

RESEARCH ARTICLE

Classifications of ovarian cancer tissues by proteomic patterns

Yi Zhu¹, Rong Wu³, Navneet Sangha³, Chul Yoo¹, Kathleen R. Cho^{3,5}, Kerby A. Shedden⁴, Hidetaka Katabuchi⁶ and David M. Lubman^{1, 2, 3, 5}

¹ Department of Chemistry, The University of Michigan, Ann Arbor, MI, USA

² Department of Surgery, The University of Michigan Medical Center, Ann Arbor, MI, USA

³ Department of Pathology, The University of Michigan Medical Center, Ann Arbor, MI, USA

⁴ Department of Statistics, The University of Michigan, Ann Arbor, MI, USA

⁵ Comprehensive Cancer Center, The University of Michigan Medical Center, Ann Arbor, MI, USA

⁶ Kumamoto University School of Medicine, Kumamoto, Japan

Ovarian cancer is a morphologically and biologically heterogeneous disease. The identification of type-specific protein markers for ovarian cancer would provide the basis for more tailored treatments, as well as clues for understanding the molecular mechanisms governing cancer progression. In the present study, we used a novel approach to classify 24 ovarian cancer tissue samples based on the proteomic pattern of each sample. The method involved fractionation according to *pI* using chromatofocusing with analytical columns in the first dimension followed by separation of the proteins in each *pI* fraction using nonporous RP HPLC, which was coupled to an ESI-TOF mass analyzer for molecular weight (MW) analysis. A 2-D mass map of the protein content of each type of ovarian cancer tissue samples based upon *pI* versus intact protein MW was generated. Using this method, the clear cell and serous ovarian carcinoma samples were histologically distinguished by principal component analysis and clustering analysis based on their protein expression profiles and subtype-specific biomarker candidates of ovarian cancers were identified, which could be further investigated for future clinical study.

Received: March 2, 2006

Revised: June 14, 2006

Accepted: July 10, 2006



Keywords:

Biomarkers / LC-MS / Multidimensional chromatography / Ovarian cancer / Proteomic profiling

1 Introduction

Ovarian carcinoma is the fifth most common cancer in American woman and the fifth most common cause of cancer-related deaths [1]. Over 70% of women with ovarian cancer are diagnosed when the disease is advanced beyond the pelvis [2]. The 5-year survival rate for patients diagnosed with

ovarian cancer has remained under 30% for many years despite the development of new treatments. Malignant ovarian epithelial neoplasms (carcinomas) are both morphologically and biologically heterogeneous. Based on morphological criteria, there are four major histological types of primary ovarian carcinomas (serous, clear cell, endometrioid, and mucinous). The serous carcinomas comprise about one-half of all ovarian carcinomas and account for the majority of ovarian cancer-related deaths. Although clear cell carcinomas are much less common, several studies have noted their particularly unfavorable prognosis, even when corrected for tumor stage [3–5]. In addition to molecular studies suggesting that the different histological types of ovarian cancers likely represent distinct disease entities [6–9], studies based on comprehensive gene expression profiling have shown that

Correspondence: Professor David M. Lubman Surgery, University of Michigan, 1150 W. Medical Center Dr., MSRB I A510B, Ann Arbor, MI, 48109, USA

E-mail: dmlubman@umich.edu

Fax: +1-734-763-6199

Abbreviations: ANXA4, Annexin IV; CF, chromatofocusing; MW, molecular weight; PCA, principal component analysis

clear cell carcinomas can be distinguished from other poor prognosis ovarian carcinomas (*e.g.* serous carcinomas) based on their distinctive gene expression signature [8]. Understanding the gene expression patterns characterizing each morphological type of ovarian cancer may eventually lead to the development of new type-specific diagnostic and/or therapeutic strategies that might significantly improve survival, particularly amongst women with serous and clear cell ovarian carcinomas.

Although cancer can be described as a disease of genes, the protein products of genes are actually the functional effector molecules that manifest the genetic information. Characterization of protein expression for each morphological type of ovarian cancer could lead to biologic insights and potential improvements in specific and effective diagnosis and treatments for this disease. LC-MS based multi-dimensional separation methods are rapidly emerging for large-scale protein expression profiling analysis to identify the relative quantities of the various proteins between samples [10–14]. We herein report a 2-D liquid phase mass mapping method to explore protein-based expression profiling for ovarian cancer classification and to search for subtype-specific biomarker candidates. In previous work, we have demonstrated that high-resolution 2-D mass maps of human cancer cell lines can be generated by this method with excellent reproducibility, allowing accurate interlysate comparison among different samples over the desired *pI* range [15, 16]. In addition, proteins in the range from 5 to 85 kDa can be identified by MALDI-TOF MS based on PMF and confirmed by LC-MS/MS with high confidence combined with *pI* and protein intact *M_r* information. In the present work, 2-D mass mapping was applied to profile the proteome of primary clear cell and serous ovarian carcinoma tissue samples. Proteins extracted from each tissue sample were fractionated by chromatofocusing (CF) based on *pI* in the first dimension, further separated by RP-HPLC in the second dimension and subsequently analyzed by online ESI-TOF to obtain intact protein *M_r*. Two protein-rich *pI* fractions (pH 5.4–5.2 and pH 5.2–5.0) were selected for interlysate comparisons of protein content between the two different histological types of ovarian tumor tissue samples (see Supplementary Figs. 1 and 3). Statistical analysis of the protein expression pattern was able to correctly differentiate clear cell from serous subtypes of ovarian cancers. In addition, significant differentially expressed proteins as type-specific marker candidates were identified, which could lead to development of new strategies for the diagnosis and treatment of these cancers.

2 Materials and methods

2.1 Tumor tissues lysis

The proteomes of 24 ovarian cancer tissue samples (16 clear cell carcinomas and 8 serous carcinomas) were profiled in this study with approval from the University of Michigan's

Institutional Review Board (IRBMED 1999-0428). Thirteen carcinoma specimens were obtained from the Cooperative Human Tissue Network/Gynecologic Oncology Group Tissue Bank, seven from the University of Michigan Health System, and four from Kumamoto University Hospital. Available clinico-pathological data associated with tumor specimens are provided in Table 1. All tumors were promptly snap frozen after resection and stored at -80°C until the time of analysis. Hematoxylin- and eosin-stained frozen sections of tumor immediately adjacent to the tissue lysed for proteomic analysis were examined to ensure the presence of at least 70% tumor cells. Tumor tissues were minced with razor blades and subsequently sealed into 2-mL screw-cap microcentrifuge vials (BioSpec Products, OK), containing hundreds of minute glass beads (BioSpec Products). Vials were subsequently filled with lysis buffer (7.5 M urea, 2.5 M thiourea, 12.5% glycerol, 50 mM Tris, 2.5% n-OG, 6.25 mM TCEP, 1.25 mM protease inhibitor, pH adjusted to 7.4) with no air bubbles. Tissue samples were homogenized for 3 min in 10-s increments at 4800 rpm in a mini-bead beater cell disruptor followed by centrifuging at 5000 rpm for 10 min at room temperature to pellet the bead mix. The supernatant containing proteins were collected and stored on ice. To avoid incomplete tissue disruption and protein

