation, in addition to defects on Wnt/β-catenin and PI3K/Pten signaling pathways, has revealed an increased number of detected proteins involved in signaling, and up-regulated expression of proteins in many signaling and metabolic pathways. p53 appears to play an important role with regards to tumor progression. Also, a number of proteins detected in tumors without mutant p53 showed increased expression levels, suggesting the possibility of improved early detection of endometrioid ovarian cancer.

Table 5.1 Top five listed associated network functions shown in IPA using proteins from two mice tumor groups.

	Group 1 mice tumor samples	**Group 2 mice tumor samples**
1	Cellular Function and maintenance	RNA Post-Transcriptional Modification,Cancer, Cellular Growth and Proliferatoin
2	Lipid Metabolism, Small Molecule Biochemistry, Cancer	Molecular Transport, RAN Trafficking, Protein Trafficking
3	Small Molecule Biochemsitry, Hematological Disease,	Cellular Assembly and Organization, Cancer, Cellular Growth and Proliferation
4	Cancer, gastrointestinal Disease, Energy Production	Skeletal and Muscular Disorders, Neurological Disease, Cellular Assembly and Organization
5	RNA Post-Transcriptional Modification	RNA Damage and Repair, Molecular Transport, Nucleic Aicd Metabolism

Table 5.2. Highly expressed protein lists (partial) in each mice tumor groups.

			fold changes			
Symbol	**Access number**	**Name**	**group 1**	**group 2**	**Location**	**Type**
Q99KI0	ACO2	aconitase 2, mitochondrial	53.6		Cytoplasm	enzyme
Q8VDN3	AHNAK	AHNAK nucleoprotein		6.8	Nucleus	other
Q3UEK9	AHSG	alpha-2-HS-glycoprotein		1.9	Extracellular Space	other
Q1L6K5	AIFM1	apoptosis-inducing factor, mitochondrion-associated, 1	3.153		Cytoplasm	enzyme
Q3TP05	AMOTL2	angiomotin like 2	19.707		Plasma Membrane	other
P48036	ANXA5	annexin A5		4.4	Plasma Membrane	other
Q99PT1	ARHGDIA	Rho GDP dissociation inhibitor (GDI) alpha		13	Cytoplasm	other
Q207D2	C3	complement component 3		1.8	Extracellular Space	peptidase
Q3UKW2	CALM1	calmodulin 1 (phosphorylase kinase, delta)	6.3		Plasma Membrane	other
P24452	CAPG	capping protein (actin filament), gelsolin-like	6.3		Nucleus	other
Q3TEY4	CAPRIN1	cell cycle associated protein 1	9.5		Plasma Membrane	other
Q544Y7	CFL1	cofilin 1 (non-muscle)		59	Nucleus	other
Q3UHW9	CFL2	cofilin 2 (muscle)	2.7		Nucleus	other
Q8BMK4	CKAP4	cytoskeleton-associated protein 4	6.3		Cytoplasm	other
Q60847	COL12A1	collagen, type XII, alpha 1		3.5	Extracellular Space	other
P12787	COX5A	cytochrome c oxidase subunit Va	21		Cytoplasm	enzyme
Q4FJX4	CSRP1	cysteine and glycine-rich protein 1		15	Nucleus	other
Q6NV50	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa		2.5	Plasma Membrane	other
Q4FJX4	CSRP1	cysteine and glycine-rich protein 1		15	Nucleus	other
Q6NV50	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa		2.5	Plasma Membrane	other
O89060	FASN	fatty acid synthase	7.4		Cytoplasm	enzyme
Q80UL3	GALK1	galactokinase 1	9.5		Cytoplasm	kinase
P26443	GLUD1	glutamate dehydrogenase 1	5.3		Cytoplasm	enzyme

Table 5.2. (Continued).

Symbol	Access number	Name	fold changes group 1	fold changes group 2	Location	Type
Q5FB19	HNRPA3	heterogeneous nuclear ribonucleoprotein A3		2	Nucleus	other
Q544Z3	HNRPAB	heterogeneous nuclear ribonucleoprotein A/B		18	Nucleus	enzyme
Q60668	HNRPD	heterogeneous nuclear ribonucleoprotein D		26	Nucleus	transcription regulator
Q80YC2	HSP90AB1	heat shock protein 90kDa alpha (cytosolic), class B member 1	5		Cytoplasm	other
Q3U6V3	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	5.2		Cytoplasm	other
P54071	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial		2.9	Cytoplasm	enzyme
Q8CGH5	IQGAP1	IQ motif containing GTPase activating protein 1	12.6		Cytoplasm	other
Q6PHM1	LAMC1	laminin, gamma 1 (formerly LAMB2)		2.9	Extracellular Space	other
P16045	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1)		21.7	Extracellular Space	other
P99027	RPLP2	ribosomal protein, large, P2		12.3	Cytoplasm	other
P07759	SERPINA3K	serine (or cysteine) peptidase inhibitor, clade A, member 3K		11.1	Extracellular Space	other
Q6A0C6	TAGLN2	transgelin 2		4.3	Cytoplasm	other
Q921I1	TF	transferrin		5.4	Extracellular Space	transporter
Q922B6	TRAF7	TNF receptor-associated factor 7		4.8	Cytoplasm	enzyme
P10639	TXN	thioredoxin		53	Cytoplasm	enzyme
Q3TFD9	VIM	vimentin		1.7	Cytoplasm	other
P10649	GSTM1	glutathione S-transferase, mu 1	not detected	detected	Unknown	other
P48774	GSTM3	glutathione S-transferase M3 (brain)	not detected	detected	Cytoplasm	enzyme
P19157	GSTP1	glutathione S-transferase pi 1	not detected	detected	Cytoplasm	enzyme
Q8C7W3	FUK	fucokinase	not detected	detected	Unknown	kinase
Q3V2Z4	ANXA6	annexin A6	not detected	detected	Plasma Membrane	other
P19091	AR	androgen receptor (dihydrotestosterone receptor)	not detected	detected	Nucleus	ligand-dependent nuclear receptor
Q99LL6	COL1A1	collagen, type I, alpha 1	not detected	detected	Extracellular Space	other
Q3TUE2	COL1A2	collagen, type I, alpha 2	not detected	detected	Extracellular Space	other
O70251	EEF1B2	eukaryotic translation elongation factor 1 beta 2	not detected	detected	Cytoplasm	translation regulator

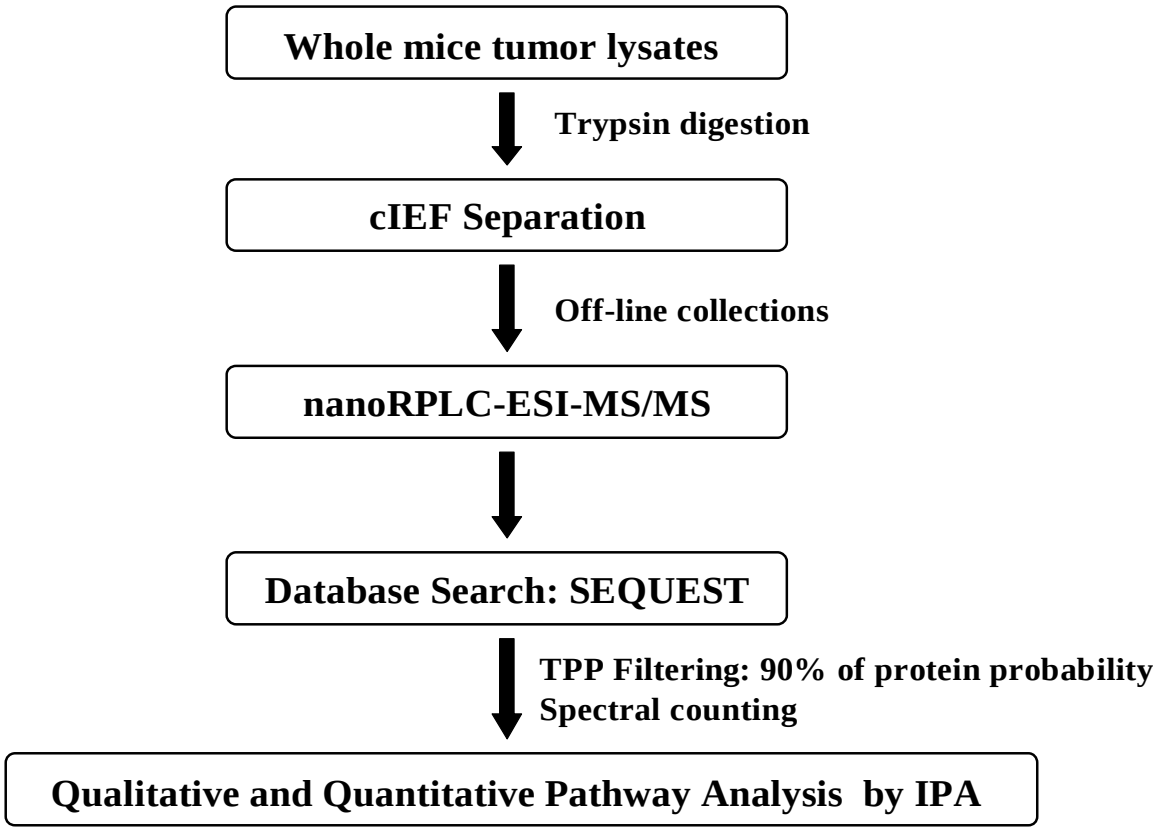


Figure 5.1. Experimental flow chart

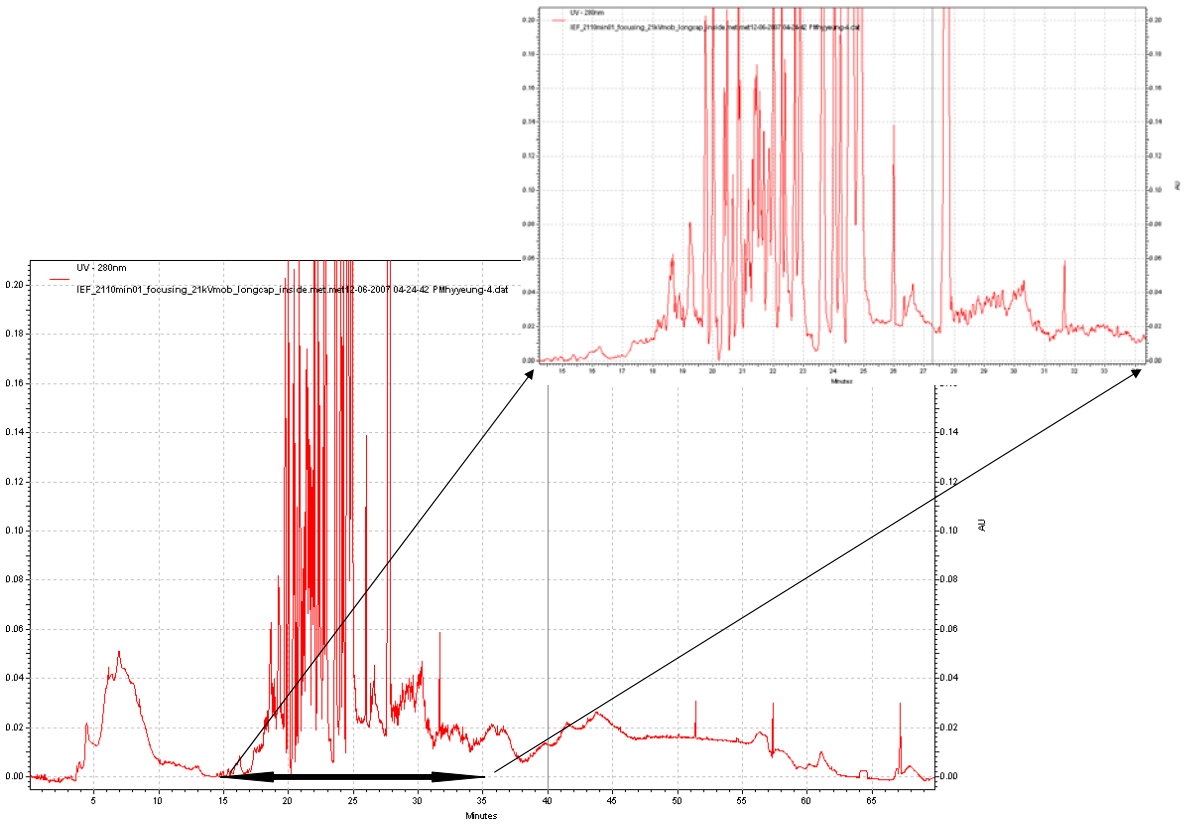


Figure 5.2. The chromatogram of cIEF separation.

