Mass spectrometry based protein profiling : taking the steps towards clinical application

Velstra, B.

6 Ultrahigh resolution profiles lead to more detailed serum peptidome signatures of pancreatic cancer

Berit Velstra, Simone Nicolardi, Bart J. Mertens, Bert Bonsing, Wilma E. Mesker, Rob A.E.M. Tollenaar, André M. Deelder, Yuri E.M van der Burgt

Translational Proteomics. 2014;2:39-51

ABSTRACT

Mass spectrometry-based (clinical) proteomics has been widely applied as a technology to find and validate disease-specific protein signatures. MALDI-based peptidome profiles provide a suitable platform for classification of body fluids or tissues, albeit at the cost of being unable to observe low abundant species. Here we show that a fully automated one-step solid-phase extraction serum sample cleanup in combination with fast MALDI acquisition and ultrahigh precision 15 tesla FTICR readout provides a powerful, fast and robust approach for obtaining biomarker signatures. This is exemplified for a cohort of pancreatic cancer patients. Specific "early cancer" symptoms such as pain, jaundice or weight loss are often not experienced, thus delaying diagnosis of the disease. Novel markers for early diagnosis of pancreatic cancer are therefore urgently needed. A total of 273 serum samples, distributed over a calibration and validation set, was processed and mass analyzed within a time frame of 24 hours. In both sets sensitivity and selectivity values were well above 85%. In these "next-generation" MALDI peptidome profiles all species up to 9 kDa were isotopically resolved. Finally, it is noted that the low ppm mass accuracy of peptides and proteins observed between 1 and 9 kDa in the FTICR profiles facilitates sequence identifications.

INTRODUCTION

Pancreatic cancer (PC) is the fourth (females) and fifth (males) leading cause of cancer death in developed countries, with a relatively low annual incidence of 5.4 cases per 100,000 females and 8.2 cases per 100,000 males.[1] Patients often die within the first half year after diagnosis, or have an extremely poor prognosis with an overall five-year survival rate of less than 5%.[2] When surgical resection is possible, five-year survival rates improve to approximately 25%. Unfortunately, when the first symptoms appear most tumors are at an advanced stage and their surgical resection would not improve the prognosis.[3;4] Molecular biomarkers that detect PC at an early stage with high sensitivity and specificity would thus be highly beneficial. At the moment, the only used blood marker for detecting and following PC in the clinic is the mucin-associated carbohydrate antigen CA 19-9. This marker, however, often fails in detecting small, resectable cancers.[5] Consequently, like in other cancer biomarker studies, serum proteomics has become a popular approach to find new markers for PC, since blood is a rich and powerful source of biomarkers in general and samples can be collected in a minimally invasive way.

The discovery of serum biomarkers is mainly performed by mass spectrometry (MS)-based proteomics methods.[6] One of these involves the comparison of serum protein profiles in a "case versus control" manner by matrix-assisted laser desorption/ionization - time of flight (MALDI-TOF) MS.[7] Such profiles (i.e. mass spectra) contain hundreds of features (or peaks), of which the presence and intensity can depend on the physiological and pathological condition of the individual. The statistical analysis of serum peptide and protein profiles obtained from both control and diseased individuals allows the identification of a set of features, or a so-called biomarker signature, that can be valuable in understanding the specific disease. Moreover, the biomarker signature may provide leads to further exploit diagnostic and therapeutic potential. Encouraging results have been obtained using profiling strategies.[8-10] Nevertheless, the route to clinically applicable protein assays faces various types of challenges.[11;12] With regard to the selected methodology, for MS-based peptide profiling approaches the problems can be categorized as follows. First of all, multiple profiling studies have shown to lack reproducibility and could not be validated. In this context, standardization of the protocols used for serum sample collection and for peptide and protein purification is pivotal.[10;13;14] The use of a fully automated high-throughput platform for sample processing based on solid-phase extraction (SPE) has been shown to minimize variation and to improve robustness of the method.[15] Secondly, previous MS-acquisitions such as performed on surface-enhanced laser desorption/ionization (SELDI) platforms were not robust and yielded poor accuracies. In addition, identification of peptides or proteins was cumbersome, or not possible at all in these early profiling studies. However, with current equipment these issues can be considered obsolete. The use of internal standards in combination with modern mass

analyzers now allows precise quantitation and detailed characterization of peptides in high-throughput profiles.[16;17] Thirdly, similar peptide profiles were found for various diseases, implying that the features were not specific. On the other hand, it has been postulated that well-defined degradation of highly abundant proteins into peptides ("degradome") can result in tumor-specific serum peptidome patterns.[18]

Recently, we reported a protein profiling study for PC performed on a fully automated SPE-based serum processing platform.[19] Proteins were first isolated with weak cation exchange (WCX) magnetic beads (MBs) using a 96-channel liquid handling robot, followed by acquisition of linear mode MALDI-TOF profiles in the range of 1 to 12 kDa, and evaluation via linear discriminant analysis with double cross-validation. This resulted in a discriminating WCX-profile for PC with a sensitivity of 78% and a specificity of 89% in the calibration set with an area under the curve (AUC) of 90%. These results were validated with a sensitivity of 74% and a specificity of 91% (AUC 90%). However, an obvious disadvantage of low resolution MS profiles is the fact that (poly)peptides and proteins are measured as broad peaks, thus leading to one of the earlier mentioned problems on peak identification. In a second profiling study using the same PC cohort, serum samples were processed with reversed-phase (RP) C18 MBs, and resulting peptides were measured with high resolution reflectron mode MALDI-TOF MS yielding isotopically resolved profiles up to 4 kDa. For statistical evaluation, a list of 42 different peptides was compiled from which a discriminating profile for PC could be defined, with an area under the curve (AUC) of 92% (98%) a sensitivity of 76% (95%) and specificity of 91% (100%) in the calibration (validation) set. Although the identity of most of these peptides was known or elucidated, it became clear that multiple peptides still overlapped at the resolving power of approximately 11,000.[20] The effect of increased resolving power was therefore further studied in MALDI-profiles obtained by Fourier transform ion cyclotron resonance (FTICR) MS, a platform that has proven to be extremely powerful for the analysis of complex mixtures, such as oil, organic matter and plasma.[21-23] With proper control, mass resolving powers higher than 100,000 (at m/z-value 1,000 with 1s transient) and low or sub-ppm mass measurement errors can be routinely obtained.[24;25]

We have previously developed a MALDI-FTICR workflow on a commercially available platform equipped with a 15 tesla magnet that allows high-throughput and fully automated profiling of human serum peptides and proteins with isotopic resolution up to 15,000 Da.[26;27] By following this approach, in comparison to high resolution TOF analyzers the spectrum alignment is more accurate and the quantification of peptides more robust due to the improved mass measurement precision. In this study this MALDI-FTICR workflow in combination with SPE-based sample cleanup with RPC18-functionalized MBs was applied for the analysis of a clinical cohort. Here, "next-generation" MALDI-FTICR peptide and protein profiles were generated using serum samples obtained from PC patients and control individuals (258 samples in total). Classification performances

of both the calibration and validation set were compared to those previously obtained from the same PC cohort, either processed with different MBs or measured on a different mass analyzer. Discriminating peaks (i.e. a biomarker signature) defined from the calibration set were validated using an independent case-control group. Finally, the low ppm mass accuracy provided by the MALDI-FTICR platform narrows the search window for de novo identifications of peptides and proteins in the profiles.

