

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>>

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/8224/>

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

COLE, Laura, MAHMOUD, Khaled, HAYWOOD-SMALL, Sarah, TOZER, Gillian M., SMITH, David and CLENCH, Malcolm (2013). Recombinant "IMS TAG" proteins : a new method for validating bottom-up matrix-assisted laser desorption/ionisation ion mobility separation mass spectrometry imaging. *Rapid Communications in Mass Spectrometry*, 27 (21), 2355-2362.

Repository use policy

Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in SHURA to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

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

4 Laura M Cole¹, Khaled Mahmoud², Sarah Haywood-Small¹, Gillian M Tozer³, David
5 P. Smith¹, and Malcolm R Clench^{*1}

6 ¹ Biomedical Research Centre, Sheffield Hallam University, Howard Street, Sheffield,
7 United Kingdom S1 1WB

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

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

11
12 * Corresponding Author Professor Malcolm R Clench, Biomedical Research Centre,
13 Sheffield Hallam University, Howard Street, Sheffield S1 1WB m.r.clench@shu.ac.uk

14 **Keywords**

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

17
18 **Abstract**

19 *Rationale*

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

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

25 *Methods*

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

33 *Results*

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

40 *Conclusions*

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

43

44 **Introduction**

45 Matrix assisted laser desorption ionisation - mass spectrometry imaging (MALDI-
46 MSI) is an advanced analytical tool that allows molecular profiling and imaging of
47 many classes of compounds including; proteins, peptides, lipids, drugs directly from
48 tissue sections. The technique was initially reported by Spengler *et al.* [1] but first
49 successfully applied to the study of biological tissue by Caprioli *et al.* in 1997 [2]. It
50 has since been improved and adapted for use in many other studies [3-5]. In the
51 most commonly used MALDI-MSI method, multiple single mass spectra are acquired
52 across a tissue section at a spatial resolution predefined by the operator (typically
53 10-200 μm). These mass spectra together generate molecular maps or images
54 which represent the distribution and the relative abundance and/or intensity of a

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

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

65 A number of MALDI-MSI methodologies have been reported for the confirmation of
66 protein identity using the analysis of on-tissue digests. These include the use of ion
67 mobility separation coupled to MS/MS [13-15], accurate mass measurement [16], the
68 use of positive controls generated from recombinant proteins [14] and the
69 combination of immunohistochemistry and MALDI-MSI [14]. Current workflows often
70 combine LC-MS/MS and MALDI-MSI approaches to give complementary information
71 [4,17].

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

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

84 Our aim was to synthesise a recombinant protein that when trypsinised yielded
85 peptide standards for identification and quantification of multiple proteins in one
86 MALDI-IMS-MSI experiment, analogous to "QconCAT" technology for LC/MS/MS
87 [18]. To construct the recombinant protein twelve peptides that we had identified as
88 tryptic fragments of target proteins in previous on-tissue digest experiments were
89 chosen (Figure 1). The standard peptide sequences were arranged in the protein
90 sequence in such a way that the charges were evenly spaced across it with the aim
91 of preventing intra cellular aggregation during expression. One of the standard
92 peptides from endothelial growth factor receptor (EGFR) incorporated a methionine
93 residue at its N-terminus and so this peptide was placed at the start of the protein
94 thus allowing the methionine amino to be incorporated during translation in *E. coli*.
95 In-order to aid purification a short peptide sequence (AWLEHHHHHH) containing a
96 six histidine tag was incorporated, tyrosine was also included to give the protein a
97 high extinction coefficient allowing tracking during purification and reliable
98 quantification of the final protein yield. A short "stuffer sequence" containing the
99 required restriction site was also incorporated into this peptide. The final protein had
100 an expected mass of 16,667.26 Da and a predicted extinction coefficient of 9,970
101 mol/cm (Figure 1). A synthetic DNA construct was produced by translating the
102 protein sequence back to the corresponding nucleotides taking into consideration
103 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 ($11 \pm 2\%$, mean \pm 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 μ g/mL fungizone. Cells were grown at 37 °C in a 100% v/v humidified

128 incubator with a gas phase of 5% v/v CO₂ and routinely screened for Mycoplasma.
129 For in vivo experiments, 1×10⁴ EMT6 cells were subcutaneously implanted into the
130 right flank region of anaesthetised (1.5% v/v halothane for <5 min) 7-week old male
131 Balb/c mice (Harlan). Mice were housed on a 12 hour light – dark illumination
132 schedule and had free access to a standard pellet diet and water. Experiments were
133 performed with UK Home Office approval (PPL 40/2343, PPL 40/2972).