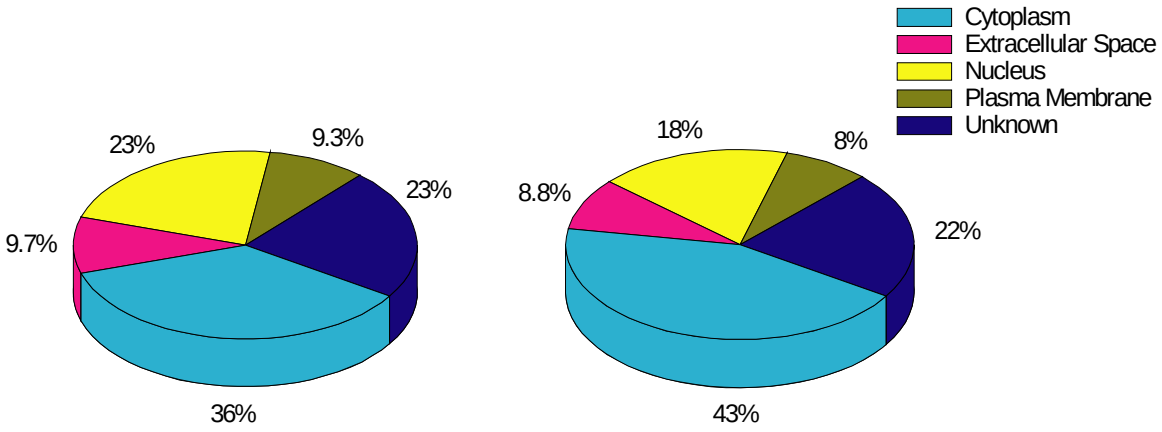


Figure 5.3. The comparison of cellular composition of identified proteins from two mice tumor samples. (Left: group 2 tumors, Right: group 1 tumors)

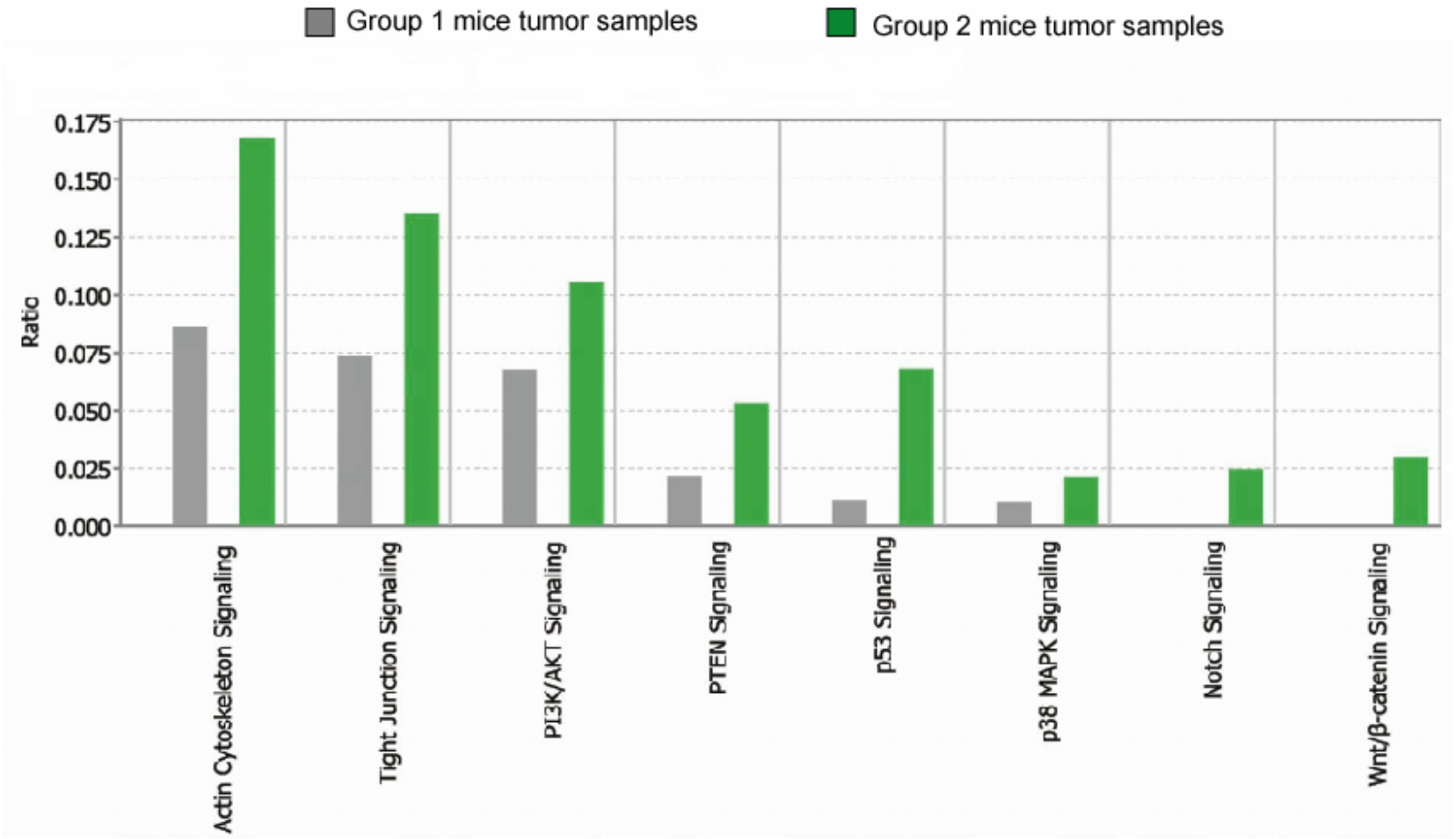


Figure 5.4. Comparison of canonical signaling pathways between two mice tumor groups.

Figure 5.5 The pathway network generated from integrative analysis using proteins from group 1 mice tumors.

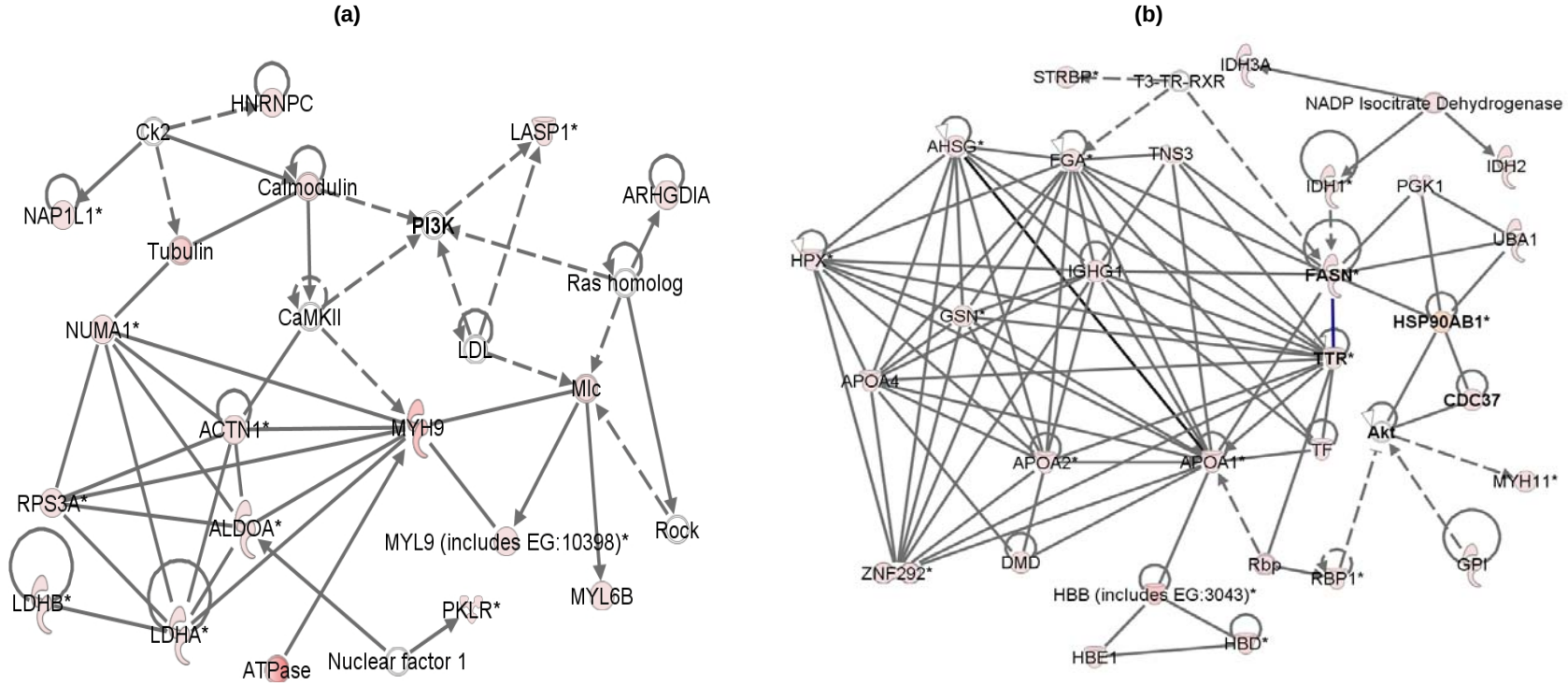
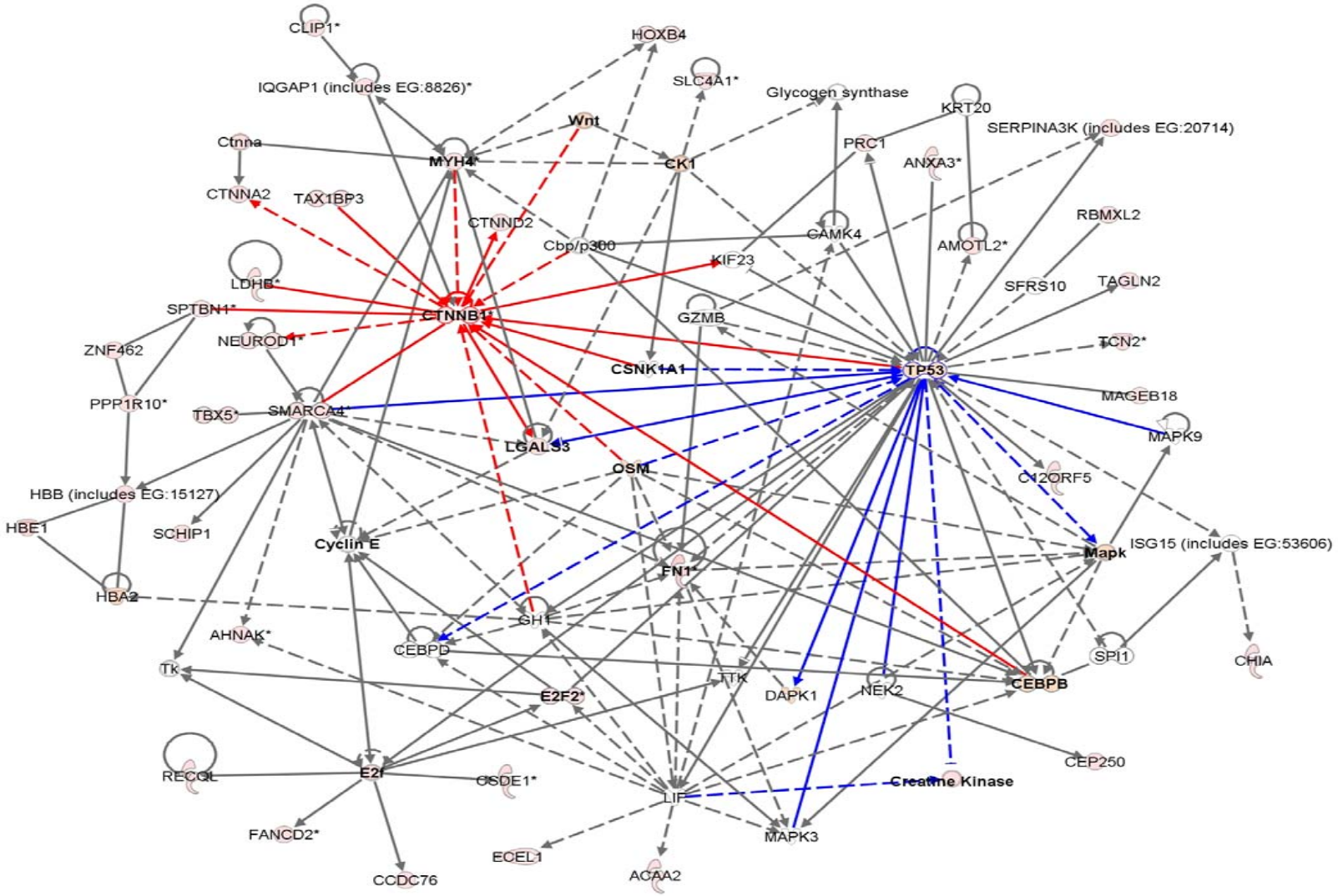