MATERIALS AND METHODS

Patients and sample collection

For the calibration set, serum samples were obtained from 49 patients with PC prior to surgery, and from 110 (age- and gender-matched) healthy volunteers ("controls") over a time period ranging from October 2002 until December 2008 at the outpatient clinic of the Leiden University Medical Center (LUMC), the Netherlands. Healthy volunteers were partners or accompanying persons of included patients. For the validation set, serum samples were obtained from 39 patients and 75 healthy (age- and gender-matched) volunteers over a time period ranging from January 2009 until July 2010. Patients were selected candidates for curative surgery, thus no patients with primary irresectable tumours were included. All surgical specimens were examined according to routine histological evaluation and the extent of the tumor spread was assessed by TNM (TNM Classification of Malignant Tumors) classification. Informed consent was obtained from all subjects and the study was approved by the Medical Ethical Committee of the LUMC. All samples were collected and processed according to a previously reported standardized protocol.[9]

Briefly, blood samples were drawn by antecubital venipuncture while the individuals, who had not been fasting prior to any invasive procedure, were seated. The samples were collected in an 8.5-cc Serum Separator Vacutainer Tube (BD Diagnostics, Plymouth, UK) and maximally within 4 h at room temperature were centrifuged at 1000 g for 10 min. Serum samples were then distributed into sterile 500-µL barcode labelled polypropylene aliquots (TrakMate; Matrix TechCorp.) and stored at -80°C. All serum samples were thawed on ice once and randomly placed in barcode labelled racks in an 8-channel Hamilton STAR® pipetting robot (Hamilton) for automated aliquotting into 60-µl daughter tubes. The aliquots were stored in 96-tubes racks at -80°C until further sample processing. Samples from the calibration and the validation set were distributed over three 96-tubes racks as following: one full 96-tube rack for both the calibration and validation set and one partially-filled 96-tube rack with 63 samples from the calibration set and 18 samples from the validation set. Identical processing steps were followed for the two sample sets.

High-throughput RPC18-MB chromatography and MALDI spotting

The isolation of peptides from human serum was performed using RPC18-functionalized MBs as previously described.[27] In short, RPC18-MBs were first activated by a three-step washing with a 0.1% TFA solution. Then, for each sample 5 µL of serum was added to the activated beads and incubated for 5 min at room temperature. The beads were washed again three times with 0.1% TFA and peptides were eluted with a 1:1 mixture of water and acetonitrile. Two microliters of each (stabilized) eluate were mixed with 10 µl of an α-cyano-4-hydroxycinnamic acid MALDI matrix solution in a 384-well PCR plate. Then, 1 µl of this mixture was spotted in quadruplicate onto a 600 µm Anchor-Chip™ MALDI-target plate (Bruker Daltonics). The so-called RPC18 eluates from the calibration and the validation set were spotted onto three 384-spots MALDI-target plate as following: 96 eluates from the calibration set and 96 eluates from the validation set were spotted in quadruplicate onto two distinct MALDI-target plates; the remaining eluates from the two sets were spotted in quadruplicate onto the same MALDI-target plate. This SPE- and MALDI-spotting procedure requires approximately 3 hours per plate of 96 samples.

MALDI-FTICR mass spectrometry and data processing

MALDI-FTICR experiments were performed on a Bruker 15 tesla solariX™ FTICR mass spectrometer equipped with a novel CombiSource (Bruker Daltonics). The MALDI-FTICR system was controlled by Compass solariXcontrol software and equipped with a Bruker Smartbeam-II™ laser system that operated at a frequency of 200 Hz. The 'medium' pre-defined shot pattern was used for the irradiation. Two acquisition settings, namely low-mass method (LM) and high-mass method (HM), respectively, were used to optimize both the sensitivity and resolving power in the mass range from 1013 to 3700 Da and in the mass range from 3500 to 10,000 Da, respectively. These methods were optimized as previously described with some modification.[27] For both methods, each mass spectrum was obtained from the sum of 10 scans of 150 laser shots each and using 512 K data points. Typically, the target plate offset was 100 V with the deflector plate set at 180 V. The ion funnels operated at 100 V and 6.0 V, respectively, with the skimmers at 15 and 5 V. The analyzer entrance was maintained at -7 V, and side kick technology was used to further optimize peak shape and signal intensity. The trapping potentials were set at 0.60 and 0.55 V for the LM and at 0.95 and 0.80 for the HM while for both LM and HM. The required excitation power was 28% with a pulse time of 20.0 µs. The two acquisition settings differentiate for the trapping potentials (LM, 0.6 and 0.55 V; HM, 0.95 and 0.80 V), the required excitation power (LM, 25%; HM, 28%) and pulse time (LM, 10 µs; HM, 20 µs), the time of flight to the ICR cell (LM, 1.350 ms; HM, 2.700 ms) and the quadrupole filter mass (LM, m/z 1300; HM, m/z 2500). For each spotted sample, two duplicate spots were measured using the LM and the other two using the HM. Approximately 4.5 hours were needed to measure 384 MALDI spots (i.e. originating from 96 different serum samples).

DataAnalysis Software 4.0 SP 5 (Bruker Daltonics) was used for the visualization and the calibration of the spectra. Prior to the measurement of each MALDI plate the FTICR system was externally calibrated using a commercially available peptide mix and a protein mix (Bruker Daltonics). The spectra obtained using the LM were internally calibrated only when used for identification purposes. The m/z-values used for the internal calibration of the LM and the HM are reported in Table S1 in the Supplementary Material section. Peaks were determined using the FTMS algorithm with a signal-to-noise threshold of 3 and using the centroid for peak position with a percentage height of 80.

Peak selection and quantification

Protein and/or peptide signals in RPC18 profiles were quantified as follows. First, based on visual inspection of the profiles, 457 and 670 peaks were selected for the LM and HM spectra, respectively, for further analysis. To this end, a so-called reference file was compiled for both types of profiles in such a way that for each selected peak the m/z-value, a peak number and an m/z-window were reported. In the LM profiles, this m/z-window ranged from 0.015 to 0.166 Da while in the HM it ranged from 0.05 to 0.31 Da reflecting the peak width along the spectra. Then, the in-house developed Xtractor tool was used to determine the intensity of each user-defined peak. This open source tool generates uniform data (peak) arrays regardless of spectral content (http://www.msutils.org/Xtractor). MALDI-FTICR profiles were exported as XY (.xy) files, all containing m/z values with corresponding intensities. Although peptide and proteins were measured up to 10,000 Da using the HM method, the peak selection was limited to 9043.3 Da. The analysis of the spectra in the m/z-range from 9043.3 to 10,000 is on-going and the results will be presented in a separate study.

Statistical analysis

Peak intensities were transformed using the logarithmic function, followed by calculation of the mean of peak intensities obtained from replicate spectra. The intensities of isotope peaks belonging to the same peptide were further summed to reduce the number of features and time needed for further analysis. For each sample, 196 and 291 peak intensity values were obtained for the LM and HM, respectively, and were used to statistical analysis. To this end, logistic regression ridge shrinkage (LRRS) analysis was applied to the calibration sets (i.e. LM and HM data from the calibration set) in order to calibrate two diagnostic rules for the classification of the serum sample either as case or control. Each sample was assigned to the group for which the probability was higher. The prediction rules obtained from the application of LRRS on the calibration sets were applied to the validation sets (i.e. LM and HM data from the validation set). Thus, each sample was classified and the results were compared with known disease status. The classification probabilities assigned to each sample using the LM and HM data from the

validation set were further combined. To this end, LRRS analysis was performed on the combination of the Logit transformed probabilities obtained for validation sets. This analysis involves the recalibration of the validated diagnostic rule. For each analysis error rate (error = the amount by which an observation differs from its expected value), sensitivity, specificity and area under the curve (AUC) were calculated. The error rates are based on the sensitivity and specificity values, assuming a prior class probability of 0.5 for each group. Receiver-operating characteristic (ROC) curves with the true-positive rate (sensitivity) were plotted in function of the false-positive rate (1-specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between groups (diseased/healthy). Univariate discriminate analysis was performed to determine which peak varied the most between case and control groups. This study was limited to peaks of which the absolute weighted discriminant coefficient was higher than 0.1 in the multivariate discriminant analysis used to calibrate the discriminant models. Finally, a t-test was performed on a selection of peaks for the calibration sets only.