134

135 Mice with EMT6 tumours reaching 300mm³ then received PBS injected intravenously
136 via the tail vein. After 24 h, mice were anaesthetised (1.5% v/v halothane for 30 min)
137 and the tumour region was depilated with hair removal cream prior to laser light
138 treatment. Tumours were treated 24 h later at a light dose of 138 J/cm² with a
139 fluence rate of 75mW/cm² using a 635 nm laser attached to a microlens fibre (23
140 min). Mice were killed by cervical dislocation at 24 h or post treatment. Tumours
141 were excised and fixed (0.1 M Tris buffer, pH 7.4, 0.5 g calcium acetate, and 5.0 g
142 zinc acetate and 5.0 g zinc chloride in 1 L of deionised water) overnight before
143 transferral to 70% v/v industrial methylated spirit, prior to processing for histological
144 analysis. Samples were fixed in 10% buffered formalin for 24 hours, dehydrated in
145 70% EtOH and paraffin embedded. 5 µm sections were cut using a microtome (Leica
146 Microsystems, Bucks, UK) and mounted onto a histological glass slide. Formalin
147 fixed paraffin embedded (FFPE) tissue sections were stored at room temperature
148 until further analysis.

149 (b) Fresh Frozen Mouse Fibrosarcoma Samples

150 Mice were injected sub-cutaneously in the dorsum with a 50 µl tumour cell
151 suspension containing 1×10⁶ cells in serum-free medium. The cells employed in this

152 study were from the mouse fibrosarcoma cell line, VEGF188. This has been
153 engineered to express only the VEGF188 isoform [21]. Tumours were allowed to
154 grow to approximately 500 mm³, before CA-4-P treatment (a single dose of 100
155 mg/kg intraperitoneal). Mice were killed and tumours excised at various times after
156 treatment.

157 (ii) Preparation of Recombinant Positive Control Sample

158 The required DNA sequence was synthesised by MWG-biotech (AG Sequencing
159 Department, Fraunhoferstr, 22 D-82152, Martinsried Germany), and sub cloned into
160 the bacterial expression vector pET23a(+) (Novagen, Merck KGaA, Darmstadt,
161 Germany). The plasmid was then transformed into the *E. coli* strain BL21 DE3
162 (Promega Corporation, Southampton, U K). All bacterial cultures were grown in luria
163 broth (LB) media with 100 µg/mL ampicillin for selection. 10 mL of a 100 mL
164 overnight starter culture was used to inoculate 1 L of LB media. Cultures were grown
165 at 37 °C with shaking until the OD₆₀₀ reached 0.5, protein expression was induced
166 with 1 mM isopropyl-β-D-thiogalactoside (IPTG). The culture was incubated for a
167 further 2.5 h before the bacterial cells were harvested by centrifugation at 10,000 g
168 for 20 min using a Sorvall RC G+ (Thermo Scientific, Hemel Hempstead UK).

169 Bacterial pellets were resuspended in binding buffer (20 mM sodium phosphate +
170 20 mM imidazole pH 7.4) 3mL per 1 g of cells before being lysed by sonication (10
171 ×20 s pulses at a 35% duty cycle) using a Vibra Cell VCX 750 (750 W).
172 Phenylmethyl sulphonyl fluoride (PMSF) (50 µg/mL), DNase (20 µg/mL), RNase (20
173 µg/mL) was added to the lysate which was then cleared of insoluble material by
174 centrifugation at 8,000 g for 1 h using an Eppendorf 5804R centrifuge. The clear
175 lysate was filtered through a 0.2 µm syringe filter and load onto a 1 mL Histrap FF

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

180 (iii) Tissue preparation

181 (a) FFPE Tissue

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

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

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

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

197 Trypsin used in this study, for on-tissue digests, was prepared in 50 mM ammonium
198 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 2 μ l/ minute, using a SunCollect™
202 automatic sprayer (SunChrom, Friedrichsdorf, Germany). Sections were then
203 incubated overnight at 37 °C (5% CO₂).

204 (v) Matrix deposition

205 5 mg/mL α -cyano-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 μ L 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 μ L/min to allow a matrix seeding process. Three subsequent
210 layers were sprayed at 3.5 μ L/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 μ m spatial resolution using variable IMS wave velocity. The

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

224 (vii) Data Processing

225 Data were processed using Waters HD Imaging software release 1.1 (Waters
226 Corporation, Manchester, UK). The parameters used were as follows: specificity type
227 (*i.e.* data type) – IMS MS, Number of most intense peaks – 1,000, resolution –
228 10,000, low energy intensity threshold – 50. The low intensity threshold was set to
229 potentially allow low abundant species to be included.

