

Metabolic profiling and imaging of CHO cells for fusion protein production

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2017

Ewa Szula

The University of Manchester
Faculty of Science and Engineering
School of Chemistry

Contents

List of Figures	7
List of Tables	17
List of Equations	21
Abbreviations	22
Abstract	24
Declaration	25
Copyright Statement	25
Acknowledgements	26
1. Introduction	27
1.1. Significance and aims of the study	27
1.2. Metabolomics	28
1.3. Mass spectrometry as a platform for metabolomic studies	30
1.3.1. Sample introduction methods	30
1.3.2. Ionisation sources	33
1.3.2.1. MALDI-MS	33
1.3.2.1.1. Principles of MALDI	33
1.3.2.1.2. Applications of MALDI-MS	35
1.3.2.2. SIMS	36
1.3.2.2.1. Principles of SIMS	36
1.3.2.2.2. Static and dynamic SIMS	38
1.3.2.2.3. Applications of SIMS	40
1.3.3. Mass analysers	41
1.3.4. MS-based metabolic imaging	42
1.4. Metabolomics in practice	45
1.4.1. Fusion proteins	46
1.4.1.1. Fc-based fusion proteins	47
1.4.1.2. EPO-Fc fusion protein	48
1.4.1.3. Mammalian cell line as a platform for protein production	48
1.4.2. Single cell based metabolomics	49

3.2.3.1.1.	Matrix coating.....	98
3.2.3.1.2.	AP-SMALDI10 imaging source.....	99
3.2.3.1.3.	Data analysis.....	101
3.2.3.2.	SIMS imaging.....	102
3.2.3.2.1.	The J105-3D chemical imager.....	102
3.2.3.2.2.	Data analysis.....	103
	References.....	104
4.	Expression of EPO-Fc fusion protein in transfected CHO cell line.....	106
4.1.	Introduction.....	106
4.2.	Sample preparation.....	108
4.3.	Methods.....	108
4.3.1.	Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	108
4.3.2.	Western Blotting.....	110
4.4.	Results and discussion.....	112
4.4.1.	Cell growth curve.....	112
4.4.2.	EPO-Fc fusion protein production.....	115
4.5.	Conclusions and future work	118
	References.....	119
5.	MALDI-MS sample preparation optimisation for metabolite profiling.....	121
5.1.	Introduction.....	121
5.2.	Materials	125
5.3.	Sample preparation.....	125
5.4.	Instrumentation.....	126
5.5.	Data analysis	127
5.6.	Results and discussion.....	128
5.6.1.	Choice of matrix and solvent conditions for MALDI metabolite detection.....	128
5.6.1.1.	Positive ionisation mode.....	128
5.6.1.2.	Negative ionisation mode.....	139
5.6.2.	Metabolite standards.....	147
5.6.3.	Optimisation of sample deposition method	158
5.6.3.1.	Positive ionisation mode.....	158
5.6.3.2.	Negative ionisation mode.....	160

5.6.4. A choice of a top MALDI method of sample preparation – cellular analytes.....	162
5.6.4.1. Positive ionisation mode.....	162
5.6.4.2. Negative ionisation mode.....	165
5.7. Conclusions and future work	169
References.....	170
6. Metabolic profiling approach to investigate transfected CHO cell line response upon EPO-Fc fusion protein production.....	178
6.1. Introduction.....	178
6.2. Sample preparation.....	182
6.3. Instrumentation.....	183
6.4. Data analysis	183
6.5. Results and discussion.....	185
6.5.1. Footprint and fingerprint profile analysis upon 1 and 3 days of the inducer exposure.....	185
6.5.2. Multivariate data analysis.....	188
6.5.3. Footprint and fingerprint metabolite identification	199
6.5.4. Pathway analysis.....	206
6.5.4.1. Mitochondrial response.....	210
6.5.4.2. Nucleotide response.....	213
6.5.4.3. Amino acid response.....	213
6.5.4.4. Lipid response.....	214
6.6. Conclusions and future work	218
References.....	220
7. Metabolic imaging of transfected CHO cells.....	225
7.1. Introduction.....	225
7.2. Sample preparation.....	229
7.3. Instrumentation.....	229
7.4. Data analysis.....	230
7.5. Results and discussion.....	231
7.5.1. MSI of transfected CHO cells under inducer exposure and EPO-Fc protein production.....	231
7.5.1.1. Cell population-based study.....	231
7.5.1.2. Single cell-based study.....	255

7.5.2. Imaging of transfected CHO cells - single cell and sub-cellular approach	261
7.5.3. Heterogeneity of transfected CHO cell population	269
7.6. Conclusions and future work	277
References	280
8. General discussion, conclusions and future work	287
References	292
Supplementary Information	293
References	321

Word count: 60 972

List of figures

CHAPTER 1: Introduction

Figure 1: Central dogma of molecular biology, where a general information flow goes from genes to mRNA transcripts to protein products that affect metabolic pathways, which generate changes in the function or phenotype of the organism. Interactions between 'omics organisation with many feedback-loops within them are also visualised.28

Figure 2: Schematic of example techniques applied in metabolomics based on instrument set-up, instrument possibilities and experimental approach. Adopted from http://www.rsc.org/images/MS2new_tcm18-102519.pdf.32

Figure 3: Schematic of the desorption-ionisation in MALDI process. Laser pulses irradiate the sample causing desorption and ionisation of multiple species.35

Figure 4: Schematic of the desorption-ionisation process in SIMS. The primary ion beam bombards the sample surface, causing in desorption of multiple species. The majority of species emitted are neutral atoms or molecules; a small part of emitted species is charged.37

Figure 5: Two operational modes of SIMS: static mode gives information about the topmost atomic layer of the sample (A), whereas dynamic mode operates with high primary ion doses to remove material to collect 3D information resulting in the creation of an etched crater (B).39

Figure 6: Snapshots of the C_{60}^{+} and Au_3^{+} impacting the sample surface (here: water ice). The C_{60}^{+} ion beam creates a crater with minimal sub-surface perturbations (A); the Au_3^{+} ion beam has a much longer energy track and hence, extensive damage is caused (B). Adopted and reproduced from (Ryan, Wojciechowski and Garrison, 2007).40

Figure 7: Data analysis workflow in metabolomics. Metabolomics investigations involve a number of steps performed before the pre-defined hypothesis is answered and include data pre-processing, pre-treatment, processing and interpretation.53

Figure 8: Examples of univariate (UVA) and multivariate (MVA) approaches in data analysis for metabolomics. Commonly used methods for UVA are *t*-test and ANOVA. MVA can be divided into unsupervised and supervised methods that include *e.g.* PCA, HCA (unsupervised) and *e.g.* DFA, PLS-DA, OPLS-DA (supervised) analyses (Gonzalez-Riano, Garcia and Barbas, 2016, Gromski *et al.*, 2014).56

Figure 9 A: Application of multivariate analysis in metabolomics studies. Example A: discriminating phospholipids from different types of bacteria using MALDI-MS; adopted and reproduced from (Ishida *et al.*, 2002).58

Figure 9 B: Application of multivariate analysis in metabolomics studies. Example B: investigation of biomarkers for hepatitis C virus from urine samples of hepatitis C virus patients using UPLC-MS; adopted and reproduced from (Zhang *et al.*, 2013).59

CHAPTER 2: Experimental

Figure 1: Schematic components of CHO-K1 T-REx EPO-Fc-pcDNA5-FRT-TO cell line used for metabolomics studies of EPO-Fc fusion protein expression. When induced, the binding between tetracycline (TET) and its repressor activates EPO-Fc promoter for EPO-Fc fusion protein production.81

Figure 2: Schematic structure of EPO-Fc fusion protein. EPO structure incorporated from Cheetham *et al.* 1998.81

Figure 3: A diagram of main stages in the footprint and fingerprint sample collection.

85

CHAPTER 3: Instrumentation

Figure 1: A schematic of a standard mass spectrometry platform. Samples are inserted to an ion source (atmospheric or vacuum pressures) through an inlet system, where ions are produced. Ions travel through mass analyser being separated according to their mass-to-charge (m/z) ratio prior the detection at the detector component. A PC is required to acquire experiments and to control system and all its parameters (Dunn 2008). 90

Figure 2: Scheme of GC-ToF-MS: sampling (1), gas chromatograph (2), electron source (3), ion focusing optics (4), ToF reflectron analyser (5) (http://cires1.colorado.edu/jimenez/CHEM-5181/Lect/2013_GCxGC_Handouts.pdf, <http://www.speciation.net/Database/Instruments/Leco-Inc/Pegasus-III-GCTOFMS-;i509>). 92

Figure 3: Scheme of MALDI Ultraflex II ToF/ToF mass spectrometer: target plate (1), electrodes (2), lens arrangement (3), Collision Induced Dissociation (CID) cell (4), timed ion selector 1 (ToF1) (5), LIFT source (6), Post Lift Metastable Suppressor (PLMS) (7), ToF 2 analyser for linear and reflectron detection (8) (Suckau et al. 2003, http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf). 94

Figure 4: Scheme of MALDI Synapt G2-Si HD mass spectrometer: target plate (1), hexapole ion guide (2), T-Wave ion guide (3), quadrupole (4), Tri-Wave: trap (5), ion mobility separation in helium cell (6), transfer (7), QuanToF: high field pusher (8), ion mirror (9) (<http://www.waters.com>). 96

Figure 5: SMALDIPrep matrix coating device (TransMIT GmbH, Giessen, Germany). Adapted from (<http://www.uni-giessen.de/faculties/f08/departments/iaac/spengler/research/instrumentations/transmit-smaldiprep>). 98

Figure 6: Scheme of AP-SMALDI imaging source connected to Q Exactive mass spectrometer: S-lens (1), injection flatapole (2), bent flatapole (3), quadrupole mass filter (4), octopole (5), C-trap (6), orbitrap (7), HCD collision cell (8) (http://planetorbitrap.com/q-exactive#.WCGx0_mLRpg, Koestler et al. 2008). 100

Figure 7: Scheme of J105-3D chemical imager: sample stage (1), quadrupole (2), electrostatic analyser (3), buncher (4), collision cell (5), detector (6), ToF reflectron (7) (Fletcher et al. 2008). 102

CHAPTER 4: Expression of EPO-Fc fusion protein in transfected CHO cell line

Figure 1: Schematic of western blotting technique: a gel electrophoresis is performed to separate proteins and is followed by a transfer of these proteins from the gel to nitrocellulose membrane using electric current. Once the proteins are transferred onto the membrane, target protein detection is completed using specific protein antibodies. 106

Figure 2: Color prestained protein standard, broad range (11-245 kDa). Adopted from (<https://www.neb.com/products/p7712-color-prestained-protein-standard-broad-range-11-245-kda>) (A). Schematic illustration of the sample organisation on the SDS-PAGE gel (B). 110

Figure 3: A schematic Western Blot configuration of EPO-Fc detection: EPO-Fc protein transferred from gel electrophoresis onto nitrocellulose membrane is recognised by primary antibody; secondary antibody that binds to primary antibody and EPO-Fc protein is conjugated with IR-dye allowing the construct for detection. 111

Figure 4: Transfected CHO cells growth curves (green) and cell viability curves (purple) in the presence ('induced') and absence ('control') of the EPO-Fc protein inducer. The

error bars represent the standard deviation calculated from the data obtained from three experiments (n=3).	113
Figure 5: Western Blots of medium samples removed from control cells (control) and cells induced for production of EPO-Fc (induced). Each blot contains a protein standard ladder with 80 kDa and 58 kDa bands marked and positive control of purified EPO-Fc (~70 kDa).	115
CHAPTER 5: MALDI-MS sample preparation optimisation for metabolite profiling	
Figure 1: The most common ways of sample deposition used in MALDI experiments: underlay (A), mixed (B) and overlay (C) deposition.	123
Figure 2: MALDI positive ion mode averaged spectra of pure (with no sample) matrices of 8 top methods of sample preparation; * - high intense matrix peaks.	138
Figure 3: MALDI negative ion mode averaged spectra of pure (with no sample) matrices of 8 top methods of sample preparation; * - the most abundant matrix peak at m/z 193.	146
Figure 4: MS spectrum of the calibration cocktail in MALDI positive ionisation mode, mass range m/z 50 – 1000 shown, * - matrix peaks.	155
Figure 5: MS spectrum of PEG600 sulfate in MALDI negative ionisation mode, mass range m/z 50 – 1000 shown, * - matrix peaks.	156
Figure 6: A comparison of intensity of ions derived from compounds creating the calibration mixture according to the sample deposition method used in MALDI positive ion mode. Expanded view of low intense ions (A). The error bars represent the standard deviation calculated from the data obtained from four experiments (n=4).	159
Figure 7: A comparison of the intensity of m/z 317 and m/z 625 ions for 5 top methods of sample preparation for MALDI negative ionisation mode. The error bars represent the standard deviation calculated from the data obtained from four experiments (n=4).	160
Figure 8: MALDI positive ion mode PC scores 1 and 2 for top sample preparation methods: DHB with mixed sample deposition (A), DHB with overlay sample deposition (B), THAP with overlay sample deposition (C); PC1 accounts for 89.7%, 93.6% and 95.1% of total variance, respectively. PCA plots illustrate the variance between fresh medium (blue) and used medium removed on day 1 (red). Loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da; n = 4, outliers removed (one repeat from B). Sample spots observed from built-in MALDI camera (magnification unknown); laser pathway seen.	163
Figure 9: A comparison of mass resolution for ‘low’ and ‘high’ mass range ions for DHB with mixed sample deposition in MALDI positive ion mode. The error bars represent the standard deviation calculated from the data obtained from four experiments (n = 4).	165
Figure 10: MALDI negative ion mode PC scores 1 and 2 for top sample preparation methods: 9AA in 70% ACN (A), 9AA in 70% EtOH (B), 9AA in 70% MeOH (C), PC1 accounts for 95.3%, 89.0% and 93.1% of total variance, respectively. PCA plots illustrate the variance between fresh medium (red) and used medium removed on day 1 (blue). Loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da; n = 4, outliers removed (one repeat from A and C). Sample spots observed from built-in MALDI camera (magnification unknown); laser marks seen.	166
Figure 11: A comparison of mass resolution for ‘low’, ‘medium’ and ‘high’ mass range ions for 9AA dissolved in: 70% ACN (A), 70% EtOH (B) and 70% MeOH (C) in MALDI negative ion mode. The error bars represent the standard deviation calculated from the data obtained from four experiments (n = 4).	167

CHAPTER 6: Metabolic profiling approach to investigate transfected CHO cell line response upon EPO-Fc fusion protein production

- Figure 1:** General view of an animal cell with cytoplasm enclosed within a cell membrane, containing a variety of organelles (Klyszejko-Stefanowicz, 2002).178
- Figure 2:** General overview of a complexity of a metabolic network in CHO cells. Metabolic network consists of major metabolic pathways highlighted in the figure. Adopted from (<http://www.genome.jp/kegg/pathway.html>).179
- Figure 3:** Carbohydrate and energy metabolic pathways highlighting the major processes occurring in cells: glycolysis (A), oxidative decarboxylation (B) and TCA cycle (C). Carbohydrate and energy pathways are directly (in blue) and indirectly linked to other metabolic pathways.180
- Figure 4:** Schematic of a cell membrane. A cell membrane consists of proteins and three major classes of lipids: phospholipids, glycolipids and cholesterol that create lipid bilayer (Klyszejko-Stefanowicz, 2002, Bretscher, 1972).181
- Figure 5:** Averaged MALDI-MS spectra in positive and negative ion modes from control (CTRL, blue) and cell-induced (IND, red) footprint samples according to inducer exposure time; m/z 20-1000, *-matrix peaks.185
- Figure 6:** Averaged MALDI-MS spectra in positive and negative ion modes from control (CTRL, blue) and cell-induced (IND, red) fingerprint samples according to inducer exposure time; m/z 20-1000, *-matrix peaks.186
- Figure 7:** Positive ion mode GC-MS metabolite footprint analysis: PCA scores plots (A) accounting for 74.8 % and 70.0 % of the total variance from the data set and PLS-DA plots (B) accounting for 58.6% and 54.2% of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores (see text for details) from Component 1 (C) show the 15 most statistically significant variables (metabolite IDs) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.191
- Figure 8:** Positive ion mode GC-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 76.9 % and 79.2 % of the total variance from the data set and PLS-DA plots (B) accounting for 75.6% and 72.0% of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (metabolite IDs) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.192
- Figure 9:** Positive ion mode MALDI-MS metabolite footprint analysis: PCA scores plots (A) accounting for 53.2 % and 57.6 % of the total variance from the data set and PLS-DA plots (B) accounting for 35.5 % and 43.9 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.193
- Figure 10:** Negative ion mode MALDI-MS metabolite footprint analysis: PCA scores plots (A) accounting for 74.6 % and 76.3 % of the total variance from the data set and PLS-DA plots (B) accounting for 71.2 % and 71.8 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C)

show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.194

Figure 11: Positive ion mode MALDI-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 53.0 % and 57.5 % of the total variance from the data set and PLS-DA plots (B) accounting for 48.8 % and 49.9 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.196

Figure 12: Negative ion mode MALDI-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 81.5 % and 73.4 % of the total variance from the data set and PLS-DA plots (B) accounting for 76.7 % and 66.0 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.197





Figure 13: Reconstructed biochemical map summarising correlations between metabolic pathways influenced by the protein production in CHO cells. The summary includes external and internal metabolome data from both time points of cell culture. Metabolic pathways are divided into 6 main metabolism paths (colour-coded). Metabolites influenced by the protein production process are highlighted with visualised abundance: high  and low . Two arrows with the same metabolite indicate changing abundance depending on the type of metabolome (external or internal). Metabolites placed in boxes are found in external (footprint) metabolome.208

Figure 14: Diagram of identified metabolites and their metabolic pathways (colour-coded) with a division to external and internal metabolome. The diagram includes data from both time points of cell culture and metabolic pathways with impact value >0.1. Metabolites influenced by the protein production are highlighted with visualised abundance: high  and low . Two arrows with the same metabolite indicate changing abundance depending on the type of metabolome (external or internal).209

CHAPTER 7: Metabolic imaging of transfected CHO cells

Figure 1: Basic workflow of cell-based MSI experiment. Cells are deposited onto a sample support followed by washing steps (A), the ionisation source rasters across the sample in an X and Y coordinate manner providing with mass spectrum for each pixel (B), data analysis provides with an image of the cells illustrating the spatial distribution of m/z values recorded (C).225

Figure 2: Averaged MALDI-MSI spectra from control (CTRL, blue) and induced cell areas (IND, red) for positive and negative ion modes; m/z 250-1000 for positive ion and m/z 280-1000 for negative ion, *-matrix peaks (A). (Inset) 2D and 3D PCA scores plots accounting for 46.1% and 51.0% of total variance (first three components) for positive and negative ionisation modes, respectively (B). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence.234

Figure 3: Averaged SIMS+ imaging spectra from control (CTRL, blue, n=3) and induced cell areas (IND, red, n=5) for the 1st layer ('top layer') and layers 2 to 10 ('lower layers'),

m/z 30-900 (A). 2D PCA scores plots accounting for 61.2% and 85.2% of total variance (first two components) for the top and lower layers (B). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence. 3D PLS-DA scores plots accounting for 65.6% and 90.4% of total variance (first three components) for the top and lower layers (C); 40 keV C_{60}^+ ion source.

235

Figure 4: SIMS images of PC head group m/z 184.06 within top layer and lower layers of control ('CONTROL') and tetracycline-induced ('INDUCED') CHO cells. The abundance of m/z 184 ion for each sample within the top and lower layers is shown. Scale bar 50 μm , pixel size $\sim 8 \mu m$; 40 keV C_{60}^+ ion source.

237

Figure 5: MALDI-MS images of control and induced cells for both positive and negative ionisation modes. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red) for positive ion mode is shown; max. count per pixel: MALDI+, CONTROL: m/z 798.54077 (2.10×10^3), m/z 782.56689 (1.05×10^3) and m/z 760.5849 (2.98×10^3); INDUCED: m/z 798.54077 (4.66×10^3), m/z 782.56689 (1.18×10^3) and m/z 760.5849 (4.51×10^3). Single ion at m/z 885.54974 for negative ion mode is shown; max. count per pixel: MALDI-, CONTROL 3.26×10^3 ; INDUCED 1.25×10^3 . Scale bar 50 μm , pixel size 5 μm .

239

Figure 6: Distribution of possible lipid ions within control and induced cell populations. Representative single ion images of Cl, SM and PC in MALDI positive ion mode are shown. Maximum count per pixel (mc) for each mass is displayed. Scale bar 50 μm , pixel size 5 μm .

245

Figure 7: Distribution of ADP/dGDP/adenosine 3'5'-diphosphate within control and induced cell population. PE ion is also shown to highlight possible cell membranes. Metabolite assignments are colour-coded: m/z 428.0368 (red), m/z 798.5407 (green). Max. count per pixel: MALDI+, CONTROL: m/z 428.0368 (1.51×10^3), m/z 798.5407 (2.10×10^3); INDUCED: m/z 428.0368 (9.02×10^3), m/z 798.5407 (9.02×10^3). Scale bar 50 μm , pixel size 5 μm .

246

Figure 8: SIMS+ spectra from the 1st (top layer) and 2-10th (lower layers) for control (CTRL, blue) and tetracycline-induced cells (IND, red) within a mass range of 700-800 Da; 40 keV C_{60}^+ beam. Different abundance of lipid species between control and induced cells is shown.

252

Figure 9: Representative SIMS images of statistically significant metabolite fragments detected in control and induced CHO cells. Images illustrate the location of each ion within the area imaged. Maximum count per pixel (mc) for each mass is displayed. Scale bar 50 μm , pixel size $\sim 8 \mu m$, 40 keV C_{60}^+ beam.

254

Figure 10: MALDI-MS images of control and induced cells in positive ionisation mode. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red) is shown; max. count per pixel: MALDI+, CONTROL: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3); INDUCED: m/z 798.54077 (6.74×10^3), m/z 782.56689 (4.56×10^3) and m/z 760.5849 (2.18×10^4). Scale bar 50 μm , pixel size 5 μm .

255

Figure 11: Single cell analysis of CHO control and induced cells for MALDI imaging in positive ion mode. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red). Enlarged single cells selected for multivariate analysis are shown. Scale bar 50 μm (zoom out image), pixel size 5 μm . (Inset) 2D and 3D PCA scores plot accounting for 48.1% of total variance (first three components). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence.

257

Figure 12: Representative MALDI+ single ion images of the most significant metabolites from single cell-based analysis of control and induced cells. A single cell highlighted from control and induced cell populations is enlarged and inset into ion images to illustrate location differences between both cell populations. Max. count per pixel: CONTROL: m/z 759.5724 (1.09×10^3), m/z 796.5251 (3.10×10^3), m/z 917.5387 (1.51×10^3); INDUCED: m/z 759.5724 (1.70×10^3), m/z 796.5251 (1.36×10^3), m/z 917.5387 (4.50×10^2). Scale bar 50 μm (zoom out image), pixel size 5 μm . 260

Figure 13: Total ion count (TIC) SIMS+ spectra and TIC images from control CHO cells from three different areas (A, B, C) within the same sample. Attached table shows instrument settings and mass range of each area. Scale bar 10 μm (green, left top corner of each image), pixel size 0.5 μm ; 25 keV Au_3^+ beam. 262

Figure 14: SIMS images of PC head group m/z 184 and its fragments of CHO cells. PC head group at m/z 184.1, fragments at m/z 86.1 and m/z 104.1 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam. 264

Figure 15: SIMS images of lipid fragments within transfected CHO cells. Fragments at m/z 165.1, m/z 166.1, m/z 189.1 and m/z 369.4 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , spatial resolution <1 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam. 265

Figure 16: SIMS images of amino acid fragments within CHO cells. Fragments at m/z 120, m/z 130, m/z 70 and m/z 81 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam. 265

Figure 17: SIMS images of nitrogenous bases' fragments within CHO cells. Fragments at m/z 119, m/z 127 and m/z 135 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam. 266

Figure 18: A comparison of MALDI and SIMS images of sub-cellular compartments of CHO cells in positive ion mode. MALDI species of PE (m/z 798.54077, m/z 782.56689) and DG (m/z 760.5849) and SIMS fragments of PC (m/z 104.1), MAG (m/z 313.3) and DG (m/z 576) are shown. Max. count per pixel, MALDI+: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3); SIMS+: m/z 104.1 (9701), m/z 313.3 (256) and m/z 576 (247). Scale bar 50 μm (zoom out image MALDI) and 10 μm (SIMS), pixel size 5 μm (MALDI) and 0.5 μm (SIMS). 267

Figure 19: Analysis of heterogeneity of CHO cell population for MALDI+. Image shows overlay of three lipid species: PE(22:5/18:1) at m/z 798.54077, PE(16:1/22:6)/PE(18:1/20:5) at m/z 782.56689, DG(22:2/0:0/20:4) at m/z 760.5849 within control cell population. Enlarged single cells with different metabolite profiles are shown (colour-coded): type A (red), type B (green) and type C (blue) and were selected for multivariate analysis ($n=5$ for each type). Max. count per pixel: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3). Scale bar 50 μm (for zoom out image), pixel size 5 μm . 269

Figure 20: A comparison of the averaged MALDI-MSI spectra from three different metabolite profiles: type A (red), type B (green) and type C (blue) (A); m/z 250-1000, *-matrix peaks. (Inset) 2D and 3D PCA scores accounting for 81.8% of total variance (first three components) (B). PCA plots illustrate the variance between analysed types of metabolic profiles with highlighted 95% regions of confidence. Heatmap shows heterogeneity and diversity of compounds within each type of metabolite profiles across whole analysed mass range (m/z 250-1000) (C). 270

Figure 21: A location of cells with different type of metabolite profile within cell population in MALDI+. Single ion images correspond to metabolite profiles with the highest abundance of each ion (colour-coded): type A (red) m/z 605.5504, type B (green) m/z 719.5416, type C (blue) m/z 465.9930. Box and Whisker plots (insets) show the abundance of ions within each type of metabolite profile. Metabolite profiles type A, B and C are illustrated in overlay image (colour-coded). Max. count per pixel: m/z 605.5504 (5.66×10^2), m/z 719.5416 (6.60×10^2) and m/z 465.9930 (2.44×10^3). Scale bar 50 μm , pixel size 5 μm . 274

Figure 22: Principal Component Analysis of cell-to-cell heterogeneity within transfected CHO cell population for SIMS+. Multivariate analysis resulted with PC2 and PC3 scores images and correlating PC2 and PC3 loadings plots. PC2 and PC3 scores images show the abundance of ions displayed in PC2 and PC3 loadings plots within cell population. The most abundant ions within PC2 and PC3 loadings plots are highlighted and representative images of these ions are shown. Maximum count per pixel (mc) is displayed. Scale bar 10 μm , pixel size 0.5 μm , m/z 70-1000. 276

SUPPLEMENTARY INFORMATION

Figure 1: A comparison of total intensity of the calibration mixture in DHB (A) for MALDI positive ion mode and PEG600 sulfate in 9AA (B *a, b*) for negative ion mode according to the laser power used. The error bars represent the standard deviation calculated from the data obtained from two experiments ($n=2$). 293

Figure 2: A comparison of S/N ratio of Arg and His ions in based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 294

Figure 3: A comparison of mass resolution of Arg and His ions based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 295

Figure 4: A comparison of the intensity of Arg and His ions based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 296

Figure 5: A comparison of S/N ratio of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 297

Figure 6: A comparison of mass resolution of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 298

Figure 7: A comparison of the intensity of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only. 299

Figure 8: A comparison of S/N ratio of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only. 300

Figure 9: A comparison of mass resolution of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.	301
Figure 10: A comparison of the intensity of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.	302
Figure 11: A comparison of S/N ratio of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.	303
Figure 12: A comparison of mass resolution of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.	304
Figure 13: A comparison of the intensity of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.	305
Figure 14: MALDI negative ion mode MS/MS spectrum of PC (18:0/0:0) with the parent ion at m/z 523 and the fragment ion at m/z 464; close-up of fragment ion $[M-N(CH_3)_3-H]^-$ (B).	306
Figure 15: A comparison of mass resolution of ions at m/z 104 and m/z 456 for fresh medium and medium samples removed from cell culture across 6 and 20 dilution factors. Data obtained from two experiments ($n = 2$); MALDI positive ion mode.	311
Figure 16: MALDI positive ion mode averaged spectra for fresh medium and medium samples removed from cell culture in three dilutions: 2x (blue), 6x (red) and 20x(pink); mass range 20-1000 Da.	312
Figure 17: A comparison of cell growth at different cell densities during 72 hours of incubation time. Densities at 0.02×10^6 cells/mL, 0.04×10^6 cells/mL and 0.14×10^6 cells/mL were not suitable due to a cell loss, changes in cell morphology, too low or too high cell density. Density at 0.07×10^6 cells/mL was therefore selected as the best.	314
Figure 18: PCA scores of a comparison of metabolite leakage from CHO cells during different fixation processes. Samples from different stages of fixation processes were compared to pure (non-used) fixation solutions and analysed in MALDI <u>positive</u> ionisation mode. Three analytical replicates of each sample are shown.	317
Figure 19: PCA scores of a comparison of metabolite leakage from CHO cells during different fixation processes. Samples from different stages of fixation processes were compared to pure (non-used) fixation solutions and analysed in MALDI <u>negative</u> ionisation mode. Three analytical replicates of each sample are shown.	318
Figure 20: A comparison of morphology of CHO cells after snap-freezing in liquid nitrogen and drying in desiccator. Swollen cells with different morphology to normal growing cells (image in a middle, 'CHO cells') were seen after liquid nitrogen treatment.	319

Figure 21: The analysis of the ratio of m/z 184 ion (organic origin) and m/z 168 (silicon) across all layers detected by SIMS. Content of m/z 184 differs for the 1st layer and is similar within layers 2-10. Different colours were used to differentiate samples.....319

Figure 22: PCA scores plots generated for the full SIMS data set for the top (1st) and lower (2-10th) layers and for both control (CTRL) and induced cell (IND) data sets, illustrating no separation between these two groups.....320

List of tables

CHAPTER 1: Introduction

Table 1: Metabolomics applications in different branches of studies.....	45
---------------------------------------------------------------------------------	----

Table 2: Mass spectrometry techniques applied in metabolomics studies for profiling and imaging of small molecules at the single-cell level.....	52
---------------------------------------------------------------------------------------------------------------------------------------------------------	----

CHAPTER 2: Experimental

Table 1: Complete growth cell medium for CHO cell line.....	82
--------------------------------------------------------------------	----

CHAPTER 3: Instrumentation

Table 1: SIMS experiments performed using J105-3D chemical imager, settings for each experiment are included.....	103
--------------------------------------------------------------------------------------------------------------------------	-----

CHAPTER 4: Expression of EPO-Fc fusion protein in transfected CHO cell line

Table 1: Formulation used for preparation of SDS-PAGE gels.....	109
------------------------------------------------------------------------	-----

CHAPTER 5: MALDI-MS sample preparation optimisation for metabolite profiling

Table 1: Chemical structures of DHB, THAP and CHCA with their most common applications for MALDI-MS analysis. Adopted from (http://www.sigmaaldrich.com/).	128
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 2: 27 different combinations of MALDI positive ion sample preparation with presence (+) or absence (-) of ions of interest. Thus, table summarises matrix, its concentration and solvent used in the study. Detection when S/N>6.....	132
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 3: Comparison of S/N ratio of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of S/N ratio (orange) and top scores* (green) are highlighted; empty box – no ion detected.....	134
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 4: Comparison of the intensity [a.u.] of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of the intensity (orange) and top scores* (green) are highlighted; empty box – no ion detected.	135
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 5: Comparison of mass resolution of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of mass resolution (orange) and top scores* (green) are highlighted; empty box – no ion detected.	136
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 6: Summary of the eight top methods of MALDI positive ion matrix preparation, including matrix, its concentration and solvent. Methods were selected based on the analytes detection, their S/N ratio, intensity and mass resolution. Methods listed by matrix, its concentration (from low to high) and type of solvent.....	137
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 7: Chemical structures of 9AA and DMAN with their most common applications for MALDI-MS analysis. Adopted from (http://www.sigmaaldrich.com/).	139
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 8: 30 different combinations of MALDI negative ion sample preparation with presence (+) or absence (-) of ions of interest. Thus, table also summarises matrix, its concentration and solvent used in the study. Detection when S/N>6.....	141
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 9: Comparison of S/N ratio of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of S/N ratio (orange) and top scores* (green) are highlighted; empty box – no ion detected.....	142
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Table 10: A comparison of the intensity [a.u.] of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of intensity (orange) and top scores* (green) are highlighted; empty box – no ion detected.

143

Table 11: A comparison of mass resolution of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of mass resolution (orange) and top scores* (green) are highlighted; empty box – no ion detected.

144

Table 12: A summary of 8 top methods of MALDI negative ion matrix preparation, including matrix, its concentration and solvent. Methods were selected based on the analytes detection, their S/N ratio, intensity and mass resolution. Methods listed by matrix concentration (from low to high) and type of solvent.

145

Table 13: A list of standards examined in MALDI positive and negative ionisation modes. Compounds are sorted in increasing molecular weight order. S/N ratio: very good S/N>2000 (++++), good S/N 500-2000 (+++), medium S/N 200-500 (++) , low S/N<200 (+), x – no detection (S/N <6), highlighted are compounds with the strongest signal within each ionisation mode. Table also summarises compound solvent, its stock concentration and molecular weight.

149

Table 14: Different combinations of analytes included in the calibration mixture for MALDI positive ionisation mode analysis. Table summarises compound, its molecular weight and volumes used to create each mixture.

154

Table 15: PEG600 sulfate ions and their m/z values used for MALDI calibration in negative ionisation mode.

157

Table 16: Ions taken into consideration during optimisation of sample deposition method in MALDI positive ion mode.

158

CHAPTER 6: Metabolic profiling approach to investigate transfected CHO cell line response upon EPO-Fc fusion protein production

Table 1: A number of statistically significant metabolites in CHO cells footprint and fingerprint samples according to the inducer exposure time and the technique used.

188

Table 2: Metabolite features of external metabolome of transfected CHO cells under inducer exposure analysed with GC-MS technique. The table summarises the metabolite IDs, hit and match with GC-MS library and MSI level*, HMDB ID, type of sample, inducer exposure time and abundance: high (↑) or low (↓) in control (CTRL) and cell-induced (IND) extracts. Sorted by inducer exposure time.

201

Table 3: Metabolite features of external metabolome of transfected CHO cells under inducer exposure analysed with MALDI-MS technique. The table summarises experimental m/z , metabolite hit with HMDB database, mass error, type of sample, inducer exposure time, MALDI-MS ionisation mode and abundance high (↑) or low (↓) in control (CTRL) and cell-induced (IND) samples. Sorted by inducer exposure time and ion mode.

202

Table 4: Metabolite features of internal metabolome of transfected CHO cells under inducer exposure analysed with GC-MS technique. The table summarises the metabolite IDs, hit and match with GC-MS library and MSI level*, HMDB ID, type of sample, inducer exposure time and abundance: high (↑) or low (↓) in control (CTRL) and cell-induced (IND) extracts. Sorted by inducer exposure time.

203

Table 5: Metabolite features of internal metabolome of transfected CHO cells under inducer exposure analysed with MALDI-MS technique. The table summarises experimental m/z , metabolite hit with HMDB database, mass error, type of sample, inducer exposure time, MALDI-MS ionisation mode and abundance high () or low (↓) in control (CTRL) and cell-induced (IND) samples. Sorted by inducer exposure time and ion mode.204

Table 6: A summary of metabolic pathways and influenced metabolites in cells under the inducer exposure. Data based on PLS-DA footprint and fingerprint profile analysis for both time points; data sorted by impact value*.207

CHAPTER 7: Metabolic imaging of transfected CHO cells

Table 1: Ions selected to localise cell areas within control and cell-induced imaged samples. The table summarises measured mass, mass error (ppm) and metabolite according to Human Metabolome Data Base.238

Table 2: A total of 23 significant endogenous compounds identified within control and tetracycline-induced cells based on the analysis of cell populations selected from images in MALDI positive ion mode. The table summarises measured mass, p value, mass error (ppm), HMDB ID, metabolite hit and abundance in induced cells (with comparison to control): high (↑) and low (↓). Metabolites sorted by p value.243

Table 3: A total of 16 significant metabolite fragments detected in control and tetracycline-induced CHO cells based on the PLS-DA of cell populations selected from SIMS images in positive ion mode. The table summarises measured and theoretical mass, VIP scores, mass error (ppm), proposed fragment and possible origin of fragments, layer: 1st 'top', 2-10 'lower', abundance in induced cells (when compare to control): high (↑) and low (↓). Metabolites sorted by VIP scores. Fragments were compared against current literature and ***Surface Spectra Static SIMS Library software.248

Table 4: The most significant metabolites identified within control and tetracycline-induced cells based on the analysis of single cells selected from images in MALDI positive ion mode. The table summarises measured mass, p value, mass error (ppm), HMDB ID, metabolite hit and abundance in induced cells (with comparison to control): high (↑) and low (↓). Metabolites sorted by p value.258

Table 5: Fragment ions detected in control CHO cells using SIMS+. The table summarises type of compound, measured and theoretical m/z values, mass error (ppm), proposed fragment and possible origin of fragments. Fragments were compared against current literature.263

Table 6: A representation of 5 significant metabolites for different types of metabolite profiles identified within control CHO cell population in MALDI+. The table summarises measured mass, p value, mass error (ppm), HMDB ID, metabolite hit and abundance in each type of metabolite profiles shown as Box and Whisker plots; type A-red, type B-green, type C-blue. Metabolites sorted by p value.273

SUPPLEMENTARY INFORMATION

Table 1: Combinations of the analytes contained within the calibration mixture in MALDI positive ion mode. S/N of each compound in each mixture is given; x – no detection (S/N <6).306

Table 2: A list of the standards examined in MALDI positive and negative ionisation modes. Compounds are sorted in increasing molecular weight order. S/N ratio values for each compound and each repeat ($n = 2$) are included; x – no detection (S/N <6).307

Table 3: Scheme of medium samples dilution for MALDI-MS analysis. Table presents each dilution with volumes used for sample preparation.....	310
Table 4: The most common fixation protocols used for preservation of mammalian cells for MS imaging. Each protocol is specified with all solutions used during the fixation experiment. Sampling points for investigation of metabolite leakage are also shown ('Sample').....	316

List of equations

CHAPTER 1: Introduction

Equation 1: The SIMS equation	38
--------------------------------------------	----

Abbreviations

9AA	9-aminoacridine
ACN	Acetonitrile
ADP	Adenosine diphosphate
AMDIS	Automated mass spectral deconvolution and identification system
AMP	Adenosine monophosphate
ANOVA	One-way Analysis of Variance
AP	Atmospheric pressure
APS	Ammonium persulfate
Arg	Arginine
ATP	Adenosine 5'-triphosphate/adenosine triphosphate
CDC-DG	Cytidine diphosphate diacylglycerol
CE	Cholesterol ester
CHCA	α -cyano-4-hydroxycinnamic acid
CHO	Chinese hamster ovary
CID	Collision induced dissociation
CL	Cardiolipin
CTRL	Control
DAN	1,5-diaminonaphthalene
DESI	Desorption electrospray ionisation
DFA	Discriminant function analysis
DG	Diglyceride
dGDP	Deoxyguanosine diphosphate
DGP	Diacylglycerophosphate
DHB	2,5-dihydroxybenzoic acid
DIMS	Direct-infusion MS
DIOS	Desorption/ionisation from porous silicon
DMAN	1,8-bis(dimethylamino)naphthalene
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EI	Electron ionisation
EPO-Fc	Erythropoietin-Fc fragment
ESI	Electrospray ionisation
EtOH	Ethanol
FBS	Foetal Bovine Serum
FT	Fourier transform
FWHM	Full width of the peak at half of its maximum height
GC-MS	Gas chromatography mass spectrometry
GPI	Glycosylphosphatidylinositol
GSH	Glutathione
GSSG	L-glutathione oxidized
HCA	Hierarchical clustering analysis
HCD	Higher-energy collision dissociation
His	Histidine
HMDB	The Human Metabolome Database
HUSERMET	The Human Serum Metabolome project
IND	Induced
IUPAC	International union of pure and applied chemistry
kNN	Means of nearest neighbours
LC	Liquid chromatography
M:A	Matrix-to-analyte

MAG	Monoacylglycerol
MALDI-MS	Matrix-assisted laser/desorption ionisation mass spectrometry
MeOH	Methanol
MS	Mass spectrometry
MSDS	Material safety data sheet
MSI	Mass spectrometry imaging
MSI	Metabolomics standards initiative
MSTFA	<i>N</i> -acetyl- <i>N</i> -[trimethylsilyl]-trifluoroacetamide
MVA	Multivariate data analysis
NADH	Reduced nicotinamide adenine dinucleotide
Nd:YAG 35	Neodymium-doped yttrium aluminium garnet
NEDN	<i>N</i> -(1-naphthyl) ethylenediamine dinitrate
OPLS-DA	Orthogonal projections to latent structures discriminant analysis
PA	Phosphatidic acid
PBS	Phosphate-buffered saline
PC	Phosphatidylcholine
PC	Principal component
PCA	Principal component analysis
PE	Phosphatidylethanolamine
PEG600	Polyethylene glycol
PG	Phosphatidylglycerol
PGP	Phosphatidylglycerolphosphate
PhCCAA	4-phenyl- α -cyanocinnamic acid amide
P _i	Phosphate group
PLMS	Post lift metastable suppressor
PLS-DA	Partial least squares regression discriminant analysis
PNA	<i>para</i> -nitroaniline
PS	Phosphatidylserine
Q	Quadrupole
QC	Quality control
RMSE	Root mean square error
RNA	Ribonucleic acid
S/N	Signal-to-noise
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIMS	Secondary ion mass spectrometry
TCA cycle	Citrate cycle/tricarboxylic acid cycle
TEMED	<i>N, N, N', N'</i> -tetramethylethylenediamine
TET	Tetracycline
TFA	Trifluoroacetic acid
TG	Triglyceride
THAP	2,4,6-trihydroxyacetophenone
TIC	Total ion count
TOF	Time-of-flight
Trp	Tryptophan
UPLC-MS	Ultra- performance liquid chromatography mass spectrometry
UVA	Univariate data analysis
VIP	Variable importance in projection

Abstract

The University of Manchester

Candidate's name: Ewa Szula

Degree title: Doctor of Philosophy

Thesis title: Metabolic profiling and imaging of CHO cells for fusion protein production

June 2017

Fc-fusion proteins (*e.g.* EPO-Fc) are the most often created fusion proteins due to their beneficial biological and pharmacological properties. The economic success of Fc-fusion proteins and other biopharmaceuticals production however, greatly depends on a robust, low-cost and highly effective protein mammalian cell extraction system. Understanding of how cells respond to a protein production environment based on metabolic profiles provides new goals for bioengineering of cell lines for best performance in biomanufacturing. Furthermore, insights on how individual cell metabolism and therefore phenotype, respond to cell microenvironment allows the underlying biological mechanisms to be explored in greater detail.

This study focused on the application of mass spectrometry (MS) technologies, combining the analysis of metabolic profiles of cells extracts by GC-MS and MALDI-MS and spatial visualisation and distribution of metabolites within cells producing the fusion protein by MALDI-MSI and SIMS imaging.

The analysis of external and internal metabolome profiles of cells producing the protein showed an extended effect of EPO-Fc fusion protein production on cell metabolism. The findings indicate that changes observed in EPO-Fc producing cells are related to enhanced protein and lipid synthesis highlighting that these cells are in a state of increased metabolic activity with the protein exocytosis into growth medium. Moreover, the composition of lipid bilayer of induced cells seemed to be different to non-induced cells.

These findings were confirmed with the analysis of EPO-Fc induced cells using MS metabolic imaging. Multivariate analysis highlighted a number of metabolites that were significantly influenced by the protein expression when compared to control cells. The major metabolic changes in induced cells were those related to lipid metabolism. The information about metabolic changes in tetracycline-induced cells obtained from the analysis of cell populations was further supported with the analysis based on single-cell studies. Single-cell based studies also proved that investigations of individual cells provide additional insights about changes in metabolism of induced cells that can be referred to a unique, single cell and its phenotype. The analysis of CHO cells revealed a high level of heterogeneity within a cell population. Different cell phenotype and hence, metabolite content allowed for correlation between cell locations and their metabolite characteristics, specific for each type of cells.

This project has successfully shown combination of bio-analytical techniques to investigate external and internal metabolome changes related to a fusion protein production in mammalian cells. Additionally, the significance of single cell approaches in metabolomics has also been highlighted, providing insights into the sub-cellular distribution of metabolites in cells producing EPO-Fc and information on the level of heterogeneity within a cell population.

A multidimensional approach for metabolic profiling and future technological improvements of single-cell platforms are required to provide improved data acquisition and data analysis in order to better understand unknown processes involved in cell metabolism.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University's policy on Presentation of Theses.

Acknowledgements

First and foremost I would like to thank my supervisors Dr Nicholas Lockyer, Professor Royston Goodacre and Professor Alan Dickson for their encouragement and great support throughout those 4 years I have spent under their guidance. I would also like to acknowledge BBSRC DTP for funding my PhD studies that I would not have been able to undertake without.

Furthermore, I would like to thank my collaborators I had the pleasure of working with. Professor Bernard Spengler for giving me the opportunity to work in his lab and Dr Dhaka Bhandari for his enormous help and guidance with experiments. I would also like to thank Ionoptika Ltd. for providing J105 data to this thesis.

I am very grateful for all the support received from present and former group members of the Lockyer group: Irma, Giles, Jo, Afnan, Huriyyah, Samar, Sadia, Andres, Taylor, Danielle and John. Special thanks to Dr Alex Henderson for all the advice and help provided with data processing and Dr Kat Hollywood for great help and consultation throughout the project.

I would like to thank all members of Goodacre group for their support and conversations on the project. Specifically, I would like to acknowledge Dr Najla Almasoud, Dr Drupad Trivedi, Dr Abdu Subaihi and Dr Piotr Gromski for their advice.

It has been also a pleasure to work with all members of Dickson group, who are a great source of help, especially Samia Akhtar and Claire Gaffney. The project would not be possible without the CHO hosts for EPO-Fc fusion protein production created and provided by Tom Leach. I would also like to thank all other people, who gave me their guidance and support: Dr Christopher Gray and Reynard Spiess.

I would like to thank all my lovely friends in Poland and those who I met in Manchester: Ewa, Dominika, Agnieszka, Kasia, Maria, Naza, Antonios, Adriana and all the others I cannot remember at the moment, for many evenings of fun filled conversations, laughter, food, drinking and dancing. Special thanks for Amadou, who has been there for good and bad and has helped me through my personal challenges over those years.

Finally, I would like to thank my extraordinary family: Grandma, Mum, Dad, Brother, also Uncle, Auntie and Cousins for everything they have done for me that I will never be able to repay and who this thesis is dedicated.

CHAPTER 1: Introduction

1.1. Significance and aims of the study

The goal of the work described within this thesis is to combine the use of various bio-analytical mass spectrometry (MS) platforms in metabolomic studies related to fusion protein expression research. In addition, the application of mass spectrometry techniques in the analysis of metabolites from small populations of mammalian cells and, ultimately single cells, is another objective of the study.

The significance of such an approach is that novel insights on how single cell metabolism and therefore phenotypes, respond to the cell's microenvironment will be achievable, allowing the underlying biological mechanisms to be explored in greater detail. This knowledge is crucial in designing new types of cell platforms with specific functions such as production of fusion proteins.

The project will hence have a biotechnology focus. Cell platforms for fusion protein expression and the metabolic response of the cells to this condition will be investigated with a view to providing goals for bioengineering of cell lines for best performance in biomanufacturing.

The analysis of cell extracts was performed using gas chromatography MS (GC-MS) and matrix-assisted laser/desorption ionisation MS (MALDI-MS). Spatial visualisation and distribution of metabolites within cells producing the fusion protein was assessed by MALDI mass spectrometry imaging (MSI) and secondary ion mass spectrometry (SIMS) imaging. As discussed further below, MSI of individual cells is an area that with further development and research could be significantly beneficial for metabolomics experiments and consequently could bring new insights into characterisation of cellular variability giving a better understanding about unknown processes involved in cell metabolism.

1.2. Metabolomics

The central dogma of molecular biology declares a general information flow from genes to mRNA transcripts to protein products. The protein products, for example enzymes, affect metabolic pathways and therefore generate changes in the function or characteristics of the organism such as morphology, development or biochemical properties (so called phenotype) (**Figure 1**) (Nicholson *et al.* 1999, Roberts *et al.* 2013). Such cellular communication however is not unidirectional but is represented as a dynamic network with many feedback-loop interactions between DNA, RNA and metabolites within different 'omics of molecular biology. Nonetheless, the understanding of these interactions, their relationships and the way they are controlled presents a huge challenge for understanding cellular function (Fell 1996, Hollywood *et al.* 2006).

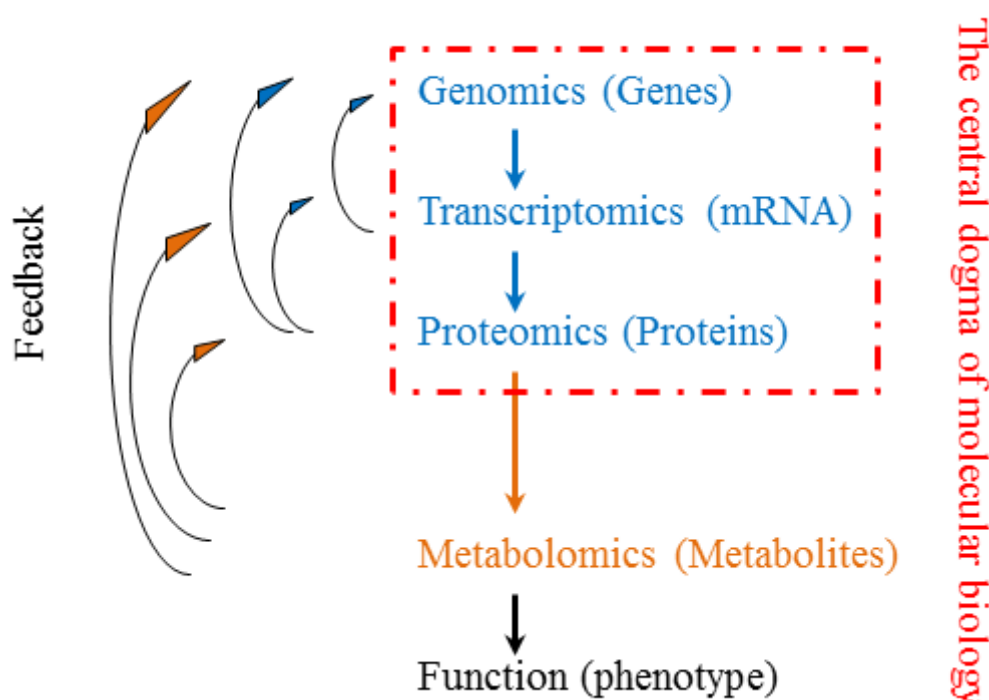


Figure 1: Central dogma of molecular biology, where information flow goes from genes to mRNA transcripts to protein products and when these protein have enzymatic properties that affect metabolic pathways, which generate changes in the function or phenotype of the organism. Interactions between 'omics organisation with many feedback-loops within them are also visualised.

Metabolomics was introduced in the 1990s to report methods aimed at assessing the metabolites present within cells and tissues (Nicholson, Lindon and Holmes, 1999). However, the idea that individuals might have 'metabolic profiles' that would reflect the

composition of their biological fluids was proposed in the late 1940s, where characteristic patterns in urine and saliva were linked to diseases such as schizophrenia (Gates & Sweeley 1978, Preti 2005).

The metabolome is considered as the final downstream product of the genome and refers to the set of low molecular weight (<1500 Da) molecules (metabolites) (Dunn *et al.*, 2011) present in living organisms, needed for their growth, development and normal function (Goodacre *et al.*, 2004). Metabolites are therefore described as a group of small molecules with a high level of chemical diversity, having a number of functions including structural, cell signalling, energy source, defence effects and stimulatory or inhibition impact on enzymes (Klyszejko-Stefanowicz, 2002). This diversity covers molecules such as carbohydrates, amino acids, lipids and nucleotides (Zenobi, 2013). Apart from chemical diversity, metabolites also have different physical properties due to the variations in atomic arrangements (*e.g.* sugars and amino acids) as well as different polarities (*e.g.* polar ethanol, non-polar lipids) (Lahner *et al.*, 2003). Additionally, depending on the cell and type of metabolites, the spectrum of metabolite concentrations can differ within a range from picomolar (10^{-12} mol/m³) to millimolar (10^{-3} mol/m³) (Dunn, Bailey and Johnson, 2005). The total amount of existing metabolites is estimated to be from a few molecules in one cell up to 10^{10} molecules for main metabolites in bigger cells. The size of the metabolome is often a subject for a discussion as emerging data show also the importance of metabolic products of the microbiome, regardless of the metabolic makeup of an individual (Roberts *et al.*, 2013). The potential size of the metabolome is extensive and it varies within species: *S. cerevisiae* ~600 metabolites (Forster *et al.*, 2003), plant kingdom up to 200,000 metabolites (Fiehn, 2001), human metabolome ~42,000 (<http://www.hmdb.ca/>).

Consideration of the large variety of physical and chemical properties that create the metabolome and the dynamic scale of metabolite concentrations across considerable orders of magnitude, makes it evident that a broad employment of analytical technologies in metabolomics research is required.

1.3. Mass spectrometry as a platform for metabolomic studies

The goal of metabolomics investigations is to analyse a large fraction or all of the metabolites present in a given sample and correlate their activity with metabolic pathways and specific physiological state (Dunn, Bailey and Johnson, 2005). However, due to physico-chemical heterogeneity of the metabolome as well as complexity of the living organisms, no single analytical technology is able to supply information on all the metabolites occurring in biological samples.

Since no single technology can compensate the coverage of the entire spectrum of the metabolome, a multiplatform analytical approach is required to broaden the detection and identification of metabolites in a biological system. A number of analytical methods have been commonly applied in metabolic studies (Bujak *et al.* 2015, Lindon & Nicholson 2008, Boughton *et al.* 2016, Zenobi 2013). Additionally, a classification of different metabolomics strategies *i.e.* targeted and non-targeted analysis, (Gonzalez-Riano *et al.* 2016, Hollywood *et al.* 2006), metabolic imaging approach (Lanni *et al.* 2012, Vickerman 2011) as well as strengths and limitations of different methods and the aims of the study define the technology being selected for metabolomics research (Villas-Bôas *et al.*, 2005). MS is a highly sensitive platform for detection, quantification and structure investigations of hundreds of metabolites in a single acquisition. At the same time, the accuracy and sensitivity of detection using MS depend on the experimental approach and the instrument set-up *e.g.* sample separation, ionisation and detection (**Figure 2**) (Gowda and Djukovic, 2014).

1.3.1. Sample introduction methods

Due to the complex nature of biological matrices, it is often essential to separate metabolites of interest prior to MS analysis (Lindon and Nicholson, 2008). However, the application of direct-infusion MS (DIMS) of crude mixtures without prior separation has been used, providing high-throughput coverage of samples, along with reduced time for the analysis that is beneficial for huge analyte numbers *e.g.* screening research or clinical trials (Gonzalez-Dominguez, R Sayago and Fernandez-Recamales, 2017). For example, direct-infusion MS allowed for successful metabolic profiling of transgenic mice models of Alzheimer's disease from urine samples (González-Domínguez *et al.*, 2014). The applicability of direct-infusion methodology is enlarged by coupling with advanced instrumentation *e.g.* orbitrap MS in order to improve mass accuracy measurements (Lei, Huhman and Sumner, 2011) however, despite the improvements, this technique is less

common than chromatographic separation of the samples, especially for complex biological analytes (Agin *et al.*, 2016).

A number of analytical separation techniques such as liquid chromatography (LC) or gas chromatography (GC) became extremely effective tools for metabolite investigations and, along with the development of new software, allow for quantitative analysis of hundreds of metabolites in automation mode (Lindon and Nicholson, 2008). In this work, the GC separation method was applied and therefore this technique will be described further.

GC is one of the most widely used separation techniques in metabolomic studies, which after coupling to an MS device, becomes extremely effective in metabolite profiling and targeted metabolic investigations, allowing for quantitative analysis of hundreds of metabolites in automation mode (Lindon and Nicholson, 2008). GC-MS has been applied in various metabolomic investigations including metabolic profiling of mammalian cells with different growth characteristics (Dietmair *et al.*, 2012) or different feeding regimes (Christopher A. Sellick *et al.*, 2011) and drug metabolism (Hande *et al.*, 1988) and up to date is used as a 'gold standard' within MS techniques in metabolomics (Yinan Chen *et al.* 2016, Xu *et al.* 2014, Dunn *et al.* 2011).

The GC-MS technique carries good metabolite separation and generally avoids ion suppression, because of the use of gas phase and the nature of analyte ionisation in this method (*i.e.* typically electron ionisation). GC-MS is only useful for volatile and thermally stable samples therefore, it requires chemical derivatization of non-volatile metabolites prior to their analysis. The derivatization procedure is crucial for GC-MS analysis; it is a complex and time-consuming process and can result in undesirable losses of metabolites (Dunn *et al.*, 2011). GC-MS spectra are highly reproducible and contain putative identification that can be confirmed with structural and mass spectral databases (Kind *et al.*, 2009). Development of the GC technique by multidimensional approaches defined as GC×GC-MS has enhanced sensitivity, resolution and robustness of the metabolic analyses, when compare to the traditional GC–MS method. This multidimensional approach employs two columns: the first column, where analytes are separated according their volatility and the second column, where compounds are separated based on their polarity (Lenz and Wilson, 2007).

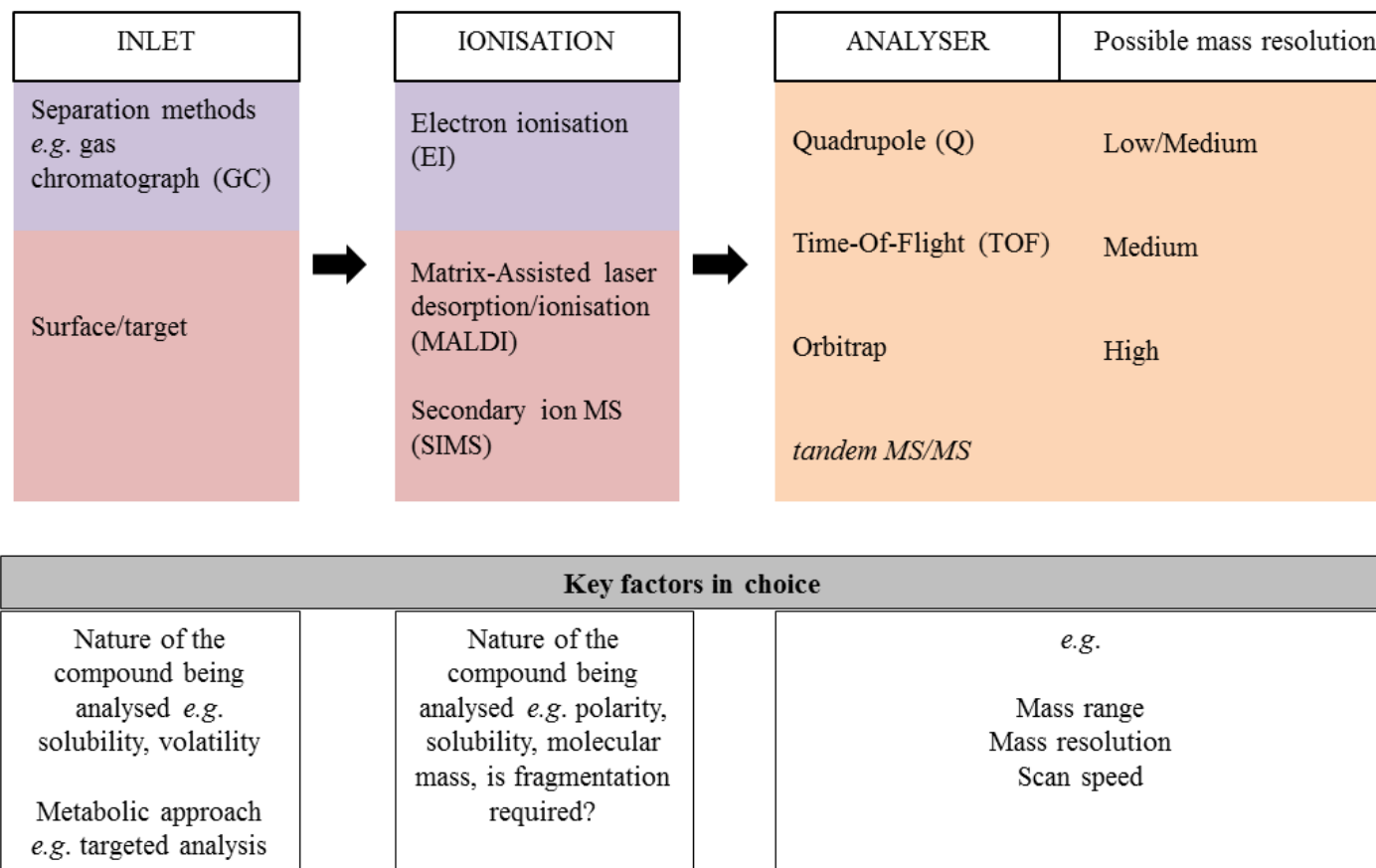


Figure 2: Schematic of example techniques applied in metabolomics based on instrument set-up, instrument possibilities and experimental approach. Adopted from http://www.rsc.org/images/MS2new_tcm18-102519.pdf.

1.3.2. Ionisation sources

Mass spectrometry by definition is “the study of systems by the formation of gaseous ions, with or without fragmentation, which are then characterized by their mass-to-charge ratios (m/z) and relative abundances” (Todd, 1991). The molecules analysed have to be charged because the mass separation relies on the movement of these charged compounds under the influence of magnetic and/or electric fields. The term ‘MS’ covers a large range of methods that differ in principles and performances (Lei *et al.* 2011, Baldwin 2005). Ionisation is one of the most important stages in MS-based metabolite analysis. The efficiency of analyte ionisation dictates the ability to detect and quantify a metabolite (Tian *et al.*, 2013).

For the GC-MS platform, the most common ionisation method available in metabolomics is electron ionisation (EI) (Hoffmann and Stroobant, 2007). This ionisation technique is a preferable method of choice when screening samples and is used to ionise volatile and thermally stable molecules that can be analysed by GC-MS. The main advantages of EI are good sensitivity and reproducibility that allow for straightforward comparison between spectra, identification of the specific profiles of the analytes or establishment of spectral databases for untargeted metabolomics. However, EI being a 'hard' ionisation method causes extensive fragmentation of metabolites, making the identification of the molecules more problematic. It is possible to decrease the energy of ionising electrons in order to lower the degree of fragmentation, nonetheless, the sensitivity of the method declines considerably (Dass 2006, Halket *et al.* 2005, Begley *et al.* 2009).

In this study, the main attention was given to another two types of ionisation modes *i.e.* MALDI-MS and SIMS and therefore, these will be described in a greater detail.

1.3.2.1. MALDI-MS

Matrix-Assisted Laser/Desorption Ionisation (MALDI) was introduced in 1988 (Karas, Bachmann and Hillenkamp, 1985) and it has since become a commonly used method in studies of broad range of molecules, including large, non-volatile proteins, oligonucleotides and metabolites (Lewis *et al.* 2000, Korte & Lee 2014, Edwards & Kennedy 2005).

1.3.2.1.1. Principles of MALDI

The first step of the MALDI technique involves sample preparation, where the analyte is mixed with a so-called ‘matrix’ (Dreisewerd, 2003). The matrix is a chromophoric, low

molecular weight compound that absorb the laser energy, protecting the analyte from fragmentation (Hoffmann and Stroobant, 2007). The mixture of analyte and matrix is then spotted on a target plate and the plate is introduced to the instrument.

The second step occurs in the instrument, inside the mass spectrometer and includes desorption and ionisation of bulk portions of the analyte-matrix mixture by intense laser pulses over a short duration time (Dreisewerd, 2003).

Laser irradiation causes rapid heating of the sample crystals due to energy accumulation in the condensed phase through excitation of matrix molecules. This rapid heating leads to sublimation of the matrix crystals, *desorption* of a crystal portion and expansion of the matrix into the gas phase, ionising intact matrix molecules in the expanding matrix plume. The expanding matrix plume contains many species: ionised matrix molecules, protonated and deprotonated matrix molecules, matrix clusters and neutral matrix molecules that are then thought to ionise the molecules of analyte (Dreisewerd 2003, Franz Hillenkamp & Karas 2007).

The MALDI mechanism that leads to the creation of charged analyte and matrix ions is not fully understood (Zenobi & Knochenmuss 1998, Dreisewerd 2003). Two most widely accepted models for *ion formation* during MALDI process have been proposed. The first model involves proton transfer to the analyte in the solid phase before desorption, whereas the second model assumes gas phase proton transfer to the analyte from photo-ionised matrix molecules in the expanding plume (Franz Hillenkamp and Karas, 2007). As a result of MALDI ionisation process, a number of ions are observed. Typically, single charged ions occur: protonated ions $[M+H]^+$, deprotonated ions $[M-H]^-$ or salt adducts *e.g.* sodium $[M+Na]^+$, where 'M' is a matrix or an analyte molecule. Rarely, multiply charged ions $[M+nH]^{n+}$ and radicals of *e.g.* matrix molecules can also appear (Hoffmann and Stroobant, 2007) (**Figure 3**).

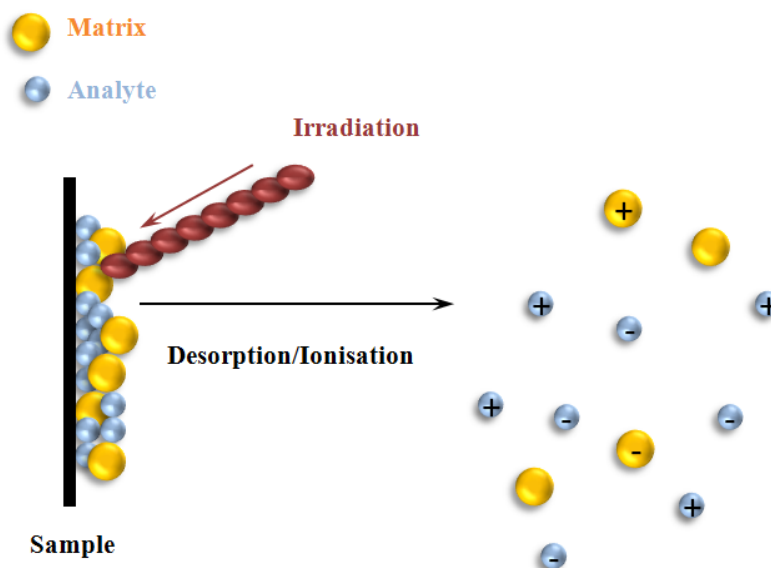


Figure 3: Schematic of the desorption-ionisation in MALDI process. Laser pulses irradiate the sample causing desorption and ionisation of multiple species.

1.3.2.1.2. Applications of MALDI-MS

MALDI-MS is a powerful analytical method that enables a parallel analysis of a large number of samples. MALDI-MS is fast and sensitive (up to 10^{-15} mol) and it can be applied for the analysis of analytes with a wide molecular weight range *i.e.* from very low (< 1000 Da) up to very high (300 000 Da) (Hoffmann and Stroobant, 2007, Duncan, Roder and Hunsucker, 2008).

In metabolomics, MALDI-MS is one of the most common ionisation techniques applied in the detection of known (targeted metabolomics) and unknown (non-targeted metabolomics) small molecules (Kiss and Hopfgartner, 2016). This technique has been used to obtain metabolite profiles from bacteria (Calvano *et al.* 2011, AlMasoud *et al.* 2016), tissue sections (Bhandari *et al.*, 2015) and cells (Povey *et al.*, 2014), providing with good quality spectra with broad mass range of analytes detected. Additionally, MALDI-MS can also be applied in the analysis of metabolic profiles of serum, urine or tissue samples and used to determine physiological changes of the organism caused by a toxic agent. Subsequently, these changes could be linked to specific metabolic pathways (Lagarigue, Caprioli and Pineau, 2016). MALDI-MS has also been applied in a number of targeted toxicity investigations *e.g.* from water samples (Cao *et al.* 2011).

One of the widest applications of MALDI-MS in metabolomics are drug discovery studies that can provide activity information about potential drug candidate, suggesting targets for further drug development and indicating binding acceptors of the tested compound (Cuperlovic-Culf and Culf, 2016). There are a number of studies, where the suitability of

MALDI-MS technique in drug metabolism analysis have been described (Muddiman *et al.* 1994, Jirasko *et al.* 2014)

The potential of MALDI-MS has also been shown in biomarker investigations (Nielsen *et al.* 2016, Bag *et al.* 2016). One of the approaches of metabolomics is to discover, validate and quantify biological markers for *e.g.* disease (Khamis, Adamko and El-Aneed, 2017). Changes at the metabolite level originate from the disease-associated metabolome adaptation and often occur in biological fluids before the appearance of clinical symptoms. Accordingly, the detection of new biomarkers can improve the understanding and diagnostics of human diseases such as cancer, diabetes or cardiovascular disorders and eventually enhance the quality of life of patients (Gowda *et al.* 2008, Cisek *et al.* 2016). Another advantage of the MALDI method is its application for quantitative studies that has been shown previously several times, for example in the study of glucose flux in mice (Sugiura *et al.* 2014), cyclosporin A concentrations in human blood (Muddiman *et al.* 1994) and investigations of potentially hazardous substances in water samples (Cao *et al.* 2011).

Due to a number of advantages of this technique *e.g.*: fast data acquisition, small analyte volumes, tolerance to salts and buffers, ease of use and potential for high-throughput analyses (Hoffmann and Stroobant, 2007, Muddiman *et al.* 1994), MALDI has started to be considered as a powerful tool in metabolomics studies.

1.3.2.2. SIMS

Developed in the 1960s (Rouberol *et al.*, 1969), the secondary ion mass spectrometry (SIMS) technique has shown the potential to become a powerful tool in metabolomic studies in *e.g.* lipid profiling in embryos (Tian *et al.*, 2014) and cancer cells (Denbigh and Lockyer, 2015) or studies of biomarkers for hypoxia (Armitage *et al.*, 2013). The great advantage of this technique is, along with chemical information, its ability to image the distribution of metabolites within cell or tissue with higher spatial resolution than the MALDI-MS technique and quantitative information (Lechene *et al.* 2007, Chandra *et al.* 2008).

1.3.2.2.1. Principles of SIMS

SIMS is a surface-sensitive method applied in the analysis of secondary ions desorbed from the sample surface after primary ion bombardment (Vickerman and Briggs, 2001). Primary ion bombardment is initiated by primary ion guns that can be monoatomic *e.g.* Au^+ , clusters *e.g.* Au_3^+ or polyatomic *e.g.* C_{60}^+ . The secondary particles that are produced

are desorbed and ionised and allow for chemical analysis of the sample surface (**Figure 4**) (Shard & Gilmore 2013, Adriaensen *et al.* 2004, Vickerman 1997). Detectable ions are usually restricted to a limited mass range of approximately a few hundred Da. Additionally, due to a high level of fragmentation of molecules, spectra analysis and data interpretation is challenging (Vickerman 2011, Henderson 2013).

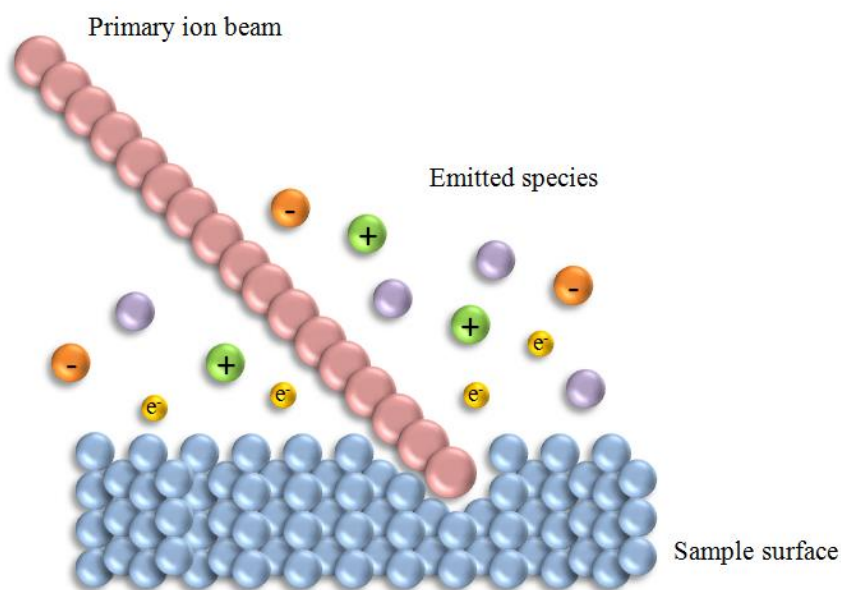


Figure 4: Schematic of the desorption-ionisation process in SIMS. The primary ion beam bombards the sample surface, causing in desorption of multiple species. The majority of species emitted are neutral (purple) atoms or molecules; a small part of emitted species is charged: cations (green), anions (orange), electrons (yellow).

The secondary ion formation depends on the primary ion beam, the nature of a sample and its electronic state and involves two processes: sputtering and ionisation of particles.

Firstly, when primary ions bombard a sample, particles *i.e.* molecules and atoms from a sample surface are desorbed (sputtered). The emission (sputtering) of molecules and atoms occurs, when the energy of the beam bombarding this surface is high enough (*e.g.* 25 keV) to overcome the surface binding energies (eV). Secondly, the sputtered particles are then ionised, resulting in the creation of neutral and charged atoms and molecules as well as fragments derived from a sample surface (Sigmund, 1981).

Only a small fraction of sputtered particles is ionised (typically <1 %) and the ionisation process can occur through: gaining or losing charged species (typically proton, also

sodium, potassium and chlorine ions), gaining or loosing electrons and through dissociation of the parent ions to produce fragment ions (Cooks and Busch, 1983).

As the SIMS technique is based on the analysis of secondary ions, the relationship between factors influencing the secondary ion generation is described by the SIMS equation (**Equation 1**) (Vickerman and Swift, 1997).

Equation 1 The SIMS equation

$$I_m = I_p \gamma_m \alpha^+ \theta_m \eta$$

Where: I_m is a positive secondary ion current from the sample species m ;

I_p – the primary ion flux;

γ_m – sputter yield of species m ;

α^+ - ionisation probability of the species m to positive ions;

θ_m –fractional concentration of species m in the sample layer

η – transmission of the analysis system including detector efficiency

The sputter yield and ionisation probability are two crucial parameters, which determine a sample behaviour under SIMS. The sputter yield is the total yield desorbed from a sample surface (including ions and neutrals) per primary ion impact. The sputter yield is dependent on the energy, charge and mass of the primary ion beam and the nature of the sample analysed (Vickerman and Swift, 1997).

1.3.2.2.2. Static and dynamic SIMS

With regard to the primary ion flux used for the analysis, there are two SIMS operational modes that can be used: static and dynamic mode (**Figure 5**).

In static SIMS, the flux of primary ions is restricted to $\leq 10^{13}$ ions per cm^2 (*i.e.* static limit), resulting in a bombardment of <1% of the atoms from the topmost layer of the sample (**Figure 5 A**). Such conditions decrease the damage to the molecular structure of the sample and limits the area of desorption to a maximum of 10 nm^2 per primary ion impact. For surface investigations, this guarantees that sputtered particles are emitted from the areas, where no previous bombardment occurred (Oran 2005, Shard & Gilmore 2013, Vickerman 1997, Lockyer 2014).

Unlike the static SIMS mode, in the dynamic SIMS mode the sub-surface of a sample is investigated. This mode involves the analysis of the material under ion irradiation as a function of depth, as shown in **Figure 5 B**.

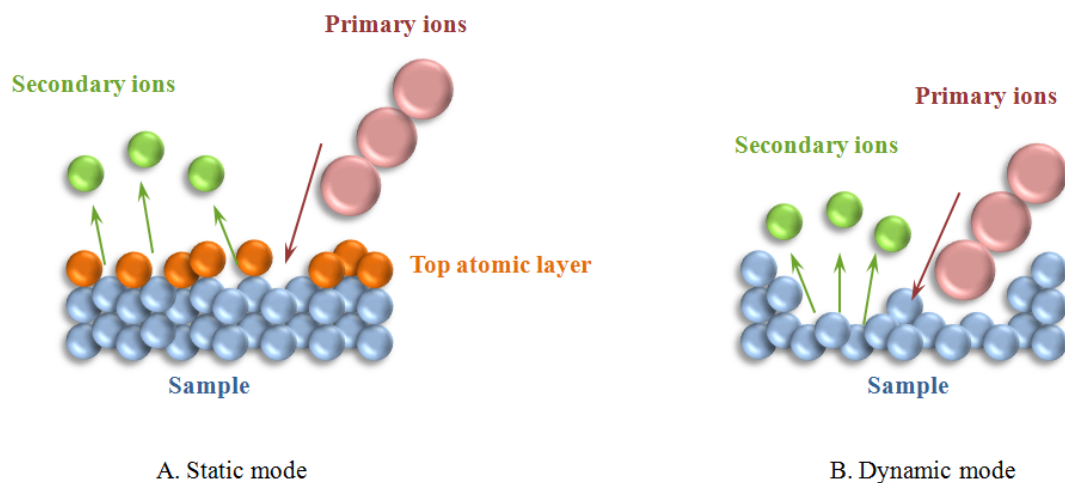


Figure 5: Two operational modes of SIMS: static mode gives information about the topmost atomic layer of the sample (A), whereas dynamic mode operates with high primary ion doses to remove material to collect 3D information resulting in the creation of an etched crater (B).

The dynamic mode allows for the characterisation of 3D chemical distribution and structural changes under high primary ion beam exposure (McPail & Dowsett 2009, Oran 2005).

The choice of a primary ion beam for SIMS analysis is important as it determines the type and amount of material detected. Only polyatomic primary ion beams *e.g.* C_{60}^{+} allow for the dynamic analysis of molecular species due to the reduced chemical damage under these beams. The interaction of the C_{60}^{+} ion with the sample is described as a formation of a crater at the surface of this sample with little perturbation of the sub-surface. On the other hand, the Au_3^{+} ion separates into single Au^{+} ions that continue their movement through the sample, leading to considerable mixing and chemical damage (tens of nanometres) (**Figure 6**) (Fletcher and Szakal, 2013).

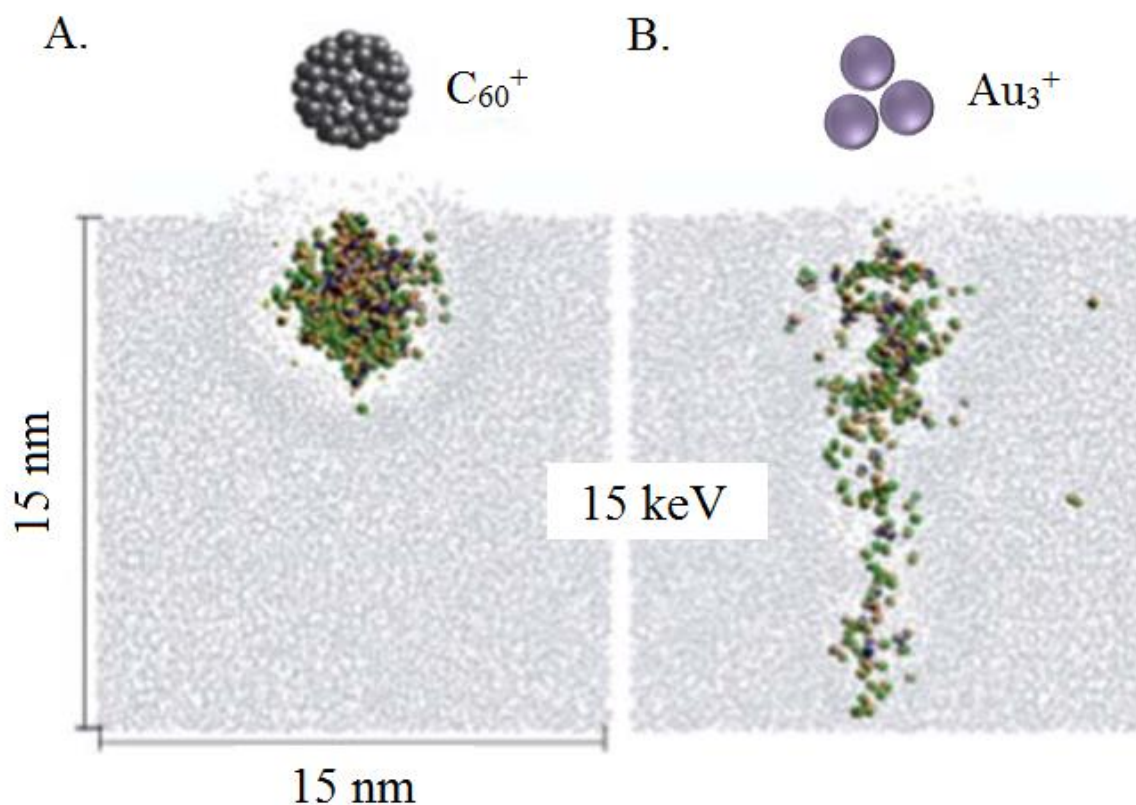


Figure 6: Snapshots of the C_{60}^+ and Au_3^+ impacting the sample surface (here: water ice). The C_{60}^+ ion beam creates a crater with minimal sub-surface perturbations (A); the Au_3^+ ion beam has a much longer energy track and hence, extensive damage is caused (B). Adopted and reproduced from (Ryan, Wojciechowski and Garrison, 2007).

1.3.2.2.3. Applications of SIMS

Several applications of the SIMS technique have been described (Lanni *et al.* 2012, Vickerman 2011). Due to their enhanced secondary ion yield the application of cluster ion sources enables the investigations of lipid content in biological systems involved in *e.g.* non-alcoholic fatty liver disease (Debois *et al.*, 2009), Alzheimer's disease (Solé-Domènech *et al.*, 2013), Fabry's disease (Touboul *et al.*, 2007) or Duchene muscular dystrophy (Touboul *et al.*, 2005). SIMS can also be applied in the analysis of organic and mineral compounds associated to pigments in *e.g.* paint samples, allowing for the determination of the pigments' content (Richardin *et al.*, 2011). Additionally, other applications of SIMS such as cosmo-chemistry (Stephan *et al.*, 2003) and analysis of human remains in cultural heritage studies (Cersoy *et al.*, 2012) have also been described.

Nevertheless, most applications of the SIMS platform is related to imaging studies due to the possibility of tightly focusing the primary ion beam, which makes this technique well suited for the creation of high spatial resolution ion images (Vanbellingen *et al.*, 2015).

1.3.3. Mass analysers

Metabolite detection with high mass resolution ($> 10\,000$), mass accuracy (< 5 ppm) and sensitivity is desired. Nevertheless, achieving these aims in a single MS analysis is challenging because higher resolution generally leads to lower sensitivity and *vice versa* (Gowda and Djukovic, 2014). There is a variety of mass analysers *e.g.* quadrupole (Q), time-of-flight (TOF) or Orbitrap and analyses alternatives such as single (MS) or tandem (MS/MS), each of which has different resolution and sensitivity performance (Hoffmann and Stroobant, 2007).

Quadrupole analysers offer good sensitivity for selected ions but limited mass resolution, whereas TOF and Orbitrap offer high mass resolution. Generally, commonly used quadrupole instruments can reach a resolving power up to 4000, *i.e.* about 3-4 times less mass resolving power than instruments with TOF, while *e.g.* Orbitrap can reach a resolving power up to 150 000 and excellent mass accuracy (< 2 ppm) (Balogh 2004, Perry *et al.* 2008, Drexler *et al.* 2011). Moreover, quadrupole instruments are only useful for analysis ions with masses up to 3,000 Da (Mishur and Rea, 2012).

A higher mass resolution simplifies a differentiation between closely localised mass-to-charge signals. Hence, the development of MS instruments to improve their performance has been implemented (Moco *et al.*, 2007). For the analysis of small molecules, a resolving power of TOF instruments has been enhanced by introducing a reflectron component to TOF tubes (Hoffmann and Stroobant, 2007). Additionally, the confidence in assignments of m/z signals to specific metabolites can be improved by tandem MS (Moco *et al.*, 2007). Mass analysers organised in a tandem configuration *e.g.* Q-TOF or Fourier transform (FT) Orbitrap (FT-Orbitrap) are widely used in metabolomics (Scheltema *et al.* 2014, Hao *et al.* 2012, Ghaste *et al.* 2016) thanks to their high mass resolving power (10 000 and more) and high mass accuracy (< 10 ppm). Fourier transform instruments *i.e.* FT-Orbitrap-MS have the highest mass resolving power ($\sim 500\,000$) and mass accuracy within 1 ppm (Brown, Kruppa and Dasseux, 2005). The appearance of high mass accuracy instruments with a broad dynamic range (*i.e.* a range over which ion signal is linear with the analyte concentration) can enhance identification effectiveness of molecules in complex mixtures (Hu *et al.* 2005, Moco *et al.* 2007).

Mass analysers commonly used with GC are single quadrupoles or TOFs, but some current platforms are supplied with Q-TOF MS. Q-TOF and FT-Orbitrap analysers are more typical for metabolite profiling and metabolite identification, because of their enhanced mass resolving power (Moco *et al.* 2007, Gowda & Djukovic 2014). TOF-TOF, Q-TOF or FT-Orbitrap can be coupled with MALDI devices (http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf, Römpf & Spengler 2013), whereas TOF mass analysers are typically used for SIMS instruments (Sheraz *et al.*, 2013). Additionally, recent development of SIMS instrument coupled to Orbitrap mass analyser (<http://www.npl.co.uk/news/3d-nanosims-label-free-molecular-imaging>), gives new possibilities for metabolomic approaches.

1.3.4. MS-based metabolic imaging

Understanding the multiple biochemical processes that occur within living organisms requires not only the investigation of the molecular entities associated with these processes but also their spatial visualisation and distribution within the individuals. The analysis of morphological and molecular cell or tissue features is the foundation of creating knowledge about *e.g.* biomarkers for disease or changed metabolism of cells, diagnostics, prediction and prognosis (Norris and Caprioli, 2013).

In metabolomics, mass spectrometry imaging (MSI) enables one to investigate spatial distribution of biological compounds present in cells and tissues by their direct ionisation and detection. MSI has also opened new insights for cell and tissue characterisation in two and three dimensions (Römpf *et al.* 2010, Hill *et al.* 2011) and promises extensive utility for *e.g.* clinical applications (Denbigh and Lockyer, 2015, Bunch, Clench and Richards, 2004) and drug discovery (Lou *et al.*, 2016, Swales *et al.*, 2016). Modern MS imaging platforms have seen major technical advances that have enhanced the applicability and adoption of the technology in other research areas *e.g.* single-cell and sub-cellular imaging (Schober *et al.*, 2012). New instrumentation provides high spatial resolution ($< 1 \mu\text{m}$) (Rabbani *et al.*, 2011), high mass resolution ($\sim 100\,000$), high mass accuracy capabilities ($< 3 \text{ ppm}$) as well as the ability to measure a wide range of analytes (Bhandari *et al.*, 2015).

Three of the most commonly used ionisation platforms for MS imaging studies are MALDI, SIMS and desorption electrospray ionisation (DESI) (Römpf *et al.* 2010, Cobice *et al.* 2015, Clench, 2016).

SIMS platforms currently give the highest spatial resolution ($< 1\ \mu\text{m}$) (Rabbani *et al.* 2011) and can provide quantitative chemical information (Chandra *et al.*, 2008). However, the ionisation of the analytes requires high vacuum (10^{-8} mbar) and due to the high energy of the primary ion beams, the analyte fragmentation is intense (Rabbani *et al.* 2011, Vickerman 2011, Hoffmann & Stroobant 2007).

The approach of SIMS imaging is frequently dedicated towards cellular and sub-cellular analyses, including 3D imaging (Graham *et al.*, 2016, Hill *et al.*, 2011, Winograd, 2015) and analyses of tissue sections (Fletcher *et al.*, 2013, Touboul and Brunelle, 2015).

The use of static SIMS for cellular MS imaging has been applied to a range of detectable endogenous compounds such as lipids and other membrane-localised small species *e.g.* cholesterol (Jerigova *et al.*, 2011). The effectiveness of SIMS in imaging the localisation of membrane contents, lipids in particular, is highly beneficial in bio-analytical investigations, especially since only a few techniques can do this, without the use of other labelling agents (Lanni, Rubakhin and Sweedler, 2012). Lipid profiles have been obtained from a frog egg during early stages of development, opening the possibility for developmental studies using high resolution imaging MS and biochemistry of the sample (Tian *et al.* 2014). Other studies demonstrated for visualisation of membrane cholesterol within cultured macrophages and their relative quantification by correlation with fluorescence imaging, allowing for separation between cholesterol-enriched and control cellular populations (Ostrowski *et al.*, 2007).

On the other hand, MALDI technique is a soft ionisation method and offers a wide mass range of analytes, high sensitivity and recently, high spatial resolution *i.e.* $1.4\ \mu\text{m}$ (Kompauer, Heiles and Spengler, 2017). Additionally, the development of new MALDI technology allows for desorption and ion formation process to take place in atmospheric pressure (AP). The limitation of MALDI with low mass resolution instruments is however, the interference of matrix-derived signals with signals coming from metabolites. This disadvantage can be overcome by application of high-resolution and high mass accuracy mass analysers *e.g.* Orbitrap (Bhandari *et al.*, 2015). Consequently, a number of studies aimed at imaging of metabolites in different types of cells and tissues using the MALDI-MSI platform have been described (Takahashi *et al.* 2015, Shanta *et al.* 2012, Fülöp *et al.* 2016, Swales *et al.*, 2015).

For example, various small molecules have been detected and localised within *Arabidopsis thaliana* dried plant tissues with the application of different MALDI matrices (Takahashi *et al.*, 2015). This technique has also been applied in metabolic imaging of drugs, drug metabolites (Shanta *et al.*, 2012, Swales *et al.*, 2015) and liposomes in mammalian tissues,

providing information about transport of drug molecules from blood into brain area (Fülöp *et al.*, 2016). Additionally, with constant development of MALDI-MSI devices, this technique can be used for metabolite characterisation in tissue samples with efficient ion yield and high spatial resolution (1.4 μm), allowing for mapping cellular and sub-cellular distribution of lipids (Kompauer, Heiles and Spengler, 2017).

DESI is a soft ionisation method that uses charged solvent droplets (*i.e.* electrospray) in order to release the molecules from a sample surface (Takáts *et al.*, 2004). This technique operates under atmospheric pressure, giving a possibility to access the sample during acquisition and it does not require the application of a matrix hence, does not experience analyte-matrix co-crystallisation issues (Cooks *et al.*, 2006). Additionally, DESI can also be an alternative for compounds that are difficult to ionise using the MALDI technique (*e.g.* inks) (Ifa *et al.*, 2007).

The major limitation of the DESI technique is its spatial resolution because ‘focusing’ the charged solvent droplets is challenging (Wiseman *et al.*, 2008). The spatial resolution can vary from approximately 150 μm (Girod *et al.*, 2011) to recently improved 35 μm that has been applied in the imaging of biomarker for disease within mouse brain tissue (Campbell *et al.*, 2012) and drug distribution studies (Ferreira *et al.*, 2014).

Along with further improvements of techniques, MSI is becoming an emerging platform for a diverse range of research fields, providing high quality spatial distribution and chemical information of visualised compounds (Lanni *et al.* 2012, Aichler & Walch 2015). Additionally, the development of validation and calibration protocols as well as information about a spatial resolution will progress fully quantitative metabolomics approaches and a realisation of the full potential of MSI will be achieved (Trim and Snel, 2016).

The most expected goal of all science studies however, is the integration of entire ‘omics’ data sets of the living organisms, where the complete control of their function is based on mutual interactions between genomic, transcriptomic, proteomic and metabolomic activity (Weckwerth 2003, Dunn *et al.* 2005). As more and more researchers become aware of the applicability and benefits of the metabolomics studies, the application area will enlarge further.

1.4. Metabolomics in practice

Metabolomics approaches have grown rapidly and become widely employed in many different branches of analyses. The application of metabolomic investigations has been shown on a number of different studies and is broadly described in the literature (**Table 1**).

Table 1: Metabolomics applications in different branches of studies.

Application	Examples	References
<u>Functional genomics</u> - data sets produced by genomics, transcriptomics and proteomics are combined with the attempt of understanding how the DNA sequence is translated into information in a cell	Phenotype changes as a result of a genetic manipulation <i>e.g.</i> deletion of a gene; <i>e.g.</i> <i>Arabidopsis</i> phenotype discrimination based on metabolic data between two related lines of this plant and their offspring	(Oliver <i>et al.</i> 1998, Pevsner 2009, Taylor <i>et al.</i> 2002)
<u>Nutrigenomics</u> - study of the relationships between genes and nutrients found in food and their influence on different phenotypes and personal health	Positive effect of polyunsaturated fatty acid supplementation on lipoprotein metabolism in Caucasian population	(Müller & Kersten 2003, Rudkowska <i>et al.</i> 2013).
<u>Pharmaceutical industry</u> <i>i.e.</i> toxicological profiling, biomarker studies for disease, drug metabolism, environmental stress	Fatty acid metabolism affected by toxicities found <i>e.g.</i> in drinking water; biological activity of a potential drug candidate	(Shi <i>et al.</i> 2014, van Ravenzwaay <i>et al.</i> 2014, Cisek <i>et al.</i> 2016, du Preez <i>et al.</i> 2017, Dinis-Oliveira 2016, Hasler-Sheetal <i>et al.</i> 2016, Cuperlovic-Culf & Culf 2016).
<u>Metabolic engineering</u> – optimisation of metabolic interactions within cells to enhance production of a substance of interest on an industrial scale and in a cost effective manner	Production of pharmaceuticals and other biotechnology products <i>e.g.</i> biofuels	(Kulkarni 2016, Divol & Bauer 2010, Trosset & Carbonell 2015, Wilkens & Gerdtzen 2015, George <i>et al.</i> 2015)

Metabolomics strategies are divided into two different approaches *i.e.* targeted analysis and untargeted analysis, sometimes referred to ‘metabolic profiling’ (Brown *et al.*, 2009). Targeted metabolomics is the analysis of defined groups of chemically characterised metabolites. This strategy can be beneficial with extensive understanding of these groups, *e.g.* metabolic enzymes, their products and biochemical routes, which they contribute. Additionally, since all investigated molecules in targeted approach are defined, sample preparation can be more easily optimised, decreasing highly abundant species in the analysis as well as detecting any analytical artefacts that can be further withdrawn. Moreover, with the application of internal standards, *i.e.* chemical substances used for comparison, the targeted metabolomics approach can be undertaken in a semi- quantitative and quantitative manner (Roberts *et al.*, 2013). When employing targeted analysis, novel relationships between metabolites can be identified in the context of specific physiological states *e.g.* disease (Zhao *et al.*, 2010).

By contrast, untargeted metabolomics is a broad study of all the measurable species present in a sample, including the unknowns. Due to its broad nature, untargeted analysis pipeline normally employs a number of analytical methods and advanced data analysis techniques in order to reduce the large data set into a smaller collection of manageable signals. These signals then need annotation using metabolite libraries and databases (Khakimov, Bak and Engelsen, 2014). Untargeted metabolomics allows for novel target detection, as metabolome coverage is only limited by the sample preparation procedures and attributes of analytical techniques used, *e.g.* sensitivity and specificity (Roberts *et al.* 2013).

1.4.1. Fusion proteins

One of the recent area in where metabolomics has been applied is the production of pharmaceuticals such as Fc-fusion proteins that create the most rapidly expanding group of bio-products and currently dominate the biopharmaceutical market due to their suitability in applications in diverse clinical settings (Johari *et al.*, 2015).

Fusion proteins are proteins constructed through assembling two or more genes that originally code individual proteins. Some fusion proteins consist of complete polypeptides and hence, have all functional sites of the original proteins. Nevertheless, other fusion proteins consist of only parts of coding sequences (especially those naturally existing) and consequently, do not preserve the functionality of the original genes that formed them. Fusion proteins combined from partial gene sequences are called mutant fusion proteins and are often found in cancer cells (*e.g.* bcr-abl fusion protein) (<http://chippi.md.biu.ac.il/>).

Both types of fusion proteins *i.e.* combined from complete or partial gene sequences can experience interactions between the two joined peptides that can alter their original functions. Additionally, for partial gene combinations a variety of interactions between different active sites and binding domains has potential to give novel proteins with new functions (<http://chippi.md.biu.ac.il/>, Czajkowsky *et al.* 2012). Individual cell analysis is also crucial due to the fact that cell populations are, in most cases, not homogeneous, where cells behave non-synchronously. Therefore, cells must be studied independently in order to obtain relevant and accurate metabolic information. Heterogeneity occurring within cell populations would be lost when analysing averaged cell populations data (Emara *et al.*, 2017) and therefore, single-cell investigations are essential to create unbiased metabolic models.

1.4.1.1. Fc-based fusion proteins

Fc-based fusion proteins are recombinant proteins consisting of an antibody Fc domain directly joined to another peptide. Recombinant fusion proteins are constructed artificially through genetic modifications.

Fc-based fusion proteins are the most often created fusion proteins due to their beneficial biological and pharmacological properties (Czajkowsky *et al.*, 2012). The presence of Fc domain considerably increases plasma half-life of the protein, which extends therapeutic activity (Roopenian and Akilesh, 2007). The Fc domain in Fc-based fusion proteins also allows for interactions of these proteins with Fc receptors present on immune cells. The ability to interact with receptors on immune cells is crucial for Fc-based fusion proteins' use in oncological therapy and vaccines (Nimmerjahn and Ravetch, 2008). In addition, Fc domain is believed to fold independently and can enhance the solubility and stability of the whole fusion protein both *in vitro* and *in vivo*. Moreover, from a technological point of view, the Fc domain enables for easy and low-cost purification of Fc-based fusion protein that is highly beneficial for manufacturers (Carter, 2011).

There are a number of Fc-fusion proteins that have been introduced and are applied in different health disorders, for example: Enbrel (Amgen/Pfizer) for rheumatoid arthritis, Amevive (Astellas Pharma) for psoriasis and transplant rejection or Nulojix (Bristol-Meyers Squibb) for organ rejection (Czajkowsky *et al.*, 2012).

1.4.1.2. EPO-Fc fusion protein

One of the Fc-based fusion proteins that has recently gained attention is erythropoietin-Fc fusion protein (EPO-Fc) (Taschwer *et al.*, 2012).

Human erythropoietin (EPO) is a glycoprotein hormone that controls red blood cells production. This protein is produced naturally in kidneys and is up-regulated in hypoxia environment, resulting in increased generation of red blood cells (Nagels *et al.*, 2012).

EPO protein has a molecular mass of 30.4 kDa (Davis *et al.*, 1987) with four glycosylation sites that accounts for approximately 40% of the molecular mass of this protein (Sasaki *et al.*, 1987). Recombinant human EPO has been produced and become extensively applied as a therapeutic protein for the treatment of anaemia, AIDS or chemotherapy (Jelkmann, 2007). However, EPO administration techniques other than injections are limited, mainly due to the inefficient penetration EPO through epithelial barriers.

Consequently, unique approaches have been undertaken to overcome this and other limitations and EPO-Fc fusion protein have been introduced. EPO-Fc allows for a recognition by neonatal Fc receptor (FcRn) on the epithelial cells of the lungs and this interaction is then used to transfer EPO-Fc across the epithelium barrier to the blood stream (Bitonti *et al.*, 2004). Additionally, a number of attempts have been made to modify structure of native EPO protein in order to improve its therapeutic properties *e.g.* by increasing its molecular weight to slow down its *in vivo* metabolism or prolong half-life (Wang *et al.*, 2010).

1.4.1.3. Mammalian cell lines as a platform for protein production

Recombinant Fc-based fusion proteins are now a well-established class of therapeutics (Czajkowsky *et al.*, 2012). However, economic success greatly depends on a robust, low-cost and high effective protein production systems (Birch and Racher, 2006). Dominant platforms used for production of biopharmaceuticals are based on mammalian cell culture due to their good effectiveness and post-translational modifications *i.e.* glycosylation that is crucial in a protein mechanism of action (Wurm, 2004).

Chinese hamster ovary (CHO) cells represent a common expression platform for the production of biopharmaceutical products hence, much effort is given into further development of the CHO-based production system (Trummer *et al.*, 2008). CHO cells are epithelial (from tissue) cells derived from Chinese hamster, first isolated in 1957. Since the original CHO cell line was characterised (Tjio & Puck, 1958), a number of variants of this

cell line have been evolved for various purposes (Taschwer *et al.* 2012, He *et al.* 2012, O’Callaghan *et al.* 2010, Kyriakopoulos *et al.* 2013).

CHO cells are a dominant platform used for the development of therapeutic proteins and have a proven history of regulatory approval for expression of recombinant proteins. They are considered safe, mainly due to their ‘resistance’ for most human pathogenic viruses also, they grow relatively fast and robust (Wurm and Hacker, 2011). Additionally, CHO cell line can be cultured either as adherent cells or in suspension and can be genetically manipulated to improve production of proteins (Xu *et al.*, 2012).

The fundamental issues affecting the choice of a cell host for protein production, apart from the ability to produce high protein concentrations and the low cost of manufacturing, are the capability to produce high quality product with desired characteristics and the time in which these high product concentrations can be accessed (Birch and Racher, 2006). Consequently, the improvement of chemically defined cell media and development of feeding strategies have contributed to increased cell populations, resulting in higher product titres.

Nevertheless, a major limiting factor for further improvements in process performance of mammalian cell culture and protein expression is inefficient cell metabolism (Taschwer *et al.*, 2012). Therefore, better understanding of cellular metabolism pathways and cellular physiology under protein production can offer improved longevity of cells and large-scale manufacturing of desired proteins.

1.4.2. Single cell based metabolomics

In life sciences, the single cell is believed to be the smallest functional unit and its analysis is certainly crucial. Single-cell based metabolic investigations represent the analysis of a wide range of molecular information carried on a variety of bio-molecules at the individual cell level. Such information can give understanding about unknown processes involved in *e.g.* cellular evolution, communication or adaptation (Emara *et al.*, 2017). Additionally, novel metabolomic investigations at the single-cell level can bring new insights to phenotypic differences within one cell population and in a very sensitive manner (Heinemann and Zenobi, 2011).

Individual cell analysis is also crucial due to the fact that cells are not necessarily synchronised (*e.g.* different growth stage) and hence, the cell population is not homogeneous. Therefore, cells must be studied independently in order to obtain relevant and accurate metabolic information. Heterogeneity occurring within cell populations would

be lost when analysing averaged cell populations data and therefore, single-cell investigations are essential to create unbiased metabolic models (Emara *et al.*, 2017).

There are a number of reasons for cellular heterogeneity: cells might be genetically different, can live in a different microenvironment or happen to be a different age or in a different cell-cycle stage (Heinemann and Zenobi, 2011). For instance, metabolomic investigations on yeast cell cycle and on cell population level revealed significant changes in gene regulation and metabolite levels during different stages of cell cycle (Murray *et al.* 2007). With regard to CHO cells, it has been shown that CHO cell populations often show diversity in spreading and adhesion during the cell growth, resulting in variations in cell morphology. It is believed that adhesion and hence morphological changes, link to phenotypic variation that can be of potential biological significance. The importance of adhesion variability has been investigated by Dao *et al.* revealing the adhesion is non-genetic and cell cycle-independent but is scaled with the membrane receptors responsible for adhesion (Dao *et al.*, 2012).

Another source for cellular heterogeneity is phenotypic heterogeneity that is believed to be stochastically (*i.e.* randomly) induced. Due to a low copy numbers of specific molecules within cells, some processes at the gene and protein production level are stochastic and can result in random variations in the abundance of bio-molecules. These variations can be then introduced to biological pathways *e.g.* regulatory processes, causing the generation of multiple cell phenotypes even within the same environment (Raj & van Oudenaarden 2008, Veening *et al.* 2008).

