Recombinant "IMS TAG" proteins : a new method for validating bottom-up matrix-assisted laser desorption/ionisation ion mobility separation mass spectrometry imaging

COLE, Laura <http://orcid.org/0000-0002-2538-6291>, MAHMOUD, Khaled, HAYWOOD-SMALL, Sarah <http://orcid.org/0000-0002-8374-9783>, TOZER, Gillian M., SMITH, David <http://orcid.org/0000-0001-5177-8574> and CLENCH, Malcolm <http://orcid.org/0000-0002-0798-831X>

Recombinant "IMS TAG" Proteins - A New Method for Validating Bottom Up Matrix Assisted Laser Desorption - Ion Mobility Separation - Mass Spectrometry Imaging (MALDI-IMS-MSI)

Laura M Cole[1] , Khaled Mahmoud[2,] , Sarah Haywood-Small[1],Gillian M Tozer[3], David P. Smith[1], and Malcolm R Clench[*1]

[1] Biomedical Research Centre, Sheffield Hallam University, Howard Street, Sheffield, United Kingdom S1 1WB

[2] Leeds Institute of Molecular Medicine, University of Leeds, LS2 9JT, UK

[3] Department of Oncology, The Medical School, University of Sheffield, Beech Hill Road, Sheffield, S10 2RX

[*] Corresponding Author Professor Malcolm R Clench, Biomedical Research Centre, Sheffield Hallam University, Howard Street, Sheffield S1 1WB m.r.clench@shu.ac.uk

Keywords

MALDI-MSI, mass spectrometry imaging, recombinant protein, IMS-TAG, *in situ* tryptic digest.

Abstract

Rationale

Matrix assisted laser desorption ionisation mass spectrometry imaging (MALDI-MSI) provides a methodology to map the distribution of peptides generated by in situ tryptic digestion of biological tissue. It is challenging to correlate these peptides to

the proteins from which they arise because of the many potentially overlapping and hence interfering peptide signals generated.

Methods

A recombinant protein has been synthesised that when cleaved with trypsin yields a range of peptide standards for use as identification and quantification markers for multiple proteins in one MALDI-IMS-MSI experiment. Mass spectrometry images of the distribution of proteins in fresh frozen and formalin fixed paraffin embedded tissue samples following in situ tryptic digestion were generated by isolating signals on the basis of their m/z value and ion mobility drift time which were correlated to matching peptides in the recombinant standard.

Results

Tryptic digestion of the IMS-TAG protein and MALDI-MS analysis yielded values for m/z and ion mobility drift time for the signature peptides included in it. MALDI-IMS-MSI images for the distribution of the proteins HSP 90 and Vimentin, in FFPE EMT6 mouse tumours and HSP-90 and Plectin in a fresh frozen mouse fibrosarcoma were generated by extracting ion images at the corresponding m/z and drift time from the tissue samples.

Conclusions

The IMS-TAG approach provides a new means to confirm the identity of peptides generated by in situ digestion of biological tissue.

Introduction

Matrix assisted laser desorption ionisation - mass spectrometry imaging (MALDI-MSI) is an advanced analytical tool that allows molecular profiling and imaging of many classes of compounds including; proteins, peptides, lipids, drugs directly from tissue sections. The technique was initially reported by Spengler *et al.* [1] but first successfully applied to the study of biological tissue by Caprioli *et al.* in 1997 [2]. It has since been improved and adapted for use in many other studies [3-5]. In the most commonly used MALDI-MSI method, multiple single mass spectra are acquired across a tissue section at a spatial resolution predefined by the operator (typically 10-200 µm). These mass spectra together generate molecular maps or images which represent the distribution and the relative abundance and/or intensity of a

specific ion signal detected within the tissue section. MALDI-MSI has been shown to be a powerful technique for direct protein analysis within tissue sections [6].

MALDI-MSI has been extensively applied in the study of tumour tissue and has been used for discrimination between tumour and non tumour regions with no requirement for predefined targets [6-9]. A relatively recent and exciting development in the technique is the use of "*on-tissue*" tryptic digestion in order to achieve direct identification of proteins within a tissue section [10-14]. Such molecular profiling and imaging could be described as a bottom-up shotgun approach to protein identification, and the technique has been successfully applied to the analysis of both fresh frozen and formalin fixed paraffin embedded tissues (FFPE) [10-14]

A number of MALDI-MSI methodologies have been reported for the confirmation of protein identity using the analysis of on-tissue digests. These include the use of ion mobility separation coupled to MS/MS [13-15], accurate mass measurement [16], the use of positive controls generated from recombinant proteins [14] and the combination of immunohistochemistry and MALDI-MSI [14]. Current workflows often combine LC-MS/MS and MALDI-MSI approaches to give complementary information [4,17].

We have previously reported the use of a single recombinant protein for validation in a MALDI-MSI experiment studying the distribution of the glucose regulator protein GRP78 in pancreatic tumour samples [14]. In this work a sample of recombinant GRP78 was digested with trypsin and spotted on the sample slide containing the tumour section to be analysed (which had been prepared by on-tissue digestion) The sample slide was then analysed by MALDI-MSI in an experiment which incorporated ion mobility separation. Images were generated from ion intensity information of ions

with a selected *m/z* value and ion mobility drift time. The criterion for positive identification of GRP78 was that a signal with a particular combination of *m/z* and drift time was observed with high ion intensity in both the spot of recombinant digest and in the tissue section, i.e. a tryptic peptide of GRP78. Peptides identified in this manor were validated by in silico digestion of the GRP78 sequence.

Our aim was to synthesise a recombinant protein that when trypsinised yielded peptide standards for identification and quantification of multiple proteins in one MALDI-IMS-MSI experiment, analogous to "QconCAT" technology for LC/MS/MS [18]. To construct the recombinant protein twelve peptides that we had identified as tryptic fragments of target proteins in previous on-tissue digest experiments were chosen (Figure 1). The standard peptide sequences were arranged in the protein sequence in such a way that the charges were evenly spaced across it with the aim of preventing intra cellular aggregation during expression. One of the standard peptides from endothelial growth factor receptor (EGFR) incorporated a methionine residue at its N-terminus and so this peptide was placed at the start of the protein thus allowing the methionine amino to be incorporated during translation in *E. coli*. In-order to aid purification a short peptide sequence (AWLEHHHHHH) containing a six histidine tag was incorporated, tyrosine was also included to give the protein a high extinction coefficient allowing tracking during purification and reliable quantification of the final protein yield. A short "stuffer sequence" containing the required restriction site was also incorporated into this peptide. The final protein had an expected mass of 16,667.26 Da and a predicted extinction coefficient of 9,970 mol/cm (Figure 1). A synthetic DNA construct was produced by translating the protein sequence back to the corresponding nucleotides taking into consideration codon optimisation for the *E. coli* expression system. The synthetic DNA construct

104 was then sub cloned into the expression vector pET23a(+) under the control of a T7
105 promoter and subsequently transformed into an *E. coli* BL21 DE3 strain. The protein
106 expression was induced during the log phase of cell growth and subsequently
107 purified by affinity chromatography.
108
109 The utility of the recombinant protein that we are terming an "IMS-TAG" protein has
110 been tested for two applications: the study of the distribution of proteins associated
111 with cell death and the stress response in formalin fixed paraffin embedded mouse
112 tumour tissue [19] and the study of proteins associated with response to vascular
113 disrupting agents in fresh frozen tissue [20].
114
115 **Experimental**

116 (i) Samples

117 (a) Formalin Fixed Paraffin Embedded (FFPE) tissue.

