

Development and application of multiple isotopes-assisted untargeted metabolomics in human health and disease

Sergey Evseev MPharm (Pharmacy)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy in Pharmacy

September 2018

Acknowledgements

First of all, I would like to take this opportunity to express my sincere gratitude to my PhD supervisors, Dr Dong-Hyun Kim and Professor David Barrett, for their constant guidance, encouragement and valuable suggestions during my PhD journey from laboratory work up until my thesis writing. I am very grateful to both of them as this thesis would not have been possible without their help, support and patience throughout the study period.

I would also like to thank Dr Ruman Rahman, at the Faculty of Medicine and Health Sciences, Queen's Medical Centre, University of Nottingham, for giving me an opportunity to work with human brain tissues for low grade glioma tumour study, as well as for his kind support and guidance on data interpretation. I am particularly thankful to his PhD students, James Wood and Mohammed Diksin, for helping me with extraction and sample preparation of brain tissues.

I would also like to thank Dr Kostas Tsintzas and Dr Marie-Christine Pardon for providing me with clinical samples for my validation and application studies, without which I would not complete my thesis.

A massive thank you goes to Ferran Casbas, a fellow PhD colleague from my department, who has helped me a lot with a very complicated data processing stage by designing a Python computer code that saved me a lot of valuable time (and nerves). Without him and his valuable code, it would have been very difficult to achieve the results that I got!

I wish to express my gratitude to all the colleagues in my department for their support and guidance throughout my PhD studies, particularly to Dr Salah Abdelrazig and Dr Petros Pousinis for their teaching of various statistical software programmes, Dr Catherine Ortori for her guidance on how to use Exactive. I am thankful to Dr Nigel Halliday for his early help with culturing *E. coli* bacteria.

And finally I would like to express my love to my parents and all my family, who were always there for me and provided me with their support, care, encouragement and, most importantly, love. I am very grateful to all of you in believing in me to succeed and finish my PhD studies.

List of Abbreviations

1. APCI – Atmospheric Pressure Chemical Ionisation
2. ANOVA – Analysis Of Variance
3. ASCA – ANOVA Simultaneous Component Analysis
4. AMW – Acetonitrile : Methanol : Water
5. ADP – Adenosine Diphosphate
6. AMP – Adenosine Monophosphate
7. ATP – Adenosine Triphosphate
8. AGEs – Advanced Glycation End-products
9. BMI – Body Mass Index
10. ^{13}C – Carbon 13 isotope
11. CNS – Central Nervous System
12. CT scan – Computed Tomography scan
13. CV – Coefficient of Variance
14. DNA – Deoxyribonucleic Acid
15. *E. coli* – Escherichia Coli
16. EMDB – E. coli Metabolome Database
17. ESI – Electrospray Ionisation
18. ESI-MS – Electrospray Ionisation Mass Spectrometry
19. FDA – Food Drug Administation
20. FDR – False Discovery Rate
21. FT-IR – Fourier-Transform Infrared Spectroscopy
22. GBM – Glioblastoma Multiforme
23. GC – Gas Chromatography
24. GDP – Guanosine Diphosphate
25. GMP – Guanosine Monophosphate
26. HbA1c – glycated haemoglobin
27. HDL – High Density Lipoprotein
28. HILIC – Hydrophilic Interaction Liquid Chromatography
29. HMDB – Human Metabolome Database
30. HPLC – High Performance Liquid Chromatography
31. ITH – Intra-Tumour Heterogeneity
32. LB – Lysogeny Broth
33. LC-MS – Liquid Chromatography Mass Spectrometry
34. LDL – Low Density Lipoprotein
35. LGGs – Low Grade Gliomas
36. MAPK – Mitogen Activated Protein Kinase
37. MGMT – Methyl Guanine Methyltransferase
38. MHC – Major Histocompatibility Complex
39. MRI – Magnetic Resonance Imaging
40. MS – Mass Spectrometry

41. MS/MS - Tandem Mass Spectrometry
42. MSI – Metabolomics Standards Initiative
43. MSTUS – Mass Spectrum Total Useful Signals
44. NICE – National Institute for Health and Care Excellence
45. NMR – Nuclear Magnetic Resonance
46. OD – Optical Density
47. OPLS-DA – Orthogonal Partial Least Square Discriminant Analysis
48. PCA – Principal Component Analysis
49. PI3K – Phosphoinositide 3 Kinase
50. PTEN – Phosphatase and Tensin Homolog
51. QC – Quality Control
52. RNA – Ribonucleic Acid
53. ROS – Reactive Oxygen Species
54. RP – Reversed Phase
55. RSD – Relative Standard Deviation
56. SIL – Stable Isotope Labelling
57. TB – Terrific Broth
58. TCA – Tricyclic Acid cycle or Krebbs cycle
59. TIC – Total Ion Counts
60. TP53 – Tumour Protein 53
61. TSB – Trypticase Soy Broth
62. U-^{13}C labelled – Uniformly ^{13}C labelled
63. UV – Ultra Violet
64. VIP – Variable Influence on Projection
65. WHO – World Health Organisation

Abstract

Global metabolite profiling, also known as untargeted metabolomics, is constantly used in the qualitative and quantitative assessment of a wide range of metabolites in human metabolomics research, where scientists are able to monitor changes in metabolite concentrations in human biofluids and tissues affected by complex clinical diseases including rheumatoid arthritis, Alzheimer's, Parkinson's and variety of cancers, to name a few. These metabolite profiling studies are regularly applied in human clinical samples in order to detect potential biomarkers, assist in drug discovery, monitor disease onset and its progress and many other bioanalytical areas. The application of the LC-MS analytical tool has been most widely used in comprehensive metabolic studies, due to its high throughput, soft ionisation and good metabolite coverage, in comparison to other analytical instruments such as NMR, FT-IR or Raman.

LC-MS based metabolomics studies are currently facing challenges due to non-linear responses derived from matrix effects and biological variations within samples, which can result in biased quantitative analysis affecting true measurement of metabolite levels and their biological relevance. To overcome these issues, stable multiple ^{13}C isotopically labelled metabolites acting as internal standards can be employed in metabolomics studies to reduce metabolome data variation and improve its accuracy, also known as a normalisation technique. The main limitation of this normalisation technique in global metabolite profiling studies, however, is a lack of available ^{13}C labelled standards to cover a wide range of metabolites, as well as their high cost to produce. To solve this problem, uniformly (U) ^{13}C labelled bacterial organisms, such as *E. coli* or *Spirulina*, can be used instead, acting as a source of multiple labelled internal standards. Currently, in metabolomics research, there is a lack of established validated ^{13}C normalisation techniques that can be applied to a wide range of metabolites in mammalian samples using a global metabolite profiling method. In this thesis the proposed LC-MS method, involving a (U)-^{13}C labelled bacterial organism as a source of internal standards for normalisation, has been developed and applied to a range of untargeted clinical studies to demonstrate the effectiveness of normalisation and improvement in data accuracy to answer biological questions.

E. coli and *Spirulina* extracts were analysed using LC-MS-based metabolite profiling to select the appropriate source of internal standards. In *E. coli* samples, around 780 putative metabolites were detected with high peak signal response, compared to approximately 600 putative metabolites in the *Spirulina* bacterium with poor peak signal response. *E. coli* appeared to have more metabolites in common with human biofluid or tissue metabolomes than *Spirulina*, fully

confirming *E. coli* to be a suitable bacterial organism to use for internal standards. The chosen (U)-^{13}C labelled *E. coli* showed a large proportion of metabolites labelled with ^{13}C isotope (77%), with only 23% of the metabolome unlabelled.

To validate the proposed normalisation method, (U)-^{13}C labelled *E. coli* was applied in human urine, human brain tissue and mouse plasma samples. ^{13}C-labelled *E. coli* extract was added to the extraction solvent (methanol) and mixed with the samples of interest in a 1:1 ratio. In all three studies, percentage RSD of peak height intensities was calculated for detected metabolites, along with constructed PCA and OPLS-DA plots, to assess the efficiency of normalisation. In human urine and brain studies, approximately 70% of identified metabolites in each group had their percentage RSD reduced, while in the mouse plasma study the result was observed to be even better with 90% of metabolites successfully normalised. When compared to other normalisation techniques such as MSTUS, TIC and creatinine, the ^{13}C normalisation has shown better results with percentage RSD range being less variable. In all three validation studies, PCA and OPLS-DA score plots showed clearer separation between sample groups with their replicates clustered and high Q2 scores in normalised metabolomes, compared to non-normalised. Finally, the normalisation technique has helped to detect more statistically significant metabolites in all three studies, compared to non-normalised datasets. This method has shown to be fully validated, reducing metabolite variation significantly and improving the accuracy of data by detecting a high number of statistically significant metabolites.

A fully validated LC-MS method with proposed normalisation technique has been applied in two clinical investigation studies, to obtain a highly accurate metabolome from clinical samples and answer the main biological questions set up by the studies: mainly searching for potential biomarkers and the effect of a disease on metabolic pathways. One clinical study investigated the effect of a fatty meal (diabetes condition) on the human urine and plasma metabolome of healthy volunteers, while the other study performed metabolomics analysis on human low-grade glioma affected brain tissue. In both studies, normalised metabolome data have helped to detect a number of statistically significant metabolites which were observed to affect certain metabolic pathways involved in the investigated diseases, showing potential for being biomarkers.

Overall, the proposed normalisation technique using multiple ^{13}C labelled internal standards, with the assistance of a bacterial organism as their source and a powerful analytical LC-MS method, has achieved great results in reducing metabolite data variation and improving data

accuracy of a wide range of metabolites in mammalian biofluid and tissue samples, analysed by untargeted metabolomics, and has shown great promise in the search for potential biomarkers in future clinical untargeted studies.

Contents

Chapter 1. General Introduction....12

1.1 Untargeted metabolite profiling in human biofluids and cell based tissues....12

1.1.1 Global metabolite profiling....12

1.1.2 Human metabolomics and biomarker discovery....15

1.2 Technical aspects of metabolite analysis....17

1.2.1 Analytical instruments in metabolomics....17

1.2.2 LC-MS based untargeted metabolomics....19

1.3 Sample preparation and extraction aspects....22

1.3.1 Sample preparation in LC-MS metabolomics....22

1.3.2 Extraction of biofluid and tissue based samples....23

1.4 Data processing and current issues....25

1.4.1 Data pre-processing and statistical analysis....25

1.4.2 Issues with metabolite data....28

1.5 Normalisation of metabolomics data....30

1.5.1 Concept of normalisation....30

1.5.2 Pre- and post-processing data normalisation....31

1.5.3 Comparison of normalisation techniques....35

1.5.4 ^{13}C normalisation approach and current limitations....36

1.6 General aims and objectives of the thesis....40

Chapter 2. LC-MS method development of uniformly-^{13}C-labelled organism extract as a source of internal standards for normalisation....44

2.1 Introduction....44

2.2 Materials and Methods....52

2.2.1 Materials, chemicals and sample preparation....52

2.2.2 Analytical LC-MS methodology....56

2.2.3 Data processing and analysis....58

2.3 Results and Discussion....60

2.3.1 Optimisation of the metabolite extraction protocol for *E. coli*....60

2.3.2 Optimisation of the metabolite extraction protocol for *Spirulina*....64

2.3.3 Comparison of unlabelled *E. coli* and *Spirulina* metabolomes with human biofluid metabolites....66

2.3.4 LC-MS method validation....69

2.3.5 Generation of a ^{13}C-labelled source of internal standards and testing its quality....72

Chapter 3. Validation of the normalisation method using human and animal biological samples....76
3.1 Introduction....76
3.1.1 Concept of validation and its purpose....76
3.1.2 Planning validation experiments....78
3.1.3 Aims and objectives....79
3.2 Materials and methods....81
3.2.1 Materials, instruments and samples collection....81
3.2.2 Design of validation experiments....82
3.2.3 LC-MS analytical run of the samples....84
3.2.4 Data processing and multivariate analysis....85
3.3 Results and Discussion....87
3.3.1 Human Urine study....87
3.3.2 Human brain tissue study....94
3.3.3 Mouse plasma study....100
3.4 Conclusion....106
Chapter 4. Urine and plasma metabolomics analysis following an oral saturated fat meal intervention study in healthy volunteers....109
4.1 Introduction....109
4.1.1 Dietary intervention: its purpose and complications....109
4.1.2 Effect of dietary intervention on human metabolism in complicated diseases....110
4.1.3 Justification of the dietary intervention study....112
4.2 Materials and methods....113
4.2.1 Study design and sample collection....113
4.2.2 Materials, chemicals and LC-MS analysis....116
4.2.3 Data processing and statistical analysis....116
4.3 Results....118
4.3.1 Metabolomics analysis of urine and plasma samples....118
4.3.2 LC-MS analytical performance and normalisation in urine and plasma samples....123
4.3.3 Tentative identification of statistically significant metabolites in urine and plasma samples....131
4.4.1 Quality of data....137
4.4.2 Key metabolites and metabolic pathways....138
Chapter 5. Metabolic signatures of low-grade glioma in human brain tissues....143
5.1 Introduction....143

5.1.1 Glioma: definition and classification 143
5.1.2 Low grade glioma: epidemiology, aetiology and symptoms 144
5.1.3 Low grade glioma: current treatment strategies 147
5.1.4. Low grade glioma: molecular pathology, metabolic pathways and potential biomarkers 148
5.1.5 The importance of intra-tumour heterogeneity and central core 152
5.1.6 Aims and objectives 155
5.2 Materials and methods 156
5.2.1 Tissue sample collection and information 156
5.2.2 Sample preparation and extraction 158
5.2.3 LC-MS analysis 158
5.2.4 Data processing and statistical analysis 159
5.3 Results 160
5.3.1 Demographic data on chosen participants 160
5.3.2 Analytical performance and metabolomics data quality 160
5.3.3 Metabolomics analysis of LGG brain tissue 164
5.3.4 Tentative identification of key significant metabolites (potential biomarkers) in brain tissue samples affected by LGG 168
5.4 Discussion 177
5.5 Conclusion 180
Chapter 6. Conclusions and future work 183
6.1 Method development and validation 183
6.2 The clinical applications of the developed LC-MS normalisation method 186
6.3 Future Work 187
Chapter 7. References 189
Appendixes 218

Chapter One

General Introduction

1.1 Untargeted metabolite profiling in human biofluids and cell based tissues

1.1.1 Global metabolite profiling

Since the 1990s, metabolomics has become a new and rapidly emerging field of bioanalytical research through technological advancements across several scientific disciplines.[1,2] It involves the comprehensive analysis of unique chemical fingerprints, also known as metabolites, which are end products of gene expression of specific cellular processes in biological systems. Metabolites are small molecules which have a low molecular weight of ≤ 1800 Daltons and are usually classified according to their metabolic pathway, biochemical function, retention time and physicochemical properties including isomerism and polarity.[3,4] The analysis of metabolites helps to discover potential biomarkers involved in specific diseases[5], monitor drug treatments and therapeutic outcomes[6], monitor toxicological studies in humans and the environment[7,8], predict biochemical pathways[9], as well as providing the biological status of particular metabolites qualitatively and quantitatively in metabolic reactions.[10]

Metabolomics is part of a wider group of studies called "omics" that aims to describe the full complexity of functional cells, tissues, organs and organisms.[11,12] Omics includes other studies such as "lipidomics" (analysis of lipids), "genomics" (analysis of genome), "transcriptomics" (analysis of RNA molecules) and "proteomics" (analysis of proteins), which collectively gather the full information about biological molecules and their functions.[13] Figure 1.1 schematically shows the relationship between the different omics studies and their application. The understanding of a living organism at the molecular system level is still very complicated, even with the use of available advanced omics platforms. It is evident that comprehensive omics studies have an important central role and will help us to better understand biochemical mechanisms in complex organisms.[14] Nowadays, researchers focus on integrating data from several omics studies to acquire a comprehensive picture of the biological process, rather than relying on a single omics approach.[15] An example of a commonly used combination is an integration of genomics with proteomics to identify the functions of genes in the regulation of certain proteins in a living organism, in order to understand more about the physiology of certain organ or tissues.[16] A combination of metabolomics with lipidomics allows scientists to acquire compounds of a polar and non-polar nature.[17] Unlike other omics studies where the functions of lipids, genes, RNAs and proteins are subject to post-translational modifications, metabolomics provides very accurate information about the biochemistry of an organism, as

metabolites are found to be easier to correlate with the phenotype of a sample – analysed metabolites are end products of multiple enzymatic reactions, so serve as direct signatures of biochemical activity.[18,19]

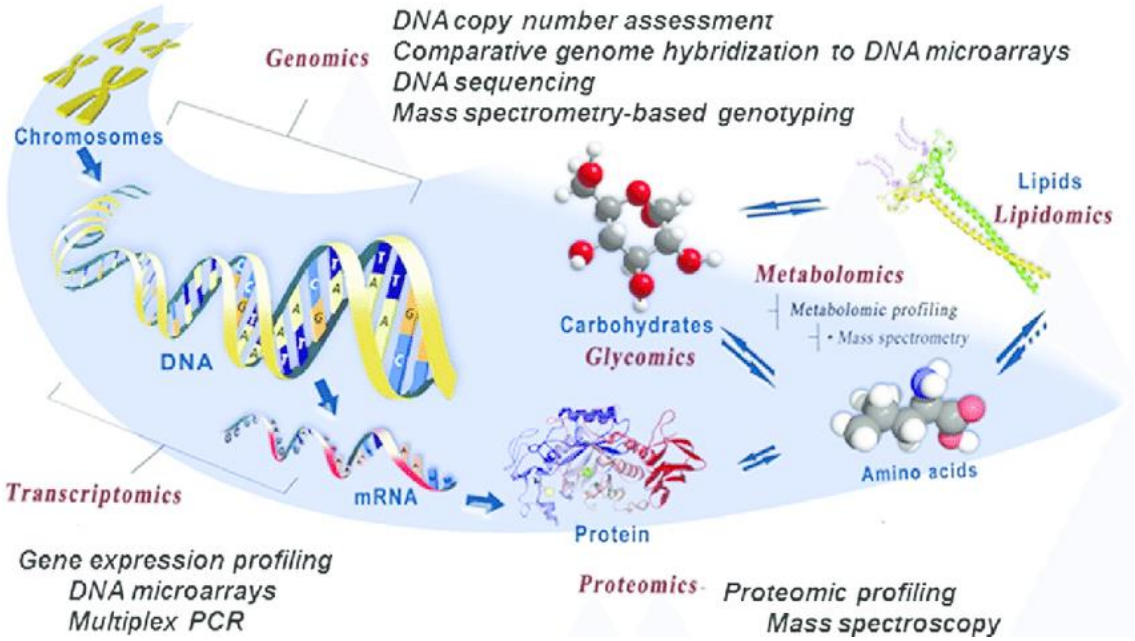


Figure 1.1 Illustration of "Omics" approaches used in identification and characterisation of variety of biological molecules. **Wu et al (2011), Molecular approaches to studying the soil biota.**

Metabolomics has a variety of applications, mainly in biomedical, pharmaceutical and environmental sciences.[20] Comprehensive analysis of metabolites from biological tissues and biofluid samples in living organisms enables researchers to monitor their concentration changes in different conditions, e.g. healthy vs disease states, understand cell physiology in more detail and apply the metabolome data (information on metabolites) in different contexts.[21,22] This is known as a global metabolite profiling.[23,24] It is widely used in the determination of potential biomarkers and diagnosis of a disease. For example, a recent study in metabolite profiling of prostate cancer cells revealed the presence of a very high concentration of lactate due to abnormal conversion of glucose, as well as the inhibition of the mitochondrial enzyme m-aconitase involved in citrate oxidation.[25] These findings can help researchers to further investigate the role of lactate and m-aconitase in the development of prostate cancer – as diagnostic tools and potential biomarkers for disease treatment. Another research study investigated the effect of an agrarian food diet on the gut microbiota using an analytical metabolomics technique – the results showed that diet influenced the metabolome of gut microbiota, therefore affecting the health of an individual.[26] Metabolomics can also be applied

in environmental sciences to study unusual biotic and abiotic perturbations on the environment and its surroundings.[27] For example, over the years metabolite profiling analysis has been used to study the effect of accidental spillage of crude oil on the aquatic environment – one study investigated the toxic effects of oil on salmon, revealing changes in the salmon's metabolome (rapid decrease in certain amino acids and organic osmolytes).[28] As well as monitoring human metabolite changes in disease states, metabolomics approaches also observe the pharmacokinetics and pharmacology of medications by monitoring human metabolites in response to administration of a pharmaceutical drug. This field of metabolomics is also referred to as "pharmacometabolomics".[29] The study of the effect of antihypertensive drugs (atenolol) on the biochemical profile of metabolites in plasma samples of Caucasian and African American individuals revealed a decrease in certain amino acids and fatty acids, helping pharmaceutical research to decide on the appropriate strengths of medications.[30]

There are usually two approaches used in any type of metabolomics analysis: targeted and untargeted techniques. When scientists are trying to find the identity of all measurable analytes (metabolites, lipids and unknown compounds) in a given biological system, they use an untargeted approach, such as global profiling, mentioned earlier.[31,32] An untargeted approach can often be applied in biomarker discovery for serious clinical conditions[33], environmental sciences in monitoring toxicity of microorganisms[34], in plant metabolomics to identify useful compounds for medicinal purposes[35], to name a few examples.[36] In a recent study this technique has been applied to investigate the metabolite content of pleural effusions taken from lung cancer sufferers. The technique was able to detect a large quantity of amino acids and lipids which were thought to be responsible for this disorder.[37] Once the identity of the metabolites is determined, further studies can be done by quantifying the selected metabolites of interest and finding their chemical properties – known as the targeted approach. This approach allows researchers to quantitatively analyse and validate the potential biomarkers to confirm their role in clinical disorders.[38] For example, in one research study this technique has been applied to analyse complex mixtures of cytokines from human plasma samples of rheumatoid arthritis sufferers, showing how the concentrations of those glycoproteins were affected according to the severity of the condition, to validate them as potential biomarkers.[39] Another study showed determination of the full metabolism of the HIV drug saquinavir in mice, allowing identification of the drug's potential pharmacological effect in the human body.[40] Together, these two methods can provide a full metabolic profile of the chosen biological organism with valid verification of metabolites.

1.1.2 Human metabolomics and biomarker discovery

As mentioned earlier, metabolomics research is largely involved in monitoring cellular concentration changes of metabolites between healthy (control) and disease (altered) states to determine potential biomarkers for the treatment of a variety of diseases including Alzheimer's[41], Parkinson's[42], rheumatoid arthritis[43] and various cancers[25]. An untargeted approach, such as global metabolic profiling has, in particular, been used in these studies, giving an opportunity to detect a large range of potential biomarkers, followed by targeted analysis for validation. Research has concentrated on metabolite profiling analysis of human body fluid samples, such as blood (plasma)[44], urine[45], saliva[46], cerebrospinal fluid[47] and sweat[48]. Human biofluids contain essential metabolites, which are involved in the important physiological processes and development of the human organism. Examples of such metabolites include nucleotides, sugars, organic acids, amino acids, peptides and other classes.[49] Human tissues such as brain, liver, heart, and kidney affected by diseases are also invaluable sources of biomarkers discovery.[50,51]

Over the years metabolomics studies, performed in several countries, have found a large number of significant and important metabolites identified in human biological fluids and tissues, as part of the "Human Metabolome Project".[52] Therefore, there has been an increasing demand from the Metabolomics society to create the reference databases of clinically important human biofluid and tissue metabolomes, in order to assist with comprehensive metabolite profiling experiments and improve the accuracy of metabolite identification. The "Human Metabolome Database" (HMDB) was created using information gathered from computer-aided literature resources (books, journal articles and electronic databases), containing detailed information about the structure, physicochemical properties and presence of human metabolites in certain biofluids and tissues.[53,54] The database also provides information on the clinical aspects of metabolites, metabolic pathways that each metabolite is involved in and, in some instances, experimental data in the form of mass spectra to aid scientists with analytical queries.[55]

Potential biomarkers are usually identified by running an accurate, untargeted analysis of biofluid and tissue samples, processing analytical data through specialised software helping to identify and validate significantly changed metabolites (more details about statistical analysis is described in chapter three).[56] A recent study in chronic kidney disease has successfully detected an increase in uremic toxins in the plasma of human individuals suffering from renal impairment. These toxins were thought to play a significant role in the disease and could therefore act as

potential biomarkers.[57] Parkinson's disease has also been investigated by the metabolomics approach. Several studies were performed using metabolic profiling of urine, blood and saliva samples taken from disease sufferers and successfully determined the status of protein and small molecule biomarkers involved in the development of the condition.[58] An untargeted metabolite profiling study of urine samples from patients with type 2 diabetes also revealed the presence of endogenous metabolites such as tryptophan, kynurenic acid and kynurenine, which were identified as potential biomarkers related to the condition.[59] Other studies showed that in disease cases such as cystic fibrosis, body sweat could be an appropriate future biofluid for metabolite profiling analysis as a diagnostic tool – sweat produces a large amount of chloride ions, which can be easily detected and can confirm the presence of cystic fibrosis.[60]

Human urine is a particularly useful biofluid, as it can be collected in a completely non-invasive manner over long periods of time from the same test subjects.[61] There have been numerous metabolite profiling experiments performed with urine samples for clinical and physiological studies.[10,62,63] One study used the human urine metabolome in diagnosing a bladder cancer. During the analysis of urine samples taken from cancer patients, there were three upregulated significant metabolites (nicotinuric acid, trehalose and asparagine) detected along with downregulated small molecules such as inosinic acid and glycine. Once tumours were removed from the patients, those bladder cancer metabolic features were shown to be reverted to their normal state, confirming that they were playing a significant role in bladder cancer development.[64] Another research study has examined changes in urine metabolites in response to disease onset and drug effects on renal glomerular filtration. The metabolite profiling assay showed that concentrations of glucose, sorbitol and trimethylamine oxide were elevated in renal transplant patients after the initial doses of the transplant drug tacrolimus, suggesting that these metabolites might be explored as potential biomarkers.[65] Some studies showed that human sweat can also be a non-invasive biofluid easily measured and analysed for metabolomic profiling, however, more studies need to be performed in this area.[61,66]

Most studies investigating the human metabolome are normally limited to a single biofluid, most commonly blood or urine. However, in order to analyse complex organisms or metabolic disorders in full detail, it is sometimes advised to take into account an analysis of multiple biofluids simultaneously. Recent metabolomics studies on diabetes, for example, were integrating metabolite profiling data from several human biofluid samples (blood, urine and saliva) taken from diabetes type 2 sufferers.[67-69]

Therefore, there is a great need to analyse human biofluid and tissue samples for metabolite profiling, in search of reliable biomarkers and to link them to metabolic pathways – this will help to accurately diagnose, and measure disease progression and response to therapy. Therefore, a successful metabolomics study using a thorough experimental design is crucial.

1.2 Technical aspects of metabolite analysis

1.2.1 Analytical instruments in metabolomics

Different highly sensitive and modern instruments, analysing a variety of metabolites, have been consistently used for global metabolite profiling and targeted analysis. Examples of common analytical applications in metabolomics include high performance liquid chromatography (HPLC) or gas chromatography (GC), mass spectrometry (MS), nuclear magnetic resonance (NMR) spectroscopy, enzymatic assays, as well as vibrational spectroscopy including Raman spectroscopy and Fourier-Transform Infrared spectroscopy (FT-IR).[31,70] Metabolomics analysis requires highly sensitive methods, usually capable of handling small volumes of samples and providing reliable accurate measurements of a large set of metabolites. Therefore, the selection of analytical techniques is usually based on the type and the amount of samples being analysed, as well as on the speed of analytical run, sensitivity and selectivity.[31,71]

The great advantage of using NMR in metabolomics, for example, is for its high reproducibility, a minimum requirement for sample preparation and the non-destructive nature of the technique.[72] NMR can generate useful structural information about metabolites involved in complex biological systems and their molecular interactions within biofluids and tissues. The technique also has advantages of quantifying endogenous metabolites without setting up precise analytical conditions. However, not all metabolites can be detected via this instrument, especially the complex mixtures, mainly due to sample clustering and signal multiplicity caused by spin-spin couplings. The low sensitivity of NMR, compared to MS, is a particular issue and limits its use in metabolomics. Nevertheless, NMR has been used in the past, mainly applied in targeted metabolomics studies.[73,74] In one NMR study, important metabolites were detected with significant clinical changes in plasma samples of cancer individuals. However, severe overlap of signals and long acquisition time were observed during the analysis, affecting the metabolite data. Another NMR based metabolomics study investigated the effect of green tea on the human urine metabolome, and although the results showed some important changes in

metabolite concentration levels, the resultant spectra of samples were very hard to interpret due to signal multiplicity and observed resonant frequencies.[75]

FT-IR spectroscopy is rarely used as a quantitative technique due to the complexity of spectra and poor selectivity, even though the technique is cheaper and more highly reproducible than NMR. As with NMR, FT-IR spectroscopy does not require a complicated sample preparation before analysis. However, it is not always easy to interpret spectra and discriminate metabolites coming from the same biological class. As the technique has limitations in quantifying metabolites, it is very hard to make precise, accurate calculations and detect which metabolites cause specific changes between different conditions i.e. potential biomarkers. Raman spectroscopy can produce better resolution of spectra peaks than FT-IR due to less interference of the aqueous phase (water produces weak Raman spectra) but, as with FT-IR, its spectra does not provide full information on metabolite origin (poor biochemical class distinction) and specific metabolite changes for biomarker detection.[76] Even though Raman spectroscopy has demonstrated a rapid non-invasive analysis of metabolites, it provides poor metabolite identification.[77,78]

The MS analysis, on the other hand, shows higher sensitivity and selectivity, compared to NMR, Raman or FT-IR, making the detection of hundreds or thousands of metabolites more accurately, even in complex samples.[79,80] A variety of MS instruments have shown high resolution of peak signals and suitability for high-throughput analysis of metabolites, providing higher accuracy in metabolite identification than other instruments. This is because the MS technique is measuring metabolite masses, rather than proton chemical shifts (as in NMR) or vibrational energy (IR). MS analysis has also been demonstrated to generally reduce the background signal noise and overlapping peaks in spectra of the metabolome.[81] Different MS and NMR techniques have been applied in biofluid, tissue untargeted and targeted metabolomics in the past, due to their clear advantages over other instruments.[82-84] Compared to NMR, however, MS techniques have provided more detailed information on quantity of metabolites and relative abundance of certain elements in their structures, especially when coupled with LC or GC instruments. MS has shown higher sensitivity, selectivity and being able to detect compounds that are usually not traced by NMR, e.g. sulphates.[82] Therefore, MS systems have been established as primary techniques in untargeted and targeted metabolomics. LC-MS and GC-MS have already shown excellent analysis of biofluid and tissue polar and non-polar metabolites, in a variety of clinical studies, detecting and quantifying a wide range of metabolites.[85,86]

1.2.2 LC-MS based untargeted metabolomics

Rapid development in the field of MS, coupled to separation techniques such as HPLC or GC, has become increasingly popular in untargeted global metabolite profiling. The use of chromatography helps to reduce the complexity of the sample, separating metabolites and making it easier to analyse them effectively.[87,88] Metabolites are separated in a time dimension according to their physicochemical properties (polar or non-polar) via isocratic or gradient elution techniques with appropriate mobile phases (see the details of gradients used in chapter two). Separation is usually performed before the ionisation stage in MS, and therefore can hugely affect the ionisation efficiency of metabolites later on in the ionisation chamber, so it is vital to choose the right chromatographic conditions from the very beginning.[89,90] GC is found to be more suitable for thermally stable volatile and non-polar compounds, while LC is used for a wide range of metabolites including polar and non-polar molecules.[86] With HPLC, there is a smaller chance of sample degradation and better separation of metabolites with liquid solvents, as well as better detection of wider range of metabolites, compared to GC. Overall, HPLC-MS is considered the primary method in untargeted metabolomics, in contrast to GC-MS, as it delivers adequate selectivity and high sensitivity for a wide range of metabolites.[85,91]

Since chromatographic separation depends on the nature of metabolites, different types of columns are applied for specific conditions. “Hydrophilic Interaction Liquid Chromatography” (HILIC) column[92] is most commonly applied for polar metabolites, while reversed-phase (RP) type, such as the C18 column, is mainly used for organic non-polar or semi-polar metabolites.[93,94] Mobile phases are also of great importance when it comes to analysis of samples, where they are carefully chosen depending on the type of analysed metabolites and chromatographic column. Aqueous phases (methanol, ethanol and water) are constantly used alongside organic phase solvents (acetonitrile), with chosen gradient systems depending on the purpose of the study.[95] When the appropriate mobile phase carries the analytes through the column, the analyte molecules travel at different speeds, causing them to separate and retain on the column at different times. Therefore, the analytes have different “retention times” (one of the parameters for their identification).[96] In HILIC conditions, polar analytes are retained on a column by partitioning between organic mobile phase (acetonitrile) and the aqueous layer formed at the silica-based surface of the column.[97] HILIC has been increasingly applied in recent untargeted HPLC-MS metabolomics studies, demonstrating superior retention of hydrophilic compounds and improved ionisation of metabolites.[98] HILIC uses the same composition of mobile phases as reversed column, which are MS compatible.[99] Good chromatographic

separation improves the sensitivity of MS spectra detection and results in better MS data quality due to reduced background noise. In some HPLC-MS experiments, it was also observed that increasing the temperature of the column allowed for the decrease of mobile phase viscosity, causing faster separation of analyte molecules, however this is not always the case as some metabolites can be degraded at high temperatures. Therefore, it is essential to monitor the chromatographic parameters throughout the experiment to find the optimum conditions for efficient analyte separation, part of the LC-MS based method development.[100,101] In chapter two, optimisation of the conditions will be discussed in more detail. The chromatographic process is shown in figure 1.2.

The MS consists of several stages including ionisation of sample molecules and mass analysis of ions by electromagnetic, electric or radio frequency fields. These help to provide highly accurate mass data. During the ionisation step, the sample molecules are bombarded by various ionising agents such as high-energy electrons (in the case of the "Electron Ionisation" approach)[102] or by gaseous ions ("Atmospheric Pressure Chemical Ionisation" approach)[103], forming unstable ions. However, in "Electrospray Ionisation" (ESI), ions are formed in a different way, involving extensive solvent evaporation (desolvation).[104] In an ESI ion source, a narrow metal capillary is held at high voltage, and a potential difference is produced between the capillary and the orifice of the mass spectrometer. A solution containing analyte ions (LC eluent) is sprayed from the capillary, and this capillary high voltage is applied to the solution to create an aerosol. This is known as nebulisation. During nebulisation, analyte molecules are believed to undergo redox reactions at metal interface of capillary tip. The formed ions in liquid droplets (aerosol) become de-solvated as solvent molecules leave the droplet surface. Solvent molecules start evaporating from larger charged droplets, shrinking droplets in size. The charge density on the shrinking droplets builds up until surface tension is balanced by repulsion at so called "Rayleigh" limit. At this limit, only de-solvated ions remain. After rapid solvent evaporations, the final ions appear in a gas-phase state and highly-charged.[105] Figure 1.3 demonstrates the process of ion formation through ESI technique. ESI is the preferred method to use in the majority of metabolic experiments, as it is shown to produce a large number of intact precursor molecular ions, therefore considered to be a "soft ionisation" type and suitable for polar metabolomics, i.e. human biofluids. ESI has shown to reduce intermolecular fragmentation, while APCI has shown to induce little ion fragmentation, mainly analysing lipids and non-polar compounds. The produced ions are then separated according to their mass-to-charge (m/z) ratios by a mass analyser, either by being trapped in three-dimensional orbits (in the case of "Orbitrap" mass

analyser) or by being accelerated towards the detector depending on their kinetic energies proportional to their m/z values ("Time-Of-Flight" mass analyser).[106,107]

In summary, LC-MS is an appropriate analytical tool for untargeted metabolite profiling of human biofluid and tissue samples, providing detailed information on quantity and chemical structure of metabolites.[108] LC-MS methods with appropriate chromatography, e.g. HILIC and electrospray soft ionisation (ESI), have been found suitable for biofluid sample metabolite analysis – urine and plasma samples can be analysed with minimal sample pre-treatment, reduced intermolecular fragmentation due to soft ionisation and better molecule separation.[109] However, this analytical technique can cause metabolite losses during the experiment, as well as producing biased results due to matrix effects.[109] So, there is a need for a suitable metabolite sample preparation and extraction procedure for certain polar metabolites, as well as correct processing of raw data, in order to produce highly accurate metabolite information.

