Searching for Potential Ovarian Cancer Biomarkers with Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Feng Qiu¹, Hong-Ying Liu¹, Zhen-Nan Dong¹, Ying-Ji Feng³, X-J Zhang¹,²*, Ya-Ping Tian¹*

¹ Department of Clinical Biochemistry, Chinese PLA General Hospital, 28 Fu-Xing Road, Beijing 100853, China.
² Department of Chemistry, University of South Florida, 4202 E. Fowler Ave., CHE 205A, Tampa, FL 33620-5250, USA.
³ Department of Chemistry, Tsinghua University, Beijing 100084, China.

*Corresponding Authors:
Department of Clinical Biochemistry, Chinese PLA General Hospital.
Phone: +86 10 66939374
Fax: +86 10 88217385
E-mail address: tianyp61@gmail.com or xzhang@cas.usf.edu

Received: 2 January 2009; | Revised: 18 January 2008; | Accepted: 19 January 2009

Abstract

Ovarian cancer is a common gynecological malignant disease, causing more deaths among women. The key objective in the treatment of ovarian cancer is early diagnosis. The objective of our study was to seek new ovarian cancer biomarkers based on a serum protein profile with the aim of discriminating ovarian cancer patients from healthy controls. An MB-WCX kit was used to analyze serum samples obtained from 20 ovarian cancer patients and 20 healthy controls and then we generated MALDI-TOF protein profiles from the analysis. After pre-processing of the spectra, linear analysis with ClinProTools bioinformatics software was used to classify protein profiles and search for prominent peaks that could be used as potential ovarian cancer biomarkers. Using ClinproTools bioinformatics and statistical software, we found 5 prominent expressed proteins in the ovarian cancer and healthy control groups. The mass to charge ratio were 4648.21(m/z), 9294.03(m/z), 3886.1(m/z), 9066.38(m/z) and 4254.71(m/z), respectively, and the former four proteins were expressed higher in the ovarian cancer patients, but the later one was expressed at lower levels in the cancer patients. The sensitivity and specificity were both more than 90%. From our study, we found that MALDI-TOF MS is a high-throughput sample preparation method and is a new potential tool for the diagnosis of human disease, not only to search for new early detection biomarkers in the ovarian cancer patients’ serum samples, but also with a potential use for routine clinical work.

Keywords: ovarian cancer; proteomic analysis; MALDI-TOF; MS biomarkers
1. Introduction

Ovarian cancer is one of the three most malignant diseases in the female reproductive system, with gynecological malignancies a leading cause of death [1]. It remains the third most frequent gynecological neoplasm, and is associated with the highest mortality rate in the developed countries. There has been limited improvement in patient renews despite therapeutic advances. The incidence of ovarian cancer has risen in recent years. In ordinary, healthy women, there is about a 1.4% occurrence of this disease, but in some cases the prevalence is significantly higher. Women with early-stage ovarian cancer have a 5-year survival rate of over 80%. The poor prognosis is mainly attributed to the detection of cases existing at a late stage [2], as a result of the nonspecific symptoms. Patients are usually not diagnosed until the disease has spread beyond the pelvis or has involved in other organs to cause serious discomfort. At a more advanced stage of increased ovarian disease, as the mortality rate reaches, the prognosis is correlated with late diagnosis [3]. Up to 70% of the cases are detected at mortality for this reason [4,5]. Currently, there is no early diagnostic test with high sensitivity, specificity, and positive predictive value, which can be used as a routine clinical screening tool. Therefore, there is a need for new biomarkers for ovarian cancer that can improve early mortality to 70% within two years and about 90% within five years [6]. Thus, early detection and early diagnosis remain to be the most appropriate ways to reduce disease-related diagnosis, to monitor disease progression, to observe therapeutic responses and to detect disease recurrence [7].

Recently, proteomic and bioinformatics approaches have been shown to be able to investigate the serum proteome and identify signature biomarker patterns of different origin cancers such as blood, breast, and prostate etc [8-14]. The application of this approach, if validated for its sensitivity and reproducibility, may influence the early diagnostic, detection, and therapeutic decisions.

