ARTICLE

A Well-Characterised Peak Identification List of MALDI MS Profile Peaks for Human Blood Serum

Ali Tiss[1], Celia Smith[1], Usha Menon[2], Ian Jacobs[2], John F. Timms[2] and Rainer Cramer[1]*
[1] The BioCentre and Department of Chemistry, University of Reading, Reading, UK
[2] Institute for Women's Health, UCL, London, UK

[*] Corresponding Author: Prof Rainer Cramer, The BioCentre and Department of Chemistry, University of Reading, Harborne Building, Whiteknights, PO Box 221, Reading, RG6 6AS, UK; Telephone.: +44-118-378-4550; Fax: +44-118-378-4551; E-mail: r.k.cramer@rdg.ac.uk.

Abstract
MALDI MS profiling has attracted considerable interest from the clinical community as a potentially valuable alternative to other diagnostic and prognostic tools by analysing biomarkers from easily available body fluids such as blood serum. Despite the numerous reports on MALDI MS profiling of human serum there is only scarce information on the identity of the biomolecular species making up these profiles, particularly in the mass range of larger peptides. Here we provide a methodology for MALDI MS profiling peak identification and an extensive list of peak identities up to 10 kDa obtained from human blood serum. The list of identities obtained so far comprises more than 90 entries. In addition to peptides published in other lists, we have also identified high-molecular weight peptides with masses above 5 kDa. Some of the identifications were related to endogenous lipids. Amongst the peptide identifications various modifications such as phosphorylation were detected. The overlap with the few other peptide lists published so far was found to be limited and hence our list significantly extends the number of identified peptides commonly found in MALDI MS profiling of human blood serum. With the data from this study it is further shown that lists of endogenous serum peptides can only be useful for MALDI MS profiling if the peak identifications can be directly linked to the MALDI MS profiles. A list of identifications of MALDI MS profile peaks from human blood serum, compiled from published data and this work, is presented.

Keywords: MALDI, mass spectrometry, MS peak identification, serum peptidome, serum profiling

Nonstandard abbreviations: UKCTOCS, United Kingdom Collaborative Trial of Ovarian Cancer Screening; UKOPS, United Kingdom Ovarian Cancer Population Study.

This is a non-final version of an article published in final form in:
Tiss A, Smith C, Menon U, Jacobs I, Timms JF and Cramer R. A well-characterised peak identification list of MALDI MS profile peaks for human blood serum. Proteomics 2010 10(18): 3388-3392

Introduction

There has been great interest in exploiting simple MALDI MS profiling for the analysis of crude or little purified human serum or plasma samples for clinical diagnostics [1-3]. However, through the many studies that have been conducted in this area it has become apparent that

there are far more challenges related to its applicability to disease diagnosis than initially thought. The technical bias in the analytical read-out and the ultimate sensitivity and diagnostic accuracy of MALDI MS profiling are of major concern [4-7]. The wide range of abundances and the complexity of the endogenous molecular components in blood further aggravate the analytical situation [2]. However, there have been now further reports substantiating the claim that MALDI MS profiling can indeed be a tool for disease (outcome) classification[8] and biomarker discovery[9].

Thus, it is surprising that little is known about most of these potential biomarkers as typically no structural identification is undertaken. This presents a serious limitation for alternative quantification and validation of potentially discriminatory peaks and restricts biomarker assay development, including immunoassays, that might potentially overcome some of the above analytical limitations of MALDI MS profiling.

MALDI MS profiling focuses on the reproducible, rapid and inexpensive acquisition of MS peak patterns that can be used for discriminative analysis as needed in population screening or clinical triage. Thus, it is different to in-depth analyses of the human blood serum or plasma where peak identification is an objective and many thousands of peptides and proteins have been identified [10-12].

As with in-depth analyses, peptide identification of MS profile peaks is ideally undertaken at the same time as MS profiling with the same sample on the same instrument, using the MS profile for precursor ion determination and directly switching to MS/MS mode for peptide sequencing. Employing MALDI, however, peptide sequencing is more challenging and usually less sensitive and informative for identification purposes than ESI-LC-MS/MS. Thus, the reported long lists of identified peptides or proteins present in serum or plasma are typically based on ESI-LC-MS/MS (mostly in combination with extensive pre-fractionation) and without any relation to MALDI MS profiles, including the lists that provide identifications in the commonly recorded low molecular weight range of these profiles, i.e. <10 kDa [13-15]. Indeed, only a few identifications of MALDI MS profile peaks have been reported so far. The most comprehensive list of peptide identifications obtained from human blood serum and related to MALDI MS profile peaks was reported by Villanueva et al. [16], although it provides only identification of peaks below 4 kDa. For blood plasma, Koomen et al. reported a list of 249 sequenced plasma profile peptides with a maximum mass of 5.2 kDa [17].

Materials and methods

Venous blood samples were collected from healthy persons recruited at different centres as part of the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) or United Kingdom Ovarian Cancer Population Study (UKOPS) (for details see Timms et al., Clin. Chem., in press). The entire study was approved by the local ethics committee and written informed consent was obtained from all donors.

For UKOPS, blood was collected in 10mL-Vacutainer ('red-top') tubes (Becton Dickinson, Oxford, UK; #367985). After phlebotomy, the tubes were inverted 5 times and left at room temperature in a vertical position to clot for 1 hr. Samples were then placed on ice for no more than 2 hrs, centrifuged at 1,500 g for 10 min at room temperature, aliquoted into sterile 10mL-tubes (Sarstedt, Leicester, UK) and frozen at -80°C. For UKCTOCS, different consumables were used and the clotting time was between 8 and 56 hours before spinning, strawing and freezing. All samples were transported to the Institute for Women's Health Laboratory at UCL on dry ice and thawed at room temperature for 30 min. The samples were then mixed by gentle inversion and aliquoted into bar-coded straws, re-frozen at -80°C for 24 hours and then transferred to liquid N_2 for long-term storage.