Figure 5.6. The pathway network generated from integrative analysis using proteins from group 2 mice tumors



5.5 References

[1] Jemal, A.; Murray, T.; Ward, E.; Samuels, A.; Tiwari, R. C., et. al., *Cancer Statistics*. Ca-A *Cancer Journal for Clinicians* **2006**, *55,* 10-30.

[2] Bell, D. A., *Mod. Pathol*. **2005**, *18*, Suppl 2, S19-32.

[3] Ichikawa, Y.; Nishida, M.; Suzuki, H.; Yoshida, S.; Tsunoda, H., Kubo, T.; Uchida, K.; Miwa, M. *Cancer Research* **1994**, *54*, 33-35.

[4] Cuatrecasas, M.; Villanueva, A.; Matias-Guiu, X.; Prat, J. *Cancer*, **1997**, 79, 1581-1586.

[5] Sagae, S.; Kobayashi, K.; Nishioka, Y.; Sugimura, M.; Ishioka, S., Nagata, M.; Terasawa, K.; Tokino, T.; Kudo, R. *Jpn. J. Cancer Res*., **1999**, 90, 510-515.

[6] Wright, K.; Wilson, P.; Morland, S.; Campbell, I.; Walsh, M.; Hurst, T., Ward, B.; Cummings, M.; Chevenix-Trench, G. *Int. J. Cancer*, **1999**, 82, 625-629

[7] Vogelstein, B.; Kinzler, K. W. *Nature Medicine*. **2004**, 10, 789-799.

[8] Flesken-Nikitin A.; Choi, K. C.; Eng, J. P.; Shimidt, E. N.; Nikitin, A. Y. *Cancer Research*, **2003**, 63, 3459-3463.

[9] Orsulic, S.; Li, Y.; Soslow, R. A.; Vitale-Cross, L. A.; Varmus, H. E. *Cancer Cell* **2002**, *1*, 53-62.

[10] Dinulescu, D. M.; Ince, T. A.; Quade, B. J.; Shafer, S. A.; Crowley, D.; Jacks, T. *Nat. Med*. **2005**, *11*, 63-70.

[11] Wu. R.; Zhai, Y.; Fearon, E. R.; Cho, K. R. *Cancer Research*, **2001**, *61*, 8247-8255.

[12] Wu, R.; Hendrix-Lucas, N. ; Kuick, R.; Zhai, Y.; Schwartz, D. R.; Akyol, A.; Hanash, S.; Misek, D. E.; Katabuchi, H.; Williams, B. O.; Fearon, E. R.; Cho, K. R. *Cancer Cell* **2007**, *11*, 321-333.

[13] Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. *Anal. Chem*. **2003**, *75*, 4646.

[14] Uriarte, S. M.; Powell, D. W.; Luerman, G. C. ; Merchant, M. L. ; Cummins, T. D. ; Jog, M. R. ; Ward, R. A. ; Mcleish, K. R.. et al, *J. Immunol*. **2008**, *180*, 5575 – 5581.

[15] Polakis, P. *Genes Dev*. **2000**, 14, 1837-1851.

[16] Temburni, M. K.; Rosenberg, M. M.; Pathak, N.; McConnell, R.; Jacob, M. H. *J. Neurosci.* **2004**, *24*, 6776- 6784.

[17] Breitman, M.; Zilberberg, A.; Caspi, M.; Rosin-Arbesfeld, R. *Biochimica et Biophysica Acta* **2008**, *1783*, 1792-1802.

[18] Akiyama, T.; Kawasaki, Y. *Oncogene* **2006**, *25*, 7538–7544.

[19] Schlessinger K.; McManus E. J.; Hall A. *J Cell Biol*. **2007**, *178(3)*, 355-361.

[20] Watanabe, T.; Wang, S.; Noritake J.; Sato, K.; Fukata, M.; Takefuki, M.; Nakagawa, M.; Izumi, N.; Akiyama, T.; Kaibuchi, K. *Dev. Cell* **2004**, *7(6),* 871-883.

[21] Williams, S. T.; Smith, A. N.; Cianci, C. D.; Morrow, J. S.; Brown, T. L. *Apoptosis* **2003**, *8*, 353-361.

[22] Maehama, T.; Dixon, J. E. *J. Biol. Chem*.**1998**, *273*, 13375-13378.

[23] Jaehama, T.; Dixon, J. E. *Trends Cell Biol*. **1999**, *9*, 125-128.

[24] Di Cristofano, A.; Kotsi, P.; Peng, Y. F.; Cordon-Cardo, C.; Pandolfi, P. P. *Nat. Genet*. **1998**, *19*, 348.

[25] Cairns, P.; Okami, K.; Halachmi, S.; Halachmi, N.; Esteller, M.; Herman, J. G.; Jen, J.; Issacs, W. B.; Bova, G. S.; Sidransky, D. *Cancer Research* **1997**, *57*, 4997-5000.

[26] Suzuki, H.; Freije, D.; Nusskern, D. R.; Okami K.; Cairns, P.; Sidransky, D.; Issacs, W. B.; Bova, G. S. *Cancer Research* **1998**, *58*, 204-209.

[27] Risinger, J. I.; Hayes,. A. K.; Berchuck, A.; Barrett, J. C. *Cancer Research* **1997**, *57*,

3935-3940.

[28] Tamura, M.; Gu J.; Matsumoto, K.; Aota, S.; Parsons, R.; Yamada, K. M. *Science* **1998**, *280*, 1614-1617.

[29] Chang, H. W.; Aoki, M.; Fruman, D.; Auger, K. R.; Bellacosa, A.; Tsichlis, P. N.; Canley, L. C.; Roberts, T. M.; Vogt, P. K. *Science* **1997**, *276*, 1848-1850.

[30] Bellacosa, A.; Testa, J. R.; Staal, S. P.; Tsichlis, P. N. *Science* **1991**, *254*, 274-277.

[31] Loughran G.; Huigsloot M.; Kiely P. A.; Smith L. M.; Floyd S.; Ayllon V.; O'Connor R. *Oncogene* **2005**, *24*, 6185–6193.

[32] Basso A. D.; Solit D. B.; Chiosis G.; Giri B.; Tsichlis P.; Rosen N. *J. Biol. Chem*. **2002**, *277*, 39858-39866.

[33] Theodoraki M. A.; Kunjappu M.; Sternberg D. W.; Caplan A. J. *Experimental Cell Research*, **2007**, *313*, 3851-3858.

[34] Perreault N.; Sackett S. D.; Katz J. P.; Furth E. E.; Kaestner K. H. *Genes & Dev*. **2005**, *19*, 311-315.

[35] Dalle-Donne I.; Scaloni, A.; Butterfield, A. (Ed). *Redox Proteomics*, John Wiley & Sons, Inc., New York, **2006**.

[36] Vogelstein B.; Kinzler, K. W. *Nature Medicine* **2004**, *10*, 789-799.

[37] Vakhrusheva O.; Smolka C.; Gajawada P.; Kostin S.; Boettger T.; Kubin T.; Braun T.; Bober, *Circ. Res*. **2008**, *102*, 703-710.

[38] Harris, S. L.; Levine, A. J. *Oncogene*, **2005**, 24, 2899-2908.

[39] Gottlieb, T. M.; Leal, J. F. M.; Seger, R.; Taya, Y.; Oren, M. *Oncogene* **2002**, *21*, 1299-1303.

[40] Wee K. B.; Aguda, B. D. *Biophysical Journal* **2006**, *91*, 857-865.

[41] Coussens, L. M.; Werb, Z. *Nature* **2003**, *422(6932)*, 559.

[42] Ciocca D. R.; Calderwood S. K. *Cell Stress Chaperone.* **2005**, *10(4)*, 263-264.

[43] Geisler J. P.; Tammela J. E.; Manahan K. J.; Geisler H. E.; Miller G. A.; Zhou Z.; Wiemann M. C. *Eur. J. Gynaecol. Oncol*. **2004**, *25(2)*, 165-168.

[44] Elpek G. O.; Karaveli S.; Simşek T.; Keles N.; Aksoy N. H. *APMIS*. **2003**. *111(4)*, 523-530.