230

231 **Results and Discussion**

232 Figure 1 shows (a) the MALDI peptide mass fingerprint obtained from an in-solution
233 digest of the recombinant protein (b) the amino acid sequence of the recombinant
234 protein synthesised and (c) a list of "signature" peptides generated and detected
235 from the protein. As can be seen from Figure 1a the $[M+H]^+$ ions from each of the
236 peptides listed in 1c were discernible. Ion intensities of peptides can be very
237 different even when peptides are present in equimolar amounts [23] This
238 phenomenon was also observed for the signature peptides, for example m/z 1168.5
239 (LGIHEDSQNR) from HSP90 alpha was 5 times more intense than the signal at m/z
240 983.4 (CEVGYTGVR) from epiregulin.

241

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

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

259

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

267

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

279

280 In our work examining the response of tumours to vascular disrupting agents we
281 have become interested in the distribution of HSP90 and Plectin in tumour tissue
282 following administration of the vascular disrupting agent combretastatin-4-AP.
283 Therefore the recombinant protein was engineered to contain the signature peptide
284 sequences LGIHEDSQNR ($[M+H]^+$ m/z 1168.5) representing HSP90-alpha,
285 GVVDSLEDLELNISR ($[M+H]^+$ m/z 1513.7) representing HPS90-beta, and
286 AQAELEAQELQR ($[M+H]^+$ m/z 1385.7) and DSQDAGGFGPEDR ($[M+H]^+$ (m/z
287 1350.5) representing Plectin. Figure 3 a-c and d-f show data from the analysis of a
288 section of a fresh frozen mouse fibrosarcoma genetically engineered to express only
289 the VEGF188 isoform (21), 72 hours after administration of combretastatin-4-AP.
290 Figures 3a-c show the distribution of m/z 1168.5 isolated by m/z and drift time as
291 previously described. The m/z and drift time were optimised using signal from
292 recombinant standard with sequence LGIHEDSQNR observed in the reference spot.

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

298

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

311

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

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

329 **Conclusions**

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

340

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

346

347 **Acknowledgements**

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

351 **References**

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

354 2. R. M. Caprioli, T. B. Farmer and J. Gile. Molecular imaging of biological samples:
355 localization of peptides and proteins using MALDI-TOF MS *Anal. Chem.* 1997; **69**,
356 4751.

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

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

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

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

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

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

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

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

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

- 384 12. M. R. Groseclose , M. Andersson , W. M. Hardesty , R. M. Caprioli.
385 Identification of proteins directly from tissue: *in situ* tryptic digestions coupled with
386 imaging mass spectrometry. Journal of Mass Spectrometry. 2007, **42**, 254.
- 387 13. M-C. Djidja , S. Francese , P. M. Loadman, C. W. Sutton, P. Scriven , E.
388 Claude , M. F. Snel, J. Franck , Salzet M, Clench MR. "Detergent addition to tryptic
389 digests and Ion Mobility separation prior to MS/MS improves peptide yield and
390 Protein Identification for *in situ* Proteomic Investigation of Frozen and FFPE
391 Adenocarcinoma tissue sections." Proteomics 2009. **9**, 2750.
- 392 14. M-C. Djidja, E. Claude, M. F. Snel, P. Scriven, S. Francese, V. A. Carolan, M. R.
393 Clench. MALDI-Ion Mobility Separation-Mass Spectrometry Imaging of Glucose-
394 Regulated Protein 78 kDa (Grp78) in Human Formalin Fixed Paraffin Embedded
395 Pancreatic Adenocarcinoma Tissue Sections, J Proteome Res 2009, **8**, 4876..
- 396 15. J. Stauber, L. MacaLeese, J. Franck, E. Claude, M. F. Snel, B. K. Kaletas, I. M.
397 Wiel, M. Wisztorski, I. Fournier , R. M. Heeren. *OnTissue* Protein Identification and
398 Imaging by MALDI-Ion Mobility Mass Spectrometry, J Am Soc Mass Spectrom,
399 2010, **21**, 338.
- 400 16. Y. Schober, T. Schramm, B. Spengler, *et al.*. Protein identification by accurate
401 mass matrix-assisted laser desorption/ionization imaging of tryptic peptides, Rapid
402 Commun. Mass Spectrom. 2011, **25**, 2475.
- 403
- 404 17. O. J. R. Gustafsson, J. S. Eddes, S. Meding, S. R. McColl, M. K. Oehler, P.
405 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. Krosi, J. Krosi , 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