Single cell metabolomics has a number of challenges. Firstly, individual cell approach has to consider very low quantities of analytes. For example, glycolytic intermediates are present at the low millimolar concentration range in *E. coli* bacterium (in 1fL of volume), attomolar concentrations occur in yeast (in 65 fL of volume), whereas in mammalian cells, metabolites are present within a concentration range of femtomoles (in 500 fL of volume). Such concentrations are approximately 10^9 times lower, when compared to the amount of metabolites that are usually used (*i.e.* nanomoles) in metabolomic studies at the population level (Schmid *et al.* 2010). Moreover, unlike *e.g.* genomic or proteomics, amplification of analyte is not possible, making metabolomics even more challenging (Heinemann & Zenobi 2011). Another challenge in metabolomic investigations is a sample preparation process, where a cell metabolism has to be stopped rapidly and ideally without a loss or degradation of metabolites (so-called quenching). Hence, the appropriate extraction procedure should be based on the greatest number of metabolites recovered and also, on the reproducibility of the method (Sellick *et al.* 2011). Lastly, because individual cell

metabolomics platforms are expected to reveal differences between single cells, these platforms also need to generate high quality measurements to enable for the meaningful statistical data analyses and interpretation (Heinemann and Zenobi, 2011).

Due to a limited number of techniques (**Table 2**) allowing for metabolic studies at the single-cell level, only a few experiments have been performed (Heinemann and Zenobi, 2011). However, single-cell analytical platforms have recently undergone rapid development and reached an important level, where they have grown into powerful tools used for characterisation of cellular variability (Emara *et al.*, 2017).

Table 2: Mass spectrometry techniques applied in metabolomics studies for profiling and imaging of small molecules at the single-cell level.

Technique	Description	References
SIMS <i>e.g.</i> ToF-SIMS, NanoSIMS	Matrix-free; semi-quantitative; MS imaging (pixel size < 1 μm), sub-cellular level; mass range ~1-1000 Da; hard ionisation; low throughput operation	Passarelli and Ewing, 2013
DESI-MS	Soft ionisation; matrix-free; MS imaging (pixel size ~35 μm); alternative for MALDI-MS; semi-quantitative; mass range ~50-70,000 Da; medium throughput operation	Cooks <i>et al.</i> , 2006, Campbell <i>et al.</i> , 2012
MALDI-MS	Broad mass range ~100-200,000 Da; soft ionisation; quantitative; MS imaging (pixel size < 10 μm), sub-cellular level; matrix application; possible application of plates (so called sample arraying) with 'anchors' for single-cell deposition; low/medium throughput operation	Heinemann and Zenobi, 2011
DIOS-MS	Matrix-free; semi-quantitative; soft ionisation; suitable for small molecules; alternative for MALDI-MS; MS imaging (pixel size ~25 μm); medium throughput operation	Thomas <i>et al.</i> , 2001, Liu, Guo and He, 2007
NanoESI-MS	Electrospray needle placed into a cell under video microscopic observation and electro-spraying the cell contents <i>via</i> this needle; allows for sub-cellular studies; low throughput operation; direct analysis of cells; suitable for larger cells; better sensitivity (than ESI) and reduced droplet size can be obtained	H Mizuno <i>et al.</i> , 2008, El-Faramawy, Siu and Thomson, 2005
Live single-cell MS	Curved fiber tip to point the focused UV laser beam onto a cell that ionises cell content; analysis by LC/ESI-MS; low throughput operation; potential for detailed qualitative and quantitative studies	Izumi <i>et al.</i> , 2009, Svatos, 2011

DIOS – desorption/ionisation from porous silicon; LC – liquid chromatography

1.5. Metabolomics and data analysis

Metabolite data generated by MS are in general rich in information and therefore, chemometric statistical tools are often required to reduce the data dimensionality from such complex datasets. The main approaches in analysing metabolite datasets are to obtain detailed insight into the molecular mechanisms of cellular metabolic routes and to identify molecules that are involved in these pathways (Gonzalez-Riano, Garcia and Barbas, 2016). Complete metabolomics investigations involve a number of steps that are performed before the pre-defined hypothesis can be answered and include: i) pre-processing-selecting features from raw data to a suitable form; ii) pre-treatment-for example scaling to put all samples and variables on a comparable scale; iii) processing-statistical modelling and data visualization; iv) interpretation (**Figure 7**) (Sumner *et al.* 2007, Gromski *et al.* 2015, Sussulini 2017).

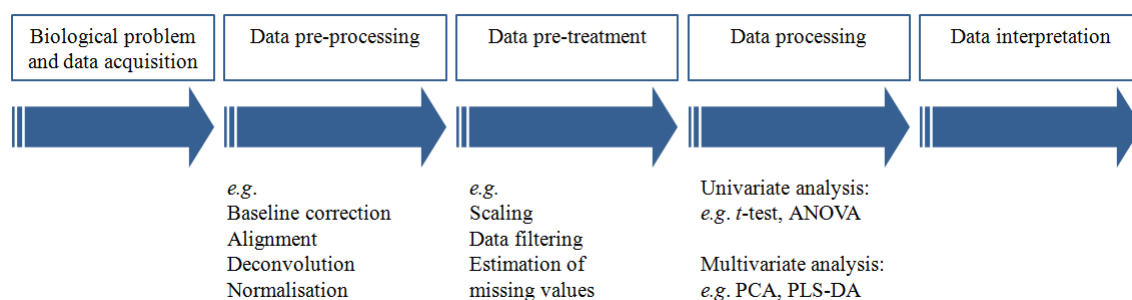


Figure 7: Data analysis workflow in metabolomics. Metabolomics investigations involve a number of steps performed before the pre-defined hypothesis is answered and include data pre-processing, pre-treatment, processing and interpretation.

1.5.1. Pre-processing

The goal of pre-processing of metabolomic data is to create 2D datasets of features, where often the rows correspond to the samples analysed and the columns to m/z values, retention time or peak area, depending on the technique used. There is a number of pre-processing steps applied such as baseline correction, alignment, deconvolution and normalisation (Sussulini 2017, Khakimov *et al.* 2014).

Baseline correction compromises randomly and uniformly generated variations caused by artefacts, which can appear during the measurement. Similarly, the aim of data alignment is to remove any shifts present and to group detected peaks across all samples according to their m/z value and/or retention time window. Consequently, the grouped peaks are then

integrated as peak intensity or peak area and are assigned to a specific feature in the data table (Sussulini, 2017).

The sources of artefacts and shifts between m/z values and/or peak retention times occurring during the experiment are variations in *e.g.* pH, temperature, pressure or effects coming from the sample matrix. However, the majority of chromatographic and mass spectrometry data visualisation and processing software such as DataAnalysis from Bruker or MassLynx from Waters provide built-in baseline correction and alignment functions (Khakimov *et al.* 2014). Additionally, other software *e.g.* MATLAB also enables fast baseline correction, alignment and further processing across many samples simultaneously (Johnsen *et al.*, 2013).

Deconvolution is related to the separation of the true profiles of closely eluted and/or overlapping metabolite peaks. This process is a powerful tool for enhancing selectivity of offered technologies and is important for quantitative investigations. Deconvolution is often performed in automated software packages provided with most GC-MS and LC-MS instruments (*e.g.* Pegasus, Leco, St. Josephs, USA) but can also be achieved with the application of other tools, for example Automated Mass Spectral Deconvolution and Identification System (AMDIS) (for GC-MS) (Hansen, 2006).

Normalisation is applied to the dataset to account for any non-sample related variations caused *e.g.* during sample preparation or data acquisition and is particularly important for quantitative investigations. Normalisation corrects variations by identifying some factor that is alike in all samples and fixing the scale of each variable using the factor characteristics (*e.g.* internal standard) (van den Berg *et al.*, 2006). In addition, normalisation is applied to each spectrum separately rather than across a group of spectra. The most commonly applied method of normalisation is sum normalisation, where each mass channel (for example) is divided by a sum of the intensities from all mass channels in the spectrum (Henderson, 2013).

1.5.2. Pre-treatment

Within all steps in the data analysis, data pre-treatment is believed to be a key step due to its ability to make the data clearer and suited for the further analysis (Goodacre *et al.*, 2007). During the pre-treatment step, the data are converted to minimise variable redundancy, resulting in all variables being more comparable with regard to size (Bro and Smilde, 2003).

There are a number of pre-treatment methods that have been applied in metabolomic data analysis, for example scaling (*e.g.* auto-scaling, Pareto scaling), data filtering (*e.g.* by

standard deviation (SD), median) and estimation of missing values resulting in the imputation of *e.g.* zero values or values using *k* nearest neighbours (kNN) (Povey *et al.* 2014, Gromski *et al.* 2015, Khakimov *et al.* 2014). All pre-treatment methods have their advantages and limitations, so careful consideration has to be given while selecting which method to use (van den Berg *et al.*, 2006).

Scaling allows for an equal importance to all variables, removes any scale differences and is usually performed after normalisation step (Sussulini, 2017). During scaling process, for each variable (column) its mean from the whole dataset is taken (also known as centering) followed by dividing by a scaling factor. Depending on the type of scaling *e.g.* auto-scaling, Pareto scaling *etc.* the scaling factor is different. For example, during auto-scaling the standard deviation of a given variable (column) is used, whereas Pareto scaling instead of standard deviation uses the square root of the standard deviation (van den Berg *et al.*, 2006).

Other common steps in data pre-treatment are data filtering and estimation of missing values. Data filtering is a process where the signal-to-noise ratio for the spectra can be improved by taking out information that is confusing *e.g.* outliers. The majority of the existing data filtering techniques are based on pre-defined window filters, based on *e.g.* standard deviation values, where standard deviation for each variable from all dataset is calculated and checked against given window filter (Hansen, 2006).

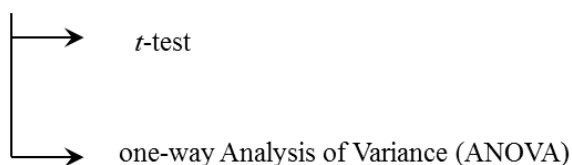
Additionally, it is advised to apply the process of replacing missing data with substituted values (also known as imputation methods) in order to improve data analysis. Quite commonly, about 10%-20% data from *e.g.* GC-MS experiments are missing (Gromski *et al.* 2014, Hrydziuszko & Viant 2012). There is a variety of sources for missing values such as measurement error, failure in the identification of signals from the background or deconvolution, resulting in false negative during segregation of overlapping signals (Hrydziuszko & Viant 2012, Steuer *et al.* 2007). Likewise, a number of imputation methods to estimate missing values can be applied, for example replacing missing value by means of nearest neighbours (kNN) or replacing missing values with zeros (Weckwerth, 2007). Regarding recent studies (Gromski *et al.* 2015, Hrydziuszko 2012), Pareto scaling and kNN imputation method have been found as optimal for the dataset generated using mass spectrometry techniques.

All above pre-treatment methods have their advantages and disadvantages and can greatly influence the outcome from the data analysis. Therefore, a consideration has to be given, which pre-treatment approach is best to select for the dataset, according to the study and type of sample (Gromski *et al.*, 2014).

1.5.3. Univariate and multivariate approaches in data analysis for metabolomics

Two main approaches of data analysis are typically applied in metabolomics investigations: univariate data analysis (UVA) and multivariate data analysis (MVA) (Figure 8).

Univariate data analysis (UVA)



Multivariate data analysis (MVA)

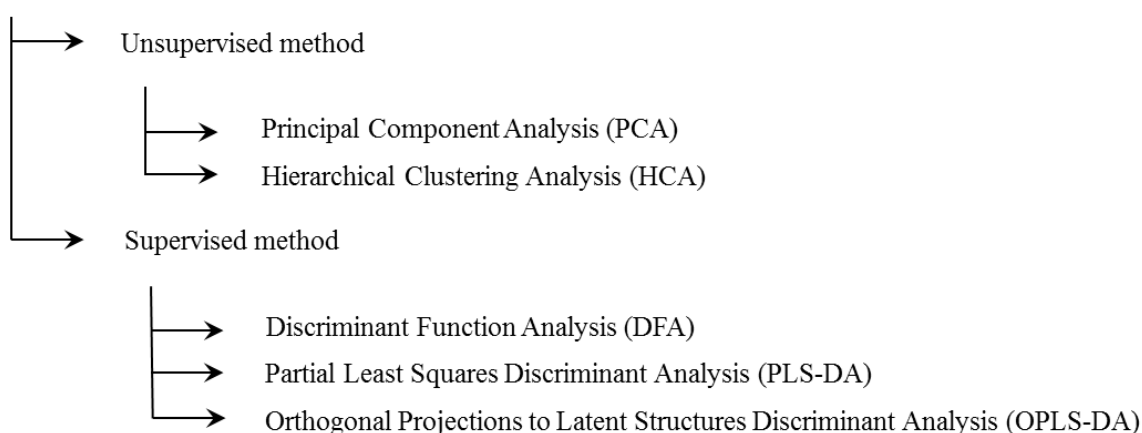


Figure 8: Examples of univariate (UVA) and multivariate (MVA) approaches in data analysis for metabolomics. Commonly used methods for UVA are *t*-test and ANOVA. MVA can be divided into unsupervised and supervised methods that include *e.g.* PCA, HCA (unsupervised) and *e.g.* DFA, PLS-DA, OPLS-DA (supervised) analyses (Gonzalez-Riano, Garcia and Barbas, 2016, Gromski *et al.*, 2014).

In this study, two multivariate analytical tools *i.e.* PCA and PLS-DA were applied and therefore, these will be described in a greater detail.

The UVA approach enables analysis to determine whether or not, the values measured for a specific variable (*e.g.* peak) differ significantly across collection of samples. There are two commonly used methods of UVA, both with an assumption that the datasets follows a normal distribution: *t*-test, applied to determine a comparison between two groups of sample and, one-way Analysis of Variance (ANOVA) test, when more than two groups of samples are given (Gonzalez-Riano, Garcia and Barbas, 2016). Both methods provide a *p*-value that corresponds to the probability of getting by chance a result as that observed (Bartel, Krumsiek and Theis, 2013).

UVA is typically used along with MVA that takes into account more than one variable at the same time. MVA includes principal component analysis (PCA) that is usually applied for a preliminary assessment of original clustering (grouping) of the samples. PCA, an unsupervised method, determines clustering of samples without *a priori* information of a sample class and is based on the differences between signals in the metabolite profiles. PCA is often applied as a starting point in the data analysis (Khakimov, Bak and Engelsen, 2014).

Another method, partial least squares regression discriminant analysis (PLS-DA) is applied for classification between groups but also to evaluate which variables cause the variation and is employed to expose this variation among a collection of samples (Gonzalez-Riano, Garcia and Barbas, 2016).

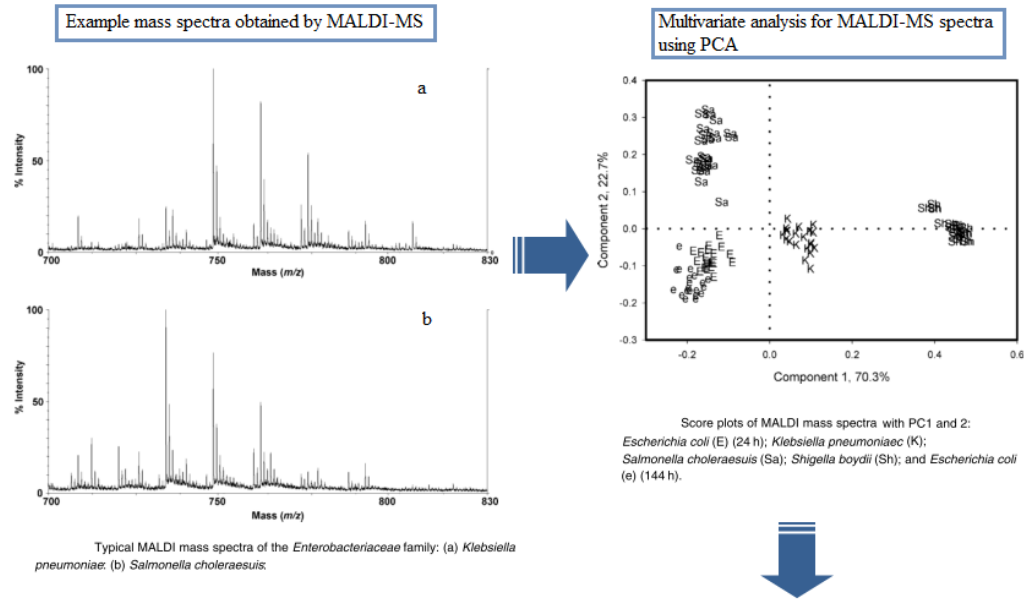
PLS-DA, a supervised method, proposes the construction of predictive models based on the estimated relationships among variables (*i.e.* regression) of the data within the dataset against the class information within this dataset (*e.g.* control, treated). PLS-DA allows for a selection of the statistically significant metabolites that contribute most to the arranged model (Gromski, Muhamadali, *et al.*, 2015).

One way to present statistically significant metabolites is *e.g.* Variable Importance in Projection (VIP), where the most meaningful metabolites are assumed to be the ones with a VIP score ≥ 1 (Godzien *et al.*, 2013).

The data analysis work-flow will result with a set of statistically significant metabolites, allowing for the identification of the metabolic pathways influenced by the investigated factor or certain environmental conditions.

A number of metabolomic studies use the application of univariate and multivariate data analysis work-flows. For instance, PCA was used in discriminating phospholipids from different types of bacteria with the employment of the MALDI-MS technique (**Figure 9 A**) (Ishida *et al.*, 2002). Another example incorporates the application of PCA and PLS-DA multivariate analyses followed by the analysis of metabolites IDs and their metabolic pathways in the study of urine samples from hepatitis C virus (HCV) patients by ultra-performance liquid chromatography MS (UPLC–MS) for biomarker discovery (**Figure 9 B**) (Zhang *et al.*, 2013).

Example A



Discrimination of different bacteria types based on phospholipids content

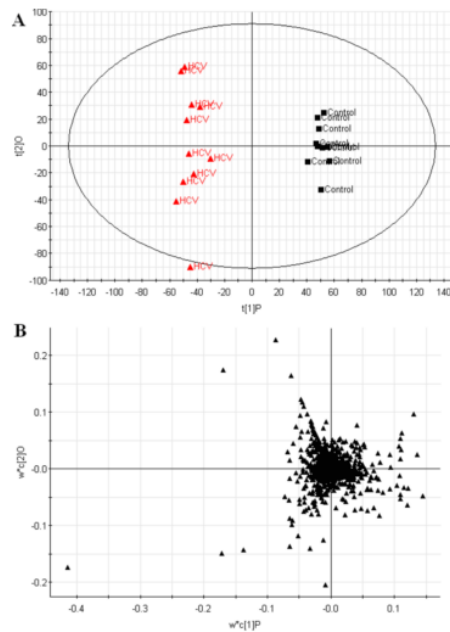
Relative peak intensities of phospholipids observed on the MALDI mass spectra of bacteria					
Class of phospholipids*	Abbreviation	Relative peak intensity (%)			
		<i>E. coli</i>	<i>K. pneumoniae</i>	<i>S. choleraesuis</i>	<i>S. boydii</i>
Phosphatidyl ethanolamines (PEs)	C30:0	3.4	6.2	4.1	5.2
	C31:1	2.5	2.4	1.4	6.2
	C32:1	9.3	14.5	5.9	24.1
	C32:0	5.4	4.8	4.8	6.0
	C33:2	1.7	2.2	1.4	2.4
	C33:1	33.2	26.0	24.1	18.7
	C34:2	3.3	5.1	4.5	6.6
	C34:1	11.7	11.6	17.8	12.5
	C35:2	4.6	5.0	5.9	2.9
	C35:1	11.2	7.5	11.6	2.8
	C36:2	3.2	3.4	3.3	2.9
	(subtotal)	(87.9)	(88.7)	(84.7)	(90.3)
Phosphatidyl glycerols (PGs)	C33:1	4.2	3.7	4.8	3.3
	C34:1	3.8	3.7	4.6	4.0
	C35:1	2.7	2.3	4.4	1.1
	C36:2	1.4	1.6	1.5	1.3
	(subtotal)	(12.1)	(11.3)	(15.3)	(9.7)
Total		100.0	100.0	100.0	100.0

* Observed as $[M + 2Na^+ + H^+]$ and $[M + Na^+]$ ions for PEs and PGs, respectively.

Figure 9 A: Application of multivariate analysis in metabolomics studies. Example A: discriminating phospholipids from different types of bacteria using MALDI-MS; adopted and reproduced from (Ishida *et al.*, 2002).

Example B

Multivariate analysis for UPLC-MS data using PCA and PLS-DA



Metabolomic profiling of HCV patient and matched healthy control in positive ionization mode by using UPLC-MS system: (A) PCA model results for the HCV group in positive mode; (B) loading plot of OPLS-DA of HCV in positive mode

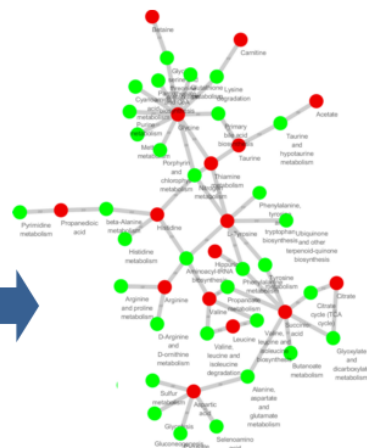
Identification of urinary biomarkers

Identification of urinary biomarkers in HCV cases.

No.	Rt	Name	Trend	VIP
1	9.38	Glycine	↑	18.0218
2	4.49	Succinic acid	↓	11.79101
3	9.17	Acetate	↑	7.47369
4	8.46	Hippuric	↑	7.36361
5	5.46	Carnitine	↓	6.28862
6	3.95	Citrate	↑	6.00245
7	7.9	Ethyl citrate	↓	5.91743
8	3.41	Arginine	↑	5.70033
9	2.58	Betaine	↑	5.28267
10	6.47	histidine	↑	5.26471
11	5.53	Indolebutyric acid	↓	5.02477
12	5.46	Leucine	↑	9.01045
13	3.39	Valine	↓	8.20191
14	3.60	N-Benzoylglycine ethyl ester	↑	7.23011
15	8.08	Phthalic acid	↑	6.71478
16	3.77	Octanedioic acid	↑	6.31235
17	8.23	Propanedioic acid	↑	5.72351
18	5.58	Aspartic acid	↑	5.6574
19	8.43	Taurine	↓	5.34362
20	1.72	L-Tyrosine	↓	5.03689

Note: ↑, content increased; ↓, content decreased

Metabolic pathways influenced by HCV based on metabolite IDs highlighted in multivariate analysis



Metabolites altered in HCV subjects compared to healthy controls map to multiple biosynthetic pathways. Altered metabolites with KEGG from the merged data set were mapped to KEGG and SMPDB reference pathways, and interaction networks were generated in Cytoscape. Green and red nodes represent pathways and related metabolites detected, respectively (details in Table S2 of the Supporting Information).

Figure 9 B: Application of multivariate analysis in metabolomics studies. Example B: investigation of biomarkers for hepatitis C virus from urine samples of hepatitis C virus patients using UPLC-MS; adopted and reproduced from (Zhang *et al.*, 2013).

1.5.4. Processing and analysis of MSI data

The very large size and complex nature of MSI datasets has led to the development of a variety of different pre-processing and data analysis tools (Jones *et al.*, 2012) however, many of them are not yet available commercially. In general, MSI datasets can be processed at the spectral level and at the image level. Pre-processing of spectral dataset can involve background subtraction, alignment and normalisation (Norris *et al.*, 2007). Due to MSI spectra typically contain high-frequency noise and unresolved background, a background subtraction method is important for accurate representation of molecular distribution in MS images. Alignment and normalisation of spectra are applied to correct for some variations in ionisation due to *e.g.* suppression effects or matrix effect (for MALDI technique). One commonly used normalisation is the total ion count (TIC) normalisation that accounts for total effects across the sample. In addition to spectral noise, image noise can also appear due to the stochastic events (*i.e.* random) during the analysis. However, recent development of MSI data analysis packages have introduced algorithms to minimise the impact of noise within an MS image. They also provide tools for data pre-processing, image manipulation *e.g.* region of interest selection and means for export of MS spectra from individual pixels to other software for further analysis (Jones *et al.* 2012, Gessel *et al.* 2014).

Ultimately, a consideration must be given to the pre-processing methods and the analysis of MSI results in order to ensure the most accurate representation of the data (Gessel *et al.* 2014).

1.6. Summary

One of the areas of application for metabolomics is the production of biopharmaceuticals. One example are Fc-fusion proteins, which are the largest group of currently created fusion proteins due to their beneficial biological and pharmacological properties. The manufacturing of Fc-fusion proteins is dependent on an effective, robust and low-cost protein mammalian cell extraction system. Insights into of how cells react to a protein expression process based on metabolic profiles, brings new aims for bioengineering of cell platforms for best performance in bio-manufacturing. Additionally, information on how single cell metabolism and hence, phenotype, responds to the cell microenvironment will allow for biological mechanisms to be examined in detail.

Different mass spectrometry techniques applied to metabolomics studies offer various advantages and limitations in terms of sample preparation procedures, mass range detected, mass accuracy and spatial resolution. At the same time, it is very unlikely that one individual method will give complete set of the answers to untargeted or targeted biological questions. Multidimensional investigations seem to be the best approach to obtain as much of such information as possible from one sample type. With this approach, data sets collected from different mass spectrometry techniques *e.g.* GC-MS, MALDI-MS and SIMS as well as combined chemical and spectral information can be used to explore cell metabolism and its changes in greater detail.

The work in this thesis attempts to investigate the effects of the Fc-based fusion protein production in mammalian cells using a metabolomic approach and with the application of mass spectrometry techniques *i.e.* GC-MS, MALDI-MS and SIMS for profiling and imaging.

Chapter 2 describes the cell line used for Fc-based fusion protein expression, its maintenance and procedures for footprint and fingerprint sample collection.

Chapter 3 describes in detail the GC-MS, MALDI-MS and SIMS instruments used to obtain the metabolomics data.

Chapter 4 presents results of preliminary analysis of a response of CHO cell line exposed to the EPO-Fc fusion protein induction and production.

- *Aims:* to check the stability of the cell line under inducer exposure; to confirm the EPO-Fc fusion protein production in cells; to choose time-points collection for footprint and fingerprint samples

Chapter 5 presents results of optimisation of the sample preparation procedure for the MALDI-MS technique used for a detection of metabolites derived from footprint and fingerprint cell samples.

- *Aims:* to choose a MALDI-MS matrix, matrix concentration and its solvent for metabolite profiling; to create a calibration mixture for MALDI-MS metabolite profiling; to choose a sample deposition method for MALDI-MS metabolite profiling; to confirm the suitability of chosen MALDI-MS sample preparation method for metabolite samples

Chapter 6 presents results of the analysis of footprint and fingerprint metabolite profiles of cells producing the EPO-Fc fusion protein obtained by GC-MS and MALDI-MS. The analysis attempts to investigate cellular activity of CHO cells and their metabolic state upon the protein production.

- *Aims:* to analyse the effect of EPO-Fc fusion protein production on metabolism of cells based on GC-MS and MALDI-MS data; to find the most significant metabolite changes in cells producing EPO-Fc using UVA and MVA approaches; to attempt metabolite assignment and metabolic pathway analysis; to compare data of metabolic profiles collected using GC-MS and MALDI-MS techniques

Chapter 7 presents metabolite MS imaging approach of CHO cells with the application of MALDI-MSI and SIMS imaging in order to investigate the cell response to the EPO-Fc fusion protein expression. The analysis attempts to characterise cells at the population level as well as single cell and sub-cellular levels.

- *Aims:* to analyse the effect of EPO-Fc fusion protein production in cells based on metabolite images from cell populations; to find the most significant metabolite changes in cells producing EPO-Fc using UVA and MVA approaches; to attempt metabolite visualisation and metabolite assignment; to attempt investigation of metabolite changes under protein production at the single-cell level; to attempt investigation of metabolite distribution at the sub-cellular level; to compare data obtained using MALDI-MSI and SIMS imaging

Chapter 8 is a general discussion of the data presented within this thesis and includes a summary of results and ideas for the future work.

References:

- 3D OrbiSIMS (2017). Available at: <http://www.npl.co.uk/news/3d-nanosims-label-free-molecular-imaging> (Accessed: 20 April 2017).
- Adriaensen, L., Vangaever, F. and Gijbels, R. (2004) 'Metal-assisted secondary ion mass spectrometry: Influence of Ag and Au deposition on molecular ion yields', *Anal Chem*, 76, pp. 6777–6785.
- Agin, A., Heintz, D., Ruhland, E., Chao de La Barca, J. M., Zumsteg, J., Moal, V., Gauchez, A. S. and Namer, I. J. (2016) 'Metabolomics – an overview. From basic principles to potential biomarkers', *Médecine Nucléaire*, 40, pp. 4–10.
- Aichler, M. and Walch, A. (2015) 'MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice', *Lab Invest*, 95, pp. 422–431.
- AlMasoud, N., Xu, Y., Trivedi, D. K., Salivo, S., Abban, T., Rattray, N. J. W., Szula, E., AlRabiah, H., Sayqal, A. and Goodacre, R. (2016) 'Classification of Bacillus and Brevibacillus species using rapid analysis of lipids by mass spectrometry', *Anal Bioanal Chem*, 408, pp. 7865–7878.
- Armitage, E. G., Kotze, H. L., Fletcher, J. S., Henderson, A., Williams, K. J., Lockyer, N. P. and Vickerman, J. C. (2013) 'Time-of-flight SIMS as a novel approach to unlocking the hypoxic properties of cancer', *Surf Interface Anal*, 45, pp. 282–285.
- Bag, S., Dutta, D., Chaudhary, A., Chandra Sing, B., Banerjee, R., Pal, M., Paul, R. R., Basak, A., Das, A. K., Ray, A. K. and Chatterjee, J. (2016) 'NanoLC MALDI MS/MS based quantitative metabolomics reveals the alteration of membrane biogenesis in oral cancer', *RSC Adv*, 6, pp. 62420–62433.
- Baldwin, M. A. (2005) 'Mass Spectrometers for the Analysis of Biomolecules', *Methods Enzymol*, 402, pp. 3–48.
- Balogh, M. P. (2004) 'Debating resolution and mass accuracy in mass spectrometry', *LC-GC N. Am.*, 22, p. 118.
- Bartel, J., Krumsiek, J. and Theis, F. J. (2013) 'Statistical methods for the analysis of high-throughput metabolomics data', *Comput Struct Biotechnol J*, 4, pp. 1–9.
- Begley, P., Francis-McIntyre, S., Dunn, W. B., Broadhurst, D. I., Halsall, A., Tseng, A., Knowles, J., Goodacre, R. and Kell, D. B. (2009) 'Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum', *Anal Chem*, 81, pp. 7038–7046.
- van den Berg, R., Hoefsloot, H., Westerhuis, J., Smilde, A. and van der Werf, M. (2006) 'Centering, scaling, and transformations: improving the biological information content of metabolomics data', *BMC Genomics*, 7, pp. 1–15.
- Bhandari, D. R. am, Schott, M., Römpf, A., Vilcinskas, A. and Spengler, B. (2015) 'Metabolite localization by atmospheric pressure high-resolution scanning microprobe matrix-assisted laser desorption/ionization mass spectrometry imaging in whole-body sections and individual organs of therove beetle *Paederus riparius*', *Anal Bioanal Chem*, 407, pp. 2189–2201.
- Birch, J. R. and Racher, A. J. (2006) 'Antibody production', *Adv Drug Deliv Rev*, 58, pp. 671–685.

- Bitonti, A. J., Dumont, J. A., Low, S. C., Peters, R. T., Kropp, K. E., Palombella, V. J., Stattel, J. M., Lu, Y., Tan, C. A., Song, J. J., Garcia, A. M., Simister, N. E., Spiekermann, G. M., Lencer, W. I. and Blumberg, R. S. (2004) 'Pulmonary delivery of an erythropoietin Fc fusion protein in non-human primates through an immunoglobulin transport pathway.', *Proc Natl Acad Sci U S A*, 101, pp. 9763–8.
- Boughton, B. A., Thinagaran, D., Sarabia, D., Bacic, A. and Roessner, U. (2016) 'Mass spectrometry imaging for plant biology: a review', *Phytochem Rev*, 15, pp. 445–488.
- Bro, R. and Smilde, A. K. (2003) 'Centering and scaling in component analysis', *J Chemometrics*, 17, pp. 16–33.
- Brown, M., Dunn, W. B., Dobson, P., Patel, Y., Winder, C. L., Francis-McIntyre, S., Begley, P., Carroll, K., Broadhurst, D. I., Tseng, A., Swainston, N., Spasic, I., Goodacre, R. and Kell, D. B. (2009) 'Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics', *The Analyst*, 134, pp. 1322–32.
- Brown, S., Kruppa, G. and Dasseux, J. (2005) 'Metabolomics applications of FT-ICR mass spectrometry', *Mass Spectrom Rev*, 24, pp. 223–231.
- Bruker (2016). Available at:
http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf (Accessed: 28 October 2016).
- Bujak, R., Struck-Lewicka, W., Markuszewski, M. J. and Kaliszan, R. (2015) 'Metabolomics for laboratory diagnostics', *J Pharma Biomed Anal*, 113, pp. 108–120.
- Bunch, J., Clench, M. and Richards, D. (2004) 'Determination of pharmaceutical compounds in skin by imaging matrix-assisted laser desorption/ionisation mass spectrometry', *Rapid Commun Mass Spectrom*, 18, pp. 3051–3060.
- Calvano, C. D., Zamboni, C. G. and Palmisano, F. (2011) 'Lipid fingerprinting of Gram-positive lactobacilli by intact cells - Matrix-assisted laser desorption/ionization mass spectrometry using a proton sponge based matrix', *Rapid Comm Mass Spectrom*, 25, pp. 1757–1764.
- Campbell, D. I., Ferreira, C. R., Eberlin, L. S. and Cooks, R. G. (2012) 'Improved spatial resolution in the imaging of biological tissue using desorption electrospray ionization', *Anal Bioanal Chem*, 404, pp. 389–398.
- Cao, D., Wang, Z., Han, C., Cui, L., Hu, M., Wu, J., Liu, Y., Cai, Y., Wang, H. and Kang, Y. (2011) 'Quantitative detection of trace perfluorinated compounds in environmental water samples by matrix-assisted laser desorption/ionization-time of flight mass spectrometry with 1,8-bis(tetramethylguanidino)-naphthalene as matrix', *Talanta*, 85, pp. 345–352.
- Carter, P. (2011) 'Introduction to current and future protein therapeutics: a protein engineering perspective.', *Exp Cell Res*, 317, pp. 1261–1269.
- Cersoy, S., Richardin, P., Walter, P. and Brunelle, A. (2012) 'Cluster TOF-SIMS imaging of human skin remains: analysis of a South-Andean mummy sample', *J Mass Spectrom*, 47, pp. 338–346.
- Chandra, S., Tjarks, W., Lorey, D. R. and Barth, R. F. (2008) 'Quantitative Subcellular Imaging of Boron Compounds in Individual Mitotic and Interphase Human Glioblastoma Cells With Imaging Secondary Ion Mass Spectrometry (SIMS)', *J Microsc*, 229, pp. 92–

Chen, Y., Zhang, J., Guo, L., Liu, L., Wen, J., Xu, L., Yan, M., Li, Z., Zhang, X., Nan, P., Jiang, J., Ji, J., Zhang, J., Cai, W., Zhuang, H., Wang, Y., Zhu, Z. and Yu, Y. (2016) 'A characteristic biosignature for discrimination of gastric cancer from healthy population by high throughput GC-MS analysis.', *Oncotarget*, 7, pp. 87496–87510.

Chimeric protein-protein interaction server (2017). Available at: <http://chippi.md.biu.ac.il/> (Accessed: 20 April 2017).

Cisek, K., Krochmal, M., Klein, J. and Mischak, H. (2016) 'The application of multi-omics and systems biology to identify therapeutic targets in chronic kidney disease', *NDT*, 31, pp. 2003–2011.

Clench, M. R. (2016) 'Advances in mass spectrometry imaging', *Proteomics*, 16, pp. 1605–1606.

Cobice, D. F., Goodwin, R. J. A., Andren, P. E., Nilsson, A., Mackay, C. L. and Andrew, R. (2015) 'Future technology insight: Mass spectrometry imaging as a tool in drug research and development', *Br J Pharmacol*, 172, pp. 3266–3283.

Cooks, R. G. and Busch, K. L. (1983) 'Matrix effects, internal energies and MS/MS spectra of molecular ions sputtered from surfaces', *Int J Mass Spectrom Ion Phys*, 53, pp. 111–124.

Cooks, R., Ouyang, Z., Takats, Z. and Wiseman, J. (2006) 'Detection Technologies. Ambient mass spectrometry.', *Science*, 311, pp. 1566–1570.

Cuperlovic-Culf, M. and Culf, A. (2016) 'Applied metabolomics in drug discovery', *Expert Opin Drug Discov*, 11, pp. 759–770.

Czajkowsky, D. M., Hu, J., Shao, Z. and Pleass, R. J. (2012) 'Fc-fusion proteins: new developments and future perspectives', *EMBO Mol Med*, 4, pp. 1015–1028.

Dao, L., Weiland, U., Hauser, M., Nazarenko, I., Kalt, H., Bastmeyer, M. and Franz, C. M. (2012) 'Revealing non-genetic adhesive variations in clonal populations by comparative single-cell force spectroscopy', *Exp Cell Res*, 318, pp. 2155–2167.

Dass, C. (2006) 'Modes of Ionization', in *Fundamentals of Contemporary Mass Spectrometry*. John Wiley & Sons, Inc., pp. 15–65.

Davis, J. M., Arakawa, T., Strickland, T. W. and Yphantis, D. A. (1987) 'Characterization of Recombinant Human Erythropoietin Produced in Chinese Hamster Ovary Cells?', *Biochemistry*, 26, pp. 2633–2638.

Debois, D., Bralet, M.-P., Le Naour, F., Brunelle, A. and Laprévote, O. (2009) 'In Situ Lipidomic Analysis of Nonalcoholic Fatty Liver by Cluster TOF-SIMS Imaging', *Anal Chem*, 81, pp. 2823–2831.

Denbigh, J. L. and Lockyer, N. P. (2015) 'ToF-SIMS as a tool for profiling lipids in cancer and other diseases', *Mat Sci Technol*, 31, pp. 137–147.

Dietmair, S., Hodson, M. P., Quek, L. E., Timmins, N. E., Chrysanthopoulos, P., Jacob, S. S., Gray, P. and Nielsen, L. K. (2012) 'Metabolite profiling of CHO cells with different growth characteristics', *Biotechnol Bioeng*, 109, pp. 1404–1414.

Dinis-Oliveira, R. (2016) 'Metabolomics of methadone: clinical and forensic toxicological

- implications and variability of dose response', *Drug Metabol Rev*, 48, pp. 568–576.
- Divol, B. and Bauer, F. (2010) 'Metabolic engineering of wine yeast and advances in yeast selection methods for improved wine quality', in Reynolds, A. G. (ed.) *Managing wine quality: Oenology and Wine Quality*. Cambridge: Woodhead Publishing Ltd., pp. 34–59.
- Dreisewerd, K. (2003) 'The Desorption Process in MALDI', *Chem Rev*, 103, pp. 56–65.
- Drexler, D. M., Reily, M. D. and Shipkova, P. A. (2011) 'Advances in mass spectrometry applied to pharmaceutical metabolomics', *Anal Bioanal Chem*, 399, pp. 2645–2653.
- Duncan, M. W., Roder, H. and Hunsucker, S. W. (2008) 'Quantitative matrix-assisted laser desorption/ionization mass spectrometry', *Briefings Funct Genomic Proteomics*, 7, pp. 355–370.
- Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005) 'Measuring the metabolome: current analytical technologies', *The Analyst*, 130, pp. 606–625.
- Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., Nicholls, A. W., Wilson, I. D., Kell, D. B. and Goodacre, R. (2011) 'Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry', *Nat prot*, 6, pp. 1060–1083.
- Edwards, J. L. and Kennedy, R. T. (2005) 'Metabolomic analysis of eukaryotic tissue and prokaryotes using negative mode MALDI time-of-flight mass spectrometry', *Anal Chem*, 77, pp. 2201–2209.
- El-Faramawy, A., Siu, K. W. M. and Thomson, B. A. (2005) 'Efficiency of Nano-Electrospray Ionization', *J Am Soc Mass Spectrom*, 16, pp. 1702–1707.
- Emara, S., Amer, S., Ali, A., Abouleila, Y., Oga, A. and Masujima, T. (2017) 'Single-Cell Metabolomics', *Adv Exp Med Biol*, 965, pp. 323–343.
- Fell, D. A. (1996) *Understanding the Control of Metabolism*. London: Portland Press.
- Ferreira, C., Wu, L., Vogt, F., Bornancini, E. and Cooks, R. (2014) 'Fiducial markers for distribution of drug and excipient on tablet surfaces by multimodal desorption electrospray ionisation mass spectrometry (DESI-MS) imaging.', *Anal Letters*, 47, pp. 91–101.
- Fiehn, O. (2001) 'Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks', *Comp Function Genom*, 2, pp. 155–168.
- Fletcher, J. S., Rabbani, S., Barber, A. M., Lockyer, N. P. and Vickerman, J. C. (2013) 'Comparison of C60 and GCIB primary ion beams for the analysis of cancer cells and tumour sections', *Surf Interface Anal*, 45, pp. 273–276.
- Fletcher, J. S. and Szakal, C. (2013) 'Cluster and polyatomic primary ion beams.', in Vickerman, J. C. and Briggs, D. (eds) *ToF-SIMS: Materials analysis by mass spectrometry*. 2nd edn. West Sussex, Manchester: IM Publications LLP and SurfaceSpectra Limited, pp. 291–310.
- Forster, J., Famili, I., Palsson, B. O. and Nielsen, J. (2003) 'Genome-Scale Reconstruction of the *Saccharomyces Cerevisiae* Metabolic Network', *Genome Res*, 13, pp. 244–253.
- Franz Hillenkamp and Karas, M. (2007) 'The MALDI Process and Method', in Franz Hillenkamp and Jasna Peter-Katalinic (ed.) *MALDI MS. A Practical Guide to*

Instrumentation, Methods and Applications. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Fülöp, A., Sammour, D. A., Erich, K., von Gerichten, J., van Hoogevest, P., Sandhoff, R. and Hopf, C. (2016) 'Molecular imaging of brain localization of liposomes in mice using MALDI mass spectrometry', *Sci Rep*, 6, p. 33791.

Gates, S. C. and Sweeley, C. C. (1978) 'Quantitative metabolic profiling based on gas chromatography', *Clin Chem*, 24, pp. 1663–1673.

George, K., Alonso-Gutierrez, J., Keasling, J. and Lee, T. (2015) 'Isoprenoid Drugs, Biofuels, and Chemicals-Artemisinin, Farnesene, and Beyond', *Biotech Isoprenoids*, 148, pp. 355–389.

Gessel, M. M., Norris, J. L. and Caprioli, R. M. (2014) 'MALDI imaging mass spectrometry: Spatial molecular analysis to enable a new age of discovery', *J Proteomics*, 107, pp. 71–82.

Ghaste, M., Mistrik, R. and Shulaev, V. (2016) 'Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics', *Inter J Mol Sci*, 17, pp. 56–66.

Girod, M., Shi, Y., Cheng, J.-X. and Cooks, R. G. (2011) 'Mapping Lipid Alterations in Traumatically Injured Rat Spinal Cord by Desorption Electrospray Ionization Imaging Mass Spectrometry', *Anal Chem*, 83, pp. 207–215.

Godzien, J., Ciborowski, M., Angulo, S. and Barbas, C. (2013) 'From numbers to a biological sense: How the strategy chosen for metabolomics data treatment may affect final results. A practical example based on urine fingerprints obtained by LC-MS', *Electrophoresis*, 34, pp. 2812–2826.

González-Domínguez, R., Castilla-Quintero, R., García-Barrera, T. and Gómez-Ariza, J. L. (2014) 'Development of a metabolomic approach based on urine samples and direct infusion mass spectrometry', *Analytical Biochemistry*, 465, pp. 20–27.

Gonzalez-Dominguez, R Sayago, A. and Fernandez-Recamales, A. (2017) 'Direct infusion mass spectrometry for metabolomic phenotyping of diseases', *BIOANALYSIS*, 9(1).

Gonzalez-Riano, C., Garcia, A. and Barbas, C. (2016) 'Metabolomics studies in brain tissue: A review', *J Pharma Biomed Anal*, 130, pp. 141–168.

Goodacre, R., Broadhurst, D., Smilde, A. K., Kristal, B. S., Baker, J. D., Beger, R., Bessant, C., Connor, S., Capuani, G., Craig, A., Ebbels, T., Kell, D. B., Manetti, C., Newton, J., Paternostro, G., Somorjai, R., Sjöström, M., Trygg, J. and Wulfert, F. (2007) 'Proposed minimum reporting standards for data analysis in metabolomics', *Metabolomics*, 3, pp. 231–241.

Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. and Kell, D. B. (2004) 'Metabolomics by numbers: Acquiring and understanding global metabolite data', *Trends Biotechnol*, 22, pp. 245–252.

Gowda, G. A. N. and Djukovic, D. (2014) 'Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges', *Methods Mol Biol*, 1198, pp. 3–12.

Gowda, G. N., Zhang, S., Gu, H., Asiago, V., Shanaiah, N. and Raftery, D. (2008) 'Metabolomics-based methods for early disease diagnostics', *Exp Revi Mol Diagn*, 8, pp.

Graham, D., Wilson, J., Lai, J., Stayton, P. and Castner, D. (2016) 'Three-dimensional localization of polymer nanoparticles in cells using ToF-SIMS', *Biointerphases*, 11, pp. 1–11.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L. and Goodacre, R. (2015) 'A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding', *Anal Chim Acta*. Elsevier B.V., 879, pp. 10–23.

Gromski, P. S., Xu, Y., Hollywood, K. A., Turner, M. L. and Goodacre, R. (2015) 'The influence of scaling metabolomics data on model classification accuracy', *Metabolomics*. Springer US, 11, pp. 684–695.

Gromski, P. S., Xu, Y., Kotze, H. L., Correa, E., Ellis, D. I., Armitage, E. G., Turner, M. L. and Goodacre, R. (2014) 'Influence of missing values substitutes on multivariate analysis of metabolomics data', *Metabolites*, 4, pp. 433–52.

Halket, J. M., Waterman, D., Przyborowska, A. M., Patel, R. K. P., Fraser, P. D. and Bramley, P. M. (2005) 'Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS', *J Exp Bot*, 56, pp. 219–243.

Hande, K., Anthony, L., Hamilton, R., Kenneth, I., Bennett, R., Sweetman, B. and Branch, R. (1988) 'Identification of etoposide glucuronide as a major metabolite of etoposide in the rat and rabbit', *Cancer Res*, 48, pp. 1829–1834.

Hansen, M. A. E. (2006) 'Data Analysis', in *Metabolome Analysis: An Introduction*. John Wiley & Sons, Inc., pp. 146–187. doi: 10.4135/9781412983846.

Hao, Z., Zhang, Y., Eliuk, S. and Blethrow, J. (2012) *A Quadrupole-Orbitrap Hybrid Mass Spectrometer Offers Highest Benchtop Performance for In-Depth Analysis of Complex Proteomes*. Bremen, Germany.

Hasler-Sheetal, H., Castorani, M., Glud, R., Canfield, D. and Holmer, M. (2016) 'Metabolomics Reveals Cryptic Interactive Effects of Species Interactions and Environmental Stress on Nitrogen and Sulfur Metabolism in Seagrass', *Env Sci Technol*, 50, pp. 11602–11609.

He, L., Winterrowd, C., Kadura, I. and Frye, C. (2012) 'Transgene copy number distribution profiles in recombinant CHO cell lines revealed by single cell analyses', *Biotechnol Bioeng*, 109, pp. 1713–1722.

Heinemann, M. and Zenobi, R. (2011) 'Single cell metabolomics', *Cur Opin Biotechnol*, 22, pp. 26–31.

Henderson, A. (2013) 'Multivariate analysis of SIMS spectra', in Vickerman, J. C. and D. Briggs (eds) *ToF-SIMS: Materials analysis by mass spectrometry*. 2nd edn. West Sussex, Manchester: IM Publications LLP and SurfaceSpectra Limited, pp. 449–483.

Hill, R., Blenkinsopp, P., Thompson, S., Vickerman, J. and Fletcher, J. S. (2011) 'A new time-of-flight SIMS instrument for 3D imaging and analysis', *Surf Interface Anal*, 43, pp. 506–509.

HMDB (2017). Available at: <http://www.hmdb.ca/> (Accessed: 5 January 2017).

- Hoffmann, E. de and Stroobant, V. (2007) *Mass spectrometry. Principles and applications*. 3rd edn. Chichester: John Wiley & Sons Ltd.
- Hollywood, K., Brison, D. R. and Goodacre, R. (2006) 'Metabolomics: Current technologies and future trends', *Proteomics*, 6, pp. 4716–4723.
- Hrydziusko, O. and Viant, M. R. (2012) 'Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline', *Metabolomics*, 8, pp. 161–174.
- Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M. and Cooks, R. G. (2005) 'The Orbitrap: A new mass spectrometer', *J Mass Spectrom*, 40, pp. 430–443.
- Ifa, D. R., Gumaelius, L. M., Eberlin, L. S., Manicke, N. E. and Cooks, R. G. (2007) 'Forensic analysis of inks by imaging desorption electrospray ionization (DESI) mass spectrometry', *The Analyst*, 137, pp. 461–467.
- Ishida, Y., Madonna, A. J., Rees, J. C., Meetani, M. A. and Voorhees, K. J. (2002) 'Rapid analysis of intact phospholipids from whole bacterial cells by matrix-assisted laser desorption/ionization mass spectrometry combined with on-probe sample pretreatment', *Rapid Commun Mass Spectrom*, 16, pp. 1877–1882.
- Izumi, Y., Kajiyama, S., Nakamura, R., Ishihara, A., Okazawa, A., Fukusaki, E., Kanematsu, Y. and Kobayashi, A. (2009) 'High-resolution spatial and temporal analysis of phytoalexin production in oats', *Planta*, 229, pp. 931–943.
- J Pevsner (2009) 'Functional Genomics', in Hoboken and Wiley-Blackwell, N. (eds) *Bioinformatics and functional genomics*. 2nd edn. John Wiley & Sons, Inc, pp. 460–514.
- Jelkmann, W. (2007) 'Erythropoietin after a century of research: younger than ever', *Eur J Haematol*, 78, pp. 183–205.
- Jerigova, M., Biro, C., Kirchnerova, J., Chorvatova, A., Chorvat, D., Lorenc, D. and Velic, D. (2011) 'Chemical Imaging of Cardiac Cell and Tissue by Using Secondary Ion Mass Spectrometry', *Mol Imag Biol*, 13, pp. 1067–1076.
- Jirasko, R., Holcapek, M., Kunes, M. and Svatos, A. (2014) 'Distribution study of atorvastatin and its metabolites in rat tissues using combined information from UHPLC/MS and MALDI-Orbitrap-MS imaging', *Anal Bioanal Chem*, 406, pp. 4601–4610.
- Johari, Y. B., Estes, Y. B., Alves, S. D., Sinacore, C. S. and James, M. S. (2015) 'Integrated Cell and Process Engineering for Improved Transient Production of a "Difficult-to-Express" Fusion Protein by CHO Cells', *Biotechnol Bioeng*, 112, pp. 256–265.
- Johnsen, L. G., Skov, T., Houlberg, U., Bro, R. and Massart, D. L. (2013) 'An automated method for baseline correction, peak finding and peak grouping in chromatographic data', *The Analyst*, 138, pp. 3502–3510.
- Jones, E. A., Deininger, S.-O., Hogendoorn, P. C. W. and Deelder, A. M. (2012) 'Imaging mass spectrometry statistical analysis', *J Proteomics*, 75, pp. 4962–4989.
- Karas, M., Bachmann, D. and Hillenkamp, F. (1985) 'Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules', *Anal Chem*, 57, pp. 2935–2939.

- Khakimov, B., Bak, S. and Engelsens, S. B. (2014) 'High-throughput cereal metabolomics: Current analytical technologies, challenges and perspectives', *J Cer Sci*, 59, pp. 393–418.
- Khamis, M. M., Adamko, D. J. and El-Aneed, A. (2017) 'Mass spectrometric based approaches in urine metabolomics and biomarker discovery', *Mass Spectrom Rev*, 36, pp. 115–134.
- Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S. and Fiehn, O. (2009) 'FiehnLib – Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry', *Anal Chem*, 81, pp. 10038–10048.
- Kiss, A. and Hopfgartner, G. (2016) 'Laser-based methods for the analysis of low molecular weight compounds in biological matrices', *Methods*, 104, pp. 142–53.
- Klyszejko-Stefanowicz, L. (2002) *Cytobiochemia. Biochemia niektórych struktur komórkowych*. Warszawa: PWN.
- Kompauer, M., Heiles, S. and Spengler, B. (2017) 'Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- μ m lateral resolution', *Nat Met*, 14, pp. 1–10.
- Korte, A. R. and Lee, Y. J. (2014) 'MALDI-MS analysis and imaging of small molecule metabolites with 1,5-diaminonaphthalene (DAN)', *J Mass Spectrom*, 49, pp. 737–741.
- Kulkarni, R. (2016) 'Metabolic Engineering: Biological Art of Producing Useful Chemicals', *Resonance*, 21, pp. 233–237.
- Kyriakopoulos, S., Polizzi, K. M. and Kontoravdi, C. (2013) 'Comparative analysis of amino acid metabolism and transport in CHO variants with different levels of productivity', *J Biotechnol*, 168, pp. 543–551.
- Lagarrigue, M., Caprioli, R. M. and Pineau, C. (2016) 'Potential of MALDI imaging for the toxicological evaluation of environmental pollutants', *J Proteomics*, 144, pp. 133–139.
- Lahner, B., Gong, J., Mahmoudian, M., Smith, E. L., Abid, K. B., Rogers, E. E., Guerinot, M. L., Harper, J. F., Ward, J. M., McIntyre, L., Schroeder, J. I. and Salt, D. E. (2003) 'Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*', *Nat Biotechnol*, 21, pp. 1215–1221.
- Lanni, E. J., Rubakhin, S. S. and Sweedler, J. V. (2012) 'Mass spectrometry imaging and profiling of single cells', *J Proteomics*, 75, pp. 5036–5051.
- Lechene, C. P., Luyten, Y., McMahon, G. and Distel, D. L. (2007) 'Quantitative Imaging of Nitrogen Fixation by Individual Bacteria Within Animal Cells', *Science*, 317, pp. 1563–1566.
- Lei, Z., Huhman, D. V. and Sumner, L. W. (2011) 'Mass Spectrometry Strategies in Metabolomics', *J Biol Chem*, 286, pp. 25435–42.
- Lenz, E. M. and Wilson, I. D. (2007) 'Analytical Strategies in Metabonomics', *J Prot Res*, 6, pp. 443–458.
- Lewis, J. K., Wei, J. and Siuzdak, G. (2006) *Matrix-assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis, Encyclopedia of Analytical Chemistry*. USA: The Scripps Research Institute.

- Lindon, J. C. and Nicholson, J. K. (2008) 'Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery', *Trends Anal Chem*, 27, pp. 194–204.
- Liu, Q., Guo, Z. and He, L. (2007) 'Mass spectrometry imaging of small molecules using desorption/ionization on silicon', *Anal Chem*, 79, pp. 3535–3541.
- Lou, S., Balluff, B., Cleven, A. H. G., Bovée, J. V. M. G. and McDonnell, L. A. (2016) 'Prognostic Metabolite Biomarkers for Soft Tissue Sarcomas Discovered by Mass Spectrometry Imaging', *J Am Soc Mass Spectrom*, 52, pp. 376–383.
- McPail, D. and Dowsett, M. (2009) 'Dynamic SIMS', in Vickerman, J. and Gilmore IS (eds) *Surface analysis—the principal techniques*. 2nd edn. Chichester: Wiley, pp. 207–268. doi: 10.1002/9780470721582.
- Mishur, R. and Rea, S. L. (2012) 'Applications of mass spectrometry to metabolomics and metabonomics: detection of biomarkers of aging and of age-related diseases', *Mass Spectrom Rev*, 31, pp. 70–95.
- Mizuno, H., Tsuyama, N., Date, S., Harada, T. and Masujima, T. (2008) 'Live single-cell metabolomics of tryptophan and histidine metabolites in a rat basophil leukemia cell', *Anal Sci*, 24, pp. 1525–1527.
- Mizuno, H., Tsuyama, N., Harada, T. and Masujima, T. (2008) 'Live single-cell video-mass spectrometry for cellular and subcellular molecular detection and cell classification', *J Mass Spectrom*, 43, pp. 1692–1700.
- Moco, S., Bino, R. J., Vos, R. C. H. De and Vervoort, J. (2007) 'Metabolomics technologies and metabolite identification', *Trends Anal Chem*, 26, pp. 963–972.
- Muddiman, D. C., Gusev, A. I., Proctor, A., Hercules, D. M., Venkataramanan, R. and Diven, W. (1994) 'Quantitative measurement of cyclosporin A in blood by time-of-flight mass spectrometry', *Anal Chem*, 66, pp. 2362–2368.
- Müller, M. and Kersten, S. (2003) 'Nutrigenomics: goals and strategies', *Nat Rev Genet*, 4, pp. 315–322.
- Nagels, B., Van Damme, E. J. M., Callewaert, N., Zabeau, L., Tavernier, J., Delanghe, J. R., Boets, A., Castilho, A. and Weterings, K. (2012) 'Biologically active, magnICON-expressed EPO-Fc from stably transformed *Nicotiana benthamiana* plants presenting tetra-antennary N-glycan structures', *J Biotechnol*, 160, pp. 242–250.
- Nicholas P. Lockyer (2014) 'Secondary Ion Mass Spectrometry Imaging of Biological Cells and Tissues', in Kuo, J. (ed.) *Electron Microscopy: Methods and Protocols, Methods in Molecular Biology*. New York: Springer Science+Business Media. doi: 0.1007/978-1-62703-776-1_32.
- Nicholson, J. K., Lindon, J. C. and Holmes, E. (1999) '"Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data', *Xenobiotica*, 29, pp. 1181–1189. doi: 82599238047.
- Nielsen, M. M. B., Lambertsen, K. L., Clausen, B. H., Meyer, M., Bhandari, D. R., Larsen, S. T., Poulsen, S. S., Spengler, B., Janfelt, C. and Hansen, H. S. (2016) 'Mass spectrometry imaging of biomarker lipids for phagocytosis and signalling during focal cerebral ischaemia', *Sci Rep*, 6, pp. 39571–79.

- Nimmerjahn, F. and Ravetch, J. (2008) 'Fcγ receptors as regulators of immune responses.', *Nat Rev Immunol*, 8, pp. 34–47.
- Norris, J. and Caprioli, R. (2013) 'Imaging mass spectrometry: a new tool for pathology in a molecular age', *Proteomics Clin Appl*, 7, pp. 733–738.
- Norris, J. L., Cornett, D. S., Mobley, J. A., Andersson, M., Seeley, E. H., Chaurand, P. and Caprioli, R. M. (2007) 'Processing MALDI mass spectra to improve mass spectral direct tissue analysis', *Int J Mass Spectrom*, 260, pp. 212–221.
- O'Callaghan, P. M., McLeod, J., Pybus, L. P., Lovelady, C. S., Wilkinson, S. J., Racher, A. J., Porter, A. and James, D. C. (2010) 'Cell line-specific control of recombinant monoclonal antibody production by CHO cells', *Biotechnol Bioeng*, 106, pp. 938–951.
- Oliver, S. G., Winson, M. K., Kell, D. B. and Baganz, F. (1998) 'Systematic functional analysis of the yeast genome', *Tibtech*, 8, pp. 373–378.
- Oran, U. (2004) 'Surface Chemical Characterization of Plasma-Chemically Deposited Polymer Films by Time of Flight Static Secondary Ion Mass Spectrometry', *Plasma Process Polym*, 1, pp. 123–133.
- Ostrowski, S. G., Kurczy, M. E., Roddy, T. P., Winograd, N. and Ewing, A. G. (2007) 'Secondary Ion MS Imaging To Relatively Quantify Cholesterol in the Membranes of Individual Cells from Differentially Treated Populations', *Anal Chem*, 79, pp. 3554–3560.
- Passarelli, M. K. and Ewing, A. G. (2013) 'Single-cell imaging mass spectrometry', *Curr Opin Chem Biol*, 17, pp. 854–859.
- Perry, R. H., Cooks, R. G. and Robert J. Noll (2008) 'Orbitrap mass spectrometry: instrumentation, ion motion and applications', *Mass Spectrom Rev*, 27, pp. 661–699.
- Povey, J. F., O'Malley, C. J., Root, T., Martin, E. B., Montague, G. A., Feary, M., Trim, C., Lang, D. A., Alldread, R., Racher, A. J. and Smales, C. M. (2014) 'Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling', *J Biotechnol*, 184, pp. 84–93.
- du Preez, I., Luies, L. and Loots, D. (2017) 'Metabolomics biomarkers for tuberculosis diagnostics: current status and future objectives', *Biomarkers Med*, 11, pp. 179–194.
- Preti, G. (2005) 'Metabolomics comes of age?', *The Scientist*. Available at: <http://www.the-scientist.com/?articles.view/articleNo/16506/title/Metabolomics-comes-of-age-/>.
- Rabbani, S., Fletcher, J. S., Lockyer, N. P. and Vickerman, J. C. (2011) 'Exploring subcellular imaging on the buncher-ToF J105 3D chemical imager', *Surf Interface Anal*, 43, pp. 380–384.
- Raj, A. and van Oudenaarden, A. (2008) 'Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences', *Cell*, 135, pp. 216–226.
- van Ravenzwaay, B., Montoya, G., Fabian, E., Herold, M., Krennrich, G., Looser, R., Mellert, W., Peter, E., Strauss, V., Walk, T. and Kamp, H. (2014) 'The sensitivity of metabolomics versus classical regulatory toxicology from a NOAEL perspective', *Toxicol Lett*, 227, pp. 20–28.

- Richardin, P., Mazel, V., Walter, P., Lapr v te, O. and Brunelle, A. (2011) ‘Identification of Different Copper Green Pigments in Renaissance Paintings by Cluster-TOF-SIMS Imaging Analysis’, *J Am Soc Mass Spectrom*, 22, pp. 1729–1736.
- Roberts, L. D., Souza, A. L., Gerszten, R. E. and Clish, C. B. (2013) ‘Targeted Metabolomics’, *Curr Protoc Mol Biol*, 6, pp. 1–34.
- R m pp, A., Guenther, S., Schober, Y., Schulz, O. and Takats, Z. (2010) ‘Histology by Mass Spectrometry: Label-Free Tissue Characterization Obtained from High-Accuracy Bioanalytical Imaging’, *Angew Chem Int Ed Engl*, 49, pp. 3834–3838.
- R m pp, A. and Spengler, B. (2013) ‘Mass spectrometry imaging with high resolution in mass and space’, *Histochem Cell Biol*, 139, pp. 759–783.
- Roopenian, D. and Akilesh, S. (2007) ‘Fc n: the neonatal Fc receptor comes of age.’, *Nat Rev Immunol*, 7, pp. 715–725.
- Rouberol, J.-M., Guernet, J., Deschamps, P., et Dagnot, J.-P. and de la Berge, J.-M. G. (1969) ‘Microanalyseur par  mission ionique secondaire’, in *Vth International Congress on X-Ray Optics and Microanalysis*. Berlin, Heidelberg, pp. 311–318.
- RSC MS schematic (2017). Available at: http://www.rsc.org/images/MS2new_tcm18-102519.pdf (Accessed: 25 April 2017).
- Rudkowska, I., Paradis, A. ., Thifault, E., Julien, P., Tchernof, A., Couture, P., Lemieux, S., Barbier, O. and Vohla, M.-C. (2013) ‘Transcriptomic and metabolomic signatures of an n-3 polyunsaturated fatty acids supplementation in a normolipidemic/normocholesterolemic Caucasian population’, *J Nutrition Biochem*, 24, pp. 54–61.
- Ryan, K. E., Wojciechowski, I. A. and Garrison, B. J. (2007) ‘Reaction dynamics following keV cluster bombardment.’, *J. Phys. Chem. C*, 111, pp. 12822–12826.
- Sasaki, H., Bothner, B., Dell, A. and Fukuda, M. (1987) ‘Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA’, *J Biol Chem*, 262, pp. 12059–76.
- Scheltema, R. A., Hauschild, J.-P., Lange, O., Hornburg, D., Denisov, E., Kuehn, A., Makarov, A., Mann, M., Performance, H., Chromatography, L. and Damoc, E. (2014) ‘The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High Performance Quadrupole and an Ultra-High Field Orbitrap Analyzer’, *Mol Cell Proteomics*, 13, pp. 3698–3708.
- Schmid, A., Kortmann, H., Dittrich, P. S. and Blank, L. M. (2010) ‘Chemical and biological single cell analysis’, *Curr Opin Biotechnol*, 21, pp. 12–20.
- Schober, Y., Guenther, S., Spengler, B. and R m pp, A. (2012) ‘Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging’, *Anal Chem*, 84, pp. 6293–7.
- Sellick, C. A., Croxford, A. S., Maqsood, A. R., Stephens, G., Westerhoff, H. V., Goodacre, R. and Dickson, A. J. (2011) ‘Metabolite profiling of recombinant CHO cells: Designing tailored feeding regimes that enhance recombinant antibody production’, *Biotechnol Bioeng*, 108, pp. 3025–3031.
- Sellick, C. A., Hansen, R., Stephens, G. M., Goodacre, R. and Dickson, A. J. (2011) ‘Metabolite extraction from suspension-cultured mammalian cells for global metabolite

profiling TL - 6', *Nat Prot*, 6, pp. 1241–1249.

Shanta, S. R., Kim, T. Y., Hong, J. H., Lee, J. H., Shin, C. Y., Kim, K.-H., Kim, Y. H., Kim, S. K. and Kim, K. P. (2012) 'A new combination MALDI matrix for small molecule analysis: application to imaging mass spectrometry for drugs and metabolites', *The Analyst*, 137, pp. 5757–62.

Shard, A. G. and Gilmore, I. S. (2013) *TOF-SIMS: Materials analysis by mass spectrometry*. 2nd edn. Manchester: IM Publications.

Sheraz, S., Barber, A., Fletcher, J. S., Lockyer, N. P. and Vickerman, J. C. (2013) 'Enhancing secondary ion yields in ToF-SIMS using water cluster primary beams . supporting information', *Anal Chem*, pp. 2–3.

Shi, P., Jia, S., Zhang, X., Zhao, F., Chen, Y., Zhou, Q., Cheng, S. and Li, A. (2014) 'A cross-omics toxicological evaluation of drinking water treated with different processes', *J Hazard Mater*, 271, pp. 57–64.

Sigmund, P. (1981) 'Sputtering by Ion Bombardment: Theoretical Concepts', in R. Behrisch (ed.) *Topics in Applied Physics: Sputtering by Particle Bombardment I*. Springer Series Topics in Applied Physics, pp. 47–49.

Solé-Domènech, S., Sjövall, P., Vukojević, V., Fernando, R., Codita, A., Salve, S., Bogdanović, N., Mohammed, A. H., Hammarström, P., Nilsson, K. P. R., LaFerla, F. M., Jacob, S., Berggren, P.-O., Giménez-Llort, L., Schalling, M., Terenius, L. and Johansson, B. (2013) 'Localization of cholesterol, amyloid and glia in Alzheimer's disease transgenic mouse brain tissue using time-of-flight secondary ion mass spectrometry (ToF-SIMS) and immunofluorescence imaging', *Acta Neuropathol*, 125, pp. 145–157.

Stephan, Th., Jessberger, E. K., Heiss, C. H. and Rost, D. (2003) 'TOF-SIMS analysis of polycyclic aromatic hydrocarbons in Allan Hills 84001', *Meteorit Planet Sci*, 38, pp. 109–116.

Steuer, R., Morgenthal, K., Weckwerth, W. and Selbig, J. (2007) 'A gentle guide to the analysis of metabolomic data', *Methods Mol Biol*, 358, pp. 105–126.

Sumner, L. W., Samuel, T., Noble, R., Gmbh, S. D., Barrett, D., Beale, M. H. and Hardy, N. (2007) 'Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)', *Metabolomics*, 3, pp. 211–221.

Sussulini, A. (2017) 'Metabolomics: From Fundamentals to Clinical Applications', in Sussulini, A. (ed.) *Advances in Experimental Medicine and Biology. Proteomics, Metabolomics, Interactomics and Systems Biology*. Switzerland: Springer International Publishing AG.

Svatos, A. (2011) 'Single-cell metabolomics comes of age new developments in mass spectrometry profiling and imaging', *Anal. Chem.*, 83, pp. 5037–5044.

Swales, J. G., Strittmatter, N., Tucker, J. W., Clench, M. R., Webborn, P. J. H. and Goodwin, R. J. A. (2016) 'Spatial Quantitation of Drugs in tissues using Liquid Extraction Surface Analysis Mass Spectrometry Imaging', *Sci Rep*, 6, pp. 1–9.

Swales, J., Tucker, J., Spreadborough, MJ Iverson, S., Clench, M., Webborn, P. and Goodwin, R. (2015) 'Mapping Drug Distribution in Brain Tissue Using Liquid Extraction Surface Analysis Mass Spectrometry Imaging', *Anal Chem*, 87, pp. 10146–10152.

- Takahashi, K., Kozuka, T., Anegawa, A., Nagatani, A. and Mimura, T. (2015) 'Development and application of a high-resolution imaging mass spectrometer for the study of plant tissues', *Plant Cell Physiol*, 56, pp. 1329–1338.
- Takáts, Z., Wiseman, J. M., Gologan, B. and Cooks, R. G. (2004) 'Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization', *Science*, 306, pp. 471–473.
- Taschwer, M., Hackl, M., Hernandez Bort, J. A., Leitner, C., Kumar, N., Puc, U., Grass, J., Papst, M., Kunert, R., Altmann, F. and Borth, N. (2012) 'Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth', *J Biotechnol*, 157, pp. 295–303.
- Taylor, J., King, R. D., Altmann, T. and Fiehn, O. (2002) 'Application of metabolomics to plant genotype discrimination using statistics and machine learning', *Bioinformatics*, 18, pp. 241–248.
- Thomas, J. J., Shen, Z., Crowell, J. E., Finn, M. G. and Siuzdak, G. (2001) 'Desorption/ionization on silicon (DIOS): A diverse mass spectrometry platform for protein characterization', *Proc Natl Acad Sci U S A*, 98, pp. 4932–4937.
- Tian, H., Bai, J., An, Z., Chen, Y., Zhang, R., He, J., Bi, X., Song, Y. and Abliz, Z. (2013) 'Plasma metabolome analysis by integrated ionization rapid-resolution liquid chromatography/tandem mass spectrometry', *Rapid Commun Mass Spectrom*, 27, pp. 2071–2080.
- Tian, H., Fletcher, J. S., Thuret, R., Henderson, A., Papalopulu, N., Vickerman, J. C. and Lockyer, N. P. (2014) 'Spatiotemporal lipid profiling during early embryo development of *Xenopus laevis* using dynamic ToF-SIMS imaging', *J Lipid Res*, 55, pp. 1970–1980.
- Tjio, J. and Puck, T. T. (1958) 'Genetics of somatic mammalian cells. II. Chromosomal constitution of cells in tissue culture', *J Exp Med*, 108, pp. 259–268.
- Todd, J. F. J. (1991) 'Recommendations for nomenclature and symbolism for mass spectroscopy', *Pure Appl Chem*, 63, pp. 1541–1566.
- Touboul, D. and Brunelle, A. (2015) 'TOF-SIMS Imaging of Lipids on Rat Brain Sections', in *Methods in Molecular Biology*. Totowa, New Jersey 07512: Humana Press Inc, pp. 21–27. doi: 10.1007/978-1-4939-1357-2.
- Touboul, D., Brunelle, A., Halgand, F., De La Porte, S. and Laprévote, O. (2005) 'Lipid imaging by gold cluster time-of-flight secondary ion mass spectrometry: application to Duchenne muscular dystrophy', *J Lipid Res*, 46, pp. 1388–95.
- Touboul, D., Roy, S., Germain, D. P., Chaminade, P., Brunelle, A. and Laprévote, O. (2007) 'MALDI-TOF and cluster-TOF-SIMS imaging of Fabry disease biomarkers', *Int J Mass Spectrom*, 260, pp. 158–165.
- Trim, P. J. and Snel, M. F. (2016) 'Small molecule MALDI MS imaging: Current technologies and future challenges', *Methods*, 104, pp. 127–141.
- Trosset, J. and Carbonell, P. (2015) 'Synthetic biology for pharmaceutical drug discovery', *Drug Design Dev Theraphy*, 9, pp. 6285–6302.
- Trummer, E., Ernst, W., Hesse, F., Schriebl, K., Latenmayer, C., Kunert, R., Vorauer-Uhl, K., Katinger, H. and Müller, D. (2008) 'Transcriptional profiling of phenotypically

different Epo-Fc expressing CHO clones by cross-species microarray analysis', *J Biotechnol*, 3, pp. 924–937.

Vanbellinghen, Q. P., Elie, N., Eller, M. J., Della-Negra, S., Touboul, D. and Brunelle, A. (2015) 'Time-of-flight secondary ion mass spectrometry imaging of biological samples with delayed extraction for high mass and high spatial resolutions', *Rapid Commun Mass Spectrom*, 29, pp. 1187–1195.

Veening, J.-W., Smits, W. K. and Kuipers, O. P. (2008) 'Bistability, Epigenetics, and Bet-Hedging in Bacteria', *Annu Rev Microbiol*, 62, pp. 193–210.

Vickerman, J. and Briggs, D. (2001) *ToF-SIMS: surface analysis by mass spectrometry*. 2nd edn. Manchester: Surface Spectra. doi: 10.1255/tof2ch1.

Vickerman, J. C. (1997) 'Surface Analysis: The principal techniques'.

Vickerman, J. C. (2011) 'Molecular imaging and depth profiling by mass spectrometry--SIMS, MALDI or DESI?', *The Analyst*, 136, pp. 2199–2217.

Vickerman, J. C. and Swift, A. J. (1997) *Secondary Ion Mass Spectrometry in Surface Analysis the Principle Techniques*. Edited by J. Vickerman. Chichester: John Wiley and Sons Ltd.

Villas-Bôas, S. G., Mas, S., Åkesson, M., Smedsgaard, J. and Nielsen, J. (2005) 'Mass spectrometry in metabolome analysis', *Mass Spectrometry Reviews*, 24(5), pp. 613–646. doi: 10.1002/mas.20032.

Wang, Y.-J., Hao, S.-J., Liu, Y.-D., Hu, T., Zhang, G.-F., Zhang, X., Qi, Q.-S., Ma, G.-H. and Su, Z.-G. (2010) 'PEGylation markedly enhances the in vivo potency of recombinant human non-glycosylated erythropoietin: A comparison with glycosylated erythropoietin', *J Controlled Release*, 145, pp. 306–313.

Weckwerth, W. (2003) 'Metabolomics in systems biology', *Annu Rev Plant Biol*, 54, pp. 669–689.

Weckwerth, W. (2007) 'Metabolomics. Methods and Protocols', in Weckwerth, W. (ed.) *Methods in Molecular Biology*. Totowa, New Jersey 07512: Humana Press Inc, pp. 1–209.

Wilkens, C. A. and Gerdtzen, Z. P. (2015) 'Comparative metabolic analysis of CHO cell clones obtained through cell engineering, for IgG productivity, growth and cell longevity', *PLoS One*, 10, pp. 1–15.

Winograd, N. (2015) 'Imaging Mass Spectrometry on the Nanoscale with Cluster Ion Beams', *Anal Chem*, 87, pp. 328–333.

Wiseman, J. M., Ifa, D. R., Zhu, Y., Kissinger, C. B., Manicke, N. E., Kissinger, P. T. and Cooks, R. G. (2008) 'Desorption electrospray ionization mass spectrometry: Imaging drugs and metabolites in tissues.', *Proc Natl Acad Sci U S A*, 105, pp. 18120–18125.

Wurm, F. and Hacker, D. (2011) 'First CHO genome', *Nat Biotechnol*, 29, pp. 718–720.

Wurm, F. M. (2004) 'Production of recombinant protein therapeutics in cultivated mammalian cells', *Nat Biotechnol*, 22, pp. 1393–1398.

Xu, Q., Liu, Y., Zhang, Q., Ma, B., Yang, Z., Liu, L., Yao, D., Cui, G., Sun, J. and Wu, Z. (2014) 'Metabolomic analysis of simvastatin and fenofibrate intervention in high-lipid diet-induced hyperlipidemia rats', *Acta Pharmacol Sinica*, 35, pp. 1265–73.

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Chen, W., Xie, M., Wang, W., Hammond, S., Mikael, R., Neff, N., Passarelli, B., Koh, W., Fan, H. C., Gui, Y., Lee, K. H., Betenbaugh, M. J. and Quake, S. R. (2012) 'The Genomic Sequence of the Chinese Hamster Ovary (CHO) K1 cell line', *Nat Biotechnol*, 29, pp. 735–741.

Zenobi, R. (2013) 'Single-Cell Metabolomics: Analytical and Biological Perspectives', *Acta Neuropathol*, 342, pp. 1201–1211.

Zenobi, R. and Knochenmuss, R. (1998) 'Ion formation in MALDI mass spectrometry', *Mass Spectrom Rev*, 17, pp. 337–366.

Zhang, A., Sun, H., Han, Y., Yan, G., Yuan, Y., Song, G., Yuan, X., Xie, N. and Wang, X. (2013) 'Ultraperformance liquid chromatography-mass spectrometry based comprehensive metabolomics combined with pattern recognition and network analysis methods for characterization of metabolites and metabolic pathways from biological data sets', *Anal Chem*, 85, pp. 7606–12.

Zhao, X., Fritsche, J., Wang, J., Chen, J., Rittig, K., Schmitt-Kopplin, P., Fritsche, A., Häring, H. U., Schleicher, E. D., Xu, G. and Lehmann, R. (2010) 'Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits', *Metabolomics*, 6, pp. 362–374.

CHAPTER 2: Experimental

2.1. Introduction

The goal of this work was to incorporate a metabolomics-based approach to investigate the effects of EPO-Fc fusion protein expression in mammalian cells.

Fc-based fusion proteins are the most often generated fusion proteins because of their beneficial bio-pharmacological properties (Czajkowsky *et al.*, 2012) such as increased plasma half-life of the protein (Roopenian and Akilesh, 2007) and presence of Fc domain that enables for easy and low-cost purification of Fc-based fusion protein that is highly beneficial for manufacturers (Carter, 2011).

One of the Fc-based fusion proteins that has gained attention is erythropoietin-Fc fusion protein (EPO-Fc) (Taschwer *et al.*, 2012). Human erythropoietin (EPO) is a glycoprotein hormone that controls red blood cells production (Nagels *et al.*, 2012) and its recombinant version has been used as a therapeutic protein for the treatment of anaemia, AIDS or chemotherapy (Jelkmann, 2007). EPO-Fc is one of some unique approaches that have been undertaken to improve EPO therapeutic properties (Wang *et al.*, 2010).

A Chinese hamster ovary (CHO) cell line was used as a platform for the EPO-Fc fusion protein production. CHO cells are an epithelial (from tissue) cells derived from a Chinese hamster, first isolated in 1957 (Tjio & Puck, 1958) and represent a common expression system for the production of biopharmaceutical products (Trummer *et al.*, 2008). This cell line is a dominant platform used for the development of therapeutic proteins and has a proven history of regulatory approval for expression of recombinant proteins. CHO cells have been considered safe, mainly due to their ‘resistance’ for most human pathogenic viruses also, they grow relatively fast and robust (Wurm and Hacker, 2011).

Metabolite extraction was performed according to the protocol developed previously (Teng *et al.* 2009).

Current metabolite extraction procedures involve two stages: the initial quenching of the cells and their metabolic activity, followed by extraction of the metabolites (Kronthaler *et al.* 2012, Teng *et al.* 2009, Wiendahl *et al.* 2007). A standard approach for quenching of cells requires the use of pre-cooled solution in order to stop cell metabolism by rapid cooling of the cells and to preserve the characteristics of the metabolome. There is a number of quenching solutions that have been used, including: 60%-80% methanol in extraction of metabolites from normal (Sellick *et al.*, 2009) and cancer (Li *et al.*, 2017) cells as well as bacteria (*Bacillus* and *Brevibacillus*) (AlMasoud *et al.*, 2016), 0.9% sodium chloride in the analysis of Chinese hamster ovary cell line (Chen, Bennett and Kontoravdi,

2013) or phosphate-buffered saline (PBS) in the analysis of mammalian cells (Kronthaler, Gstraunthaler and Heel, 2012). Buffers maintain the membrane of cells however, they do not stop enzyme activity (Teng *et al.* 2009, Wiendahl *et al.* 2007). Up to date, methanol solutions have been applied for quenching metabolism of cells and bacteria the most often, due to its low quenching temperatures (freeze point at -48°C), good metabolite recovery and possibility of buffer supplementation that helps keeping integrity of cellular membranes (Hernandez Bort *et al.* 2014, Japelt *et al.* 2015, Martano *et al.* 2015, Ser *et al.* 2015, AlMasoud *et al.* 2016).

The second step of the sample preparation protocol for metabolite profiling involves the extraction of metabolites from cells. Firstly, cells need to be detached from their growth surface, for example by scraping or application of trypsin, which allows for collection of sufficient number of biomass (Villas-Bôas SG *et al.*, 2007). The extraction is then performed by the disruption of cell walls (lysation) by *e.g.* the application of liquid nitrogen followed by removal of the cellular debris (Sellick *et al.* 2009, Sellick *et al.* 2011). The extraction procedure however, should be carefully designed as it can provide a number of complications, including damage and/or decrease in metabolite levels when rapidly frozen (Wittmann *et al.*, 2004) or irreversible binding of metabolites to cell walls while freeze-dry cycles are applied (Dunn, Bailey and Johnson, 2005).

Accordingly, metabolite extraction in this study was performed with the application of cold 100% methanol followed by detaching cells by scraping and interruption of cell walls (lysation) using liquid nitrogen (Sellick *et al.* 2009, Sellick *et al.* 2011, Teng *et al.* 2009).

Metabolic profiling was conducted using GC-MS and MALDI-MS analysis of both external (footprint) and internal (fingerprint) metabolome of CHO cells.

2.2. Materials and methods

2.2.1. Materials

All cell culture consumables and solvents were purchased from Sigma-Aldrich (Gillingham, UK) or Life Technologies/Gibco (Invitrogen Group, Paisley, UK), unless stated otherwise. Plasticware was purchased from Fisher Scientific (Fisher Scientific, Fair Lawn, NJ) or Dutscher Scientific (Essex, UK). Cell culture reagents included HAM F12 nutrient mix, L-glutamine, blasticidin, tetracycline (Sigma-Aldrich, Gillingham, UK), Foetal Bovine Serum (FBS) (Life Technologies/Gibco) and hygromycin B in PBS (Invitrogen Group, Paisley, UK).

2.2.2. Methods

All cell culture work was conducted within a Class II biological safety cabinet (Mach-Aire Limited, UK) and all work areas were carefully cleaned with 70% ethanol before use.

2.2.2.1. Characterisation of cell line

For this study, the CHO-K1 adherent cell line transfected with T-REx system was used (**Figure 1**). T-REx system is a tetracycline-regulated mammalian expression system, where tetracycline regulation is based on one of the T-REx system components - pcDNA6/TR regulatory vector. The pcDNA6/TR regulatory vector expresses the tetracycline (TET) repressor and tetracycline regulation is dependent on the binding of tetracycline to this repressor. This bond results in activation of the promoter for expression of the gene with encoded EPO-Fc fusion protein. Additionally, pcDNA6/TR vector encodes blasticidin (antibiotic) resistance that allows for selection of stable transfectants only (Kimura *et al.* 1994, Hillen & Berens, 1994). CHO-K1 T-REx cells were transfected with EPO-Fc-pcDNA5-FRT-TO plasmid that encodes hygromycin (antibiotic) resistance and erythropoietin (EPO) Fc fragment recombinant protein (~70 kDa). Accordingly, EPO-Fc is under control of tetracycline repressor. Hygromycin, similarly to blasticidin, plays a role of selectable marker and is essential for selecting cells that contain these plasmids. In the absence of hygromycin, the cell population might lose this plasmid, therefore losing the ability to produce EPO-Fc (<http://www.thermofisher.com>, <https://www.addgene.org>).

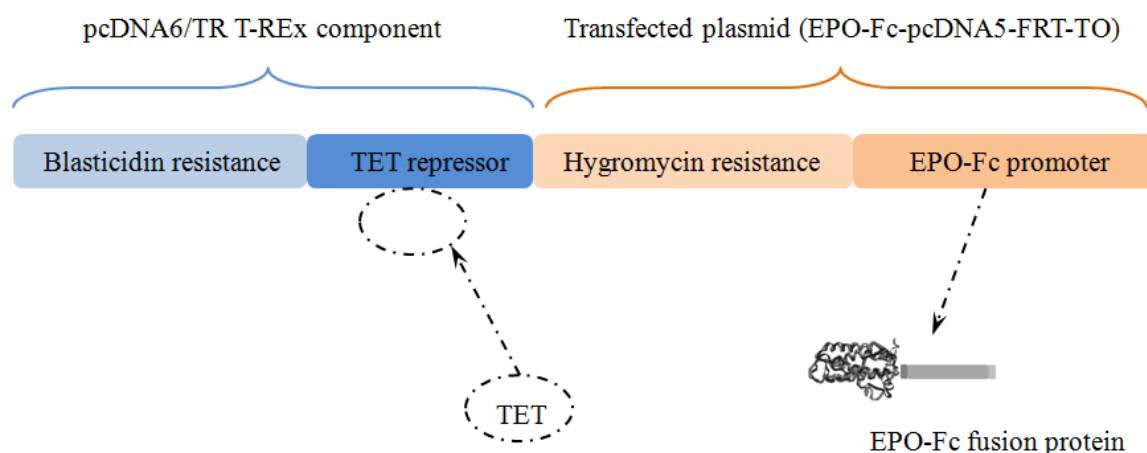


Figure 1: Schematic components of CHO-K1 T-REx EPO-Fc-pcDNA5-FRT-TO cell line used for metabolomics studies of EPO-Fc fusion protein expression. When induced, the binding between tetracycline (TET) and its repressor activates EPO-Fc promoter for EPO-Fc fusion protein production.

A EPO-Fc fusion protein schematic structure is shown in **Figure 2**. The construct is built from EPO protein and Fc fragment with six residues of histidine attached. Additionally, an enzyme enterokinase site between EPO protein and Fc fragment is assembled.

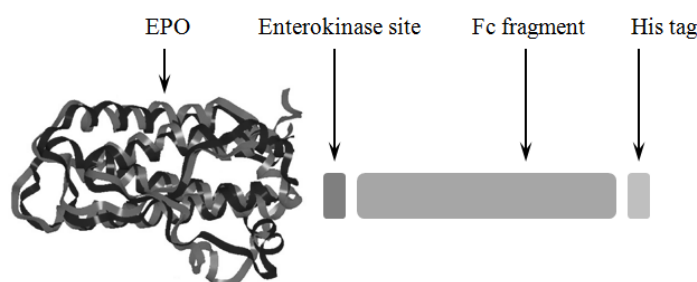


Figure 2: Schematic structure of EPO-Fc fusion protein. EPO structure incorporated from Cheetham *et al.* 1998.

An aliquot of a frozen CHO-K1 T-REx cell line containing EPO-Fc-pcDNA5-FRT-TO plasmid (further called ‘CHO’ cell line) was obtained from Prof. Alan Dickson and his PhD student Thomas Leach at the University of Manchester.

2.2.2.2. Resurrection of frozen cells

A cryogenic tube containing frozen CHO cells in liquid nitrogen was warmed in a water bath heated to 37 °C to thaw the cells. Once thawed, 1 mL of cell suspension was added to 9 mL of pre-warmed complete growth medium (**Table 1**) in a conical tube and centrifuged (Rotofix32A Hettich Centrifuges, UK) for 10 min at 500 *g* to remove dimethyl sulfoxide

(DMSO). The supernatant was removed and cells were suspended in 8 mL of fresh medium and seeded on T-25 flask. Cells were counted in a cell counting chamber (Thoma) using Trypan blue according to standard protocol and placed in an incubator at 37 °C in 5% of CO₂ atmosphere. Cell culture was examined microscopically at 24 h periods to check for any morphological changes and nutrient requirements. Also, cell culture was checked for *Mycoplasma* contamination monthly.

Table 1: Complete growth cell medium for CHO cell line.

Content	Concentration in cell medium	Company	Catalogue number
HAM F12 nutrient mix	n/a*	Sigma-Aldrich	N6658
L-glutamine	1.8 mM	Sigma-Aldrich	G7513
Foetal Bovine Serum (FBS)	10%	Life Technologies/Gibco	10500-064
Blasticidin	10 µg/mL	Sigma-Aldrich	15205
Hygromycin B in PBS	250 µg/mL	Invitrogen	10687010

*complete formulation available on company website

2.2.2.3. Sub-culturing (passaging) of cells

Cells were grown as a monolayer in T-25 or T-75 flasks containing respectively 8 mL and 15 mL of complete growth medium. Flasks were assessed under inverted phase contrast microscopy (Model Pim, World Precision Instruments WPI, Sarasota, FL, USA) to determine the level of confluency. Cells were sub-cultured every 3-4 days at which point 70–90% confluence of cells was obtained. The old cell medium was aspirated and 5-10 mL of pre-warmed calcium and magnesium free phosphate buffered saline (PBS) was added to wash the cell monolayer and remove any remaining FBS that can inhibit the activity of the proteolytic enzymes to be used. The PBS was removed and 2.5–5 mL of trypsin was added to the flask and the flask was placed in the incubator for 5 min to boost the performance of the enzyme. After the incubation time had finished, 5–10 mL of pre-warmed cell medium was added and cell suspension was agitated to disattach cells. Cells were counted in a cell counting chamber (Thoma) using Trypan blue according to standard protocol (Sellick *et al.*, 2009) and seeded in fresh medium at the density as necessary, usually 0.1×10^6 cells/mL for 3 days of cell culture. The flask was placed in the incubator at 37 °C in 5% of CO₂ atmosphere (Galaxy 170S, New Brunswick, Eppendorf, UK). Viability of cells was checked with each passage.

2.2.2.4. Cryogenic preservation (freezing) of cells

Cryogenic preservation is a practice of retaining cell culture in a stable environment for a prolonged period of time without the need for maintenance of the cell conditions *i.e.* feeding and passaging procedures detailed above. The sub-culturing procedure was followed to the stage where the cells were incubated with trypsin and 5–10 mL of pre-warmed cell medium was added. Cell suspension was transferred to a conical tube and centrifuged for 10 min at 500 g. Supernatant was removed and the cell pellet was re-suspended to a density of 10×10^6 cells/mL in fresh cell medium containing 10 % DMSO. DMSO acted as a cryoprotective agent in the freezing process. 1 mL of cell suspension was aliquoted to labelled cryotubes and placed in a -20 °C freezer for 24 h. After 24 h, the cryotubes were transferred to a -80 °C freezer (U570-86, New Brunswick, UK) for another 48 h. Once frozen, the ampoules were transferred to liquid phase nitrogen storage vessels for long term storage.

2.2.2.5. Inducer exposures

Tetracycline stock solution was prepared in water in a concentration of 1 mg/mL, added to cells to a final concentration of 1 µg/mL and incubated for 3 days. Stock solution was applied to cells when cells were examined under the microscope after 24 h from seeding to exclude contamination and to control the morphology of cells.

2.2.2.6. Metabolome sample preparation

A development of a suitable metabolite extraction protocol was required that would be optimum for metabolome analysis. The basic practice of harvesting adherent cells includes the use of trypsin to disassemble cells from flask surface. This however, brings some major complications: trypsin does not inactivate enzymes, it can potentially change the metabolite profile because of its interactions with membrane proteins and, time needed for centrifugation of samples during such protocol creates a great gap between harvesting and quenching of metabolism that might result in the loss of significant metabolites. In order to minimise changes in metabolite profile, cells were harvested by cell scraping. Metabolite extraction was performed according to the method developed previously (Teng *et al.* 2009), however the use of room temperature PBS rather than ice cold PBS during the washing step was applied in order to reduce cell shock and thus minimise potential metabolome leakage (Hollywood *et al.*, 2015). **Figure 3** illustrates the diagram of stages in the collection of footprint and fingerprint samples. Cells were seeded and grown as

described previously and in T-225 flasks with 25 mL of growth medium. After the addition of tetracycline and thus, induction of the protein production, cell culture was maintained for 3 days. Sample collection of footprint and fingerprint samples was performed on day 1 and day 3 of cell culture. Control flasks were grown along treated cells and there were 4 biological replicates of control and tetracycline-induced cells. All used cell media, PBS, FBS, antibiotics and solvents had the same batch numbers to reduce introduction of variation.

Control cells

Cells treated with the inducer



Footprint

1. Medium aspirated and centrifuged.
2. Snap frozen in liquid nitrogen.
3. Lyophilised samples were stored at -80°C until analysis.



Fingerprint

4. Cells washed with PBS (twice).
5. Quenching with 100% MeOH (7mL, -48°C , about 5 minutes).
6. Cells harvested by cell scraping.
7. Cellular biomass removed and collected into centrifuge tubes.
8. Metabolite extraction through 3 freeze thaw cycles. Samples snap frozen in liquid nitrogen and thawed on dry ice.
9. Samples centrifuged to pellet cell debris; supernatant collected and dried.
10. Stored at -80°C until analysis.

Figure 3: A diagram of main stages in the footprint and fingerprint sample collection.

2.2.2.7. Metabolomic footprint sample collection

The expired media were aspirated from each culture flask, collected into 15 mL conical tubes and centrifuged 10 min at 1000 g at -8 °C to remove any remaining cellular debris. The centrifuged metabolic footprint samples were aliquoted into smaller volumes and immediately snap frozen in liquid nitrogen. Same samples were lyophilised (Eppendorf Vacufuge Concentrator 5301, Eppendorf, UK). All samples were stored at -80 °C until further analysis.

2.2.2.8. Metabolomic fingerprint sample collection

After removal of the expired media, PBS (10 mL, room temperature) was applied to each flask to wash cellular monolayer from any remaining medium. PBS was aspirated, collected to 15 mL conical tubes and immediately snap frozen in liquid nitrogen for further analysis of any potential metabolite leakage. Washing step with PBS was repeated twice, however PBS after first washing step only was collected for metabolite leakage analysis. Following this, the cellular metabolism was quenched by addition of 100 % methanol (7 mL, -48 °C, about 5 min) and harvested by scraping the cell monolayer from the flask with a disposable cell scraper (Corning, UK). The collected biomass was removed by pipette and placed in 15 mL conical tubes. The samples were snap frozen in liquid nitrogen and allowed to thaw on dry ice in order to break cell membrane and extract metabolites. This process was repeated three times to ensure sufficient extraction. The samples were then centrifuged 10 min at 5000 g at -8 °C to pellet the cell debris. The supernatant was collected and aliquoted into smaller volumes and lyophilised (Eppendorf Vacufuge Concentrator 5301, Eppendorf, UK) in order to concentrate metabolite extracts. PBS samples were also lyophilised. All samples were stored at -80 °C until further analysis.

References:

- AddGene vector database* (no date) 2016. Available at: <https://www.addgene.org> (Accessed: 4 November 2016).
- AlMasoud, N., Xu, Y., Trivedi, D. K., Salivo, S., Abban, T., Rattray, N. J. W., Szula, E., AlRabiah, H., Sayqal, A. and Goodacre, R. (2016) 'Classification of *Bacillus* and *Brevibacillus* species using rapid analysis of lipids by mass spectrometry', *Anal Bioanal Chem*, 408, pp. 7865–7878.
- Carter, P. (2011) 'Introduction to current and future protein therapeutics: a protein engineering perspective.', *Exp Cell Res*, 317, pp. 1261–1269.
- Cheetham, J. C., Smith, D. M., Aoki, K. H., Stevenson, J. L., Hoeffel, T. J., Syed, R. S., Egrie, J. and Harvey, T. S. (1998) 'NMR structure of human erythropoietin and a comparison with its receptor bound conformation.', *Nat Struct Biol*, 5, pp. 861–866.
- Chen, N., Bennett, M. H. and Kontoravdi, C. (2013) 'Analysis of Chinese hamster ovary cell metabolism through a combined computational and experimental approach', *Cytotechnology*, 66, pp. 945–966.
- Czajkowsky, D. M., Hu, J., Shao, Z. and Pleass, R. J. (2012) 'Fc-fusion proteins: new developments and future perspectives', *EMBO Mol Med*, 4, pp. 1015–1028.
- Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005) 'Measuring the metabolome: current analytical technologies', *The Analyst*, 130, pp. 606–625.
- Hernandez Bort, J. A., Shanmukam, V., Pabst, M., Windwarder, M., Neumann, L., Alchalabi, A., Krebiehl, G., Koellensperger, G., Hann, S., Sonntag, D., Altmann, F., Heel, C. and Borth, N. (2014) 'Reduced quenching and extraction time for mammalian cells using filtration and syringe extraction', *J Biotechnol. Elsevier B.V.*, 182–183, pp. 97–103.
- Hillen, W. and Berens, C. (1994) 'Mechanisms underlying expression of TN10 encoded tetracycline resistance', *Annual Rev Microbiol*, 48, pp. 345–69.
- Hollywood, K. A., Winder, C. L., Dunn, W. B., Xu, Y., Broadhurst, D. I., Griffiths, C. E. and Goodacre, R. (2015) 'Exploring the mode of action of dithranol therapy for psoriasis: a metabolomic analysis using HaCaT cells', *Mol Biosyst*, 11, pp. 2198–2209.
- Japelt, K., Christensen, J. and Villas-Boas, S. (2015) 'Metabolic fingerprinting of *Lactobacillus paracasei*: the optimal quenching strategy', *Microbi Cell Fact*, 14, pp. 132–139.
- Jelkmann, W. (2007) 'Erythropoietin after a century of research: younger than ever', *Eur J Haematol*, 78, pp. 183–205.
- Kimura, M., Takatsuki, A. and Yamaguchi, I. (1994) 'Blasticidin S deaminase gene from *Aspergillus terreus* (BSD): a new drug resistance gene for transfection of mammalian cells', *Biochim Biophys Acta*, 1219, pp. 653–659.
- Kronthaler, J., Gstraunthaler, G. and Heel, C. (2012) 'Optimizing High-Throughput Metabolomic Biomarker Screening: A Study of Quenching Solutions to Freeze Intracellular Metabolism in CHO Cells', *J Integ Biol*, 16, pp. 90–97.
- Li, X., Wong, C. C., Tang, Z., Wu, J., Li, S., Qian, Y., Xu, J., Yang, Z., Shen, Y., Yu, J. and Cai, Z. (2017) 'Determination of amino acids in colon cancer cells by using UHPLC-

MS/MS and [U-13C5]-glutamine as the isotope tracer', *Talanta*, 162, pp. 285–292.

Life Technologies (2016). Available at:

<https://www.thermofisher.com/uk/en/home/references/protocols/proteins-expression-isolation-and-analysis/protein-expression-protocol/inducible-protein-expression-using-the-trex-system.html> (Accessed: 4 November 2016).

Martano, G., Delmotte, N., Kiefer, P., Christen, P., Kentner, D., Bumann, D. and Vorholt, J. (2015) 'Fast sampling method for mammalian cell metabolic analyses using liquid chromatography-mass spectrometry', *Nat Protoc*, 10, pp. 1–11.

Nagels, B., Van Damme, E. J. M., Callewaert, N., Zabeau, L., Tavernier, J., Delanghe, J. R., Boets, A., Castilho, A. and Weterings, K. (2012) 'Biologically active, magnICON-expressed EPO-Fc from stably transformed *Nicotiana benthamiana* plants presenting tetra-antennary N-glycan structures', *J Biotechnol*, 160, pp. 242–250.

Roopenian, D. and Akilesh, S. (2007) 'FcRn: the neonatal Fc receptor comes of age.', *Nat Rev Immunol*, 7, pp. 715–725.

Sellick, C. A., Hansen, R., Stephens, G. M., Goodacre, R. and Dickson, A. J. (2011) 'Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling TL - 6', *Nat Prot*, 6, pp. 1241–1249.

Sellick, C. a, Hansen, R., Maqsood, A. R., Warwick, B., Stephens, G. M., Goodacre, R., Dickson, A. J. and Dunn, W. B. (2009) 'Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells', *Anal Chem*, 81, pp. 174–183.

Ser, Z., Liu, X., Tang, N. N. and Locasale, J. W. (2015) 'Extraction parameters for metabolomics from cultured cells', *Anal Biochem*, 475, pp. 22–28.

Taschwer, M., Hackl, M., Hernandez Bort, J. A., Leitner, C., Kumar, N., Puc, U., Grass, J., Papst, M., Kunert, R., Altmann, F. and Borth, N. (2012) 'Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth', *J Biotechnol*, 157, pp. 295–303.

Teng, Q., Huang, W., Collette, T. W., Ekman, D. R. and Tan, C. (2009) 'A direct cell quenching method for cell-culture based metabolomics', *Metabolomics*, 5, pp. 199–208.

Tjio, J. and Puck, T. T. (1958) 'Genetics of somatic mammalian cells. II. Chromosomal constitution of cells in tissue culture', *J Exp Med*, 108, pp. 259–268.

Trummer, E., Ernst, W., Hesse, F., Schriebl, K., Latenmayer, C., Kunert, R., Vorauer-Uhl, K., Katinger, H. and M?ller, D. (2008) 'Transcriptional profiling of phenotypically different Epo-Fc expressing CHO clones by cross-species microarray analysis', *J Biotechnol*, 3, pp. 924–937.

Villas-Bôas SG, Roessner, U., A., E. H. M., Smedsgaard, J. and Nielsen, J. (2007) *Metabolome analysis: an introduction*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/0470105518.

Wang, Y.-J., Hao, S.-J., Liu, Y.-D., Hu, T., Zhang, G.-F., Zhang, X., Qi, Q.-S., Ma, G.-H. and Su, Z.-G. (2010) 'PEGylation markedly enhances the in vivo potency of recombinant human non-glycosylated erythropoietin: A comparison with glycosylated erythropoietin', *J Controlled Release*, 145, pp. 306–313.

Wiendahl, C., Brandner, J. J., Kuppers, C., Lu, B., Schygulla, U., Noll, T. and Oldiges, M. (2007) 'A microstructure heat exchanger for quenching the metabolism of mammalian cells', *Chem Eng Technol*, 30, pp. 322–328.

Wittmann, C., Krömer, J., Kiefer, P., Binz, T. and Heinzle, E. (2004) 'Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria', *Anal Biochem*, 327(1).

Wurm, F. and Hacker, D. (2011) 'First CHO genome', *Nat Biotechnol*, 29, pp. 718–720.

CHAPTER 3: Instrumentation

3.1. Introduction

Biological experiments require specialised instrumentation for the acquisition of data with high mass accuracy, high mass and spatial resolution. The application of different analytical tools suitable for the characterisation of small molecules offers the advantage of gaining complementary information about biological samples and leads to better understanding of metabolism of living organisms (Bujak et al. 2015).

Since its invention in 1912, mass spectrometry (MS) is well-established platform for a number of scientific disciplines (Dunn, 2008). A standard arrangement of mass spectrometry instrument is shown in **Figure 1**.

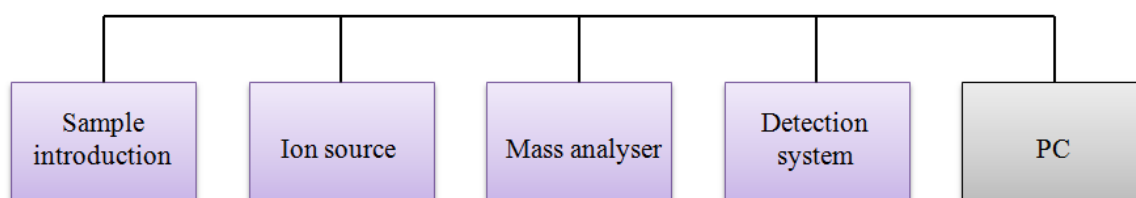


Figure 1: A schematic of a standard mass spectrometry platform. Samples are inserted to an ion source (atmospheric or vacuum pressures) through an inlet system, where ions are produced. Ions travel through mass analyser being separated according to their mass-to-charge (m/z) ratio prior the detection at the detector component. A PC is required to acquire experiments and to control system and all its parameters (Dunn, 2008).

