Oral Cancer Plasma Tumor Marker Identified with Bead-Based Affinity Fractionated Proteomic Technology

Ann-Joy Cheng,1* Li-Chiu Chen,1 Kun-Yi Chien,2 Yin-Ju Chen,1 Joseph Tung-Chieh Chang,3 Hung-Ming Wang,4 Chun-Ta Liao,5 and I-How Chen5

Background: There is no plasma marker for detecting oral cancer, one of the most frequent cancers in the world. We developed a bead-based affinity-fractionated proteomic method to discover a novel plasma marker for oral cancer.

Methods: Affinity purification of heparinized plasma with magnetic beads and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis were used to screen potential oral cancer markers. We compiled MS protein profiles for 57 patients with oral cancer and compared them with profiles from 29 healthy controls. The spectra were analyzed statistically using flexAnalysisTM and ClinProtTM bioinformatic software. In each MS analysis, the peak intensities of interest were normalized with an internal standard (adrenocorticotropic hormone 18–39). For identification, affinity bead–purified plasma protein was subjected to MALDI TOF/TOF analysis followed by Mascot identification of the peptide sequences and a search of the National Center for Biotechnology Information protein database.

Results: To optimize MALDI-TOF analysis based on the best discriminator of the cancer and control spectra, copper-chelated beads were used for plasma protein profiling. The within- and between-run CVs for assays were <4% and 7%, respectively. Six markers that differentiated between cancer and control spectra were found, with mean (SD) molecular masses of 2664 (1), 2850 (1), 3250 (1), 7735 (2), 7927 (2), and 9240 (2) Da. The 2664-Da marker, identified as a fragment of the fibrinogen α-chain, had the highest sensitivity (100%) and specificity (97%) for cancer.

Conclusion: The high specificity and sensitivity of the fibrinogen α-chain fragment suggest that it may be a clinical useful tumor marker.

Cancer may be accompanied by the production and release into the blood of a substantial number of proteins and/or hormones that could serve as useful markers for assessing prognosis, monitoring treatment, and detecting malignant disease at an early stage. Identification of clinically useful markers is a key challenge; currently used markers include carcinoembryonic antigen for colorectal cancer, α-fetoprotein for liver cancer, and prostate-specific antigen for prostate cancer (1). Methods for parallel analysis of the expression of a large number of genes at the protein level are evolving rapidly, providing opportunities for the identification of previously unrecognized disease markers (2). Matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-TOF MS)6 sensitively and precisely separates target proteins according to their mass-dependent velocities (m/z). This approach makes possible the discovery of new disease-associated biomarkers by comprehensively examining the protein expression profiles in individuals with and without a disease (3), e.g., hepatocellular carcinoma (4). The clinical and diagnostic use of this technology has been
in early-stage cancer (stage I or II), disease relapses, although standard care is frequently initially successful. The recurrence rate in advanced-stage disease is ~50%–60%, and distant metastases occur in 20%–35% of cases (24). Even with good treatment response, patients with advanced disease often suffer substantial functional and cosmetic morbidity, which decreases the quality of life. Although improved therapy may lead to better cancer control and long-term survival, a circulating tumor marker useful for diagnosis and monitoring of disease progression may offer a chance to decrease the morbidity of this disease. We used affinity bead purification and MALDI-TOF MS to assess protein expression profiles to identify potential plasma tumor markers in ORC.

Materials and Methods

Patients and Blood Sample Preparation

A total of 57 consecutive patients (53 men and 4 women) with ORC seen at the Otorhinolaryngology or Head and Neck Surgery Clinics at Chang Gung Memorial Hospital (Taoyuan, Taiwan) were recruited for this study. Their mean age was 40 years (range, 30–74 years). Eighteen patients had early stage (stage I or II) and 39 advanced stage (stage III or IV) ORC. Written informed consent was obtained from all participating patients. All participants underwent a series of clinical evaluations, including an assessment of cancer extent before treatment and response to therapy. Pathology data and cancer staging were carefully recorded. The clinical staging was based on the American Joint Committee on Cancer TNM classification of malignant cancers.