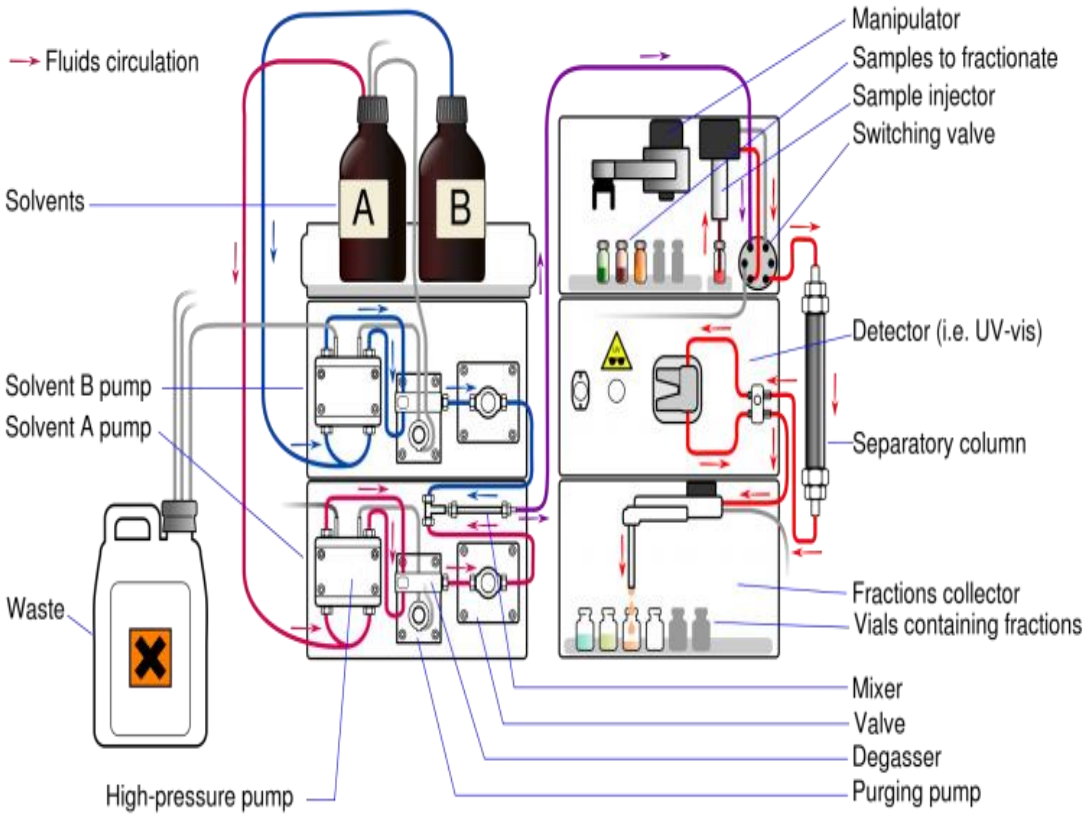


Figure 1.2. Schematic diagram of HPLC system typically used in LC-MS analysis experiments. **Peng WP et al, Measuring Masses of Large Biomolecules and Bioparticles Using Mass Spectrometric Techniques, 2014.**

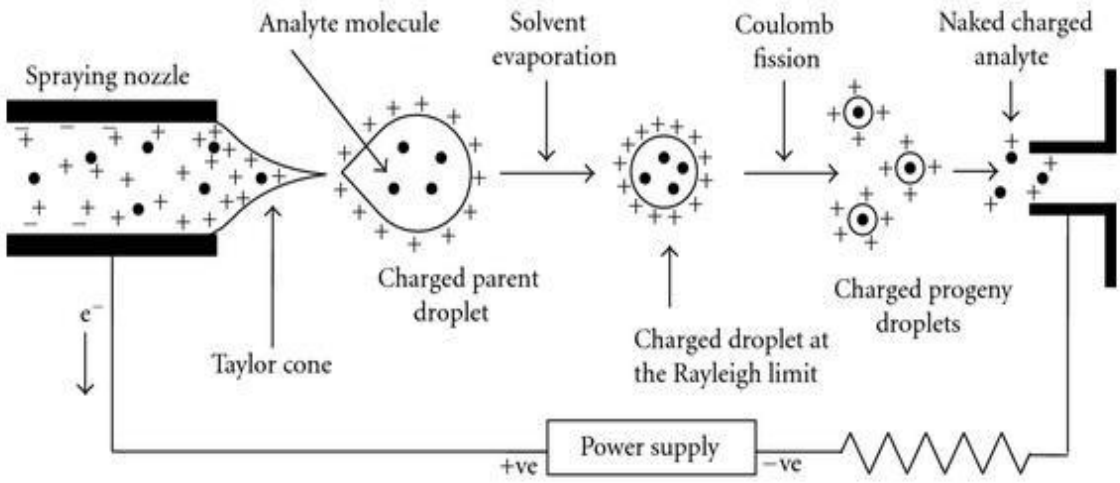


Figure 1.3. Schematic diagram of ESI source operated in positive mode in MS system. **L Konerman et al, Unraveling the mechanism of Electrospray Ionisation, 2012.**

1.3 Sample preparation and extraction aspects

1.3.1 Sample preparation in LC-MS metabolomics

When designing an LC-MS metabolomics study, an appropriate number and type of biological test samples should be considered, as well as correct procedures on how to handle them effectively through extraction and sample preparation. In untargeted metabolomics, it is particularly important to extract and prepare samples effectively, in order to obtain a large number of endogenous metabolites for analysis, and also to have a wider range of metabolites detected – that is an aim of untargeted analysis. Generally, sample acquisition in any untargeted metabolomics study is primarily driven by experimental design – if human samples are involved, the influences of gender, diet, age and genetic factors should be considered before collecting samples.[110,111] To obtain statistically validated metabolite data from human and animal samples, sometimes a so-called "statistical power analysis" or "sample size distribution" can be performed to ensure a sufficient number of samples are acquired for the study to reduce the influence of biological variability later on.[112,113] These statistical analyses are normally conducted before sample collection to help researchers to determine the smallest sample number required for the study to detect a desirable level of significance. One study demonstrated the use of these techniques with certain modifications of statistical calculations on human and *C. elegans* microorganism samples, to decide on sample number for the experiments, and has achieved a high power of 0.8 for 200 samples, making a decision to run a large number of samples.[114]

During preparation of biofluid and tissue based biological samples, extra care should be taken to minimise the possible degradation of metabolites due to remaining enzyme activity or oxidation processes in biological cells. Specific techniques have been put in place to inhibit enzyme or oxidation activity, such as freezing in liquid nitrogen[115], acid treatment[116] or quenching in salt with low temperature methanol.[117] In several studies, quenching samples in ice cold methanol was shown to minimise the enzyme metabolism in cells. However, this method demonstrated that cold shock quenching may cause metabolite leakage from cells, and therefore cause the loss of important metabolites.[118] Quenching is used occasionally in biofluid and tissue sample preparation, but the method is often not well controlled and may introduce variability in the dataset. Therefore, extreme care and appropriate quenching methods have to be employed for particular types of sample.[119]

For biofluid and tissue metabolomics studies, each experiment has to cover a sufficient amount of biological and analytical sample replicates, quality controls (QCs) and standard reference samples.[120] This is needed for a study to have the desired statistical power to make the data more reliable. According to standard requirements proposed by the Metabolomics Society[121], a minimum of three biological and five analytical replicates of each sample should be tested in any LC-MS experiment. QC samples are created to assess the analytical variability by mixing equal amounts of pooled samples together and repeatedly analysing throughout the LC-MS experiment.[120] The design should also include the appropriate sample storage and preparation conditions. For most of the metabolites, the storage temperature should be around -80◦C to avoid their degradation.[122] Recent HPLC-MS studies in untargeted biofluid metabolomics showed that consistent sample preparation, handling and correct storage minimises variability in the later analysis.[31]

1.3.2 Extraction of biofluid and tissue based samples

Over the years, many standard protocols have been established for efficient extraction of polar and non-polar metabolites from human and animal biofluids and tissues. Incorrect sample preparation has a great influence on experimental reproducibility, so it is vital to consider the right preparation technique. Human biofluids and tissues consist of complex fluid and matrix proteins, making it harder to prepare the samples for experiment using standard protocols. Therefore, to fully optimise the preparation methods, metabolite extraction is of paramount importance.[123,124]

Various extraction procedures are involved in metabolomics methodology to obtain a good metabolite recovery for LC-MS analysis. The most common technique used for biofluids and tissues is "liquid-liquid" extraction[125], where polar and non-polar metabolites are extracted by shaking and vortexing with organic solvents which are chosen depending on metabolite chemical properties – normally methanol or ethanol for hydrophilic metabolites, chloroform or ethyl acetate for the lipophilic (hydrophobic) type.[126,127] The main advantage of this technique is the generation of a biphasic sample, fractioning metabolites into aqueous and lipophilic organic layers, which can be analysed separately. Changing the temperature and composition of the extraction solvent can sometimes improve metabolite recovery but not significantly, as observed in a few studies.[128] It is currently a challenging task to get a sufficient metabolite recovery in biofluid and, especially, tissue samples, to show high reproducibility. This is mainly due to the dynamic nature of metabolite turnover, enzymatic activities occurring inside the cells and a complex mixture of metabolites, as mentioned earlier.[117] Quenching as a part of the extraction procedure helps to deal with some of these matters, however, additional steps in extraction are required to obtain better metabolite recoveries. Optimised extraction parameters for each type of sample can lead to more robust metabolomics studies.[129]

Extraction steps in addition to organic solvents (which separate polar from non-polar metabolites), will primarily depend on the nature of the metabolites, type of biofluids, and compatibilities of solvents and temperatures with the LC-MS apparatus.[130] For human plasma samples, for example, the extraction procedure would usually involve centrifugation, in order to separate serum from blood plasma, and deproteinisation to remove unnecessary proteins from the samples which may interfere with further analysis. Some studies used an anticoagulant to separate plasma from the whole blood sample.[131] The application of acidic treatments such as nitric acid has also been used to separate plasma from blood but caused a severe reduction in the number of detected metabolites, and so are not commonly used. For urine samples, additional extraction steps usually involve centrifugation and a "solid-phase" technique using specific sorbents.[132] One study involved the extraction of disease-control human urine samples with an extraction solvent of hydrogen chloride in water (1:1 ratio), followed by evaporation to dryness. The extracts were successfully recovered, and later GC-MS analysis showed a high peak response of metabolites. Another study described the method of solid-phase extraction on urine metabolites, which were successfully separated from the matrix and liquid phase by retaining them on an alkylated silica sorbent.[133]

For tissue samples such as liver, kidney and brain, advanced extraction steps are required in addition to the use of appropriate solvents which separate polar from non-polar metabolites.

This is because extraction of tissues is the most labour intensive and error prone method in metabolomics, so requires extra care steps. Methods usually involve homogenisation of tissues[134], freeze-thawing using liquid nitrogen[135], mortar grinders, mixing with beads and using hot water baths.[136] These procedures demonstrated effective breaking of cell membranes, helping to release intracellular metabolites. One study used bead-based homogenisation in combination with a standard organic solvent extraction for several animal tissues (liver, kidney and muscles), demonstrating high metabolite recoveries and improved extraction yields, compared to standard (solvent-only) techniques.[137]

It is recommended that the proposed minimum standards for extraction and sample preparation, outlined by Metabolomics Standards Initiative (MSI), are followed.[138,139] These extraction procedures must be used with care and knowledge of the likely impact on metabolomes in the samples.

1.4 Data processing and current issues

1.4.1 Data pre-processing and statistical analysis

Data pre-processing

After an analytical run of biological samples on an LC-MS instrument, the 3D of LC-MS data produced (in the form of peak intensities, mass-to-charge ratios, retention time of analytes) is usually obtained with a specifically written programme for the instrument, e.g., "XCalibur" for LC-MS Exactive, visually showing the spectra peaks (raw data). The raw data files are then pre-processed to generate 2D output files, which can later be used for further statistical analysis.[140] Data pre-processing has played an important role in metabolomics research so far, with several bioinformatics tools assisting the accurate processing of data.[141] Computational platforms, such as "metAlign", "XCMS" and "mzMatch," were developed for a variety of pre-processing purposes, transferring raw data into a more organised, tabulated and accurate format.[142] "MetAlign" is intentionally designed to remove any background noise and outliers from the raw data, correct baseline, perform peak matching and align retention times.[143] "XCMS" processing is found to be similar to "metAlign", but it also helps scientists with the identification of putative metabolites, structural characterisation and statistical analysis. "XCMS" has been used more regularly in human biofluid and tissue untargeted LC-MS metabolomics and metabolite profiling, as well as targeted analysis. The main advantage of "XCMS" over other pre-processing

software is that it can correlate specific metabolites to their biological origin and incorporate non-linear retention time alignment, helping to calculate relative metabolite ion intensities – useful in identifying potential biomarkers later on.[144] "mzMatch" helps with data pre-processing further by matching and aligning peaks.[145] All pre-processing stages applied on the raw data are normally reported in detail when the studies are published, so that other scientists can replicate those steps again. This has been strongly recommended as a proposed minimum requirement by MSI.[146]

However, even with advanced pre-processing tools being used, limitations still exist with the presence of background noise or artefact peaks often incorrectly recognised as metabolites. Metabolite identification algorithms, applied in those programmes, are hard to use without special training. Recently, a group of scientists in Molecular Parasitology, Glasgow University has developed a bioanalytical programme called "IDEOM" that comprises several advanced pre-processing computational programmes, including "mzMatch" and "XCMS" tools, and metabolite identification steps.[147] This was performed to improve comprehensive data processing further, particularly paying attention to noise filtering, correct identification of metabolites and data visualisation. Putative metabolites can be identified by matching their masses and retention times to authentic standards. Retention times of those authentic standards are uploaded into IDEOM before the main processing steps of the sample data. Therefore, the final list of putative metabolites is produced with identification confidence levels related to standards. Retention times of authentic standards uploaded into the programme help to make a retention time prediction model, assisting with confidence identification of putative metabolites.[148] IDEOM is designed as an Excel user-friendly interface, allowing for a rapid data generation without the need for complicated bioinformatics skills. Additionally, IDEOM provides dual polarity data merging, targeted analysis and isotope tracking for normalisation.[147]

Statistical analysis

Once the pre-processing has been performed, statistical analysis can further help to detect significantly altered metabolites between distinct biological groups, helping scientists to find potential biomarkers or major changes in metabolic pathways in certain clinical conditions. Specific choice of statistical analysis methods depends on the design and aim of the study.[149] There are usually two main statistical approaches involved in metabolomics and other omics studies: univariate and multivariate analysis. The univariate approach assesses the statistical significance of each metabolite peak separately, using techniques such as t-test, sum test, false discovery rate (FDR)[150] and analysis of variance (ANOVA)[151]. T-test compares the means of two

groups, for example control vs treatment/disease, calculating p values for each metabolite to determine their significance – if the p value is less than 0.05, the null hypothesis stating a lack of statistical significance is rejected and instead accepts that there is significant difference for particular metabolites between the studied groups. In a large, untargeted metabolomics study, with a wide range of detected metabolites, a multiple hypothesis t-test for a large dataset may cause high false discoveries, even with a small p value threshold. Therefore, an FDR correction is usually performed on the metabolite dataset straight after the t-test to control and prevent a large total number of false discoveries.[152] Univariate statistical tests help to reduce a large number of measured metabolites to only those which show the strongest response under the investigated conditions. In one study, a two-way ANOVA test investigated medication-induced level changes of individual metabolites, and has managed to demonstrate several statistically significant metabolites from patients with metastatic disease.[25] However, univariate analysis fails to discriminate between biological groups properly if there are only minor differences present at the single molecule level. Multivariate analysis, on the other hand, discriminates between groups based on changes in single molecule levels, as well as on dependency structures between the individual molecules.[153,154]

Multivariate analysis is usually used to reduce dimensionality (random variables) of metabolomics data and to see some general trends between groups in highly complex datasets, by dealing with multiple variables, rather than one variable as in univariate analysis. Multivariate analysis allows us to identify biologically relevant spectral features and biological class differences from the multivariate dataset.[154] This type of analysis is further categorised into supervised and unsupervised methods. The “Principal Component Analysis” (PCA) technique belongs to the unsupervised category that explains the variance of a dataset by analysing principal components (vectors of metabolite contributions), usually without a knowledge of metabolite classes or groups – therefore called unsupervised. The technique focuses more on differences between samples, rather than between groups. PCA is often used as a starting step in multivariate data analysis, and has been extensively applied in untargeted metabolomics studies where it showed good data visualisation, clustering of sample points and sample group discrimnation.[42,155] One of the developed unsupervised techniques, ANOVA-Simultaneous Component Analysis (ASCA), combines ANOVA and PCA methods, and is normally applied in datasets with complex designs, allowing variation in the data to be directly related to various design factors. This technique successfully showed the sample clustering and distribution of saliva metabolome in one of the studies, but the method is still under investigation.[156,157] A supervised technique, such as “Orthogonal-partial least square-discriminant analysis” (OPLS-

DA), builds on the PCA model and uses the sample group information for classification purposes. This classification helps to discriminate between sample groups, improving clustering and separation between data groups in the model, helping to identify statistically significant metabolites.[158] However, this approach often misses some low concentration metabolites which are significantly changed, and which can be detected by univariate analysis. The issue of data over-fitting is also of a major issue in supervised technique.[159] So, it is very important to choose the appropriate samples for this type of analysis – human biofluid metabolite studies normally involve observation and analysis of several variables at a time, therefore they benefit from using a multivariate analysis in search of biomarkers.[160]

Both univariate and multivariate analysis approaches are regularly used in human metabolomics studies, in particular untargeted, producing highly accurate data with clear observations between sample groups. This helps metabolomics research in the search for biomarkers and metabolic pathways associated with clinical diseases and other types of perturbation.[161,162]

1.4.2 Issues with metabolite data

Despite the advances of efficient LC-MS methodology approaches in untargeted metabolic studies, the majority of metabolomics data are still being affected by the presence of technical, instrumentational and biological effects.[4,163] Because of those issues present in LC-MS analysis, it is common to observe a so called “non-linear peak signal response” for metabolite concentrations, where the peak signal response is not proportional to the metabolite concentration increase.[164,165] Accurate quantification of metabolites relies heavily on this proportional relationship, important in both untargeted and targeted metabolomics. However, it is hard to achieve such a proportional relationship in every single metabolomics study with all types of biofluid and tissue samples. Therefore, metabolite data pre-processing steps, such as peak detection and integration, may introduce an error in metabolite data, significantly affecting the statistical analysis and quality of data.

This can happen due to several analytical issues such as signal saturation at high analyte concentration during ionisation, matrix effects, unwanted in-source fragmentation, metabolite degradation during sample preparation, all of which contribute to analytical variation between sample replicates.[166,167] Matrix effects hugely influence metabolite data; the composition of evaporated sample in the ionisation chamber of the MS can affect the ionisation efficiency of the sample molecules and this, in turn, can lead to either ion suppression or enhancement, with

a huge impact on signal intensities, as well as precision and coverage of metabolome.[168,169] Matrix effects commonly happen with all types of ionisation, even with ESI.[170] During in-source fragmentation in MS, metabolites can break down into more than one ion species, generating unwanted by-products of identical mass to common metabolites, complicating metabolite identification during the detection stage.[171] Instrumentational issues include background noise and analytical drift in retention times, which can be easily resolved with correct solvents, column and sample preparation. These analytical and instrumental issues are also known as "obscuring" sources of variability in data, introduced during experimental studies, mainly coming from human or instrument error. When an unwanted analytical variability occurs between sample replicates, it becomes difficult to quantify metabolite concentrations and identify them accurately. It is recommended by MSI that QC samples should be analysed repeatedly throughout the experiment, to evaluate and control analytical variability. [172,173]

Biological variation is a separate and more complicated issue to deal with. This type of variation arises between biological samples and replicates (usually in human and animal models). This variation is caused by dietary, environmental and physiological states, and genetic factors influencing the composition of metabolomes.[174] Different biological cell sizes, varying sample measurements in the experiments (sample weight, volume) and, most importantly, varying metabolite concentration levels in biofluid and tissue samples all contribute towards the variation in metabolite data. Urine samples, in particular, show a wide variation in metabolite concentrations, due to biological homeostasis occurring in the human body and chemical complexity of metabolite breakdown products.[175] Personal hydration status, large variations in physiological pH and osmolality, as well as unpredictable dilution of metabolite concentrations, makes the urine metabolome very varied. Human and animal tissues (brain, kidney and liver) may also introduce huge variability in metabolite data, usually because of multiple step extraction techniques causing errors and low reproducibility between sample replicates. Tissue extracts can also have different total numbers of metabolites, depending on the composition of cells, cell lysis efficiency and cell density within the tissue sample. Understanding the causes and removal of unwanted biological variation poses a great challenge in metabolomics research. This is further complicated by not knowing how to measure or quantify this type of variation. In most clinical metabolomics studies, complications may arise from a combination of biological and analytical data variations happening in the same sample, making it difficult to detect the origin of unwanted biological variation.[176]

All of the mentioned obstacles could significantly reduce the quality of metabolite quantification and identification, particularly in low abundant metabolites. The lack of validated quantitative data and non-linear responses in metabolite peaks affect the true measurement of metabolites, and hence their biological relevance. Therefore, there is a need to reduce these obstacles, particularly large metabolome data variation, with a standardised analytical method.

1.5 Normalisation of metabolomics data

1.5.1 Concept of normalisation

To overcome the analytical, biological and instrumentational obstacles mentioned above, the metabolomics data can be normalised. The purpose of normalisation is to identify and remove both the observed and unobserved factors which trigger non-linear response between metabolite concentration and peak signal intensity, improving data accuracy overall.[177] Normalisation forms an integral part of metabolomics data analysis nowadays, with a variety of approaches being employed in untargeted and targeted studies.[172,178] This includes most common techniques such as normalisation to creatinine (commonly used for urine metabolomics)[179], "osmolality"[180], "total ion counts" (TIC)[181], "mass spectrum total useful signals" (MSTUS)[182] and the use of isotopically labelled chemical compounds as the source of internal standards, also known as "stable isotope labelling" (SIL) normalisation.[183] Other less common techniques, which have been used in several metabolomics studies, include "between batch signal intensity"[184], "specific gravity"[185], "UV absorbance"[186], "cell count"[187] and "matrix normalisation factor"[188] methods. However, these methods were used in a specific type of study and were not thoroughly investigated in all types of samples, so are not regularly used. Because researchers want to determine metabolite concentration changes caused only by biological or clinical factors (e.g. disease, perturbation), rather than non-biological factors, metabolite normalisation is critical and necessary in LC-MS metabolomics.[189,190]

Normalisation mainly targets systematic variation in LC-MS data, as it appears to be the biggest issue arising from errors in sample preparation, extraction and other biological fluctuations.[191] The technique detects and reduces this variation, and thus helps to improve the linear response and overall accuracy of data. The choice of appropriate normalisation technique depends on experimental design (untargeted or targeted study), statistical aims of the study and data being analysed (nature of samples). For biofluid or tissue samples, it is very hard to find a perfect

common compound or method that can be recognised as a reference for normalisation and effectively normalise a wide range of metabolites in those samples in untargeted metabolomics. Metabolite concentrations vary a lot in this type of sample, as mentioned earlier, and the degree of variation is different from one type of biofluid or tissue to another. Therefore, for every study, a different normalisation approach may be appropriate – for example, normalisation to creatinine is more suitable for urine metabolomics, while MSTUS normalisation can be used for all types of samples/studies.[178,192]

It is also important to decide how to assess normalisation effectiveness accurately. In developing a normalisation method, overall performance is assessed by the ability of the technique to minimise the variations within the same sample groups. However, more complicated techniques, for example those involving internal standards, which are able to remove interfering variation better than simple strategies, may also remove part of the useful variation and even increase the degree of variation in some metabolites. For this reason, it is important that normalisation effectiveness is not only monitored through reduction of variability, but also by measuring the amount of differentially abundant metabolites (increased or decreased after normalisation). While developing a normalisation method, it is also important to establish correlation between a new method and already existing normalisation approaches, making comparisons of normalised data between those techniques as part of the method validation. It is assumed that under normal conditions (healthy human samples), normalisation results obtained from different techniques should have positive correlations with each other.

1.5.2 Pre- and post-processing data normalisation

Pre-processing normalisation

The current normalisation methods used in metabolomics studies can either be the pre- or post-processing type. Pre-processing normalisation occurs during the sample preparation stage, before the metabolite data is acquired by LC-MS, while post-processing techniques are applied after the metabolome data has been acquired, without affecting the sample preparation stage.[184] In most pre-processing techniques volumes of biological samples (or metabolite concentrations) are adjusted according to measured quantities of specific internal standard(s). A suitable internal standard should show similar behaviour to a measured analyte, for the normalisation process to be effective. The calculated ratios of metabolite ion intensities adjusted to intensities of internal standard(s) are used later in statistical analysis, helping to reduce unwanted variation in sample metabolite concentrations. One of the advantages of pre-

analysis normalisation is that similar instrumental responses can be observed in all the samples, because the same amount of internal standard is used in every sample. This advantage is very important, especially in ESI-MS analysis where the signal response of individual metabolites in all the samples is not always uniform at different concentrations. As a result of this normalisation, the metabolite concentrations to signal response may be more linearly scaled. This leads to another advantage of improving the overall metabolite information recovery, with fewer metabolites showing a non-linear response. However, one of the biggest disadvantages is that the technique cannot use every single compound for each metabolite of interest. Examples of pre-analysis techniques include SIL, creatinine value, specific gravity, normalisation to mass and Ultra Violet (UV) absorbance.[193]

SIL normalisation is perhaps the most common pre-processing technique that can be used in different areas of metabolomics and pharmaceutical research, mainly used so far in targeted studies. In this technique, isotopically labelled internal standards are introduced right at the beginning of sample preparation or during metabolite extraction. The technique usually employs stable isotopes, such as carbon (^{13}C), oxygen (^{18}O), nitrogen (^{15}N) and deuterium (^{2}H).[194] Because physicochemical properties of isotopically labelled internal standards are similar to measured analytes, the signal changes caused by matrix effects can be effectively calibrated, improving quantification of endogenous metabolites – a big advantage in targeted metabolomics. Most targeted studies showed that using an isotopically labelled standardisation method successfully reduces background noise and matrix effects, improving metabolite recovery.[195] For example, one study using ^{13}C and ^{15}N isotopically labelled internal standards for phase 2 drug plasma metabolites in a pharmacokinetic study measuring glutathione conjugates, successfully observed an improved absolute quantification of these metabolites, compared to the non-normalised (without internal standards) state. This normalisation technique helped to obtain highly accurate pharmacokinetic data of the drugs and better understand their toxicity.[196] However, recent SIL normalisation publications have demonstrated that ^{2}H labelled internal standards, for example, failed to demonstrate accurate quantification of metabolite concentrations, leading to unacceptable matrix effects even in the presence of labelled internal standards.[197] The study suggested that the replacement of a hydrogen atom with a deuterium isotope on an analyte's molecule slightly altered its hydrophilic properties, and so affected its retention time, making it different from the unlabelled version of the metabolite – this is also known as the isotope effect.[198,199] Isotopes ^{15}N and ^{18}O have not been used in metabolomics that frequently, because, as mentioned earlier, not all metabolites contain those elements in their

organic structures. Therefore, in untargeted metabolomics, it would not be convenient to use them in SIL normalisation.

Normalisation to creatinine also falls into the pre-processing category and is only used in urinary metabolomics, expressing each metabolite level relative to creatinine concentration.[200] This normalisation is suitable for the urine metabolome, because the rate of excretion of creatinine through glomerular filtration is relatively constant in human organisms under normal (healthy) physiological conditions.[163] Therefore, this technique helps to achieve highly accurate results with reduced variation in LC-MS data, although SIL normalisation has also showed similar results in a variety of studies. Creatinine normalisation, however, is a limited technique and cannot be used for other types of biofluid or tissue samples, showing a clear limitation. Also, creatinine levels may be affected during clinical diseases, especially in renal impairment disorder, and therefore would not work as a normalisation technique in such a case.[201]

Osmolality normalisation is another example of pre-processing technique only applied in urinary metabolomics. Osmolality is a direct measure of total urine solute concentration and therefore is also a good valid scaling factor. In some studies, it was even demonstrated that application of osmolality normalisation showed better separation between different biological groups, compared to creatinine normalisation.[180,202] Normalisation to creatinine, osmolality and cell count can all be applied as post-processing methods as well. The limiting factor of these methods being post-processing is that they are applied after LC-MS analysis and so they perform poorly for metabolites that do not follow a linear dilution pattern – for example, for low abundance metabolites which can be diluted below limit of detection or for highly concentrated metabolites exceeding dynamic range of mass spectrometer. Therefore, these methods being pre-processing can perform better and are more commonly than post-processing.[180,203]

Normalisation using UV absorbance has been applied in a variety of biofluid samples such as plasma, urine and even body sweat and saliva. In this technique, each metabolite concentration is adjusted to the total concentration of solutes which absorb UV at a specific wavelength. This way, the technique is more representative of overall sample composition, compared to the use of a single internal standard such as creatinine. One study demonstrated that this method was preferred over creatinine normalisation in urine metabolomics and expressed highly accurate results. However, the biggest limitation to this normalisation technique is that it requires an LC-UV instrument, rather than an LC-MS, and this is a very expensive instrument to use simply to improve data accuracy.[204]

Post-processing normalisation

In post-processing normalisation, the samples are not experimentally normalised and therefore do not require additional time-consuming steps during extraction and sample preparation, showing a clear advantage over the pre-processing method. This is especially important to consider when a large number of samples are analysed. However, any post-processing normalisation does not involve internal standards, so individual metabolite signals are adjusted based on a certain criterion or common denominator, not necessarily closely related to specific metabolites. Therefore, the analyte signals may be less linearly scaled with metabolite concentrations in complex biological studies, showing lower data accuracy than pre-processing techniques. Common examples of post-processing techniques are MSTUS, TIC, normalisation to osmolality, cell count and "median fold change".[184]

The MSTUS technique uses the sum of all peak height intensities of metabolites within a sample chromatogram as a common denominator, calculating ratios which are used in statistical analysis. This technique is similar to TIC normalisation where all mass spectra are divided by their TIC, assuming that there are comparable numbers of signals present in each spectrum for datasets, and all spectra in a dataset have the same integrated area. However, when comparing widely different tissue or cell types, they may express different ion distributions due to their heterogeneity, so TIC and MSTUS may not be appropriate normalisation in such cases.[180,182]

As is the case with creatinine normalisation, normalisation to osmolality is only common in urine metabolomics, because urine osmolality is a direct measure of total urinary solute concentrations.[205] Osmolality also provides a true physiological measure of urinary metabolite concentration. But, as with creatinine normalisation, this technique's disadvantage is that it is limited to urine metabolomics, and during renal disease or impairment, osmolality can be hugely affected. Also, this type of normalisation requires a special clinical instrument to measure osmolality that is not easily available.

It is clear that post-processing normalisation has more limitations than advantages when compared to pre-processing normalisation. However, there are circumstances when pre-processing normalisation might not be suitable for a particular set of samples or is too expensive.

1.5.3 Comparison of normalisation techniques

So far, it is clear that the decision as to which is the most suitable normalisation technique in LC-MS metabolomics is based upon several factors: type and number of biological samples (biofluid, tissue), normalisation accuracy, cost, convenience (time consuming or not), and availability of specific instruments or conditions. Table 1.1 summarises the advantages and disadvantages of common pre- and post-processing methods, with a focus on instrument availability, time and convenience to perform measurement, as well as the method validity.

From the general comparison in table 1.1, it can be seen that SIL normalisation technique using isotopically labelled internal standards has major advantages when compared to other methods. It helps to produce higher quality data by assigning individual metabolites to specific internal standards, which are the same metabolites but isotopically labelled and have similar physicochemical properties. It has been proven by a variety of untargeted and targeted metabolomics studies that using isotopically labelled compounds for normalising data reduces variations between analytical and biological replicates, improving accuracy and precision for the analysis of small and large metabolites. The SIL normalisation technique is not limited to a particular type of sample, but can be used with every type of biofluid and tissue. It does not require any additional instruments such as UV, freeze-dryer or osmometer (for osmolality measurements), making the sample preparation method easier. And finally, as it is a pre-processing normalisation method, it helps to correct for sample preparation variations introduced during the extraction steps and compensates for variability in MS detection.[199]

Normalisation technique	Pre or post processing type	Advantages	Disadvantages	Suitable sample type
Creatinine	Pre-processing Post-processing	- Constant levels during normal conditions - Relatively inexpensive - Quick and easy to perform	- Limited to urine metabolomics only - Creatinine levels may not be constant during perturbations/renal impairment	Urine
Cell count	Pre-processing Post-processing	- Accurate technique - Direct measure of cell numbers	- Larger variations observed, especially with small cell numbers	Tissues
MSTUS	Post-processing	- More universally applicable - Easy to make calculations - Contribution from artefacts and xenobiotics is minimised	- Not entirely accurate as it doesn't use a specific internal standard for each metabolite	All types
Normalisation to mass (weight)	Pre-processing Post-processing	- Samples are adjusted to the same mass concentration to reduce matrix effects	- Extraction volumes are not the same for each sample, so may have effect on extraction recovery - Time consuming process	Tissues

Osmolality	Pre-processing Post-processing	- Provides closest measure to physiology of urinary concentration	- Limited to urine metabolomics - Requires specific instrument	Urine
SIL (isotopes)	Pre-processing	- Normalisation of individual metabolite to specific internal standard (higher accuracy obtained) - Internal standards having similar physicochemical properties as analyte metabolites - Easy to purchase (especially ^{13}C)	- Cannot cover a wide range of metabolites (availability limited) - Expensive to purchase individually labelled standards for use with each metabolite - Some isotopes (deuterium) might express isotope effect	All types
TIC	Post-processing	- More universally applicable - Easy to make calculations	- Not entirely accurate as it doesn't use a specific internal standard for each metabolite - Involves all sorts of signals including those from artefacts and xenobiotics	All types
UV absorbance	Pre-processing	- High accuracy measurement - Low sample consumption	- Requires expensive LC-UV instrument - Accuracy may be reduced in disease state	All types

Table 1.1 Advantages and disadvantages of various pre- and post-processing normalisation techniques applied to biofluid and tissue samples

The most common isotope used in SIL normalisation is usually ^{13}C, since the majority of analysed metabolites contain carbon atoms in their structure. So, a wide range of metabolites can be normalised using this isotope in comparison to other isotopes such as ^{5}N or ^{18}O. The ^{13}C isotope is found to be more abundant and stable than ^{2}H, as ^{2}H can be easily vaporised in the ionisation chamber of the mass spectrometer. In contrast, ^{2}H would be preferred over the ^{13}C isotope in NMR based metabolomics due to sensitivity issues. Unlike deuterium, ^{13}C labelled internal standards do not cause the isotope effect when their carbon atoms are replaced with isotope on the molecular structure.[194]

1.5.4 ^{13}C normalisation approach and current limitations

During ^{13}C normalisation, single or multiple internal standards can be used, with the latter type usually favoured. Using a single ^{13}C labelled metabolite standard to normalise each metabolite peak makes the technique very limited and assumes that all compounds experience the same variation as the internal standard, which is not an accurate assumption. In some respects, it is similar to a creatinine normalisation, where every metabolite's concentration is dependent on

one standard. A single standard cannot sufficiently represent specific analytical features or discriminations for all of the analytes.[195] The decision to use multiple standards, on the other hand, may be a more reasonable choice; they are applied to analytes which elute within a specified retention time window.[206] Preferably, each internal standard should be assigned to normalise its unlabelled pair (analyte metabolite peak) – for example, a ^{13}C labelled amino acid L-alanine would normalise unlabelled L-alanine. The analyte of interest and its labelled internal standard are affected by the same analytical conditions during sample preparation, extraction and analysis, so the ratios between them should not be influenced by any analytical discriminations or by a certain type of metabolite class, e.g., lipophilic or hydrophilic compounds.[198] When the exact internal standard could not be found for a particular metabolite, a different internal standard with similar chemical properties, such as the same retention time, could be used instead (this could be termed "co-elution" normalisation).[207] However, this may present a potential issue – retention time does not fully reflect all matrix and chemical properties of a metabolite, and it might not be very accurate to use one internal standard on several metabolites which happen to co-elute with the same retention time.[194] In some untargeted and targeted metabolomics studies, applying ^{13}C isotopically labelled internal standards, even co-eluted ones, has shown an accurate normalisation and reduction in data variation for most of the metabolites through statistical analysis.[189,206,208] Other studies expressed concern with regards to using "co-eluted" internal standards for metabolite normalisation, for the reasons mentioned earlier, and have demonstrated other ways of normalising metabolites which could not be matched with their labelled versions – one method was to normalise them by their own unlabelled versions from quality control (QCs) samples, while other study normalised such metabolites by the sum of all internal standards in the samples. Both techniques showed some progress but had their own limitations and potential inaccuracy of normalising data.[209]

Normalisation calculations and assessment

This type of normalisation usually involves calculating the ratios between peak areas or height intensities of unlabelled metabolites of interest and their ^{13}C labelled internal standards.[86] These calculated ratios are referred to as normalised data, which are later used in univariate and multivariate statistical analysis for obtaining highly accurate data. The ratios are usually calculated by simple division of labelled by unlabelled peak areas or intensities for each metabolite ($^{13}C/^{12}C$). To assess the normalisation technique, two methods are normally used: "relative standard deviation" (RSD) comparison and multivariate statistical analysis via PCA and OPLS-DA plots. Normalised ratios are calculated for every metabolite in each

analytical/biological replicate and then are expressed as percentage RSD values. They are then compared to percentage RSDs of unlabelled metabolites, to see if there is a significant improvement or not.[173] If metabolites are found to be normalised successfully, their normalised percentage RSD would be significantly reduced compared to unlabelled percentage RSDs, showing the reduction in inter-individual sample variation. However, an increase in percentage RSD would indicate unsuccessful normalisation, where either the wrong ^{13}C peak had been used or there were no ^{13}C peaks which could be used to normalise the metabolite – even "co-eluted" ones. Calculated ratios are also uploaded into a multivariate programme to produce PCA and OPLS-DA plots for further normalisation assessment, looking at the degree of separation between data groups and the cluster effect of replicate points. Successful normalisation would indicate a large degree of separation between the groups and tight clustering between replicate points.[210]

Limitations of the technique and a possible solution

Despite the regular use of ^{13}C labelled internal standards in normalising metabolite data, research scientists still face issues such as matrix effects and ion suppression problems. In one study, a slight change in retention time of certain metabolites such as carvedilol was noticed after using a ^{13}C labelled standard, affecting the calculated ratio between labelled and unlabelled carvedilol and causing inaccuracy.[199] Other studies demonstrated that internal standards had an effect on analyte structure and concentration, and flow rate in chromatography, all of which could cause ion suppression. However, the main advantages of using ^{13}C internal standards, as mentioned earlier, outweigh these limitations.