Samples can be introduced to the mass spectrometer in liquid (*e.g.* atmospheric pressure MALDI), solid (*e.g.* MALDI, SIMS) or gas phase (*e.g.* GC-MS), depending on the type and arrangement of the system. Subsequently, a sample is transferred to an ion source where positively or negatively charged ions are produced. Ions then are separated according to their mass-to-charge (m/z) ratios in a mass analyser. The separated ions go through a detector component for required detection and data is then collected and stored on PC (Li & Gross 2004, Dunn 2008).

3.2. Platforms for metabolic footprint and fingerprint analysis

3.2.1. GC-MS

3.2.1.1. Agilent 6890N GC LECO Pegasus III EI ToF-MS

Analysis was performed using Agilent 6890N gas chromatograph oven (Wokingham, UK) coupled to a LECO Pegasus III electron ionisation (EI) time-of-flight mass spectrometer (ToF-MS) (LECO Corporation, St Joseph, USA) controlled by ChromaTOF software ver. 2.32 (**Figure 2**). The GC oven was operated with 6 N helium as the carrier gas with a flow rate of 0.8 mL per minute, in a 1:3 split mode with a start temperature of 70 °C. A VF5-MS column (Supelco, Gillingham, UK, 30 m × 0.25 mm × 0.25 µm film thickness) was applied with the transfer line and source temperatures held at 230 °C and 200 °C respectively. The temperature of the GC oven was ramped up at a speed of 20 °C per minute until approaching the final temperature of 320 °C. Subsequently, the final temperature was held for 5 min with the total run time of 20.5 min including a solvent delay at the start. MS scan acquisition rate was set to 15 Hz, detector voltage was 1650 V and samples were analysed within a mass range of 40–600 Da.

Once the sample is injected to a gas chromatograph (**Figure 2, (1) and (2)**), it enters the column, through which different compounds of a sample pass in a carrier gas at different rates. Different retention times of compounds are recorded once they exit the end of the column. Subsequently, the gas phase compounds are introduced to the mass spectrometer, where they interact with an electron beam to produce ions (**Figure 2 (3)**). Ions then are separated according to their mass-to-charge ratio in the ToF analyser, detected and results are recorded (**Figure 2 (5)**) (http://cires1.colorado.edu/jimenez/CHEM-5181/Lect/2013_GCxGC_Handouts.pdf).

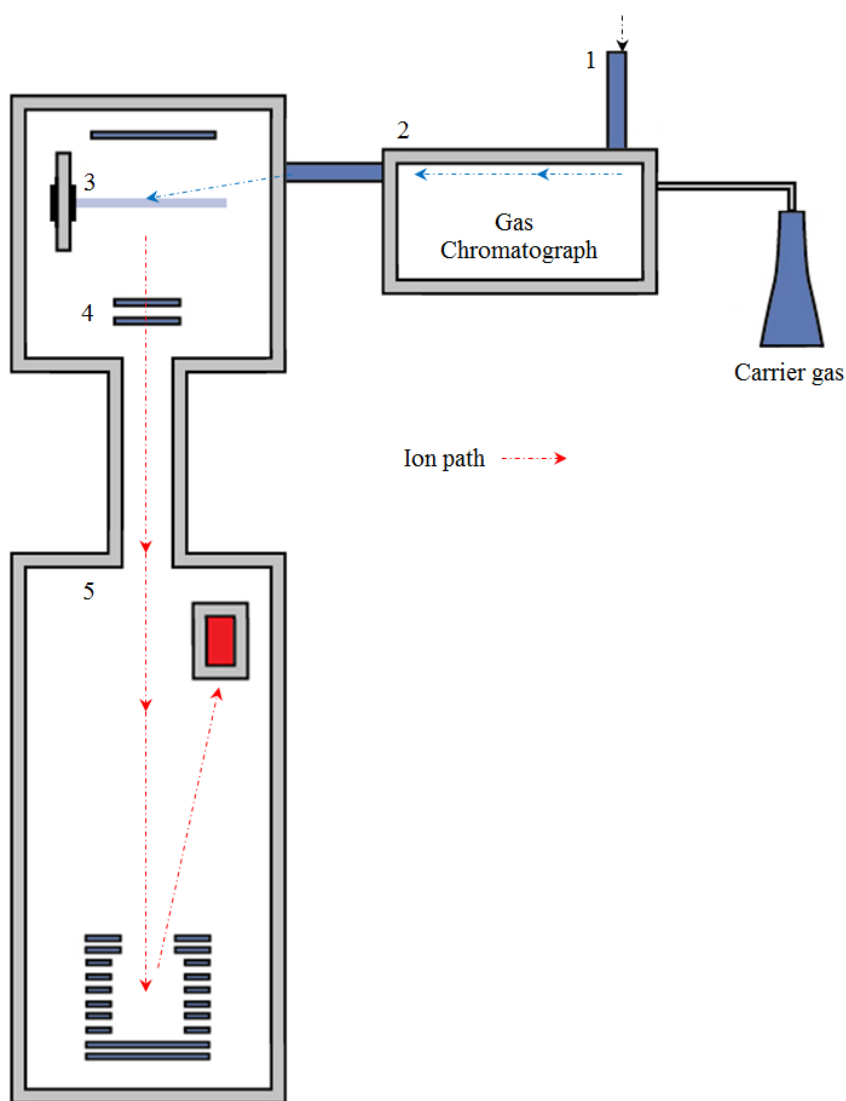


Figure 2: Scheme of GC-ToF-MS: sampling (1), gas chromatograph (2), electron source (3), ion focusing optics (4), ToF reflectron analyser (5) (http://cires1.colorado.edu/jimenez/CHEM-5181/Lect/2013_GCxGC_Handouts.pdf, <http://www.speciation.net/Database/Instruments/Leco-Inc/Pegasus-III-GCTOFMS-;i509>).

3.2.1.2. Data analysis

Deconvolution of raw data was performed using LECO ChromaTof ver. 2.32 software package (LECO Corporation, St Joseph, USA). Deconvolution included baseline correction (minimising noise), peak picking based on peak width, library search from 4 databases (2 instrument-dependent and 2 National Institute of Standards and Technology (NIST) databases) and calculating the peak area and retention time based on retention index of known hydrocarbons. QCs for footprint and fingerprint samples were also deconvolved and used to generate metabolite IDs. The final output was a retention time vs.

mass data matrix with related metabolite IDs and peak areas linked to each sample injection. Metabolite identification strictly followed the Metabolomics Standards Initiative guidelines described previously (Sumner *et al.*, 2007). The data matrix was exported from GC-MS software to Microsoft Excel ver. 2007 (Microsoft Corporation, USA) files and further analysed with MATLAB ver. R2012a (The MathWorks Inc., UK) and MetaboAnalyst 3.0. (<http://www.metaboanalyst.ca/faces/home.xhtml>).

3.2.2. MALDI-MS

3.2.2.1. MALDI–ToF–MS Ultraflex II ToF/ToF mass spectrometer

MALDI-MS optimisation was performed using MALDI–ToF–MS Ultraflex II ToF/ToF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) equipped with 337 nm wavelength nitrogen pulsed UV laser (**Figure 3**). Spectral acquisition was conducted using FlexControl 3.0.183.0 software (Bruker Daltonics, Billerica, MA, USA). The instrument was operated in high vacuum (10^{-7} mbar), the ion source voltage (ISV)1 and ISV2, lens voltage (LV), reflectron voltage (RV) 1 were 25 kV, 21 kV, 9.5 kV, 26.3 kV respectively for positive ionisation mode and -20 kV, 17.45 kV, 7.5 kV and 21 kV respectively for negative ionisation mode. The mass range of 0 to 1000 Da was analysed in reflectron positive and negative ionisation modes at pulse repetition rate of 100 Hz. A number of 1500 shots was acquired using ‘random’ pattern within a sample spot. Samples were spotted on a MPT 384-well ground steel target plate (Bruker Daltonics, Billerica, MA, USA).

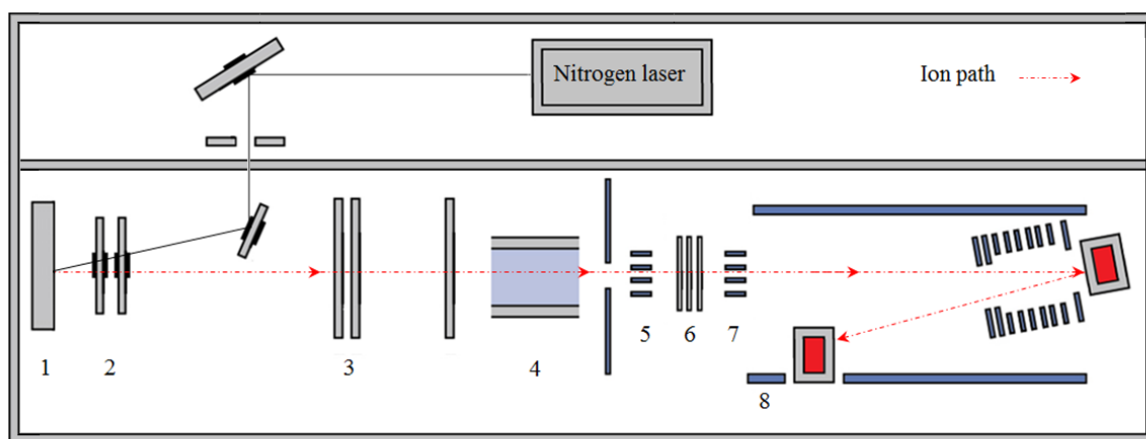


Figure 3: Scheme of MALDI Ultraflex II ToF/ToF mass spectrometer: target plate (1), electrodes (2), lens arrangement (3), Collision Induced Dissociation (CID) cell (4), timed ion selector 1 (ToF1) (5), LIFT source (6), Post Lift Metastable Suppressor (PLMS) (7), ToF 2 analyser for linear and reflectron detection (8) (Suckau et al. 2003, http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf).

Firstly, the sample is placed on a target plate and is introduced to the instrument (**Figure 3** (1)). Once the laser irradiates the sample, it triggers ablation, desorption and ionisation of sample and matrix, producing different ions. The ions are then directed to the ToF mass analyser and are separated according to their m/z ratio (**Figure 3** (8)). For MS/MS analysis, the ions are directed to Collision Induced Dissociation (CID) cell, where they are

dissociated by collisions during a passage with gas inside the CID cell (**Figure 3 (4)**) before they enter the mass analyser. Subsequently, all ions are introduced to timed ion selector 1 (**Figure 3 (5)**). Time ions selector 1 is a mass filter used to separate a particular precursor ion and related products from all the other ions for MS/MS study. Selected precursor and product ions then go through the LIFT source, where they gain particular energies in order to be simultaneously detected in MS/MS spectrum (**Figure 3 (6)**). The parent and fragment ions travel through ToF mass analyser and hit the detector (**Figure 3 (8)**). Additional function can be applied with Post Lift Metastable Suppressor (PLMS) (**Figure 3 (7)**), where the parent ion and daughter ions formed after the LIFT process are deflected, resulting in a spectrum with fragments formed between the source and the LIFT cell only (Suckau et al. 2003, http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf).

3.2.2.2. MALDI Synapt G2–Si HDMS mass spectrometer

MALDI-MS optimisation, footprint and fingerprint profiling were performed with MALDI Synapt G2–Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK) equipped in Nd:YAG 355 nm wavelength nitrogen UV laser and repetition rate of 1000 Hz (**Figure 4**). The spectra were acquired using MassLynx ver. 4.1 SCN941 build 18 (Waters Corporation/Micromass, Manchester, UK) in the mass range of 20 to 1000 Da and in the “sensitivity mode”, with expected mass resolution of 10 000. The instrument was operated in high vacuum ($\geq 10^{-8}$ mbar) in reflectron positive and negative ionisation modes with 0 V on the sample plate and 10 V extraction voltages. The total acquisition time was 30 second at a rate of 1 scan per second using default pattern of spectra collection. Samples were spotted on a MPT 96-well steel target plate (Waters Corporation/Micromass, Manchester, UK).

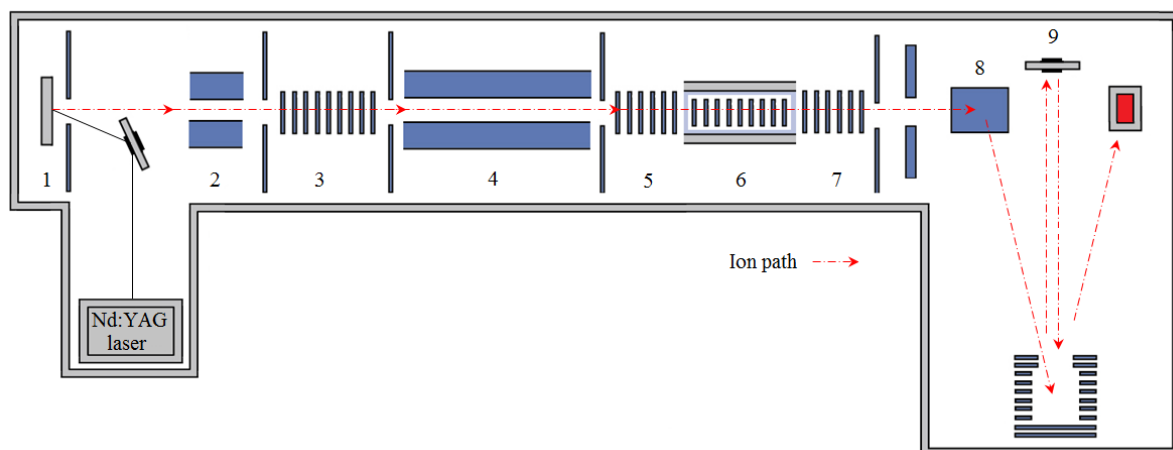


Figure 4: Scheme of MALDI Synapt G2-Si HD mass spectrometer: target plate (1), hexapole ion guide (2), T-Wave ion guide (3), quadrupole (4), Tri-Wave: trap (5), ion mobility separation in helium cell (6), transfer (7), QuanToF: high field pusher (8), ion mirror (9) (<http://www.waters.com>).

Firstly, a target plate with the sample on is introduced to the instrument (**Figure 4 (1)**). Once the laser is fired and hits the sample, different ions are produced. The ions then travel through hexapole ion guide (**Figure 4 (2)**) that focuses beam of ions and moves them further with little or no loss of signal through Travelling Wave (T-Wave) (**Figure 4 (3)**), quadrupole (**Figure 4 (4)**), ion mobility Tri-Wave cell (**Figure 4 (5), (6), (7)**) until they reach QuanToF mass analyser, where the ions are analysed according to their mass-to-charge ratio. With applied ion mirror present in QuanToF the resolution can be enhanced (**Figure 4 (9)**). The separated ions then go through a detector component for required detection and data is then collected.

In the Tri-Wave ion guide (**Figure 4 (5-7)**) the ions can also be additionally dissociated (CID) for MS/MS studies. Subsequently, the parent and daughter ions are mass analysed in QuanToF mass analyser. Additionally, ion mobility experiments can be performed using Tri-Wave component with the trap cell, mobility separation in helium cell and transfer cell (**Figure 4 (5), (6), (7)**). In the Tri-Wave device the ions are separated according to their mobility (drift time), providing structural (collisional cross-sectional) information on the sample (http://www.waters.com/waters/en_GB/SYNAPT-G2-Si-High-Definition-Mass-Spectrometry/nav.htm?cid=134740622&locale=en_GB).

3.2.2.3. Data analysis

Spectra were analysed with FlexAnalysis 3.0.96.0 (Bruker Daltonics, Billerica, MA, USA) or MassLynx ver. 4.1 SCN941 build 18 software (Waters Corporation/Micromass, Manchester, UK), depending on the instrument used for experiments. The data was combined, saved and exported in Bruker-specific files or .raw files (for Synapt G2-Si) and further processed using MATLAB ver. R2012a. The results were assessed in Microsoft Excel ver. 2007 and Origin Pro ver.8.1 (OriginLab, Northampton, MA, USA) for statistical calculation.

3.2.3. Mass spectrometry imaging (MSI)

3.2.3.1. MALDI-MSI

3.2.3.1.1. Matrix coating

Cell samples on poly-L microscope glass slides were transferred to Giessen, Germany in order to conduct matrix coating and MALDI imaging experiments.

Matrix was applied on the cells using a robot matrix preparation system (SMALDIPrep, TransMIT GmbH, Giessen, Germany) (**Figure 5**). Sheath gas (grade 5.0 N₂) was regulated to 80 *psi* at the output of the nitrogen regulator to have a flow rate of 5 liters/minute. The spray nozzle was placed 11 mm above the sample and a syringe pump was used to apply 100 μ L of 30 mg/mL DHB in 50% acetone with 0.1% TFA and 80 μ L of 30 mg/mL pNA in 50% acetone at a rate of 10 μ L per minute for positive and negative ionisation modes, respectively. This combination of gas flow, nozzle height, deposition rate and matrix solutions has been previously optimised by the Spengler group and used to acquire high quality imaging data (Römpp & Spengler 2013, Römpp et al. 2010, Bhandari et al. 2015, Khalil et al. 2015).



Figure 5: SMALDIPrep matrix coating device (TransMIT GmbH, Giessen, Germany). Adapted from (<http://www.uni-giessen.de/faculties/f08/departments/iaac/spengler/research/instrumentations/transmit-smaldiprep>).

3.2.3.1.2. AP-SMALDI10 imaging source

Experiments were performed in collaboration with Professor Bernhard Spengler at the Justus Liebig University Giessen in Germany. The research was carried out using an Atmospheric-Pressure Scanning microprobe Matrix-Assisted Laser Desorption/Ionisation (AP-SMALDI10) imaging source (Trans-MIT GmbH, Giessen, Germany) connected to a Fourier transform orbital trapping mass spectrometer (Thermo Q Exactive, San Jose, CA, USA) equipped with a nitrogen pulsed UV laser using wavelength of 337 nm and repetition rate of 60 Hz (**Figure 6**). The laser beam was focused to a spatial resolution of 5 μm and energy of 0.06 μJ (for attenuator at 40°) per spot. The stage raster steps were equal to the laser spot size. The sample was placed on the stage and was moved in X, Y dimensions to coordinate the area of interest. MS imaging data was acquired in continuous mode using the Tune software (Thermo Scientific, Bremen, Germany) in the mass range of 250 to 1000 Da and in the “profiling mode”, giving a mass resolution of 140 000 (FWHM) at m/z 200. The instrument was operated in positive and negative ionisation modes with +4 kV and -4 kV on the sample plate, respectively. Internal calibration was applied using matrix peaks as reference masses for both positive and negative ionisation modes: $[\text{5DHB-4H}_2\text{O+NH}_4]^+$ at m/z 716.12461 for positive and $[\text{3PNA-5H}]^-$ at m/z 409.09021 for negative ionisation mode. Mass accuracy of ≤ 2 ppm Root Mean Square Error (RMSE) was achieved. Laser, sample stage and mass spectrometer were synchronised and controlled by a microcontroller and in-house built software. More information on the instrument details can be found in the recent literature (Bhandari et al. 2015, Koestler et al. 2008, Römpp et al. 2010).

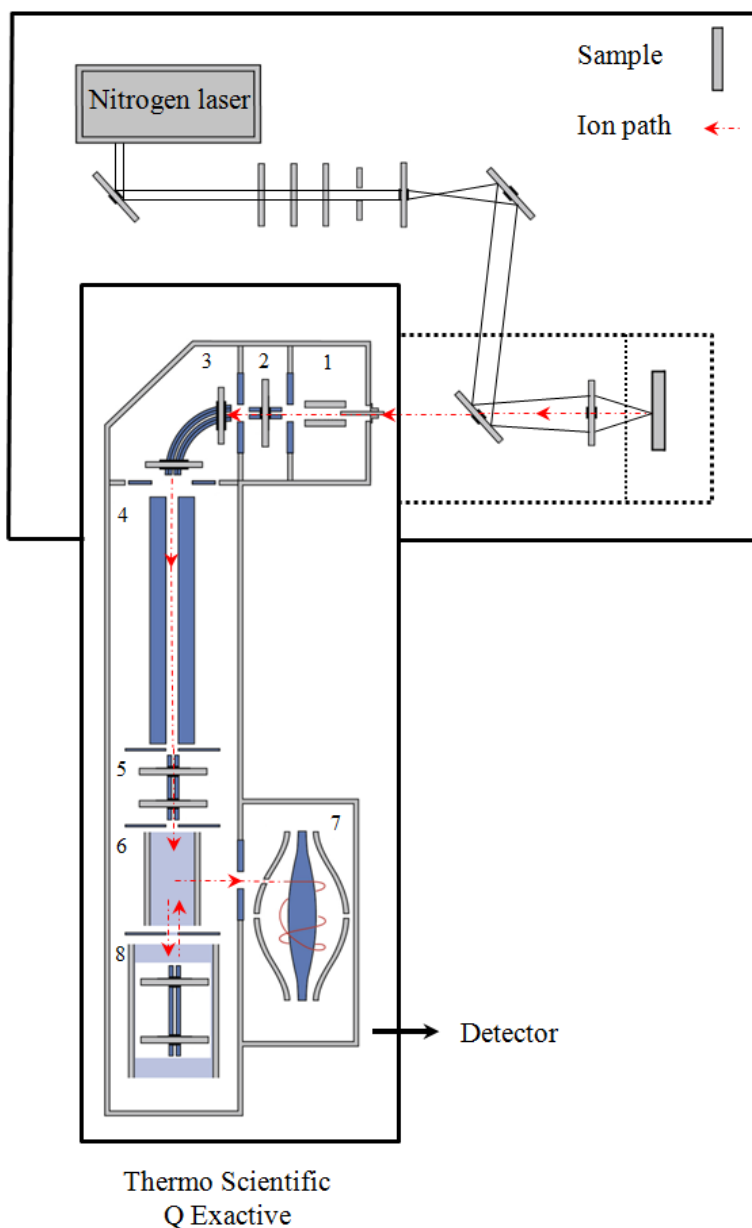


Figure 6: Scheme of AP-SMALDI imaging source connected to Q Exactive mass spectrometer: S-lens (1), injection flatapole (2), bent flatapole (3), quadrupole mass filter (4), octopole (5), C-trap (6), orbitrap (7), HCD collision cell (8) (http://planetorbitrap.com/q-exactive#.WCGx0_mLRpg, Koestler et al. 2008).

Once the sample is fixed on a sampler stage and the laser is fired, different ions are produced. The ions are transferred to the mass spectrometer, where they travel through the S-lens that increases the transmission of the ions (**Figure 6 (1)**). Subsequently, the ions go through injection flatapole and bent flatapole (**Figure 6 (2), (3)**). Both compartments prevent unwanted molecules (*i.e.* neutrals and high-velocity clusters) from getting into the system, therefore increasing experiment robustness and reducing the noise. The ions then

pass a quadrupole mass filter and octapole (**Figure 6 (4), (5)**) and enter the C-trap (**Figure 6 (6)**). The quadrupole mass filter allows movement of ions of only a specified mass range into the C-trap for accumulation, therefore enhancing sensitivity of MS and MS/MS experiments. The ions are then pushed into the orbitrap mass analyser (**Figure 6 (7)**) that traps the ions. The oscillations of the ions between the left-right halves of the trap relate to the m/z values. The image current of this left-right oscillation is then measured and Fourier-transform processed to record a mass spectrum. For MS/MS analysis, the ions are fragmented in Higher-energy Collision Dissociation (HCD) cell (**Figure 6 (8)**). The ions selected in the quadrupole pass through the C-trap and then are transferred to the HCD cell, where multiple collisions take place. The ions and their fragments are then returned to the C-trap and are injected to the orbitrap for mass analysis (Hao et al. 2012, Scheltema et al. 2014, Hu et al. 2005).

3.2.3.1.3. Data analysis

Spectra were saved in .raw files and converted to .imzML files using online free .raw to .imzML converter (<http://ms-imaging.org/wp/raw-to-imzml-converter/>). The data was then accessed in an online free MS imaging software (MSiReader ver. 0.09, <http://www4.ncsu.edu/~dcmuddim/msireader.html>). Results were exported in .xlsx format and further processed with MATLAB ver. R2012a software and free online MetaboAnalyst ver. 3.0 software (<http://www.metaboanalyst.ca/faces/home.xhtml>).

3.2.3.2. SIMS imaging

3.2.3.2.1. The J105-3D chemical imager

Samples with control CHO cells (non-induced) grown on silicon wafers were sent to Ionoptika Limited and analysed using a J105-3D chemical imager (Ionoptika Limited, Hampshire, UK) (**Figure 7**).

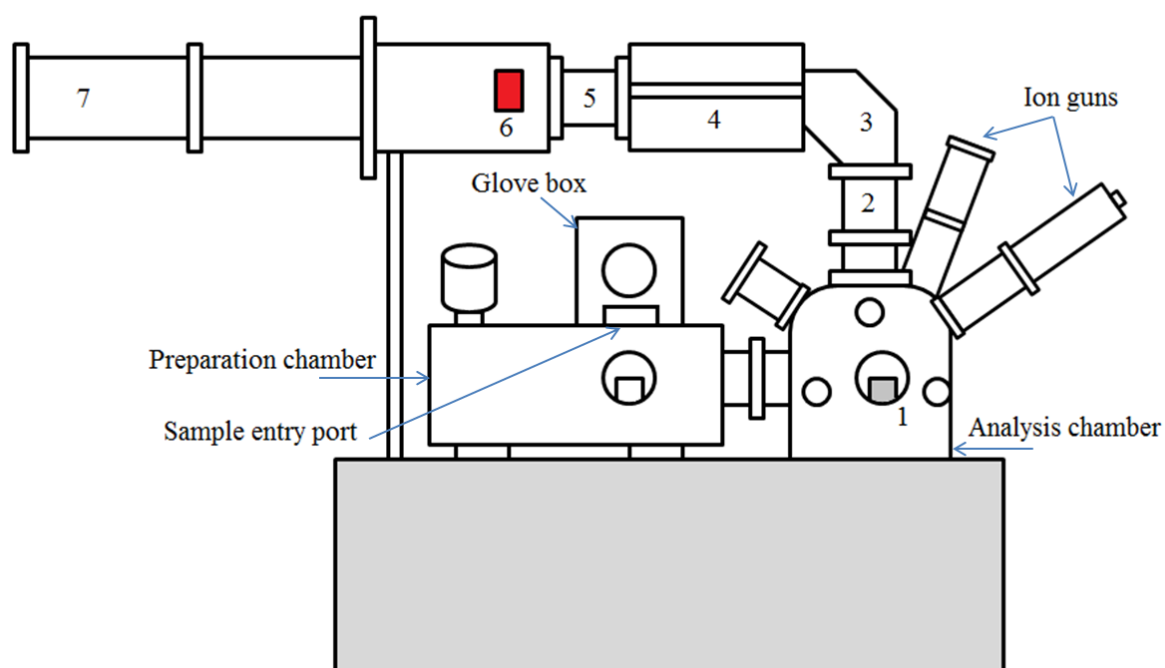


Figure 7: Scheme of J105-3D chemical imager: sample stage (1), quadrupole (2), electrostatic analyser (3), buncher (4), collision cell (5), detector (6), ToF reflectron (7) (Fletcher *et al.*, 2008).

Firstly, the sample is introduced to the glove box and further to the instrument through the sample entry port and the preparation chamber. The experimental setup is initiated by manipulating the sample, using camera-assisted sample positioning. The sample is then introduced to the sample analysis chamber by software-controlled automated sample handling. When the sample is placed on the sample stage (**Figure 7 (1)**), it is sputtered with the primary ion beam, resulting in a creation of secondary ions originating from the sample. The secondary ion beam enters the quadrupole element (**Figure 7 (2)**) and then go through an electrostatic analyser (**Figure 7 (3)**). Any excess energy ions are rejected by an electrostatic analyser, resulting in decrease of the secondary ion energy spread before the ions get injected to the buncher (**Figure 7 (4)**). Here, the secondary ions are collected and temporally condensed through ejection in a secondary ion pulse. The secondary ions are

then accelerated into the ToF mass analyser (**Figure 7 (7)**), where the ions are analysed according to their m/z and further detected (**Figure 7 (6)**). The instrument configuration allows for MS/MS experiments too. In MS/MS mode, before the secondary ions are injected to the ToF mass analyser, they enter the collision cell (**Figure 7 (5)**), where a selected precursor ion is dissociated. Precursor and product ions are then introduced to the ToF mass analyser and detected (Fletcher et al. 2008, Hill et al. 2011). Other ions are blanked before entry to the ToF by a timed ion gate.

The sample was sputtered with 25 keV Au_3^+ and 40 keV C_{60}^+ ion beams. Experiments included imaging of a range of different cell samples (**Table 1**). The instrument was operated in positive ionisation mode and spatial resolution of $<1\ \mu\text{m}$ and $\sim 8\ \mu\text{m}$. Experiments were also performed using ‘low quad’ and ‘high quad’ settings in order to record spectra with extra focus on low (‘low quad’) or high (‘high quad’) mass range.

Table 1: SIMS experiments performed using J105-3D chemical imager, settings for each experiment are included.

Sample	Settings	Primary ion gun
Control cells, area A	128 FOV, 256x256 pixels, mass range 20-800 Da, 6pA current	25 keV Au_3^+
Control cells, area B	128 FOV, 256x256 pixels, low quad, mass range 20-800 Da, 6pA current	25 keV Au_3^+
Control cells, area C	128 FOV, 256x256 pixels, high quad, mass range 70-1000 Da, 6pA current	25 keV Au_3^+
Control and induced CHO cells (day1)	64x64 pixels, mass range 30-900 Da, 110pA current, 10 layers	40 keV C_{60}^+
Control and induced CHO cells (day2)	64x64 pixels, mass range 30-900 Da, 70pA current, 10 layers	40 keV C_{60}^+

3.2.3.2.2. Data analysis

Spectra were saved in .dat files and further processed with MATLAB ver. R2012a software and J105 Image Analyser software (Ionoptika Limited, Hampshire, UK).

References:

- Bhandari, D. R. am, Schott, M., Römpf, A., Vilcinskas, A. and Spengler, B. (2015) 'Metabolite localization by atmospheric pressure high-resolution scanning microprobe matrix-assisted laser desorption/ionization mass spectrometry imaging in whole-body sections and individual organs of the rove beetle *Paederus riparius*', *Anal Bioanal Chem*, 407, pp. 2189–2201.
- Bruker* (2016). Available at: http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf (Accessed: 28 October 2016).
- Bujak, R., Struck-Lewicka, W., Markuszewski, M. J. and Kaliszan, R. (2015) 'Metabolomics for laboratory diagnostics', *J Pharma Biomed Anal*, 113, pp. 108–120.
- Dunn, W. B. (2008) 'Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes', *Phys Biol*, 5, p. 24.
- Fletcher, J. S., Rabbani, S., Henderson, A., Blenkinsopp, P., Thompson, S. P., Lockyer, N. P. and Vickerman, J. C. (2008) 'A New Dynamic in Mass Spectral Imaging of Single Biological Cells', *Anal Chem*, 80, pp. 9058–9064.
- GC-MS* (2016). Available at: http://cires1.colorado.edu/jimenez/CHEM-5181/Lect/2013_GCxGC_Handouts.pdf (Accessed: 17 November 2016).
- GC-MS Leco* (2016). Available at: <http://www.speciation.net/Database/Instruments/Leco-Inc/Pegasus-III-GCTOFMS-;i509> (Accessed: 17 November 2016).
- Hao, Z., Zhang, Y., Eliuk, S. and Blethrow, J. (2012) *A Quadrupole-Orbitrap Hybrid Mass Spectrometer Offers Highest Benchtop Performance for In-Depth Analysis of Complex Proteomes*. Bremen, Germany.
- Hill, R., Blenkinsopp, P., Thompson, S., Vickerman, J. and Fletcher, J. S. (2011) 'A new time-of-flight SIMS instrument for 3D imaging and analysis', *Surf Interface Anal*, 43, pp. 506–509.
- Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M. and Cooks, R. G. (2005) 'The Orbitrap: A new mass spectrometer', *J Mass Spectrom*, 40, pp. 430–443.
- Imzml converter* (2016). Available at: <http://ms-imaging.org/wp/raw-to-imzml-converter/> (Accessed: 7 November 2016).
- Khalil, S. M., Römpf, A., Pretzel, J., Becker, K. and Spengler, B. (2015) 'Phospholipid Topography of Whole-Body Sections of the *Anopheles stephensi* Mosquito, Characterized by High-Resolution Atmospheric-Pressure Scanning Microprobe Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging', *Anal Chem*, 87, pp. 11309–11316.
- Koestler, M., Kirsch, D., Hester, A., Leisner, A., Guenther, S. and Spengler, B. (2008) 'A high-resolution scanning microprobe matrix-assisted laser desorption/ionization ion source for imaging analysis on an ion trap/Fourier transform ion cyclotron resonance mass spectrometer', *Rapid Commun Mass Spectrom*, 22, pp. 3275–3285.
- Li, Y. L. and Gross, M. L. (2004) 'Ionic-liquid matrices for quantitative analysis by MALDI-TOF mass spectrometry', *J Am Soc Mass Spectrom*, 15, pp. 1833–1837.
- MetaboAnalyst 3.0* (2016). Available at: <http://www.metaboanalyst.ca/faces/home.xhtml>

(Accessed: 7 November 2016).

MSiReader (2016). Available at: <http://www4.ncsu.edu/~dcmuddim/msireader.html>
(Accessed: 7 November 2016).

Römpp, A., Guenther, S., Schober, Y., Schulz, O. and Takats, Z. (2010) 'Histology by Mass Spectrometry: Label-Free Tissue Characterization Obtained from High-Accuracy Bioanalytical Imaging', *Angew Chem Int Ed Engl*, 49, pp. 3834–3838.

Römpp, A. and Spengler, B. (2013) 'Mass spectrometry imaging with high resolution in mass and space', *Histochem Cell Biol*, 139, pp. 759–783.

Scheltema, R. A., Hauschild, J.-P., Lange, O., Hornburg, D., Denisov, E., Kuehn, A., Makarov, A., Mann, M., Performance, H., Chromatography, L. and Damoc, E. (2014) 'The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High Performance Quadrupole and an Ultra-High Field Orbitrap Analyzer', *Mol Cell Proteomics*, 13, pp. 3698–3708.

SMALDIPrep (2016). Available at: <http://www.uni-giessen.de/faculties/f08/departments/iaac/spengler/research/instrumentations/transmit-smaldiprep> (Accessed: 18 November 2016).

Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J. and Holle, A. (2003) 'A novel MALDI LIFT-ToF/ToF mass spectrometer for proteomics', *Anal Bioanal Chem*, 376, pp. 952–965.

Sumner, L. W., Samuel, T., Noble, R., Gmbh, S. D., Barrett, D., Beale, M. H. and Hardy, N. (2007) 'Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)', *Metabolomics*, 3, pp. 211–221.

Thermo Scientific, Q Exactive (2016). Available at: http://planetorbitrap.com/q-exactive#.WCGx0_mLRpg (Accessed: 8 November 2016).

Waters Corporation (2016). Available at:
http://www.waters.com/waters/en_GB/SYNAPT-G2-Si-High-Definition-Mass-Spectrometry/nav.htm?cid=134740622&locale=en_GB (Accessed: 28 October 2016).

CHAPTER 4: Expression of EPO-Fc fusion protein in transfected CHO cell line

4.1. Introduction

Changes in cell metabolism can be caused by a variety of different factors, for example cell culture exposure to specific inducers, resulting in protein expression in these cells (Wilson, 2013). A number of preliminary techniques are applied in order to evaluate such changes and investigate their potency.

Western Blot is an analytical technique used to detect specific protein in *e.g.* tissue homogenate or cell medium and involves two steps: gel electrophoresis to separate proteins according to the polypeptide length followed by a transfer of the proteins from the gel to a nitrocellulose membrane and their detection by a specific detection antibody (**Figure 1**) (Gilda & Gomes 2013, Kar *et al.* 2015, Lee 2008).

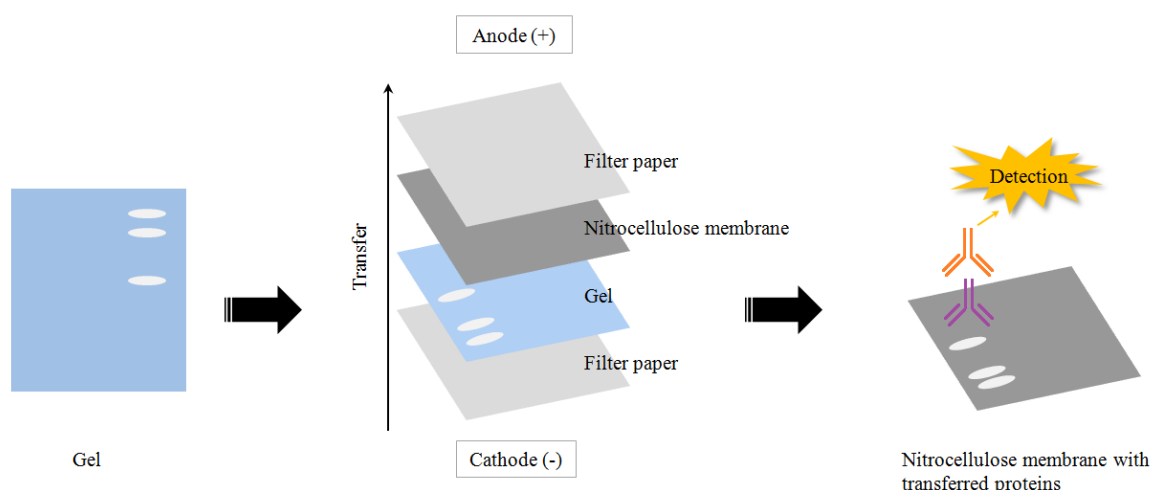


Figure 1: Schematic of western blotting technique: a gel electrophoresis is performed to separate proteins and is followed by a transfer of these proteins from the gel to nitrocellulose membrane using electric current. Once the proteins are transferred onto the membrane, target protein detection is completed using specific protein antibodies.

SDS-PAGE is a method used to separate macromolecules such as proteins according to their electrophoretic mobility that depends on the length of the polypeptides. The first step of this procedure involves a sample preparation, where the protein samples are mixed with buffer containing anionic detergent *i.e.* SDS. Commonly used buffer for protein sample preparation for SDS-PAGE is Laemmli buffer, containing bromophenol blue and glycerol.

The bromophenol blue acts as a dye front of the gel that runs with proteins, therefore helping in sample localisation. It also helps to observe the sample during loading onto a gel. The glycerol raises the density of the sample so that it stays in the sample well. The beta-mercaptoethanol is used to reduce disulfide bonds within proteins (Kurien & Scofield 2006, Gilda & Gomes 2013). The addition of SDS causes denaturation of the proteins and imparts a negative charge to denaturated proteins. Thus, negatively charged proteins can migrate from the negative electrode towards the positive electrode within polyacrylamide gels and be separated based on their mass only. Depending on their mass, each protein migrates differently through the gel: large molecules move slower, small molecules move faster (Lee 2008, Kurien & Scofield 2006).

The following step in SDS-PAGE procedure is the preparation of stacking and resolving buffers. Stacking and resolving buffers are used in order to stabilise the pH value of the experiment set up to the desired pH value (Jensen, 2012). Moreover, the stacking gel is a gel that has lower concentration of polyacrilamide than resolving gel, which is more concentrated. The stacking gel is usually fixed on the top of the resolving gel and it is used to increase the electrophoresis resolution due to its concentrating effect on the sample, right at the beginning of the resolving gel (Lee 2008, Jensen 2012). In addition, both stacking and resolving gels consists of ammonium persulfate (APS) that plays a source of free radicals that initialise formation of a gel and *N, N, N', N'*-tetramethylethylenediamine (TEMED), which stabilises free radicals and enhances polymerisation (Jensen 2012, Kurien & Scofield 2006). The last step of SDS-PAGE process is preparation of the electrophoresis buffer containing glycine in order to maintain electric neutrality within the buffer and gels (Kurien & Scofield 2006, Gilda & Gomes 2013).

Once an electrophoresis is performed, proteins from the gel are transferred onto a nitrocellulose membrane using an electric current. This is usually completed with blotting buffer containing alcohol that promotes binding of proteins to the membrane (Kurien and Scofield, 2006).

This chapter presents a preliminary analysis of a response of CHO cell line exposed to a tetracycline as an inducer for EPO-Fc fusion protein expression.

4.2. Sample preparation

Cells were grown as a monolayer in a 6-well dish at a density of 0.05×10^6 cells/mL in 2.5 mL/well of complete growth medium for 4 days (3 biological replicates). Cell dishes were assessed every day under inverted phase contrast microscopy to observe cell morphology and to determine level of confluency. On each data point (*i.e.* day 1, 2, 3 and 4) the old medium was removed and collected and cells were washed with 3 mL of pre-warmed PBS in order to remove medium and FBS remains. The PBS was aspirated and 1.5 mL of trypsin was added to each well and dishes were placed in the incubator for 5 min. Subsequently, 3 mL of pre-warmed cell medium was added and cell suspension was shaken to disattach cells. Cells were then counted in a Thoma cell counting chamber (1:1, v/v cell suspension with Trypan blue). The number of cells and viability were recorded to obtain a cell growth curve. Previously removed medium was then spun (5 min, 1000 g) to remove cell debris and 50 μ L of the supernatant was collected to Eppendorf tubes and kept frozen (-20 °C) for Western Blot analysis.

4.3. Methods

The following experiments were performed by Thomas Leach, who kindly agreed to help with the analysis. A gel electrophoresis and western blotting were accomplished according to standard procedures (Lee 2008, Li *et al.* 2002, Kurien & Scofield 2006) and protocols used in Prof. Alan J. Dickson lab. All reagents were purchased from Sigma-Aldrich (Gillingham, UK), unless stated otherwise.

4.3.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was completed using the Mini-PROTEAN Tetra Cell system (Bio-Rad Laboratories Ltd., Hertfordshire, UK). Protein molecular masses were extrapolated by measuring against the Color prestained protein standard, broad range 11-245 kDa (BioLabs Inc., Ipswich, MA, USA). Firstly, stacking and resolving buffers were prepared: stacking buffer 0.5 M Tris-HCl, pH 6.8, resolving buffer 1.5 M Tris-HCl, pH 8.8. Subsequently, stacking and resolving gels were made according to the formulation presented in **Table 1**.

Table 1: Formulation used for preparation of SDS-PAGE gels.

	Stacking gel	Resolving gel
H ₂ O	9 mL	5 mL
30 % [w/v] Acrylamide 0.8 % <i>N, N'</i> -bis-acrylamide	2 mL	6 mL
Buffer*	3.75 mL	3.75 mL
10% SDS	150 µL	150 µL
10% APS*	75 µL	75 µL
TEMED*	15 µL	7.5 µL

* Added immediately before use; APS: ammonium persulfate, TEMED: *N, N, N', N'*-tetramethylethylenediamine.

Resolving gel was fixed with the addition of an isopropanol layer on the top to prevent air bubbles formation. The isopropanol layer was removed once the resolving gel had fixed, and stacking gel was then fixed on the top of the resolving gel. This was followed with a preparation of SDS-PAGE samples. Medium samples that had been removed from cell culture were mixed 1:1 (v/v) with 2x Laemmli buffer (0.01 % (w/v) bromophenol blue, 20 % (v/v) glycerol, 0.125 mM Tris pH 6.8) and 1.8 % (v/v) beta-mercaptoethanol and then boiled at 95 °C for 5 min prior to loading on a gel. Laemmli buffer was with the addition of anionic detergent (SDS) in order to denature the proteins and to impart a negative charge to transformed proteins. 25 µL of the medium samples, the standard protein mass ladder (**Figure 2 A**) and positive control for Epo-Fc were then loaded on a gel and the electrophoresis in a electrophoresis buffer (15 mM Tris, 192 mM glycine and 0.1 % (w/v) SDS solution, pH 8.3) was performed. Initial electrophoresis was set at 80 V until the bromophenol had passed to the resolving gel. Voltage was then set to 180 V until the bromophenol had begun to run off the bottom of the gel. Gels were then stained with InstantBlue, proprietary Coomassie stain (Expedeon Inc., San Diego, CA, USA) for < 1 h in order to visualise the proteins. There were two SDS-PAGE gels: gel with medium samples removed from non-induced cells and gel with medium samples removed from induced cells. Schematic illustration of the gel is presented in **Figure 2 B**. Gels were stored in deionised water until further analysis.

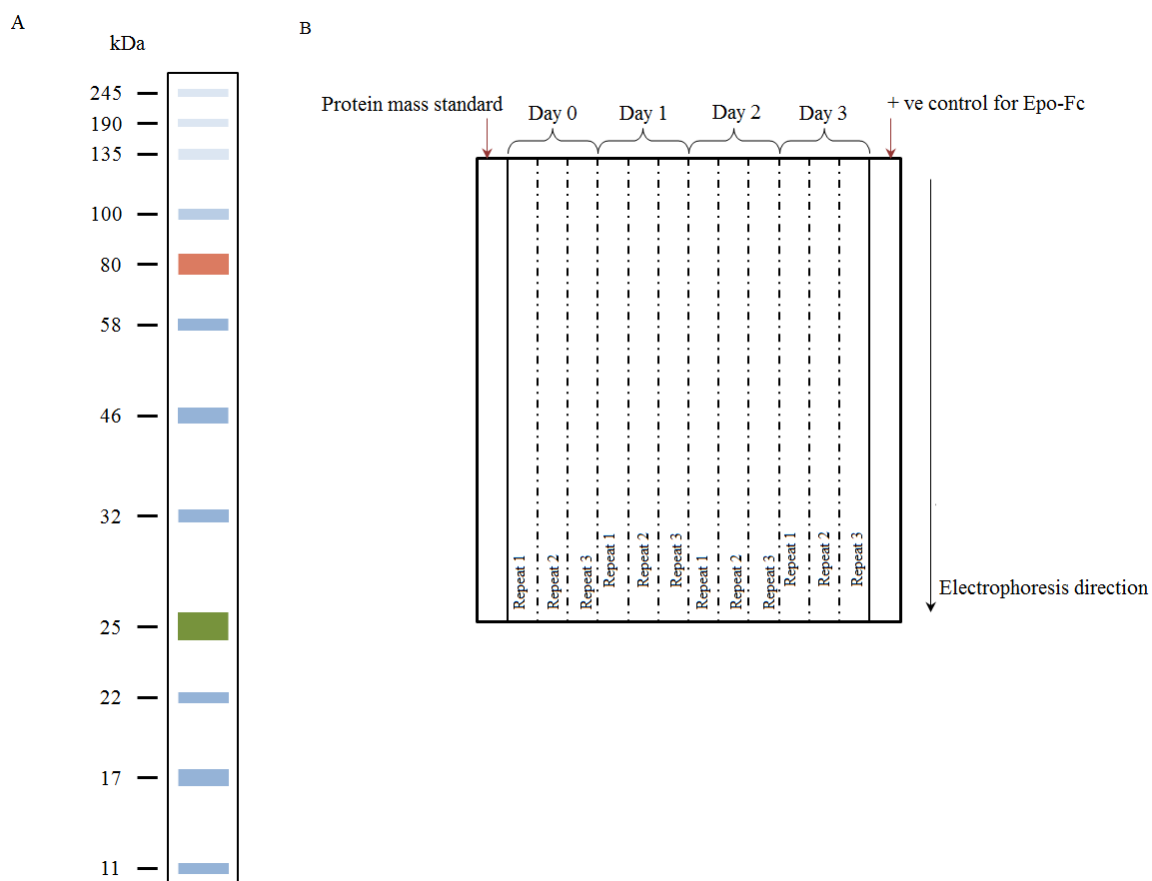


Figure 2: Color prestained protein standard, broad range (11-245 kDa). Adopted from (<https://www.neb.com/products/p7712-color-prestained-protein-standard-broad-range-11-245-kda>) (A). Schematic illustration of the sample organisation on the SDS-PAGE gel (B).

4.3.2. Western Blotting

Western Blot with SDS-PAGE gels prepared previously was performed using the TE 22 Mini Tank Transfer Unit (Amersham Biosciences Ltd., Buckinghamshire, UK). SDS-PAGE gels, nitrocellulose membrane and filter papers were soaked in blotting buffer (25 mM Tris, 192 mM glycine, 20 % (v/v) MeOH; pH 7.4) prior the transfer. The transfer was performed and run overnight at 70 V at 4 °C. In order to check the transfer efficiency, Ponceau stain (0.1 % (w/v) Ponceau in 5 % (v/v) acetic acid) was added to the nitrocellulose membrane and was incubated for 30 s. Once the protein bands were visualised, the stain was washed off with PBS (with 0.075 % (v/v) Tween20 to prevent non-specific protein-protein interactions (Kurien & Scofield 2006, Lee 2008).

Once the proteins were transferred on a nitrocellulose membrane, the membrane was incubated for 1 h at room temperature in blocking protein solution (3 % (w/v) milk PBS with 0.075 % Tween20). After blocking step, 10 µL of primary mouse anti-His antibody

(Abcam, UK) (2 $\mu\text{g/mL}$ in PBS) was added to the membrane and membrane was incubated for 1 h at room temperature. The primary antibody binds to His tag on Fc fragment of EPO-Fc fusion protein. The incubation was then followed by washing the membrane with 3 % (w/v) milk PBS with 0.075 % Tween20 for 5 min. The washing step was repeated 3 times to remove excess of unbound primary antibody. After washing, the membrane was incubated for 1 h at room temperature with secondary anti-mouse-IgG conjugated IR-dye antibody (Biosciences, USA) (dilution 1:15000 in 3 % (w/v) milk PBS with 0.075 % Tween20). The secondary antibody has bispecificity function and binds to both: the primary antibody and IgG portion of the EPO-Fc construct. The incubation was then followed by washing step with 3 % (w/v) milk PBS with 0.075 % Tween20 for 5 min and was repeated 3 times to remove excess of secondary antibody. A schematic Western Blot configuration of EPO-Fc detection is presented in **Figure 3**. The blots were then observed under IR (LI-COR, USA) and the images were recorded.

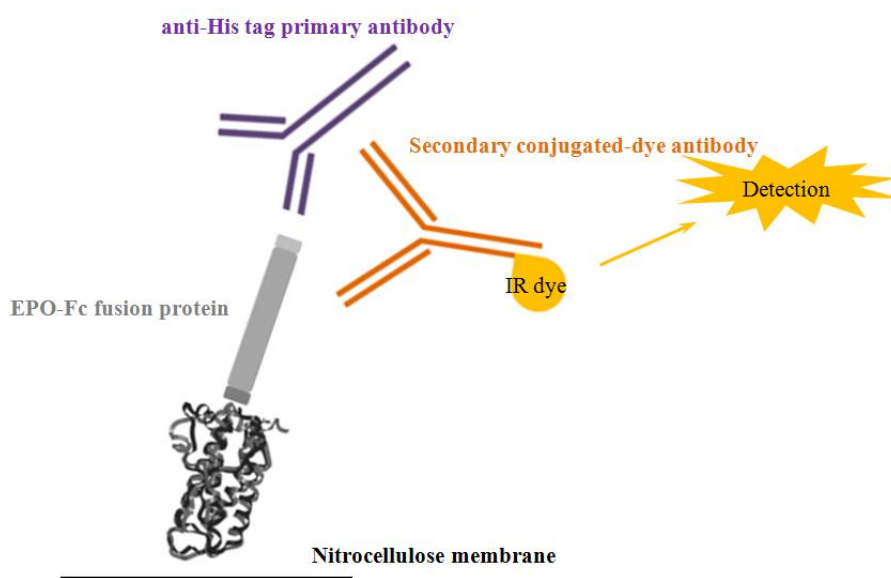


Figure 3: A schematic Western Blot configuration of EPO-Fc detection: EPO-Fc protein transferred from gel electrophoresis onto nitrocellulose membrane is recognised by primary antibody; secondary antibody that binds to primary antibody and EPO-Fc protein is conjugated with IR-dye allowing the construct for detection.

4.4. Results and discussion

4.4.1. Cell growth curve

Plotting cell growth curves and measuring viability of cells are the basic analyses routinely performed on mammalian cell cultures (European Medicines Agency, 2008). In order to evaluate a general cell line behaviour and response to the inducer, both non-induced ('control') cells and cells with the inducer present in the cell medium ('induced') were grown. The transfected CHO cell growth curves for control and induced cells present relatively common trend of growing cell culture (Porter *et al.*2010). However, some differences of induced cells to controls cells can be observed. The control and induced cells have similar growth during first two days of the cell culture (**Figure 4, green**). Both control and induced cells show a 24 h adaptation phase to growth conditions. After the adaptation phase (> 24 h), control and induced cells enter the early exponential phase and mid exponential phase after > 48 h. On both day 1 and day 2, a small decrease in a number of cells (0.07 mln cells) in the induced cell culture is seen, when compared to control cells. On the other hand, on the third day of the incubation, a noticeable drop in a number of cells (0.8 mln cells) in induced cell culture is observed, when compared to control cells. However, despite a large decrease of a number of induced cells on day 3, the cells return to the density similar to the control cells on day 4 of cell culture. Also, after the third day of cell culture, a difference between growth phases of control cells and induced cells could be seen: after day 3, the control cells seem to enter a decline phase, whereas the induced cells are still in exponential phase or start to enter the stationary phase.

Figure 4 implies that the addition of tetracycline to the transfected CHO cell culture and the protein secretion might cause a decrease in a number of induced cells, when compare to the control cell population. Differences in cell line behaviour due to diversity in environmental surrounding have previously been described. Lack of control of parameters like pH, temperature, pressure of CO₂ or other agents (*e.g.* drugs) will result in dissimilarities in the microenvironment of cell lines, deriving with a heterogeneous product secretion (Jain & Kumar 2008, O'Callaghan *et al.*2010), changes in cell culture longevity and cell metabolism (Porter *et al.*2010). It was shown that engineered cells carry a different metabolism and are more energy-demanding with decreased lactate production per glucose consumed than wild-type cells (Wilkens and Gerdtzen, 2015). Therefore, the lower numbers of induced cells in this study might be due to increased metabolism of these cells as well as faster depletion of nutrients within growth cell medium (Castillo *et al.*2005) when compare to control cells.

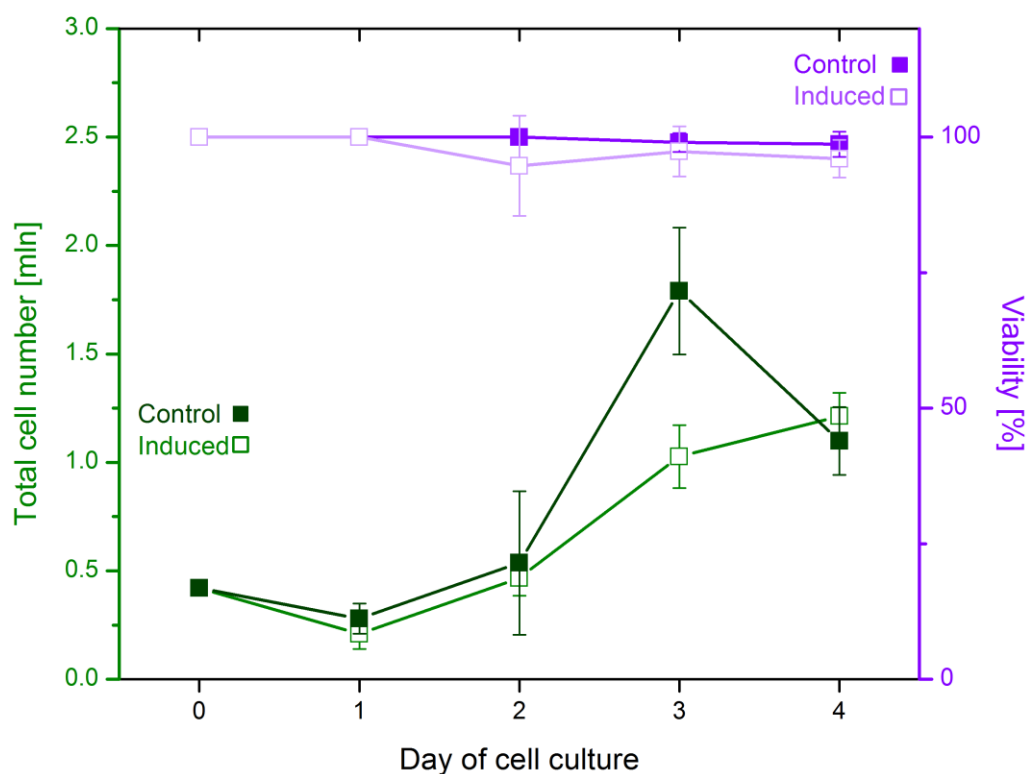


Figure 4: Transfected CHO cells growth curves (green) and cell viability curves (purple) in the presence (‘induced’) and absence (‘control’) of the EPO-Fc protein inducer. The error bars represent the standard deviation calculated from the data obtained from three experiments ($n=3$).

According to the protocols for growth and maintenance of T-REx CHO cell lines, a fresh growth medium should be added to cells every 3-4 days of cell culture. In order to avoid introducing additional fluctuations of compounds and metabolite concentration in the cell culture, four-day cell growth cell culture was therefore conducted.

In this study, cells were induced with tetracycline in order to produce EPO-Fc fusion protein. Tetracycline is commonly used as a wide spectrum antibiotic. In T-REx CHO cell line, tetracycline is applied as an inducer. The antibiotic activates the protein production within 24 hours of incubation with cells hence, after this time cells start producing EPO-Fc construct, which is then released into culture medium. A tetracycline concentration was also adjusted to stimulate cells for EPO-Fc production only, without influencing their normal growth (<https://www.thermofisher.com/uk/en/home/references/protocols/proteins-expression-isolation-and-analysis/protein-expression-protocol/inducible-protein-expression-using-the-trex-system.html>). According to the above, two different time-points

were chosen in order to investigate metabolic changes occurring in cells under inducer exposure with early stages of EPO-Fc production (day 1) and a stable phase of EPO-Fc production (day 3). The selected time-points (*i.e.* day 1 and day 3) fall within the exponential phase of CHO cells growth and therefore, are suitable for metabolomics investigations. The fourth day of cell culture was not considered for further investigations due to the decline of cell growth. During decline phase, cells die because of a number of different factors such as lack of nutrients, too high concentration of waste metabolites or changes in microenvironment. Unlike the decline phase, growth during exponential phase is fairly constant and reproducible for a given set of cell growth conditions (ATCC 2014, Szuster & Kosz-Vnenchak 2005, Sellick *et al.*2010).

Simultaneously, the viability of control and induced cells was recorded (**Figure 4, purple**) and was in line with common viability curves for growing CHO cells (Xie *et al.*2016, Porter *et al.*2010). Despite the biggest decline of cell viability of induced cells appearing on day 2 of cell culture (5 % less than control cells), the overall viability of induced cells did not differ greatly when compared to the control cell population and accounted for > 90% during all days of cell culture.

4.4.2. EPO-Fc fusion protein production

In order to investigate the tetracycline-induced protein production in transfected CHO cells, growth cell media were removed from control and induced cells (3 days inducing time prior sample collection) on days 0, 1, 2 and 3 of cell culture. **Figure 5** presents Western Blots of medium samples removed from control cells (control) and cells induced for production of EPO-Fc (induced).

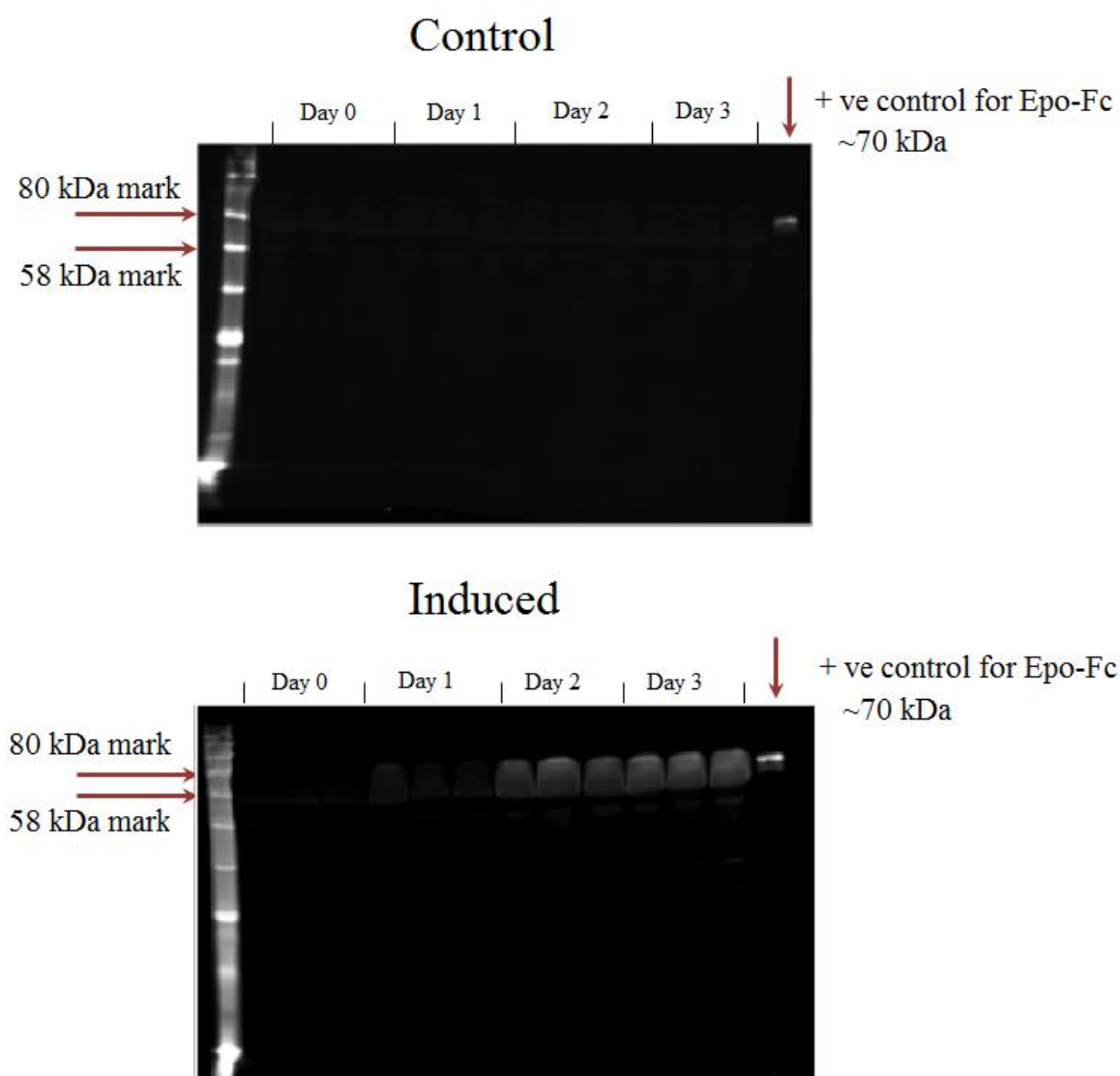


Figure 5: Western Blots of medium samples removed from control cells (control) and cells induced for production of EPO-Fc (induced). Each blot contains a protein standard ladder with 80 kDa and 58 kDa bands marked and positive control of purified EPO-Fc (~70 kDa).

Each blot contains a protein standard ladder with 80 kDa and 58 kDa bands marked and positive control of purified EPO-Fc (~70 kDa). Since control cells were not induced, no EPO-Fc expression was expected. Consequently, the blot appears to show very low to no expression in the control samples. This slight expression of the protein might be caused by the presence of fetal bovine serum (FBS) in the cell culture medium. Many FBS lots contain low amounts of tetracycline as FBS is separated from cows that have been fed a diet containing tetracycline (<https://www.thermofisher.com/uk/en/home/references/protocols/proteins-expression-isolation-and-analysis/protein-expression-protocol/inducible-protein-expression-using-the-trex-system.html>). Therefore, this might be a possible reason for a low, basal expression of the EPO-Fc in medium samples from control cells.

Another explanation for the low EPO-Fc expression in control cells might be due to tetracycline promoters being 'leaky'. As mentioned previously, CHO cell line consists of EPO-Fc plasmid that is linked to tetracycline promoter. In the presence of tetracycline (inducer), the promoter is activated and EPO-Fc is produced. Apart from tetracycline being an inducer for EPO-Fc production, there are additional activators that regulate tetracycline-dependent protein production for example so called tetracycline-responsive element (TRE). Studies on TRE have shown that this activator happens to bind weakly to tetracycline repressor in the absence of tetracycline or its analogue doxycycline, resulting in a very low expression of the encoded protein within a transfected plasmid (Hosoda *et al.* 2011, Uchida *et al.* 2006).

By contrast, the blot with medium samples from induced cells confirms an increasing concentration of EPO-Fc during the days of incubation. As shown in **Figure 5**, the expression of EPO-Fc seems to be very low on day 1 of cell culture whereas visible increase of the protein concentration can be observed on day 2 and 3. This also confirms the incubation time needed for tetracycline to induce EPO-Fc production estimated to be ~24 hours.

The EPO-Fc bands as well as positive control for EPO-Fc lie between the 80 kDa and 58 kDa marks on the protein standard ladder. The bands of EPO-Fc detected in medium samples from induced cells appear as wide smudges, possibly due to the serum proteins (present in growth medium) are in high concentrations. The serum proteins interfere with EPO-Fc and its mobility on gel electrophoresis, resulting in coverage of EPO-Fc region on blots. The blot with wide bands of EPO-Fc seen in medium samples of induced cells can be however improved by the protein purification. Protein purification would allow for separation of the EPO-Fc from all other parts of the mixture (Jain and Kumar, 2008) that

would enable for more precise EPO-Fc identification. Nevertheless, since the EPO-Fc protein secretion has been confirmed in induced CHO cell line, these results were sufficient and further analysis was not essential.

The wide bands of EPO-Fc could also be a result of different glycosylation profiles of this protein. Protein glycosylation is a post-translational process, where a number of glycans are coupled to the protein. Due to sugar residues take part in many biological pathways and mechanisms such as protein recognition, biological efficacy of secreted protein and its serum half-life, glycosylation is considered as a critical step (Butler and Spearman, 2014). Additionally, because of diverse functionality of sugar residues taking part in many biological processes, glycosylation is considered as a quite heterogeneous mechanism and differs notably, depending on the specific growth conditions of the cells producing certain protein (Woo *et al.*2008, Lee *et al.*2012). Endogenous erythropoietin is a highly glycosylated protein, of which ~40 % of its total molecular weight (34 kDa) is made up of different N- and O-glycan bonds (Wojchowski *et al.*1987). Similarly, recombinant EPO has been shown to have various glycosylation profiles that have been described previously (Hua *et al.*2015, Taschwer *et al.*2012). Accordingly, the diverse molecular mass of EPO-Fc fusion protein seen in **Figure 5** (blot for induced) can be therefore due to a variety of glycoforms attached to this protein. Nevertheless, cellular glycosylation of recombinant proteins can be influenced by an additional number of factors such as a host cell line, cell metabolism, medium components, cell culture conditions or a protein structure and therefore, should be strictly controlled to keep a quality and biological activity of a desired glycoprotein (Reinhart *et al.*2015, Del Val *et al.*2013, Butler & Spearman 2014).

4.5. Conclusions and future work

The work presented here has demonstrated that EPO-Fc protein secretion into cell growth medium increased with progressing growth of cell culture. The results also confirmed the EPO-Fc fusion protein production in tetracycline-induced CHO cells that appear to be after ~24 h from applying the inducer. Therefore, in order to investigate metabolism changes in CHO cells under inducer exposure as well as during early and stable phase of the protein production process, day 1 and day 3 of cell culture were chosen as time-points for metabolome sample collection for further analyses. The results could be however improved by protein purification that would enable for more precise EPO-Fc protein identification. Moreover, it would also be beneficial to perform the analysis of the cell line that has not been transfected with EPO-Fc-pcDNA5-FRT-TO plasmid in order to observe the influence of this element on cell morphology, growth and CHO cell metabolism.

References:

- ATCC (2014) *Animal Cell Culture Guide*. Manassas, VA. Available at: www.atcc.org.
- Butler, M. and Spearman, M. (2014) 'The choice of mammalian cell host and possibilities for glycosylation engineering', *Curr Opin Biotechnol*, 30, pp. 107–112.
- Castillo, A. J., Vítores, S., Faife, E., Rabasa, Y. and Luz, K. R. D. E. L. A. (2005) 'Development of an Integrated Strategy for Recombinant Cell Line Selection', in Gòdia, F. and Fussenegger, M. (eds) *Animal Cell Technology Meets Genomics. ESACT Proceedings*. Dordrecht: Springer, pp. 505–507.
- Color prestained protein standard, broad range* (2016). Available at: <https://www.neb.com/products/p7712-color-prestained-protein-standard-broad-range-11-245-kda> (Accessed: 25 November 2016).
- European Medicines Agency (2008) *Guideline on human cell-based medicinal products*. England. doi: EMEA/CHMP/410869/2006.
- Gilda, J. E. and Gomes, A. V (2013) 'Stain Free Total Protein Staining is a Superior Loading Control to β -Actin for Western Blots', *Anal Biochem*, 440, pp. 186–188.
- Hosoda, H., Miyao, T., Uchida, S., Sakai, S. and Kida, S. (2011) 'Development of a tightly-regulated tetracycline-dependent transcriptional activator and repressor co-expression system for the strong induction of transgene expression', *Cytotechnology*, 63, pp. 211–216.
- Hua, S., Oh, M. J., Ozcan, S., Seo, Y. S., Grimm, R. and An, H. J. (2015) 'Technologies for glycomic characterization of biopharmaceutical erythropoietins', *Trends Anal Chem*, 68, pp. 18–27.
- Jain, E. and Kumar, A. (2008) 'Upstream processes in antibody production: Evaluation of critical parameters', *Biotechnol Adv*, 26, pp. 46–72.
- Jensen, E. C. (2012) 'The basics of Western blotting', *Anatom Record*, 295, pp. 369–71.
- Kar, P., Agnihotri, S. K., Sharma, A., Sachan, R., Bhatt, M. L. and Sachdev, M. (2015) 'A novel methodology for stripping and reprobing of Western blots originally developed with colorimetric substrate TMB', *Methods Mol Biol*, 1314, pp. 349–355.
- Kurien, B. T. and Scofield, R. H. (2006) 'Western Blotting: A Guide to Current Methods', *Methods*, 38, pp. 283–293.
- Lee, C. (2008) 'Western Blotting', in Rosato, E. (ed.) *Circadian Rhythms. Methods and Protocols*. Humana Press, pp. 391–399.
- Lee, J. S., Ha, T. K., Lee, S. J. and Lee, G. M. (2012) 'Current state and perspectives on erythropoietin production', *Appl Microbiol Biotechnol*, 95, pp. 1405–1416.
- Li, W., Murai, Y., Okada, E., Matsui, K., Hayashi, S., Horie, M. and Takano, Y. (2002) 'Modified and simplified western blotting protocol: Use of intermittent microwave irradiation (IMWI) and 5% skim milk to improve binding specificity', *Pathology Intern*, 52, pp. 234–238.
- Life Technologies* (2016). Available at: <https://www.thermofisher.com/uk/en/home/references/protocols/proteins-expression->

isolation-and-analysis/protein-expression-protocol/inducible-protein-expression-using-the-trex-system.html (Accessed: 4 November 2016).

O'Callaghan, P. M., McLeod, J., Pybus, L. P., Lovelady, C. S., Wilkinson, S. J., Racher, A. J., Porter, A. and James, D. C. (2010) 'Cell line-specific control of recombinant monoclonal antibody production by CHO cells', *Biotechnol Bioeng*, 106, pp. 938–951.

Porter, A. J., Dickson, A. J. and Racher, A. J. (2010) 'Strategies for selecting recombinant CHO cell lines for cGMP manufacturing: realizing the potential in bioreactors', *Biotechnol Prog*, 26, pp. 1446–1454.

Reinhart, D., Damjanovic, L., Kaisermayer, C. and Kunert, R. (2015) 'Benchmarking of commercially available CHO cell culture media for antibody production', *Appl Microbiol Biotechnol*, 99, pp. 4645–4657.

Sellick, C. A., Hansen, R., Jarvis, R. M., Maqsood, A. R., Stephens, G. M., Dickson, A. J. and Goodacre, R. (2010) 'Rapid monitoring of recombinant antibody production by mammalian cell cultures using Fourier transform infrared spectroscopy and chemometrics', *Biotechnol Bioeng*, 106, pp. 432–442.

Szuster, A. and Kosz-Vnenchak, M. (2005) 'Autocrine growth regulation of W12 and GCA cells in culture', *Folia Histochem Cytobiol*, 43, pp. 91–102.

Taschwer, M., Hackl, M., Hernandez Bort, J. A., Leitner, C., Kumar, N., Puc, U., Grass, J., Papst, M., Kunert, R., Altmann, F. and Borth, N. (2012) 'Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth', *J Biotechnol*, 157, pp. 295–303.

Uchida, S., Sakai, S., Furuichi, T., Hosoda, H., Toyota, K., Ishii, T., Kitamoto, A., Sekine, M., Koike, K., Masushige, S., Murphy, G., Silva, A. J. and Kida, S. (2006) 'Tight regulation of transgene expression by tetracycline-dependent activator and repressor in brain', *Genes Brain Behav*, 5, pp. 96–106.

Del Val, I. J., Kyriakopoulos, S., Polizzi, K. M. and Kontoravdi, C. (2013) 'An optimized method for extraction and quantification of nucleotides and nucleotide sugars from mammalian cells', *Anal Biochem*. Elsevier Inc., 443, pp. 172–180.

Wilkens, C. A. and Gerdtzen, Z. P. (2015) 'Comparative metabolic analysis of CHO cell clones obtained through cell engineering, for IgG productivity, growth and cell longevity', *PLoS One*, 10, pp. 1–15.

Wilson, D. F. (2013) 'Regulation of cellular metabolism: programming and maintaining metabolic homeostasis', *J Appl Physiol*, 115, pp. 1583–8.

Wojchowski, D. M., Orkin, S. H. and Sytkowski, A. J. (1987) 'Active human erythropoietin expressed in insect cells using a baculovirus vector: a role for N-linked oligosaccharide', *Biochim Biophys Acta*, 910, pp. 224–232.

Woo, S. A., Jeon, J. J., Jeong, Y. R., Seung, J. L. and Sung, K. Y. (2008) 'Effect of culture temperature on erythropoietin production and glycosylation in a perfusion culture of recombinant CHO cells', *Biotechnol Bioeng*, 101, pp. 1234–1244.

Xie, P., Niu, H., Chen, X., Zhang, X., Miao, S., Deng, X., Liu, X., Tan, W. S., Zhou, Y. and Fan, L. (2016) 'Elucidating the effects of pH shift on IgG1 monoclonal antibody acidic charge variant levels in Chinese hamster ovary cell cultures', *Appl Microbiol Biotechnol*, 100, pp. 10343–10353.

CHAPTER 5: MALDI-MS sample preparation optimisation for metabolite profiling

5.1. Introduction

Since the invention of the MALDI-MS (Karas, Bachmann and Hillenkamp, 1985) and its application to the analysis and identification of large molecules (Serafín-Higuera *et al.* 2016, van Linde *et al.* 2016), MALDI-MS has developed to become a powerful tool for the study of small molecules (Korte & Lee 2014, Trim & Snel 2016). Although there is a number of sample preparation procedures well suited for metabolite measurements, a few challenges remain to be overcome such as interferences between the matrix and the analyte ions, ionisation selectivity (Lou *et al.*, 2009), sample complexity (Trim and Snel, 2016) and sample homogeneity, especially when applied to biological imaging (Schober *et al.* 2012, Römpp & Spengler 2013).

In MALDI-MS of small molecules, the choice of a matrix is an essential aspect as it can greatly influence the quality of the collected data. A matrix is a low molecular weight compound of crystallised molecules that encourages ionisation of the analyte of interest. The inaccurate choice of matrix not only can hinder the ionisation of the analytes of interest (Kiss and Hopfgartner, 2016) but also it can overlap signals in low mass m/z regions (<500 Da) with the same nominal mass of the target molecules (Shanta *et al.*, 2012) or introduce heterogenic distribution of desired molecules across the sample (Liang *et al.*, 2013).

The most common matrices used for the analysis of low molecular weight compounds are organic molecules that show strong ultraviolet absorbance. Depending on their chemical structure, there are two main groups of matrices that are used in MALDI-MS studies: derivatives of aromatic carbonyl compounds *e.g.* 2,5-dihydroxybenzoic acid (DHB), 2,4,6-trihydroxyacetophenone (THAP) or 9-aminoacridine (9AA) and derivatives of cinnamic acid *e.g.* α -cyano-4-hydroxycinnamic acid (CHCA) (Krause, Stoeckli and Schlunegger, 1996). Due to some drawbacks of employing crystalline matrices for small molecule analysis such as poor sample-to-sample and shot-to-shot reproducibility and uneven distribution of the analyte within a sample spot, liquid matrices such as glycerol or 3-nitrobenzyl alcohol have been introduced as alternatives (Cornett, Duncan and Amster, 1993). Despite their advantages such as higher signal reproducibility and lower fragmentation of analytes (Fukuyama, 2011), the applicability of liquid matrices is still linked with low resolution of mass spectra, poor ionisation efficiency and potential instrument contamination (Karbach *et al.* 1998, Cohen *et al.* 2007).

Other important factors in MALDI sample preparation methodology are the matrix concentration and its solvent. The main role of MALDI matrix is to provide effective ionisation of the analyte with a minimum matrix background. Therefore, a matrix concentration has to be high enough to assist good mobility of the analyte but simultaneously, the matrix solution cannot be too concentrated to avoid sample suppression by matrix ions (Cohen, Go and Siuzdak, 2007). Also, the type and quality of the matrix solvent can affect MALDI analysis in different ways. For example, inappropriate solvent can influence on analyte incorporation in matrix crystals, resulting in heterogeneous sample spot and causing poor shot-to-shot reproducibility (Hillenkamp *et al.* 2009, Fujita *et al.* 2010). A similar result is observed when the solvent does not allow for good solubility of either the matrix or the analyte. Different polarities of solvent and the analyte may result in precipitation of the analyte making the sample spot uneven and the analysis highly complicated (Mechref & Novotny 1998, Zenobi & Knochenmuss 1998). Matrix and analyte solvents are therefore matched according to their solubility, yet, an additional consideration should be made with regard to the solvent evaporation process (Sarkar *et al.*, 2009). Quicker crystallisation creates smaller crystals and a more uniform distribution of the analyte into the matrix crystals. To make the crystallisation process faster, the matrix is often formulated in a volatile solvent and/or the target is placed in a vacuum and/or elevated temperature (Nicola *et al.*, 1995).

There is no single MALDI sample preparation protocol that would suit to all challenges and type of analytes in MALDI technique. Apart from a suitable matrix, its concentration and solvent, attention to the way of sample deposition onto MALDI plate has to be also considered (Korte and Lee, 2014).

A standard and most common method of sample deposition is “dry droplet” or “mixed” deposition. Here, the matrix and the analyte are first dissolved, mixed together in suitable matrix-to-analyte ratio and spotted on a MALDI plate (**Figure 1**). Yet, over the years, a number of alternative methods of sample deposition to “dry droplet” technique have been developed. These variations are often related to personal preferences of MALDI users as well as they relate to the type of the analyte being analysed or the matrix being used (Cohen, Go and Siuzdak, 2007). The most common alternatives to “mixed” method are “overlay” and “underlay” methods of sample spotting. In “overlay” method the matrix is applied on top of the analyte previously spotted on a MALDI plate, whereas in “underlay” method it is the matrix that is spotted first on a MALDI plate (AlMasoud *et al.* 2014, Helmelt *et al.* 2014). **Figure 1** illustrates the most common methods of sample deposition used for MALDI experiments.

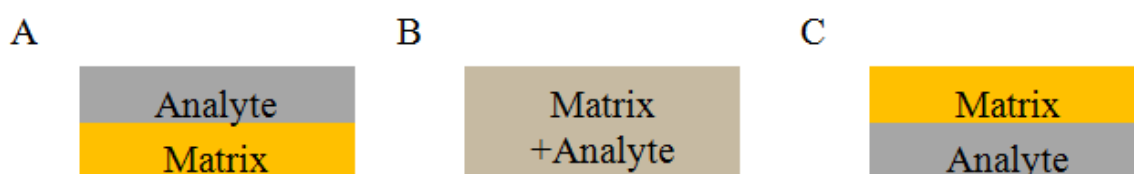


Figure 1: The most common ways of sample deposition used in MALDI experiments: underlay (A), mixed (B) and overlay (C) deposition.

A crucial consideration when choosing a matrix is the solvent to use as well as a method of sample deposition is the ability of creating a homogenous crystallisation with small crystal size, especially for quantitative and imaging experiments (Kiss & Hopfgartner 2016, Römpp & Spengler 2013, Duncan *et al.* 2008). Crystallisation process can be equated to the process of purification, where impurities (in this case analyte) are being removed from the slowly forming matrix crystal, resulting in heterogeneous sample spot (Hensel *et al.* 1997, Sarkar *et al.* 2009). While applying matrix–analyte solution onto a MALDI target, analyte is absorbed into the matrix crystals to a varying grade (Cohen and Chait, 1996), resulting in forming so called “hot/sweet spots” that are certain locations of crystals where the signal of the analyte is stronger than in other parts of the sample. Hence, inadequate sample preparation and thus, crystallisation process often leads to poor sample reproducibility and sample – to – sample repeatability. However, development of MALDI instruments using higher laser frequencies along with the automation of the entire analysis have improved concerns of heterogeneity of the sample to some extent (Cohen, Go and Siuzdak, 2007).

A good choice of MALDI sample preparation method is crucial to obtain satisfying MS profiles of samples of interest with acceptable sensitivity and mass resolution. Metabolic profiling of biological samples is based on the analysis of extracellular and intracellular metabolites and their mutual relations that can provide insights to the behaviour of cells and their physiological processes (Dorries and Lalk, 2013). The presence of chemical overlapping at the same or similar mass as the analyte of interest in cellular samples can be a limiting factor in detecting analytes with good sensitivity, mass resolution and mass accuracy (Trim & Snel 2016, Edwards & Kennedy 2005). Therefore, an accurate and reliable calibration of the instrument is important for exact mass determination of metabolites of interest (Weidmann *et al.*, 2013) as well as it is crucial for quantitative studies, where calibration standards are used to spike the sample (Trim & Snel 2016, Duncan *et al.* 1993). A selection of a mixture of standard compounds is critical in achieving accurate analyte recognition and quantification. The use of a calibration mixture

can minimise both shot-to-shot and sample-to sample variability (Muddiman *et al.* 1994). However, standard compounds have a number of requirements *i.e.* their physico-chemical properties should be similar to those of the analytes of interest in order to provide similar ionisation behaviour. Additionally, standards must not overlap with a mass of interest to ensure a good mass separation between analyte and standard as well as clear measurement of peak intensity (Cohen, Go and Siuzdak, 2007).

The goal of this chapter was to optimise the sample preparation procedure for a detection of small molecules derived from extracellular and intracellular samples of CHO cell line using MALDI-MS profiling in positive and negative ionisation modes.

5.2. Materials

MALDI matrices, solvents *i.e.* acetonitrile (ACN), ethanol (EtOH), methanol (MeOH) and water (liquid chromatography grade) and all compounds used in this section were purchased from Sigma–Aldrich. Fresh cell growth medium was purchased from BD Biosciences (Oxford, UK). Medium sample from cells was removed on day 1 of cell culture according to the procedure described previously (**Chapter 2**).

5.3. Sample preparation

CARE: 9AA has mutagenic effects; TFA causes severe burns, harmful by inhalation; work performed in a fume cupboard.

For section 5.6.1.: DHB, CHCA, THAP, 9AA and 1,8-bis(dimethylamino)naphthalene (DMAN) matrix solutions and mixture of arginine (Arg) and histidine (His) as well as tryptophan (Trp) and adenosine 5'-triphosphate (ATP) disodium salt hydrate were made fresh. Stock solutions of Arg and His were made in deionised water at concentration of 200 mM each and in a matrix-to-analyte (M:A) ratio of 1:1 (v/v). Stock solutions of Trp and ATP were made in 90% EtOH and water at concentration of 40 mM and 30 mM, respectively and in M:A ratio of 2:3 (v/v), respectively. 1 µL of each sample was spotted on a MALDI plate using mixed deposition. Each sample spot was measured 3 times ($n=3$) and the average and standard deviation were calculated. Experiments were performed using MALDI–ToF–MS Ultraflex II ToF/ToF in positive and negative ionisation modes.

For section 5.6.2.: The analytes were prepared and mixed in a M:A ratio of 1:1 (v/v) with freshly made DHB and 9AA matrices and 1 µL of a sample was spotted on a MALDI plate. Each compound was examined in two repeats ($n=2$) and the mean of S/N ratio was calculated. Experiments were performed using MALDI-ToF-MS Ultraflex II ToF/ToF in positive and negative ionisation modes.

For section 5.6.3.: DHB, THAP and CHCA were prepared in a solution of ACN/water 1:1 (v/v) with 0.1% trifluoroacetic acid (TFA) at a concentration of 10 mg/mL. 9AA matrix was prepared in a concentration of 5 mg/mL and in a solution of five solvents: 70% and 100% ACN, 70% and 100% MeOH and 70% EtOH. The calibration cocktail and PEG600 sulfate used from previous section (**Section 5.6.2.**) (called ‘analyte’) were used for positive and negative ionisation modes, respectively. Analytes and matrices were mixed and samples were deposited on a steel MALDI plate in three different ways: i) overlay deposition: 1 µL of metabolite mixture was first deposited on the plate and left to dry

followed by deposition of 1 μ L of matrix; ii) mixed deposition: the metabolite mixture and matrices were mixed 1:1 (v/v) and 1 μ L of a sample was deposited on the plate and left to dry; iii) underlay deposition: 1 μ L of a matrix was first deposited on the plate and left to dry followed by deposition of 1 μ L of metabolite mixture. Experiments were performed using a MALDI Synapt G2-Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK) in positive and negative ionisation modes. Each sample was run four times ($n=4$, independent repeats).

For section 5.6.4.: Experiments were carried out including fresh (non-used) cell growth medium sample as a control and cell growth medium sample removed from a flask with cell culture on day 1. DHB and THAP matrices were freshly prepared at a concentration of 10 mg/mL and dissolved in a solution of ACN/water 1:1 (v/v) with 0.1% TFA. Two ways of sample deposition were applied: mixed deposition and overlay deposition (M:A ratio 1:1). 9AA matrix was prepared in a concentration of 5 mg/mL and in a solution of three solvents: 70% ACN, 70% MeOH and 70% EtOH. Matrix solutions and medium samples were mixed (M:A 1:1) and spotted onto a MALDI plate. Each sample was run four times ($n=4$, independent repeats) using MALDI Synapt G2-Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK) in positive and negative ion modes. The average of all repeats was taken into account for mass resolution and intensity calculations. Instrument was calibrated to the calibration cocktail or to PEG600 sulfate for positive or negative ion modes, respectively.

5.4. Instrumentation

Experiments were performed using MALDI-ToF-MS Ultraflex II ToF/ToF mass spectrometer (Bruker Daltonics, Billerica, MA, USA). The instrument was operated in reflectron positive and negative modes with laser repetition rate of 100 Hz and with a mass range of 0–1000 Da. A number of 600 laser shots within a whole sample spot was collected from each sample. Laser power was adjusted to each sample while collecting spectra, typically in a range of 40–100% of total laser power. Further experiments were performed using MALDI Synapt G2-Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK). The spectra were acquired in reflectron “sensitivity mode” for both positive and negative ionisation modes in the mass range of 20 to 1000 Da. The acquisition time was 30 s for each spot, composed of 30 laser shots at rate of 1 scan per s using default pattern of sample collection. Laser diameter was fixed to 60 μ m and laser power was previously optimised (**Supplementary Information, Figure 1**)

and was set to 250 and 200 for positive and negative ionisation modes, respectively for further investigations, unless stated otherwise.

5.5. Data analysis

MALDI data were analysed in Microsoft Excel ver. 2007 and MATLAB ver. R2012a software. MS spectra, PCA scores plots and loading plots were generated using sum normalisation and m/z 0.2 binning.

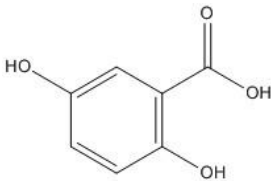
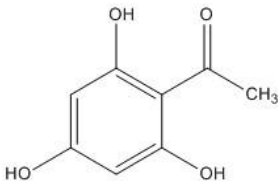
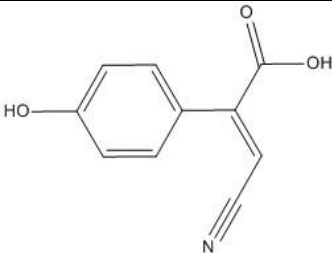
5.6. Results and discussion

5.6.1. Choice of matrix and solvent conditions for MALDI metabolite detection

5.6.1.1. Positive ionisation mode

Three most often used MALDI matrices for metabolomics studies: DHB, THAP and CHCA (Trim & Snel 2016, Kiss & Hopfgartner 2016) were verified to check their suitability for metabolite profiling of transfected CHO cells (**Table 1**). The matrices were prepared in three different concentrations (5, 10, 20 mg/mL) and in three different solvents *i.e.* ACN/water 1:1 (v/v) with 0.1% of TFA, 70% MeOH and 70% EtOH in order to optimise the most effective conditions of sample preparation for metabolite detection. Following this, a mixture of arginine and histidine (1:1, v/v) was added to matrices solutions to check the influence of matrix concentration and matrix solvent on signal-to-noise ratio (S/N), intensity and mass resolution (Full Width of the peak at Half of its Maximum height – FWHM) of mentioned compounds.

Table 1: Chemical structures of DHB, THAP and CHCA with their most common applications for MALDI-MS analysis. Adopted from (<http://www.sigmaaldrich.com/>).

Matrix	Structure	Application	References
DHB (MW 154 g/mol)		Lipids, peptides, amino acids, sugars, vitamins	(AlMasoud <i>et al.</i> 2016, Bae <i>et al.</i> 2014, Korte & Lee 2014)
THAP (MW 186 g/mol)		Glycans, polysaccharides, lipids	(Sić <i>et al.</i> 2016, Fujita <i>et al.</i> 2010, Zhang <i>et al.</i> 2014)
CHCA (MW 189 g/mol)		Peptides, amino acids, sugars, vitamins	(Bae <i>et al.</i> 2014, Korte & Lee 2014)

Matrices chosen for this experiment have been reported to be well suited for profiling of small molecules in positive ionisation mode (Zhang *et al.* 2014, Lou *et al.* 2009). Depending on the application however, the selection of the most relevant matrix for analytes of interest is highly important in order to collect good quality mass spectra. Eibisch *et al.* have investigated the changes in phospholipids detected from lipid extracts of human erythrocytes using DHB in both positive and negative ionisation modes (Eibisch and Schiller, 2011). Similar application of DHB was shown in profiling of lipids, where researchers were aiming to classify different strains of bacteria based on bacterial extracts (AlMasoud *et al.*, 2016). Additionally, another example of DHB and CHCA application was presented, where spectral reproducibility and quantification of peptides have been studied (Bae *et al.*, 2014). The effectiveness of DHB and CHCA for examination of not only peptides or lipids, but also different classes of metabolites such as amino acids, sugars and vitamins has been shown in a study of Korte and Lee using soy extracts (Korte and Lee, 2014). Also, a good analysis of polysaccharides (Fujita *et al.*, 2010) and lipids (Zhang, Smith and Purves, 2014) with a usage of THAP was applied, where successful measurements of the analyte of interest without any matrix suppression were presented. Targeted profiling however, differs to a great extent from untargeted metabolomics, where all the measurable analytes in a sample including targeted metabolomics, assigned metabolites and all chemical unknowns are taken into consideration (Roberts *et al.*, 2013). Hence, the easiest and the most reliable way to investigate the suitability of different matrices is to analyse them with regard to intensity and signal-to-noise ratio of the analyte of interest as well as mass resolution and overall quality of obtained mass spectra (Cohen, Go and Siuzdak, 2007).

With respect to the above, DHB, CHCA and THAP matrices were analysed with Arg and His mixture. The choice of the analytes for initial experiments was based on their good ionisation properties, low molecular weight fitting in a mass range of interest (< 1000 Da) as well as their wide usage as standards for mass spectrometry investigations (Alterman *et al.* 2004, Zenobi & Knochennuss 1998). Also, Arg and His are basic amino acids that take part in cell metabolism and therefore, this class of compounds were to be expected in fingerprint and footprint samples of CHO cells (Mohmad-Saberi *et al.* 2013, Sellick *et al.* 2011).

The initial study was performed with three different matrix concentrations and three different matrix solvents. A concentration of the matrix varies according to a matrix type and character of the analysed sample. For matrices chosen for the experiments *i.e.* DHB, CHCA and THAP a wide range of concentrations is applied for profiling of small

molecules: from 5 mg/mL up to 30 mg/mL (Zhang *et al.* 2014, Alterman *et al.* 2004, Cao *et al.* 2011). Yet, the most common matrix concentration used for analysis of low molecular weight compounds stands at 10 mg/mL (Kim *et al.* 2011, Griffiths & Bunch 2012). Usually, lower matrix concentration range (< 20mg/mL) is applied for metabolic profiling *e.g.* footprint and fingerprint samples, whereas higher matrix concentrations are used for MALDI imaging of cells and tissues due to the fragile nature of such samples (Becker *et al.* 2014, Phan *et al.* 2016, Schober *et al.* 2012). Hence, in order to evaluate a suitable matrix preparation for metabolite profiling of CHO cells, 5, 10 and 20 mg/mL of DHB, CHCA and THAP concentrations were investigated.

Due to a presence of a vast number of different solvents and their combinations for matrices and analytes of interest being applied to MALDI-MS experiments (Lou *et al.* 2009, Hillenkamp *et al.* 2009, Eibisch & Schiller 2011, Edwards & Kennedy 2005, Korte & Lee 2014, Calvano *et al.* 2013), this study involved three, the most common solvents used for measurements of small molecules.

Firstly, ACN/H₂O 1:1 (v/v) + 0.1% TFA, widely applied for analysis of metabolites was chosen (Paz *et al.* 2011, Fujita *et al.* 2010, Goodwin *et al.* 2012). TFA is typically used in MALDI-MS as a promoter of protonated ($[M+H]^+$) species, thus improving the intensity of the analyte (Schiller *et al.*, 1999). Studies have shown improved S/N ratio of analytes when TFA was applied (Schiller *et al.*, 2003). Also, the TFA percentage in this study was chosen based on the research, where different concentrations of this acid were investigated on bacterial samples (Helmel, Marchetti-Deschmann and Allmaier, 2014). It was shown that collected mass spectra did not change significantly in a range of 0.1% to 2.5% of TFA used for experiments. Since it has been presented there are no critical changes in mass spectra between different concentrations of TFA, the lowest (*i.e.* 0.1%) concentration of TFA was applied for further experiments. Apart from protonated ions for positive ionisation mode and deprotonated ions ($[M-H]^-$) in negative ionisation mode, sodium adducts ($[M+Na]^+$), potassium adducts ($[M+K]^+$), ammonium adducts ($[M+NH_4]^+$) and chlorine adducts ($[M+Cl]^-$) are often observed - depending on the ionisation mode - during analysis of biological samples (Zaima *et al.* 2010, AlMasoud *et al.* 2016). This is due to a presence of salts in organic samples. Thanks to the existence of adducts in MALDI spectra, it was possible to investigate a number of additives such as previously mentioned TFA, ammonium acetate (Stübiger *et al.*, 2010) or potassium chloride (Stubiger *et al.* 2009, Garrett *et al.* 2007) that are additionally included to a matrix solvent in order to promote formation of a particular adduct or act as a 'cationising agent' (Suarez *et al.*, 2011). However, addition of additives is often used for targeted studies such as lipid profiling,

rather than untargeted metabolomics (Griffiths and Bunch, 2012). Hence, 0.1% TFA was selected to be implemented in ACN/water mixture in order to promote protonation across the whole content of our sample.

The remaining two chosen matrix solvents were 70% EtOH, applied for analysis *e.g.* peptides and sugars (Hillenkamp *et al.* 2009, Sugiura *et al.* 2014) and 70% MeOH, oftentimes used for metabolic analysis (Edwards and Kennedy, 2005) *inter alia*, drugs (Chakrabarty *et al.*, 2015). Additionally, a choice of MeOH for this study was based on the fact that metabolism in CHO cells was quenched by methanol, accordingly to extraction protocols for fingerprint samples of CHO cells (Kronthaler *et al.* 2012, Sellick *et al.* 2011). The choice of these three solvents was also made according to the comparability of the solvents normally used for preparation of calibrants *i.e.* water being used to dissolve peptide standards (Dashtiev *et al.*, 2007), a mixture of chloroform/MeOH being used for preparation of lipid standards (AlMasoud *et al.*, 2016), 50-70% methanol being used as nucleotide solvent (Edwards and Kennedy, 2005), whereas ethanol can be applied as a solvent for fatty acids and peptides samples (Shroff and Svatos, 2009). Moreover, all sample spots were spotted by dried-droplet method that is the easiest and the quickest method for MALDI sample preparation (Sarkar *et al.* 2009, Paz *et al.* 2011, Franz Hillenkamp & Karas 2007). **Table 2** presents a summary of matrices, their concentrations and type of solvents used for the initial optimisation of this study.

The data were collected and analysed for presence of arginine and histidine peaks in each sample. There were three peaks: arginine peak [Arg+H]⁺ at *m/z* 175.120 and histidine peaks [His+H]⁺ at *m/z* 156.077 and [His+K]⁺ at *m/z* 194.033 with S/N ratio higher than 6 taken into consideration. Out of 27 methods of sample preparation (3 *matrices* X 3 *concentrations* X 3 *solvents*) presented in **Table 2**, 3 of them failed with detection on histidine peak either [His+H]⁺ or [His+K]⁺ and therefore, were removed from further analysis. In the remain 24 methods of sample preparation all 3 peaks [Arg+H]⁺, [His+H]⁺ and [His+K]⁺ could be seen.

Table 2: 27 different combinations of MALDI positive ion sample preparation with presence (+) or absence (-) of ions of interest. Thus, table summarises matrix, its concentration and solvent used in the study. Detection when S/N>6.

Matrix	Concentration	Solvent	[Arg+H] ⁺	[His+H] ⁺	[His+K] ⁺
DHB	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	-	+
		70% EtOH	+	+	+
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	+	+
		70% MeOH	+	+	+
		70% EtOH	+	-	-
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	+	+
		70% MeOH	+	-	-
		70% EtOH	+	+	+
CHCA	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	-	+
		70% EtOH	+	-	+
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	+	+
		70% MeOH	+	-	+
		70% EtOH	+	-	+
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	-	+
		70% EtOH	+	+	-
THAP	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	+	+
		70% EtOH	+	-	+
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	-	+
		70% EtOH	+	-	-
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA	+	-	+
		70% MeOH	+	-	+
		70% EtOH	+	-	+

The following step included a comparison of peak identification, S/N ratio, mass resolution and intensity of [Arg+H]⁺, [His+H]⁺ and [His+K]⁺ ions. The comparison of these 24

methods was made by 3 different solvents used for sample preparation and by 3 different matrix concentrations (**Supplementary Information, Figure 2-7**). A summary of these tested methods based on S/N ratio and intensity of target ions as well as their mass resolution is presented in **Table 3, 4, 5**, respectively. These data show that the S/N ratio and intensity of $[\text{Arg}+\text{H}]^+$ ion are higher than histidine peaks in all screened methods. In terms of both histidine ions, $[\text{His}+\text{K}]^+$ ion occurs more often to protonated form of this amino acid and generally appears with greater intensity and S/N ratio than $[\text{His}+\text{H}]^+$. Additionally, a relationship between matrix concentration, matrix solvent, S/N ratio and intensity of arginine and histidine ions *i.e.* $[\text{Arg}+\text{H}]^+$, $[\text{His}+\text{H}]^+$ and $[\text{His}+\text{K}]^+$ can be noticed: both parameters (S/N ratio and intensity) seem to be the highest for these ions when a concentration of 10 mg/mL and ACN/H₂O with 0.1% TFA as a solvent are applied (**Table 3 and 4**). Mass resolution of all three ions appears to be similar and is between a range of 2010 ± 101 to 4437 ± 487 for arginine and between 3131 ± 146 to 5474 ± 426 for histidine ions across all screened methods of sample preparation. Opposite to intensity and S/N ratio comparison, a correlation between mass resolution and matrix concentration and its solvent cannot be found. However, a general trend shows that mass resolution of mentioned ions seems to be better in solvents such as EtOH or MeOH rather than ACN/H₂O with 0.1% TFA (**Table 5**).

Table 3: Comparison of S/N ratio of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of S/N ratio (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Arg+H] ⁺	[His+H] ⁺	[His+K] ⁺	Score
DHB	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	20		10	0
		70% MeOH	63		10	0
		70% EtOH	1150	13	24	3
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	595	13	136	3
		70% MeOH	175	9	10	1
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	593	12	10	2
		70% EtOH	262	7	9	1
CHCA	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	43		19	0
		70% MeOH	85		24	1
		70% EtOH	133		15	0
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	1195	9	77	3
		70% MeOH	155		17	0
		70% EtOH	958		8	1
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	19		7	0
		70% MeOH	97		22	1
		70% EtOH	11369	7		2
THAP	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	49		26	1
		70% MeOH	149	7	11	1
		70% EtOH	205		18	0
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	197		14	0
		70% MeOH	91		11	0
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	199		9	0
		70% MeOH	209		16	0
		70% EtOH	138		11	0

*scores-the summary of the highest values for a given method

Table 4: Comparison of the intensity [a.u.] of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of the intensity (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Arg+H] ⁺	[His+H] ⁺	[His+K] ⁺	Score
DHB	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	110		55	0
		70% MeOH	355		55	0
		70% EtOH	6634	78	140	3
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3646	78	861	3
		70% MeOH	1020	51	67	1
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3534	72	54	2
		70% EtOH	1438	41	52	0
CHCA	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	242		99	0
		70% MeOH	475		129	0
		70% EtOH	779		85	0
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	7720	63	537	3
		70% MeOH	998		121	0
		70% EtOH	5480		45	1
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	159		42	0
		70% MeOH	699		170	1
		70% EtOH	71518	44		2
THAP	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	285		155	1
		70% MeOH	821	40	63	0
		70% EtOH	1385		144	1
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	1226		91	0
		70% MeOH	523		67	0
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	1094		52	0
		70% MeOH	1356		107	0
		70% EtOH	832		70	0

*scores-the summary of the highest values for a given method

Table 5: Comparison of mass resolution of Arg and His ions according to the matrix used, its concentration and solvent in MALDI positive ion mode. The highest values of mass resolution (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Arg+H] ⁺	[His+H] ⁺	[His+K] ⁺	Score
DHB	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3237		3724	0
		70% MeOH	4283		3341	1
		70% EtOH	3863	4725	4255	2
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3936	3763	3131	1
		70% MeOH	4344	4516	3713	2
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3306	3599	4739	2
		70% EtOH	4242	4707	4633	3
CHCA	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3716		3800	0
		70% MeOH	4370		3952	1
		70% EtOH	3892		4135	0
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3180	3205	3234	0
		70% MeOH	3934		3301	0
		70% EtOH	3206		5362	1
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	4437		3238	1
		70% MeOH	3946		3663	0
		70% EtOH	2010	3598		1
THAP	5 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3360		3582	0
		70% MeOH	3772	3186	3694	0
		70% EtOH	4105		3839	0
	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3808		3669	0
		70% MeOH	4232		5474	2
	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1%	3779		5138	1
		70% MeOH	3287		3680	0
		70% EtOH	3693		3335	0

*scores-the summary of the highest values for a given method

Based on the comparison of S/N ratio, intensity and mass resolution of Arg and His ions, the highest 6 values for each ion, within each of the discussed properties were chosen (highlighted in orange). This was then followed by evaluation of the best methods of sample preparation ('score' column highlighted in green in **Table 3, 4 and 5**). Consequently, there were 8 preferred methods of matrix preparation with both arginine and histidine ions detected taken into account (**Table 6**).

Table 6: Summary of the eight top methods of MALDI positive ion matrix preparation, including matrix, its concentration and solvent. Methods were selected based on the analytes detection, their S/N ratio, intensity and mass resolution. Methods listed by matrix, its concentration (from low to high) and type of solvent.