We collected ~3 mL of heparinized plasma from each patient before treatment and with the same collection procedure obtained control blood samples from 29 healthy individuals (27 males and 2 females, mean age, 39 years; range, 23–71 years). To obtain fresh plasma, we centrifuged the blood specimens at 3000 × g for 5 min within 4 h of collection. All samples were stored at −70 °C until use and were assayed within 1 year of frozen storage.

Removal of Albumin

Plasma albumin was removed with a convenient Montage Albumin Depletion Kit according to the manufacturer’s instructions (Millipore Corporation). Briefly, 100 µL of 10-fold–diluted plasma was applied to the affinity column insert. After brief incubation, the sample was centrifuged for 2 min at 3000g. The unbound plasma proteins were recovered from the collection tube. To evaluate the success of albumin removal, 10 µL of the plasma samples with and without albumin depletion were subjected to 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis analysis. We then stained the gel with Coomassie blue, to make the protein pattern visible, and photographed it.
PROTEOMIC ANALYSIS TO PROFILE PLASMA PROTEINS
All samples were assayed in duplicate. Plasma samples were thawed and purified with a reagent set with chemically coated magnetic beads, including a carbon series (C5, C8, C18), copper, and cation beads (ClinProt™; Bruker Daltonics Company). For each sample, 5 µL of plasma was mixed with 5 µL of beads. Samples were purified through 3 steps, binding, washing, and elution, according to the manufacturer’s suggested protocol, in which the binding incubation time took 1 min. A total of 5 µL of each sample was eluted, and the purified plasma samples were further diluted 4-, 8-, and 16-fold. Carbon-series bead-purified samples were diluted with 500 mL/L acetonitrile, whereas copper- or cation-purified samples were diluted with copper- or cation-specific elution solutions. We prepared the various diluted solutions for MALDI-TOF MS by mixing 1 µL of the diluted sample with 0.5 µL of matrix solution containing 2 g/L α-cyano-4-hydroxycinnamic acid, 10 nmol/L angiotensin II, 10 nmol/L adrenocorticotropic hormone (ACTH) 18–39 (Bruker Daltonics), and 10 mL/L formic acid in 500 mL/L acetonitrile and allowing the droplet to dry on the MALDI sample plate. The profiling spectra were obtained by use of an Ultraflex MALDI-TOF mass spectrometer (Bruker Daltonics) operated in positive-ion linear or reflectron mode. Briefly, the profiling data were acquired in linear-mode geometry, and the mass maps were acquired in reflectron mode. All spectra were obtained randomly over the surface of the matrix spot. Internal calibration during reflectron mode data acquisition was performed with 2 calibrators, angiotensin II (1046.73 Da) and the ACTH 18–39 polypeptide (2465.20 Da), to achieve a mass accuracy of 30 ppm. The profiling spectra were calibrated externally with a mixture of protein/peptide calibrators (Bruker Daltonics). A ±5 Da mass accuracy for each spectrum was observed and was probably attributable to the geometric factor of varied sample position on the sample plate. Such mass shifts can be further corrected by the flexAnalysis™ software after alignment of 2 internal standards. Briefly, all spectra were processed by automatic baseline subtraction, peak detection, recalibration, and peak-area calculation according to the predefined settings. The criteria for peak detection were as follows: signal-to-noise ratio >5, a 2-Da peak width filter, and a maximum peak number of 200. The intensities of the peaks of interest were normalized with the peak intensity of the ACTH internal standard.