[45] Nanbu K.; Konishi I.; Mandai M.; Kuroda H.; Hamid A. A.; Komatsu T.; Mori T. *Cancer Detect Prev*. **1998**, *22(6)*, 549-555.

[46] Chambon M.; Rochefort H.; Vial H. J.; Chalbos D. *J. Steroid Biochem.* **1989**, *33*, 915–922.

[47] Swinnen J. V.; Esquenet M.; Goossens K.; Heyns W.; Verhoeven G.. *Cancer Research* **1997**, *57*, 1086–1090.

[48] Bandyopadhyay S.; Pai S. K.; Watabe M.; Gross S. C.; Hirota, S.; Hosobe, S.; Tsukada, T.; Miura, K.; Saito K.; Markwell, S. J.; Wang, Y.; Huggenvik, J.; Pauza, M. E.; Iiizumi, M.; Watabe K. *Oncogene* **2005**, *24*, 5389–5395.

Chapter 6

The application of newly designed plate to proteomics study using MALDI-QIT-MS.

6.1 Introduction

Matrix Assisted Laser Desorption Ionization (MALDI) was introduced by Hillenkamp, Karas et al.[1-3] and commercialized in the early 1990's. MALDI is a ionization technique that utilizes a short pulsed laser as source, and matrix as a material for desorption of energy from laser to deliver to analyte in the mixture of analyte and matrix, generating analyte ion $[M+H]^+$ ion (where M is the mass of the analyte molecules) [4].

The most common of matrixes are α-cyano-4-hydroxycinnamic acid (α-CHCA) and 2, 5-dihydroxybenzoic acid (DHB). The matrix is generally a low molecular weight molecule that vaporizes easily but not low enough to evaporate during sample preparation. The matrix is acidic to provide a proton for enhancing ionization of analyte and generally an acid such as trifluoroacetic acd (TFA) was added to the matrix solution to facilitate protonation. Those matrixes are functionalized with a polar group to utilize them in aqueous solution.

The most popular mass analyzer for MALDI is the time-of-flight mass spectrometer (TOF-MS). The TOF-MS is simple, inexpensive and a non-scanning instrument. TOF-

MS has the advantages of high speed, wide mass range, and high sensitivity [5] although the main difficulty of using TOF-MS is the operating condition which is under pulsed mode to establish a start time to accomplish time resolution. In order to solve this problem, pulsed orthogonal extraction geometry has been adopted [6-8] and a high pulsed extraction repetition rate (>2000Hz) to achieve a high duty cycle. The coupling of MALDI with TOF-MS has been an ideal technique in proteomics for analyzing biomolecules such as oligopeptides, proteins [9-11].

The quadruple ion trap has been coupled with TOF-MS to perform MS/MS experiment [12]. The major feature of ion trap in mass spec is the ability to trap ion in a three dimensional electrical field. The ion trap which is composed of a ring electrode and end cap electrodes is capable of isolating and retaining specific ions for fragmentation upon collision with an inert gas in the cell. With variable energy CID (collision-induced dissociation) control, QIT-TOF-MS/MS provides high resolution precursor selection, high resolution detection and constant accuracy across MS and MS_n modes for structural studies of biological samples.

The application of MALDI-QIT-TOF-MS/MS has been popular in proteomics and numerous methods of sample preparation for these techniques have been suggested including dried droplet, crushed-crystal [13], thin layer [14], and sandwich methods [15]. Although MALDI has shown relatively high tolerance to impurities such as salt, pre-concentration /desalting step before utilizing mass spec has been highly recommended. One of the popular methods to achieve purification is using microcolumn [16] or Zip-Tip [17]. Though there were various methods to prepare sample to improve quality of spectra obtained from mass spec, the prepared samples often suffered from heterogeneity, in

which specific components were detected in certain positions of the sample but not in others. This phenomenon has been shown to significantly influence the MALDI-TOF-MS or MALDI-QIT-TOF-MS/MS response of peptide/proteins.

In this study, the newly developed plate for MALDI-QIT-TOF-MS and MS_n called the μ-focusing plate was tested to observe the possibility to use this new plate in proteomics when the heterogeneity problem could be reduced, and for the studying of the comparison of performance between this new plate and common stainless steal plate. Also the new plate was utilized to measure the availability for the on-plate digestion method which was studied in Dr. Lubman's lab [18].

6.2 Experimental Section

6.2.1Materials

Acetonitrile, Methanol, Ethanol, Water were purchased from Sigma-Aldrich in HPLC grade. Trifluoroacetic acid (TFA) was purchased from VWR International (West Chester, PA). α-CHCA (α-cyano-4-hydroxycinnamic acid) was purchased from Sigma-Aldrich. DHB (2, 5-dihydroxybenzoic acid) and synthetic serine phosphopeptide were purchased from Waters Corporation. Standard protein / peptide materials (beta-casein) were purchased from Sigma-Aldrich and Focusing plate was a gift from Hudson Surface Technology, Inc. (Newark, NJ).

6.2.2 Stock solution for standard protein solution

2-5mg/1ml stock solutions were prepared in water to make a solution with a final concentration of 1-100ppm.

6.2.3. Digestion

1) regular digestion procedure: standard proteins solution was placed in a speedvac concentrator to reduce the volume of solvent and NH_4HCO_3, 100 mM, and 10 mM DTT were added to these samples. Then modified trypsin was added at an enzyme-to-substrate ratio of 1:50, gently stirred, and incubated at 37 °C for 24 h. TFA, 2.5%, was used to end the digestion. ZipTips (Millipore, Inc.) were used to clean the sample before spotting the MALDI plate. The peptides were eluted from ZipTips and concentrated in 50% acenonitrile with 0.1% TFA.

2) on-plate digestion: The MALDI plate was precoated with 0.5 μL of trypsin stock solution of ~0.15 μg/μL and 0.5 μL of 50 mM ammonium bicarbonate was added to the top layer of each spot. The plate was kept in a humidifier chamber for 5 min at room temperature for digestion. Then, 0.5 μL of 0.1% TFA was added to each spot to stop the digestion, followed by adding 0.5 μL of α-CHCA and/or DHB matrix solution prepared by diluting saturated α-CHCA with 60% acetonitrile/ 0.1% TFA at a 1:4 ratio and DHB with 50% acetonitrile/0.1% TFA.

6.2.4 MALDI-QIT-TOF-MS

MS and MS2 spectra of peptide samples were obtained from a Shimadzu Axima QIT MALDI quadrupole ion trap-TOF (MALDI-QIT; Manchester, UK). Data acquisition and

processing during MALDI-QIT-TOF analysis were controlled by Kratos "Launchpad" software, with the standard instrument setting for optimum transmission at medium mass. A pulsed N_2 laser light (337 nm) with a pulse rate of 5 Hz was used for ionization. Each profile resulted from two laser shots. Argon was used as the collision gas for CID, and helium was used for cooling the trapped ions. The TOF was externally calibrated using 500 fmol/μL bradykinin fragment 1-7 (757.40 m/z), angiotensin II (1046.54 m/z), P14R (1533.86 m/z), and ACTH (2465.20 m/z) (Sigma). The MALDI-QIT-TOF was also externally calibrated using fullerite deposited directly onto the stainless steel plate. The data acquired from MALDI-TOF MS analysis was searched in MS-fit (http://prospector.ucsf.edu/ucsfhtml13.4/msfit.htm) for protein identification.

6.3 Results and Discussion

MALDI-QIT-TOF-MS/MS has been one of the popular instruments used in proteomics. In the process of preparing a sample for MALDI-QIT-TOF–MS/MS, the mixture of bio-sample (peptide/protein) and matrix (α-CHCA/DHB) was spotted onto the plate and dried. The dried sample/matrix mixture on the plate was typically measured as 5-15mm^2, where of only a small portion (0.002-0.03mm^2) [18] was irradiated by a pulsed laser, generating analyte ions. Since a liquid sample, purified using Zip-Tip [19], was dropped on the plate and dried, the prepared samples often suffered from heterogeneity, in which specific components were detected in certain positions of the sample but not in other positions. This occurrence is a critical disadvantage of this sample preparation method for mass spectrometry during spectral acquisition and accumulation of spectra

per sample. There have been improvements such as reducing sample diameter [20, 21] and using sample support [22].

Recently, a new plate has been developed using lithography techniques, resulting in hydrophobicity on the plate. This plate is called a μ-focusing plate. The μ-focusing plate offered several advantages such as convenience due to disposability, fewer contamination problems due to minimizing the chances of cross-contamination between samples, and increased MS sensitivity based on reduced contact angle to the plate compared to that of a stainless plate [23]. However, there are fewer recognized experiments using this plate to the study of peptide/protein detection. In this study, the μ-focusing plate was tested with standard peptide/protein for its use in proteomics, compared to that of a stainless steel plate for several aspects such as the direct comparison of signal intensity, usage of two matrixes for MS/MS, solvent effect, and on-plate washing method.

Figure 6.1 shows the result of using the hydrophobic plate versus the stainless steel plate in the MALDI-QIT-TOF-MS. The digested β-casein dissolved in water has been used as the sample with DHB as the matrix. These spectra were taken under similar conditions for both plates, and showed that the peaks on the hydrophobic plate showed almost double the intensity compared to that from a regular stainless steel plate in Figure 6.1(a). It is evident in Figure 6.1b that the use of the hydrophobic plate improves resolution and the relative intensity of the phosphorylated β-casein peaks where the two quadruply, and a singly phosphorylated peaks appear in the spectrum at m/s 3477.50, 3122.25, 2106.17 respectively. This result may confirm the crystal formed on the spot on the plate. The Figure 6.2 shows the image of crystal which is the mixture of

peptide/matrix on a regular stainless plate and new plate using the camera attached to the MALDI-QIT instrument. The crystal of mixture of peptide/matrix on the μ-focusing shows less heterogeneity than that on the stainless steel plate.

An important issue in the use of the MALDI-QIT-TOF-MS/MS has been that this device has been limited to the use of the DHB matrix, which is a relatively cold matrix. However, α-CHCA has often proved of greater utility for detection of peptides in complex digest mixtures. α-CHCA has generally not been used successfully as a matrix for external injection of ions into the QIT-TOF. We have found that the use of the hydrophobic plate allows the use of the α-CHCA matrix with results that are very comparable to that obtained using the DHB matrix. A direct comparison of the MALDI spectrum obtained for three standard peptides with the use of DHB or α-CHCA is shown in Figure 6.3(a). Figure 6.3(b) presents MS2 spectrum of synthetic serine phosphopeptide with α-CHCA as the matrix on the new plate. The choice of matrix for MS/MS could be wider, For the particular sample the MALDI-MS spectrum may be optimized using a mixture of DHB and α-CHCA where a ratio of DHB: α-CHCA of 3:1 was found to optimize the detection of the quadruply phosphorylated molecular ion peak. This is shown in Figure 6.4 for β-casein for a mixture of DHB and α-CHCA in a ratio of 3:1 using the hydrophobic plate and acetonitrile as the solvent in the matrix preparation.