Matrix	Concentration	Solvent
DHB	5 mg/mL	70% EtOH
DHB	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA
DHB	10 mg/mL	70% MeOH
DHB	20 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA
DHB	20 mg/mL	70% EtOH
CHCA	10 mg/mL	ACN/H ₂ O 1:1 (v/v) + 0.1% TFA
CHCA	20 mg/mL	70% EtOH
THAP	10 mg/mL	70% MeOH

In order to restrict the investigation of the best matrix preparation method, MS profiles of DHB, CHCA and THAP matrices were analysed (**Figure 2**). To achieve a good quality and reproducibility of the data an appropriate MALDI matrix is characterised by reduced background (*i.e.* matrix signals ought to be very small) in order to avoid matrix-analyte ions interference and hence, allowing the identification of analyte peaks. Also, an optimal matrix should produce only a single adducts of the target molecule and should not tend to form matrix clusters as such clusters can cause complications in analyte identification and analysis of the data (Shanta *et al.* 2012, Shariatgorji *et al.* 2012).

As seen in **Figure 2**, CHCA gives the most intense peaks (10^4 a.u) when compared to both DHB and THAP (< 400 a.u.), independently on the solvent used. Also, there are much more matrix peaks from CHCA than from DHB or THAP. CHCA peaks that show the highest abundance in ACN/water with TFA at 10 mg/mL are at m/z 172.041 [CHCA-H₂O+H]⁺ and at m/z 379.094 [2CHCA+H]⁺, whereas the most intense matrix peaks for CHCA in 70% EtOH at 20 mg/mL are observed at m/z 172.041, m/z 212.024 [CHCA+Na]⁺

and at m/z 234.010 $[\text{CHCA-H+Na}]^+$. Unlike CHCA, DHB and THAP appear to be superior matrices for analysis of small molecules because of their low intensity and small number of peaks. There are three DHB peaks that are the most abundant within the spectra and are seen in all screened methods: m/z 137.023 $[\text{DHB+H-H}_2\text{O}]^+$, m/z 155.024 $[\text{DHB+H}]^+$ and m/z 177.012 $[\text{DHB+Na}]^+$. However, THAP profile presents the lowest intensity when compared to remain two matrices with peaks at m/z 169.051 $[\text{THAP+H}]^+$ and m/z 381.179 $[\text{2THAP+2Na-H}]^+$.

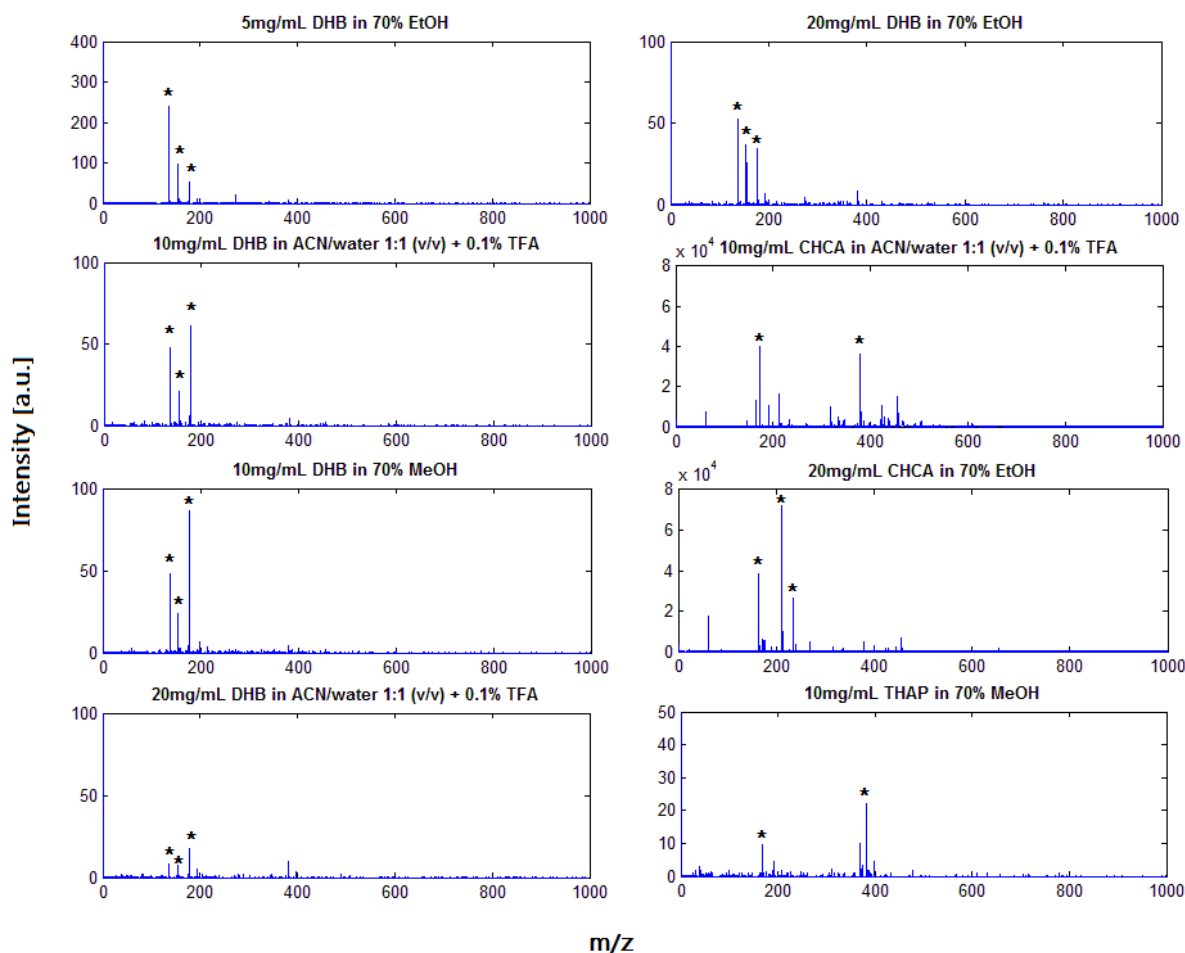


Figure 2: MALDI positive ion mode averaged spectra of pure (with no sample) matrices of 8 top methods of sample preparation; * - high intense matrix peaks.

According to **Table 6**, due to mainly DHB matrix presence within 8 combinations shown, this matrix seems to be the most suitable for MALDI analysis in both 10 mg/mL and 20 mg/mL concentrations. However, after the analysis of spectra of pure DHB prepared in different solvents, it was confirmed that a mixture of ACN/water with 0.1% TFA gives less matrix noise than 70% EtOH or 70% MeOH and therefore, this solvent system appears as the most appropriate one. Additionally, based on S/N ratio and intensity comparison of Arg

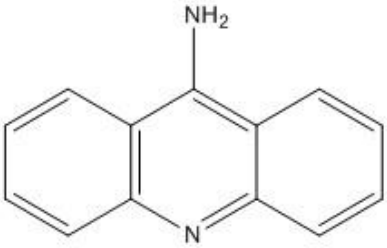
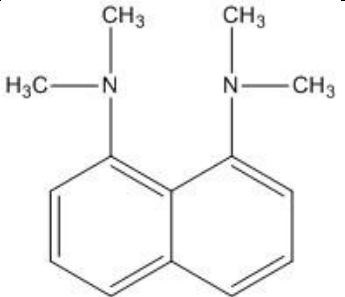
and His ions the concentration of 10 mg/mL seems to be better as it gives higher analyte signals.

5.6.1.2. Negative ionisation mode

Ionisation of the analyte depends on its chemical structure. Several studies show that oligonucleotides, phospholipids and steroids are more likely to create negative ions due to the presence of groups such as $[-OH]$, $[=O]$, $[PO_4]^{3-}$ or $[-COOH]$ (Vermillion-Salsbury & Hercules 2002, Amantonico *et al.* 2008, Calvano *et al.* 2013, Chen *et al.* 2012). Negative ionisation mode is also used to enrich and supplement data of positive ionisation mode with more sample information, hence allowing direct correlation of the collected data (Ellis *et al.*, 2016). Due to this and because of the fact that nucleotides, phospholipids and many other compounds were expected to be in CHO cell extracts (Dietmair *et al.*, 2012), dual polarity MALDI optimisation of sample preparation was performed.

Initial optimisation of matrix preparation for negative ionisation mode included two matrices: 9AA and DMAN (**Table 7**). Matrices were prepared in three different concentrations (5, 10, 20 mg/mL) and in five solvents: 70% and 100% ACN, 70% and 100% MeOH and 70% EtOH. Following matrix preparation, a mixture of Trp and ATP was added to each sample in order to check the influence of different sample preparation on mentioned analytes.

Table 7: Chemical structures of 9AA and DMAN with their most common applications for MALDI-MS analysis. Adopted from (<http://www.sigmaaldrich.com/>).

Matrix	Structure	Application	References
9AA (MW 194 g/mol)		Lipids, oligonucleotides, metabolites	(Sun <i>et al.</i> 2008, Vermillion-Salsbury & Hercules 2002, Vaidyanathan & Goodacre 2007)
DMAN (MW 214 g/mol)		Fatty acids, phosphopeptides, metabolites	(Shroff & Svatoš 2009, Zhang & Yao 2012, Ye <i>et al.</i> 2013)

Several organic matrices such as 9AA, DMAN (Vermillion-Salsbury & Hercules 2002, Shroff & Svatoš 2009), N-(1-naphthyl) ethylenediamine dinitrate (NEDN) (Chen *et al.*, 2012), 1,5-diaminonaphthalene (DAN) (Dong *et al.*, 2013) as well as binary matrices for example CHCA/9AA (Guo and He, 2007) have been applied for profiling of metabolites in negative ionisation mode. Although low matrix background can be attained with these matrices, several issues such as low ionisation efficiency or lack of uniform surface morphology are still to be overcome (Cao *et al.* 2011, Thomas *et al.* 2012). Therefore, two matrices 9AA and DMAN (**Table 7**), commonly used for metabolite profiling in negative ion mode, were compared in three concentrations of 5, 10 and 20 mg/mL in different solvents: 70% and 100% ACN, 70% and 100% MeOH and 70% EtOH. Similarly to positive ionisation mode, concentrations and solvents for matrices were chosen based on the most frequent matrix preparations in regard to profiling of small molecules, giving high quality data (Korte & Lee 2014, Dong *et al.* 2013, Calvano *et al.* 2011, Zhang & Yao 2012). As previously section, a selection of matrix concentration and solvent are crucial for its good solubility, ionisation efficiency for the analyte, creation of homogeneous film and low matrix background (Hillenkamp *et al.* 2009, Fujita *et al.* 2010). Studies have reported a good suitability of MeOH as matrix solvent for 15 mg/mL 9AA in profiling of oligonucleotides (Vermillion-Salsbury and Hercules, 2002) and 10 mg/mL 9AA and DMAN in lipid (Calvano *et al.*, 2013) and phosphopeptide fingerprinting (Zhang and Yao, 2012). On the other hand, others have applied EtOH as a DMAN solvent in fatty acid analysis (Shroff and Svatos, 2009) and 50% ACN for 10 mg/mL 9AA solution in small molecule studies (Guo and He, 2007). With respect to the above, 9AA and DMAN were studied with ATP and Trp mixture. These analytes are widely used as standards for mass spectrometry analysis and have good ionisation properties (Blatherwick *et al.* 2013, Vaidyanathan & Goodacre 2007).

The initial analysis of matrix preparation methods included a comparison of detection ($S/N > 6$) of ions of interest *i.e.* $[\text{Trp-H}]^-$ and $[\text{ATP-2Na+H}]^-$ (**Table 8**). Out of 30 possibilities (2 matrices \times 3 concentrations \times 5 solvents) there were 15 methods that failed with a detection of one of the ions of interest, either $[\text{Trp-H}]^-$ or $[\text{ATP-2Na+H}]^-$ and therefore, these matrix-solvent combinations were excluded from further investigations. In the remaining 15 methods both Trp and ATP ions were detected.

Table 8: 30 different combinations of MALDI negative ion sample preparation with presence (+) or absence (-) of ions of interest. Thus, table also summarises matrix, its concentration and solvent used in the study. Detection when S/N>6.

Matrix	Concentration	Solvent	[Trp-H] ⁻	[ATP-2Na+H] ⁻
9AA	5 mg/mL	70% ACN	+	+
		100% ACN	+	+
		70% MeOH	+	+
		100% MeOH	+	+
		70% EtOH	+	+
	10 mg/mL	70% ACN	+	+
		100% ACN	-	+
		70% MeOH	+	+
		100% MeOH	+	+
		70% EtOH	+	+
	20 mg/mL	70% ACN	-	+
		100% ACN	+	+
		70% MeOH	+	+
		100% MeOH	+	+
		70% EtOH	+	+
DMAN	5 mg/mL	70% ACN	+	-
		100% ACN	+	+
		70% MeOH	+	-
		100% MeOH	+	-
		70% EtOH	+	-
	10 mg/mL	70% ACN	+	-
		100% ACN	+	+
		70% MeOH	+	-
		100% MeOH	+	-
		70% EtOH	+	-
	20 mg/mL	70% ACN	+	-
		100% ACN	+	-
		70% MeOH	+	-
		100% MeOH	+	-
		70% EtOH	+	-

Subsequently, S/N ratio, intensity and mass resolution of those ions were compared. Observation of these 15 methods was made by 3 different matrix concentrations and 3 different matrix solvents used for sample preparation (**Supplementary Information, Figure 8-13**). A summary of screened methods based on S/N ratio, intensity and mass resolution of [Trp-H]⁻ and [ATP-2Na+H]⁻ ions is presented in **Table 9, 10 and 11**.

Table 9: Comparison of S/N ratio of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of S/N ratio (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Trp-H] ⁻	[ATP-2Na+H] ⁻	Scores
9AA	5 mg/mL	70% ACN	302	823	2
		100% ACN	97	64	0
		70% MeOH	1350	661	2
		100% MeOH	180	134	1
		70% EtOH	256	104	2
	10 mg/mL	70% ACN	182	115	2
		70% MeOH	92	43	0
		100% MeOH	88	17	0
		70% EtOH	134	43	0
	20 mg/mL	100% ACN	76	12	0
		70% MeOH	17	2992	1
		100% MeOH	59	7	0
		70% EtOH	57	26	0
DMAN	5 mg/mL	100% ACN	1232	9	1
	10 mg/mL	100% ACN	2647	6	1

*scores-the summary of the highest values for a given method

Table 10: A comparison of the intensity [a.u.] of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of intensity (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Trp-H] ⁻	[ATP-2Na+H] ⁻	Scores
9AA	5 mg/mL	70% ACN	3091	5871	2
		100% ACN	681	365	0
		70% MeOH	10503	4582	2
		100% MeOH	1670	650	2
		70% EtOH	1555	533	2
	10 mg/mL	70% ACN	1162	592	1
		70% MeOH	482	177	0
		100% MeOH	514	84	0
		70% EtOH	904	222	0
	20 mg/mL	100% ACN	376	43	0
		70% MeOH	291	19696	1
		100% MeOH	391	33	0
		70% EtOH	292	106	0
DMAN	5 mg/mL	100% ACN	11659	39	1
	10 mg/mL	100% ACN	19909	28	1

*scores-the summary of the highest values for a given method

Table 11: A comparison of mass resolution of Trp and ATP ions according to the matrix used, its concentration and solvent in MALDI negative ion mode. The highest values of mass resolution (orange) and top scores* (green) are highlighted; empty box – no ion detected.

Matrix	Concentration	Solvent	[Trp-H] ⁻	[ATP-2Na+H] ⁻	Scores
9AA	5 mg/mL	70% ACN	3488	4273	0
		100% ACN	4597	8291	1
		70% MeOH	3008	4689	0
		100% MeOH	3900	8495	1
		70% EtOH	4151	8578	2
	10 mg/mL	70% ACN	4123	8079	0
		70% MeOH	4259	8295	2
		100% MeOH	4442	10357	2
		70% EtOH	4072	7823	0
	20 mg/mL	100% ACN	4925	13375	2
		70% MeOH	1564	2152	0
		100% MeOH	3977	8153	0
		70% EtOH	4616	8109	1
DMAN	5 mg/mL	100% ACN	3058	8279	0
	10 mg/mL	100% ACN	2765	11346	1

*scores-the summary of the highest values for a given method

These data show that for both matrices tested, S/N ratio and the intensity of both ions of interest were higher for ACN and MeOH when compared to S/N ratio and intensity values for EtOH.

However, these results suggest that 9AA appears to be far superior to DMAN as only two of matrix preparation methods screened *i.e.* 5 mg/mL and 10 mg/mL in 100% ACN for DMAN were suitable for ATP and Trp detection. Also, no strong correlation between mass resolution and choice of solvent was seen: mass resolution of [Trp-H]⁻ ion fits in the range between 1564±453 and 4442±139 for MeOH, 2765±1110 and 4925±428 for ACN and 4072±253 and 4616±216 for EtOH, whereas mass resolution of [ATP-2Na+H]⁻ occurs between 2152±337 and 10357±1111 for MeOH, 4273±189 and 13375±811 for ACN and 7823±519 and 8578±548 for EtOH across all methods tested, depending on matrix concentration (**Table 11**). Additionally, a relationship between matrix concentration, S/N ratio and the intensity of ATP and Trp ions for 9AA matrix can be observed: both parameters give higher values for a concentration of 5 mg/mL (S/N between 64±7 -

1350±128, intensity between 365±20 - 10503±277 a.u.) than for 10 mg/mL (S/N between 17±5 - 182±27, intensity 84±22 - 1162±116 a.u.) or 20 mg/mL (S/N between 7 - 2992±241, intensity between 33 - 19696±1738 a.u.) (**Table 9 and 10**).

Based on a comparison of S/N ratio, intensity and mass resolution of Trp and ATP ions, 6 values for each ion, within each of the discussed properties were chosen (highlighted in orange). This was then followed by evaluation of the best methods of sample preparation ('score' column highlighted in green in **Table 9, 10, 11**). Consequently, there were 8 top methods of matrix preparation with both Trp and ATP ions detected taken into consideration (**Table 12**).

Table 12: A summary of 8 top methods of MALDI negative ion matrix preparation, including matrix, its concentration and solvent. Methods were selected based on the analytes detection, their S/N ratio, intensity and mass resolution. Methods listed by matrix concentration (from low to high) and type of solvent.

Matrix	Concentration	Solvent
9AA	5 mg/mL	70% ACN
9AA	5 mg/mL	70% MeOH
9AA	5 mg/mL	100% MeOH
9AA	5 mg/mL	70% EtOH
9AA	10 mg/mL	70% ACN
9AA	10 mg/mL	70% MeOH
9AA	10 mg/mL	100% MeOH
9AA	20 mg/mL	100% ACN

According to a summary presented in **Table 12**, 9AA turns out to be a better matrix than DMAN as well as a concentration of 5 and 10 mg/mL seems to be more relevant for small molecule profiling than a concentration of 20 mg/mL. However, further investigation of matrix background for each concentration and each solvent from **Table 12** was performed to decide on a final matrix preparation protocol for this study (**Figure 3**).

All MS spectra from pure (with no sample) 9AA matrix show similar profiles: the most abundant peak at m/z 193.091 that corresponds to $[M-H]^-$ 9AA ion can be observed. Also, the second most abundant peak is seen at m/z 96.938 that potentially corresponded to phosphoric acid $[M-H]^-$, sulphuric acid $[M-H]^-$ or phosphate $[H_2PO_4]^-$ contamination.

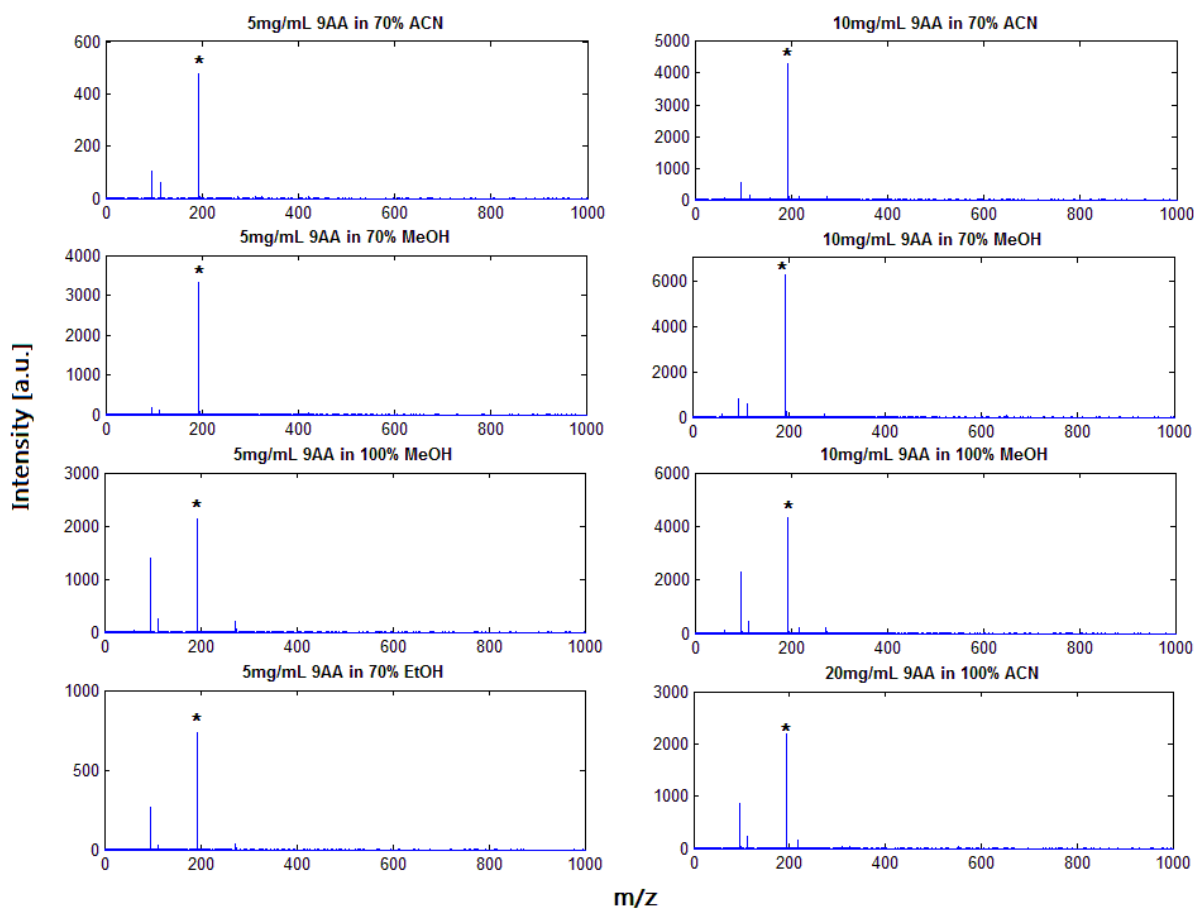


Figure 3: MALDI negative ion mode averaged spectra of pure (with no sample) matrices of 8 top methods of sample preparation; * - the most abundant matrix peak at m/z 193.

All 9AA MS profiles appear to be very similar, regardless of the solvent used for the experiment. Yet, the difference in intensity of m/z 193 peak can be seen: the lowest signal is found for 5 mg/mL 9AA in 70% ACN and 70% EtOH and 10 mg/mL 9AA in 100% ACN (<1000 a.u.), whereas 5 and 10 mg/mL 9AA in MeOH and 10 mg/mL 9AA in 70% ACN give higher m/z 193 intensity (>1000 a.u.).

5.6.2. Metabolite standards

The objective of this section was to assess MALDI-MS experiments in order to screen a number of standard compounds and their mixtures to create a calibration cocktail for further metabolomics investigations.

Table 13 presents a list of compounds that were screened in both MALDI positive and negative ionisation modes that were selected to investigate their suitability to build a standards mixture for this study. The preference was made according to the molecular weight to be between 0 and 1000 Da, a type of metabolite for instance amino acids (*e.g.* proline), organic acids (*e.g.* citric acid), nucleosides (*e.g.* cytidine), fatty acids (*e.g.* palmitic acid) and sugars (*e.g.* glucose-6-phosphate), price and their availability. The choice of standards was also made in regard to the ionisation effectiveness and their usage as calibration compounds according to available literature (Shroff *et al.* 2007, Vaidyanathan & Goodacre 2007, Xu *et al.* 2009, Dashtiev *et al.* 2007, Johnson & Lewis 2005, Sun *et al.* 2004, Helsper *et al.* 2013). The selection of solvents for standard compounds was made according to their solubility based on Material Safety Data Sheet (MSDS). Also, a consideration was given to a choice of solvents that would be similar to those used for matrices preparation *i.e.* ACN, MeOH and water to avoid introducing additional source of signals and to prevent precipitation of the analyte due to disagreement of solvent polarities. Moreover, compounds were prepared in excess concentration to ensure high MS signal. The detection of each compound is described according to their S/N ratio: very good S/N > 2000 (++++), good S/N 500–2000 (+++), medium S/N 200–500 (++) and low S/N < 200 (+).

Many studies have shown that MS analysis, identification and quantification of cells and tissue metabolites provide a fingerprint of different classes of compounds, separating different levels of biological conditions *e.g.* health and abnormal regions (Nemes *et al.* 2011, Korte & Lee 2014). This concept is applied in regard to the detection and quantification of specific classes of metabolites for example drugs. However, the approach of untargeted metabolite profiling is much more complicated as it is aimed to investigate the widest possible range of compounds from different classes (Forcisi *et al.*, 2013). Hence, a real challenge of metabolomics research is correlated to chemical complexity of sample that represents different classes of endogenous and exogenous metabolites. The vast diversity of small molecules (<1000 Da), including their physical and chemical properties such as polarity and stereochemistry as well as different biological functions and concentrations, makes the detection of the whole set of metabolites an ambitious approach (Zhao *et al.* 2010, Lenz *et al.* 2004). It is therefore necessary to optimise a standard

compounds mixture to fit within a mass range as well as different chemical and physical properties of the analyte of interest.

In total, there were 26 compounds screened in positive and negative ionisation modes. The comparison of compounds is made according to their detection and signal strength and is presented as S/N ratio. As seen in **Table 13**, out of 25 tested compounds in positive ionisation mode, there were 6 metabolites that gave very strong signal ($S/N > 2000$), 4 metabolites with good (S/N 500-2000) and medium (S/N 200-500) signal, 8 metabolites with low ($S/N < 200$) signal and 3 compounds that were not detected ($S/N < 6$) as protonated form. Most of the investigated metabolites (16 compounds) gave protonated ions $[M+H]^+$ however, sodium $[M+Na]^+$ and potassium $[M+K]^+$ adducts were also seen. Typical MALDI-MS spectra (*i.e.* with no advanced sample purification) are normally abundant with protonated species as well as sodium and potassium adducts. Protonated forms are often dominated by $[M+Na]^+$ and $[M+K]^+$ adducts and this phenomenon is seen for example in lipid analysis. The hypothesis states that a free $-OH$ group in the diacylglycerols greatly interacts with metal ions resulting in enhanced formation of such adducts (Schiller *et al.*, 1999). Similarly, there were sodium and potassium adducts of lipid species tested in this study *i.e.* DPPC and tripalmitin. Additionally, there were metal adducts detected from non-lipid species, but rich in $-OH$ group, such as atenolol, glucose-6-phosphate, adenosine 5'-triphosphate, leucine enkephalin, etoposide, GSSG, erythromycin A and bradykinin fragment 2-9. Detection of metal adducts of those compounds was reported previously (Langley *et al.* 2007, Nordstrom *et al.* 2006, Edwards & Kennedy 2005, Calvano *et al.* 2011).

Similar observation can be made for detection of compounds tested in negative ionisation mode. Most of the metabolites (13 compounds) gave low ($S/N < 200$) signal; only 2 metabolites gave very good ($S/N > 2000$) and good (S/N 500-2000) signal. Additionally, none of them showed medium (S/N 200-500) detection and there were 9 metabolites that gave no signal ($S/N < 6$) for deprotonated form. The majority of species detected (14 compounds) were deprotonated $[M-H]^-$. Analogous to positive ionisation mode, in negative ion mode the creation of adducts is equally possible (Boudah *et al.*, 2014). This could be seen for erythromycin A, where chlorine $[M+Cl]^-$ adduct was detected.

Table 13: A list of standards examined in MALDI positive and negative ionisation modes. Compounds are sorted in increasing molecular weight order. S/N ratio: very good S/N>2000 (++++), good S/N 500-2000 (+++), medium S/N 200-500 (++) , low S/N<200 (+), x – no detection (S/N <6), highlighted are compounds with the strongest signal within each ionisation mode. Table also summarises compound solvent, its stock concentration and molecular weight.

Compound	Solvent	Stock concentration [mM]*	Molecular weight [g/mol]	Positive ions	Signal	Negative ions	Signal
Glycine	50% MeOH	225.0	75.1	$[M+H]^+$	+	$[M-H]^-$	+
2-methylimidazole	Water	300.0	82.1	$[M+H]^+$	++++	$[M-H]^-$	x
Proline	Water	300.0	115.1	$[M+H]^+$	+	$[M-H]^-$	x
Nicotinic acid	90% EtOH	40.0	123.1	$[M+H]^+$	+	$[M-H]^-$	++++
Histidine	Water	200.0	155.2	$[M+H]^+$	++	$[M-H]^-$	+
Pyridoxine (vitamin B6)	Water	200.0	169.2	$[M+H]^+$	+	$[M-H]^-$	+
Arginine	Water	200.0	174.2	$[M+H]^+$	+++	$[M-H]^-$	x
Citric acid	50% MeOH	300.0	192.1	$[M+H]^+$	x	$[M-H]^-$	+++
Caffeine	Water	50.0	194.2	$[M+H]^+$	++++	$[M-H]^-$	+
Tryptophan	90% EtOH	40.0	204.2	$[M+H]^+$	+	$[M-H]^-$	+++
Cytidine	30% EtOH	130.0	243.1	$[M+H]^+$	++	$[M-H]^-$	+

Palmitic acid	Chloroform/ MeOH 1:1 (v/v)	40.0	256.4	$[M+H]^+$	x	$[M-H]^-$	+
Atenolol	25% MeOH	100.0	266.3	$[M+H]^+$	+++	$[M-H]^-$	x
				$[M+Na]^+$	+++		
				$[M+K]^+$	++		
Glucose-6-phosphate sodium salt	50% MeOH	300.0	282.1	$[M+K]^+$	+	$[M-H]^-$	x
Palmityl palmitate	Chloroform/ MeOH 0.67:0.3 (v/v)	31.3	480.8	$[M+H]^+$	x	$[M-H]^-$	+
Verapamil hydrochloride	MeOH/ EtOH 1:1 (v/v)	56.6	491.1	$[M+H]^+$	+	$[M+Cl]^-$	+
				$[M-HCl+H]^+$	++++		
				$[M-HCl-150+H]^+$	++++		
PC (18:0/0:0) 1-stearoyl-sn-glycero-3-phosphocholine	Chloroform/ MeOH 1:13 (v/v)	1.1	523.7	$[M+H]^+$	+	$[M-H]^-$	+
						$[M-N(CH_3)_3-H]^-$	+
Adenosine 5'-triphosphate disodium salt hydrate	Water	30.0	551.1	$[M+H]^+$	++	$[M-2Na+H]^-$	+
				$[M+Na]^+$	+		

				$[M-2Na+3H]^+$	++++		
Leucine enkephalin acetate salt hydrate	Water	1.8	555.6	$[M+Na]^+$	+	$[M-H]^-$	+
Polyethylene glycol 600 (PEG 600) sulfate	100% MeOH	~1 μ L in 500 μ L of solvent	570-630			$[SO_4]^{2-}$	++++
						$[H(OCH_2CH_2)_7SO_4]^-$	++
						$[H(OCH_2CH_2)_9SO_4]^-$	+++
						$[H(OCH_2CH_2)_{11}SO_4]^-$	+++
						$[H(OCH_2CH_2)_{13}SO_4]^-$	+++
Etoposide	Chloroform/ MeOH 1:1 (v/v)	109.1	588.6	$[M+K]^+$	+	$[M-206-H]^-$	+
				$[M+Na]^+$	+++		
L-Glutathione oxidized (GSSG)	Water	82.3	612.6	$[M+H]^+$	++	$[M-H]^-$	x
				$[M+Na]^+$	+++		
Erythromycin A	EtOH	54.7	733.9	$[M-H_2O+Na]^+$	++++	$[M-H]^-$	+
				$[M+Na]^+$	++	$[M+Cl]^-$	+
Dipalmitoylphosphatidylcholine (DPPC)	Chloroform/ MeOH 1:1 (v/v)	127.0	734.0	$[M+K]^+$	+	$[M-H]^-$	x
				PC headgroup	++		

Tripalmitin	Chloroform/ MeOH 9:1 (v/v)	46.0	807.3	$[M+Na]^+$	++	$[M-H]^-$	x
Bradykinin fragment 2-9	Water	23.4	904.0	$[M+H]^+$	++++	$[M-H]^-$	x
				$[M+K]^+$	+		
				$[M+Na]^+$	+++		

*- concentration unit excluding PEG600 sulfate

As soft/mild ionisation technique, MALDI-MS produces few or little fragmentation (Bergman, Shevchenko and Bergquist, 2014). Most fragments are believed to be generated as the ions travel across the drift tube, by ‘metastable’ (of a state of equilibrium) breakdown and by contact with the background gas. This mechanism is called “post – source decay” (Spengler 1997, Spengler & Kaufmann 1992). However, the fragmentation is very limited due to a presence of matrix that plays important role in controlling fragmentation process by absorbing majority of the laser energy and passing it to the analyte (Zenobi and Knochenmuss, 1998). Yet, fragmentation of species such as lipids is commonly seen in MALDI experiments (Al-Saad *et al.* 2003, Zhang *et al.* 2014, Anderson *et al.* 2013). This is believed to be caused by two reactions occurring upon lipid analysis: the quartetnary ammonium group fragmentation (Marto *et al.*, 1995) and a cleavage of the double bond in fatty acid residues (Schiller *et al.*, 1999). In this study, PC head-group at m/z 184.074 of DPPC and PC (18:0/0:0) $[M-N(CH_3)_3-H]^-$ fragment at m/z 463.619 could be detected in positive and negative ionisation modes, respectively. This has also been shown in another studies of DPPC, where PC head-group ($H_2PO_4(CH_2)_2N(\text{methyl group})_3^+$) at m/z 184.074 was detected (Jaskolla, Onischke and Schiller, 2014). Fragmentation of parent ion of PC (18:0/0:0) was confirmed with MS/MS spectra shown in **Supplementary Information, Figure 14**.

Furthermore, other fragments of non-lipid compounds could be seen as well: verapamil fragments $[M-HCl+H]^+$ and $[M-HCl-150+H]^+$ in positive ion mode and etoposide fragment $[M-206-2H]^-$ in negative ion mode. These findings are in line with reported MS/MS studies of verapamil, where fragments at m/z 303.207 ($[M-HCl-150+H]^+$) and m/z 455.291 ($[M-HCl+H]^+$) could also be detected (Sun *et al.* 2004, Cohen *et al.* 2007). Another study has confirmed etoposide fragment at m/z 381.593 for $[M-206-H]^-$ (Hande *et al.*, 1988).

Except for PEG600 sulfate, there were only 3 compounds with very good (S/N >2000) or good (S/N 500-2000) signal within all scanned calibrants in negative ion mode. These 3 compounds *i.e.* nicotinic acid, citric acid and tryptophan fit the mass range only up to 204 Da. Due to lack of full mass range coverage (0-1000 Da) and low signals from the rest of the compounds, PEG600 sulfate was selected to be the calibrant for negative ion mode in this study. Its performance as MS calibrant has been previously shown (Shroff *et al.* 2007, Shroff & Svatoš 2009).

Polymers such as polyethylene glycols (PEGs) are suitable calibrants for small molecule applications. PEGs produce a collection of peaks that are usually observed to be repeated a number of its subunits (monomers) and are accessible to assign (Garofolo, 2004). Accordingly, very good (S/N >2000) and good (S/N 500-2000) signal from a number of

PEG600 sulfate ions was seen in this study. Representative PEG600 ions: sulfate ion $[\text{SO}_4]^-$ and polymers ions ($n=7, 9, 11, 13$) are presented in **Table 13**.

Based on the detection and signal strength of the screened analytes, 6 of them were selected to create a calibration cocktail across a mass range from 0 to 1000 Da. For positive ionisation mode these analytes are (in order of increasing molecular weight): 2-methylimidazole (82 g/mol), caffeine (194 g/mol), verapamil hydrochloride (491 g/mol), adenosine 5'-triphosphate disodium salt hydrate (551 g/mol), erythromycin A (734 g/mol), bradykinin fragment 2-9 (904 g/mol) (highlighted compounds in **Table 13**). In order to evaluate a final concentration of each of these compounds in a calibration cocktail, different concentrations of these analytes were examined (**Table 14**). Based on presence of ions of interest and their S/N ratio in each mixture a final mixture C was chosen as the most suitable one and was used as the calibration cocktail for further experiments. Full data with S/N values can be found in **Supplementary Information, Table 1**). Representative MALDI spectrum of the best mixture C is presented in **Figure 4**.

Table 14: Different combinations of analytes included in the calibration mixture for MALDI positive ionisation mode analysis. Table summarises compound, its molecular weight and volumes used to create each mixture.

Compound	Molecular weight [g/mol]	Mixture		
		A	B	C
		Volume [μL] from stock solution		
2-methylimidazole	82	1.0	1.5	2.0
Caffeine	194	20.0	80.0	60.0
Verapamil hydrochloride	491	12.0	25.0	14.0
Adenosine 5'-triphosphate disodium salt hydrate	551	5.0	10.0	20.0
Erythromycin A	734	2.0	4.0	10.0
Bradykinin fragment 2-9	904	0.5	0.5	0.5

Figure 4 presents a spectrum of the calibration cocktail for MALDI positive ion experiments with ions assigned for each compound: 2-methylimidazole $[M+H]^+$ at m/z 83.061, caffeine $[M+H]^+$ at m/z 195.088, verapamil $[M-HCl-150+H]^+$ at m/z 303.207 and $[M-HCl+H]^+$ at m/z 455.291, ATP $[M+Na]^+$ at m/z 573.949, erythromycin A $[M-H_2O+Na]^+$ at m/z 738.874 and bradykinin 2-9 $[M+H]^+$ at m/z 904.468 and $[M+Na]^+$ at m/z 926.450. Additionally, three matrix peaks are detected: $[DHB+H-H_2O]^+$ at m/z 137.023, $[DHB+H]^+$ at m/z 155.024 and $[2DHB+H-2H_2O]^+$ at m/z 273.036 and can also be used for mass correction. Verapamil, erythromycin and bradykinin ions are the most intense in a calibration mixture with intensities about 9000 a.u. and higher. 2-methylimidazole, caffeine and ATP ions seem to be less intense with intensities about 2000 a.u. for 2-methylimidazole and caffeine and <1000 a.u. for ATP. In addition, a difference in the response of used compounds in the calibration cocktail to their response when analysed separately can be observed. The signal of all ions is observed to be lower (about 7 times in average) in the mixture of compounds when compared to their signal when tested separately (**Supplementary Information, Table 1 and 2**). This may be due to a different affinity of compounds to be protonated, resulting in suppression of ions with lower affinity by those with higher affinity (Cohen, Go and Siuzdak, 2007).

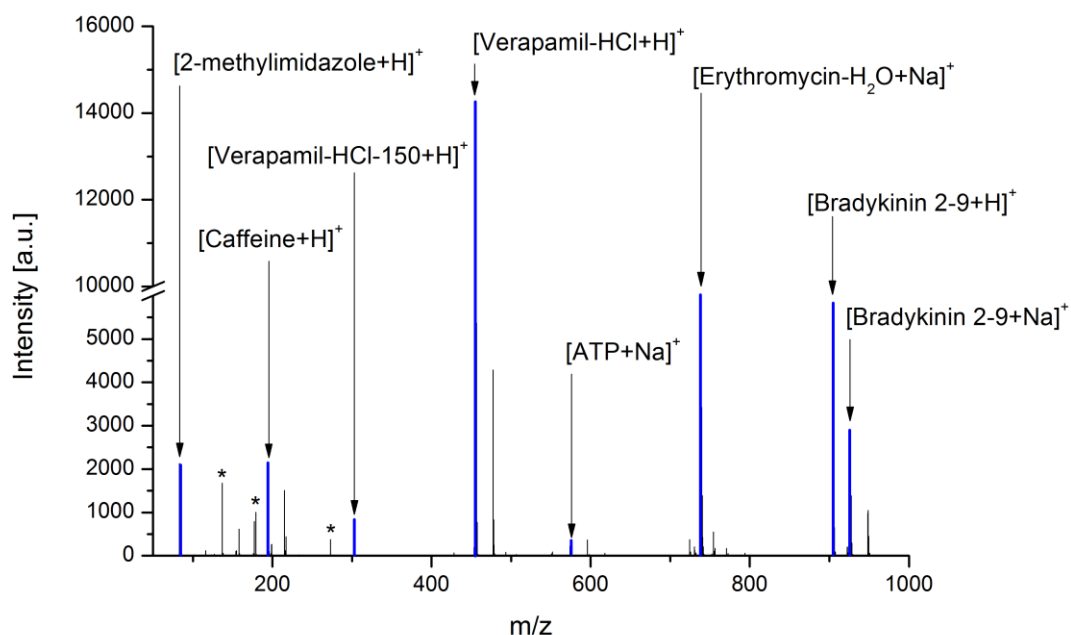


Figure 4: MS spectrum of the calibration cocktail in MALDI positive ionisation mode, mass range m/z 50 – 1000 shown, * - matrix peaks.

Similarly, based on the detection and the signal strength of analytes screened for negative ionisation mode, there were only 2 compounds showing the highest response: nicotinic acid (123 g/mol) and polyethylene glycol 600 (PEG600) sulfate (570-630 g/mol). Since PEG600 gave good signal across the whole mass range of interest, nicotinic acid was therefore excluded from further experiments. Representative MALDI spectrum of PEG600 is presented below (**Figure 5**).

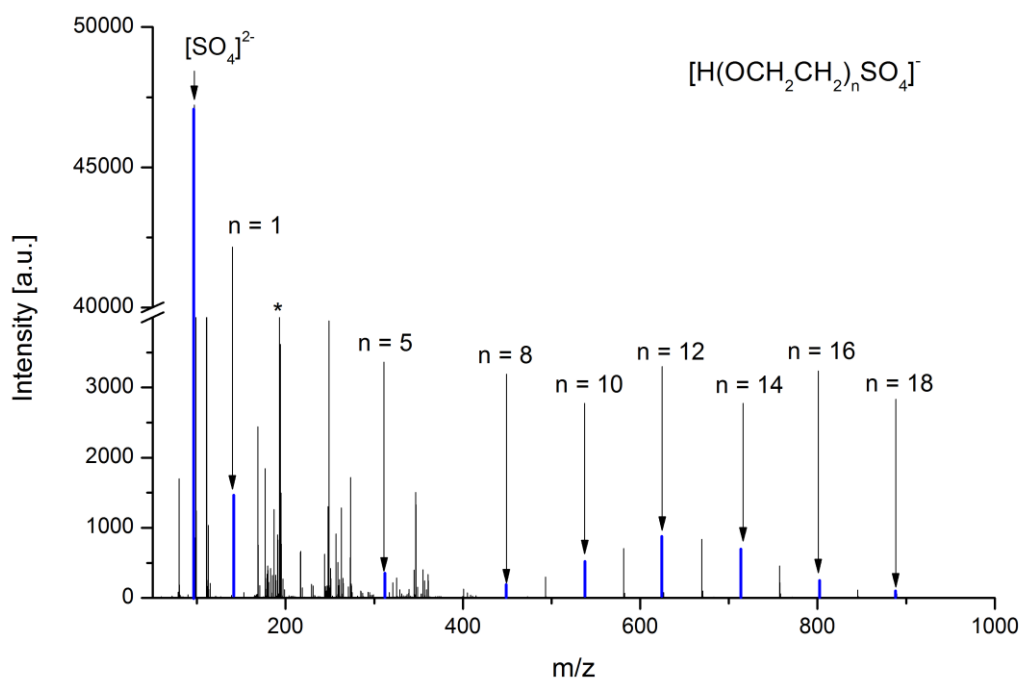


Figure 5: MS spectrum of PEG600 sulfate in MALDI negative ionisation mode, mass range m/z 50 – 1000 shown, * - matrix peaks.

Figure 5 shows an example spectrum of PEG600 for MALDI negative ion experiments with ions assigned, depending on the number of monomers repeated. 9 calibration points across a the mass range from m/z 96 for sulfate $[\text{SO}_4]^-$ to m/z 889 for 18 monomers $[\text{H}(\text{OCH}_2\text{CH}_2)_{18}\text{SO}_4]^-$ were selected and are shown in **Table 15**.

Table 15: PEG600 sulfate ions and their m/z values used for MALDI calibration in negative ionisation mode.

$[\text{H}(\text{OCH}_2\text{CH}_2)_n\text{SO}_4]^-$	m/z
$[\text{SO}_4]^-$	95.9517
$n = 1$	140.9858
$n = 5$	317.0906
$n = 8$	449.1693
$n = 10$	537.2217
$n = 12$	625.2741
$n = 14$	713.3266
$n = 16$	801.379
$n = 18$	889.4314

5.6.3. Optimisation of sample deposition method

The objective of this section was to evaluate the most suitable method of sample deposition for MALDI fingerprint and footprint profiling experiments in positive and negative ionisation modes.

5.6.3.1. Positive ionisation mode

The optimisation included the application of three commonly used ways of sample deposition: overlay, mixed and underlay (AlMasoud *et al.* 2014, Cohen *et al.* 2007). Spectra were collected and analysed for presence of 8 ions from the compounds included in the calibration cocktail. In all of 9 methods of sample preparation tested (3 *matrix-solvent combinations X 3 ways of sample deposition*), the detection of all ions of interest is seen (Table 16).

Table 16: Ions taken into consideration during optimisation of sample deposition method in MALDI positive ion mode.

Compound	<i>m/z</i>
2-methylimidazole	83.061
Caffeine	195.088
Verapamil hydrochloride	303.207, 455.291
Adenosine 5'-triphosphate disodium salt hydrate	574.949
Erythromycin A	738.874
Bradykinin fragment 2-9	904.468, 926.450

Therefore, the following step included a comparison of intensity of these ions according to different ways of sample deposition. The intensity threshold was set to 5% of a total intensity of each spectrum to refine the results. It turned out that 6 of sample deposition methods tested failed with the detection of the intensity higher than 5% of total intensity of the spectrum for one or more ions of interest thus, there were only 3 remaining combinations taken into account *i.e.* DHB with mixed and overlay method of sample deposition and THAP with overlay method of sample deposition. **Figure 6** presents the comparison of the ions intensity of three sample deposition methods applied. Inset figure illustrates zoom-in of ions with lower intensity *i.e.* *m/z* 83.061 and *m/z* 195.088.

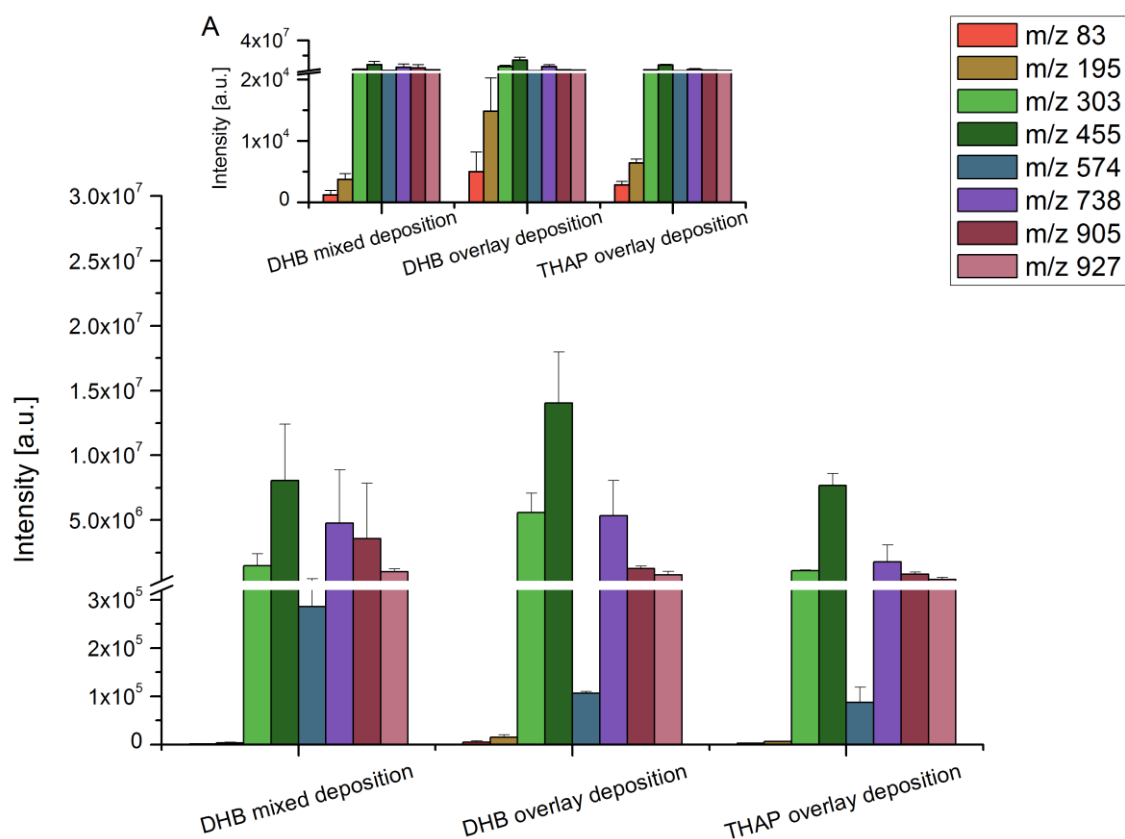


Figure 6: A comparison of intensity of ions derived from compounds creating the calibration mixture according to the sample deposition method used in MALDI positive ion mode. Expanded view of low intense ions (A). The error bars represent the standard deviation calculated from the data obtained from four experiments ($n=4$).

The comparison shows that the most intense compounds are verapamil (m/z 455.291 and m/z 303.207), erythromycin (m/z 738.874) and bradykinin 2-9 (m/z 904.468 and m/z 926.450) with intensity of 10^6 - 10^7 a.u. magnitude. Ions at m/z 83.061 and m/z 195.088 are much low intense (intensity of 10^3 - 10^4 a.u. magnitude) (**Figure 6 A**). In addition, the ion at m/z 573.949 appear to be more intense (intensity of 10^5 a.u. magnitude) when compared to the spectrum of the calibration cocktail from previous section, where its abundance was lower. A similar trend however, is observed for all three tested methods. Different dependency in intensity of these ions is most likely related to a different instrument being used for experiments in this section *i.e.* MALDI Synapt G2-Si HDMS mass spectrometer. In comparison to MALDI-ToF-MS Ultraflex II ToF/ToF mass spectrometer that uses 337 nm nitrogen laser, (http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf), MALDI

Synapt G2-Si HDMS mass spectrometer is equipped with 355 nm Neodymium - doped Yttrium Aluminium Garnet (Nd:YAG) laser (http://www.waters.com/waters/en_GB/SYNAPT-G2-Si-High-Definition-Mass-Spectrometry/nav.htm?cid=134740622&locale=en_GB).

According to **Figure 6**, there is no significant differences between intensity of observed ions within three tested methods of sample deposition. Therefore, all three methods *i.e.* DHB with mixed and overlay sample deposition and THAP with overlay sample deposition were applied for further optimisation of MALDI sample preparation protocol.

5.6.3.2. Negative ionisation mode

Similarly to positive ionisation mode, spectra for negative ion mode were collected and analysed for presence of PEG600 sulfate ions. In each of 15 methods of sample preparation (*3 ways of sample deposition X 5 different solvents*) all ions of interest could be detected. Therefore, in order to refine the choice of the best sample preparation method, two PEG600 ions *i.e.* m/z 317.0906 and m/z 625.2741 were selected. The intensity of these ions was compared for each solvent and according to different ways of sample deposition. To restrict the results, the intensity of both m/z 317.0906 and m/z 625.2741 ions within all 15 methods of sample preparation was analysed and 5 methods with the highest intensity for these ions were chosen (**Figure 7**).

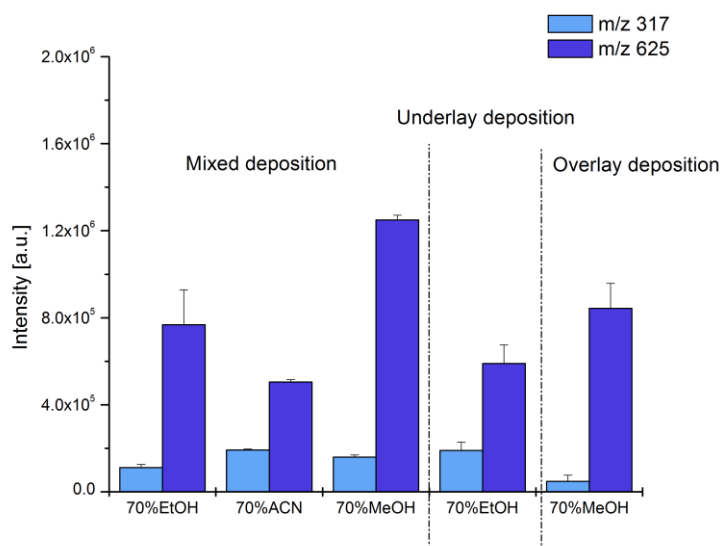


Figure 7: A comparison of the intensity of m/z 317 and m/z 625 ions for 5 top methods of sample preparation for MALDI negative ionisation mode. The error bars represent the standard deviation calculated from the data obtained from four experiments ($n=4$).

The comparison of the intensity of m/z 317.0906 and m/z 625.2741 ions shows that m/z 625.2741 ion is more abundant (4×10^5 a.u. and more) than m/z 317.0906 ion ($> 2 \times 10^5$ a.u.) within all methods of sample deposition tested. Although there was no significant difference between m/z 317.0906 and m/z 625.2741 ion intensity, there was some consistency in appearance of preferable sample deposition *i.e.* mixed deposition. Also, taking into consideration 9AA solvents used for the experiments, it was clearly shown that the concentration of 70 % was better than the pure (100 %) solvent. This might be because of too high volatility of such solutions that can cause too heterogeneous film hence, different signal response across the sample spot. The influence of the solvents on sample preparation and creation of uniform sample spots was discussed in previous section.

Within matrix solvents used for this experiment *i.e.* ACN, EtOH and MeOH, none of them seems to be outstanding. Therefore, all three solvents *i.e.* 70% EtOH, 70% ACN and 70% MeOH and mixed sample deposition were selected for further MALDI optimisation experiments.

5.6.4. A choice of a top MALDI method of sample preparation - cellular analytes

In this section, the final optimisation of MALDI sample preparation for positive and negative ionisation modes based on real cellular analytes is shown.

5.6.4.1. Positive ionisation mode

Initially, the investigation of the differences and clustering between non-used and used growth medium samples was conducted. In order to identify the level of separation between these analytes, the Principal Component Analysis (PCA) was applied. PCA scores plots were also used to visualise repeatability between analysed samples.

Figure 8 illustrates PCA scores plots with PC1 loadings plots for three sample preparation methods: DHB with mixed and overlay deposition (**Figure 8 A, B**) and THAP with overlay deposition only (**Figure 8 C**). PCA scores plots account for 89.7 %, 93.6 % and 95.1 % of total variance, respectively, within the data set. PC1 loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da.

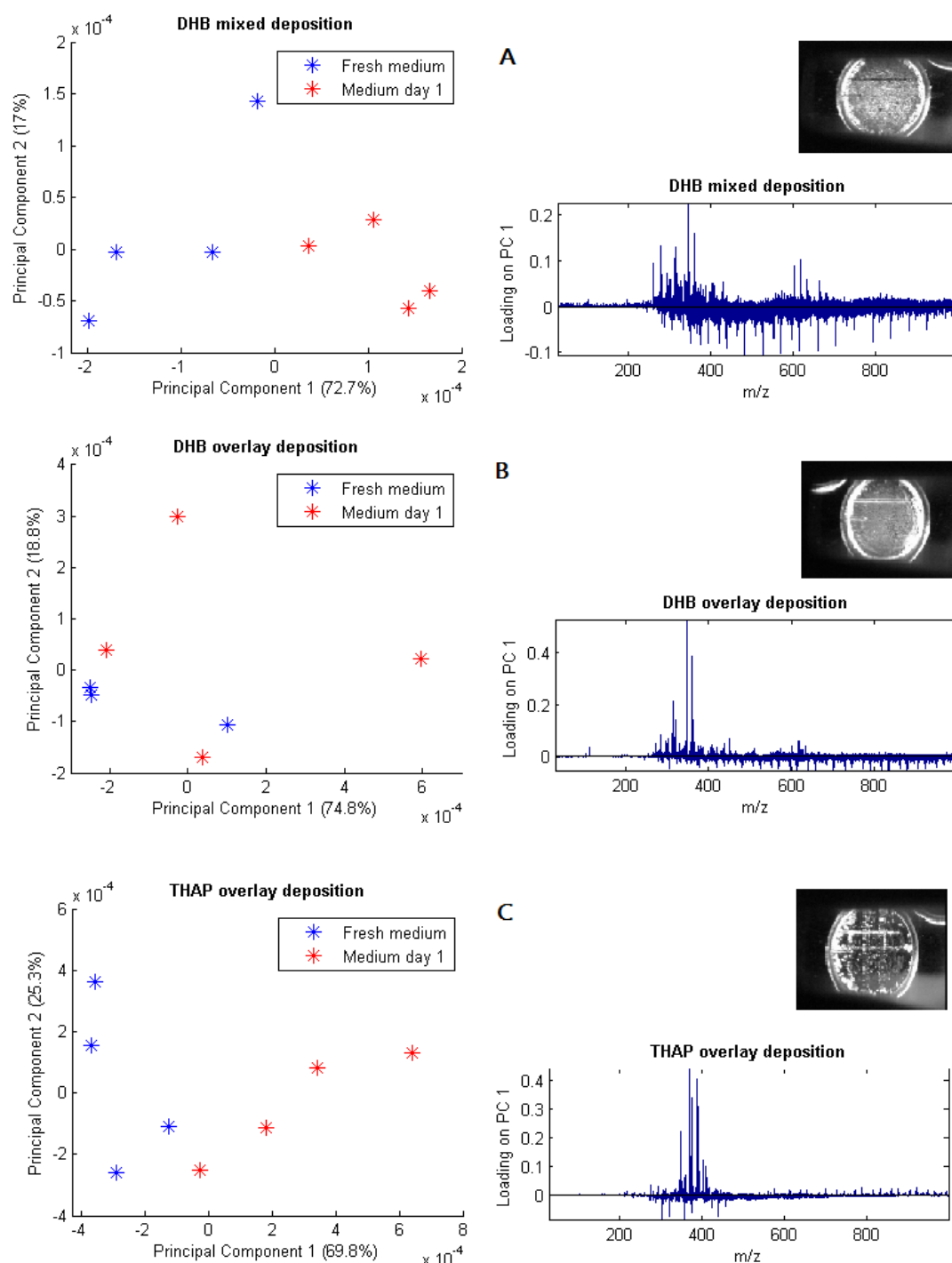


Figure 8: MALDI positive ion mode PC scores 1 and 2 for top sample preparation methods: DHB with mixed sample deposition (A), DHB with overlay sample deposition (B), THAP with overlay sample deposition (C); PC1 accounts for 89.7%, 93.6% and 95.1% of total variance, respectively. PCA plots illustrate the variance between fresh medium (blue) and used medium removed on day 1 (red). Loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da; $n = 4$, outliers removed (one repeat from B). Sample spots observed from built-in MALDI camera (magnification unknown); laser pathway seen.

Two medium samples used for the experiment are with different metabolite content: cell growth medium removed from cells on day 1 of cell culture contains diverse metabolite profile to the fresh medium, resulting from biological processes ongoing in cells. According to PC1, the separation between fresh medium and medium removed from cell culture was better for two methods: DHB with mixed deposition (PC1 72.7 %) and DHB with overlay deposition (PC1 74.8 %). The separation for THAP with overlay deposition method was also relatively good and accounted for PC1 69.8 %. However, the reproducibility and consistency of the repeats for DHB with mixed deposition and THAP with overlay deposition methods seem to be better when compared to the remaining method with DHB and overlay sample deposition.

Additionally, when analysing PC1 loading plots, the ion information carried in spectra appears to be very rich and complex, making the interpretation of loadings plots difficult. Nevertheless, the same complex information of loadings plots is delivered from all the solvent used for matrix preparation. The possible reason of such wealth in information in MS spectra is that cell media normally contain a vast number of metabolites and other defined compounds to support cell growth, viability, product formation (for example an antibody) and its quality (Butler 2005, Luo *et al.* 2012, Reinhart *et al.* 2015). Therefore, the dilution of the medium is required to obtain more readable and better resolved MS spectra. The optimisation of medium dilutions was performed, using dilution factors of 6 or 20 (depending on ionisation mode and laser power used). Full optimisation is described in **Supporting Information, Table 3, Figure 15 and 16**.

A comparison of spot surfaces was also made in order to evaluate the most homogeneous sample deposition method. According to the images shown in **Figure 8**, DHB creates good quality uniform surfaces while THAP seems to be forming more heterogeneous films with possible “hot spots” areas. Taking into consideration separation and sample repeatability compared *via* PCA as well as homogeneity of the sample surface, the DHB matrix with mixed sample deposition method appears to be the most preferable one for further studies. To support this assumption, the mass resolution analysis for ‘low’ mass region (m/z 277.3) and ‘high’ mass region (m/z 964.6) of spectra was performed (**Figure 9**).

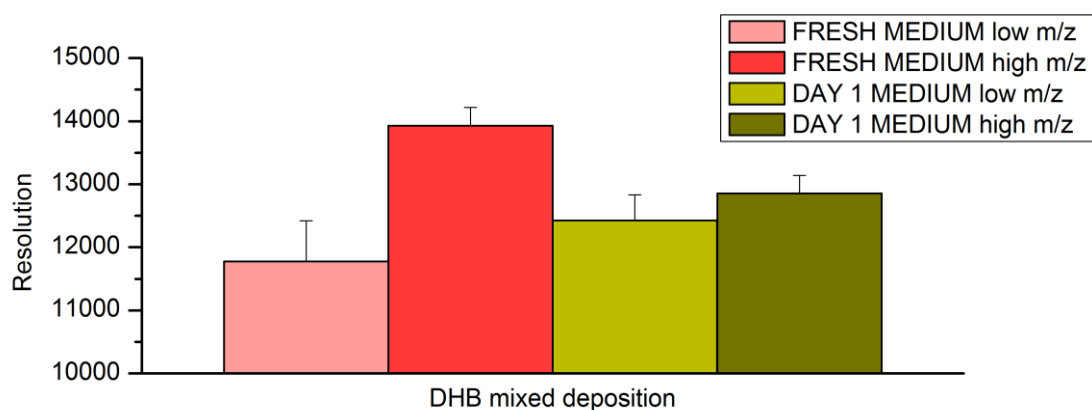


Figure 9: A comparison of mass resolution for ‘low’ and ‘high’ mass range ions for DHB with mixed sample deposition in MALDI positive ion mode. The error bars represent the standard deviation calculated from the data obtained from four experiments ($n = 4$).

As seen in **Figure 9**, the mass resolution of ‘low’ and ‘high’ mass ions for both medium samples is high (expected values ~10 000 according to the instrument specification). For the ‘low’ mass ions the mass resolution is similar and falls between >11 000 for fresh medium sample and >12 000 for medium sample removed from cells on day 1. Higher mass resolution is observed for ‘high’ mass ions (between 12 000-14 000) for both tested medium samples.

5.6.4.2. Negative ionisation mode

Spectra for MALDI negative ionisation mode were compared and analysed using PCA. PCA scores plots with PC1 loadings plots for top sample preparation methods: 9AA in 70% ACN, 9AA in 70% EtOH and 9AA in 70% MeOH are presented in **Figure 10 A-C**. PCA scores plots account for 95.3 %, 89.0 % and 93.1 % of total variance, respectively, within the data set. PC1 loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da.

Based on PC1 of PCA scores plots, the separation between fresh medium and medium removed from cell culture on day 1 is better for 70% ACN (PC1 90.0 %) and 70% MeOH (PC1 87.2 %), whereas for 70% EtOH PC1 accounts for 75.6 % of total variance. Analogously to separation, the repeatability of samples for each of 9AA solvents used within this study was compared.

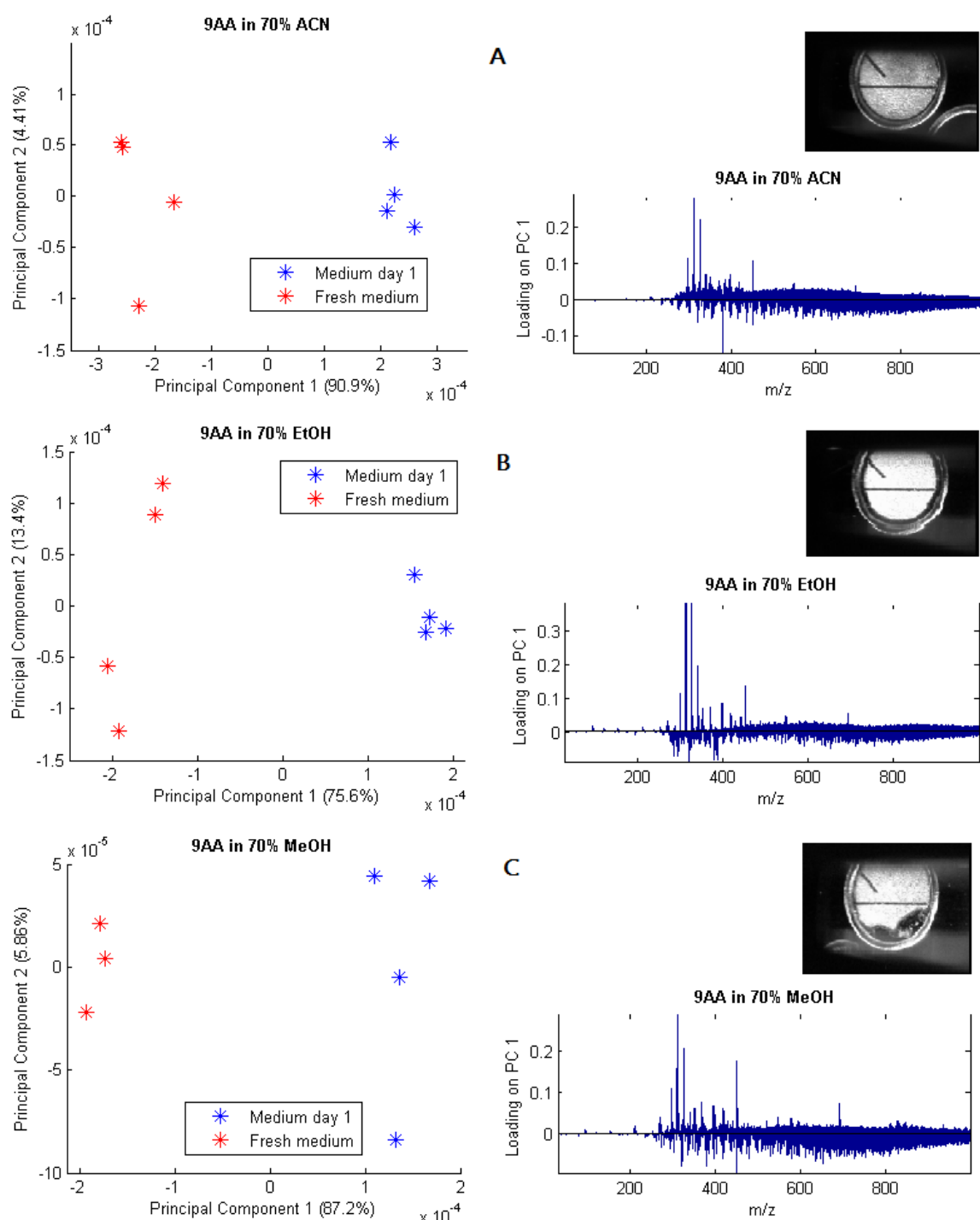


Figure 10: MALDI negative ion mode PC scores 1 and 2 for top sample preparation methods: 9AA in 70% ACN (A), 9AA in 70% EtOH (B), 9AA in 70% MeOH (C), PC1 accounts for 95.3%, 89.0% and 93.1% of total variance, respectively. PCA plots illustrate the variance between fresh medium (red) and used medium removed on day 1 (blue). Loadings plots illustrate abundance of ions detected within the mass range of 20-1000 Da; $n = 4$, outliers removed (one repeat from A and C). Sample spots observed from built-in MALDI camera (magnification unknown); laser marks seen.

More consistent sample repeats are observed for 70% ACN and 70% EtOH for medium samples removed from cell culture and for 70% in MeOH for fresh medium samples. However, the overall repeatability across all methods is similar, suggesting that there is no outstanding 9AA solvent that could be used for MALDI profiling.

Similarly, to PC1 loading plots from positive ionisation mode, PC1 loading plots in negative ion mode show the same complex information for each of 9AA solvents used, confirming the requirement of medium samples dilutions for MALDI analysis.

Additionally, a comparison of spot surfaces was also performed. Based on the images presented in **Figure 10**, the spots look homogeneous and similar for each of matrix solvent applied. No “hot spots” areas can be seen for any of the methods tested. Therefore, the analysis of mass resolution of ‘low’ (m/z 385.2), ‘medium’ (m/z 544.7) and ‘high’ (m/z 894.5, m/z 916.6, m/z 934.5 or m/z 956.5 depending on the sample) mass ions for all of 9AA solvents was performed (**Figure 11**). The ‘high’ mass ions were selected based on their presence in samples (if one ion was absent then another one was chosen for comparison).

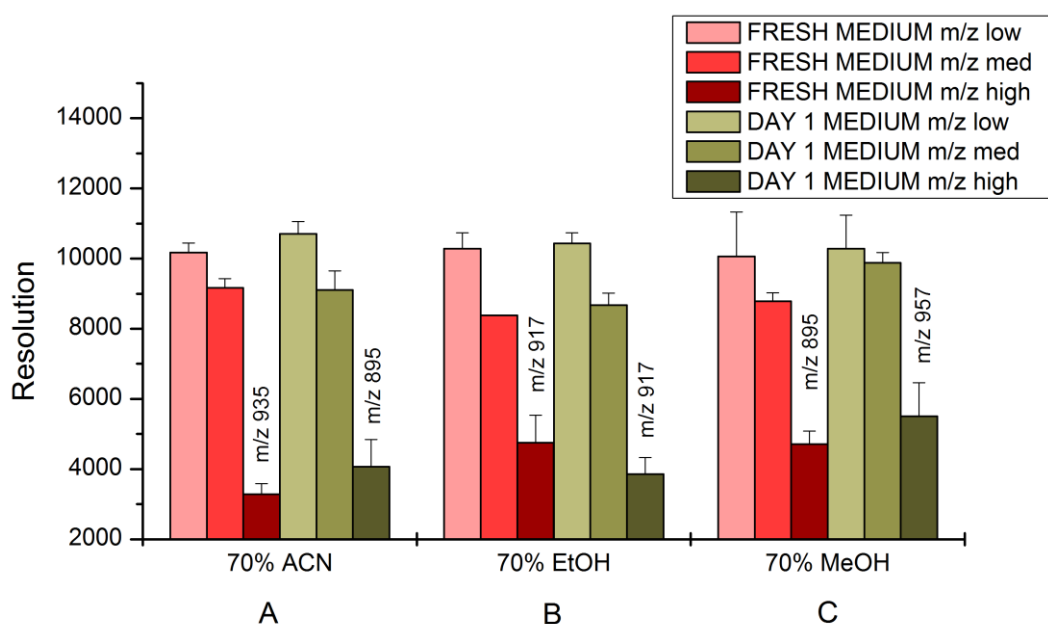


Figure 11: A comparison of mass resolution for ‘low’, ‘medium’ and ‘high’ mass range ions for 9AA dissolved in: 70% ACN (A), 70% EtOH (B) and 70% MeOH (C) in MALDI negative ion mode. The error bars represent the standard deviation calculated from the data obtained from four experiments ($n = 4$).

As seen in **Figure 11**, the comparison of mass resolution of selected ions is lower than in positive ionisation mode. Mass resolution of ‘low’ and ‘medium’ mass ions is between 8000 and 11 000 for all solvents used and for both medium samples. The highest mass resolution for ‘low’ and ‘medium’ mass range was observed for 70% MeOH for medium sample removed from cell culture on day 1 (**Figure 11 C**) and for 70% ACN for fresh medium sample (**Figure 11 A**). Mass resolution of ‘high’ mass ions appears to be lower than 10 000 and falls below 6000 for each solvent and for both samples. The comparison illustrates that the mass resolution of ‘high’ mass ions for medium sample removed from cell culture was the highest for 70% MeOH (>5000) and for fresh medium sample for 70% EtOH (between 4000 and 5000) (**Figure 11 B**).

5.7. Conclusions and future work

The goal of this chapter was to optimise the sample preparation procedure for the detection of small molecules derived from extracellular and intracellular samples of CHO cell line using MALDI-MS profiling in both positive and negative ionisation modes. The experiments involved the study of different matrices, their various concentrations as well as solvents and were used to obtain MS spectra of the mixture of arginine and histidine in positive ion mode or tryptophan and adenosine 5'-triphosphate disodium salt hydrate in negative ion mode. The experiments also included the analysis of MS spectra derived from extracellular analytes of medium from a CHO cell line.

DHB, CHCA and THAP prepared in ACN/water with TFA, EtOH or MeOH for positive ion mode and 9AA and DMAN prepared in ACN, EtOH or MeOH for negative ion mode at 5, 10 and 20 mg/mL (M:A 1:1) were analysed with the analytes spotted on the MALDI plate in three different ways: mixed, overlay or underlay deposition. The analysis of the data included the observation of the detection of the analytes of interest, their S/N ratio, intensity and mass resolution as well as the analysis of PCA separation and reproducibility between control and extracellular medium samples.

Based on the data analysis, 10 mg/mL DHB in ACN/H₂O 1:1 (v/v) with 0.1% TFA, 5 mg/mL 9AA in 70% MeOH and mixed sample deposition turned out to be the best sample preparation methods for metabolite profiling in positive and negative ion modes, respectively. In addition, the calibration cocktail for external calibration was optimised to be a selection of 6 compounds *i.e.* 2-methylimidazole, caffeine, verapamil hydrochloride, adenosine 5'-triphosphate disodium salt hydrate, erythromycin A and bradykinin fragment 2-9 for positive ion mode and PEG600 sulfate for negative ion mode.

The selected sample preparation protocols for MALDI profiling resulted with good resolved MS spectra and high intensity of detected metabolites. The calibration mixtures used for MALDI experiments resulted in high accuracy of detected ions (≤ 2 ppm).

Further investigation needs to be performed with the application of different matrix-to-analyte ratio that could increase the detection of a number of small molecules or their specific class. Moreover, the optimisation of the internal calibration mixture would be also beneficial in order to perform quantitative investigations in the future.

References:

- Al-Saad, K. A., Zabrouskov, V., Siems, W. F., Knowles, N. R., Hannan, R. M. and Hill, H. H. (2003) 'Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of lipids: Ionization and prompt fragmentation patterns', *Rapid Commun Mass Spectrom*, 17, pp. 87–96.
- AlMasoud, N., Xu, Y., Nicolaou, N. and Goodacre, R. (2014) 'Optimization of matrix assisted desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) for the characterization of Bacillus and Brevibacillus species', *Anal Chim Acta*, 840, pp. 49–57.
- AlMasoud, N., Xu, Y., Trivedi, D. K., Salivo, S., Abban, T., Rattray, N. J. W., Szula, E., AlRabiah, H., Sayqal, A. and Goodacre, R. (2016) 'Classification of Bacillus and Brevibacillus species using rapid analysis of lipids by mass spectrometry', *Anal Bioanal Chem*, 408, pp. 7865–7878.
- Alterman, M. A., Gogichayeva, N. V. and Kornilayev, B. A. (2004) 'Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry-based amino acid analysis', *Anal Biochem*, 335, pp. 184–191.
- Amantonico, A., Oh, J. Y., Sobek, J., Heinemann, M. and Zenobi, R. (2008) 'Mass spectrometric method for analyzing metabolites in yeast with single cell sensitivity', *Angew Chemi Intern Edit*, 47, pp. 5382–5385.
- Anderson, D. M. G., Mills, D., Spraggins, J., Lambert, W. S., Calkins, D. J. and Schey, K. L. (2013) 'High-resolution matrix-assisted laser desorption ionization-imaging mass spectrometry of lipids in rodent optic nerve tissue', *Mol Vision*, 19, pp. 581–592.
- Bae, Y. J., Park, K. M., Ahn, S. H., Moon, J. H. and Kim, M. S. (2014) 'Spectral reproducibility and quantification of peptides in MALDI of samples prepared by micro-spotting', *J Am Soc Mass Spectrom*, 25, pp. 1502–1505.
- Becker, L., Carre, V., Poutaraud, A., Merdinoglu, D. and Chaimbault, P. (2014) 'MALDI mass spectrometry imaging for the simultaneous location of resveratrol, pterostilbene and viniferins on grapevine leaves', *Molecules*, 19, pp. 10587–10600.
- Bergman, N., Shevchenko, D. and Bergquist, J. (2014) 'Approaches for the analysis of low molecular weight compounds with laser desorption/ionization techniques and mass spectrometry', *Anal Bioanal Chem*, 406, pp. 49–61.
- Blatherwick, E. Q., Svensson, C. I., Frenguelli, B. G. and Scrivens, J. H. (2013) 'Localisation of adenine nucleotides in heat-stabilised mouse brains using ion mobility enabled MALDI imaging', *Intern J Mass Spectrom*, 345–347, pp. 19–27.
- Boudah, S., Olivier, M. F., Aros-Calt, S., Oliveira, L., Fenaille, F., Tabet, J. C. and Junot, C. (2014) 'Annotation of the human serum metabolome by coupling three liquid chromatography methods to high-resolution mass spectrometry', *J Chromatogr B*, 966, pp. 34–47.
- Bruker (2016). Available at: http://maldi.ch.pw.edu.pl/pomiary/Artykuly/ultraflex_III_User_Manual.pdf (Accessed: 28 October 2016).
- Butler, M. (2005) 'Animal cell cultures: Recent achievements and perspectives in the production of biopharmaceuticals', *Appl Microbiol Biotechnol*, 68, pp. 283–291.
- Calvano, C. D., Monopoli, A., Ditaranto, N. and Palmisano, F. (2013) '1,8-Bis(dimethylamino)naphthalene/9-aminoacridine: A new binary matrix for lipid fingerprinting of intact bacteria by matrix assisted laser desorption ionization mass

spectrometry', *Anal Chim Acta*, 798, pp. 56–63.

Calvano, C. D., Zambonin, C. G. and Palmisano, F. (2011) 'Lipid fingerprinting of Gram-positive lactobacilli by intact cells - Matrix-assisted laser desorption/ionization mass spectrometry using a proton sponge based matrix', *Rapid Comm Mass Spectrom*, 25, pp. 1757–1764.

Cao, D., Wang, Z., Han, C., Cui, L., Hu, M., Wu, J., Liu, Y., Cai, Y., Wang, H. and Kang, Y. (2011) 'Quantitative detection of trace perfluorinated compounds in environmental water samples by matrix-assisted laser desorption/ionization-time of flight mass spectrometry with 1,8-bis(tetramethylguanidino)-naphthalene as matrix', *Talanta*, 85, pp. 345–352.

Chakrabarty, S., DeLeeuw, J. L., Woodall, D. W., Jooss, K., Narayan, S. B. and Trimpin, S. (2015) 'Reproducibility and Quantification of Illicit Drugs Using Matrix-Assisted Ionization (MAI) Mass Spectrometry', *Anal Chem*, 87, pp. 8301–8306.

Chen, R., Chen, S., Xiong, C., Ding, X., Wu, C. C., Chang, H. C., Xiong, S. and Nie, Z. (2012) 'N-(1-naphthyl) ethylenediamine dinitrate: A new matrix for negative ion MALDI-TOF MS analysis of small molecules', *J Am Soc Mass Spectrom*, 23, pp. 1454–1460.

Cohen, L., Go, E. P. and Siuzdak, G. (2007) 'Small-Molecule Desorption / Ionization Mass Analysis', in Hillenkamp, F. and Peter-Katalinić, J. (eds) *MALDI MS A Practical Guide to Instrumentation, Methods and Applications*. Münster: WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, pp. 299–337.

Cohen, S. L. and Chait, B. T. (1996) 'Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins', *Anal Chem*, 68, pp. 31–37.

Cornett, D. S., Duncan, M. A. and Amster, I. J. (1993) 'Liquid mixtures for matrix-assisted laser desorption', *Anal Chem*, 65, pp. 2608–2613.

Dashtiev, M., Waflera, E., Rohling, U., Gorshkovb, M., Franz Hillenkamp and Zenobi, R. (2007) 'Positive and negative analyte ion yield in matrix-assisted laser desorption/ionization', *Int J Mass Spectrom*, 268, pp. 122–130.

Dietmair, S., Hodson, M. P., Quek, L. E., Timmins, N. E., Chrysanthopoulos, P., Jacob, S. S., Gray, P. and Nielsen, L. K. (2012) 'Metabolite profiling of CHO cells with different growth characteristics', *Biotechnol Bioeng*, 109, pp. 1404–1414.

Dong, W., Shen, Q., Baibado, J. T., Liang, Y., Wang, P., Huang, Y., Zhang, Z., Wang, Y. and Cheung, H. Y. (2013) 'Phospholipid analyses by MALDI-TOF/TOF mass spectrometry using 1,5-diaminonaphthalene as matrix', *Int J Mass Spectrom*. Elsevier B.V., 343, pp. 15–22.

Dorries, K. and Lalk, M. (2013) 'Metabolic footprint analysis uncovers strain specific overflow metabolism and D-isoleucine production of *Staphylococcus aureus* COL and HG001', *PLoS One*, 8, pp. 81500–81510.

Duncan, M. W., Matanovic, G. and Cerpa-Poljak, A. (1993) 'Quantitative analysis of low molecular weight compounds of biological interest by matrix-assisted laser desorption ionization', *Rapid Comm Mass Spectrom*, 7, pp. 1090–1094.

Duncan, M. W., Roder, H. and Hunsucker, S. W. (2008) 'Quantitative matrix-assisted laser desorption/ionization mass spectrometry', *Briefings Funct Genomic Proteomics*, 7, pp. 355–370.

Edwards, J. L. and Kennedy, R. T. (2005) 'Metabolomic analysis of eukaryotic tissue and prokaryotes using negative mode MALDI time-of-flight mass spectrometry', *Anal Chem*, 77, pp. 2201–2209.

- Eibisch, M. and Schiller, J. (2011) 'Sphingomyelin is more sensitively detectable as a negative ion than phosphatidylcholine: A matrix-assisted laser desorption/ionization time-of-flight mass spectrometric study using 9-aminoacridine (9-AA) as matrix', *Rapid Comm Mass Spectrom*, 25, pp. 1100–1106.
- Ellis, S. R., Cappell, J., Potočník, N. O., Balluff, B., Hamaide, J., Van der Linden, A. and Heeren, R. M. A. (2016) 'More from less: high-throughput dual polarity lipid imaging of biological tissues', *The Analyst*, 141, pp. 3832–3841.
- Forcisi, S., Moritz, F., Kanawati, B., Tziotis, D., Lehmann, R. and Schmitt-Kopplin, P. (2013) 'Liquid chromatography-mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling', *J Chromatogr A*, 1292, pp. 51–65.
- Franz Hillenkamp and Karas, M. (2007) 'The MALDI Process and Method', in Franz Hillenkamp and Jasna Peter-Katalinic (ed.) *MALDI MS. A Practical Guide to Instrumentation, Methods and Applications*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- Fujita, T., Fujino, T., Hirabayashi, K. and Korenaga, T. (2010) 'MALDI mass spectrometry using 2,4,6-trihydroxyacetophenone and 2,4-dihydroxyacetophenone with cyclodextrins: suppression of matrix-related ions in low-molecular-weight region', *Anal Sci*, 26, pp. 743–748.
- Fukuyama, Y. (2011) 'Applications of Ionic Liquids in Science and Technology', in Handy, S. (ed.) *Applications of Ionic Liquids in Science and Technology*. InTech, pp. 361–374. doi: 10.5772/1769.
- Garofolo, F. (2004) 'LC-MS instrument calibration', in Chow, C., Chan, H. L., Lee, Y. C., and Zhang, X.-M. (eds) *Analytical Method Validation and Instrument Performance Verification*. John Wiley & Sons, Inc, pp. 197–220.
- Garrett, T. J., Prieto-Conaway, M. C., Kovtoun, V., Bui, H., Izgarian, N., Stafford, G. and Yost, R. A. (2007) 'Imaging of small molecules in tissue sections with a new intermediate-pressure MALDI linear ion trap mass spectrometer', *Int J Mass Spectrom*, 260, pp. 166–176.
- Goodwin, R. J. A., Iverson, S. L. and Andren, P. E. (2012) 'The significance of ambient-temperature on pharmaceutical and endogenous compound abundance and distribution in tissues sections when analyzed by matrix-assisted laser desorption/ionization mass spectrometry imaging', *Rapid Comm Mass Spectrom*, 26, pp. 494–498.
- Griffiths, R. L. and Bunch, J. (2012) 'A survey of useful salt additives in matrix-assisted laser desorption/ionization mass spectrometry and tandem mass spectrometry of lipids: Introducing nitrates for improved analysis', *Rapid Comm Mass Spectrom*, 26, pp. 1557–1566.
- Guo, Z. and He, L. (2007) 'A binary matrix for background suppression in MALDI-MS of small molecules', *Anal Bioanal Chem*, 387, pp. 1939–1944.
- Hande, K., Anthony, L., Hamilton, R., Kenneth, I., Bennett, R., Sweetman, B. and Branch, R. (1988) 'Identification of etoposide glucuronide as a major metabolite of etoposide in the rat and rabbit', *Cancer Res*, 48, pp. 1829–1834.
- Helmel, M., Marchetti-Deschmann, M. and Allmaier, G. (2014) 'Improved sample preparation for intact cell mass spectrometry (biotyping) of mycelium samples taken from a batch fermentation process of *Penicillium chrysogenum*', *Rapid Comm Mass Spectrom*, 28, pp. 957–964.
- Helsper, J. P. F. G., Peters, R. J. B., Brouwer, L. and Weigel, S. (2013) 'Characterisation

- and quantification of liposome-type nanoparticles in a beverage matrix using hydrodynamic chromatography and MALDI-TOF mass spectrometry', *Anal Bioanal Chem*, 405, pp. 1181–1189.
- Hensel, R. R., King, R. C. and Owens, K. G. (1997) 'Electrospray sample preparation for improved quantitation in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry', *Rapid Comm Mass Spectrom*, 11, pp. 1785–1793.
- Jaskolla, T. W., Onischke, K. and Schiller, J. (2014) '2,5-Dihydroxybenzoic acid salts for matrix-assisted laser desorption/ionization time-of-flight mass spectrometric lipid analysis: Simplified spectra interpretation and insights into gas-phase fragmentation', *Rapid Comm Mass Spectrom*, 28, pp. 1353–1363.
- Johnson, R. D. and Lewis, R. J. (2005) *Simultaneous Quantitation of Atenolol, Metoprolol and Propranolol in Biological Matrices Via LC / MS*. Oklahoma City, OK 73125.
- Karas, M., Bachmann, D. and Hillenkamp, F. (1985) 'Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules', *Anal Chem*, 57, pp. 2935–2939.
- Karbach, V., Knochenmuss, R. and Zenobi, R. (1998) 'Matrix-Assisted Filament Desorption/Ionization Mass Spectrometry', *J Am Soc Mass Spectrom*, 9, pp. 1226–1228.
- Kim, S.-J., Jung, H.-J., Chung, B.-C. and Choi, M.-H. (2011) 'Screening Analysis of 10 Adrenal Steroids by Matrix-Assisted Laser Desorption Ionization-Tandem Mass Spectrometry', *Mass Spectrom Letters*, 2, pp. 69–72.
- Kiss, A. and Hopfgartner, G. (2016) 'Laser-based methods for the analysis of low molecular weight compounds in biological matrices', *Methods*, 104, pp. 142–53.
- Korte, A. R. and Lee, Y. J. (2014) 'MALDI-MS analysis and imaging of small molecule metabolites with 1,5-diaminonaphthalene (DAN)', *J Mass Spectrom*, 49, pp. 737–741.
- Krause, J., Stoeckli, M. and Schlunegger, U. P. (1996) 'Studies on the selection of new matrices for ultraviolet matrix-assisted laser desorption/ionization time-of-flight mass spectrometry', *Rapid Comm Mass Spectrom*, 10, pp. 1927–1933.
- Kronthaler, J., Gstraunthaler, G. and Heel, C. (2012) 'Optimizing High-Throughput Metabolomic Biomarker Screening: A Study of Quenching Solutions to Freeze Intracellular Metabolism in CHO Cells', *J Integ Biol*, 16, pp. 90–97.
- Langley, G. J., Herniman, J. M. and Townell, M. S. (2007) '2B or not 2B, that is the question: further investigations into the use of pencil as a matrix for matrix-assisted laser desorption/ionisation', *Rapid Comm Mass Spectrom*, 21, pp. 180–190.
- Lenz, E. M., Bright, J., Knight, R., Wilson, I. D. and Major, H. (2004) 'A metabonomic investigation of the biochemical effects of mercuric chloride in the rat using ¹H NMR and HPLC-TOF/MS: time dependent changes in the urinary profile of endogenous metabolites as a result of nephrotoxicity', *The Analyst*, 129, pp. 535–541.
- Liang, C. W., Lee, C. H., Lin, Y. J., Lee, Y. T. and Ni, C. K. (2013) 'MALDI mechanism of dihydroxybenzoic acid isomers: Desorption of neutral matrix and analyte', *J Phys Chem B*, 117, pp. 5058–5064.
- van Linde, M. E., van der Mijn, J. C., Pham, T. V., Knol, J. C., Wedekind, L. E., Hovinga, K. E., Aliaga, E. S., Buter, J., Jimenez, C. R., Reijneveld, J. C. and Verheul, H. M. W. (2016) 'Evaluation of potential circulating biomarkers for prediction of response to chemoradiation in patients with glioblastoma', *J Neuroncol*, 129, pp. 221–230.
- Lou, X., van Dongen, J. L. J., Vekemans, J. A. J. M. and Meijer, E. W. (2009) 'Matrix suppression and analyte suppression effects of quaternary ammonium salts in matrix-

- assisted laser desorption/ionization time-of-flight mass spectrometry: an investigation of suppression mechanism', *Rapid Comm Mass Spectrom*, 23, pp. 3077–3082.
- Luo, J., Vijayasankaran, N., Autsen, J., Santuray, R., Hudson, T., Amanullah, A. and Li, F. (2012) 'Comparative metabolite analysis to understand lactate metabolism shift in Chinese hamster ovary cell culture process', *Biotechnol Bioeng*, 109, pp. 146–156.
- Marto, J. A., White, F. M., Seldomridge, S. and Marshall, A. G. (1995) 'Structural Characterization of Phospholipids by Matrix-Assisted Laser Desorption/Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry', *Anal Chem*, 67, pp. 3979–3984.
- Mechref, Y. and Novotny, M. V. (1998) 'Matrix-assisted laser desorption/ionization mass spectrometry of acidic glycoconjugates facilitated by the use of spermine as a co-matrix', *J Am Soc Mass Spectrom*, 9, pp. 1293–1302.
- Mohmad-Saberi, S. E., Hashim, Y. Z. H. Y., Mel, M., Amid, A., Ahmad-Raus, R. and Packeer-Mohamed, V. (2013) 'Metabolomics profiling of extracellular metabolites in CHO-K1 cells cultured in different types of growth media', *Cytotechnology*, 65, pp. 577–586.
- Muddiman, D. C., Gusev, A. I., Proctor, A., Hercules, D. M., Venkataramanan, R. and Diven, W. (1994) 'Quantitative measurement of cyclosporin A in blood by time-of-flight mass spectrometry', *Anal Chem*, 66, pp. 2362–2368.
- Nemes, P., Knolhoff, A. M., Rubakhin, S. S. and Sweedler, J. V. (2011) 'Metabolic differentiation of neuronal phenotypes by single-cell capillary electrophoresis-electrospray ionization-mass spectrometry', *Anal Chem*, 83, pp. 6810–6817.
- Nicola, A. J., Gusev, A. I., Proctor, A., Jackson, E. K. and Hercules, D. M. (1995) 'Application of the Fast-evaporation Sample Preparation Method for Improving Quantification of Angiotensin II by Matrix-assisted Laser Desorption/Ionization', *Rapid Comm Mass Spectrom*, 9, pp. 1164–1171.
- Nordstrom, A., Apon, J. V., Uritboonthai, W., Go, E. P. and Siuzdak, G. (2006) 'Surfactant-Enhanced Desorption/Ionization on Silicon Mass Spectrometry', *Anal Chem*, 78, pp. 272–278.
- Paz, B., Riobó, P. and Franco, J. M. (2011) 'Preliminary study for rapid determination of phycotoxins in microalgae whole cells using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.', *Rapid Comm Mass Spectrom*, 25, pp. 3627–39.
- Phan, N. T. N., Mohammadi, A. S., Dowlathahi Pour, M. and Ewing, A. G. (2016) 'Laser Desorption Ionization Mass Spectrometry Imaging of Drosophila Brain Using Matrix Sublimation versus Modification with Nanoparticles', *Anal Chem*, 88, pp. 1734–1741.
- Reinhart, D., Damjanovic, L., Kaisermayer, C. and Kunert, R. (2015) 'Benchmarking of commercially available CHO cell culture media for antibody production', *Appl Microbiol Biotechnol*, 99, pp. 4645–4657.
- Roberts, L. D., Souza, A. L., Gerszten, R. E. and Clish, C. B. (2013) 'Targeted Metabolomics', *Curr Protoc Mol Biol*, 6, pp. 1–34.
- Römpf, A. and Spengler, B. (2013) 'Mass spectrometry imaging with high resolution in mass and space', *Histochem Cell Biol*, 139, pp. 759–783.
- Sarkar, P. K., Prajapati, P. K., Shukla, V. J., Ravishankar, B. and Choudhary, A. K. (2009) 'Toxicity and recovery studies of two ayurvedic preparations of iron', *Ind J Exp Biol*, 47, pp. 987–992.
- Schiller, J., Arnhold, J., Benard, S., Müller, M., Reichl, S. and Arnold, K. (1999) 'Lipid analysis by matrix-assisted laser desorption and ionization mass spectrometry: A

methodological approach', *Anal Biochem*, 267, pp. 46–56.

Schiller, J., Müller, K., Süß, R., Arnhold, J., Gey, C., Herrmann, A., Leßig, J., Arnold, K. and Müller, P. (2003) 'Analysis of the lipid composition of bull spermatozoa by MALDI-TOF mass spectrometry - A cautionary note', *Chem Phys Lipids*, 126, pp. 85–94.

Schober, Y., Guenther, S., Spengler, B. and Römpf, A. (2012) 'Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging', *Anal Chem*, 84, pp. 6293–7.

Sellick, C. A., Hansen, R., Stephens, G. M., Goodacre, R. and Dickson, A. J. (2011) 'Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling TL - 6', *Nat Prot*, 6, pp. 1241–1249.

Serafín-Higuera, I., Garibay-Cerdenares, O. L., Illades-Aguilar, B., Flores-Alfaro, E., Jiménez-López, M. A., Sierra-Martínez, P. and Alarcón-Romero, L. D. C. (2016) 'Differential proteins among normal cervix cells and cervical cancer cells with HPV-16 infection, through mass spectrometry-based Proteomics (2D-DIGE) in women from Southern México', *Proteome Sci*, 14, p. 10.

Shanta, S. R., Kim, T. Y., Hong, J. H., Lee, J. H., Shin, C. Y., Kim, K.-H., Kim, Y. H., Kim, S. K. and Kim, K. P. (2012) 'A new combination MALDI matrix for small molecule analysis: application to imaging mass spectrometry for drugs and metabolites', *The Analyst*, 137, pp. 5757–62.

Shariatgorji, M., Nilsson, A., Goodwin, R. J. A., Svenningsson, P., Schintu, N., Banka, Z., Kladni, L., Hasko, T., Szabo, A. and Andren, P. E. (2012) 'Deuterated Matrix-Assisted Laser Desorption Ionization Matrix Uncovers Masked Mass Spectrometry Imaging Signals of Small Molecules', *Anal Chem*, 84, pp. 7152–7157.

Shroff, R., Muck, A. and Svatos, A. (2007) 'Analysis of low molecular weight acids by negative mode matrix-assisted laser desorption/ionization time-of-flight mass spectrometry', *Rapid Comm Mass Spectrom*, 21, pp. 3295–3300.

Shroff, R. and Svatos, A. (2009) 'Proton Sponge: A Novel and Versatile MALDI Matrix for the Analysis of Metabolites Using Mass Spectrometry', *Anal Chem*, 81, pp. 7954–7959.

Sić, S., Maier, N. M. and Rizzi, A. M. (2016) 'Quantitative profiling of O-glycans by electrospray ionization- and matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry after in-gel derivatization with isotope-coded 1-phenyl-3-methyl-5-pyrazolone', *Anal Chim Acta*, 935, pp. 187–196.

Sigma-Aldrich (2017). Available at: <http://www.sigmaaldrich.com/> (Accessed: 10 March 2017).

Spengler, B. (1997) 'Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules', *J Mass Spectrom*, 32, pp. 1019–1036.

Spengler, B. and Kaufmann, R. (1992) 'Gentle Probe for Tough Molecules: MALD Mass Spectrometry.', *Analysis*, 20, pp. 91–101.

Stübiger, G., Belgacem, O., Rehulka, P., Bicker, W., Binder, B. R. and Bochkov, V. (2010) 'Analysis of oxidized phospholipids by MALDI mass spectrometry using 6-Aza-2-thiothymine together with matrix additives and disposable target surfaces', *Anal Chem*, 82, pp. 5502–5510.

Stubiger, G., Pittenauer, E., Belgacem, O., Rehulka, P., Widhalm, K. and Allmaier, G. (2009) 'Analysis of human plasma lipids and soybean lecithin by means of high-performance thin-layer chromatography and matrix-assisted laser desorption/ionization

mass spectrometry', *Rapid Comm Mass Spectrom*, 23, pp. 2711–2723.

Suarez, E., Nguyen, H. P., Ortiz, I. P., Lee, K. J., Kim, S. B., Krzywinski, J. and Schug, K. A. (2011) 'Matrix-assisted laser desorption/ionization-mass spectrometry of cuticular lipid profiles can differentiate sex, age, and mating status of *Anopheles gambiae* mosquitoes', *Anal Chim Acta*, 706, pp. 157–163.

Sugiura, Y., Honda, K., Kajimura, M. and Suematsu, M. (2014) 'Visualization and quantification of cerebral metabolic fluxes of glucose in awake mice', *Proteomics*, 14, pp. 829–838.

Sun, G., Yang, K., Zhao, Z., Guan, S., Han, X. and Gross, R. W. (2008) 'Matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis of cellular glycerophospholipids enabled by multiplexed solvent (Supplementary Material) dependent analyte-matrix interactions', *Anal Chem*, 80, pp. 7576–7585.

Sun, L., Zhang, S.-Q. and Zhong, D.-F. (2004) 'In vitro identification of metabolites of verapamil in rat liver microsomes', *Acta Pharmacol Sinica*, 25, pp. 121–128.

Thomas, A., Charbonneau, J. L., Fournaise, E. and Chaurand, P. (2012) 'Sublimation of new matrix candidates for high spatial resolution imaging mass spectrometry of lipids: Enhanced information in both positive and negative polarities after 1,5-diaminonaphthalene deposition', *Anal Chem*, 84, pp. 2048–2054.

Trim, P. J. and Snel, M. F. (2016) 'Small molecule MALDI MS imaging: Current technologies and future challenges', *Methods*, 104, pp. 127–141.

Vaidyanathan, S. and Goodacre, R. (2007) 'Quantitative detection of metabolites using matrix-assisted laser desorption/ionization mass spectrometry with 9-aminoacridine as the matrix', *Rapid Comm Mass Spectrom*, 21, pp. 2072–2078.

Vermillion-Salsbury, R. L. and Hercules, D. M. (2002) '9-Aminoacridine as a matrix for negative mode matrix-assisted laser desorption/ionization', *Rapid Comm Mass Spectrom*, 16, pp. 1575–1581.

Waters Corporation (2016). Available at:

http://www.waters.com/waters/en_GB/SYNAPT-G2-Si-High-Definition-Mass-Spectrometry/nav.htm?cid=134740622&locale=en_GB (Accessed: 28 October 2016).

Weidmann, S., Barylyuk, K., Nespovitaya, N., Mädler, S. and Zenobi, R. (2013) 'A new, modular mass calibrant for high-mass MALDI-MS', *Anal Chem*, 85, pp. 3425–3432.

Xu, F., Zou, L., Lin, Q. and Choon Nam Ong (2009) 'Use of liquid chromatography/tandem mass spectrometry and online databases for identification of phosphocholines and lysophosphatidylcholines in human red blood cells', *Rapid Comm Mass Spectrom*, 23, pp. 3243–3254.

Ye, H., Gemperline, E., Venkateshwaran, M., Chen, R., Delaux, P. M., Howes-Podoll, M., Ané, J. M. and Li, L. (2013) 'MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula*-*Sinorhizobium meliloti* symbiosis', *Plant J*, 75, pp. 130–145.

Zaima, N., Hayasaka, T., Goto-Inoue, N. and Setou, M. (2010) 'Matrix-assisted laser desorption/ionization imaging mass spectrometry', *Int J Mol Sci*, 11, pp. 5040–5055.

Zenobi, R. and Knochennuss, R. (1998) 'Ion formation in MALDI mass spectrometry', *Mass Spectrom Rev*, 17, pp. 337–366.

Zhang, H., Smith, M. A. and Purves, R. W. (2014) 'Optimization of Triacylglycerol-estolide Analysis by Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry', *J Ame Oil Chem Soc*, 91, pp. 905–915.

Zhang, S. and Yao, Z.-P. (2012) 'Improved detection of phosphopeptides by negative ion matrix-assisted laser desorption/ionization mass spectrometry using a proton sponge co-matrix', *Anal Chim Acta*, 711, pp. 77–82.

Zhao, X., Fritsche, J., Wang, J., Chen, J., Rittig, K., Schmitt-Kopplin, P., Fritsche, A., Häring, H. U., Schleicher, E. D., Xu, G. and Lehmann, R. (2010) 'Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits', *Metabolomics*, 6, pp. 362–374.

CHAPTER 6: Metabolic profiling approach to investigate transfected CHO cell line response upon EPO-Fc fusion protein production

6.1. Introduction

Characterisation of metabolites present within a cell ('fingerprint' or 'internal metabolome', 'endometabolome') and metabolites that occur in the surrounding cell microenvironment *e.g.* cell medium ('footprint' or 'external metabolome', 'exometabolome') is beneficial for the assessment of cellular activity and metabolic state of a cell (Sellick *et al.*, 2011, Kell *et al.*, 2005).

A cell is the smallest, basic, structural and functional unit of living organisms that belongs to one of the two groups: containing a nucleus *i.e.* eukaryotic, such as CHO cell line used in this study (ATCC, 2014), or cells absent in nucleus *i.e.* prokaryotic. Animal cells consist of cytoplasm enclosed within a cell membrane, containing a variety of organelles *e.g.* lysosome, mitochondria, centrosome, ribosome, endoplasmic reticulum, Golgi apparatus and nucleus (**Figure 1**). The cytoplasm is a fluid containing water and solutes that holds cell organelles. The organelles are highly organised sub-cellular arrangements within a cell, with each having its specific shape and function (Klyszejko-Stefanowicz, 2002).