Many new technologies have emerged over the past few years, which can supply and enhance the proteomic research. Proteomic pattern analysis by mass spectrometry is one of the most promising new approaches for the identification of potential blood biomarkers and to distinguish health vs. disease [15-16]. The discovery of biomarkers in biological fluids has been advanced by the use of MS-based screening methods, such as surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) MS and matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS [17].

The mass spectrometer can be considered as a highly accurate weighing scale for extremely low mass particles such as proteins or peptides. This makes protein profiling with high-throughput sample preparation become to be true, and MALDI-TOF MS analysis becomes a new potential tool for the diagnosis of human disease [18]. In MALDI-TOF MS analysis, the proteins or peptides are liberated in an ionized form from the target surface through firing a laser pulse at the crystallized form. The ionized proteins are accelerated through a vacuum tube by an electrical field and reach a detector. Smaller proteins are accelerated to higher velocities than heavier proteins, and the “time of flight” is proportional to \( m/z \), which means the mass to charge ratio.

The objective of this study was to assess the feasibility of applying mass spectrometry for the profiling of serum proteins in ovarian cancer patients from healthy control individuals, to find some new early potential biomarkers, to assist in carrying out the early detection of ovarian cancer.

2. Patients and methods

2.1 Patients.
Serum samples were obtained from a total of 20 ovarian cancer patients the day before surgery. Surgical specimens and pathological validation historically confirmed ovarian cancer. The median age of the female patient group was 52 years old, with patients admitted from November 2006 until May 2007 in our hospital. The healthy control volunteers group consisted of 20 individuals. The median age of healthy female control group was 50 years old, and all were recruited from our hospital healthy check center.
2.2 Chemicals and Calibrators

Gradient-grade alcohol and acetone were obtained from J.T. Baker; p.a. trifluoroacetic acid and Ammonium acetate were purchased from Sigma-Aldrich. The peptide calibrator containing Angiotensin II, the protein calibrator, and the matrix α-Cyano-4-hydroxycinnamic acid (HCCA) were purchased from Bruker Daltonics GmbH. For magnetic bead preparations, we used 0.2 mL of polypropylene tubes (8-tube strips) from Axygen Company. Multifly needle sets and polypropylene vialute without anticoagulants (EDTA, heparin, and citrate) were obtained from BD Company.

2.3 Blood samples

The blood samples were processed according to a standardized protocol. After sample collection, the vacuette serum (collected in a red-top glass tube containing no preservatives or anticoagulant) were allowed to clot or to sediment at room temperature (about 25°C) for at least half an hour and then centrifuged at 1500g for 15 min. Then, serum samples were divided in aliquots immediately and frozen (100μl at -80°C) until use; so for proteomic fractionation, samples were thawed at room temperature (about 25°C) for 15 min and processed immediately.

2.4 Proteomic fractionation (Isolation of peptides)

Serum samples were subjected to fractionate by using functionized magnetic bead-based MB-WCX kits (ClinProt Kits, Bruker Daltonics Inc., Fremont, CA). These magnetic beads (particle size, 1 mm), exhibited with weak cation exchange chromatography (WCX). These were designed for capturing of proteins and peptides from biological samples in order to obtain specific profiles and were used according to the manufacturer’s protocols. As recommended in the manufacturer’s protocol, we added 5 μL WCX magnetic beads with 10μL of a binding solution in a 0.2mL polypropylene PCR tube, then added 5μL of serum sample and mixed thoroughly by pipetting up and down several times; incubated the tube for 5 minutes, and separated the unbound solution using a magnetic bead separator. After magnetic bead separation and three times washing, the bound proteins/peptides were eluted from the magnetic beads according to the manufacturer’s instructions (Profiling Kit WCX, Bruker Daltonics GmbH, Germany). A portion of the eluted sample was diluted 1:10 in a solution containing HCCA (0.6g/L in 2:1 ethanol: acetone). Then, 1μL of the resulting mixture was spotted onto the AnchorChip target (Bruker Daltonics Inc., CA), and allowed to air dry for approximately 5 min at room temperature. Usually, each sample was conducted three times.

2.5 MS analysis (Protein profiling)

For the proteomic analysis, we used a linear MALDI-TOF mass spectrometer (Microflex; Bruker Daltonics) with the following settings: ion source 1, 20 kV; ion source 2, 18.40kV; lens, 7.50 kV; pulsed ion extraction, 120 ns; nitrogen pressure, 1700-2000 mbar. Ionization was achieved by irradiation with a nitrogen laser (λ=337 nm) operating at 25 Hz.