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

431

432 **Legends for Figures and Tables**

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

439 **Figure 2:** MALDI-IMS-MSI Data generated from the Analysis of FFPE EMT6 mouse
440 tumour tissue (a) MALDI-IMS-MSI image for the distribution of m/z 1168.5 selected
441 such that both its m/z and ion mobility drift-time matched that of the peptide
442 LGIHEDSQNR (from HSP 90) produced by in-solution digest of the recombinant
443 IMS-TAG protein, the distribution of the peptide in both the tissue and the positive
444 control spot of the recombinant standard are clearly visible. (b) Partial MALDI mass
445 spectrum of the most abundant 1,000 peaks in imaging data set the generated by
446 the Waters HD Imaging Software (c) Display of drift time vs m/z for the 1,000 most
447 intense peaks in the imaging data set generated by the Waters HD Imaging Software.
448 In figures 2(b) and 2(c) the peak of interest has been highlighted. (d) MALDI-IMS-
449 MSI image for the distribution of m/z 1093.5 selected such that both its m/z and ion
450 mobility drift-time matched that of the peptide FADLSEAANR (from Vimentin)
451 produced by in-solution digest of the recombinant IMS-TAG protein, the distribution
452 of the peptide in both the tissue and the positive control spot of the recombinant
453 standard are clearly visible. (e) Partial MALDI mass spectrum of the most abundant
454 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software
455 (f) Display of drift time vs m/z for the 1,000 most intense peaks in the imaging data
456 set generated by the Waters HD Imaging Software. In figures 3(b) and 3(c) the peak

457 of interest has been highlighted.(g) MALDI-IMS-MSI image for the distribution of m/z
458 944.5 believed to be AGLQFPVGR (from Histone H2A) (b) Partial MALDI mass
459 spectrum of the most abundant 1,000 peaks in imaging data set the generated by
460 the Waters HD Imaging Software (h) Display of drift time vs m/z for the 1,000 most
461 intense peaks in the imaging data set generated by the Waters HD Imaging Software.
462 In figures 4(b) and 4(i) the peak of interest has been highlighted. As can be seen,
463 since this peptide was not included in the recombinant IMS-TAG standard although it
464 is clearly visible in the tissue it is not highlighted in the region of the image covering
465 the spotted recombinant standard.

466 **Figure 3:** MALDI-IMS-MSI Data Generated from the Analysis of a fresh frozen
467 mouse fibrosarcoma genetically engineered to express only the VEGF188 isoform
468 taken 72 hours after the administration of combretastatin-4-AP. (a) MALDI-IMS-MSI
469 image for the distribution of m/z 1168.6 selected such that both its m/z and ion
470 mobility drift-time matched that of the peptide LGIHEDSQNR (from HSP90 alpha)
471 produced by in-solution digest of the recombinant IMS-TAG protein, the distribution
472 of the peptide in both the tissue and the positive control spot of the recombinant
473 standard are clearly visible. (b) Partial MALDI mass spectrum of the most abundant
474 1,000 peaks in imaging data set the generated by the Waters HD Imaging Software
475 (c) Display of drift time vs m/z for the 1,000 most intense peaks in the imaging data
476 set generated by the Waters HD Imaging Software. In figures 5(b) and 5(c) the peak
477 of interest has been highlighted.

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

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

487

488 **Figure 4:** MALDI positive ion product ion spectrum obtained from the ion at m/z
489 1168.5 arising from the peptide LGIHEDSQNR and present in the tryptic digest of
490 the recombinant "IMS-TAG" protein. These data were obtained using the "transfer
491 fragmentation" feature of the Synapt instrument and hence the precursor ion and
492 product ions have the same ion mobility drift-time (inset box).