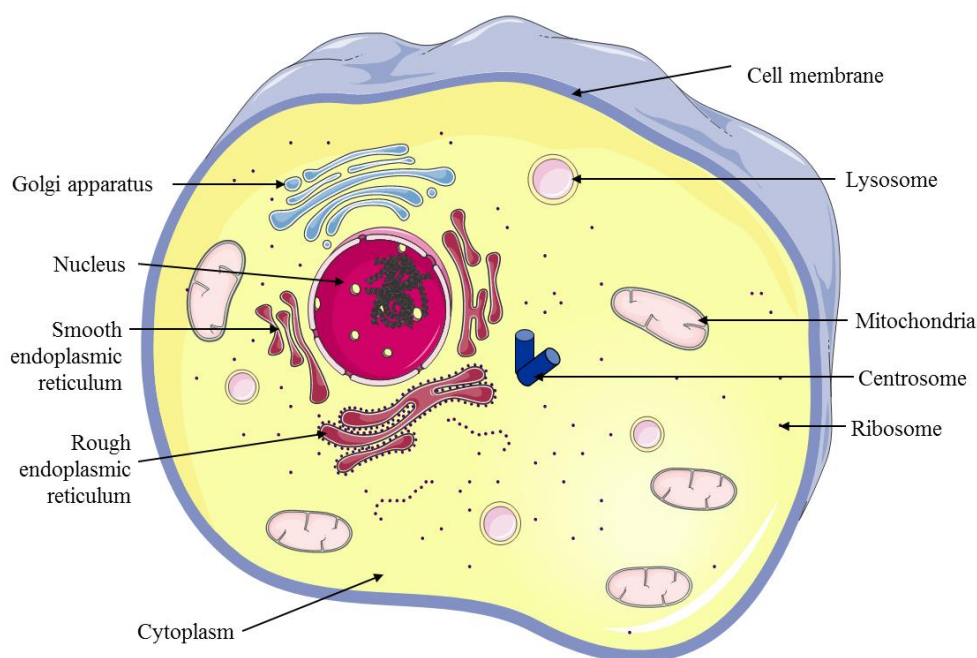


Figure 1: General view of an animal cell with cytoplasm enclosed within a cell membrane, containing a variety of organelles (Klyszejko-Stefanowicz, 2002).

In order to grow, develop and function, cells maintain a set of chemical transformations that build a large and complex metabolic network consisting of multiple interactions *i.e.* metabolic pathways (Falk Schreiber, 2003). **Figure 2** illustrates a general overview of a

metabolic network of CHO cells with major metabolic pathways highlighted: glycan biosynthesis and metabolism; nucleotide metabolism; metabolism of cofactors and vitamins; lipid metabolism; carbohydrate metabolism; amino acid metabolism (*e.g.* alanine, glycine, cysteine *etc.*); metabolism of other amino acid (*e.g.* glutathione, glutamate *etc.*); energy metabolism; biosynthesis of other secondary metabolites; metabolism of terpenoids and polyketides and xenobiotics biodegradation and metabolism (http://www.genome.jp/kegg/pathway.html). Each metabolic pathway is built from a number of biochemical reactions that are linked by their intermediates (metabolites): the products of one reaction are the substrates for following reactions *etc.* (Falk Schreiber, 2003). Therefore, the analysis of metabolites of a particular organism can be linked to metabolic pathways and their specific reactions, allowing for in-depth insight into the molecular mechanisms of cells (Sellick *et al.*, 2011).

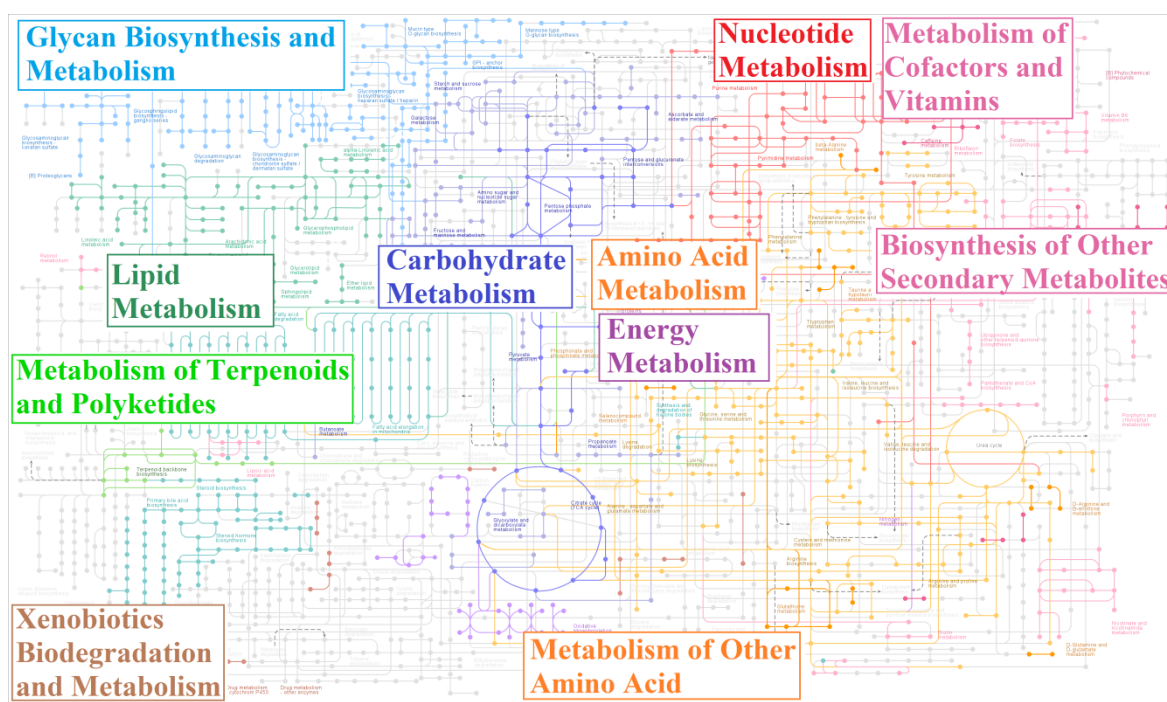


Figure 2: General overview of a complexity of a metabolic network in CHO cells. Metabolic network consists of major metabolic pathways highlighted in the figure. Adopted from (http://www.genome.jp/kegg/pathway.html).

Glycolysis and citric cycle metabolic processes are two of the main chains of reactions occurring in living cells that belong to carbohydrate and energy metabolic pathways. Glycolysis is a metabolic process, where glucose is converted into pyruvate with a release of free energy. During glycolysis, the free energy is released in a form of adenosine triphosphate (ATP) from a transfer of a phosphate group (P_i) to adenosine diphosphate

(ADP) and in a form of reduced nicotinamide adenine dinucleotide (NADH) from NAD^+ molecules **Figure 3 A** (Klyszejko-Stefanowicz, 2002).

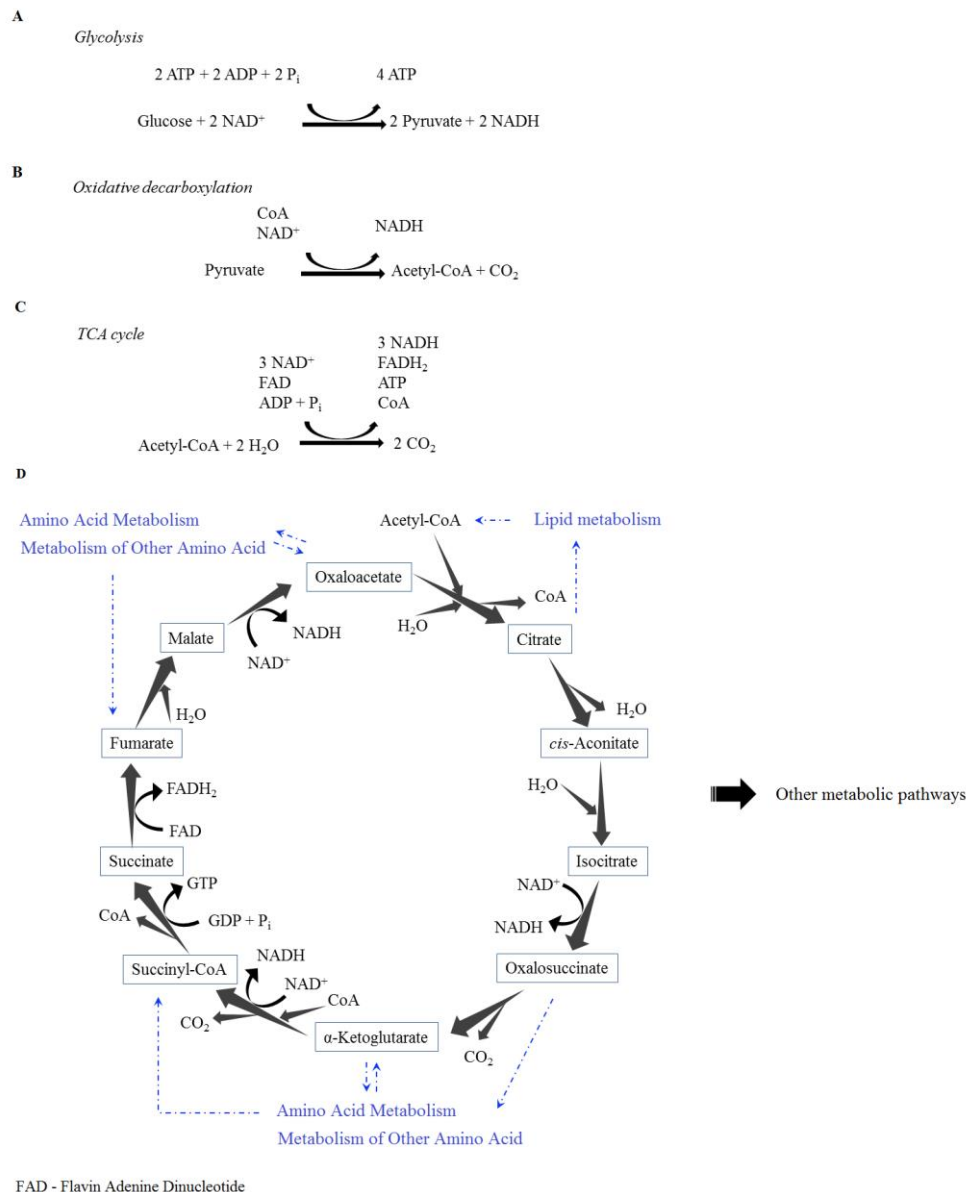


Figure 3: Carbohydrate and energy metabolic pathways highlighting the major processes occurring in cells: glycolysis (A), oxidative decarboxylation (B) and TCA cycle (C, D). Carbohydrate and energy pathways are directly (in blue) and indirectly linked to other metabolic pathways (Klyszejko-Stefanowicz, 2002, <http://www.genome.jp/kegg/pathway.html>).

Glycolysis takes place in a cytoplasm of a cell and occurs before another important metabolic pathway *i.e.* citrate cycle, also known as tricarboxylic acid (TCA) cycle ongoing in mitochondria (**Figure 3 C, D**). Once pyruvate is created, it is further transformed to acetyl-CoA during oxidative decarboxylation process (**Figure 3 B**) and then introduced to the TCA cycle. TCA cycle involves multiple reactions (**Figure 3 D**), where the energy

stored in a form of carbohydrates, lipids and proteins is converted to chemical energy *i.e.* ATP and NADH (**Figure 3 C**) (Klyszejko-Stefanowicz, 2002).

Carbohydrate and energy metabolisms are also linked to other metabolic pathways, such as amino acid metabolism or lipid metabolism (<http://www.genome.jp/kegg/pathway.html>).

Lipid metabolism is an important pathway that includes degradation or synthesis of lipids that play a crucial role in a number of biological functions *e.g.* storage of energy, signalling or they act as structural components of cell membranes (Klyszejko-Stefanowicz, 2002).

A cell membrane (also plasma membrane) forms the outside barrier of the cell that separates the internal environment from outside and controls the transport of substances, in and out of a cell. A cell membrane consists of proteins and three major classes of lipids *i.e.* phospholipids, glycolipids and cholesterol that create a lipid bilayer (**Figure 4**) (Klyszejko-Stefanowicz, 2002, Bretscher, 1972).

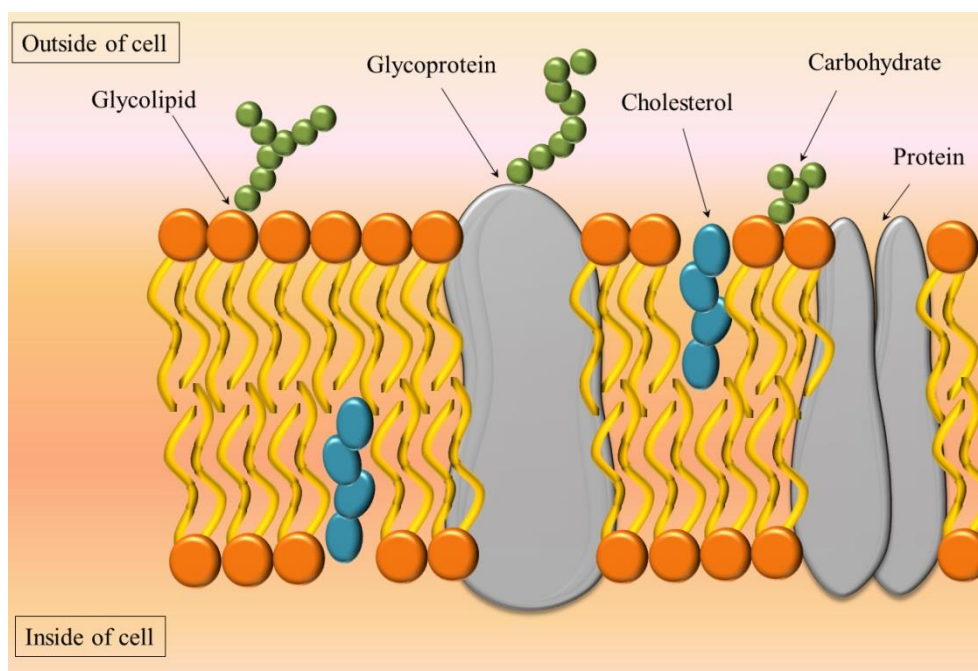


Figure 4: Schematic of a cell membrane. A cell membrane consists of proteins and three major classes of lipids: phospholipids, glycolipids and cholesterol that create lipid bilayer (Klyszejko-Stefanowicz, 2002, Bretscher, 1972).

This chapter presents a profiling approach of metabolites present within CHO cells (fingerprint) and metabolites that occur in the cell medium (footprint) in order to investigate the cell response to EPO-Fc fusion protein expression. Characterisation of metabolites allows for the assessment of cellular activity of these cells and their metabolic state upon protein production.

6.2. Sample preparation

Footprint and fingerprint samples were collected according to the procedure described in **Chapter 2**. In total, there were 4 biological replicates of control cells (non-induced) and 4 biological replicates of tetracycline-induced cells. Samples were collected at two time-points: day 1 and day 3 of cell culture. These time-points were chosen in order to investigate metabolism changes in transfected CHO cells under inducer exposure (day 1) as well as during early and stable phase of the protein production process (day 1, day 3) (see **Chapter 4** for more details).

GC-MS derivatization was performed based on The Human Serum Metabolome project (HUSERMET) standard practice described previously (Dunn *et al.* 2011, Begley *et al.* 2009, Halket *et al.* 2005) and a procedure used within the group. All compounds were purchased from Sigma-Aldrich, Gillingham, UK. The application of quality control (QC) samples is proposed in order to assure that the experiment is valid and results are reliable. A QC sample can be described as a ‘mean’ of all tested metabolites from a given study, usually obtained by pooling all samples within the experiment (Dunn *et al.*, 2012). The employment of QC samples within entire experiment at regular intervals can also improve detection of variations during data acquisition *e.g.* abundance or retention time. A further quality measure in the experimental design is sample randomisation within the analysed sequence of samples. This practice decreases the bias occurring when preparing and testing samples together (Bouhifd *et al.*, 2015).

An internal standard solution 1 (IS1) containing 10 mg of succinic- d_4 acid, malonic- d_2 acid, glycine- d_5 in 10 mL of water and IS2 (IS1 in water 0.7:9.3 v/v) were prepared. A 200 μ L of footprint samples and footprint quality control samples (QCs), 3 mL of fingerprint samples and fingerprint QCs and 3 mL of PBS samples previously lyophilised were then spiked with 100 μ L of IS2. A two-stage chemical derivatization was applied on the dried samples, where 50 μ L of 20 mg/mL O-methoxyamine hydrochloride in dry pyridine solution was added. Samples were vortexed for 10 s, spun at 17000 g for 1 min in room temperature and heated at 65 °C for 40 min. This was then followed by addition of 50 μ L *N*-acetyl-*N*-[trimethylsilyl]-trifluoroacetamide (MSTFA), vortexing (10 s), spinning (17000 g, 1 minute at room temperature) and heating at 65 °C for 40 min. A retention index solution 1 (IR1) (30 mg docosane, nonadecane, 40 μ L of decane, dodecane, pentadecane in 10 mL hexane) and retention index solution 2 (IR2) (IR1 in dry pyridine 1:9 v/v) were prepared and 20 μ L of IR2 was added to each sample. Samples were vortexed for 10 s and spun at 17000 g for 15 min at room temperature. Around 90 μ L of each sample was placed

into GC vials; GC vials were randomised and samples were taken for analysis. In total there were 72 samples analysed (*4 biological replicates X 2 control/induced cells X 2 time-points X 2 footprint/fingerprint + 24 QCs + 16 PBS samples*).

Optimisation of sample preparation method for MALDI experiments was performed and it has been described in **Chapter 5**. For MALDI-MS experiments, a 50 μ L footprint, fingerprint samples as well as footprint quality control and fingerprint quality control samples were defrosted and used for experiments. Footprint samples were diluted 6 times and 20 times in water, for positive and negative ionisation modes, respectively. Diluted footprint samples and defrosted, lyophilised fingerprint samples were then mixed with 10 mg/mL of DHB in ACN/water 1:1 (v/v) with 0.1% TFA (M:A, 1:1) for positive ionisation mode and in 5 mg/mL of 9AA in 70% MeOH (M:A, 1:1) for negative ionisation mode and spotted on a MALDI plate using mixed deposition. In total there were 360 samples analysed in positive and negative ionisation modes (*4 biological replicates X 2 control/induced cells X 2 time-points X 2 footprint/fingerprint X 10 technical replicates + 40 QCs*).

6.3. Instrumentation

GC-MS analysis was performed using an Agilent 6890N gas chromatograph oven (Wokingham, UK) coupled to a LECO Pegasus III electron ionisation time-of-flight mass spectrometer (EI-ToF-MS) (LECO Corporation, St Joseph, USA) controlled using ChromaTOF software ver. 2.32. MALDI-MS experiments were performed using MALDI Synapt G2–Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK). Detailed settings and experiment conditions for GC-MS and MALDI analyses are described in **Chapter 3**.

6.4. Data analysis

Deconvolution of raw GC-MS data was performed using LECO ChromaTof ver. 2.32 software package (LECO Corporation, St Joseph, USA) and is described in **Chapter 3**. The data matrix was exported from GC-MS software to Microsoft Excel ver. 2007 files and was weight corrected according to biomass for each sample. MALDI-MS spectra were collected as .raw files, combined and converted to .txt files using MassLynx ver. 4.1 SCN941 build 18 software (Waters Corporation/Micromass, Manchester, UK).

All collected data were then processed and further analysed using MATLAB ver. R2012a and MetaboAnalyst 3.0. 'Fish-plots' of averaged MS spectra were generated using sum normalisation and m/z 0.02 binning. MetaboAnalyst 3.0 analysis was conducted using m/z

0.25 binning, KNN missing value estimation, SD data filtering, sum normalisation and Pareto scaling; outliers removed. *t*-test analysis was performed. Metabolite peaks were chosen as significant with a *p* value threshold $p < 0.05$. Metabolite identifications were assigned through inspecting and matching against an in-house constructed libraries (for GC-MS data), NIST metabolome libraries or The Human Metabolome Database (HMDB) (<http://www.hmdb.ca/>). A *definite* match implies the retention index and mass value meet the authentic standard analysed on the same instrument. A *possible* hit means that the mass value matches only to non in-house library and cannot be confirmed *via* in-house comparison (for GC-MS). Metabolomics Standards Initiative (MSI) was used for *definite* match (level 1) and *possible* match (level 2) (Sumner *et al.*, 2007). For HMDB metabolite identification a threshold of 5 ppm of mass error was used. Principal component analysis (PCA) and Partial least square-discriminant analysis (PLS-DA) were further performed. PCA, PLS-DA scores plots and Variable Importance on Projection (VIP) scores were generated. Pathway analysis was conducted using MetaboAnalyst 3.0 software and was based on *t*-test performance. The PBS and QC samples were used for the pre-analysis of full sets of data and were applied for the assessment data quality of the experiments and acted as an internal reference within the individual samples (Dunn 2008, Dunn *et al.* 2011, Brown *et al.* 2009, Kronthaler *et al.* 2012, Sellick *et al.* 2009, Sumner *et al.*, 2007).

6.5. Results and discussion

External (footprint) and internal (fingerprint) metabolite profiles investigation included the analysis of MS spectra derived from growth medium samples and metabolite extracts collected from CHO cell culture and analysed using GC-MS and MALDI-MS techniques.

6.5.1. Footprint and fingerprint profile analysis upon 1 and 3 days of the inducer exposure

At first glance, the MALDI-MS spectra derived from control and cell-induced samples appear to have a very similar pattern within both external and internal metabolome in the mass range of m/z 20-1000 and for both MALDI+ and MALDI- ion modes (**Figure 5** and **Figure 6**).

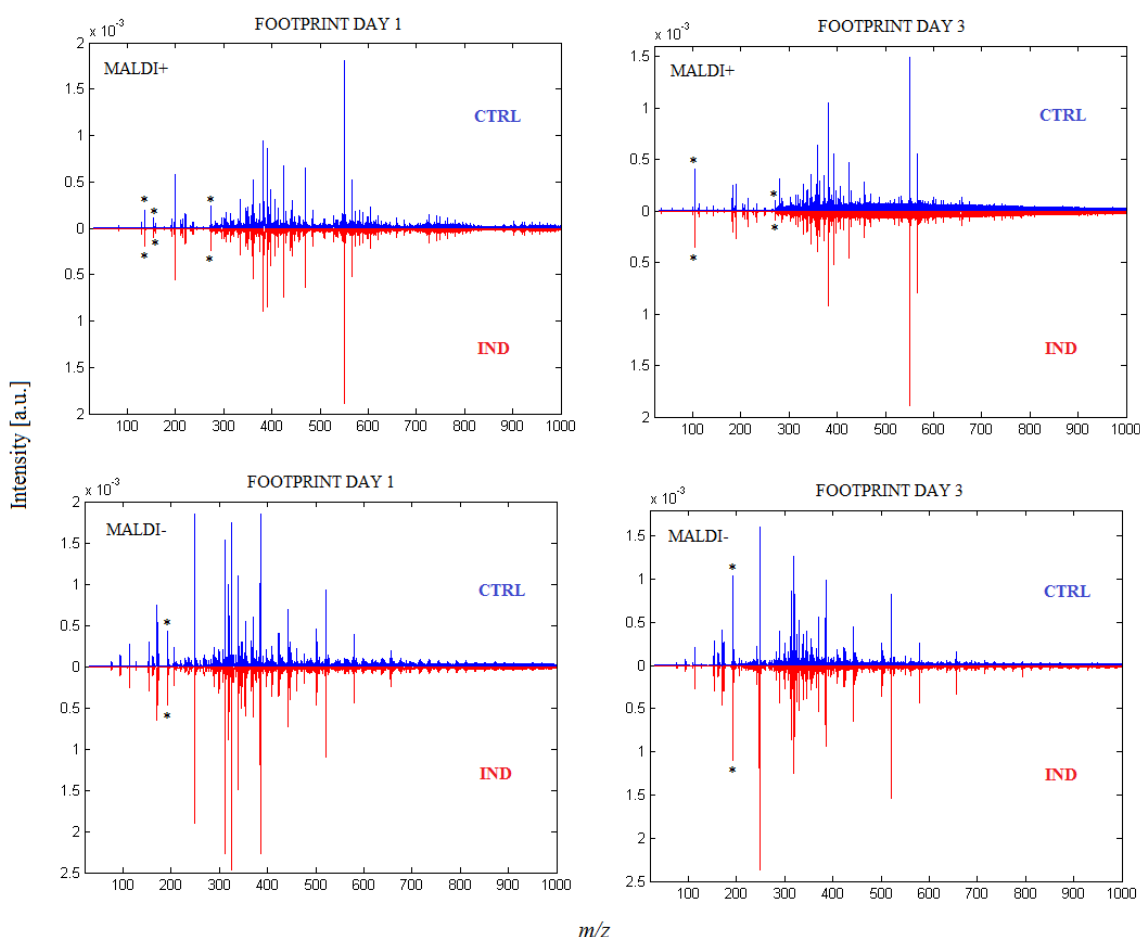


Figure 5: Averaged MALDI-MS spectra in positive and negative ion modes from control (CTRL, blue) and cell-induced (IND, red) footprint samples according to inducer exposure time; m/z 20-1000, *-matrix peaks.

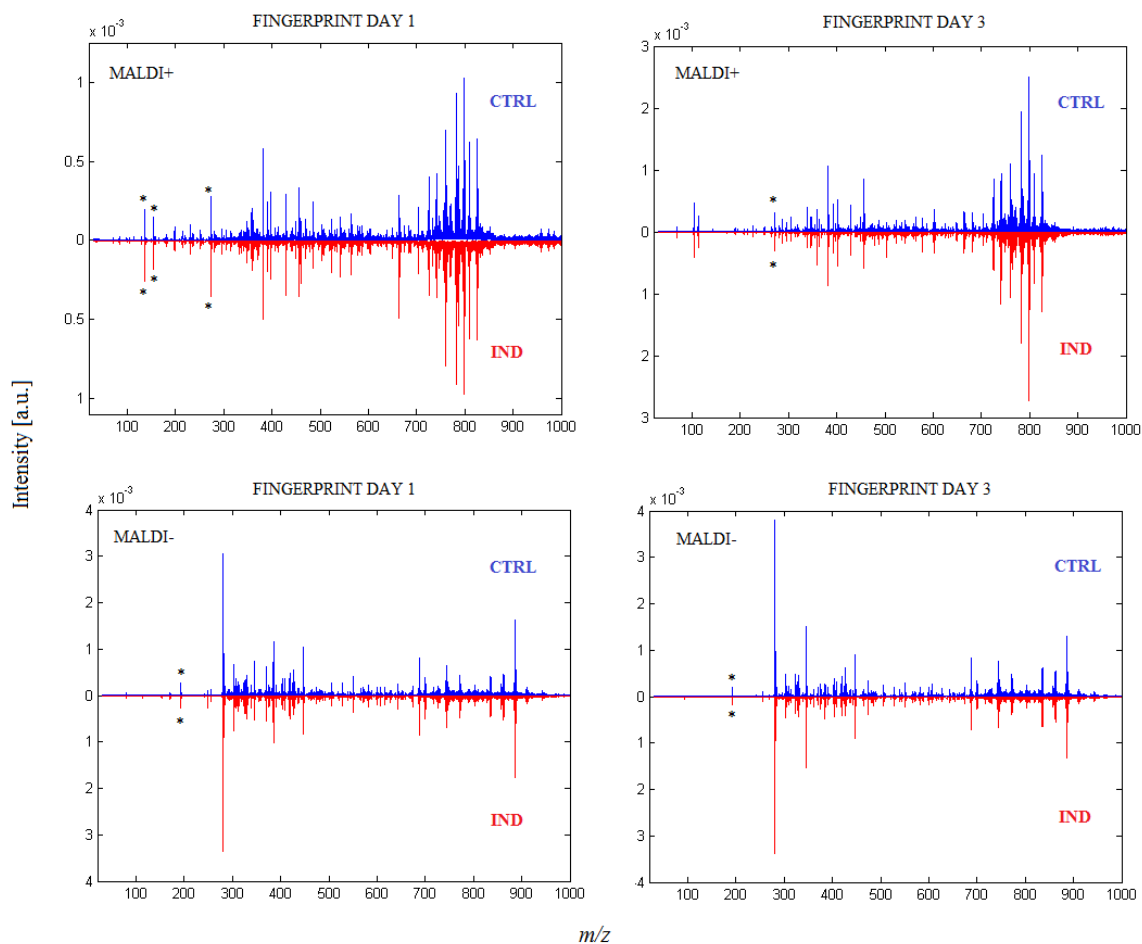


Figure 6: Averaged MALDI-MS spectra in positive and negative ion modes from control (CTRL, blue) and cell-induced (IND, red) fingerprint samples according to inducer exposure time; m/z 20-1000, *-matrix peaks.

The spectra derived from footprint samples are rich in information and metabolites are detected across the complete mass range screened, mainly below m/z 800. Some parts of the spectra have similar metabolites detected but with different abundance, for example the mass region between m/z 300–500 for MALDI+ on day 1. Also, some of the spectral regions remain the same for samples from both time points: the mass range between m/z 300-400 and m/z 500-600 for MALDI+ and the mass range between m/z 200-300 and m/z 400-700 for MALDI-. On the other hand, other peak areas of the spectra are unique to each sample: the mass range $m/z > 700$ for MALDI+ and the mass range between m/z 300-400 for MALDI-.

As with the footprint profile, fingerprint MS spectra are rich in information and metabolites can be seen across the complete mass range. Some spectral regions have little information with low abundance of peaks according to time of the inducer exposure: below m/z 200 for MALDI+ and below m/z 250 for MALDI-, where matrix peaks can mainly be

observed. There are some parts of the spectra that by eye remain the same for both control and cell-induced samples and in both time points: between m/z 300-500 and m/z 700-900 for MALDI+ and between m/z 400-900 for MALDI-. Yet, there are some peaks that differ between the inducer exposure times: m/z 900 and above for MALDI+. Also, the abundance of some peaks is visibly greater than others. The mass region of m/z 700-900 is strongly abundant in both fingerprint samples and in both time points for MALDI positive ion mode. In MALDI negative ion mode, only some peaks are visibly intense *i.e.* $\sim m/z$ 281, $\sim m/z$ 448, $\sim m/z$ 688, $\sim m/z$ 745 and $\sim m/z$ 886 for control and cell-induced samples and for both times of inducer exposure. Additionally, matrix peaks within footprint and fingerprint MS profiles were detected: $[\text{DHB}+\text{H}-\text{H}_2\text{O}]^+$ at m/z 137, $[\text{DHB}+\text{H}]^+$ at m/z 155, $[\text{2DHB}+\text{H}-2\text{H}_2\text{O}]^+$ at 273 and $[\text{9AA}-\text{H}]^-$ at m/z 193 (**Figure 6**).

The analysis of transfected CHO external and internal metabolic profiles suggests that the exposure of cell culture to the protein inducer triggers changes in the metabolome of these cells. Footprint MALDI-MS spectra illustrate the overall metabolic profiles of both control and cell-induced medium samples, highlighting detection of metabolites released from cells to medium across the whole analysed mass range, especially above m/z 300 for MALDI+ and below m/z 800 for MALDI-. These mass regions with the highest abundance are well known for the detection of nucleotides, lipids, amino acids, waste metabolites such as lactate and others (Sellick *et al.*, 2011, Hollywood *et al.*, 2015, Dorries & Lalk, 2013, Section, 2013). Similarly, fingerprint MALDI-MS spectra show presence of different types of metabolites in cell extract across the whole of the mass range tested. In contrast to footprint profiles, the fingerprint MS spectra are dominated by the peaks in the mass regions of m/z 700-900 for MALDI+, suggesting high appearance of lipids in the membrane and internal membranes of transfected CHO cells (Ellis *et al.* 2013, Jaskolla *et al.* 2014, Son *et al.* 2014).

On visual inspection of the data, only some changes in fingerprint profile of cell-induced when compare to control cells are seen and no significant differences between control and cell-induced samples within external metabolome are observed. Therefore, further multivariate data analysis was performed, in order to expose the most statistically significant variations of metabolites that occur in the external and internal metabolome in the protein-induced cells.

6.5.2. Multivariate data analysis

Metabolite peaks were chosen as significant through the application of a *t*-test ($p < 0.05$) based analysis for both GC-MS and MALDI-MS data. The number of metabolites that were recognised as significantly different in footprint and fingerprint profiles of cell-induced samples varied greatly between time of the inducer exposure and also between used MS techniques and polarities (Table 1).

Table 1: A number of statistically significant metabolites in CHO cells footprint and fingerprint samples according to the inducer exposure time and the technique used.

Number of metabolite features recognised as significant ($p < 0.05$)			
Technique/Inducer exposure time	Day 1	Day 3	Sample
GC-MS	2	3	footprint
MALDI +	54	658	
MALDI -	611	272	
GC-MS	15	11	fingerprint
MALDI +	799	1059	
MALDI -	1253	1132	

In GC-MS analysis of footprint profile, there are only 2 (day 1) and 3 (day 3) significant metabolites found that corresponds to approximately 4 % and 6 % of all detected metabolites from footprint samples, respectively. Unlike GC-MS analysis, the MALDI-MS study highlights 54 (day 1) and 658 (day 3) for positive ion mode and 611 (day 1) and 272 (day 3) for negative ion mode statistically significant metabolites from control and cell-induced samples that corresponds to 3 % and 31 % for positive ion mode and 30 % and 13 % for negative ion mode of all deemed metabolites, respectively. The data also show the increasing number of statistically significant metabolites delivered from footprint samples with the longer inducer exposure time.

On the other hand, the GC-MS analysis of fingerprint profiles results in 15 (day 1) and 11 (day 3) significant metabolites found that account for 30 % and 22 % of all detected metabolites in fingerprint samples, respectively. A higher number of significant metabolites is found with MALDI-MS technique application. 799 (day 1) and 1059 (day 3) significant metabolites are seen in MALDI positive ion mode that correspond to 38 % and 51 % of the total number of detected compounds in fingerprint samples. In MALDI negative ion mode, there are over 1000 statically significant metabolites observed in both

days of inducer exposure that corresponds to 61 % (day 1) and 55 % (day 3) of all detected metabolites from control and cell-induced fingerprint samples.

Within this study, MALDI-MS offered a rapid approach (~ 1 min/sample spot) for identification of metabolite differences between control and tetracycline-induced cells, while with GC-MS the time of the sample analysis increased (~ 30 min/sample) due to the presence of the separation step. Due to the increased time of the sample analysis as well as the high volume of the sample used for GC-MS technique (~ 200 µL/sample), it was not practical to evaluate a high number of sample replicates that is beneficial for better data analysis. Unlike GC-MS, MALDI-MS as a fast-acquisition technique allowed for the profiling of a larger number of footprint and fingerprint samples. Nevertheless, both MS methods showed great capability for generation of information-rich metabolomics data sets, in which metabolites from different classes were identified.

The following step of the data analysis incorporated the use of PCA and PLS-DA to identify the fundamental relations between two groups of samples and to find any trends within the data that could potentially be linked to the inducer exposure and EPO-Fc fusion protein production. PCA and PLS-DA scores plots were also used to visualise repeatability between analysed samples.

Two dimensional (2D) PCA scores plots generated from GC-MS footprint analysis shown in **Figure 7 A** account for 74.8 % (day 1 of inducer exposure) and 70.0 % (day 3 of inducer exposure) of the total variance using first two components. The PCA plots illustrate no separation between control and induced samples on day 1 yet, some separation between these samples on day 3 can be observed. In order to determine the level of variation between control and cell-induced samples, PLS-DA method was applied. 2D PLS-DA plots generated for the set of data shown in **Figure 7 B** highlight good separation of both control and cell-induced samples however, the plots also show variance in sample repeatability within these groups. PLS-DA plots generated for the data set using first two components account for 58.6 % and 54.2 % of the total variance according to the inducer exposure time (day 1 and day 3), respectively.

Likewise, the analysis of the data set from GC-MS fingerprint profile shows some separation between control and cell-induced fingerprint samples for both time points tested shown in PCA scores plots in **Figure 8 A**. The total variance between these samples account for 76.9 % (day 1) and 79.2 % (day 3) using first two principal components. Improved separation between control and cell-induced samples is observed in PLS-DA plots (**Figure 8 B**). The PLS-DA plots also show some variance in repeatability in the majority of fingerprint samples. The total explained variance for fingerprint profile based

on Component 1 and 2 for accounts for 75.6 % (day 1) and 72 % (day 3) of the inducer exposure (**Figure 8 B**).

The most significant variables in the PLS-DA analysis *i.e.* the 15 most statistically significant variables are presented on VIP scores diagrams shown in **Figure 7 C** for external metabolome and **Figure 8 C** for internal metabolome for GC-MS data. For GC-MS analysis, VIP scores plots are constructed using metabolite IDs taken from GC-MS libraries as variables (Y axis, left) and the VIP scores corresponding to each metabolite ID (X axis). VIP scores also show the abundance of particular metabolite ID in footprint or fingerprint samples (red for high abundance and green for low).

In footprint analysis, VIP scores from day 1 of inducer exposure illustrate 9 metabolites more abundant in the control group and 6 metabolites more abundant in the cell-induced group. VIP scores from day 3 of inducer exposure reveal a different trend, where more metabolites (11) become more abundant in induced cells. Additionally, 8 metabolites occur on both days of the inducer exposure. Some of them remain in high concentrations over both days of inducer exposure: metabolite IDs 68 and 42 in control footprint samples and metabolite IDs 39, 93 and 28 in cell-induced footprint samples, whereas some of them change over time: the abundance of metabolite IDs 88 and 15 in control samples decreases with time and increases in cell-induced samples. The abundance of metabolite IDs 52 decreases in cell-induced samples over time and increases in control samples over time.

VIP scores of GC-MS fingerprint data from day 1 shown in **Figure 8 C** highlight that the majority of significant metabolites (12 out of 15) are found to be more abundant in control cells. This observation changes on day 3 of inducer exposure, where 8 metabolites are more abundant in control samples and 7 metabolites with the higher intensity in cell-induced samples are observed. Also, 7 metabolites are present in day 1 and day 3 fingerprint samples and occur either in control samples or cell-induced samples: metabolite IDs 17, 14, 63, 18, 94, 9, and 77.

Similar analysis was performed for MALDI-MS data acquired for footprint and fingerprint samples of CHO cells in both positive (MALDI+) and negative (MALDI-) ionisation modes.

Three-dimensional (3D) PCA scores plots generated from MALDI-MS footprint analysis shown in **Figure 9 A** and **Figure 10 A** illustrate almost no separation between control and cell-induced samples for both ionisation modes and account for; MALDI+: 53.2 % (day 1 of inducer exposure) and 57.6 % (day 3 of inducer exposure) (**Figure 9 A**); MALDI-: 74.6 % (day 1 of inducer exposure) 76.3 % (day 3 of inducer exposure) (**Figure 10 A**) of the total variance using first three components.

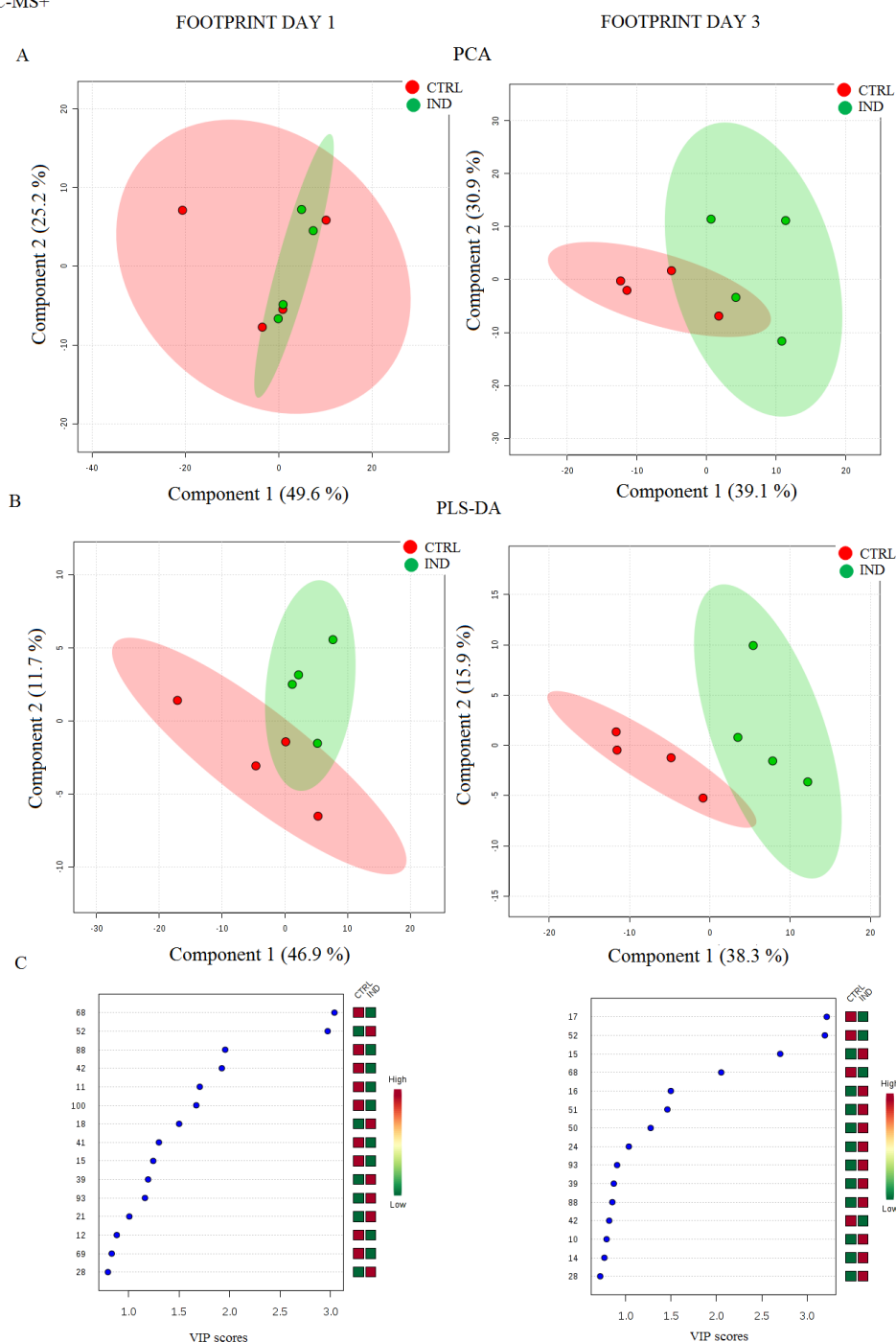


Figure 7: Positive ion mode GC-MS metabolite footprint analysis: PCA scores plots (A) accounting for 74.8 % and 70.0 % of the total variance from the data set and PLS-DA plots (B) accounting for 58.6% and 54.2% of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores (see text for details) from Component 1 (C) show the 15 most statistically significant variables (metabolite IDs) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

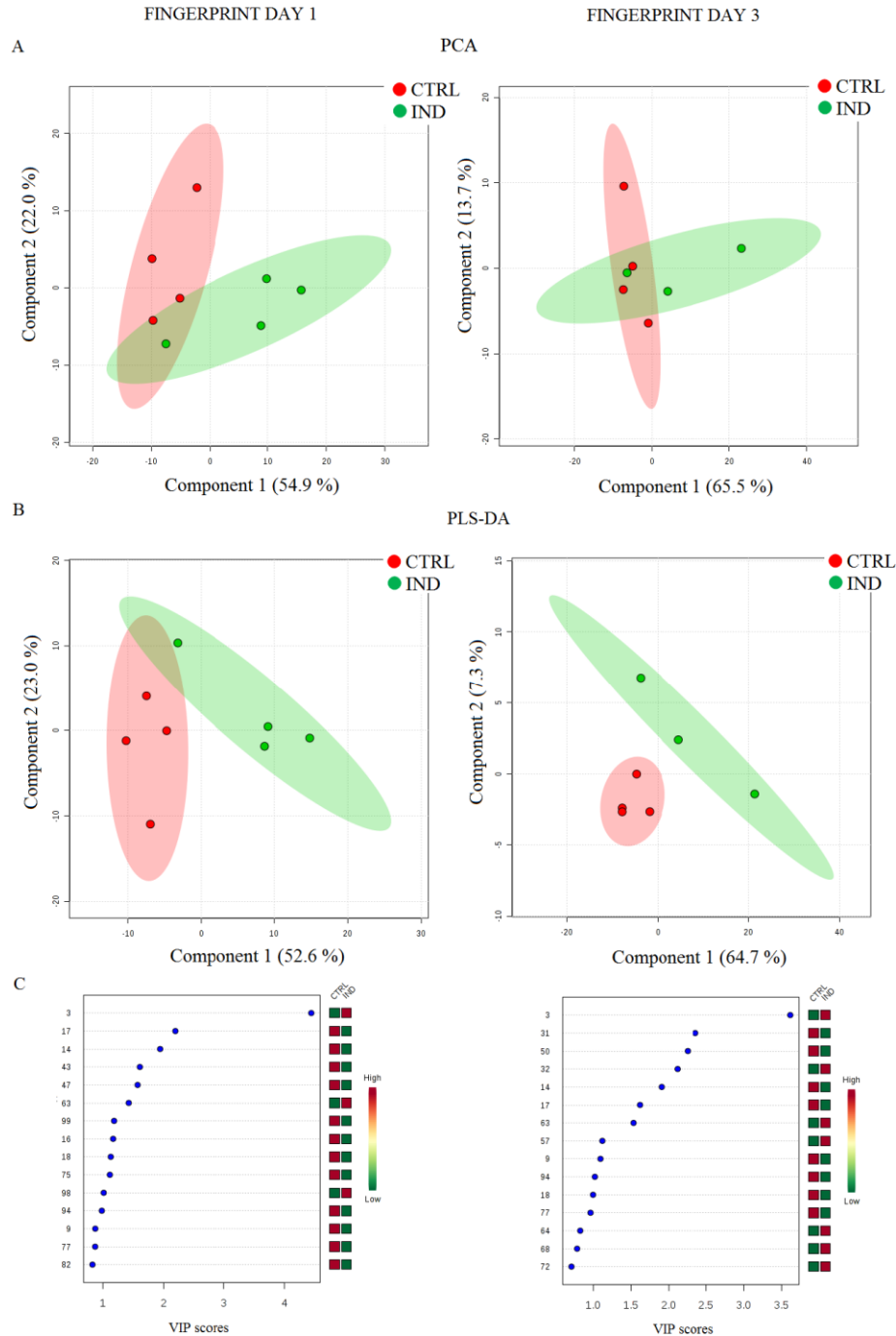


Figure 8: Positive ion mode GC-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 76.9 % and 79.2 % of the total variance from the data set and PLS-DA plots (B) accounting for 75.6% and 72.0% of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (metabolite IDs) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

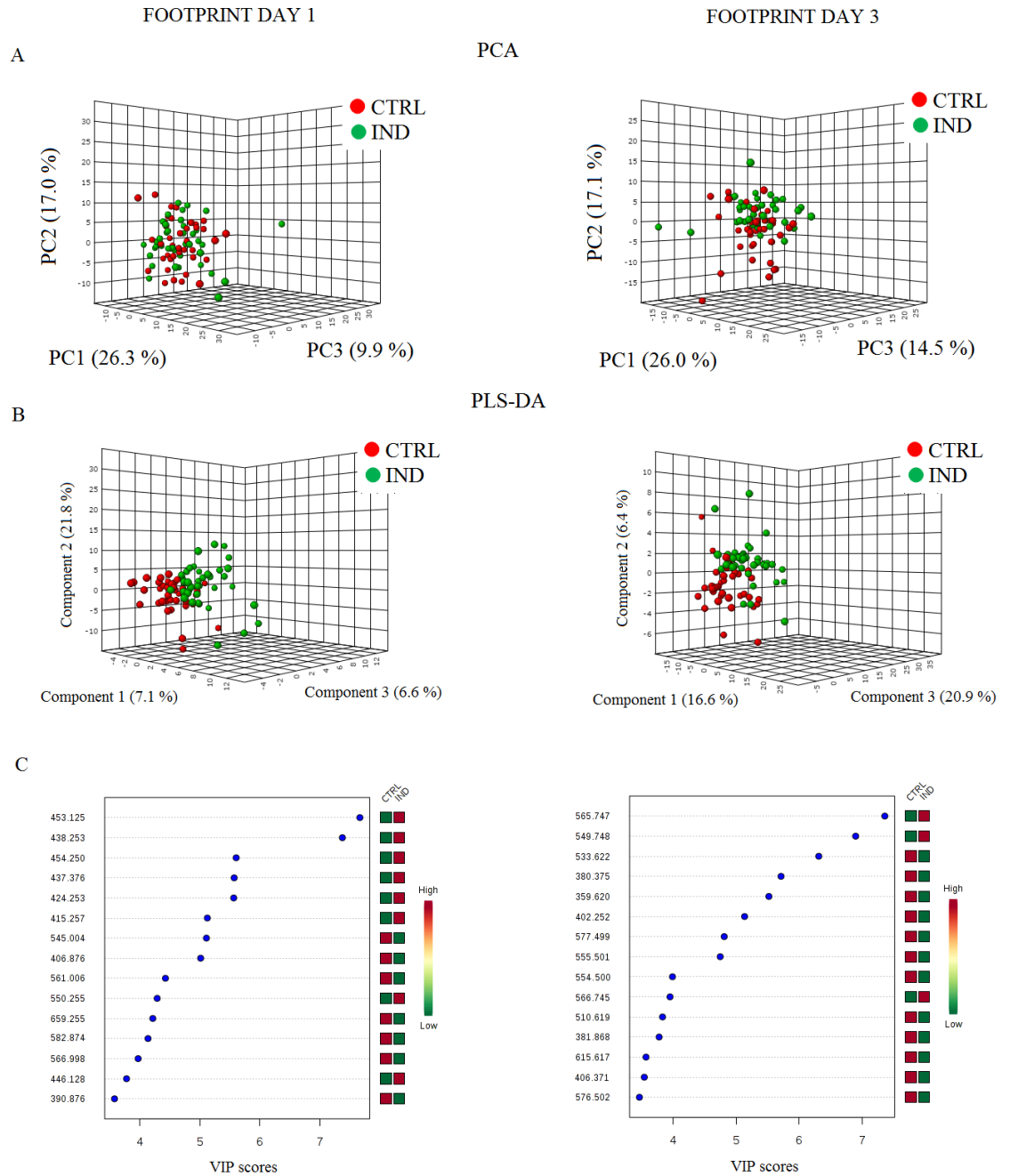


Figure 9: Positive ion mode MALDI-MS metabolite footprint analysis: PCA scores plots (A) accounting for 53.2 % and 57.6 % of the total variance from the data set and PLS-DA plots (B) accounting for 35.5 % and 43.9 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

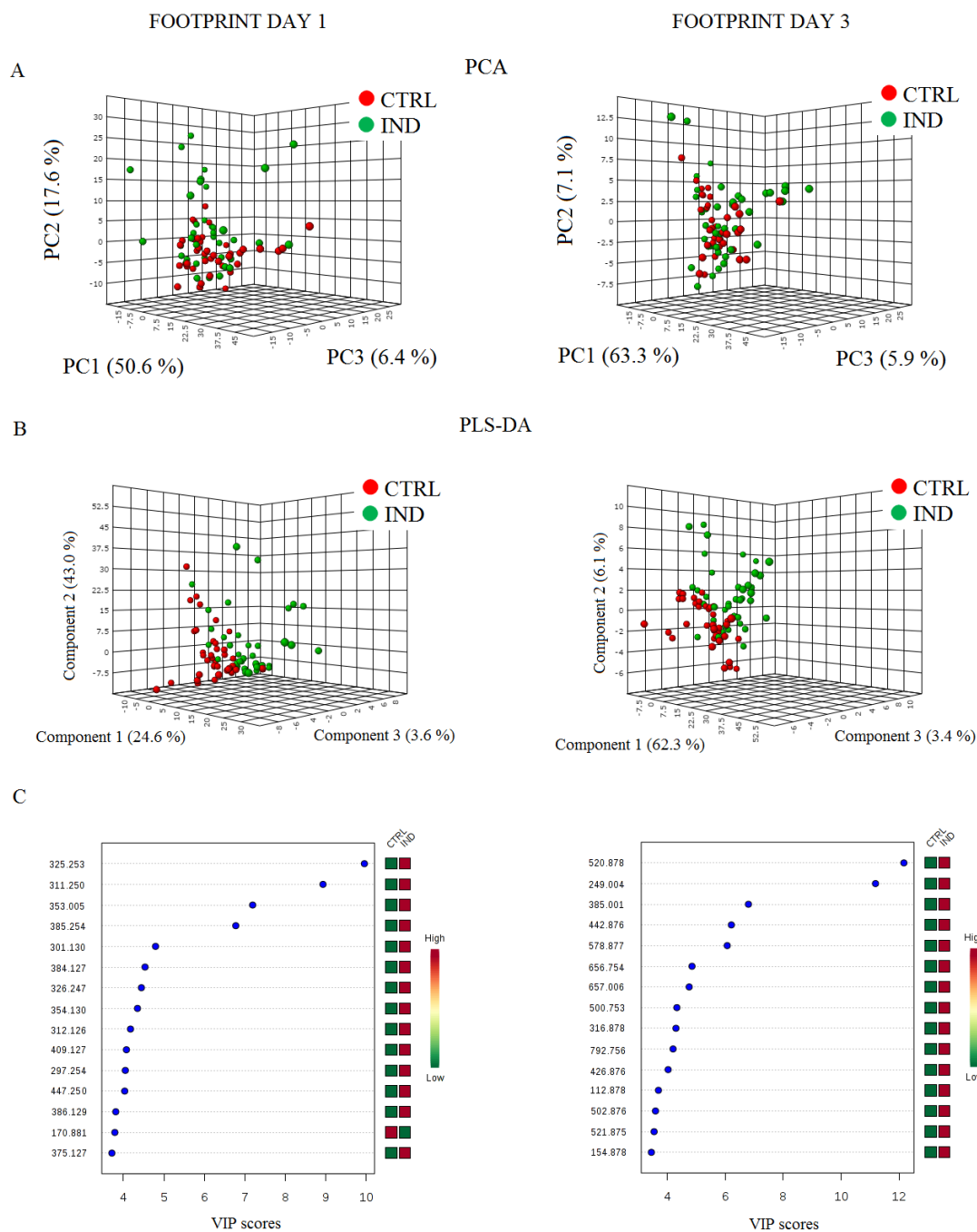


Figure 10: Negative ion mode MALDI-MS metabolite footprint analysis: PCA scores plots (A) accounting for 74.6 % and 76.3 % of the total variance from the data set and PLS-DA plots (B) accounting for 71.2 % and 71.8 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control footprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

Unlike PCA scores plot for footprint profile, 3D-PCA plots generated from MALDI-MS fingerprint analysis **Figure 11 A** and **Figure 12 A** show some level of separation between control and induced samples (for positive ion mode and day 1 samples for negative ion mode) and account for; MALDI+: 53.0 % (day 1 of inducer exposure) and 57.5 % (day 3 of inducer exposure) (**Figure 11 A**); MALDI-: 81.5 % (day 1 of inducer exposure) and 73.4 % (day 3 of inducer exposure) (**Figure 12 A**) of the total variance using three first components.

Generated for MALDI-MS footprint and fingerprint data sets 3D-PLS-DA plots illustrate some level of separation between control and cell-induced samples (**Figure 9-12, B**). In external metabolome, PLS-DA plots suggest little separation between control and cell-induced footprint profiles *ie.* 7.1 % (Component 1) on day 1 and 16.6 % (Component 1) on day 3 for MALDI+ in **Figure 9 B**. PLS-DA plots in MALDI negative ion mode however, show higher level of separation between control and cell-induced footprint profiles *ie.* 24.6 % (Component 1) on day 1 and 62.3 % (Component 1) on day 3 of explained variance (**Figure 10 B**). PLS-DA plots account for; MALDI+: 35.5 % (day 1) and 43.9 % (day 3) and, MALDI-: 71.2 % (day 1) and 71.8 % (day 3) of total explained variance with respect to time of the inducer exposure using first three components.

PLS-DA plots for internal metabolome data suggest that there is a limited separation between control and cell-induced samples for day 1 and day 3 of the inducer exposure in MALDI+, accounting for 19.6 % (Component 1) and 20 % (Component 1), respectively (**Figure 11 B**). A total explained variance from three components between these two groups and on both time points is also shown and corresponds to 48.8 % and 49.9% for day 1 and day 3, respectively. The data analysis of fingerprint profiles in MALDI- shown in **Figure 12 B** illustrates that the variance between control and cell-induced fingerprint profiles accounts for 43.3 % (day 1) and 32 % (day 3) according to Component 1 of PLS-DA model, suggesting good separation of these two. **Figure 12 B** also shows total variance between fingerprint control and cell-induced samples based on Component 1 and 2 that correspond to 76.7 % and 66.0% for day 1 and day 3, respectively.

PCA and PLS-DA plots also show a level of repeatability of samples, which seems to be slightly better for fingerprint measurements (**Figure 11-12, A**) when compared to footprint measurements (**Figure 9-10, A**), where sub-sets are visible outliers.

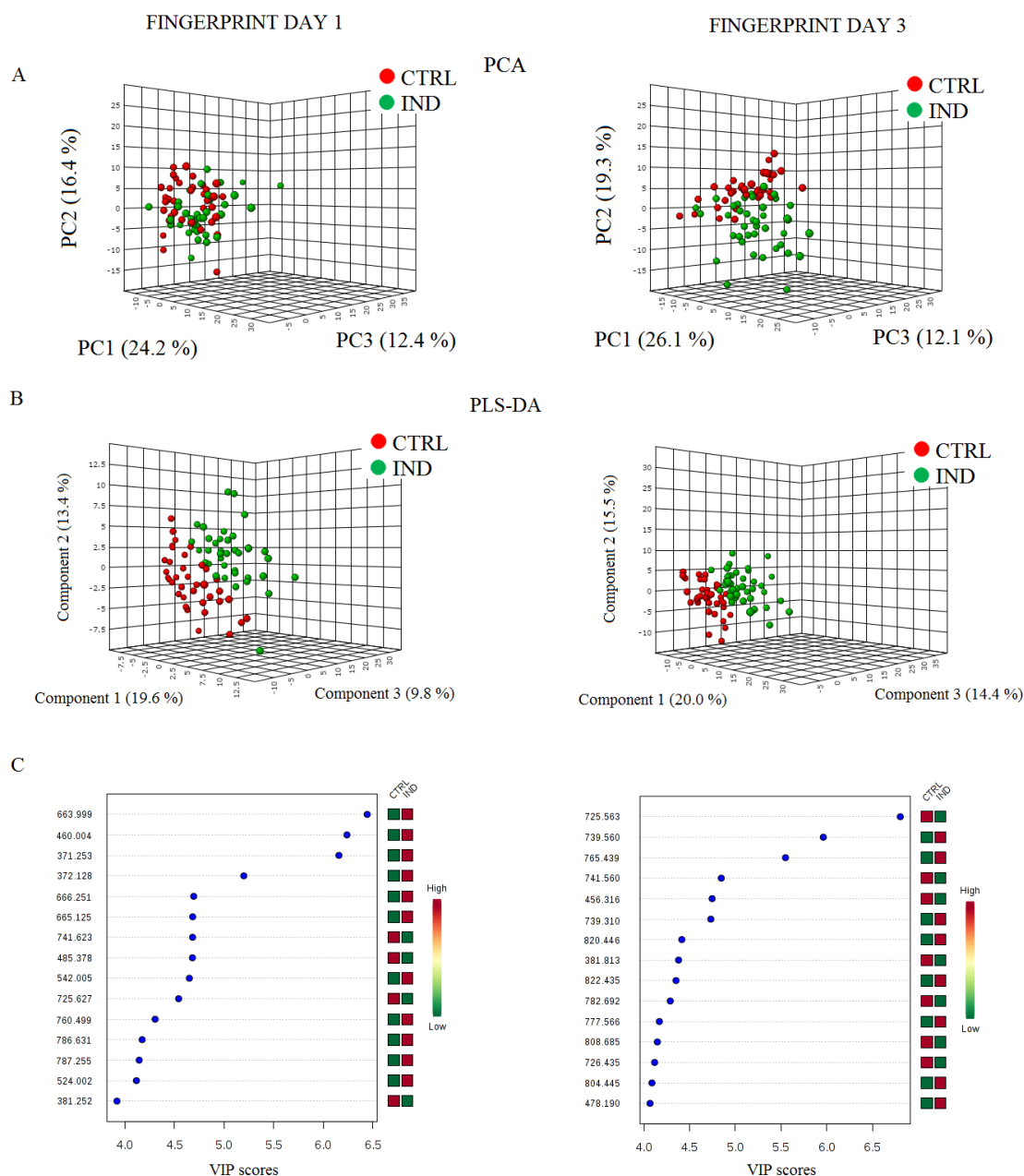


Figure 11: Positive ion mode MALDI-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 53.0 % and 57.5 % of the total variance from the data set and PLS-DA plots (B) accounting for 48.8 % and 49.9 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

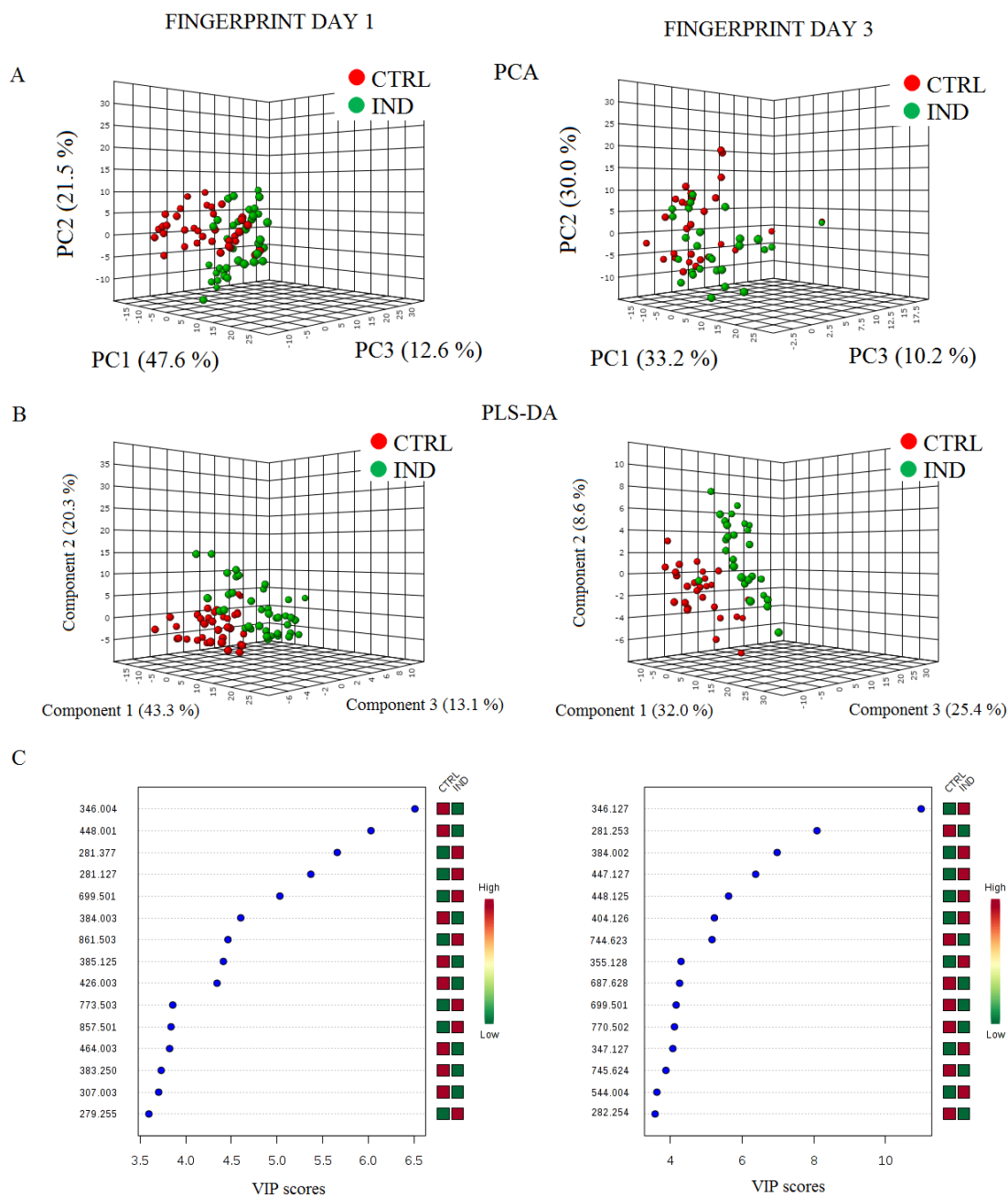


Figure 12: Negative ion mode MALDI-MS metabolite fingerprint analysis: PCA scores plots (A) accounting for 81.5 % and 73.4 % of the total variance from the data set and PLS-DA plots (B) accounting for 76.7 % and 66.0 % of the total variance from the data set on day 1 and day 3 of inducer exposure, respectively. VIP scores from Component 1 (C) show the 15 most statistically significant variables (m/z) in the PLS-DA model. PCA and PLS-DA plots illustrate the variance between control fingerprint samples (CTRL, red) and cell-induced footprint samples (IND, green). Highlighted areas show 95% confidence regions of given samples.

VIP scores for MALDI-MS data from footprint and fingerprint samples from Component 1 shown in **Figure 9-10 and 11-12**, C represent the 15 most statistically significant variables delivered from control and cell-induced samples in the PLS-DA analysis of MALDI+ and MALDI- data, respectively. Here, VIP scores are constructed using m/z values taken from mass spectra of tested samples as variables (Y axis, left) and the VIP scores corresponding to each m/z value (X axis).

For MALDI-MS external metabolome data, VIP scores from day 1 of the inducer exposure of MALDI+ analysis presents 7 out of 15 statistically significant metabolites that are more abundant in control medium samples and another 8 metabolites that are more abundant in cell-induced samples. VIP scores from day 3 of inducer exposure shows another 12 metabolites that are more abundant in control samples and only 3 that are present in cell-induced samples in higher concentrations (**Figure 9 C**). VIP scores from Component 1 from MALDI- analysis illustrate that the majority of statistically significant metabolites are abundant more in cell-induced footprint samples for both days of inducer exposure (**Figure 10 C**).

Generated VIP scores for internal metabolome data from Component 1 shown in **Figure 11 C** present the 15 most significant metabolites found in fingerprint profile in MALDI+. The majority of these metabolites can be observed in high intensities in cell-induced samples: 11 metabolites on day 1 and 8 metabolites on day 3. Similarly, for MALDI-fingerprint analysis, there are 7 metabolites (day 1) and 8 metabolites (day 3) that are abundant more in cell-induced samples (**Figure 12 C**). Additionally, in MALDI-fingerprint profiling, there are 2 possible metabolites that seem to be changing according to exposure time of the inducer: the abundance of m/z 699.501 ion from high in induced cells on day 1 decreases with time and shows low intensity on day 3 in these cells, simultaneously with high abundance in control cells on day 3. The high intensity of another metabolite at m/z 384.0 found in control cells on day 1 seems to be decreasing with time and increasing in cell-induced samples at the same time.

6.5.3. Footprint and fingerprint metabolite identification

The following step of the data analysis requires the metabolites of interest to be identified. The metabolite assignments were completed through investigation against libraries, as outlined in the data analysis section of this chapter (**Section 6.4.**). A total of 23 metabolites within the footprint profile and 31 metabolites within the fingerprint profile have been identified and are shown in **Table 2, 3 and Table 4, 5**. The metabolite identification is made according to a method used to access the metabolic profiles and is based on metabolite IDs and m/z values previously used within the t -test significance test. The metabolite hits with HMBD and abundance in control and cell-induced samples with respect to the inducer exposure time for all metabolites are provided. Additionally, the tables (**Tables 2-5**) define the location of each metabolite within metabolome: footprint (external) or fingerprint (internal).

The number of metabolite assignments is higher for the MALDI-MS technique for both metabolome profiles, when compared to GC-MS hits: 13 and 19 for MALDI-MS footprint and fingerprint, respectively and 10 and 12 for GC-MS footprint and fingerprint, respectively. One of the possible reasons for this might be a different mass range studied using both techniques: samples were analysed within mass range of 40-600 Da using GC-MS and within 20-1000 Da using MALDI. Additionally, both instruments use different ionisation methods *i.e.* MALDI and electron ionisation (EI) for GC-MS. Due to MALDI is considered as a soft ionisation technique, it causes very little or no fragmentation (Hoffmann and Stroobant, 2007), whereas EI, being a hard ionisation method, generates high fragmentation of the analyte, resulting in more complex spectra however, it gives a possibility to metabolite identification by fragments generated (Dass 2006, Halket et al. 2005, Begley et al. 2009). Moreover, the accuracy and sensitivity of both techniques vary, which can also be correlated to the generated data by these two platforms.

The majority of metabolite assignments are accessed for MALDI positive ionisation mode within footprint and fingerprint data. The greater part of more abundant metabolites from external metabolome is found in cell-induced samples (16 out of 23). In contrast to the external metabolome, the majority of metabolites from the internal metabolome with higher abundance is found in control samples (18 out of 31). A similar trend is seen when the time of inducer exposure is considered. The majority of high abundant metabolites are found on day 1 of inducer exposure in external metabolome and on day 3 in internal metabolome. In addition, two metabolites within the external metabolome (**Table 2, 3**) and six metabolites within internal metabolome (**Table 4, 5**) are identified on both days of the

inducer exposure. The percentage of the 15 top statistically significant metabolites (for each technique and each MALDI ion mode) that are successfully identified accounts for ~51 % of the external metabolome and ~69 % of the internal metabolome. Metabolite identification accounts for 67 % of external and 80 % of internal metabolome for GC-MS technique and 43 % of external and 63 % of internal metabolome using MALDI-MS technique. The assignment includes a variety of compounds including amino acids, lipids, carbohydrates, nucleotides and their analogues.

Summarising, the investigation of the inducer treated CHO cells reveals 667 metabolite features after 1 day and 933 metabolite features after 3 days of the inducer exposure from the external metabolome (footprint) (**Table 1**), of which 26 are exclusively assigned (15 after day 1 and 11 after day 3 of inducer exposure) (**Table 2 and 3**). For the internal metabolome (fingerprint), 2067 metabolite features after day 1 and 2202 metabolite features after 3 days of the inducer exposure were detected (**Table 1**), of which 37 are uniquely identified (22 after day 1 and 15 after day 3 of inducer exposure) (**Table 4 and 5**). Accordingly, more significant metabolites are detected after 3 days of the inducer exposure in both external and internal metabolome, suggesting that the cell metabolism changes more considerably with longer exposure to the inducer. Due to a huge number of metabolites that are detected in both metabolic profiles of the protein-induced cells, it would be unrealistic to interpret each single metabolite feature. Therefore, only a part of metabolites are taken into consideration for determination of changes in metabolites in these cells *i.e.* 90 of the most significant metabolites from external and internal metabolome each (shown on VIP scores; 2×15 VIP scores for MALDI+, 2×15 VIP scores for MALDI-, 2×15 VIP scores for GC-MS+).

In addition, according to the number of statistically significant metabolites recognised for footprint and fingerprint profiles, there are 2.6 times more assigned metabolites for internal metabolome, when compared to external metabolome. Consequently, there are 3 times more and 2.3 times more statistically significant metabolites recognised in fingerprint profiles, after day 1 and day 3 of the inducer exposure respectively. This suggests that while there is a great effect of the inducer and the protein production on cellular metabolism, only some part of detectable signal is secreted outside cells into the cell medium.

Table 2: Metabolite features of external metabolome of CHO cells under inducer exposure analysed with GC-MS technique. The table summarises the metabolite IDs, hit and match with GC-MS library and MSI level*, HMDB ID, type of sample, inducer exposure time and abundance: high (↑) or low (↓) in control (CTRL) and cell-induced (IND) extracts. Sorted by inducer exposure time.

Metabolite ID	Hit	Library match, (MSI level)*	HMDB ID	Sample	Inducer exposure time	Abundance	
						CTRL	IND
68	D-Glucose	Possible (Level 2)	HMDB00122	Footprint	Day 1, day 3	↑	↓
88	Heneicosane	Possible (Level 2)	HMDB61782	Footprint	Day 1	↑	↓
					Day 3	↓	↑
18	Glycine_1348_3TMS	Definite (Level 1)	HMDB00123	Footprint	Day 1	↓	↑
15	Propanedioic acid	Possible (Level 2)	HMDB00691	Footprint	Day 1	↑	↓
					Day 3	↓	↑
39	Pyroglutamic acid_1724_2TMS	Definite (Level 1)	HMDB00267	Footprint	Day 1, day 3	↓	↑
21	Phosphate_1372_4TMS	Definite (Level 1)	HMDB01429	Footprint	Day 1	↓	↑
12	Ethanedioic acid	Possible (Level 2)	HMDB02329	Footprint	Day 1	↑	↓
17	2-Butanone	Possible (Level 2)	HMDB00474	Footprint	Day 3	↑	↓
10	Pyruvic acid_1175_1TMS	Definite (Level 1)	HMDB00243	Footprint	Day 3	↓	↑
14	Glycerol	Possible (Level 2)	HMDB00131	Footprint	Day 3	↓	↑

*MSI-Metabolomics Standards Initiative based on library match

Table 3: Metabolite features of external metabolome of CHO cells under inducer exposure analysed with MALDI-MS technique. The table summarises experimental m/z , metabolite hit with HMDB database, mass error, type of sample, inducer exposure time, MALDI-MS ionisation mode and abundance high (↑) or low (↓) in control (CTRL) and cell-induced (IND) samples. Sorted by inducer exposure time and ion mode.



Experimental m/z	Hit	Error [ppm]	HMDB ID	Sample	Inducer exposure time	Ion mode	Abundance	
							CTRL	IND
453.125	3-trans,5-cis-Octadienoyl-CoA	0.2	HMDB02185	Footprint	Day 1	Pos	↓	↑
437.376	1,26-Hexacosanediol	1.1	HMDB36581	Footprint	Day 1	Pos	↓	↑
390.876	DG(22:0/24:1(15Z)/0:0)	1.0	HMDB07616	Footprint	Day 1	Pos	↑	↓
384.127	2-(3-Carboxy-3-(methylammonio)propyl)-L-histidine	2.0	HMDB11654	Footprint	Day 1	Neg	↓	↑
326.247	Eicosapentaenoyl Ethanolamide	4.3	HMDB13649	Footprint	Day 1	Neg	↓	↑
409.127	8-Epidiosbulbin E acetate	0.3	HMDB35110	Footprint	Day 1	Neg	↓	↑
447.25	(4-Hydroxybenzoyl)choline	0.1	HMDB29559	Footprint	Day 1	Neg	↓	↑
375.127	Phaseolic acid	0.6	HMDB31897	Footprint	Day 1	Neg	↓	↑
549.748	Ganglioside GM1 (18:1/22:0)	3.7	HMDB04858	Footprint	Day 3	Pos	↓	↑
359.620	CDP-DG(16:0/22:3(10Z,13Z,16Z))	0.2	HMDB06975	Footprint	Day 3	Pos	↑	↓
555.501	DG(15:0/16:0/0:0)	4.9	HMDB07069	Footprint	Day 3	Pos	↑	↓
381.868	Ginsenoside Rs1 Ginsenoside Rs2	2.4	HMDB39548 HMDB39552	Footprint	Day 3	Pos	↑	↓
249.004	2-Amino-5-phenylpyridine	2.9	HMDB29747	Footprint	Day 3	Neg	↓	↑























Mass error <5 ppm; DG-diglyceride; CDP-DG-cytidine diphosphate diacylglycerol.

Table 4: Metabolite features of internal metabolome of CHO cells under inducer exposure analysed with GC-MS technique. The table summarises the metabolite IDs, hit and match with GC-MS library and MSI level*, HMDB ID, type of sample, inducer exposure time and abundance: high (↑) or low (↓) in control (CTRL) and cell-induced (IND) extracts. Sorted by inducer exposure time.

Metabolite ID	Hit	Library match, (MSI level)*	HMDB ID	Sample	Inducer exposure time	Abundance	
						CTRL	IND
3	Lactic acid	Possible (Level 2)	HMDB03328	Fingerprint	Day 1, day 3	↓	↑
17	Glycine	Definite (Level 1)	HMDB00123	Fingerprint	Day 1, day 3	↑	↓
14	Propanedioic acid	Possible (Level 2)	HMDB00691	Fingerprint	Day 1, day 3	↑	↓
43	Aspartic acid_1581_2TMS	Definite (Level 1)	HMDB06483	Fingerprint	Day 1	↑	↓
47	Glutamine	Possible (Level 2)	HMDB03423	Fingerprint	Day 1	↑	↓
63	Galactose_1854_5TMS	Definite (Level 1)	HMDB00143	Fingerprint	Day 1, day 3	↓	↑
99	Cholesterol	Possible (Level 2)	HMDB00067	Fingerprint	Day 1	↑	↓
75	Hexadecanoic acid_2101_1TMS	Definite (Level 1)	HMDB00220	Fingerprint	Day 1	↑	↓
98	5'-Adenylic acid (adenosine monophosphate, AMP)	Possible (Level 2)	HMDB00045	Fingerprint	Day 1	↓	↑
77	Heneicosane	Possible (Level 2)	HMDB61782	Fingerprint	Day 1, day 3	↑	↓
82	Octadecanoic acid_2292_1TMS	Definite (Level 1)	HMDB00827	Fingerprint	Day 1	↑	↓
50	Pyroglutamic acid_1731_3TMS	Definite (Level 1)	HMDB00267	Fingerprint	Day 3	↑	↓

*MSI-Metabolomics Standards Initiative based on library match

Table 5: Metabolite features of internal metabolome of CHO cells under inducer exposure analysed with MALDI-MS technique. The table summarises experimental m/z , metabolite hit with HMDB database, mass error, type of sample, inducer exposure time, MALDI-MS ionisation mode and abundance high () or low ()  in control  (CTRL) and cell-induced (IND) samples. Sorted by inducer exposure time and ion mode.

Experimental m/z	Hit	Error [ppm]	HMDB ID	Sample	Inducer exposure time	Ion mode	Abundance	
							CTRL	IND
371.253	6-Hydroxypentadecanedioic acid	2.8	HMDB31885	Fingerprint	Day 1	Pos		
372.128	Methyl (Z)-2-decene-4,6,8-trienoate	0.1	HMDB33765	Fingerprint	Day 1	Pos		
665.125	Nicotinamide adenine dinucleotide (NAD ⁺)	1.2	HMDB00902	Fingerprint	Day 1	Pos		
542.005	TG(24:1(15Z)/24:1(15Z)/18:2(9Z,12Z)) TG(24:1(15Z)/20:2n6/24:1(15Z))	2.0	HMDB52142 HMDB52189	Fingerprint	Day 1	Pos		
760.499	CL(18:1(11Z)/18:2(9Z,12Z)/18:2(9Z,12Z)/18:1(11Z)) CL(18:1(9Z)/18:0/20:4(5Z,8Z,11Z,14Z)/18:1(9Z))	4.8	HMDB58120 HMDB58368	Fingerprint	Day 1	Pos		
787.255	Farnesyl pyrophosphate	4.7	HMDB00961	Fingerprint	Day 1	Pos		
346.004	Urothion	2.9	HMDB02377	Fingerprint	Day 1	Neg		
448.001	Adenosine 3',5'-diphosphate ADP dGDP	4.6	HMDB00061 HMDB01341 HMDB00960	Fingerprint	Day 1	Neg		
281.127	Leucyl-Leucine Leucyl-Isoleucine	1.1	HMDB28933 HMDB28932	Fingerprint	Day 1	Neg		
861.503	PGP(18:0/20:3(5Z,8Z,11Z)) PGP(18:0/20:3(8Z,11Z,14Z))	1.9	HMDB13510 HMDB13511	Fingerprint	Day 1	Neg		
699.501	PA(18:1(9Z)/18:1(11Z)) PA(18:0/18:2(9Z,12Z))	4.2	HMDB07864 HMDB07861	Fingerprint	Day 1, Day 3	Neg		

739.560	3-Hexadecanoyloleanolic acid	1.6	HMDB36967	Fingerprint	Day 3	Pos	↑	↓
456.316	PS(18:2(9Z,12Z)/18:0) PS(18:1(9Z)/18:1(9Z))	1.6	HMDB12400 HMDB12390	Fingerprint	Day 3	Pos	↑	↓
777.566	PG(18:1(11Z)/18:0) PG(18:0/18:1(9Z))	2.6	HMDB10617 HMDB10604	Fingerprint	Day 3	Pos	↓	↑
808.685	PE(24:1(15Z)/20:2(11Z,14Z)) PE(24:0/20:3(8Z,11Z,14Z))	0.1	HMDB09759 HMDB09728	Fingerprint	Day 3	Pos	↑	↓
726.435	PS(14:1(9Z)/16:1(9Z)) PS(16:1(9Z)/14:1(9Z))	4.6	HMDB12344 HMDB12364	Fingerprint	Day 3	Pos	↑	↓
355.128	4,8,12,15-Octadecatetraenoic acid (E,E)-11,13-Octadecadien-9-ynoic acid	0.5	HMDB32672 HMDB34382	Fingerprint	Day 3	Neg	↓	↑
687.628	DG(24:0/0:0/18:1n7) DG(18:0/0:0/24:1n9)	1.7	HMDB56114 HMDB56047	Fingerprint	Day 3	Neg	↑	↓
745.624	CE(22:0)	4.1	HMDB06727	Fingerprint	Day 3	Neg	↑	↓

Mass error <5 ppm; TG-triglyceride; CL-cardiolipin; PGP-phosphatidylglycerolphosphate; PA-phosphatidic acid; PS-phosphatidylserine; PG-phosphatidylglycerol; PE-phosphatidylethanolamine; DG-diglyceride; CE-cholesterol ester.

6.5.4. Pathway analysis

The final part of the metabolite analysis process is to attempt to understand and interpret the function of metabolite features assigned within this study and how they can be linked to changes in cell metabolism under the inducer exposure and EPO-Fc production.

As previously mentioned, the expression of EPO-Fc protein is low on day 1 of cell culture and increases remarkably with the growth of cell culture (**Chapter 4, Figure 5**). Consequently, the different external and internal metabolic changes in tetracycline-induced cells according to their incubation time with the inducer, can be a result of a different efficiency of EPO-Fc protein production. Despite the different external and internal metabolic profiles of cells on both incubation times, the EPO-Fc protein production is already noticeable after 24 hours of the inducer exposure. Therefore, the metabolic pathway analysis is performed as a summary of both incubation times (*i.e.* day 1 and day 3), focusing on changes in the protein-producing cells *vs.* control cells. Additionally, the low number of significant metabolites identified on day 1 and day 3 of cell culture when considered separately is not enough to perform the metabolic pathway analysis that gives no hits.

Pathway analysis is applied using statistically significant metabolites ($p < 0.05$) that are identified within the external and internal metabolomes. The pathway analysis is used to investigate cell response and changes in cell metabolism under the exposure of the inducer. **Table 6** shows the metabolic pathways and metabolites that are influenced the most in induced cells according to the impact value calculated for each pathway and considering both external and internal metabolome. Out of all 54 metabolites identified, there are 15 metabolites found to be significant in 16 metabolic pathways of mammalian cells. Consequently, because only 54 metabolites are assigned and only 15 out of these metabolites are correlated to specific metabolic pathways, the analysis of the data become limited.

Figure 13 illustrates a reconstructed biochemical map, summarising correlations between metabolic pathways influenced by the inducer exposure and EPO-Fc production in CHO cells. This summary is based on both external and internal metabolome data from both time points. For instance, purine metabolism only compromises adenosine 3',5'-diphosphate, adenosine diphosphate (ADP), deoxyguanosine diphosphate (dGDP) and adenosine monophosphate (AMP) and not all the intermediates existing in this pathway, as these were not found in the samples. At the same time, changing levels of metabolites within purine metabolism can influence glycine, serine and threonine metabolism as these pathways are correlated to each other.

Table 6: A summary of metabolic pathways and influenced metabolites in cells under the inducer exposure. Data based on PLS-DA footprint and fingerprint profile analysis for both time points; data sorted by impact value*.

Metabolite	Pathway	Sample	Impact*
Glycerol, glucose, galactose	Galactose metabolism	Fingerprint	0.39407
Glycerol	Glycerolipid metabolism	Footprint	0.30018
PA	Glycerophospholipid metabolism	Fingerprint	0.28920
Glycine, pyruvic acid	Glycine, serine and threonine metabolism	Footprint, Fingerprint	0.26884
NAD ⁺	Nicotinate and nicotinamide metabolism	Fingerprint	0.20833
Pyruvic acid	Pyruvate metabolism	Footprint	0.18375
dGDP, ADP, AMP, adenosine 3',5' - diphosphate	Purine metabolism	Fingerprint	0.10086
Pyruvic acid	Glycolysis or Gluconeogenesis	Footprint	<0.1
Pyruvic acid	Citrate cycle (TCA cycle)	Footprint	
Cholesterol, Glycine	Primary bile acid biosynthesis	Footprint, Fingerprint	
Cholesterol, farnesyl pyrophosphate	Steroid biosynthesis	Fingerprint	
PA	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	Fingerprint	
Glucose	Starch and sucrose metabolism	Footprint	
Pyruvic acid	Cysteine and methionine metabolism	Footprint	
Glycine, pyroglutamic acid	Glutathione metabolism	Footprint	
Cholesterol	Steroid hormone biosynthesis	Fingerprint	

*Impact value – a cumulative percentage from the matched metabolites

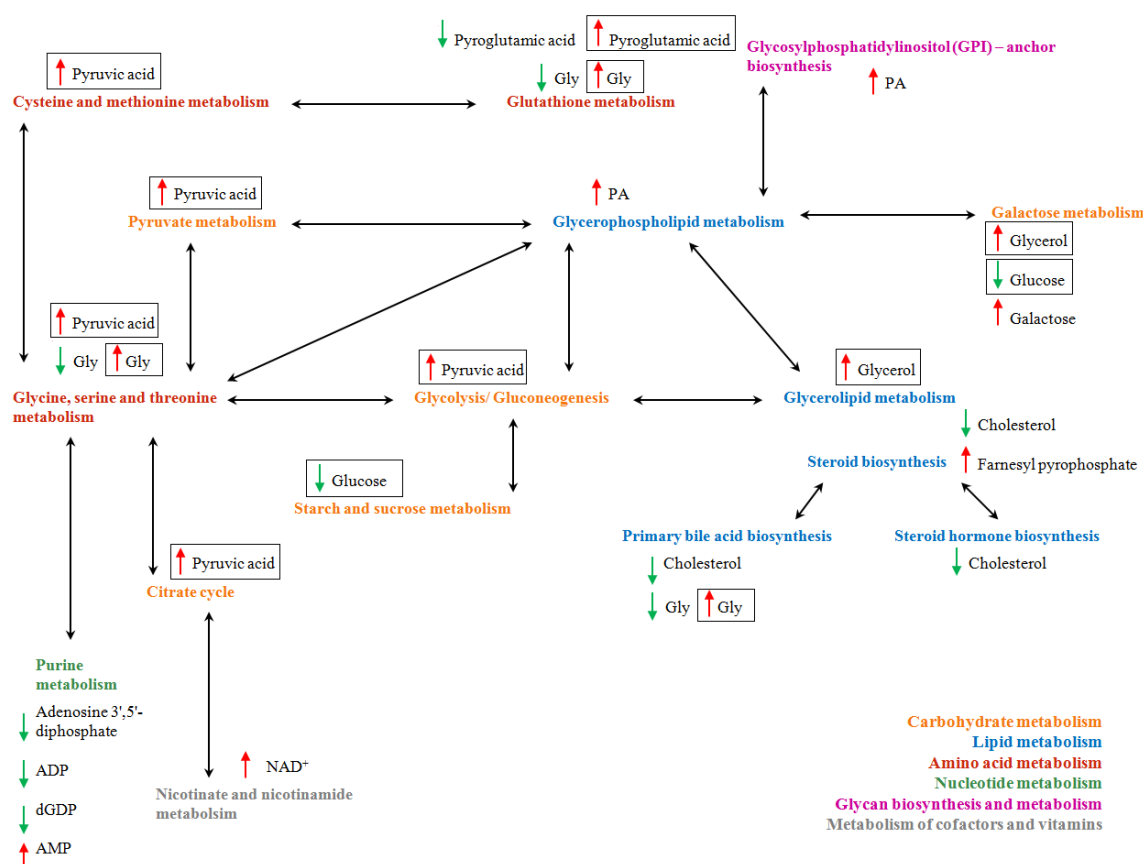


Figure 13: Reconstructed biochemical map summarising correlations between metabolic pathways influenced by the protein production in transfected CHO cells. The summary includes external and internal metabolome data from both time points of cell culture. Metabolic pathways are divided into 6 main metabolism paths (colour-coded). Metabolites influenced by the protein production process are highlighted with visualised abundance: high (↑) and low (↓). Two arrows with the same metabolite indicate changing abundance depending on the type of metabolome (external or internal). Metabolites placed in boxes are found in external (footprint) metabolome.

According to the KEGG pathway database (<http://www.genome.jp/kegg/pathway.html>), 16 metabolite pathways are found to be affected by the protein production and are divided into 6 main groups (colour-coded), correlated to general metabolic pathways. Metabolites identified using *t*-test that are found to be influenced by the inducer exposure are also highlighted. Arrows visualise the abundance (high or low) of each metabolite in induced cells. Also, two arrows next to the same metabolite indicate changing abundance depending on the type of metabolome (external or internal). Metabolites placed in boxes are found in external (footprint) metabolome. In order to highlight the origin of influenced

metabolites by EPO-Fc protein production, a diagram illustrating the identified compounds and their metabolic routes with a division to external and internal metabolome is also shown (**Figure 14**). The summary includes data from both time points of the cell culture and the top metabolic pathways that are influenced by the inducer exposure and the protein production (impact value >0.1 according to **Table 6**). Metabolites influenced by protein production are also highlighted with visualised abundance (high or low). The number of metabolites recognised as significant for the metabolic pathways is somewhat disappointing and might show a limiting factor in the investigation of the results.

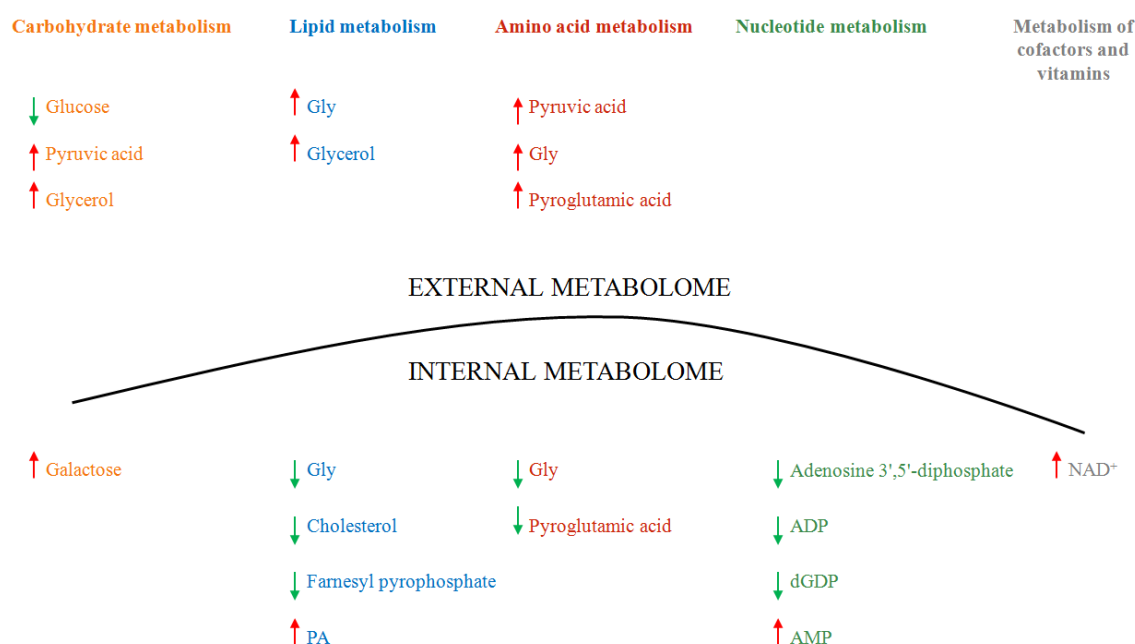


Figure 14: Diagram of identified metabolites and their metabolic pathways (colour-coded) with a division to external and internal metabolome. The diagram includes data from both time points of cell culture and metabolic pathways with impact value >0.1. Metabolites influenced by the protein production are highlighted with visualised abundance: high (↑) and low (↓). Two arrows with the same metabolite indicate changing abundance depending on the type of metabolome (external or internal).

It is evidently seen from the results discussed that the inducer influences on metabolome of cells during the production of EPO-Fc. Therefore, it is possible to couple *a priori* literature based information with the present outcome to establish an understanding of the changes in metabolism of CHO cells producing EPO-Fc fusion protein.

The most influenced by the inducer exposure and the protein production metabolic pathways (impact value >0.1 according to **Table 6**) are galactose metabolism, glycerolipid metabolism, glycerophospholipid metabolism, glycine, serine and threonine metabolism, nicotinate and nicotinamide metabolism, pyruvate metabolism and purine metabolism. They belong to different, main metabolism groups: carbohydrate metabolism (galactose and pyruvate metabolism), lipid metabolism (glycerolipid and glycerophospholipid metabolism), amino acid metabolism (glycine, serine and threonine metabolism), metabolism of cofactors and vitamins (nicotinate and nicotinamide metabolism) and nucleotide metabolism (purine metabolism).

The objective of metabolomic investigations therefore, is to analyse as many as possible of the metabolites present in a given sample and correlate their activity with metabolic pathways and specific physiological state (Dunn, Bailey and Johnson, 2005) such as EPO-Fc protein production.

6.5.4.1. Mitochondrial response

Mitochondria are important cell compartments that are responsible for cellular differentiation, cell cycle control as well as cell growth maintenance. A dominant role of these organelles however, is to produce energy in a form of adenosine triphosphate (ATP) through a number of reactions including a variety of compounds such as glucose and pyruvic acid (Klyszejko-Stefanowicz, 2002). Increased levels of pyruvic acid and decreased levels of glucose are found in external metabolome of cells producing EPO-Fc protein. A formation of pyruvic acid (pyruvate) occurs from glucose molecules through glycolysis process (Klyszejko-Stefanowicz, 2002) hence, the lower abundance of glucose in induced cells may be correlated to enhanced glycolysis route.

Glycolysis is one of the main biochemical reactions existing in cells, taking place before another important cycle of reactions *i.e.* citrate cycle, also known as tricarboxylic acid (TCA) cycle ongoing in mitochondria. Both pathways are closely related: during glycolysis, glucose is enzymatically converted into pyruvate by pyruvate kinase, which also transfers a phosphate group to ADP, resulting in energy formation in a form of ATP. The process of glycolysis also requires high levels of NAD^+ used for a formation of nicotinamide adenine dinucleotide reduced form (NADH). Then, once pyruvate is created, it is further transformed to acetyl-CoA and is introduced to the TCA cycle. TCA cycle involves multiple reactions, where the energy stored in form of carbohydrates, lipids and proteins is converted to chemical energy in a form of adenosine triphosphate (ATP), precursors of some amino acids and NADH (**Figure 3**) (Klyszejko-Stefanowicz, 2002).

In the internal metabolome of induced cells, higher abundance of NAD^+ and lower abundance of some nucleotides *i.e.* adenosine 3',5' -diphosphate is observed, when compared to control cells. This may confirm enhanced cell activity and simultaneous maintenance of intercellular balance between metabolites, where some TCA cycle intermediates are formed in order to be precursors for the others.

Another possibility for pyruvate to be processed is to enter the route of so called 'anaerobic glycolysis'. Reactions during anaerobic glycolysis occur in the absence/partial absence of oxygen and are ongoing in bacteria. However, anaerobic glycolysis is also possible in mammalian cells under hypoxic conditions. In these conditions (hypoxia), pyruvate is converted to lactic acid (lactate) with a release of NAD^+ derived from NADH (Gerriets & Rathmell 2012, Pearce & Pearce 2013, Donnelly & Finlay 2015, Wang *et al.* 1976, Lane *et al.* 2009). In this study, increased levels of lactic acid are detected in internal metabolome of cells producing EPO-Fc that would suggest that cells might undergo a process of anaerobic glycolysis. Similar interpretation of the data was made on antibody-producing CHO cell system, where the hypothesis established that higher lactate concentrations in the internal metabolome could be a result of an adaptation to lack of TCA cycle intermediates. As a consequence, glucose is redirected to alternative metabolic pathways (Sellick *et al.*, 2011). Enhanced rates of anaerobic glycolysis and therefore greater levels of lactic acid can also be caused by the activation of B lymphocytes that was described previously and was correlated to increased expression of the glucose transporters on the cell surface (Donnelly and Finlay, 2015).

One more assumption with regard to high pyruvate as well as glycerol levels in the external metabolome of cells producing EPO-Fc protein can be made based on the correlation between cell growth phase and concentrations of glycolysis and TCA cycle metabolites. It was shown that the shift from exponential to stationary growth phase and full stationary growth phase in CHO cells were marked with exhaustion of pyruvate, glutamate and glycerol metabolites. Also, during both exponential and stationary phases, lactate production inside cells was also detected (Sellick *et al.* 2011, Taschwer *et al.* 2012). Therefore, enhanced levels of pyruvate and glycerol imply that CHO cells stay in exponential phase for at least three days of cell culture, actively producing EPO-Fc.

During anaerobic glycolysis, the ratio of NAD^+/NADH alters, where molecules of NAD^+ are regenerated (Wang *et al.* 1976, Lane *et al.* 2009). In this study, enhanced levels of NAD^+ are present in internal metabolome of induced transfected CHO cells, which supports the hypothesis that cells might undergo a shift to anaerobic glycolysis. Not only can anaerobic glycolysis be caused by the cells adaptation to lack of TCA cycle

intermediates, but also it can be a result of maintaining a regeneration of NAD^+ taking part in the intracellular reduction-oxidation balance between NADH/NAD^+ (Sellick *et al.*, 2011). The balance between NADH and NAD^+ is an important factor for cell function and plays an important role in many biological processes such as antioxydation and oxidative stress, calcium homeostasis, aging, carcinogenesis or energy metabolism and mitochondrial functions (Ying, 2008). Therefore, the relationship between NADH/NAD^+ has been largely investigated for bacteria as well as cancer cells (Koeblmann *et al.* 2008, Briehl *et al.* 2014, Matthews *et al.* 2012, Cho *et al.* 2014, Hugenholtz *et al.* 2000).

In cells producing EPO-Fc protein, decreased levels of glucose in medium can be also correlated to the increased metabolic activity of these cells, resulting in greater consumption of glucose present in the cell medium. The TCA cycle is weakly active when cells are in exponential growth phase with glucose present in growth medium. It is when cells become activated to be more effective, glucose from medium starts to be consumed greatly (Liebeke *et al.* 2011, Wang *et al.* 1976). Therefore, due to enhanced TCA cycle activity, the decreased concentrations of glucose in medium of induced cells can be seen. Similar trends of increased uptake of glucose in metabolically active cells have been shown multiple times in previous studies of CHO cells producing monoclonal antibody (Correa *et al.* 2016) as well as lymphocytes undergoing differentiation into different cell types (Donnelly & Finlay 2015).

An increased abundance of galactose in cells producing EPO-Fc is also observed. Because galactose can be converted into glucose (Sunehag *et al.*, 2002) that may therefore suggest, the production of this monosaccharide is enhanced in order to supply a cell with increased need for glucose. On the other hand, high galactose concentrations may be a result of ongoing post-translational modifications of EPO-Fc *i.e.* glycosylation (see **Chapter 1**). Glycosylation is an extremely important for protein structural integrity and its therapeutic efficacy process. This mechanism, in which a carbohydrate is attached to a target molecule *i.e.* lipid or protein (Jefferis 2005) has been widely investigated by researchers (Wacker *et al.* 2011, Gramer *et al.* 2011, Ye *et al.* 2011, Dent *et al.* 2016). Glycosylation has also been examined with regard to cell culture pH in CHO cells producing monoclonal antibody. It was shown that a pH downshift in the stationary phase of cell growth caused significant changes of galactose content in glycosylation profiles of produced antibody (Xie *et al.*, 2016).

6.5.4.2. Nucleotide response

As seen in **Figure 13**, changes within nucleotide metabolism in tetracycline-induced cells are also observed. The decreased abundance of adenosine 3'5' – diphosphate, ADP and dGDP within cells producing EPO-Fc is observed and can be correlated to the higher activity of these cells. Adenosine 3'5' –diphosphate is a nucleotide closely related to ADP (has phosphate group at 5' position instead pyrophosphoric acid) (Klyszejko-Stefanowicz, 2002). As previously stated, during glycolysis when glucose is converted into pyruvate, the enzyme catalyzing this conversion transfers a phosphate group to ADP, resulting in a release of ATP. Consequently, decreased levels of ADP and adenosine 3'5' –diphosphate imply enhanced formation of ATP hence, greater glycolysis rates in induced cells. This hypothesis may also be confirmed by the increased levels of AMP detected in internal metabolome of induced transfected CHO cells. AMP can be a substrate for ADP in a reaction with phosphate group as well as it can be a precursor for ADP formation (in a reaction with ATP), required for glycolysis process (Klyszejko-Stefanowicz, 2002). Moreover, low abundance of dGDP also verifies enhanced metabolic state of induced cells. dGDP is a derivative of GDP (lack of hydroxyl group). During the TCA cycle, GDP in a reaction with phosphate group gives high energy GTP nucleoside that is further used in the synthesis of RNA and plays essential role in cell signalling (Klyszejko-Stefanowicz, 2002).

6.5.4.3. Amino acid response

The reconstructed biochemical map highlights that external and internal amino acid abundance is affected by EPO-Fc production in transfected CHO cells. In the medium of these cells, glycine and pyroglutamic acid concentrations are increased, whereas the same metabolites with low abundance in internal metabolome are detected. Additionally, cellular levels of aspartic acid and glutamine are also decreased.

Glutamine can be converted to glutamic acid, which is further converted to pyroglutamic acid, an uncommon amino acid derivative found in cells (Klyszejko-Stefanowicz, 2002). Hence, low internal glutamine concentrations can be a result of a formation of pyroglutamic acid. Pyroglutamic acid however, is then pumped outside a cell, where higher abundance of this amino acid is detected in medium (and low in internal metabolome), when compare to non-induced cells. This suggests that too high internal concentration of pyroglutamic acid is not favourable for a cell and might cause “metabolic acidosis” that has been widely studied with regard to genetic disorders and metabolism changes after paracetamol exposures (Kumar & Bachhawat 2012, Liss *et al.* 2013, Hunter

et al. 2016). Moreover, glutamine has been found to act as an energy source that can be converted to pyruvate or lactate, particularly in cancer cells (Li *et al.* 2017, Zhang *et al.* 2017). It suggests that the protein producing cells desire more intermediates for TCA cycle for their enhanced metabolism. On the other hand, it has been also acknowledged that transfected CHO cell culture has been grown in glutamine-enriched medium and so, the intake of the amino acids into cells might be complicating the observed metabolic response.

Lower levels of aspartic acid are detected in the internal metabolome of induced cells when compare to control cell cultures. Aspartic acid is known to be a precursor for several amino acids such as methionine, lysine, threonine and isoleucine, and hence is involved in protein biosynthesis. Also, it can participate in the generation of glucose (in a process called gluconeogenesis) that is further used in glycolysis for energy production (Klyszejko-Stefanowicz, 2002). Consequently, this confirms that induced cells are more metabolically active and require enhanced synthesis of bio-elements to maintain desired EPO-Fc production. Also, it is in line with increased up-take of glucose from cell medium in induced cell culture observed in this study. Similar results with decreased amino acid content in internal metabolome of CHO cells producing antibodies have been published (Sellick *et al.*, 2011). Decreased levels of glycine are also detected in internal metabolome of induced cells that suggests enhanced protein and lipid synthesis. Glycine is a simple amino acid that acts as a precursor for a number of proteins and is involved in phospholipids production (Honda *et al.* 2016, Al Suleimani & Al Mahruqi 2017, Klyszejko-Stefanowicz 2002). Increased export of glycine outside cells found in this study is, however, not fully understood.

6.5.4.4. Lipid response

Despite a relatively high number of pathways recognised to be influenced by the inducer exposure and EPO-Fc production, the number of metabolites linked to each pathways is relatively low. This however, can be a result of matching only primary metabolites and not including secondary metabolites into the analysis of metabolic pathways.

Primary and secondary metabolites are closely correlated to each other. Secondary metabolites are non-fundamental for the normal growth and development of the organism and require many genes to be expressed. These metabolites are created by pathways derived from primary metabolic routes and are often of poorly known function but can play a role, for instance, a signal compound that enhance survival of some cells within a given cell population (Roessner, 2006).