It has been noticed that in untargeted metabolomics, where a wide range of metabolites are detected, it is hard to achieve this sort of normalisation for all the analysed metabolites. Regrettably, it is difficult to obtain isotopically labelled standards for individual metabolites, as ^{13}C labelled compounds are very expensive to purchase, as well as it being difficult to synthesise a wide range of labelled standards for each metabolite due to chemical diversity. Therefore, currently there is reduced availability of commercially available isotopes.[190] To address these issues, isotopically labelled microorganisms could be applied instead as a source of internal standards, because they contain a wide range of polar and non-polar metabolites – this makes them potential internal standards.[190,208] Bacteria, fungi and yeast organisms, for example, are found to have a similar metabolome to human biofluids and tissues[211-214], making them a very effective biological source to label a wide range of metabolites in the samples of interest. Using

a uniformly (U) labelled microorganism in untargeted metabolomics can be a simple, cheap and effective method.

Several bacterial cells such as *Escherichia Coli* (*E. coli)*[189] and *Spirulina*[190], fungi *Fusarium graminearum*[194] and yeast *Pichia pastoris*[215] have already been used as sources of (U)-^{13}C labelled internal standards in untargeted studies, focusing on reducing data variation and improving accurate quantification of intracellular metabolites in plants, wheat or parasites.[190,216,217] In all of these studies, the microorganism was grown in (U)-^{13}C labelled glucose limited culture media, in order to generate successfully labelled extracts. The quality of the normalisation method in each study was assessed by an absolute quantification technique, calculating the ratios between the unlabelled (U)-^{12}C and labelled (U)-^{13}C peaks, and constructing calibration curves to measure absolute concentrations of metabolites. These studies managed to show excellent results with a large number of identified metabolites showing reduced variability between replicates and accurately measured absolute concentrations.[189,190] However, these untargeted studies were not used with clinical human or animal samples. Scientists identified that insufficient research has been performed with (U)-^{13}C labelled microorganisms for normalisation in untargeted human and mammalian metabolomics, thus opening an opportunity for further research. (U)-^{13}C normalisation also has a long history in targeted metabolomics[218,219], flux studies[220] and in vivo metabolic studies[221], all of which were involved with human, animal and other types of living organisms.

When it comes to choosing the most suitable microorganism as a source of internal standards, it is important to choose the one that has the closest resemblance to the mammalian (human) metabolome. Comparing metabolomes of different microorganisms (bacteria, e.g., *E. coli;* yeast, e.g., *Saccharomyces cerevisiae;* fungi, e.g., *Aspergillus genus*)[213,214,222] to the human metabolome (HMDB)[53] using the published research approved databases, one can see that yeast *Saccharomyces cerevisiae* and bacterium *E. coli* have large metabolomes very similar to human, making them potential candidates for a source of internal standards. Both organisms are easy to grow in laboratory conditions, with yeast cultures usually growing faster and to a higher density. Yeast and human organisms are both eukaryotic, while bacteria are prokaryotic. However, to date, more ^{13}C normalisation studies have been performed using bacterial organisms as a source of internal standards, as opposed to yeast. This suggests that bacterial organisms might work very well in untargeted normalisation of mammalian samples, so should be tried first, before testing yeast organisms.

More investigations need to be done with regard to (U)-^{13}C normalisation in untargeted mammalian (human, animal) metabolomics. This will benefit research by showing that a wide range of metabolites in mammalian samples can be normalised by a highly accurate method involving ^{13}C isotope standardisation. Reducing data variability in concentrations of a wide range of metabolites in human and mammalian samples, as well as reducing associated matrix effects, can help scientists to improve data accuracy in large clinical metabolomics studies. This will help metabolomics research in finding potential biomarkers and metabolic pathways in clinical diseases with highly accurate identification.

1.6 General aims and objectives of the thesis

So far, very limited research has been done in mammalian untargeted metabolomics in search of potential biomarkers, where any type of normalisation method including SIL technique was involved in order to help improve the data accuracy. Metabolome data accuracy in this context can be defined as a clear identification of metabolites, accurate quantification and reduction in data variability between analytical sample replicates.

The main issue for that is the struggle to normalise such a wide range of polar and non-polar metabolites, usually found in mammalian samples, by using highly accurate method. Some methods including MSTUS, TIC have been applied in the past in untargeted studies but mainly in environmental, toxicological and pharmaceutical experiments, as mentioned earlier. The use of ^{13}C normalisation and especially the use of bacteria in the method has not been applied in detail in mammalian studies. The main focus of current metabolomics normalisation research is to come up with an accurate method, which would help to normalise most of metabolites in the samples, improving data accuracy significantly. This gap in research has been addressed before in mammalian untargeted studies by using several normalisation techniques such as MSTUS and TIC, however it did not demonstrate a good reduction in data variability. ^{13}C normalisation has never been used before in untargeted mammalian studies, therefore it is a clear opportunity for metabolomics research to apply this method now. Targeted metabolomics using ^{13}C normalisation has already shown a good progress in improving data accuracy but only in specific metabolites, not a wide range.

This thesis, for the first time, presents a potential LC-MS based method, applying (U)-^{13}C isotopically labelled bacterial organism as a source of internal standards, to normalise human and animal samples in untargeted metabolomics. In this thesis, the detailed experimental work

will demonstrate how this method has been developed using an appropriate bacterial organism labelled with ^{13}C isotope, validated using mammalian samples and applied in clinical untargeted studies to improve metabolite data accuracy and to answer the clinical question of the studies.

Chapter two will describe the method used to develop (U)-^{13}C labelled bacterial extract and the LC-MS analytical method. It will compare two bacterial organisms, *E. coli* and *Spirulina*, in terms of their metabolome, extraction optimisation and quality, to determine the suitable source for normalisation. The chosen organism will be uniformly labelled with ^{13}C isotope and tested for isotopic labelling purity. The LC-MS analytical method will also be optimised for an efficient analytical run of samples.

Chapter three will focus on applying the proposed normalisation method on human brain, urine and mouse plasma samples, in order to observe the reduction in statistical variation between biological and analytical replicates through statistical tests – percentage RSD comparison, univariate and multivariate analysis – and to fully validate the proposed method.

Chapters four and five will focus on the application of a validated normalisation method on fully clinical biofluid and tissue based samples, improving the accuracy of metabolome data and helping to search for potential biomarkers and metabolic pathways. The clinical conditions which will be used are: “diabetes” (chapter four) and “low grade glioma cancer” (chapter five).

Chapter six will make the final conclusions on the effectiveness of the proposed method in clinical metabolomics, evaluating the advantages and disadvantages, as well as suggesting the future path for the method.

So, to summarise, the main specific research aims of this work are the following:

1. To develop a novel LC-MS based method for metabolite profiling of mammalian biofluid and tissue based samples using SIL normalisation approach:
    - To develop a (U)-^{13}C labelled bacterial organism as a source of internal standards by choosing an appropriate bacterial organism, extraction and labelling parameters (chapter two)
    - To apply and validate the normalisation method with a small number of mammalian samples to observe the effect of the method (chapter three)

2. To apply the validated normalisation method in clinical mammalian samples in search of potential biomarkers and metabolic pathways (chapters four and five). For each clinical chapter:

- To incorporate the U-^{13}C labelled bacterial extract in animal and human biofluid/tissue samples, run LC-MS analysis and normalisation
- To identify key statistically significant metabolites between control and intervention groups from highly accurate data
- To identify which metabolic pathways were affected during the intervention

In figure 1.3, a schematic workflow diagram shows a typical untargeted metabolomics study used in identification of putative metabolites and potential biomarkers.

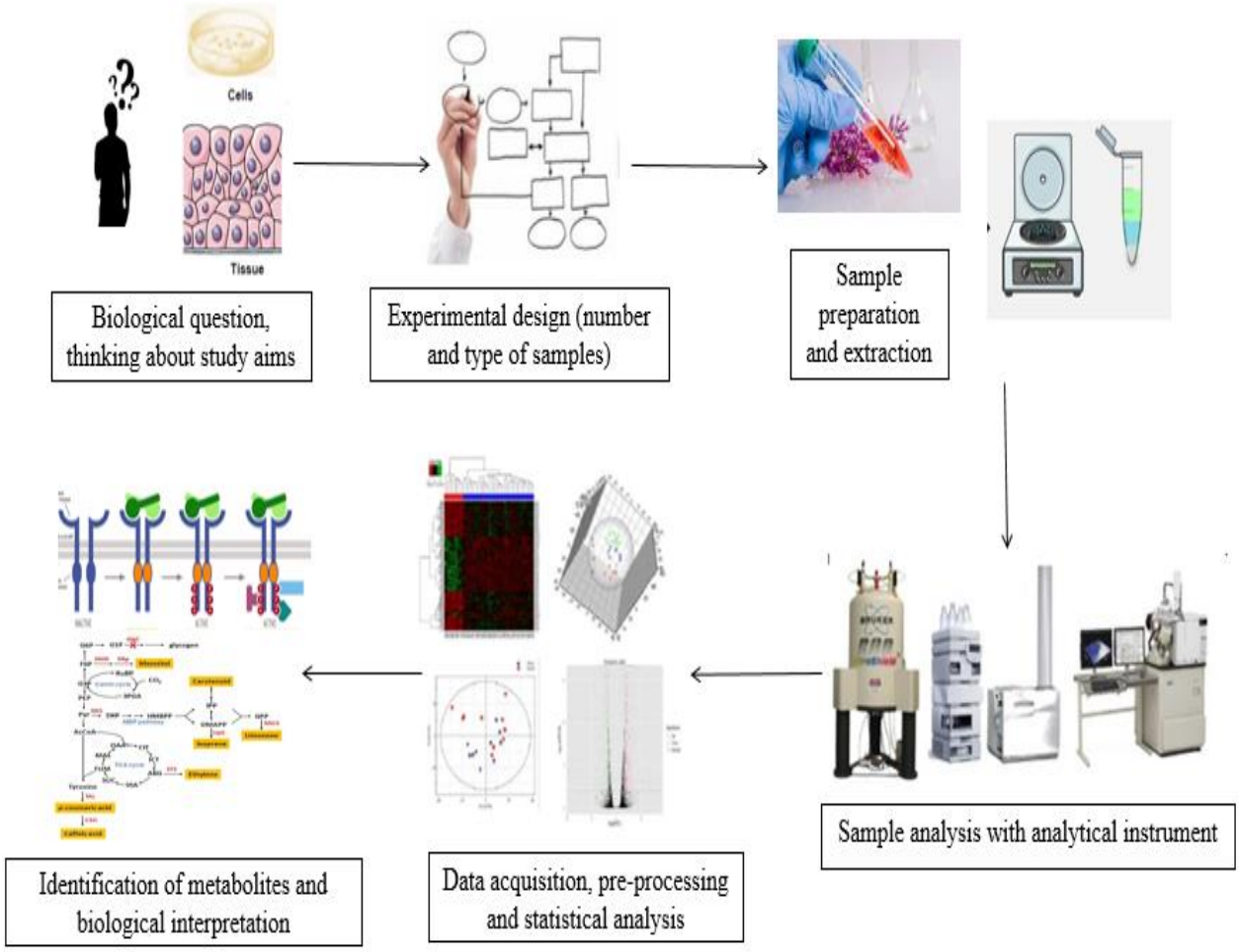


Figure 1.3 Typical untargeted metabolomics workflow in identification of putative metabolites and potential biomarkers (general schematic diagram)

Chapter Two

LC-MS method development of uniformly-^{13}C-labelled organism extract as a source of internal standards for normalisation

2.1 Introduction

2.1.1 Bacterial growth parameters

As proposed in chapter one, a good quality bacterial extract can act as a potential source of a wide range of U-^{13}C labelled metabolites (internal standards), having a similar central metabolome to those normally found in human body fluids and tissues.[223,224] Bacterial extracts acting as isotopically labelled sources of internal standards, have previously been involved in metabolomics and related scientific disciplines, mainly normalising data in targeted studies by improving metabolite quantification and accurately calculating metabolite fluxes.[189,190,195] In several studies, bacterial organisms such as *Escherichia Coli (E. coli)*[177], *Spirulina*[225], *Trypanasoma brucei (T. brucei)*[226] have been shown to grow with certain isotope (^{13}C in most cases) and used later in variety of biochemical, environmental and metabolomics normalisation experiments. The novelty in this thesis, however, is an application of a similarly developed isotopically labelled bacterial extract for the first time in mammalian body fluid and tissue samples, as a means to normalise their metabolome in clinical untargeted studies, for accurate determination of potential biomarkers or statistically significant metabolites between control and disease states, as stated in chapter one.

In order to produce such a high quality source, several considerations need to be taken into account. The first step is to choose the appropriate bacterium that has clear advantages over other bacterial microorganisms, such as having more metabolites in common with human biofluids or tissues, and cost-effective growth and extraction parameters. Once an appropriate bacterium is chosen, the next step would be to grow and label it with a ^{13}C source covering a wide range of metabolites, assess the quality of extract produced and its stability in long-term storage conditions. Already some of the research has been performed in this area, where variety of bacterial extracts were designed for different metabolomics purposes and these considerations had to be taken into account.[227-229] These considerations will ensure the production of an excellent source of internal standards for normalisation. Appropriate bacterial growth conditions achieving high density cultures, suitable cost-effective extraction technique, excellent extract spectrum quality and a large number of detected metabolites in common with the human metabolome, are all important factors when selecting the right bacterium as a source of internal standards for this project.

Suitable growing conditions involve culturing bacteria in the right medium with appropriate nutrients, pH, temperature and length of incubation.[230] Particular types of bacterial culture methods already exist in microbiology and are usually selected based on the agent being

cultured.[231] Broth liquid cultures such as Lysogeny Broth (LB), Terrific Broth (TB), Trypticase Soy Broth (TSB) and 2x YT medium are commonly used to grow and cultivate a variety of *Escherichia coli (E. coli)* strains, *Spirulina*, *Streptobacillus*, *Staphylococcus* and other common bacteria.[232] Therefore, it is important from the beginning to choose the appropriate medium for bacteria to multiply and reproduce efficiently, as each medium contains its own essential macronutrients including carbon, phosphates, nitrogen and organic acids.[233] LB medium, for example, is widely used in microbiology research, as it allows bacteria such as *E. coli* (wide range of strains) to grow fast and in large amounts in laboratory conditions up to a cell density of 1g/1dcw.[234] This is a great advantage over other broth liquid cultures which do not show fast growth of large amounts of bacteria. Another advantage of LB is that it is readily available and can be directly purchased from the suppliers, although sometimes, the chemical nature of certain LB components (tryptone, peptides) is hard to control leading to wide variations of LB media from batch to batch.[235,236] For instance, components such as riboflavin and reducing sugars can generate reactive oxygen species which can hugely decrease the stability of other medium components. However, this issue has been addressed by the addition of catalase, pyruvate and alpha-ketoglutaric acid components to LB medium, which suppress reactive oxygen species, thus enhancing the survival rate of bacterial growth. These agents also reduce precipitation issues which might arise during bacterial incubation.[237] Other methods of bacterial culturing include the use of agar plates and stab cultures, which may be used for bacteria grown at lower or human body (37°C) temperatures but these techniques are used for limited bacterial range and for short-term storage.[238] It is also important to set up the right pH for a good bacterial growth, depending on the species, with the majority of bacterial organisms growing well around pH 6.5 to 7.0. The chosen culture medium for bacterial growth is usually set at a specific pH that is optimal for good growth. For instance, *E. coli* grows best around pH 7 and can even grow in acidic conditions, so LB is an appropriate culture medium for its growth where the pH can be easily adjusted.[239,240] *Spirulina*, on the other hand, requires relatively high pH values between 9.5 to 9.8, and therefore a different culture medium would be applicable. For example, Zarrouk's medium (SM) consists of phosphate, sodium bicarbonate and ammonium nitrate, all of which would help to increase the pH to the optimal value for *Spirulina*.[241,242]

Temperature control during bacterial incubation is also an important parameter, which is chosen depending on the type of bacteria, with some requiring low optimal temperatures (ranging from -20°C to +10°C), moderate temperatures (+20°C to +45°C) and others requiring higher (above 45°C).[243] Most bacteria used in laboratories, such as *E. coli, Spirulina, Salmonella* and *Bacillus subtilis*, are found to adapt to moderate temperatures and are also known as

"mesophiles".[244] The required temperature, carbon dioxide and oxygen levels are normally kept constant in the incubator throughout bacterial growth.[245] The length of incubation, however, is controlled by the user and bacteria can be removed from the incubator at any time. It is important to note that exposing bacteria to oxygen and carbon dioxide for a long period of time e.g. 48 hours, can have a significant inhibiting effect on their growth rate – carbon dioxide is believed to inhibit enzymatic reactions necessary for growth[246,247], while oxygen may react negatively with anaerobic bacteria (*E. coli*, *Staphylococcus* genus, *Spirulina*) causing an organism to die.[248] Therefore, controlling the incubation period of bacteria is of a great importance to avoid a large amount of gases interacting with their growth.

To assess the successful growth of bacteria, several methods are usually employed in microbiology – plate counts that involve counting the number of actively growing bacterial cells in a sample [249], flow cytometry analysis[250], radiometry[251], determination of biomass concentration by dry weight technique[252], as well as measuring optical density (OD) values to follow bacterial population growth in real time.[253] Methods such as plate counting and dry weight biomass determination are found to be time-consuming and not as accurate as other techniques. Calculating bacterial OD can give an estimate of the concentration of bacterial cells grown in a culture solution over a period of time, and perhaps is one of the best tools to monitor the bacterial growth. OD measurements are performed under an assumption that OD value is proportional to the bacterial cell number.[254] OD values are usually taken after the completion of the bacterial incubation stage but can also be taken at any time point.[255] Knowing the concentration from the OD reading can determine the stage of the final cultured bacteria i.e. whether a bacterium is still adapting to growth conditions (lag phase), growing and dividing (exponential phase) or has reached its maximum growing point (stationary phase).[232,256] When a bacterium reaches the stationary phase, the rate of cell growth starts to match the rate of cell death, and it is important at this point to remove the culture from the incubator or stop any further growing process, to prevent the death phase. The decline in bacterial cells usually happens due to growth limiting factors such as depletion of essential nutrients or inhibitory products affecting the growth rate.[257] From previous studies, *E. coli* is expected to grow to an optical density of 1 up to maximum value of 7 at 600nm, where the growth rate reaches saturation.[258] For *Spirulina*, on the other hand, an optical density can reach up to a value of 30 before reaching saturation (stationary phase), as shown in one study.[259] The growing process is schematically shown in figure 2.1. Biosafety precautions when dealing with a variety of pathogenic and non-pathogenic bacteria should always be considered when choosing to work with particular bacteria. Level one bacteria such as *E. coli*, *Spirulina and Bacillus subtilis* are non-

pathogenic and commonly used in microbiological laboratories.[260] Bacterial organisms such as *Staphylococcus aureus, Salmonella, Mycobacterium tuberculosis* belong to higher biosafety levels (two–three), and so are restricted to certain laboratory personnel with permission to access those pathogenic bacteria.[261]

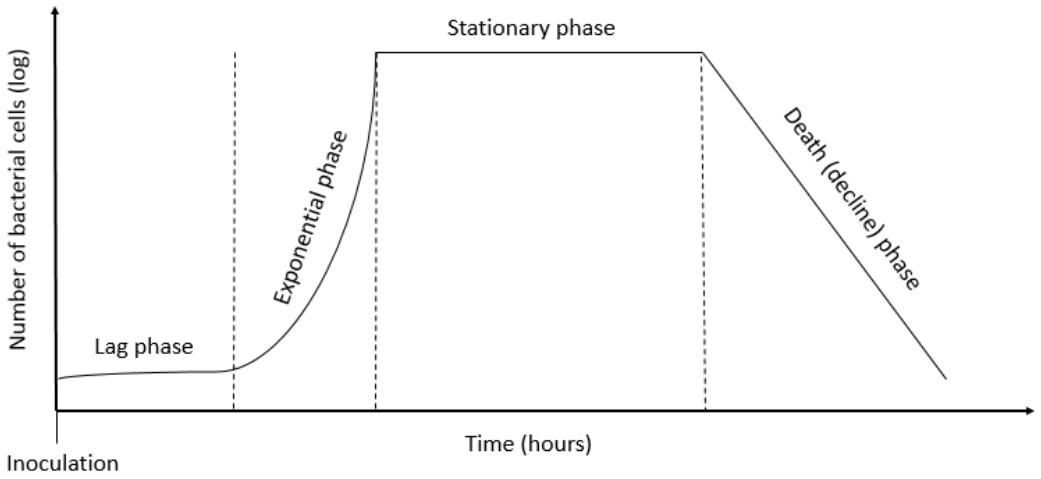


Figure 2.1 Bacterial growth curve over a period of time during incubation of any bacteria

2.1.2 Bacterial extraction and isotope labelling

Depending on bacterial species, a variety of extraction methods exist to obtain polar, non-polar (lipid) and protein compounds. These techniques involve similar extraction solvents, as discussed in chapter one – usually methanol/ethanol for polar metabolites and isopropanol/chloroform for lipid fractions.[262,263] For obtaining intracellular bacterial metabolites, the additional extraction steps may involve the use of liquid nitrogen or heating baths to help initiate cell wall lysis or peptidoglycan destruction (in outer membrane), thus assisting with extraction of intracellular metabolites.[264] Other extraction techniques may involve the use of filtration systems such as pore-size pre-sterilised filter paper to separate bacterial cells from medium, or the use of glass beads and homogeniser to lyse the cells more efficiently.[265] Quenching methods are hugely involved in bacterial extractions nowadays, as they rapidly cool the temperature of culture media, usually using ice-cold methanol. These methods are applied to bacterial cells in order to release intracellular metabolites and inactivate enzymatic reactions.[266] Different bacterial cell types (gram-positive and -negative) have different cell wall compositions, and therefore the amount of leaked intracellular metabolites caused by various quenching methods can differ between organisms, so specific methods exist for each bacterium.[119] In some instances, the released intracellular metabolites may start to

interact with the extracellular environment or can be diluted by medium, so metabolite detection by analytical instruments can be affected. Therefore, it is important to note that any of the mentioned extraction and quenching techniques can be destructive to the actual bacterial metabolome, so the techniques need to be carefully optimised for a particular bacterium. The process of bacterial cell wall lysis and the release of intracellular metabolites is shown in figure 2.2. For *E. coli* bacterium it was found that the best extraction approaches were to either use ice-cold (-40°C) methanol followed by a freeze-thaw cycle (liquid nitrogen, thawing on ice and rapid mixing), or apply hot ethanol with vacuum centrifugation. Cold methanol can permeabilise the bacterial cell wall for metabolites access, as well as rapidly quenching enzymatic changes in metabolites.[267] Extractions using strong acids or alkali (NaOH or KOH), on the other hand, indicated low extraction performance resulting in destruction of important metabolites.[268] In one *E. coli* study, various extraction methods were compared to find the most suitable technique to extract the largest number of metabolites. Extraction with cold (-40°C) methanol has shown to generate a good quality bacterial extract, showing high resolution for more than 95 metabolites after chromatographic analysis, while other extraction techniques including acidic and alkaline treatment generated poorer quality of extract.[269] *Spirulina* bacteria are normally extracted using a solvent extraction technique (ethanol and water) or using "superficial fluid extraction" with carbon dioxide and ethanol. This latter technique has been showing clear advantages over other methods as it was easier and faster than the solvent extraction technique for isolating compounds.[241,270]

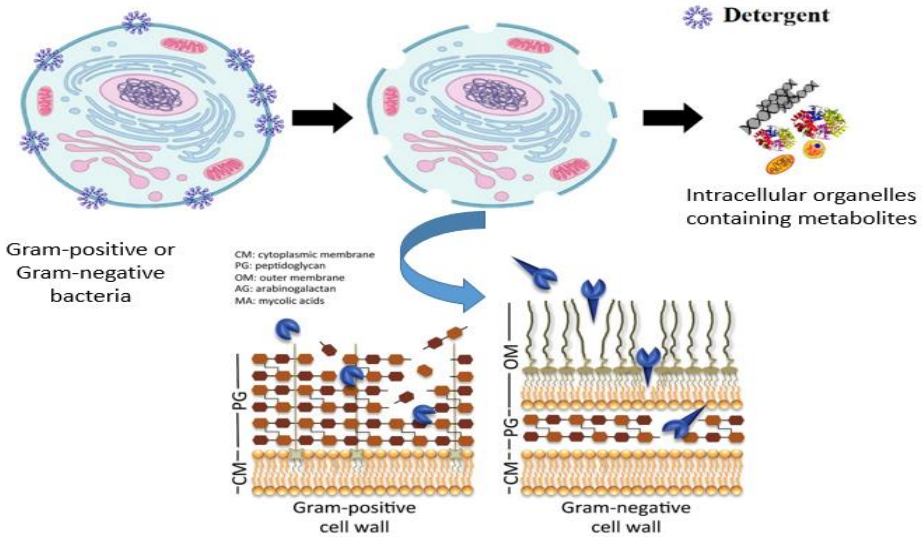


Figure 2.2 Bacterial cell wall lysis and the release of intracellular metabolites during extraction and quenching techniques (detergent referred to in this image is extraction/quenching solvent)

When it comes to ^{13}C-labelling of a bacterium, it is important to introduce ^{13}C isotopes at the beginning of bacterial growth, as part of an ingredient in the medium, so that isotopes would be fully incorporated in the formed bacterial metabolites.[271,272] As described in chapter one, so far, isotopically labelled bacteria have been used in quantitative studies to accurately measure the concentration of certain metabolites, where the extracts were successfully labelled with isotopes. In one study, uniformly ^{13}C-labelled *E. coli* extract was used for large scale quantification of metabolite levels in human African *Trypanosoma brucei* species[189], whereas another study involved ^{13}C-labelled *Spirulina* to quantify intracellular metabolites in *Clostridium autoethanogenum* species.[190] In order to incorporate ^{13}C isotopes into a bacterial metabolome, the right medium has to be chosen. In the case of growing *E. coli* with ^{13}C isotopically labelled glucose, for example, it has been shown through several studies that an M9 minimal medium is the most suitable one for generating a wide range of labelled metabolites, having a limited amount of ^{12}C isotopes in its ingredients. This medium is usually chosen to avoid a huge incorporation of ^{12}C isotopes and to ensure that the majority of metabolites would be labelled with ^{13}C isotope instead.[189,273] However, the growth rate for *E. coli* in minimal medium is generally slower compared to other media. Growing bacterial cell mass first on unlabelled rich media, such as LB, allows rapid growth of bacteria achieving a higher OD. Following growth in unlabelled medium, bacterial cells can then be exchanged into isotopically defined minimal media for isotopic incorporation.[274] This procedure using two media (unlabelled, followed by ^{13}C-labelled) has been successfully demonstrated in a study of labelling recombinant proteins with ^{13}C isotope, achieving a higher yield of labelled proteins (threefold) compared to the yield when only M9 minimal medium has been employed.[275] Therefore, to ensure a good growth rate for bacteria with high OD and a wide range of metabolites isotopically labelled, using both unlabelled (LB) and minimal isotopically labelled media together could be the right step.

2.1.3 Justification of the method development study

The overall aim of this chapter was to develop an LC-MS based method using high quality U-^{13}C-labelled bacterial extract for further normalisation studies presented later in this thesis. Two bacterial species have been chosen for comparative analysis to determine the best extract: *E. coli MG1655* and lyophilised *Spirulina*. Both of these bacterial organisms were chosen for a variety of reasons: their regular use in microbiology research and related scientific disciplines; their non-pathogenic profile (biosafety level 1); quick availability in laboratories or easy to

purchase and grow; and both were successfully labelled with ^{13}C isotopes before, showing promising results.[189,190]

E. coli and *Spirulina* have been consistently analysed in metabolomics studies, for medical and pharmaceutical research, over several years.[276,277] Variety of E. coli strains have been investigated to design antibiotics[277], treatment interventions and understand the pathogenesis of the bacteria.[278] Spirulina has been used in clinical practice for a long time, with multiple studies showing its efficacy in treating several diseases and shown to have anticancer, antiviral and antiallergic effects.[279-281] As mentioned previously in chapter one, both bacteria have been used in metabolomics, particularly for normalisation studies, proving a uniformly ^{13}C labelled extracts. These studies have demonstrated the effectiveness of normalisation and the usefulness of bacteria.[189,190]

E. coli bacteria will be grown in microbiological conditions using already established protocol with chosen optimised growing conditions mentioned above: in LB medium first, followed by M9 minimal medium over the period of 17 hours at 37°C to reach an OD of 1.0 or above.[269,282] *Spirulina* bacteria will be purchased directly from a manufacturer, as it is an easier and faster process to obtain readily available bacteria than growing it in the laboratory. Before ^{13}C-labelling will take place, both bacteria will be analysed in an unlabelled environment (without ^{13}C isotope), to decide which bacterium could be labelled later – to prevent expensive unnecessary experiments. Details of *E. coli* growth is described fully in the Methods section.

As described earlier, these bacteria will require specific extraction and quenching methods, some of which will be tested and optimised to develop high quality extracts with large numbers of detected metabolites. Techniques will involve the use of liquid nitrogen, ice-cold methanol or acetonitrile solvents for polar metabolite extraction, centrifugation, vacuum centrifugation, and hot water baths to aid bacterial cell wall lysis and release of intracellular metabolites, all of which are mentioned above and applicable for *E. coli*[269] and *Spirulina*.[283] Details of those procedures are described in the Methods section.

In addition to bacterial labelled extract development, the LC-MS analytical method has also been optimised in this chapter. Two LC-MS methodologies with different gradient elution programmes (18 and 24 minutes running time) were compared to each other to determine the best optimum method for use in future experiments, in order to obtain a strong peak signal response and detection of a wide range of metabolites. Both proposed methods are similar in mobile phase composition, column chromatography, MS conditions and type of gradient, and have been previously used in metabolomics analysis in several studies.[189,284] Overall analytical

running time and percentage of mobile phase composition will be different in the two proposed methods. Details are described in the Methods section.

2.1.4 Aims and objectives

The following aims and objectives describe what has to be done to develop such a method:

1. Choose the appropriate bacteria for isotopic labelling

- Grow unlabelled *E. coli* bacteria using the established protocol
- Optimise the extraction parameters for *E. coli* bacteria and perform LC-MS analysis
- Optimise the extraction parameters for *Spirulina* bacteria and perform LC-MS analysis
- Observe the spectra quality of both bacteria. Perform data processing using IDEOM
- Compare the detected metabolomes of both bacteria with the HMDB to determine the appropriate bacterium

2. Generate ^{13}C-labelled bacteria extract

- Grow the chosen bacteria with the ^{13}C-labelled source (glucose), extract it with chosen optimised parameters and perform LC-MS analysis (repeating steps in Aim 1, above)
- Test the quality of labelled extract by measuring the amount of labelled metabolites

3. Develop an efficient cost-effective LC-MS method with optimised conditions

- Compare two analytical LC-MS methods with different flow gradients and total running times (18 vs 24 minute methods) by running standard group mixtures of known metabolites, and observe the quality of spectra and number of detected metabolites
- Run a chosen (unlabelled) bacterial extract with the chosen LC-MS method to fully validate it

4. Generate a series of ^{13}C-labelled extracts and observe their quality consistency (performed throughout the PhD project).

2.2 Materials and Methods

2.2.1 Materials, chemicals and sample preparation

Materials and chemicals

Mobile phase solvents used for this study (to wash zic-pHilic column, needle and other LC parts) were 20mM of ammonium carbonate in purified water (A) and HPLC-grade acetonitrile (B), both purchased from "Sigma-Aldrich", Gillingham. Solvents used for extraction of bacteria were methanol, acetonitrile ("Sigma-Aldrich") and a mixture of acetonitrile, methanol and purified water (AMW, 2:2:1 ratio) prepared in the laboratory. Positive and negative MS calibration solutions "Calmix" (mixtures of glycine, acetic acid, caffeine and ultramark) have been prepared by analytical group in the laboratory and stored in a -20°C fridge. Zic-pHilic column (4.6x150 mm, 5µm particle size) was purchased from "Merck SeQuant" and tested by analytical group.

^{12}C *Spirulina* whole cells (lyophilised powder), unlabelled, were purchased from "CK Isotopes Limited" Leicestershire.

E coli MG1655 strain was initially cultured on agar plate by the bacteriology group laboratories, Centre of Biomolecular Sciences. The growth of *E coli* MG1655 was supported by LB culture medium (prepared by the same department) and M9 culture medium (prepared by mixing 100mL of 5xM9 salts, 1mL of 1M magnesium sulphate, 1mL of 1M calcium chloride, 3mL of 2M thiamine, 5mL of either ^{13}C-labelled or unlabelled glucose 20% solution, 445mL aseptically purified water per litre). All M9 minimum medium chemicals were ordered from "Sigma Aldrich", with the exception of purified water (provided by the bacteriology group) and uniformly ^{13}C-labelled (U-^{13}C6, 99%) glucose (ordered from "CK Isotopes Limited", Cambridge).

A set of 268 authentic standards (10mM and 100mM concentrations of amino acids, carbohydrates and peptides) were used in an analytical run of the samples and were prepared from chemicals ordered from "Sigma-Aldrich" (see Appendix A1 table).

The samples and standards were stored in Eppendorf tubes, and later transferred into 200µL thread vials ("Chromatography Direct") for the LC-MS run.

Growth of *E. coli* MG1655 bacteria (unlabelled and ^{13}C-labelled)

Over the period of time, several *E. coli* extracts (first unlabelled, later ^{13}C-labelled) were produced in the same conditions, with the growing procedures taking place in the bacteriology group in the Centre of Biomolecular Sciences, University of Nottingham. Each produced extract had a volume of 10mL, as the equipment for growing bacteria was limited, so only a small amount of extract could be produced at a time. In order to produce 10mL of unlabelled *E. coli* extract, several steps were performed. Bacteria was first taken from agar plate and inoculated into a small amount (3mL) of LB medium in a tube, which was left overnight in an "Allegra" incubator at 37◦C. 500µL of the grown *E. coli* was then mixed with 500mL of M9 minimal medium in several flasks under aseptic conditions (using a standard Bunsen burner technique) and put into the same incubator for around 17 hours. The cultures in the flasks were allowed to reach a density of 1x10^6 cells/mL or more, measured by optical density reader "Biomate" set at 600nm wavelength. Bacterial cells were then centrifuged at 3000g, 4°C for 10 minutes ("Harrier" 18/80 centrifuge) to obtain the bacterial pellets for extraction. Labelled ^{13}C *E. coli* were grown under the same conditions, with 20% U-^{13}C glucose instead of unlabelled glucose.[189]

***E. coli* MG1655 extraction optimisation**

Several extraction methods were tested on *E. coli* pellets during this study, in order to decide the most effective technique in obtaining the bacterial extract with wider metabolite coverage. Extraction tests were done using a "freeze-thaw" procedure, combinations of "freeze-thaw" and "vacuum drying" procedures (drying extraction), as well as using different extraction solvents such as acetonitrile: methanol: purified water (AMW) in 2:2:1 ratio, or pure methanol, in various volumes. The "freeze-thaw" procedure involved freezing the *E. coli* pellet solutions in liquid nitrogen for one minute (extra safety precautions had to be considered to avoid the hazard of liquid nitrogen), defrosting (thawing) them on the surface of the ice and vortex mixing them for 30 seconds. The "freeze-thaw" procedure was repeated at least four times to ensure the bacterial cell wall had been broken in the *E. coli* pellets. Mixtures were then centrifuged using the "Harrier 18/80" centrifuge for ten minutes at 13000g at 4◦C to obtain the supernatants (upper layer solutions).[189] The extra steps in the drying extraction method involved adding cool extraction solvent into the tubes again (according to the optical densities of bacteria), repeating the "freeze-thaw" procedure four times, centrifuging the tubes in similar conditions described above, as well as putting the tubes into a centrifugal evaporator ("Joan") for several hours to evaporate the solvents and dry the *E. coli* pellets. These dried pellets were later reconstituted with extraction solvent before LC-MS analysis. The extraction of ^{13}C-labelled *E. coli* extracts was

performed using several extraction methods producing various samples (described later in Results and Discussion sections). All the tubes with *E. coli* extracts (unlabelled or labelled) were stored at -80◦C until LC-MS analysis.