493

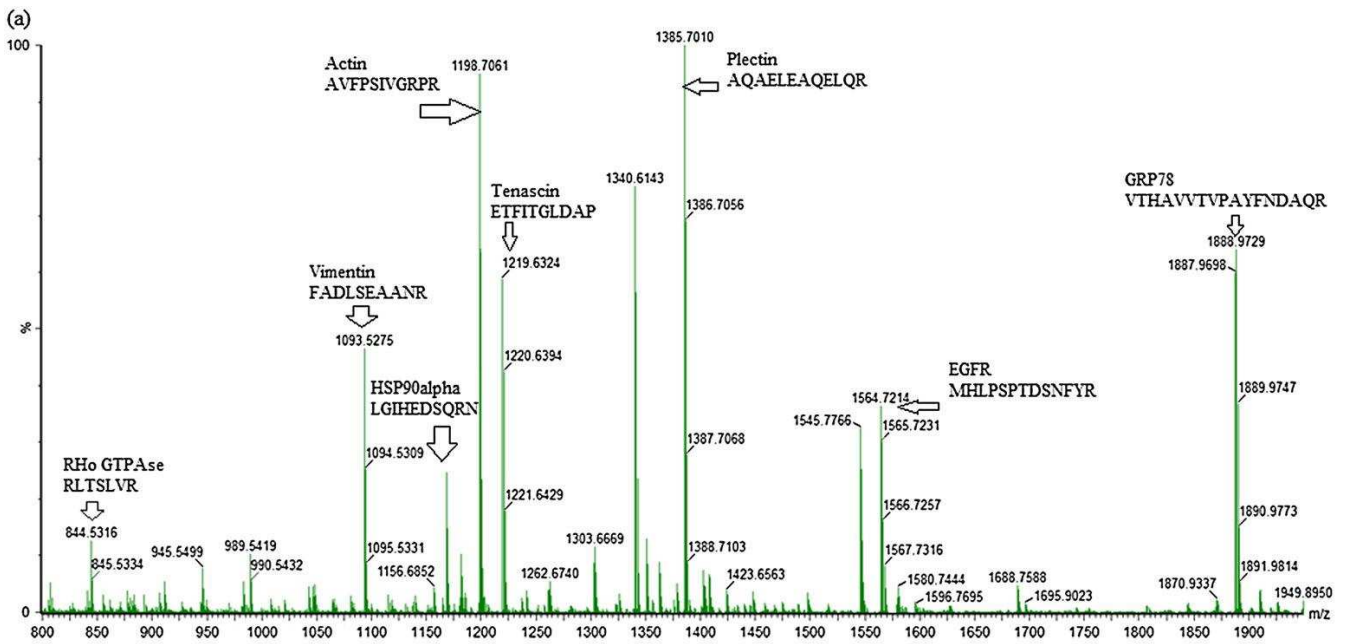
494 **Figure 5:** MALDI Positive Ion product ion spectra obtained from the ion at m/z
495 1198.7 arising from the Actin tryptic peptide AVFPSIVGRPR obtained (a) directly
496 from EMT6 tumour tissue (b) from the recombinant standard. The y_2 ion (m/z 272)
497 is clearly visible in each spectrum and the inset box shows the presence of the y_3
498 (m/z 428) ion. It is suggested that even though these spectra are of low quality the
499 combination of product ion mass and drift time would be sufficient to allow the
500 distribution of this protein to be imaged with confidence.

501

502

503

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	GVVDSLELNSR
Plectin	1385.7	AQAELEAQELQR
Haemoglobin beta chain	1350.5	DSQDAGGFPEDR
	1302.6	VNSDEVGGEALGR
Tenascin	1219.6	ETFITGLD
Actin	1198.7	AVFPSIVGRPR
HSP-90 alpha	1168.5	LGIHEDSQNR
Vimentin	1093.5	FADLSEAAANR
Epiregulin	983.4	CEVGYTGVR
Rho GTPase activating protein 2	844.5	RLTSLVR
Tag reporter		A WLE HHH HHH