Table 1. Summary of ovarian carcinoma samples

Sample ID	Age at diagnosis	Tumor grade	Stage at Dx
KU-OC-002	48	3 ^{a)}	2B
KU-OC-003	43	3 ^{a)}	1A
KU-OC-006	55	3 ^{a)}	1A
KU-OC-007	65	3 ^{a)}	1C
CHTN-OC-003	74	3 ^{a)}	3
CHTN-OC-004	48	3 ^{a)}	3C
CHTN-OC-006	67	3 ^{a)}	1
CHTN-OC-012	51	3 ^{a)}	4
CHTN-OC-025	58	3 ^{a)}	3
CHTN-OC-032	54	3 ^{a)}	1
CHTN-OC-034	62	3 ^{a)}	1
CHTN-OC-043	55	3 ^{a)}	1
UM-OC-1T	56	3 ^{a)}	3
UM-OC-2T	60	3 ^{a)}	1C
UM-OC-3T	55	3 ^{a)}	1
UM-OC-4T	78	3 ^{a)}	3
CHTN-OS-002	83	3	3C
CHTN-OS-018	57	2	3C
CHTN-OS-020	44	3	3C
CHTN-OS-046	53	2	2C
CHTN-OS-081	74	3	3C
UM-OS-07	54	3	3C
UM-OS-10	61	1	3
UM-OS-11	43	3	3C

a) Clear cell carcinomas cannot be graded using the WHO or FIGO systems (Y. Shimizu *et al.*, Cancer, 1998) and were classified as grade 3 as recommended by the NCCN Practice Guidelines for Ovarian Cancer.

extraction, the vials were filled with fresh lysis buffer again and homogenized for 2 min. These two lysis solutions were combined in 10-mL polycarbonate centrifuge tubes and insoluble material was precipitated by centrifugation at 35 000 rpm for 1 h (80Ti Beckman Ultracentrifuge). Tissue lysates were stored at -80°C .

2.2 Chromatofocusing

CF is a weak anion exchange technique for separating proteins according to their *pI*. In our work, CF was performed on an HPCF-1D column (250×2.1 mm) (Eprogen, Darien, IL). A pH gradient was formed using start buffer (SB) and eluent buffer (EB) on the column. The pH of the SB, which contains 25 mM bis-Tris (Sigma), 6 M urea (Sigma) and 0.2% OG, was set slightly above the upper limit of the pH gradient (pH 7.2). The pH of the EB, which contained 1:10 diluted polybuffer 74, 6 M urea and 0.2% OG, was adjusted to the value determined for the lower limit of the pH gradient (4.0). The pH of SB and the EB were adjusted with iminodiacetic acid. The 2.5-mL lysates were exchanged with 3 mL SB using a PD-10 column (Amersham Pharmacia, Piscataway, NJ) to avoid the effect of salt on the weak anion exchange column.

The column was equilibrated with the SB until the pH of the effluent was the same as the pH of SB and the A of UV (280 nm) was stable. The pH of the eluent was monitored online by a post-UV detector pH flow cell (Lazar Research Laboratories, Los Angeles, CA). The buffer exchanged tissue lysates (7.1 mg) were loaded onto the CF column by multiple injections to prevent overloading. No replicate CF runs were performed due to the limited amount of tumor available for each sample. Proteins with *pI* under pH 7.2 will bind to the column. Once both the pH of the effluent and the A of UV were stable, the column was eluted with the EB to start the gradient. Proteins elute out in order of their *pI* at flow rate of 0.2 mL/min. The proteins eluted out were collected from pH 7.2 to pH 4.0 in 0.2 pH-unit intervals and protein fractions were subsequently stored at -80°C for later use.

2.3 LC/ESI-TOF-MS

The LC separation was performed at a flow rate of 0.4 mL/min on an ODS IIIIE ($33 \text{ mm} \times 4.6 \text{ mm}$) column packed with 1.5 μm non-porous silica beads derivatized with C18 (Eprogen). As solvent A, 0.1% TFA and 0.3% formic acid (FA) in H_2O was used. The percentage of TFA and FA in ACN was used as solvent B in the gradient. The column was equilibrated with 5% of solvent B until the A of UV (214 nm) (Thermo separation products sp UV 200, FL) became stable and the protein fraction (pH 5.4–5.2) from CF separation was applied to the column, which was placed in a column heater to maintain the column temperature at 65°C . The gradient was run from 5 to 25% in 2 min, 25 to 35% in

5 min, 35 to 45% in 10 min, 45 to 75% in 10 min, and 75 to 100% in 2 min and from 100% back to 5% in 1 min. A post-column splitter was used to split the effluent into half. Half of the effluent was delivered into the source of ESI-TOF (LCT, Waters, US). The other half of the effluent was passed through a Beckman model 166 UV absorption detector (214 nm) and was collected using a Beckman SC-100 fraction collector for MALDI-TOF MS analysis. The LCT parameters were set as follows: desolvation temperature 300°C , source temperature 110°C , desolvation gas 970 L/h, capillary voltage +3200 V, sample cone +35 V, extraction cone +2 V, and rf voltage 750 V. Before sample injection, the external calibration was utilized with the direct infusion of NaI-CsI solution by a syringe pump. Bovine insulin (0.2 μg) was mixed with samples and injected into the LC-MS for calibration and quantification.

Maxent1 software was used for LCT data analysis. The total ion chromatogram (TIC) was scanned for regions that contain redundant multiple-charged peaks, and those regions were combined for deconvolution (as described in Supplementary Figs. 2 and 4). The deconvoluted peaks were subsequently combined into a single mass spectrum, which was converted to a text file for input into the 2-D mapping software (Proteovue) written in-house.

2.4 Tryptic digestion of the LCT fractions and peptide preparation for MALDI-MS analysis

After collection, fractions from the LCT separation were concentrated down to 20 μL using a Speedvac concentrator (Labconco, Kansas City, MO). To the reduced volume fractions, 10% v/v 10 mM DTT (Sigma), 10 μL 1 M NH_4HCO_3 (Sigma) and 0.25 μg of TPCK-treated trypsin (Promega, Madison, WI) were added. The fractions were then incubated at 37°C for 24 h. After 24-h incubation, 2.5% v/v TFA was added into each fraction to stop digestion. Prior to MALDI analysis, the peptide mixtures were desalted and concentrated using 2 μm C₁₈ Ziptip (Millipore, Billerica, MA) with a final elution of 5 μL of 50% ACN with 0.1% TFA. Sample spotting was performed by layering 0.5 μL of matrix on top of a 0.5 μL portion of purified peptide solutions. The MALDI matrix was prepared by diluting saturated CHCA (Sigma) solution made in 50% v/v ACN and 1% v/v TFA with the same solution at a 1:4 ratio v/v. For internal calibration, 50 fmol/mL of Angiotensin I, Adrenocorticotrophic hormone (ACTH) 1–17 and ACTH 18–39 (Sigma) were introduced in the matrix as internal standards.