In numerous studies, the protein from separated samples was digested using tryptic enzyme followed by purification using a C18 Zip-Tip before subjecting the sample to mass spectrometry. This process typically occurs overnight and sample purification is time consuming with a high risk of losing the sample during the processes. Various methods have been proposed to improve this critical sample preparation step

before using mass spectrometry, including microwave-digestion [24, 25], and on-line digestion [26-28]. Another new approach suggested, is called on-plate digestion [17, 29-30]. This method was able to reduce preparation time and sample loss. In this study, the on-plate digestion method was conducted onto the μ-focusing plate. Before performing direct on-plate digestion, the solvent effect was studied. The on-plate digestion method was suggested for the application of the on-line LC-MALDI-MS experiment and most of protein/peptide from biological samples was separated using HPLC based on their hydrophobicity using different types of solvent. The fractionated liquid sample using LC was directly spotted onto the plate containing a different amount of solvents which may affect signal intensity. The most common solvent used for separation based on hydrophobicity in reverse phase chromatography is acetonitrile.

The signal intensity was compared using different amounts of solvent with two matrixes. This is shown in Figure 6.5(a) for detection of β-casein peptides mixed with the DHB matrix in a different percent of acetonitrile using the hydrophobic plate in the sample preparation. The solvent used with the DHB matrix is changed from 75% acetonitrile to pure water where the higher molecular weight peaks are enhanced at the expense of the lower molecular peaks as the percent acetonitrile increases. The amount of 75% acetonitrile is usable to detect peaks at 3121.9, 2925.6 which correspond to quadruply phosphorylated peptide and H_3PO_4 loss. This new plate showed the availability for usage of both matrixes (DHB, α-CHCA). The usage of matrix DHB in the new plate presents a higher suitability for acetonitrile with higher contents up to 75% compared to that with α-CHCA which showed a lower suitability up to 60% in Figure 6.5(b). It is shown that β-casein can be detected using α-CHCA and the detection of the quadruply

phosphorylated peak is optimized versus the lower molecular weight peaks for a percent acetonitrile of 35% compared to pure water and higher acetonitrile content. The spectra obtained can be further modified depending on the percent acetonitrile used in the matrix preparation, and physical properties of samples (peptide/protein) which was from cell line or tissues. Considering noise peaks from the solvent, it could be recommended to use the amount of solvent up to 60-65%. The separation of peptides based on hydrophobicity has been performed mainly under 50% so that this result showed quite good signal detection using on-plate digestion directly after separation.

The quality of the MALDI mass spectra could be further improved using a combination of the hydrophobic plate and an on-plate washing procedure. In the proteomics study, a biological sample (proteins from cell line or tissues) was digested using enzymes, and then purified to remove remaining salts or impurities using Zip-Tips which is time-consuming and causes sample loss. In order to increase efficiency of enzymatic digestion, the pH was adjusted to basic pH using NH_4HCO_3. This salt could affect signal intensity and detection limits. Since the on-plate digestion is performed on the plate using the pre-coated tryptic enzyme under basic conditions, it is necessary to remove extra impurities or salt on the plate. The procedure of washing and re-crystallization was followed using the protocol described in the experimental section. It seems that usage of the Zip-Tip produces a cleaner spectrum. However, the on-plate digestion method produced a spectrum similar to the spectrum from using the ZipTip procedure, but in much less time using the on-plate digestion method. Another on-plate washing after digestion on the plate was tested under the condition with a higher content of solvent (45%) as shown in Figure 6.6a. It seems that the signal in the higher contents

of solvent produced a spectrum with a lower noise level when using on-plate washing. Since β-casein contains phosphopeptides, it may provide strong ion signal in the negative mode due to the phosphate group. Negative mode also was performed and Figure 6.6b shows similar results obtained from the negative mode compared to the positive mode using on-plate washing procedure. A signal of the quadruply phosphoyrlated β-casein peptide (marked as *) without conditioning mass spectrometry for the high mass range of peptides showed effectiveness using this washing in both the positive and negative modes in this instrument.

Since the on-plate digestion and washing procedure was performing well with β-casein, on-plate digestion with mouse tumor sample was performed. Less than 2μg of a lysed liquid sample solution was separated using reverse–phase chromatography based on hydrophobicity and directly spotted on the μ-focusing plate which was pre-coated with trypsin. After digesting and washing of the spot on the plate, a spectrum was obtained using MALDI-QIT-TOF-MS/MS. Figure 6.7 shows an example of MALDI spectrum obtained from one of the spots on the plate, identifying four proteins, RING finger protein 181, Poly [ADP-ribose] polymerase 1, Lymphoid-restricted membrane protein and Ephrin type-B receptor 2(data not shown). There was a limitation of performing MS_3 experiments, unlike a well-performed MS_3 with standard materials. The current study presented good test results for using the new plate in proteomics application with a low amount of sample usage. It will be necessary to perform the experiment with the best optimized condition in order to analyze small amounts of sample since sample-loss may be the one of main concerns while performing multi-steps (separation, purification, etc.) in proteomics using a very limited amount of sample.

6.4 Conclusion

The newly developed μ-focusing plate for MALDI-QIT-TOF-MS/MS experiment was tested. This plate has shown to have hydrophobicity and showed an excellent possibility for usage in proteomics. The new plate showed an increased signal compared to that from the stainless steel plate for standard materials such as β-casein, as well as a higher tolerance for solvent (acetonitrile), up to 75% with a proper matrix. Also an uncommon matrix, α-CHCA for MALDI-QIT-TOF-MS/MS was shown to give a good signal for MS2. The best condition for using a matrix combination between DHB: α-CHCA is a 3:1 ratio. When the new plate was subjected to on-plate digestion, more digested peaks were likely shown in spectra after the on-plate washing step, and several signals was detected in one spot. This study may suggest this hydrophobicity plate will be well suited in the proteomics field, and that coupling with micro-separation may produce improved performance in terms of detection of limited amounts of sample.

.

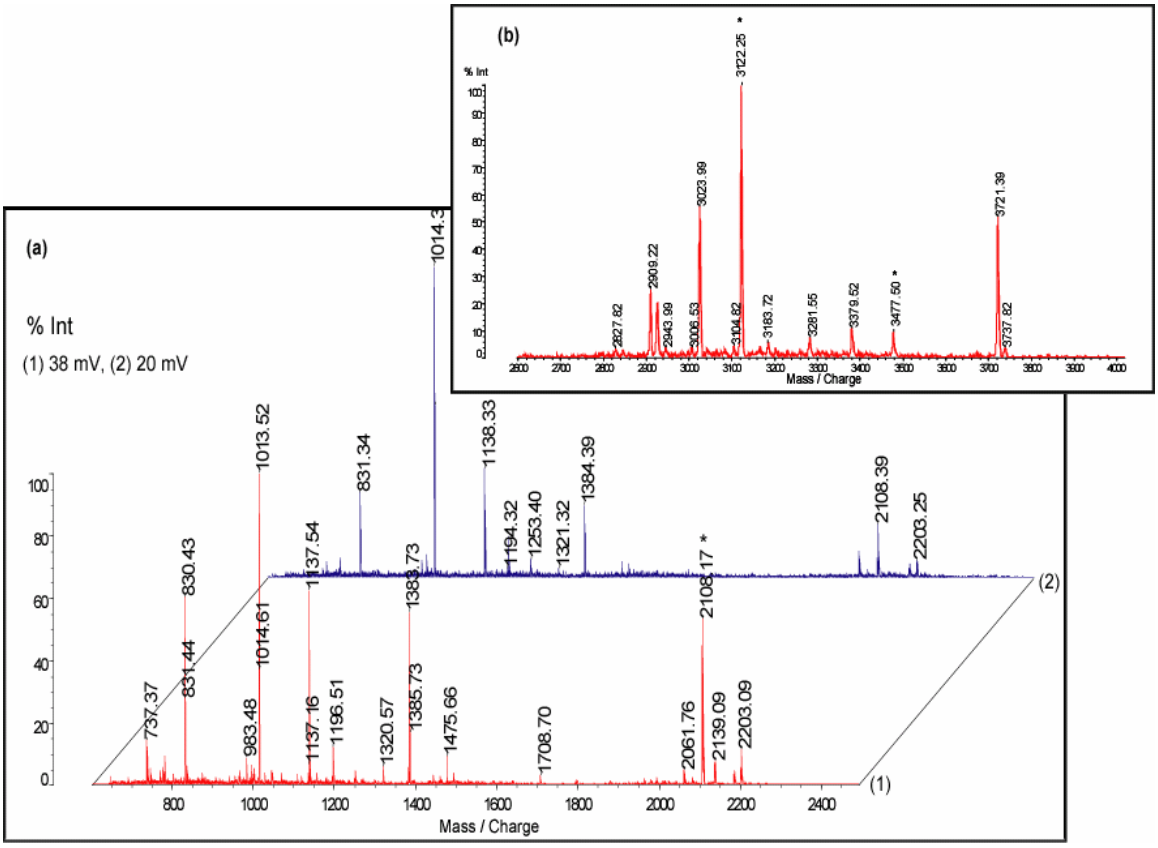
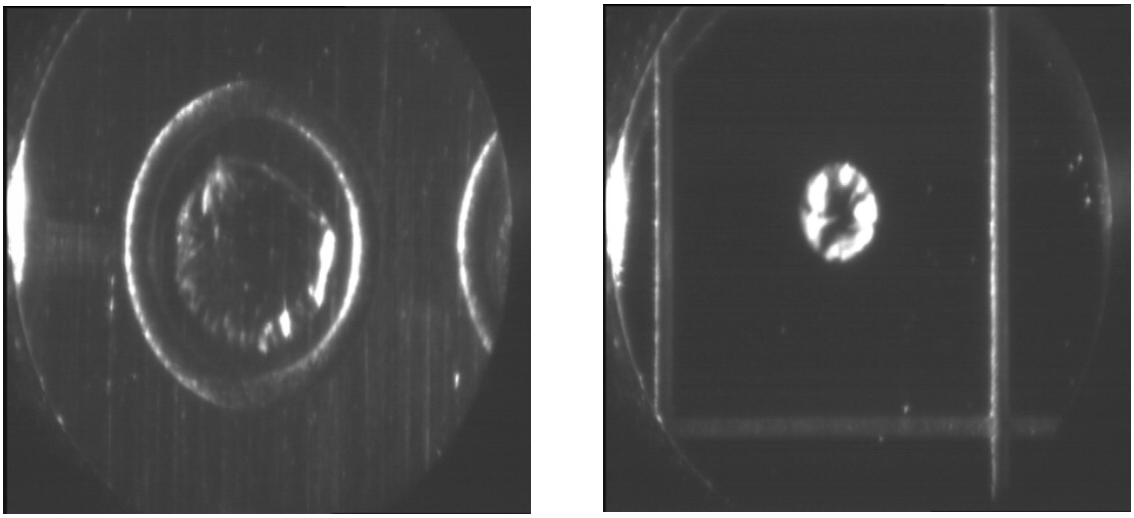


Figure 6.1. The spectra of digested β-casein peptide obtained from μ-focusing plate (1) and stainless plate (2) in Figure 1(a). The high m/z range is zooming-in in the Figure 1(b). DHB was used as matrix, and phosphorylated peaks marked as *



(a) (b)

Figure 6.2. The image of crystal which is the mixture of sample and matrix on the regular stainless steel plate (a, left) and μ-focusing plate (b, right). DHB was used as matrix.

