Serum peptidome patterns of hepatocellular carcinoma based on magnetic bead separation and mass spectrometry analysis

Ke Li, Wei-Hua He, Nai-Jun Fan, Chun-Fang Gao, Guang Zhao, Gao-Bin Sun, Guo-Bao Zheng
Institute of Anal-colorectal Surgery, No. 150 Central Hospital of PLA, Luoyang, China

ABSTRACT

Background/Aims: The only hope for a cure from hepatocellular carcinoma (HCC) rests on early diagnosis. The present study aims to determine serum peptidome patterns for early diagnosis of HCC.

Materials and Methods: To identify novel peptidome patterns for diagnosing HCC, serum from 31 healthy volunteers and 32 HCC patients were subjected to a comparative proteomic analysis using a ClinProt Kit combined with mass spectrometry (MS). This approach allows the determination of peptidome patterns that are able to differentiate the HCC from healthy volunteers. For further validation, the diagnostic and differential diagnostic capabilities of the peptidome patterns were verified blindly by an independent group of sera consisted of 31 HCC, 23 liver fibrosis and 33 healthy volunteers.

Results: A Quick Classifier Algorithm was used to construct the peptidome patterns for the identification of HCC from the control samples. One of the identified peaks at m/z 7771 was used to construct the peptidome patterns with almost 100% accuracy. Furthermore, the peptidome patterns could also differentiate the validation group with high accuracy.

Conclusion: These results suggest that the ClinProt Kit combined with MS achieves significantly high accuracy for HCC diagnosis and differential diagnosis.

Keywords: Hepatocellular carcinoma, proteomics, diagnosis, mass spectrometry

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most commonly diagnosed cancer in worldwide (1). Despite advances in surgical therapy for HCC, the overall prognosis of patients has not improved markedly during the past few decades due to the fact that most patients have locally-advanced or disseminated diseases at diagnosis. Currently, no satisfactory biomarkers are available to HCC screening. Early diagnosis of HCC is therefore critical for guiding the therapeutic management and improving prognosis.

Proteomics, which concerns comprehensive protein profile changes caused by multiple gene alterations, has emerged as a valuable tool in scientific medicine (2-6). Human serum contains thousands of proteolytically derived peptides called peptidomes, which may provide a robust correlation with the physiologic and pathologic processes in the entire body (7,8). The panels of peptidome markers might be more sensitive and specific than conventional biomarker approaches (9). Proteomic pattern by mass spectrometry is one of the most promising new approaches for early identification of disease from health volunteers. Advances in mass spectrometry (MS) now permit the display of hundreds of small- to medium-sized peptides using only microliters of serum (10,11). Preliminary studies have shown that great interest has been focused on the low-molecular-weight region, particularly on peptides smaller than 20 kDa, which may provide a novel means of diagnosing cancer and other diseases (8,12,13).

Matrix-assisted laser desorption/ionization time-of-flight MS (MALDI-TOF MS) can detect peptides with low molecular weights with the necessary sensitivity and resolution, which makes it a useful technique for serum peptide profiling. Furthermore, for accurate MS analyses, the peptidome fractionation procedure and
the preanalytical conditions of peptidome mapping must be carefully assessed (14). Magnetic beads (MBs), based on nanomaterials, have been developed and considered as a promising material for convenient and efficient enrichment of peptides and proteins in biological samples (15,16). The combination of MALDI-TOF MS enables the high throughput and sensitive investigation of peptides and proteins.

In the current study, a well-defined novel technology platform called ClinProt (Bruker Daltonics, Ettlingen, Germany) was used for early screening of HCC (17,19). A diagnostic model, which consists of one differentially expressed peptides, was established and validated by the QC Algorithm, by which different groups were effectively discriminated. Then, the diagnostic model was further verified using blinded samples from HCC and healthy volunteers. Thus, the preliminary work for the early and differential diagnoses of HCC was completed from an integrated perspective of peptide mass patterns.

MATERIALS AND METHODS

Reagents and instruments
The AutoFlex III MALDI-TOF mass spectrometer, MTP 384 target plate polished steel, α-cyano-hydroxycinnamic acid (CHCA), MB-WCX kit, and peptide calibration standard were purchased from Bruker Daltonics (Leipzig, Germany). The trifluoroacetic acid (TFA) and acetonitrile (ACN) were purchased from Alfa Aesar (Ward Hill, MA, USA), Sigma (St. Louis, MO, USA), and Roche Diagnostics GmbH (Sandhofer Strasse, Germany), respectively.

Patients and sample collection
With their consent, 64 healthy subjects (blood donor volunteers), 23 liver fibrosis patients and 63 HCC patients were enrolled into the study, and the blood samples were collected. Endoscopy was performed on all healthy subjects to exclude the presence of incidental digestive tract masses. The HCC patients underwent clinical staging, surgical lesion excision, and follow-up sessions.

