RESEARCH ARTICLE

Utility of Serum Peptidome Patterns of Esophageal Squamous Cell Carcinoma Patients for Comprehensive Treatment

Qing-Lian Wan[1]*, Xiang-Sheng Hou[1], Guang Zhao[2]

Abstract

Esophageal cancer (EC) is one of the most common malignant tumors, and the incidence of esophageal squamous cell carcinoma (ESCC) is highest in China. Early diagnosis and effective monitoring are keys to comprehensive treatment and discovering tumor metastases and recurrence in time. The aim of this study was to confirm serum peptidome pattern utility for diagnosis of ESCC, and assessment of operation success, postoperative chemotherapy results, tumor metastasis and recurrence. Serum samples were collected from 61 patients treated with surgery and chemotherapy and 20 healthy individuals. Spectral data generated with weak cationic-exchanger magnetic beads (WCX-MB) and MALDI-TOF MS by a support vector machine (SVM), were used to construct diagnostic models and system training as potential biomarkers. A pattern consisting of 11 protein peaks, separated ESCC (m/z 650.75), operated (m/z 676.61, 786.1, 786.58), postoperative chemotherapy (m/z 622.77, 650.66, 676.46) and tumor metastasis and recurrence (m/z 622.63, 650.56, 690.77, 676.12) from the healthy individuals with a sensitivity of 100.0% and a specificity of 100.0%. These results suggested that MALDI-TOF MS combined with MB separation yields significantly higher sensitivity and specificity for the detection of serum protein in patients with EC patients treated with surgery and chemotherapy.

Keywords: Esophageal cancer (EC) - serum peptidome patterns - treatment - recurrence

Introduction

Esophageal cancer (EC) is one of the most common malignant tumors, the 6th cause of death in the worldwide (Duan et al., 2009). The incidence of esophageal squamous carcinoma was highest in China, but in the western countries had high incidence of esophageal adenocarcinoma. At present, the conventional esophageal cancer diagnosis is through the endoscope and X-ray inspection, diagnosis depends on pathological. But these invasive methods limits the early census and after treatment monitoring, and the satisfactory biomarkers are available to screen for esophageal cancer patients treated with surgery and chemotherapy, and recurrence, which is rare reports. So it is essential to hit effective biomarker for diagnosis of EC, Operated, Postoperative chemotherapy and tumor metastasis and recurrence.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI- TOF MS) can detect peptides with low molecular weights at necessary sensitivity and resolution, which make it a useful technique for serum peptide profiling, Magnetic bead (MB), based on nanomaterials, has been developed and was considered as a promising material for convenient and efficient enrichment of peptides and proteins in biological samples (Yao et al., 2008; Sun et al., 2011; Shao et al., 2012). MALDI-TOF MS and MS can outcome peptides high throughput and sensitive investigation ClinProt (Bruker Daltonics, Ettlingen, Germany) which comprise a weak cationic-exchanger magnetic beads- (WCX-MB) based sample separation, MALDI-TOF MS which collected peptide profiling, and created "disease-specific" peptidome pattern models, which can diagnose cancer for a powerful tool (Maurer et al., 2011; Fan et al., 2012). Many studies have shown that low molecular weight region, particularly peptides smaller than 20 kD, which may provide a novel means of diagnosing cancer and other diseases (Dai et al., 2010; Liu et al., 2010; Sui et al., 2010). In this study, we used the software ClinProt 2.2 and patterns recognition SVM Algorithm construct serum peptidome patterns for EC, operated, postoperative chemotherapy and tumor metastasis and recurrence diagnostic model, different groups were discriminated EC, surgery, chemotherapy, and recurrence, from health volunteer samples effectively.

Materials and Methods

Patients and sample collection
Twenty health volunteers, twenty-one EC patients,

[1]Institute of Esophageal Cancer of Zhengzhou City, Department of Thoracic Surgery, The People's Hospital of Zhengzhou, Zhengzhou, [2]Centeral Laboratory, The 150th Hospital of Chinese People's Liberation Army, Luoyang, Henan Province, China *For correspondence: qinglianwan@126.com

sixteen operated patients (1 week after surgery), fifteen postoperative chemotherapy (With docetaxel, cisplatin, fluorouracil treatment after a period), and seven tumor metastasis and recurrence were recorded with the permission of the Local Ethical Commission, and blood sample was collected after the patients permission. Esophageal cancer was diagnosed by gastroscopy and pathologic diagnosis. Serum samples were collected in Department of Thoracic Surgery, The People's Hospital of Zhengzhou. This study was conducted in accordance with the declaration of Helsinki. This study was conducted with approval from the Ethics Committee of the People's Hospital of Zhengzhou. Written informed consent was obtained from all participants.

