

Identification of two new serum protein profiles for renal cell carcinoma

JUDITH Y.M.N. ENGWEGEN¹, NIVEN MEHRA², JOHN B.A.G. HAANEN³,
JAN H.M. SCHELLENS^{3,4}, EMILE E. VOEST² and JOS H. BEIJNEN^{1,4}

¹Department of Pharmacy & Pharmacology, The Netherlands Cancer Institute/Slotervaart Hospital, Louwesweg 6, 1066 EC Amsterdam; ²Department of Medical Oncology, University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX Utrecht; ³Department of Medical Oncology, The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam; ⁴Utrecht University, Faculty of Science, Department of Pharmaceutical Sciences, Division of Biomedical Analysis, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands

Received March 23, 2009; Accepted May 18, 2009

DOI: 10.3892/or_00000451

Abstract. Renal cell carcinoma (RCC) is the most lethal urological cancer, and survival greatly depends on early diagnosis. Therefore, reliable, new biomarkers for detection of RCC are required. We assessed serum protein profiles of samples from two institutes with SELDI-TOF MS in duplicate on CM10 chips at pH 6.0 (set 1: 37 RCC + 32 healthy controls (HC), set 2: 20 RCC + 25 HC). Mean peak intensities of detected proteins were compared between RCC and HC with non-parametric testing. Classification trees were built with discriminating peaks using one sample set as training set and the other as independent validation set. We found 15 peaks significantly different ($p < 0.01$) between RCC and HC. Two classification trees could be built with these peaks. Tree A achieved 75% sensitivity and 85% specificity for cross-validation and 76 and 65% for independent validation. Tree B had 71 and 62% sensitivity and specificity for cross-validation and 83 and 82% for independent validation. Although two serum protein profiles comprising 5 protein peaks were found that could separate RCC from HC, the sensitivity and specificity is not sufficient to recommend large scale use. Upon structural identification and quantitative validation, however, these proteins might prove suitable markers in the follow up of RCC patients.

Introduction

Renal cell carcinoma (RCC) is the most lethal of all urological cancers, with a 5-year survival rate of only about 10% (1). Conventional chemotherapy and immunotherapy are generally

considered ineffective for RCC. Recently approved drugs such as the multi-targeted kinase inhibitors sorafenib and sunitinib have significantly improved survival. However, these treatments can not cure advanced disease and are not void of side effects. Thus, early diagnosis and resection of localised disease remains the best option for treatment. Unfortunately, a lack of clear clinical symptoms prevents diagnosis at early stages and one third of patients will eventually relapse after previous nephrectomy (1,2). Although translational research has come up with several molecular markers for RCC (3,4), thus far, none of them has proven sufficiently effective. Therefore, reliable, new biomarkers that can be used for early diagnosis, prognosis and follow-up of RCC are still required.

The potential of the human genome and proteome to provide new biomarkers has been greatly acknowledged, as witness the number of publications in this field (5-11). Although gene expression profiling can yield new biomarkers as well (6), it relies on the acquisition of biopsies to provide cellular (tumour) material. Circulating proteins that are shed or proteolytically processed by tumour cells can be easily picked-up in accessible biological matrices such as urine, serum or plasma, making proteomic analysis especially suitable for the search for new biomarkers. The introduction of Surface Enhanced Laser Desorption Ionisation-Time of Flight Mass Spectrometry (SELDI-TOF MS) (12,13) has facilitated gross analysis of the proteome, thereby adding to the search for new and better tumour markers. The technology enables the analysis of large sample cohorts at the time, with a minimum of sample preparation. Thus far, it has been used to find new diagnostic biomarkers for e.g. ovarian (14), breast (15), prostate (16) and colorectal cancer (17).

Although tissue profiling with SELDI-TOF MS has been performed for RCC in a few studies (18-21), the search for new (non-invasive) RCC biomarkers has concentrated around profiling of urine and serum/plasma (22-26). Collection of urine is not invasive and the sample complexity is lower than that of serum or plasma due to the glomerular filtration of only smaller proteins in healthy persons. However, the proteome of urine samples may suffer from more variation (24,27) and protein profiling of urine likely requires even more stringent

Correspondence to: Dr J.Y.M.N. Engwegen, Atrium Medical Centre, Department of Pharmacy, P.O. Box 4446, 6401 CX Heerlen, The Netherlands
E-mail: JEN01@atriummc.nl

Key words: biomarkers, renal cell carcinoma, proteomics

sample collection procedures than that of blood samples, which may not always be achievable or controllable in clinical practice. In addition, albumin, which is not filtered into urine, can act as a carrier molecule for smaller proteins and protein fragments (28), and as such, profiling of urine will not take into account these potentially interesting, smaller proteins. Lastly, protein changes in urine are not easily interpretable, as they may result from altered production, altered glomerular filtration, or altered tubular re-uptake. However, previous RCC protein profiles obtained from serum have not all been validated, which is a prerequisite considering the potential problems with reproducibility or over-fitting of data (29-32). In fact, attempts by our group to validate some of these profiles with samples obtained from two other institutes had limited success. Only the increased expression of one protein was confirmed in our samples (33). Therefore, we assessed serum protein profiles for RCC with other assay conditions with which we have seen more discriminating peaks, in order to derive potentially better biomarker candidates for RCC.

Materials and methods

Chemicals. All used chemicals were obtained from Sigma, St. Louis, MO, USA, unless stated otherwise.