For initial MALDI-TOF MS profiling, serum samples were taken from long-term storage and prepared according to a protocol we have reported elsewhere (see [5] for details). Briefly,

polypeptides were enriched from 5 µL of serum using a semi-automated protocol based on reversed phase pre-packed tips (C18 ZipTips®). A CyBi®-Disk robot (CyBio AG, Jena, Germany) equipped with a 96-piston head for 25µL-tips was adapted and used for this purpose. After C18 ZipTip® purification, enriched polypeptides were eluted from the ZipTips®, and 2 µL from the eluate were mixed with CHCA matrix and spotted onto a 600µm-AnchorChip target plate (Bruker Daltonics, Bremen, Germany) for analysis on an Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics), using FlexAnalysis v3.0 and ClinProTools v2.1 software (Bruker Daltonics) for further data analysis. The remaining 5 µL of the eluate were immediately stored at -80°C until being used for further analysis.

Peak identifications were obtained from the same sample eluates used for MALDI MS profiling, by pooling 96 single ZipTip® eluates and using three different methods: (i) MALDI MS/MS, (ii) ESI MS/MS and (iii) off-line LC fractionation and ESI MS/MS with or without tryptic digestion. Additional MALDI MS profiles were taken throughout the sample preparation after each essential step in order to link the MALDI MS profile peaks with the obtained peak identification.

For peak identification using MALDI MS/MS, the pooled eluates were mixed in a ratio of 1:1 (v:v) with either 5 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) or 50 mg/mL 2,5-dihydroxybenzoic acid (DHB) matrix prepared in 50% acetonitrile (ACN) / 0.1% trifluoroacetic acid (TFA) and spotted onto a MALDI target plate for analysis on a Q-TOF Premier (Waters, Manchester, UK) equipped with a 337nm-UV-MALDI source.

For peak identification using ESI MS/MS, the pooled eluates were first dried, then resuspended in 50% ACN / 0.1% formic acid (FA) and sprayed using a Nanomate HD (Advion Biosciences, Ithaca, USA) as nanoESI ion source in front of the Q-TOF Premier. A positive voltage of 1.7 kV was applied to the chip and the flow rate was kept constant at about 80 nL/min. At this flow rate, a sample volume of 10 µL provided stable static electrospray for at least 2 hrs.

For peak identification using the more complex sample processing of off-line LC fractionation and ESI MS/MS, serum fractionation was first carried out, using an Agilent 1100 LC system (Agilent Technologies inc., Santa Clara, USA). A volume of 100 µL of the pooled serum sample eluates was loaded onto a HiChrom ACE C18 column (2.1mm ID, 150 mm length) (HiChrom Ltd., Reading, UK). The column was initially equilibrated with 5% ACN / 0.1% FA. The elution was achieved using a binary solvent gradient of 5% ACN / 0.1% FA to 70% ACN / 0.1% FA at a flow rate of 100 µL/min over 90 min, and finally up to 95% ACN / 0.1% FA in 10 min. The UV detection was set at 214 nm and fractions were collected at a rate of 1 per minute. The fractions were then dried and resuspended in 40 µL of 50% ACN / 0.1% FA. From each resuspended fraction a volume of 15 µL was taken for chip-based nanoESI analysis using the Nanomate and Q-TOF Premier, 2 µL were used for MALDI-TOF analysis using the Ultraflex and the remaining volume was dried before being resuspended with a 15µL-trypsin solution (100 ng trypsin in 100 mM ammonium bicarbonate, pH 8.5). The resuspended solutions were thoroughly mixed and incubated overnight at 37°C. Finally, the digested fractions were subjected to automatic cleaning using C18 ZipTips® according to the protocol detailed above for MALDI-TOF MS profiling except that the elution step which was performed using 10 µL of 50% ACN / 0.1% FA. The purified digested fractions were then loaded on the Nanomate system and nanoESI MS/MS data were collected using data dependent acquisition. Identification of peptides was performed by searching against the human protein sequence library in NCBInr (v20081128, 216937 sequences) using the Mascot search engine (Matrix Science, London, UK). The searches were performed choosing “none” for digestion enzyme and with a mass tolerance of 0.1 Da for parent ions and 0.2 Da for fragment ions. “Ammonia-loss (N-term C)”, “Deamidated (NQ)”, “Dehydrated (N-term C)”, “Oxidation (M)”, “Phospho (ST)”, and “Phospho (Y)” were set as variable modifications.

GPMAW software v7.10 (Lighthouse data, Odense, Denmark) was used to match accurate mass data obtained from undigested samples with theoretical peptide masses of protein hits obtained from the MASCOT analysis of the MS/MS data of the digested fractions. Isotope pattern software v3.0 (Bruker Daltonics) was used for the comparison of experimental isotopomer distributions with the theoretical distribution from putative peptide identifications and the UniProt and NCBI protein databases were used to extract additional information for substantiating or rejecting putative identities, e.g. information with regard to alternative splicing, prosequence cleavage, disulfide bridges and post-translational modifications.

Results and Discussion

MALDI-TOF MS profiling of 60 serum samples revealed 108 aligned peaks within the combined UKCTOCS and UKOPS serum sample set, using an S/N ratio above 5 (Supplemental Table S1). With the threshold set at S/N > 3, the number of common peaks was 156. Despite the high number of common peaks, the different sample collection and handling protocols for the UKOPS and UKCTOCS sample sets resulted in markedly different MALDI MS profiles (Figure 1). Most of the peak intensities of the 108 common peaks were significantly different between the UKCTOCS and UKOPS samples (Supplemental Table S1). Thus, serum samples from both sample sets were employed in an effort to identify as many profile peaks as possible using the higher mass accuracy and sequencing capability of MALDI Q-TOF instrumentation.

Many of the peak identifications were obtained by direct switching from MALDI MS to MALDI CID MS/MS, analysing the precursor ions determined in the MALDI MS profile. This approach provided a total of 36 identities through MASCOT searching (Supplemental Table S2). The average mass accuracy of the precursor ion mass measurements was 13 ppm. The next approach was using the same samples for nanoESI MS/MS analysis without complex sample processing such as enzymatic digestion or fractionation. A straightforward change of the solvent system made the samples amenable for nanoESI analysis via direct infusion using the chip-based Nanomate system. Similarly to the MALDI analysis, MS profile spectra were first acquired and MS/MS analysis of the precursor ions of the MS profile provided their identity. In this approach, the MALDI and ESI MS profiles were compared to match peaks unambiguously and a total of 16 peaks were identified, of which 6 could not be identified by earlier MALDI MS/MS analysis.