The first set of experiments was performed to observe if drying extraction conditions (freeze-thaw with vacuum drying using the centrifugal evaporator) could be of additional help to a standard freeze-thaw (non-drying) extraction of *E. coli*. For this experiment, two sets of bacterial cultures (500µL of bacteria in 500mL media volume) were grown and each was extracted by a different method (either drying or non-drying). Two sets of samples were produced and analysed by LC-MS – five analytical replicates of dried extract and another five of non-dried. The extraction solvent used in this experiment was ice cold methanol; the same volume (4mL) was applied to all bacterial extracts in the falcon tubes. Figure 2.3 shows a schematic drawing of the sample preparation.

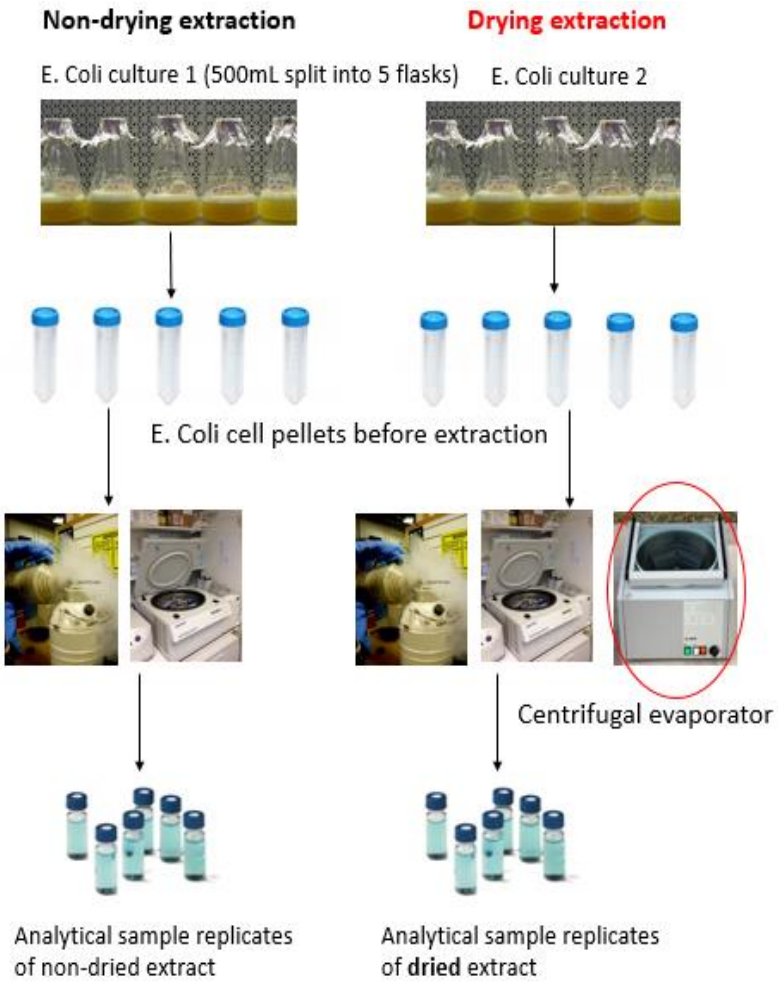


Figure 2.3 *E. coli* sample preparation using two extraction techniques in the first set of experiments

The second set of experiments of *E. coli* extraction looked at the solvent choice of AMW (2:2:1 ratio) or pure methanol, and several volumes, to find efficient extraction parameters. AMW and methanol were chosen, as they seemed to be appropriate for extraction of polar metabolites and have been regularly used in the past.[285] The volumes tested were 2mL, 3mL and 4mL – the reason for this choice was that a 4mL volume of methanol solvent, for example, has already shown some promising results in previous experiments, so using even smaller volumes (resulting in more highly concentrated *E. coli* extracts) could generate better quality extract. A small volume of *E. coli* (100µL) has been used for each extraction experiment, instead of 500µL, because a variety of extraction techniques were tested simultaneously, so growing large quantities of bacteria for each experiment would have been too much. Twelve types of *E. coli* extracts were generated: non-dried extracted in 2mL, 3mL and 4mL volumes of **AMW** solvent; non-dried extracted in 2mL, 3mL and 4mL volumes of **methanol** solvent; **dried** extracted in 2mL, 3mL and 4mL volumes of **AMW** solvent; and **dried** extracted in 2mL, 3mL and 4mL volumes of **methanol** solvent (with three analytical replicates for each type).

The third set of experiments for *E. coli* extraction was testing the chosen extraction parameters from the first two sets by growing a large amount of bacteria (500 µL) in volume.

^{12}C *Spirulina* extraction optimisation

In order to extract metabolites from the unlabelled whole cells *Spirulina* powder in methanol or acetonitrile solvents, various techniques were applied. One technique involved a simple mix of a weighed amount of *Spirulina* powder (10mg) in either 2mL, 3mL or 4mL of ice-cold methanol or acetonitrile solvents in volumetric vials using vortex and shaking, before pipetting different volumes of mixed solution into 200µL thread vials for LC-MS analysis (creating several replicates). The optimised, improved technique tested several concentrations of *Spirulina* solutions in ice-cold methanol and acetonitrile solvents (10mg/2mL, 10mg/3mL, 10mg/4mL) and more thorough solubilisation extraction: the samples were vortex mixed for 30 seconds, put into a ball mill "Retsch MM301" at 25.0 1/s speed for three minutes, put into a water bath at 45◦C for 20 minutes, then centrifuged ("Harrier") at 4200rpm for three minutes and vortexed vigorously again several times. The supernatants were then pipetted into fresh tubes, and later transferred into 200µL thread vials. The procedures followed a recommended FDA report on *Spirulina* whole cells solubilisation.[241] All samples from different extraction techniques were stored at -80◦C and were analysed with the LC-MS instrument.

2.2.2 Analytical LC-MS methodology

The LC-MS instrument used in this study to run the bacterial extract samples was a combination of Orbitrap ESI-MS Exactive coupled with Accela autosampler LC (Thermo Fisher Scientific). Bacterial samples and standards were transferred into 200µL thread vials specific for LC-MS runs and filled up to the volume before being placed on autosampler for a run. *E. coli* and *Spirulina* optimisation extract samples were run with the 24-minute methodology.[189] Once a suitable bacterial organism was chosen, it was then run on the 18-minute method for full validation. Results of the two methods were later compared.

Chromatography was carried out using a ZIC-pHILIC column (Merck SeQuant), which was maintained at 30◦C throughout the run. In the **24-minute method,** samples were eluted with a linear gradient decreasing from 80% mobile phase B to 5% within 15 minutes, followed by an increase of gradient back to 80% within two minutes and maintenance until the 24th minute, with a flow rate of 300µL per minute. In the **18-minute method**, the linear gradient elution decreased from 80% mobile phase B to 5% within the first 12 minutes, followed by a rapid increase back to 80% within a minute and maintenance until the 18th minute, with the same flow rate. The injection volume from a sample was 10µL and samples were maintained at 4°C inside the autosampler.

MS "Exactive" was operating in both positive and negative polarity switching modes throughout the run. Electrospray ionisation (ESI) was the type of ionisation used in the instrument. The following settings were applied: resolution 50,000; mass-to-charge (m/z) ratio 70-1400; capillary voltage 40V; capillary temperature 275°C; sheath gas 40 units. Mass calibration has been applied before the run, for each polarity, using "Calmix" preparation (positive and negative mode mixtures of glycine, ultramark, caffeine), covering low molecular weight molecules up to 70m/z.

Analytical details:	**Methodology:**
Instrument:	Exactive Orbitrap ESI-MS and Accela Autosampler LC
Calibration mixture:	"Calmix" (positive and negative)
Polarity:	Switching (positive and negative)
Mass resolution:	500,000 FWHM at m/z 200
ACG target:	Balanced
Max inject time (seconds):	250
Low mass limit (m/Z):	70
High mass limit (m/Z):	1,400
Spray voltage (kV)	4.5
Capillary voltage (V)	40
Tube voltage (V)	70
Skimmer voltage (V)	20
HESI temp (◦C)	150
Capillary temp (◦C)	275
Sheath gas (units)	40
Aux gas (units)	5
Sweep gas (units)	1
Autosampler temp (◦C)	4
Injection volume (µL)	10
LC Solvent A (Aqueous phase)	20mM Ammonium Carbonate in purified water
LC Solvent B (Organic phase)	HPLC-grade Acetonitrile 99%
LC-MS column:	ZIC-pHILIC
Column diameter (mm)	4.6
Column length (mm)	150
Pore size (µm)	5
Column temp (◦C)	45
Flow rate (µL/min)	300

Table 2.1 Methodology parameters chosen for all LC-MS experiments throughout the study.

The calibration 268 standards were prepared by diluting stock solutions of various unlabelled metabolites (including common amino acids, organic acids and nucleotides – similar to the ones found in human biofluids) to the appropriate concentrations with water or methanol. The standards were split into five main groups according to their isomers, physicochemical properties and masses, so that metabolites of similar chemical properties would not overlap and affect the peak response on the spectra (see Appendix A1 table). The standard mixtures samples were analysed together with *Spirulina* and then separately with *E. coli* samples. Methanol was

used as a reagent blank in every analytical run. Table 2.1 (above) summarises the LC-MS conditions which were applied throughout the analysis.

2.2.3 Data processing and analysis

The raw data from all the test samples (in the "Xcalibur" format) were pre-processed with "XCMS" software for untargeted peak-picking[144], followed by conversion into "mzXML" and "mzMatch" formats for peak matching and annotation of related peaks.[141] "IDEOM" software (Welcome Trust Centre for Molecular Parasitology, University of Glasgow) was chosen as a main data pre-processing and processing software to filter the background noise and putatively identify metabolites.[147] The programme selected the most relevant metabolites and rejected the ones with poor signal or unmatched retention times, reducing the general noise background. Positive and negative modes of data were also combined to see the full presence of all metabolites in each sample. Statistical analysis was applied to the processed data to improve its accuracy and significance.[142] Metabolite annotation in "IDEOM" was achieved by matching accurate masses and retention times of observed peaks in the data to metabolites in the included database, inside IDEOM programme, which incorporated all likely metabolites from wide range of biological databases. Metabolite annotation was also done by matching peaks to authentic standards but not for all metabolites. The final lists of identified and rejected peaks were annotated with confidence scores regarding the identification of each metabolite, based on retention times and molecular masses – from score 1 (least confident) to 10 (most one). If measured retention time was within 5% of predicted one (either from database or authentic standard), the confidence score would be given as 9 or 10. If measured retention time was within 35% or more, the confidence score would be 7 or less. Metabolites with confidence scores of 5 or above would be moved to the list of identified metabolites, while the ones with lower scores would be rejected. "Related peaks" were also annotated by IDEOM as possible ESI artefacts, isotopes, adducts, fragments based on their retention times, peak shape and correlation of peak intensities. Raw data from standards were processed alongside, converting raw "XCalibur" data into "TOX-ID" files and uploading them into the "IDEOM" programme to create a calibration graph.[286]

Unlabelled *E. coli* and *Spirulina* processed metabolome data were then compared to the existing human metabolome database (HMDB)[53] to see which bacterium provides the most number of metabolites in common with human biofluids (urine, blood, saliva). The spectra quality of each bacterium sample has also been compared. It was decided at this point that whichever extract

provides the highest number of putative metabolites was enough to make a decision on which bacterial organism to use. The chemical nature of metabolites detected was not as important as the total number of metabolites in the extract overall (preference was given to more hydrophilic compounds). Most important aspect of choosing an extract is for its rich wide range metabolome, most of which matches with metabolites from HMDB. *E. coli* metabolome database, EMDB, has been used as well, to verify that detected metabolites in *E. coli* extract was correct and indeed belonged to *E. coli* metabolism, but otherwise the EMDB was not used in data processing – when comparing metabolomes of *E. coli* and Spirulina. As the analysis is untargeted and there is no aim of detecting particular metabolites or classes of metabolites, the use of EMDB and HMDB have only be used for direct comparisons. Attention to isomers and other physicochemical properties was not very significant.

For U-^{13}C-labelled *E. coli* data, additional processing steps have been applied, using "mzMatch-ISO" function and "R" software programme. Data generated from "IDEOM" were passed through these programmes to see which metabolites were successfully labelled. PDF and Excel files were created showing the chromatogram graphs of isotope peaks of labelled metabolites, the ratios of unlabelled and ^{13}C-labelled peaks indicating the fraction of labelling efficiency, as well as the actual peak height intensity values for unlabelled and labelled peaks to aid the statistical analysis.[287]

Data pre-scaling has been achieved through pre-processing stages using IDEOM software, where unwanted signals, background noise, artefacts have been removed, as well as scaling the data in order to put the samples and variables on a comparable scale. In the next chapter, normalisation stage contributed to pre-scaling of the data. Data pre-scaling was important to minimise variable redundancy and make all variables more comparable in size.[288]

2.3 Results and Discussion

2.3.1 Optimisation of the metabolite extraction protocol for *E. coli*

A series of metabolite profiling experiments have been performed on unlabelled *E. coli* extracts, which were produced (extracted) under different conditions (see methods section for more details). The aim was to determine the optimised extraction method for *E. coli* to generate an extract with the highest number of putative metabolites and excellent peak signal response.

First set of experiments: comparing drying and non-drying technique

The optical density of bacteria in each sample replicate (before extraction) was observed to be very close to 1.0 or above, indicating a good growth. After extraction and LC-MS analysis, IDEOM results revealed totals of 513 and 588 putative metabolites detected in unlabelled dried and non-dried *E. coli* extracts, respectively, showing similar results (see figure 2.4 a and b). In both extracts, the majority of putative metabolites detected were of amino acid metabolism, followed by carbohydrates and "others" (polar metabolites). When comparing metabolomes of non-dried and dried extracts, there was a lot of similarity in the content and nature of metabolites, but it was noticed that in non-dried conditions the number of putative metabolites was found to be greater (mainly amino acids and carbohydrates). Also, the peak signal response was observed to be higher in non-dried extract replicates (TICs of E^8 for all of the samples), compared to dried extract replicates (TICs of E^7 or even E^6) which showed fewer peaks in their spectra. This suggests that the non-drying extraction technique presented slightly better results than drying, with a larger number of detected metabolomes, as well as greater signal response for metabolite peaks. The drying technique generally presents an issue of sample loss during the evaporation of a solvent, hence it could have led to potential loss of some important metabolites in this experiment. It is also a time-consuming process. From this experiment it could be seen that non-drying extraction works better for *E. coli* than the drying technique. However, the overall number of putative metabolites was lower than expected (in previous *E. coli* metabolite profiling studies there were 700–800 metabolites), so there was a need to look into further extraction parameters such as solvent choice and volumes being used.

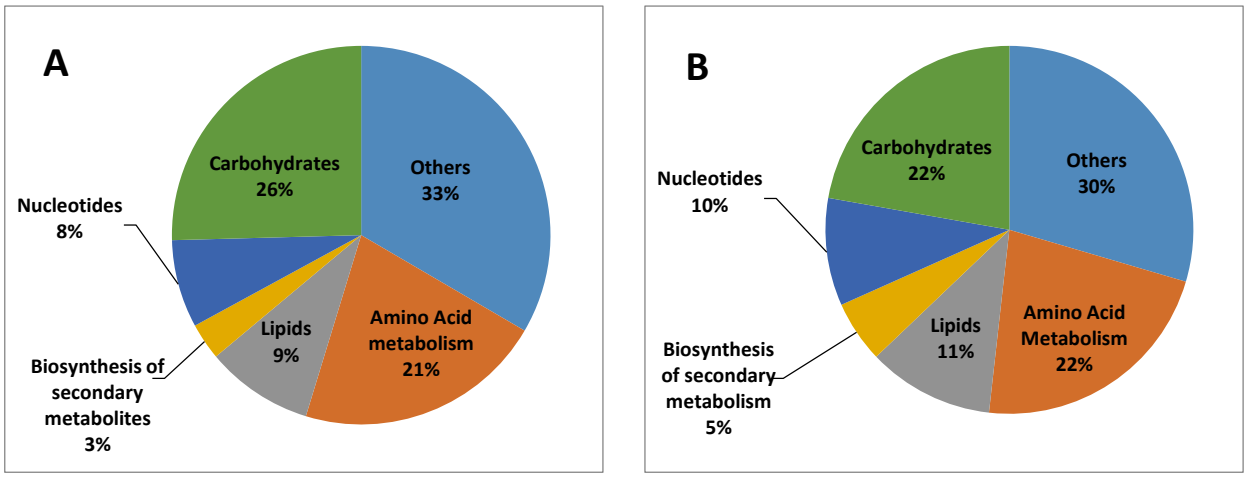


Figure 2.4. The metabolite composition of *E. coli* extracts in drying extraction (A) and non-drying (B). Percentages represent the number of metabolites, split by biochemical class.

Second set of experiments: optimisation of volume and solvent

The small volumes of bacteria in each flask had an optical density of 1.0 or above. The results, in the form of total number of detected putative metabolites for each type of extract, are presented in figure 2.5. The highest number was observed in both non-dried extracts by 2mL volume of methanol and AMW solvents (547 in methanol and 504 in AMW). Compared to the other volumes (3mL and 4mL), it seems that extraction in 2mL of either solvent in non-drying conditions produced a better extract with slightly higher number of detected metabolites. This can be explained that the smaller the volume of extraction solvent used, the higher the concentration of an extract was produced, hence a higher number of putatively identified metabolites. However, the reduction in number of metabolites was not observed to be big when extraction solvent was increased. The possible explanation for this could have been due to a small amount of extraction volumes used. By increasing extraction solvent by 1mL might not have been enough to make bacterial intracellular metabolites become highly diluted in extracellular medium (solvent). Even with 4mL volume used, the extract's concentration would still be very high and so a high number of metabolites were detected.[119] The choice of solvent did not make a huge difference, with methanol showing slightly better results than AMW solvent, possibly because it was not diluted by purified water as is the case with AMW solvent. Similar results were shown in drying conditions, where the 2mL volume extraction of methanol and AMW solvents showed the highest numbers (463 and 455, respectively) compared to the other volumes. The peak signal response for non-dried extracts (both solvents, any volume) was relatively high (E^8), compared to dried extracts which showed a poorer quality of spectra. This suggests that the most effective extraction parameters for *E. coli* would be the non-drying

technique, with a 2ml volume of methanol solvent. To confirm these chosen extraction parameters, a large volume of *E. coli* (500µL) was tested next.

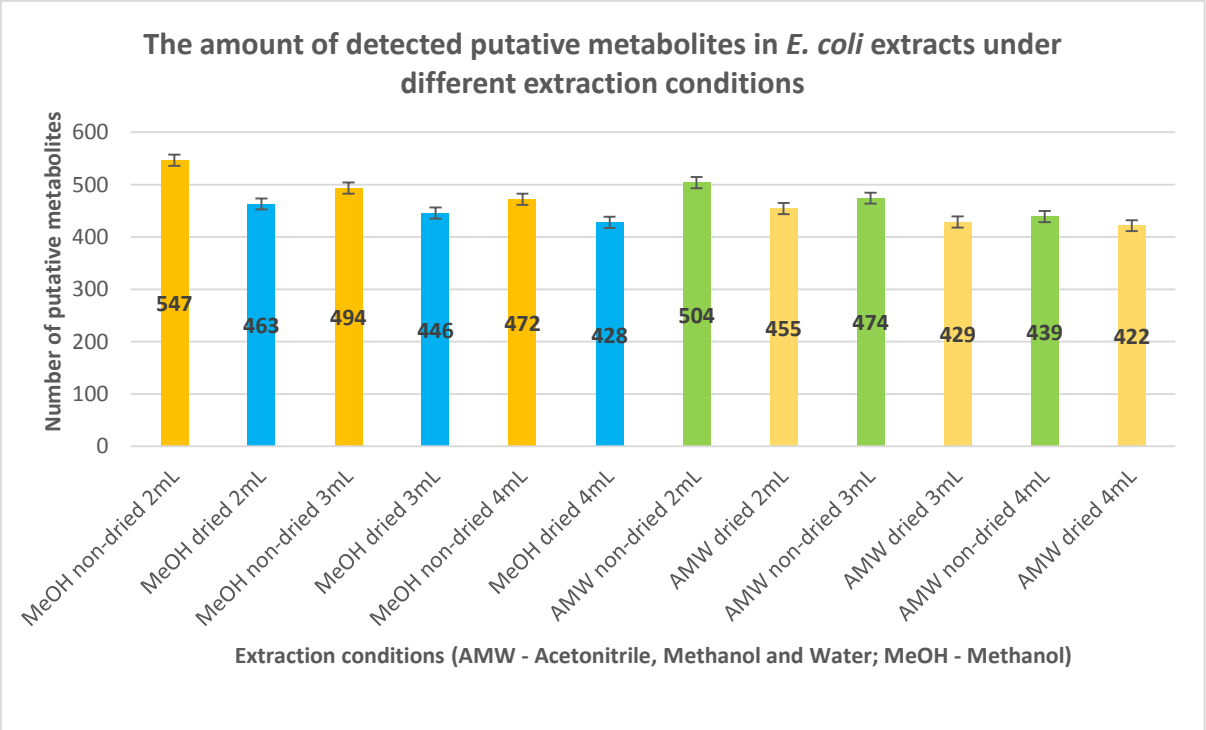


Figure 2.5. The number of putatively identified metabolites in *E. coli* extracts produced by various extraction techniques

Third set of experiments: large scale growth of *E. coli*

The chosen extraction parameters were later applied on larger-scale *E. coli* (500µL inoculums), with 10mL of ice-cold methanol solvent used (proportional to the earlier experiment where 2mL of solvent was applied for 100µL bacterium), to see if comparable results can be achieved than the small-scale optimisation experiment. This large-scale experiment was repeated two times, growing fresh bacteria for each experiment, to see if we could get consistent results in metabolome and spectra quality of grown and extracted *E. coli* under the chosen optimised methods. In the first experiment, bacteria were grown in five small 500mL flasks, while in the second experiment all of the bacteria were grown in a large 1L flask. (All other conditions stayed the same.) This was done to see whether growing bacteria in several flasks separately can produce a better growth rate (OD). Figure 2.6 presents the metabolomes of *E. coli* in those experiments (763 and 787 metabolites). The majority of detected metabolites in both extracts were of amino acid, carbohydrate and "other" biochemical classes. It was also observed, in those experiments, that growing large-scale *E. coli* in several 500mL volumetric flasks, rather than

inside a large 1L volume flask, significantly improved the growth of bacteria, and therefore the production of extract – a higher OD value was reported in the first experiment. This suggests that growing bacteria in several flasks provided them with a larger surface area for growth and expansion – in a 500mL volume flask, it will be a lot easier for a small bacterial culture to grow, than for 500µL of bacteria in a 1L volume flask (because it will be more difficult for the bacteria to be mixed with the culture). The peak signal response of samples was observed to be very high (TICs ≥ E^8) in both experiments – see figure 2.7. These final experiments confirmed that the chosen extraction parameters were appropriate for *E. coli*.

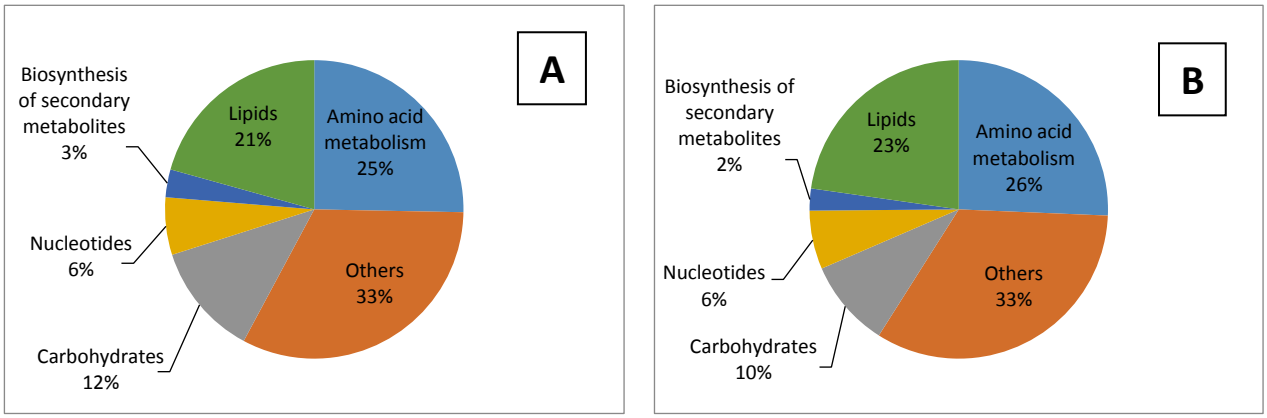


Figure 2.6. Metabolite profiling of large-scale *E. coli* (500µL) in two experiments. Percentages represent the amount of metabolites split by biochemical class. First experiment (A) Second experiment (B).

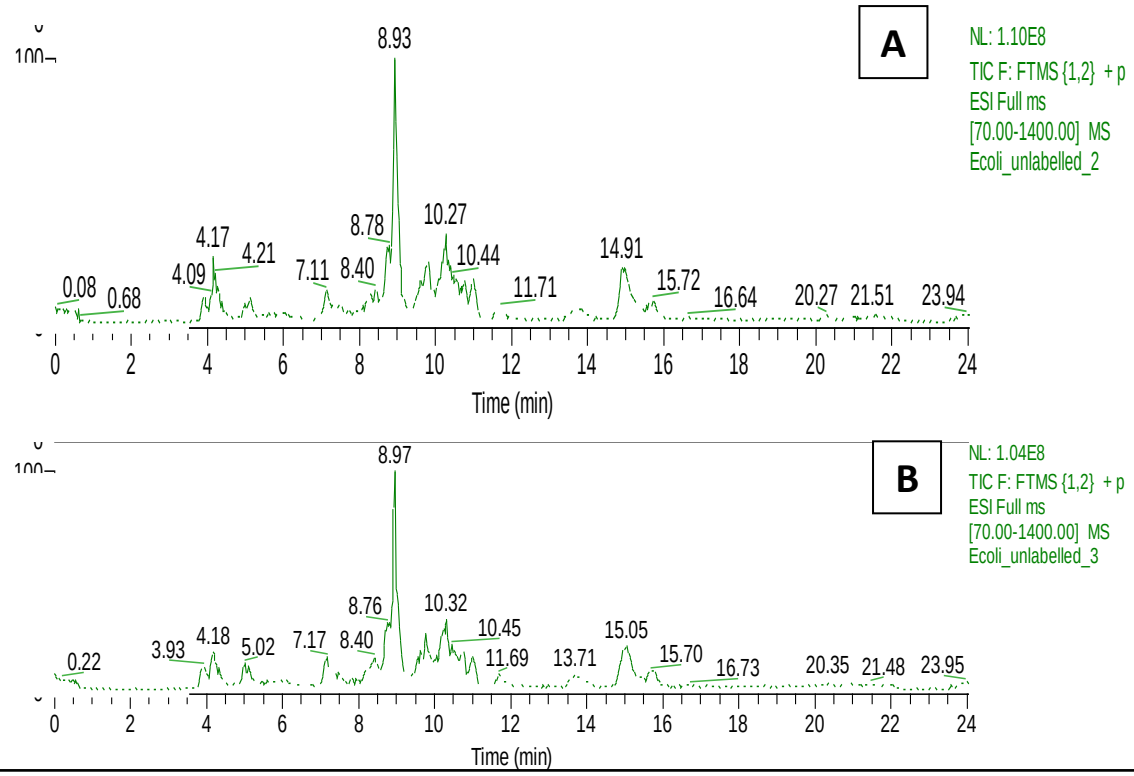


Figure 2.7. Example of two spectra of *E. coli* 500µL/10mL sample analysed by "XCalibur" software showing metabolite peaks during a 24-minute run. The signal response (TIC) in both is recorded as 1.13 and 1.04 E^8. (A) First experiment (B) Second experiment

2.3.2 Optimisation of the metabolite extraction protocol for *Spirulina*

Several untargeted LC-MS Exactive experiments were performed on generated *Spirulina* extracts to identify the suitable extraction conditions, and the results were compared to those of *E. coli*.

First set of optimisation experiments: Volume test

Firstly, it was decided to see if *Spirulina* (10mg) could be extracted with either ice-cold acetonitrile or methanol solvents, in certain volumes (2mL, 3mL and 4mL), without any additional steps (e.g. heating baths and ball mill) to break the cell wall, as a few experiments in the past showed that it was possible to extract *Spirulina* in a simple mix with solvents. Three analytical replicates were made for each type of sample. Simple mixing of *Spirulina* powder with acetonitrile solvent, followed by vortex mixing, was unfortunately not sufficient to solubilise all of the powder at any volume, even with 4mL. Some of the powder managed to solubilise, however, a large amount of powder remained undissolved at the bottom of the Eppendorf tubes. When the small solubilised *Spirulina* in all of the tested sample replicates was analysed by LC-MS, the detected number of total putative metabolites was observed to be lower than in *E. coli* (around 300–350 putative metabolites in each extraction experiment). The produced samples generated very weak peak signal responses of metabolites (TICs of E^4 on average) – lower than expected. When mixing with methanol, the solubilisation of samples was slightly better but the results were still of a very poor quality. (Some experiments were not run due to a risk of column blockage because the powders were not dissolved at all.) Although *Spirulina* is a gram-negative bacterium (genus cyanobacterium) and consists of a thinner peptidoglycan layer in its cell wall compared to gram-positive, it was not sufficient for methanol or acetonitrile to dissolve the cell wall thoroughly without additional extraction steps. It was decided at that stage to improve the extraction techniques further.

Second set of optimisation experiments: additional techniques

With modified techniques applied (ball mill, water bath and centrifugation), the metabolite profiling results improved. Metabolite profiling results of *Spirulina* extract in methanol (10mg/2mL, three analytical replicates) showed a wide range of detected metabolites (596), compared to the other two concentrations (574 and 568 putative metabolites for 3mL and 4mL, respectively). The nature and content of metabolites from all three extraction concentrations were very similar. The results from the acetonitrile solvent extraction (same method, same samples) were not significantly different (608 putative metabolites for 10mg/2mL concentration, and 563 and 549 for 10mg/3mL and 10mg/4mL, respectively). The

majority of identified metabolites in all three experiments were of amino acid, carbohydrate and other biochemical classes. This demonstrated that *Spirulina* was successfully solubilised using the advanced extraction techniques without any dissolving issues, as well as showing that in the highest concentration it can produce a higher range of metabolites. The solvent choice did not show a big difference. In all of the experiments, however, it was noticed that peak signal response for metabolites was still observed to be lower than expected (TICs E^7) with a high background noise. Figure 2.8 shows the metabolite profiling results of *Spirulina* extracted in concentrations of 10mg/2mL for methanol and acetonitrile solvents. Figure 2.9 (A and B) shows the spectra of those extracted samples, demonstrating high background noise and low TIC values.

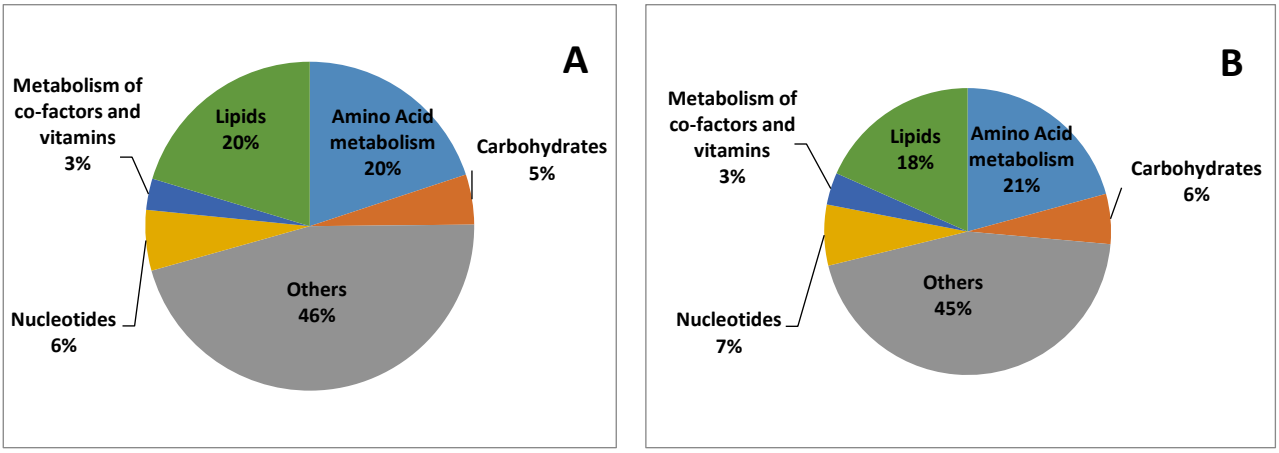


Figure 2.8 The metabolite composition of *Spirulina* extracts (10mg/2mL) in methanol (A) and acetonitrile (B) solvents. Percentages represent the amount of metabolites split by biochemical class.