STATISTICAL METHODS, EVALUATION OF ASSAY PRECISION, AND DIAGNOSTIC EFFICACY
We analyzed each spectrum obtained from MALDI-TOF MS with flexAnalysis and ClinProt™ software (Bruker Daltonics Company), the former to detect the peak intensities of interest and the latter to compile the peaks across the spectra obtained from all samples. This allowed for differentiation between the cancer and control samples. To evaluate the precision of the assay, we determined within-run and between-run variations by use of multiple analyses of bead fractionation and MS for 2 plasma samples. For within- and between-run variation, we examined 3 peaks with various intensities. We determined within-run imprecision by evaluating the CVs for each sample, using 12 assays within a run; we determined between-run imprecision by performing 8 different assays over a period of 10 days. To assess diagnostic efficacy, we calculated the means (SD) of the peaks of interest in the control group. The cutoff value was set as the value of the mean plus 2 SD in the control samples. The sensitivity (ratio of the cancer samples with a mass intensity greater than the cutoff value to all samples in the cancer group) and specificity (ratio of control samples with a mass intensity less than the cutoff value to all samples in the control group) were analyzed accordingly.

BIOINFORMATICS AND IDENTIFICATION OF PROTEIN MARKERS
Selected peptides were further purified by use of C8 beads and serially eluted and concentrated with 150, 250, 350, and 500 mL/L acetonitrile. These peptides were identified directly by MALDI TOF/TOF analysis to obtain the peptide sequences. Peptide mass fingerprinting was performed with the Mascot search engine (Matrix Science) and a search of the National Center for Biotechnology Information (NCBI) protein-protein BLAST database (http://www.ncbi.nlm.nih.gov/BLAST/).

Results
SAMPLE PREPARATION FOR MS
Albumin is the most abundant protein in blood. To determine whether albumin interfered with protein profiling, we first compared the plasma protein profiles before and after removal of albumin, using 2 control and 2 cancer plasma samples. The results showed that the removal process significantly depleted albumin in both control and cancer plasma samples (see Fig. 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol51/issue12). Samples were then subjected to proteomic analysis with C8 beads, MALDI-TOF MS, and ClinProt software analysis, which compiled protein spectra from the 4 plasma samples (see Fig. 2 in the online Data Supplement). In general, the mass signals in the albumin-depleted samples were weaker than those from whole plasma. In addition, the differences between the mass spectra of the control and cancer samples in the albumin-depleted group were smaller than the differences from the whole-plasma samples, possibly because of removal of albumin-binding or other nonspecific proteins along with albumin. Because these albumin-binding proteins may include a tumor marker, we believe that removal of albumin is unnecessary.
BEAD SELECTION AND ASSAY OPTIMIZATION
To optimize the MALDI-TOF MS analysis, we used different amounts of plasma with equal amounts of MS matrix solution. Each type of chemical affinity bead was tested, with the results from the C8 and copper beads presented as examples. After purification by C8 or copper beads, the fractionated plasma was diluted with specific elution solutions: 500 mL/L acetonitrile for C8 beads or copper elution solution for copper beads. After 4-, 8-, and 16-fold dilutions with MS matrix solution, the samples were analyzed directly by MALDI-TOF MS with a total of 100 laser shots. In general, the peak complexity in the plasma samples was greatly reduced, which facilitated further identification by direct MS analysis. For the MS spectra analyzed by flexAnalysis, we found that a 16-fold dilution for C8 and 8-fold for copper beads yielded the best combination of low background noise and high peak intensity (see Fig. 3 in the online Data Supplement).

To determine which types of beads were suitable for plasma protein profiling, we first used plasmas from 6 control individuals and 6 cancer patients for screening. In general, the intensities of background noise for each bead type were all <20, so that a specific peak was easily distinguishable when the intensity was >40. The peak pattern and intensity varied for each sample and were dependent on the different sample and bead types used. Approximately 15–25 peaks were produced after fractionation by C8, copper, or cation beads. The carbon-series beads yielded similar profiles with certain minor differences (Fig. 1, A to C). C3 beads yielded profiles with the most peaks but the highest background noise, whereas C18 beads gave the fewest peaks. C8 beads afforded the best differentiation between control and cancer samples and had low background noise. Results with the copper and cation beads are shown in panels D and E, respectively, of Fig. 1. Both yielded good differentiation. Iron beads did not produce adequate protein profiles, possibly because of the high specificity of iron beads for plasma proteins. On the basis of our results, we chose copper beads for further study of our samples.