Here, some secondary metabolites such as ginsenoside Rs1 and ginsenoside Rs2 are identified (**Table 3**). Ginsenoides are believed to have antioxidant properties by increasing the levels of internal antioxidant enzymes *i.e.* glutathione (GSH) and its oxidized form (GSSG) that are required to prevent cellular damage that may occur during oxidative stress (Klyszejko-Stefanowicz, 2002). Oxidative stress, during which reactive oxygen species are produced, can influence the regulation of a cell growth and its function (Taschwer *et al.*, 2012). Decreased release of ginsenoides to growth medium of cells producing the protein is observed in this study. Lower abundance of these compounds in growth medium may suggest their increased intracellular consumption in order to raise the concentrations of GSH to limit cell destruction. This assumption is in line with another observation, where internal pyroglutamic acid levels are decreased in cells producing EPO-Fc. As mentioned previously, pyroglutamic acid is an uncommon amino acid that is produced from glutamic acid conversions (Klyszejko-Stefanowicz, 2002). Therefore, low intracellular concentration of pyroglutamic acid may suggest an increased consumption of glutamic acid that is known to be an important precursor for GSH synthesis (Taschwer *et al.*, 2012). Another suggested mechanism of the activity of ginsenoides is the interaction of cell membranes thus, changes in their properties (Leung & Wong 2010, Lu *et al.* 2009). Decreased levels of ginsenoides are detected in footprint profiles of induced transfected CHO cells that can suggest lower secretion of these compounds outside the cells. Consequently, this may imply that the membrane of cells producing EPO-Fc changes and there may be an increased concentration of reactive oxygen species in cells exposed to the inducer activity. This finding would be in line with another observation, where lipid species in both footprint and fingerprint profiles are detected (**Table 3, 5 and Figure 9**). The majority of lipid species from different class are found in internal metabolome: triglycerides (TG), cardiolipins (CL), phosphatidylglycerolphosphate (PGP), phosphatidylserines (PS), phosphatidylglycerol (PG), phosphatidylethanoalamines (PE), diglyceride (DG) and cholesterol esters (CE) (**Table 5**). The abundance of some of these lipids is increased (*e.g.* triglycerides or cardiolipins) and the abundance of others is decreased (*e.g.* phosphatidylserines or phosphatidylethanoalamines) in cells producing EPO-Fc protein. Apart from hypothesis of increased levels of reactive oxygen species, the changes in lipid content of cells producing EPO-Fc may be correlated to the secretion of this protein into cell medium.

In cells, proteins are transported outside the cell by an exocytosis process. Exocytosis is a type of an active transport, where a cell transports molecules outside a cell membrane by expelling them into the exterior. Exocytosis is a process that involves fusion of a vesicle

containing transported protein with the cell membrane. The fusion of a vesicle with a cell membrane causes an increase of a cell membrane surface, by the surface of the fused vesicle (Roman-Fernandez & Bryant 2016, Bruns & Jahn 2002). Consequently, this can suggest that the vesicle-cell membrane interaction may cause changes and different distribution of lipids present in a membrane of cells producing EPO-Fc. It also implies that the size and morphology of these cells may be changed by increased surface of cells.

A cell membrane consists of proteins and different types of lipids that create a lipid bilayer (**Figure 4**). Lipids and lipid-like precursors for lipid creation found in a cell membrane, belong to diverse classes such as phospholipids, glycolipids, sterols and glycerolipids, where the most abundant are phospholipids (Ali *et al.* 2017, Di Bartolomeo *et al.* 2017). The majority of lipids detected in the metabolome of cells producing EPO-Fc that differ from lipid content of non-induced cells are hence a part of phospholipid groups: phosphatidylglycerolphosphate (PGP), phosphatidylserine (PS), phosphatidylglycerol (PG), phosphatidylethanolamine (PE) and cardiolipin (CL). Other classes of lipids are also found: glycolipids (ganglioside), sterols (cholesterol ester (CE), cholesterol) and glycerolipids (triglyceride (TG), diglyceride (DG)).

PS, PE, DG and CE lipids show decreased abundance while TG, CL, PGP, phosphatidic acid (PA), PG show increased abundance in internal metabolome of induced cells.

The decrease of DG can be correlated to a production of TG, which increased levels are detected in cells producing EPO-Fc protein. DG acts as a substrate for addition of the third fatty acid to create TG. The function of TG is to provide the energy source and storage as well as lipid transport (Vasconcellos *et al.*, 2016). Therefore, the increased levels of TG may be a result of higher demand of cells for lipid production and raised activity of cell metabolism.

This is in line with other findings within this study, where lower levels of PS and PE are detected in induced cells, which may suggest that cells activate the synthesis of lipids. The synthesis of lipids can be due to increased or changed cell metabolism as well as due to ongoing exocytosis process, where lipids are used for vesicle creation in order to transport the protein outside a cell (Bhattacharya *et al.* 2016, Flis *et al.* 2015).

Moreover, decreased levels of both cholesterol and CE in cells producing the protein are observed, when compare to non-induced cells. Cholesterol and cholesterol esters do not contribute to cell membranes directly but are an essential component of lipids creating bilayers and act as a precursor for hormones (Klyszejko-Stefanowicz, 2002). This may also confirms that a synthesis of lipids can be significantly increased in tetracycline-induced cells.

Another correlation between CL, PG and PGP lipids can be found, where their abundance in induced cells is increased. PGP and PG can be synthesised from cytidine diphosphate diacylglycerol (CDC-DG) as an intermediate that reacts, indirectly, with glycerol. PGP and PG are then used for the synthesis of CLs. In mammalian cells, cardiolipins are a crucial component of the inner mitochondrial membrane, where they maintain the optimal energy metabolism of a cell and function of this organelle. It is also known that CLs are large in number in metabolically active cells (Tatsuta & Langer 2017, Martensson *et al.* 2017). Hence, this may suggest that induced cells require more energy to maintain protein formulation therefore, the increased production of mitochondrial lipids is desired. Increased levels of mitochondrial lipids indirectly generated from CDC-DG may be a cause for lower abundance of CDC-DG in medium of induced cells, when compared to control cells. This dependency implies that a reduced extraction of CDC-DG into medium of induced cells results by the increased metabolic activity of cells hence, enhanced consumption of CDC-DG. Increased levels of glycerol and ganglioside GM1 (18:1/22:0) (glycolipid) are detected in external metabolome of tetracycline-induced transfected CHO cells. Lipids, apart from maintaining a membrane integrity and stability as well as acting as the energy source and storage, are crucial components in cell signalling (Vasconcellos *et al.*, 2016). Cell-to-cell interactions are important for cellular responses, function, recognition and growth. Glycolipids are the main cell membrane components that maintain cellular signalling and are localised on the outer surface of a cell membrane. Thanks to a presence of a carbohydrate or glycerol attached to a lipid part, glycolipids serve as recognition sites for cell-to-cell communication (Shvets *et al.* 2015, Villas-Bôas S.G. *et al.* 2007, Owen & Gaus 2013). Enhanced abundance of glycolipids in medium of induced cells can imply that cells release some parts of their membrane into their exterior, while transferring the protein outside a cell during an exocytosis process, which can also have cell-signalling function. However, such assumptions have to be investigated further to elucidate wider knowledge of the specific metabolism changes caused by EPO-Fc fusion protein production in CHO cells. A collection of additional metabolites that are changed under the protein production in tetracycline-induced CHO cells is also observed (**Table 2-5**). These metabolites belong to a variety of other groups of metabolites for example cofactors (*e.g.* urothion) and are intermediates involved in metabolic pathways, cell signalling, cell aggregation and many others (<http://www.hmdb.ca/>). Nevertheless, their specific role and function within tetracycline-induced cells metabolism need to be further explored.

6.6. Conclusions and future work

The application of a non-targeted metabolomics approach for the analysis of metabolite changes in tetracycline-induced CHO cells with EPO-Fc fusion protein secretion can be considered successful. The employment of GC-MS and MALDI-MS techniques has provided information about the response of cells to the protein production and changes in key metabolites within internal and external cell metabolomes. These techniques have also been complementary to each other, providing different types of information. Data have shown an extended effect of EPO-Fc production on cell metabolism.

According to the multivariate analyses, a great number of metabolites was identified to be significantly influenced by the protein induction when compared to non-induced cells. Correlation of these significant metabolites and reconstruction of the metabolite map, classified metabolome changes into metabolic pathways influenced by the production of EPO-Fc. A differentiation between metabolites derived from internal and external metabolome was also seen. Thanks to the pathway map and the correlation analysis it was possible to observe individual trends of metabolites within metabolic routes under the protein production in CHO cells. Biological interpretation of a list of metabolites highlighted five main metabolic pathways the most influenced by EPO-Fc production: carbohydrate, lipid, amino acid, nucleotide metabolisms and metabolism of cofactors and vitamins.

The findings indicate that changes observed in EPO-Fc producing cells are related to enhanced protein and lipid synthesis highlighting that these cells are in a state of increased metabolic activity with the protein exocytosis into growth medium. Results also suggest that some lipid changes can be correlated to mitochondrial membranes specifically. Moreover, the composition of lipid bilayer of induced cells seems to be different to non-induced cells. Thus, it implies that the targeted investigation of the lipids would deliver further knowledge of the effect of the fusion protein production on cell metabolism. Results also suggest that, due to the increased metabolic activity, cells may adapt to the lack of crucial substrates for energy-producing pathways, but still support the protein production. Also, the process of EPO-Fc glycosylation remains active.

With the above findings and insights into CHO metabolism under the fusion protein production, targeted metabolite profiling studies could be applied to further explore the influence of EPO-Fc on these cells. This would help with the detection of a larger number of TCA cycle intermediates as well as lipid species therefore, expanding and validating correlations between specific pathways. Also, the exact identification and quantification of these metabolites would validate the biochemical interpretations of this study.

In addition to metabolome changes discussed above, this study has shown a comparison of two techniques, giving complementary information of metabolism of induced CHO cells. Moreover, MALDI-MS experimental conditions for the detection of metabolites from external and internal metabolome have been verified. The MALDI-MS method used in this study was fast and reproducible, giving good quality spectra, with good mass resolution and rich in biological information.

References:

- Ali, A. H., Zou, X., Lu, J., Abed, S. M., Yao, Y., Tao, G., Jin, Q. and Wang, X. (2017) 'Identification of phospholipids classes and molecular species in different types of egg yolk by using UPLC-Q-TOF-MS', *Food Chem*, 221, pp. 58–66.
- ATCC (2014) *Animal Cell Culture Guide*. Manassas, VA. Available at: www.atcc.org.
- Di Bartolomeo, F., Wagner, A. and Daum, G. (2017) 'Cell biology, physiology and enzymology of phosphatidylserine decarboxylase', *Biochim Biophys Acta*, 1862, pp. 25–38.
- Begley, P., Francis-McIntyre, S., Dunn, W. B., Broadhurst, D. I., Halsall, A., Tseng, A., Knowles, J., Goodacre, R. and Kell, D. B. (2009) 'Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum', *Anal Chem*, 81, pp. 7038–7046.
- Bhattacharya, S., McElhanon, K. E., Gushchina, L. V and Weisleder, N. (2016) 'Role of phosphatidylinositol-4, 5-bisphosphate 3-kinase signaling in vesicular trafficking', *Life Sci*, 167, pp. 39–45.
- Bouhifd, M., Beger, R., Flynn, T., Guo, L., Harris, G., Hogberg, H., Kaddurah-Daouk, R., Kamp, H., Kleensang, A., Maertens, A., Odwin-DaCosta, S., Pamies, D., Robertson, D., Smirnova, L., Sun, J., Zhao, L. and Hartung, T. (2015) 'Quality Assurance of Metabolomics', *Altex*, 32, pp. 319–326.
- Bretscher, M. (1972) 'Asymmetrical lipid bilayer structure for biological membranes', *Nature*, 236, pp. 11–12.
- Briehl, M. M., Tome, M. E., Wilkinson, S. T., Jaramillo, M. C. and Lee, K. (2014) 'Mitochondria and redox homeostasis as chemotherapeutic targets.', *Biochem Soc Trans*, 42, pp. 939–44.
- Brown, M., Dunn, W. B., Dobson, P., Patel, Y., Winder, C. L., Francis-McIntyre, S., Begley, P., Carroll, K., Broadhurst, D. I., Tseng, A., Swainston, N., Spasic, I., Goodacre, R. and Kell, D. B. (2009) 'Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics', *The Analyst*, 134, pp. 1322–32.
- Bruns, D. and Jahn, R. (2002) 'Molecular determinants of exocytosis', *Pflugers Arch EJP*, 443, pp. 333–338.
- Cho, S. W., Park, J. S., Heo, H. J., Park, S. W., Song, S., Kim, I., Han, Y. M., Yamashita, J. K., Youm, J. B., Han, J. and Koh, G. Y. (2014) 'Dual modulation of the mitochondrial permeability transition pore and redox signaling synergistically promotes cardiomyocyte differentiation from pluripotent stem cells', *J Am Heart Assoc*, 13, pp. 693–701.
- Correa, A. L., Senna, J. P. M. and de Sousa, A. P. B. (2016) 'Effects of passage number on growth and productivity of hybridoma secreting MRSA anti-PBP2a monoclonal antibodies', *Cytotechnology*, 68, pp. 419–427.
- Dass, C. (2006) 'Modes of Ionization', in *Fundamentals of Contemporary Mass Spectrometry*. John Wiley & Sons, Inc., pp. 15–65.
- Dent, M., Hurtado, J., Paul, A. M., Sun, H., Lai, H., Yang, M., Esqueda, A., Bai, F., Steinkellner, H. and Chen, Q. (2016) 'Plant-produced anti-dengue virus monoclonal antibodies exhibit reduced antibody-dependent enhancement of infection activity', *J Gen Virol*, 97, pp. 3280–3290.
- Donnelly, R. P. and Finlay, D. K. (2015) 'Glucose, glycolysis and lymphocyte responses', *Mol Immunol*, 68, pp. 513–519.

- Dorries, K. and Lalk, M. (2013) 'Metabolic footprint analysis uncovers strain specific overflow metabolism and D-isoleucine production of *Staphylococcus aureus* COL and HG001', *PLoS One*, 8, pp. 81500–81510.
- Dunn, W. B. (2008) 'Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes', *Phys Biol*, 5, p. 24.
- Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005) 'Measuring the metabolome: current analytical technologies', *The Analyst*, 130, pp. 606–625.
- Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., Nicholls, A. W., Wilson, I. D., Kell, D. B. and Goodacre, R. (2011) 'Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry', *Nat prot*, 6, pp. 1060–1083.
- Dunn, W. B., Wilson, I. D., Nicholls, A. W. and Broadhurst, D. (2012) 'The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans', *Bioanalysis*, (18), pp. 2249–2264.
- Ellis, S. R., Brown, S. H., In Het Panhuis, M., Blanksby, S. J. and Mitchell, T. W. (2013) 'Surface analysis of lipids by mass spectrometry: More than just imaging', *Prog Lipid Res*, 52, pp. 329–353.
- Falk Schreiber (2003) 'Visual comparison of metabolic pathways', *J Visual Lang Comp*, 14, pp. 327–340.
- Flis, V. V., Fankl, A., Ramprecht, C., Zellnig, G., Leitner, E., Hermetter, A. and Daum, G. (2015) 'Phosphatidylcholine supply to peroxisomes of the yeast *Saccharomyces cerevisiae*', *PLoS One*, 10, pp. 1–19.
- Gerriets, V. A. and Rathmell, J. C. (2012) 'Metabolic pathways in T cell fate and function', *Trends Immunol*, 33, pp. 168–172.
- Gramer, M. J., Eckblad, J. J., Donahue, R., Brown, J., Shultz, C., Vickerman, K., Priem, P., van den Bremer, E. T. J., Gerritsen, J. and van Berkel, P. H. C. (2011) 'Modulation of antibody galactosylation through feeding of uridine, manganese chloride, and galactose', *Biotechnol Bioeng*, 108, pp. 1591–1602.
- Halket, J. M., Waterman, D., Przyborowska, A. M., Patel, R. K. P., Fraser, P. D. and Bramley, P. M. (2005) 'Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS', *J Exp Bot*, 56, pp. 219–243.
- HMDB (2017). Available at: <http://www.hmdb.ca/> (Accessed: 5 January 2017).
- Hoffmann, E. de and Stroobant, V. (2007) *Mass spectrometry. Principles and applications*. 3rd edn. Chichester: John Wiley & Sons Ltd.
- Hollywood, K. A., Winder, C. L., Dunn, W. B., Xu, Y., Broadhurst, D. I., Griffiths, C. E. and Goodacre, R. (2015) 'Exploring the mode of action of dithranol therapy for psoriasis: a metabolomic analysis using HaCaT cells', *Mol Biosyst*, 11, pp. 2198–2209.
- Honda, T., Soeda, S., Tsuda, K., Yamaguchi, C. and Aoyama, K. (2016) 'Protective role for lipid modifications of Src-family kinases against chromosome missegregation', *Sci Rep*, 12, pp. 38751–38759.
- Hugenholtz, J., Kleerebezem, M., Starrenburg, M., Delcour, J., De Vos, W. and Hols, P. (2000) 'Lactococcus lactis as a cell factory for high-level diacetyl production', *Appl EnviMicrobiol*, 66, pp. 4112–4114.
- Hunter, A. R. W., Lawson, A. C., Galitsiou, B. E., Gifford, F. and Neary, J.J. (2016)

- ‘Pyroglutamic acidosis in association with therapeutic paracetamol use Cases’, *Clin Med*, 16, pp. 524–530.
- Jaskolla, T. W., Onischke, K. and Schiller, J. (2014) ‘2,5-Dihydroxybenzoic acid salts for matrix-assisted laser desorption/ionization time-of-flight mass spectrometric lipid analysis: Simplified spectra interpretation and insights into gas-phase fragmentation’, *Rapid Comm Mass Spectrom*, 28, pp. 1353–1363.
- Jefferis, R. (2005) ‘CCE IX : Review Glycosylation of Recombinant Antibody Therapeutics’, *Biotechnol Prog*, 21, pp. 11–16.
- KEGG pathway database* (2017). Available at: <http://www.genome.jp/kegg/pathway.html> (Accessed: 11 January 2017).
- Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B., Spasic, I. and Oliver, S. G. (2005) ‘Metabolic footprinting and systems biology: the medium is the message’, *Nat Rev Microbiol*, 3, pp. 557–565.
- Klyszejko-Stefanowicz, L. (2002) *Cytobiochemia. Biochemia niektórych struktur komórkowych*. Warszawa: PWN.
- Koebmann, B., Blank, L. M., Solem, C., Petranovic, D., Nielsen, L. K. and Jensen, P. R. (2008) ‘Increased biomass yield of *Lactococcus lactis* during energetically limited growth and respiratory conditions’, *Biotechnol Appl Biochem*, 50, pp. 25–33.
- Kronthaler, J., Gstraunthaler, G. and Heel, C. (2012) ‘Optimizing High-Throughput Metabolomic Biomarker Screening: A Study of Quenching Solutions to Freeze Intracellular Metabolism in CHO Cells’, *J Integ Biol*, 16, pp. 90–97.
- Kumar, A. and Bachhawat, A. K. (2012) ‘Pyroglutamic acid: Throwing light on a lightly studied metabolite’, *Curr Sci*, 102, pp. 288–297.
- Lane, A. N., Fan, T. W.-M. and Higashi, R. M. (2009) ‘Metabolic acidosis and the importance of balanced equations’, *Metabolomics*, 5, pp. 163–165.
- Leung, K. W. and Wong, A. S.-T. (2010) ‘Pharmacology of ginsenosides: a literature review’, *Chinese Med*, 5, p. 20.
- Li, X., Wong, C. C., Tang, Z., Wu, J., Li, S., Qian, Y., Xu, J., Yang, Z., Shen, Y., Yu, J. and Cai, Z. (2017) ‘Determination of amino acids in colon cancer cells by using UHPLC-MS/MS and [U-13C5]-glutamine as the isotope tracer’, *Talanta*, 162, pp. 285–292.
- Liebeke, M., Dörries, K., Zühlke, D., Bernhardt, J., Fuchs, S., Pané-Farré, J., Engelmann, S., Völker, U., Bode, R., Dandekar, T., Lindequist, U., Hecker, M. and Lalk, M. (2011) ‘A metabolomics and proteomics study of the adaptation of *Staphylococcus aureus* to glucose starvation’, *Mol Biosyst*, 7, pp. 1241–1253.
- Liss, D. B., Paden, M. S., Schwarz, E. S. and Mullins, M. E. (2013) ‘What is the clinical significance of 5-oxoproline (pyroglutamic acid) in high anion gap metabolic acidosis following paracetamol (acetaminophen) exposure?’, *Clin Toxicol*, 51, pp. 817–27.
- Lu, J.-M., Yao, Q. and Chen, C. (2009) ‘Ginseng Compounds: An Update on their Molecular Mechanisms and Medical Applications’, *Curr Vasc Pharmacol*, 7, pp. 293–302.
- Martensson, C. U., Doan, K. N. and Becker, T. (2017) ‘Effects of lipids on mitochondrial functions’, *Biochim Biophys Acta*, 1862, pp. 102–113.
- Matthews, G. M., Howarth, G. S. and Butler, R. N. (2012) ‘Short-chain fatty acids induce apoptosis in colon cancer cells associated with changes to intracellular redox state and glucose metabolism’, *Chemotherapy*, 58, pp. 102–109.
- Owen, D. M. and Gaus, K. (2013) ‘Imaging lipid domains in cell membranes: the advent of

super-resolution fluorescence microscopy', *Front Plant Sci*, 4, p. 503.

Pearce, E. and Pearce, E. (2013) 'Metabolic pathways in immune cell activation and quiescence', *Immunity*. Elsevier Inc., 38(4), pp. 633–643. doi: 10.1016/j.immuni.2013.04.005.

Roessner, U. (2006) 'The Chemical Challenge of the Metabolome', in *Metabolome Analysis: An Introduction*. John Wiley & Sons, Inc., pp. 15–38.

Roman-Fernandez, A. and Bryant, D. M. (2016) 'Complex Polarity: Building Multicellular Tissues Through Apical Membrane Traffic', *Traffic*, 17, pp. 1244–1261.

Sellick, C. A., Croxford, A. S., Maqsood, A. R., Stephens, G., Westerhoff, H. V., Goodacre, R. and Dickson, A. J. (2011) 'Metabolite profiling of recombinant CHO cells: Designing tailored feeding regimes that enhance recombinant antibody production', *Biotechnol Bioeng*, 108, pp. 3025–3031.

Sellick, C. a, Hansen, R., Maqsood, A. R., Warwick, B., Stephens, G. M., Goodacre, R., Dickson, A. J. and Dunn, W. B. (2009) 'Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells', *Anal Chem*, 81, pp. 174–183.

Shvets, E., Bitsikas, V., Howard, G., Hansen, C. G. and Nichols, B. J. (2015) 'Dynamic caveolae exclude bulk membrane proteins and are required for sorting of excess glycosphingolipids', *Nat Comm*, 6, p. 6867.

Son, J., Lee, G. and Cha, S. (2014) 'Direct analysis of triacylglycerols from crude lipid mixtures by gold nanoparticle-assisted laser desorption/ionization mass spectrometry', *J Am Soc Mass Spectrom*, 25, pp. 891–894.

Al Suleimani, Y. M. and Al Mahruqi, A. S. (2017) 'The endogenous lipid N-arachidonoyl glycine is hypotensive and nitric oxide-cGMP-dependent vasorelaxant', *Eur J Pharmacol*, 794, pp. 209–215.

Sumner, L. W., Samuel, T., Noble, R., Gmbh, S. D., Barrett, D., Beale, M. H. and Hardy, N. (2007) 'Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)', *Metabolomics*, 3, pp. 211–221.

Sunehag, A. L., Louie, K., Bier, J. L., Tigas, S. and Haymond, M. W. (2002) 'Hexoneogenesis in the human breast during lactation', *J Clin Endocrinol Metab*, 87, pp. 297–301.

Taschwer, M., Hackl, M., Hernandez Bort, J. A., Leitner, C., Kumar, N., Puc, U., Grass, J., Papst, M., Kunert, R., Altmann, F. and Borth, N. (2012) 'Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth', *J Biotechnol*, 157, pp. 295–303.

Tatsuta, T. and Langer, T. (2017) 'Intramitochondrial phospholipid trafficking', *Biochim Biophys Acta*, 1862, pp. 81–89.

Vasconcellos, R., Alvarenga, E. C., Parreira, R. C., Lima, S. S. and Resende, R. R. (2016) 'Exploring the cell signalling in hepatocyte differentiation', *Cell Signal*, 28, pp. 1773–1788.

Villas-Bôas SG, Roessner, U., A., E. H. M., Smedsgaard, J. and Nielsen, J. (2007) *Metabolome analysis: an introduction*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/0470105518.

Wacker, C., Berger, C. N., Girard, P. and Meier, R. (2011) 'Glycosylation profiles of

therapeutic antibody pharmaceuticals', *Eur J Pharm Biopharm*, 79, pp. 503–507.

Wang, T., Marquardt, C. and Foker, J. (1976) 'Aerobic glycolysis during lymphocyte proliferation', *Nature*, 261, pp. 702–705.

Xie, P., Niu, H., Chen, X., Zhang, X., Miao, S., Deng, X., Liu, X., Tan, W. S., Zhou, Y. and Fan, L. (2016) 'Elucidating the effects of pH shift on IgG1 monoclonal antibody acidic charge variant levels in Chinese hamster ovary cell cultures', *Appl Microbiol Biotechnol*, 100, pp. 10343–10353.

Ye, J., Ly, J., Watts, K., Hsu, A., Walker, A., McLaughlin, K., Berdichevsky, M., Prinz, B., Sean Kersey, D., D'Anjou, M., Pollard, D. and Potgieter, T. (2011) 'Optimization of a glycoengineered *Pichia pastoris* cultivation process for commercial antibody production', *Biotechnol Prog*, 27, pp. 1744–1750.

Ying, W. (2008) 'NAD⁺/NADH and NADP⁺/NADPH in Cellular Functions and Cell Death: Regulation and Biological Consequences', *Antioxid Redox Signal*, 10, pp. 179–206.

Zenobi, R. (2013) 'Single-Cell Metabolomics: Analytical and Biological Perspectives', *Acta Neuropathol*, 342, pp. 1201–1211.

Zhang, Q., Liang, Z., Gao, Y., Teng, M. and Niu, L. (2017) 'Differentially expressed mitochondrial genes in breast cancer cells: Potential new targets for anti-cancer therapies', *Gene*, 596, pp. 45–52.

CHAPTER 7: Metabolic imaging of CHO cells

7.1. Introduction

Mass spectrometry imaging (MSI) is a powerful tool that offers a unique perspective on biological systems, allowing for spatial visualisation and distribution of biochemical species within individual cells (*e.g.*, Lanni et al. 2012, Vickerman 2011, Burrell, Earnshaw and Clench, 2007).

The basic workflow of cell-based MSI experiment is shown in **Figure 1** and requires that cells are deposited onto a sample support followed by washing steps in order to remove debris and, depending on the technique used, coated with a layer of matrix, which assists ionisation of analyte of interest (**Figure 1 A**) (Weaver and Hummon, 2013, Lanni et al. 2012).

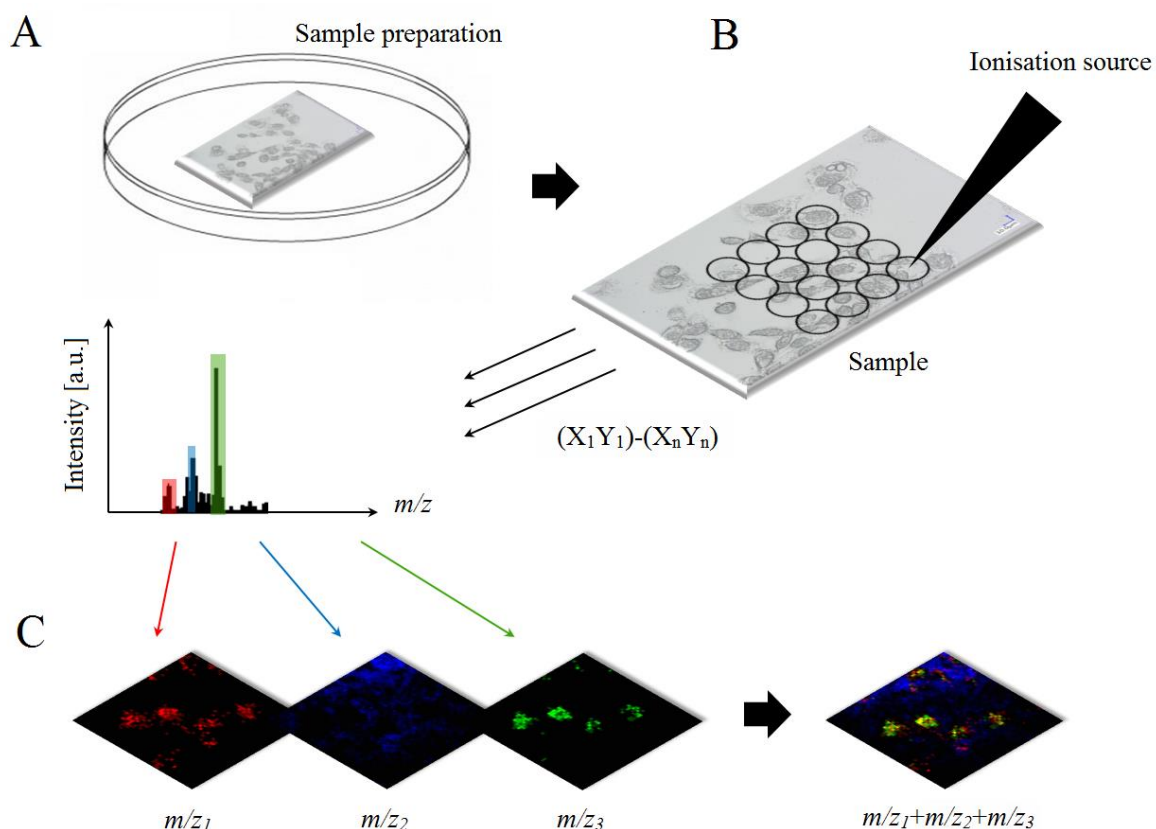


Figure 1: Basic workflow of cell-based MSI experiment. Cells are deposited onto a sample support followed by washing steps (A), the ionisation source rasters across the sample in an X and Y coordinate manner providing with mass spectrum for each pixel (B), data analysis provides with an image of the cells illustrating the spatial distribution of m/z values recorded (C).

Once the target is introduced to the instrument, the ionisation source rasters across the sample in an X and Y coordinate manner, while a mass spectrometer records chemical information for each position *i.e.* pixel. As a result, each pixel is a mass spectrum from a specific coordinate position on the cells and contains all of the ionised species detected within a chosen mass range (**Figure 1 B**). Data acquisition is then followed by application of data analysis software that combines chemical information with an image of the cells, generating heat maps illustrating the spatial distribution of any m/z values recorded by the mass spectrometer (**Figure 1 C**) (Walch *et al.*, 2008).

A number of challenges have to be taken into consideration in MSI experiments and, specifically, MSI cell-based studies. Apart from the challenge of designing appropriate ionisation sources at micron and submicron scale, for this purpose, sensitivity becomes a major concern. Therefore, longer acquisition times or increasing volume sampled (*e.g.* deeper, as a voxel) can be applied in order to overcome this challenge (Lanni *et al.* 2012). Additionally, the sensitivity issue can also be related to the suppression effects by *e.g.* matrix in MALDI-MSI due to ionisation properties as well as the size of the matrix crystals (Zhu *et al.*, 2011).

Other considerable challenges in cell-based MSI experiments are the development of suitable sample preparation procedures and data analysis methods that are needed to reliably visualise the distributions of species at micron and submicron resolutions (Spengler, 2015). Appropriate sample preparation is crucial to maintain sample integrity and must preserve the native state morphology, biochemical composition and localisation of the analytes and be reproducible (Schaepe *et al.*, 2015, Fletcher *et al.*, 2011, McDonnell and Heeren, 2007).

Unlike for profiling analyses, samples for MS imaging require some additional steps to be considered prior to the experiment, including *e.g.* washing and fixation procedures of the analyte (*i.e.* cells). Washing procedures are required in order to remove cell debris, salts or buffers that would impair the detection of the species of interest. During these washing protocols, considerable care must be taken in order to prevent delocalisation of the molecules within the sample (Sparvero *et al.*, 2012). Washing protocols vary, depending on the target analyte. For example, for peptide imaging cold mixtures of aqueous alcohols are used that reduce the solubility and delocalisation of these molecules (Yang and Caprioli, 2011, Angel and Caprioli, 2014). On the other hand, for MSI analysis of drugs and other small molecules it is advised to minimise washing steps to avoid compounds being washed away (Cornett *et al.*, 2008, Weaver and Hummon, 2013).

Additionally, fixation protocols are applied for the MSI analysis of cell cultures (Schaepe *et al.*, 2015). The fixation of the sample is applied in order to preserve the biochemical and structural integrity of cells and tissues. Fixation prevents cell internal structures from destruction that may lead to delocalisation of biomolecules as well as changes in osmotic pressure, resulting in cell shrinkage or cell swelling causing autolysis process (Lanni *et al.* 2012, Rabbani *et al.* 2011, Denbigh & Lockyer 2015). During autolysis, the cell's internal enzymes initiate the digestion of biomolecules such as proteins, lipids and nucleotides, leading to the destruction of biochemical cell composition and integrity (Klyszejko-Stefanowicz, 2002).

A number of fixation protocols used in metabolomics for imaging of different classes of biomolecules have been developed. Several of different organic solvents such as methanol or ethanol have been applied in lipidomic studies. It has been shown that alcohols act as coagulant agents, decreasing the solubility of proteins, precipitating and accumulating them within the cells (Vekemans *et al.* 2004, Liu *et al.* 2007). Despite the broad usage of alcohols as fixatives, many studies have reported their effects on cell dehydration and restrictions with lipid solubility and removal (Vekemans *et al.* 2004, Liu *et al.* 2007, Didonato & Brasaemle 2003, Hanrieder *et al.* 2011).

Another choice apart from alcohol fixation is aldehyde fixation (*e.g.* glutaraldehyde, paraformaldehyde or formalin), where proteins are cross-linked within the cells with no harm to protein structure and lipids can be fixed by the reactions of formalin with double bonds of unsaturated lipids (Schober *et al.* 2012, Anderson *et al.* 2013, Schaepe *et al.* 2015, Lee *et al.* 2006, Nagata *et al.* 2014). Whilst, the above selection of fixation methods have been applied for MS imaging of targeted biomolecules, it is important to employ an appropriate fixation treatment suitable for non-targeted investigations which allows for a wide diversity of intracellular molecules to be analysed. The combination of above challenges effectively limits cell-based MS imaging to a set of the most abundant species, hence leaving much room for improvement of MSI experiments (Lanni *et al.* 2012).

With recent development of MSI instruments *e.g.* improved spatial resolution in the SMALDI platform with 1.4 μm pixel size reported lately (Kompauer *et al.*, 2017) and improved mass accuracy in OrbiSIMS (<http://www.npl.co.uk/news/3d-nanosims-label-free-molecular-imaging>), single cell-based MS imaging has gained even more attention. Single-cell based analysis, providing molecular information from within individual cells, can give more understanding and new insights in processes involved in *e.g.* cellular evolution, communication or adaptation (Emara *et al.*, 2017). Moreover, metabolic

investigations at the single-cell level can provide with new observations to phenotypic differences within one cell population (Heinemann and Zenobi, 2011).

Cell-based level MSI analyses using MALDI and SIMS platforms have been reported in a number of studies, where *e.g.* a wide mass range of analytes was detected by MALDI such as large macromolecules at 100 μm spatial resolution (Llombart *et al.*, 2016), proteins at 10 μm spatial resolution (Yang and Caprioli, 2011), and lipids at 5 μm (Schober *et al.*, 2012) and down to 1.4 μm spatial resolution (Kompauer, Heiles and Spengler, 2017). The SIMS platform has been applied in more limited mass ranges (*i.e.* up to ~ 600 Da) yet, the high spatial resolution within a range from ~ 8 μm to <1 μm , has given possibilities to image individual cells and sub-cellular compartments with three-dimensional visualisation (Kotze *et al.*, 2013, Rabbani *et al.*, 2011, Touboul *et al.*, 2005).

This chapter presents a metabolite MS imaging approach of CHO cells in order to investigate the cell response to EPO-Fc fusion protein expression. Characterisation of metabolites allows for the assessment of cellular activity of these cells and their metabolic state under protein production at cell population level as well as single cell and sub-cellular levels.

7.2. Sample preparation

CHO cells in medium and with or without inducer were grown directly on microscopic glass slides and on 5 mm × 5 mm silicon wafers at the density of 0.07×10^6 cells/mL and 0.03×10^6 cells/mL for MALDI-MS and SIMS imaging, respectively. Optimisation of cell density for MS imaging is further described in **Supplementary Information, Figure 17**. After incubation time (day 3), media were aspirated and cells were washed with PBS (twice), followed by fixation with 150 mM ammonium formate for 1 min (three times). Slides and wafers were then placed in a desiccator with silica gel until fully dry. Dry samples were kept in a desiccator until analysis. Determination of suitable fixation method for MS imaging is described further in **Supplementary Information, Table 4, Figure 18, 19 and 20**. For MALDI-MSI experiments, matrix coating with 100 μ L of 30 mg/mL DHB in 50% acetone with 0.1% TFA and 80 μ L of 30 mg/mL *para*-nitroaniline (PNA) in 50% acetone for positive and negative ionisation modes respectively was applied.

7.3. Instrumentation

MALDI-MSI experiments were performed in collaboration with Professor Bernhard Spengler at the Justus Liebig University Giessen in Germany. The research was carried out using an AP-SMALDI10 imaging source (Trans-MIT GmbH, Giessen, Germany) connected to a Thermo Q Exactive (San Jose, CA, USA) Fourier transform orbital trapping mass spectrometer equipped in 337 nm wavelength nitrogen pulsed UV laser and repetition rate of 60 Hz. The laser beam was focused to a spatial resolution of 5 μ m and energy of 0.06 μ J (for attenuator at 40°). The rastering step size was equal to the laser spot size. The sample was placed on the stage and moved by an X, Y dimensions to coordinate an area of interest. MS imaging data was acquired in continuous mode using Tune software (Thermo Scientific, San Jose, CA, USA) in the mass range of 250 to 1000 Da and in the “profiling mode”, giving a mass resolution of 140 000 (FWHM) at m/z 200. The instrument was operated in positive and negative ionisation modes with +4 kV and -4 kV on the sample plate respectively. Internal calibration was applied using matrix peaks as reference masses for both positive and negative ionisation modes: $[5\text{DHB}-4\text{H}_2\text{O}+\text{NH}_4]^+$ at m/z 716.12461 for positive and $[3\text{PNA}-5\text{H}]^-$ at m/z 409.090209 for negative ionisation mode. Mass accuracy of 2 ppm Root Mean Square Error (RMSE) was found. Laser, sample stage and mass spectrometer were synchronised and controlled by a microcontroller and in-house built software. More information on the instrument details can be found in the recent literature (Tsai *et al.* 2016, Khalil *et al.* 2015, Bhandari *et al.* 2015). In this study there were 8 experiments in total (*1 repeat* × *1 slide* × *2 samples*

(positive ion mode) + 2 repeats \times 1 slide \times 2 samples (positive ion mode) + 1 repeat \times 1 slide \times 2 samples (negative ion mode)).

SIMS imaging experiments were conducted at Ionoptika Limited (Hampshire, UK) and University of Manchester, UK using a J105-3D chemical imager. Samples were sputtered with 25 keV Au_3^+ and 40 keV C_{60}^+ primary ion beams in positive ionisation mode and spatial resolution of $\sim 0.5\ \mu\text{m}$ and $\sim 8\ \mu\text{m}$ respectively. Experiments with 25 keV Au_3^+ source were performed using ‘low quad’ and ‘high quad’ settings in order to record spectra with extra focus on low (for ‘low quad’) or high (for ‘high quad’) mass range. Detailed settings for each SIMS experiment are shown in **Table 1, Chapter 3**. There were 8 experiments in total (1^{st} Si wafer \times 2 repeats (ctrl sample) + 2^{nd} Si wafer \times 1 repeat \times (ctrl sample) + 1^{st} Si wafer \times 2 repeats (induced sample) + 2^{nd} Si wafer \times 3 repeats (induced sample)).

7.4. Data analysis

MALDI-MSI data were saved in .raw files and converted to .imzML files. The data were then opened in free online MS imaging MSiReader ver. 0.09 software. Data were normalised to a cell derived ion at m/z 798.541 (Römpf *et al.*, 2010, Goodwin *et al.*, 2011) for positive ion mode and to matrix ion at m/z 409.090 for negative ion mode. Results were exported in .xlsx format and further processed with MATLAB ver. R2012a and free online MetaboAnalyst ver. 3.0. SIMS data was saved in .dat files and further processed with MATLAB ver. R2012a and J105 Image Analyser software. Averaged MS spectra were generated using sum normalisation and m/z 0.05 binning (for MATLAB). J105 Image Analyser (Ionoptika Limited, Hampshire, UK) was used for SIMS data processing. MetaboAnalyst 3.0 analysis was conducted using m/z 0.25 binning, KNN missing value estimation, SD data filtering, sum normalisation and Pareto scaling. Analysis of variance (ANOVA) and *t*-test analysis were performed. Metabolite peaks were chosen as significant with a *p* value threshold $p < 0.05$. Box-and-Whisker plots, PCA and PLS-DA were generated based on *t*-test or ANOVA test. Metabolite identifications were assigned through inspecting and matching against The Human Metabolome Database (HMDB) (<http://www.hmdb.ca/>) or Surface Spectra Static SIMS library. Metabolite identification threshold of 5 ppm of mass error was used (for MALDI).

7.5. Results and discussion

7.5.1. MSI of CHO cells under inducer exposure and EPO-Fc protein production

7.5.1.1. Cell population-based study

The analysis of MS images of CHO cell populations suggests that the exposure of cell culture to the EPO-Fc inducer triggers great changes in metabolic profiles of these cells **Figure 2 and Figure 3**.

Figure 2 A illustrates the averaged mass spectra collected by AP-SMALDI in positive and negative ionisation modes from control and tetracycline-induced CHO cells in a range of m/z 250 to 1000. The spectra were collected from 10 different locations of cell populations within control and induced cells each and for both ionisation modes.

The overall MALDI-MSI spectra show good alignment, but also highlight different mass information and different intensities between positive and negative ion modes. Spectra in positive ion mode are rich in information and metabolites are detected across the whole mass range. The abundance of species across the whole mass range is similar. A great number of peaks is noted for a mass range between m/z 700-900, suggesting a number of lipid and lipid-like species detected that are well known to appear within this mass range (Sellick *et al.*, 2011, Hollywood *et al.*, 2015, Dorries & Lalk, 2013). The high content of lipid species was expected as these compounds are remarkably abundant in cell membranes (Klyszejko-Stefanowicz, 2002). A number of cell MS imaging studies show a great content of fatty acids, lipids and their fragments in collected MALDI and SIMS spectra from different type of samples. This was shown in several investigations of healthy and cancer mammalian cell lines (Robinson *et al.* 2016, Schober *et al.* 2012) and tissues (Phan *et al.* 2016, Vanbellingen *et al.* 2015, Tian *et al.* 2014). In this study, MALDI spectra from positive ion mode also present many peak areas that differ between control and induced cells across the whole mass range, for instance $\sim m/z$ 300, $\sim m/z$ 500 and m/z 920-1000, indicating specific changes in cell metabolome occurring under protein production.

By contrast, spectra collected in negative ion mode provide less information, as the amount of metabolites detected is limited, when compare to positive ion mode. Relatively high ion intensities are seen within mass area of m/z 280-550 however, the most peaks are dominated by a matrix peak at m/z 273 (excluded from the spectrum in **Figure 2 A**). Nevertheless, the metabolite detection at the region around $m/z > 300$ might suggest of the presence of species such as nucleotides, appearing at m/z 322 (uridine 5'-monophosphate), m/z \sim 426 (adenosine 5'-diphosphate), m/z \sim 506 (adenosine triphosphate) and m/z \sim 522 (guanosine triphosphate) in negative ion MALDI-MSI studies (Amantonico *et al.* 2010,

Steinhoff *et al.* 2014). Because of the higher content of phosphate and hydroxyl groups than amino groups within these compounds (Klyszejko-Stefanowicz, 2002), the detection of such metabolites in MSI negative ion mode experiments is also expected. Similarly, due to the presence of phosphate and hydroxyl groups in lipids (Klyszejko-Stefanowicz, 2002), these species represent a major content of negative ion mass spectra in a mass range $> m/z$ 700. This has been confirmed in a number of MALDI imaging studies of cells and tissues (Zhang & Yao 2012, Schober *et al.* 2012, Anderson *et al.* 2013). However, in this study low peak content in the lipid mass range is noticeable.

The majority of regions within spectra from control and induced cells have similar metabolites detected but with different abundance, for example m/z 350-550, where higher intensities for induced cells are noted. However, some peak areas of spectra are unique, such as mass region of m/z 700-800, suggesting specific metabolome changes occurring in CHO cells under protein production. Additionally, for both ionisation modes, matrix peaks were also identified: m/z 273 $[2\text{DHB}+\text{H}-2\text{H}_2\text{O}]^+$, m/z 290 $[2\text{DHB}+\text{NH}_4-2\text{H}_2\text{O}]^+$, m/z 409 $[3\text{DHB}+\text{H}-3\text{H}_2\text{O}]^+$, m/z 444 $[3\text{DHB}^*-\text{H}_2\text{O}]^+$, m/z 447 $[3\text{DHB}+\text{K}-3\text{H}_2\text{O}]^+$, m/z 580 $[4\text{DHB}^*-2\text{H}_2\text{O}]^+$, m/z 273 $[2\text{PNA}-3\text{H}]^-$.

Despite all the advantages of MSI techniques *e.g.* label-free, direct detection of m/z values over areas of interest, clear correlation m/z values to mapped regions (Jirasko *et al.*, 2014), MALDI platforms introduce an additional challenge in terms of matrix application during sample preparation (Seeley and Caprioli, 2012). Matrix application is required in order to protect the sample and improve its ionisation (Schober *et al.* 2012). Specifically for MALDI-MS imaging experiments the additional complications such as uniform distribution of the matrix, its small crystal size, limited redistribution of analyte molecules and uniform ionisation enhancement need to be also recognised (Kiss & Hopfgartner 2016, Shanta *et al.* 2012, Schober *et al.* 2012). Hence, the right choice of matrix that would be suitable for specific cell or tissue sample plays a crucial role in MALDI imaging. Various sample preparation protocols for negative ion MALDI-MSI have been recently incorporated, including the application of 9-aminoacridine (9AA) in the study of sarcoma biomarker identification (Lou *et al.*, 2016), alpha-cyano-4-hydroxycinnamic acid (CHCA) derivative (4-phenyl-alpha-cyanocinnamic acid amide, PhCCAA) in imaging liposomes in brain tissues (Fülöp *et al.*, 2016) and *para*-nitroaniline (PNA) in the study of biomarker lipids for phagocytosis in brain sections (Nielsen *et al.*, 2016). The stability of PNA in different vacuum conditions was also shown, demonstrating additional utility of this matrix in both positive and negative ion modes (Steven, Race and Bunch, 2013). Despite satisfying metabolite detection with the application of above matrices for imaging studies,

challenges regarding poor matrix solubility and hence, heterogeneously coated samples are still to be overcome (Miura *et al.*, 2010). Other matrices such as DHB and 1,5-diaminonaphthalene (1,5-DAN) are also used but, despite their advantages *e.g.* uniform film creation or versatility, some disadvantages regarding sensitivity of lipids in negative ion mode have been reported (Spengler 2015, Boughton *et al.* 2016, Korte & Lee 2014, Korte *et al.* 2015). In this study, the application of PNA resulted in low spectral information with limited number of metabolites detected. A number of approaches could be introduced to improve this result *e.g.* the use of different matrix, the employment of co-matrices or matrix additives. Co-matrices have been proposed to improve the quality of MALDI data by generating uniformly coated films, increasing detection sensitivity *e.g.* for phosphopeptides (Li-Hua Zhou *et al.* 2009, Shanta *et al.* 2011) and expanding the detection range regarding chemistry of analytes (Calvano *et al.*, 2013).

Recent studies have presented better co-crystallisation, low matrix interferences and high detection sensitivity of a broad range of small molecules such as amino acids, oligosaccharides and nucleosides with the application of carbon dots and 9AA as a binary matrix in profiling of urine in MALDI positive ion mode (Chen *et al.*, 2016). Additionally, several research groups have applied matrix additives, for example DHB with nitrates (Griffiths and Bunch, 2012) or sodium and potassium acetate (AlMasoud *et al.*, 2016) in lipid extracts analysis, where improved detection of these species was presented. Also, the application of DHB matrix with sodium acetate in brain tissue imaging in positive ion mode resulted in low (overall) mass spectra intensity but high intensity of phospholipids region (Gill *et al.*, 2017). Nonetheless, there is no current literature available that would confirm these findings for negative ion mode MALDI imaging.

The investigation of the influence of EPO-Fc fusion protein production on CHO cells was also performed with the use of the J105 3D Chemical Imager (further called SIMS) (**Figure 3**). The experiments included imaging the first 10 top ‘layers’ of control and tetracycline-induced CHO cells. A layer relates to a single 2D image plane corresponding to a specific ion dose. Stacking these image planes provides subsurface chemical information limited only by the residual damage caused by the primary ion beam. The analysis of the ratio of m/z 184 ion (lipid origin) and m/z 168 (silicon) (Touboul *et al.*, 2005) across all layers revealed that the m/z 184 content is different for the 1st layer and similar within layers 2-10 (**Supplementary Information, Figure 21**). This suggested that the chemical information of the top (1st) layer (information mostly from intact cell membrane) might differ to the rest of the layers (layers from 2 to 10, information also from

cell interior) and therefore, further analysis of the data included two groups: the 1st layer further called ‘top layer’ and layers from 2 to 10 further called ‘lower layers’.

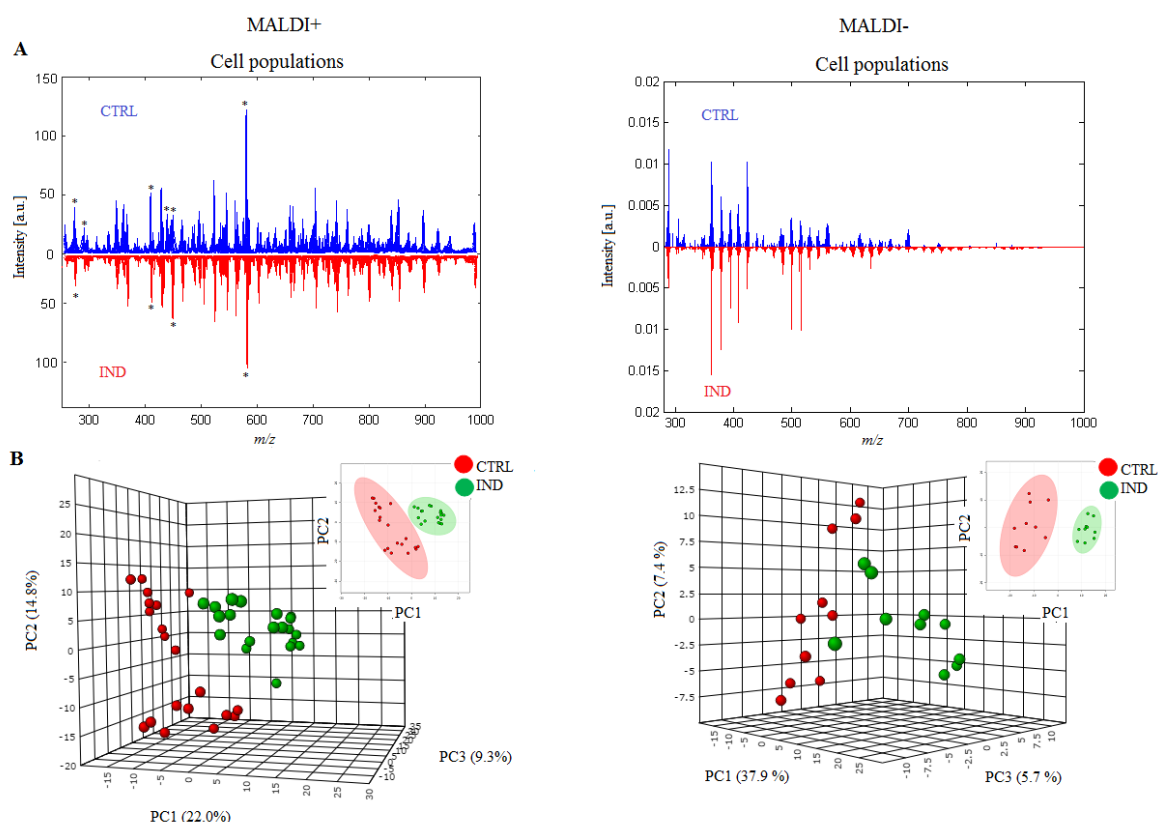


Figure 2: Averaged MALDI-MSI spectra from control (CTRL, blue) and induced cell areas (IND, red) for positive and negative ion modes; m/z 250-1000 for positive ion and m/z 280-1000 for negative ion, *-matrix peaks (A). (Inset) 2D and 3D PCA scores plots accounting for 46.1% and 51.0% of total variance (first three components) for positive and negative ionisation modes, respectively (B). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence.

Figure 3 A illustrates the averaged mass spectra collected by SIMS (40 keV C_{60}^+ ion source) in positive ion mode from control and tetracycline-induced CHO cells in a range of m/z 30 to 900. The spectra were collected from 2 different locations of cell populations within control and induced cells for each experiment, based on the presence of m/z 184 ion. The overall SIMS spectra are rich in information with metabolite detection from ~100 Da to 500 Da. The spectra of the top and lower layers data sets illustrate a number of ions detected with high abundance from m/z 100 to about m/z ~230. Ions detected in a mass range of m/z <100 and m/z ~230 are much less abundant. Also, almost no ions above m/z

350 are seen. Additionally, spectra from both top and lower layers data sets look alike suggesting similar chemical content across all 10 layers.

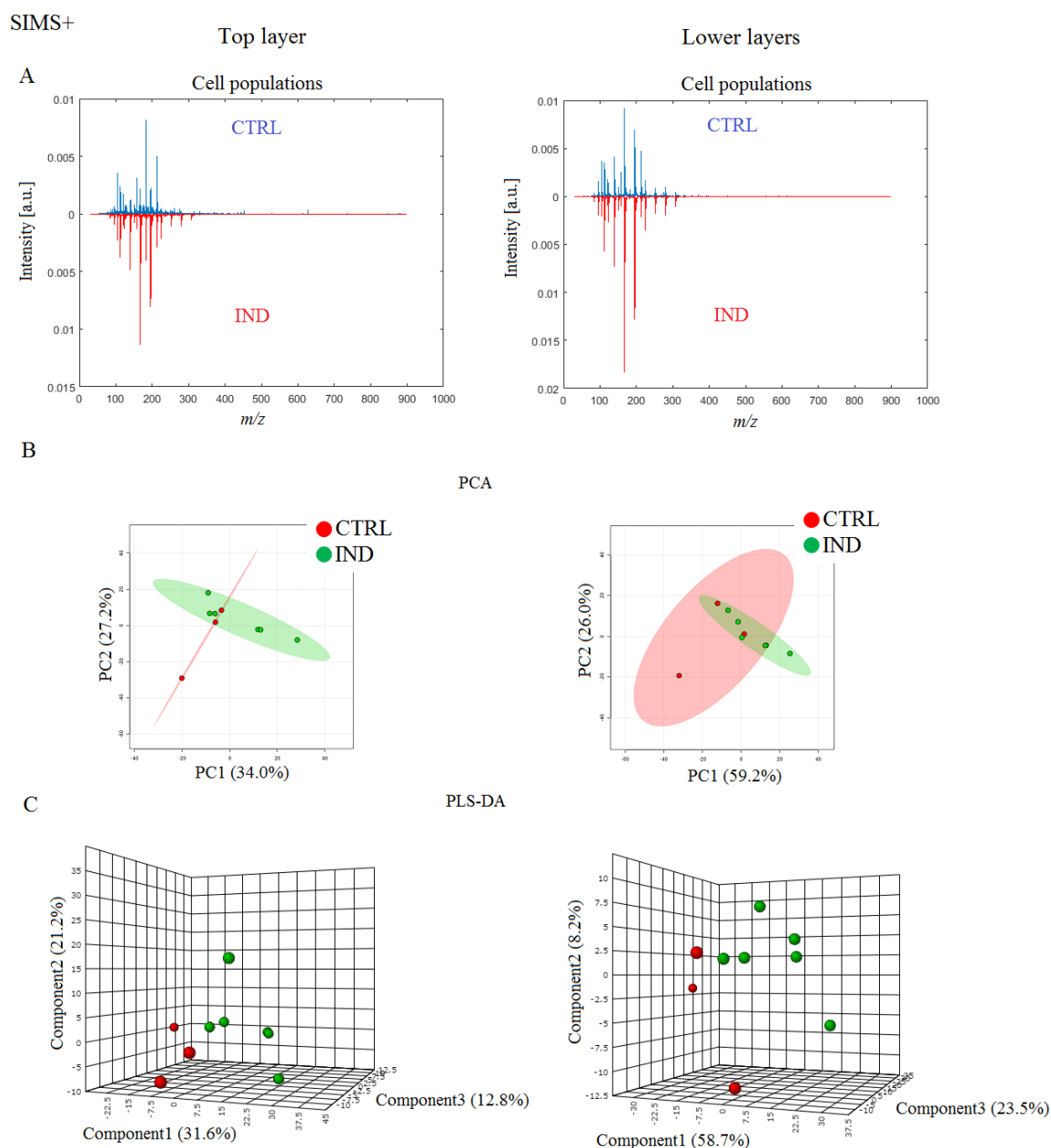


Figure 3: Averaged SIMS+ imaging spectra from control (CTRL, blue, $n=3$) and induced cell areas (IND, red, $n=5$) for the 1st layer ('top layer') and layers 2 to 10 ('lower layers'), m/z 30-900 (A). 2D PCA scores plots accounting for 61.2% and 85.2% of total variance (first two components) for the top and lower layers (B). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence. 3D PLS-DA scores plots accounting for 65.6% and 90.4% of total variance (first three components) for the top and lower layers (C); 40 keV C_{60}^+ ion source.

The SIMS technique is known as a “hard ionisation” method because of the high level of sample damage around the area, where the initial impact of high-energy primary ions occurs. Due to the high energy involved, the secondary ions are usually small fragments of the biomolecules that represent the surface of the sample around the region of primary beam impact (Vickerman 2011). Therefore, a high level of fragmentation within SIMS spectra is seen (**Figure 3 A, Table 3**).

Unlike MALDI technique where the application of matrix improves ionisation of molecules and protects them from degradation (Shanta *et al.* 2012, Fukuyama 2011), MALDI spectra contain no or little fragmentation with high mass ions detected (**Figure 2 A**). Because of the lack of matrix improving analyte ionisation during SIMS experiments as well as the extensive fragmentation of surface biomolecules, the mass limit of SIMS measurement is around 2000 Da (Lanni *et al.* 2012, Weaver & Hummon 2013).

The following step incorporated the application of multivariate analysis on data sets obtained using MALDI-MSI and SIMS imaging in order to investigate metabolic changes in cells producing the protein. Multivariate analysis was performed on the cell areas chosen from control and induced cells. PCA and PLS-DA scores plots were used to visualise relations between two groups of samples (**Figure 2 B and Figure 3 B, C**). For MALDI-MSI, 3D PCA scores plots generated for data sets using first three components account for 46.1% and 51.0% of total variance for positive and negative ion modes respectively. 2D PCA scores plots illustrate some separation between control and induced cells, accounting for 22.0% (PC1) for positive ion mode and 37.9% (PC1) for negative ion mode. Also, all areas chosen for analysis show good clustering within both groups and fit in the 95% area of confidence (**Figure 2 B**).

On the other hand, 2D PCA scores plots generated for SIMS data using first two components show little separation and account for 61.2% and 85.2% of total variance for the top layer and lower layers respectively. Therefore, in order to investigate differences between control and induced cells further analysis with the application of PLS-DA was performed (**Figure 3 C**). 3D PLS-DA scores plots illustrate good separation of these two groups, accounting for 65.6% and 90.4% of total variance (first three components) for the top and lower layers respectively.

The next step of data analysis incorporated the choice of groups of lipids and their fragments commonly imaged in cells and tissues (Tian *et al.* 2014, Schober *et al.* 2012, Khalil *et al.* 2015, Touboul *et al.*, 2005) in order to map cell areas within the imaged field of view. Lipids are the major organic constituent of living cells and are relatively easily

ionised, when compared to other biomolecules (Tian *et al.*, 2014) therefore, lipid fragments were first selected to image their localisation within the cell population.

In SIMS experiments, the highly abundant phosphocholine head group ion at m/z 184 is often used as a marker for cells and is commonly imaged in cells and tissues (Denbigh & Lockyer 2015, Rabbani *et al.* 2011, Touboul *et al.*, 2005). **Figure 4** presents the distribution of PC head group at m/z 184 within top layer and lower layers of control and tetracycline-induced CHO cells.

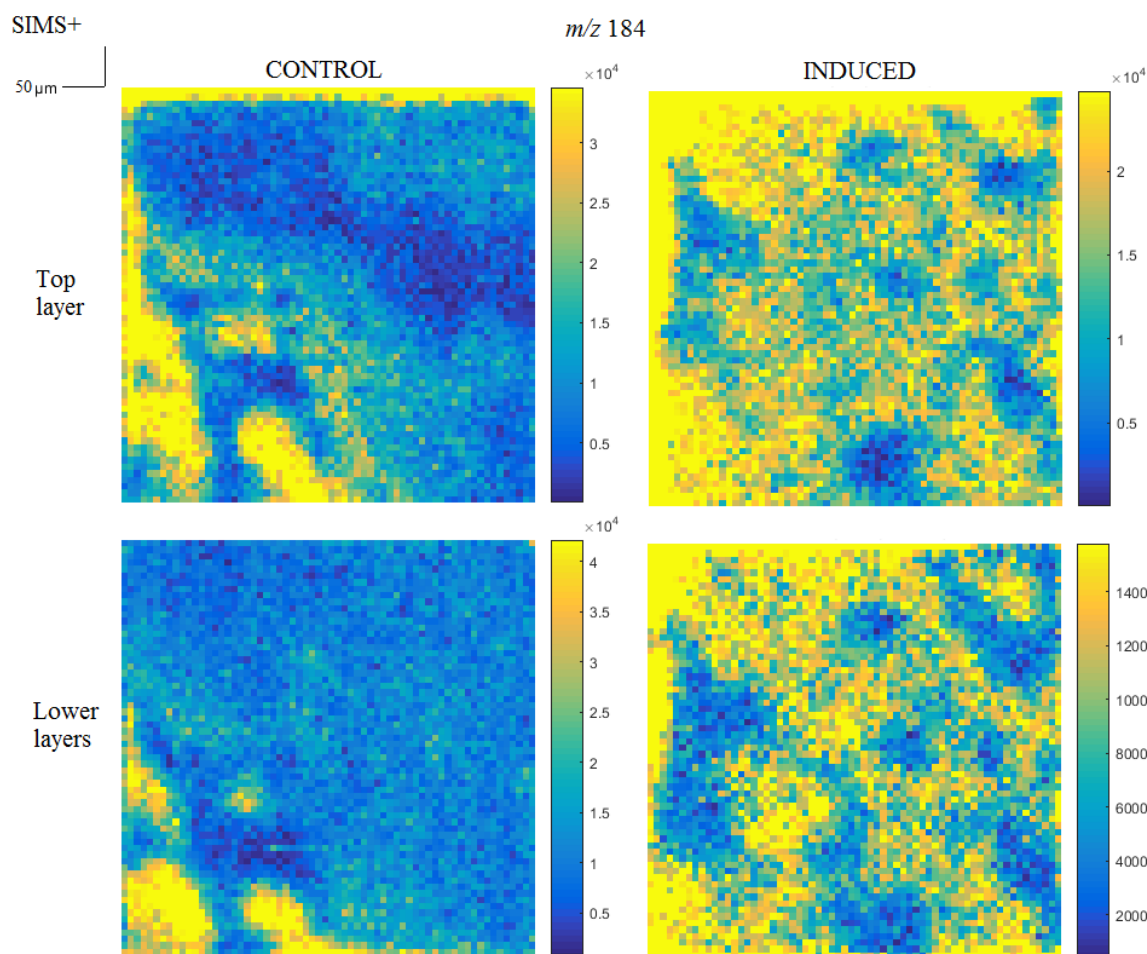


Figure 4: SIMS images of PC head group m/z 184.06 within top layer and lower layers of control (‘CONTROL’) and tetracycline-induced (‘INDUCED’) CHO cells. The abundance of m/z 184 ion for each sample within the top and lower layers is shown. Scale bar 50 μm , pixel size $\sim 8 \mu\text{m}$; 40 keV C_{60}^+ ion source.

For MALDI-MSI, the following ions were selected: m/z 798.54077, m/z 782.56689, m/z 760.5849 for positive ion and m/z 885.54974 for negative ion mode (**Table 1**). The chosen ions were then used for generation of the images (**Figure 5**). Localised cell areas in **Figure 5** were also compared to optical images of samples taken before the measurements.

Table 1: Ions selected to localise cell areas within control and cell-induced imaged samples. The table summarises measured mass, mass error (ppm) and metabolite according to Human Metabolome Data Base.

Measured mass	Δ ppm	Metabolite
798.54077	0.05	PE(22:5/18:1); PE(22:6/18:1)
782.56689	0.05	PE(16:1/22:6); PE(18:1/20:5)
760.58490	0.19	DG(22:2/0:0/20:4)
885.54974	0.13	PI(22:4/16:0)

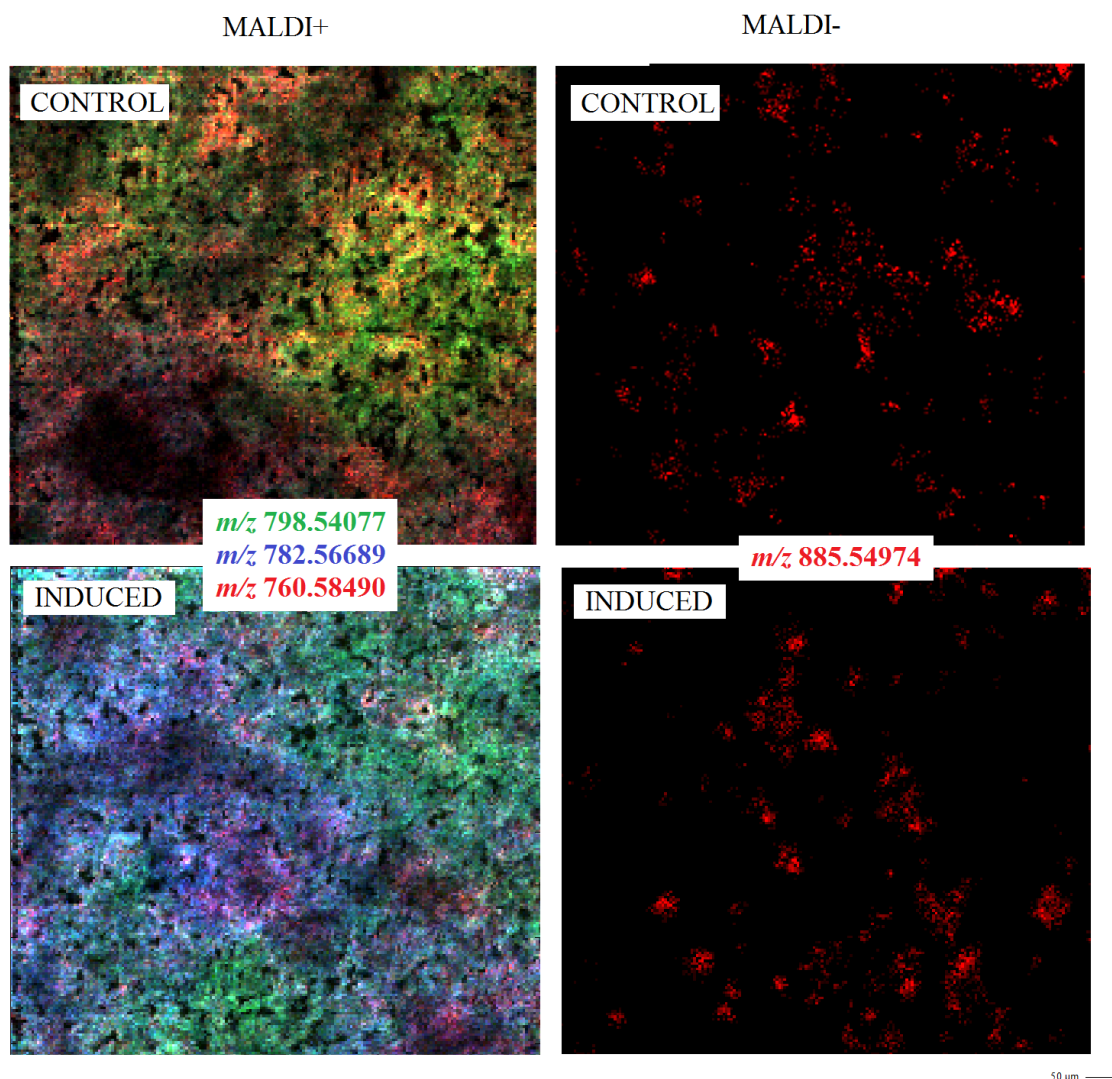


Figure 5: MALDI-MS images of control and induced cells for both positive and negative ionisation modes. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red) for positive ion mode is shown; max. count per pixel: MALDI+, CONTROL: m/z 798.54077 (2.10×10^3), m/z 782.56689 (1.05×10^3) and m/z 760.5849 (2.98×10^3); INDUCED: m/z 798.54077 (4.66×10^3), m/z 782.56689 (1.18×10^3) and m/z 760.5849 (4.51×10^3). Single ion at m/z 885.54974 for negative ion mode is shown; max. count per pixel: MALDI-, CONTROL 3.26×10^3 ; INDUCED 1.25×10^3 . Scale bar 50 μm , pixel size 5 μm . NOTE: different cell coverage for MALDI+ and MALDI-

The results of the spectral maps provide detailed biochemical information of CHO cells that is not readily accessible from other methods.

In SIMS images of control and induced cells, the location of PC head group at m/z 184 is noticeable, indicating the location of cell areas (**Figure 4**). The difference in the abundance

of this ion between top and lower layers for both control and induced cells is very little and seems to be higher within the top layer. This could be explained with the bigger exposure of the cell membrane in the top layer and therefore, higher content of lipid species. **Figure 4** also illustrates ion images with cellular shapes preserved. However, images present visualisation of population of cells, where discrimination of single cells is limited. At this time, optimisation of the C_{60}^+ ion beam for high-resolution imaging was not possible so image resolution was greatly compromised in favour of sensitivity. The C_{60}^+ ion beam is capable of a minimum spot size of <500 nm.

According to database assignments (**Table 1**), MALDI images show regions of lipid content with different abundance within control and induced CHO cells (**Figure 5**). For MALDI positive ion mode, three species were highlighted illustrating a location of cell areas. Similarly to SIMS data, MALDI images show visualisation of population of cells however, single cells within these populations are impossible to notice that might be due to the high cell coverage within the imaged areas.

A comparison of images of tetracycline-induced cells against control cells show differences in abundance of these species in cells under protein production. Induced cells present increased abundance of all highlighted ions: m/z 760.58490, m/z 798.54077 and m/z 782.56689, when compare to control cells. Additionally, the heterogeneity for both cell populations *i.e.* control and tetracycline-induced cells can be seen, where high level of cell diversity is detected. Cells producing EPO-Fc seem to have all three lipid species *i.e.* m/z 798.54077, m/z 782.56689 and m/z 760.58490 present, whereas only two lipid species are seem to be the most abundant within control CHO cells species *i.e.* m/z 798.54077 and m/z 760.58490. Moreover, the morphology of both cell populations appears to be well preserved: the shape of cells is consistent however, as mentioned previously, single cells are impossible to differentiate.

Images from MALDI negative ion mode show ion at m/z 885.54974 highlighted. The abundance of this ion is lower in tetracycline-induced cells (max. count per pixel 1.25×10^3) to cells from a control group (max. count per pixel 3.26×10^3). The analysis of data from negative ion mode was limited due to a lack of species corresponding to cell locations. It was possible to identify only one lipid that would correlate to a location of CHO cells. Considering the distribution of this ion within cell areas, similarly to images from positive ion mode, induced cells show different morphology to control cells (cells are bigger). However, metabolite differences between control and induced cells are not possible to determine at this stage due to limited signals coming from cells in negative ion mode.

In order to perform further investigation of collected data, a number of statistically significant metabolite peaks were chosen based on the application of *t*-test analysis ($p < 0.05$) on cell populations selected from MALDI-MSI and SIMS data sets (control and cell-induced).

There were 16 statistically significant metabolite fragments detected in SIMS positive ion mode in the mass range of 30-900 Da (**Table 3**). There were 16 fragments that corresponded to cell metabolites and were successfully correlated to cell locations within imaged areas. The fragment identification of detected ions is based using m/z values from PLS-DA. **Table 3** presents 16 metabolite fragments detected in CHO cells, of which the majority is unknown (11 fragments) and the rest of fragments can be identified as species from molecules with different origin *i.e.* lipids (6 fragments), amino acids (3 fragments) and nucleotides (4 fragments). The fragment identification is not unique and therefore, some fragments can be assigned to two groups of molecules *e.g.* fragment at m/z 178.06 could be either nucleotide or amino acid derived species.

Table 2 shows metabolite species detected using the MALDI-MSI technique and presents 23 compounds found in CHO cells, of which the majority is classified as lipids (17 metabolites). Remaining metabolites were categorised as nucleotides (1 metabolite) or unknown (5 metabolites).



All of the metabolites assigned (23 metabolites) were detected with higher intensity in induced cells, when compare to control cells. Analysis of MALDI-MSI data showed that there were 229 statistically significant endogenous compounds found in positive ion mode and 117 in negative ion mode in the mass range of 250-1000 Da after excluding matrix peaks also detected in the blank sample with DHB and PNA sprayed onto glass slides without cells. There were 23 metabolites that were successfully identified and correlated to the cell locations in positive ion mode (**Table 2**). In negative ion mode, none of 117 metabolite peaks could be successfully linked to the cell locations therefore, further analysis of images in this mode was not performed. The metabolite assignment of ions from positive ion mode is based using m/z values previously used within the *t*-test significance test. The metabolite hits with HMDB IDs and abundance in induced cells are provided. The correlation between different lipid content in cells producing EPO-Fc to control cells can be linked to the protein secretion into cell medium during exocytosis process. During exocytosis, a cell membrane surface is expanded during a fusion process with a vesicle (Roman-Fernandez & Bryant 2016, Bruns & Jahn 2002), suggesting changes in lipid metabolism and their distribution in induced cells. Similar results with increased lipid abundance in induced cells were also shown in this study in profiling of external and













internal metabolome of these cells (**Chapter 6**). Different content of PE and PC species detected in CHO cells and identified as significant could also confirm a change in lipid metabolism in cells producing EPO-Fc. Lipids play an important role in cell signalling, energy source and storage as well as membrane component and integrity (Klyszejko-Stefanowicz, 2002). PE and PC are two metabolically linked lipids that can be found in membranes of mammalian cells (Vance, 2008). One of the functions of PC is as a precursor for PE creation, which is the second most abundant (~20% of total phospholipids) phospholipid in mammalian cells (after PC) (Vance and Tasseva, 2013).

A number of PC and PE ions with different abundance in control and induced cells were detected. Spectral distribution of *e.g.* PC at m/z 824.5564 in **Figure 6** implies the presence of this species in outer parts of cells that might be recognised as possible cell membranes. In addition, the abundance of this ion is different within both control and induced cell populations and is much higher in induced cells, when compare to control population. The abundance of sphingomyelin (SM)/palmitoyl sphingomyelin ion at m/z 741.5306 is clearly enhanced in cells producing EPO-Fc (**Table 2**). Additionally, its high abundance within both cell populations and its distribution shown in **Figure 6** confirms the high outer cell membrane content in SM species. Each 5 μm pixel represents the intensity of a highlighted single ion hence, the intensity difference indicates the chemical difference between various cell locations within the image. **Figure 6** also presents ion images, illustrating variance in localisation of lipids such as CL, SM or PC within control and induced cell populations.

An important function of PE has been described, highlighting a high content of this phospholipid in the inner membrane of mitochondria. (Vance and Tasseva, 2013). Studies have shown that a decrease in the mitochondrial content of PE greatly alters mitochondrial morphology in mammalian cells (Steenbergen *et al.*, 2005). Also, another studies support these findings, where PE played a role in membrane fusion and modulated a shape of cell membrane (Martens and McMahon, 2008). Therefore, changes in lipid metabolism related to exocytosis and enhanced intracellular transport of the protein to external environment may explain higher abundance of PE (also PC) detected in protein-induced cells.

Another metabolite at m/z 758.5691 assigned as DG species is recognised to be more abundant in cells producing EPO-Fc (**Table 2**). DGs and their phosphorylated forms, for example diacylglycerophosphate (DGP), have been shown to have an influence on regulation of secretion events within cells, where DGP has been suggested to modify membrane shape and promote vesicle creation during *e.g.* exocytosis (Lev 2006, Munro 2005, Fagone & Jackowski 2009). Therefore, this would also explain enhanced synthesis of lipids in cells producing EPO-Fc.

Table 2: A total of 23 significant endogenous compounds identified within control and tetracycline-induced cells based on the analysis of cell populations selected from images in MALDI positive ion mode. The table summarises measured mass, *p* value, mass error (ppm), HMDB ID, metabolite hit and abundance in induced cells (with comparison to control): high  and low . Metabolites sorted by *p* value.

Measured mass	P value	Δ ppm	HMDB ID	Metabolite	Abundance in induced cells
824.5564	1.04E-15	0.26	HMDB00593; HMDB07888	PC(18:1/18:1); PC(14:0/22:2)	
826.5724	1.15E-11	0.12	HMDB07919; HMDB07978; HMDB08010	PC(14:1/22:0); PC(16:0/20:1); PC(16:1/20:0)	
798.54077	1.79E-09	0.05	HMDB09677; HMDB09709	PE(22:5/18:1); PE(22:6/18:1)	
822.5405	7.61E-08	0.43	HMDB07921; HMDB08136; HMDB08012	PC(14:1/22:2); PC(18:2/18:1); PC(16:1/20:2)	
785.5645	1.73E-06			Unknown	
772.5253	2.69E-06	0.35	HMDB08031; HMDB09219	PC(18:0/14:0); PE(20:0/15:0)	
741.5306	4.29E-06	0.13	HMDB13464; HMDB61712	SM(18:0/16:1); Palmitoyl sphingomyelin	
812.5199	2.46E-05	0.26	HMDB09045; HMDB09397	PE(18:1/22:6); PE(20:4/20:3)	
895.5885	3.35E-05			Unknown	
825.5596	3.78E-05	0.69	HMDB57229	CL(18:0/18:2/22:5/22:5)	
744.4942	3.84E-05	0.35	HMDB07869; HMDB08892	PC(14:0/16:0); PE(15:0/18:0)	
797.5281	3.34E-04			Unknown	

682.45703	1.01E-03			Unknown	↑
773.5284	1.30E-03	0.56	HMDB56446	CL(16:0/16:0/18:2/22:6)	↑
742.534	1.32E-03			Unknown	↑
799.5441	3.00E-03	0.69	HMDB57225 HMDB58520	CL(18:0/18:2/20:4/20:4); CL(18:1/18:1/22:6/18:2);	↑
784.5602	6.20E-03	0.32	HMDB09087; HMDB07999	PE(18:2/15:0); PC(16:1/14:1)	↑
428.0367	8.35E-03	0.15	HMDB00061; HMDB00960; HMDB01341	Adenosine 3',5'-diphosphate; dGDP; ADP	↑
745.4973	1.02E-02	0.59	HMDB10570	PG(16:0/16:0)	↑
756.5513	1.56E-02	0.23	HMDB09219; HMDB08031	PE(20:0/15:0); PC(18:0/14:0)	↑
758.5691	2.00E-02	0.87	HMDB56348; HMDB56324	DG(22:2/0:0/20:5); DG(20:3/0:0/22:4)	↑
522.3556	2.10E-02	0.55	HMDB02815	LysoPC(18:1)	↑
796.5249	2.24E-02	0.24	HMDB09710; HMDB07912; HMDB07973	PE(22:6/18:1); PC(14:1/20:1); PC(16:0/18:2)	↑

SM-sphingomyelin, PE-phosphatidylethanolamine, CL-cardiolipin, PC-phosphatidylcholine, DG- diglyceride.

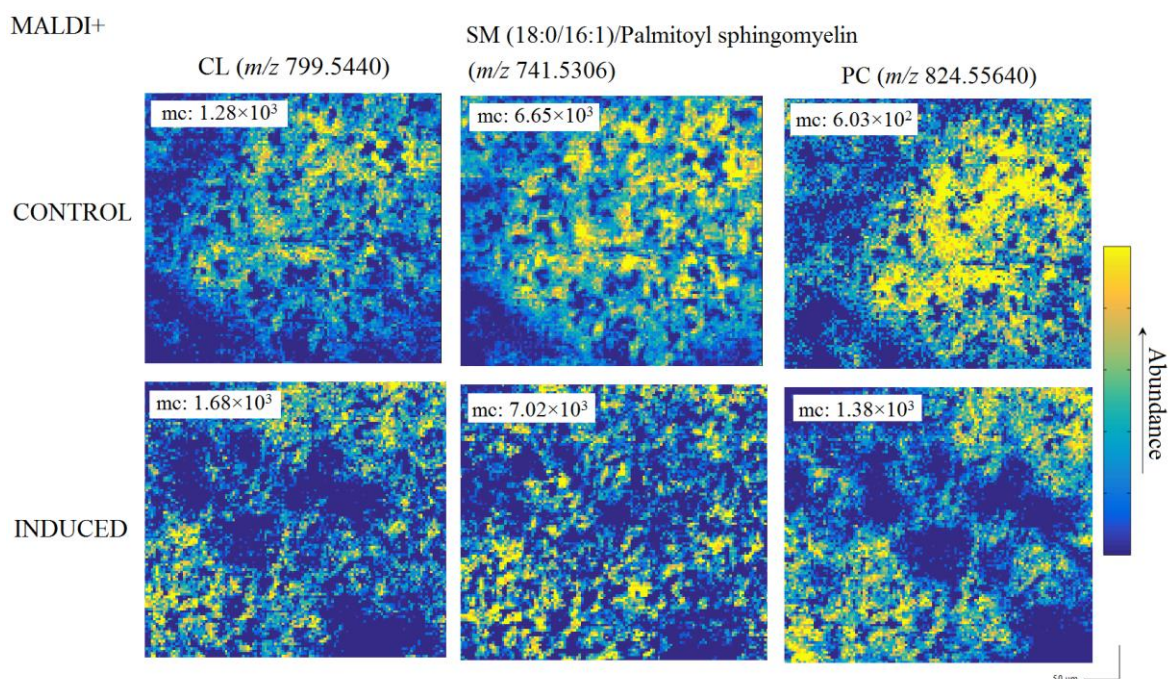


Figure 6: Distribution of possible lipid ions within control and induced cell populations. Representative single ion images of CL, SM and PC in MALDI positive ion mode are shown. Maximum count per pixel (mc) for each mass is displayed. Scale bar 50 μm , pixel size 5 μm .

The hypothesis that protein production causes cells expansion in size and their lower structural stability (due to exocytosis process) may also be explained, according to higher abundance of nucleotides (*i.e.* ADP, dGDP, adenosine diphosphate) detected in the induced cell population (**Table 2**). Unlike fingerprint profiling results where these species were detected in lower concentrations in induced cells, MALDI imaging data shows higher abundance of these ions in cells producing EPO-Fc. As discussed previously, ADP, dGDP and adenosine diphosphate take part in glycolysis and TCA cycle that occur in cytosol (glycolysis) or mitochondria (TCA cycle) (Klyszejko-Stefanowicz, 2002).

MALDI spectral maps in **Figure 7** imply the presence of these species outside cells, indicating possible leak and some level of cell destruction. **Figure 7** illustrates the ion distribution of ADP/dGDP/adenosine diphosphate at m/z 428.0368 corresponding to $[M+H]^+$ within control and induced cell populations. Spectral maps indicate higher abundance of m/z 428.0368 ion in induced cells, when compared to control cell populations, suggesting less nucleotide leakage in control cells. However, the detection of this ion can also be seen for control cells, which might imply some level of cell destruction caused by *e.g.* sample preparation process. In order to highlight these differences, PE ion at m/z 798.5407 is also shown to help in cell localisation. Additionally, the leak displays a

loss of spatial distribution of these species to some extent. Therefore, one of the biggest challenges of MS imaging of cells is the sample preservation before the analysis.

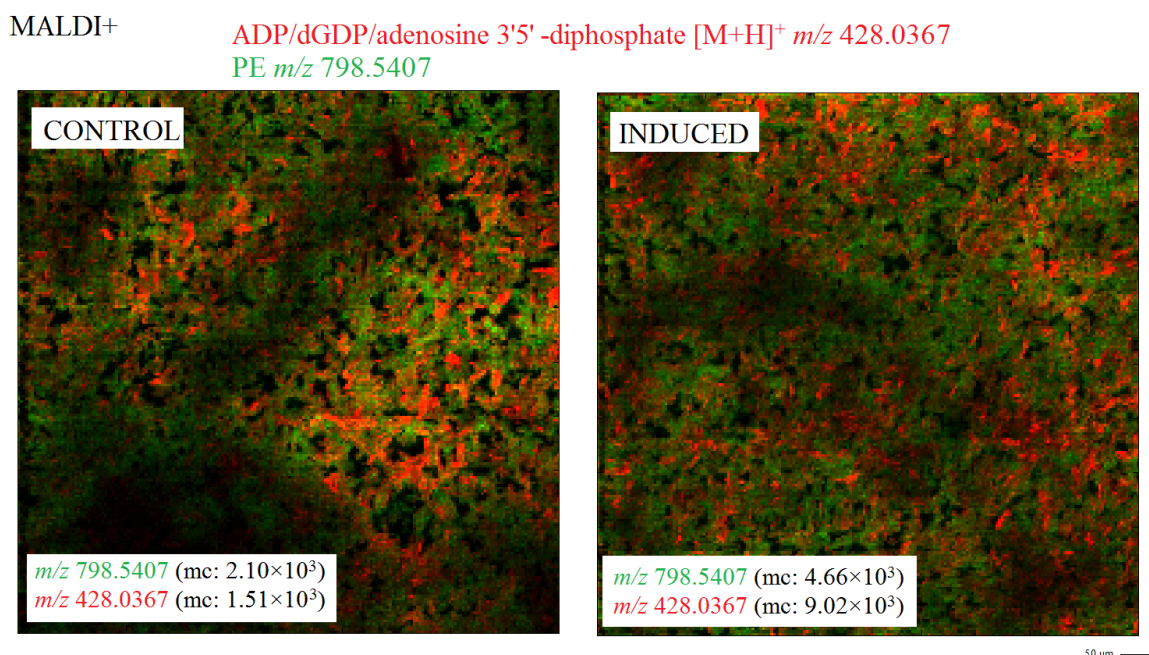


Figure 7: Distribution of ADP/dGDP/adenosine 3'5'-diphosphate within control and induced cell population. PE ion is also shown to highlight possible cell membranes. Metabolite assignments are colour-coded: *m/z* 428.0368 (red), *m/z* 798.5407 (green). Max. count per pixel: MALDI+, CONTROL: *m/z* 428.0368 (1.51×10^3), *m/z* 798.5407 (2.10×10^3); INDUCED: *m/z* 428.0368 (9.02×10^3), *m/z* 798.5407 (9.02×10^3). Scale bar 50 μm , pixel size 5 μm .

The application of ammonium formate as a fixation agent for CHO cells can be considered successful, as the shape and integrity of cell membranes seem to be preserved. Spectral distribution and localisation of lipids shown in **Figure 5** confirms that this sample preparation of cells is effective for lipid imaging. Based on a good correlation between lipid ions detected and their spectral distributions within CHO cell populations (**Figure 5 and 6**) it is therefore believed that the metabolite leak was due to the inducer exposure and protein secretion rather than fixation treatment.



The complexity and size of data generated in SIMS experiments is huge so that the correlations and trends within data sets can be very difficult to find. The interpretation of SIMS data has become even more challenging in recent years due to instrumental improvements that have given rise to enhanced sensitivity and resolution. In SIMS data, the majority of signal within spectra comes from relatively low mass fragment ions (**Figure 3 A**). It is common to see inorganic species derived from the substrate (*e.g.* silicon) occurring close to the species of interest and additionally, many isobaric overlaps







appearing from species with the same nominal but different exact masses (Fletcher *et al.*, 2011, Robinson *et al.*, 2016). Another difficulty when handling SIMS data is their large size. A common practise is the application of ‘binning’ process of the mass spectra in order to reduce data dimensions by lowering mass resolution. Lowering the mass resolution leads to a lower number of variables (or channels per pixel) that extremely hinders peak identification, mass accuracy as well as decreases contrast of the image (Fletcher *et al.*, 2011).

Based on PLS-DA scores plots for the data set generated using the SIMS technique (**Figure 3 C**), there are a number of statistically significant fragments that have been differentiated between control and tetracycline-induced cells (**Table 3**). These fragments were then compared against the Surface Spectra Static SIMS library and available literature.


One of the limitations of SIMS is that currently, there is no web-based database available. For other analytical techniques such as MALDI or GC-MS, a number of well-established libraries *e.g.* METLIN, HMDB or MassBank are accessible and are widely applied in identifying metabolic features of biological systems (Kind *et al.*, 2009, Brown *et al.*, 2009). For the TOF-SIMS platform, there is only one commercial database available, the ‘Surface Spectra Static SIMS library’ (Vickerman and Briggs, 2001) that allows for searching specific m/z values or comparing a full mass spectrum against the library in order to find matching species. This database however, is more suitable to the industrial applications of the platform (see **Chapter 1**) with only some information available for biological applications. Therefore, an additional approach used within this study was to compare metabolite fragments detected using SIMS with the available literature.

Data show that there are 6 metabolite hits according to SIMS library at: m/z 157.06, m/z 112.81, m/z 115.06, m/z 178.06, m/z 159.06 and m/z 128.06 (**Table 3**). The majority of these species might be fragments of nucleotides that would agree with MALDI data, where a nucleotide ion has been also detected (**Table 2**). However, the abundance of nucleotide species in induced cells is different for SIMS and MALDI imaging data, where higher abundance of these ions is seen in MALDI, unlike SIMS results.

Table 3: A total of 16 significant metabolite fragments detected in control and tetracycline-induced CHO cells based on the PLS-DA of cell populations selected from SIMS images in positive ion mode. The table summarises measured and theoretical mass, VIP scores, mass error (ppm), proposed fragment and possible origin of fragments, layer: 1st ‘top’, 2-10 ‘lower’, abundance in induced cells (when compare to control): high  and low . Metabolites sorted by VIP scores. Fragments were compared against current literature and *Surface Spectra Static SIMS Library software.**

Measured <i>m/z</i>	*VIP scores	**Theoretical <i>m/z</i>	Δ ppm	Proposed fragment	Origin	Layer	Abundance in induced cells	Reference
184.06	14.691, 3.4294	184.074	76.1	$[\text{C}_5\text{H}_{15}\text{NPO}_4]^+$	PC head-group	Top, Lower		(Touboul <i>et al.</i> , 2005, Jerigova <i>et al.</i> , 2011, Liu, Guo and He, 2007, Fletcher, Rabbani, <i>et al.</i> , 2013)
212.81	2.2484, 1.8095				Unknown	Top, Lower		
157.06	2.1526, 1.2989	157			2'-Deoxyuridine	Top, Lower		SIMS Library
		157.06	0.0		Unknown			(Robinson, 2013)
112.81	1.9082	113		$[\text{C}_4\text{H}_5\text{N}_2\text{O}_2]^+$	2'-Deoxyuridine	Top		SIMS Library
274.81	1.5968, 1.4818				Unknown	Top, Lower		
130.06	1.3602, 1.2622	130.066	46.1	$[\text{C}_9\text{H}_8\text{N}]^+$	Glutamic acid, glutamine or tryptophan	Top, Lower		(Kempson <i>et al.</i> 2010, Canavan <i>et al.</i> 2007, Rabbani, 2010, Fletcher, Kotze, <i>et al.</i> , 2013)

		130.06	0.0	$[\text{C}_9\text{H}_8\text{N}]^+$	Unknown			(Robinson, 2013)
115.06	1.2716, 1.695	115			Phosphatidylethanolamines	Top, Lower	↓	SIMS Library
		115.05	-86.9		Unknown			(Robinson, 2013)
189.06	1.1259, 1.5505	189.1	211.5		Glycerophosphoglycerol	Top, Lower	↑	(Tian <i>et al.</i> , 2014, Fletcher, Rabbani, <i>et al.</i> , 2013, Jerigova <i>et al.</i> , 2011)
		189.04	- 105.8		Unknown			(Robinson, 2013)
174.81	2.1745				Unknown	Lower	↓	
158.56	2.0704				Unknown	Lower	↓	
103.81	1.5402				Unknown	Lower	↓	
178.06	1.5008	178			2'-Deoxyguanosine, 2'- deoxyguanosine 5'- monophosphate, adenosine or aspartic acid	Lower	↓	SIMS Library
		178.05	-56.1		Unknown			(Robinson, 2013)
158.81	1.4834				Unknown	Lower	↓	
159.06	1.3719	159		$[\text{C}_{10}\text{H}_{11}\text{N}_2]^+$	Tryptophan	Lower	↓	SIMS Library
128.06	1.3183	128	-78.1		2'-Deoxythymidine	Lower	↓	SIMS Library
		128.05		$[\text{C}_6\text{H}_8\text{O}_3]^+$	Unknown			(Robinson, 2013)

165.06	1.1732	165.1	242.3		Glycerophospholipids	Lower		(Tian <i>et al.</i> , 2014, Fletcher, Rabbani, <i>et al.</i> , 2013, Jerigova <i>et al.</i> , 2011)
		165.09	181.7	[C ₁₀ H ₁₃ O ₂] ⁺	Vitamin E			(Robinson, 2013)

*-based on PLS-DA Component 1; **-theoretical values calculated based on fragment formulas or based on experimentally measured m/z from SIMS Library or literature; ***-more info: <http://surfacespectra.com/simslibrary>

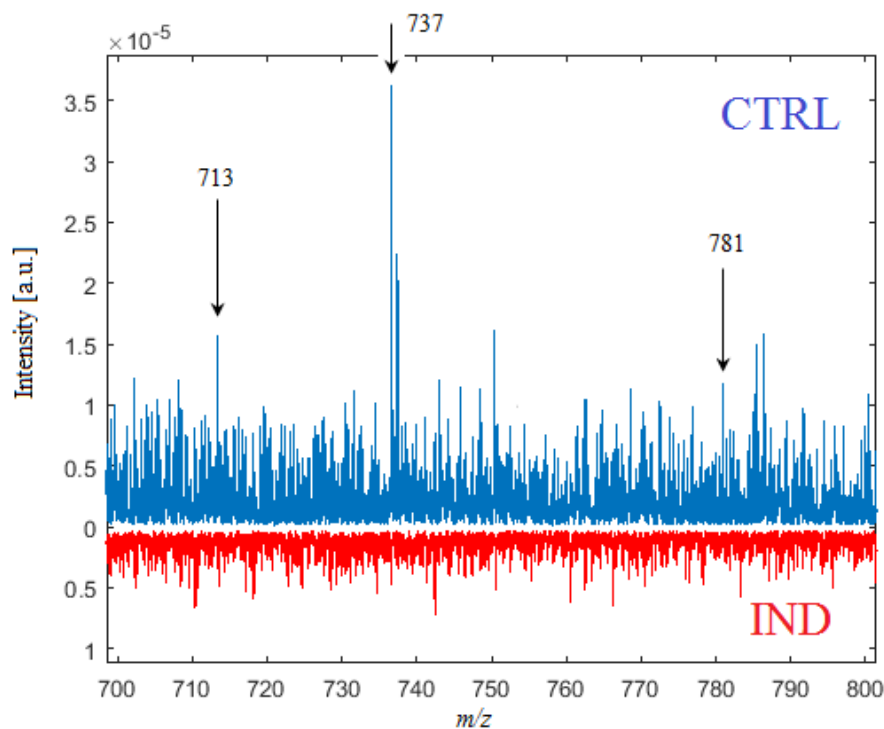
On the other hand, SIMS imaging data, where low abundance of nucleotide species is observed, are supported by results of metabolite profiling of CHO cells producing EPO-Fc studies (**Chapter 6**), where abundance of these ions was also decreased in cells producing the protein. In addition, the abundance of lipid-origin fragments at m/z 184.06, m/z 115.06 and m/z 165.06 observed in SIMS data is lower in induced cells when compared to control cells, where similar results have also been seen in metabolite profiles of these cells, presented in **Chapter 6**. On the contrary, the abundance of possible lipid fragment at m/z 189.06 is higher in cells producing EPO-Fc protein that would agree with MALDI data, where high abundant lipid species have been detected. Additionally, some amino acid fragments at m/z 130.06 and m/z 159.06 are also observed using the SIMS technique.

Fragments detected using SIMS and derived from cells can also be confirmed with other studies available in the literature. There have been a number of metabolite fragments detected in breast cancer tumors, HeLa cells and fibroblasts (Robinson, 2013, Kempson *et al.* 2010, Canavan *et al.* 2007, Rabbani, 2010, Fletcher, Kotze, *et al.*, 2013). Some of these ions have been successfully assigned to metabolite IDs, for example the ion at m/z 130.06 as a possible fragment of glutamic acid, glutamine or tryptophan (Canavan *et al.* 2007), m/z 189.06 as fragment of glycerophosphoglycerol (Fletcher, Rabbani, *et al.*, 2013, Jerigova *et al.*, 2011) and m/z 165.05 as fragment of glycerophospholipids (Fletcher, Rabbani, *et al.*, 2013) or vitamin E (Robinson, 2013).

The presence of lipid fragments gave a rise for further investigation of SIMS spectra in order to analyse possible differences between control and induced cells in terms of lipid content. **Figure 8** presents SIMS spectra from the top (1st) and lower (2-10th) layers for control and tetracycline-induced cells within a mass range of 700-800 Da with some high mass species highlighted. Some high mass species at m/z 713, m/z 737 and m/z 781 have been detected and are found to be less abundant in induced cells when compared to control cells. This would also confirm the lower abundance of possible lipid fragments in induced cells seen in **Table 3**. These high mass species have not been yet described in the literature, however, in this study they might be important (high abundance). Additionally, a low number of significant ions (according to *t*-test) within a mass range of m/z 700-800 suggest that some chemical damage has already occurred under the primary ions dose used (2.8×10^{13} or 4.4×10^{13} ions/cm² depending on the day of the analysis). Nevertheless, in order to investigate lipid content of tetracycline-induced cells, more experiments need to be performed, particularly with a targeted approach, where lipid and nucleotide standards would give the confidence of species detected.

SIMS+

Top
layer



Lower
layers

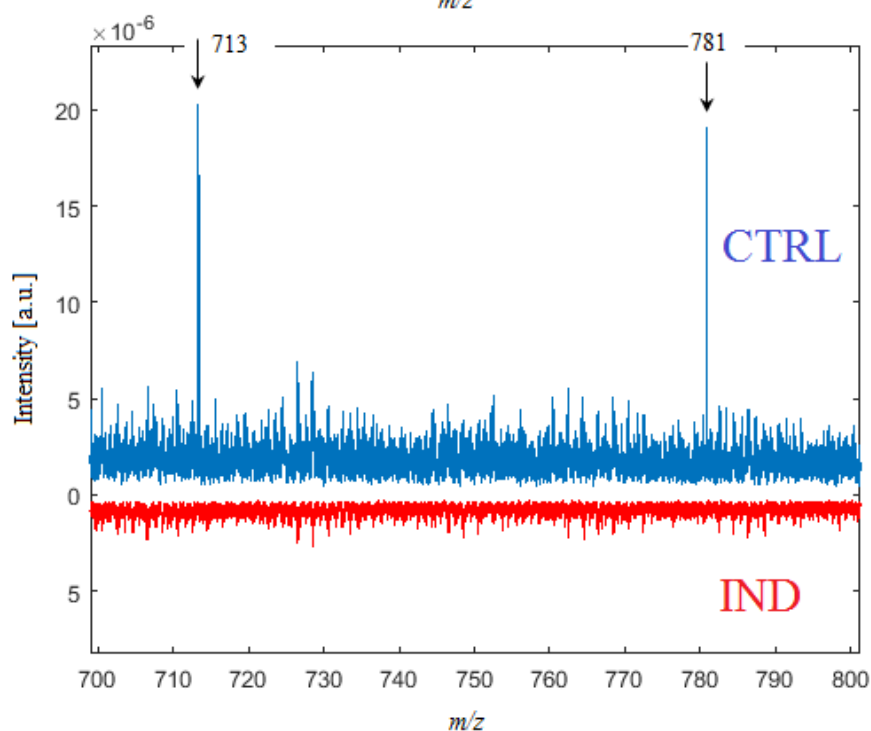


Figure 8: SIMS+ spectra from the 1st (top layer) and 2-10th (lower layers) for control (CTRL, blue) and tetracycline-induced cells (IND, red) within a mass range of 700-800 Da; 40 keV C_{60}^+ beam. Different abundance of lipid species between control and induced cells is shown.

Polyatomic ion sources such as C_{60}^+ allow for the analysis of biological samples in three dimensions as a function of depth profiles and with low sub-surface damage due to the

lower energy density than *e.g.* the Au_3^+ source. However, because the C_{60}^+ ion beams are also difficult to focus, the imaging with C_{60}^+ sources is more likely limited to 1-2 μm per pixel, depending on sensitivity requirements (Lanni *et al.* 2012, Fletcher *et al.*, 2013).

In this study, imaging with the C_{60}^+ primary ion beam was applied in order to image 10 layers of CHO cells. Despite the fact that **Figure 8** shows some differences in lipid content between the top and lower layers of imaged cells, the overall chemical difference between the 1st layer and 2-10th layer is not significant. Considering the full mass range that was investigated (30-900 Da), the difference between chemical content for the top and lower layers of both control and induced cell data sets is very little as shown in PCA scores plots, where almost no separation can be seen (**Supplementary Information, Figure 22**).

Representative single ion images of fragments detected within control and induced CHO cells are presented in **Figure 9**, where a location of cell populations within the area imaged is shown. Images present the location of cell populations however, it is impossible to distinguish single cells within the imaged area due to high cell coverage of both samples as well as limited spatial resolution ($\sim 8 \mu\text{m}$). Therefore, further investigation with the application of the SIMS technique with more focused C_{60}^+ primary ion beam could be performed that would allow for single cell analysis. Additionally, more experiments should be executed in order to increase number of repeats to enhance the confidence of statistical analysis.