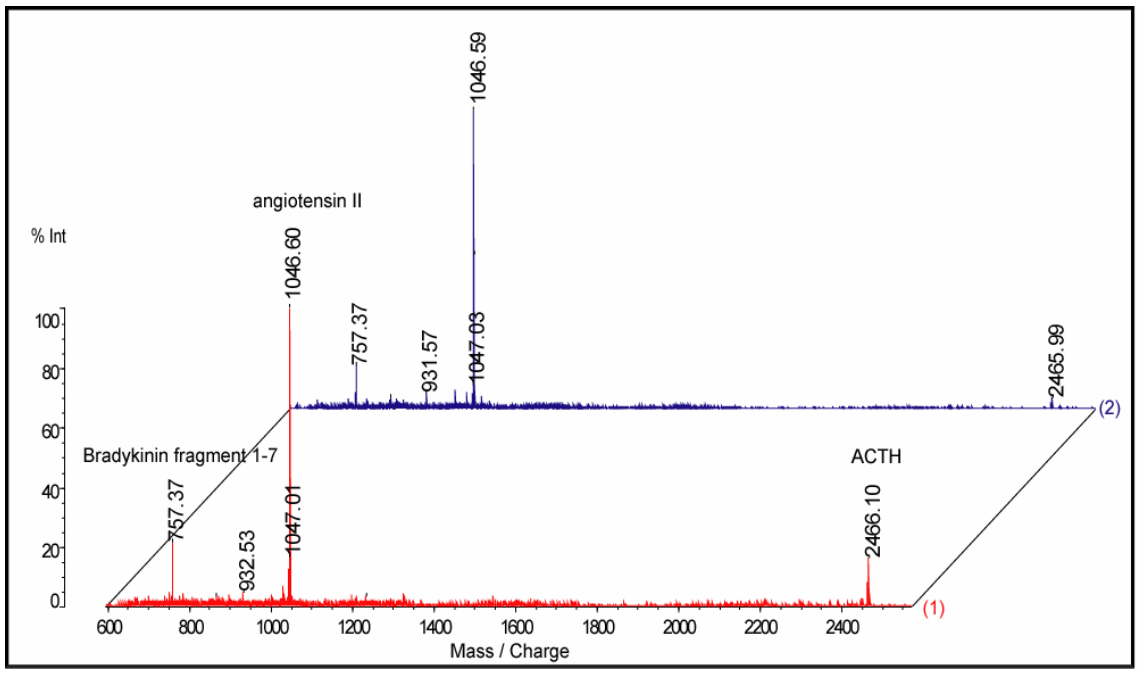


Figure 6.3 (a). The spectra obtained using different matrixes on the μ-focusing plate. The spectrum (1) on the bottom: DHB matrix, Top spectrum (2) on the top: α-CHCA matrix

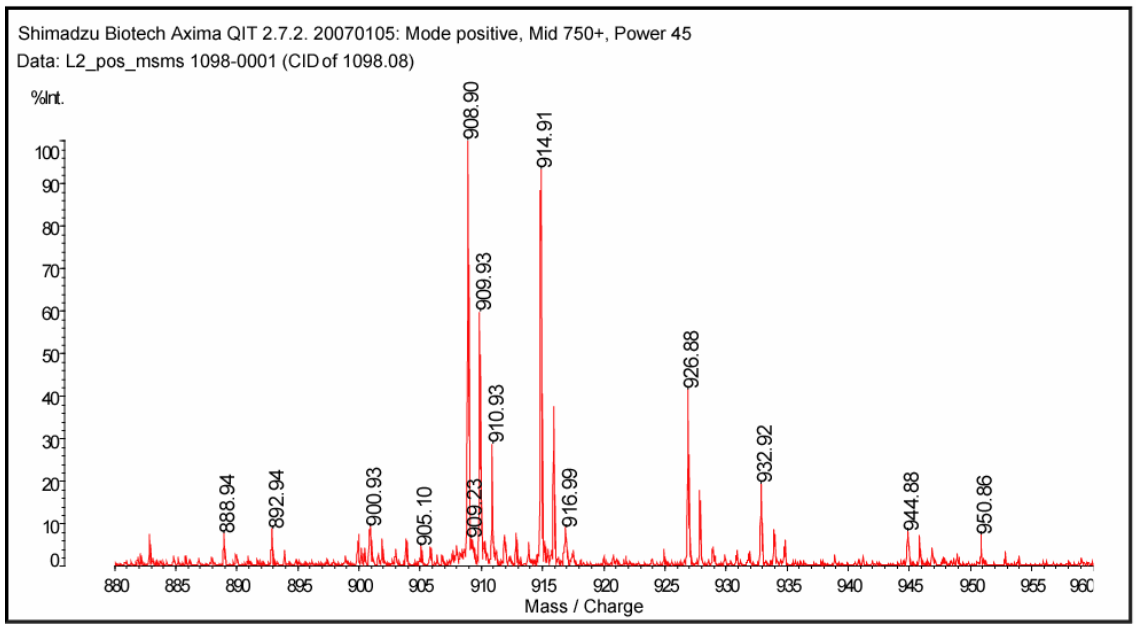


Figure 6.3(b) The MS/MS spectrum of synthetic serine phosphopeptide.

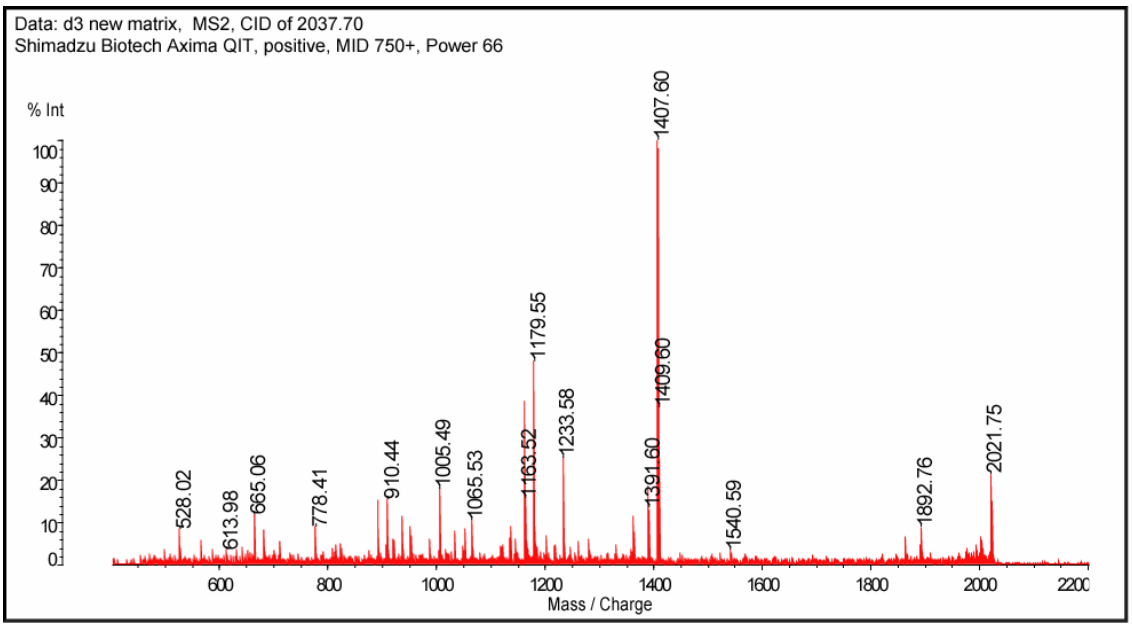


Figure 6.4. The MS2 spectrum of beta-casein in positive mode using matrix mixture (DHB: α-CHCA) on the μ-focusing plate.

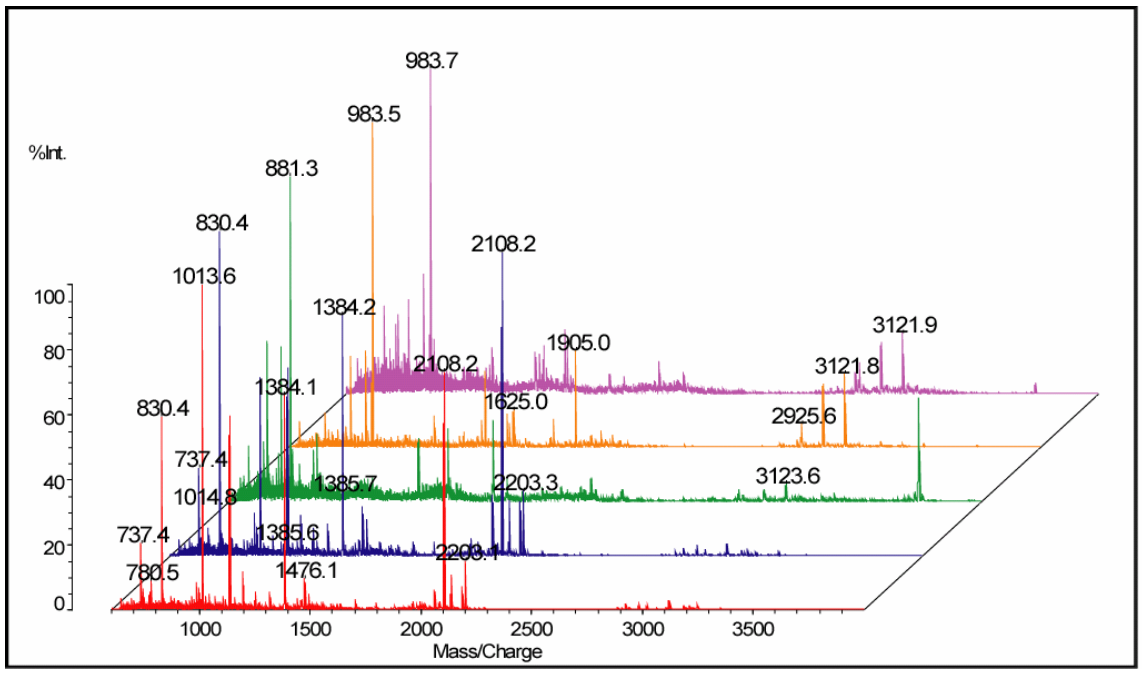


Figure 6.5(a). The comparison of signal intensity with DHB in the different amount of solvents.