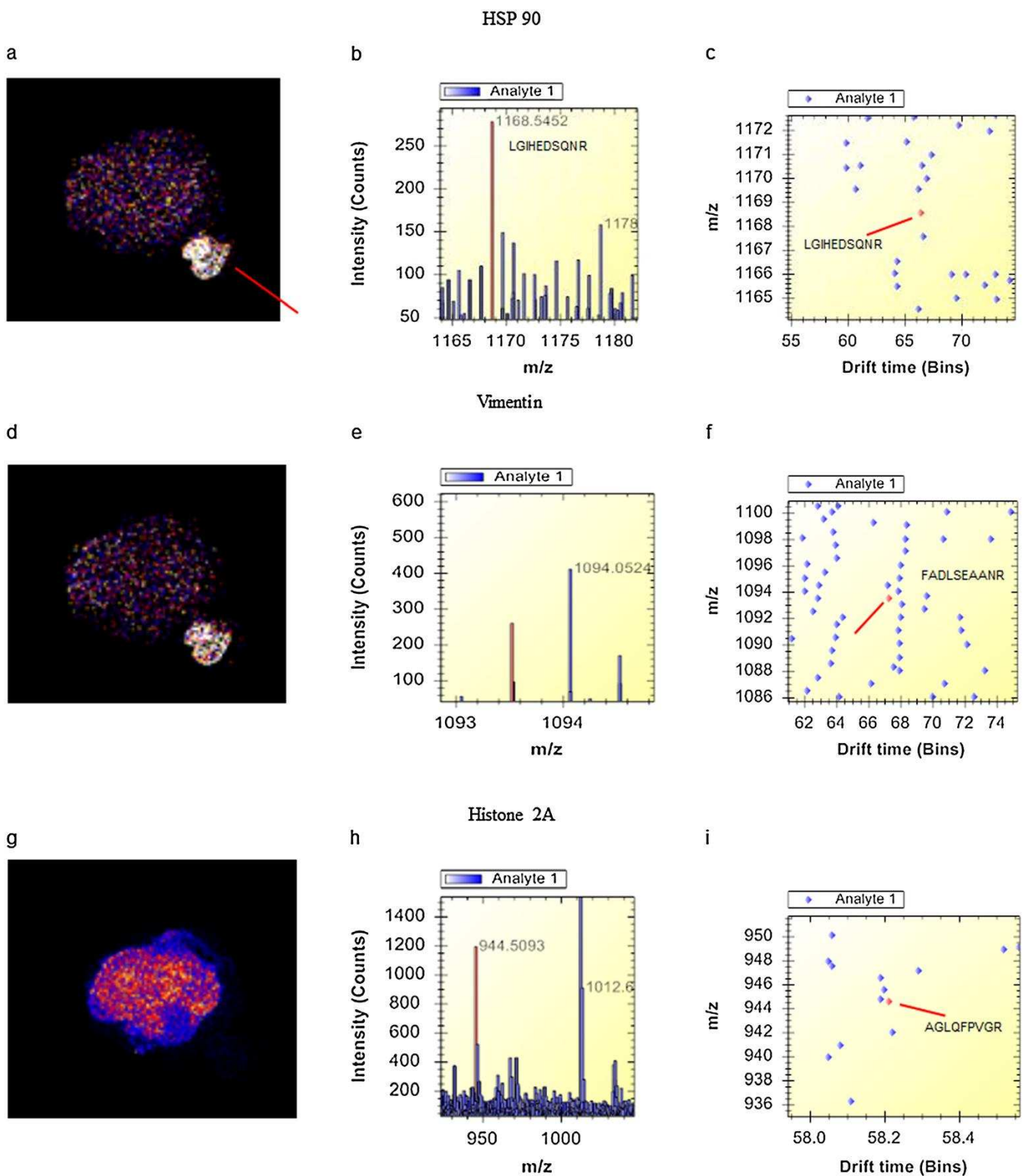
505

506

507

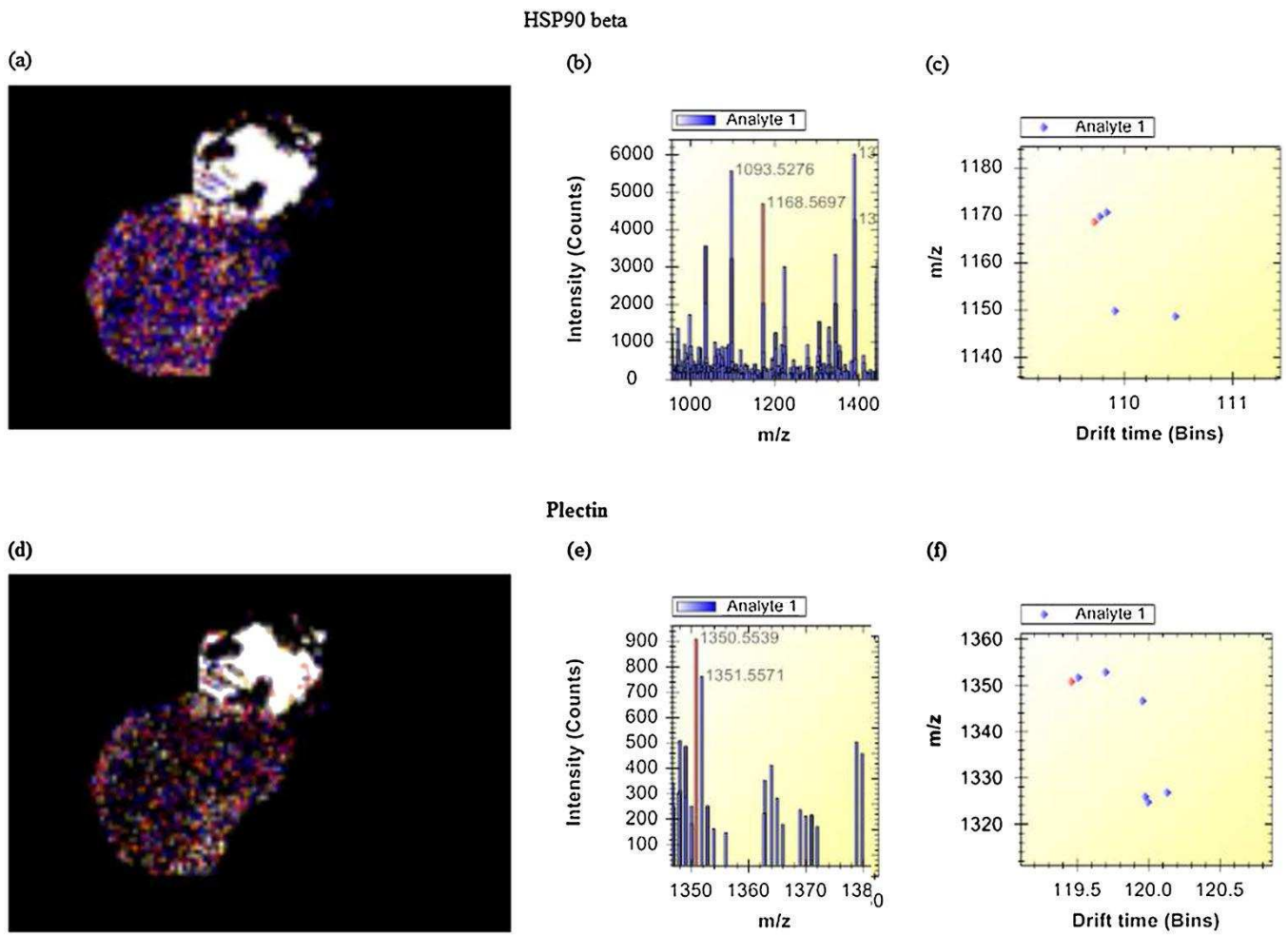
508

509



511

512