Patient samples. Two independent serum sample sets (Table I) were analysed at separate time-points in our laboratory using standardised analytical procedures. Set 1 consisted of samples from 37 patients with renal cell carcinoma and 32 healthy controls obtained at the Netherlands Cancer Institute, Amsterdam, The Netherlands. Set 2 consisted of samples from 20 patients with renal cell carcinoma and 25 healthy controls (HC) obtained at the University Medical Centre Utrecht, Utrecht, The Netherlands. All serum samples originated from a serum bank, where they had been collected according to institutional protocols. Sample collection was performed with individuals' informed consent after approval by the institutional review boards of both institutes.

Protein profiling. Protein profiling was performed using SELDI-TOF MS (Biorad Laboratories, Hercules, CA, USA). We screened Q10, CM10, H50 and IMAC-Cu chip surfaces with suitable binding conditions for discriminative mass-to-charge ratios (m/z) between patients with RCC and healthy controls. CM10 chips (weak cation exchange moiety) were ultimately selected with a binding buffer containing 50 mM sodium phosphate and 0.05% Triton X-100 at pH 6.0. In brief, samples were denatured by adding 10 μ l of 20 mM Tris-HCl pH 7.5, containing 9.5 M urea, 2% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and 2% dithiothreitol (DTT) to 5 μ l of serum. Chips were assembled in a 96-well bioprocessor (Biorad Laboratories). During all steps of the protocol, the bioprocessor was placed on a platform shaker at 350 rpm. Chips were equilibrated twice with 350 μ l of binding buffer for 5 min. Following, 45 μ l of binding buffer were added to 5 μ l of denatured serum. After vortexing and centrifuging, the supernatant was applied to the chip for a 45-min incubation. Subsequently, chips were washed twice for 5 min with 350 μ l binding buffer, followed by two 5-min washes with 350 μ l of 1:20 diluted binding buffer. Chips were

air-dried, after which 2 times 0.5 μ l of a 50% solution of sinapinic acid (SPA) in 50% acetonitrile (ACN) and 0.5% trifluoroacetic acid (TFA) was applied to the spots. Protein chips were analysed using the PBS-IIc ProteinChip Reader (Biorad Laboratories). Data were collected between 0 and 50,000 Da, optimising detection of discriminating peaks with an average of 65 laser shots per spectrum at laser intensity 147 and detector sensitivity 6.

To ensure reproducibility of results, all samples were reanalysed on a separate occasion more than one year later with the same assay conditions. Quality control samples, consisting of a pooled sample from HC for each sample set, were included in this analysis to assess intra-assay reproducibility. M/z values for the detected proteins were calibrated externally with a standard peptide mixture (Biorad Laboratories) containing [Arg8] vasopressin (1084.3 Da), somatostatin (1637.9 Da), dynorphin (2147.5 Da), ACTH (2933.5 Da), insulin β -chain (bovine) (3495.9 Da), insulin (human recombinant) (5807.7 Da), and hirudin recombinant (6963.5 Da). The complete analysis was replicated on a separate occasion to ensure robustness of detected expression differences.

Statistics and bioinformatics. Data were analysed with the ProteinChip Software package, version 3.1 (Biorad Laboratories). Per sample set, all acquired spectra were compiled, baseline subtracted and normalised to the total ion current from 1,500 to 50,000 Da. Spectra with normalisation factors higher than 2.00 or lower than 0.50 were excluded from further data analysis. We used the Biomarker Wizard (BMW) software application (Biorad Laboratories) to search for discriminative peaks. Peaks were autodetected when occurring in at least 5% of spectra and when having a signal-to-noise ratio (S/N) of at least 3. Peak clusters were completed with peaks with an S/N of at least 1.5 in a 0.3% cluster mass window. Group differences were calculated comparing mean intensities of all detected peaks between groups with non-parametric statistical tests. A $p < 0.01$ was considered statistically significant. Influence of patient characteristics on peak intensities were assessed similarly. Intra-assay reproducibility was evaluated by means of the peaks' coefficient of variation (CV) in replicate measurements of the quality control samples. Another estimate of the reproducibility was made calculating the CV of the peak intensity ratios for RCC vs. HC of discriminating peaks across the first and second analysis. Biomarker Patterns Software (BPS, Biorad Lab.) was used to generate classification trees from the BMW files. The method was set to 'gini', and the best tree within 1 standard error of the minimal cost tree was built. Performance was assessed by 10-fold cross validation or by usage of the remaining sample set as independent validation set.

Tentative protein identification. Tentative identities were based on protein mass and iso-electric point (pI) using the UniProtKB/Swiss-Prot and UniProtKB/TrEMBL database within the TagIdent option from www.expasy.org. Only identities of proteins known to be secreted in blood were considered. Protein pI was estimated by profiling on CM10 chips using identical procedures but binding buffers with successively higher pH, starting with the original binding



	Sample set 1		Sample set 2	
	RCC	HC	RCC	HC
Gender				
Female	14	12	5	19
Male	23	20	15	6
Mean age (years \pm SD)	56.7 \pm 11.1	56.0 \pm 8.8	63.9 \pm 11.2	43.6 \pm 10.7 ^a
RCC type				
Clear cell carcinoma	25		12	
Non-clear cell carcinoma	4		2	
Mixed papillary/clear cell	1		1	
No PA histology done	7		5	
AJCC stage				
1	3		0	
2	10		0	
3	4		0	
4	20		20	
Surgical status				
Pre-surgery/no surgery	33		7	
Surgery	4		13	
Sample handling	30 min coagulation at RT. Centrifugation at 1900 x g. Storage at -30°C		2-6 h coagulation at RT and overnight at 4°C. Centrifugation at 1500 x g. Storage at -80°C	
Sample storage duration (years \pm SD)	2.99 \pm 0.55	1.90 \pm 0.03 ^b	1.07 \pm 0.55	1.35 \pm 0.2