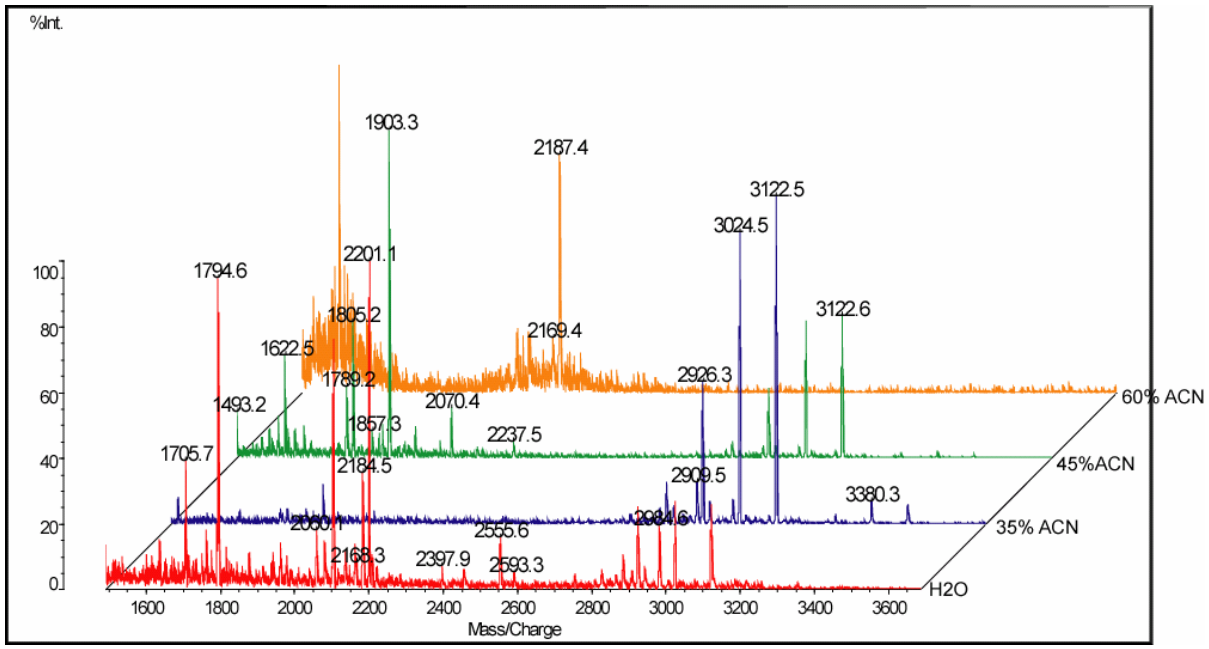


Figure 6.5(b). The comparison of signal intensity with α-CHCA in the different amount of solvents.

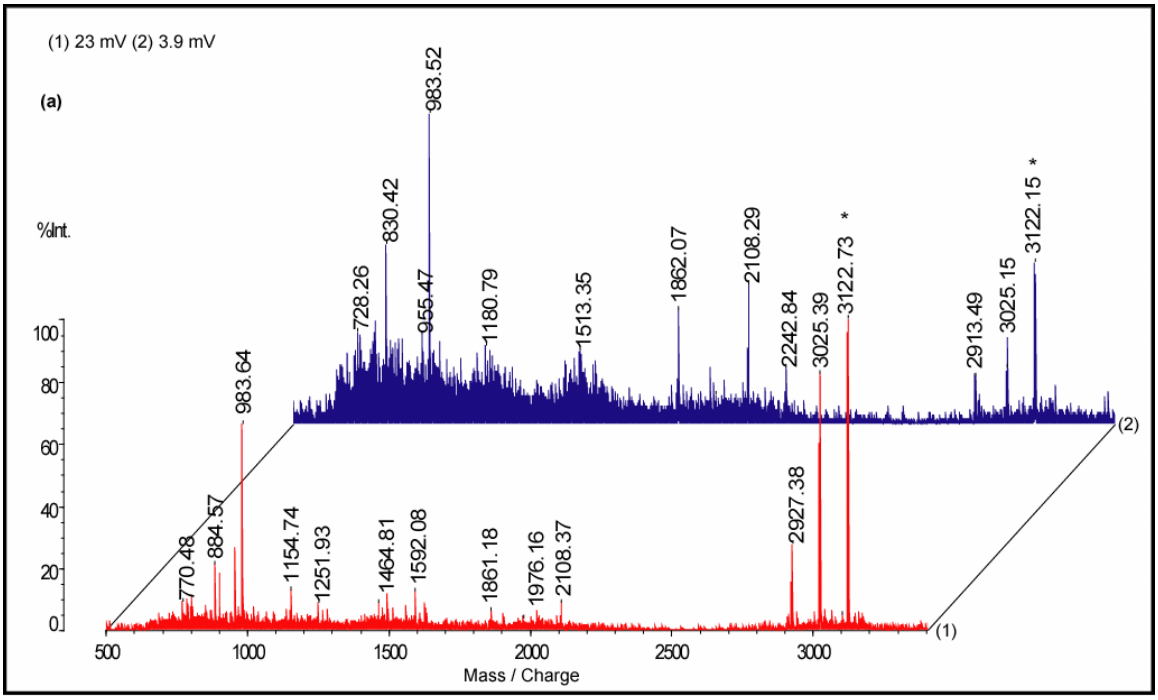


Figure 6.6(a). The comparison of signal using on-plate washing after on-plate digestion (1) and using zip-tip(2) in 45% ACN in positive mode.

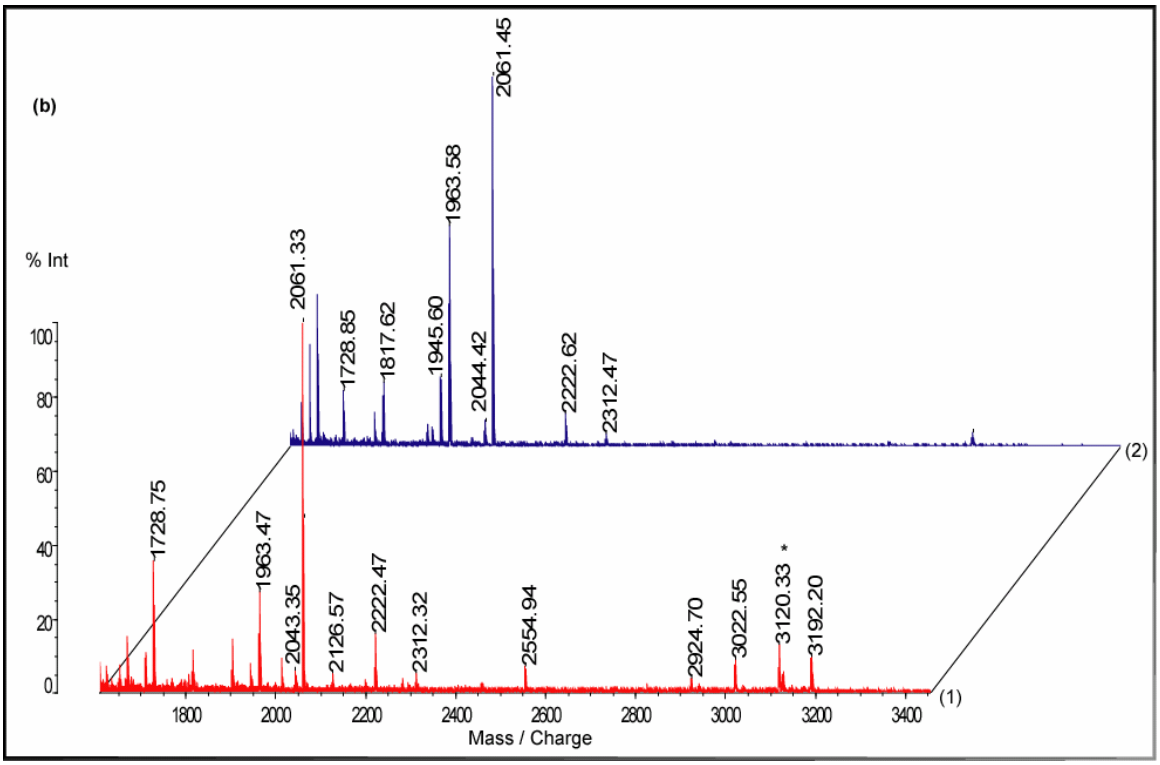


Figure 6.6(b). The comparison of signal using on-plate washing after on-plate digestion (1) and using zip-tip(2) in 30% ACN in positive mode.

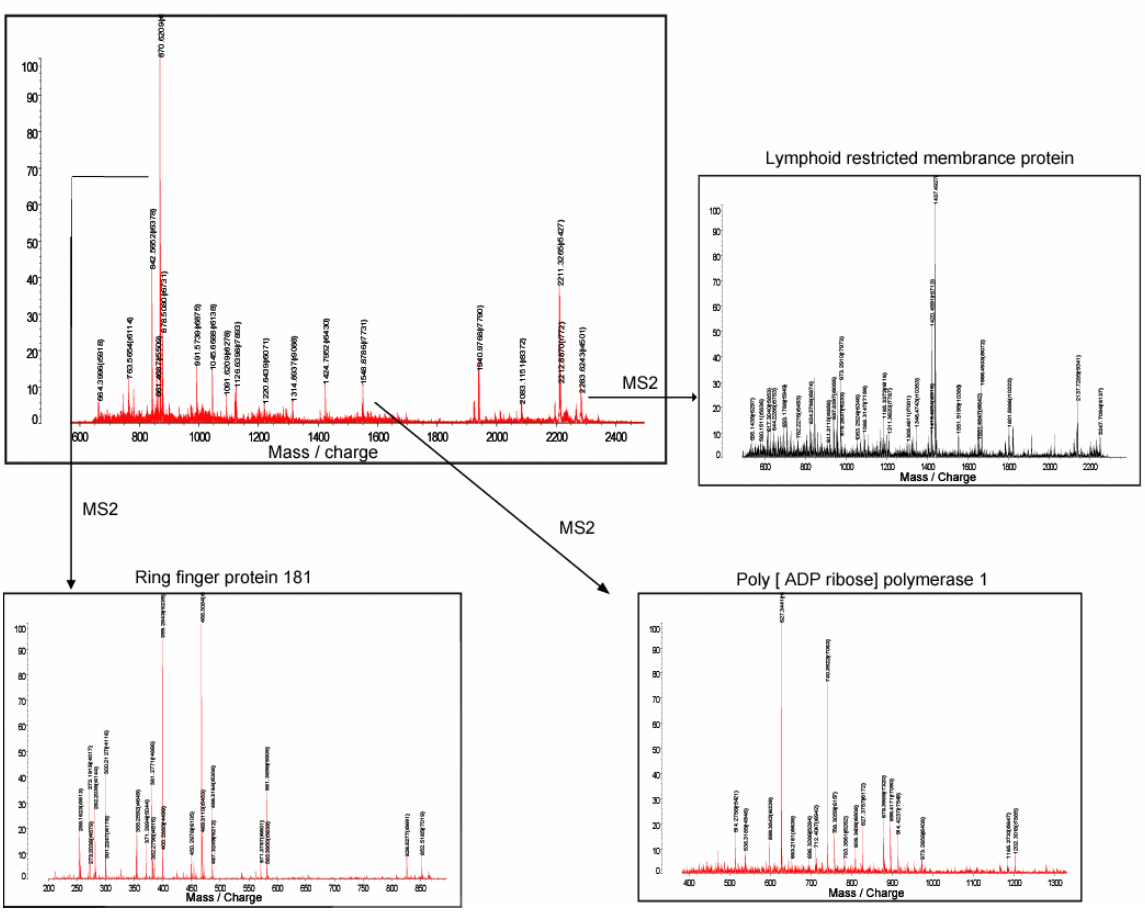


Figure 6.7. On-plate digestion and identification of mouse tumor samples with MALDI-QIT-TOF-MS/MS in positive mode. DHB was used as matrix.

6.5 References

[1] Karas,M.; Bachmann, D.; Bahr, U.; Hillenkamp, F.; *Int. J. Mass Spectrom. Ion Process* **1987**, *78*, 53–68.

[2] Karas, M.; Hillenkamp, F. *Anal. Chem*. **1988**, *60*, 2299–2301.

[3] Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal. Chem*. **1991**, *63*, 1193-1200.