RESULTS

MALDI-FTICR-MS peptide and protein profiling

Serum samples of PC patients as well as control individuals were processed simultaneously using a previously described fully automated and standardized SPE-based RPC18-MB protocol.[15] Thus obtained MB eluates were spotted onto a MALDI target plate in quadruplicate. Two types of ultrahigh resolution peptide and protein profiles were then acquired applying an automated acquisition procedure on the MALDI-FTICR system (see Materials and methods section). Two out of four spots were used to obtain a so-called low mass (LM) profile (m/z-values from 1,013 to 3,700) and the remaining two spots were used to generate a so-called high mass (HM) profile (m/z-values from 3,500 to 10,000). In total, 273 serum samples were analyzed in this way, thus yielding 1,092 profiles. A typical example of both an LM and HM MALDI-FTICR profile is depicted in Figure 1A. It was verified that all peptides and (small) proteins were measured with isotopic resolution through all the spectra, with typical resolving powers varying from 130,000 (m/z 1039.6727) to 46,000 (m/z 3523.7664) in the LM spectra and from 150,000 (m/z 3680.8709) to 33,000 (m/z 9744.6054) in the HM spectra (as plotted in Figure 2A).

As a result, a large number of peptides or proteins that would overlap in high resolution MALDI-TOF MS were measured as distinct features by MALDI-FTICR MS. Two examples of resolved species are shown in Figure 2B, one for the LM and one for the HM profiles. The ultrahigh resolving power allowed the accurate quantification of the

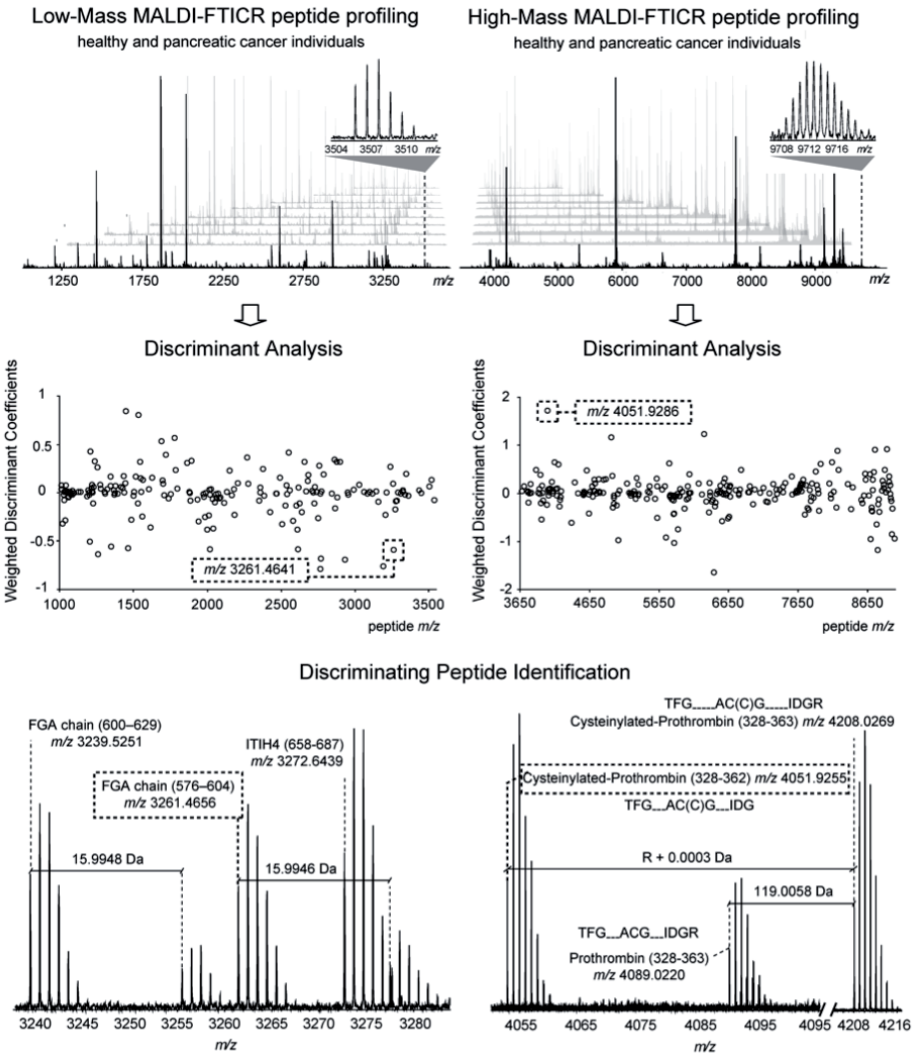

Figure 1. Ultrahigh resolution 15T MALDI-FTICR peptide and protein profiles obtained from human serum after RPC18 magnetic bead-based SPE (upper panel). Using two different acquisition settings, optimized for the low-mass and high-mass range, peptides and proteins were isotopically resolved up to 10 kDa. For statistical analysis, logistic regression ridge shrinkage analysis was used to find changes in the peptide and protein profiles obtained from healthy individuals and pancreatic cancer patients. A weighted discriminant coefficient was assigned to each detected peptide or protein according to its discriminant property in the way that the higher the value of the discriminant coefficient the higher the case probability (middle panel). Identifications of the most discriminating peptides and proteins were based on previously reported peptide IDs or on accurate mass measurement of mass differences in the spectra (lower panel). Note that these identifications need further confirmation by MS/MS-data.

selected peptides and proteins in all the spectra. After manual inspection of the profiles, 457 and 670 peaks for the LM and HM, respectively, were selected for statistical analysis. After taking into account isotopic peaks from the same species, 196 peptides remained from the 457 selected peaks in LM spectra and 291 peptides or proteins remained from the 670 selected peaks in HM spectra. Peptides and proteins were detected with signal intensities that typically ranged over two orders of magnitude. For example, Fibrinopeptide alpha chain (2–16) (at m/z-value 1465.6554) was often observed as the most intense peptide, and was 304 times more intense than Complement C4-A (1337–1350) (at m/z-value 1626.8459) detected with a signal-to-noise ratio (S/N) of 6.6, in a typical spectrum. Thus, peptides observed with low S/N were also evaluated. For example, the peptide identified as oxidized Fibrinogen beta chain (45-71) (m/z 2898.5334) (see Serum peptide identification by accurate mass difference measurement section) was observed in the spectra in the calibration set with an averaged S/N 9.6 with a standard deviation (SD) of 6.4, while the highly intense Complement C3f fragment peptide (at m/z-value 2021.1039) was observed with an averaged S/N of 2035 with an SD of 345. As a final remark, from 12 out of 1,032 profiles the quality was insufficient for further statistical analysis, most likely because of failed MALDI spotting.

Statistical analysis

The signal intensities of all selected peaks were determined in all serum profiles using the Xtractor tool described in the Materials and methods section. As shown in Figure 2A, the m/z-windows in the reference files were fine-tuned according to the resolving power calculated for each m/z-value. The presence of different peptides with close masses was also taken into account as well as the mass measurement precision (see Figure 2B). The optimization of this m/z-window allowed the accurate quantification of all peaks selected from the spectra. Thus obtained peak intensity values were then used for statistical analysis. To this end, a discriminate model was first calibrated and then validated using LRRS analysis on the calibration and validation sets, respectively.

The ROC curves resulted from this analysis are shown in Figure 3. Error rates of 0.136 and 0.104, sensitivities of 88% and 91% and specificities of 96% and 94% with AUC of 0.987 and 0.980 were obtained for the LM and HM validation sets, respectively. The LRRS analysis performed on the combination of the logit of the classification probabilities obtained for the LM and HM validation sets resulted in an error of 0.0784, a sensitivity of 89% and a specificity of 100% with an AUC of 0.989. The logit transformation involves a recalibration of the discriminant models obtained using the validation sets. The discriminant analysis performed on the recalibrated validation sets resulted in errors of 0.098 and 0.088, sensitivities of 88% and 90% and specificities of 96% and 93% with AUC of 0.987 and 0.977 for the LM and HM validation sets, respectively.