Next, the serum sample eluates were fractionated by off-line C18 reversed phase chromatography and the fractions were analysed by chip-based nanoESI as before. Although LC fractionation was employed, no further advanced sample processing was applied such as depletion or multi-dimensional separation strategies, which have been typically used in previous reports where 200 to 400 peptides were identified [13-15, 17]. In contrast to these reports, the objective of our fractionation strategy was to enhance mass spectrometric analysis of the profile peaks without changing the overall composition to a degree that would make the assignment of the identities too uncertain. Thus, we also monitored the change of the MALDI MS profile after fractionation by comparing the pre-fractionation MALDI MS profile with the MALDI MS (and ESI MS) profiles of each fraction. The fractions were then analysed by chip-based nanoESI MS/MS and the MS/MS data of each precursor ion was searched against the human subset of the NCBInr protein database using MASCOT, which resulted in an additional 28 identities and another 9 confirmations of identities obtained by the two earlier approaches without LC fractionation.

Unfortunately for peaks above m/z 4,500, none of these approaches led to unambiguous peptide identification. The collected LC fractions were therefore further processed by performing tryptic in-solution digestion with subsequent purification as employed for MALDI-TOF MS profiling. The digested fractions were analysed by chip-based nanoESI MS/MS and the MS/MS data of each fraction was combined and searched against the human subset of the NCBInr protein database using MASCOT. For each fraction, the protein hits from

these searches were then used to search for peptide sequences with a theoretical mass that was within 50 ppm of the experimental value of the previously recorded profile peaks from the undigested fractions. In addition, any putative matches were further interrogated by their isotopomer distribution, their known post-translational modifications and the likelihood of occurrence derived from their structural and functional information in protein databases. For instance, any peptide sequences that stretched over a pro-sequence cleavage site were dismissed. Where possible, fragment ion spectra from the previous methods were also used to substantiate putative identities. Overall, this approach led to an additional 11 peak identities.

In total, more than 80 polypeptides were identified with high confidence, representing peptides from 13 different but highly abundant proteins (see Supplemental Table S2). Among these there were three full length proteins, Apoliprotein CII (Apo CII), Apolipoprotein CIII (Apo CIII) and Connective tissue-activating peptide III (CTAP-III). The peaks at m/z 1547 and 1616 were clearly identified by both MALDI MS/MS and ESI MS/MS without LC fractionation, and correspond to Fibrinopeptide A fragments phosphorylated at serine (see Supplemental Figure S1). Interestingly, many of the Fibrinopeptide A/B peaks and others show a loss of 17 Da (ammonia) and sometimes of water. Some of the identified peptides/proteins were found to correspond to oxidised forms already identified. These were excluded from the list since many of these did not appear in the MALDI MS profiles and were most likely products obtained from the sample handling. This clearly shows that the MS peaks/profiles can substantially change with complex sample preparation methods, underlining the need for careful profile peak identification that is different to peptide identification of in-depth proteomic serum analyses.

A detailed comparison of our peak list with the peak lists reported by Villanueva *et al.* [16] and Peng *et al.* [18] shows that the lists are complementary rather than redundant (see Figure 2 and Table 1). It is worth noting that although the number of identified peaks from this study is comparable to the number reported by Villaneuva *et al.* [16], only about 30% (25 peaks) are common between the two lists (Figure 2). Furthermore, the present study was able to identify nine polypeptides in the mass range of 4 – 10 kDa, three times as many as reported so far. However, within this mass range there were many more polypeptides for which the parent protein was identified after digestion but the data was inconclusive for assigning single identities to the MALDI MS profile peaks.

A comparison with a list of plasma peptide identifications reported by Koomen et al. shows that again only about 30% (25 peaks) are common. This, however, constitutes only about 10% of the earlier published plasma peptide list which comprises 249 peptide identifications, reflecting the inherent difference between blood plasma and serum. Interestingly, none of the identifications up to a mass of 1,200 Da are shared. In the serum profiles, these are mostly Fibrinogen peptides. The higher signal intensities of Fibrinogen peptides in serum profiles have been reported previously [19].

Combining this peak list with the lists from Villanueva et al. [16] and Peng et al. [18] increases the number of identified polypeptides of MALDI MS serum profiles to 139. This provides an extensive list of peptide identifications obtained from human blood serum that can be used for provisional identification of MALDI MS profiling peaks below 10 kDa through simple comparison. Furthermore, this will allow, by quick comparison of MS profile and/or MS/MS data, the identification of many potential biomarkers commonly listed in reports without identification and thus, facilitate biomarker validation.

MALDI and ESI MS/MS analysis of some of the MALDI MS profiling peaks resulted in MS/MS spectra different to those usually found for peptides (Figure S2). These fragmentation spectra were typical for phospholipids and their identification was carried out by comparing their accurate masses (within 10 ppm) to the calculated masses obtained from

the LipidMaps database (www.lipidmaps.org). The presence of ion fragments at m/z 86.10, 104.11 and 184.07 in all MS/MS spectra is characteristic for choline, H_2O-choline and phosphocholine, respectively (see Figure S2 and Table S3), and confirmed that these peaks are associated with phosphocholine lipids.