And the mass calibration was performed with the calibration mixture of peptides and proteins in a mass range of 1000–10000 Da. We measured 3 MALDI preparations (MALDI spots) from each magnetic bead fraction. For each MALDI spot, 400 spectra were acquired in analysis (50 laser shots at 8 different spot positions).

To increase the detection sensitivity, we usually use higher energy to shot the spot positions at first and then lower energy to shot again and then save the spectra, which can remove excess matrix with higher laser power before data acquisition.

2.6 Data processing

The spectra of all signals with a signal-to-noise (S/N) ratio >5 in a mass range of 1000–10000 Da obtained from all of the samples, were analyzed by using ClinProTool 2.1 software for spectrum processing and generation of proteomic fingerprints. Only the spectra obtained from WCX magnetic beads were used in the analysis of this report. We used the ClinProTools bioinformatics software (Ver.2.1; Bruker Daltonics) for proteomic pattern recognition.

To increase the accuracy, we used the average of eight spots represent one serum sample. Basically, all spectra were normalized to their own
total ion count (summation of peak areas) and recalibrated using the most prominent peaks, followed by baseline subtraction, peak defining, and calculation of peak intensities and areas. Then, we calculated the mean value of peak intensity, SD and CV (%) for each corresponding peak among the ovarian cancer and healthy control groups. The degree of variation on the basis of the whole spectrum was determined by calculating the CV values for some prominent peaks of the samples. We defined p<0.01 as a significant difference.

3. Results

3.1 Unique proteins or peptides analysis between ovarian cancer group and healthy control group

In the mass range 1000-10,000 (m/z), we used the model established with ovarian cancer patients and healthy controls, compared the mass spectra, and drew the conclusions:

a. 152 peaks have been detected by ClinProTools software;

b. In which we found 5 differential prominent peaks with great statistical significance that is P<0.000001, and the mass to charge ratio were 4648.21 (m/z), 9294.03 (m/z), 3886.1 (m/z), 9066.38 (m/z), and 4254.71 (m/z), respectively (See Table 1).

<table>
<thead>
<tr>
<th>Peak index</th>
<th>m/z</th>
<th>PTTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>4648.21</td>
<td>&lt; 0.000001</td>
</tr>
<tr>
<td>150</td>
<td>9294.03</td>
<td>&lt; 0.000001</td>
</tr>
<tr>
<td>84</td>
<td>3886.1</td>
<td>&lt; 0.000001</td>
</tr>
<tr>
<td>149</td>
<td>9066.38</td>
<td>&lt; 0.000001</td>
</tr>
<tr>
<td>94</td>
<td>4254.71</td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

Table 1: Unique proteins or peptides in ovarian cancer and healthy control group

Figure 1: The average expression of 4648.21 (m/z) (red: ovarian cancer, green: healthy control)

In all five prominent peaks, the mass range of 4648.21 (m/z) (see Figure 1), 9294.03 (m/z) (see Figure 2), 3886.1 (m/z), and 9066.38 (m/z) were higher expressed in the ovarian cancer group, but the mass range of 4254.71 (m/z) (see Figure 3) was lower expressed in the ovarian cancer group (see Table 2) compared with the healthy control group.
Figure 2: The average expression of 9294.03 (m/z) (red: ovarian cancer, green: healthy control)

Figure 3: The average expression of 4254.71 (m/z) (red: ovarian cancer, green: healthy control)
### Table 2: Different expression proteins or peptides between ovarian cancer and healthy control group

<table>
<thead>
<tr>
<th>m/z</th>
<th>Average peak areas in ovarian cancer group</th>
<th>Average peak areas in healthy control group</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4648.21</td>
<td>447.16</td>
<td>117.08</td>
<td>0.00000927</td>
</tr>
<tr>
<td>9294.03</td>
<td>1936.02</td>
<td>610.95</td>
<td>0.000985</td>
</tr>
<tr>
<td>3886.1</td>
<td>37.96</td>
<td>11.64</td>
<td>0.0213</td>
</tr>
<tr>
<td>9066.38</td>
<td>122.09</td>
<td>40.96</td>
<td>0.00128</td>
</tr>
<tr>
<td>4254.71</td>
<td>15.49</td>
<td>39.39</td>
<td>0.0101</td>
</tr>
</tbody>
</table>

We analyzed the two group data with the genetic algorithm method (GA), compared the data between them, and got the sensitivity and specificity (see Table 3).