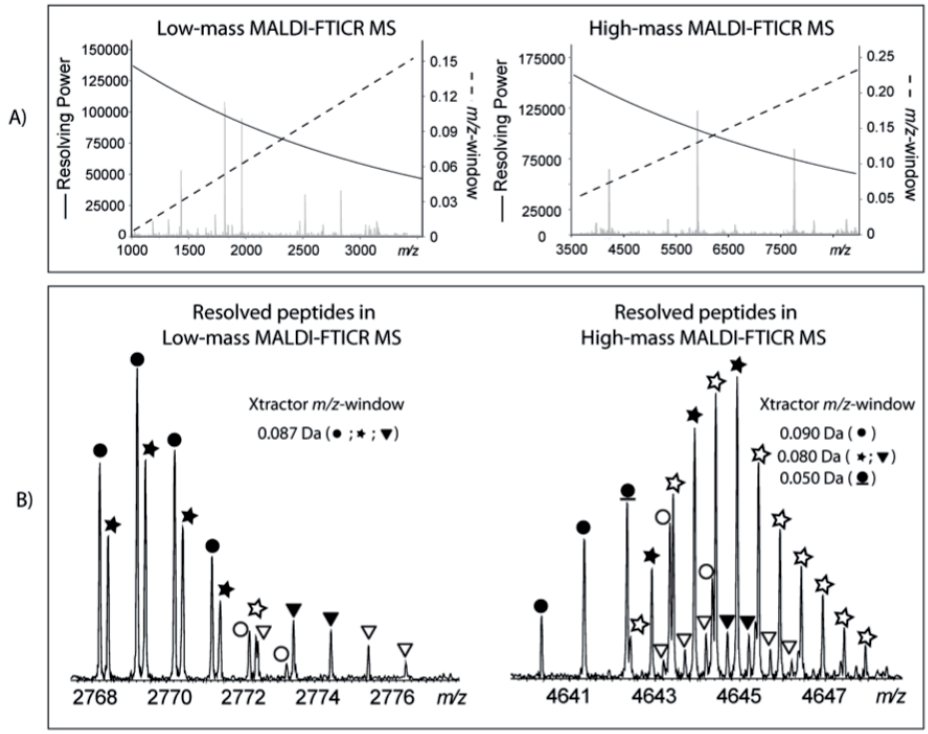

Figure 2. For each isotopically resolved peptide or protein signal a specific m/z-window was defined to allow accurate quantification. The m/z-window was based on the resolving power at a specific m/z-value (A) and in the case of overlapping species further optimized (B). Note that the resolving power at m/z 3,500 is higher in the High-mass measurements (i.e. ~ 150,000) than in the Low-mass measurements (i.e. ~ 46,000) as a result of the specific broadband frequency-sweep waveform that was used to excite the ions into the ICR cell.

A sequential analysis was performed by sub-typing the PC cases into cases without any metastasis (i.e. regional lymph node-negative (LN-) and no distant metastasis (DM-)) versus cases that were lymph node-positives (LN+) and/or showed distant metastasis (DM+), based on TNM-classification summarized in Table 1. This sub-typing resulted in a box plot (see Figure 3) with clear separation between controls and cases, and in addition good separation between cases with and without metastasis (Wilcoxon Mann-Whitney test with a p-value of 7.7293e-05 for controls versus "(LN-)and(DM-)", and a p-value of 0.015844 for "(LN+)and/or(DM+)" versus "(LN-)and(DM-)"). Patient characteristics, number of serum samples, and the results of the classification methods set are shown in Table 1.

A logistic regression coefficient weighted by the standard deviation of the peak intensity was assigned to each peak as determined from multivariate analysis on the calibration set (i.e. the calibration of the discriminating rule). These discriminant weights

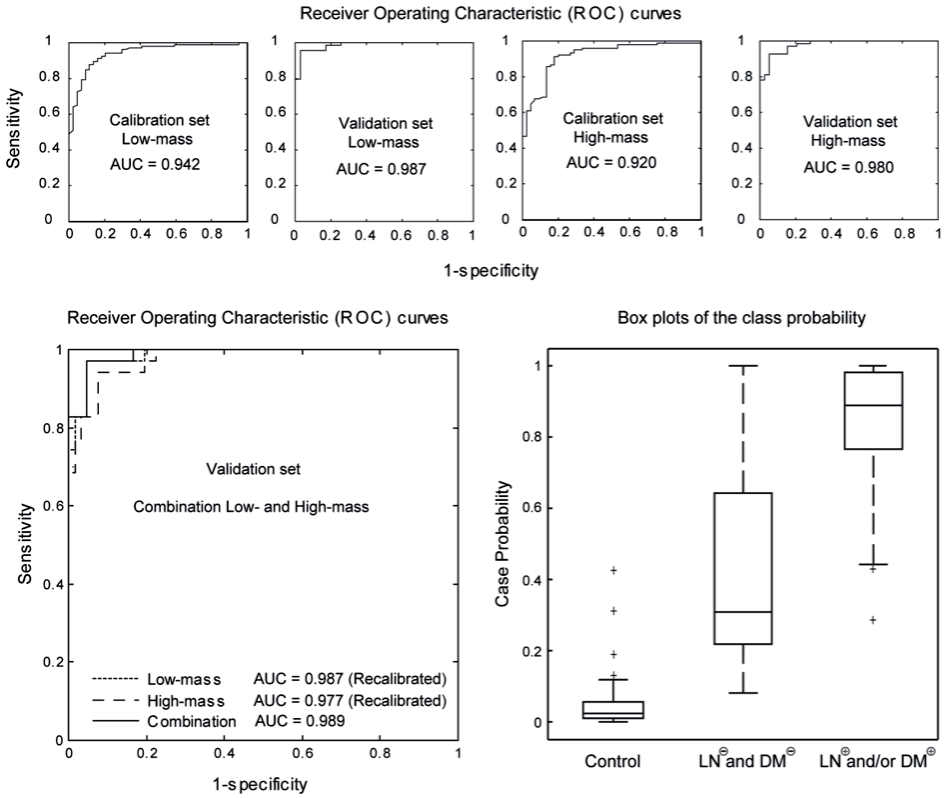

Figure 3. Receiver Operating Characteristic (ROC) curves generated from the case-control classification of the serum samples from both the calibration sets and the validation sets using Low-mass (LM) and High-mass (HM) data (four upper panels). The larger the area under the ROC curve (AUC), the better two groups (diseased/healthy) can be distinguished. The lower panel on the left hand-side shows ROC curves generated from the combination of the classification results obtained from the validation set. This combination involved the recalibration of the validated diagnostic rules, resulting in recalibrated LM and HM data. The lower panel on the right hand-side shows box plots that are obtained after sub-typing the PC cases into cases without any metastasis (i.e. regional lymph node-negatives (LN-) and no distant metastasis (DM-)) versus lymph node-positives (LN+) and/or distant metastasis (DM+) based on TNM-classification.

denote the conditional effect associated with each peak, after taking into account the variation in expression across the other selected peaks. Thus, the higher the value of the discriminant weight the higher the case probability. Note that the reverse applies to control samples. The plots with the weighted discriminant coefficients vs the m/z-values are shown in Figure (middle panel). A t-test was performed on peaks with absolute discriminant coefficient higher than 0.1 in the calibration set. A P-value smaller than 0.001 was considered as significant.

Peaks that satisfied these criteria are reported in Table 2 with corresponding protein names, t-test values, standard deviations (SD), P-values, 95%-confidence interval and

Table 1. Patient characteristics for the calibration and validation sets.