Serum samples were prepared by collecting blood in a vacuum tube and allowing it to clot for 30 min at room temperature. About 1 mL of serum was obtained after centrifugation at 2000 rpm for 10 min and stored in small aliquots at -80 °C until analysis.

Study design
The data set, including 64 control and 63 HCC patients, was randomly split into two groups. The 23 liver fibrosis patients were recruited in the second group. The clinical characteristics of the HCC patients are shown in Table 1. The first group (model construction data set: 33 healthy volunteers, 23 liver fibrosis patients and 31 HCC patients) was used for the blind independent pattern validation of the cluster.

The gender ratios (male/female) of the healthy volunteers and HCC patients were 1.46 and 2.44, respectively. The mean ages (years) of the healthy volunteers and HCC patients were 54.63±1.37 and 58.48±10.60, respectively. The age and gender ratio differences between the healthy volunteers in the model construction group and those in the external evaluation data set were not significant. No significant differences were also observed for the ages and gender ratios of the HCC patients and healthy volunteers, as well as for the TNM stages of the HCC patients in the model construction and external evaluation groups.

Sample purification
WCX-MBs were used for the peptidome separation of samples following the manufacturer’s standard protocol (20). First, 10 μL of WCX-MB binding solution and 10 μL of WCX beads were combined in a 0.5 mL microfuge tube after thoroughly vortexing both reagents. Second, 5 μL of serum sample was added and mixed by pipetting up and down. Third, the microfuge tubes were then placed in an MB separator (MBS) and agitated 10 times. The beads were collected from the tube walls 1 min later. Fourth, the supernate was carefully removed using a pipette. Fifth, 100 μL of WCX-MB wash buffer was added into tubes, which were again agitated 10 times in the MBS. The beads were then collected from the tube walls, and the supernate was carefully removed using a pipette. After three washes, 5 μL of the WCX-MB elution buffer was added to disperse the beads in tubes by pipetting up and down. The beads were collected from the tube walls after 2 min, and the clear supernate was transferred into fresh tubes. Then, 5 μL of WCX-MB stabilization solutions were added to the collected supernate and mixed intensively using a pipette. The mixture was then ready for spotting onto MALDI-TOF MS targets and measurement. Finally, prior to the MALDI-TOF MS analysis, the targets were prepared by spotting 1 μL of the proteome fraction on the polished steel target (Bruker Daltonics, Bremen, Germany).

Table 1. Clinical characteristics of hepatocellular carcinoma (HCC) patients recruited in model construction group and external validation group

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>Model construction group (n=60)</th>
<th>Evaluation group (n=59)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: male/female</td>
<td>34/26</td>
<td>35/24</td>
<td>0.937</td>
</tr>
<tr>
<td>Age (years, X±SD)</td>
<td>58.24±8.21</td>
<td>57.46±9.39</td>
<td>0.912</td>
</tr>
<tr>
<td>TNM stage</td>
<td></td>
<td></td>
<td>0.381</td>
</tr>
<tr>
<td>I</td>
<td>18</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>42</td>
<td>43</td>
<td></td>
</tr>
</tbody>
</table>

§Pearson Chi-Square test.
#Independent-Sample T Test.
samples were taken as a training set in the class predictor algo-

rithm. Second, by designing a double-blind test, the samples
were provided as a class prediction for a cross-validation set to

pile the peaks across the spectra recorded from all samples. This
interest and with ClinProt™ software (Bruker Daltonics) to com-

Statistical methods and evaluation of assay precision
Each spectrum recorded using the MALDI-TOF MS was ana-
alyzed with Autoflex Analysis to detect the peak intensities of
interest and with ClinProt™ software (Bruker Daltonics) to com-
pile the peaks across the spectra recorded from all samples. This
setup allowed differentiation between the cancer and the con-
rol samples. To evaluate the precision of the assay, the within-
and between-run variations were determined using multiple
analyses of bead fractionation and MS for two plasma samples.
For the within- and between-run variations, three peaks with
various intensities were examined. The within-run imprecision
was determined by evaluating the coefficient variations (CVs)
for each sample, using eight assays within a run, and then the
between-run imprecision was determined by performing eight
different assays over a period of seven days. SPSS 16.0 was used
to analyze the clinical characteristics of the volunteers using a
χ² test or a t-test. The significance was set at \( p < 0.05 \).

RESULTS
For the reproducibility of the protein profiling, the within- and
between-run reproducibility of two samples was determined
via WCX-MB fractionation and MALDI-TOF MS analysis. In each
profile, three peaks with different molecular masses were se-
lected to evaluate assay precision. Despite varying peptide
masses and spectral intensities, the peak CVs were all <4% and
<10% in the within- and between-run assays, respectively.
These values were consistent with the reproducibility data for
the Protein Biology System reported by Bruker Daltonics.