### Table 3: Data analysis with genetic algorithm method

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>sensitivity</th>
<th>specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>98.86%</td>
<td>100%</td>
</tr>
</tbody>
</table>

3.2 The distribution of ovarian cancer and healthy control group samples

With the use of two prominent proteins Pk101, 4648 (m/z) and Pk150, 9294 (m/z), we established the ovarian cancer model, by which we analyzed the sample distribution of two groups (see Figure 4).

Then, we performed all the samples distribution map of Principle Component Analysis (see Figure 5).

We analyzed the two group data with the genetic algorithm method (GA), compared the data between them, and got the sensitivity and specificity (see Table 3).

Figure 4: The distribution of ovarian cancer and healthy control samples (red: ovarian cancer, green: healthy control)
ratio of them were $P{\text{k}}101 \cdot 4648.21 \text{(m/z)}$, $P{\text{k}}84 \cdot 3886.1 \text{(m/z)}$, $P{\text{k}}31 \cdot 1213.79 \text{(m/z)}$, $P{\text{k}}127 \cdot 6635.1 \text{(m/z)}$ and $P{\text{k}}113 \cdot 5340.49 \text{(m/z)}$, respectively. The expression differences in two groups were as follows:

We find that $P{\text{k}}101 \cdot 4648.21 \text{(m/z)}$ is higher expressed in the ovarian cancer group than in the healthy control group, which means that it was abundant and highly expressed in the ovarian cancer disease group.

And we also analyzed the other four prominent proteins or peptides, and the results show that $P{\text{k}}84 \cdot 3886.1 \text{(m/z)}$ and $P{\text{k}}113 \cdot 5340.49 \text{(m/z)}$ expressed higher in the ovarian cancer group than in the healthy control group; but $P{\text{k}}113 \cdot 1213.79 \text{(m/z)}$ and $P{\text{k}}127 \cdot 6635.1 \text{(m/z)}$ were lower expressed in the ovarian cancer group (see Table 4).

Figure 5: Samples distribution of a Principle Component Analysis (red: ovarian cancer, green: healthy control)

<table>
<thead>
<tr>
<th>m/z</th>
<th>Ovarian cancer average peak areas</th>
<th>Healthy controls average peak areas</th>
<th>DAVe</th>
</tr>
</thead>
<tbody>
<tr>
<td>4648.21</td>
<td>447.16</td>
<td>117.08</td>
<td>+330.08</td>
</tr>
<tr>
<td>3886.1</td>
<td>37.96</td>
<td>11.64</td>
<td>+26.33</td>
</tr>
<tr>
<td>1213.79</td>
<td>3.02</td>
<td>6.71</td>
<td>−3.7</td>
</tr>
<tr>
<td>6635.1</td>
<td>31.71</td>
<td>107.7</td>
<td>−75.99</td>
</tr>
<tr>
<td>5340.49</td>
<td>425.49</td>
<td>136.68</td>
<td>+288.81</td>
</tr>
</tbody>
</table>

Table 4: Higher and lower expressed differential prominent proteins in genetic algorithm method.

The data results analyzed with the supervised neural network method were similar to those with the genetic algorithm method, so we did not show them again here. However, the results analyzed with the quick classifier method were of higher sensitivity and specificity, and are described as follows:

Unlike the genetic algorithm and supervised neural network method, analysis with quick classifier method uses 11 proteins as differential prominent ones to classify the two group data. The
added proteins were Pk93、4235.41 m/z, Pk94、4254.71 m/z, Pk115、5526.05 m/z, Pk134、7011.77 m/z, Pk139、7769.86 m/z, Pk147、8868.13 m/z, Pk149、9066.38 m/z and Pk150、9294.03 m/z. By adding more proteins or peptides, we can differentiate the ovarian cancer and healthy control group data better than before, and give more information for further optimization of the ovarian cancer model.