	Calibration Set			Validation Set		
	HM [Correct Classif.]	**LM [Correct Classif.]**	**Combination HM and LM**	**HM [Correct Classif.]**	**LM [Correct Classif.]**	**Combination HM and LM**
N. of samples	46	45		39	35	
Age median (min - max)	65.6 (41-80)	65.2 (41-80)		63.4 (38-81)	62.8 (38-81)	
Male	23 [16]	22 [17]	16	17 [15]	15 [11]	12
Female	23 [15]	23 [15]	17	22 [17]	20 [15]	15
Localisation (pancreas)						
Head	38 [28]	36 [27]	33	35 [28]	31 [22]	23
Body	3 [1]	3 [2]	1	1 [1]	1 [1]	1
Tail	3 [1]	4 [2]	2	1 [1]	1 [1]	1
Unknown/Other	2 [1]	2 [1]	1	2 [2]	2 [2]	2
Stage						
IA	7 [7]	7 [6]	7	2 [1]	2 [0]	0
IB	5 [3]	5 [3]	3	2 [2]	1 [1]	1
IIA	3 [2]	3 [1]	2	3 [1]	3 [0]	1
IIB	19 [13]	17 [14]	13	21 [18]	20 [17]	17
III	5 [3]	5 [4]	3	2 [2]	2 [2]	2
IV	7 [3]	8 [4]	5	9 [8]	7 [6]	6
Tumor differentation						
Unknown (irresectable tumor)	12 [7]	11 [6]	7	11 [10]	10 [9]	9
Grade 1	10 [5]	11 [5]	6	6 [2]	6 [3]	3
Grade 2	11 [9]	11 [9]	9	15 [14]	13 [10]	10
Grade 3	13 [10]	12 [12]	11	7 [6]	6 [4]	5

the weighted discriminant coefficients. Note that the P-values here reported ranged from 6.0x10-4 to 4.0x10-9 indicating a high statistical significance.

Serum peptide identification by accurate mass difference measurement

A list of serum peptides and proteins that are commonly observed in MALDI-TOF profiles obtained after RPC18-based sample cleanup has been compiled previously and this was used for statistical evaluations.[9] However, a number of peptides remained unidentified in this list, and moreover in the current MALDI-FTICR ultrahigh resolution profiles many RPC18-MB serum eluate peaks are unknown. Likely, a large number of these degradome peptides originate from the same high abundant proteins after proteolytic cleavage as was reported earlier.[18;28;29] New peptide assignments were performed based on matching accurate mass measurements of m/z-differences between peaks in 15T MALDI-FTICR spectra with possible decreased or increased sequences ("degradome"). Thus, a search

Table 2. Peptides present in MALDI-FTICR precision profiles with low P-values after comparative analysis of serum samples from pancreatic cancer patients and control individuals.

Observed *m/z* [M+H]+	Protein Name	T-test	SD	P value	CI lower	CI upper	Bweighted
Low-Mass MALDI-FTICR profiles							
1206.5753	FPA chain (5-16)	3.7	3.4	2.7E-04	1.1	3.5	-0.51
1211.6535	Complement C3f fragment (7–16)	-4.0	2.0	1.1E-04	-2.1	-0.7	0.42
1263.5965	FPA chain (4–16)	5.0	3.0	1.3E-06	1.6	3.7	-0.64
1348.7118	Complement C3f fragment (6–16)	-3.6	1.5	4.1E-04	-1.5	-0.4	0.16
1350.6282	FPA chain (3–16)	4.4	3.5	2.4E-05	1.5	3.9	-0.56
1449.7597	Complement C3f fragment (5–16)	-4.0	3.3	1.1E-04	-3.5	-1.2	0.84
1561.7265	Protrombin (315-327)	4.5	1.9	1.4E-05	0.8	2.2	-0.25
1616.6584	FPA chain 1P	5.0	3.4	1.8E-06	1.8	4.2	-0.36
1626.8458	Complement C4-A (1337–1350)	-3.7	1.1	3.2E-04	-1.1	-0.3	0.15
1698.7450	n.i.	-3.5	0.9	6.0E-04	-0.9	-0.3	0.12
1718.9430	Complement C3f fragment (4–16)	-4.2	1.6	4.4E-05	-1.7	-0.6	0.39
1786.8545	ITIH4 (671–687)	-4.7	1.7	6.5E-06	-2.0	-0.8	0.23
1984.9879	n.i.	3.6	1.2	4.8E-04	0.3	1.2	-0.25
2685.3746	Thrombin light chain (342-363)	-5.8	1.1	3.9E-08	-1.5	-0.7	0.18
2768.2293	FGA chain (576–600)	3.9	3.5	1.5E-04	1.2	3.7	-0.79
2898.5334	FGB chain (45-71) + 15.9952 Da	-6.3	1.3	4.0E-09	-2.0	-1.0	0.32
2931.2909	FGA chain precursor (576–601)	4.9	3.8	2.2E-06	2.0	4.7	-0.70
3190.4279	FGA chain (576–603)	5.9	3.1	2.9E-08	2.1	4.3	-0.76
3206.4243	FGA chain (576–603) [Met-ox]	5.2	2.1	7.1E-07	1.2	2.7	-0.23
3261.4664	FGA chain (576–604)	5.1	3.5	1.3E-06	1.9	4.4	-0.59
High-Mass MALDI-FTICR profiles							
3679.8665	n.i.	-4.3	0.7	3.7E-05	-0.8	-0.3	0.14
3806.8902	n.i.	-4.2	0.7	5.4E-05	-0.7	-0.3	0.19
4051.9297	Thrombin light chain (328-363) (Cysteinylated)	-5.8	2.8	4.4E-08	-3.8	-1.9	1.72
4108.9479	Thrombin light chain (328-363) (Cysteinylated+Gly)	-4.6	1.5	7.6E-06	-1.8	-0.7	0.24
4394.0803	n.i.	3.8	1.6	2.1E-04	0.5	1.7	-0.62
4854.2768	Apolipoprotein CIII2 (21-99) *	-3.9	2.5	1.7E-04	-2.6	-0.8	0.28
4961.4906	FGA chain (529–574) or (513–558)	-4.5	2.9	1.3E-05	-3.3	-1.3	1.16
4979.4945	FGA chain (529–574) or (513–558) + 15.9952 Da	-5.0	2.1	1.7E-06	-2.6	-1.1	0.36
4985.4806	n.i.	-3.7	1.7	2.9E-04	-1.7	-0.5	-0.30
5802.6370	FGA chain (576-627)	-3.7	3.2	2.7E-04	-3.2	-1.0	0.62
6223.3096	n.i.	-4.5	1.4	1.2E-05	-1.6	-0.6	0.19
7151.4191	n.i.	-3.5	1.6	5.9E-04	-1.5	-0.4	0.11
8205.0930	Apolipoprotein CII (23-101)	-3.8	1.8	2.4E-04	-1.9	-0.6	0.88
8781.2819	Apolipoprotein-CIII (21-99) + 15.9952 Da	3.8	2.8	1.8E-04	0.9	2.9	-1.19

***This species was observed as a doubly-charged ion.**

for consecutive mass differences corresponding to one amino acid was performed, starting from a previously identified peptide in the spectrum with relatively highest signal intensity. In this way, new peptides with one or more additional amino acids at the N-terminus or/and the C-terminus or modified peptides (i.e. oxidized, cysteinylated) were identified. Following this strategy the amino acid sequence of 34 new peptides was derived and these are reported in Table 3. In general, the LM and HM profiles provided sub- and low-ppm mass measurement errors for these identifications, respectively.