In the pilot study, the differences between the serum pro-
teome profiles of HCC patients and healthy subjects were eval-
uated. The mass spectra from 1 kDa to 20 kDa were obtained
using MALDI-TOF MS in linear mode. The representative mass
spectra of the prefractionated sera of the model construction
group are reported in Figure 1. On average, about 169 signals
common to the two groups have been detected in this mass
range and about 16 were identified by the ClinProt software
with a statistically different area (\( p < 0.05 \) using the t-test) in the
model construction population, including 1 upregulated and
15 downregulated peptides, respectively (Table 2).

Classification models were developed to classify between the
HCC and healthy samples. The use of individual peaks as di-
agnostic biomarkers for HCC was addressed using QC algo-
rithm analysis. First, the HCC patients and healthy volunteers
were compared. Second, all detected peaks were analyzed us-
ing ClinProt 2.2 to generate the cross-validated classification
models. The optimized model resulted in the following correct
sample classification. One peptide ion signatures (m/z 7771)
were provided as a class prediction for a cross-validation set to

Figure 1. View of the aligned mass spectra of the serum protein profiles of
the model construction group (red represents healthy subjects, and blue
denotes HCC patients) obtained using MALDI-TOF after purification with
WCX-MBs.
Ke et al. Serum peptidome patterns of hepatocellular carcinoma

Table 2. Statistics of the 16 dysregulated proteins/peptides in hepatocellular carcinoma (HCC) patients compared with healthy individuals

<table>
<thead>
<tr>
<th>Mass (Average±SD)</th>
<th>Healthy volunteers ▲ (Average±SD)</th>
<th>Regulation in HCC</th>
<th>p value ▲</th>
</tr>
</thead>
<tbody>
<tr>
<td>7643.01 0.25±0.12</td>
<td>0.92±0.58</td>
<td>↓</td>
<td>0.000202</td>
</tr>
<tr>
<td>7927.69 0.34±0.27</td>
<td>1.09±0.61</td>
<td>↓</td>
<td>0.000202</td>
</tr>
<tr>
<td>7833.44 3.72±3.73</td>
<td>16.79±11.58</td>
<td>↓</td>
<td>0.000289</td>
</tr>
<tr>
<td>4397.32 2.96±1.94</td>
<td>9.68±6.18</td>
<td>↓</td>
<td>0.0004</td>
</tr>
<tr>
<td>7771.17 ▲ 29.82±13.68</td>
<td>111.35±17.93</td>
<td>↓</td>
<td>0.000494</td>
</tr>
<tr>
<td>5027.39 2.58±1.35</td>
<td>7.50±4.86</td>
<td>↓</td>
<td>0.000636</td>
</tr>
<tr>
<td>7896.09 0.85±0.81</td>
<td>3.31±2.48</td>
<td>↓</td>
<td>0.000773</td>
</tr>
<tr>
<td>8147.79 0.62±0.54</td>
<td>1.86±1.19</td>
<td>↓</td>
<td>0.000773</td>
</tr>
<tr>
<td>9359.98 3.28±4.60</td>
<td>11.72±8.40</td>
<td>↓</td>
<td>0.00146</td>
</tr>
<tr>
<td>4965.18 7.25±6.49</td>
<td>29.35±24.79</td>
<td>↓</td>
<td>0.00165</td>
</tr>
<tr>
<td>6636.05 0.34±0.18</td>
<td>1.12±0.45</td>
<td>↓</td>
<td>0.00165</td>
</tr>
<tr>
<td>1082.55 16.22±3.70</td>
<td>10.13±6.01</td>
<td>↑</td>
<td>0.00165</td>
</tr>
<tr>
<td>4272.16 50.13±14.56</td>
<td>101.95±38.12</td>
<td>↓</td>
<td>0.00165</td>
</tr>
<tr>
<td>9524.77 0.05±0.04</td>
<td>0.12±0.08</td>
<td>↓</td>
<td>0.00177</td>
</tr>
<tr>
<td>2510.81 8.59±5.01</td>
<td>27.28±12.89</td>
<td>↓</td>
<td>0.00326</td>
</tr>
<tr>
<td>2354.46 6.69±3.53</td>
<td>18.35±4.61</td>
<td>↓</td>
<td>0.00416</td>
</tr>
</tbody>
</table>

*The peptide selected for model construction.
▲ Peak area.
• p value calculated with the t-test; value lower than 0.05 suggest statistical relevance.

discriminate the HCC patients from healthy volunteers, which achieved 100% recognition and 90.42% cross-validation accuracy. The areas of the peak (m/z 7771) in the spectra of HCC patients were statistically different from those of the healthy volunteers (Figure 2).