CHARACTERIZATION AND EVALUATION OF PROTEIN PROFILING USING COPPER BEADS
To determine the range of spectra obtained with copper bead fractionation in MALDI-TOF MS profiling, we examined the profiles of proteins smaller than 20 kDa obtained from 4 plasma samples. The mass spectra of the overlapping intensities of these 4 plasma samples were analyzed by flexAnalysis, and the protein profiles across all spectra were compiled with ClinProt software (see Fig. 4 in the online Data Supplement). Most of the protein peaks from the 4 samples were <10 kDa, suggesting that searching in this range is sufficient for analysis.

We next examined the precision of this technology. Within- and between-run reproducibility of 2 samples were determined with copper-bead fractionation and MALDI-TOF MS analysis. Example data of 3 replicates (within-run) or experiments (between-run) are shown in Fig. 5 in the online Data Supplement. In each profile, 3 peaks with different molecular masses were selected to analyze the precision of the assay. CVs from 12 within-run and 8 between-run assays are summarized in Table 1. Despite varying peptide masses and spectrum intensities, the peak CVs were all <5% in the within-run and <9% in...
the between-run assays. The mean CVs of the runs represent the overall imprecision, <4% in the within-run and 7% in the between-run assays, indicating that results produced by this method are highly reproducible for both plasma samples.

Search for an ORC Marker
To determine proteomic patterns and search for a specific ORC tumor marker, we used copper-bead fractionation and MALDI-TOF MS to profile the mass spectral patterns of plasma samples from 29 healthy controls and from 57 oral cancer patients. We used ClinProt software to bin the peaks across the spectra obtained from the samples and found 6 peaks that differed significantly between the 2 groups (see Fig. 6 in the online Data Supplement). The mean (SD) masses were 2664 (1), 2850 (1), 3250 (1), 7735 (2), 7927 (2), and 9240 (2) Da after alignment calibration and were used to designate the peaks as ORC markers A to F, respectively. We used flexAnalysis software to determine the actual intensities of these 6 peaks of interest in every specimen after normalizing them by the ACTH internal standard. As shown in Fig. 2, the peak intensities in the ORC samples were all stronger than in the control samples.

To assess the diagnostic efficacy, the mean (SD) masses of the 6 peaks in the control group were calculated. Results for the 6 markers are summarized in Table 2. The sensitivities of ORC-specific markers A to F were 100%, 82.5%, 75.4%, 43.9%, 87.7%, and 75.4%, respectively, and the specificities were 96.6%, 96.6%, 96.6%, 93.1%, 93.1%, and 96.6%.

Identification of the Cancer Markers
With this bead-based proteomic technology, we found several potential ORC tumor markers. The most sensitive (100%) and specific (96.6%) was marker A [2664 (1) Da], which we further identified. After fractionation by copper beads, the eluted ORC plasma samples were further purified by C8 beads and serially eluted with 150, 250, and 350 mL/L acetonitrile. Samples were then subjected to MALDI-TOF MS analysis. An example of the various C8 elution fractions analyzed by flexAnalysis software is shown in Fig. 3. Marker A was significantly enriched in the 250 mL/L acetonitrile fraction with minimal background noise. This 250 mL/L acetonitrile eluate was further subjected to MALDI-TOF MS MS analysis. The MS fingerprint was subjected to Mascot searching for peptide sequence and further to NCBI database for protein identification. The Mascot score for marker A was 170, with a mass accuracy of 1 ppm, indicating a high confidence in identification of the peptide sequences. The sequence was determined to be DEAGSEADHEGTHSTKRGHAKSRPV, which matches 25 of the 200 amino acids (residues 175–200) in the C-terminal fragment of fibrinogen α-chain in NCBI BLAST protein-protein database (locus number AAB26584), as originally identified by Rixon et al. (25).