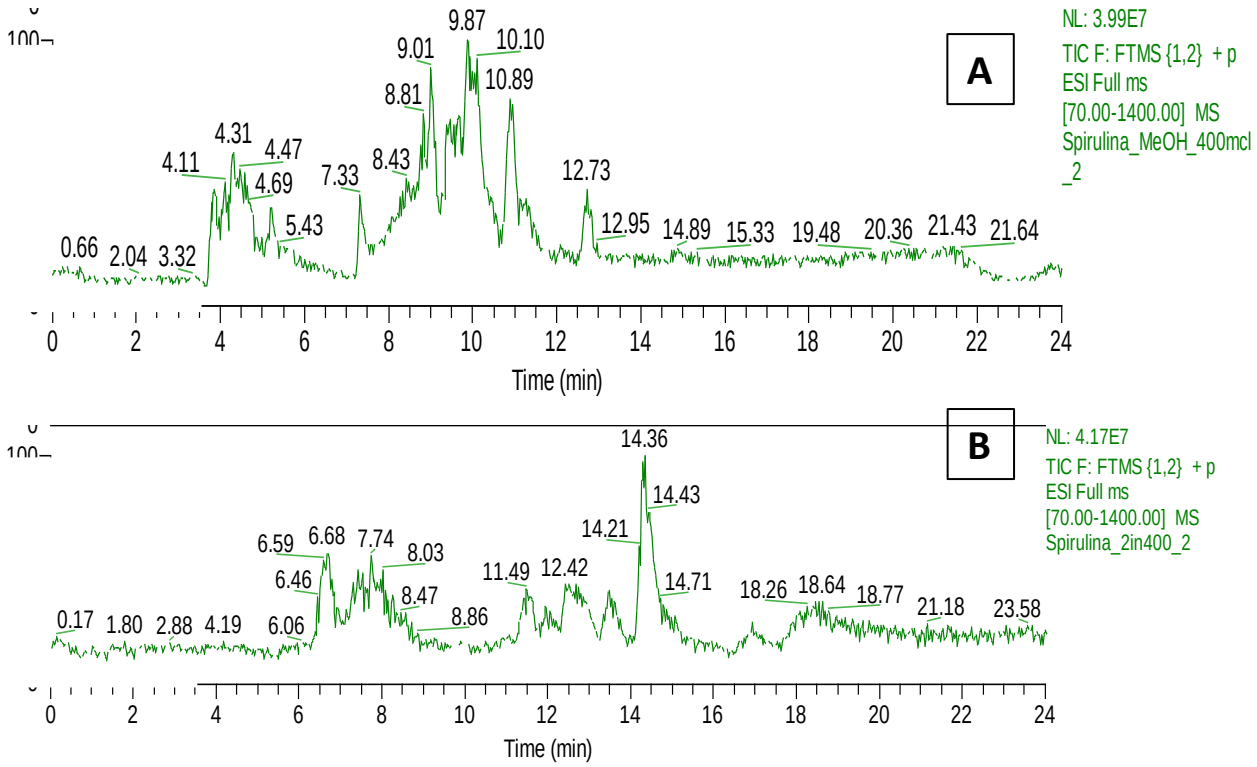


Figure 2.9 The full spectrum of *Spirulina* samples (10mg/2mL) extracted in methanol (A) and acetonitrile (B) analysed by "XCalibur" software showing metabolite peaks during a 24-minute run. The signal response (TIC) is recorded to be $3.00E^{7}$

2.3.3 Comparison of unlabelled *E. coli* and *Spirulina* metabolomes with human biofluid metabolites

Comparison

High quality data of unlabelled *E. coli* and *Spirulina* bacterial metabolomes, with the largest number of putative metabolites taken from previous experimental data (787 for *E. coli* and 596 for *Spirulina*), were directly compared to the online human metabolome database (HMDB)[53], in order to assist in finding a suitable bacterial extract with a metabolome closely related to human biofluids. The comparison was only performed with human blood and urine metabolites, as those biofluids are the most common sources of human metabolites and are usually investigated in metabolomics clinical studies.[54,61,289]

The metabolome of *E. coli* was initially verified with the "*E. coli* metabolome database"[214], to confirm the nature of metabolites detected in the experiments with metabolites already registered on the online database (a reliable source from past research studies on *E. coli* metabolite profiling). For *Spirulina*, such an online database did not exist, so full verification of

detected metabolites was not possible (the metabolome had to be checked with published research papers describing detection of *Spirulina* metabolites).[290,291] From direct comparison with the human metabolome database, *E. coli* appeared to have more metabolites in common with biofluids (plasma and urine) than *Spirulina,* as shown in figure 2.10 (A) (658 *E. coli* metabolites, 492 *Spirulina*). When both bacterial metabolomes were compared to each other, 450 metabolites were found to be in common, out of which 367 metabolites were in common with human biofluid metabolites from the database. These included common amino acids (L-alanine, L-proline, L-phenylalanine); nucleotides (AMP, adenine and cytidine); carbohydrates (D-glycerol); common organic acids and other biochemical entities.[212] This was an expected observation, as both bacteria would require such important metabolites for their growth and development, and would contain a wide range of polar metabolites, as established earlier in this chapter. Metabolites which were absent in *Spirulina* but were present in *E. coli* data (and were found to be in common with human biofluids) were important amino acids (L-adrenaline, L-cysteine, L-lysine), carbohydrates (D-glucose, D-glycerate), organic acids (maleic acid, orotate, picolinic acid, pyruvate), nucleotides (ATP, ADP, GMP), all of which are found to be essential for the general growth of an organism. Detailed comparison tables of *E. coli*, *Spirulina* and human biofluid metabolomes are found in appendix (Table A2).

Overall, these results have demonstrated that both *E. coli* and *Spirulina* have a lot of metabolites in common with human urine and plasma biofluids, but *E. coli* has shown better results with more metabolites in common with biofluids. The metabolome of *E. coli* extract showed the presence of some key metabolites which were missing in the *Spirulina* extract – amino acids L-cysteine and L-lysine and several important nucleotides. The absence of those metabolites in the *Spirulina* extract could have been due to the preparation of the lyophilised dry powder by the manufacturer. It appears that one of the drying methods using hot drums can significantly reduce the level of sulphur in amino acids such as L-cysteine, L-methionine and L-lysine, as investigated by FDA research.[292] Therefore, it is important to know how the *Spirulina* powder was produced and by which drying method. This aspect is difficult to control, as most *Spirulina* powders are produced industrially. As mentioned earlier, solubilisation and extraction of *Spirulina* were also found to be challenging aspects, even with all the recommended extraction equipment used. Ineffective solubilisation could have led to a loss of some key metabolites.

Figure 2.10 (B) shows the number of metabolites detected in each biofluid (human blood and urine) for each bacterial extract metabolome. Similar numbers and identities of *E. coli* metabolites were found in both biofluids (523 and 563), as was the case for *Spirulina* (383 and 411), suggesting that there was no significant variation in metabolome between the biofluids when analysing individual extracts. Important metabolites, including common amino acids, organic acids, carbohydrates and lipids, which are essential components for normal human development and physiological function, would be expected to be present in human urine and plasma biofluids.

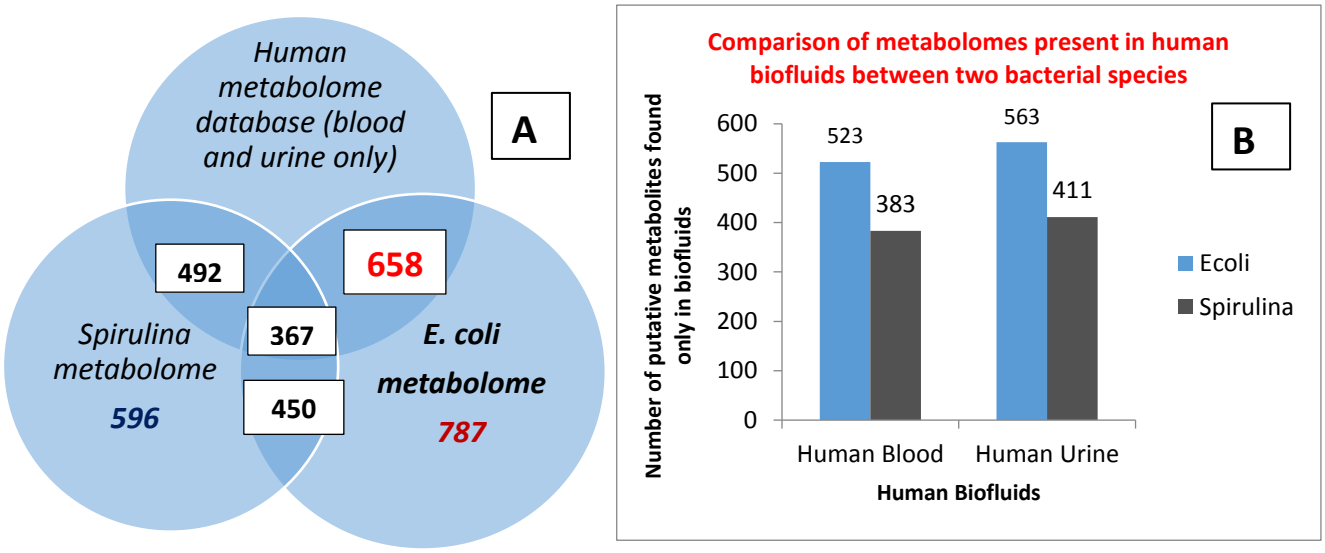


Figure 2.10 Comparison of metabolomes of unlabelled *E. coli* and *Spirulina* extracts with metabolites of human biofluids – blood and urine. General comparison (A) and comparison with each biofluid (B).

Deciding on the most suitable bacterial extract

Based on all the bacterial metabolite profiling experiments conducted so far, it has been concluded that the *E. coli* extract appears to be a more suitable bacterial microorganism for the development of internal standards than *Spirulina*. *E. coli* has shown a larger number of detected putative metabolites and better quality sample spectra (with higher TIC values). In addition to this, the *E. coli* metabolome was found to have more metabolites in common with human biofluids (blood and urine), and some key metabolites that are not detected in *Spirulina*, thus giving *E. coli* a big advantage. It was also observed that extraction of *E. coli* was found to be a lot easier and more effective in generating a high quality extract, as *Spirulina* has shown poor solubility during extraction (even with improved techniques) and involved multiple steps.[293] *E. coli* bacteria have, therefore, been chosen for future normalisation experiments.

2.3.4 LC-MS method validation

Two LC-MS methods (18 and 24 minutes – described in the methodology section earlier in this chapter) were compared to find the most efficient method of analysing the samples without losing important metabolite information. The aim was to make the method as time-efficient as possible. Reducing the overall running time of the samples from 24 to 18 minutes would help to save time overall (sometimes from days to hours) when running a large number of samples in future clinical experiments. To fully validate the new 18-minute method, in the first series of experiments, 268 authentic standards (split into five groups) were analysed with both the 24 and 18-minute methods, in turn, using the same LC-MS conditions described in the methodology section of this chapter. Those 5 groups (samples) were run three times each (3 analytical replicates) first by the 24-minute method, followed by the 18-minute method, in one big sequence. In the second set of experiments, an *E. coli* extract (500µ in 10mL, three replicates) were run with both the 24 and 18-minute methods using the same LC-MS conditions, to observe whether it was possible to detect a similar number of putative metabolites and the same quality of spectra in each case. Data was analysed by TOX-ID and IDEOM software, as described in the methodology section.

Authentic standards analytical run

After the 24-minute method analysis, 170 metabolites out of 268 tested authentic standards (68%) managed to show full detection with clear peaks and high extracted ion chromatogram values (E^7 or E^8). Their presence was detected in all three replicate samples. These results were similar to the data obtained in previous experiments where authentic standard mixtures were run alongside *E. coli* and *Spirulina* bacterial samples, and they showed a similar percentage of detected standards. In all the experiments where standards were analysed with the 24-minute method, TOX-ID software showed that the detected standard metabolite peaks were very close to expected retention times of authentic standards, confirming metabolite identification. IDEOM graph of "expected retention times vs experimental (obtained) retention times" of standards was showing a good proportional relationship with a high R2 score (gradient) of 0.674. In most of the experiments, the R2 score was between 0.6–0.7, indicating a very good run of the standards.

When the **18-minute** method was analysed using the same number of authentic standard samples, the overall results were found to be very similar to the ones obtained in the 24-minute method. The number of detected authentic standard metabolites was slightly lower (163 out of 250, 65.2%) with clear high extracted ion chromatogram peaks (E^7 or E^8) and a high R2 score obtained in IDEOM (0.656), indicating once again that detected metabolite peaks had retention times that were very close to the theoretical values. In figure 2.11, the examples of spectra of authentic standards mixture group 1 (containing 103 metabolites), analysed by the 24 (A) and 18 (B) minute methods, are presented. It can be seen that in both methods the peaks were similar in number and shape profile, with a very high TIC (E^8). However, it was noticed that there was a presence of retention time drift, with some of the peaks in the 18-minute spectrum eluting later than in the 24-minute one. However, this did not affect the identification of standards metabolites significantly.

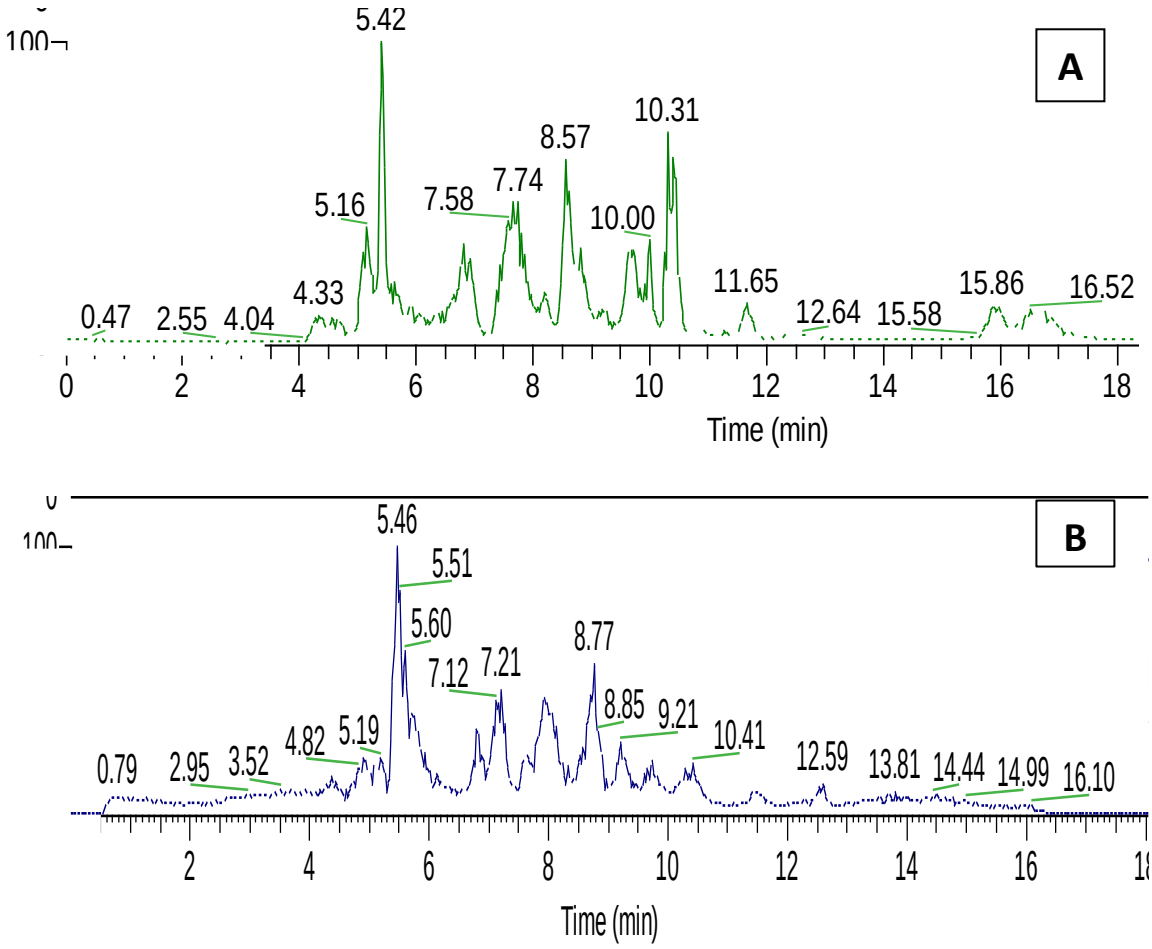


Figure 2.11 Spectra of authentic standard group 1 mixture sample in 24-minute LC-MS method (A) and 18-minute LC-MS method (B)

***E. coli* extract analytical run**

To fully validate the proposed 18-minute LC-MS method, the recently developed *E. coli* extract has also been analysed with 18 and 24-minute methods in turn. After data processing by IDEOM, the total number of detected putative metabolites in the *E. coli* extract from the 24 and 18-minute methods did not differ significantly (745 and 740 metabolites, respectively). When looking at the spectra quality in figure 2.12, there was a slight difference in the number of peaks and shape profiles between the 24 (A) and 18 (B) minute spectra, with some peaks in the 18-minute spectrum co-eluting at different retention times to the 24-minute one. Similar retention time drift was observed in a previous experiment with standard mixtures. It appears that reducing the analytical running time of a sample has an effect on the elution of peaks but not to a great extent – metabolite identification with IDEOM was still accurate and unaltered. The TIC values were reported to be high (E^8) in both spectra.

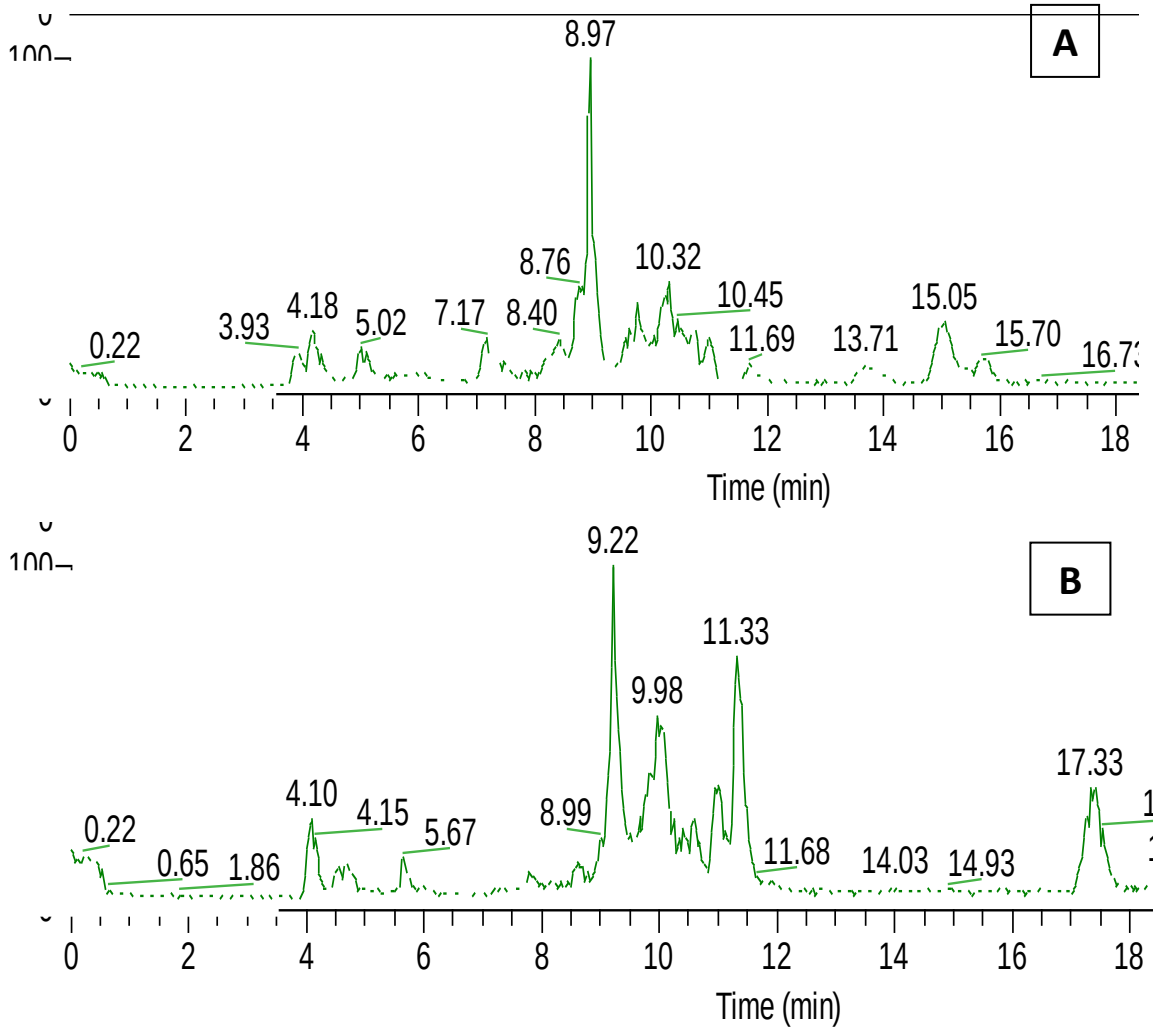


Figure 2.12 Spectra of *E. coli* extract sample in 24 minutes LC-MS method (A) and 18 minutes LC-MS method (B)

Based on the results from both experiments (with authentic standards and *E. coli*), it was concluded that using the 18-minute LC-MS method in future experiments for normalisation and clinical applications is still accurate and allows detection of a wide range of polar and non-polar metabolites.

2.3.5 Generation of a ^{13}C-labelled source of internal standards and testing its quality

Labelling test

Since *E. coli* bacterial organism has been established as a potential candidate for a source of internal standards, it was ready to be uniformly labelled with ^{13}C isotopes. U-^{13}C-labelled *E. coli* extract was produced under the same conditions as unlabelled extract, with the exception of U-^{13}C glucose as one of the growing constituents, and analysed with LC-MS Exactive, alongside an unlabelled extract, under the proposed 18-minute LC-MS method conditions. The total number of putative metabolites detected in unlabelled extract was 708, the majority of which were amino acids and carbohydrates, with the peak signal response of extract sample being very high (TICs of E^8).

After the processing of those detected metabolites with "mz-Match-Iso" software, it was confirmed that 548 putative metabolites (77%) were successfully labelled with ^{13}C isotope. The majority of the labelled metabolites were found to be amino acids, nucleotides, organic acids, carbohydrates and a small proportion of lipids, as was expected. The lack of ^{13}C isotope in the remaining unsuccessfully labelled metabolites (23%) could possibly have resulted from the contribution of ^{12}C from carbon dioxide that was introduced while *E. coli* was growing, or from *Lysogeny Broth* (LB) culture medium that had been used for initial *E. coli* growth prior to using M9 minimal medium. LB medium was found to contain numerous peptides such as tryptone, which contains ^{12}C isotopes.[235]

Quality of the U-^{13}C-labelled *E. coli* extracts

The labelled *E. coli* extracts which were produced for all the normalisation and clinical experiments (see chapters three, four, five and six) were checked for their quality – specifically, the amount and nature of ^{13}C-labelled metabolites. After processing the data with "mz-Match-Iso" software, it was confirmed that on average around 80% of putative metabolites in each extract were successfully labelled with ^{13}C isotope (fully or partially labelled), with the remaining

20% not showing any presence of ^{13}C isotope in their structure. As in the first labelling experiment, the majority of the ^{13}C-labelled metabolites were found to be of common amino acids, carbohydrates and common organic acids. The majority of the ^{13}C-labelled metabolites in each extract had isotopic purity of more than 85–90%, with some even expressing 100% ^{13}C purity. Figure 2.13 (A) shows the examples of several successfully labelled metabolites presented in one of the extracts, analysed by the "mz-Match-Iso" programme, which calculated the percentage of ^{13}C isotope purity. In figure 2.13 (B), the whole metabolome of *E. coli* extract is presented showing proportions of fully labelled, partially labelled and unlabelled metabolites.

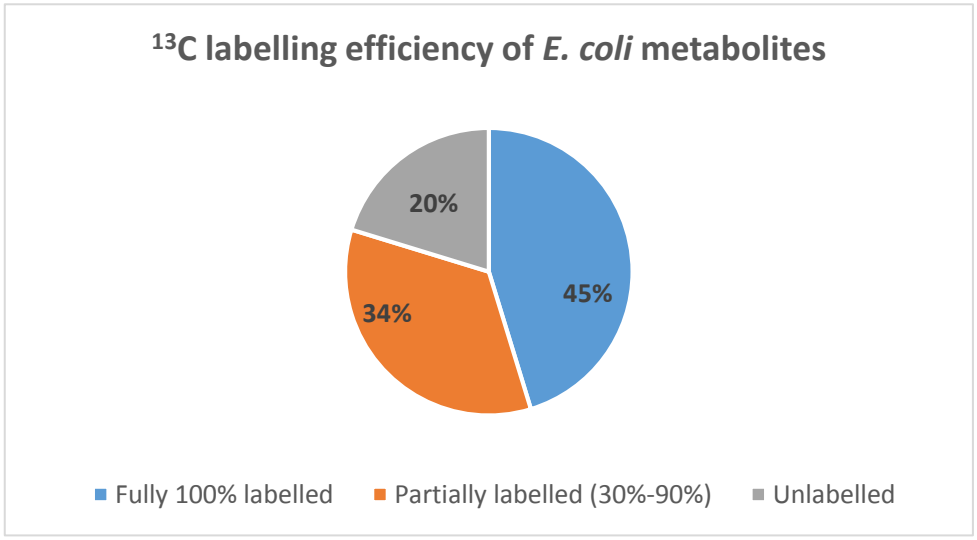


Figure 2.13 Isotopic purity of U-^{13}C labelled E. coli extract, chosen as an example – isotope purity of labelled extract (A) and labelling efficiency of whole metabolome (B)

It was concluded that *E. coli* extracts were successfully labelled with ^{13}C isotopes, showing a wide range of polar labelled metabolites with high isotopic purity. A certain percentage of metabolites could not be labelled and would not be used as internal standards but this issue can be rectified using another method (see chapter three for more details).

2.4 Conclusion

In this chapter, the development of an isotopically labelled bacterial organism as a source of internal standards for normalisation, was presented. The use of this labelled extract in a proposed normalisation approach would help to make an advanced step towards accurate data processing by correcting technical variability and matrix effects, and providing an unbiased method for the analysis of metabolites in human clinical samples.

E. coli appears to be a suitable bacterial organism as a source of internal standards. *E. coli* extract has shown a wide range of metabolites, most of which were in common with human biofluid metabolomes (mainly human blood and urine), in comparison to *Spirulina* results. The peak signal response, in the form of TICs in spectra for each extract, was found to be larger and clearer for *E. coli*, without any background noise. The most suitable extraction parameters have been chosen and they helped to generate a successful *E. coli* extract, while in *Spirulina* the extraction and optimisation techniques failed to generate an excellent extract. Finally, growing *E. coli* was a lot cheaper than purchasing *Spirulina* powder. These results confirmed that *E. coli* was suitable choice.

U-^{13}C-labelled *E. coli* extracts managed to show very good isotope labelling of a large number of putatively identified metabolites (around 80%) with high isotopic purity. Among them were common amino acids, organic acids, carbohydrates and lipids. The amount of labelled metabolites is high and similar results were observed in previous studies, where *E. coli* metabolome was labelled at high isotopic purity.[189,208] Other microorganisms, such as fungi, have also been involved in labelling experiments, with their metabolome successfully labelled with ^{13}C isotope.[194]

As U-^{13}C-isotope labelled *E. coli* extract has been used in normalisation in the past (mainly in targeted studies), it appears that it would be a promising source of internal standards for normalisation of human biofluid and tissue metabolites in untargeted metabolite profiling experiments.

Chapter Three

Validation of the normalisation method using human and animal biological samples

3.1 Introduction

3.1.1 Concept of validation and its purpose

In chapter one, the concept of normalisation has been introduced as the means of reducing unwanted analytical and technical LC-MS data variation in metabolite concentrations of samples that normally has a huge effect on the quality of metabolite results. With metabolites being normalised effectively, it can help to obtain a highly accurate data and to only focus on biologically relevant differences from the analysed samples.[193,294] Variety of normalisation methods have been applied to metabolomics data in the past, as discussed in chapter one, demonstrating their advantages and limitations. Stable ^{13}C isotope labelling technique has particularly been used in research, utilising single or multiple internal standards (labelled ^{13}C metabolites) to normalise the data.[181,188,295] In order to show its accuracy and advantage in metabolomics, this technique needs to be validated.

Different LC-MS methods together with variety of normalisation techniques have already been used in metabolomics in the past, as discussed in chapter 1. To ensure their accuracy, they were required to be fully validated. Validation concept was first introduced in 1970s by Food Drug Administration (FDA) for the purposes of assessing the quality of pharmaceuticals.[296] FDA defined validation as "Establishing documented evidence that provides a high degree of assurance that a specific process will consistently produce a product meeting its pre-determined specifications and quality attributes".[297] In other words, validation confirms that a properly designed system, such as normalisation technique in metabolomics, can provide a high degree of assurance that it is properly evaluated before its implementation and suitability for its intended use.[298] Validation has been used in metabolomics before. Apart from assessing normalisation techniques, validation tested variety of methods used for confirming potential biomarkers, metabolic pathways and other relevant information.[299]

In order to confirm that the proposed ^{13}C normalisation method is effective and normalised biological samples well, it requires a thorough validation process. As mentioned in chapter 1, two main ways of assessing normalised data is by looking at %RSD change in metabolite data and by running a multivariate analysis through PCA and OPLS-DA models to observe the improvement in quality of data visually. Validation is particularly important for this type of normalisation, because it is difficult to use available labelled ^{13}C peaks to normalise all the metabolome in the samples – for some metabolites, there is no ^{13}C peak to match it to. The so

called “co-elution” technique, mentioned in chapter 1, where ^{13}C peak of one metabolite being matched to another with similar retention time, may present an issue rather than solve a problem – metabolite might not be normalised accurately, as its co-eluted metabolite does not fully reflect all the chemical and matrix properties of the metabolite of interest. Therefore, metabolite data variation might get even worse after using “co-elution” normalisation. An assurance is needed to check that this type of normalisation is performed as accurately as possible, hence why validation is required.

Quite often, there is no proper validation or quality assessment of the published metabolomics methods, in particular normalisation techniques, present in the studies.[300] In the past 5 years, more than 900 papers have been published in metabolomics for biomarker discovery or testing normalisation technique, but there was a small number of them containing a high level validation assessment with certain criteria.[178] In one study, for example, the proposed normalisation technique using ^{13}C labelled internal standards, did a validation of the method on a previously published dataset to show that it succeeded in removing systematic variation in metabolite data, but accurately reproducing same results from that published dataset.[301] However, most of normalisation studies performed so far in metabolomics did not show that kind of validation, especially for untargeted studies.

Validation of normalisation method can be achieved by assessing the technique on a variety of clinical, biological (biofluids and tissues) samples. By demonstrating an effective normalisation in a variety of biological samples, reducing inter and intra-variation in the data, can fully validate the method. A good validation would also test the method on a large number of samples and replicates, to ensure that normalisation can be effective in a big clinical study in the future.

3.1.2 Planning validation experiments

The produced U-^{13}C-labelled extract (described in chapter two) is going to be applied in three validation experiments in this chapter – two projects with biofluid samples (human urine and mouse plasma) and one with human brain tissue – to assess the proposed normalisation technique by demonstrating the reduction in **analytical** metabolite data variation, possible reduction in biological variation (even though it is not essential) and comparing the number of statistically significant metabolites between normalised and non-normalised data. The reason for choosing biofluid samples, particularly human urine, was to observe the normalisation effect on wide-ranging intra and inter-individual metabolite variance usually observed in these samples, and to determine to what degree it could be reduced (mainly analytical variance).[163,302] The proposed method would also be compatible with hydrophilic metabolites, very large quantities of which are found in urine, as the labelled extract was created in hydrophilic conditions. The normalisation method may also be compatible with hydrophobic metabolites (lipids), normally detected in plasma.[303] Brain tissue was chosen for validation studies to show that the method can also work in complex tissues, where a large data variation is particularly observed and extraction of polar or lipophilic metabolites is not easy.[304] A sufficient number of analytical replicates (n=3) will be used in each validation case study to mirror the natural metabolite variation normally present when conducting metabolomics research. Biological replicates have also been used (n=6 or 10), creating a clinical variation in addition to analytical one. Normalisation effect will be observed on both analytical and biological variation. In each case study, an equal number of unlabelled and ^{13}C-labelled samples will be analysed, in order to do an accurate comparative normalisation analysis. In each study, there will also be a comparison of two groups of samples (for example, control vs treatment), fully completing the normalisation analysis.

Finding statistically significant metabolites

After normalising the metabolites with the proposed method and assessing the quality of data, it is reasonable to perform multivariate and univariate statistical analysis on normalised data in order to find a list of statistically significant metabolites.[153] This could be an additional step in validation experiments to fully confirm the efficiency of normalisation. Univariate analysis methods including T-test, FDR and Bonferroni corrections and ANOVA performed on metabolite data, are used in order to assist in finding significantly changed metabolites between groups of data (e.g. control and treatment groups), as mentioned in chapter 1. The T-test is used to test the null hypothesis which states there is no significant difference between

two sets of data, and to produce p values for metabolites. The FDR correction further processes the data and checks whether the metabolites that had a p value less than 0.05 (statistically significant) are correct or not. Statistically processed data can then be displayed via "volcano plots", commonly used in omics studies, which graphically show how many metabolites are statistically significant and have a large magnitude of changes (fold changes) between data groups.[152,305] Multivariate analysis can further help to show the difference between data group sets and to obtain VIP scores for metabolites from the OPLS-DA model, which helps to measure a metabolite's importance in the model. It is usually done in parallel to univariate analysis, producing lists of statistically significant metabolites.[306,307] The lists of statistically significant metabolites from univariate and multivariate analyses are then compared to see which metabolites they have in common (in other words, passed both univariate and multivariate tests) and the final list of these significant metabolites is produced.

3.1.3 Aims and objectives

The following aims and objectives of this chapter outline the planned validation experiments to fully show the effectiveness of the proposed normalisation method using stable ^{13}C isotopes as internal standards:

1. Conduct validation experiments and ^{13}C normalisation data processing (%RSD calculations and multivariate analysis) in:

- Human urine study
- Human brain tissue study
- Mouse plasma study

2. Find and compare statistically significant metabolites in each study between normalised and non-normalised datasets, using multivariate analysis (VIP scores), T-test, FDR correction and volcano plots.

3. Perform other post-processing data normalisation tests on the metabolome data from the above studies and compare the results to ^{13}C normalisation to confirm the proposed method:

- TIC normalisation
- MSTUS normalisation
- Normalisation to creatinine (in urine study only)

Figure 3.1 below shows the summary of data processing workflow that will be used for normalisation analysis in each validation study, using the software programmes mentioned earlier for different purposes.

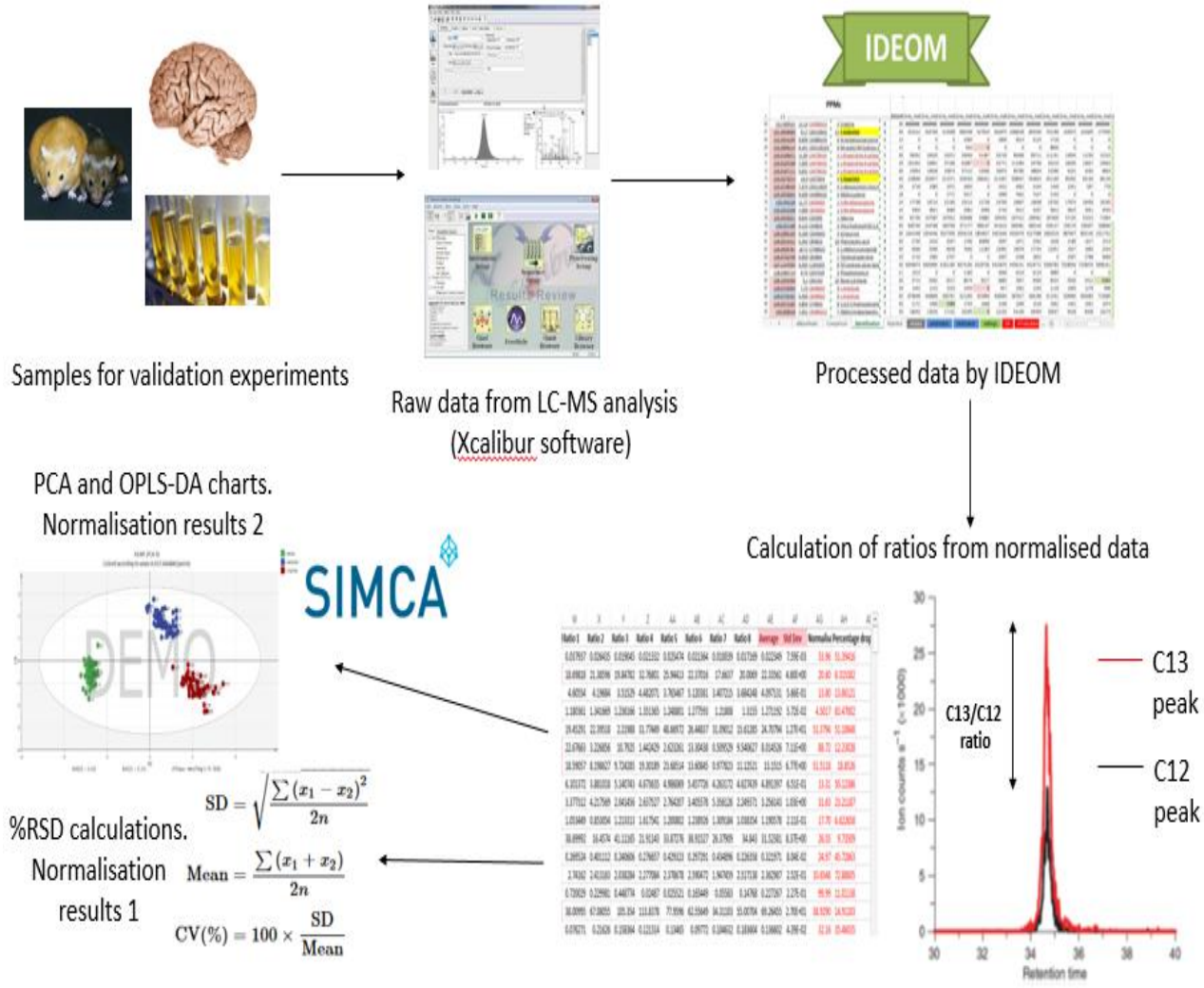


Figure 3.1. Normalisation data processing workflow applied in validation studies

3.2 Materials and methods

3.2.1 Materials, instruments and samples collection

All materials and reagents used in analytical experiments in this chapter were the same as described in chapter 2. They included LC-MS mobile phase solvents such as 20mM ammonium carbonate in purified water and HPLC-grade acetonitrile, zic-pHilic column, system blank (acetonitrile), "Calmix" MS calibration mixture and 268 authentic standards. Their details and supplier sources are described in chapter two.