118 Formalin Fixed Paraffin Embedded (FFPE) subcutaneous EMT6 tumours were
119 grown in syngeneic BALB/c mice. This is an established immunogenic tumour model
120 which can be effectively cured in some circumstances [19]. Laser treatment was
121 employed to induce a small amount of thermal necrosis (11 ± 2%, mean ± S.E (n=6))
122 which was confirmed and monitored by haematoxylin and eosin (H & E) staining.

123 EMT6 adenocarcinoma cells (ATCC No CRL-2755) were purchased from LGC
124 Promochem, Teddington, UK. Cells were cultured in Dulbecco's modified Eagle's
125 medium (DMEM) supplemented with 10% v/v fetal bovine serum, antibiotics (1×10^4
126 units/mL penicillin G sodium, 10 mg/mL streptomycin sulphate in 0.85% w/v saline)
127 and 250 μg/mL fungizone. Cells were grown at 37 °C in a 100% v/v humidified

incubator with a gas phase of 5% v/v CO_2 and routinely screened for Mycoplasma. For in vivo experiments, 1×10^4 EMT6 cells were subcutaneously implanted into the right flank region of anaesthetised (1.5% v/v halothane for <5 min) 7-week old male Balb/c mice (Harlan). Mice were housed on a 12 hour light – dark illumination schedule and had free access to a standard pellet diet and water. Experiments were performed with UK Home Office approval (PPL 40/2343, PPL 40/2972).

Mice with EMT6 tumours reaching 300mm^3 then received PBS injected intravenously via the tail vein. After 24 h, mice were anaesthetised (1.5% v/v halothane for 30 min) and the tumour region was depilated with hair removal cream prior to laser light treatment. Tumours were treated 24 h later at a light dose of 138 J/cm2 with a fluence rate of 75mW/cm2 using a 635 nm laser attached to a microlens fibre (23 min). Mice were killed by cervical dislocation at 24 h or post treatment. Tumours were excised and fixed (0.1 M Tris buffer, pH 7.4, 0.5 g calcium acetate, and 5.0 g zinc acetate and 5.0 g zinc chloride in 1 L of deionised water) overnight before transferral to 70% v/v industrial methylated spirit, prior to processing for histological analysis. Samples were fixed in 10% buffered formalin for 24 hours, dehydrated in 70% EtOH and paraffin embedded. 5 μm sections were cut using a microtome (Leica Microsystems, Bucks, UK) and mounted onto a histological glass slide. Formalin fixed paraffin embedded (FFPE) tissue sections were stored at room temperature until further analysis.

(b) Fresh Frozen Mouse Fibrosarcoma Samples

Mice were injected sub-cutaneously in the dorsum with a 50 μl tumour cell suspension containing 1×10^6 cells in serum-free medium. The cells employed in this

study were from the mouse fibrosarcoma cell line, VEGF188. This has been engineered to express only the VEGF188 isoform [21]. Tumours were allowed to grow to approximately 500 mm^3, before CA-4-P treatment (a single dose of 100 mg/kg intraperitoneal). Mice were killed and tumours excised at various times after treatment.

(ii) Preparation of Recombinant Positive Control Sample

The required DNA sequence was synthesised by MWG-biotech (AG Sequencing Department, Fraunhoferstr, 22 D-82152, Martinsried Germany), and sub cloned into the bacterial expression vector pET23a(+) (Novagen, Merck KGaA, Darmstadt, Germany). The plasmid was then transformed into the *E. coli* strain BL21 DE3 (Promega Corporation, Southampton, U K). All bacterial cultures were grown in luria broth (LB) media with 100 μg/mL ampicillin for selection. 10 mL of a 100 mL overnight starter culture was used to inoculate 1 L of LB media. Cultures were grown at 37 °C with shaking until the OD_{600} reached 0.5, protein expression was induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG). The culture was incubated for a further 2.5 h before the bacterial cells where harvested by centrifugation at 10,000 g for 20 min using a Sorvall RC G+ (Thermo Scientific, Hemel Hampstead UK).

Bacterial pellets where resuspended in binding buffer (20 mM sodium phosphate + 20 mM imidazole pH 7.4) 3mL per 1 g of cells before being lysed by sonication (10 ×20 s pulses at a 35% duty cycle) using a Vibra Cell VCX 750 (750 W). Phenylmethyl sulphonyl fluoride (PMSF) (50 μg/mL), DNase (20 μg/mL), RNase (20 μg/mL) was added to the lysate which was then cleared of insoluble material by centrifugation at 8,000 g for 1 h using an Eppendorf 5804R centrifuge. The clear lysate was filtered through a 0.2 μm syringe filter and load onto a 1 mL Histrap FF

column (GE) at a flow rate of 2 mL/min. Protein was eluted over 20 column volumes using a 0 to 100% gradient into elution buffer (20 mM sodium phosphate + 500 mM imidazole pH 7.4). Fractions containing the recombinant protein were then desalted into 50 mM ammonium acetate using PD10 column (GE Healthcare, Bucks, UK Ltd).

(iii) Tissue preparation

(a) FFPE Tissue

Antigen retrieval was performed by heating in a microwave oven for ~4 min at 50% power in 0.01 M tri-sodium citrate buffer (pH = 6.3). The section was cooled to room temperature, rinsed with water and then allowed to dry. It was then washed in 70% and 90% EtOH for 1 min each before being immersed in chloroform for 15 s. It was then allowed to air dry prior to trypsin and matrix deposition.*(b) Fresh Frozen Tissue*

Frozen tissue sections were cut into ~10 μm, sections using a Leica CM3050 cryostat (Leica Microsystems, Milton Keynes UK). The sections were then freeze thaw mounted on poly-lysine glass slides by warming the underside of the slide gently by application of a fingertip. Mounted slides were either used immediately or stored in an airtight tube at -80 °C for subsequent use.

(iv) In-solution and *In situ* digestion

Trypsin (sequence grade modified, Promega Corporation, SO16 7NS, United Kingdom) digestion was performed on a 50 μl sample of the desalted recombinant protein using 1 μl of trypsin (1mg/mL) added for 1 h plus a further 1 μl of trypsin overnight.

Trypsin used in this study, for on-tissue digests, was prepared in 50 mM ammonium bicarbonate buffer (pH = 8.4) containing 0.5% octyl-α/β-glucoside (OcGlc) (Sigma,

199 UK). *In situ* digestion was performed on both FFPE placental tissue and fresh frozen
200 tumours sections under humid conditions; the trypsin solution, was sprayed onto the
201 sections in a series of 5 layers, at a flow rate 2µl/ minute, using a SunCollect^{TM}
202 automatic sprayer (SunChrom, Friedrichsdorf, Germany). Sections were then
203 incubated overnight at 37 ^{o}C (5% CO_2).