2.5 MALDI-TOF-MS

MALDI-MS was performed on a Micromass ToFSpec2E (Micromass/Waters, Milford, MA) with delayed extraction in the reflectron mode using a nitrogen laser (337 nm). The operating voltage was 20 kV with 18-kV extraction voltage. The pulse voltage was set in the range of 2100–2500 V depending on the peptide masses under investigation. The

delay time was 520 ns and the reflectron was at a 1:3 ratio to the source voltage. The tryptic peptide masses were measured over a range of 800–4000 Da. Three internal standards were used to calibrate the instrument with typical 50 ppm mass accuracy or less. The monoisotopic peptide mass list was submitted to the Swiss-Prot database using the MS Fit search engine (<http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm>) for protein identification. The search was performed allowing one missed cleavage and possible modifications such as oxidation and phosphorylation. The molecular weight ranged from 5000 to 100 000 and there was no restriction on the pH range. A strong hit (an identification with high confidence) meets the following criteria: (i) MOSWE score is at least on the order of 10^4 with five minimum matched peptides; (ii) the protein is ranked first or second by the database search engine; (iii) the percentage of coverage is at least 30%; and (iv) the matched peptides include the most abundant peak in the MALDI spectrum.

2.6 LC-MS/MS

In order to confirm the MALDI-TOF results, tandem MS experiments were performed using the linear IT MS (LTQ, Thermo Finnigan, San Jose, CA). The capillary transfer tube was set at 200°C and ESI voltage at 4.0 kV. A sheath gas flow of 12 arbitrary units was used. The ion activation was achieved by utilizing helium at normalized collision energy of 35%. All MS/MS data obtained were analyzed using the TurboSequest feature of Bioworks 3.1 SR1 (Thermo Finnigan). By allowing up to two missed cleavages, peptide ions were automatically assigned with the Xcorr values to consider >3.5 for +3 ions, >2.5 for +2 ions, and >1.5 for +1 ions and the ΔC_n of 0.1 or higher. Data and results for the identification of Annexin IV (ANXA4) are included in Supplementary Fig. 5 and Table 1.

2.7 Immunohistochemistry

An existing ovarian carcinoma tissue microarray [8] containing 10 clear cell and 25 serous ovarian carcinomas from the Surgical Pathology archives of the University of Michigan Health System was immunohistochemically stained with anti-ANXA4 antibody (1:200 dilution, Santa-Cruz Biotech). Antigen-antibody complexes were detected with the avidin-biotin peroxidase method using 3,3'-diaminobenzidine as a chromogenic substrate (Vector Laboratories). Immunostained sections were lightly counterstained with hematoxylin and then examined by light microscopy.

2.8 Statistical analysis

Protein abundances for each sample were first log (base 2) transformed and summed with consecutive, non-overlapping windows of width 2×10^4 PPM. That is, starting at a lower mass bound of 5000 Da, the first window spanned from 5000 to $5000(1 + 2 \times 10^4/10^6)$ Da = 5100 Da, the second

window spanned from 5100 Da to $5100(1 + 2 \times 10^4/10^6)$ Da = 5202 Da, and so on. At a typical protein mass of 50 kDa, the mass window was 1000 Da wide, with wider windows at greater masses and narrower windows at lesser masses. The reason for constructing the windows in this way is that absolute measurement error is greater for larger proteins, and larger proteins are presumed to have greater mass variation due to PTM.

Next, two-sample *t*-tests were performed within each window, to compare the mean expression level in the clear cell group from the mean expression level in the serous group. A significant difference was defined when the *t*-test *p*-value was less than 0.01 and when the levels within each group had standard deviation greater than 10^{-6} . For display purposes (e.g. Fig. 1), bands with significantly greater expression in clear cell samples were colored pink and bands with significantly greater expression for serous samples were colored cyan. Bands in which a nonzero abundance was detected in every sample were colored yellow.

Principal component analysis (PCA) and complete 5000 PPM-linkage hierarchical clustering were performed using the covariance matrix of the window-level abundance data as described above. Pearson correlation coefficients between the window-level aggregated masses were used in the hierarchical clustering. All mass windows were used in both analyses.

3 Results and discussion

3.1 2-D liquid phase mass mapping

A 2-D liquid phase mass mapping method has been developed in our laboratory and for the first time applied to profile protein expression in primary tumor tissues (Fig. 2). Proteins extracted from two different histological types of ovarian carcinoma samples (clear cell and serous carcinomas) were separated by utilizing this 2-D liquid phase separation method in order to identify proteins differentially expressed in serous versus clear cell carcinomas. This analytical approach fractionated proteins by CF in the first dimension based on *pI*, and each pH fraction was further separated by nonporous (NPS)-RP-HPLC in the second dimension. The protein eluents from NPS-RP-HPLC were interfaced online to ESI-TOF to obtain an accurate and reproducible protein intact molecular weight. The 2-D mass maps of each tissue were finally generated by the 2-D liquid phase separation method, which also allows for highly accurate protein profile comparisons between the two different histological types of ovarian cancer. Figure 1 shows the 2-D comparison mass map of 16 clear cell and 8 serous ovarian cancer samples generated from a selected pH fraction (5.4–5.2), where the clear cell samples are labeled with a -C and the serous with a -S. The horizontal and vertical dimensions of the 2-D protein content map are plotted in terms of the selected *pI* range and protein intact MW, respectively. Each vertical lane on the