The LC-MS instrument (Orbitrap "Exactive" coupled with "Accela" autosampler) used was the same as described in chapter two. Extraction apparatus used in these validation experiments were centrifuge ("Harrier 18/80 R", UK 2008), liquid Nitrogen dewar (25L "Agar Scientific", UK 2008), variable speed vortex mixer ("Cole-Palmer", 90 – 240 VAC, UK 2010), portable hand-held homogeniser ("Bibby Scientific Stuart", SHM1, 2012).

Human urine samples were collected from ten healthy volunteers in Boots Science Building, School of Pharmacy, University of Nottingham. Human brain tissue samples were collected from volunteers having clinical seizures but otherwise healthy, provided by the Children's Brain Tumour department, Queen's Medical Centre, University of Nottingham. Mouse plasma samples (healthy but having pre-disposition to Alzheimer's disease) were provided by the School of Life Sciences, Queen's Medical Centre, University of Nottingham. All samples did not contain any major diseases, with urine samples completely healthy.

The produced U-^{13}C-labelled extract (see chapter two) was applied to biofluid samples (human urine and mice plasma) and to human brain tissue, to assess the normalisation technique. In each case study, an equal number of "unlabelled" (extracted with ice-cold methanol, no presence of ^{13}C isotopes) and "labelled" (extracted with U-^{13}C labelled extract in methanol) samples were prepared, in order to do an accurate comparative normalisation analysis. The extract with internal standard was added in a 1:1 ratio to the sample in each experiment – in human urine project 500 µL of extract was mixed with 500 µL urine sample by pipetting and vortex mixing; in mouse plasma project the same volumes were applied as in urine; in human brain tissues project the volume of extract was added proportional to weight of the tissue (30 µL per 1mg). Analytical and biological replicates have been produced as well. QC samples were prepared by mixing an equal amount of each sample – this was done in each validation experiment. See next section for more details.

3.2.2 Design of validation experiments

Human urine study

This study investigated the effect of caffeine-rich black tea ("Earl Grey", Twinings) on the human urine metabolome, collected from healthy individuals by the researcher, in order to validate the proposed normalisation method. Two sets of human urine samples: **control** (urine sample before ingestion of tea) and **treatment** (after ingestion of tea), were collected from 10 healthy male volunteers (n=10, BMI 20–30, age 25–35) into large 50mL Falcon tubes. A small amount of urine (500 μL) from each collected sample was extracted (either with ice cold methanol or U-^{13}C labelled extract in methanol) in a 1:1 ratio in Eppendorf tubes, vortex-mixed for 1–2 minutes and centrifuged by "Harrier" centrifuge (5000g, 10 minutes, 4◦C) to obtain a supernatant containing polar metabolites. The supernatant from each tube was collected into fresh pre-cooled Eppendorf tubes and stored at -80◦C conditions. There were 40 **extracted** samples produced in total – 20 of which were controls (10 ^{13}C labelled and 10 unlabelled) and 20 taken after tea (10 ^{13}C labelled, 10 unlabelled). Urine samples from subjects after ingestion of tea and controls were analysed in a single analytical run using the developed LC-MS method, randomised with QCs and reagent blanks as stated earlier. Recruitment of volunteers and collection of raw urine samples were performed in the School of Pharmacy, ethically approved by the School's Ethics Board Committee (approval license number 021-2016). The project study was titled "**Global metabolite profiling of human urine samples treated with caffeine-rich tea and the application of normalisation methodology on the metabolome data**".

Volunteers were asked to answer a series of healthcare questions (health questionnaire) before taking part in the experiment, in order to determine whether they were suitable for the study or not. The criteria for selecting volunteers depended on the volunteer's health status, having no renal or hormonal diseases, being a non-smoker and having no medications prescribed. The presence of these criteria might have had a significant effect on individual's urine metabolome, and thus might have affected the true measurement of urine metabolites when comparing the control and post-tea datasets.[308] Renal impairment especially affects metabolite concentrations, elevating or reducing levels of certain amino acids and organic acids, as well as hugely affecting concentrations of creatinine, since the kidney is an important organ for urine glomerular filtration.[309] Therefore, it was necessary to have healthy volunteers participating in this study. On the day of the experiment, volunteers were asked to avoid any caffeine-containing drinks in the morning before participating, so their urine wouldn't be affected by additional caffeine, other than coming from the "Earl Grey" tea. During the

experiment, each volunteer produced a sample of urine before consuming the tea (control), then he was asked to drink one cup of tea and give urine samples at several time points: two hours and four hours after tea consumption. These were later pooled in one Falcon tube (see figure 3.2 for more details). The documents describing participant's information sheet, participant's consent form, health questionnaire and ethical approval form are found in the Appendix (A3). One of the volunteers (number 10) was later removed from the study because they failed to meet the specified requirements (the volunteer drank a caffeine product in the morning before the study began). The other participants fully complied with the study requirements and their samples were accepted for analysis. The data has been kept confidential and secure.

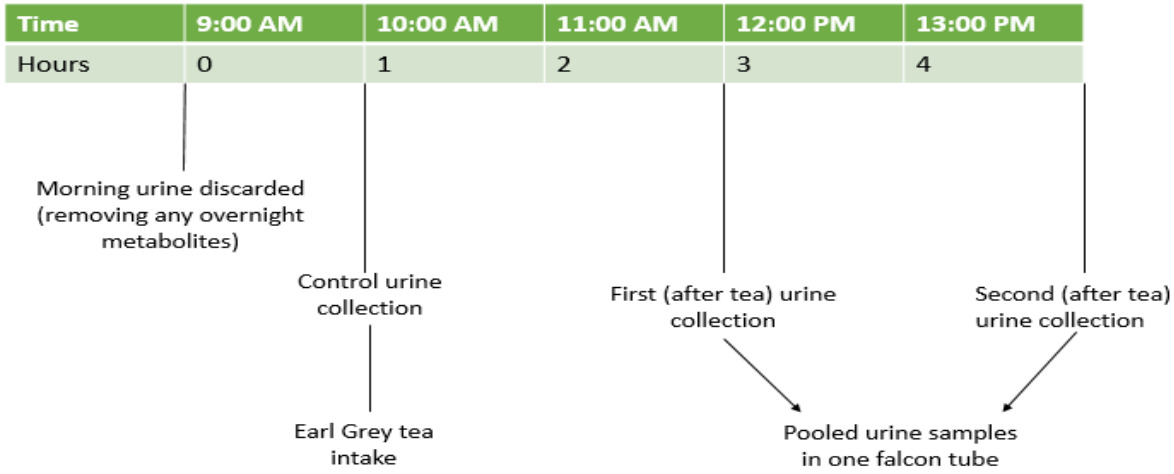


Figure 3.2 Schematic overview of urine sample collection from a volunteer during the experiment

Human brain tissue study

Human brain tissue (grey and white matter) were cut down into the required weights (≈30 mg), using a scalpel on a dry ice surface, and were stored in Eppendorf tubes – there were 24 samples produced in total – 12 for each matter, 6 of which were ^{13}C labelled and other 6 unlabelled (no presence of ^{13}C, extracted in methanol). The extraction procedure involved several steps including homogenisation of tissues and addition of various solvents to separate the lipid, aqueous and protein layers. The tissues were homogenised using a portable hand-held "Bibby Scientific Stuart" homogeniser in the appropriate extraction solvents (either ice cold methanol or ^{13}C labelled extract in methanol). The volumes of extraction solvent were calculated according to the exact weight of a tissue (30µl per 1mg weight). After homogenisation, 4°C chloroform (HPLC-grade) and 4°C purified water (HPLC-grade) were added to the tissues in a 2:1 ratio,

followed by vortex-mixing and centrifugation by "Harrier" centrifuge (13000g, 10 minutes, 4°C). The upper layer (aqueous phase) was collected separately from the protein (middle) and lipid (lower) layers.[176] The extracted samples were stored at -80°C. The study has been granted an ethical approval from National Research Ethics Service Committee from East Midlands, UK.

Mouse plasma study

Mouse plasma samples (500µL volume, female gender, aged 3 months and 18 months) were extracted either with ice cold methanol or labelled extract in a 1:1 ratio, following the same procedure as described for urine samples. There were 24 samples produced in total – 12 of each age, 6 of which were ^{13}C labelled and other 6 unlabelled (no presence of ^{13}C, extracted in methanol). The extracted samples were stored at -80°C. Mouse plasma samples were collected under ethical approval from National Research Ethics Service Committee from East Midlands, UK.

3.2.3 LC-MS analytical run of the samples

LC-MS analysis of human urine, human brain tissue and mice plasma samples were performed with Exactive Orbitrap-Accela Autosampler LC-MS with the same validated 18 minutes method, as described in chapter two, using zic-pHilic column chromatography. This time, the analytical sequence run of samples was randomised in each experiment, with QCs, analytical replicates (n=3 for each sample) and reagent blank (methanol) samples run alongside the main samples, after a certain interval. Randomisation is usually performed to minimise the analytical drift on samples that is normally present in the analysis of large sample cohorts. Analytical drift often introduces a huge bias, obscuring the analysis of data and its interpretation. Constrained independent randomisation between and within sample pairs (e.g. control vs treatment) minimises the instrument drift bias and increases the accuracy of the data.[210] Pooled QC samples were interspaced with the rest of the samples to monitor the stability, robustness and performance of chromatographic analysis.

3.2.4 Data processing and multivariate analysis

Normalisation data processing

Putative metabolite identification in biofluid and tissue samples was performed by "XCMS", "mzMatch" and "IDEOM" programme which were used earlier (see chapter two for more details).[147] An unlabelled metabolite peak height intensity value was obtained from IDEOM, while ^{13}C-labelled peaks for every metabolite were located in a pre-processed file called "mzMatch_output" with the application of Python code (described later). The ratios were calculated by division of the ^{13}C-labelled peak by the unlabelled peak intensities of each metabolite ($^{13}C/^{12}C$) in each replicate (separately for analytical and then biological replicates) – the whole procedure is discussed in chapter 1. For those metabolites which were not matched with their ^{13}C peaks, different "co-eluted" ^{13}C-labelled internal standards were used instead (^{13}C peaks of metabolites with very similar retention time to metabolite, window range -0.09 to +0.09 minutes). The calculated normalised ratios were uploaded into "Simca 14" multivariate software to produce PCA and OPLS-DA score plots, to assess normalised data. Original IDEOM ^{12}C (unlabelled) peak intensities of the same metabolites, non-normalised data, were also uploaded to "Simca 14". Direct comparison of normalised and non-normalised data through OPLS-DA and PCA score plots was observed. See figure 3.1 for overall normalisation data processing method and assessment of data quality. Those metabolites which did not have a carbon atom in their structures or were not detected in several replicates, generating zero peak height intensity, were not used for normalisation at all, as the ratios would not be calculated accurately. Once normalised metabolite data was checked for the quality, it was run through several univariate and multivariate statistical programmes in order to find any statistically significant metabolites between the two comparison groups in each study, and to compare these results with any statistically significant metabolites from unlabelled samples. Univariate analysis via T-test and FDR correction has been used to generate statistically significant metabolites, showing data through volcano plots. Multivariate analysis using OPLS-DA was used to produce VIP scores for metabolites to see which generated a score of 1.0 or above. The lists from univariate and multivariate analyses were compared, and the final list of significantly changed metabolites presented.

Other post-analysis normalisation approaches, such as MSTUS and TIC, have been applied to the data sets of human brain, urine and mice plasma metabolomes as well, in order to compare with the ^{13}C strategy and to determine the most effective normalisation technique. These strategies were not involved during sample preparation but were applied on already processed data – post-

processing techniques. Normalisation to creatinine has also been used but only on human urine data, used on already processed metabolome data (as a post-processing normalisation). As mentioned in chapter 1, MSTUS normalisation uses the sum of all peak height intensities of metabolites within a sample chromatogram as a common denominator. For each sample (replicate), the sum of all peak intensities of every metabolite present in that sample was calculated, then each peak intensity of every metabolite was divided by that sum, generating the ratios (normalised MSTUS data).[180] The calculations for TIC normalisation was done in a similar manner to MSTUS, where each peak intensity of metabolite in every sample was divided by the sum of all peak intensities. However, before the calculation, original metabolome dataset (peak intensities) had to be changed into TIC format peak intensities, which then could be used in calculation of ratios (final TIC normalised data). That transformation of original metabolome dataset into TIC format was done automatically by IDEOM function (TIC normalisation).[147] And finally, Creatinine normalisation was performed by dividing peak intensity of each metabolite in a sample replicate by the peak intensity of creatinine metabolite in that sample replicate. The creatinine normalised ratios were calculated.[200] In all three types of normalisation, the calculated ratios (normalised data) went through the same statistical analysis as ^{13}C normalised data. Normalisation assessment was done for analytical and biological replicates separately.

Identification of ^{13}C peaks by Python programme (just for ^{13}C normalisation only)

To assist with finding the correct ^{13}C peak height intensities for each metabolite, the computer generated Python code (List-mass-c13-check-ACT.py) was specifically written to assist locating peak intensities in a raw pre-processed data file ("mzMatch output"), acquired in the early stages of data pre-processing. The code was generated by a colleague (Mr Ferran Casbas) from Centre for Analytical Biosciences group, School of Pharmacy, University of Nottingham. The code Python has been chosen for its accurate comprehensive data analysis commonly used in bioinformatics.[310] The code works by checking the "mzMatch output" file for ^{13}C peaks which match the specified molecular weight and retention time (RT) parameters set by the user. Users must edit the script directly to specify the ranges of ppm (5 ppm to set an exact mass) and RT (small range e.g. -0.09 to +0.09 minutes). Where the exact ^{13}C peaks for metabolites could not be detected, the ^{13}C peaks from co-eluted metabolites having the same retention times were used instead. In this case, the ppm range was chosen to be high with a very strict, small RT range, allowing the programme to choose any ^{13}C peaks having the same retention time as the investigated metabolite. This Python program required the library, *xlrd*, to read Excel spreadsheets and the common Python libraries *re* and *NumPy*. On average, the Python code was running for less than a minute, saving a huge amount of time. A manual check verification was

still performed in each data set to confirm the accuracy of the code. The description of the code can be found in the appendix (Table A4).

3.3 Results and Discussion

3.3.1 Human Urine study

Urine metabolome and analytical performance (before normalisation)

A wide range of polar compounds including amino acids, carbohydrates and peptides were detected in urine samples (overall 723 putatively identified metabolites in common for control and "after tea" groups). Chromatograms of urinary metabolomes from control and "after tea" samples were found to be similar in peak shape characteristics. However, the TIC for the controls were slightly higher than for the other group. This might suggest that urine samples taken after ingestion of the tea were more diluted than the control samples, due to a diuretic effect of caffeine in the tea.[311] The fold changes in metabolites produced by IDEOM between the two groups demonstrated a certain difference in metabolite concentrations, showing either an increase or decrease in peak intensity values. Metabolites expressing a huge concentration difference were later shown through multivariate analysis (in OPLS-DA models, see figure 3.4). The QC samples demonstrated almost consistent chromatograms with very similar TICs (>E^8), with an average %RSD less than 30% for most metabolites. This was within the acceptable limits according to the recommended threshold for metabolomics analysis.[312,313]

Normalisation analysis

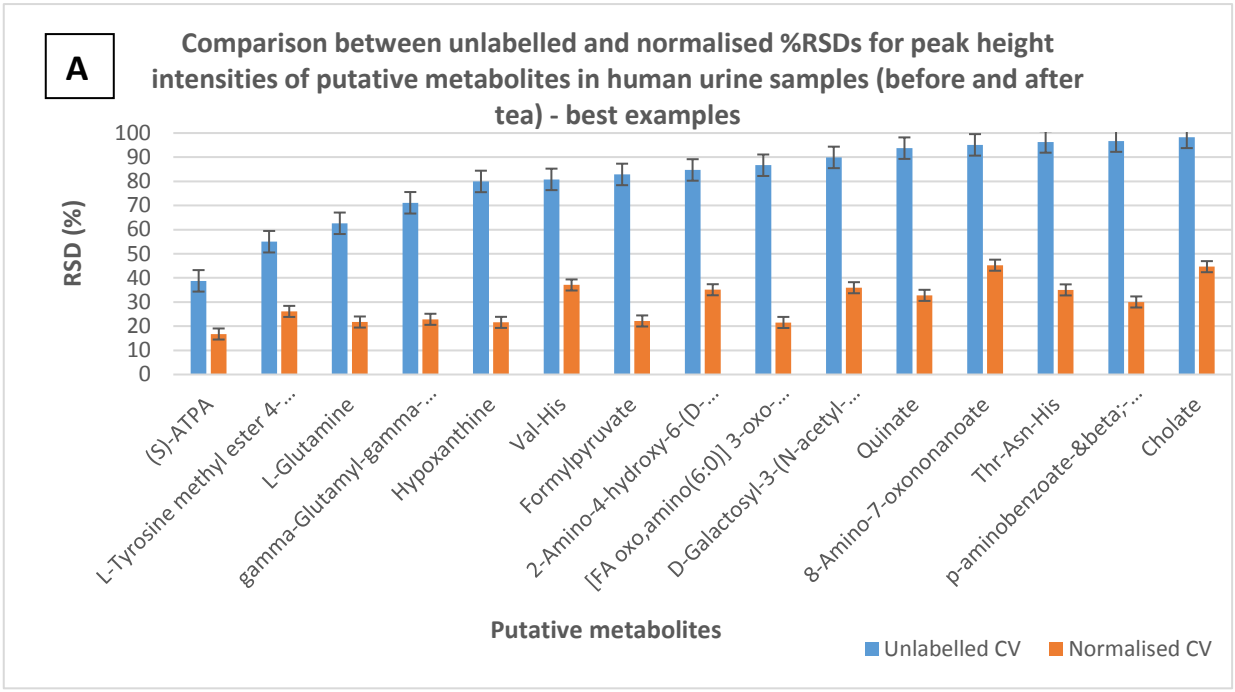
Before normalisation, almost all of the non-normalised (unlabelled) metabolites were expressing high %RSD variation (≥30%) in peak intensities between their biological replicates, and even between analytical replicates too. It is known that urine samples vary significantly in total solute concentration, presenting a challenge of achieving a high quality dataset.[61,314] There is usually a wide intra- and inter-individual variation observed in metabolite concentrations between human or animal control replicates, shown in most urine metabolomics studies.[163,313] After ^{13}C normalisation has been applied to the dataset, most metabolites (more than 670 in both groups) showed a reduction in %RSD **analytical** variation (94% of metabolites in control group and 95% in "tea" group), being successfully normalised either by their ^{13}C paired peaks or by co-eluted labelled metabolites from the extract. Similar

observation has been demonstrated when looking at **biological** variation, with high number of metabolites expressing reduced %RSD (71% in the control group, 68% in the "tea" group). Appendix (Table A5) shows information on %RSD values for analytical replicates only.

Figure 3.3 (A) shows several examples of those successfully normalised metabolites (with both analytical and biological %RSD reduction). They included common amino acids (such as l-alanine, l-lysine and l-histidine), a wide range of organic acids and carbohydrates. Poorly normalised metabolites, which expressed a higher %RSD than before, or stayed the same, were found to be mainly lipids and several hydrophilic compounds. They were poorly normalised because they did not have corresponding ^{13}C paired peaks or appropriate co-eluted metabolites which could have reduced the %RSD effectively. In figure 3.3 (B), box and whisker plots show the distribution of %RSD (biological replicates only) values for peak intensities of non-normalised metabolomes and metabolomes normalised by ^{13}C, creatinine, MSTUS and TIC methods. Box and whisker plots demonstrated distribution of %RSD values (black dots), from lowest to highest values (ends of whiskers of the plots) observed for particular normalisation method, with upper and lower quartiles representing medians of upper and lowest halves of data. It can be seen that %RSD range for the ^{13}C normalised metabolome was less variable (shorter) than non-normalised range, with lower %RSD values generally present. Even though the median (horizontal line in box and whisker plot) for the normalised range was above 30% (36.8% in the control group, 53.4% in the "after tea" group), normalisation of the urine metabolome significantly improved the accuracy of the data and managed to reduce peak intensity variation between replicates for most metabolites, confirming that the method is working well. Normalisation by MSTUS and TIC also showed a decrease in %RSD in peak intensities of metabolites but not to such a good extent as the ^{13}C or creatinine methods (as shown in figure 3.3 (B), the control and "after tea" treatment groups median values for MSTUS and TIC ranges were similar and a bit lower than in non-normalised). Normalisation to creatinine managed to reduce %RSDs significantly (from 68.3% RSD to 46.7% in the control group, and from 85.9% RSD to 64.3% in the tea group), however, the ^{13}C method managed to reduce the %RSD even further, proving it to be the better method. There were several outliers (20 and 31 metabolites in control and tea groups) which can be observed for ^{13}C normalisation – those outliers belonged to some of lipophilic and "others" class metabolites whose %RSD got worse after normalisation (black dots above the error bars). The reason for this observation remains unclear, however there were no matched ^{13}C peaks found for those metabolites, as well as some of them were lipophilic in nature, therefore would not be normalised effectively.

Therefore, it seems that ^{13}C normalisation has successfully reduced analytical variation in majority of metabolites in comparison to other techniques and non-normalisation. It has also successfully reduced biological data variation in most of metabolites. Among the successfully normalised metabolites, only around 35% in both groups had their %RSD below 30%. Even though normalisation significantly reduced metabolite data variation for most of metabolites, %RSD still stayed above 30% for majority of them. This demonstrated that normalised urine is still presenting a challenge of obtaining highly accurate data. This could have happened due to factors such as human age, genetic background, diurnal variation, activity and diet, all of which contribute towards huge data variation.[163,181]

The multivariate analysis in figure 3.4 showed a clear separation between the two data groups (control vs tea treatment) in normalised PCA and OPLS-DA plots (OPLS-DA Q2 score = 0.887). The clustering of replicate points on the normalised OPLS-DA model was also observed, indicating that normalisation worked effectively by reducing variation between biological replicates. In unlabelled PCA there was no clear separation between the groups, as the model was unsupervised and metabolites were not normalised. QC sample replicates in normalised dataset have shown lower %RSD values than non-normalised, on average, for majority of metabolites, showing high reproducibility of data and stability of analytical methodology throughout the analytical run of urine samples.[312]



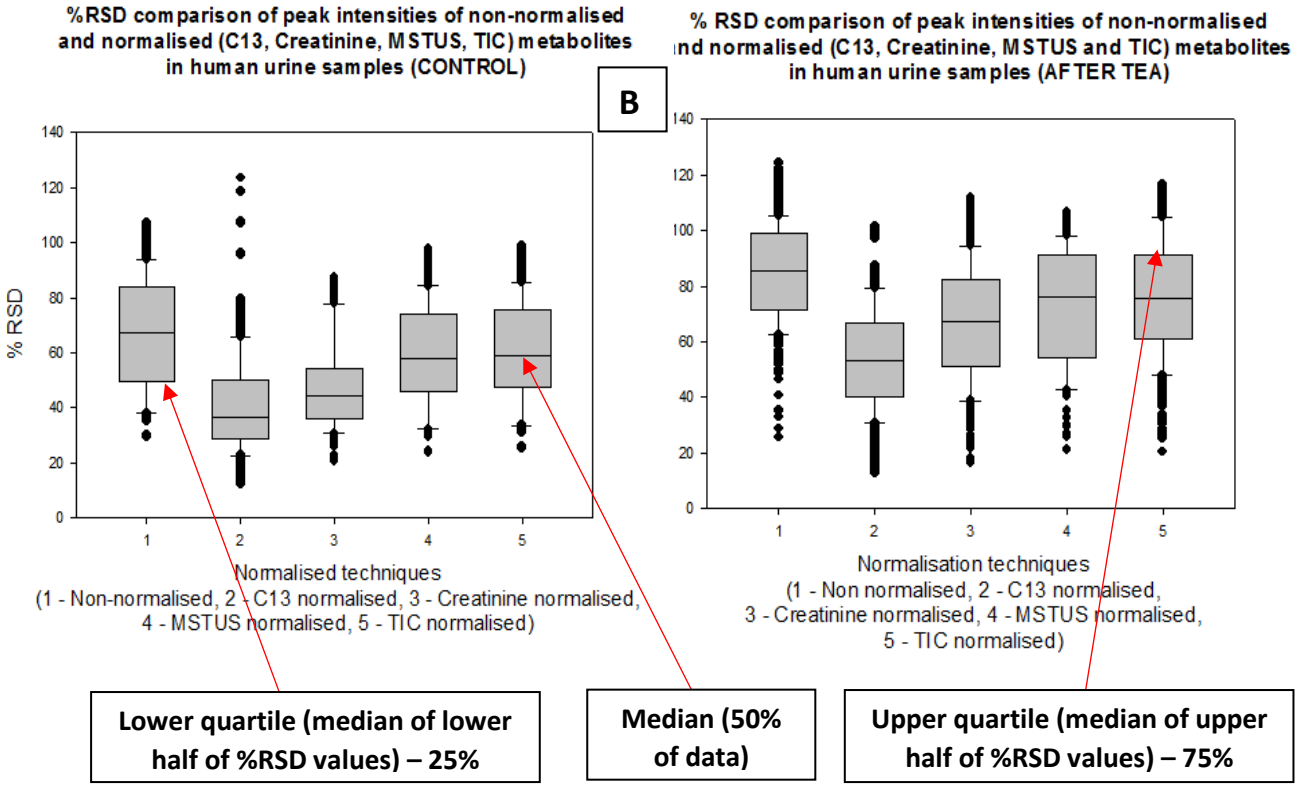


Figure 3.3. Comparison of %RSD between unlabelled and normalised metabolomes. Comparison with best metabolite examples (A). Comparison of %RSD range (biological replicates only) across the whole metabolome (B).

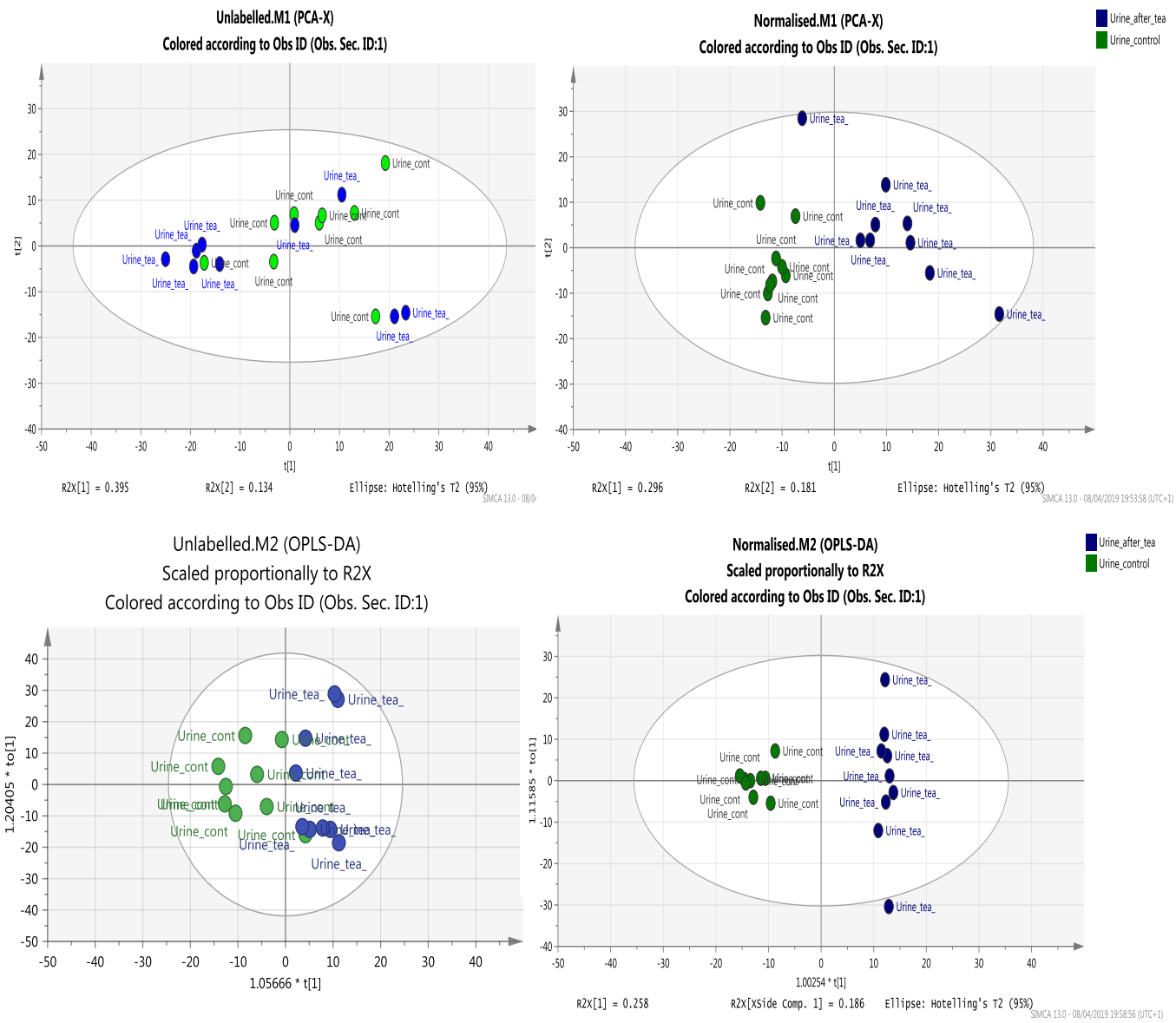


Figure 3.4. Multivariate analysis of urine control and tea treatment data groups using PCA and OPLS-DA score plots for non-normalised (unlabelled) and normalised datasets (green replicate points represent control, blue replicate points represent treatment (tea). PC1 and PC2 % variance for non-normalised plot = 75% and 65%, for normalised plot = 82% and 46%, respectively. OPLS-DA Q2 and R2 scores for non-normalised plot = 0.62 and 0.72, for normalised = 0.887 and 0.93, respectively.

Statistically significant metabolites

Univariate analysis through T-test, followed by FDR statistical correction, has shown that 175 normalised metabolites were observed as statistically significant with their p value being less than 0.05 and large fold changes, while for non-normalised data there were only 84 statistically significant metabolites detected. This once again demonstrated that using normalisation and helping to reduce analytical variation in metabolite data improves the data accuracy and increases the chance of identifying significantly changed metabolites. Volcano plots in figure 3.5 demonstrate the amount of statistically significant metabolites in unlabelled and normalised samples, showing that among normalised metabolites there were more statistically significant metabolites (above the log p value line – p less than 0.05) and a large number of green points indicating a large number of metabolites with higher fold changes, confirming the statistical significance of those metabolites.

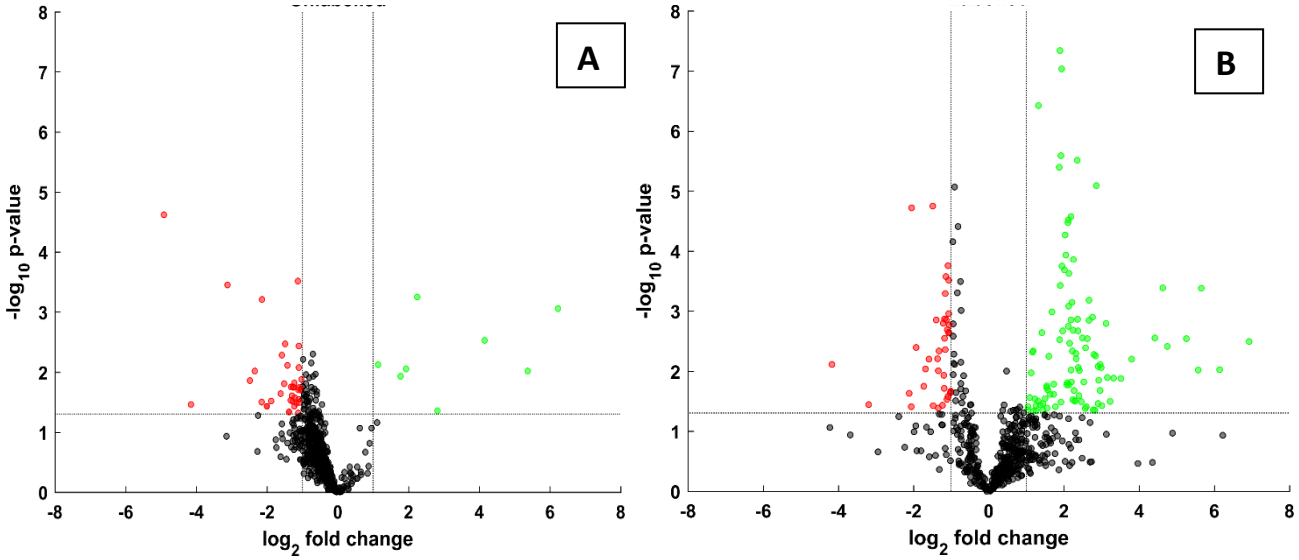


Figure 3.5. Volcano plots showing the amount of statistically significant metabolites with their fold changes in non-normalised samples (A) and normalised samples (B).

From multivariate analysis using a normalised OPLS-DA scores plot, a list of 342 metabolites showed a VIP score of 1.0 or above. When this list was compared to 175 statistically significant normalised metabolites from univariate analysis, there were 127 metabolites found to be in common (top 10 statistically significant metabolites are presented in table 3.1). Most of these metabolites were of hydrophilic compound class (amino acids, peptides, carbohydrates) and expressed large fold changes, indicating that “Earl grey” tea (treatment) elevated metabolite concentrations, compared to the control. This was an expected observation, as the tea contains a large amount of caffeine and caffeine related products would have a strong diuretic effect on

urine, increasing urine volume and its excretion. Therefore, the concentration of excreted urinary metabolites would be elevated.[315]

The number of statistically significant metabolites (which passed both univariate and multivariate analysis) in non-normalised dataset was a lot lower (76 metabolites), demonstrating how ^{13}C normalisation method can improve the data accuracy and detection of significantly changed metabolites.

Metabolite	Biochemical Class	Molecular mass	P value	VIP score	Fold change (Tea/Control)
Betaine	Amino acids	117.079	1.07E-05	1.279	2.563
Citrate	Carbohydrates	192.027	2.15E-04	1.275	1.225
D-Gluconic acid	Carbohydrates	196.058	3.97E-07	1.207	3.677
DL-2-Aminooctanoic acid	Others	159.126	<1.00E-15	1.206	1.664
Glutarylcarnitine	Others	275.137	9.35E-05	1.176	1.673
Hippurate	Amino acids	179.058	1.86E-08	1.173	2.924
Indoxylsulfate	Others	213.015	1.79E-06	1.181	1.823
L-Carnitine	Amino acids	161.105	1.85E-05	1.152	2.745
N-acetyl-(L)-arginine	Others	216.122	8.26E-08	1.124	3.485
Val-Val	Peptides	216.147	8.63E-05	1.108	1.225

Table 3.1. Top 10 statistically significant metabolites in normalised dataset between control and treatment groups. Positive fold-changes show the elevation of metabolites in treatment group compared to control.

3.3.2 Human brain tissue study

The metabolome of human brain tissue samples (both white and grey matter) showed a wide range of metabolites from various biochemical classes including amino acids, carbohydrates and lipids. There were 692 putatively identified metabolites found in common between grey and white matter, which were detected in all replicates and could therefore be used further for normalisation. Before normalisation, only 35% of metabolites (in both grey and white matter) expressed their peak intensity %RSD variation between **biological** replicates as less than 30%, with the rest showing a huge variation (>30% RSD) – this observation was expected to observe, as the extraction of intracellular metabolites from human brain tissue is often a complicated process involving several steps, leading to low metabolite reproducibility and issues in their detection in LC-MS. Different amount of water content in wet tissues and possible metabolite degradation during extraction contribute to low metabolite reproducibility. Hence, the variation in peak intensities between the replicates was observed to be high.[137] Variation among analytical replicates was also observed but not to such a great extent. QC replicates have shown that for 420 metabolites out of 692, %RSD was below 30%.

After normalisation, around 70% of analysed metabolites (506 in grey matter, 474 in white) have been successfully normalised by either their ^{13}C paired peaks or by the presence of other co-eluted ^{13}C peaks from *E. coli* extract, with their %RSD (among **biological** replicates) reduced significantly – for some metabolites by more than half. Of the successfully normalised metabolites, around 75% (380 metabolites for grey, 362 for white) had their %RSD below 30%. For **analytical** variation, %RSD was reduced in almost all of metabolites (95% in grey and 92% in white matter group). The detailed %RSD values for analytical replicates are shown in table A6 (appendix). Examples of some successfully normalised metabolites, which showed reduction in both analytical and biological data variation, are presented in figure 3.6 (A). Most of these normalised metabolites were of polar origin, such as common amino acids (l-arginine, l-proline, l-citrulline), nucleotides (cytosine, guanine, ATP) and other important polar compounds. This showed that the proposed method was working again, effectively normalising the metabolome data and reducing high variation in metabolite peak intensities between sample replicates.