204 (v) Matrix deposition

205 5 mg/mL α-cynano-4-hydroxycinnamic acid (α-CHCA) mixed with aniline .Equimolar
206 amounts of aniline were added to the CHCA solution, i.e. one millilitre of 5 mg/ml
207 CHCA solution contained 2.4 µL of aniline. This matrix solution was then sprayed
208 onto the section using the SunCollect automatic sprayer. The first and second layers
209 were sprayed at 3 µL/min to allow a matrix seeding process. Three subsequent
210 layers were sprayed at 3.5 µL/min.

211 (vi) Imaging Mass Spectrometry

212 Peptide mass fingerprints and images were acquired by MALDI/IMS/MSI using a
213 SYNAPT™ G2 HDMS system (Waters Corporation, Manchester, UK). This
214 instrument has been described in detail by Giles *et al.* [22] It is fitted with a variable
215 repetition rate Nd:YAG laser which was set to 1 kHZ for these experiments.
216 Instrument calibration was performed using standards consisting of a mixture of
217 polyethylene glycol (Sigma-Aldrich, Gillingham, UK) ranging between *m/z* 100 to
218 3000 Da prior to MALDI-IMS-MSI analysis. Imaging data were acquired in positive
219 ion sensitivity mode at a mass resolution of 10,000 FWHM with ion mobility
220 separation enabled and over the mass range 800 Da to 3,000 Da. Image acquisition
221 was performed at 100 µm spatial resolution using variable IMS wave velocity. The

IMS wave velocity parameters applied were ramped over the full IMS cycle with a start velocity (m/s) of 800 and end velocity (m/s) of 200.

(vii) Data Processing

Data were processed using Waters HD Imaging software release 1.1 (Waters Corporation, Manchester, UK). The parameters used were as follows: specificity type (*i.e.* data type) – IMS MS, Number of most intense peaks – 1,000, resolution – 10,000, low energy intensity threshold – 50. The low intensity threshold was set to potentially allow low abundant species to be included.

Results and Discussion

Figure 1 shows (a) the MALDI peptide mass fingerprint obtained from an in-solution digest of the recombinant protein (b) the amino acid sequence of the recombinant protein synthesised and (c) a list of "signature" peptides generated and detected from the protein. As can be seen from Figure 1a the $[M+H]^+$ ions from each of the peptides listed in 1c were discernible. Ion intensities of peptides can be very different even when peptides are present in equimolar amounts [23] This phenomenon was also observed for the signature peptides, for example m/z 1168.5 (LGIHEDSQNR) from HSP90 alpha was 5 times more intense than the signal at m/z 983.4 (CEVGYTGVR) from epiregulin.

In order to use the recombinant protein as a positive control for the validation of MALDI-IMS-MSI data, use was made of the Waters (Manchester, UK) HDI Imaging software. This software allows simultaneous viewing of the mass spectrum, ion

mobility drift time separation and selected mass spectral image. The ion to be imaged can be selected on the basis of its *m/z* and its drift time. Figure 2 shows an example of such an application. Figure 2a shows the image of the distribution of an HSP-90 "signature" peptide, *m/z* 1168.5 (LGIHEDSQNR) in a sample of FFPE mouse EMT6 tumour tissue. The image is constructed by selecting the appropriate *m/z* value in the plot of the ion mobility separated mass spectral data (Figure 2c). In Figure 2c the ion selected is shown highlighted in red, the corresponding peak in the mass spectrum (Figure 2b) (taken from a selected area) is also then highlighted in red. Using this methodology to select the appropriate ion mobility drift time for the peptide of interest allows increased specificity and hence confidence in ion selection compared to the m/z only scenario (i.e. without MS/MS data). As can be seen the *m/z* value of interest gives a clear signal in the region of the image containing the spot of digested recombinant standard (Figure 2a bottom left hand corner) and the tissue..

A similar approach to that described above was used to image the distribution of Vimentin in the FFPE mouse EMT6 tumour tissue (Figure 2 d-f). Figure 2d shows the image of the distribution of the Vimentin "signature" peptide *m/z* 1093.5 (FADLSEAANR) selected on the basis of its *m/z* ratio and ion mobility drift time (Figures 2e and respectively). Again the distribution of the peptide in the tissue and its presence in the region of the image containing the spot of digested recombinant standard are clearly observable.

In contrast *m/z* 944.5 was also selected for imaging. This is a signal we have previously reported as being readily observable in *in situ* digests of tissue. The signal corresponds to the [M+H]+ ion of the peptide sequence AGLQFPVGR, a tryptic peptide of Histone H2A [14]. This peptide sequence was not incorporated into the recombinant protein sample and hence is employed here as a negative control. Figure 2g shows the distribution for this signature peptide of Histone H2A in the tissue and recombinant standard spot; Figure 2h its abundance in the peptide mass fingerprint and Figure 2i an expanded region of the ion mobility drift time plot for this sample. As can be seen, whilst in Figure 2g the signal is clearly observable in the tissue, it is not observable in that region of the image containing the spot of digested recombinant standard (compare these data with Figures 2a and 2d).

In our work examining the response of tumours to vascular disrupting agents we have become interested in the distribution of HSP90 and Plectin in tumour tissue following administration of the vascular disrupting agent combretastatin-4-AP. Therefore the recombinant protein was engineered to contain the signature peptide sequences LGIHEDSQNR ($[M+H]^+$ *m/z* 1168.5) representing HSP90-alpha, GVVDSEDLELNISR ($[M+H]^+$ *m/z* 1513.7) representing HPS90-beta, and AQAELEAQELQR ($[M+H]^+$ *m/z* 1385.7) and DSQDAGGFGPEDR ($[M+H]^+$ (*m/z* 1350.5) representing Plectin. Figure 3 a-c and d-f show data from the analysis of a section of a fresh frozen mouse fibrosarcoma genetically engineered to express only the VEGF188 isoform (21), 72 hours after administration of combretastatin-4-AP. Figures 3a-c show the distribution of m/z 1168.5 isolated by m/z and drift time as previously described. The m/z and drift time were optimised using signal from recombinant standard with sequence LGIHEDSQNR observed in the reference spot.

Figures 3d-f show the distribution of m/z 1350.5 similarly optimised. The distribution of the peptides in the tumour tissue and the corresponding signal in the spot of recombinant IMS-TAG standard are again clearly visible, giving good confirmation that the signals identified in the tissue are arising from the same species as that in the recombinant IMS-TAG standard.