Discussion
We directly profiled protein/peptide patterns from affinity bead–purified plasma samples with MALDI-TOF MS and determined several markers that differentiated ORC from control samples with high sensitivity (>90%) and specificity (>90%). Statistically, there is at least 95% confidence that the selected tumor marker is a significant discriminator. However, the number of specimens analyzed in this study was relatively small, which may limit the validity of the results. Further independent validation studies with a larger sample size are needed to determine the utility of this marker for diagnosis.

Two general categories of cellular proteins are released into the circulation, large proteins that are actively secreted, and low–molecular-weight (LMW) proteins that enter the blood passively from cellular degradation or cleavage (14,17). There is great interest in the LMW region as a source of diagnostic information, particularly substances smaller than 20 kDa. In our study, after alignment by 2 internal standards, the protein peaks were generally <10 kDa, and our results were in agreement.

Table 1. Reproducibility of mass spectra profiled by copper beads and MALDI-TOF analysis.a

<table>
<thead>
<tr>
<th></th>
<th>Within-run</th>
<th></th>
<th></th>
<th></th>
<th>Between-run</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>MI</td>
<td>ICV,a %</td>
<td>MCV,f %</td>
<td>n</td>
<td>MI</td>
<td>ICV,b %</td>
</tr>
<tr>
<td>Sample 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>261.4</td>
<td>3.7</td>
<td>3.9</td>
<td>8</td>
<td>259.0</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2646 (1)</td>
<td>37.5</td>
<td>5.0</td>
<td></td>
<td>39.5</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5883 (1)</td>
<td>87.1</td>
<td>3.0</td>
<td></td>
<td>90.2</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>170.7</td>
<td>4.3</td>
<td>3.7</td>
<td>8</td>
<td>163.9</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>2016 (1)</td>
<td>1508.9</td>
<td>2.3</td>
<td></td>
<td>1532.4</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3230 (1)</td>
<td>66.9</td>
<td>4.5</td>
<td></td>
<td>63.5</td>
<td>6.2</td>
<td></td>
</tr>
</tbody>
</table>

a Reproducibility was determined by calculating the mean CV of the normalized peak amplitudes for each of the 3 peptides (molecular masses) with the highest consistent amplitudes.

b Masses of the 3 peptide peaks with the highest consistent amplitudes.

c Numbers of tests.
d MI, mean intensity of each peak.
e ICV, individual CV of each peak.
f MCV, mean CV of the runs.
with findings in the LMW region of the blood proteome in other reports (14, 17). Degradation by endoprotease activity is a proposed mechanism to account for the abundance of LMW peptides found by direct MS analysis of plasma samples (26). This mechanism is an unlikely explanation of the results of our study, however, because all of the samples were processed similarly and within a 4-h time period during which no change in the assay results was demonstrated (data not shown).

Our finding that the mean absolute intensity of the internal standard was slightly lower for the cancer group than for the control group was interesting and unlikely to have resulted from experimental error. The lower intensity of the internal standard can be attributed to an the suppression effect, commonly observed in MS, attributable to the presence of other, more abundant components in the cancer group samples. Although other studies using total intensities for spectrum normalization have been published (27), only the relative difference in marker abundance between diseased and nondiseased populations was reported. We used total peak intensity for normalization, given similar results, and normalized spectra with an internal standard of known concentration to get more information on the absolute abundances of markers and its correlation with their biological significance. Nevertheless, it is advisable to include a calibration sample throughout experimental runs to act as an internal standard for the different stages of the process.