The remaining metabolites (around 30%) did not manage to show the reduction in biological replicates data variation, either showing no change in %RSD or instead, an unexpected increase, indicating unsuccessful normalisation. These metabolites mostly included lipids and non-polar compounds, for which there was a limited number of internal standards from *E. coli* extract in the first place, as the bacterium was extracted in polar conditions, therefore it was expected that

poor normalisation in lipophilic metabolites would be observed. Other poorly normalised metabolites included several nucleotides (FAD, GDP, 3'-AMP, ATP, d-TMP), few amino acids and carbohydrates. For these metabolites, there were also no ^{13}C paired peaks or co-eluted internal standards detected which could have reduced the %RSD. As so, those unsuccessfully normalised metabolites were left non-normalised. This demonstrated that this method has its limitations.

In figure 3.6 (B), the following box and whisker plots show the whole range of % RSD values (biological replicates) of metabolites (grey and white matter) normalised by the proposed ^{13}C technique, as well as by MSTUS and TIC methods. The ranges were compared to each other and to a non-normalised range. From both plots it can be seen that the %RSD for a detected metabolome has been effectively reduced by ^{13}C normalisation, compared to non-normalised range, as was the expected observation. The median and upper/lower quartile values for the ^{13}C normalised range were lower than for the non-normalised range (median reduced from 37.5% to 22.3%) with the overall range being less variable. Compared to MSTUS and TIC normalisations, their %RSD ranges were observed to be less variable than the non-normalised range, with median values reported to be 30.1% and 32.4% for MSTUS and TIC, respectively. It can be seen that the MSTUS technique normalised metabolites more effectively than TIC (having less variable range). This demonstrated that the proposed ^{13}C normalisation technique managed to reduce the %RSD for most metabolites (in both grey and white matter) more effectively than both MSTUS and TIC methods. Several %RSD values in ^{13}C normalised data (13 and 7 metabolites in grey and white matter respectively) were shown to be above 100%, higher than the non-normalised maximum value, which showed that some metabolites were not normalised effectively and therefore displayed worse percentage values (outliers). Majority of those metabolites were of a lipid metabolism and classes including few carbohydrates and biosynthesis of secondary metabolism. See pages 88-90 for box and whisker plots explanation.

A normalisation effect has also been observed by PCA and OPLS-DA. In figure 3.7, both non-normalised (unlabelled) and ^{13}C normalised OPLS-DA plots show clear separation between the two data groups (white and grey matter). This separation was the expected observation due to different metabolic profiling of the two brain matter types.[316] In the non-normalised OPLS-DA plot, the separation between the two data sets was not as clear as in the "normalised" case, with lower Q2 value observed (Q2= 0.85 for non-normalised). Analysis of ^{13}C-normalised brain sample metabolites markedly improved the sample clustering on the OPLS-DA score plot, increasing the Q2 score to 0.969. In the non-normalised PCA model, the two data groups were separated clearly, indicating that there was a huge variation present in metabolite data throughout the samples. The QC replicate samples for normalised metabolome have shown that overall 514

metabolites have shown %RSD below 30%, compared to 420 metabolites in non-normalised samples.

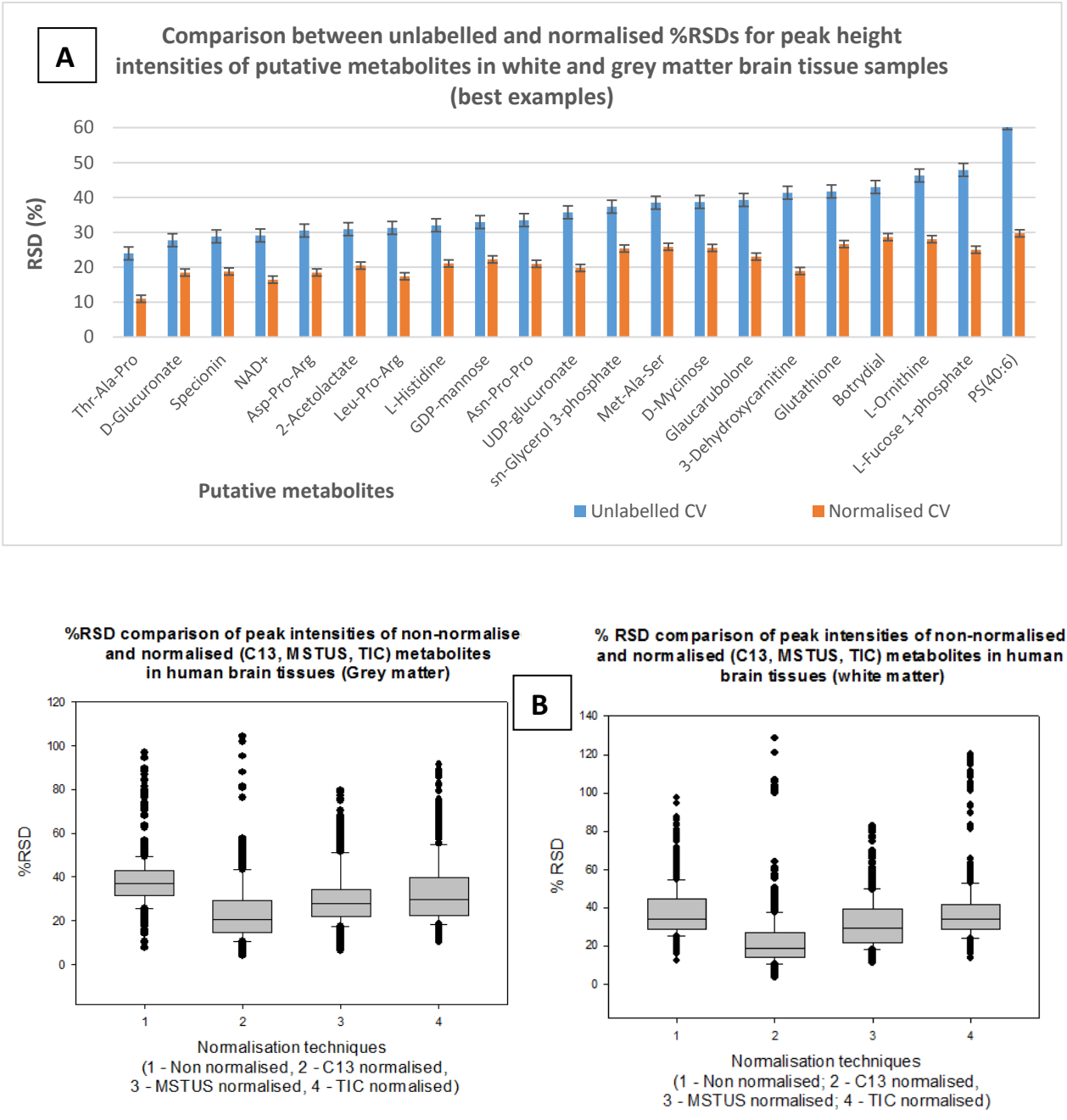


Figure 3.6. Comparison of %RSD between unlabelled and normalised metabolomes. Comparison with best metabolite examples (A). Comparison of %RSD range across the whole metabolome (B).

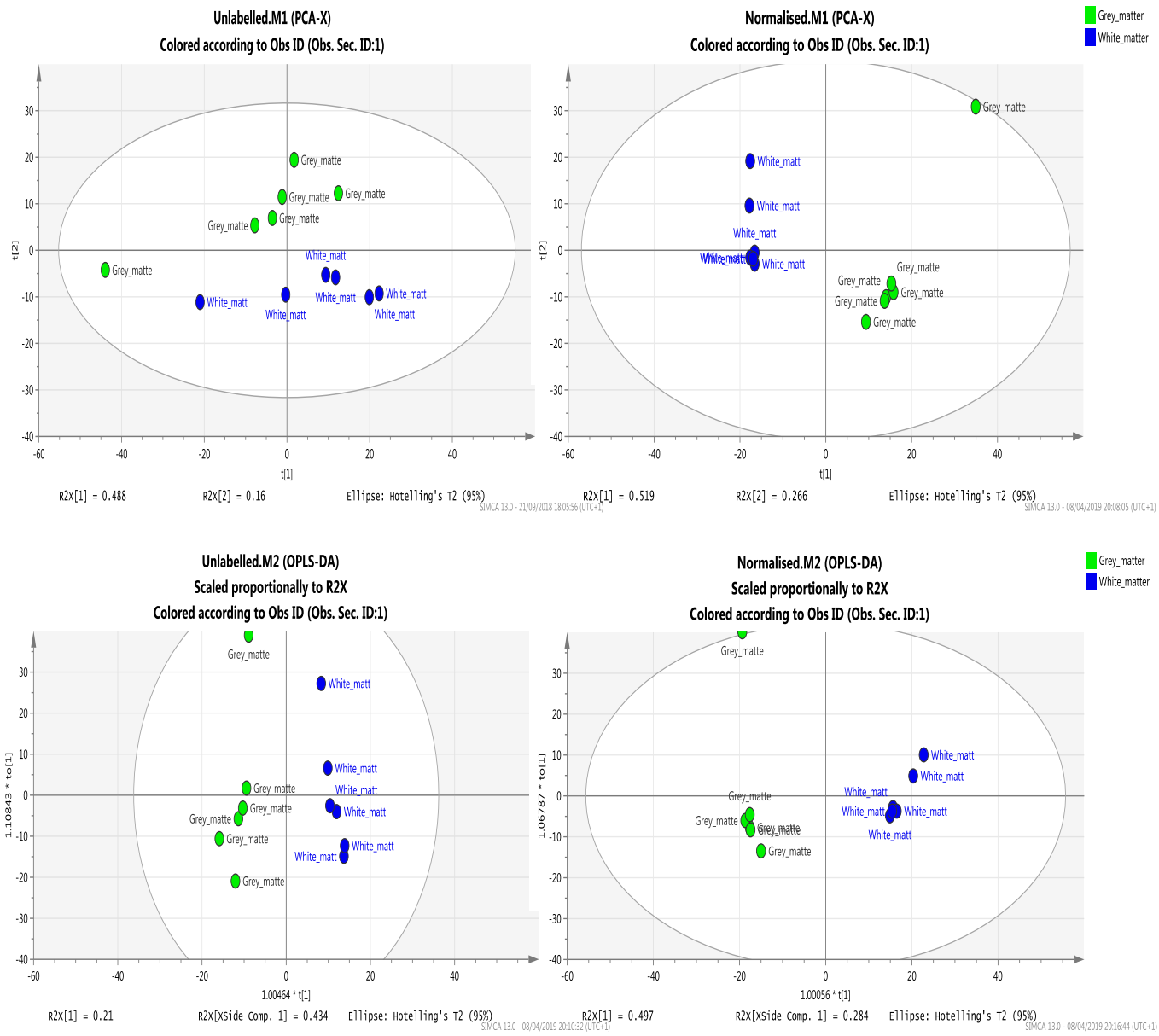


Figure 3.7. Multivariate analysis of human brain white and grey matter data groups using PCA and OPLS-DA score plots for non-normalised (unlabelled) and normalised datasets (green replicate points represent grey matter, blue replicate points represent white matter. PC1 and PC2 % variance for non-normalised plot = 48% and 46%, for normalised plot = 52% and 26%, respectively. OPLS-DA Q2 and R2 scores for non-normalised plot = 0.85 and 0.93, for normalised plot = 0.969 and 0.99, respectively.

Statistically significant metabolites

Univariate analysis through T-test followed by FDR statistical correction, has shown that 585 normalised metabolites were statistically significant with their p value being less than 0.05 and large fold changes observed. Almost the same result was observed for non-normalised data, with 560 statistically significant metabolites detected. This time normalisation did not have a huge effect on the detection of significantly changed metabolites, even though the data variation was shown to be reduced. The nature of statistically significant metabolites in non-normalised and normalised sets were very similar (a lot of common metabolites). This showed that even by reducing the metabolite data variation, it did not help to increase the number of statistically significant metabolites, possibly because the rest of normalised metabolites did not show a difference in fold changes or was any difference between grey or white matter. Volcano plots in figure 3.8 demonstrate the number of statistically significant metabolites in non-normalised and normalised samples, showing no real difference in the number of statistically significant metabolites (above the log p value line – p value less than 0.05).

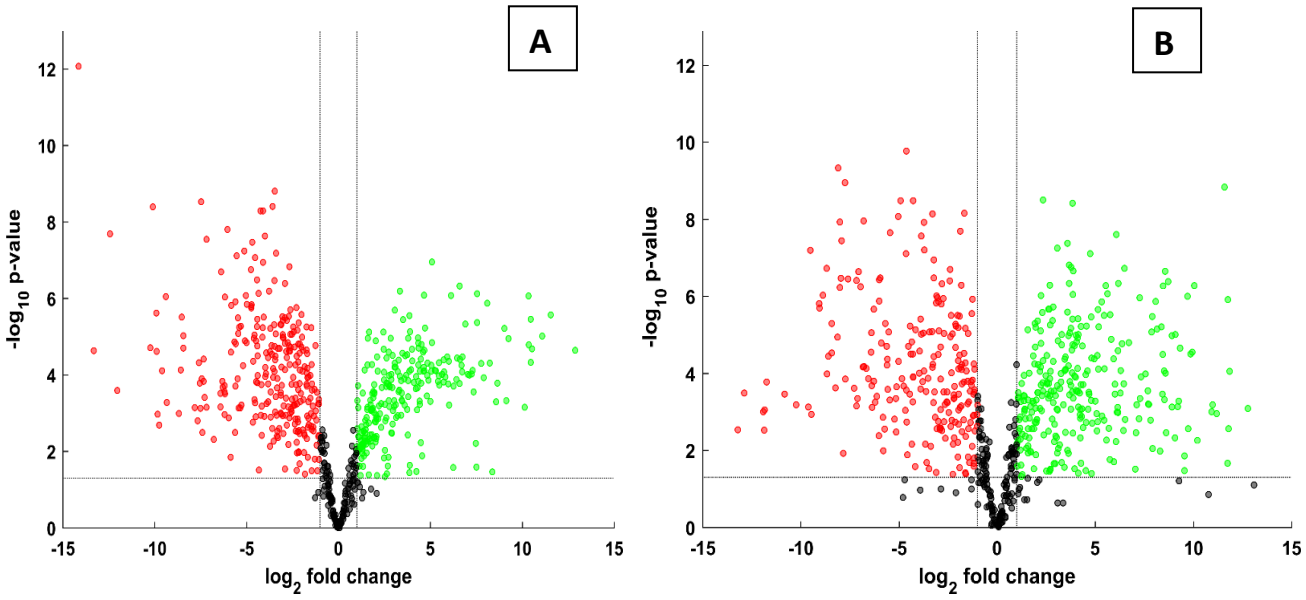


Figure 3.8. Volcano plots showing the number of statistically significant metabolites with their fold changes in non-normalised samples (A) and normalised samples (B).

From multivariate analysis, a large number of metabolites from the normalised OPLS-DA model were observed to have a VIP score of 1.0 or above – 361 metabolites. When compared to 585 normalised statistically significant metabolites from univariate analysis, the common list of statistically significant metabolites turned out to be 290. The majority of these significantly

changed metabolites were lipids, peptides and amino acids, all of which are normally present in large quantities in white and grey brain matter, especially lipophilic compounds. The fold changes between grey and white matter were mixed, with half of the metabolites showing high fold changes and half showing low fold changes (a decrease in concentration). This indicates that the comparison of white and grey matter metabolites is complex. As this study did not contain a real treatment group and a control group was difficult to assign (white and grey matter were treated as equal), the conclusion on brain matter effect is complex to interpret. Top 10 statistically significant metabolites are presented in table 3.2.

The number of statistically significant metabolites (which passed both univariate and multivariate analysis) in non-normalised dataset was 153 metabolites, which was a lot lower than for normalised dataset, showing once again that ^{13}C normalisation could achieve higher accuracy results.

Metabolite	Biochemical Class	Molecular mass	P value	VIP score	Fold change (Grey/White)
Deoxyinosine	Nucleotides	252.085	<1.00E-06	1.794	0.982
Triethanolamine	Lipids	149.105	<1.00E-06	1.463	0.672
5-Oxopentanoate	Amino acids	116.047	9.516E-05	1.134	- 1.019
Ala-Pro-Pro	Peptides	283.153	4.535E-08	1.523	0.879
Asp-Gln-Gln-Pro	Peptides	486.205	<1.00E-06	1.309	- 1.286
Berberastine	Others	351.113	1.252E-10	1.748	- 0.925
2-Oxoglutaramate	Amino acids	145.035	6.545E-08	1.434	1.074
[FA (17:1/2:0)] 8E-Heptadecenedioic acid	Lipids	298.214	4.264E-04	1.532	- 0.473
Oxalosuccinate	Carbohydrates	190.011	2.35E-05	1.788	1.613
ADP	Nucleotides	116.047	<1.00E-05	1.223	1.102

Table 3.2. Top 10 statistically significant metabolites in normalised dataset between white and grey matter groups. Positive fold-changes show the elevation of metabolites in grey matter group compared to white.

3.3.3 Mouse plasma study

The final validation study was performed on animal plasma samples (female mouse), in order to observe whether normalisation is working as effectively as in human models. Initial results with IDEOM showed the detection of 750 putatively identified metabolites, in common between the 3 and 18 month age groups. A large proportion of those metabolites were amino acids, carbohydrates and other polar biological classes. After normalisation, 85% of metabolites in the 3-month group and 92% in the 18-month group were successfully normalised by their ^{13}C paired or co-eluted labelled metabolites, with an observed decrease in %RSD between **biological** replicates. The remaining metabolites (15% and 8% in 3 and 18 months groups) had their biological %RSD increased or stayed the same, the majority of which were lipids, for which there was no ^{13}C paired or appropriate co-eluted metabolites to reduce variation. Almost all metabolites (94% in 3-months and 97% in 18-months) had their %RSD between analytical replicates also dropped – see appendix (Table A7). Examples of some successfully normalised metabolites (for both analytical and biological replicates) are shown in figure 3.9 (A). Although a large number of metabolites was successfully normalised, demonstrating that the proposed normalisation method was working once again, the metabolite's %RSD decrease (among analytical and biological replicates) was relatively small. Before normalisation, around 500 metabolites in both groups already had %RSD of < 30%, indicating that most non-normalised metabolites did not have a huge variation in their peak intensity values. After normalisation, 100 more metabolites expressed %RSD of < 30% (around 600 in each group). This showed that the normalisation technique did not reduce %RSD as effectively as in previous studies, possibly due to the low data variation between non-normalised metabolites in the first place, which was not observed in brain tissue or urine studies. Therefore, after normalisation, the effect was not very significant. Such a small metabolite variability detected in the non-normalised state could be due to the nature of the samples (mice plasma with similar genotype, weight and environment throughout their lives) and the extraction technique that did not involve many steps, reducing possible variation at the beginning.[317]

In figure 3.9 (B), %RSD range (for **biological** replicates) is shown to be different between non-normalised and ^{13}C normalised metabolomes, with median values of 18% after normalisation and 25% before (for the 3-month group); 18% after normalisation and 23% before (for the 18-month group). The outliers indicated those metabolites which were not successfully normalised and had their %RSD increased instead of decreased, however there were not a lot

of outliers detected compared to previous two studies. Among outliers were few lipids, carbohydrates and two peptides. Even though a significant difference between non-normalised and ^{13}C normalised data could not be seen, ^{13}C normalisation has still shown better results than MSTUS or TIC methods. See pages 88-90 for box and whisker plots explanation.

Figure 3.10 shows multivariate analysis of mice plasma data. A clear separation between data groups (3 and 18 months of age) was observed in non-normalised and normalised OPLS-DA models with high Q2 values (0.927 and 0.997, respectively). This was the expected observation, as the %RSD of non-normalised metabolites was already small, so the cluster of replicate points on both OPLS-DA models was similar. In all four plots (non-normalised and normalised PCA and OPLS-DA models) the separation between the groups was clear. In this validation project, normalisation did not have much influence on data variation between replicates or improve the accuracy of the results.

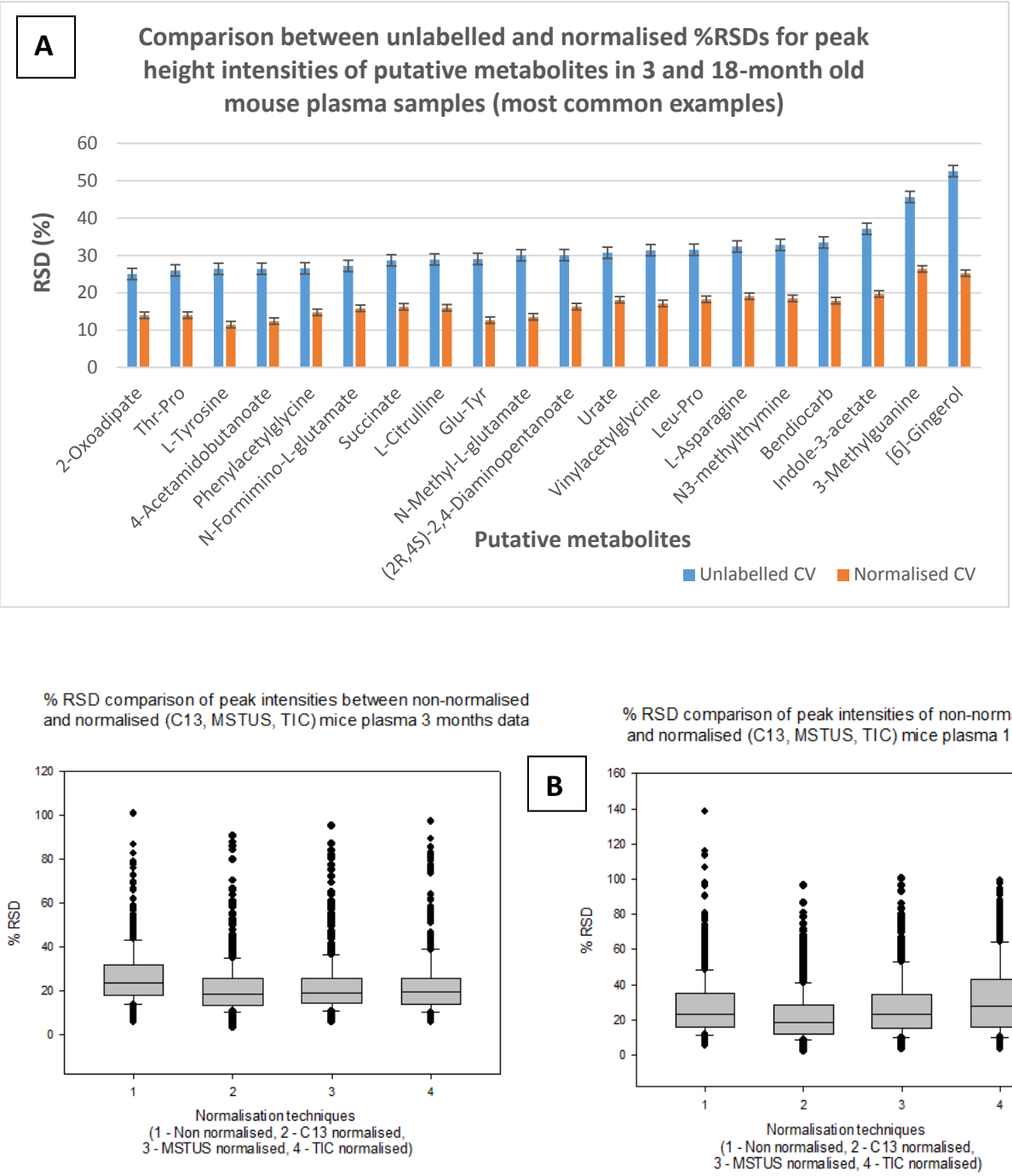


Figure 3.9. Comparison of %RSD between unlabelled and normalised metabolomes (A). Comparison with best metabolite examples (B). Comparison of %RSD range across the whole metabolome

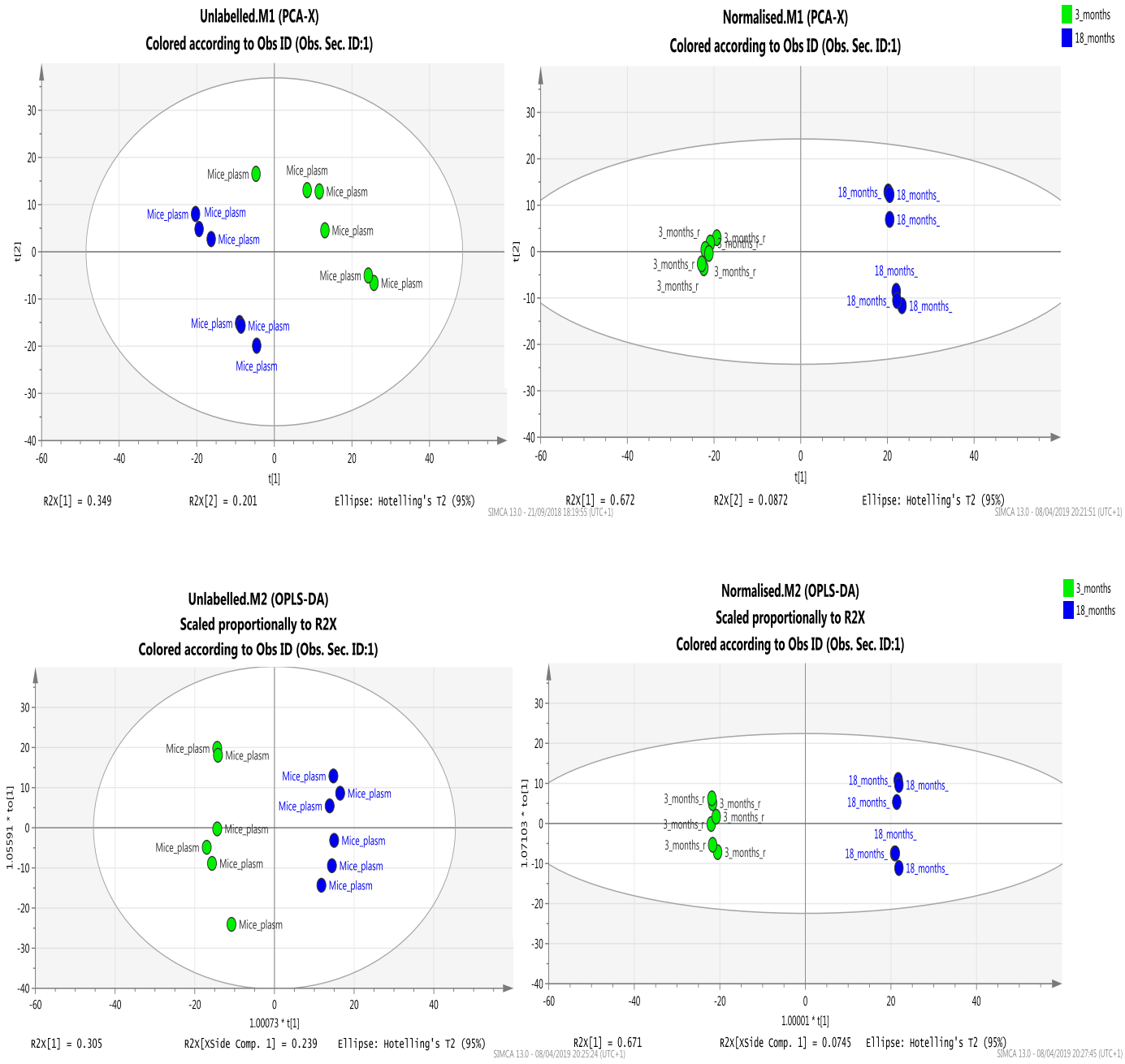


Figure 3.10. Multivariate analysis of mouse plasma 3 and 18 months age data groups using PCA and OPLS-DA score plots for non-normalised (unlabelled) and normalised datasets (green replicate points represent 3 months age, blue replicate points represent 18 months age. PC1 and PC2 % variance for non-normalised plot = 49% and 21%, for normalised plot = 55% and 32%, respectively. OPLS-DA Q2 and R2 scores for non-normalised plot = 0.92 and 0.95; for normalised = 0.97 and 0.99, respectively.

Statistically significant metabolites

Univariate analysis through T-test, followed by FDR test, has shown that 514 normalised metabolites were observed as statistically significant, with their p values less than 0.05 and large fold changes. Non-normalised data has shown similar number of statistically significant metabolites (465). Once again with the help of normalisation, a rise in number of significantly changed metabolites was detected. The volcano plots in figure 3.11 demonstrate the number of statistically significant metabolites in non-normalised and normalised samples.

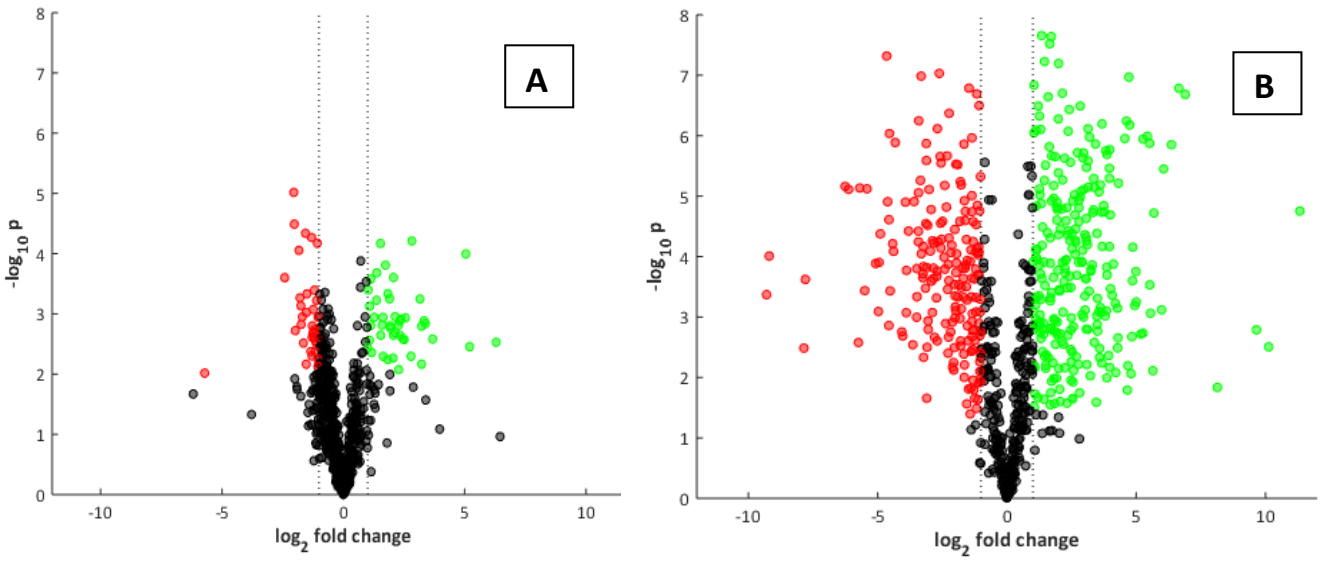


Figure 3.11. Volcano plots showing the number of statistically significant metabolites with their fold changes in unlabelled samples (A) and normalised samples (B)

From multivariate analysis, a large number of metabolites from the normalised OPLS-DA model were observed to have a VIP score of 1.0 or above – 487 metabolites. When compared to 514 normalised statistically significant metabolites from univariate analysis, the common list contained 338 metabolites. The majority of those significantly changed metabolites were lipids, carbohydrates and other biochemical classes, all of which are commonly present in plasma samples. Even though, with the normalisation, a larger number of significantly changed metabolites could be seen, it was difficult to interpret the difference in fold changes. There doesn't seem to be a clear conclusion as to why or how metabolites were affected by the different age groups of mice, or if age of mice has had any effect on metabolite concentrations. Table 3.3 shows the top 10 statistically significant metabolites with variety of fold changes. The number of statistically significant metabolites (which passed both univariate and multivariate

analysis) in non-normalised dataset was 276, lower than in ^{13}C normalised dataset, showing that normalisation technique helped to detect more statistically significant metabolites.

Metabolite	Biochemical Class	Molecular mass	P value	VIP score	Fold change (18 months/3 months)
[PR] Vitamin E	Lipids	459.477	1.288E-04	1.153	1.029
15-Keto-prostaglandinE2	Lipids	370.275	1.34E-06	1.077	1.514
2-Methylcitrate	Carbohydrates	213.066	7.79E-07	1.109	- 0.802
Hydroxytetradecanedioic acid	Lipids	288.224	<1.00E-06	1.015	1.678
D-Glucono-1,5-lactone	Carbohydrates	184.067	<1.00E-06	1.109	1.281
Cer(d18:1/23:0)	Lipids	676.757	3.31E-05	1.126	- 0.616
Glucosylceramide (d18:1/16:0)	Lipids	739.698	<1.00E-06	1.148	0.936
Glycerophosphoglycerol	Others	252.073	3.68E-05	1.113	- 0.698
LPA(0:0/18:1(9Z))	Lipids	457.328	3.92E-07	1.152	1.032
Myristoylglycine	Others	301.283	2.18E-05	1.141	3.778

Table 3.3. Top 10 statistically significant metabolites in normalised dataset between 3 and 18 months age groups. Positive fold-changes show the elevation of metabolites in 18 months age group compared to 3 months age.

3.4 Conclusion

The normalisation technique, using U-^{13}C-labelled *E. coli* extract as a source of internal standards, has fully demonstrated the effectiveness of reducing **analytical** variation in peak height intensity data for a large range of metabolites in three different types of samples. Reduction in **biological** data variation has also been observed with excellent results, mainly in human urine and brain projects. In all three validation experiments, a large number of putative metabolites were successfully normalised – their %RSD of peak height intensities between analytical and biological replicates were reduced significantly, in some cases by more than 50%, showing a reduction in whole data variation. The experiments also demonstrated improvement of within-individual sample clustering on normalised OPLS-DA models with high Q2 scores, mostly shown in human urine and brain tissue studies. When statistical analysis involving univariate and multivariate techniques were applied to the non-normalised and normalised data, a large number of statistically significant metabolites were detected, especially in normalised data, proving once again that normalisation helped to improve the data accuracy and detect more significantly changed metabolites. In comparison to different normalisation techniques such as MSTUS, TIC and creatinine, the proposed ^{13}C technique was shown to reduce data variability much more effectively.

With the help of ^{13}C normalisation technique in reducing analytical and even biological data variation, it may serve a huge benefit when applied in untargeted metabolomics to a wide range of different mammalian tissues and biofluid samples. As demonstrated through validation experiments so far, this method has achieved excellent results in improving data quality of human urine, mouse plasma and human brain tissues, and even managed to help detect more number of significantly changed metabolites than in non-normalised data.

In the past, ^{13}C normalisation has been more applied in targeted studies for absolute quantification of metabolites and measuring metabolic fluxes, showing successful results.[189,196,208] Very rarely that the studies employed ^{13}C isotopes in untargeted metabolomics, and if so, for different purposes rather than for normalisation of a wide range of mammalian metabolites. One study has used a software, X^{13}CMS, to track ^{13}C isotopic labelling patterns of a wide range of metabolites in untargeted study with rat brain tumour samples, however this technique was not used as a normalisation but rather like a detection study, revealing significant labelling patterns of metabolites to help monitor metabolic pathways.[318] Another study demonstrated the use of ^{13}C normalisation in a wide range of metabolites in marine

environmental samples for ecological metabolite profiling, rather than in mammalian samples, normalising data effectively.[206]

Therefore, there is a gap in metabolomics research for an effective accurate normalisation method using stable ^{13}C labelled internal standards in **untargeted mammalian** metabolomics that can be applied in clinical studies for biomarker discovery and other related biological questions. The validated proposed ^{13}C technique, using an isotopically labelled bacterial extract as a source of labelled metabolites, has shown a promise of an effective normalisation in a wide range of metabolites in human urine and brain tissues, as well as in mouse plasma, so is now ready to be applied in future clinical untargeted studies.