Figure 4 shows a MALDI positive ion product mass spectrum obtained from the ion at *m/z* 1168.5 (LGIHEDSQNR) present in the tryptic digest of the recombinant "IMS-TAG" protein. The spectrum is of high quality and searching the spectrum against the SwissProt protein sequence database using the MASCOT MS/MS search engine (http://www.matrixscience.com) correctly identified the peptide as a tryptic fragment of mouse HSP90 alpha with a MASCOT score of 65 ($p<0.05$ >19). The spectrum was recorded using the "transfer fragmentation" feature of the Synapt instrument i.e. CID occurred after the travelling wave ion mobility device but before the ions entered the time of flight mass analyser. Using this mode of fragmentation precursor and product ions have the same ion mobility drift time. This is shown in the inset portion of Figure 4 which shows a plot of ion intensity against ion mobility drift time as a heat map. As can be seen all ions in the spectrum have the same drift time.

The use of transfer fragmentation on Synapt type instruments in combination with recombinant "signature" peptides opens up a number of possibilities for further experimentation. We have previously demonstrated that it is possible to carry out pseudo-MRM type experiments for the targeted detection of proteins in tissue following *in situ* digests [24]. In this work we were able to image the distribution of

eight proteins in a pancreatic tumour section via the use of product ions from "signature peptides" and high speed MRM, monitoring 26 transitions over a 3 s cycle time. This was possible since the instrument used had been modified to incorporate a 20 kHz laser and utilised the Applied Biosystems "Dynamic Pixel" feature to move the target plate within the boundaries of each defined pixel area to improve sensitivity. Figure 5 (a-b) illustrates the feasibility of this approach; shown are product ion spectra for the Actin tryptic peptide AVFPSIVGRPR obtained (a) from the recombinant standard and (b) directly from tissue. For each of these spectra although the spectral quality is too poor to allow unambiguous database searching, the y2 and y3 product ions at m/z 272.1 and 428.2 are clearly visible. Hence these mass to charge values (along with their corresponding drift times) are the targets for MS^e type imaging. This will be the subject of further investigation.

Conclusions

A recombinant protein has been synthesised to contain a range of "signature" peptides previously identified as arising from proteins via *in situ* tryptic digestion of mammalian tissue. Tryptic digests of this new type of recombinant protein, that we are terming an "IMS-TAG" protein, have been used to confirm the identity of peptides in *in situ* tryptic digests of fresh frozen and FFPE tissue. This was achieved by using the combination of the *m/z* and ion mobility drift time identified for the "signature" peptide in the tryptic digest of the recombinant standard and extracting from the MALDI-imaging data set images to represent the distribution of only that combination of *m/z* and drift time. This minimises the risk of isobaric interferences in the MALDI-MSI images.

IMS-TAG proteins can be synthesised to contain any desired peptides and we are currently investigating the utility of the IMS-TAG approach for MS^e pseudo MRM experiments for targeted protein imaging and the incorporation of multiple isotopically labelled peptides as a potential source of internal standards for protein quantification in MALDI-MSI.

Acknowledgements

Laura Cole is funded by Programme Grant C1276/A10345 from Cancer Research UK and EPSRC with additional funding from MRC and Department of Health England.

References

1. B. Spengler, M. Hubert, R. Kaufmann, in Proc. 42nd Conf. Mass Spectrometry and Allied Topics, Chicago, Illinois, 1994, p. 1041.

2. R. M. Caprioli, T. B. Farmer and J. Gile. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS *Anal. Chem.* 1997; **69**, 4751.

3. L. A. McDonnell, R. M. A. Heeren. Imaging mass spectrometry. Mass Spectrom. Rev. 2007, **26**, 606

4. S. Francese and M. R. Clench. “MALDI Mass Spectrometry Imaging, a New Frontier in Biostructural Techniques: Applications in Biomedicine”. in "Mass Spectrometry for Microbial Proteomics" Ed Shah and Garbia, Wiley, London (2010).

5. H. Gagnon, J. Franck, M. Wisztorski, R. Day, I. Fournier , M. Salzet. Targeted Mass spectrometry Imaging: Specific Targeting Mass Spectrometry imaging technologies from history to perspective Prog Histochem Cytochem 2012,**47**, 133.

6. L. A . McDonnell, G. L. Corthals, S. M. Willems , A. Van. Remoortere, R. J. M. Van Zeijl, A. M. Deelder. Peptide and protein imaging mass spectrometry in cancer research. J Proteom. 2010, **73.** 1921.

7. K. Schwamborn, R. C. Krieg, M. Reska, G. Jakse , R. Knuechel , A. Wellmann. Identifying prostate carcinoma by MALDI-Imaging. International Journal of Molecular Medicine. 2007, **20**, 155.

8. S.O. Deininger, M.P. Ebert, A. Futterer, M. Gerhard, C. Rocken. MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers J Proteome Res, 2008, **7,** 5230.

9. M.C. Djidja, E. Claude, M. F. Snel, S. Francese, P. Scriven, V. Carolan, M. R. Clench. Novel molecular tumour classification using MALDI-mass spectrometry imaging of tissue micro-array Anal Bioanal Chem, 2010, **397**, 587.

10. S. Shimma , M. Furuta , K. Ichimura , Y. Yoshida , M. Setou. A Novel Approach to *In situ* Proteome Analysis Using Chemical Inkjet Printing Technology and MALDI-QIT-TOF Tandem Mass Spectrometer. Journal of Mass Spectrometry Society Japan.2006, **54**, 133.

11. R. Lemaire , A. Desmons , J. C. Tabet , R. Day , M. Salzet , I. Fournier. I Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. Journal of Proteome Research 2007, 6**,** 1295.

12. M. R. Groseclose , M. Andersson , W. M. Hardesty , R. M. Caprioli. Identification of proteins directly from tissue: *in situ* tryptic digestions coupled with imaging mass spectrometry. Journal of Mass Spectrometry. 2007, **42**, 254.

13. M-C. Djidja , S. Francese , P. M. Loadman, C. W. Sutton, P. Scriven , E. Claude , M. F. Snel, J. Franck , Salzet M, Clench MR. "Detergent addition to tryptic digests and Ion Mobility separation prior to MS/MS improves peptide yield and Protein Identification for *in situ* Proteomic Investigation of Frozen and FFPE Adenocarcinoma tissue sections." Proteomics 2009. **9**, 2750.

14. M-C. Djidja, E. Claude, M. F. Snel, P. Scriven, S. Francese, V. A. Carolan, M. R. Clench. MALDI-Ion Mobility Separation-Mass Spectrometry Imaging of Glucose-Regulated Protein 78 kDa (Grp78) in Human Formalin Fixed Paraffin Embedded Pancreatic Adenocarcinoma Tissue Sections, J Proteome Res 2009, **8**, 4876..

15. J. Stauber, L. MacaLeese, J. Franck,E. Claude, M. F. Snel, B. K. Kaletas, I. M. Wiel, M. Wisztorski, I. Fournier , R. M. Heeren. *OnTissue* Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry, J Am Soc Mass Spectrom, 2010, **21**, 338.

16. Y. Schober, T. Schramm, B. Spengler, *et al*.. Protein identification by accurate mass matrix-assisted laser desorption/ionization imaging of tryptic peptides, Rapid Commun. Mass Spectrom. 2011, **25**, 2475.