Serum samples were collected in a vacuum tube and cloted for 30 min at room temperature. After whizzing at 3000 g for 10 min and about 1 ml of serum stored in small aliquots at −80°C until analysis.

Study design

The data set including 20 health controls and Twenty-one EC patients, 16 operated patients, Fifteen Postoperative chemotherapy, and seven tumor metastasis and recurrence was randomly divided into two groups. The first group (80% health volunteers and model construction group) was used for the identification of signals related to peptides expressed differentially in EC Comprehensive treatment and recurrence patients. The second group (20% health volunteers, and random model construction data patients) was used for the independent patterns validation of the cluster blindly.

The gender ratio (male/female) of health volunteers, EC patients, operated patients, postoperative chemotherapy and tumor metastasis and recurrence patients was 1.62, 1.25, 2.0, and 2.5, respectively. The mean age (years) of health volunteers, EC patients, Operated patients, Postoperative chemotherapy and Tumor metastasis and recurrence patients was 61.23±3.87, 61.51±5.06, 62.16±5.94, 61.60±4.9, 59.42±4.073, respectively. The difference of age and gender of health volunteers in model construction group were not significant. No significant differences were observed for age and gender between EC patients, Operated patients, Postoperative chemotherapy, Tumor metastasis or recurrence patients and health volunteers.

We used MB-WCX for peptidome separation of samples following the standard protocol by the manufacturer (An et al., 2004). Serum ice bath next thaw, take 5 µl join has been installed with buffer (BS) and WCX magnetic beads sample tube, the setting of magnetic separation on 5 min; Carefully blotting suspension liquid, with magnetic bead cleaning buffer (WS) cleaning three times, each time cleaning in magnetic bead separator and two adjacent between hole and move the sample tube ten times (note magnetic beads in tube of moving), liquid should be clear, completely blotting suspension liquid and avoid sucked magnetic beads; Add 5 µl magnetic bead elution buffer (ES), elution adherent magnetic beads; In the magnetic bead separator, magnetic bead attached 2 min; The supernatant fluid shift into dry clean 0.5 ml sample tube, add 5 µl magnetic bead stable buffer (SS),

careful suction play blending can be used to direct mass spectrometry. With known relative molecular quality standard product to MADLI mass spectrometry system relative molecular mass error correction. All the samples random in the 384 general target, each and every sample add a standard serum experiment process quality control. The first point 1 µl samples, put dry at room temperature, then point 1 µl base mass (3 mg/ml CHCA ACN, 50%, 2% TFA), room temperature dry and put computer detection.

Mass spectrometry analysis

Mass spectrum data acquisition methods ClinProt linear cation model, parameter Settings are as follows: the first ion source and 20.00kv, the second ion source 18197 kv, detection range 0-20000 da. Each sample point four targets, different crystallization point, more collection, each point 200 shots, choose 10 different position accumulated 20000 shots. To obtain better mass spectrogram, first with high energy laser bombardment, then with low energy laser acquisition map.

Data processing and analysis using The ClinProt Tools software 2.2 (Bruker Daltonik, Bremen, Germany), the collection and partial homogenization processing, denoising smoothing, filter signal-to-noise ratio is less than 2 peak, for each charge to mass ratio peak do wilconxon rank and inspection, there are differences about the charge to mass ratio peak. Through SVM Algorithm operation to establish esophageal cancer diagnosis model, in order to keep a method for cross validation. Twenty percent of model construction group were randomly selected sample as a test set, and the rest samples were taken as a training set in the class predictor algorithm.