[4] Zenobi R.; Knochenmuss R. *Mass Spectrom. Rev*. **1999**, *17*, 337-366.

[5] Cotter, R. T. *Anal. Chem*. **1992**, *64*, 1027A-1039A.

[6] Fang, L.; Zhang, R.; Williams, E. R.; Zare, R. N. *Anal. Chem*. **1994**, *66*, 3696-3701.

[7] Boyle J.G.; Whitehouse, C. M. *Anal. Chem*. **1992**, *64*, 2084-2087.

[8] Mirgorodskaya, O. A.; Shevchenko, A. A.; Chernushevich, I. V.; Dodonov, A. F.; Miroshinikov, A. I. *Anal. Chem*. **1994**, *66*, 99-107.

[9] Burlingame, A. L.; Carr, S. A. *Mass Spectrometry in the Biological Sciences*, Human Press, Totowa, NJ, **1996**.

[10] Burlingame, A. L.; Boyd, R. K.; Gaskell, S. J. *Anal. Chem*. **1996**, *68*, 599R-651R.

[11] Baillie, T. A. *Int. J. Mass Spectrom. Ion Proc.* **1992**, *118*, 289-314.

[12] Krutchinsky, A. N.; Kalkum, M.; Chait, B. T. *Anal. Chem*. **2001**, *73*, 5066-5077.

[13] Xiang, F.; Beavis, R. C. *Rapid Commun. Mass Spectrom*. **1994**, *8*, 199–204.

[14] Vorm, O.; Roepstorff, P.; Mann, M. *Anal. Chem* **1994**, *66*, 3281–3287.

[15] Li, L.; Golding, R. E.; Whittal, R. M. *J. Am. Chem. Soc*. **1996**, *118*, 11662–11663.

[16] Gobom, J.; Nordhoff, E.; Mirgorodskaya, E.; Ekman, R.; Roepstorff, P. *J. Mass Spectrom*. **1999**, *34*, 105–116.

[17] Schuerenberg, M.; Luebbert, C.; Eickhoff, H.; Kalkum, M.; Lehrach, H.; Nordhoff, E. *Anal. Chem*. **2000**, *72*, 3436-3442.

[18] Zheng, S.; Yoo,C.; Delmotte, N.; Miller, F. R.; Christian G.; Huber, C. G.; Lubman, D. M. *Anal. Chem*. *2006*, **78**, 5198-5204.

[19] Bagshaw, R. D.; Callahan, J. W.; Mahuran, D. J. *Anal. Biochem*. **2000**, *284*, 432–435.

[20] Little, D. P.; Cornish, T. J.; ODonnell, M. J.; Braun, A.; Cotter, R. J.; Koster, H. *Anal. Chem*. **1997**, *69*, 4540-4546.

[21] Gobom, J.; Nordhoff, E.; Mirgorodskaya, E.; Ekman, R.; Roepstorff, P. *J. Mass Spectrom*. **1999**, *34*, 105-116.

[22] Hung, K. C.; Ding, H.; Guo, B. C. *Anal. Chem*. **1999**, *71*, 518-521.

[23] Park, K.; Ha, M.; Cha, N.; Park, C.; Lim, H.; Lee, E.; Kim, Y. Development of Disposable Microfocusing Plate for MALDI-TOF. ASMS, San Antonio, Texas, USA. **2005**.

[24] Pramanik, B. N.; Mirza, U. A.; Ing, Y. H.; Liu, Y. H.; Bartner, P. L.; Weber, P. C.; Bose, A. K. *Protein Sci*. **2002**, *11*, 2676-2687.

[25] Lin, S. S.; Wu, C. H.; Sun, M. C.; Sun, C. M.; Ho, Y. P. *J. Am. Soc. Mass Spectrom*. **2005**, *16*, 581-588.

[26] Marie, G. ; Serani, L. ; Laprevote, O. *Anal. Chem*. **2000**, *72*, 5423-5430.

[27] Vecchione, G.; Casetta, B.; Santacroce, R.; Margaglione, M. *Rapid Commun. Mass Spectrom*. **2001**, *15*, 1383-1390.

[28] Calleri, E.; Temporini, C.; Perani, E.; De Palma, A.; Lubda, D.; Mellerio, G.; Sala, A.; Galliano, M.; Caccialanza, G.; Massolini, G. *J. Proteome Res*. **2005**, *4*, 481-490.

[29] Warscheid, B.; Fenselau, C. *Proteomics* **2004**, *4*, 2877-2892.

[30] Harris, W. A.; Reilly, J. P. *Anal. Chem.* **2002**, *74*, 4410-4416.

Chapter 7

Conclusion

The study of the proteome has great challenges due to its complexity and dynamic range of biological samples to be analyzed in cancer studies. Analytical technologies for proteomics have improved to overcome these problems and posses the ability to detect low concentrations of proteins or peptides with biological significance. In addition to continuously developing techniques including mass spectrometry, improved software allowed researchers to study proteins not only highly expressed in cancer but also involved in signal pathways related to tumor progressions.

This dissertation has described the development and application of an integrated liquid separation, protein microarray and tandem mass spectrometry strategy for global screening of proteins in ovarian adenocarcinomas and pancreatic ductal adenocarcinoma.

The mass mapping approach using liquid 2D separation followed by mass spectrometry to detect molecular weights of proteins was investigated for comparative proteins in low-stage (FIGO stage 1 or 2) versus high-stage (FIGO stage 3 or 4) human OEAs. Two-dimensional liquid-based separation/mass mapping techniques to elucidate molecular weight and pI measurements of the differentially expressed intact proteins was done. These mass maps (over a pI range of 5.6-4.6) revealed that the low-stage OEAs demonstrated protein over-expression at the lower pI ranges (pI 4.8-4.6) in comparison to

the high-stage tumors, which demonstrated protein over-expression in the higher pI ranges (pI 5.4-5.2). These data suggest that both low- and high-stage OEAs have characteristic pI signatures of abundant protein expression probably reflecting, at least in part, the different signaling pathway defects that characterize each group. In Chapter II, the low-stage OEAs were distinguishable from high-stage tumors based upon their proteomic profiles. Interestingly, when only high-grade (grade 2 or 3) OEAs were included in the analysis, the tumors still tended to cluster according to stage, suggesting that the altered protein expression was not solely dependent upon tumor cell differentiation. Further, these protein profiles clearly distinguish OEA from other types of ovarian cancer at the protein level.

Another global screening of protein in cancer was studied for humoral response in cancer, using protein microarrays. The idea that there is an immune response to cancer in humans has been demonstrated by the identification of autoantibodies against a number of intracellular antigens in patients with various tumor types. This phenomenon is known as the humoral response and the detection of such autoantibodies has been shown to be of great potential diagnostic and prognostic value in the detection of cancer and the ability to predict the course of the disease. The autoantibody response to pancreatic cancer has been explored using a natural liquid-based microarray approach. A Panc-1 cell line was fractionated using a 2-D liquid separation method into over 1029 proteins as expressed in the cell and spotted onto nitrocellulose-coated glass slides. Each slide was probed with samples from 38 pancreatic cancer sera, 23 pancreatitis sera and 25 normal sera and probed for the humoral response against each protein spot. The response data obtained from protein microarrays for each sample over the protein set was analyzed

by four different statistical methods (COPA, OS, Wilcoxon, Pamr) to determine the proteins that showed the greatest differential response against serum. Each statistical method generated a list of differential response proteins in a comparison between cancer vs normal, cancer vs pancreatitis and pancreatitis vs normal. Among identified proteins using LC-MS/MS were annxein 2 and cytoplasmic, malate dehydrogenase which showed a higher response in cancer sera versus normal sera. Further work was performed using recombinant proteins and a blotting method on microarrays for the proteins identified in the discovery set using sera from cancer, pancreatitis, type-2 diabetes and normals. This pre-validation experiment identified proteins that could potentially identify cancer from pancreatitis, normal controls or type-2 diabetes.

An important issue in current proteomic analysis is the ability to work with small-volume samples. This becomes particularly important in situations where only a limited amount of sample may be available, such as tissue or fluid samples extracted from aspirates, laser-capture micro-dissection, tumor micro-environment experiments or stem-cell research. In these cases, there may be <100,000 cells available or fluids in the amount of 50μl or less. This number may correspond to only several micrograms of total protein where, if there are a thousand proteins present, then on average there may only be tens of femtomoles of each protein available. It thus becomes critical to find methods capable of separating and analyzing such small amounts of sample with the ability to identify the presence of large numbers of proteins.

A method has been developed using a micro-chromatofocusing (micro-CF) procedure with a weak-anion exchanger packed in a capillary column. This method is a scaled-down version of the CF separation currently used in the Beckman PF2D

instrument and described in Chapter III. Instead of milligram amounts of the starting material, 2-30μg of material can be fractionated based upon pH using the micro-CF method. The intact proteins can be collected according to pH value and further analyzed by a second dimension. Digestion of the eluted fraction of intact proteins from micro-CF was demonstarated and followed by nano-RPLC-MS/MS. This micro-proteomic procedure is demonstrated for analysis of 700-800 proteins from only 10 μg of two ovarian cell lines (MDAH 2774 and TOV 112D).

Another analytical method to utilize limited amounts of sample is the use of the cIEF/nano-LC/MS-MS technique. Using this technique, profiling of proteins based on defects in signaling pathways and gene mutation in two groups of tumor samples has facilitated the study of a genetically engineered mouse-model of human endometrioid ovarian cancer in Chapter V. Mouse tumors with the p53 mutation, in addition to defects on Wnt/β-catenin and PI3K/Pten signaling pathways, has revealed an increased number of detected proteins involved in signaling, and up-regulated expression of proteins in many signaling and metabolic pathways. p53 appears to play an important role with regard to tumor progression. Also, a number of proteins detected in tumors without mutant p53 showed increased expression-levels, suggesting the possibility of improved early detection of endometrioid ovarian cancer.

The development of new methods can help to increase the ability to study proteomics. The newly developed plate for the MALDI-QIT-TOF-MS/MS experiment was tested in Chapter VI. This plate has been shown to have hydrophobicity and showed an excellent possibility for usage in proteomics. The new plate showed an increased signal compared to that from stainless steel for standard materials such as β-casein, as

well as a higher tolerance for solvent (acetonitrile), up to 75% with a proper matrix. Also an uncommon matrix, α-CHCA for MALDI-QIT-TOF-MS/MS was shown to provide a good signal for MS_2. The best condition for using a matrix combination between DHB: α-CHCA is a 3:1 ratio. When the new plate was subjected to on-plate digestion, more digested peaks were likely shown in spectra after the on-plate washing step, and several signals detected in one spot. This study may suggest this hydrophobicity plate will be well-suited in the proteomics field, and that coupling with good micro-separation may produce better performance in terms of detection of limited amounts of sample.

The strategy described herein can be successfully utilized to study a variety of cancers with different amounts of sample-scale (milligram to microgram scale) in order to detect proteins associated tumor progression / genetic defects. The described techniques are applicable to answer biological questions that can help to understand cancer pathology as well as to detect proteins for early diagnosis in cancer. However further work needs to be done to stretch the limits of use of sample pools in the GEM study of ovarian endometrioid cancer and the microarray study of pancreatic cancer in order to assess the quality of the markers highlighted and described in this dissertation.