17. O. J. R. Gustafsson, J. S. Eddes, S. Meding, S. R. McColl, M. K. Oehler, P. Hoffmann. Matrix-assisted laser desorption/ionization imaging protocol for *In situ*

406 characterization of tryptic peptide identity and distribution in formalin-fixed tissue
407 Rapid Commun Mass Spectrom, 2013, **27,** 655.

408 18. R. J. Beynon, M. Doherty, J. M . Pratt, S. J. Gaskell..“Multiplexed absolute
409 quantification in proteomics using artificial QCAT proteins of concatenated signature
410 peptides” Nat Meth, 2005, **2**, 587.

411 19. M. Korbelik , G. Krosl, J. Krosl , G.J. Dougherty, The role of host lymphoid
412 populations in the response of mouse EMT6 tumour to photodynamic therapy,
413 Cancer Research, 1996, **56**, 5647.

414 20. L. M. Cole , M-C Djidja, J. Bluff , E. Claude , V. A. Carolan , M. Paley, G. M.
415 Tozer and M. R. Clench, "Investigation of Protein Induction in Tumour Vascular
416 Targeted Strategies by MALDI-MSI" Methods 2011, **54**, 442.

417 21. G. M. Tozer, S. Akerman, N. A. Cross, P. R. Barber, M. A. Björndahl, O. Greco,
418 S. Harris , S. A. Hill, D. J. Honess, C. R. Ireson, K. L. Pettyjohn, V. E. Prise, C. C.
419 Reyes-Aldasoro, C. Ruhrberg, D. T. Shima, C. Kanthou . Blood Vessel Maturation
420 and Response to Vascular-Disrupting Therapy in Single Vascular Endothelial Growth
421 Factor-A Isoform–Producing Tumours, Cancer Res, 2008, **68,** 2301.

422 22. K. Giles, J. P. Williams, I. Campuzano . Enhancements in Travelling Wave Ion
423 Mobility Resolution. Rapid Commun Mass Spectrom 2011, **25**, 559.

424 23. M-L Valero, E Giralt, D Andreu, An Investigation of Residue Specific
425 Contributions to Peptide Desorption in MALDI-TOF Mass Spectrometry. Lett. Pept.
426 Sci. 1999, **6**, 109.

427 24. P.J. Trim, M-C. Djidja, S.J. Atkinson, K. Oakes, L.M. Cole, D.M.G. Anderson,
428 P.J. Hart, S. Francese, M.R. Clench, A Hybrid Quadrupole Time of Flight Mass

Spectrometer Modified to Incorporate a 20kHz Nd:YVO4 Laser for MALDI-MS Imaging, Anal Bioanal Chem 2010, **397**, 3409.

Legends for Figures and Tables

Figure 1: (a) MALDI-MS peptide mass fingerprint generated from an in-solution digest of the recombinant "IMS-TAG" protein produced to assist in the interpretation of MALDI-MSI images generated from *in situ* tryptic digests. The protein was designed to contain signature peptides from 10 important proteins. (b) amino acid sequence of the "IMS-TAG" protein and (c) List of peptides generated following trypsin digest of the "IMS-TAG" protein.

Figure 2: MALDI-IMS-MSI Data generated from the Analysis of FFPE EMT6 mouse tumour tissue (a) MALDI-IMS-MSI image for the distribution of *m/z* 1168.5 selected such that both its *m/z* and ion mobility drift-time matched that of the peptide LGIHEDSQNR (from HSP 90) produced by in-solution digest of the recombinant IMS-TAG protein, the distribution of the peptide in both the tissue and the positive control spot of the recombinant standard are clearly visible. (b) Partial MALDI mass spectrum of the most abundant 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software (c) Display of drift time vs *m/z* for the 1,000 most intense peaks in the imaging data set generated by the Waters HD Imaging Sofware. In figures 2(b) and 2(c) the peak of interest has been highlighted. (d) MALDI-IMS-MSI image for the distribution of *m/z* 1093.5 selected such that both its *m/z* and ion mobility drift-time matched that of the peptide FADLSEAANR (from Vimentin) produced by in-solution digest of the recombinant IMS-TAG protein, the distribution of the peptide in both the tissue and the positive control spot of the recombinant standard are clearly visible. (e) Partial MALDI mass spectrum of the most abundant 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software (f) Display of drift time vs *m/z* for the 1,000 most intense peaks in the imaging data set generated by the Waters HD Imaging Sofware. In figures 3(b) and 3(c) the peak

of interest has been highlighted.(g) MALDI-IMS-MSI image for the distribution of m/z 944.5 believed to be AGLQFPVGR (from Histone H2A) (b) Partial MALDI mass spectrum of the most abundant 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software (h) Display of drift time vs *m/z* for the 1,000 most intense peaks in the imaging data set generated by the Waters HD Imaging Sofware. In figures 4(b) and 4(i) the peak of interest has been highlighted. As can be seen, since this peptide was not included in the recombinant IMS-TAG standard although it is clearly visible in the tissue it is not highlighted in the region of the image covering the spotted recombinant standard.

Figure 3: MALDI-IMS-MSI Data Generated from the Analysis of a fresh frozen mouse fibrosarcoma genetically engineered to express only the VEGF188 isoform taken 72 hours after the administration of combretastatin-4-AP. (a) MALDI-IMS-MSI image for the distribution of *m/z* 1168.6 selected such that both its *m/z* and ion mobility drift-time matched that of the peptide LGIHEDSQNR (from HSP90 alpha) produced by in-solution digest of the recombinant IMS-TAG protein, the distribution of the peptide in both the tissue and the positive control spot of the recombinant standard are clearly visible. (b) Partial MALDI mass spectrum of the most abundant 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software (c) Display of drift time vs *m/z* for the 1,000 most intense peaks in the imaging data set generated by the Waters HD Imaging Sofware. In figures 5(b) and 5(c) the peak of interest has been highlighted.