Evaluation of assay precision

Experimental results of quality control as external standard calibration standard product contains 11 polypeptide, every eight sample data collection before do external standard calibration, calibration average molecular weight deviation is less than 100 PPM. As internal standard calibration standard serum mass spectrum diagram and the database for the corresponding WCX magnetic beads the same sampling method standard serum data comparison, calculating coefficient of variation (CV) is 12%. Each spectrum obtained from MALDI-TOF MS were analyzed by Autoflex and ClinProt TM software (Bruker Daltonics, Bremen, Germany), Autoflex
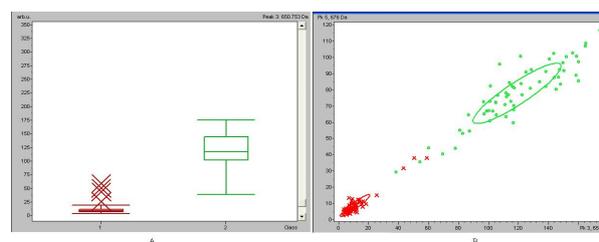
Figure 1. A. Box-and-whiskers plot calculated from the areas of the two signals used in the cluster for the two studied populations. Red represents EC, green represents healthy volunteers. B. Two-dimensional peak distribution view of the two peaks selected for the diagnostic model. The peak area and the m/z values are indicated on the x- and y-axes. The ellipses represent the standard deviation of the class average of the peak areas/intensities

Table 1. Statistic of the Candidate Biomarker Signals Selected for the Diagnostic Model for Identify from Health Individuals

Mass	PTTA	PAD	Ave (operated)	Ave (H) (operated)	SD	SD (H)
786.1	0.00001	0.000001	665.31	133.28	68.38	55.26
786.58	0.00001	0.00003	526.61	111.38	75.27	51.72
676.61	0.000001	0.000001	27.2	385.72	8.2	40.1

Table 2. Statistic of the Candidate Biomarker Signals Selected for the Diagnostic Model for Identify from Health Individuals

Mass	PTTA	PAD	Ave (chemotherapy)	Ave (H) (chemo)	SD	SD (H)
622.77	<0.00001	<0.000002	6.09	105.69	4.4	21.64
650.66	<0.00001	<0.000001	37.4	588.4	22.29	60.01
676.46	<0.000001	<0.000001	27.8	385.87	10.7	40.05

detect the peak intensities of interest and ClinProt TM software compile the peaks across the spectra obtained from all samples. This distinguished the cancer, operated, Postoperative chemotherapy and tumor metastasis and recurrence from control samples.

Statistical analysis

SPSS 16.0 (SPSS Inc., Chicago, IL, USA) was used for analysis of the clinical characteristics of volunteers using t-test. A difference at $P < 0.05$ was considered statistically significant. Also, SPSS 16.0 was used to compare the accuracy of the peptidome models.

Results

We evaluated the differences of the serum proteome profiles of EC in comparison to health subjects. The mass spectra from 0 to 20 kDa were obtained by MALDI-TOF MS in linear mode. The representative mass spectra of prefractionated serum of model construction group are reported .On average about 123 signals common to the two groups have been detected in this mass range and about 57 were identified by the ClinProt software with a statistically different area ($P < 0.01$ by t-test) in model construction population, including 8 upregulated and 44 downregulated peptides.

Classification models were developed to classify samples between EC and health volunteers. The use of individual peaks as diagnostic biomarker for EC was addressed using SVM algorithm analysis. First, we conducted comparison between EC and health volunteers. Second, all detected peaks were analyzed by ClinProt 2.2 to generate cross-validated classification models. The optimized model resulted in the following correct classification of samples. One peptide ion signatures (m/z 650) was provided as a class prediction for a cross-validation set to discriminate EC from health volunteers, which achieved a recognition capacity of 100% and a cross-validation of 100%. Preliminary statistical analysis was carried out for each single marker and for the cluster of signals by the receiver operating characteristic curve analysis. Area under curve (AUC) of peak 3 at m/z 650 ($P<0.000001$) was 0.9989, which corresponds to a highly
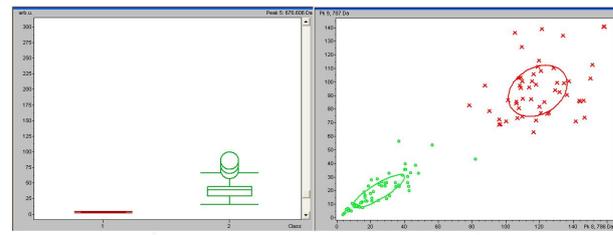
Figure 2. A. Box-and-whiskers plot calculated from the areas of the two signals used in the cluster for the two studied populations. Red represents operated, green represents healthy volunteers. B. Two-dimensional peak distribution view of the two peaks selected for the diagnostic model. The peak area and the m/z values are indicated on the x- and y-axes. The ellipses represent the standard deviation of the class average of the peak areas/intensities
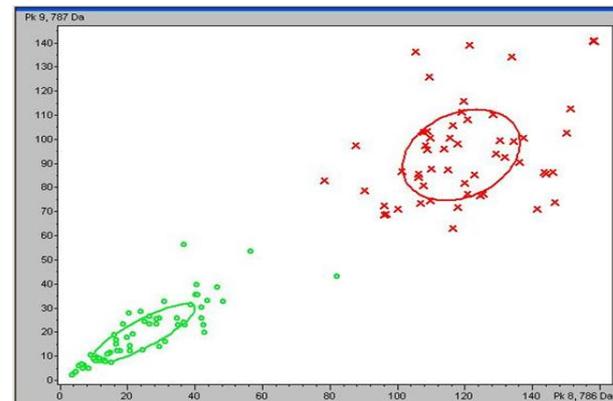
Figure 3. Two-dimensional Peak Distribution View of the Two Peaks Selected for the Diagnostic Model. The peak area and the m/z values are indicated on the x- and y-axes. The ellipses represent the standard deviation of the class average of the peak areas/intensities