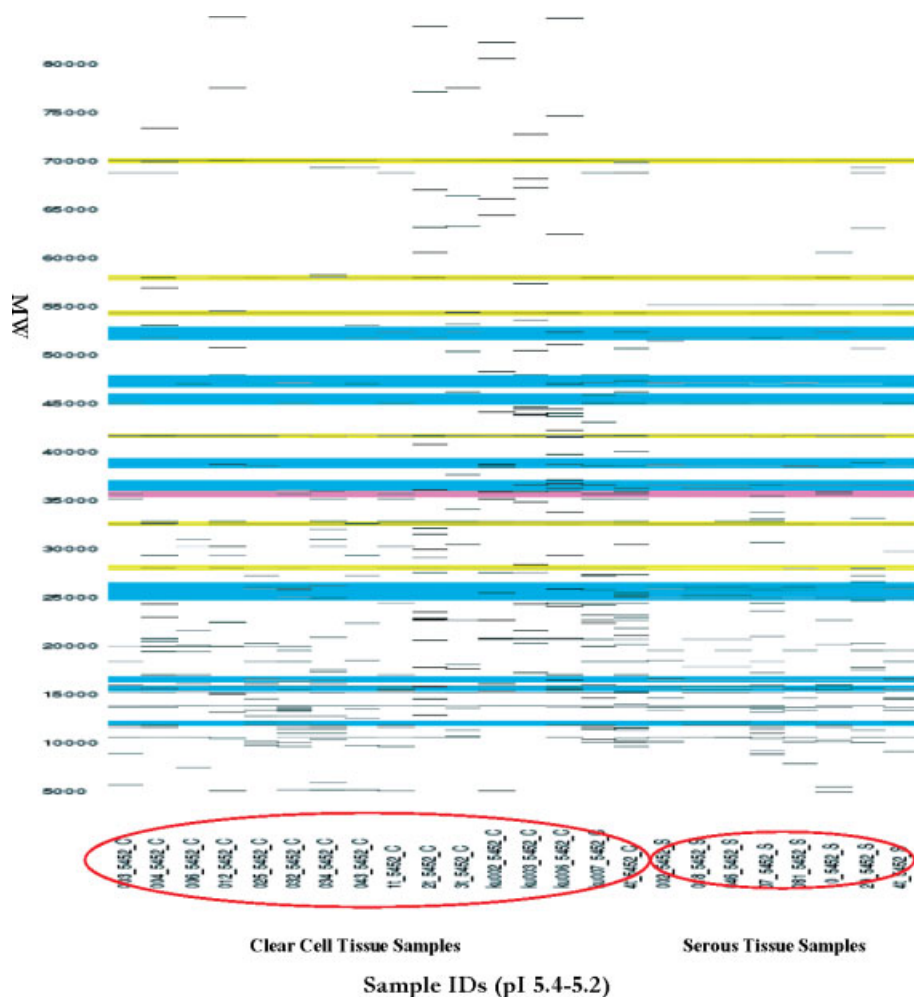


Figure 1. Comparison of a selected *pI* range (pH 5.4–5.2) in 2-D mass map of 16 clear cell and 8 serous ovarian carcinomas. Each lane represents a sample and the intensity of each band is proportional to the protein's concentration at a given MW value. Differentially expressed proteins are highlighted in different colors.

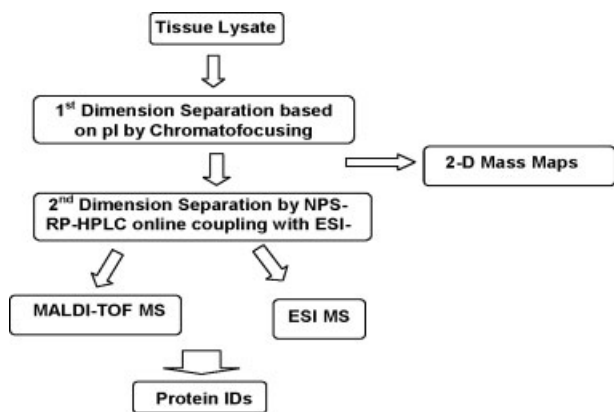


Figure 2. Experimental work flow of 2-D liquid phase mass mapping method followed by protein identifications with PMF analysis and LC-MS/MS sequencing.

map represents the protein content of the same selected pH fraction collected from each sample. Each band represents each protein eluted from RP-HPLC and detected by

ESI-TOF. For identification of differentially expressed proteins between the clear cell and serous sample groups, the protein mass range was divided into windows of 20000 PPM width. That is, a window that starts at K Daltons ends at $K + K/20$ Daltons. This window was chosen to include possible isoforms due to posttranslational modifications. Other windows of 10000 PPM and 5000 PPM were also tested and, generally, similar results were obtained. A non-overlapping sequence of windows was constructed covering the entire measured range. Within each of these windows, the total measured protein amount was summed for each sample, and these sums were compared between the 8 serous samples and the 16 clear cell samples using two-sample *t*-tests. Figure 1 shows the result of this analysis for one selected *pI* range. Pink bands in the figure correspond to mass ranges that are significantly more abundant in clear cell samples. Blue bands correspond to mass ranges that are significantly more abundant in serous samples. Yellow bands correspond to mass ranges in which all 24 samples have some level of measured protein expression.

Taking a more detailed look at differential expression, Fig. 3 shows the total measured protein per sample for two selected windows identified as differentially expressed using the *t*-test analysis. The expression level of each protein is calculated based on the total ion count of each protein peak detected by ESI-TOF. The two adjacent pH fractions were checked and the peak areas of any proteins that may cross over two or three fractions were combined. Normalization was employed by two adjustments based on sample size (the total amount of protein injected from each tissue sample) and internal standard response (the quantity of bovine insulin internal standard added). One of the selected windows shows a protein band that is more abundant in clear cell samples, and the other shows a protein band that is more

abundant in serous samples. Note that in both cases, the sample class with lower expression actually has zero expression for all but one sample, which expresses at an intermediate level compared to the other group.

3.2 Comparison of global expression patterns distinguishes clear cell from serous ovarian carcinomas

Previously, Schwartz *et al.* [8] reported that the clear cell subtype of ovarian carcinoma displays a distinct gene expression pattern among various subtypes of ovarian cancers. The protein expression patterns in 16 clear cell and 8 serous ovarian cancer samples were compared using the 2-D mass