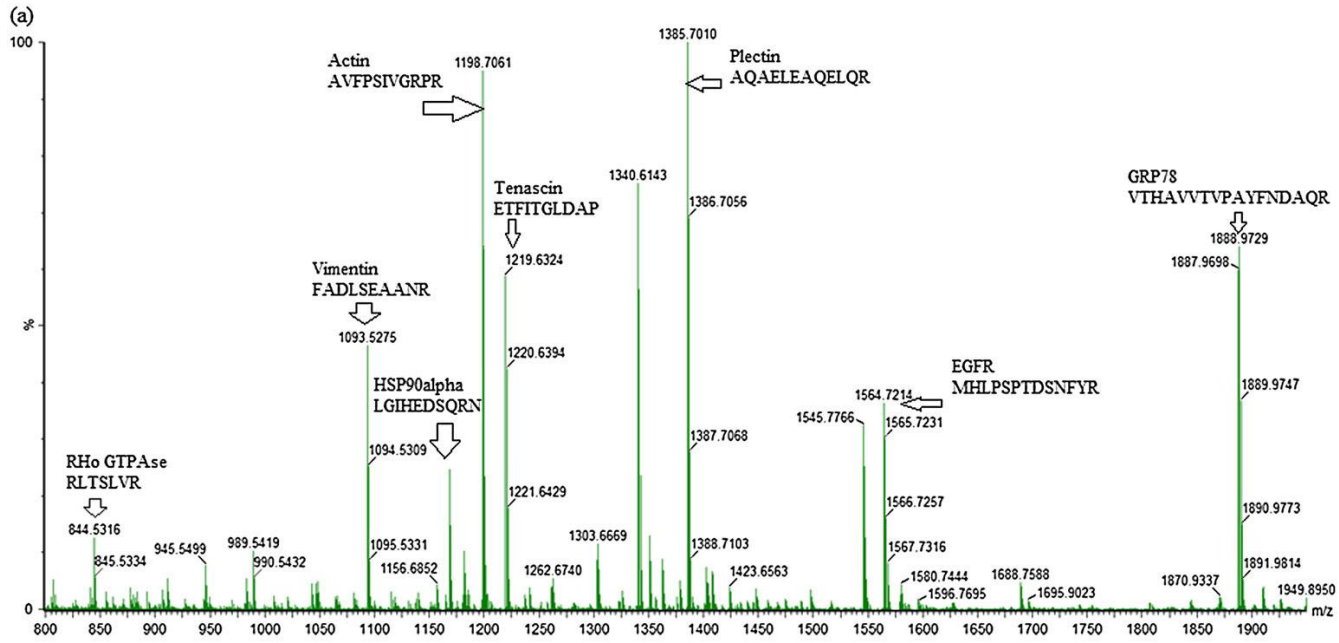
(d) MALDI-IMS-MSI image for the distribution of *m/z* 1350.5 selected such that both its *m/z* and ion mobility drift-time matched that of the peptide DSQDAGGFGPEDR (from Plectin) produced by in-solution digest of the recombinant IMS-TAG protein, the distribution of the peptide in both the tissue and the positive control spot of the

recombinant standard are clearly visible. (e) Partial MALDI mass spectrum of the most abundant 1,000 peaks in imaging data set the generated by the Waters HDI Imaging Software (f) Display of drift time vs *m/z* for the 1,000 most intense peaks in the imaging data set generated by the Waters HDI Imaging Sofware. In figures 3(e) and 3(f) the peak of interest has been highlighted.

Figure 4: MALDI positive ion product ion spectrum obtained from the ion at *m/z* 1168.5 arising from the peptide LGIHEDSQNR and present in the tryptic digest of the recombinant "IMS-TAG" protein. These data were obtained using the "transfer fragmentation" feature of the Synapt instrument and hence the precursor ion and product ions have the same ion mobility drift-time (inset box).

Figure 5: MALDI Positive Ion product ion spectra obtained from the ion at *m/z* 1198.7 arising from the Actin tryptic peptide AVFPSIVGRPR obtained (a) directly from EMT6 tumour tissue (b) from the recombinant standard. The y2 ion (*m/z* 272) is clearly visible in each spectrum and the inset box shows the presence of the y3 (*m/z* 428) ion. It is suggested that even though these spectra are of low quality the combination of product ion mass and drift time would be sufficient to allow the distribution of this protein to be imaged with confidence.

504 Figure 1



(b) Amino acid sequence

MHL PSP TDS NFY RVN SDE VGG EAL GRA VFP SIV GRP RRL TSL VRE TFI TGL DAP RGV VDS EDL ELN ISR LGI HED SQN RFA DLS EAA NRA QAE LEL QEL QRD SQD AAG FGP EDR CEV GYT GVR VTH AVV TVP AYF NDA QRA WLE HHH HHH

(c) Peptide List

Peptide	m/z	Peptide Sequence
GRP-78	1887.9	VTHAVVTVPAYFNDAQR
EGFR	1564.7	MHLPSPTDSNFYR
HSP- 90 beta	1513.7	GVVDSEDLELNISR
Plectin	1385.7	AQAELEAQELQR
	1350.5	DSQDAGGFGPEDR
Haemoglobin beta chain	1302.6	VNSDEVGGEALGR
Tenascin	1219.6	ETFITGLD
Actin	1198.7	AVFPSIVGRPR
HSP-90 alpha	1168.5	LGIHEDSQNR
Vimentin	1093.5	FADLSEAANR
Epiregulin	983.4	CEVGYTGVR
Rho GTPase activating protein 2	844.5	RLTSLVR
Tag reporter		AWLE HHH HHH