^ap<0.001 non-parametric Mann-Whitney U-test, for RCC vs. HC in set 2; ^bp<0.001 independent samples t-test, for RCC v. HC in set 1. RT, room temperature.

condition, pH 6.0. Furthermore, protein profiles further acquired under similar conditions as the original ones where compared between samples reduced with DTT or not reduced and for spectra generated with SPA or CHCA (α -cyano-4-hydroxycinnamic acid) as energy absorbing matrix.

Results

Patient samples. Patient and sample characteristics are summarised in Table I. The first sample set mainly consisted of samples from newly diagnosed patients prior to nephrectomy, whereas the second sample set comprised mostly relapsed patients with stage-IV disease. Also, sample handling conditions were different for set 1 and 2. However, the differences between set 1 and 2 enabled the selection of protein expression changes robust to these differences.

Protein profiling. Comparing the results from a pilot screening on Q10, CM10, H50 and IMAC-Cu chip surfaces, we found the most peaks and best discriminating peaks on CM10 chips at pH 6.0. Therefore, subsequent analyses were performed with these conditions. Only one spectrum (from set 2) was not assessable due to an aberrant normalisation

factor. Peaks below 2000 Da were discarded, as in this region there was interference of the SPA matrix. To ensure that observed expression differences were reproducible, all samples were reanalysed on a separate occasion more than one year later, after storage of samples at -30°C, one additional freeze-thaw cycle and maintenance of the PBS-IIc mass spectrometer, including renewal of the laser. We only considered peaks that proved detectable and significantly different in the combined sample sets both in the first and second analysis for further evaluation (Table II). M/z 2743, 3774, 3891, 3920, 4157, 4282, 4677, 7623, 7981, 8758, 10838, 13149, 15113, 15940 and 16120 met these criteria. The mean peak intensities in patients and controls for these peaks were quite similar across the two sample sets, however, in HC, several significant differences between set 1 and 2 for peaks at m/z 3774, 3891, 4157, 4282 and 15113 were observed, as well as for m/z 4677 in RCC (Fig. 1A; p<0.01). In the second analysis, a similar discrepancy between the two sample sets in peak intensities for m/z 3744, 3891 and 7620 was apparent in HC (Fig. 1B; p<0.001) and for m/z 4282 in RCC. However, for most of these peaks the ratio of mean peak intensities for patients vs. controls was roughly the same in both sample sets (Table II). Intra-assay reproducibility of peak intensities calculated as

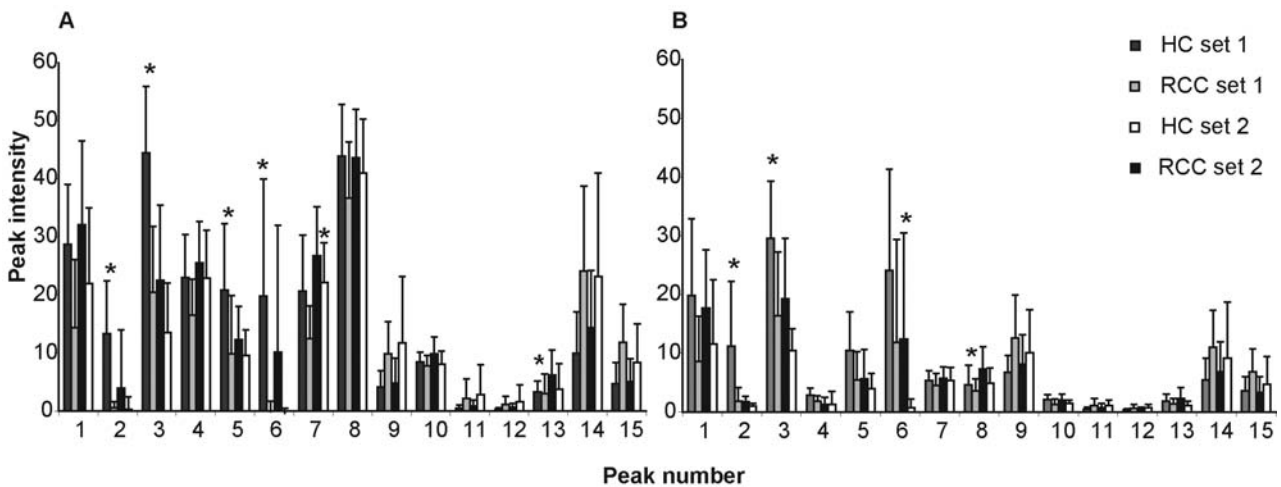


Figure 1. Mean peak intensities and SD in patients and controls from sample set 1 and 2. The results for the first and second analysis are shown in part A and B, respectively. Several peaks show large differences between the sample sets of each institute (significant differences are marked with *). Numbers correspond to the peaks in Table II.

Table II. Discriminative peaks with reproducible expression differences.