Two examples of this approach are shown in Figure 1C. The first one is the identification of an oxidized form of the peptide Fibrinogen alpha chain (576–604) that was statistically evaluated with a discriminant weight factor of -0.59 (see Table 2). In the second example the accurate mass-based identification of the species observed at m/z-value 4051.9255 is depicted, a peptide that was found to be the best predictor (i.e. highest absolute discriminant weight) of healthy and disease individuals in HM profiles (see Table 2). The mass difference between this peptide and a peptide previously MS/MS-identified as cysteinylated-Prothrombin (328-363), observed at m/z-value 4208.0269, was 156.1014 Da. This mass difference corresponds to an arginine residue with an error of only 0.3 milliDa. In addition, the accurate measurement of mass differences allowed the identification of peptides containing a single amino acid mutation. For example, a peptide from coagulation factor XIII (Factor XIIIa) alpha chain with a previously reported Val35Leu mutation corresponding to a mass difference of 14.0156 Da between "normal" and mutant fragment peptides was indeed observed (see Table 3). Here, the species at m/z-value 2602.3113 corresponds to a previously identified peptide from Factor XIIIa (14-38), whereas the species at m/z-value 2531.2735 and m/z-value 2545.2883 both lack an alanine residue but differ at the site of mutation (i.e. Val35 Factor XIIIa (15-38) and Leu35 Factor XIIIa (15-38), respectively). It is emphasized that isobaric peptides containing modifications such as oxidation cannot be uniquely characterized by the accurate measurement of mass differences. For this purpose additional MS/MS-experiments are needed to confirm the identifications and localize for instance modified amino acids in the sequence. As a final remark, the accurate and precise MALDI-FTICR mass measurements will allow a reliable match between the MS/MS-data obtained using other MS techniques such as LC-ESI-MS/MS and the peptides observed in the MALDI-FTICR spectra.

DISCUSSION

The past decade, MS-based profiling studies have been carried out to determine disease-specific serum peptidome signatures in a "case-control" setting. Due to the relatively high biological variability of the serum peptidome (and proteome) a large number of samples

Table 3. Proposed sequences of (large) peptides present in MALDI-FTICR precision profiles based on accurate mass measurements.

Protein Name	Peptide Sequence	Mass difference	Calculated *m/z* [M+H]+	Observed *m/z* [M+H]+	Mass Measur. Error (ppm)	MALDI-FTICR method
ITIH4 UniProt Q14624	M.NFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	+(NRF + 0.0033)	3141.6017	3141.6047	0.96	LM
	N.FRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	+(RF + 0.0009)	3027.5588	3027.5593	0.18	LM
	R.PGVLSSRQLGLPGPPDVPDHAAYHPFR or RPGVLSSRQLGLPGPPDVPDHAAYHPF.R	+(R - 0.0004)	2880.4904	2880.4896	-0.26	LM
	R.PGVLSSRQLGLPGPPDVPDHAAYHPF.R	**-**	**2724.3893**	**2724.3889**	**-0.13**	**LM**
	P.GVLSSRQLGLPGPPDVPDHAAYHPF.R	-(P + 0.0002)	2627.3365	2627.3360	-0.19	LM
Thrombin light chain UniProt P00734	L.FEKKSLEDKTERELLESYIDGR	+(FEKK - 0.0011)	2685.3730	2685.3720	-0.36	LM
	F.EKKSLEDKTERELLESYIDGR	+(EKK - 0.0008)	2538.3046	2538.3039	-0.28	LM
	E.KKSLEDKTERELLESYIDGR	+(KK - 0.0005)	2409.2620	2409.2615	-0.19	LM
	K.KSLEDKTERELLESYIDGR	+(K + 0.0004)	2281.1670	2281.1676	0.23	LM
	K.SLEDKTERELLESYIDGR	**-**	**2153.0721**	**2153.0722**	**0.04**	**LM**
	S.LEDKTERELLESYIDGR	-(S + 0.0012)	2066.0400	2066.0390	-0.53	LM
	L.EDKTERELLESYIDGR	-(SL + 0.0001)	1952.9560	1952.9560	0.02	LM
	E.DKTERELLESYIDGR	-(SLE + 0.0010)	1823.9134	1823.9124	-0.52	LM
	D.KTERELLESYIDGR	-(SLED + 0.0001)	1708.8864	1708.8865	0.01	LM
Thrombin light chain UniProt P00734	TFGSGEADC(CysGly)GLRPLFEKKSLEDKTERELLESYIDG	+(G + 0.0032)	4265.0489	4265.0520	0.73	HM
	TFGSGEADC(Cys)GLRPLFEKKSLEDKTERELLESYIDGR	**-**	**4208.0275**	**4208.0274**	**-0.02**	**HM**
	TFGSGEADC(CysGly)GLRPLFEKKSLEDKTERELLESYIDG	-(R-Gly-0.0001)	4108.9478	4108.9479	0.02	HM
	TFGSGEADCGLRPLFEKKSLEDKTERELLESYIDGR	- (Cysteinylation + 0.0010)	4089.0234	4089.0222	-0.27	HM
	TFGSGEADC(Cys)GLRPLFEKKSLEDKTERELLESYIDG	-(R - 0.0034)	4051.9263	4051.9297	0.82	HM
	TFGSGEADC(CysGly)GLRPLFEKKSLEDKTERELLESYIDG	+(G + 0.0032)	4265.0489	4265.0520	0.73	HM
Fibrinogen alpha chain UniProt P02671	**K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPV.R**	**-**	**5901.70298**	**5901.70641**	**0.58**	**HM**
	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRP.V	-(V + 0.0010)	5802.63456	5802.6370	0.42	HM

Table 3. Proposed sequences of (large) peptides present in MALDI-FTICR precision profiles based on accurate mass measurements. (continued)

Protein Name	Peptide Sequence	Mass difference	Calculated *m/z* [M+H]+	Observed *m/z* [M+H]+	Mass Measur. Error (ppm)	MALDI-FTICR method
Fibrinogen alpha chain UniProt P02671	**K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHA.K**	**-**	**5334.3536**	**5334.3572**	**0.67**	**HM**
	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKR.G	-(GHA + 0.0002)	5069.2362	5069.2395	0.65	HM
	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTK.R	-(RGHA + 0.0099)	4913.1351	4913.1288	-1.28	HM
	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHST.K	-(KRGHA - 0.0005)	4785.0401	4785.0442	0.86	HM
Platelet Factor 4 UniProt P02776	FASAEAEEDGDLQCLCVKTTSQVRPRHITSLEVIKAPHCPTAQLIATLKNGRKICLDLQAPLYKKIIK-KLLES	+(FASA + 0.0385)	8141.3692	8141.3932	2.94	HM
	F.SAEAEEDGDLQCLCVKTTSQVRPRHITSLEVIKAPHCPTAQLIATLKNGRKICLDLQAPLYKKIIK-KLLES	+(SA - 0.0027)	7923.2636	7923.2464	-2.17	HM
	S.AEAEEDGDLQCLCVKTTSQVRPRHITSLEVIKAPHCPTAQLIATLKNGRKICLDLQAPLYKKIIKK LLES	+(A + 0.0297)	7836.2316	7836.2468	1.93	HM
	A.EAEEDGDLQCLCVKTTSQVRPRHITSLEVIKAGPHCPTAQLIATLKNGRKICLDLQAPLY KKIIKKLLES	**-**	**7765.1945**	**7765.1799**	**-1.88**	**HM**
	E.AEEDGDLQCLCVKTTSQVRPRHITSLEVIKAPHCPTAQLIATLKNGRKICLDLQAPLYKKIIKKLLES	-(E + 0.0132)	7636.1518	7636.1240	-3.64	HM
	A.EEDGDLQCLCVKTTSQVRPRHITSLEVIKAPHCPTAQLIATLKNGRKICLDLQAPLYKKIIKKLLES	-(EA + 0.068)	7565.1147	7565.0933	-2.83	HM
HMW Kininogen UniProt P01042	**H.NLGHGHKHERDQGHGHQ**	**-**	**1943.9080**	**1943.9071**	**-0.47**	**LM**
	HNLGHGHKHERDQGHGHQ	+(H + 0.0001)	2080.9669	2080.9661	-0.39	LM
Complement C3f fragment UniProt P01024	**SSKITHRIHWESASLLR**	**-**	**2021.1039**	**2021.10398**	**0.04**	**LM**
	S.SKITHRIHWESASLLR	-(S + 0.0005)	1934.0719	1934.0715	-0.20	LM
	S.KITHRIHWESASLLR	-(SS + 0.0013)	1847.0399	1847.03869	-0.63	LM
	H.RIHWESASLLR	-(SSKITH + 0.0003)	1367.75424	1367.75405	-0.14	LM
	I.HWESASLLR or RIHWESASL.L	-(SSKITHRI/L - 0.0005)	1098.56907	1098.56968	0.56	LM
Factor XIIIa UniProt P00488	**R.AVPPNNSNAAEDDLPTVELQGVVPR.G**	**-**	**2602.3107**	**2602.3113**	**0.22**	**LM**
	A.VPPNNSNAAEDDLPTVELQGLVPR.G	-(A - 14.0141)	2545.2893	2545.2883	-0.38	LM
	A.VPPNNSNAAEDDLPTVELQGVVPR.G	-(A - 0.0006)	2531.2736	2531.2735	-0.04	LM