4. Discussion

For quite some time, there have been poor screening and early diagnosis methods for ovarian cancer. Patients are usually not diagnosed until the disease has spread beyond the pelvis or involved other organs to cause significant discomfort. A high survival rate of more than 85% has been seen with stage 1 disease. Therefore, we made great efforts to develop ovarian cancer screening tools in order to pick up the disease at an early stage. The completion of gene sequences necessitates an understanding of altered protein function. In addition to gene mutations, transcription or post-translation modifications may be reflected as changes in protein levels. Early tumors will express little but important changes in protein level. So, we employed a proteomics approach for the early diagnosis of malignant tumors. MALDI-TOF MS is a promising tool and being applied to discover disease-related proteomic patterns in complex mixtures of proteins derived from tissue samples or from easily obtained biological fluids such as serum, urine, nipple aspirate fluid, etc. Proteomic patterns can thus be used for early diagnosis, to predict prognosis, to monitor disease progression or response to treatment, or even to identify which patients are most likely to benefit from a particular treatment.

The fractionation of proteins and peptides from complex mixtures, for example serum, is the very important and fundamental step in the proteomic analysis. For this reason, at the very beginning of the research, we did the optimization of magnetic beads for our future analysis, and chose MB-WCX as our right kind [20].

By using ovarian cancer patients and healthy controls serum peptide mass fingerprints (PMF), and with the help of ClinProTools bioinformatics software and relevant statistical software, we found that only by using these two peptides Pk101、4648.21 m/z and Pk150、9294.03 m/z could we distinguish the disease group and healthy control group. We also found that the two differential prominent peptides are both expressed at a higher level in the ovarian cancer group. Based on these findings, we can enlarge our sample size in future research, and these two peptides may become new potential biomarkers for the early diagnosis of ovarian cancer.

Then, we made data analysis with the genetic algorithm method. By the use of one, two, or several combinations of proteins or peptides, we can distinguish the disease group and the healthy control group well. This result is in accordance with the usual proteomics research results. With the use of several proteins or peptides as a combination chip, the sensitivity and specificity will be higher than just the use of one or two.

The potential prominent proteins or peptides, though we have not yet to characterize them, through the mass to charge ratio (m/z), are combinations of 35-84 amino acids of poly-peptides.

Serum proteomics profiling with high-throughput methods, such as MALDI-TOF MS or SELDI-TOF MS, is a useful tool in cancer research. Both methods can enable rapid analysis of large-scale clinical individual serum samples, but the pre-analytical and analytic steps require thorough validation before clinical implementation can be warranted. In our study, we used MALDI-TOF MS instead of SELDI-TOF MS to analyze our cancer and healthy control samples. Several laboratories have demonstrated the feasibility of using mass spectrometric proteomic pattern analysis for the diagnosis of several categories of tumors, including breast, lung, pancreas and prostate cancer [21-38]. There has been considerable controversy concerning the SELDI profiling approach. The major concern is the platform reproducibility and the sensitivity to detect low-abundance serum protein markers. Similar to published findings, we found that MALDI-TOF MS seems to be more sensitive than SELDI-TOF MS. Although the process for calibrating the SELDI-TOF-MS instrument is explained in the manufacturer’s manual, the
importance of the process is not made very clear. An inadequate calibration can lead to a significant shift of the m/z values for the peak maximum. If this had occurred during our analysis, the sites using this poor calibration would have failed to classify proteins correctly. So before beginning the research, we have finished the reproducibility study with our MALDI-TOF MS, and confirmed the high reproducibility to continue the further cancer research.

In summary, we have established a reproducible, validated, and standardized approach to pre-analyze serum samples and optimized the conditions for serum sampling and storage, with the aim of obtaining high sensitivity MALDI-TOF MS measurements of altered proteins in diseases such as ovarian cancer.

But the disease and healthy control samples are so limited in our study; the pre-fractionation serum method is only using magnetic beads and analysis of mass spectra data, not separating the real makeup in the serum mixture samples. This method shows the use of routine clinical work has much to be examined and to validate down the road.

References

12. Adam, P. J., Boyd, R., Tyson, K. L., Fletcher, G. C., Stamps, A., Hudson, L., Poyser, H. R., Redpath, N., Grziñths, M., Steers, G., Harris,