Peak number	M/z (Da)	First analysis		Second analysis		p-value in set 1+2
		Expression ratio RCC vs. HC set 1	Expression ratio RCC vs. HC set 2	Expression ratio RCC vs. HC set 1	Expression ratio RCC vs. HC set 2	
1	2743	0.50	0.68	0.43	0.65	2.61×10^{-7}
2	3774	0.047	0.082	0.17	0.58	2.01×10^{-6}
3	3891	0.46	0.60	0.56	0.54	7.61×10^{-8}
4	3920	0.72	0.90	0.65	1.06	1.47×10^{-4}
5	4157	0.47	0.78	0.52	0.71	3.02×10^{-7}
6	4282	0.03	0.004	0.49	0.06	1.48×10^{-8}
7	4677	0.60	0.83	0.83	0.94	3.93×10^{-5}
8	7623	0.84	0.94	0.89	0.67	2.79×10^{-3}
9	7981	2.38	2.45	1.88	1.25	3.73×10^{-8}
10	8758	0.92	0.82	0.62	0.74	4.34×10^{-3}
11	10838	4.49	3.13	2.41	1.46	4.10×10^{-8}
12	13149	3.57	2.39	1.78	1.57	1.901×10^{-6}
13	15113	0.91	0.60	0.75	0.46	2.45×10^{-3}
14	15940	2.43	1.62	2.00	1.35	2.86×10^{-6}
15	16120	2.53	1.63	1.90	1.40	2.18×10^{-6}

the median CV of all clustered peaks in the quality control samples was 43%. The reproducibility of the peak intensity ratios (RCC/HC) of discriminating peaks across the first and second analysis was 31% for sample set 1 and 26% for set 2. For m/z 3774, 4282, 7981, 8758, 10838 and 13149 intensity ratios differed between the first and second analysis, generally to the same extent for sample set 1 and 2.

The significantly different peaks were used to generate classification trees. Two classification trees were obtained having a suitable sensitivity and specificity for both sample sets (Fig. 2). For tree A cross-validation sensitivity and specificity were 75 and 85% and sensitivity and specificity for independent validation with set 2 respectively 76 and 65%. For tree B cross-validation sensitivity and specificity were 71

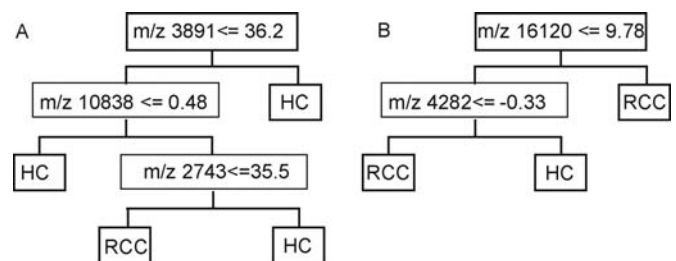


Figure 2. Classification trees. Tree A was generated with set 1 as training and set 2 as validation set. Tree B was generated with set 2 as training and set 1 as validation set.

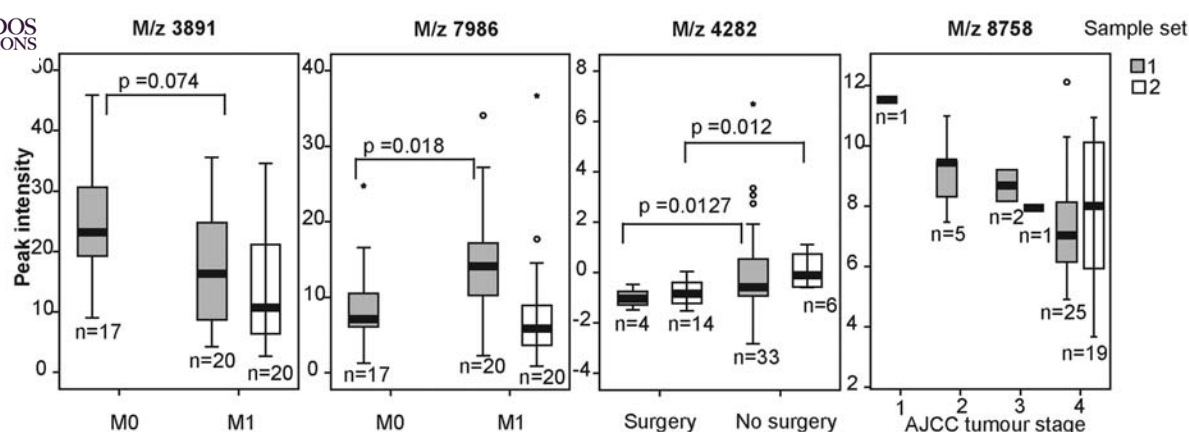


Figure 3. Influence of patient characteristics on peak intensities in RCC sera from both sample sets. (M0, no distant metastases; M1, distant metastases).

Table III. Tentative protein identities.