accurate test, according to the criteria suggested by Swets. Moreover areas of these peaks in the spectra of EC were statistically different from those of the health volunteers (Figure 1A). Combination of the two peaks allowed to yielding a specificity of 100%, and a sensitivity of 100% for EC (Figure 1B).

113 signals common to the Operated and health groups have been detected and about 49 were identified by the ClinProt software with a statistically different area ($P < 0.05$ by t analysis) in model construction population, including 8 upregulated and 41 downregulated peptides. Three peptides selected for model construction were shown in Table 1. Area under curve (AUC) of peak 5 at m/z 676 ($P < 0.000001$) was 1.0, which corresponds to a highly accurate test, according to the criteria suggested by Swets. Moreover areas of these peaks in the spectra of operated were statistically different from those of the health volunteers (Figure 2A). Combination of the two peaks allowed to yielding a specificity of 100%, and a sensitivity of 100% for operated (Figure 2B).

The postoperative chemotherapy mass spectra from 0 to 20 kDa were obtained by MALDI-TOF MS in linear mode. On average about 105 signals common to the two groups have been detected in this mass range and about 58 were identified by the ClinProt software with a statistically different area ($P < 0.05$ by t analysis) in model construction population, including 8 upregulated and 50 downregulated

Table 3. Statistic of the Candidate Biomarker Signals Selected for the Diagnostic Model for Identify from Health Individuals

Mass	PTTA[1]	PAD[2]	Ave[3] (recurrence)	Ave (H)[4] (recurrence)	SD[5]	SD (H)[6]
622.63	<0.00001	<0.0004	4.43	104.73	1.75	20.86
650.56	<0.00001	<0.000001	28.36	588.4	5.12	60.01
690.77	<0.00001	0.00003	12.46	152.16	2.3	23.46
676.12	<0.000001	<0.000001	7.53	186.5	2.35	18.24

[1]P value calculated with the t-test; values lower than 0.05 suggest statiscal; [2]P value calculated with the t- test, values lower than 0.05 suggest statiscal; [3]Average area of peaks for EC subjects; [4]Average area of peaks for health subjects; [5]Standard deviation of peaks for EC subjects; [6]Standard deviation of peaks for health and EC subjects
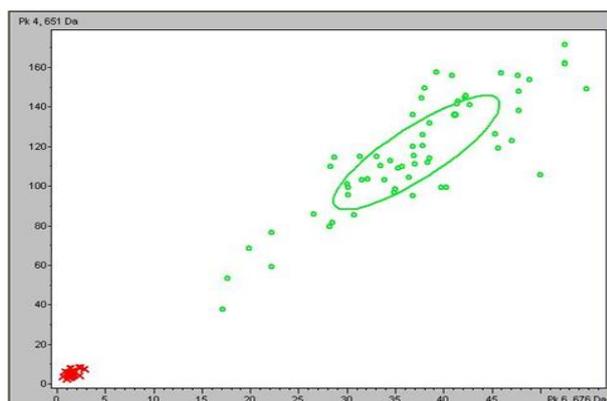
Figure 4. Two-dimensional Peak Distribution View of the Two Peaks Selected for the Diagnostic Model. The peak area and the m/z values are indicated on the x- and y-axes. The ellipses represent the standard deviation of the class average of the peak areas/intensities

peptides. Three peptides selected for model construction were shown in Table 2. Area under curve (AUC) of peak 1, at m/z 622 ($P < 0.000001$) was 1.0, which corresponds to a highly accurate test, according to the criteria suggested by Swets. Moreover areas of these peaks in the spectra of postoperative chemotherapy were statistically different from those of the health volunteers. Combination of the two peaks allowed to yielding a specificity of 100%, and a sensitivity of 100% for postoperative chemotherapy (Figure 3).