514

515

516

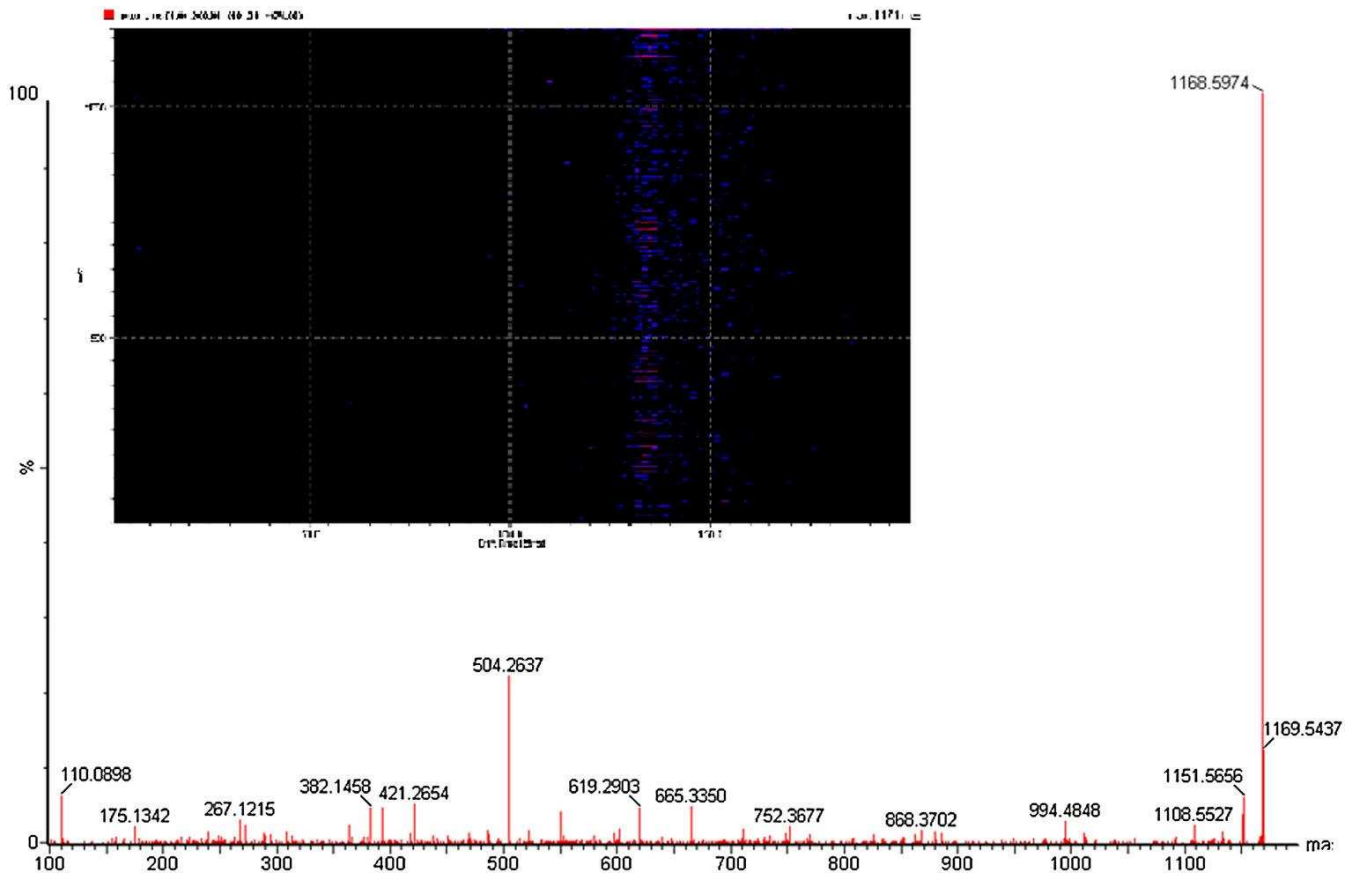
517

518

519

520

521 Figure 4



522

523

524

525

526