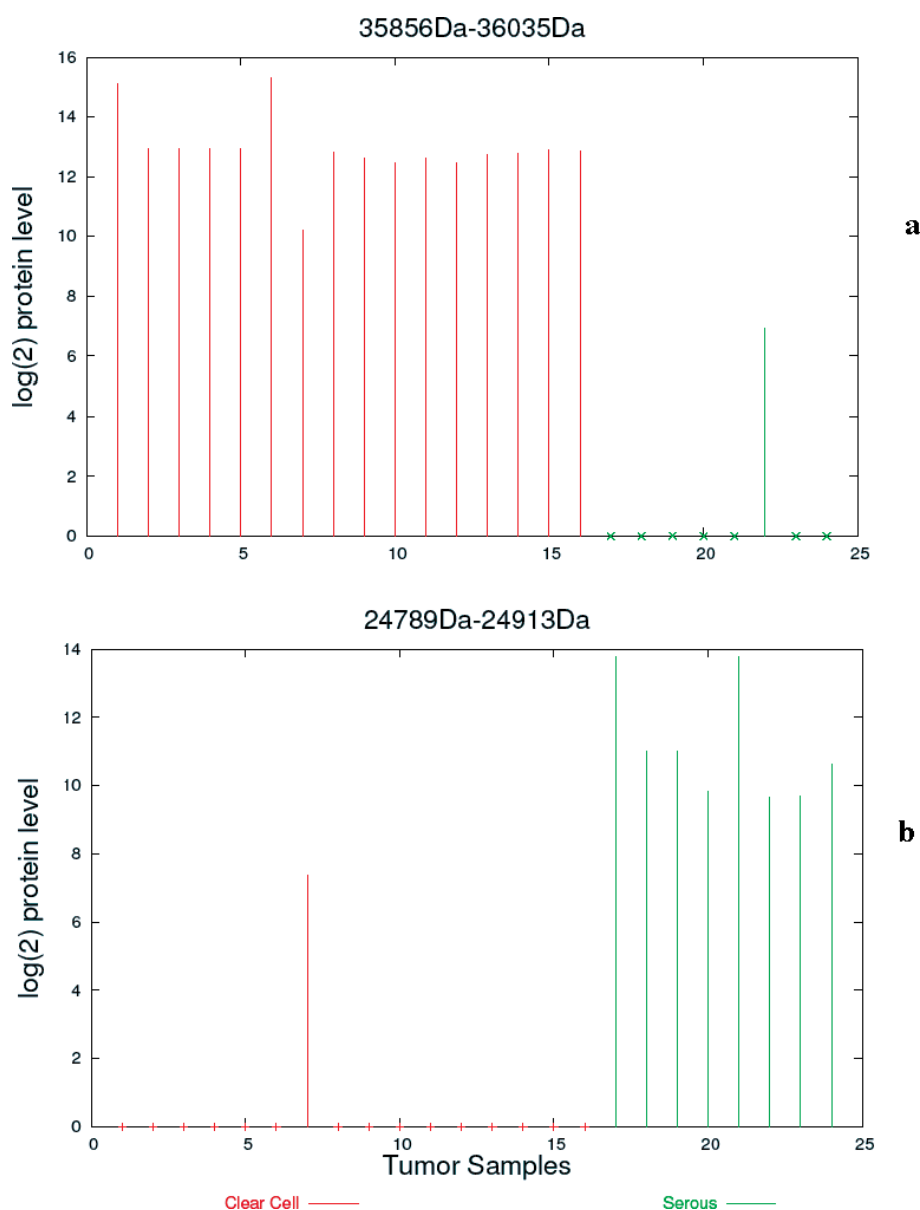


Figure 3. Comparisons of selected differentially expressed proteins (Annexin A4 (a) and Ras-related protein Rab-3D (b)) between clear cell and serous ovarian carcinoma samples.

mapping method to search for signatures that might differentiate the two groups. An alternative way of looking at the data is to focus on the interrelationships among the samples in terms of their overall patterns of protein expression, without aiming to identify specific bands of interest. To do this, again the protein amounts are summed within 20 000-PPM bands, and then log-transformed to reduce the skew in the data. In-house developed software based on PCA is then used to reduce the data to a set of points in the plane that can be displayed as a scatter plot (Fig. 4). Each point in this plot corresponds to one of the samples, and the points are distributed such that similar sample pairs fall closer together in the plot compared to less similar sample pairs (where similarity is based on all protein expression levels in all measured fractions). The points are colored to indicate the tissue type. The fact that the points corresponding to serous (green) and clear cell (red) samples are well separated indicates that the major source of variation in the data can be attributed to whether a sample is of clear cell or serous origin. Note that the PCA procedure does not have access to the tissue type information. Therefore, this finding is not due to over-fitting of the data. Obviously, differences in protein expression pattern between clear cell and serous ovarian tumor samples are able to discriminate two different histological types of ovarian cancer. As expected, this result is consistent with our previously published mRNA expression study, which indicated that the clear cell subtype of ovarian carcinomas displays a distinct gene expression signature in comparison with other types of ovarian cancers [8].

3.3 Hierarchical clustering analysis

An alternative method that has similar goals as PCA is hierarchical clustering, as shown in Fig. 5. This method aims to define a sequence of clusters of the samples, starting with a partition in which each sample is in its own cluster, and ending with a partition in which every sample is in the same

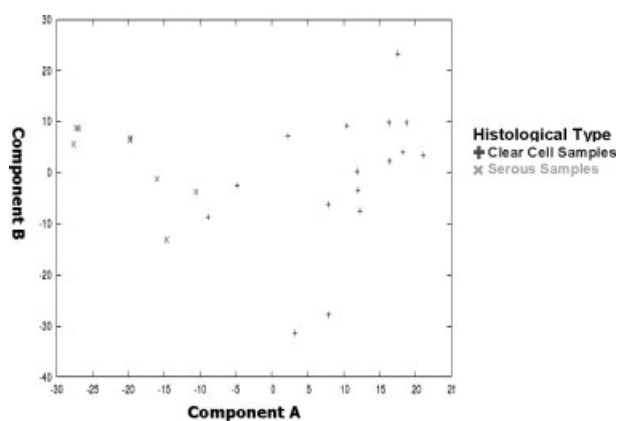


Figure 4. Principal component analysis (PCA) scatter plot with samples annotated with histological type. The plot clearly shows that clear cell and serous samples segregate out based upon the 2-D mapping information.

cluster. Greatest interest focuses on the large clusters obtained by making a vertical cut toward the right side of Fig. 5. It can be seen that the final step of the clustering algorithm (the rightmost join in the figure) is between a homogeneous group of serous samples and a homogeneous group of clear cell samples. This indicates, consistently with PCA, that the most important source of variation in the ovarian cancer data is histological type, and hierarchical clustering based on the protein expression of the 16 clear cell and 8 serous ovarian tissue samples results in complete separation of the two ovarian cancer subtypes. This result is similarly in concert with previously published statistical results based on a gene expression comparison among different histological types of ovarian cancers [17].

3.4 Protein identification

The 2-D liquid phase protein separation is readily interfaced to protein identification by mass spectrometry. Protein solution eluting from RP-HPLC is split into two streams. One is directly interfaced to ESI-TOF MS to obtain an intact protein MW value. The remainder is collected online by a fraction collector and further identified by PMF using MALDI-TOF MS and protein IDs of interest (fold change difference higher than 3) are also confirmed by LC/MS/MS. Proteins identified as potential histological-type-specific biomarkers and associated proteins from *pI* fractions pH 5.4–5.2 and 5.2–5.0 are summarized in Tables 2 and 3, respectively. These proteins may serve as type-specific diagnostic markers and some may ultimately allow treatment to be better tailored to each subtype of ovarian cancer. Of particular interest is ANXA4, which belongs to a large family of glycoproteins that bind both Ca^{2+} ions and negatively charged phospholipids. The annexins are involved in several important biological functions, such as exocytosis, anticoagulation, and anti-inflammation [18]. Furthermore, ANXA4 is a substrate for protein kinase C, which may act as regulator of membrane fusion and possesses the structural properties necessary to form ion channels [19–22]. These properties suggest that it may be involved in processes such as growth, differentiation and transformation [23]. It has also been shown that ANXA4 plays an important role in the morphological diversification and dissemination of clear cell renal carcinomas [24]. We found that ANXA4 was highly expressed in all of 16 clear cell carcinoma samples, but in very few serous tumor samples. In order to validate our proteome profiling data, we used immunohistochemical staining to evaluate expression of ANXA4 in an independent set of ovarian carcinomas represented in an ovarian cancer tissue microarray (details of the TMA provided in reference [8]). As expected, immunohistochemical staining confirmed markedly increased expression of ANXA4 in clear cell compared to serous carcinomas (Fig. 6).