are required for statistical evaluation. Thus, high-throughput analytical methodologies have been adopted in combination with MS, pioneered by SELDI-TOF platforms. In the same period, high-throughput robotic platforms with more flexible and user-defined sample preparation protocols were combined with MALDI-TOF read-out. Both low-resolution TOF-profiles with a wide m/z-range and high-resolution profiles with smaller m/z-windows were reported for proteins and peptides, respectively.[7;30;31] However, single- or even multi-step protein fractionations still yield highly complex samples and the low resolving powers in linear mode SELDI- or MALDI-TOF profiles do not allow accurate quantification of the profiled species. Peptides up to m/z-values of 4,500 can be routinely analysed with isotopic resolution using TOF-analysers in reflectron mode, but at the cost of restricting the analyzed m/z-range and thus excluding proteins from the evaluation. Moreover, reflectron mode profiles still contain a significant number of overlapping peptides, as we previously demonstrated in ultrahigh resolution MALDI-FTICR profiles.[20]

In this study the ultrahigh resolving power provided by a 15 tesla MALDI-FTICR system was exploited in terms of discriminative power of case-control peptidome profiles and identification of observed species. This is the first profiling study that reports on the application of such ultrahigh resolution profiles exemplified by a clinical cohort of serum samples from healthy individuals and PC patients. Aiming for cancer-specific peptide and protein signatures, these serum samples were first fractionated on a fully automated SPE-platform based on functionalized MBs and then profiled using a 15T MALDI-FTICR mass spectrometer. In total, 487 peptides or small proteins (i.e. 196 and 291 in LM and HM spectra, respectively) were measured with isotopic resolution in the m/z-range

1-9 kDa and quantified with high accuracy and precision. The ultrahigh resolving power allowed the correct quantification of peptides or proteins that previously were observed to suffer from overlapping isotopic distributions in lower resolution profiles (see Figure2). Note that the total number of detectable peptides was higher, i.e. several peptides were detected only in few particular samples, probably due to a higher expression of a particular protein or an elevated protease activity.

Two different MALDI-FTICR acquisition methods, namely a low mass and a high mass method, were used to generate peptide and protein profiles from two independent groups of serum samples. A calibrated and validated discriminating rule built on the combination of the data obtained from the two MALDI-FTICR methods resulted in a sensitivity of 89% and a specificity of 100% with an AUC of 0.989. These results corroborate classification numbers from our previous MALDI-TOF studies.[19;32] The t-test analysis performed on the peptides with absolute discriminant weights higher than 0.1 resulted in the identification of 34 peptides that (i.e. p-value lower than 0.001) differentiate between case and control groups (see Table 2). The high precision and accuracy of the mass measurements allowed the identification of 26 of these peptides either by com-

parison with previously reported peptides or by accurate mass measurement of mass differences in the spectra (see the Materials and Method section). Application of the latter approach resulted in the identification of peptides generated through proteolysis of the same protein. In fact, starting from a previously identified peak (i.e. peptide) it was found that accurate measurement of the difference between that specific m/z-value and the m/z-value of a new peak matched to a similar peptide with either one amino acid more or less at the C- or the N-terminus, corresponding to the "overall" protein sequence. Thus, up to 8 new peptides could be identified starting from the fragment peptide K.SLEDKTERELLESYIDGR of thrombin light chain (UniProt P00734) (see Table 3). Nevertheless, the presence of isobaric peptides cannot be excluded and MS/MS experiments are required to further validate the identifications. In conclusion, using the two identification approaches described above, we are now able to further expand the total number of identified peptides, especially at higher m/z-values. Other MALDI-profiling methods that so far have been used for the characterization of human serum peptides were not suitable for the identification of high molecular weight peptides or proteins, because these lacked sensitivity and resolving power.[28;29]

As a final remark, it should be noted that at this stage the peptidome profiles were not evaluated for the m/z-range from 9,000 to 10,000. Here, both the high density of peaks and the relatively lower resolving power do not permit binning of the data points. The most abundant peaks present in this range were identified as apolipoprotein-CIII isoforms and these data will be evaluated in a separate study using a different quantification method.[26] In this study, we have shown that high quality human serum peptide and protein profiles can be generated using a standardized and robust protocol for the sample preparation and ultrahigh resolution 15T MALDI-FTICR MS for the mass measurements. The use of this mass analyzer allowed the isotopic resolution and the accurate and precise mass measurement of a high number of peptides and small proteins in a wide m/z-range. Notably, recent innovation in ICR-cell technology potentially provides similar performance at a lower magnetic field strength.[33] The statistical analysis of profiles generated from a clinical cohort of samples allowed the discrimination between healthy individuals and PC patients with sensitivity and specificity comparable with those reported by other authors using MALDI-TOF MS. A total of 273 serum samples was processed and mass analyzed within a time frame of 24 hours and the high quality of the data both facilitated the interpretation and evaluation of the generated profiles. These ultrahigh resolution mass spectra represent a "next-generation" of MS-based peptidome profiles and provide a new tool for a more detailed description of the high-abundant proteins in clinical serum sample cohorts aiming for new diagnostic leads.

Supplementary Table 1. Peptides present in MALDI-FTICR precision profiles that were used for internal calibration.

Protein (Uniprot Entry)	Peptide sequence	*m/z*-value	Method
Fibrinopeptide alpha chain 5-16 (P02671)	G.EGDFLAEGGGVR	1206.5749	LM
Fibrinopeptide alpha chain 3-16 (P02671)	D.SGEGDFLAEGGGVR	1350.6284	LM
Phosphorylated Fibrinopeptide alpha chain (P02671)	ADSpGEGDFLAEGGGVR	1616.6594	LM
Complement C3f fragment (P01024)	SSKITHRIHWESASLLR	2022.1067	LM
Factor XIIIa 14–38 (P00488)	R.AVPPNNSNAAEDDLPTVELQGVVPR.G	2603.3136	LM
Fibrinogen alpha chain precursor 576–601 (P02671)	K.SSSYSKQFTSSTSYNRGDSTFESKSY.K	2932.2944	LM
Fibrinogen alpha chain 76–604 (P02671)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMA.D	3262.4670	LM
Prothrombin 328-363 (P00734)	TFGSGEADC(Cys)GLRPLFEKKSLEDKTERELLESYIDGR	4208.0275	HM
Fibrinogen alpha chain 577–624 (P02671)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHA.K	5337.3618	HM
Fibrinogen alpha chain 577–629 (P02671)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPV.R	5904.7111	HM
Apolipoprotein-CIII (P02656)	SEAEDASLLSFMQGYMKHATKTAKDALSSVQESQVAQQARGWVTDGFSSLKDYW STVKDKFSEFWDLDPEVRPTSVAA	8765.2387	HM
Glycosylated Apolipoprotein-CIII (P02656)	SEAEDASLLSFMQGYMKHATKTAKDALSSVQESQVAQQARGWVTDGFSSLKDYW STVKDKFSEFWDLDPEVRPT(GalNAcGalneu5Ac)SVAA	9421.4664	HM