Interestingly, a comparison with a recently published list of endogenous serum peptides that were obtained by extensive prefractionation and ESI-MS/MS shows that more than 40% of the fibrinogen (P02671) peptides in Table 1 were not identified despite the much higher number of identified peptides (>1000) obtained in that study[20]. Other comparisons including lists of endogenous plasma peptides[21] and other proteins showed similar results. In addition, none of these lists include any post-translational modifications (PTMs), and only recently a few phosphorylated endogenous serum peptides were reported[22]. The presented list is based on MALDI MS profile peak identification and as such includes PTMs as well as other non-peptidic components. The focus on MALDI MS profile peak identification also provides an additional and decisive advantage. While other lists might provide many more endogenous peptides, these are not linked to MALDI MS profiling. As a result entries cannot be ranked according to their signal intensity in MALDI MS profiles and many of the entries will have the same mass value within a given mass tolerance. For instance, the above comparison (based only on fibrinogen peptides) reveals that the profile peak at m/z 1518.7 would have been wrongly assigned using the more extensive list of endogenous serum peptides published by Bakun et al.[20]. In the latter, the peptide at the corresponding mass was identified as NRGDSTFESKSYK (cf. Table 1). In the same list, there is no entry for the serine-phosphorylated peptide DSGEGDFLAEGGGV or for its associated peptide that has lost phosphoric acid, which has the same mass value as NRGDSTFESKSYK within a tolerance of 25ppm. In addition, the peptide entries for the profile peak at m/z 1309.6 would have led to an ambiguous identification of either SSSYSKQFTSST or DSGEGDFLAEGGGV, again assuming a mass accuracy of around 25ppm or more. Even at 10ppm most fibrinogen peptides are not unique for their respective mass value. For instance, there are seven fibrinogen peptides with exactly the same m/z value of 1206.57. The above analysis clearly shows that any list of endogenous serum peptides can only be useful for MALDI MS profiling if the peak identifications can be directly linked to the MALDI MS profiles.

In conclusion, the dataset presented provides the identification of more than 90 serum MALDI MS profiling peaks with little overlap with already published datasets. The more than 80 identified polypeptides originate from only 13 proteins, mainly Fibrinogen alpha and beta chain (total of 48). Complementary to other published lists this dataset also provides the identification of very low molecular weight (< 1 kDa) peaks often associated with phospholipids and peaks above 5 kDa. Together with previously published data the most comprehensive MALDI MS serum profiling list consisting of 148 peak identifications is now available (see Table 1).

References

(1) Callesen, A. K.; Madsen, J. S.; Vach, W.; Kruse, T. A.; Mogensen, O.; Jensen, O. N. *Proteomics* **2009**, *9*, 1428-1441.

(2) Anderson, N. L.; Anderson, N. G. *Mol Cell Proteomics* **2002**, *1*, 845-867.

(3) Petricoin, E. F.; Ardekani, A. M.; Hitt, B. A.; Levine, P. J.; Fusaro, V. A.; Steinberg, S. M.; Mills, G. B.; Simone, C.; Fishman, D. A.; Kohn, E. C.; Liotta, L. A. *Lancet* **2002**, *359*, 572-577.

(4) Diamandis, E. P. *Mol Cell Proteomics* **2004**, *3*, 367-378.

(5) Tiss, A.; Smith, C.; Camuzeaux, S.; Kabir, M.; Gayther, S.; Menon, U.; Waterfield, M.; Timms, J.; Jacobs, I.; Cramer, R. *Proteomics* **2007**, *7 Suppl 1*, 77-89.

(6) Villanueva, J.; Philip, J.; Chaparro, C. A.; Li, Y.; Toledo-Crow, R.; DeNoyer, L.; Fleisher, M.; Robbins, R. J.; Tempst, P. *J Proteome Res* **2005**, *4*, 1060-1072.

(7) West-Norager, M.; Kelstrup, C. D.; Schou, C.; Hogdall, E. V.; Hogdall, C. K.; Heegaard, N. H. *J Chromatogr B Analyt Technol Biomed Life Sci* **2007**, *847*, 30-37.

(8) Taguchi, F.; Solomon, B.; Gregorc, V.; Roder, H.; Gray, R.; Kasahara, K.; Nishio, M.; Brahmer, J.; Spreafico, A.; Ludovini, V.; Massion, P. P.; Dziadziuszko, R.; Schiller, J.; Grigorieva, J.; Tsypin, M.; Hunsucker, S. W.; Caprioli, R.; Duncan, M. W.; Hirsch, F. R.; Bunn, P. A., Jr.; Carbone, D. P. *J Natl Cancer Inst* **2007**, *99*, 838-846.
(9) Fiedler, G. M.; Leichtle, A. B.; Kase, J.; Baumann, S.; Ceglarek, U.; Felix, K.; Conrad, T.; Witzigmann, H.; Weimann, A.; Schutte, C.; Hauss, J.; Buchler, M.; Thiery, J. *Clin Cancer Res* **2009**, *15*, 3812-3819.
(10) Anderson, N. L.; Polanski, M.; Pieper, R.; Gatlin, T.; Tirumalai, R. S.; Conrads, T. P.; Veenstra, T. D.; Adkins, J. N.; Pounds, J. G.; Fagan, R.; Lobley, A. *Mol Cell Proteomics* **2004**, *3*, 311-326.
(11) Omenn, G. S.; States, D. J.; Adamski, M.; Blackwell, T. W.; Menon, R.; Hermjakob, H.; Apweiler, R.; Haab, B. B.; Simpson, R. J.; Eddes, J. S.; Kapp, E. A.; Moritz, R. L.; Chan, D. W.; Rai, A. J.; Admon, A.; Aebersold, R.; Eng, J.; Hancock, W. S.; Hefta, S. A.; Meyer, H.; Paik, Y. K.; Yoo, J. S.; Ping, P.; Pounds, J.; Adkins, J.; Qian, X.; Wang, R.; Wasinger, V.; Wu, C. Y.; Zhao, X.; Zeng, R.; Archakov, A.; Tsugita, A.; Beer, I.; Pandey, A.; Pisano, M.; Andrews, P.; Tammen, H.; Speicher, D. W.; Hanash, S. M. *Proteomics* **2005**, *5*, 3226-3245.
(12) Schenk, S.; Schoenhals, G. J.; de Souza, G.; Mann, M. *BMC Med Genomics* **2008**, *1*, 41.
(13) Harper, R. G.; Workman, S. R.; Schuetzner, S.; Timperman, A. T.; Sutton, J. N. *Electrophoresis* **2004**, *25*, 1299-1306.
(14) Tirumalai, R. S.; Chan, K. C.; Prieto, D. A.; Issaq, H. J.; Conrads, T. P.; Veenstra, T. D. *Mol Cell Proteomics* **2003**, *2*, 1096-1103.
(15) Zheng, X.; Baker, H.; Hancock, W. S. *J Chromatogr A* **2006**, *1120*, 173-184.
(16) Villanueva, J.; Shaffer, D. R.; Philip, J.; Chaparro, C. A.; Erdjument-Bromage, H.; Olshen, A. B.; Fleisher, M.; Lilja, H.; Brogi, E.; Boyd, J.; Sanchez-Carbayo, M.; Holland, E. C.; Cordon-Cardo, C.; Scher, H. I.; Tempst, P. *J Clin Invest* **2006**, *116*, 271-284.
(17) Koomen, J. M.; Li, D.; Xiao, L. C.; Liu, T. C.; Coombes, K. R.; Abbruzzese, J.; Kobayashi, R. *J Proteome Res* **2005**, *4*, 972-981.
(18) Peng, J.; Stanley, A. J.; Cairns, D.; Selby, P. J.; Banks, R. E. *Proteomics* **2009**, *9*, 492-498.
(19) Tammen, H.; Schulte, I.; Hess, R.; Menzel, C.; Kellmann, M.; Mohring, T.; Schulz-Knappe, P. *Proteomics* **2005**, *5*, 3414-3422.
(20) Bakun, M.; Karczmarski, J.; Poznanski, J.; Rubel, T.; Rozga, M.; Malinowska, A.; Sands, D.; Hennig, E.; Oledzki, J.; Ostrowski, J.; Dadlez, M. *Proteomics Clinical Applications* **2009**, *3*, 932-946.
(21) Antwi, K.; Hostetter, G.; Demeure, M. J.; Katchman, B. A.; Decker, G. A.; Ruiz, Y.; Sielaff, T. D.; Koep, L. J.; Lake, D. F. *J Proteome Res* **2009**, *8*, 4722-4731.
(22) Hu, L.; Zhou, H.; Li, Y.; Sun, S.; Guo, L.; Ye, M.; Tian, X.; Gu, J.; Yang, S.; Zou, H. *Anal Chem* **2009**, *81*, 94-104.