Observed m/z (Da)	Observed pI	Tentative identity	Theoretical mass (Da)	Theoretical pI
2743	>9	Fetuin A/ α -2HS glycoprotein chain B	2740	10.86
3891	>9	Islet amyloid polypeptide	3906	8.90
4282	>9	Inter- α trypsin inhibitor heavy chain 4 (ITIH4) fragment (34;35): QAGAAGSRMNFPRPGLVSSRQLGLPGPPDVPDHAAYHPFRR	4285	
10838	>9	Ig heavy chain V-I region V35	10880	9.59
		Ig heavy chain V-I region HG3	10803	9.23
		Ig kappa chain V-II region MIL	10825	9.39
16120	6-7	Haemoglobin- β chain SPA adduct	15867+206=16073	6-7

pI, iso-electric point; SPA, sinapinic acid.

and 62% and for independent validation with set 1 respectively 83 and 82%. All RCC patients misclassified by tree A and B had stage-IV clear cell or non-clear cell RCC. Stage I-III patients were all correctly classified. Other classification trees that were built were only suitable for classification of one of the two sample sets. M/z 3774 was a good classifier in set 1 (sensitivity 91.7 and specificity 78.8%). Validating this tree with sample set 2 yielded similar sensitivity (96.2%), but unacceptable specificity (15.4%). A similar result was seen for a classification tree with m/z 3891 as a single classifier. For sample set 1 sensitivity and specificity were 80-90%, whereas the specificity dropped to 15% for sample set 2.

Influence of patient characteristics. We assessed whether differences in patient or sample characteristics between institutes influenced the observed peak intensities of the significantly different peaks (Fig. 3). In set 1 a statistically non-significant increase in peak intensity was seen in patients without distant metastases for m/z 3891 ($p=0.074$) and a statistically non-significant decrease for m/z 7986 ($p=0.018$), but also in these subgroups these peaks were still lower, respectively higher than in controls. We saw slightly lower peak intensities in RCC sera drawn after nephrectomy for m/z 4282 ($p=0.127$ and 0.012 for sample set 1 and 2).

However, all RCC peak intensities for m/z 4282, including those in patients before/without nephrectomy were around zero and well below those of healthy controls. As pre-surgery patients mostly had loco-regional disease (Pearson correlation $r=0.278$, $p=0.036$), we also assessed whether m/z 4282 was different for metastasised and loco-regional disease. No difference was observed, however ($p=0.619$; set 1). M/z 8758 was associated with tumour stage in sample set 1 ($p=0.027$; Kruskal-Wallis test), showing lower peak intensities with progressive stage. Peak intensities were similar for men and women and did not correlate with the age of the patient or the storage duration of samples.

Tentative protein identification. Attempts to purify and collect biomarker candidates for identification by peptide mapping or MS/MS analysis were unsuccessful. M/z 16120 could not be readily separated from concomitant 15-kDa peaks and low abundance as well as mass limitations for direct MS/MS analysis hampered structural identification of the other peaks. Therefore, tentative identification was done by indirect methods. Protein pIs were estimated for each of the biomarker peaks in the classification trees. Two samples with either high abundance or low abundance for m/z 2743, 3891, 4282, 10838 and 16120 were profiled on CM10 chips

using binding buffers with successively higher pH, starting with the original binding condition, pH 6.0. All peaks except m/z 16120 bound to the CM10 chips up to pH 9.0, indicating a pI >9. M/z 16120 was present when profiled at pH 6.0, but did not bind at higher pH, indicating a pI between 6 and 7. When no DTT was used for denaturation, m/z 3891 and 4288 remained present in the spectra, whereas m/z 2743, 16120 and concomitant 15-kDa peaks were not, implying that the latter peaks are generated by decomposition of a protein multimer. In addition, a change in mass of 57 Da was seen when CHCA was used as matrix instead of SPA (m/z 16120→16066), indicating that the peak was an adduct of SPA (mass 206 Da) and protein. Using the TagIdent option from www.expasy.org, we postulated haemoglobin- β chain as the identity for this peak (m/z 15867+206=16073).

As m/z 2743 was generated by reduction with DTT, this protein is likely to result from a multimer. Fetuin A (α -2HS glycoprotein chain B) has a theoretical mass of 2740 Da and pI of 10.86. Furthermore, it is combined by a disulphide bond with fetuin B. Therefore, we propose this identity for m/z 2743. For m/z 4282 no database protein complied with our observed m/z and pI. In the literature a similar protein mass has been assigned to a fragment of inter- α trypsin inhibitor heavy chain 4 (34,35). The theoretical pI of this fragment, 11.42 as estimated with the 'compute pI/molecular weight' tool from www.expasy.org, corresponded to our observed pI. The observed mass and pI of m/z 3891 correspond to that of islet amyloid polypeptide (amylin), but it cannot be excluded that this peak is a fragment of a larger protein. The only proteins with masses around 10838 Da and a pI >9 in the Swiss-Prot and TrEMBL databases, were immunoglobulin chains (Table III).

Discussion

In the current study we analysed sera of patients with renal cell carcinoma in search of potential biomarker proteins discriminating them from healthy controls. Two sample sets from different institutes with different patient characteristics were used. Protein m/z 2743, 3891, 4282, 10838 and 16120 were shown to be useful in classifying RCC and HC in classification trees with a sensitivity and specificity ranging from 62 to 85% (both independent validation and cross-validation). The intra-assay CV of our peak intensities, including the discriminating peaks, was higher than what is generally reported (~20-30%) (36,37). This might be caused by the manual instead of automated processing of the samples. However, we chose to report the CV of all detected peak clusters, and not only of a subset of peaks. The latter is commonly done and could account for a lower CV if only peaks with high intensities are selected. Despite the variability in absolute peak intensities, we found several peaks with reproducible expression differences in both sample sets upon duplicate analysis and were able to validate the classification from each tree with the remaining sample set. For some peaks the intensity ratios in sample set 1 and 2 were two- to ten-fold (e.g. for m/z 4282) different in the second analysis compared to the first, indicating an effect of the extra freeze-thaw cycle, longer storage time, or technical variation in the second analysis.