These findings suggest that clear cell and serous subtypes of ovarian cancers have different pathogenesis pathways and that ANXA4 should be examined further to deter-

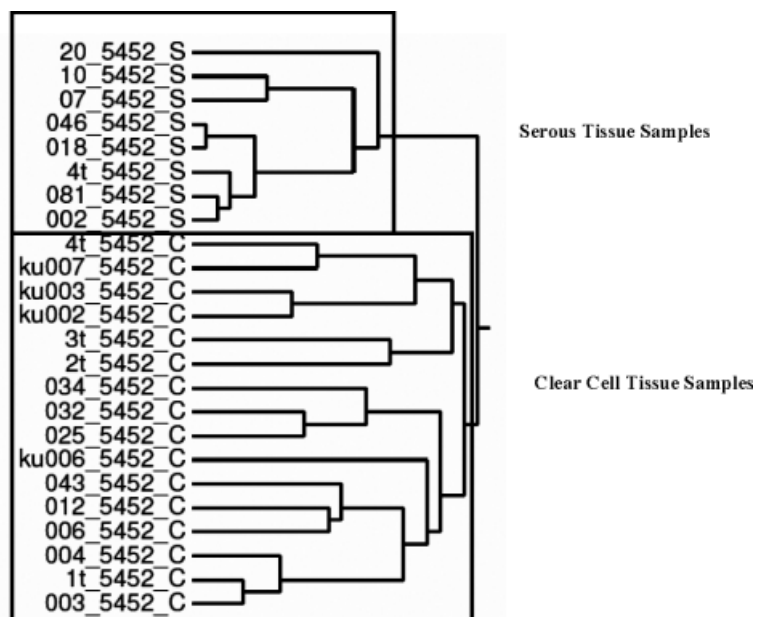


Figure 5. Hierarchical clustering dendrogram of clear cell and serous ovarian carcinomas.

Table 2. Differentially expressed proteins from a selected *pI* fraction of pH 5.4–5.2 are identified by PMF with MALDI-TOF

Protein	Access no.	Theoretical MW	Exp. MW	% Change in expression (Clear cell/serous)	MOWSE score	Sequence coverage %
Annexin A4 (Annexin IV)	P09525	35 883	35 995	+300 (\pm 55)	1.51E+09	58
ATP synthase beta chain (mitochondrial precursor)	P06576	56 560	55 199	-289 (\pm 81)	9.85E+08	42
Alpha enolase	P06733	47 169	47 123	-154 (\pm 47)	5.22E+09	58
L-Lactate Dehydrogenase M chain	P00338	36 689	36 600	-95 (\pm 32)	2.83E+05	40
Extracellular superoxide dismutase [Cu-Zn] precursor	P08294	25 881	26 120	-199 (\pm 53)	4.12E+06	35
Ras-related protein Rab-3D	O95716	24 758	24 900	-233 (\pm 74)	7.09E+06	32
Zinc finger protein 125 (HZF-3)	P35274	13 676	13 763	-65 (\pm 29)	7.88E+04	38

Table 3. Proteins from a selected *pI* fraction of pH 5.2–5.0, which are found highly expressed in serous ovarian carcinoma are identified by PMF with MALDI-TOF

Protein	Access no.	Theoretical MW	Exp. MW	% Change expression (Serous vs. clear cell)	MOWSE score	Sequence coverage %
Keratin, type II cytoskeletal 8	P05787	53 675	53 617	+14 (\pm 5)	2.11E+10	61
Keratin, type II cytoskeletal 7	P08729	51 418	51 841	+61 (\pm 28)	4.54E+08	55
Keratin, type I cytoskeletal 18	P05783	48 058	47 969	+9 (\pm 4)	8.14E+06	43
Wilm's tumor 1-associating protein	Q15007	44 244	41 661	+155 (\pm 21)	6.15E+08	42
Activator 1 36 kDa subunit (Replication factor C 36 kDa subunit)	P40937	38 497	36 336	+22 (\pm 10)	7.06E+06	30
Proto-oncogene C-crk (P38)	P46108	33 872	32 569	+65 (\pm 27)	3.27E+04	29
Spindlin (Ovarian cancer-related protein)	Q9Y657	27 077	25 813	+74 (\pm 11)	5.33E+04	35
ATP synthase D chain, mitochondrial (My032 protein)	O75947	18 491	18 403	+3 (\pm 2)	2.16E+07	46
Group X secretory phospholipase A2 precursor	O15496	17 132	17 214	+2.5 (\pm 1)	2.42E+03	27

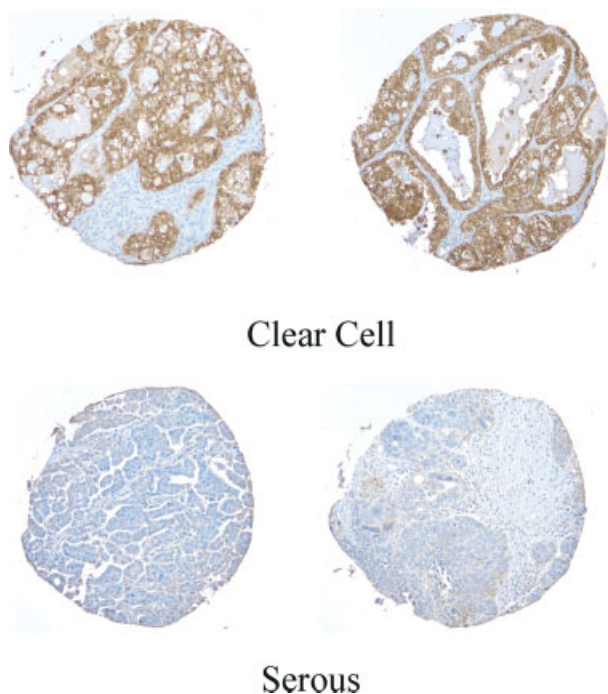


Figure 6. Immunohistochemical analysis of ANXA4 expression in primary ovarian carcinomas. Staining of representative clear cell and serous carcinomas is shown.

mine its utility as a marker of clear cell carcinomas. Clear cell adenocarcinoma of the ovary is also known as a strongly chemoresistant tumor in contrast to most of the ovarian surface epithelial carcinomas [25–28]. Notably, ANXA4 has been implicated in drug resistance after exposure of cells to paclitaxel [29].

We also identified several proteins highly expressed in the serous subtype compared with clear cell ovarian tumor samples (Table 2). WT-1 (Wilm's tumor 1), for example, has been described as a highly sensitive marker of serous ovarian carcinoma [30]. This protein was not expressed in the clear cell subtype, while it was highly expressed in the serous cases. WT-1 is a DNA-binding protein, located in the nucleus, which plays a critical role in the development of the genitourinary tract. Recent microarray gene expression studies confirmed our findings [17], and, using immunohistochemistry, we have previously confirmed [8] high expression of WT-1 protein in serous ovarian carcinomas compared to clear cell carcinomas.