Fibrinogen is a plasma glycoprotein synthesized in the liver and is composed of 3 structurally different subunits: 2 α-chains, 2 β-chains, and 2 γ-chains. Usually, the amounts of α-chains will be in a constant ratio vs total fibrinogen or the other chains. Thrombin causes limited proteolysis of the fibrinogen molecule, during which α- and β-chain peptides are released to form fibrin monomers. Fibrinogen is tightly regulated and involved in

---

**Fig. 2. Distribution of ORC-specific markers.**

flexAnalysis software was used to find the actual intensity of each marker in every sample and to normalize the results against the internal standard. The distributions of ORC-specific markers A to F are shown in panels A to F, with the peak intensity along the x axis and the specific samples/groups along the y axis.
The extracellular matrix serves as a scaffold to support proliferation, and migration during carcinogenesis. The binding of growth factors and promotes cellular adhesion, and other adhesive glycoproteins, such as fibronectin, in cancer cells found that these cells produce and secrete fibrinogen. Although it is thought that fibrinogen in tumor tissue may not originate from ORC tissue itself but originates from plasma exudates with subsequent deposition in the tumor stroma, recent studies of MCF-7 breast cancer cells found that these cells produce and secrete fibrinogen polypeptides. This finding suggests that the endogenous synthesis of this molecule may be, at least in part, the source of fibrinogen in the extracellular matrix of breast cell carcinomas. Studies in animal models have shown that inhibition of fibrinogen strongly diminishes the development of metastatic lung cancer, further demonstrating the important role of fibrinogen in sustaining invasion and survival of tumor cells. Changes in plasma fibrinogen concentrations in association with cancers have also been reported. In a series of coagulation markers tested, only the fibrinogen α-chain was significantly increased in gastric cancer, and it was correlated with tumor T, N, and M stages. Similarly, the fibrinogen α-chain peptide has been found to be significantly increased in melanoma. In a clinical trial of chemotherapy, treatment with mopidamol was associated with a statistically significant prolongation of survival in patients with lung cancer, an effect correlated with the reduction in mean plasma fibrinogen concentration. In the present study, we found that the fibrinogen α-chain peptide fragment was significantly increased in the plasma of all 57 patients with ORC. Our results and those of other investigators indicate that the presence of fibrinogen within a tumor may affect tumor progression and metastasis. The plasma concentrations of fibrinogen α-chain and/or its proteolytic fragments may thus serve as a useful marker to monitor tumor status.

To be clinically useful, biomarkers must assist in disease detection. In our study, plasma samples were distributed over every stage of ORC. Much more work needs to be done to determine when in the course of disease the plasma fibrinogen α-chain concentration increases. Questions include whether it increases at the actual initiation of malignant transformation and whether it is present in benign conditions such as leukoplakia. More detailed prospective studies are required to determine whether increased concentrations of this peptide are associated with specific ORC TNM stages, prognosis, or treatment outcome. Furthermore, high concentrations of plasma fibrinogen may not originate from ORC tissue itself but...
may be produced by other organs in response either directly to the presence of cancer or to the general condition of the patient with cancer. If the increase in fibrinogen α-chain concentrations is an epiphenomenon of cancer, this molecule may serve as a marker for a variety of malignant diseases. It is therefore worth measuring the plasma concentrations of fibrinogen α-chain in other cancers as well.

In conclusion, we have shown that a convenient, fast proteomic technique, affinity-bead purification and MALDI-TOF analysis in combination with bioinformatic software, facilitates the identification of novel biomarkers. Our assay indicates that a fibrinogen α-chain peptide fragment is significantly increased in patients with ORC. The high specificity and sensitivity for this substance in our study indicate that it has great potential as a clinically useful tumor marker. We advocate use of this proteomic approach to look for novel markers for other diseases in plasma, urine, cerebrospinal fluid, or other specimens.

This work was supported by Chang Gung Research Grants CMRPG32017 (to A.J.C.) and CMRPG32038 (to J.T.C.) and by National Science Research Grants of Taiwan (NSC 92-2314-B-182-067 and 92-2314-B-010-005; to A.J.C.). We thank Mary Jeanne Buttrey, MD, for critical reading and correction of the manuscript.

References

28. Redman CM, Xia H. Fibrinogen biosynthesis. Assembly, intracel-