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: Cancer J. Clin. 2011;61:69-90.
2. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011. CA: Cancer J. Clin. 2011;61:212-236.
3. Wong T, Howes N, Threadgold J, Smart HL, Lombard MG, Gilmore I, Sutton R, Greenhalf W, Ellis I, Neoptolemos JP. Molecular Diagnosis of Early Pancreatic Ductal Adenocarcinoma in High-Risk Patients. Pancreatology 2001;1:486-509.
4. Sener SF, Fremgen A, Menck HR, Winchester DP. Pancreatic cancer: a report of treatment and survival trends for 100,313 patients diagnosed from 1985-1995, using the National Cancer Database. J. Am. Coll. Surg. 1999;189:1-7.
5. Goonetilleke KS, Siriwardena AK. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. Eur. J. Surg. Oncol. 2007;33:266-270.
6. Nilsson T, Mann M, Aebersold R, Yates JR, Bairoch A, Bergeron JJM. Mass spectrometry in high-throughput proteomics: ready for the big time. Nat. Methods 2010;7:681-685.
7. Albrethsen J. The first decade of MALDI protein profiling: a lesson in translational biomarker research. J. Proteomics 2011;74:765-773.
8. de Noo ME, Tollenaar RA, Deelder AM, Bouwman LH. Current status and prospects of clinical proteomics studies on detection of colorectal cancer: hopes and fears. World J. Gastroenterol. 2006;12:6594-6601.
9. Velstra B, van der Burgt YE, Mertens BJ, Mesker WE, Deelder AM, Tollenaar RA. Improved classification of breast cancer peptide and protein profiles by combining two serum workup procedures. J. Cancer Res. Clin. Oncol. 2012;138:1983-1992.
10. Callesen AK, Vach W, Jørgensen PE, Cold S+, Mogensen O, Kruse TA, Jensen ON, Madsen JS. Reproducibility of Mass Spectrometry Based Protein Profiles for Diagnosis of Breast Cancer across Clinical Studies: A Systematic Review. J. Proteome Res. 2008;7:1395-1402.
11. Beretta L. Proteomics from the clinical perspective: many hopes and much debate. Nat. Methods 2007;4:785-786.
12. Anderson NG. Adventures in Clinical Chemistry and Proteomics: A Personal Account. Clin. Chem. 2010;56:154-160.
13. Kinsinger CR, Apffel J, Baker M, Bian X, Borchers CH, Bradshaw R, Brusniak MY, Chan DW, Deutsch EW, Domon B, Gorman J, Grimm R, Hancock W, Hermjakob H, Horn D, Hunter C, Kolar P, Kraus HJ, Langen H, Linding R, Moritz RL, Omenn GS, Orlando R, Pandey A, Ping P, Rahbar A, Rivers R, Seymour SL, Simpson RJ, Slotta D, Smith RD, Stein SE, Tabb DL, Tagle D, Yates JR, Rodriguez H. Recommendations for mass spectrometry data quality metrics for open access data (corollary to the Amsterdam principles). Prot. Clin. Appl. 2011;5:580-589.
14. Callesen AK, Christensen Rd, Madsen JS, Vach W, Zapico E, Cold S, Jørgensen PE, Mogensen O, Kruse TA, Jensen ON. Reproducibility of serum protein profiling by systematic assessment using solid-phase extraction and matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun. Mass Spectrom. 2008;22:291-300.
15. Bladergroen MR, Derks RJE, Nicolardi S, de Visser B, van Berloo S, van der Burgt YEM, Deelder AM. Standardized and automated solid-phase extraction procedures for high-throughput proteomics of body fluids. J. Proteomics 2012;77:144-153.
16. Percy A, Chambers A, Yang J, Domanski D, Borchers C. Comparison of standard- and

nano-flow liquid chromatography platforms for MRM-based quantitation of putative plasma biomarker proteins. Anal. Bioanal. Chem. 2012;404:1089-1101.

17. Nicolardi S, Dalebout H, Bladergroen MR, Mesker WE, Tollenaar RAEM, Deelder AM, van der Burgt,YEM. Identification of Human Serum Peptides in Fourier Transform Ion Cyclotron Resonance Precision Profiles. Int. J. Proteomics 2012; 804036.
18. Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, Fleisher M, Lilja H, Brogi E, Boyd J, Sanchez-Carbayo M, Holland EC, Cordon-Cardo C, Scher HI, Tempst P. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J. Clin. Invest. 2006;116:271-284.
19. Velstra B, Bonsing BA, Mertens BJ, van der Burgt YEM, Huijbers A, Vasen H, Mesker WE, Deelder AM, Tollenaar RAEM. Detection of pancreatic cancer using serum protein profiling. HPB 2012;n/a.
20. Nicolardi S, Palmblad M, Dalebout H, Bladergroen M, Tollenaar RAEM, Deelder AM, van der Burgt YEM. Quality control based on isotopic distributions for high-throughput MALDI-TOF and MALDI-FTICR serum peptide profiling. J. Am. Soc. Mass Spectrom. 2010;21:1515-1525.
21. Marshall J, Jankowski A, Furesz S, Kireeva I, Barker L, Dombrovsky M, Zhu W, Jacks K, Ingratta L, Bruin J, Kristensen E, Zhang R, Stanton E, Takahashi M, Jackowski G. Human Serum Proteins Preseparated by Electrophoresis or Chromatography Followed by Tandem Mass Spectrometry. J. Proteome Res. 2004;3:364-382.
22. Hagman C, Ramström M, Håkansson P, Bergquist J. Quantitative Analysis of Tryptic Protein Mixtures Using Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. J. Proteome Res. 2004;3:587-594.
23. Hertkorn N, Ruecker C, Meringer M, Gugisch R, Frommberger M, Perdue EM, Witt M, Schmitt-Kopplin P. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. Anal. Bioanal. Chem. 2007;389:1311-1327.
24. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: A primer. Mass Spectrom. Rev. 1998;17:1-35.
25. Scigelova M, Hornshaw M, Giannakopulos A, Makarov A. Fourier Transform Mass Spectrometry. Mol. Cell. Proteomics 2011; 10.
26. Nicolardi S, Yuri E.M.van der Burgt, Wuhrer M, Deelder AM. Mapping O-glycosylation of apolipoprotein C-III in MALDI-FTICR protein profiles. Proteomics 2013; 13:992-1001
27. Nicolardi S, Palmblad M, Hensbergen PJ, Tollenaar RAEM, Deelder AM, van der Burgt YEM. Precision profiling and identification of human serum peptides using Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun. Mass Spectrom. 2011;25:3457-3463.
28. Hortin GL. The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. Clin. Chem. 2006;52:1223-1237.
29. Tiss A, Smith C, Menon U, Jacobs I, Timms JF, Cramer R. A well-characterised peak identification list of MALDI MS profile peaks for human blood serum. Proteomics 2010;10:3388-3392.
30. Villanueva J, Philip J, Entenberg D, Chaparro CA, Tanwar MK, Holland EC, Tempst P. Serum Peptide Profiling by Magnetic Particle-Assisted, Automated Sample Processing and MALDI-TOF Mass Spectrometry. Anal. Chem. 2004;76:1560-1570.
31. Villanueva J, Philip J, Chaparro CA, Li Y, Toledo-Crow R, DeNoyer L, Fleisher M, Robbins RJ, Tempst P. Correcting Common Errors in Identifying Cancer-Specific Serum Peptide Signatures. J. Proteome Res. 2005;4:1060-1072.

32. Velstra B, Burgt Y, Mertens B, Mesker W, Deelder AM, Tollenaar R. Improved classification of breast cancer peptide and protein profiles by combining two serum workup procedures. J. Cancer Res. Clin. Oncol. 2012;138:1983-1992.

33. Nikolaev EN, Jertz R, Grigoryev A, Baykut G. Fine Structure in Isotopic Peak Distributions Measured Using a Dynamically Harmonized Fourier Transform Ion Cyclotron Resonance Cell at 7 T. Anal. Chem. 2012;84:2275-2283.