505

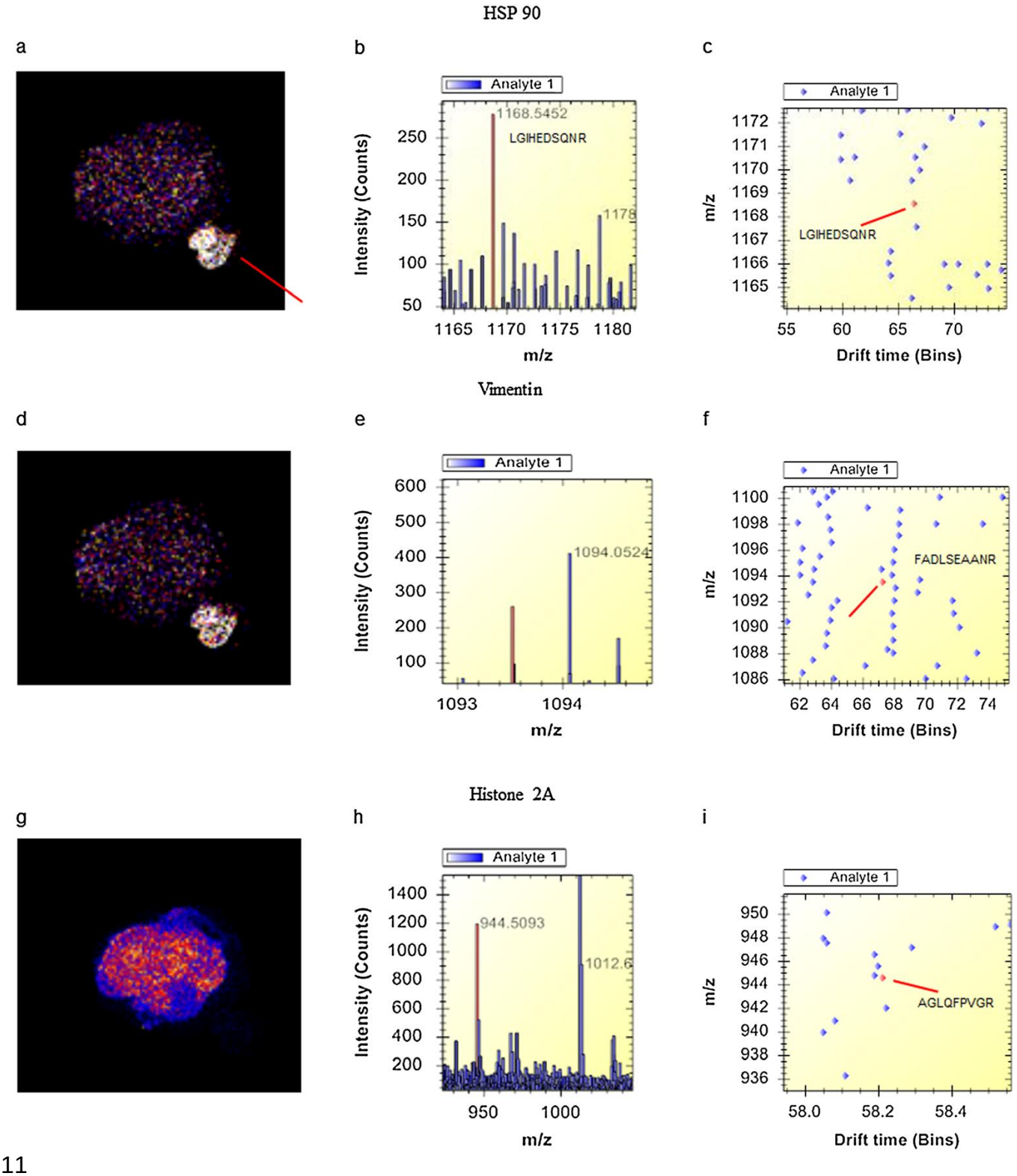
506

507

508

509

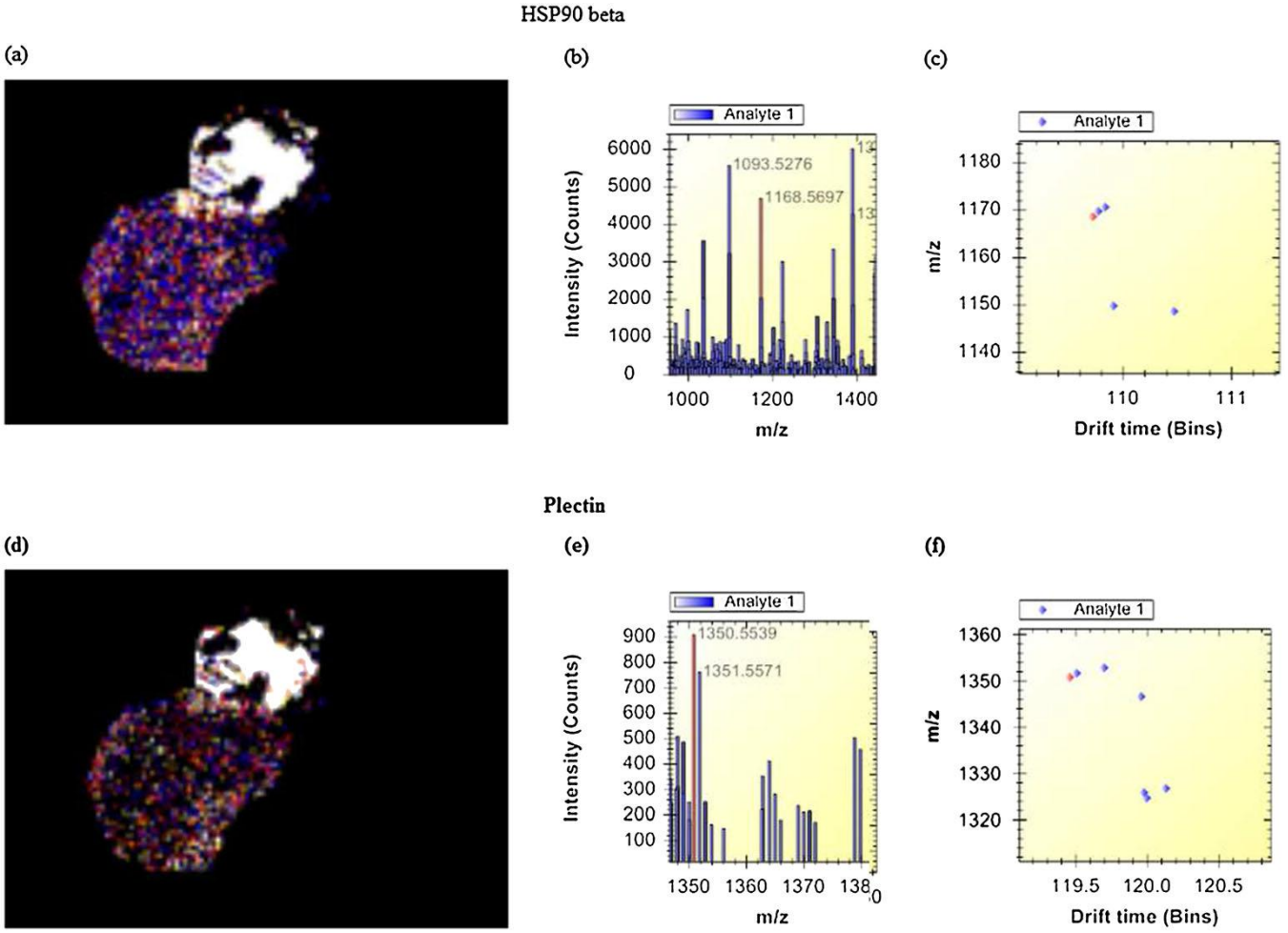
510 Figure 2



511

512

513 Figure 3



514

515

516

517

518

519

520

Figure 4

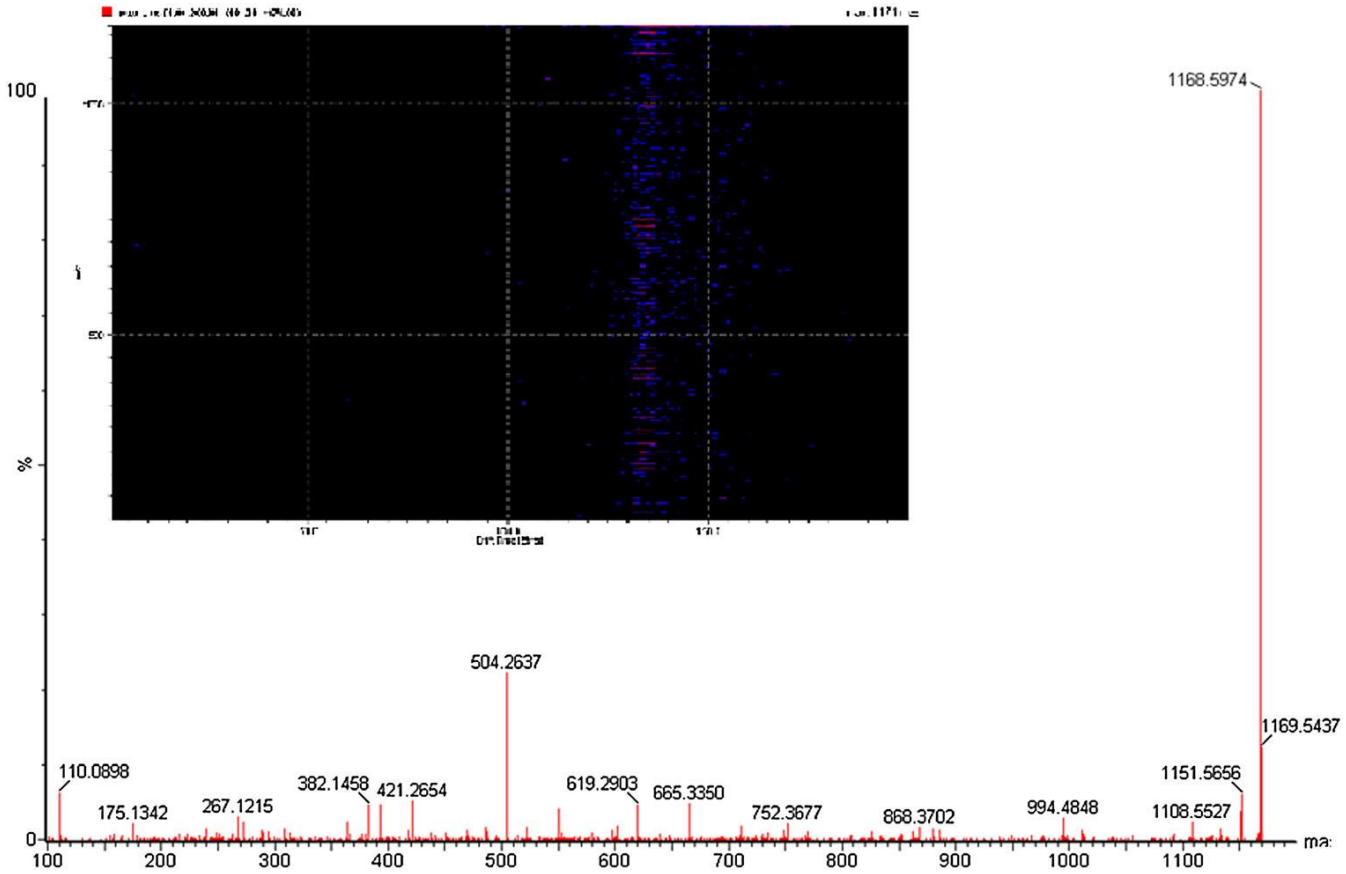
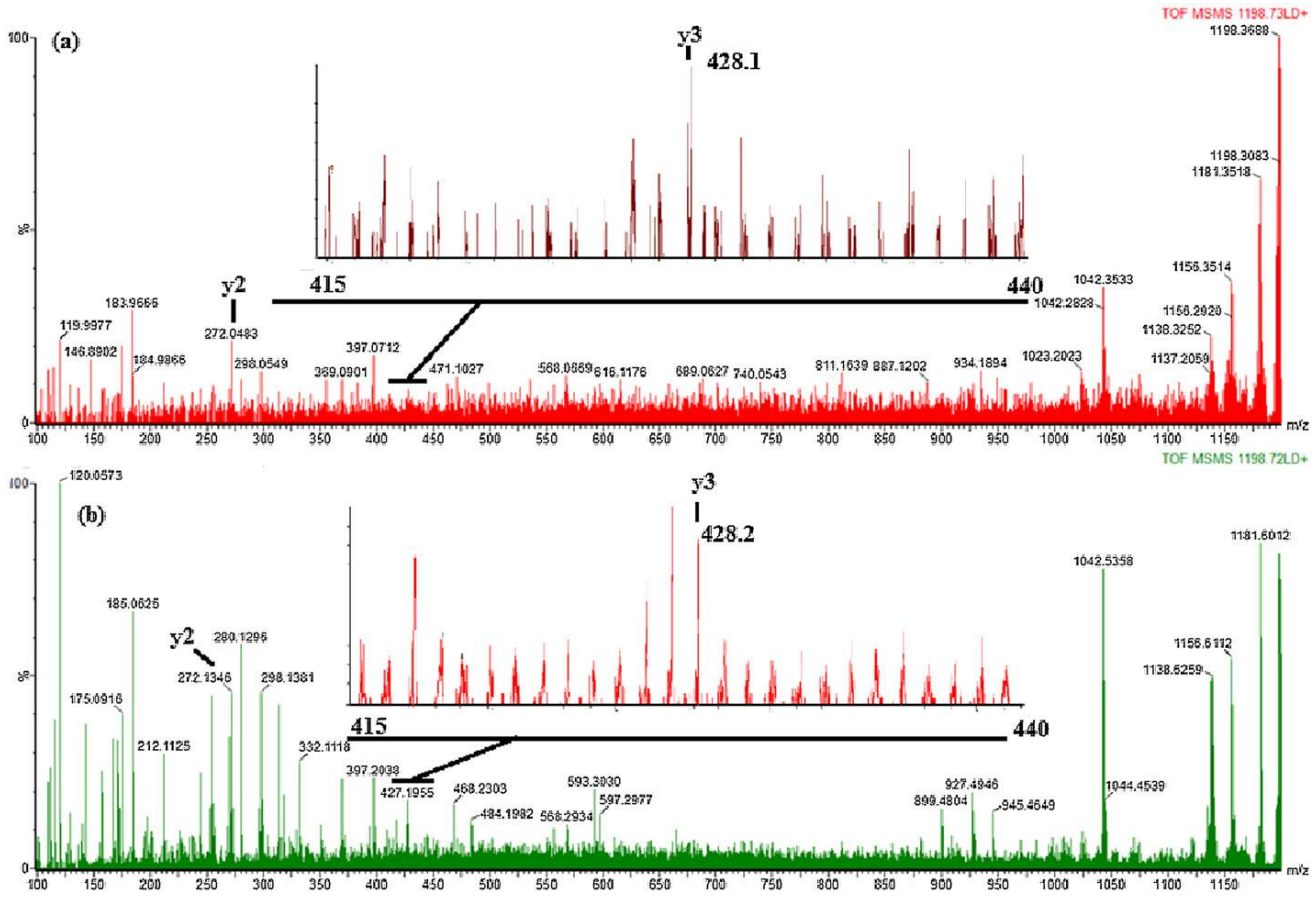

100
%
0
1168.5974
504.2637
1169.5437
110.0898
175.1342
267.1215
382.1458
421.2654
619.2903
665.3350
752.3677
868.3702
994.4848
1108.5527
1151.5656
100
200
300
400
500
600
700
800
900
1000
1100

Figure 5


(a)
TOF MSMS 1198.73LD+
y3
428.1
y2
415
440
1198.3688
1198.3083
1181.3518
1156.3514
1042.3533
1042.2828
1156.2920
1138.3252
1137.2059
1023.2023
934.1894
811.1639
887.1202
740.0543
689.0627
616.1176
568.0869
471.1027
397.0712
369.0901
298.0549
272.0483
184.9866
183.9666
146.8802
119.9977
(b)
TOF MSMS 1198.72LD+
y3
428.2
y2
415
440
1181.6012
1042.5358
1156.6112
1138.6259
1044.4539
945.4649
927.4946
899.4804
593.3030
597.2977
568.2934
484.1982
468.2303
427.1955
397.2093
332.1118
298.1381
280.1296
272.1346
212.1125
185.0625
175.0916
120.0573
m/z