All patients with stage-I to -III RCC, although the minority, were classified correctly by a classification tree generated from stage-IV patients. This indicates that at least the peaks from tree B, m/z 16120 and 4282, can be useful for detection of early-stage RCC. As the two populations investigated were quite different according to surgical status, the proteins in these classification trees with suitable sensitivity and specificity for both sample sets are likely not influenced by previous surgery of the primary tumour and reflect a general presence of tumour load. In patients with distant metastases we found different peak intensities of m/z 3891 and 7986, yet evaluation of their potential prognostic value by Cox regression analysis did not show a significant association with survival (data not shown). This was also the case for m/z 8758, which was related to tumour stage. For m/z 10838 we did see such an association ($p < 0.001$; data not shown), indicating a potential prognostic value, independent of the presence of distant metastasis.

Although we saw reproducible expression differences between RCC and HC for m/z 3774 and 3891 in both sample sets and in the duplicate analyses, specificity of classification trees with these classifiers was limited when using the remaining sample set as independent validation set. This was caused by the discrepancies in mean intensities of these peaks in HC from the two samples sets. A suitable cut-off value for both sets could therefore not be established. It is unlikely that these differences are the result of discrepancies in sample handling between the institutes, between-day technical variability, or differences in patient characteristics, since mean peak intensities were generally only different for the HC samples. Possibly, the healthy individuals differ in their characteristics, although we did not find a correlation of peak intensities with any of the known characteristics such as age. Protein identification and absolute, instead of relative, quantitation of these peaks in larger populations could define real, quantitative cut-offs for discrimination of patients and controls. Then, m/z 3774 and 3891 might also prove to be suitable biomarkers for RCC.

Our observed expression differences concord with previous results in the literature. Changes in the abundance of ITIH4 fragments in cancer have been a common result of many SELDI-TOF MS serum profiling studies (34,35,38,39). In a quantitative SELDI-TOF MS immunoassay m/z 4283 was increased in breast cancer patients and decreased in patients with pancreatic cancer (35). The combined expression profile of all ITIH4 fragments together differed between the four tested cancer types. Also in a previous RCC serum protein profile validation study we found a mass of 4289 Da decreased in RCC in the same populations with other assay procedures (33). Since this specific ITIH4 fragment is aberrantly expressed in several cancers, its specificity for RCC will depend on its combined analysis with other RCC biomarkers and/or other ITIH4 fragments.

With slightly different assay procedures Tolson *et al* showed an increase of an unknown protein at 10.84 kDa in RCC (23). The peak at 10838 Da in the current study could represent the same protein. In the Tolson study a positive relationship of their 10.84-kDa peak with progressive disease stage was suggested (23). Indeed, we saw a negative correlation of m/z 10838 with survival in a univariate Cox regression

SPANDIDOS data not shown). Whether this prognostic value will earlier disease stages has to be investigated further.

We propose that this peak represents an immunoglobulin chain, although no previous studies on the increase of these specific chains in RCC have been made. Alternatively, this mass may represent a fragment of a different protein.

The decrease of a peak at 3891 Da in RCC was also found in our previous validation study (33). A similar peak (m/z 3885) was reported to differ between pancreatic cancer patients and healthy controls, but it was not mentioned whether this was an increase or decrease (40). We suggest this protein to be islet amyloid polypeptide (amylin), which is synthesised by the beta cells of the pancreatic islets and subsequently co-secreted with insulin. This protein has been reported to be increased in pancreatic cancer patients with diabetes, but was not sensitive enough for detection of pancreatic cancer (41). Its levels are also increased in end-stage renal disease due to decreased renal clearance (42). Others have reported a decrease of islet amyloid polypeptide gene expression in patients with neuroendocrine tumours due to multiple endocrine neoplasia type 1 (43). However, no previous studies of any decreased expression in RCC or benign kidney disease have been made to our knowledge.

M/z 2743 has been previously found to be decreased in serum of patients with renal failure undergoing haemodialysis (44). We propose fetuin A as this protein's identity. It is a protease inhibitor produced by the liver and secreted into serum in high concentrations (0.5-1.0 g/l). By forming complexes with calcium and phosphate, it enhances their solubility and prevents vascular calcification. Fetuin A is known to be present in lower amounts in chronic kidney disease (45,46) and is decreased in the acute phase response (47). Other studies describe an increase of this protein in acute renal injury, however (48).

The identity of m/z 16120 is likely an SPA adduct of the β -chain of haemoglobin. This same protein chain has also been found elevated in ovarian cancer patients compared to controls (49). In contrast to haemoglobin- β levels, haemoglobin- α levels (although not formally identified, the 15.1-kDa peak most likely represents this protein) were decreased in RCC. Low (total) haemoglobin levels are related to poor prognosis in RCC (50). The presence of free haemoglobin in serum suggests the haemolysis of erythrocytes, which could have occurred during serum collection. No overt haemolysis was visible in these samples, however. As the appearance of these haemoglobin chains in our mass spectra was the result of reduction of serum with DTT, they should reflect the abundance of the haemoglobin tetramer (of two α - and two β -chains) released into serum. It is therefore surprising that haemoglobin- α and - β show opposite regulation in RCC. We currently do not have a satisfactory explanation for this observation.