Legends

Figure 1:
Averaged MALDI-TOF mass spectra of human serum samples collected for UKCTOCS (upper panel) and UKOPS (lower panel) and obtained from purifications of 5µL-aliquots using pre-packed C18 ZipTips®. Averaged mass spectra were generated by ClinProTools software v2.1 after baseline subtraction, smoothing, normalisation and peak realignment.

Figure 2:
Venn diagram showing the number of overlapping and total number of peptides identified in this study and those of the studies of Villanueva *et al.* [16] and Peng *et al.* [18].

Figure S1:
Fragmentation mass spectra of the MALDI MS profile peaks at the average mass of 1466.51 Da and 1617.55 Da acquired by MALDI MS/MS (a, b) and ESI MS/MS (c, d). The MS/MS spectra were acquired using the precursor ion m/z value of 1465.66 and 1616.66 for MALDI, and 733.33 and 808.83 for ESI, respectively. The peaks could be assigned to the following sequences from Fibrinopeptide A: DSGEGDFLAEGGGVR (a, c) and ADpSGEGDFLAEGGGVR (b, d) with Mowse scores of 72 (a), 83 (b), 88 (c) and 122 (d).

Figure S2:
MALDI MS/MS spectra of the precursor ions at m/z 991.67 and 1015.67, corresponding to phosphatidylcholine-containing phospholipids (see Supplemental Table S3 for details).

Table 1. List of MALDI MS profile peaks

Calculated Monoisot. $[M+H]^+$	Calculated Average $[M+H]^+$	Protein/Peptide or Lipid		Peptide Sequences Identified	Study		
		UniProt Entry	Name (Fragment)		Tiss et al.	Villanueva et al.	Peng et al.
445.2518	445.4944	P02671	FPA (12-16)	E.GGGVR	√		
496.3426		NA	PC(16:0/0:0)	NA	√		
518.3236		NA	PC(18:3/0:0)	NA	√		
520.3393		NA	PC(18:2/0:0)	NA	√		
522.3590		NA	PC(18:1/0:0)	NA	√		
524.3742		NA	PC(18:0/0:0)	NA	√		
574.2944	574.6087	P02671	FPA (11-16)	A.EGGGVR	√		
645.3315	645.6867	P02671	FPA (10-16)	L.AEGGGVR	√		
740.2846	740.7172	P02675	FPB (1-7) – NH_3 – H_2O	$Q_{(-NH3)}$GVND$_{(-H2O)}$NE.E	√		
758.4155	758.8522	P02671	FPA (9-16)	F.LAEGGGVR	√	√	
758.5701		NA	PC(16:0/18:2 or 18:0/16:2)	NA	√		
760.5915		NA	PC(16:0/18:1 or 18:0/16:1)	NA	√		
822.4297	822.9731	P10909	Clusterin precursor (217-222)	P.HFFFPK.S		√	
842.3944	842.9214	Q14624	ITIH4 (681-687)	D.HAAYHPF.R		√	
864.4097	864.9299	P02671	FPA (7-15)	D.DFLAEGGGV.R	√		
869.3272	869.8334	P02675	FPB (1-8) – NH_3 – H_2O	$Q_{(-NH3)}$GVND$_{(-H2O)}$NEE.G	√		
904.4676	905.0324	P01042	Bradykinin (1-8)	RPPGFSPF.R		√	
905.4839	906.0288	P02671	FPA (8-16)	D.FLAEGGGVR	√	√	
920.4624	921.0385	P01042	Bradykinin (1-8) [Pro_Hydroxyl]	RPPGFSP$_{(Hydroxyl)}$F.R	√	√	
926.3486	926.8856	P02675	FPB (1-9) – NH_3 – H_2O	$Q_{(-NH3)}$GVND$_{(-H2O)}$NEEG.F	√		
942.4680	943.0358	P01024	Complement C3f fragment (9-16)	I.HWESASLL.R		√	
944.3592	944.8817	P02675	FPB (1-9) – NH_3	$Q_{(-NH3)}$GVNDNEEG.F	√		
991.6807		NA	PC(16:0/0:0) dimer	NA	√		
998.4955	999.1075	Q14624	ITIH4 (681-688)	D.HAAYHPFR.		√	
1015.6831		NA	PC(16:0/0:0) + PC(18:2/0:0) dimer	NA	√		
1020.5108	1021.1174	P02671	FPA (7-16)	G.DFLAEGGGVR	√	√	
1055.5520	1056.1937	P01024	Complement C3f fragment (8-16)	R.IHWESASLL.R		√	
1060.5787	1061.2185	P01042	Bradykinin	RPPGFSPFR		√	
1073.4171	1074.0634	P02675	FPB (1-10) – NH_3 – H_2O	$Q_{(-NH3)}$GVND$_{(-H2O)}$NEEGF.F	√		
1076.5736	1077.2179	P01042	Bradykinin [Pro_Hydroxyl]	RPPGFSP$_{(Hydroxyl)}$FR		√	
1077.5323	1078.1693	P02671	FPA (6-16)	E.GDFLAEGGGVR	√	√	
1091.4277	1092.0559	P02675	FPB (1-10) – NH_3	$Q_{(-NH3)}$GVNDNEEGF.F	√		
1108.4542	1109.1005	P02675	FPB (1-10)	QGVNDNEEGF.F	√		
1194.5273	1195.2274	P02671	FPA (3-15)	D.SGEGDFLAEGGGV.R	√		
1206.5749	1207.2848	P02671	FPA (5-16)	G.EGDFLAEGGGVR	√	√	
1211.6531	1212.3798	P01024	Complement C3f fragment (7-16)	H.RIHWESASLL.R		√	
1220.4855	1221.2413	P02675	FPB (1-11) – NH_3 – H_2O	$Q_{(-NH3)}$GVND$_{(-H2O)}$NEEGFF.S	√		
1238.4961	1239.2519	P02675	FPB (1-11) – NH_3	$Q_{(-NH3)}$GVNDNEEGFF.S	√		
1263.5963	1264.3367	P02671	FPA (4-16)	S.GEGDFLAEGGGVR	√	√	