The preliminary statistical analysis was performed for each single marker and signal cluster using the receiver operating characteristic curve analysis. The area under curve (AUCs) of peak at m/z 7771 was 0.81, which shows a highly accurate test, according to the criteria suggested by Swets [21] (Figure 3).

To verify the accuracy of the established QC classification model with the adopted peptides, another group of samples was introduced (not used in the model construction), which consisted of 31 HCC and 33 healthy subjects. As a result, the model correctly classified 96.77% (30/31) of the HCC (sensitivity), 82.61% (19/23) of liver fibrosis (specificity) samples and 100% (33/33) of the healthy (specificity) samples.

DISCUSSION

The usefulness of multiple markers for diagnosis, prognosis, and prediction of the risk of developing diseases or their complications is now widely recognized (7,22). Various proteomic approaches have been applied to biomarker discovery using biological fluids. Interestingly, low-molecular-weight peptides, such as S100A8 and fibrinogen, have been recognized to play important roles in physiologic and pathologic processes and could be used as relevant biomarker candidates (23,24). Recently, the mass spectrum that directly detects and differentiates short peptides has offered a promising approach for peptidomic biomarker discovery (8,12,25-27).

Compared with genomic approaches, proteomic analysis has the advantage of visualizing co- and posttranslational protein modifications, which are possibly relevant to biologic functions. Alternative approaches for measuring polypeptides, such as the surface-enhanced laser desorption and ionization, recently reported by several groups, have several disadvantages, such as low resolution and the loss of most proteins and peptides (28-30). MALDI is a soft ionization technique used in MS that allows the analysis of biomolecules such as proteins, peptide sugars, and large organic molecules. As a powerful tool for surveying the complex patterns of biologically informative molecules, MALDI-TOF MS protein/peptide profiling has been applied in proteomics biomarker research and has become a promising tool in cancer biomarker research (25,31,32).
In the present study, by integrating short peptide purification with WCX-MBs, peak intensity detection with MALDI-TOF MS, and profile analysis with ClinProt Tools software 2.2, a series of differentially expressed short peptides in the sera of HCC patients has been successfully detected. A comparative case control analysis between HCC and healthy subjects was performed. Peptidomic maps associated with the disease were drawn. The results show that compared with the normal controls, the HCC patients share 16 significantly differentiated peptides, including 1 upregulated and 15 downregulated peptides. The current knowledge on cellular regulation indicates that many networks operate at the epigenetic, transcriptional, and translational levels. Genomic and proteomic technologies will further help us understand the intracellular signaling and gene transcription systems, as well as the protein pathways that connect the extracellular microenvironment to the serum or plasma macroenvironment of cancer (33). These 16 interesting significantly differentiated peptides may provide further evidence for understanding the occurrence and progress of HCC. In particular, the prominent peptides that have a greater than twofold change in intensity, such as m/z 7833, 4397 and 9359 may be defined as the leading differential peptides associated with HCC, which are worthy of further sequence determination and functional analysis.

Using QC algorithm analysis, classification models were developed to classify samples between normal controls and HCC. The peptide at m/z 7771 achieved a recognition capacity and a cross-validation of close to 100% (83.87% specificity and 93.75% sensitivity) to discriminate HCC from normal volunteers. The blinded verification of the QC classification model proved the correct classification of 96.77% (30/31) of the HCC patients, 82.61% (19/23) of liver fibrosis patients and 100% (33/33) of the healthy volunteers. This result demonstrates that the QC Algorithm is effective in facilitating the construction of a sensitive and specific diagnostic and different diagnostic model. This study is the focused on screening for HCC-related short peptides in sera by combining WCX-MBs and MALDI-TOF MS. The classification model that we have set up has applications in providing alternatives for HCC diagnosis or differential diagnosis and may provide a better understanding of HCC pathogenesis or help in tailoring the use of chemotherapy for each patient, finally resulting in improved patient outcomes.

In conclusion, peptidome patterns from WCX-MB-purified serum samples were directly profiled with MALDI-TOF MS and a peptidome model that differentiated HCC from the control samples was constructed with high sensitivity and specificity. Despite the high sensitivity and specificity, the number of specimens analyzed in this study was relatively small, which may limit the validity of the results. The next step in our study will be to analyze larger patient cohorts and to run blinded samples to confirm the usefulness of the currently identified peptides for HCC diagnosis. After this confirmation, the biomarkers of the interest will then be isolated and identified and their biological role in HCC pathogenesis will be studied.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** This research was supported by The Medical Science Research Project of The Chinese People's Liberation Army (No. 082006).

**REFERENCES**