The Tumor metastasis and recurrence spectra from 0 to 20 k Da were obtained by MALDI-TOF MS in linear mode. On average about 84 signals common to the two groups have been detected in this mass range and about 55 were identified by the ClinProt software with a statistically different area ($P < 0.01$ by t analysis) in model construction population, including 6 upregulated and 49 downregulated peptides. Four peptides selected for model construction were shown in Table 3. Area under curve (AUC) of peak 1, at m/z 622 ($P < 0.000001$) was 1.0, which corresponds to a highly accurate test, according to the criteria suggested by Swets. Moreover areas of these peaks in the spectra of tumor metastasis and recurrence were statistically different from those of the health volunteers. Combination of the two peaks allowed to yielding a specificity of 100%, and a sensitivity of 100% for Tumor metastasis and recurrence (Figure 4).

Discussion

At present, the conventional esophageal cancer diagnosis is through the endoscope and X-ray inspection, diagnosis depends on pathological. But these invasive method limits the early census and after treatment monitoring. Research showed that, if can be in the organization or the blood to the tumor markers for early stage esophageal cancer will help to make the right diagnosis, since it is difficult to get early esophageal cancer histological specimens, looking for serum markers for early diagnosis, staging and pathologic types of distinguish has important significance. In the past ten years, have found some of esophageal cancer markers, such as carcinoembryonic antigen (CEA), serum amyloid A (SAA), calcium phospholipid binding protein I (Annexin I) and glutamine transferase 3 (TGM 3), etc (Hu et al., 2004; Du et al., 2007; Banki et al., 2008; Uemura et al., 2009). It has been proved that some markers is able to assess the clinical treatment effect. However, all of these markers in early stage are not very good sensitivity (positive predictive value) and specificity (negative predictive value). Therefore, so far has not been found that can be used for routine monitoring of esophageal cancer tumor markers.

Tumor growth is one of the many factors and many stage, many genes involved in the complicated process, including multiple gene mutation of molecular events, such as the activation of oncogene and tumor suppressor gene function loss. Therefore, the present research to find a group of differentially expressed proteins or peptides and based on a variety of characteristics of the diagnosis model, so as to avoid using a single tumor markers or a number of tumor markers from the simple superposition detection sensitivity and specificity of contradictions. Therefore, so far has not been found that can be used for routine monitoring of esophageal cancer tumor markers. Tumor growth is one of the many factors and many stage, many genes involved in the complicated process, including multiple gene mutation of molecular events, such as the activation of oncogene and tumor suppressor gene function loss. Therefore, the present research to find a group of differentially expressed proteins or peptides and based on a variety of characteristics of the diagnosis model, so as to avoid using a single tumor markers or a number of tumor markers from the simple superposition detection sensitivity and specificity of contradictions. MALDI - TOF MS technology is widely applied to the technology of proteomics, and serology tumor markers research field has more advantages. MALDI - TOF MS method can identification of cell, tissue or the body all the protein, providing a set of protein function and its mode of information, can also reflect the intracellular genetic characteristics and the effect of external factors, so as to test specimen of a large number of proteins, is a more overall detection to provide the possibility. This study applies the MALDI ClinProt and magnetic bead sorting technical detection esophageal squamous carcinoma and healthy controls serum specimens, its purpose is to establish a Chinese esophageal squamous carcinoma diagnosis model for tumor markers. Based on EC patients