1277.7153	1278.5258	P10909	Clusterin precursor (217-226)	P.HFFFPKSRIV.R		√	
1309.5542	1310.3160	P02671	FPA (2-15)	A.DSGEGDFLAEGGGV.R	√		
1325.5281	1326.3076	P02675	FPB (1-12) – NH_3	Q$_{(-NH3)}$GVNDNEEGFFS.A	√		
1348.7120	1349.5194	P01024	Complement C3f fragment (6-16)	T.HRIHWESASLL.R		√	
1350.6284	1351.4149	P02671	FPA (3-16)	D.SGEGDFLAEGGGVR	√	√	
1396.5652	1397.4098	P02675	FPB (1-13) – NH_3	Q$_{(-NH3)}$GVNDNEEGFFSA.R	√		
1447.6447	1448.4929	P02671	FPA (2-16) – H_2O	A.DS$_{(-H2O)}$GEGDFLAEGGGVR	√		
1449.7597	1450.6235	P01024	Complement C3f fragment (5-16)	I.THRIHWESASLL.R		√	
1465.6554	1466.5035	P02671	FPA (2-16)	A.DSGEGDFLAEGGGVR	√	√	
1498.7873	1499.6562	P0C0L4	Complement C4-A (1337-1349)	R.NGFKSHALQLNNR.Q		√	
1518.6818	1519.5717	P02671	FPA – H_2O	ADS$_{(-H2O)}$GEGDFLAEGGGVR	√		
1530.8692	1531.8273	Q59FS1	IGIH4	R.RPHFFFPKSRIV.R			√
1536.6924	1537.5823	P02671	FPA	ADSGEGDFLAEGGGVR	√	√	
1542.5939	1543.4566	P02671	FGA chain (605-619)	A.DEAGSEADHEGTHST.K	√		
1545.6000	1546.4482	P02671	FPA (2-16) + P	A.DS$_p$GEGDFLAEGGGVR	√		
1562.8438	1563.7813	P01024	Complement C3f fragment (4-16)	K.ITHRIHWESASLL.R		√	
1616.6594	1617.5493	P02671	FPA + P	ADS$_p$GEGDFLAEGGGVR	√		
1626.8459	1627.7857	P0C0L4	Complement C4-A (1337-1350)	R.NGFKSHALQLNNRQ.I		√	
1690.9387	1691.9540	P01024	Complement C3f fragment (3-16)	S.KITHRIHWESASLL.R		√	
1739.9299	1740.9619	P0C0L4	Complement C4-A (1337-1351)	R.NGFKSHALQLNNRQI.R	√	√	
1751.9187	1752.9510	P01024	Complement C3f fragment (1-15)	.SSKITHRIHWESASL.L		√	
1762.9222	1763.9653	P0C0L4	Complement C4-A (1353-1368)	R.GLEEELQFSLGSKINV.K		√	
1768.8236	1769.8406	P01024	Complement C3 (1322-1337)	S.EETKENEGFTVTAEGK.G	√		
1771.8497	1772.8893	P06727	Apolipoprotein A-IV (288-303)	K.SLAELGGHLDQQVEEF.R		√	
1777.9708	1779.0314	P01024	Complement C3f fragment (2-16)	S.SKITHRIHWESASLL.R		√	
1786.8547	1787.9504	Q14624	ITIH4 (671-687)	L.GLPGPPDVPDHAAYHPF.R	√	√	
1807.9297	1808.9884	P02647	Apolipoprotein A-I (149-163)	A.ELQEGARQKLHELQE.K	√	√	
1847.8379	1848.8859	P02671	FGA chain (607-624)	E.AGSEADHEGTHSTKRGHA.K	√		
1855.8555	1856.9375	P01024	Complement C3 (1321-1337)	R.SEETKENEGFTVTAEGK.G	√		
1865.0028	1866.1295	P01024	Complement C3f fragment (1-16)	SSKITHRIHWESASLL.R	√	√	
1883.8114	1884.8701	P02671	FGA chain (605-622)	A.DEAGSEADHEGTHSTKRG.H	√		
1884.9899	1886.0639	P02671	FGA chain (606-622)	D.HEGTHSTKRGHAKSRPV.R			√
1896.0311	1897.1494	P0C0L4	Complement C4-A (1336-1351)	G.RNGFKSHALQLNNRQI.R	√	√	√
1927.9508	1929.0755	P06727	Apolipoprotein A-IV (288-304)	K.SLAELGGHLDQQVEEFR.R		√	
1943.9080	1945.0261	P01042	HMW Kininogen (440-456)	H.NLGHGHKHERDQGHGHQ.R	√	√	
1971.0433	1972.2228	P02647	Apolipoprotein A-I (251-267)	K.VSFLSALEEYTKKLNTQ		√	
2008.0788	2009.3518	Q0P5N8	TMSB4X (31-47)	K.PDMAEIEKFDKSKLKKT.E	√		
2010.9708	2012.2352	Q14624	ITIH4 (669-687) – NH_3	R.Q$_{(-NH3)}$LGLPGPPDVPDHAAYHPF.R	√		
2021.1039	2022.2950	P01024	Complement C3f fragment	SSKITHRIHWESASLLR		√	√
2027.9974	2029.2378	Q14624	ITIH4 (669-687)	R.QLGLPGPPDVPDHAAYHPF.R		√	√
2045.9172	2047.1100	P01042	HMW Kininogen (480-497)	L.DDDLEHQGGHVLDHGHKH.K	√		
2053.0812	2054.2822	P02647	Apolipoprotein A-I (220-238)	K.ATEHLSTLSEKAKPALEDL.R		√	
2067.1921	2068.4016	P02775	Platelet Basic Protein precursor (110-127)	D.APRIKKIVQKKLAGDESAD			√