Cytokeratin 7, cytokeratin 8 and cytokeratin 18 have both structural and signaling functions and have been identified as being differentially expressed in the serous cancers. Expression of cytokeratins in human tumors has been extensively studied and some researchers have demonstrated that different cell types in human breast tissues express a different array of cytokeratins [31–35]. The Ras-related protein Rab-3D, which plays a critical role in regulating the eukaryotic cell cycle as a shuttling factor through the nuclear

pore complex [36] or membrane transport mechanisms [37], was detected in high abundance only in serous samples. Proto-oncogene C-crck (P38), which is involved in phagocytosis of apoptotic cells and cell motility, was also found to be highly expressed in serous ovarian tumor samples. Various data indicate a role for Crk in growth factor-stimulation [38, 39]. Spindlin is another differentially expressed protein found in the serous carcinomas, with a potential role in cell-cycle regulation during the transition from gamete to embryo. It has been submitted to the EMBL/GenBank/DDBJ databases as an ovarian cancer-related protein by W. Yue and colleagues [40]. Notably, cytokeratin 7, cytokeratin 8, cytokeratin 18, Ras-related protein Rab-3D and proto-oncogene C-crck (P38) were relatively poorly expressed (not detected by ESI) in the clear cell cancers compared to serous tumors. These three types of keratin were consistently detected in 8 serous samples but not in most of clear cell samples, where all 24 tissue samples were processed at one time. They were previously proposed as biomarkers of ovarian cancer [13, 41, 42] and it is unlikely that they are due to contamination. Our findings suggest that these markers are subtype-specific and clear cell ovarian cancers should be considered separately in studies of these unregulated proteins as a diagnostic test for ovarian carcinomas.

Finally, we note that some of the proteins found to be differentially expressed between clear cell and serous ovarian carcinomas would not have been identified through mRNA-based gene expression profiling. We used our existing ovarian cancer oligonucleotide microarray data to examine mRNA expression corresponding to several of the differentially expressed proteins listed in Tables 2 and 3. Some genes, such as ANXA4, WT-1 and HZF-3, show concordant differential expression of the corresponding mRNA. For HZF-3, high expression in serous carcinomas appears to be driven by only a subset of the tumors, while expression of ANXA4 and WT-1 is more uniformly increased in clear cell and serous carcinomas, respectively. Notably, some genes such as SPN, C-CRK (p38), and Rab-3D do not show differential expression of the corresponding mRNA (Fig. 7). Although not unexpected, this finding underscores the notion that gene expression profiling may fail to identify a substantial number of genes whose protein products play biologically significant roles in disease pathogenesis or could serve as clinically useful biomarkers. In their comparison of the proteome and transcriptome in human platelets, McRedmond and colleagues [43] noted that messages for approximately 30% of proteins could not be detected, possibly representing RNA destruction following protein synthesis. A number of other post-transcriptional and post-translational modifications (PTM) of mRNA and proteins, respectively, could contribute to discordance between the mRNA and protein data for specific genes. In contrast to the static nature of the genome, which is essentially identical in every cell of an organism, the proteome is dynamic, constantly changing and responding to internal and external stimuli [44]. Proteomic technologies also allow

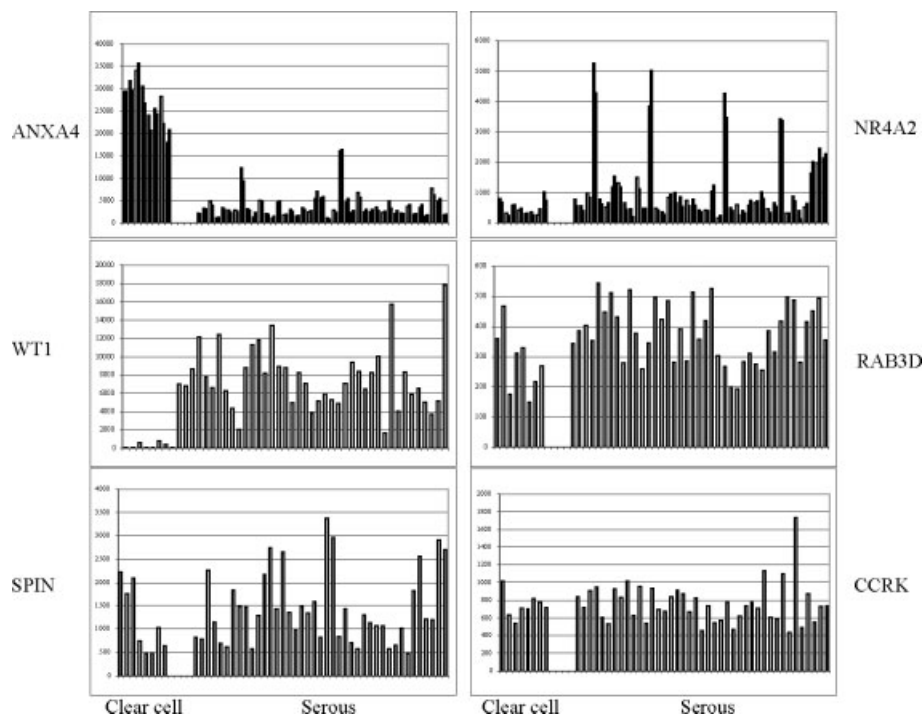


Figure 7. Relative RNA expression levels of ANXA4, HZF-3, WT1, RAB3D, SPIN and CCRK genes in primary ovarian clear cell and serous carcinoma samples using oligonucleotide microarray analysis. ANXA4 and HZF3 are represented by two probe sets in the microarray.

more accurate identification of the protein changes caused by the disease [45, 46]. At the protein level, distinct changes occur during the transformation of a healthy cell into a malignant cell, including PTM and protein activities, interactions and localization, all of which may affect cellular function. Identifying and understanding these changes is the underlying theme in cancer proteomics [47–49].

4 Concluding remarks

A 2-D liquid phase mass mapping method has been applied to explore variations in protein expression related to histological subtypes of ovarian cancer. The 2-D pI/mass maps of clear cell and serous ovarian tumor tissue samples were generated and used for interlysate comparison of protein expression between the different samples. Two histological types of ovarian tumor tissues were successfully distinguished based on protein expression patterns. In this study, differentially expressed proteins between the two subtypes were identified by PMF as promising candidates of type-specific biomarkers, which can also be utilized as a comprehensive protein expression signature for each subtype. Recognition of such a protein expression signature may ultimately allow therapeutic approaches to be better tailored to the characteristics of each tumor type. Identification of these type-specific markers should eventually facilitate future studies aimed at improving the diagnosis and treatment of ovarian cancers.

This work was supported in part by the National Cancer Institute under grant R01CA10010 (DML, KRC), R01CA90503 (DML, FRM) and the National Institutes of Health under grant R01GM49500 (DML). Eprogen, Inc also generously provided support. We would like to thank Dr. Christian G. Huber (Saarland University, Saarbrücken, Germany) for providing the monolithic capillary columns.