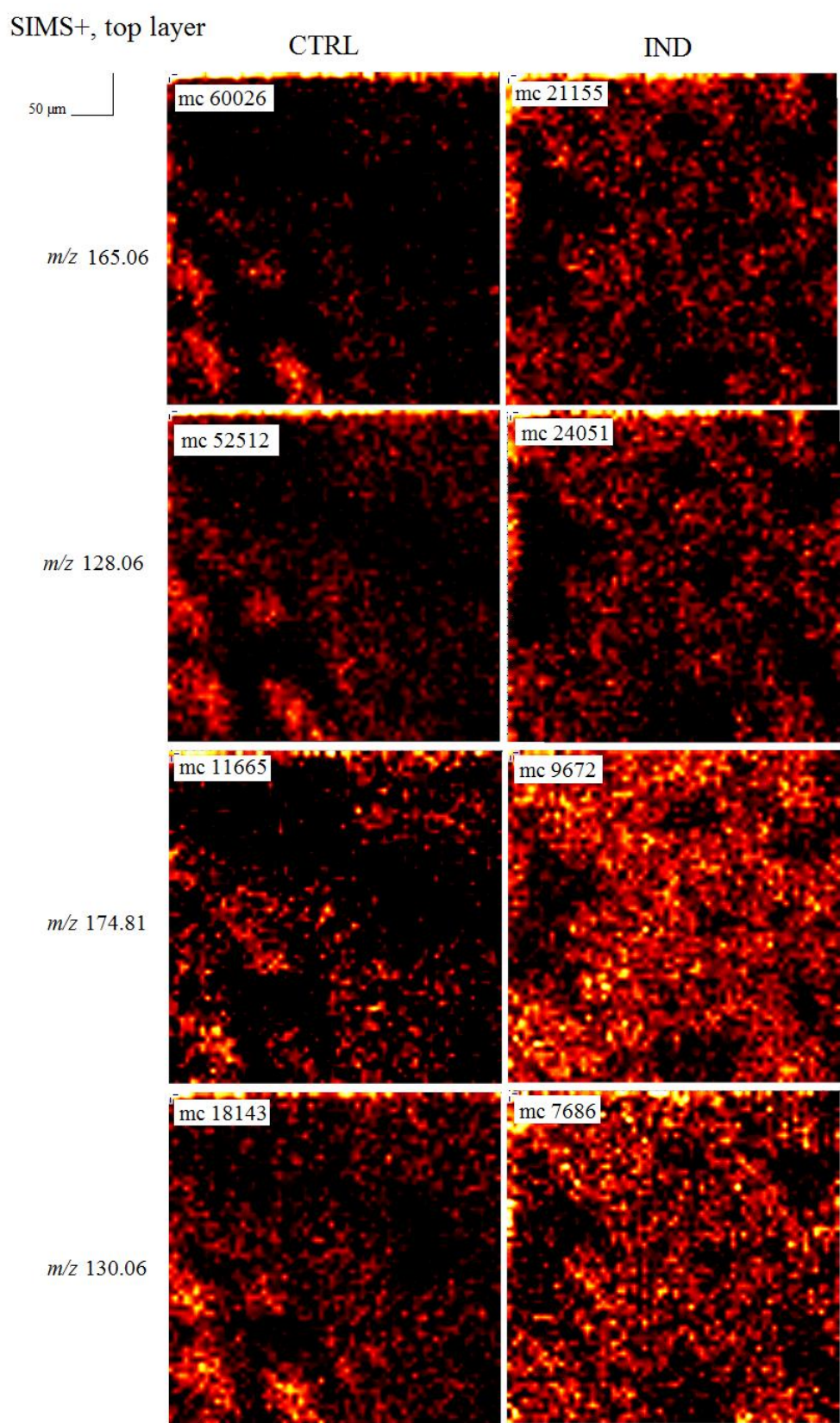


Figure 9: Representative SIMS images of statistically significant metabolite fragments detected in control and induced CHO cells. Images illustrate the location of each ion within the area imaged. Maximum count per pixel (mc) for each mass is displayed. Scale bar 50 μm , pixel size $\sim 8 \mu\text{m}$, 40 keV C_{60}^+ beam.

7.5.1.2. Single cell-based study

Further investigations focus on analysis of cell metabolism response to the protein production based on a single cell study. **Figure 10** presents the overlay image of three lipids at m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red) detected in MALDI+ within control and induced cell populations.

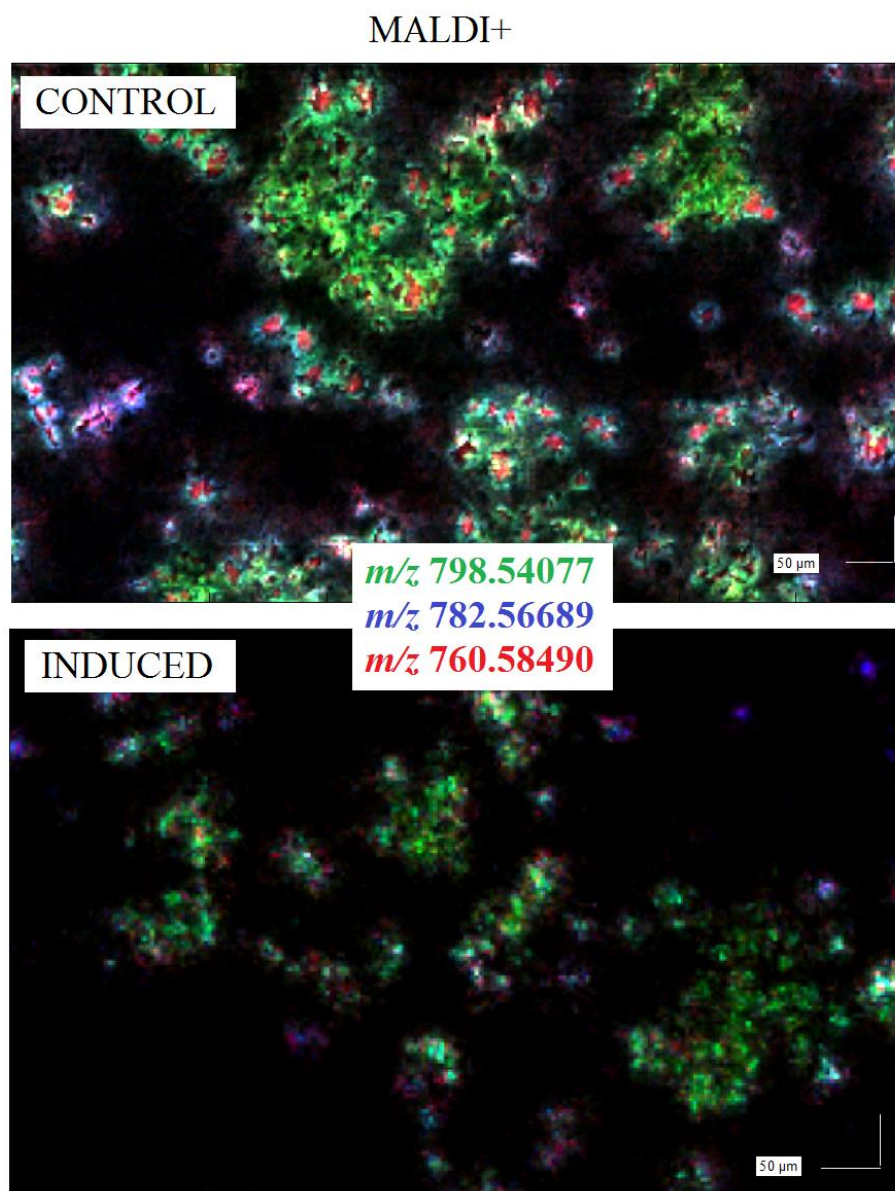


Figure 10: MALDI-MS images of control and induced cells in positive ionisation mode. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red) is shown; max. count per pixel: MALDI+, CONTROL: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3); INDUCED: m/z 798.54077 (6.74×10^3), m/z 782.56689 (4.56×10^3) and m/z 760.5849 (2.18×10^4). Scale bar 50 μm , pixel size 5 μm .

Within each control and induced cell populations, 10 single cells were chosen for multivariate analysis, of which enlarged images are displayed in **Figure 11**. In order to explore metabolic changes in cells producing protein and compare this information to results obtained based on cell populations; *t*-test and PCA were performed based on single cells data. 2D and 3D PCA scores plots were used to illustrate diversity between control cells and induced cell population. As seen in **Figure 11**, 3D PCA scores plot created for single cell data using first three components account for 48.1% of total variance between control and induced cells. (Inset) 2D PCA scores plot illustrate some level of separation between these two groups of cells, accounting for 28.1% of variance on PC1.

Comparing the control and induced cells variance from the analysis based on cell populations, the variance based on single-cell analysis is very similar with the difference of 2% (46.1% of total variance for cell populations). 2D and 3D PCA scores plots also highlight some cell diversity within control and cell-induced groups. Control cells show higher level of heterogeneity than cells producing EPO-Fc, which can also be seen in the single cell images produced from both of cell populations.

Similarly to the analysis of cell populations, the next step of single cell data examination included metabolite assignment of statistically significant compounds influenced by inducer exposure and protein production. Statistically significant metabolites were chosen based on the application of *t*-test analysis ($p < 0.05$) on the data extracted from single cells. Out of top 50 the most significant compounds found, there were only 6 metabolites successfully identified and correlated to spectral maps of control and induced cells (**Table 4**). **Table 4** summarises measured mass, *p* value, mass error (ppm), HMDB ID, metabolite hit and abundance of each metabolite in control and induced cells (high or low). Metabolites that were assigned are categorised as lipids (3 metabolites); for the remaining 3 compounds the identification was unsuccessful *i.e.* no hits were found in the database.

MALDI+ m/z 798.54077, m/z 782.56689, m/z 760.5849

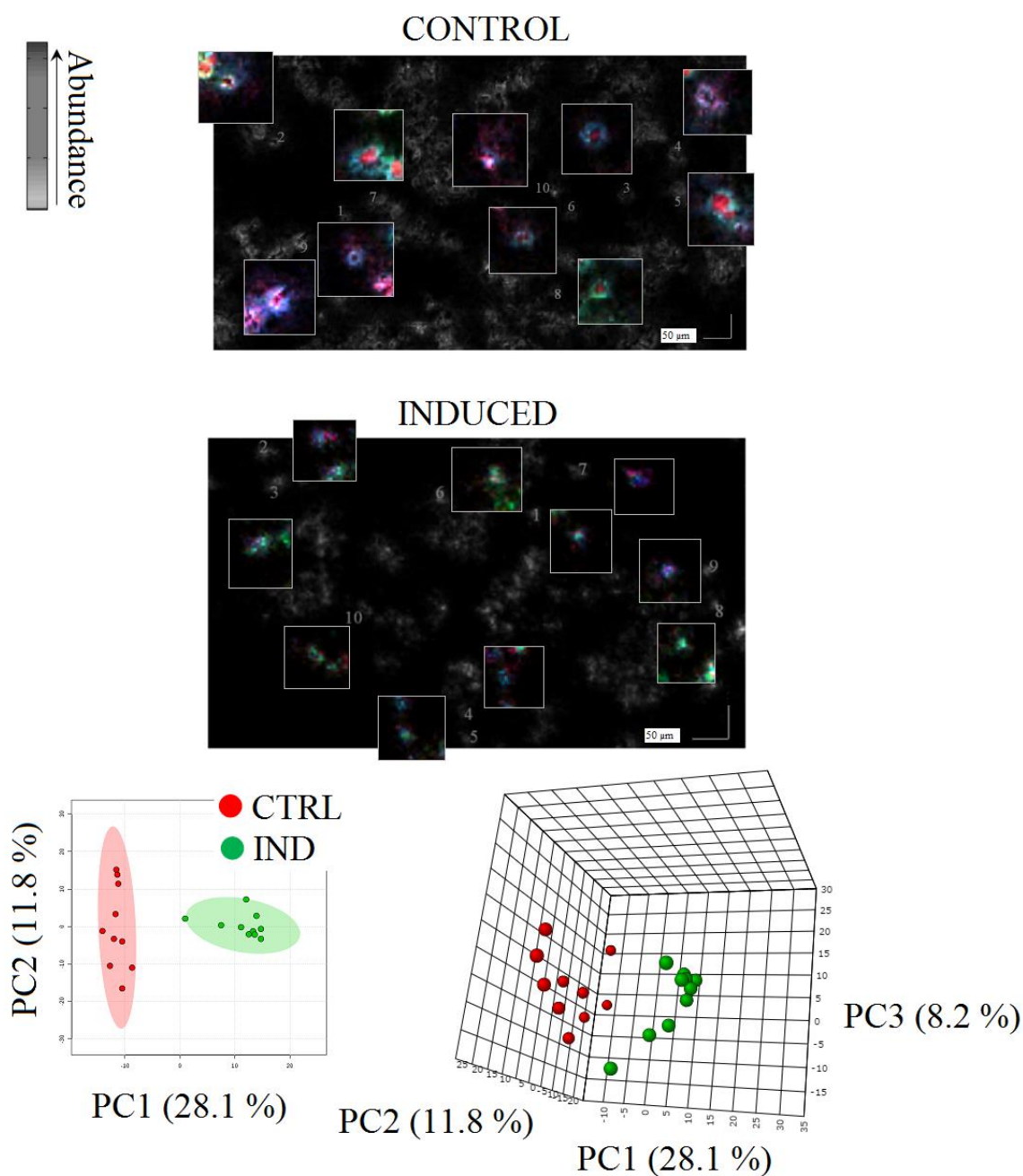










Figure 11: Single cell analysis of CHO control and induced cells for MALDI imaging in positive ion mode. Overlay (colour-coded) of three species: m/z 798.54077 (green), m/z 782.56689 (blue) and m/z 760.5849 (red). Enlarged single cells selected for multivariate analysis are shown. Scale bar 50 µm (zoom out image), pixel size 5 µm. (Inset) 2D and 3D PCA scores plot accounting for 48.1% of total variance (first three components). PCA plots illustrate the variance between control (CTRL, red) and induced cell areas (IND, green) with highlighted 95% regions of confidence.

Table 4: The most significant metabolites identified within control and tetracycline-induced cells based on the analysis of single cells selected from images in MALDI positive ion mode. The table summarises measured mass, *p* value, mass error (ppm), HMDB ID, metabolite hit and abundance in induced cells (with comparison to control): high  and low . Metabolites sorted by *p* value.

Measured mass	P value	Δ ppm	HMDB ID	Metabolite	Abundance in induced cells
673.4244	3.79E-05			Unknown	
796.5251	8.90E-05	0.09	HMDB09710 HMDB11427	PE(22:6/18:1) PE(18:1/22:6)	
759.5724	0.000103			Unknown	
771.5126	0.000108	0.67	HMDB57515 HMDB57551	CL(16:1/16:1/18:2/22:6) CL(16:1/16:1/22:6)	
760.5851	0.000141	0.00	HMDB08263 HMDB09516	PC(20:0/14:1) PE(22:1/15:0)	
917.5387	0.00018			Unknown	

PE-phosphatidylethanolamine, CL-cardiolipin, PC-phosphatidylcholine.

The abundance of all endogenous compounds listed in **Table 4** vary between control and cells producing EPO-Fc protein. Decreased intensities of PE, CL and metabolites at m/z 673.4244 and m/z 917.5387 and increased abundance of PC and metabolite at m/z 759.5724 were seen in induced cells. Increased abundance of some metabolite ions within induced cells have also been detected based on the study of cell populations shown in previous section. This suggests that single cell-based analysis provides similar information about the metabolic changes in CHO cells under protein production. What is more, single cell-based analysis provides additional insights about changes in metabolism of induced cells that can be referred to a unique, single cell. Different abundance of metabolite identities in induced cells from the single-cell analysis to the analysis of cell populations, where all of the hits are more abundant in these cells can also imply that single cell-based investigations provide with more specific knowledge that can be linked to unique cells within given cell population.

Additionally, the differences between control and induced cells can be illustrated at the single cell level, highlighting different localisation of species within a single cell (**Figure 12**). Ion images present enlarged cell with highlighted spectral differences in the distribution of three metabolites within single cells. Images of the single cells with three identified lipids previously chosen are also shown in order to support the single cells regions within cell populations. In control cells, ion at m/z 917.5387 is distributed peripherally, whereas ion at m/z 759.5724 can be distinguished within a centre of a cell. Also, the ion at m/z 796.5251 identified as PE can be seen across the whole surface of the control cell. On the other hand, both species at m/z 759.5724 and PE are spread homogeneously across the whole area of the tetracycline-induced cell. However, the ion at m/z 917.5387 varies in its distribution within the induced cell greatly, when compared to the control cell. The species is distinctly detected within a centre of the induced cell, in contrast to external localisation in the control cell. This illustrates the great advantage of single cell-based imaging providing mass correlation and spectral integrity of biological samples at sub-cellular level. The different distribution of these ions also follows different intensity of identified compounds as this is lower in induced cells (max. count per pixel (mc): 4.50×10^2) when compare to control cells ((mc): 1.51×10^3) for ion at m/z 917.5387 as well as for ion at m/z 796.5251 (mc; induced: 1.36×10^3 , control: 3.10×10^3). The intensity of ion at m/z 759.5724 seems to be similar in both, induced and control cells (**Figure 12**).

MALDI+

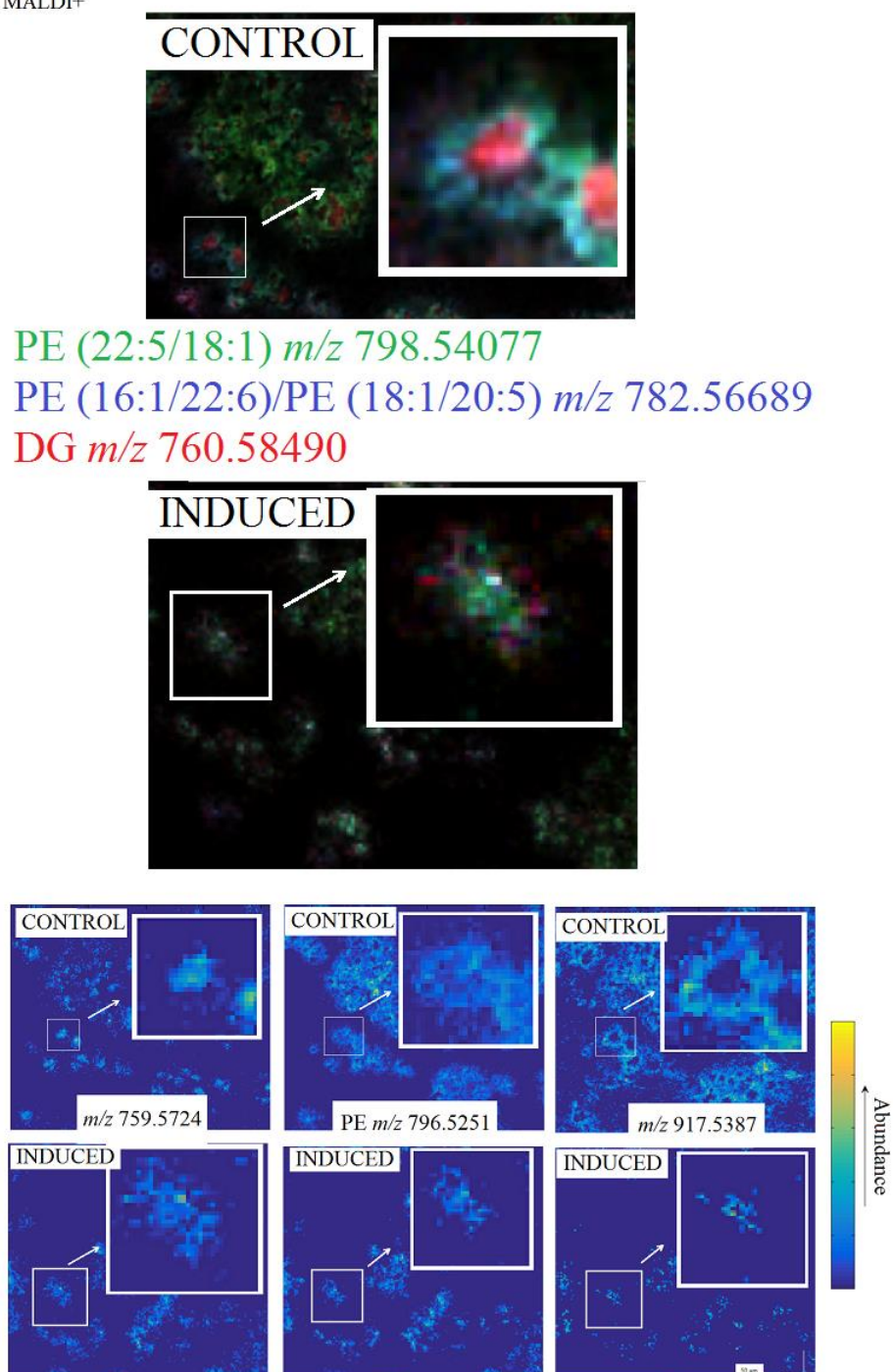


Figure 12: Representative MALDI+ single ion images of the most significant metabolites from single cell-based analysis of control and induced cells. A single cell highlighted from control and induced cell populations is enlarged and inset into ion images to illustrate location differences between both cell populations. Max. count per pixel: CONTROL: m/z 759.5724 (1.09×10^3), m/z 796.5251 (3.10×10^3), m/z 917.5387 (1.51×10^3); INDUCED: m/z 759.5724 (1.70×10^3), m/z 796.5251 (1.36×10^3), m/z 917.5387 (4.50×10^2). Scale bar 50 μm (zoom out image), pixel size 5 μm .

7.5.2. Imaging of CHO cells - single cell and sub-cellular approach

Single cell analysis with the application of SIMS and Au_3^+ primary ion gun to CHO cells in positive ionisation mode was performed. The study included imaging control (non-induced) cells from three different areas within the same sample (areas A, B and C). In order to collect extensive knowledge about the sample, different instrument settings were employed for each of the mapped areas. Detailed instrument settings are described in **Chapter 3**. **Figure 13** shows total ion count (TIC) SIMS+ spectra and TIC images collected from three areas within control CHO cells. Table presented next to the spectra (also shown in **Chapter 3, Table 1**) summarises brief instrument set-up used for each experiment performed. The overall SIMS spectra indicate rich mass information. As discussed previously, the SIMS technique is known as a “hard ionisation” method because of the high level of sample damage around the area, where the initial impact of high-energy primary ions occurs. The damage of a sample when using the Au_3^+ primary ion gun is high (when compare to *e.g.* C_{60}^+) however, this primary ion gun provides the highest spatial resolution *i.e.* 0.5 μm , (Vickerman 2011, Lanni *et al.* 2012), where a pixel size is an order of magnitude smaller than that used in the high resolution MALDI imaging instruments (in this study 5 μm) and SIMS instrument with C_{60}^+ primary ion beam with a pixel size $\sim 8 \mu\text{m}$.

The metabolites and their fragments are detected across a mass range of 20-400 Da for areas A and B and mass range of 70-500 Da for area C. Area A was scanned in default instrument settings within a mass range of 20-800 Da. The spectra obtained illustrate a number of ions detected with high abundance from m/z 20 to about m/z 230. Ions detected in a mass range of m/z ~ 230 to m/z 400 are much less abundant. No ions above m/z 400 are seen. Some similarities to the spectrum from area A can be found in a spectrum obtained from area B that was scanned with ‘low quad’ settings, focusing on low mass ions. Likewise, high abundant ions appear in a mass range below m/z ~ 230 and ions with low intensities can be noticed within a mass range of m/z ~ 230 to m/z 400. Also, there are no ions detected above m/z 400. Area C was scanned with ‘high quad’ instrument settings in a mass range of m/z 70-1000 with additional focus on high mass ions. It can be seen from **Figure 13** that the spectrum provides highly abundant (~ 220 a.u.) and rich ion information between m/z 70-300. There are a number of ions detected above m/z 300 with noticeable intensities and also, some ions can be found within higher mass ranges *i.e.* above m/z 400. **Figure 13** also presents total ion count images of each cell area. Images illustrate a highly focused ion image with cellular and sub-cellular shapes preserved. However, some instrumental artefacts appearing as horizontal lines across TIC images are seen.

CONTROL, SIMS+

TIC

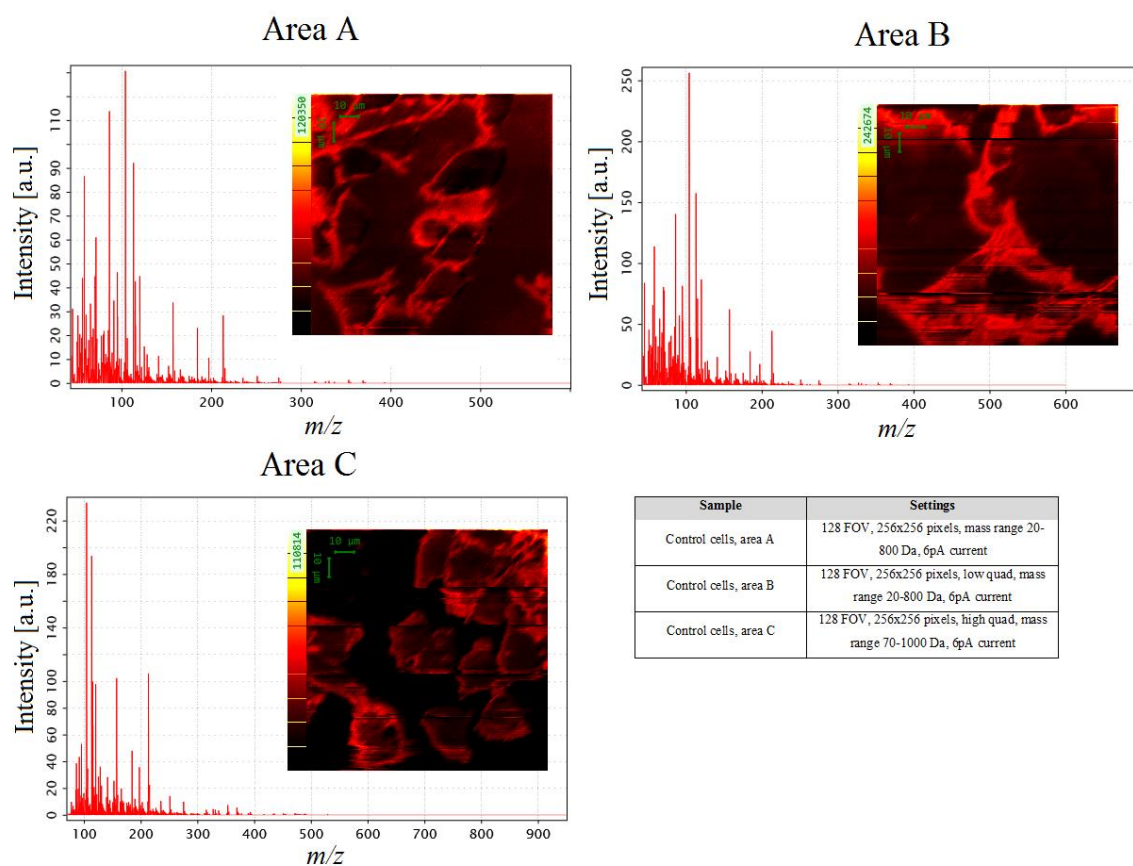


Figure 13: Total ion count (TIC) SIMS+ spectra and TIC images from control CHO cells from three different areas (A, B, C) within the same sample. Attached table shows instrument settings and mass range of each area. Scale bar 10 μm (green, left top corner of each image), pixel size 0.5 μm ; 25 keV Au_3^+ beam.

Table 5: Fragment ions detected in control CHO cells using SIMS+. The table summarises type of compound, measured and theoretical m/z values, mass error (ppm), proposed fragment and possible origin of fragments. Fragments were compared against current literature.

Compound	Measured m/z	*Theoretical m/z	Δ ppm	Proposed fragment	Origin	Reference
Lipids	184.09	184.074	87.5	$[\text{C}_5\text{H}_{15}\text{NPO}_4]^+$	PC head-group	(Touboul <i>et al.</i> , 2005, Jerigova <i>et al.</i> , 2011, Liu, Guo and He, 2007, Fletcher, Rabbani, <i>et al.</i> , 2013)
	86.10	86.097	34.8	$[\text{C}_5\text{H}_{12}\text{N}]^+$		
	104.11	104.108	24.0	$[\text{C}_5\text{H}_{14}\text{NO}]^+$		
	165.11	165.1			Glycerophospholipids	(Tian <i>et al.</i> , 2014, Fletcher, Rabbani, <i>et al.</i> , 2013, Jerigova <i>et al.</i> , 2011)
	166.07	166.063	40.3	$[\text{C}_5\text{H}_{13}\text{NPO}_3]^+$	Glycerophosphocholine	
	189.10	189.1			Glycerophosphoglycerol	
	369.37	369.352	48.5	$[\text{C}_{27}\text{H}_{45}]^+$	Cholesterol	(Tian <i>et al.</i> , 2014)
	313.25	313.274	-77.6	$[\text{C}_{19}\text{H}_{37}\text{O}_3]^+$	MAG C16:0	
	577.54	577.520	35.3	$[\text{C}_{37}\text{H}_{69}\text{O}_4]^+$	DG C34:1	
Amino acids	120.09	120.081	72.5	$[\text{C}_8\text{H}_{10}\text{N}]^+$	Phenylalanine, glycine or methionine	(Kempson <i>et al.</i> 2010, Canavan <i>et al.</i> 2007, Rabbani, 2010, Fletcher, Kotze, <i>et al.</i> , 2013)
	130.06	130.066	-43.8	$[\text{C}_9\text{H}_8\text{N}]^+$	Glutamic acid, glutamine or tryptophan	
	70.07	70.066	61.4	$[\text{C}_4\text{H}_8\text{N}]^+$	Arginine, asparagine, leucine, proline or valine	
	81.05	81.045	58.0	$[\text{C}_4\text{H}_5\text{N}_2]^+$	Histidine	
Nitrogenous bases	119.04	119.036	35.3	$[\text{C}_5\text{H}_3\text{N}_4]^+$	Adenine	(John S. Fletcher <i>et al.</i> , 2013)
	127.06	127.051	72.4	$[\text{C}_5\text{H}_7\text{N}_2\text{O}_2]^+$	Thymine	
	135.04	135.031	68.9	$[\text{C}_5\text{H}_3\text{N}_4\text{O}]^+$	Guanine	

*-theoretical values calculated based on fragment formulas; PC-phosphatidylcholine; MAG- monoacylglycerol; DG-diglyceride

Some of the species usually imaged in cells in SIMS experiments were chosen in order to create single images to map their distribution in the CHO cells (**Table 5**). As stated previously, lipids are the major element of living cells (Klyszejko-Stefanowicz, 2002), hence, lipid fragments were first selected to image their localisation within the cell population. The highly abundant phosphocholine head group at m/z 184 was used as a marker for cells. Additionally, PC head groups fragments at m/z 104 and m/z 86 (which can also be identified as leucine ion) are also seen (Denbigh & Lockyer 2015, Rabbani *et al.* 2011). **Figure 14** presents the distribution of PC head group and its fragments within the plasma membrane of CHO cells.

CONTROL, SIMS+, Area B, C

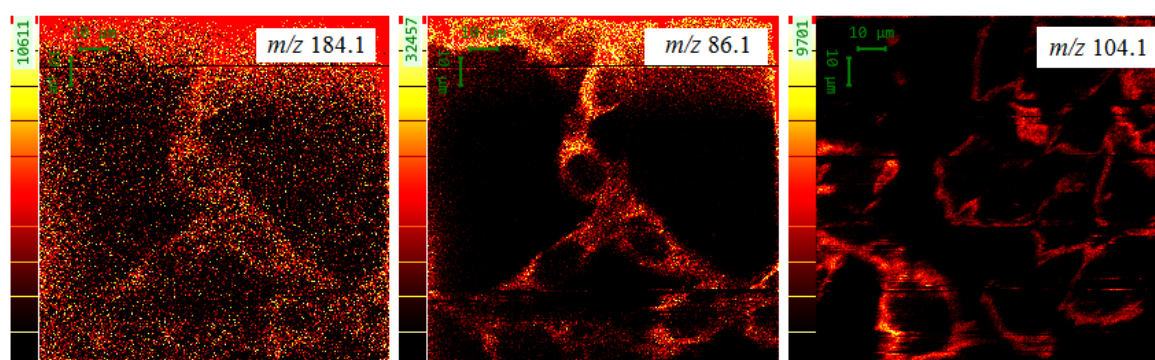


Figure 14: SIMS images of PC head group m/z 184 and its fragments of CHO cells. PC head group at m/z 184.1, fragments at m/z 86.1 and m/z 104.1 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam.

Additionally, fragments of lipids derived from cells that have been described in the literature (Tian *et al.*, 2014) were also found in CHO cells (**Figure 15**). The common fragment at m/z 165.1 for all glycerophospholipids as well as the glycerophosphocholine fragment at m/z 166.1 and glycerophosphoglycerol fragment at m/z 189.1 and cholesterol fragment at m/z 369.4 appear to be localised across the whole cell surface. However, some localisation of these ions, especially m/z 165.1, m/z 166.1 and m/z 189.1 species within the cell plasma membrane can be observed. Additionally, some differences in the abundance of these ions from different cells are also noticeable. The ion at m/z 369.4 appears to be more intense within one cell, when compare to other cells from the population (**Figure 15**).

CONTROL, SIMS+, Area C

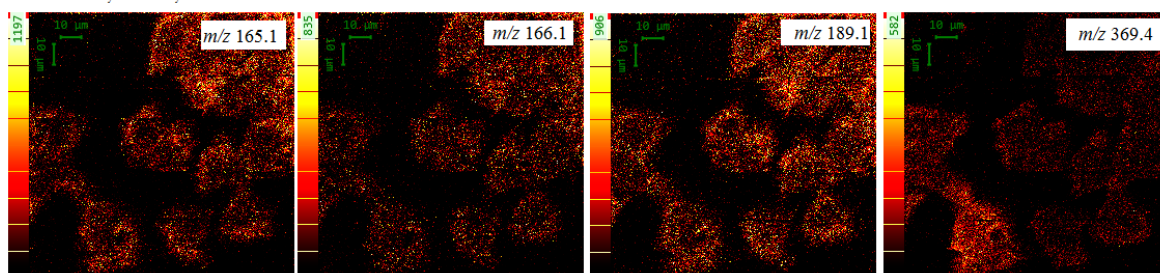


Figure 15: SIMS images of lipid fragments within CHO cells. Fragments at m/z 165.1, m/z 166.1, m/z 189.1 and m/z 369.4 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , spatial resolution <1 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam.

Additionally, other metabolites such as (potentially) amino acids were imaged (**Figure 16**). The ion at m/z 120 has been recognised as a fragment of phenylalanine, glycine or methionine (Kempson *et al.* 2010, Canavan *et al.* 2007, Rabbani, 2010) and is distributed within the whole cell surface. Similar, spread distribution is observed for other ions such as m/z 130 (glutamic acid, glutamine or tryptophan) and m/z 81 (histidine) that have also been detected in cell studies and have been identified as important elements in the cell growth (Kempson *et al.* 2010, Canavan *et al.* 2007). The ion at m/z 70 (arginine, asparagine, leucine, proline or valine) however, appears within a whole cell surface but also seems to be more condense within cell borders.

CONTROL, SIMS+, Area B, C

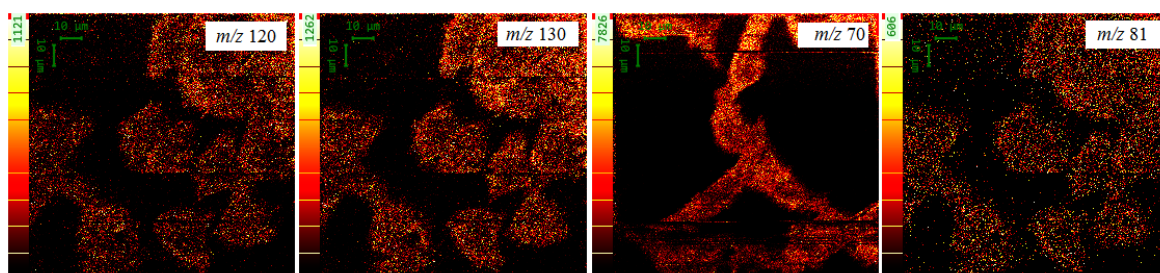


Figure 16: SIMS images of amino acid fragments within CHO cells. Fragments at m/z 120, m/z 130, m/z 70 and m/z 81 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μm , pixel size 0.5 μm ; 25 keV Au_3^+ beam.

The distribution of particular ions can be dependent on the dose of the primary ion beam per cm^2 . Described ions at m/z 120, m/z 130, m/z 70 and m/z 81 could also be derived from proteins present in the cytoplasm of cells and therefore, appear over the whole cell area. This could be confirmed with the primary ion fluence that was used in this study *i.e.* 7.5×10^{13} ions per cm^2 , suggesting that the static limit ($<10^{13}$ ions per cm^2) was exceeded.

This implies that the analysis was probably beyond the surface layer, exposing the interior of the cells to the ion beam.

The spread distribution of ions at m/z 119, m/z 127 and m/z 135 within the cells is observed, when deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are imaged (**Figure 17**). DNA and RNA are two types of nucleic acids that occur in mammalian cells. DNA is a double strands molecule that consists of a sugar (deoxyrybose), phosphate group and one of the four nucleobases *i.e.* adenine, guanine, cytosine, thymine. DNA carries genetic information and is found mainly in cell nucleus, but also in mitochondria or chloroplasts in eukaryotic cells. Similarly, RNA consists of a sugar (ribose), phosphate group and one of the four bases *i.e.* adenine, guanine, cytosine, uracil but unlike DNA, RNA is shaped as a single strand molecule. Different types of RNA play various roles in cells *inter alia*. coding and regulation of genes and can be found either in nucleus or in cytoplasm of the cell (Klyszejko-Stefanowicz 2002, Albi 2011). Fragments of nitrogenous bases adenine at m/z 119 [$C_5H_3N_4$]⁺, guanine at m/z 135 [$C_5H_3N_4O$]⁺ and thymine at m/z 127 [$C_5H_7N_2O_2$]⁺ were imaged. The distribution of these ions rather than localised to the centre of the cells where the nuclear DNA and RNA occur, is localised within the whole cell. In addition, the signal intensity of these ions appears to be low, when compared to TIC. Apart from reasons mentioned previously (*i.e.* high sub-surface damage with Au_3^+ , exceeded static limit), there might be a number of other reasons for this *e.g.* low ionisation probability that are inherent to the measurement process and do not suggest any particular conclusions about the analyte distribution or its concentration (Mouhib *et al.*, 2013, Fletcher *et al.*, 2011, Zenobi, 2013).

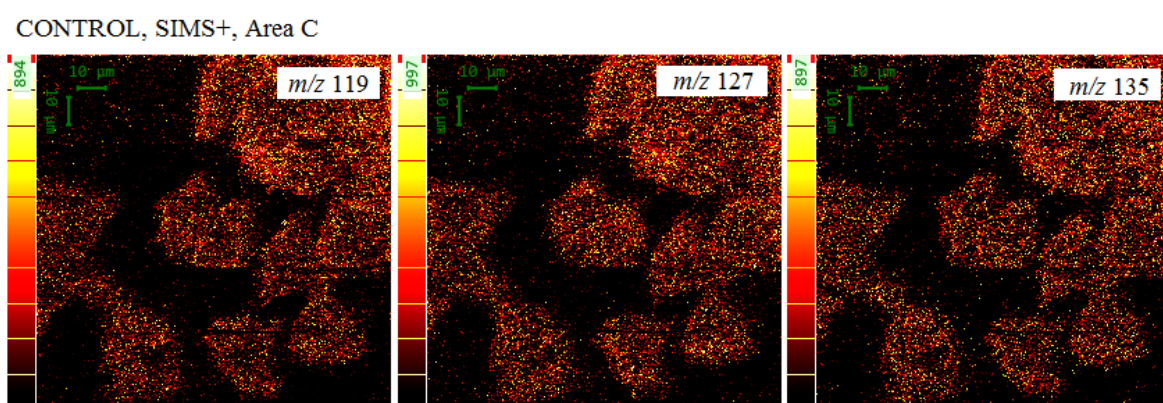


Figure 17: SIMS images of nitrogenous bases' fragments within CHO cells. Fragments at m/z 119, m/z 127 and m/z 135 are shown. Maximum count per pixel for each mass is displayed (in green). Scale bar 10 μ m, pixel size 0.5 μ m; 25 keV Au_3^+ beam.

Some consistency between MALDI imaging data and SIMS imaging data using Au_3^+ source can also be observed. Capabilities of SIMS and MALDI imaging are gathering, which can be illustrated in **Figure 18**, where the localisation of lipids in the CHO cells was compared. **Figure 18** presents sub-cellular compartments of the cells, with different class of lipids highlighted: PE and DG for MALDI and PC, DG and monoacylglycerol (MAG) for SIMS. The two experiments performed using MALDI and SIMS platforms demonstrate positive correlation of spatial features within the cells and hence, show that comparable results can be achieved with very different MS imaging methods. Phospholipids such as PC, PE and PS occur mostly in the plasma membrane, however PCs are also found in nuclear membrane along with CLs and high abundant DGs (Albi 2011, Klyszejko-Stefanowicz 2002). Similar cell structures are acquired in the low micrometer range. The distribution of diglycerides at m/z 760.5849 and m/z 576 in images acquired from MALDI and SIMS respectively, illustrates their high abundance within the nuclei. On the other hand, phospholipids such as PE at m/z 798.54077 and m/z 782.56689 (for MALDI) and PC at m/z 104.1 (for SIMS) are distinguished within outer surface of the cells, suggesting their high abundance in the plasma membrane. Another intranuclear lipid fragment *i.e.* MAG at m/z 313.3 was also detected that has been previously detected in human breast cancer cell lines analysed with ToF-SIMS, along with m/z 576 ion (Lanekoff *et al.* 2011, Robinson *et al.* 2016).

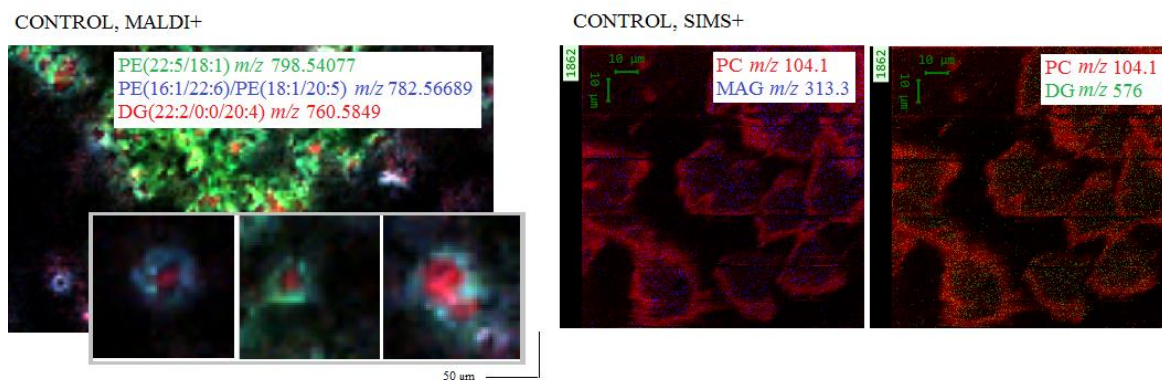


Figure 18: A comparison of MALDI and SIMS images of sub-cellular compartments of CHO cells in positive ion mode. MALDI species of PE (m/z 798.54077, m/z 782.56689) and DG (m/z 760.5849) and SIMS fragments of PC (m/z 104.1), MAG (m/z 313.3) and DG (m/z 576) are shown. Max. count per pixel, MALDI+: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3); SIMS+: m/z 104.1 (9701), m/z 313.3 (256) and m/z 576 (247). Scale bar 50 μm (zoom out image MALDI) and 10 μm (SIMS), pixel size 5 μm (MALDI) and 0.5 μm (SIMS).

Nonetheless, these data also illustrate that MALDI and SIMS imaging techniques are complementary by the high spatial resolution of SIMS and biomolecular specificity and wider mass range of the MALDI data. The complementary data acquired from different MS imaging techniques can be employed in multi-approach studies so that the rich spatial information correlated to the specific biochemical information could be obtained. Additionally, with a recent development of the SIMS instrument (*3D OrbiSIMS*, 2017) future investigations could be performed with not only high spatial resolution but also with high mass accuracy.

7.5.3. Heterogeneity of CHO cell population

Previous observations and data analysis of cell populations and single cells reported above, derived with information of cell diversity occurring within control CHO population. This section describes cell heterogeneity investigations within non-induced (control) CHO cells studied using MALDI and SIMS imaging techniques.

Figure 19 presents a MALDI overlay image of three lipids: PE(22:5/18:1) at m/z 798.54077, PE(16:1/22:6)/PE(18:1/20:5) at m/z 782.56689 and DG(22:2/0:0/20:4) at m/z 760.5849, previously identified. Without *a priori* knowledge of the cellular heterogeneity in the sampled cell populations we have used the abundance of these three lipid signals as an indication of three possible cell phenotypes (which we define as Type A, B and C). Within the sampled cell population, different types of cells based on the diverse abundance and distribution of these lipids can be observed. In order to explore this diversity, metabolite profiles from type A (red), B (green) and type C (blue) cells were selected for multivariate analysis: **Figure 20 A** shows averaged MALDI+ spectra collected from each of the cell types: A, B and C.

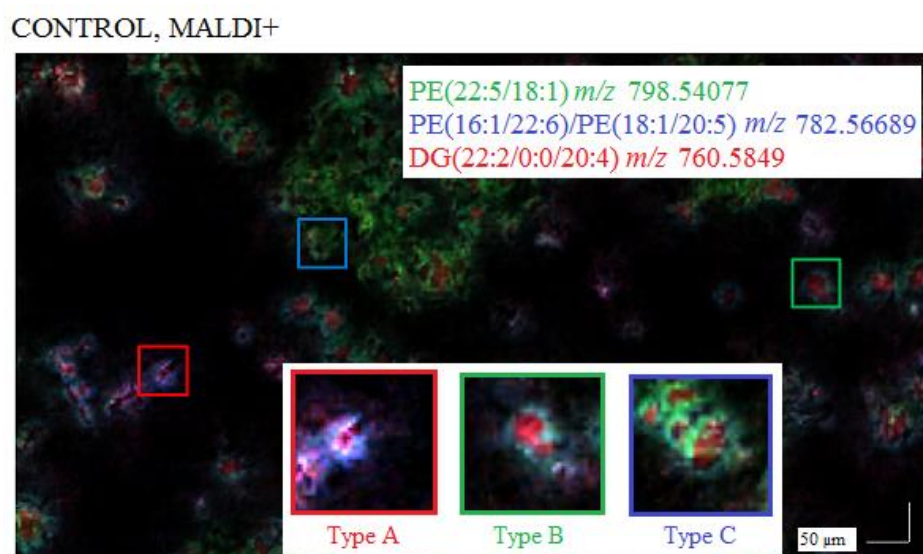


Figure 19: Analysis of heterogeneity of CHO cell population for MALDI+. Image shows overlay of three lipid species: PE(22:5/18:1) at m/z 798.54077, PE(16:1/22:6)/PE(18:1/20:5) at m/z 782.56689, DG(22:2/0:0/20:4) at m/z 760.5849 within control cell population. Enlarged single cells with different metabolite profiles are shown (colour-coded): type A (red), type B (green) and type C (blue) and were selected for multivariate analysis ($n=5$ for each type). Max. count per pixel: m/z 798.54077 (1.54×10^4), m/z 782.56689 (1.01×10^4) and m/z 760.5849 (8.27×10^3). Scale bar 50 μm (for zoom out image), pixel size 5 μm .

MALDI+

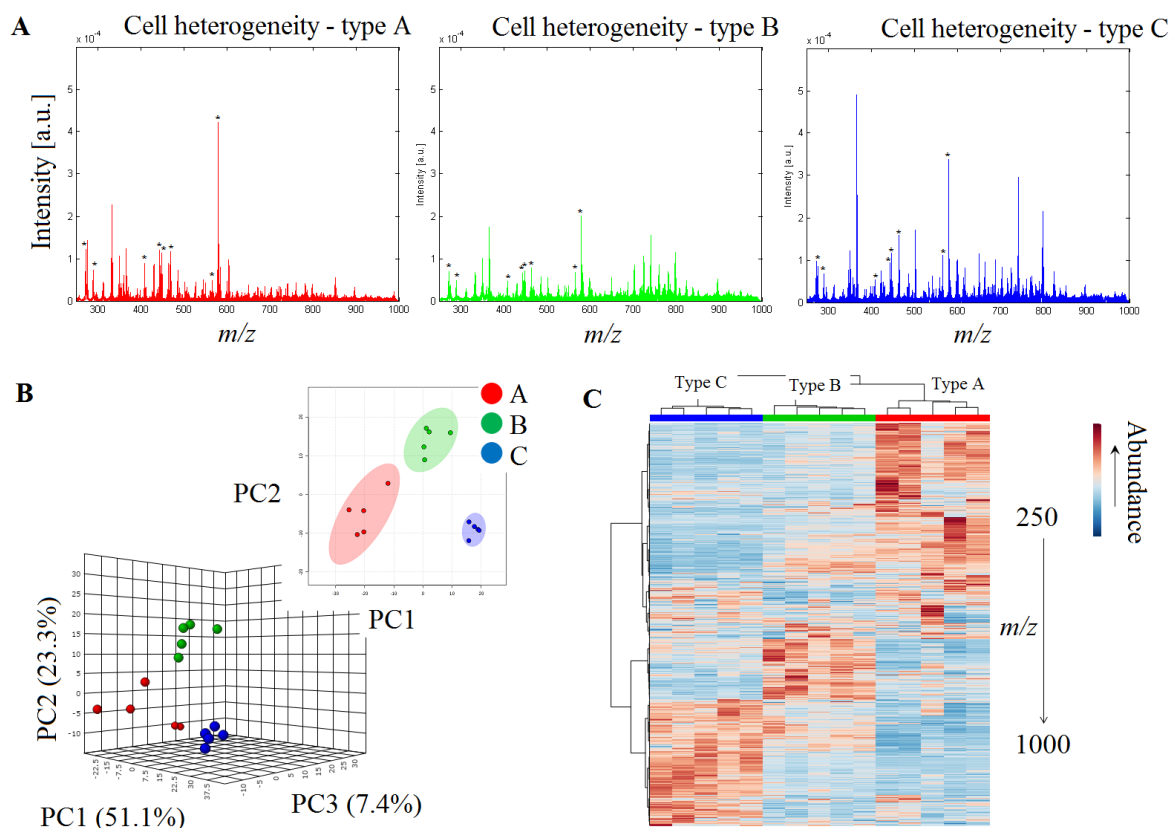


Figure 20: A comparison of the averaged MALDI-MSI spectra from three different metabolite profiles: type A (red), type B (green) and type C (blue) (A); m/z 250-1000, *-matrix peaks. (Inset) 2D and 3D PCA scores accounting for 81.8% of total variance (first three components) (B). PCA plots illustrate the variance between analysed types of metabolic profiles with highlighted 95% regions of confidence. Heatmap shows heterogeneity and diversity of compounds within each type of metabolite profiles across whole analysed mass range (m/z 250-1000) (C).

The spectra are rich in information and metabolites are detected across whole mass range of m/z 250-1000. The overall view of the spectra suggests that the majority of peak areas are very similar between all three groups of cells, as would be expected given that these three cell 'types' were chosen arbitrarily based only on the intensity of three lipid species. However, different intensities of ions for each cell group can also be seen. Additionally, matrix peaks were detected and marked on the spectra: m/z 273 [2DHB+H-2H₂O]⁺, m/z 290 [2DHB+NH₄-2H₂O]⁺, m/z 409 [3DHB+H-3H₂O]⁺, m/z 444 [3DHB*-H₂O]⁺, m/z 447 [3DHB+K-3H₂O]⁺, m/z 465 [3DHB+K-2H₂O]⁺, m/z 567 [4DHB+Na-4H₂O]⁺, m/z 580 [4DHB*-2H₂O]⁺. In order to investigate differences between selected cell types, multivariate analysis with the application of ANOVA test was performed. The analysis derived with 2D and 3D PCA scores plots as well as heatmap illustrating clustering and

describing metabolite diversity between cell profiles (**Figure 20 B and C**). 2D and 3D PCA scores plots show good clustering and high level of diversity between three metabolite profiles within control cells (**Figure 20 B**). As seen in the figure, some level of heterogeneity within sub-groups is also detected. 3D PCA scores plot generated using first three components account for 81.8% of total variance between different types of cells. Additionally, 2D PCA scores plot illustrates high level of separation, accounting for 51.1% of variance on PC1. The lowest variance is present for type C cells, whereas type A and B cells show higher variance, suggesting that further sub-classification into more types/metabolite profiles of CHO cells can be made. This result can be also confirmed on the heatmap (**Figure 20 C**), presenting each cell type (colour-coded in red – type A, green – type B, blue – type C) and its heterogeneity within sub-type across whole scanned mass range (m/z 250-1000). Heatmap also illustrates the abundance of each ion detected: from high (red) to low (blue). The majority of ions with low masses seem to occur in type A cells, whereas medium and high mass ions are observed in type B and C cells.

Heterogeneity within CHO cell populations regarding diversity in expression and stability of a protein (Pilbrough *et al.* 2009, Du *et al.* 2013), proliferation rate and production of cell biomass have been previously described (Davies *et al.* 2013). A number of key factors such as cell microenvironment, genetic modifications originating from mutations or transfected gen into a cell line can influence above processes (Avery 2006, Davidson & Surette 2008, Dorai *et al.* 2007). Moreover, the genetic drift can also generate an unwanted behaviour of *in vitro* cell line for example decrease in protein productivity during sub-culture (Kim *et al.*, 2011), decrease in proliferation rate (Barnes, Moy and Dickson, 2006) or cause variations in glycosylation profile of secreted protein (Van Berkel *et al.*, 2009). Therefore, the adaptation of cell lines to specific media and a separation of clones with the highest ability to survive in *in vitro* environment and with high rates of cell proliferation and protein expression are routinely performed in the bio-industry (O’Callaghan *et al.* 2010, Porter *et al.* 2010). In order to decrease cell-to-cell variability, a chemically defined cell medium was used as well as isolation of a parental clone with the best *in vitro* proliferation rate and the highest protein expression level was selected for this study. However, a great level of diversity within control CHO cells can still be observed.

CHO cell populations often show diversity in spreading and adhesion during the cell growth, resulting in variations in cell morphology. It is believed that adhesion and hence morphological changes, link to phenotypic variation that can be of potential biological significance. The importance of adhesion variability has been investigated by Franz *et al.*, revealing the adhesion is non-genetic and cell cycle-independent but is scaled with the

membrane receptors responsible for adhesion (Dao *et al.*, 2012). With regard to the above, the cell culture for this study was grown for 3 days prior the MS experiments as it is described in **Chapter 4**, when the stability of cell growth and the protein production were achieved.

In order to define metabolite changes observed in selected metabolite profiles, compound identification of statistically significant metabolites based on the application of ANOVA test ($p < 0.05$) was performed. Out of top 100 the most significant compounds detected, 5 of them were selected, assigned and correlated to cell locations in order to illustrate dependencies between three types of cells (**Table 6**). **Table 6** summarises measured m/z , p value, m/z error (ppm), HMDB ID, metabolite hit and Box and Whisker plots presenting the abundance of each metabolite in selected type of cells. Two metabolites were successfully assigned: m/z 465.9930 and m/z 719.5416 with the higher intensities in type C and type B cell, respectively. Ions at m/z 793.9878, m/z 950.6481 and m/z 605.5504 with the higher abundance for type C, B and A cells respectively, remained unknown. Some consistency within identification of adenosine 3',5'-diphosphate/dGDP/ADP can be observed. These compounds were assigned not only to ion at m/z 465.9930, but also to m/z 428.0368 ion discussed in previous sections, which suggests these are important factors in a cell response to changing microenvironment and cell-to-cell diversity.

Table 6 summarises variations of abundance of ions within different metabolite profiles of cells. Remaining ions belong to a mass range of >500 Da that suggest they belong to the class of lipids (Sellick *et al.*, 2011, Hollywood *et al.*, 2015, Dorries & Lalk, 2013). Some of them appear to be more intense in type A cells *i.e.* m/z 605.5504, some in type B cells *i.e.* m/z 950.6481 and m/z 719.5416 and some in type C cells *i.e.* m/z 465.9930 and m/z 793.9878.

Apart from previously described factors contributing to cell-to-cell variability (see **Chapter 1**) such as genetic modifications, cell environment, cell cycle or proliferation rate (Avery 2006, Davidson & Surette 2008, Dorai *et al.* 2007, Barnes *et al.* 2006), the diversity within cell population can also be caused by a cell surface changes and its lipid content (Du *et al.*, 2013). As discussed earlier, the protein production in cells can be related to exocytosis process that causes an increase of a cell membrane surface (Roman-Fernandez & Bryant 2016, Bruns & Jahn 2002). Consequently, it suggests that this interaction could affect the distribution of lipids present in a membrane of CHO cells.

Table 6: A representation of 5 significant metabolites for different types of metabolite profiles identified within control CHO cell population in MALDI+. The table summarises measured mass, *p* value, mass error (ppm), HMDB ID, metabolite hit and abundance in each type of metabolite profiles shown as Box and Whisker plots; type A-red, type B-green, type C-blue. Metabolites sorted by *p* value.

Measured mass	P value	Δ ppm	HMDB ID	Metabolite	Box and Whisker plot
465.9930	4.64E-09	0.98	HMDB00061 HMDB00960 HMDB01341	Adenosine 3',5'-diphosphate dGDP ADP	
793.9878	1.54E-08			Unknown	
950.6481	1.15E-07			Unknown	
719.5416	4.11E-07	0.09	HMDB13164 HMDB61638	2-Hydroxy-lauroylcarnitine 3-hydroxydodecanoyl carnitine	
605.5504	1.78E-05			Unknown	

As can be seen from the analysis of cell populations, both control and tetracycline-induced cell populations show some level of heterogeneity. It can also be observed that this heterogeneity is different between these two cell populations, where the abundance of some lipids seem to vary depending on EPO-Fc production process. Therefore, it would be

beneficial to perform more single-cell based MALDI-MSI experiments in order to investigate this diversity further.

The additional advantage of MS imaging applied to cell-to-cell heterogeneity research is the relation of spectral information to a specific mass that has been applied in this study. Consequently, the cell type-specific ions shown in **Table 6** can be used to define cells with individual metabolite characteristics. **Figure 21** presents single ions images for three chosen ions correlating to three metabolic cell types: type A at m/z 605.5504, type B at m/z 719.5416 and type C at m/z 465.9930. Additionally, the overlay image shows the application of the ion information for the localisation of cells with specific metabolic type, resulting in the ability to distinguish different metabolic profiles within CHO cell population using MALDI-MSI technique.

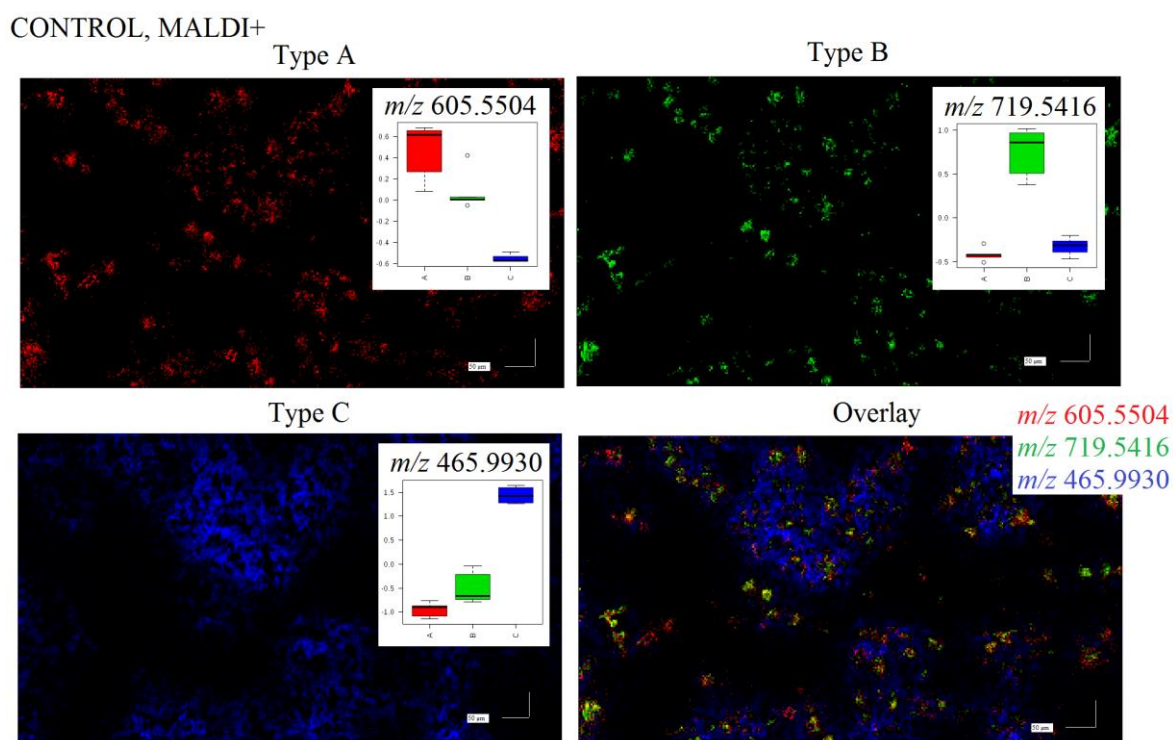


Figure 21: A location of cells with different type of metabolite profile within cell population in MALDI+. Single ion images correspond to metabolite profiles with the highest abundance of each ion (colour-coded): type A (red) m/z 605.5504, type B (green) m/z 719.5416, type C (blue) m/z 465.9930. Box and Whisker plots (insets) show the abundance of ions within each type of metabolite profile. Metabolite profiles type A, B and C are illustrated in overlay image (colour-coded). Max. count per pixel: m/z 605.5504 (5.66×10^2), m/z 719.5416 (6.60×10^2) and m/z 465.9930 (2.44×10^3). Scale bar 50 μm , pixel size 5 μm .

Comparative analysis of cell-to-cell diversity within one cell population using SIMS+ data was performed. Multivariate analysis resulted in PC2 and PC3 scores images and correlating to these images loadings plots PC2 and PC3 (**Figure 22**). PC2 and PC3 scores images show the abundance of ions displayed in PC2 and PC3 loadings plots within imaged areas. Moreover, the results highlight cell-to-cell heterogeneity within CHO cell population, where one of the cells appears to have the different biochemistry than the rest of cells in the culture. The most significant ions within PC2 and PC3 loadings plots are highlighted and representative single ion images are shown.

Multivariate analysis captured some variations in signal between areas within single cells (green and red regions). There were multiple fragments chosen based on PC2 and PC3 variance for green regions (positive going): m/z 86.4, m/z 104.0, m/z 124.8, m/z 184.0 for PC2 and 86.4, m/z 115.2, m/z 124.8, m/z 184.0 for PC3. Similarly, some of the fragments for red regions (negative going) were selected: m/z 112.9, m/z 120.0, m/z 156.9, m/z 196.9, m/z 212.9, m/z 235.2, m/z 251.2, m/z 352.8, m/z 368.8 for PC2 and additional fragment at m/z 104.0 for PC3. PC2 and PC3 scores images show good correlation of the high abundant species at m/z 104 and m/z 115 with spectral information seen in images generated for these ions. Both species *i.e.* m/z 104 and m/z 115 are present within all analysed cells however, for one cell the higher concentration within its edges is observed. From the analysis of PC2 and PC3 scores images and ion images for negative going, it is possible to distinguish more specific cell-to-cell variations. According to the PC2 loadings plot, m/z 112, m/z 196 and m/z 212 are the most abundant fragments suggesting to be the significant in the analysis of cell diversity. This is in line with images generated for these ions, where their distribution is mainly seen in one of the cells, implying its specific biochemistry when compared to the rest of the cells from this cell population. On the other hand, fragments at m/z 235, m/z 251 and m/z 369 show lower abundance (and the size of score associated with loadings plot) that suggests their smaller input to the cell-to-cell heterogeneity. This assumption is confirmed in ion images of these ions, where their localisation within the whole cell population can be seen. An analogous situation is observed for the analysis of principal component 3 for most of the ions of negative going score.

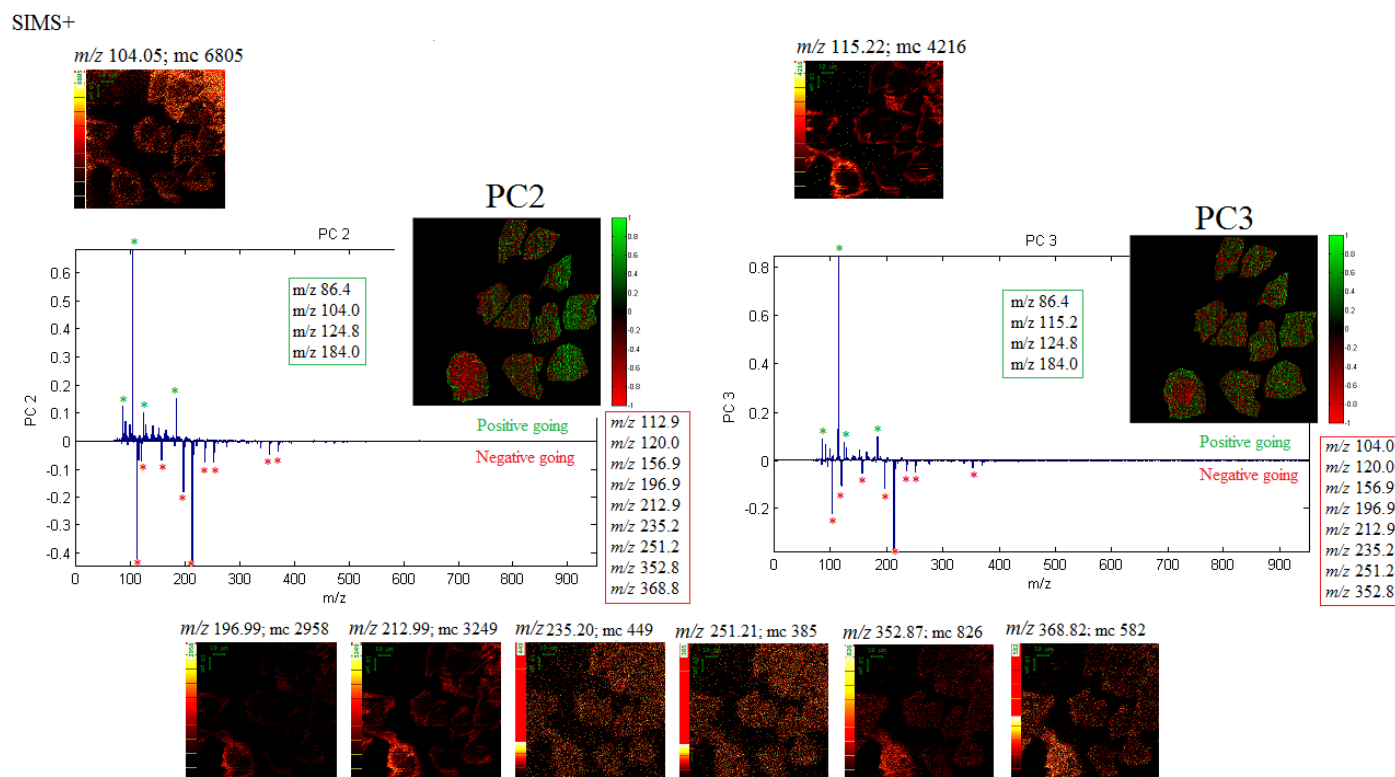


Figure 22: Principal Component Analysis of cell-to-cell heterogeneity within CHO cell population for SIMS+. Multivariate analysis resulted with PC2 and PC3 scores images and correlating PC2 and PC3 loadings plots. PC2 and PC3 scores images show the abundance of ions displayed in PC2 and PC3 loadings plots within cell population. The most abundant ions within PC2 and PC3 loadings plots are highlighted and representative images of these ions are shown. Maximum count per pixel (mc) is displayed. Scale bar 10 μm , pixel size 0.5 μm , m/z 70-1000.

7.6. Conclusions and future work

The employment of mass spectrometry imaging approaches in metabolomics of tetracycline-induced CHO cells has been presented and can be considered successful. The application of MALDI-MSI and SIMS imaging techniques has provided detailed information about the CHO cell response to the EPO-Fc fusion protein production. The great potential of these two platforms has been shown by the metabolic analysis based on cell populations and individual cells, with additional possibilities to investigate sub-cellular levels.

Multivariate analysis highlighted a number of metabolite that were significantly influenced by the protein expression when compared to control cells. According to the analysis, the major metabolic changes caused by EPO-Fc production in CHO cells were those related to lipid metabolism. In addition, amino acid metabolism as well as nucleotide metabolism were also influenced by the protein expression. These findings are consistent with the analysis of external and internal metabolic profiles of CHO cells (reported elsewhere in this thesis), showing the potential of the applications of complementary techniques to deepen insights into cellular activity. Additionally, correlation of these significant metabolites to their locations within cell populations provided with detailed biochemical information of CHO cells producing EPO-Fc fusion protein.

The information about metabolic changes in tetracycline-induced cells obtained from the analysis of cell populations was further confirmed with the analysis based on single-cell study. The analysis of individual cells also revealed other changes in metabolic content of induced cells to those seen in the cell-population based study. This therefore proves that single cell-based investigations provide additional insights about changes in metabolism of induced cells that can be referred to a unique, single cell. Furthermore, single cell analysis offers the potential to correlate metabolic profiles of cells producing EPO-Fc to their production efficiency that could bring new goals for biomanufacturing companies.

Another aim of MALDI and SIMS imaging of CHO cells was to investigate heterogeneity within cell population. The analysis showed a high level of diversity of CHO cells within one population, where a number of metabolites could have been correlated to a specific type of cells. Linking of these significant metabolites of each cell type to their location within cell population allowed for defining metabolite characteristics, specific for each type of cells.

In addition to cell population and single-cell based study, the investigation of sub-cellular content within individual cells was also performed. The results of spectral maps allowed

for a comparison of metabolite content of different cellular organelles within individual CHO cells.

The application of MALDI-MS and SIMS imaging in the study of tetracycline-induced protein production in CHO cells has been shown. These were very challenging experiments, at the forefront of current technical developments. MALDI-MS imaging was successfully applied in mapping of CHO cells on cell population and single-cell levels. In addition, this method also presented a great potential in the study of sub-cellular content of mammalian cells, where high spatial resolution and high mass accuracy allowed for detailed biochemical analysis of cells producing EPO-Fc.

Using the MALDI-MSI technique, further experiments on these cells could also be performed to bring more information about EPO-Fc fusion protein production. The application of PNA matrix for negative ionisation mode was not successful in this study. Other matrices however are available for MALDI-MSI (*e.g.* 9AA) that can potentially be used in investigation of CHO cells producing the fusion protein.

Complementary analysis with the application of SIMS imaging with C_{60}^+ and Au_3^+ primary ion guns was also conducted. The incorporation of two different primary ion sources allows for more broad analysis of chemical content of CHO cells. The use of polyatomic C_{60}^+ ion beam allowed for metabolic analysis of depth profiles of these cells. However, the spatial resolution was compromised by the current instrument configuration and so, further experiments could be performed with the use of a more focused C_{60}^+ ion beam. This would bring possibilities of creating high-resolution spectral maps with specific chemical content of CHO cells in 3D visualisation.

By contrast, the employment of the Au_3^+ primary ion source allowed for high resolved images of the top layer of CHO single cells. With the application of this ion beam, it was possible to present detailed biochemical information along with metabolite fragments' localisation within individual CHO cells. Nevertheless, more experiments for SIMS using both C_{60}^+ and Au_3^+ primary sources and MALDI-MS could also be executed in order to increase number of repeats to enhance confidentiality of statistical analysis. Additionally, with a recent development of the SIMS instrument (*3D OrbiSIMS*, 2017) future investigations could be performed with not only high spatial resolution but also with high mass accuracy.

One more possibility would be to perform imaging studies incorporating another approach – ion mobility. Ion mobility mass spectrometry can be applied to provide separation of ions according to their 3D molecular structure (*e.g.* shape) (Blatherwick *et al.*, 2013). This would enable different compound classes to be separated, giving a possibility to

distinguish matrix molecules as well as would allow for more accurate metabolite classification.

References:

- 3D OrbiSIMS (2017). Available at: <http://www.npl.co.uk/news/3d-nanosims-label-free-molecular-imaging> (Accessed: 20 April 2017).
- Albi, E. (2011) 'Role of intranuclear lipids in health and disease', *Clini Lipid*, 6, pp. 59–69.
- AlMasoud, N., Correa, E., Trivedi, D. K. and Goodacre, R. (2016) 'Fractional Factorial Design of MALDI-TOF-MS Sample Preparations for the Optimized Detection of Phospholipids and Acylglycerols', *Anal Chem*, 88, pp. 6301–6308.
- Amantonico, A., Urban, P. L., Fagerer, S. R., Balabin, R. M. and Zenobi, R. (2010) 'Single-cell MALDI-MS as an analytical tool for studying intrapopulation metabolic heterogeneity of unicellular organisms', *Anal Chem*, 82, pp. 7394–7400.
- Anderson, D. M. G., Mills, D., Spraggins, J., Lambert, W. S., Calkins, D. J. and Schey, K. L. (2013) 'High-resolution matrix-assisted laser desorption ionization-imaging mass spectrometry of lipids in rodent optic nerve tissue', *Mol Vision*, 19, pp. 581–592.
- Angel, P. M. and Caprioli, R. M. (2014) 'Matrix-Assisted Laser Desorption Ionization Imaging Mass Spectrometry: In Situ Molecular Mapping', *Biochemistry*, 52, pp. 1–21.
- Avery, S. V (2006) 'Microbial cell individuality and the underlying sources of heterogeneity', *Nat Rev Microbiol*, 4, pp. 577–87.
- Barnes, L. M., Moy, N. and Dickson, A. J. (2006) 'Phenotypic Variation During Cloning Procedures: Analysis of the Growth Behavior of Clonal Cell Lines', *Biotechnol Bioeng*, 94, pp. 530–537.
- Van Berkel, P. H. C., Gerritsen, J., Perdok, G., Valbjørn, J., Vink, T., Van De Winkel, J. G. J. and Parren, P. W. H. I. (2009) 'N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG', *Biotechnol Prog*, 25, pp. 244–251.
- Bhandari, D. R. am, Schott, M., Römpf, A., Vilcinskas, A. and Spengler, B. (2015) 'Metabolite localization by atmospheric pressure high-resolution scanning microprobe matrix-assisted laser desorption/ionization mass spectrometry imaging in whole-body sections and individual organs of the rove beetle *Paederus riparius*', *Anal Bioanal Chem*, 407, pp. 2189–2201.
- Blatherwick, E. Q., Svensson, C. I., Frenguelli, B. G. and Scrivens, J. H. (2013) 'Localisation of adenine nucleotides in heat-stabilised mouse brains using ion mobility enabled MALDI imaging', *Intern J Mass Spectrom*, 345–347, pp. 19–27.
- Boughton, B. A., Thinagaran, D., Sarabia, D., Bacic, A. and Roessner, U. (2016) 'Mass spectrometry imaging for plant biology: a review', *Phytochem Rev*, 15, pp. 445–488.
- Brown, M., Dunn, W. B., Dobson, P., Patel, Y., Winder, C. L., Francis-McIntyre, S., Begley, P., Carroll, K., Broadhurst, D. I., Tseng, A., Swainston, N., Spasic, I., Goodacre, R. and Kell, D. B. (2009) 'Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics', *The Analyst*, 134, pp. 1322–32.
- Bruns, D. and Jahn, R. (2002) 'Molecular determinants of exocytosis', *Pflugers Arch EJP*, 443, pp. 333–338.
- Burrell, M. M., Earnshaw, C. J. and Clench, M. R. (2007) 'Imaging Matrix Assisted Laser Desorption Ionization Mass Spectrometry: A technique to map plant metabolites within tissues at high spatial resolution', *J Exp Bot*, 58, pp. 757–763.
- Calvano, C. D., Monopoli, A., Ditaranto, N. and Palmisano, F. (2013) '1,8-Bis(dimethylamino)naphthalene/9-aminoacridine: A new binary matrix for lipid

fingerprinting of intact bacteria by matrix assisted laser desorption ionization mass spectrometry', *Anal Chim Acta*, 798, pp. 56–63.

Canavan, H. E., Graham, D. J., Cheng, X., Ratner, B. D. and Castner, D. G. (2007) 'Comparison of native extracellular matrix with adsorbed protein films using secondary ion mass spectrometry.', *Langmuir*, 23, pp. 50–56.

Chen, Y., Gao, D., Bai, H., Liu, H., Lin, S. and Jiang, Y. (2016) 'Carbon Dots and 9AA as a Binary Matrix for the Detection of Small Molecules by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry', *J Am Soc Mass Spectrom*, 27, pp. 1227–1235.

Cornett, D., Frappier, S. and Caprioli, R. (2008) 'MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue.', *Anal. Chem.*, 80, pp. 5648–5653.

Dao, L., Weiland, U., Hauser, M., Nazarenko, I., Kalt, H., Bastmeyer, M. and Franz, C. M. (2012) 'Revealing non-genetic adhesive variations in clonal populations by comparative single-cell force spectroscopy', *Exp Cell Res*, 318, pp. 2155–2167.

Davidson, C. J. and Surette, M. G. (2008) 'Individuality in bacteria', *Annu Rev Genet*, 42, pp. 253–268.

Davies, S. L., Lovelady, C. S., Grainger, R. K., Racher, A. J., Young, R. J. and James, D. C. (2013) 'Functional heterogeneity and heritability in CHO cell populations', *Biotechnol Bioeng*, 110, pp. 260–274.

Denbigh, J. L. and Lockyer, N. P. (2015) 'ToF-SIMS as a tool for profiling lipids in cancer and other diseases', *Mat Sci Technol*, 31, pp. 137–147.

Didonato, D. and Brasaemle, D. L. (2003) 'Fixation methods for the study of lipid droplets by immunofluorescence microscopy', *J Histochem Cytochem*, 51, pp. 773–780.

Dorai, H., Sauerwald, T., Campbell, A., Kyung, Y., Goldstein, J. and Magill, A. (2007) 'Investigation of product microheterogeneity: A case study in rapid detection of mutation in mammalian production cell lines', *Bioprocess Int*, 5, pp. 66–75.

Dorries, K. and Lalk, M. (2013) 'Metabolic footprint analysis uncovers strain specific overflow metabolism and D-isoleucine production of *Staphylococcus aureus* COL and HG001', *PLoS One*, 8, pp. 81500–81510.

Du, Z., Mujacic, M., Le, K., Caspary, G., Nunn, H., Heath, C. and Reddy, P. (2013) 'Analysis of heterogeneity and instability of stable mAb-expressing CHO cells', *Biotechnol Bioeng*, 18, pp. 419–429.

Emara, S., Amer, S., Ali, A., Abouleila, Y., Oga, A. and Masujima, T. (2017) 'Single-Cell Metabolomics', *Adv Exp Med Biol*, 965, pp. 323–343.

Fagone, P. and Jackowski, S. (2009) 'Membrane phospholipid synthesis and endoplasmic reticulum function', *J Lipid Res*, 50, pp. S311–6.

Fletcher, J. S., Kotze, H. L., Armitage, E. G., Lockyer, N. P. and Vickerman, J. C. (2013) 'Evaluating the challenges associated with time-of-flight secondary ion mass spectrometry for metabolomics using pure and mixed metabolites', *Metabolomics*, 9, pp. 535–544.

Fletcher, J. S., Rabbani, S., Barber, A. M., Lockyer, N. P. and Vickerman, J. C. (2013) 'Comparison of C60 and GCIB primary ion beams for the analysis of cancer cells and tumour sections', *Surf Interface Anal*, 45, pp. 273–276.

Fletcher, J. S., Rabbani, S., Henderson, A., Lockyer, N. P. and Vickerman, J. C. (2011) 'Three-dimensional mass spectral imaging of HeLa-M cells - Sample preparation, data interpretation and visualisation', *Rapid Comm Mass Spectrom*, 25, pp. 925–932.

Fukuyama, Y. (2011) 'Applications of Ionic Liquids in Science and Technology', in

- Handy, S. (ed.) *Applications of Ionic Liquids in Science and Technology*. InTech, pp. 361–374. doi: 10.5772/1769.
- Fülöp, A., Sammour, D. A., Erich, K., von Gerichten, J., van Hoogevest, P., Sandhoff, R. and Hopf, C. (2016) ‘Molecular imaging of brain localization of liposomes in mice using MALDI mass spectrometry’, *Sci Rep*, 6, p. 33791.
- Gill, E. L., Yost, R. A., Vedam-Mai, V. and Garrett, T. J. (2017) ‘Precast Gelatin-Based Molds For Tissue Embedding Compatible with Mass Spectrometry Imaging’, *Anal Chem*, 89, pp. 576–580.
- Goodwin, R. J. A., Pitt, A. R., Harrison, D., Weidt, S. K., Langridge-Smith, P. R. R., Barrett, M. P. and Logan MacKay, C. (2011) ‘Matrix-free mass spectrometric imaging using laser desorption ionisation Fourier transform ion cyclotron resonance mass spectrometry’, *Rapid Comm Mass Spectrom*, 25, pp. 969–972.
- Griffiths, R. L. and Bunch, J. (2012) ‘A survey of useful salt additives in matrix-assisted laser desorption/ionization mass spectrometry and tandem mass spectrometry of lipids: Introducing nitrates for improved analysis’, *Rapid Comm Mass Spectrom*, 26, pp. 1557–1566.
- Hanrieder, J., Wicher, G., Bergquist, J., Andersson, M. and Fex-Svenningsen, A. (2011) ‘MALDI mass spectrometry based molecular phenotyping of CNS glial cells for prediction in mammalian brain tissue’, *Anal Bioanal Chem*, 401, pp. 135–147.
- Heinemann, M. and Zenobi, R. (2011) ‘Single cell metabolomics’, *Cur Opin Biotechnol*, 22, pp. 26–31.
- HMDB (2017). Available at: <http://www.hmdb.ca/> (Accessed: 5 January 2017).
- Hollywood, K. A., Winder, C. L., Dunn, W. B., Xu, Y., Broadhurst, D. I., Griffiths, C. E. and Goodacre, R. (2015) ‘Exploring the mode of action of dithranol therapy for psoriasis: a metabolomic analysis using HaCaT cells’, *Mol Biosyst*, 11, pp. 2198–2209.
- Jerigova, M., Biro, C., Kirchnerova, J., Chorvatova, A., Chorvat, D., Lorenc, D. and Velic, D. (2011) ‘Chemical Imaging of Cardiac Cell and Tissue by Using Secondary Ion Mass Spectrometry’, *Mol Imag Biol*, 13, pp. 1067–1076.
- Jirasko, R., Holcapek, M., Kunes, M. and Svatos, A. (2014) ‘Distribution study of atorvastatin and its metabolites in rat tissues using combined information from UHPLC/MS and MALDI-Orbitrap-MS imaging’, *Anal Bioanal Chem*, 406, pp. 4601–4610.
- Kempson, I. M., Martin, A. L., Denman, J. A., French, P. W., Prestidge, C. A. and Barnes, T. J. (2010) ‘Detecting the presence of denatured human serum albumin in an adsorbed protein monolayer using TOF-SIMS’, *Langmuir*, 26, pp. 12075–12080.
- Khalil, S. M., Römpf, A., Pretzel, J., Becker, K. and Spengler, B. (2015) ‘Phospholipid Topography of Whole-Body Sections of the Anopheles stephensi Mosquito, Characterized by High-Resolution Atmospheric-Pressure Scanning Microprobe Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging’, *Anal Chem*, 87, pp. 11309–11316.
- Kim, M., O’Callaghan, P. M., Droms, K. A. and James, D. C. (2011) ‘A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies’, *Biotechnol Bioeng*, 108, pp. 2434–2446.
- Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S. and Fiehn, O. (2009) ‘FiehnLib – Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry’, *Anal Chem*, 81, pp. 10038–10048.

- Kiss, A. and Hopfgartner, G. (2016) 'Laser-based methods for the analysis of low molecular weight compounds in biological matrices', *Methods*, 104, pp. 142–53.
- Klyszejko-Stefanowicz, L. (2002) *Cytobiochemia. Biochemia niektórych struktur komórkowych*. Warszawa: PWN.
- Kompauer, M., Heiles, S. and Spengler, B. (2017) 'Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- μ m lateral resolution', *Nat Met*, 14, pp. 1–10.
- Korte, A. R. and Lee, Y. J. (2014) 'MALDI-MS analysis and imaging of small molecule metabolites with 1,5-diaminonaphthalene (DAN)', *J Mass Spectrom*, 49, pp. 737–741.
- Korte, A. R., Yandea-Nelson, M. D., Nikolau, B. J. and Lee, Y. J. in (2015) 'Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer', *Anal Bioanal Chem*, 407, pp. 2301–2309.
- Kotze, H. L., Armitage, E. G., Fletcher, J. S., Henderson, A., Williams, K. J., Lockyer, N. P. and Vickerman, J. C. (2013) 'ToF-SIMS as a tool for metabolic profiling small biomolecules in cancer systems', *Surf Interf Anal*, 45, pp. 277–281.
- Lanekoff, I., Kurczy, M. E., Adams, K. L., Malm, J., Karlsson, R., Sjövall, P. and Ewing, A. G. (2011) 'An in situ fracture device to image lipids in single cells using ToF-SIMS', *Surf Interface Anal*, 43, pp. 257–260.
- Lanni, E. J., Rubakhin, S. S. and Sweedler, J. V. (2012) 'Mass spectrometry imaging and profiling of single cells', *J Proteomics*, 75, pp. 5036–5051.
- Lee, T. G., Shon, H. K., Lee, K. B., Kim, J., Choi, I. S. and Moon, D. W. (2006) 'Time-of-flight secondary ion mass spectrometry chemical imaging analysis of micropatterns of streptavidin and cells without labeling', *J Vac Sci Technol A*, 24, pp. 1203–1207.
- Lev, S. (2006) 'Lipid homeostasis and Golgi secretory function', *Biochem Soc Trans*, 34, pp. 363–6.
- Liu, Q., Guo, Z. and He, L. (2007) 'Mass spectrometry imaging of small molecules using desorption/ionization on silicon', *Anal Chem*, 79, pp. 3535–3541.
- Llombart, V., Trejo, S. A., Bronsoms, S., Morancho, A., Feifei, M., Faura, J., García-Berrocso, T., Simats, A., Rosell, A., Canals, F., Hernández-Guillamón, M. and Montaner, J. (2016) 'Profiling and identification of new proteins involved in brain ischemia using MALDI-imaging-mass-spectrometry', *J Proteomics*. Elsevier B.V., 152, pp. 243–253.
- Lou, S., Balluff, B., Cleven, A. H. G., Bovée, J. V. M. G. and McDonnell, L. A. (2016) 'Prognostic Metabolite Biomarkers for Soft Tissue Sarcomas Discovered by Mass Spectrometry Imaging', *J Am Soc Mass Spectrom*, 52, pp. 376–383.
- Martens, S. and McMahon, H. T. (2008) 'Mechanisms of membrane fusion: disparate players and common principles.', *Nat Rev Mol Cell Biol*, 9, pp. 543–556.
- McDonnell, L. A. and Heeren, R. M. A. (2007) 'Imaging Mass Spectrometry', *Mass Spectrometry reviews*, 26, pp. 606–643. doi: 10.1002/mas.
- Miura, D., Fujimura, Y., Tachibana, H. and Wariishi, H. (2010) 'Highly sensitive matrix-assisted laser desorption ionization-mass spectrometry for high-throughput metabolic profiling', *Anal Chem*, 82, pp. 498–504.
- Mouhib, T., Poleunis, C., Wehbe, N., Michels, J. J., Galagan, Y., Houssiau, L., Bertrand, P. and Delcorte, A. (2013) 'Molecular depth profiling of organic photovoltaic heterojunction layers by ToF-SIMS: comparative evaluation of three sputtering beams', *The Analyst*, 138, pp. 6801–10.

- Munro, S. (2005) 'The Golgi apparatus: Defining the identity of Golgi membranes', *Curr Opin Cell Biol*, 17, pp. 395–401.
- Nagata, Y., Ishizaki, I., Waki, M., Ide, Y., Hossen, M. A., Ohnishi, K., Sanada, N. and Setou, M. (2014) 'Glutaraldehyde fixation method for single-cell Lipid analysis by time-of-flight secondary ion-mass spectrometry', *Surf Interf Anal*, 46(SI), pp. 185–188.
- Nielsen, M. M. B., Lambertsen, K. L., Clausen, B. H., Meyer, M., Bhandari, D. R., Larsen, S. T., Poulsen, S. S., Spengler, B., Janfelt, C. and Hansen, H. S. (2016) 'Mass spectrometry imaging of biomarker lipids for phagocytosis and signalling during focal cerebral ischaemia', *Sci Rep*, 6, pp. 39571–79.
- O'Callaghan, P. M., McLeod, J., Pybus, L. P., Lovelady, C. S., Wilkinson, S. J., Racher, A. J., Porter, A. and James, D. C. (2010) 'Cell line-specific control of recombinant monoclonal antibody production by CHO cells', *Biotechnol Bioeng*, 106, pp. 938–951.
- Phan, N. T. N., Mohammadi, A. S., Dowlatshahi Pour, M. and Ewing, A. G. (2016) 'Laser Desorption Ionization Mass Spectrometry Imaging of Drosophila Brain Using Matrix Sublimation versus Modification with Nanoparticles', *Anal Chem*, 88, pp. 1734–1741.
- Pilbrough, W., Munro, T. P. and Gray, P. (2009) 'Intracolonial protein expression heterogeneity in recombinant CHO cells', *PLoS ONE*, 4, pp. 8432–8439.
- Porter, A. J., Dickson, A. J. and Racher, A. J. (2010) 'Strategies for selecting recombinant CHO cell lines for cGMP manufacturing: realizing the potential in bioreactors', *Biotechnol Prog*, 26, pp. 1446–1454.
- Rabbani, S. (2010) *Advances in Time-of-Flight Secondary Ion Mass Spectrometry for the Analysis of Single Cells on Sub-Cellular Scale*. University of Manchester.
- Rabbani, S., Fletcher, J. S., Lockyer, N. P. and Vickerman, J. C. (2011) 'Exploring subcellular imaging on the buncher-ToF J105 3D chemical imager', *Surf Interface Anal*, 43, pp. 380–384.
- Robinson, M. (2013) *Chemical Analysis of Cells and Tissues with Time-of-Flight Secondary Ion Mass Spectrometry*. University of Washington.
- Robinson, M. a., Graham, D. J., Morrish, F., Hockenbery, D. and Gamble, L. J. (2016) 'Lipid analysis of eight human breast cancer cell lines with ToF-SIMS', *Biointerphases*, 11, p. 02A303.
- Roman-Fernandez, A. and Bryant, D. M. (2016) 'Complex Polarity: Building Multicellular Tissues Through Apical Membrane Traffic', *Traffic*, 17, pp. 1244–1261.
- Römpp, A., Guenther, S., Schober, Y., Schulz, O. and Takats, Z. (2010) 'Histology by Mass Spectrometry: Label-Free Tissue Characterization Obtained from High-Accuracy Bioanalytical Imaging', *Angew Chem Int Ed Engl*, 49, pp. 3834–3838.
- Schaepe, K., Kokesch-Himmelreich, J., Rohnke, M., Wagner, A.-S., Schaaf, T., Wenisch, S. and Janek, J. (2015) 'Assessment of different sample preparation routes for mass spectrometric monitoring and imaging of lipids in bone cells via ToF-SIMS', *Biointerphases*, 10, pp. 16–25.
- Schober, Y., Guenther, S., Spengler, B. and Römpp, A. (2012) 'Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging', *Anal Chem*, 84, pp. 6293–7.
- Seeley, E. H. and Caprioli, R. M. (2012) 'MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives', *Trends Biotechnol*, 29, pp. 136–143.
- Sellick, C. A., Croxford, A. S., Maqsood, A. R., Stephens, G., Westerhoff, H. V., Goodacre, R. and Dickson, A. J. (2011) 'Metabolite profiling of recombinant CHO cells:

- Designing tailored feeding regimes that enhance recombinant antibody production', *Biotechnol Bioeng*, 108, pp. 3025–3031.
- Shanta, S. R., Kim, T. Y., Hong, J. H., Lee, J. H., Shin, C. Y., Kim, K.-H., Kim, Y. H., Kim, S. K. and Kim, K. P. (2012) 'A new combination MALDI matrix for small molecule analysis: application to imaging mass spectrometry for drugs and metabolites', *The Analyst*, 137, pp. 5757–62.
- Shanta, S. R., Zhou, L. H., Park, Y. S., Kim, Y. H., Kim, Y. and Kim, K. P. (2011) 'Binary matrix for MALDI imaging mass spectrometry of phospholipids in both ion modes', *Anal Chem*, 83, pp. 1252–1259.
- Sparvero, L., Amoscato, A., Dixon, C., Long, J., Kochanek, P., Pitt, B., Bayir, H. and Kagan, V. (2012) 'Mapping of phospholipids by MALDI imaging (MALDI-MSI): realities and expectations', *Chem Phys Lipids*, 165, pp. 545–562.
- Spengler, B. (2015) 'Mass Spectrometry Imaging of Biomolecular Information', *Anal Chem*, 87, pp. 64–82.
- Steenbergen, R., Nanowski, T. S., Beigneux, A., Kulinski, A., Young, S. G. and Vance, J. E. (2005) 'Disruption of the phosphatidylserine decarboxylase gene in mice causes embryonic lethality and mitochondrial defects', *J Biol Chem*, 280, pp. 40032–40040.
- Steinhoff, R. F., Krismer, J., Eyer, K., Fagerer, S. R., Ibaez, A., Pabst, M. and Zenobi, R. (2014) 'Rapid estimation of the energy charge from cell lysates using matrix-assisted laser desorption/ionization mass spectrometry: Role of in-source fragmentation', *Anal Biochem*, 447, pp. 107–113.
- Steven, R. T., Race, A. M. and Bunch, J. (2013) 'Para-nitroaniline is a promising matrix for MALDI-MS imaging on intermediate pressure MS systems', *J Am Soc Mass Spectrom*, 24, pp. 801–804.
- Tian, H., Fletcher, J. S., Thuret, R., Henderson, A., Papalopulu, N., Vickerman, J. C. and Lockyer, N. P. (2014) 'Spatiotemporal lipid profiling during early embryo development of *Xenopus laevis* using dynamic ToF-SIMS imaging', *J Lipid Res*, 55, pp. 1970–1980.
- Touboul, D., Brunelle, A., Halgand, F., De La Porte, S. and Laprévote, O. (2005) 'Lipid imaging by gold cluster time-of-flight secondary ion mass spectrometry: application to Duchenne muscular dystrophy', *J Lipid Res*, 46, pp. 1388–95.
- Tsai, Y.-H., Bhandari, D. R., Garrett, T. J., Carter, C. S., Spengler, B. and Yost, R. A. (2016) 'Skeletal muscle fiber analysis by atmospheric pressure scanning microprobe matrix-assisted laser desorption/ionization mass spectrometric imaging at high mass and high spatial resolution', *Proteomics*, 14, pp. 2179–2189.
- Vanbellingen, Q. P., Elie, N., Eller, M. J., Della-Negra, S., Touboul, D. and Brunelle, A. (2015) 'Time-of-flight secondary ion mass spectrometry imaging of biological samples with delayed extraction for high mass and high spatial resolutions', *Rapid Commun Mass Spectrom*, 29, pp. 1187–1195.
- Vance, J. E. (2008) 'Phosphatidylserine and phosphatidylethanolamine in mammalian cells: two metabolically related aminophospholipids', *J Lipid Res*, 49, pp. 1377–87.
- Vance, J. E. and Tasseva, G. (2013) 'Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells', *Biochim Biophys Acta*, 1831, pp. 543–554.
- Vekemans, K., Rosseel, L., Wisse, E. and Braet, F. (2004) 'Immuno-localization of Fas and FasL in rat hepatic endothelial cells: Influence of different fixation protocols', *Micron*, 35, pp. 303–306.

- Vickerman, J. and Briggs, D. (2001) *ToF-SIMS: surface analysis by mass spectrometry*. 2nd edn. Manchester: Surface Spectra. doi: 10.1255/tof2ch1.
- Vickerman, J. C. (2011) 'Molecular imaging and depth profiling by mass spectrometry--SIMS, MALDI or DESI?', *The Analyst*, 136, pp. 2199–2217.
- Walch, A., Rauser, S., Deininger, S. O. and Höfler, H. (2008) 'MALDI imaging mass spectrometry for direct tissue analysis: A new frontier for molecular histology', *Histochem Cell Biol*, 130, pp. 421–434.
- Weaver, E. M. and Hummon, A. B. (2013) 'Imaging mass spectrometry: From tissue sections to cell cultures', *Adv Drug Deliv Rev*, 65, pp. 1039–1055.
- Yang, J. and Caprioli, R. M. (2011) 'Matrix Sublimation/Recrystallization for Imaging Proteins by Mass Spectrometry at High Spatial Resolution', *Anal. Chem.*, 83, pp. 5728–5734.
- Zenobi, R. (2013) 'Single-Cell Metabolomics: Analytical and Biological Perspectives', *Acta Neuropathol*, 342, pp. 1201–1211.
- Zhang, S. and Yao, Z.-P. (2012) 'Improved detection of phosphopeptides by negative ion matrix-assisted laser desorption/ionization mass spectrometry using a proton sponge co-matrix', *Anal Chim Acta*, 711, pp. 77–82.
- Zhang, Y., Li, H., Ma, Y. and Lin, J. (2014) 'Lipid profiling of mammalian cells with in situ matrix-assisted laser desorption ionization-mass spectrometry', *Sci China Chem*, 57, pp. 442–446.
- Zhou, L. H., Kang, G.-Y. and Kim, K. P. (2009) 'A binary matrix for improved detection of phospho-peptides in matrix-assisted laser desorption/ionization mass spectrometry', *Rapid Comm Mass Spectrom*, 23, pp. 2264–2272.
- Zhu, L., Stadler, J., Schmitz, T. A., Krumeich, F. and Zenobi, R. (2011) 'Atmospheric Pressure Sampling for Laser Ablation Based Nanoscale Imaging Mass Spectrometry: Ions or Neutrals?', *J. Phys. Chem. C*, 115, pp. 1006–1016.

CHAPTER 8: General discussion, conclusions and future work

The aim of the work undertaken in this project was to combine the use of various bio-analytical mass spectrometry platforms in metabolomic studies related to EPO-Fc fusion protein expression. Furthermore, the application of mass spectrometry techniques in the analysis of metabolites within small populations of mammalian cells and, ultimately single cells, was another objective of the study.

Firstly, the work began with the attempt to investigate the CHO cell line response to the inducer (*i.e.* tetracycline) application and therefore, EPO-Fc fusion protein production. It was crucial to confirm that the fusion protein production in CHO cells was due to the inducer addition as well as that the cell growth was stable. The investigation revealed that the EPO-Fc protein production started after ~24 h from applying the inducer and that the protein was secreted into cell growth medium. The result also suggested that the EPO-Fc production may have increased with progression of the cell culture growth. This however would need to be confirmed with specific productivity tests *i.e.* tests of how much protein is being produced per cell per unit time (Lloyd *et al.*, 2000). Additionally, the results could also be improved by protein purification that would enable for more precise EPO-Fc protein identification. Determining the relationship between protein productivity and cell culture growth, hence, cell cycle phase could bring new insights into metabolism of induced cells. Moreover, it would also be beneficial to perform the analysis of the cell line that has not been transfected with EPO-Fc-pcDNA5-FRT-TO plasmid. In this study, the aim was to investigate the direct influence of EPO-Fc fusion protein expression on metabolism of CHO cells. Nevertheless, the analysis of cell line without EPO-Fc-pcDNA5-FRT-TO plasmid could give additional information about the influence of this element on cell morphology, growth and CHO cell metabolism.

In this work, two MS techniques *i.e.* GC-MS and MALDI-MS were employed in order to obtain profiles of external and internal metabolomes of cells producing EPO-Fc. Unlike GC-MS that is used as a ‘*gold standard*’ for metabolomics studies, MALDI-MS requires additional attention due to its challenging sample preparation process, specifically when it comes to metabolomics research. Therefore, the optimisation of the MALDI-MS method for metabolite profiling of CHO cells in both ionisation modes was performed. The results of metabolic profiles obtained with MALDI-MS confirmed a suitability of optimised sample preparation method for this type of analyte, and gave rich spectral information. However, further investigation can be performed with the application of different matrix-to-analyte ratios that could increase the detection of a number of small molecules or their specific class. It has been previously shown (Vaidyanathan, Gaskell and Goodacre, 2006)

that decreasing matrix-to-analyte molar ratio can enhance the detection of small molecules (<500 Da) due to matrix suppression issue and therefore, this parameter can be further adjusted to improve metabolite detection.

The application of a non-targeted metabolomics approach for the analysis of metabolite changes in CHO cells under the protein production can be considered successful. The employment of GC-MS and MALDI-MS has provided information about the response of cells to the EPO-Fc production, allowing for the investigation of changes in key metabolites within internal and external cell metabolomes. Both techniques, being complementary to each other, produced a wide range of detected metabolites and, with the application of MVA, a number of metabolites were identified to be significantly influenced by the protein expression when compared to non-induced cells. Correlation of these significant metabolites and reconstruction of the metabolite map, classified metabolome changes into metabolic pathways influenced by the production of EPO-Fc. However, the analysis of the data was limited due to the low number of metabolites recognised to be correlated to specific metabolic routes.

A differentiation between metabolites derived from internal and external metabolomes was also seen. However, the interpretation of internal and external metabolome data needs to be carefully performed. This is because, for example, some metabolites can be expected to be enriched inside cells but, as they are permeable through the cell membrane, are seen outside the cell. Moreover, some metabolites present outside the cell can become very diluted and therefore, any differences may not be significant enough for detection. Therefore, targeted metabolic profiling focused on the detection of separate classes of metabolites and with the application of different MALDI-MS matrices or techniques, *e.g.* LC-MS could also be beneficial in this study. The advantage of LC-MS over GC-MS is the ability to detect a larger pool of metabolites with no chemical modification of a sample required, which makes the sample preparation easier and the analysis time shorter. In addition, high-throughput possibilities of currently used high performance LC-MS methods (HPLC-MS) allows for qualitative and quantitative analysis of a number of metabolites (Dunn, Bailey and Johnson, 2005).

Biological interpretation of internal and external metabolites that were influenced by EPO-Fc expression in CHO cells highlighted that changes observed in protein producing cells are related to enhanced protein and lipid synthesis. Similar results of changed lipid content in cells producing EPO-Fc were seen from MS imaging data thus, it implies that the targeted investigation of the lipids could be further preformed in order to bring new insights of the effect of the fusion protein production on lipid metabolism. Targeted

metabolite profiling could be carried out to further explore influence of EPO-Fc production and increased metabolic activity that was observed in CHO cells along with the adaptation to the lack of crucial substrates for energy-producing pathways. Such targeted investigations would help with the detection of a larger number of TCA cycle intermediates hence, expanding and validating correlations between specific pathways. Also, the exact identification and quantification of these metabolites would validate the biochemical interpretations of this study.

The information about metabolic changes in tetracycline-induced CHO cells obtained from the MSI analysis of cell populations was confirmed with the analysis based on single-cell studies. In addition, the analysis of individual cells highlighted other changes in metabolic content of induced cells to those seen in the cell-population based study. This therefore proves that investigations based on single-cell analysis provide additional insights about changes in metabolism of tetracycline-induced cells that can be referred to a unique, individual cell. Moreover, single cell studies offer the potential to correlate metabolic profiles of cells producing EPO-Fc to their production efficiency that could identify new goals for biomanufacturing companies.

The application of MALDI-MS and SIMS imaging in the study of CHO cells allowed for an investigation of the high level of diversity of this cell population. The analysis of CHO cells showed extended heterogeneity within one cell population with a number of significant metabolites that could have been correlated to a specific type of cells. The correlation between these metabolites for each cell type could have been used to define specific metabolic characteristic of each of these cell types and allowed for defining their localisation within whole cell population.

Heterogeneity within a cell population is a result of a number of different factors *e.g* cell cycle phase. Therefore, various approaches could be applied to improve investigation of cell populations diversity. One of the approaches could be cell cycle phase synchronisation of cultured cells in order to allow the cell cycle-independent investigations of metabolism and cellular physiology. This could be done with the application of typical techniques employed for cell cycle studies that use chemical agents for arresting cells in a specific phase. However, these methods usually modify cells' homeostasis and cell environment (Hung *et al.* 1996, Cooper 2003). Cell behaviour could then be changed by the synchronisation agent that would hinder results of the investigation of interest. Another limitation of these methods is that only several viable cells are recovered after such treatment (Mendoza-Perez *et al.*, 2016). Nevertheless, there are other approaches in cell cycle synchronisation such as the possibility to separate cells according to their cell phase

by centrifugal elutriation. This technique allows to isolate cells at different cell phase based on different cell size that is correlated to cell cycle stages. This technique however, requires several injections in order to obtain high cell numbers, resulting in increased cell concentration and cell stress (Lloyd *et al.*, 2000).