2091.9074	2093.0900	P02671	FGA chain (605-624)	A.DEAGSEADHEGTHSTKRGHA.K	√		
2113.0771	2114.3166	Q0P5N8	TMSB4X (52-70)	E.KNPLPSKETIEQEKQAGES	√		
2115.0512	2116.3462	Q14624	(ITIH4) (347-367)	R.NVHSAGAAGSRMNFRPGVLSS.R		√	
2122.8618	2124.1563	P02671	FGA chain (600-619)	K.SYKMADEAGSEADHEGTHST.K	√		
2127.0128	2128.2565	P01042	HMW Kininogen (458-477)	R.GHGLGHGHEQQHGLGHGHKF.K		√	
2162.9445	2164.1688	P02671	FGA chain (604-624)	M.ADEAGSEADHEGTHSTKRGHA.K	√		
2184.0985	2185.4238	Q14624	ITIH4 (669-688)	R.QLGLPGPPDVPDHAAYHPFR.R		√	
2209.0619	2210.3192	P01042	HMW Kininogen (438-456)	R.KHNLGHGHKHERDQGHGHQ.R		√	√
2267.1891	2268.5127	P02649	Apolipoprotein E (212-232)	A.TVGSLAGQPLQERAQAWGERL.R		√	
2271.1305	2272.5274	Q14624	ITIH4 (667-687)	S.SRQLGLPGPPDVPDHAAYHPF.R	√	√	√
2279.2718	2280.6046	P02775	Platelet Basic Protein precursor (108-127)	D.PDAPRIKKIVQKKLAGDESAD		√	√
2305.2034	2306.5524	P0C0L4	Complement C4-A (1353-1374)	R.GLEEELQFSLGSKINVKVGGNS.K		√	
2353.1518	2354.5363	Q0P5N8	TMSB4X (50-70) – NH_3	T.QEKN$_{(-NH3)}$PLPSKETIEQEKQAGES	√		
2358.1626	2359.5787	Q14624	ITIH4 (666-687)	S.SSRQLGLPGPPDVPDHAAYHPF.R		√	
2371.1624	2372.5469	Q0P5N8	TMSB4X (50-70)	T.QEKN$_{(deam)}$PLPSKETIEQEKQAGES	√		
2378.2098	2379.6342	P0C0L4	Complement C4-A (1429-1449)	K.DDPDAPLQPVTPLQLFEGRRN.R	√		
2379.0371	2380.4173	P02671	FGA (577-597)	S.SSYSKQFTSSTSYNRGDSTFE.S		√	
2409.2633	2410.6688	P02649	Apolipoprotein E (210-232)	R.AATVGSLAGQPLQERAQAWGERL.R		√	
2451.2051	2452.6608	P02766	Transthyretin precursor (101-123)	K.ALGISPFHEHAEVVFTANDSGPR.R		√	
2453.9610	2455.4140	Q0P5N8	TMSB4X (49-70) – NH_3	E.TQ$_{(-NH3)}$EKNPLPSKETIEQEKQAGES	√		
2464.0793	2465.5698	P02671	FGA chain (600-622)	K.SYKMADEAGSEADHEGTHSTKRG.H	√		
2471.2260	2472.6679	Q0P5N8	TMSB4X (49-70)	E.TQEKNPLPSKETIEQEKQAGES	√		
2508.3529	2509.7985	P06727	Apolipoprotein A-IV (256-278)	R.ISASAEELRQRLAPLAEDVRGNL.K		√	
2544.2297	2545.6650	P02671	FGA chain (606-629)	D.EAGSEADHEGTHSTKRGHAKSRPV.R			√
2551.1769	2552.7529	P0C0L4	Complement C4-A (957-979)	R.TLEIPGNSDPNMIPDGDFNSYVR.V		√	
2553.1012	2554.5991	P02671	FGA chain (576-598)	K.SSSYSKQFTSSTSYNRGDSTFES.K	√	√	
2565.3644	2566.8549	P02649	Apolipoprotein E (210-233)	R.AATVGSLAGQPLQERAQAWGERLR.A		√	
2567.3729	2568.9088	P02768	Albumin precursor (27-48)	A.HKSEVAHRFKDLGEENFKALVL.I	√		
2582.3402	2583.9185	Q14624	ITIH4 (617-639)	R.NVHSGSTFFKYYLQGAKIPKPEA.S	√		
2599.2635	2600.7734	P02647	Apolipoprotein A-IV (280-303)	K.GNTEGLQKSLAELGGHLDQQVEEF.R		√	
2602.3107	2603.8204	P00488	Factor XIIIa (14-38)	R.AVPPNNSNAAEDDLPTVELQGVVPR.G		√	
2627.3400	2628.9194	Q14624	ITIH4 (663-687)	P.GVLSSRQLGLPGPPDVPDHAAYHPF.R		√	
2659.2567	2660.7791	P02671	FGA chain (605-629)	A.DEAGSEADHEGTHSTKRGHAKSRPV.R	√	√	√
2704.4516	2706.0385	P0C0L4	Complement C4-A (1353-1378)	R.GLEEELQFSLGSKINVKVGGNSKGTL.K		√	
2724.3893	2726.0348	Q14624	ITIH4 (662-687)	R.PGVLSSRQLGLPGPPDVPDHAAYHPF.R		√	
2753.4369	2755.1053	P02768	Albumin precursor (25-48)	R.DAHKSEVAHRFKDLGEENFKALVL.I	√		
2755.3646	2756.9595	P06727	Apolipoprotein A-IV (280-304)	K.GNTEGLQKSLAELGGHLDQQVEEFR.		√	
2768.2282	2769.8514	P02671	FGA chain (576-600)	K.SSSYSKQFTSSTSYNRGDSTFESKS.Y	√	√	√
2778.4520	2780.0741	P02654	Apolipoprotein C-I (29-53)	P.DVSSALDKLKEFGNTLEDKARELIS.R		√	
2816.3234	2817.9581	P02671	FGA (548-574)	R.GSESGIFTNTKESSSHHPGIAEFPSRG.K		√	
2829.4112	2831.0626	Q0P5N8	TMSB4X (46-70)	K.KTETQEKNPLPSKETIEQEKQAGES	√		
2861.3343	2863.6505	P02671	FGA chain (603-629)	K.MADEAGSEADHEGTHSTKRGHAKSRPV.R	√		√
2931.2915	2933.0274	P02671	FGA chain precursor (576-601)	K.SSSYSKQFTSSTSYNRGDSTFESKSY.K	√	√	√