5 References

- [1] Jemal, A., Tiwari Ram, C., Murray, T., *Cancer J. Clinic.* 2004, **54**, 8–29.
- [2] Hoskins, W. J., Perez, C. A., Young, R. G. (Eds.), In: *Principles and Practice of Gynecologic Oncology*, Lippincott, Williams and Wilkins, Philadelphia 2001, pp. 165–182.
- [3] Aunoble, B., Sanches, R., Didier, E., *Int. J. Oncol.* 2000, **16**, 567–576.
- [4] Sugiyama, T., Kamura, T., Kigawa, J., Taguchi, K., *Cancer (Phila.)* 2000, **88**, 2584–2589.
- [5] Omura, G. A., Brady, M. F., Homesley, H. D., Yordan, E. *et al.*, *J. Clin. Oncol.* 1991, **9**, 1138–1150.
- [6] Tammela, J., Geisler, J. P., Eskew, P. N., Jr., Geisler, H. E., *Eur. J. Gynaecol. Oncol.* 1998, **19**, 438–440.
- [7] Feeley, K. M., Wells, M., *Histopathology* 2001, **38**, 87–95.
- [8] Schwartz, D. R., Kardia, S. L. R., Shedden, K. A., *Cancer Res.* 2002, **62**, 4722–4729.
- [9] Hough, C. D., Sherman-Baust, C. A., Pizer, E. S., *Cancer Res.* 2000, **60**, 6281–6287.
- [10] Millea, K. M., Krull, I. S., Cohen, S. A., *J. Proteome Res.* 2006, **5**, 135–146.

- [11] Conrads, T. P., Anderson, G. A., Veenstra, T. D., *Anal. Chem.* 2000, **72**, 3349–3354.
- [12] Smith, R. D., Anderson, G. A., Lipton, M. S., *Proteomics* 2002, **6**, 61–90.
- [13] Wang, H., Kachman, M. T., Schwartz, D. R., *Proteomics* 2004, **4**, 2476–2495.
- [14] Lubman, D. M., Kachman, M., Parus, S. J., *US Patent Applied Publications* 2003, 59.
- [15] Yan, F., Subramanian, B., Nakeff, A., *Anal. Chem.* 2003, **75**, 2299–2308.
- [16] Zhu, K., Kachman, M. T., Miller, F. R., Lubman, D. M., Zand, R., *Journal of Chromatography A* 2004, **1053**, 133–142.
- [17] Schaner, M. E., Ross, D. T., Ciaravino, G., *Mol. Biol. Cell* 2003, **14**, 4376–4386.
- [18] Sohma, H., Ohkawa, H., Hashimoto, E., *Alcoholism: Clin. Exp. Res.* 2002, **26**, 44S–48S.
- [19] Kaplan, R., Jaye, M., Burgess, W. H., *J. Biol. Chem.* 1988, **263**, 8037–8043.
- [20] Kourie, J. I., Wood, H. B., *Progress Biophys. Mol. Biol.* 2000, **73**, 91–134.
- [21] Ishitsuka, R., Kojima, K., Utsumi, H., *J. Biol. Chem.* 1998, **273**, 9935–9941.
- [22] Sohma, H., Creutz, C. E., Saitoh, M., *Biochem. J.* 1999, **341**, 203–209.
- [23] Bergman, A.C., Benjamin, T., Alaiya, A., *Electrophoresis* 2000, **21**, 679–686.
- [24] Zimmermann, U., Balabanov, S., Giebel, J., *Cancer Lett. (Amsterdam, Netherlands)* 2004, **209**, 111–118.
- [25] Itamochi, H., Kigawa, J., Sugiyama, T., *Obstetrics Gynecol.* 2002, **100**, 281–287.
- [26] Kigawa, J., Itamochi, H., Sugiyama, T., *Recent Res. Develop. Cancer* 2002, **4**, 493–506.
- [27] Itamochi, H., Kigawa, J., Sultana, H., *Japanese J. Cancer Res.* 2002, **93**, 723–728.
- [28] Itamochi, H., Kigawa, J., Akeshima, R., *Oncology* 2002, **62**, 349–353.
- [29] Han, E. K., Tahir, S. K., Cherian, S. P., *British J. Cancer* 2000, **83**, 83–88.
- [30] Hwang, H., Quenneville, L., Yaziji, H., *Appl. Immunohistochem. Mol. Morphol.* 2004, **12**, 122–126.
- [31] Franzen, B., Linder, S., Alaiya, A. A., *British J. Cancer* 1996, **74**, 1632–1638.
- [32] Franzen, B., Linder, S., Alaiya, A. A., *Electrophoresis* 1997, **18**, 582–587.
- [33] Williams, K., Chubb, C., Huberman, E., *Electrophoresis* 1998, **19**, 333–343.
- [34] Giometti, C. S., Tollaksen, S. L., Chubb, C., *Electrophoresis* 1995, **16**, 1215–1224.
- [35] Trask, D. K., Band, V., Zajchowski, D. A., *Proc. Natl. Acad. Sci. USA* 1990, **87**, 2319–2323.
- [36] Weis, K., *Cell* 2003, **112**, 441–451.
- [37] Rosenfeld, J. L., Knoll, B. J., Moore, R. H., *Receptors and Channels* 2002, **8**, 87–97.
- [38] Hempstead, B. L., Birge, R. B., Fajardo, J. E., *Mol. Cell. Biol.* 1994, **14**, 1964–1971.
- [39] Beitner-Johnson, D., LeRoith, D., *J. Biol. Chem.* 1995, **270**, 5187–5190.
- [40] Gao, Y. H., Qin, L. P., Yue, W., *Progress Natural Sci.* 2004, **14**, 966.
- [41] Kachman, M. T., Wang, H., Schwartz, D. R., *Anal. Chem.* 2002, **74**, 1779–1791.
- [42] Wang, H., Kachman, M. T., Schwartz, D. R., *Electrophoresis* 2002, **23**, 3168–3181.
- [43] McRedmond, J. P., Park, S. D., Reilly, D. F., *Mol. Cell. Proteomics* 2004, **3**, 133–144.
- [44] Stierum, R., Gaspari, M., Dommels, Y., *Biochimica et Biophysica Acta, Proteins and Proteomics* 2003, **1650**, 73–91.
- [45] Chaurand, P., DaGue, B. B., Pearsall, R. S., *Proteomics* 2001, **1**, 1320–1326.
- [46] Kolch, W., Mischak, H., Pitt, A. R., *Clin. Sci.* 2005, **108**, 369–383.
- [47] Yang, X. J., *Oncogene* 2005, **24**, 1653–1662.
- [48] Pawson, T., *Cell* 2004, **116**, 191–203.
- [49] Lane, C. S., *Cell. Mol. Life Sci.* 2005, **62**, 848–869.