In this study, the application of two different SIMS primary ion sources allowed for a more broad analysis of chemical content of CHO cells. With the use of a polyatomic C_{60}^+ ion beam, it was possible to investigate different layers of CHO cells, allowing for metabolic analysis of depth profiles of these cells. However, the spatial resolution was not optimum hence, further experiments could be performed with the use of more focused C_{60}^+ ion beam (up to $\sim 1\text{-}2\ \mu\text{m}$), with the loss of beam current and therefore sensitivity. This would however bring possibilities of creating high-resolution spectral maps with specific chemical content of CHO cells in 3D visualisation.

By contrast, the application of the Au_3^+ primary ion source allowed for high resolved images of the top layer of CHO single cells. With the use of this ion beam, it was possible to present detailed biochemical information along with metabolite fragments' localisation within individual CHO cells. Nevertheless, more experiments for SIMS using both C_{60}^+ and Au_3^+ primary sources and MALDI-MS could also be executed in order to increase number of repeats to enhance confidentiality of statistical analysis. Additionally, with a recent development of the SIMS instrument (*3D OrbiSIMS*, 2017) future investigations could be performed with not only high spatial resolution but also with very high mass accuracy.

Another complementary technique that could be used in the study of cells producing EPO-Fc is Raman spectroscopy. Raman spectroscopy can be applied for *in vivo* single-cell based metabolomics studies and has become a powerful tool in life sciences over past few years (Wu *et al.* 2011, Brauchle & Schenke-Layland 2013). Raman spectroscopy is a laser-based technology, allowing for quantitative molecular information from cells and tissues in a label-free and real-time (or near real-time) manner. It can provide biochemical information of biological samples such as cells and tissues with high spatial resolution ($\sim 2\text{-}5\ \mu\text{m}$) (Konorov *et al.*, 2013, Stewart *et al.*, 2012). Additionally, Raman measurements are non-destructive and do not need sample processing, making the Raman technique a robust and reliable platform for a number of applications *e.g.* in the investigations of biochemical processes occurring in cells such as growth and cell development (Konorov *et al.*, 2013) and in the studies of biochemical distribution of different metabolites within single cells along with quantitative approach (Hedegaard, Bergholt and Stevens, 2016). The Raman

technique however, suffers from weak signal intensities that lead to low sensitivity, making it problematic to measure low concentrations of the analytes (Konorov *et al.*, 2013).

One more possibility to widen the study of CHO cells would be to perform imaging studies incorporating another approach – ion mobility. Ion mobility mass spectrometry can be applied to provide separation of ions according to their molecular structure (*e.g.* shape) and is often couple to MS analysis (Blatherwick *et al.*, 2013). This would enable different compound groups to be separated, giving a possibility to distinguish matrix molecules as well as would allow for more accurate metabolite classification. In addition, the MS/MS approach could also be applied to enhance structural identification of metabolites detected. Metabolomics is a great tool to investigate changes in cellular activity but, despite of all its advantages, there is still a great deal to be done. The limitations such as *e.g.* availability of relevant databases (for SIMS) and the need of suitable bioinformatics engines and data analysis work-flows represent major challenges. In addition, further requirements specifically important for single-cell based analyses include *e.g.* more extensive coverage of the metabolome, high throughput experiments and non-destructive measurements. Nevertheless, this project has successfully shown combination of bio-analytical techniques to investigate external and internal metabolome changes related to a fusion protein production in mammalian cells. The significance of single cell metabolomics approaches has also been highlighted, providing insights of sub-cellular distribution of metabolites in cells producing EPO-Fc and information on the heterogeneity of cell populations. It has been found that the application of complementary techniques along with statistical data handling approach offer a valuable insights into unknown processes involved in cell metabolism.

References:

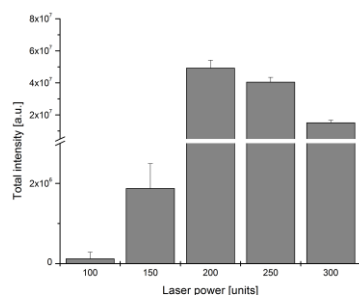
- 3D OrbiSIMS (2017). Available at: <http://www.npl.co.uk/news/3d-nanosims-label-free-molecular-imaging> (Accessed: 20 April 2017).
- Blatherwick, E. Q., Svensson, C. I., Frenguelli, B. G. and Scrivens, J. H. (2013) 'Localisation of adenine nucleotides in heat-stabilised mouse brains using ion mobility enabled MALDI imaging', *Intern J Mass Spectrom*, 345–347, pp. 19–27.
- Brauchle, E. and Schenke-Layland, K. (2013) 'Raman spectroscopy in biomedicine - non-invasive in vitro analysis of cells and extracellular matrix components in tissues', *Biotechnol J*, 8, pp. 288–297.
- Cooper, S. (2003) 'Rethinking synchronization of mammalian cells for cell cycle analysis', *Cell Mol Life Sci*, 60, pp. 1099–1106.
- Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005) 'Measuring the metabolome: current analytical technologies', *The Analyst*, 130, pp. 606–625.
- Hedegaard, M. A. B., Bergholt, M. S. and Stevens, M. M. (2016) 'Quantitative multi-image analysis for biomedical Raman spectroscopic imaging', *J Biophotonics*, 9, pp. 542–550.
- Hung, D. T., Jamison, T. F. and Schreiber, S. L. (1996) 'Understanding and controlling the cell cycle with natural products', *Chem Biol*, 3, pp. 623–639.
- Konorov, S. O., Schulze, H. G., Piret, J. M., Blades, M. W. and Turner, R. F. B. (2013) 'Label-Free Determination of the Cell Cycle Phase in Human Embryonic Stem Cells by Raman Microspectroscopy', *Anal Chem*, 85, pp. 8996–9002.
- Lloyd, D. R., Holmes, P., Jackson, L. P., Emery, A. N. and Al-Rubeai, M. (2000) 'Relationship between cell size, cell cycle and specific recombinant protein productivity', *Cytotechnology*, 34, pp. 59–70.
- Mendoza-Perez, E., Hernandez, V., Palomares, L. A. and Serrato, J. A. (2016) 'An integrated system for synchronous culture of animal cells under controlled conditions', *BioTechniques*, 61, pp. 129–136.
- Stewart, S., Priore, R. J., Nelson, M. P. and Treado, J. P. (2012) 'Raman Imaging', *Annu Rev Anal Chem*, 5, pp. 337–360.
- Vaidyanathan, S., Gaskell, S. and Goodacre, R. (2006) 'Matrix-suppressed laser desorption/ionisation mass spectrometry and its suitability for metabolome analyses', *Rapid Comm Mass Spectrom*, 20, pp. 1192–1198.
- Wu, H., Volponi, J. V, Oliver, A. E., Parikh, A. N., Simmons, B. A. and Singh, S. (2011) 'In vivo lipidomics using single-cell Raman spectroscopy', *Proc Natl Acad Sci USA*, 108, pp. 3809–3814.

SUPPLEMENTARY INFORMATION

CHAPTER 5: MALDI-MS sample preparation optimisation for metabolite profiling

Optimisation of the laser power for MALDI Synapt G2–Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK)

A.



B.

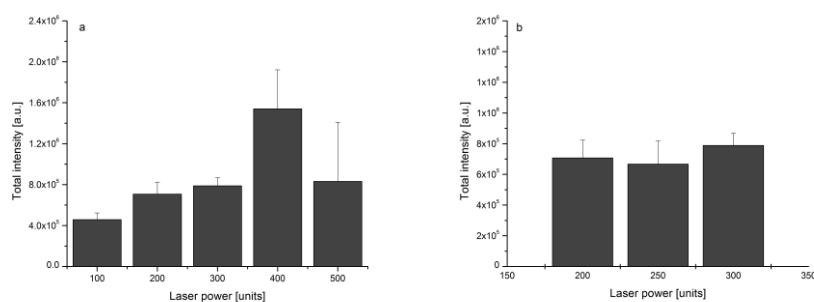


Figure 1: A comparison of total intensity of the calibration mixture in DHB (A) for MALDI positive ion mode and PEG600 sulfate in 9AA (B *a*, *b*) for negative ion mode according to the laser power used. The error bars represent the standard deviation calculated from the data obtained from two experiments ($n=2$).

Total intensity of ions was measured with different laser power applied to 10 mg/mL DHB matrix dissolved in ACN/water 1:1 (v/v) with 0.1% TFA mixed with calibration mixture (M:A, 1:1; mixed deposition). Laser power of 200 – 250 [units] was set as the most suitable one, giving the highest intensities and the lowest divergence between repeats.

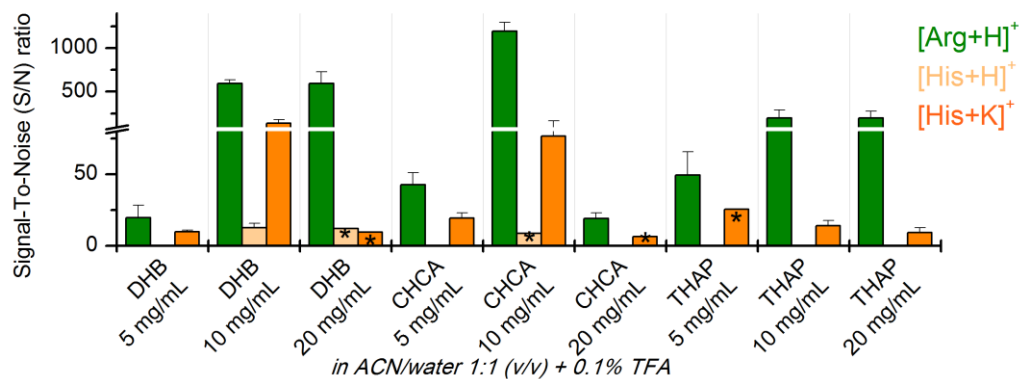
Total intensity of ions was measured with different laser power applied to 5 mg/mL 9AA matrix dissolved in 70% MeOH mixed with PEG600 sulfate (M:A, 1:1; mixed deposition). Laser power in a range of 200 – 300 [units] was set as the best one (Figure 1B-a), giving good intensities and small error between repeats. Follow up experiment with more measurements of this range (Figure 1B-b) resulted with more suitable range of the laser power between 200 – 250 [units].

A choice of matrix, its concentration and solvent for MALDI metabolite detection

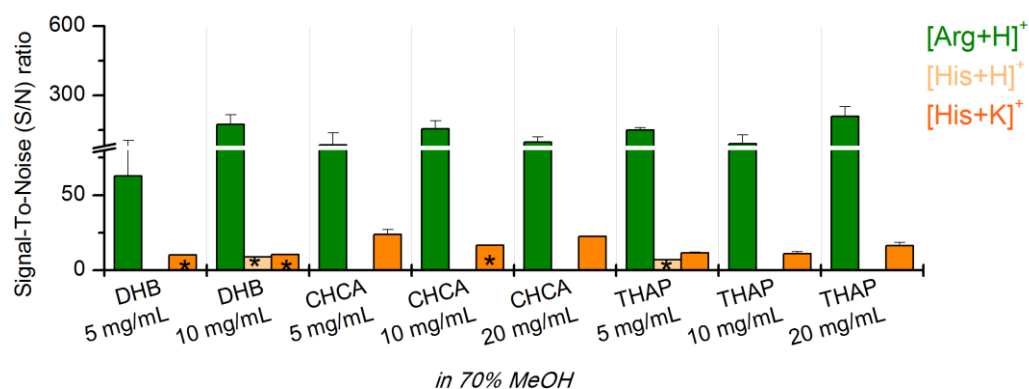
Positive ionisation mode

Comparison based on matrix solvent

A.



B.



C.

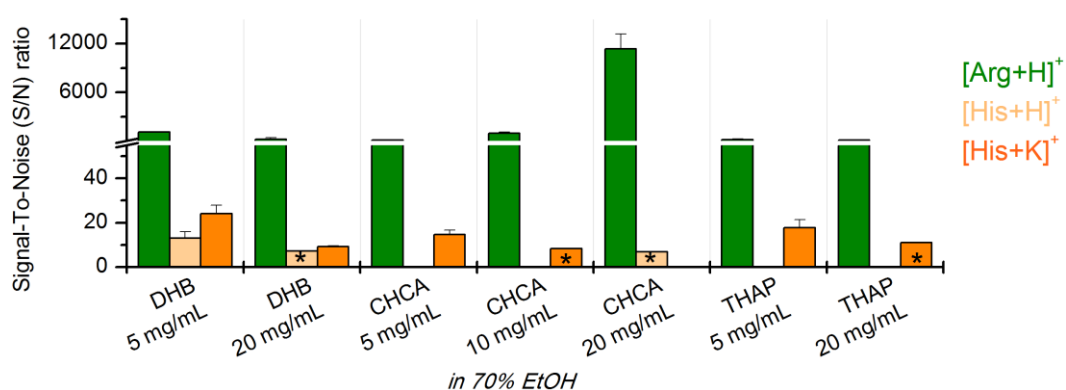
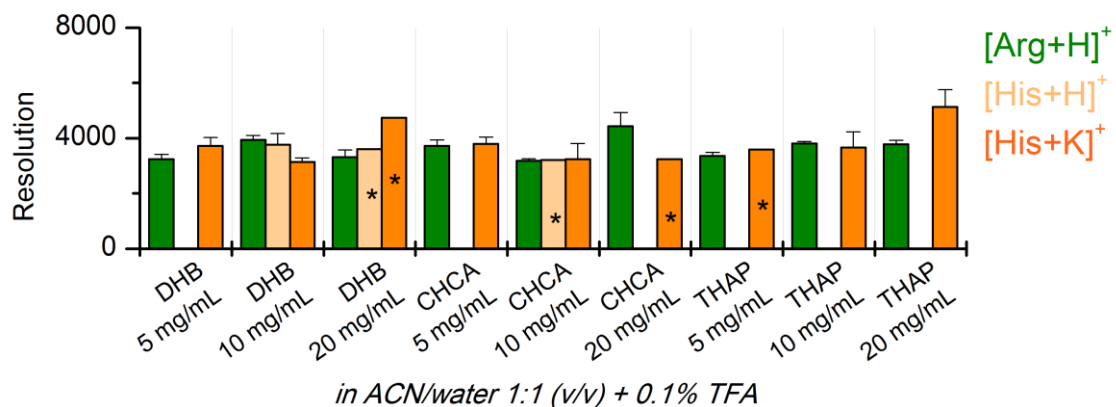
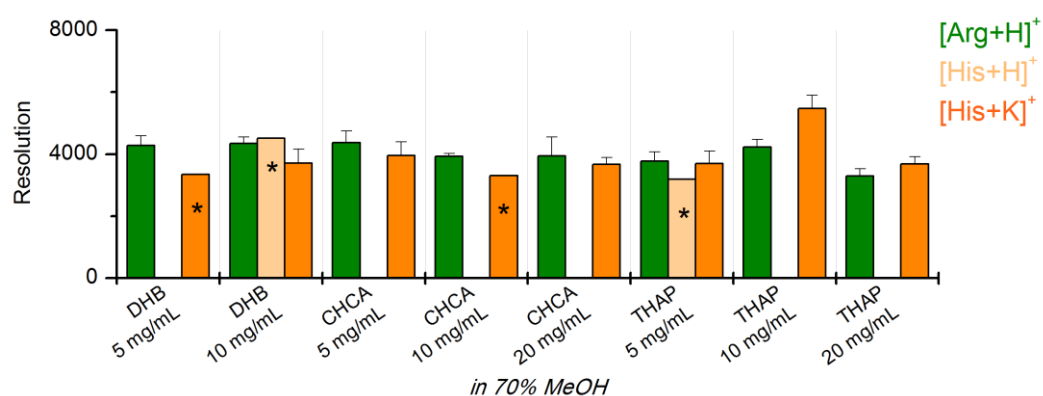


Figure 2: A comparison of S/N ratio of Arg and His ions in based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI⁺; * - one repeat only.

A.



B.



C.

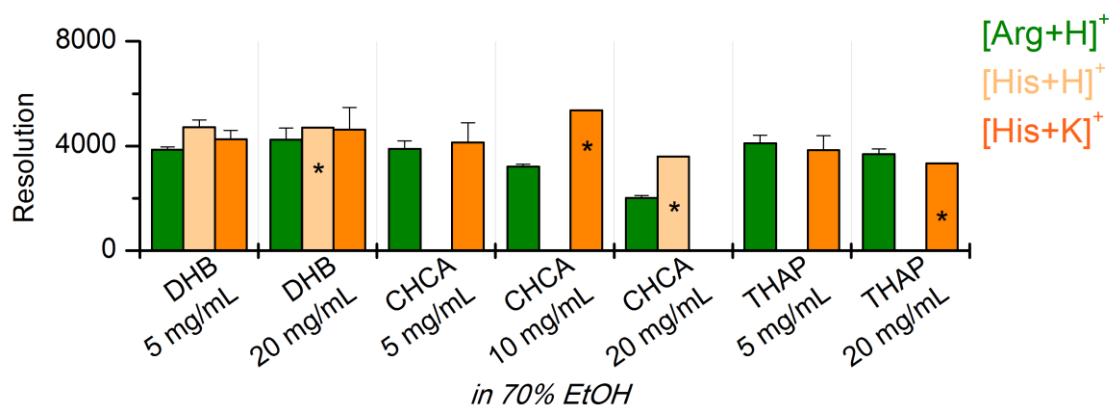
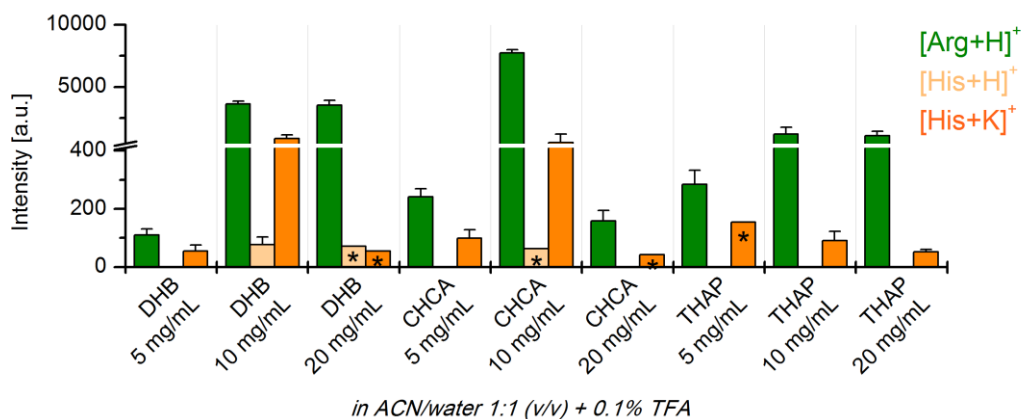
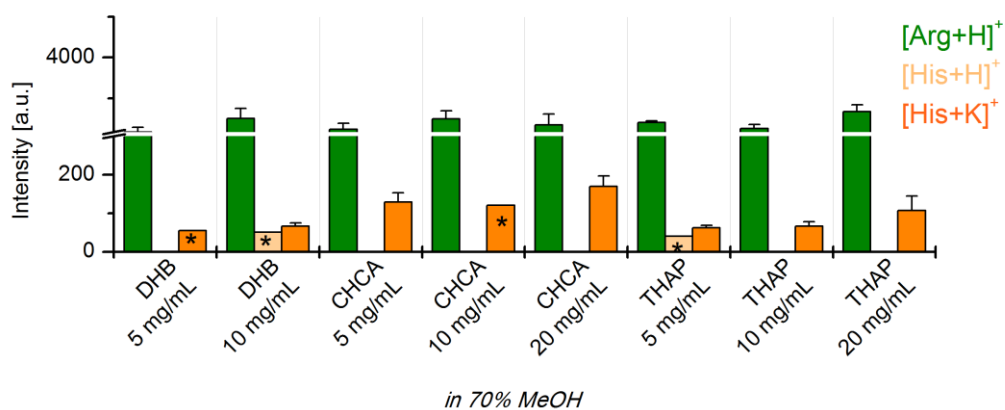


Figure 3: A comparison of mass resolution of Arg and His ions based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI⁺; * - one repeat only.

A.



B.



C.

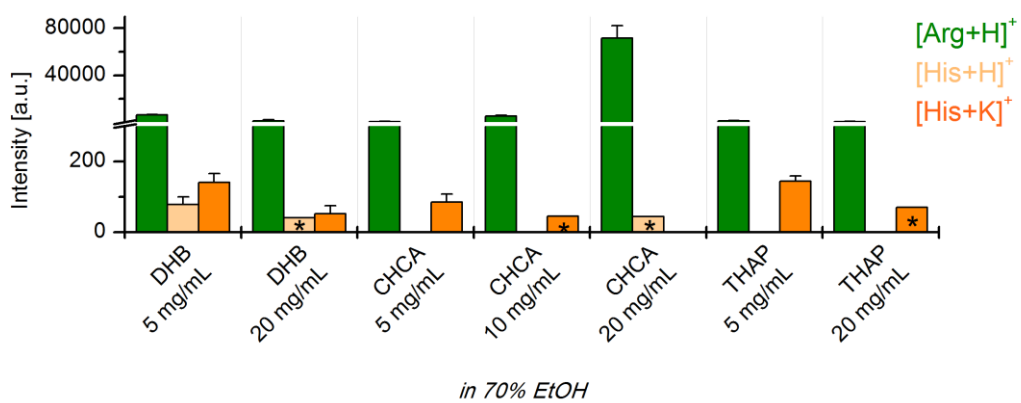
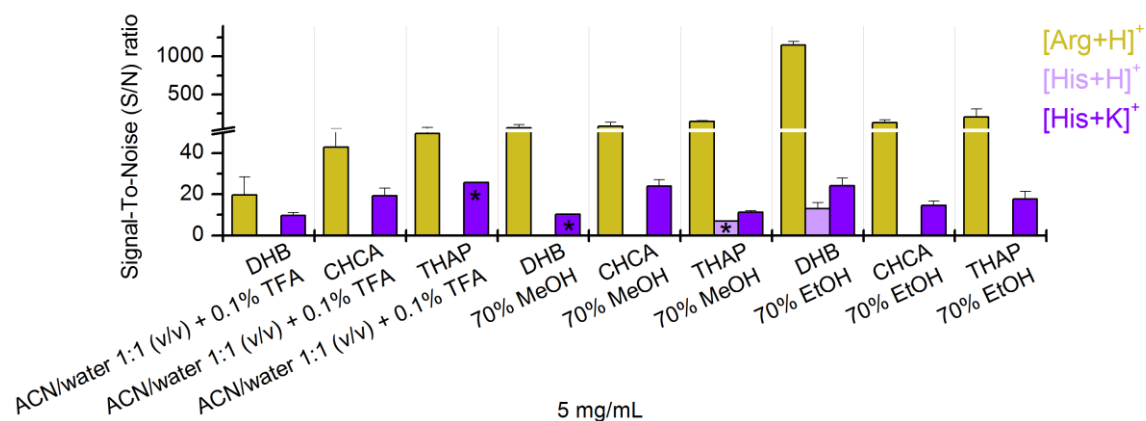


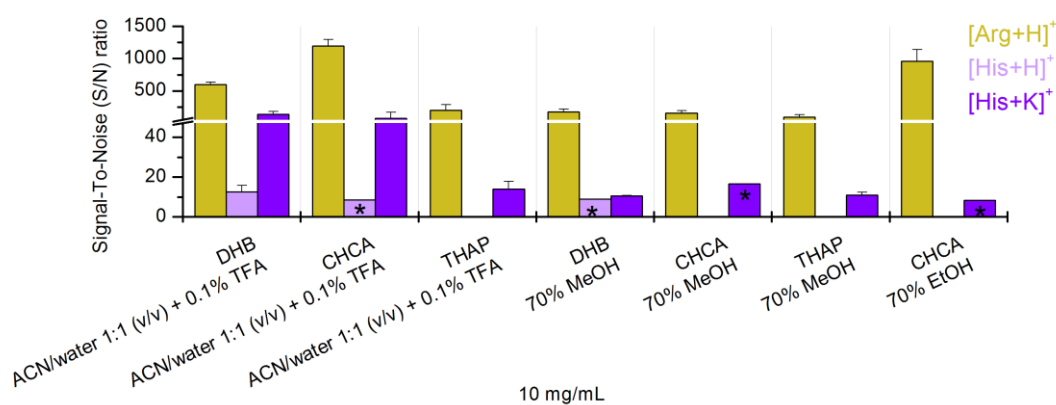
Figure 4: A comparison of the intensity of Arg and His ions based on matrix solvents: ACN/water 1:1 (v/v) with 0.1% TFA (A), 70% MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only.

Comparison based on matrix concentration

A.



B.



C.

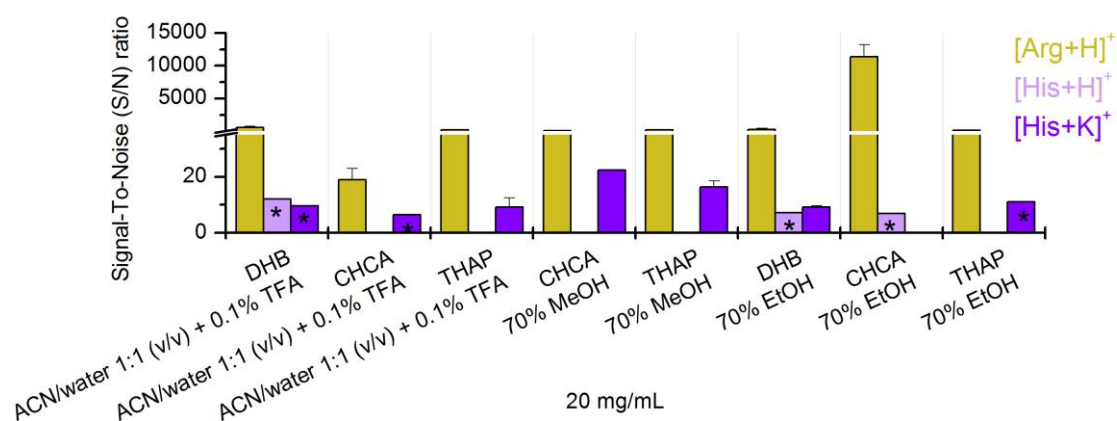
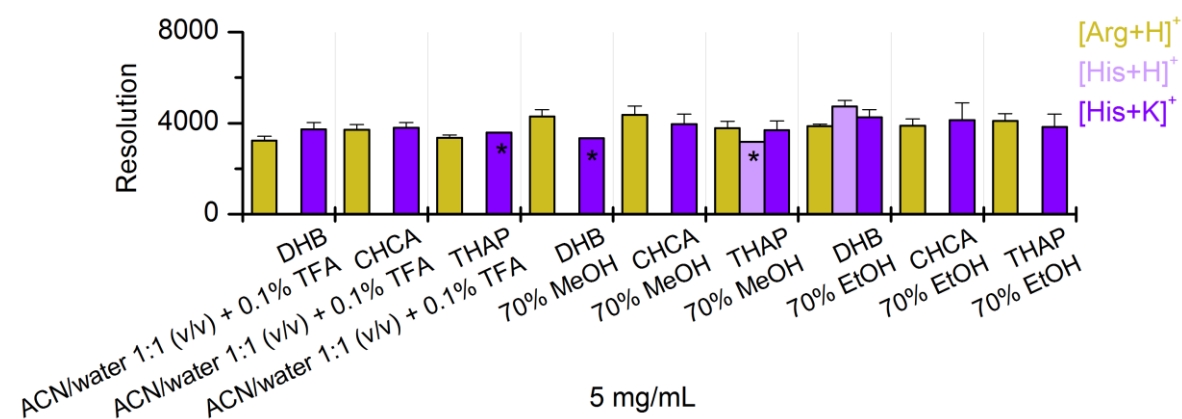
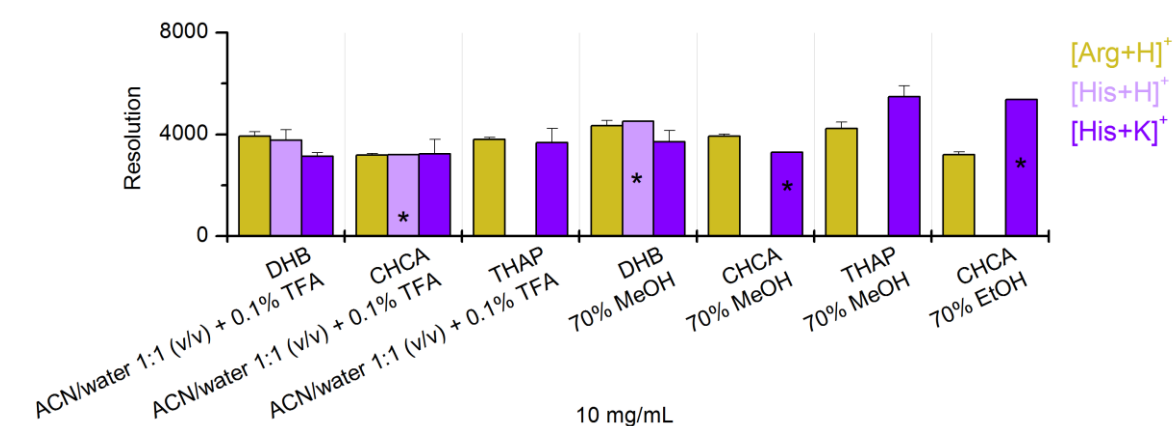


Figure 5: A comparison of S/N ratio of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only.

A.



B.



C.

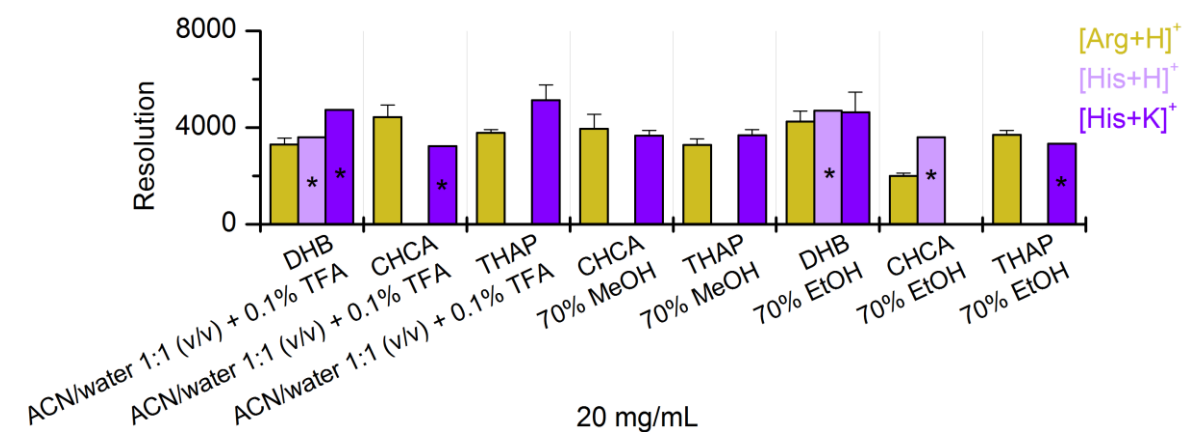
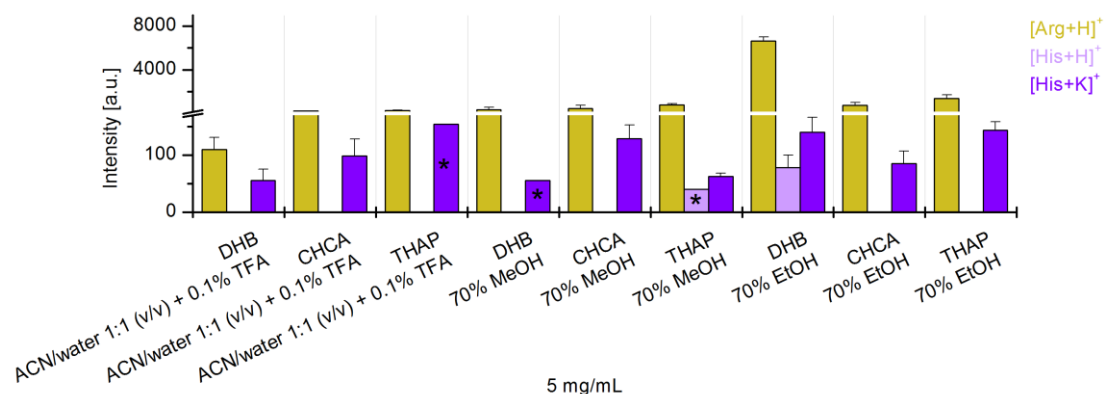
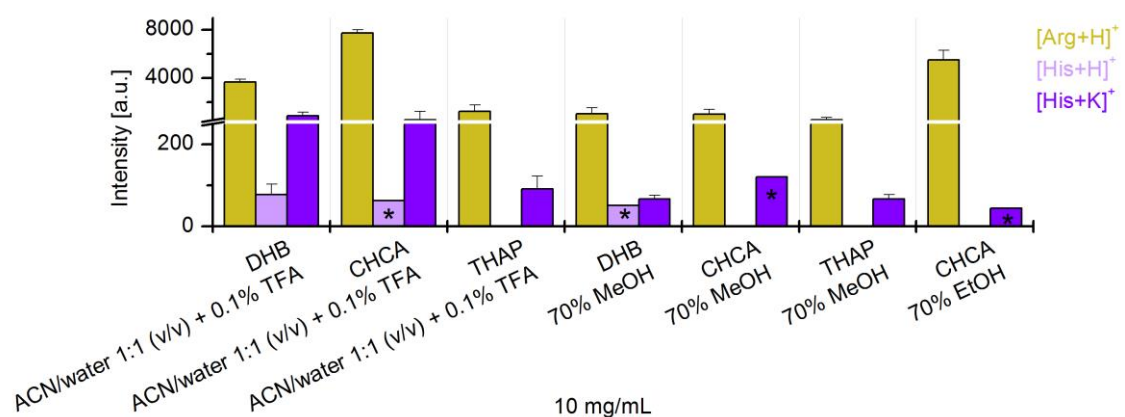


Figure 6: A comparison of mass resolution of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only.

A.



B.



C.

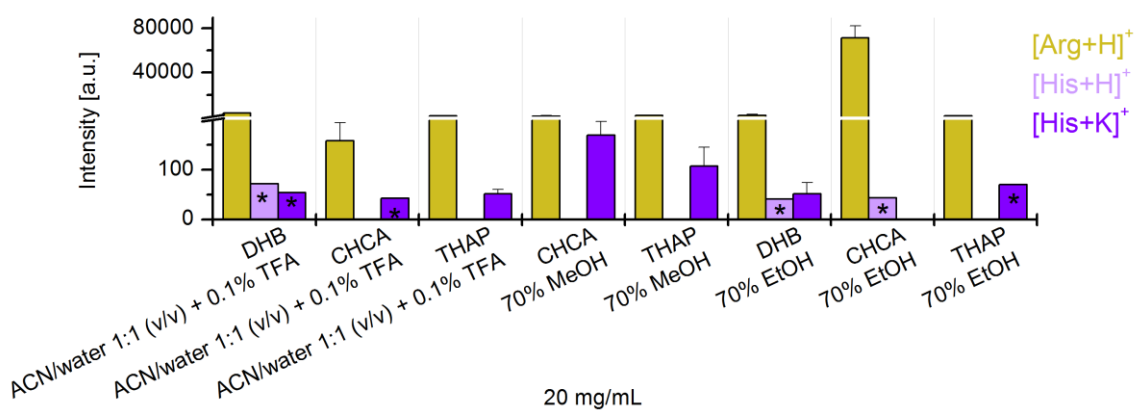
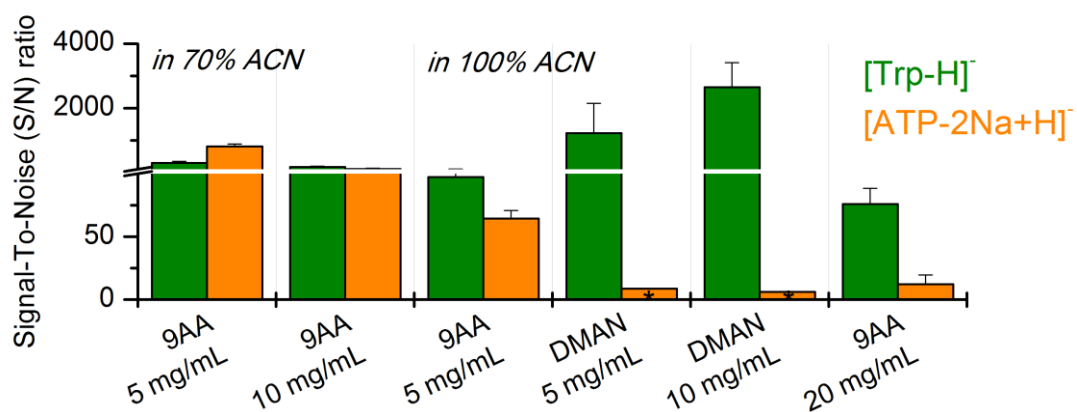


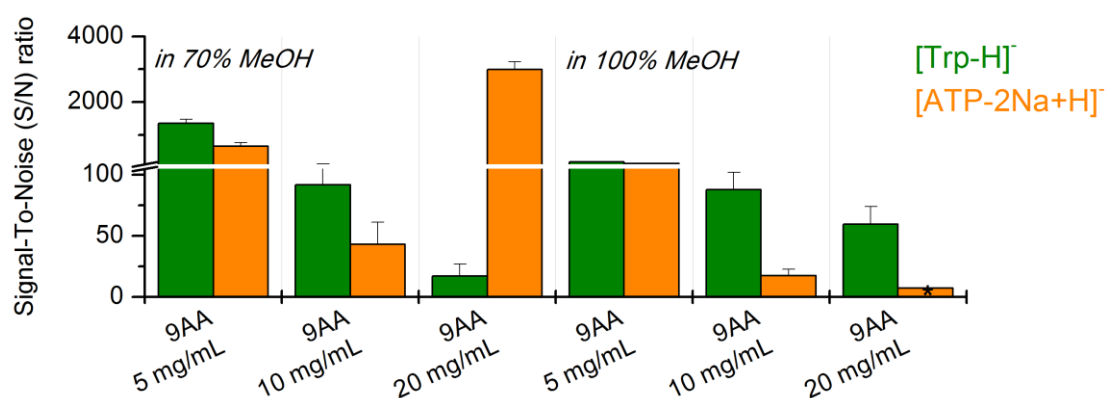
Figure 7: A comparison of the intensity of Arg and His ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI+; * - one repeat only.

Negative ionisation mode
Comparison based on matrix solvent

A.



B.



C.

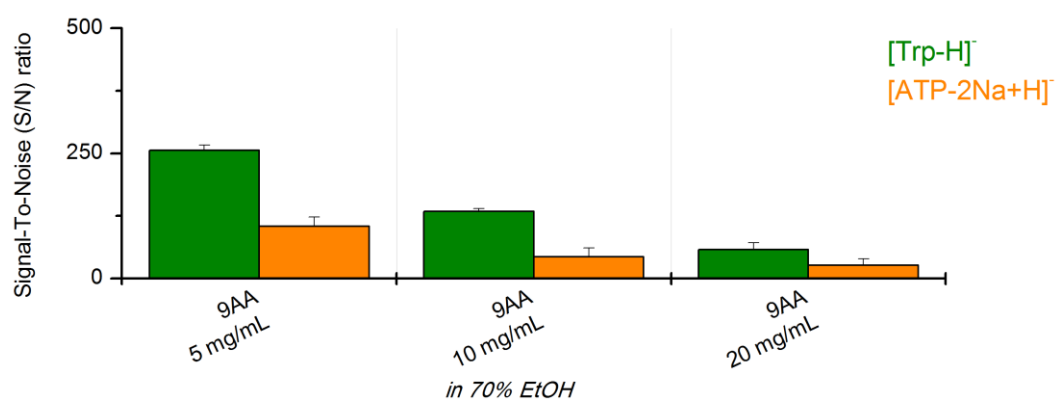
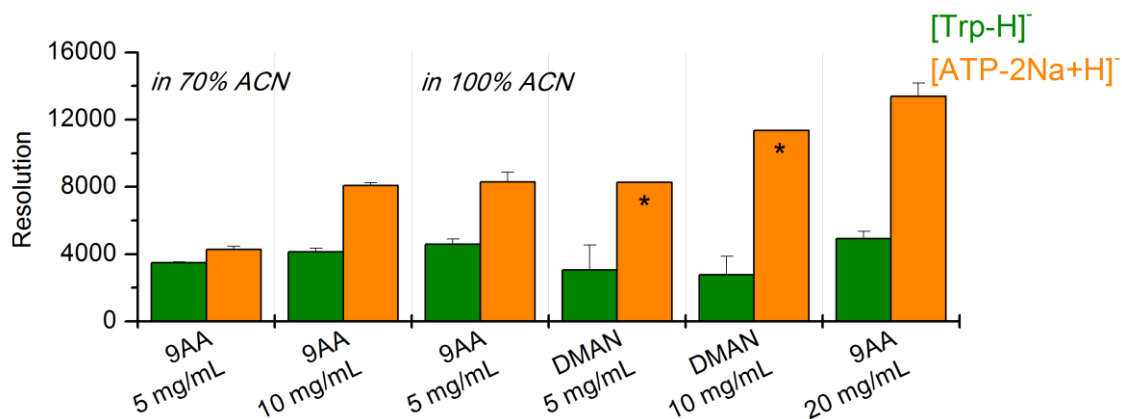
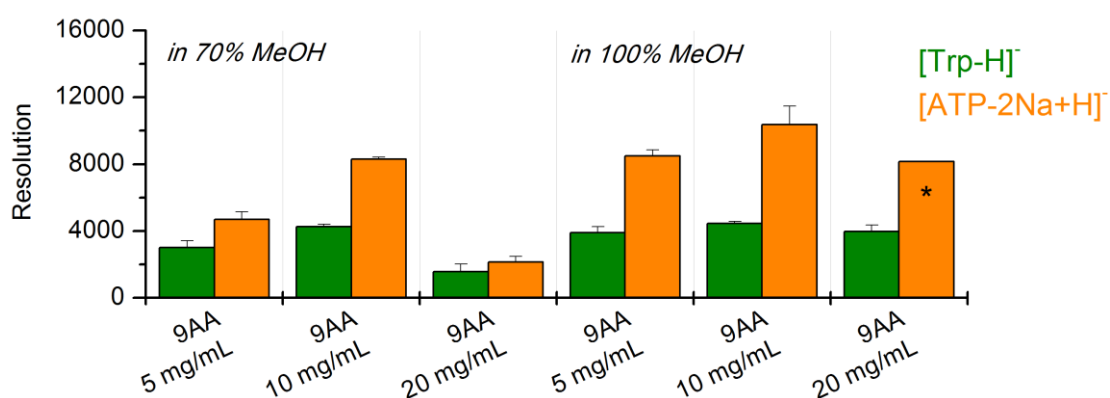


Figure 8: A comparison of S/N ratio of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

A.



B.



C.

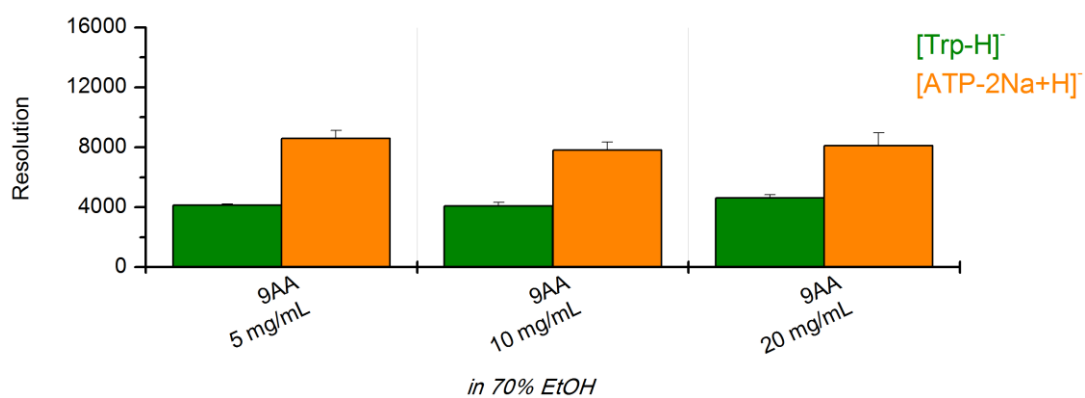
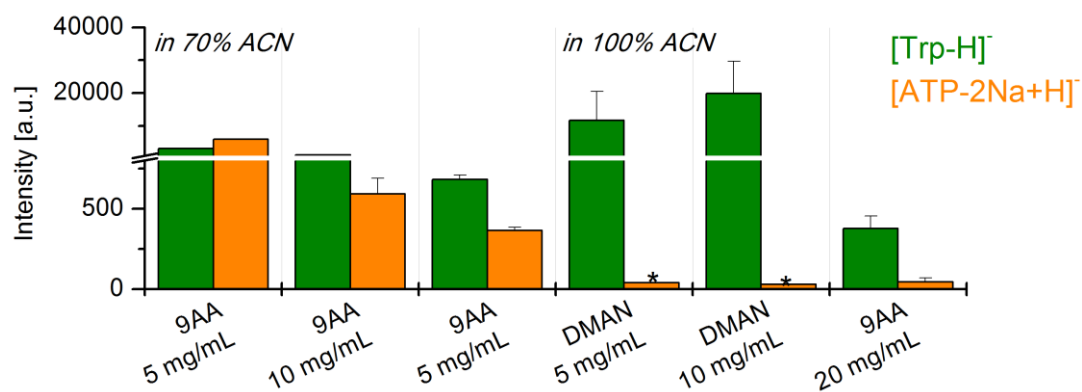
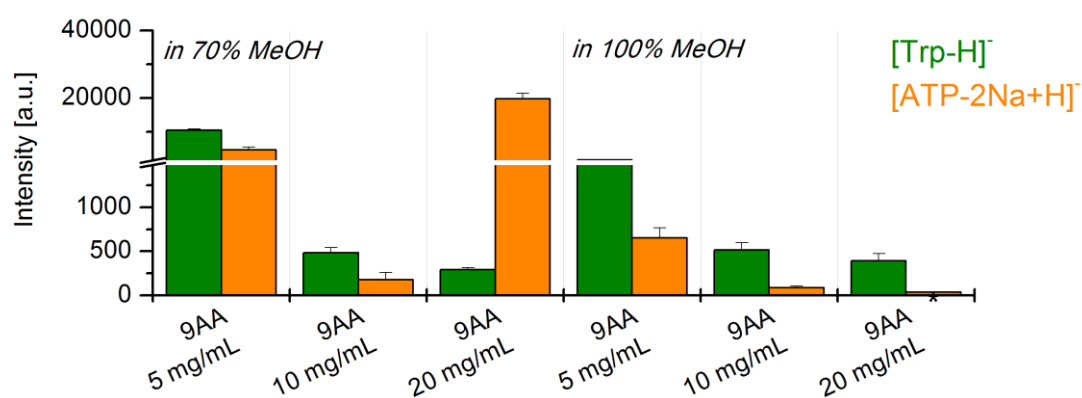


Figure 9: A comparison of mass resolution of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

A.



B.



C.

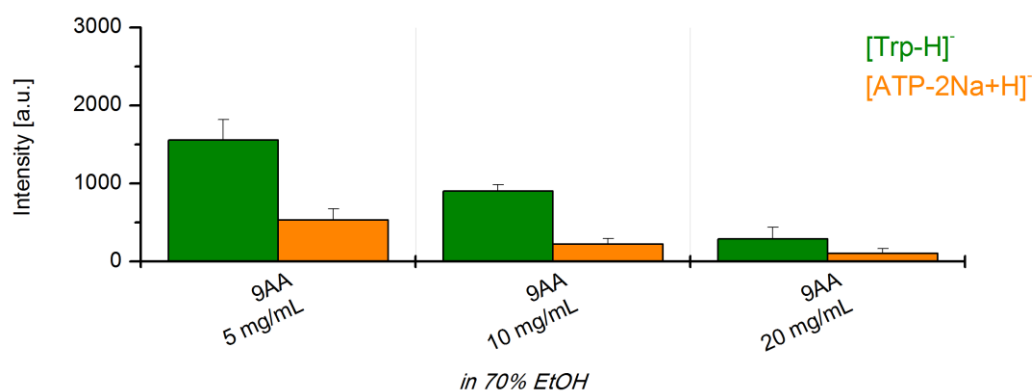
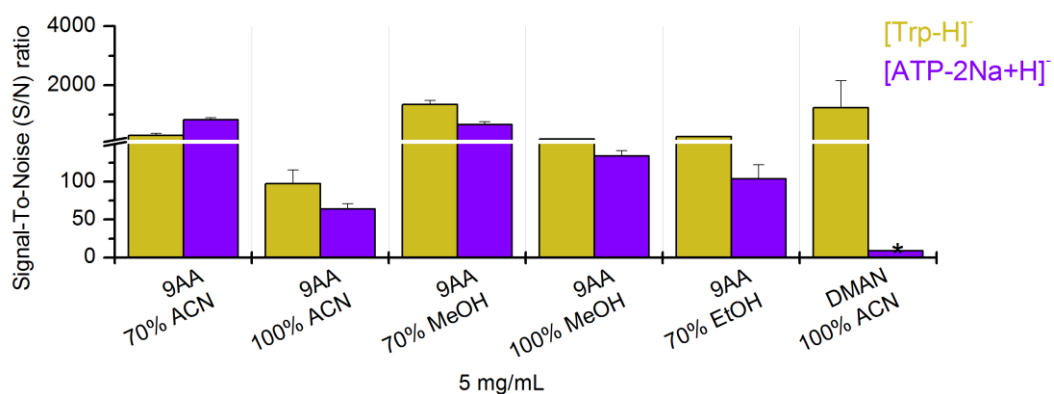


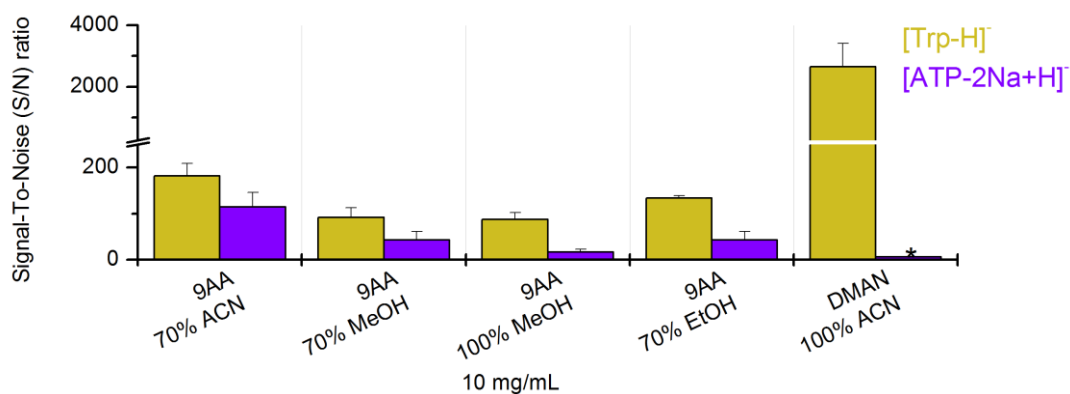
Figure 10: A comparison of the intensity of Trp and ATP ions based on matrix solvents: ACN (A), MeOH (B), 70% EtOH (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

Comparison based on matrix concentration

A.



B.



C.

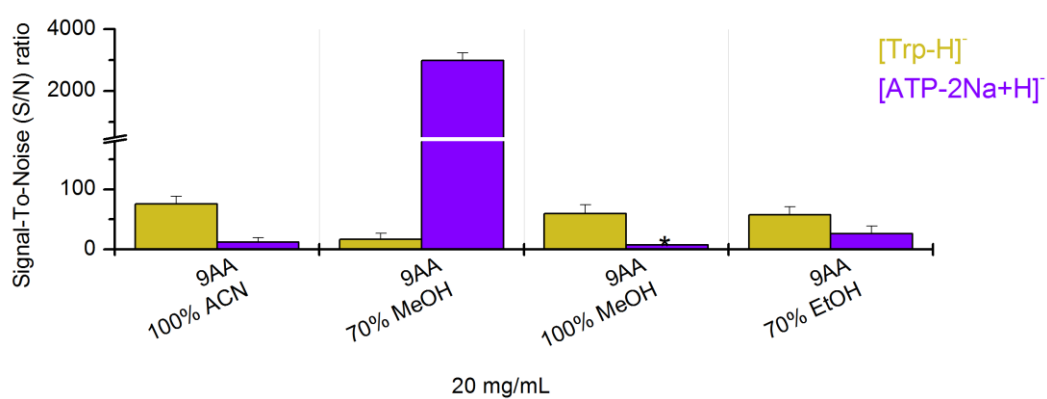
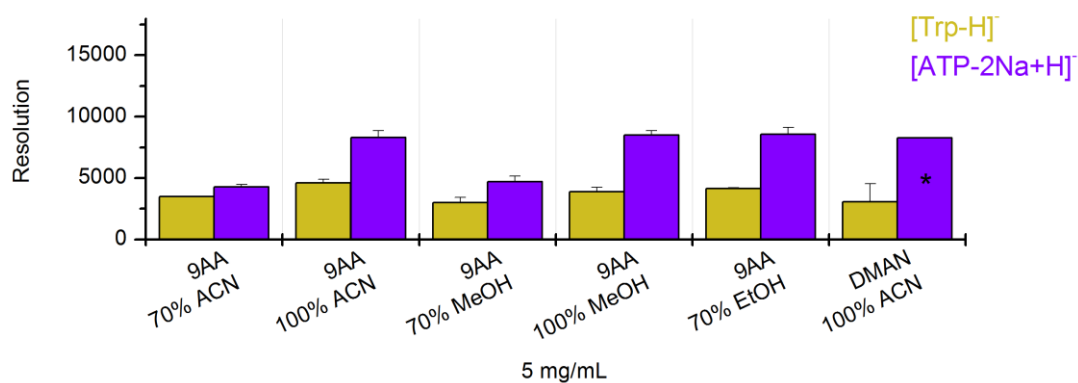
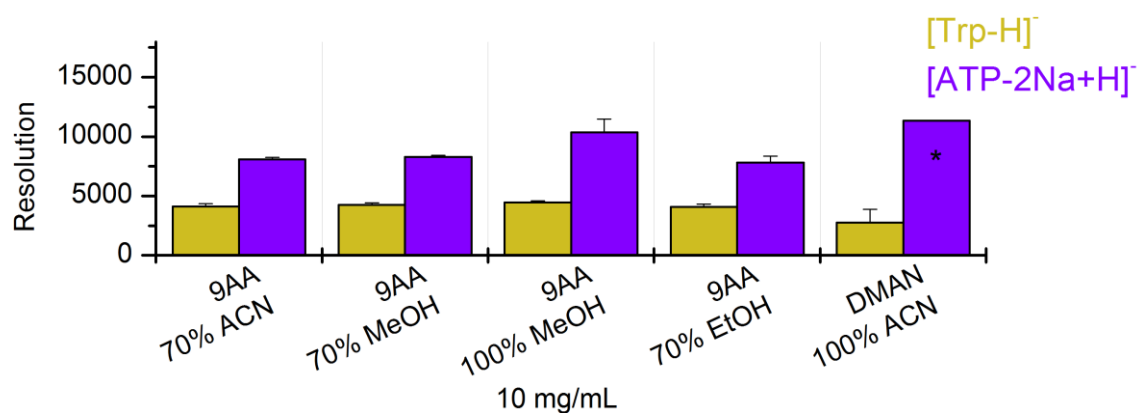


Figure 11: A comparison of S/N ratio of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

A.



B.



C.

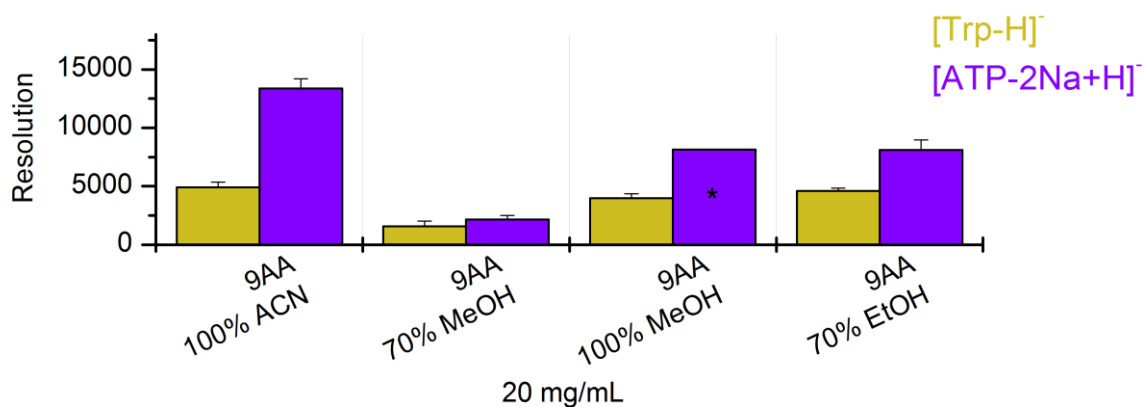
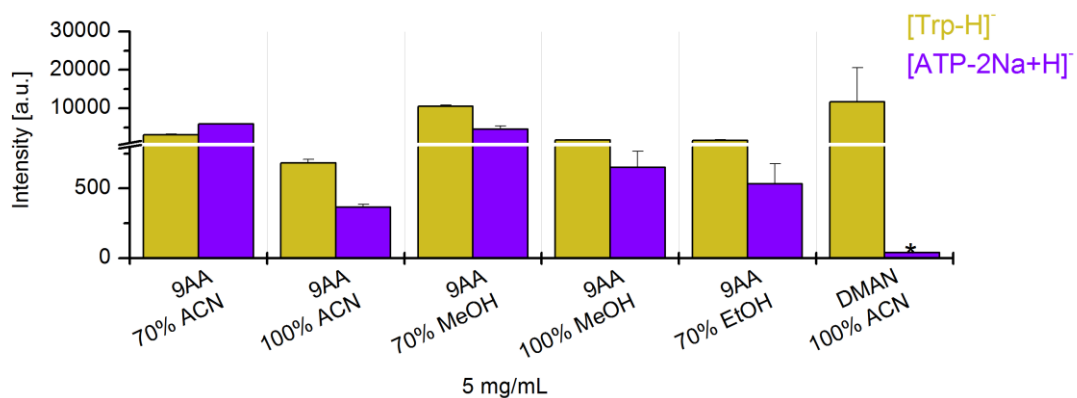
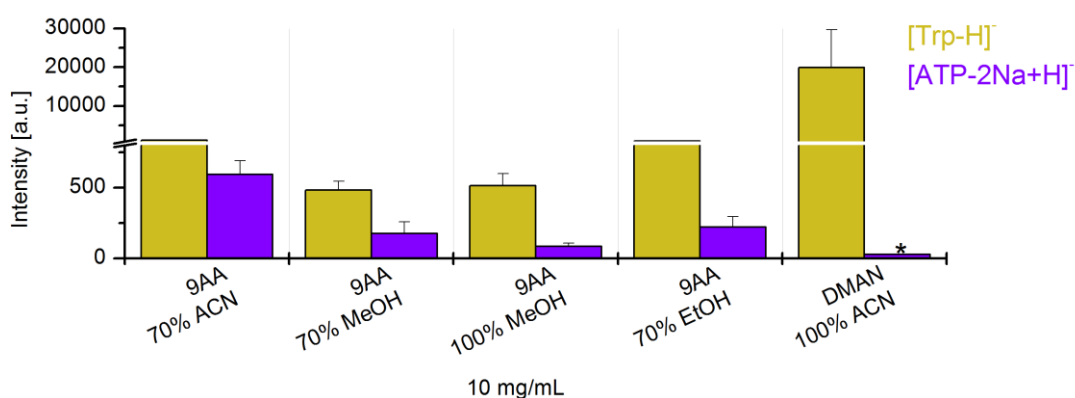


Figure 12: A comparison of mass resolution of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

A.



B.



C.

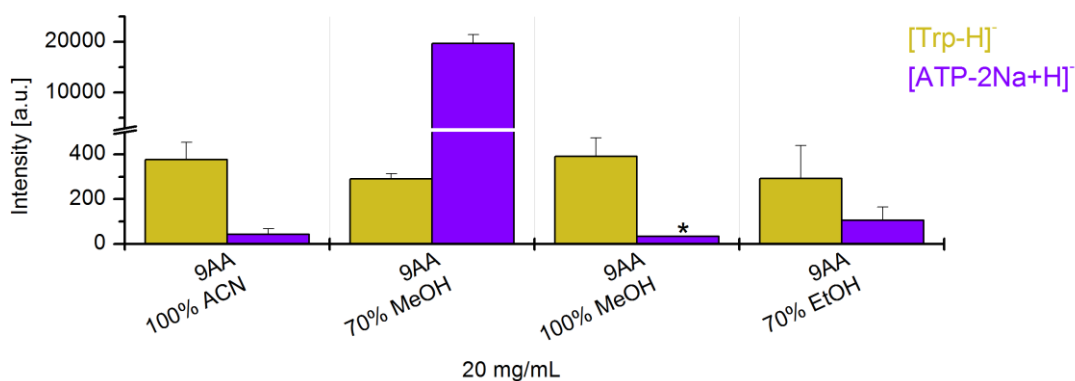


Figure 13: A comparison of the intensity of Trp and ATP ions based on matrix concentration: 5 mg/mL (A), 10 mg/mL (B), 20 mg/mL (C). The error bars represent the standard deviation calculated from the data obtained from triplicate experiments ($n=3$); MALDI-; * - one repeat only.

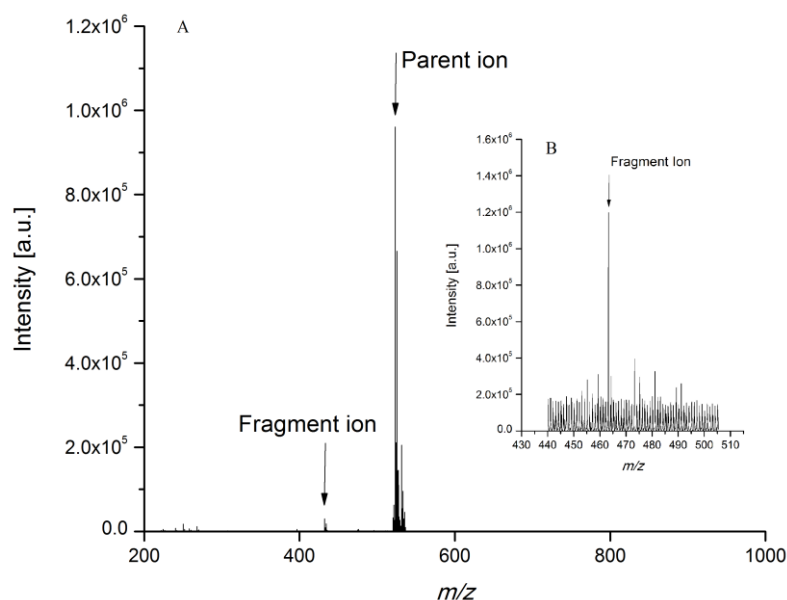


Figure 14: MALDI negative ion mode MS/MS spectrum of PC (18:0/0:0) with the parent ion at m/z 523 and the fragment ion at m/z 464; close-up of fragment ion $[M-N(CH_3)_3-H]^-$ (B).

Table 1: Combinations of the analytes contained within the calibration mixture in MALDI positive ion mode. S/N of each compound in each mixture is given; x – no detection (S/N <6).

Compound	Ion	Mixture		
		A	B	C
		S/N		
2-methylimidazole	$[M+H]^+$	x	85	402
Caffeine	$[M+H]^+$	8	547	388
Verapamil hydrochloride	$[M-HCl+H]^+$	606	3639	2877
	$[M-HCl-150+H]^+$	119	741	164
Adenosine 5'-triphosphate disodium salt hydrate	$[M+Na]^+$	x	10	39
Erythromycin A	$[M-H_2O+Na]^+$	84	342	1018
Bradykinin fragment 2-9	$[M+H]^+$	65	848	1018
	$[M+Na]^+$	24	39	495

Table 2: A list of the standards examined in MALDI positive and negative ionisation modes. Compounds are sorted in increasing molecular weight order. S/N ratio values for each compound and each repeat ($n = 2$) are included; x – no detection ($S/N < 6$).

Compound	Positive ions	S/N repeat 1	S/N repeat 2	Mean	Negative ions	S/N repeat 1	S/N repeat 2	Mean
Glycine	$[M+H]^+$	68	153	111	$[M-H]^-$	65	23	44
2-methylimidazole	$[M+H]^+$	3322	4068	3695	$[M-H]^-$	x	x	
Proline	$[M+H]^+$	124	60	92	$[M-H]^-$	x	x	
Nicotinic acid	$[M+H]^+$	84	20	52	$[M-H]^-$	10496	1155	5826
Histidine	$[M+H]^+$	450	238	344	$[M-H]^-$	54	236	145
Pyridoxine (vitamin B6)	$[M+H]^+$	174	98	136	$[M-H]^-$	17	34	26
Arginine	$[M+H]^+$	1013	1796	1405	$[M-H]^-$	x	x	
Citric acid	$[M+H]^+$	x	x		$[M-H]^-$	986	710	848
Caffeine	$[M+H]^+$	1927	3692	2810	$[M-H]^-$	38	27	33
Tryptophan	$[M+H]^+$	72	x	72	$[M-H]^-$	1142	759	951
Cytidine	$[M+H]^+$	33	549	291	$[M-H]^-$	182	16	99
Palmitic acid	$[M+H]^+$	x	x		$[M-H]^-$	10	41	26
Atenolol	$[M+H]^+$	1408	1540	1474	$[M-H]^-$	x	x	
	$[M+Na]^+$	1108	476	792				
	$[M+K]^+$	389	173	281				
Glucose-6-phosphate sodium salt	$[M+K]^+$	24	22	23	$[M-H]^-$	x	x	

Palmityl palmitate	$[M+H]^+$	x	x		$[M-H]^-$	86	x	86
Verapamil hydrochloride	$[M+H]^+$	92	250	171	$[M+Cl]^-$	257	54	156
	$[M-HCl+H]^+$	14954	7661	11308				
	$[M-HCl-150+H]^+$	2839	3689	3264				
PC (18:0/0:0) 1-stearoyl- <i>sn</i> -glycero-3-phosphocholine	$[M+H]^+$	71	72	72	$[M-H]^-$	146	14	82
					$[M-N(CH_3)_3-H]^-$	14	15	15
Adenosine 5'-triphosphate disodium salt hydrate	$[M+H]^+$	287	352	320	$[M-2Na+H]^-$	185	173	179
	$[M+Na]^+$	118	259	189				
	$[M-2Na+3H]^+$	2502	1513	2008				
Leucine enkephalin acetate salt hydrate	$[M+Na]^+$	23	26	25	$[M-H]^-$	10	x	10
Polyethylene glycol 600 (PEG 600) sulfate					$[SO_4]^{2-}$	2086	4435	3261
					$[H(OCH_2CH_2)_7SO_4]^-$	119	292	206
					$[H(OCH_2CH_2)_9SO_4]^-$	272	841	557
					$[H(OCH_2CH_2)_{11}SO_4]^-$	353	1185	769
					$[H(OCH_2CH_2)_{13}SO_4]^-$	266	832	549
Etoposide	$[M+K]^+$	169	147	158	$[M-206-H]^-$	11	9	10
	$[M+Na]^+$	2167	1589	1878				
L-Glutathione oxidized (GSSG)	$[M+H]^+$	640	64	352	$[M-H]^-$	x	x	

	$[M+Na]^+$	995	153	574				
Erythromycin A	$[M-H_2O+Na]^+$	5241	393	2817	$[M-H]^-$	7	8	8
	$[M+Na]^+$	331	153	242	$[M+Cl]^-$	25	x	25
Dipalmitoylphosphatidylcholine (DPPC)	$[M+K]^+$	x	7	7	$[M-H]^-$	x	x	
	PC headgroup	378	468	423				
Tripalmitin	$[M+Na]^+$	341	333	337	$[M-H]^-$	x	x	
Bradykinin fragment 2-9	$[M+H]^+$	7547	6906	7227	$[M-H]^-$	x	x	
	$[M+K]^+$	59	92	76				
	$[M+Na]^+$	1060	1120	1090				

Medium samples dilution for MALDI-MS analysis

Experiments were performed including fresh (non – used) medium sample and medium samples collected on day 3 from induced and non-induced cell cultures. Subsequently, all samples were diluted 2x, 6x and 20x in water according to the scheme shown in **Table 3**. 10 mg/mL DHB matrix solution in ACN/water 1:1 (v/v) with 0.1% TFA was freshly prepared and mixed (M:A 1:1 v/v) with analytes prior spotting on a MALDI plate. Each sample was run twice (n=2, independent repeats) using MALDI Synapt G2-Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK) operated in positive ion mode, a mass range of 20-1000 Da and laser power of 200 [units]. The average of two repeats was taken into account for mass resolution calculations (**Figure 15**). Instrument was calibrated to the calibration cocktail optimised previously. Spectra were established in MATLAB ver. R2012 (**Figure 16**).

Table 3: Scheme of medium samples dilution for MALDI-MS analysis. Table presents each dilution with volumes used for sample preparation.

Dilution	Sample [μ L]	Water [μ L]	Final dilution (on a plate)
2x	1	1	4x
6x	1	5	12x
20x	1	19	40x

Analysis of spectra shown in **Figure 16** demonstrates that dilution factor 2 is not enough for acquiring good quality mass spectra. For mass resolution (FWHM) calculations, dilutions by factor 6 and 20 were taken into consideration only. The mass resolution of two mass ranges (at m/z 104 and at m/z 456) is calculated, showing that there is no great difference between dilution factors of 6 and 20 (**Figure 15**). Therefore, further MALDI-MS footprint analysis is performed with dilution factor of either 6 or 20, depending on ionisation mode and laser power used.

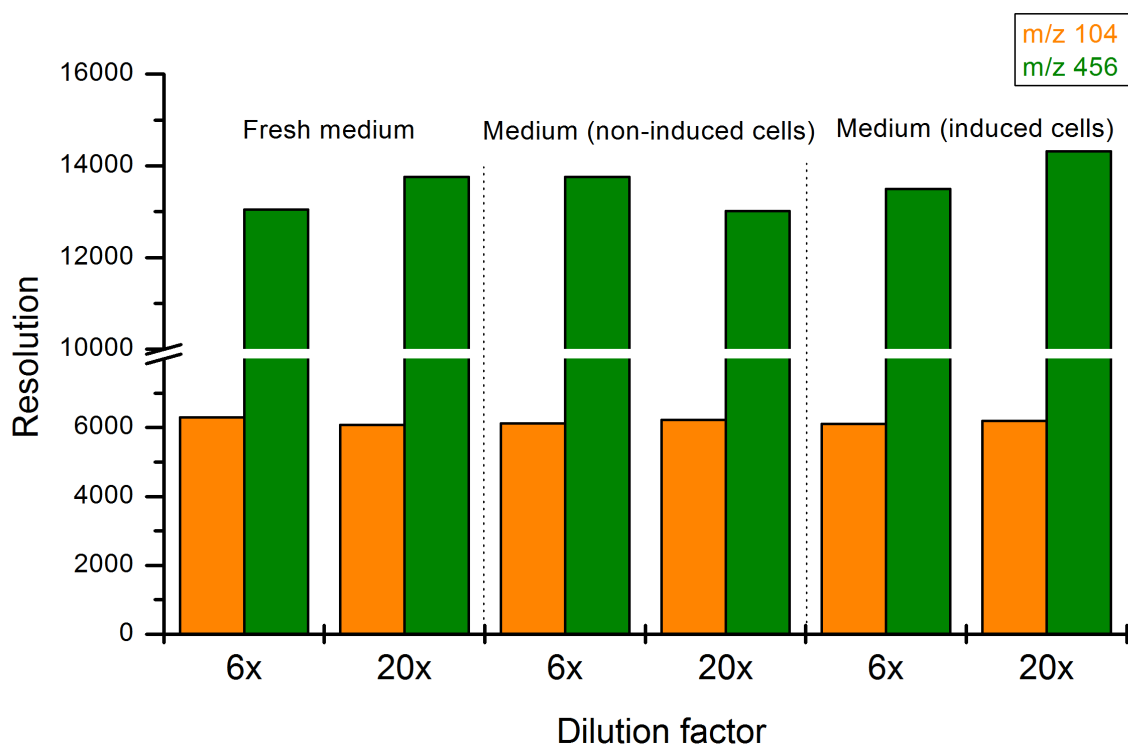


Figure 15: A comparison of mass resolution of ions at m/z 104 and m/z 456 for fresh medium and medium samples removed from cell culture across 6 and 20 dilution factors. Data obtained from two experiments ($n = 2$); MALDI positive ion mode.

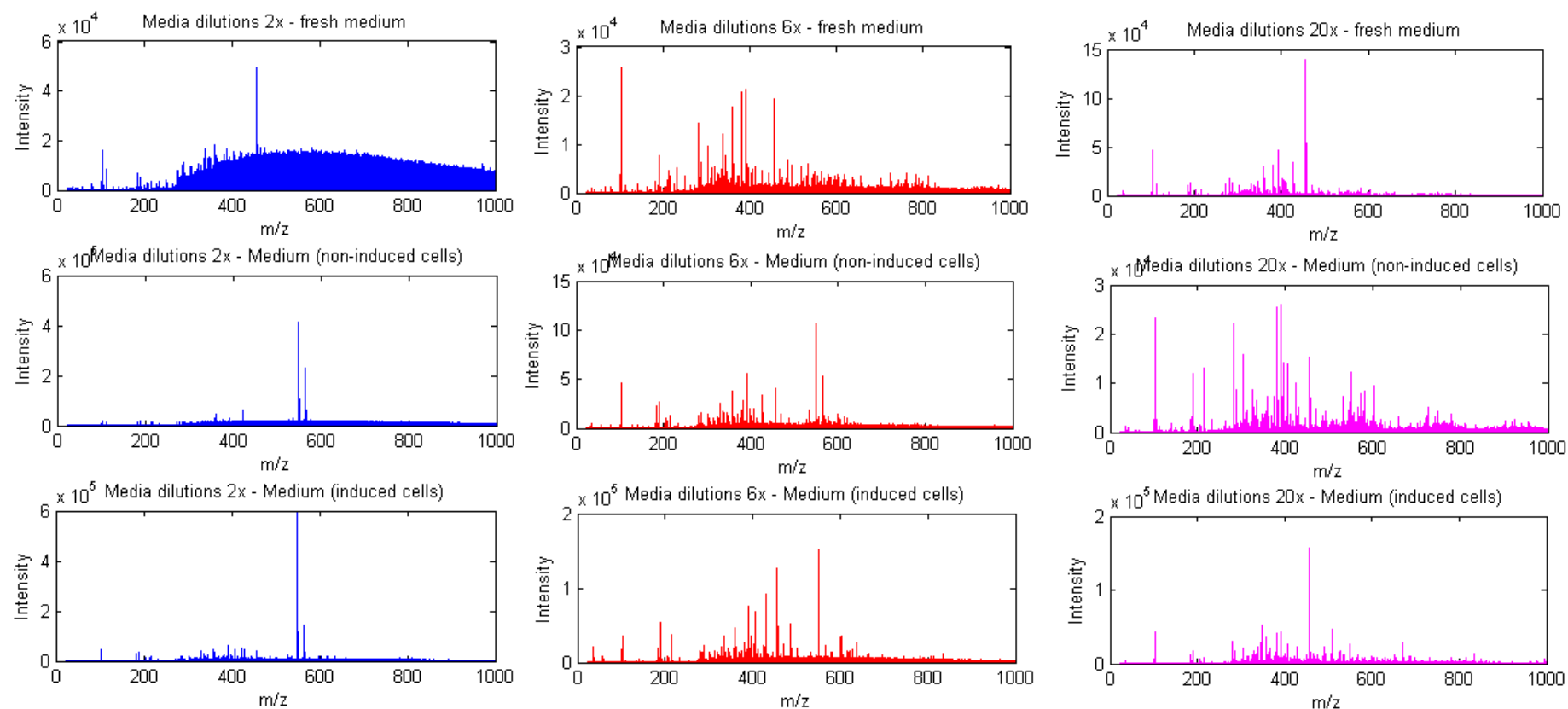


Figure 16: MALDI positive ion mode averaged spectra for fresh medium and medium samples removed from cell culture in three dilutions: 2x (blue), 6x (red) and 20x(pink); mass range 20-1000 Da.

CHAPTER 7: Metabolic imaging of CHO cells

Determination of cell density for MALDI and SIMS experiments

In order to optimise a cell density for imaging cell imaging experiments, CHO cells were grown directly on microscopic glass slides in growth medium in four different densities: at 0.02×10^6 cells/mL, 0.04×10^6 cells/mL, 0.07×10^6 cells/mL and 0.14×10^6 cells/mL for 3 days. Images were taken after 24, 48 and 72 hours of incubation time (**Figure 17**).

Images and observations showed the cell density of 0.02×10^6 cells/mL and 0.04×10^6 cells/mL were not suitable due to a cell loss during the culture growth, changes in cell morphology and the density of cells seemed to be too low after 72 hours of incubation. Additionally, the density of 0.14×10^6 cells/mL turned out to be too high and therefore, was excluded from further investigations. The most suitable cell density was obtained at 0.07×10^6 cells/mL with equal cell growth, healthy cell morphology and presence of single cells within a cell population. Consequently, the density of 0.07×10^6 cells/mL was chosen as the best for MS imaging investigations and was used for further MALDI-MS and SIMS experiments.

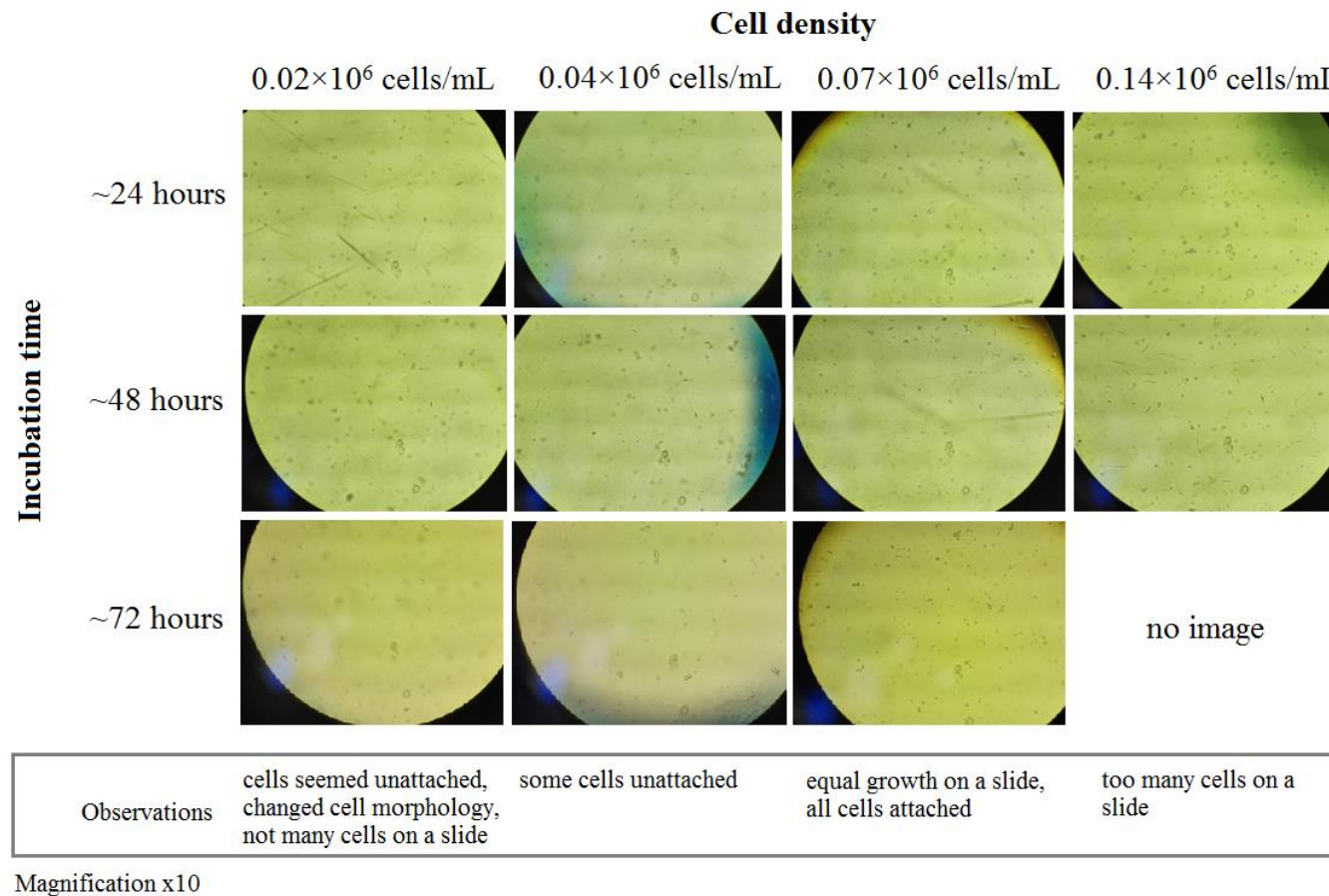


Figure 17: A comparison of cell growth at different cell densities during 72 hours of incubation time. Densities at 0.02×10^6 cells/mL, 0.04×10^6 cells/mL and 0.14×10^6 cells/mL were not suitable due to a cell loss, changes in cell morphology, too low or too high cell density. Density at 0.07×10^6 cells/mL was therefore selected as the best

Determination of suitable fixation protocol for CHO cells imaging

Experiments were performed on CHO cells suspended in growth medium (0.1×10^6 cells/mL), followed by fixation and washing steps shown in **Table 4**. In order to investigate the most suitable fixation protocol for CHO cells, the metabolite leakage from cells to outer environment during fixation procedure was analysed. Samples from different stages of fixation processes were taken and compared to pure (non-used) fixation solutions. The analysis was performed using MALDI Synapt G2-Si HDMS mass spectrometer (Waters Corporation/Micromass, Manchester, UK) with 355 nm UV laser and 1000 Hz repetition rate, in a mass range of 20 – 1000 Da in both positive and negative ionisation modes with 10 mg/mL of DHB in ACN/water (v/v) with 0.1% TFA (M:A 1:1) and 5 mg/mL of 9AA in 70% EtOH (M:A 1:10) as matrices respectively. The comparison was made based on PCA analysis of all samples, regarding to applied fixation method in both positive and negative ionisation modes (**Figure 18 and 19**). The comparison of PCA scores plots in positive ionisation mode shows that samples taken during different fixation stages differ from pure (non-used) fixation solutions across whole investigated mass range *i.e.* 20-1000 Da (loadings plots not shown). This suggests that in all examined fixation methods some metabolite leakage was present and no outstanding fixation method was found. However, a comparison of PCA scores plots from negative ionisation mode between applied fixation methods shows, the majority of these methods do not significantly differ across whole investigated mass range (loadings plots not shown). Considering no crucial differences between screened methods were found as well as taking into account both MALDI-MSI and SIMS imaging techniques and their most common sample preparation protocols for imaging experiments (Wang *et al.* 2016, Black *et al.* 2007, Lee *et al.* 2006), 150 mM ammonium formate was chosen as a fixation agent. The preference of this fixation agent also included considerations of the sample chemistry. Also, trying to avoid introducing new compounds into the cell chemistry, the preference of this fixation agent included these considerations too. Additionally, the optimisation of sample preservation with liquid nitrogen and desiccator was performed. CHO cells were grown on glass slides in growth medium (0.07×10^6 cells/mL) for 3 days and snap-frozen in liquid nitrogen or left to dry in desiccator (~2 hours). Images in **Figure 20** show the morphology of cells was not maintained when liquid nitrogen was used. Also, some cells were unattached and swollen. Therefore, better procedure was to leave cells in desiccators until dry. The chosen fixation agent with the addition of pre-washing step with PBS was applied for all MALDI-MS and SIMS imaging experiments within this project.

Table 4: The most common fixation protocols used for preservation of mammalian cells for MS imaging. Each protocol is specified with all solutions used during the fixation experiment. Sampling points for investigation of metabolite leakage are also shown ('Sample').

Method	1	2	3	4	5	6	7
References	(Lee <i>et al.</i> 2006, Anderson <i>et al.</i> 2013, Schaepe <i>et al.</i> 2015)	(Anderson <i>et al.</i> 2013, Schaepe <i>et al.</i> 2015)	(Schober <i>et al.</i> 2012, Anderson <i>et al.</i> 2013, Schaepe <i>et al.</i> 2015)	(Anderson <i>et al.</i> 2013, Ong <i>et al.</i> 2015)	(Ellis <i>et al.</i> 2016, Hanrieder <i>et al.</i> 2011)	(Fletcher <i>et al.</i> 2013, Kotze <i>et al.</i> 2013)	(Brown <i>et al.</i> , 2012)
Step I	PBS	PBS	PBS	Glycerol/ Ammonium acetate	70% EtOH	Ammonium Formate	100% MeOH
Sample				4a	5a		7a
Step II	PFA	Fixative	Glutaraldehyde	Ammonium acetate	95% EtOH	Ammonium Formate	
Sample	1a	2a	3a				
Step III	Ammonium acetate	Ammonium acetate	Ammonium acetate	Ammonium acetate	95% EtOH	Ammonium Formate	
Sample					5b	6a	
Step IV	Ammonium acetate	Ammonium acetate	Ammonium acetate	Ammonium acetate			
Sample				4b			
Step V	Ammonium acetate	Ammonium acetate	Ammonium acetate				
Sample	1b	2b	3b				

4% PFA; 150 mM ammonium acetate; 3% fixative: 2% PFA+1% glutaraldehyde in 0.1M NPP; 2.5% glutaraldehyde in 0.1M NPP; 40% glycerol in 150 mM ammonium acetate; 150 mM ammonium formate; NPP-sodium phosphate buffer pH 7.2; cells spun 5 minutes, 500 g after each step.

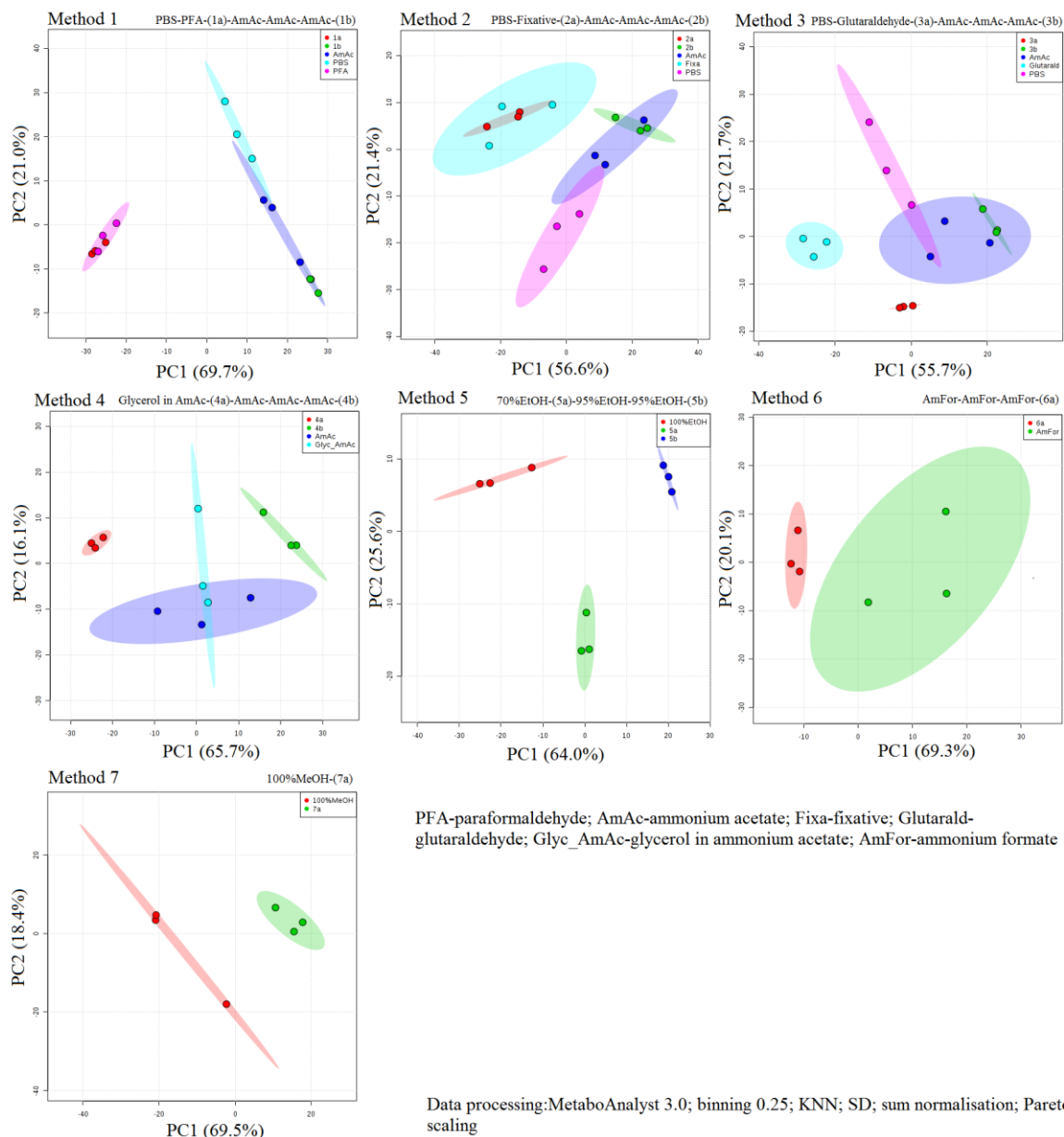


Figure 18: PCA scores of a comparison of metabolite leakage from CHO cells during different fixation processes. Samples from different stages of fixation processes were compared to pure (non-used) fixation solutions and analysed in MALDI positive ionisation mode. Three analytical replicates of each sample are shown.

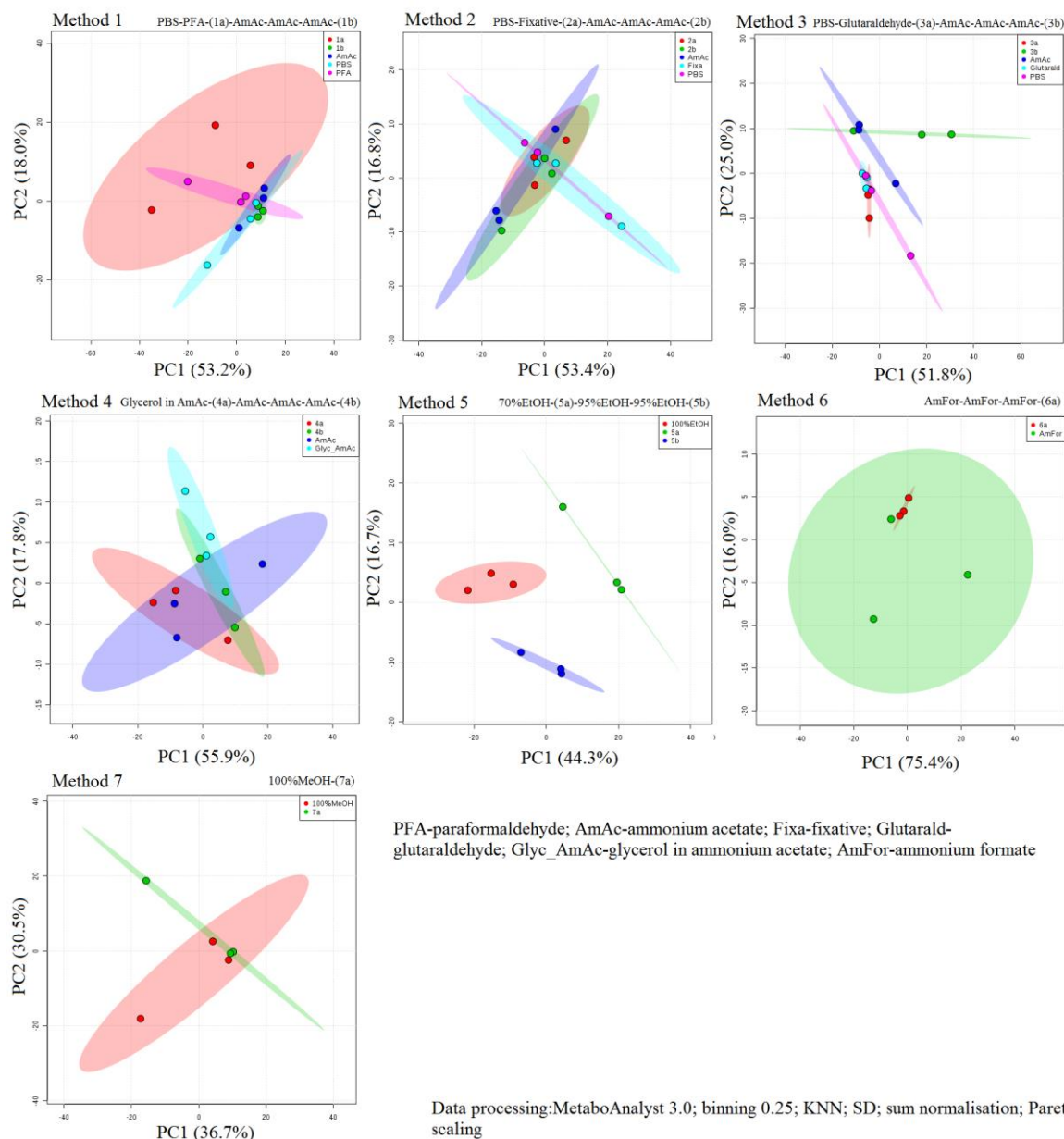


Figure 19: PCA scores of a comparison of metabolite leakage from CHO cells during different fixation processes. Samples from different stages of fixation processes were compared to pure (non-used) fixation solutions and analysed in MALDI negative ionisation mode. Three analytical replicates of each sample are shown.

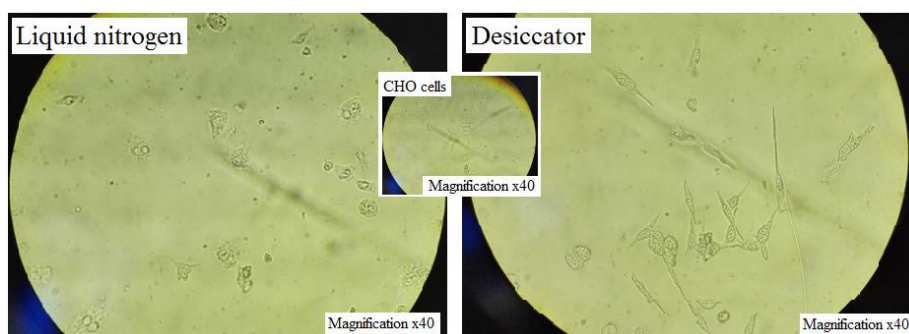


Figure 20: A comparison of morphology of CHO cells after snap-freezing in liquid nitrogen and drying in desiccator. Swollen cells with different morphology to normal growing cells (image in a middle, ‘CHO cells’) were seen after liquid nitrogen treatment.

MSI of CHO cells under inducer exposure and EPO-Fc protein production

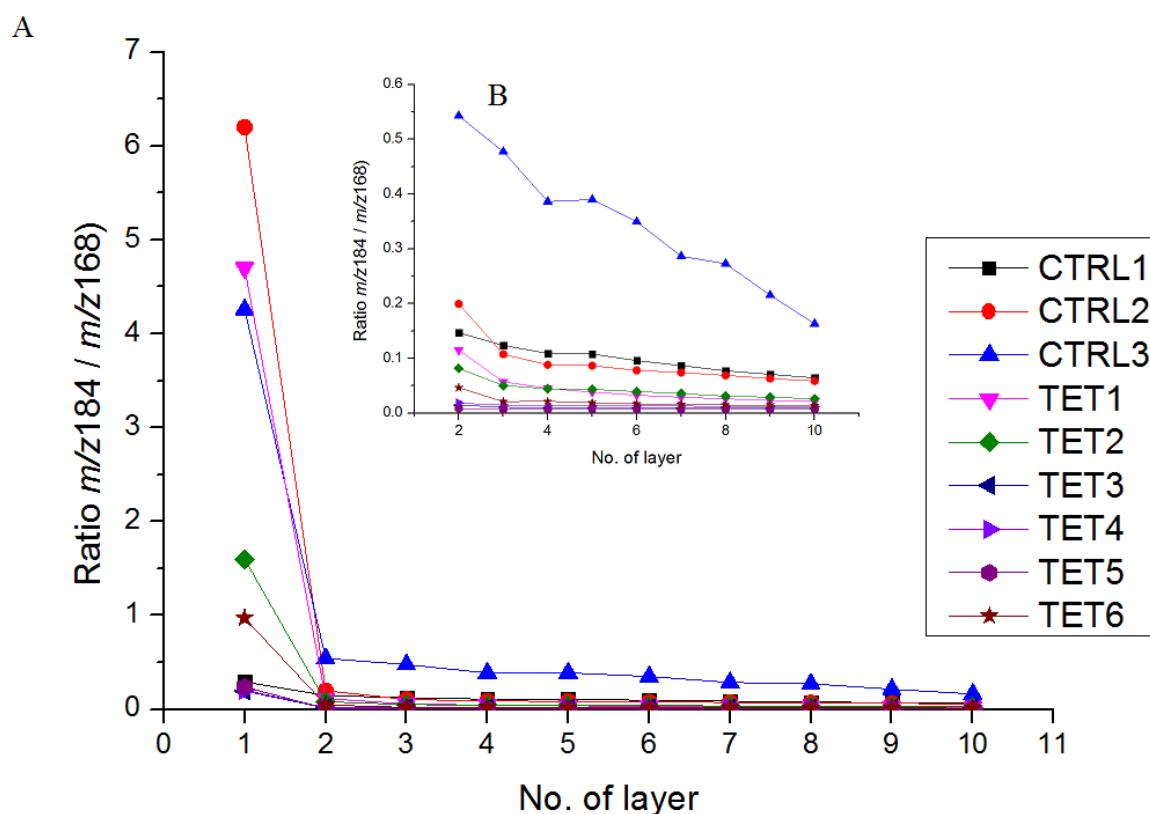


Figure 21: The analysis of the ratio of m/z 184 ion (organic origin) and m/z 168 (silicon) across all layers detected by SIMS. Content of m/z 184 differs for the 1st layer and is similar within layers 2-10. Different colours were used to differentiate samples.

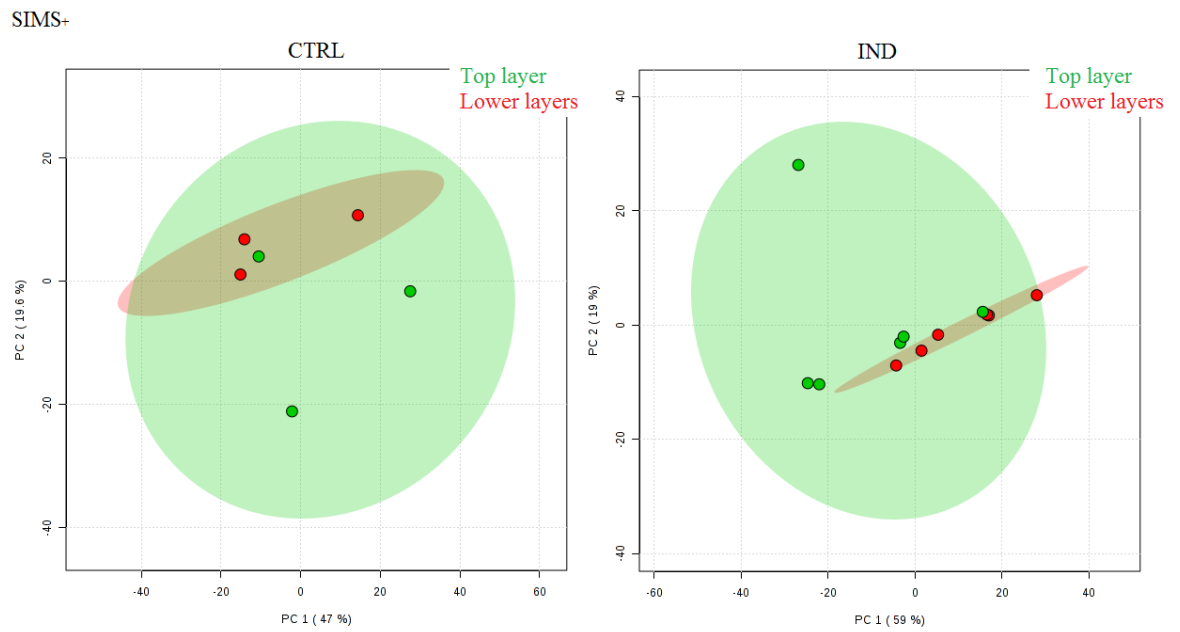


Figure 22: PCA scores plots generated for the full SIMS data set for the top (1st) and lower (2-10th) layers and for both control (CTRL) and induced cell (IND) data sets, illustrating no separation between these two groups.

References:

- Anderson, D. M. G., Mills, D., Spraggins, J., Lambert, W. S., Calkins, D. J. and Schey, K. L. (2013) 'High-resolution matrix-assisted laser desorption ionization-imaging mass spectrometry of lipids in rodent optic nerve tissue', *Mol Vision*, 19, pp. 581–592.
- Black, P., Agar, N. Y. R., Yang, H. W., Carroll, R. S., Black, P. M. and Agar, J. N. (2007) 'Matrix Solution Fixation : Histology-Compatible Tissue Preparation for MALDI Mass Spectrometry Imaging', *Anal Chem*, 79, pp. 7416–7423.
- Brown, M. V, McDunn, J. E., Gunst, P. R., Smith, E. M., Milburn, M. V, Troyer, D. a and Lawton, K. a (2012) 'Cancer detection and biopsy classification using concurrent histopathological and metabolomic analysis of core biopsies', *Genome Med*, 4, p. 33.
- Ellis, S. R., Cappell, J., Potočník, N. O., Balluff, B., Hamaide, J., Van der Linden, A. and Heeren, R. M. A. (2016) 'More from less: high-throughput dual polarity lipid imaging of biological tissues', *The Analyst*, 141, pp. 3832–3841.
- Fletcher, J. S., Rabbani, S., Barber, A. M., Lockyer, N. P. and Vickerman, J. C. (2013) 'Comparison of C60 and GCIB primary ion beams for the analysis of cancer cells and tumour sections', *Surf Interface Anal*, 45, pp. 273–276.
- Hanrieder, J., Wicher, G., Bergquist, J., Andersson, M. and Fex-Svenningsen, A. (2011) 'MALDI mass spectrometry based molecular phenotyping of CNS glial cells for prediction in mammalian brain tissue', *Anal Bioanal Chem*, 401, pp. 135–147.
- Kotze, H. L., Armitage, E. G., Fletcher, J. S., Henderson, A., Williams, K. J., Lockyer, N. P. and Vickerman, J. C. (2013) 'ToF-SIMS as a tool for metabolic profiling small biomolecules in cancer systems', *Surf Interf Anal*, 45, pp. 277–281.
- Lee, T. G., Shon, H. K., Lee, K. B., Kim, J., Choi, I. S. and Moon, D. W. (2006) 'Time-of-flight secondary ion mass spectrometry chemical imaging analysis of micropatterns of streptavidin and cells without labeling', *J Vac Sci Technol A*, 24, pp. 1203–1207.
- Ong, T.-H., Kissick, D. J., Jansson, E. T., Comi, T., Romanova, E. V., Rubakhin, S. S. and Sweedler, J. V. (2015) 'Classification of Large Cellular Populations and Discovery of Rare Cells Using Single Cell Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry', *Anal Chem*, 87, pp. 7036–7042.
- Schaepe, K., Kokesch-Himmelreich, J., Rohnke, M., Wagner, A.-S., Schaaf, T., Wenisch, S. and Janek, J. (2015) 'Assessment of different sample preparation routes for mass spectrometric monitoring and imaging of lipids in bone cells via ToF-SIMS', *Biointerphases*, 10, pp. 16–25.
- Schober, Y., Guenther, S., Spengler, B. and Römpf, A. (2012) 'Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging', *Anal Chem*, 84, pp. 6293–7.
- Wang, S., Chen, X., Luan, H., Gao, D., Lin, S., Cai, Z., Liu, J., Liu, H. and Jiang, Y. (2016) 'Matrix-assisted laser desorption/ionization mass spectrometry imaging of cell cultures for the lipidomic analysis of potential lipid markers in human breast cancer invasion', *Rapid Comm Mass Spectrom*, 30, pp. 533–542.