EC operated, EC operated suffered to chemotherapy and recurrence, and 20 healthy individuals protein map, through data analysis, Pattern consisting of 1 protein peaks, separated EC patients (m/z 650.75) from the healthy individuals with a sensitivity of 100.0% and a specificity of 100.0%. Pattern consisting of 10 protein peaks, separated EC operated (m/z 676.61, 786.1, 786.58), EC operated suffered to chemotherapy (m/z 622.77, 650.66, 676.46) and recurrence (m/z 622.63, 650.56, 690.77, 676.12) with from the healthy individuals a sensitivity of 100.0% and a specificity of 100.0%. Also, we have established a high sensitivity and specificity to distinguish different pathological staging difference model. Application of MALDI mass spectrometry technology is one of the challenges of its repeatability problem, for this, we each experimental procedures are strictly operation, establish a standardized experimental process and avoid the generation of system error (Liu et al., 2010). The second challenge is proteomics research produces a large number of data, so the data analysis is very important. In order to solve this problem, we use support vector machine (SVM) method for the operation. SVM is a based on the principle of statistics of the complicated calculation method, can solve many complicated problems, such as small sample model processing, model selection, assessment (Han et al., 2008). However, these differences of protein expression peak name and its structure and function of the unknown to us, or is the next step to differences in protein purification appraisal to further research the structure function, this paper discusses the significance of it as tumor markers. In short, this study applies magnetic bead sorting and ClinProt method can detect EC patients, EC, operation, postoperative chemotherapy and recurrence tumor markers and establish a high sensitivity and specificity of diagnosis model.

In the later research will further expand the sample size, if can identify specific esophageal cancer, esophageal cancer patients suffered to comprehensive treatment and recurrence markers will be the future research and has profound significance.

Acknowledgements

References

An JY, Fan ZM, Zhuang ZH, et al (2004). Proteomic analysis of blood level of proteins before and after operation in patients with esophageal squamous cell carcinoma at high-incidence area in Henan Province. *World J Gastroenterol*, **10**, 3365-8.

Banki F, Yacoub WN, Hagen JA, et al (2008). Plasma DNA is more reliable than carcinoembryonic antigen for diagnosis of recurrent esophageal cancer. *J Am Coll Surg*, **207**, 30-5.

Dai Y, Hu C, Wang L, et al (2010). Serum peptidome patterns of human systemic lupus erythematosus based on magnetic bead separation and MALDI-TOF mass spectrometry analysis. *Scand J Rheumatol*, **39**, 240-6.

Du XL, Hu H, Lin DC, et al (2007). Proteomic profiling of proteins dysregulted in Chinese esophageal squamous cell carcinoma. *J Mol Med (Berl)*, **85**, 863-75.

Fan NJ, Gao CF, Wang XL, et al (2012). Serum peptidome patterns of colorectal cancer based on magnetic bead separation and MALDI-TOF mass spectrometry analysis. *J Biomed Biotechnol*, **2012**, 985020.

Han M, Liu Q, Yu J, et al (2008). Detectiong and significance of serum protein markers of small-cell lung cancer. *J Clin Lab Anal*, **22**, 131-7.

Hu N, Flaig MJ, Su H, et al (2004). Comprehensive characterization of annexin I alterations in esophageal squamous cell carcinoma. *Clin Cancer Res*, **10**, 6013-22.

Liu LH, Shan BE, Tian ZQ, et al (2010). Potential biomarkers for esophageal carcinoma detected by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Chem Lab Med*, **48**, 855-61.

Maurer K, Eschrich K, Schellenberger W, et al (2013). Oral brush biopsy analysis by MALDI-ToF Mass Spectrometry for early cancer diagnosis. *Oral Oncol*, **49**, 152-6.

Shao C, Tian Y, Dong Z, et al (2012). The use of principal component analysis in MALDI-TOF MS: a powerful tool for establishing a mini-optimized proteomic profile. *Am J Biomed Sci*, **4**, 85-101.

Sui W, Huang L, Dai Y, et al (2010). Proteomic profiling of renal allograft rejection in serum using magnetic bead-based sample fractionation and MALDI-TOF MS. *Clin Exp Med*, **10**, 259-68.

Sun L, Zhao Q, Zhu G, et al (2011). Octyl-functionalized hybrid magnetic mesoporous microspheres for enrichment of low-concentration peptides prior to direct analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*, **25**, 1257-65.

Uemura N, Nakanishi Y, Kato H, et al (2009). Transglutaminase 3 as aprognositic biomarker in esophageal cancer revealed by proteomics. *Int J Cancer*, **124**, 2106-15.

Yao N, Chen H, Lin H, et al (2008). Enrichment of peptides in serum by C8- function- alized magnetic nanoparticles for direct matrix-assisted laser desorption /ionization time-of-flight mass spectrometry analysis. *J Chromatogr A*, **1185**, 93-101.