In a similar study to ours, Hara *et al* found two peaks of m/z 4151 and 8968 with SELDI-TOF MS serum protein profiling, which were increased in 21, mostly early-stage, RCC compared to 24 HC. The combined peaks had a sensitivity and specificity of 80-90% for detection of RCC upon validation in an independent population (29 RCC + 20 HC) (25). One of their masses is similar to our m/z 4157, but the expression difference opposite to our current results.

However, as they performed anion exchange fractionation on serum, this peak may well represent a different protein than ours. We saw a small peak with mass 8968 Da incompletely separated from a larger peak at 8933 Da in our spectra, which was also increased in RCC when manually selected before clustering (data not shown). The fact that we did not perform serum fractionation has likely caused this protein to be overshadowed by the 8933-Da one.

Concluding, we found two protein profiles that were able to reproducibly discriminate patients with RCC from healthy controls in two different RCC populations, each used as an independent validation set for the other population. The specificity of these profiles for RCC should be investigated further, comparing RCC profiles with those in benign kidney disease as well as other cancers. Although the current sensitivity and specificity of these profiles are insufficient for use in diagnostic screening, they might have a role in follow-up of RCC patients. Elucidation of the identities of the proteins in these profiles and subsequent absolute quantitation is needed before they can be further evaluated as biomarkers in large-scale clinical trials.

References

1. Patel PH, Chaganti RS and Motzer RJ: Targeted therapy for metastatic renal cell carcinoma. *Br J Cancer* 94: 614-619, 2006.
2. Cohen HT and McGovern FJ: Renal-cell carcinoma. *N Engl J Med* 353: 2477-2490, 2005.
3. Kashyap MK, Kumar A, Emelianenko N, *et al*: Biochemical and molecular markers in renal cell carcinoma: an update and future prospects. *Biomarkers* 10: 258-294, 2005.
4. Thongboonkerd V: Proteomic analysis of renal diseases: unraveling the pathophysiology and biomarker discovery. *Exp Rev Proteomics* 2: 349-366, 2005.
5. Kim HL, Seligson D, Liu X, *et al*: Using protein expressions to predict survival in clear cell renal carcinoma. *Clin Cancer Res* 10: 5464-5471, 2004.
6. Zhao H, Ljungberg B, Grankvist K, Rasmuson T, Tibshirani R and Brooks JD: Gene expression profiling predicts survival in conventional renal cell carcinoma. *PLoS Med* 3: E13, 2006.
7. Lovisolo JA, Casati B, Clerici L, *et al*: Gene expression profiling of renal cell carcinoma: a DNA microarray analysis. *BJU Int* 98: 205-216, 2006.
8. Jacquemier J, Ginestier C, Rougemont J, *et al*: Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res* 65: 767-779, 2005.
9. Van 't Veer LJ, Dai H, van de Vijver MJ, *et al*: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415: 530-536, 2002.
10. Barrier A, Boelle PY, Roser F, *et al*: Stage II colon cancer prognosis prediction by tumor gene expression profiling. *J Clin Oncol* 24: 4685-4691, 2006.
11. Oberthuer A, Berthold F, Warnat P, *et al*: Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. *J Clin Oncol* 24: 5070-5078, 2006.
12. Fung ET, Thulasiraman V, Weinberger SR and Dalmasso EA: Protein biochips for differential profiling. *Curr Opin Biotechnol* 12: 65-69, 2001.
13. Issaq HJ, Veenstra TD, Conrads TP and Felschow D: The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem Biophys Res Commun* 292: 587-592, 2002.
14. Zhang Z, Bast RC Jr, Yu Y, *et al*: Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 64: 5882-5890, 2004.
15. Li J, Orlandi R, White CN, *et al*: Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry. *Clin Chem* 51: 2229-2235, 2005.
16. Adam BL, Qu Y, Davis JW, *et al*: Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res* 62: 3609-3614, 2002.