2989.4292	2991.2246	P02671	FGA chain (602-629)	Y.KMADEAGSEADHEGTHSTKRGHAKSRPV.R	√		
3156.6265	3158.5491	Q14624	ITIH4 (617-644)	R.NVHSGSTFFKYYLQGAKIPKPEASFSPR.R	√	√	√
3182.7347	3184.6627	P02647	Apolipoprotein A-I (240-267)	R.QGLLPVLESFKVSFLSALEEYTKKLNTQ		√	
3190.4270	3192.3659	P02671	FGA (576-603)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKM.A		√	√
3200.8001	3202.6864	P0C0L4	Complement C4-A (1353-1382)	R.GLEEELQFSLGSKINVKVGGNSKGTLKVLR.T		√	
3206.4219	3208.3653	P02671	FGA (576-603) [Met-ox]	K.SSSYSKQFTSSTSYNRGDSTFESKSYKM$_{ox}$.A		√	
3239.5246	3241.4788	P02671	FGA chain (600-629)	K.SYKMADEAGSEADHEGTHSTKRGHAKSRPV.R	√	√	√
3261.4641	3263.4440	P02671	FGA chain (576-604)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMA.D	√	√	√
3272.6422	3274.6954	Q14624	ITIH4 (658-687)	R.MNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	√	√	
3277.4590	3279.4434	P02671	FGA (576-604) [Met-ox]	K.SSSYSKQFTSSTSYNRGDSTFESKSYKM$_{ox}$A.D		√	
3377.7230	3379.7435	P02647	Apolipoprotein A-I (148-176)	R.AELQEGARQKLHELQEKLSPLGEEM$_{ox}$RDRA.R		√	
3675.9075	3678.0877	Q0P5N8	TMSB4X (39-70)	K.FDKSKLKKTETQEKNPLPSKETIEQEKQAGES	√		
3953.9098	3956.4034	Q14624	ITIH4 (645-681)	R.RGWNRQAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDH.A	√		
3970.9882	3973.4255	Q14624	(ITIH4) (650-687)	R.QAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R		√	
4280.1319	4282.7883	Q59FS1	IGIH4 (602-642)	R.NVHSAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	√		√
4961.3111	4964.3899	P02671	FGA chain (529-574) or (513-558)	T.FPGFFSPMLGEFVSETESRGSESGIFTNTKESSSHHPGIAEFPSRG.K or R.HPDEAAFFDTASTGKTFPGFFSPMLGEFVSETESRGSESGIFTNTK.E	√		
5334.3536	5337.5402	P02671	FGA chain (577-624)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHA.K	√		√
5901.7029	5905.2293	P02671	FGA chain (577-629)	K.SSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPV.R	√		√
6627.3167	6631.4463	P02655	Apoliporotein C-II (32-90)	M.PSPTFLTQVKESLSSYWESAKTAAQNLYEKTYLPAVDEKLRDLYSKSTAAMSTYTGIFT.D	√		
8200.0508	8205.1815	P02655	Apoliporotein C-II (29-101)	Q.DEMPSPTFLTQVKESLSSYWESAKTAAQNLYEKTYLPAVDEKLRDLYSKSTAAMSTYTGIF TDQVLSVLKGEE	√		
8760.2248	8765.6732	P02656	Apolipoprotein C-III	SEAEDASLLSFMQGYMKHATKTAKDALSSVQESQVAQQARGWVTDGFSSLKDYWSTVKDKFSE FWDLDPEVRPTSAVAA	√		
8910.3855	8915.9261	P02655	Apoliporotein C-II (Chain A)	TQQPQQDEMPSPTFLTQVKESLSSYWESAKTAAQNLYEKTYLPAVDEKLRDLYSKSTAAMSTY TGIFTDQVLSVLKGEE	√		
9282.8208	9288.7182	P02775	CTAP-III	NLAKGKEESLDSDLYAELRCMCIKTTSGIHPKNIQSLEVIGKGTHCNQVEVIATLKDGRKICL DPDAPRIKKIVQKKLAGDESAD	√		

** In the MS/MS spectra of these peaks, Choline, H_2O-Choline and PhosphoCholine peaks were dominant but no corresponding phospholipids were found in the LipidMaps database.*

CTAP-III: Connective tissue-activating peptide III; FGA: Fibrinogen alpha; FPA: Fibrinopeptide A; FPB: Fibrinopeptide B; HMW Kininogen: High molecular weight kininogen; IGIH4: Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein) varaint; ITIH4: Inter-alpha-trypsin inhibitor heavy chain H4; NA: Not applicable; PC: Phophatidylcholine; TMSB4X: Thymosin, beta 4. $[M+H]^+$ values were calculated using IsotopePattern Software (Bruker Daltonics).