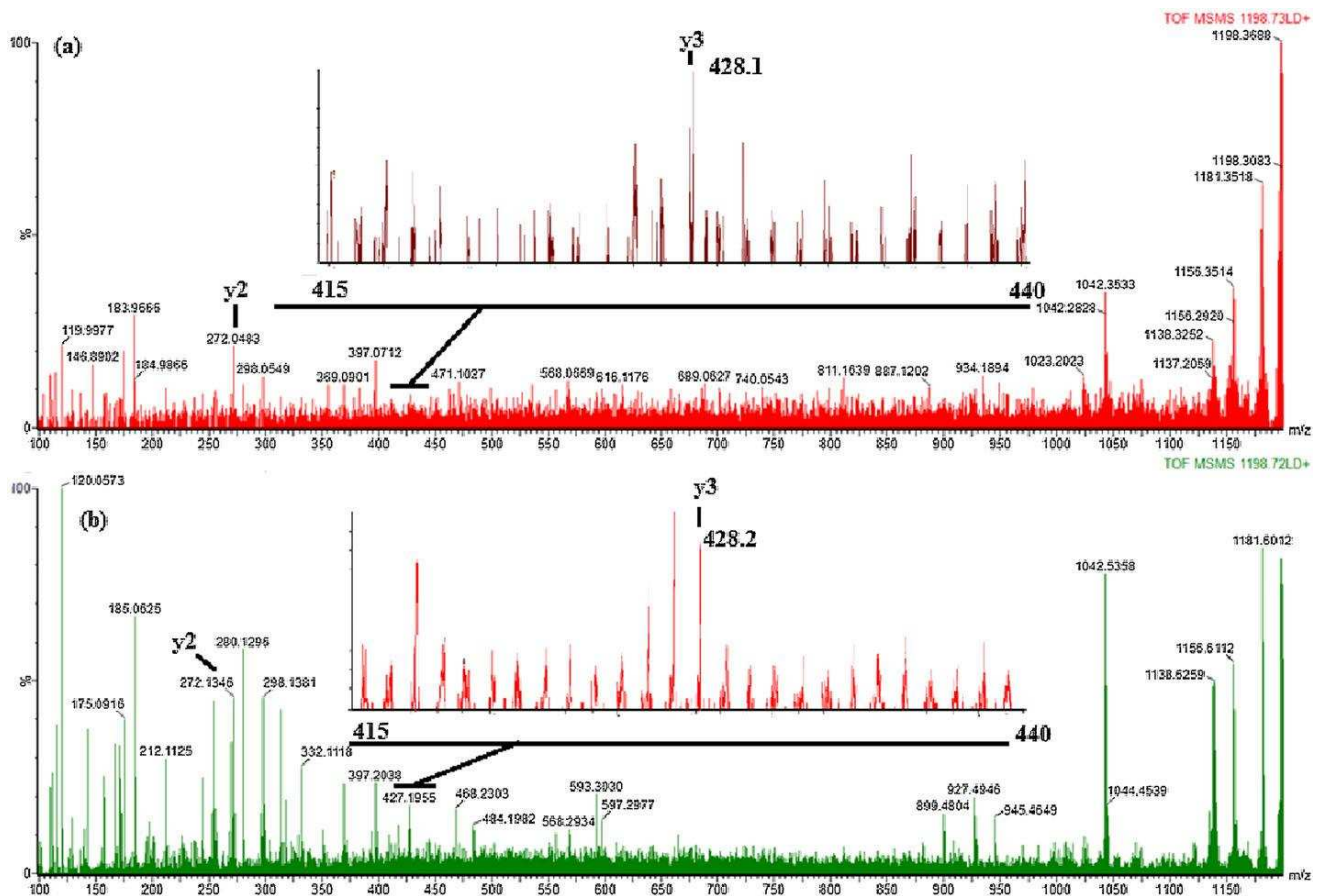
527

528

529

530

531 Figure 5



532

533

534

535

536

537

538

539

540