17. Ward DG, Suggett N, Cheng Y, *et al.*: Identification of serum biomarkers for colon cancer by proteomic analysis. *Br J Cancer* 94: 1898-1905, 2006.
18. Von Eggeling F, Junker K, Fiedle W, *et al.*: Mass spectrometry meets chip technology: a new proteomic tool in cancer research? *Electrophoresis* 22: 2898-2902, 2001.
19. Fetsch PA, Simone NL, Bryant-Greenwood PK, *et al.*: Proteomic evaluation of archival cytologic material using SELDI affinity mass spectrometry: potential for diagnostic applications. *Am J Clin Pathol* 118: 870-876, 2002.
20. Junker K, Gneist J, Melle C, *et al.*: Identification of protein pattern in kidney cancer using ProteinChip arrays and bioinformatics. *Int J Mol Med* 15: 285-290, 2005.
21. Holcakova J, Hernychova L, Bouchal P, *et al.*: Identification of alphaB-crystallin, a biomarker of renal cell carcinoma by SELDI-TOF MS. *Int J Biol Markers* 23: 48-53, 2008.
22. Won Y, Song HJ, Kang TW, Kim JJ, Han BD and Lee SW: Pattern analysis of serum proteome distinguishes renal cell carcinoma from other urologic diseases and healthy persons. *Proteomics* 3: 2310-2316, 2003.
23. Tolson J, Bogumil R, Brunst E, *et al.*: Serum protein profiling by SELDI mass spectrometry: detection of multiple variants of serum amyloid alpha in renal cancer patients. *Lab Invest* 84: 845-856, 2004.
24. Rogers MA, Clarke P, Noble J, *et al.*: Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility. *Cancer Res* 63: 6971-6983, 2003.
25. Hara T, Honda K, Ono M, Naito K, Hirohashi S and Yamada T: Identification of 2 serum biomarkers of renal cell carcinoma by surface enhanced laser desorption/ionization mass spectrometry. *J Urol* 174: 1213-1217, 2005.
26. Wu DL, Zhang WH, Wang WJ, Jing SB and Xu YM: Proteomic evaluation of urine from renal cell carcinoma using SELDI-TOF-MS and tree analysis pattern. *Technol Cancer Res Treat* 7: 155-160, 2008.
27. Schaub S, Wilkins J, Weiler T, Sangster K, Rush D and Nickerson P: Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Int* 65: 323-332, 2004.
28. Mehta AI, Ross S, Lowenthal MS, *et al.*: Biomarker amplification by serum carrier protein binding. *Dis Markers* 19: 1-10, 2003.
29. Baggerly KA, Morris JS and Coombes KR: Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. *Bioinformatics* 20: 777-785, 2004.
30. Diamandis EP and van der Merwe DE: Plasma protein profiling by mass spectrometry for cancer diagnosis: Opportunities and limitations. *Clin Cancer Res* 11: 963-965, 2005.
31. Baggerly KA, Morris JS, Edmonson SR and Coombes KR: Signal in noise: Evaluating reported reproducibility of serum proteomic tests for ovarian cancer. *J Natl Cancer Inst* 97: 307-309, 2005.
32. Liotta LA, Lowenthal M, Mehta A, *et al.*: Importance of communication between producers and consumers of publicly available experimental data. *J Natl Cancer Inst* 97: 310-314, 2005.
33. Engwegen JYMN, Mehra N, Haanen JBAG, *et al.*: Validation of SELDI-TOF MS serum protein profiles for renal cell carcinoma in new populations. *Lab Invest* 87: 161-172, 2007.
34. Koomen JM, Shih LN, Coombes KR, *et al.*: Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. *Clin Cancer Res* 11: 1110-1118, 2005.
35. Song J, Patel M, Rosenzweig CN, *et al.*: Quantification of fragments of human serum inter- α -trypsin inhibitor heavy chain 4 by a surface-enhanced laser desorption/ionization-based immunoassay. *Clin Chem* 52: 1045-1053, 2006.
36. Albrethsen J, Bogebo R, Olsen J, Raskov H and Gammeltoft S: Preanalytical and analytical variation of surface-enhanced laser desorption-ionization time-of-flight mass spectrometry of human serum. *Clin Chem Lab Med* 44: 1243-1252, 2006.
37. Rai AJ, Stemmer PM, Zhang Z, *et al.*: Analysis of human proteome organization plasma proteome project (HUPO PPP) reference specimens using surface enhanced laser desorption/ionization-time of flight (SELDI-TOF) mass spectrometry: multi-institution correlation of spectra and identification of biomarkers. *Proteomics* 5: 3467-3474, 2005.
38. Banez LL, Prasanna P, Sun L, *et al.*: Diagnostic potential of serum proteomic patterns in prostate cancer. *J Urol* 170: 442-446, 2003.
39. Fung ET, Yip TT, Lomas L, *et al.*: Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. *Int J Cancer* 115: 783-789, 2005.
40. Koopmann J, Zhang Z, White N, *et al.*: Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin Cancer Res* 10: 860-868, 2004.
41. Chari ST, Klee GG, Miller LJ, Raimondo M and DiMagno EP: Islet amyloid polypeptide is not a satisfactory marker for detecting pancreatic cancer. *Gastroenterology* 121: 640-645, 2001.
42. Leckstrom A, Bjorklund K, Permert J, Larsson R and Westermark P: Renal elimination of islet amyloid polypeptide. *Biochem Biophys Res Commun* 239: 265-268, 1997.
43. Dilley WG, Kalyanaraman S, Verma S, Cobb JP, Laramie JM and Lairmore TC: Global gene expression in neuroendocrine tumors from patients with the MEN1 syndrome. *Mol Cancer* 4: 9, 2005.
44. Langlois RG, Trebes JE, Dalmaso EA, *et al.*: Serum protein profile alterations in hemodialysis patients. *Am J Nephrol* 24: 268-274, 2004.
45. Cottone S, Nardi E, Mule G, *et al.*: Association between biomarkers of inflammation and left ventricular hypertrophy in moderate chronic kidney disease. *Clin Nephrol* 67: 209-216, 2007.
46. Mehrotra R: Disordered mineral metabolism and vascular calcification in nondialyzed chronic kidney disease patients. *J Ren Nutr* 16: 100-118, 2006.
47. Gabay C and Kushner I: Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med* 340: 448-454, 1999.
48. Janech MG, Raymond JR and Arthur JM: Proteomics in renal research. *Am J Physiol Renal Physiol* 292: F501-F512, 2007.
49. Kozak KR, Su F, Whitelegge JP, Faulk K, Reddy S and Farias-Eisner R: Characterization of serum biomarkers for detection of early stage ovarian cancer. *Proteomics* 5: 4589-4596, 2005.
50. Mekhail TM, Abou-Jawde RM, Boumerhi G, *et al.*: Validation and extension of the memorial sloan-kettering prognostic factors model for survival in patients with previously untreated metastatic renal cell carcinoma. *J Clin Oncol* 23: 832-841, 2005.