Chapter Four

Urine and plasma metabolomics analysis following an oral saturated fat meal intervention study in healthy volunteers

4.1 Introduction

4.1.1 Dietary intervention: its purpose and complications

Nutrition (or dietary) interventions can have a positive impact on long term treatment of serious human health conditions. The purpose of this class of interventions is to resolve or improve the nutrition diagnosis and associated problems by providing an advice, education or delivery of food component of specific diet tailored to the needs of the patient.[319,320] Dietary intervention usually focuses on adjustment of personal diet of eating certain types of food, to improve individual's overall health, part of so called "behavioural intervention".[321] Additional nutrients provided separately from diet, such as supplements (vitamins and minerals medications) have also assisted healthcare professionals in treatment of specific diseases in patients. Generally, poor diet in a human organism may lead to complications later on such as malnutrition[322], diabetes[323], osteoporosis[324], renal[325] and cardiovascular problems[326], as well as some sorts of cancer.[327,328] These complications may arise due to lack or gain of certain food nutrient(s), for example rise in glucose in diabetes or decrease in vitamin D in osteoporosis. According to World Health Organisation (WHO), the nutritional targets by 2025, which were endorsed by organisation in early 2013, are to greatly reduce malnutrition, diabetes and other associated complications by providing excellent dietary intervention plans – expected reduction in malnutrition in 40% world child population under 5 years of age, 50% reduction of anaemia in women of reproductive age and a stop in rise of diabetes and obesity, as well as many other aims.[329]

Dietary intervention usually belongs to a non-pharmacological treatment, where therapeutic medications are not used and instead dietary supplements together with diet change is applied. There is some evidence that specific non-pharmacological therapies, such as cognitive training, may assist in dietary intervention programme for patient, by improving cognitive function, behaviour, mood and quality of life of an individual to help to make the right decisions about balanced healthy diet.[330] Self-management of chronic illness by regulating a diet through intervention is seen as a full responsibility of a patient, and as such, presents with complications including non-compliance (not following an intervention plan thoroughly). With this in mind, patient may find it difficult to control the diet and follow the intervention plan correctly, hence affecting their treatment.[331]

4.1.2 Effect of dietary intervention on human metabolism in complicated diseases

As discussed in chapter one, human metabolism is hugely affected during complicated and long term diseases such as cancers, rheumatoid arthritis, osteoporosis, diabetes and other inflammatory conditions. Metabolites which are affected by these conditions are usually involved in important metabolic pathways such as glycolysis, Krebs cycle, immune response pathways, which regulate the physiological processes in human organisms. By changing certain dietary requirements or nutrients as a part of a therapy against disease, can influence biological levels and activity of certain metabolites in important pathways, and therefore can have an impact on the disease condition.

In diabetes type 2 condition, for example, nutritional intervention plays an important part, as it helps to regulate levels of glucose, high levels of which is believed to cause the condition. Healthy balanced diet and reduction in high saturated fats, sugars and high cholesterol may help to reduce glucose levels in blood. Diet intervention can play an important part in promoting weight loss among overweight individuals who are prone to develop diabetes or at pre-diabetes stage, therefore preventing the condition to develop.[332] Intervention involves avoidance of processed food (starches and sugars) and instead consuming low carbohydrate diet (vegetables, fruits, whole grains). However, several studies have demonstrated that nutrition is one of the difficult aspects in management of diabetes type 2, due to non-compliance from patients taking a strict diet over a long period of time. Consumption of meals with high fat content, in particular those containing saturated fat, can cause an impairment in the action of the hormone insulin to promote the use of glucose (sugar) in skeletal muscle (insulin resistance), a key element in the pathogenesis of type 2 diabetes, whereas liquid meals high in polyunsaturated fat content (such as fish oil) can be protective. However, little is known about the effects of high fat meals with different fat composition on muscle protein metabolism in sedentary middle-aged overweight/obese individuals, a population that is susceptible to type 2 diabetes and age-related decline in skeletal muscle mass.

Metabolites which are mostly affected by high fat diet are likely to be biomolecules such as enzymes, hormones and cell receptors acting in the pancreas (beta cells) and liver, regulating utilisation of glucose and insulin secretions, therefore may be potential biomarkers for related disorders such as diabetes, obesity, renal impairment. These metabolites are normally detected in mammalian urine and plasma but they might also be detected in endothelial tissues in a variety of organs, including skeletal muscles where metabolome-based pathways are altered

significantly.[333] Recent technological advances in "omics" studies have been collecting information on metabolites, lipids, peptides and other macromolecules over the years, in patients with diabetes types 1 and 2. They have managed to detect a variety of statistically significant metabolites between healthy and diabetes clinical samples in humans and animal models (urine, plasma, fat and muscle tissues), which were proposed as potential biomarkers. Some of the potential biomarkers were found to relate to vascular function or tissue damage caused by diabetes, with organ complications (impaired retina of an eye, heart blood vessels, kidneys, liver).[334,335] Therefore, it can be assumed to expect similar metabolic changes in healthy overweight individuals who are consuming high-fat meals and can easily develop pre-diabetes condition.[336]

The example of common statistically significant metabolites detected in individuals in diabetes, pre-diabetes or high-fat meal diet interventions are usually oxidised forms of molecules, known as "advanced glycation end-products" (AGEs), which cause oxidative stress in a variety of cells and tissues, resulting in molecular damage. Animal-derived foods that are high in fat and protein are generally AGE-rich and are prone to further AGE formation during cooking. Examples of AGEs include cholesterol (oxidised form of low-density lipoprotein, LDL), L-erythrulose and oxalate (oxidised products of L-ascorbate), and cis-aconitate (oxidised form of citrate). AGEs can cause further complications including cardiomyopathy, nephropathy, retinopathy and neuropathy, so their levels are constantly high in biofluid and tissue samples individuals with high glucose levels.[337] AGEs are known to interact with their cellular receptors such as RAGE, lactoferin and scavenger receptors. The activation of those receptors lead to oxidative stress and activation of mitogen-activated protein kinase (MAPK) signalling pathway.[337] In coronary arteries, glucose crosslinks with collagen protein and forms a complicated AGE, activating MAPK pathway and causing vascular stiffening of arterial walls, leading to cardiovascular issues such as stroke or coronary heart disease. It has been established that generally a high fat/glucose diet can induce metabolic syndrome, also known as constellation of metabolic disorders, and so can lead to oxidative stress and formation of AGEs, which may be present in other related conditions – dyslipidemia, obesity, hypertension. Therefore, AGEs are worth investigating as potential biomarkers in those related clinical conditions.[338]

4.1.3 Justification of the dietary intervention study

In this dietary intervention study, **"Skeletal muscle protein metabolism and insulin sensitivity in overweight individuals",** organised by the School of Life Sciences, University of Nottingham, several overweight but otherwise healthy individuals were recruited to provide their urine and plasma samples after consumption of a high fatty meal, **in order to observe the effects of this diet on metabolic activity in skeletal muscles**. The main aim of this clinical study was to investigate whether meals with a high fatty acid composition can impair the action of insulin in promoting the use of glucose in skeletal muscle, by giving healthy volunteers a liquid meal rich in saturated fat. It is thought that high saturated fat reduces the ability of muscles to synthesise new protein in response to dietary protein intake, which ultimately may compromise maintenance of muscle size and quality of life, as well as causing an impairment in the action of the hormone insulin, thus promoting conditions such as type 2 diabetes.[339] So far, little is known about the effect of saturated fat on muscle performance and insulin action in middle-aged overweight individuals, who are more predisposed to diabetes type 2. Therefore, this study was designed to answer this question and help scientists and clinicians to better understand the effect of high fat diet on skeletal muscle physiological quality.

This study involved a series of experiments, including collection of urine and plasma samples from those individuals, analysis using a variety of clinical tests (blood pressure monitoring, cholesterol levels test) and metabolic profiling experiments. For the purposes of this thesis, the proposed LC-MS method with an advanced (U)-^{13}C normalisation technique has been applied on plasma and urine samples to assist in searching statistically significant metabolites affected by this dietary intervention (saturated fat consumption) compared to control (no intervention).

The following objectives were set up for this study:

1. To extract and perform LC-MS analysis on urine and plasma samples, using (U)-^{13}C normalisation method that has been validated in previous chapter
2. To identify putative metabolites in urine and plasma samples
3. To identify urinary and plasma statistically significant metabolites which can differentiate between controls (no dietary intervention) and treatment ("intervention") samples, and give a rough idea of potential biomarkers for conditions like diabetes type 2.
4. To link those statistically significant metabolites to metabolic pathways and answer the biological question of the study with regards to effect on skeletal muscle performance.

4.2 Materials and methods

4.2.1 Study design and sample collection

Urine and plasma sample collection

As stated earlier, urine and plasma samples were collected over the course of two days from middle-aged, overweight, healthy male volunteers (40–60 years old) for this study. Before being accepted for the study, volunteers were carefully selected by undergoing medical screening (blood pressure, heart rate, weight and height measurements) and by meeting the health requirements – non-smokers and no history of diabetes or cardiovascular disorders. Following screening, six volunteers were selected. On the day of an experiment, each volunteer was asked to consume an oral standardised drink containing 30 g of milk protein and 75 g of sugar in 250 ml of water. The sugar in the drink was included to stimulate the secretion of the hormone insulin. Two hours later, each volunteer was asked to provide a urine sample (first time point collection), and then two more urine samples four hours later (second time point) and after a further four hours at the end of the trial (third time point). Blood (plasma) sample collection was done simultaneously, once every hour, including at the same time points as the urine collections. For the plasma collection, each volunteer was asked to lie on a bed throughout the experiment (about 10 hours), only getting up to produce urine samples. At the start of experiment, two small plastic needles were inserted into volunteer's vein at the back of a hand and the forearm of the contralateral arm, to collect blood samples at regular intervals – once every hour. The first time point collection was two hours after ingestion of a drink, then the collection of blood samples for four hours were pooled together to make the "second time point" collection and at the end, the "third time point" (the final four hours of pooled blood samples). For the control samples collection, each volunteer was given a drink of 250mL of water mixed with 75g sugar only (without milk proteins), and the collection of urine with plasma samples were performed in the same manner as described earlier, over a 10 hour period but on a different day. To make the study more unbiased and randomised, a volunteer did not know on which day he was given a "control" or "milk protein" drink. Each volunteer had to produce three "control" and three "dietary intervention" samples of urine and plasma – so in total, 12 samples from each volunteer. See figure 4.1 for more details of sample collection. During the study, one volunteer could not produce urine and plasma samples at the first time point when he was given a milk protein drink (intervention samples), so he only had five samples each of plasma and urine by the end of the study. The collected samples were aliquoted and kept in a -

80°C freezer. The whole study was reviewed and approved by the University of Nottingham, Medical School Ethics Committee.

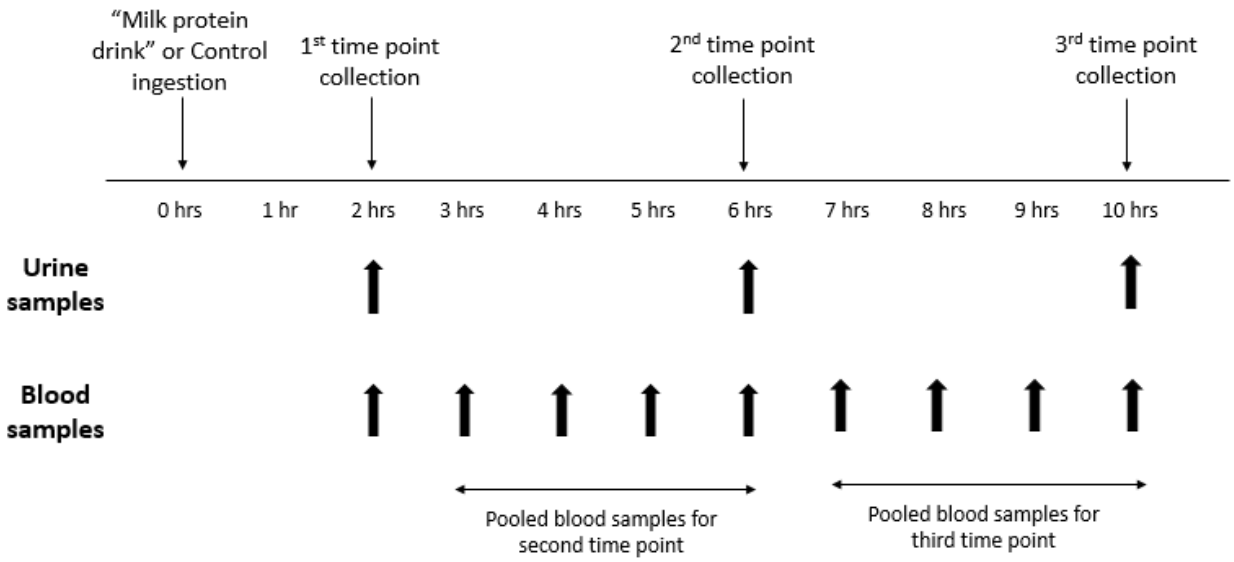


Figure 4.1. Simplified schematic diagram of urine and plasma samples collection at specified regular intervals from each volunteer, either taking control or intervention sample.

Sample extraction and preparation

For each experiment, the same number of biofluid samples and their replicates have been analysed. As mentioned earlier, one volunteer could not produce urine or plasma samples at the first time point collection in the main study, so he has only five samples of each biofluid, as opposed to the other five volunteers who managed to produce six samples of each biofluid. The collected biofluid samples were aliquoted into smaller volumes (2mL Eppendorf tubes) and kept frozen in -80°C conditions. The final samples were prepared by mixing a small amount of each aliquoted biofluid sample (0.5mL) with 0.5mL of (U)-^{13}C labelled *E. coli* extract (in 1:1 ratio) by pipetting them into fresh Eppendorf tubes, followed by vortex mixing for one minute and centrifugation at 13000g at 4°C for ten minutes. The supernatant was collected from each Eppendorf tube (sample) and transferred into fresh Eppendorf tubes (containing polar metabolites). From each of these Eppendorf tubes, a 200µL volume was transferred into an LC-MS vial, making the final sample. So in total, for each experiment (urine or plasma), there were 35 final samples, mixed with labelled extract – biological replicates in this study were the volunteers (n=6). The design of sample preparation is shown in figure 4.2. The fully prepared final samples (35 of urine, 35 of plasma) were stored in -80°C conditions before LC-MS analysis. The preparation of urine and plasma samples was done in a randomised order.

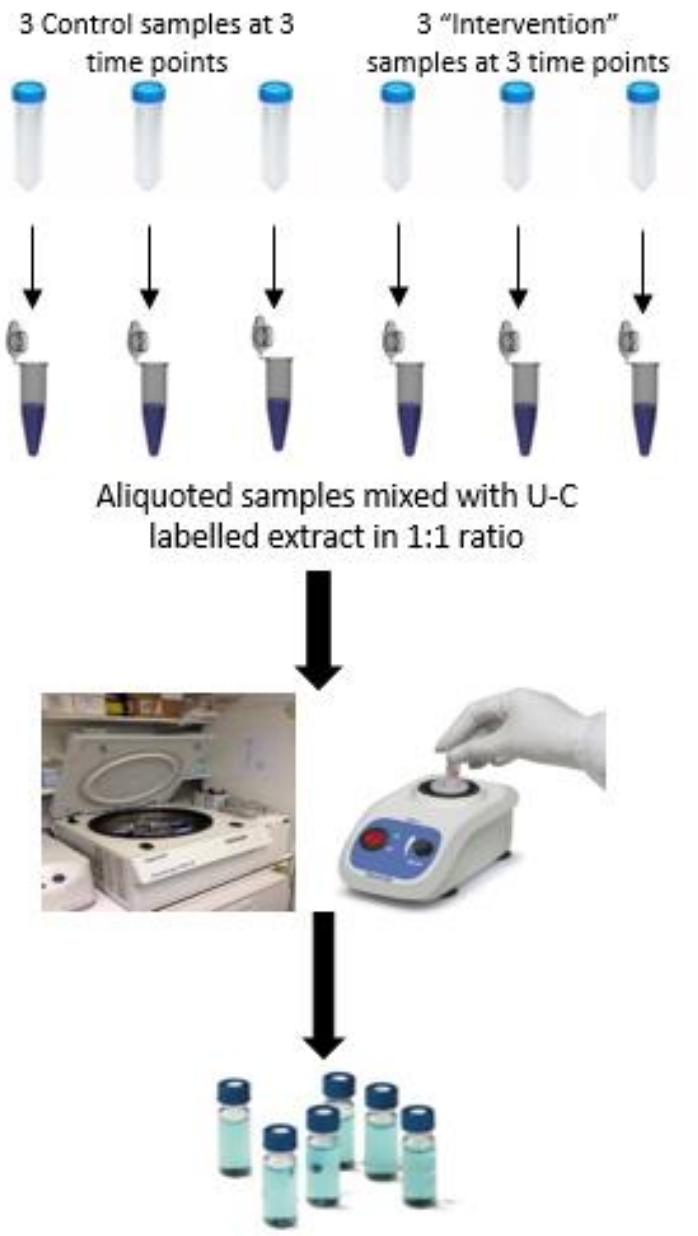


Figure 4.2. Schematic diagram of three control and three "intervention" (treatment) **urine OR plasma** samples prepared for one volunteer (biological replicate). The same procedure has been done for the other five volunteers.

4.2.2 Materials, chemicals and LC-MS analysis

Analytical LC-MS mobile phases (HPLC-grade acetonitrile, 20mM ammonium carbonate in purified water), zic-pHilic column, system blank (acetonitrile), and 268 authentic standards have all been used during the analytical run of samples. The LC-MS instrument was chosen to be Orbitrap MS ("Exactive") coupled with "Accela" autosampler LC, exactly the same as has been used in previous experiments. The description and sources of these materials, chemicals and instrument have been previously described in chapter two (sections 2.2.1 and 2.2.2).

The analytical run of urine samples was done on a separate day from the plasma samples run, with both sequences being randomised. For each metabolomics analysis (urine or plasma), a pooled QC sample was prepared by mixing an equal amount taken from each sample. In each experiment, six injections of pooled QC sample were analysed at the beginning of the run to condition (stabilise) the column before the analysis of randomised clinical samples. A pooled QC sample was then injected once after each randomised seven samples for the purpose of monitoring the stability, repeatability, robustness and performance of an analytical run. A reagent blank (methanol) was injected alongside every QC to minimise any possible carryover effect.

4.2.3 Data processing and statistical analysis

For each experiment (urine and plasma), an identical data processing method with the same software has been applied to a data set: "XCalibur", "mzMatch", "IDEOM" processing software for metabolite putative identification, a normalisation analysis using Excel and Python tools, multivariate analysis using "Simca P" (PCA and OPLS-DA plots, VIP score calculation) and univariate analysis (T-test, FDR correction, volcano plots) for key metabolite identification. The details of the software analysis are described in chapters two (section 2.2.3) and three (section 3.2.4). The overall results (particularly a list of key metabolites) from plasma and urine experiments were then compared to each other to see the similarities and draw an overall conclusion.

In each experiment, the raw data files from each sample and biological replicate were grouped according to a collected time point and condition (control or intervention). For example, in the urine experiment, in order to do a correct data processing in IDEOM, 35 raw data files (samples) were grouped into three control groups (1st, 2nd and 3rd time points) and three intervention

groups (1st, 2nd and 3rd time points) with six biological replicates (volunteers) in each group – except in one group, the "1st time point intervention" had five biological replicates, as one patient could not produce a sample. Once the list of putatively identified metabolites was generated, showing a number of detected metabolites in each condition group, further analysis was performed on metabolite peak height intensities for normalisation. The same approach as was performed in chapter three in the validation experiments, involved generating ratios of $^{13}C/^{12}C$ peak height intensities for each metabolite in each biological replicate for each group. ^{13}C peaks were identified using the proposed Python code, used in the validation experiments. The assessment of normalisation, by calculating percentage RSD values, was performed to check the quality of the data, but it was not compared to "non-normalised" or other normalisation techniques, as it is beyond the scope of this chapter.

4.3 Results

4.3.1 Metabolomics analysis of urine and plasma samples

LC-MS analytical performance

Analytical runs of both urine and plasma samples have generated clear spectra with distinctive metabolite peaks, high TIC values (≥ E^8) and no background noise or carryover effects. Spectra of QC samples in both experiments also showed clear peaks with high TIC values, demonstrating a good analytical stability of sample metabolites and reproducibility. The shape of the spectra peaks of QC samples was found to be similar to the control and treatment spectra, with some peaks identical, showing that the QC samples contained the same common metabolite peaks as the actual samples. Examples of some spectra graphs are shown on the following pages in figures 4.4 and 4.5, for urine and plasma metabolomes, respectively.

Metabolite profiling of urine samples

There were 750 putative metabolites detected overall in human urine samples (679 of them were found in both control and treatment groups). The majority of detected metabolites in all analysed urine samples were amino acids, peptides, carbohydrates and other biochemical polar classes, with a certain proportion of lipids (hydrophobic) class. The composition of biochemical classes is shown in figure 4.3, above. It was expected that a large number of hydrophilic compounds would be observed, as the extraction of urine samples was performed in hydrophilic conditions with a polar extraction solvent (U-^{13}C labelled *E. coli* in methanol).

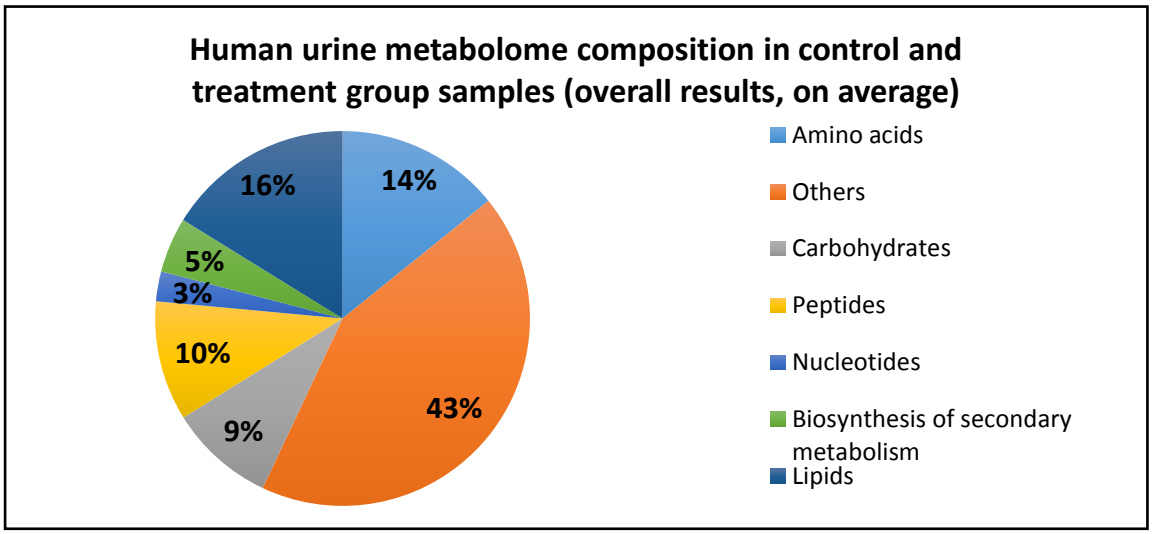


Figure 4.3 Composition of putative metabolites in human **urine** samples (control and treatment groups) according to biochemical class.

For multivariate analysis, the decision had to be made on how to compare the control and treatment (intervention) groups effectively, in each of plasma and urine experiments. It was decided to compare each time point in turn, for example "Control" 1st time point data set with "Intervention" 1st time point data set, and so on. So, there were three matrix tables generated for each time point data set, in both plasma and urine experiments. Based on these matrix tables containing normalised data (ratios), PCA and OPLS-DA plots were constructed in the same manner as in the validation experiments. VIP score analysis was done in every OPLS-DA plot, choosing metabolites with a score of 1.0 or above only.

Univariate analysis was performed on the same grouped data as used in multivariate analysis for each experiment using "Prysm" software, creating a list of statistically significant metabolites between control and treatment groups for each time point (three lists for plasma and three lists for urine experiments). The parameters for FDR and T-test were chosen to be the same as in the validation experiments in chapter three (p value of less than 0.05, FDR rate 5%). The lists were then compared to their corresponding multivariate VIP score lists to produce final lists of statistically significant metabolites. These metabolites were then compared to literature to see if any of them are normally affected by high fat dietary intervention, as well as comparing them between urine and plasma data to see the similarity. The fold changes obtained from univariate analysis and IDEOM were used to decide whether those metabolites were elevated or decreased during the "intervention" condition compared to the control.

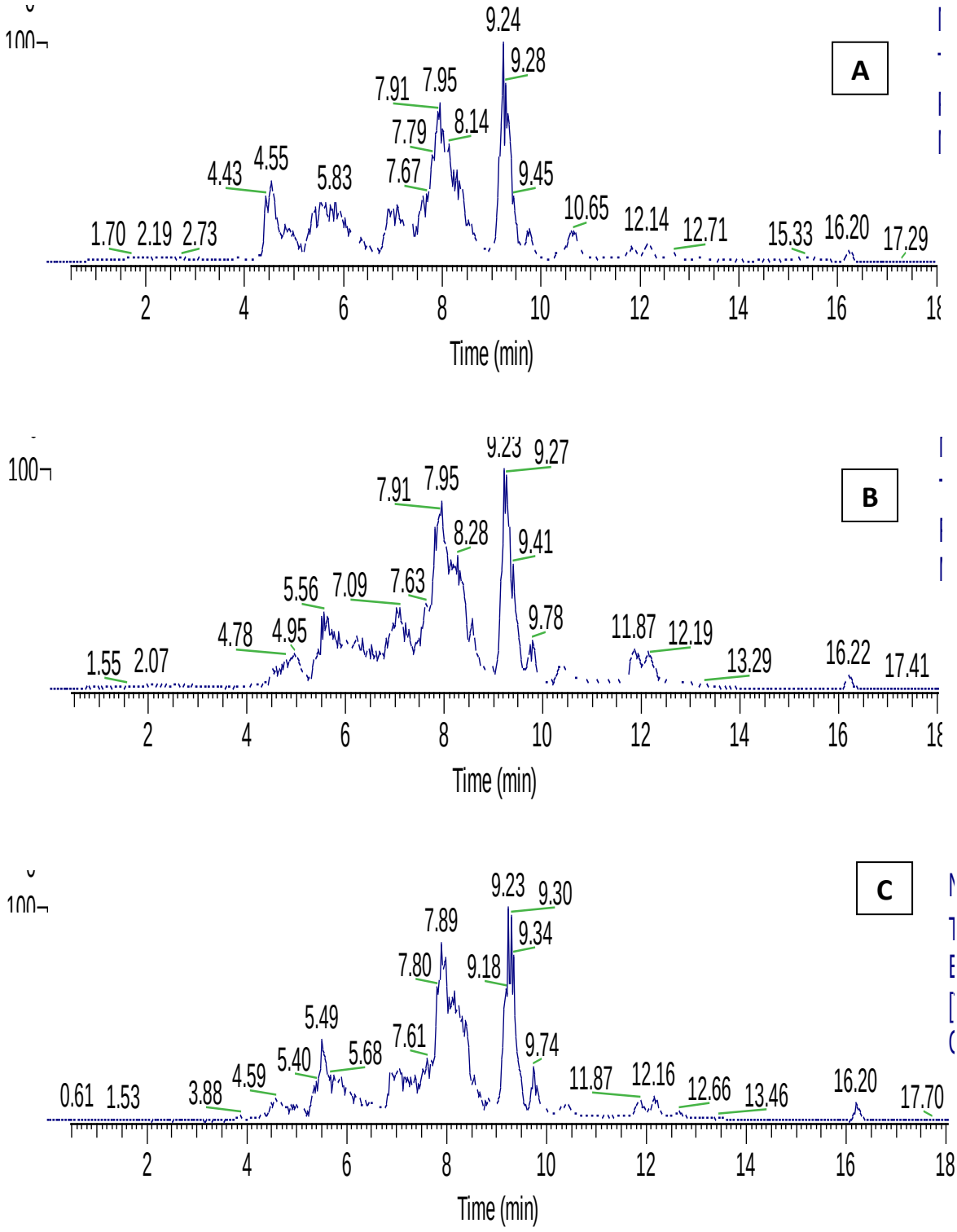


Figure 4.4 Examples of LC-MS spectra of human **urine** samples. Control sample third time point (A) Treatment (intervention) sample third time point (B) QC sample (C).

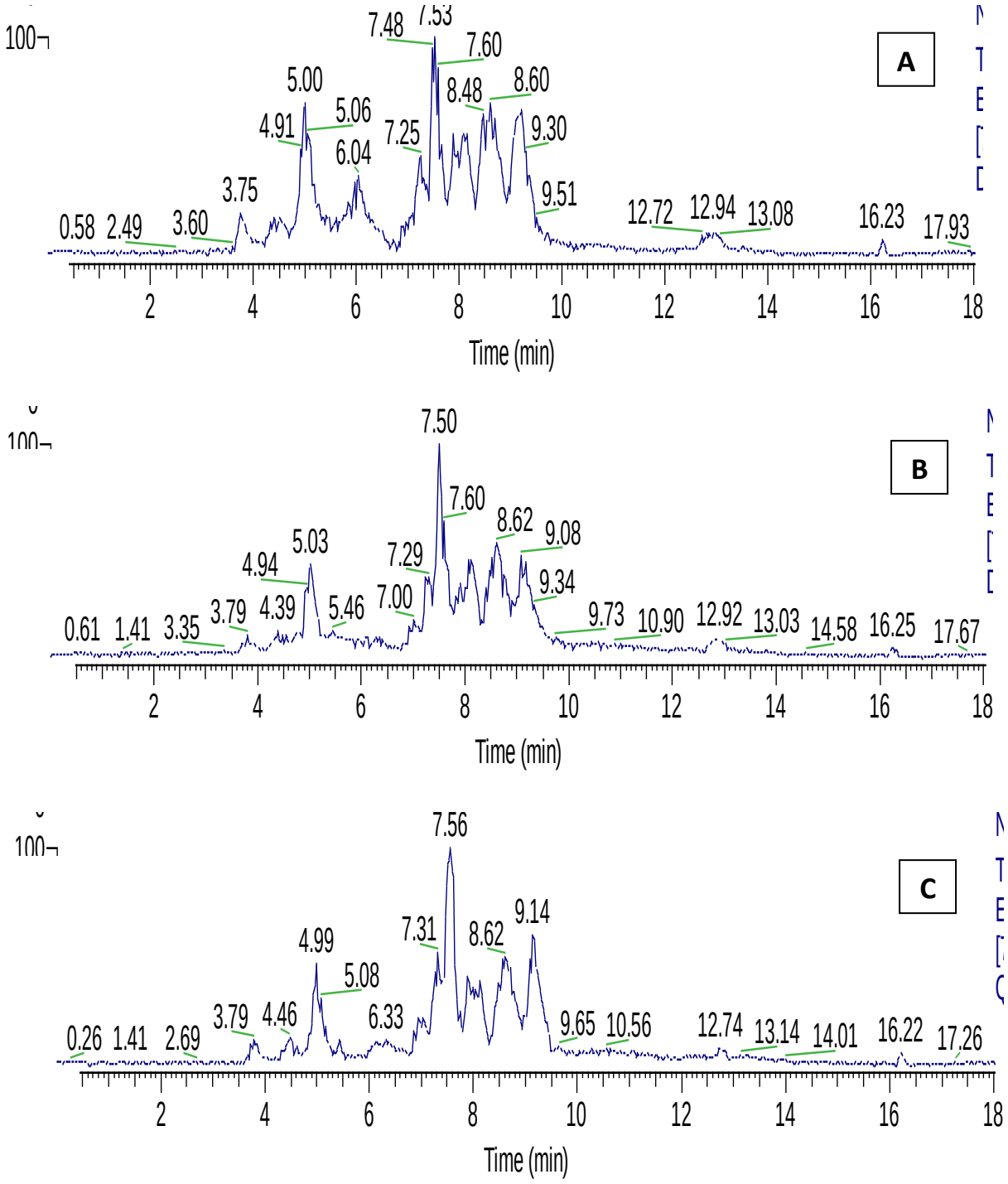


Figure 4.5 Examples of LC-MS spectra of human **plasma** samples. Control sample third time point (A) Treatment (intervention) sample third time point (B) QC sample (C).

Metabolite profiling of plasma samples

There were 580 putatively identified metabolites detected in all plasma samples (526 of which were present in both control and treatment groups), showing that extracted plasma samples contained a lot fewer metabolites compared to extracted urine samples. The majority of detected metabolites in plasma were amino acids, carbohydrates, other biochemical classes and a large proportion of lipids (see figure 4.6). Once again, it was expected that a large number of polar metabolites would be observed due to the nature of the samples and an efficient hydrophilic extraction technique. A large number of lipid species (triglycerides, sterols, phospholipids) were also detected as they are commonly found in human and animal plasma samples.[340] Metabolomes of urine and plasma control and treatment samples had a significantly large number of metabolites in common when biofluid samples were compared directly. However, there was a large difference in fold changes showed by IDEOM software in the metabolites of urine samples (control vs treatment), and the same difference was observed in plasma samples.

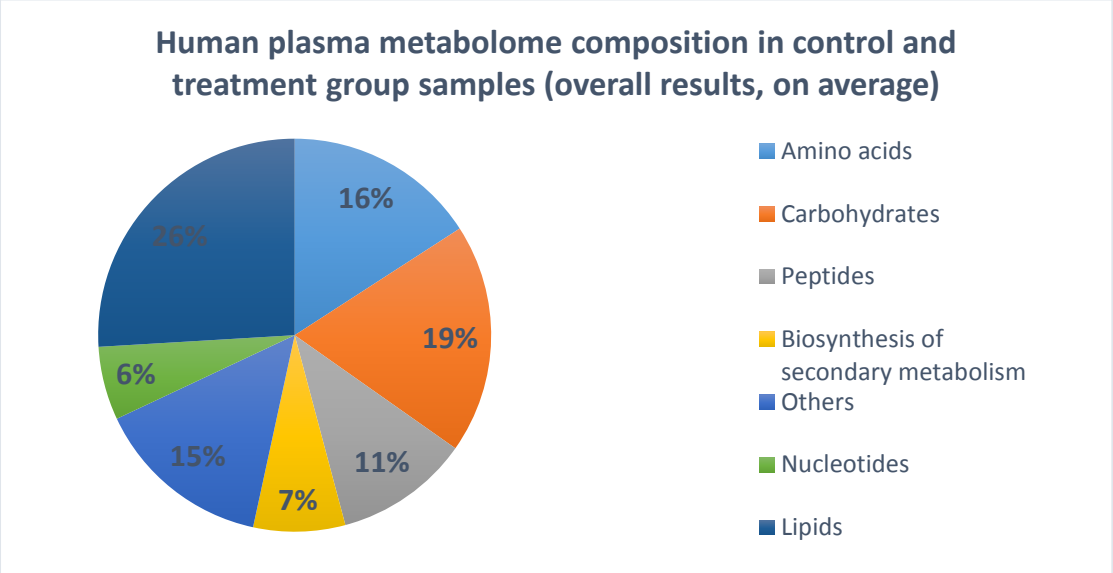


Figure 4.6 Composition of putative metabolites in human **plasma** samples (control and treatment groups) according to biochemical class

4.3.2 LC-MS analytical performance and normalisation in urine and plasma samples

Analytical performance

In the previous section, 4.3.1, the spectra graphs of urine and plasma samples (control and treatment groups) alongside their QC samples were shown to have a clear peak signal response with high TIC values and high signal to noise ratios. A relatively large number of putative metabolites were also detected in both types of biofluids, across different samples and replicates. This demonstrated a good analytical performance of an LC-MS instrument and sample preparation, without any unnecessary contamination and background noise. QC analysis had to be done in more detail to assess the sample extraction and preparation consistency, and experimental reproducibility.

In the urine experiment, the seven QC samples were analysed by IDEOM software – the first six QC samples in the analytical run were not included in data processing, as they were used to stabilise the system. The peak height intensity of every metabolite in each of the seven QC samples was calculated and percentage RSD values were determined to observe the degree of data variation of metabolites between QC samples. It has been observed that the majority of detected metabolites (604 out of 750 metabolites) expressed an RSD of less than 30% between QC samples – that is before normalisation calculations. This showed that there was not a huge degree of variation between sample metabolites, demonstrating accurate QC sample preparation, and therefore demonstrated small analytical variance in metabolite data throughout all urine sample preparations, analysis and data processing.

Plasma QC samples have demonstrated a similar result, with the majority of metabolites (520 out of 580) showing an RSD of less than 30% – before normalisation. This once again demonstrated the high accuracy of QC sample preparation and small analytical variance observed in the data.

Normalisation data analysis

Once the normalisation analysis has been applied to datasets of urine and plasma samples, percentage RSD values of peak height intensities of every metabolite between biological replicates (volunteers) were calculated in each control and treatment group, in the same way that the data was normalised in the validation experiments in chapter three. The following numbers of successfully normalised metabolites (with their percentage RSD decreased compared to non-normalised) were calculated in the urine sample groups. Out of a total number

of detected metabolites in each group (679), the number of successfully normalised metabolites were: "control 1st time point" – 617, 90.3%; "control 2nd time point" – 607, 89.3%; "control 3rd time point" – 593, 87.3%; "treatment 1st time point" – 604, 88.9%; "treatment 2nd time point" – 622, 91.6%; "treatment 3rd time point" – 615, 90.5%. One can see that in all groups there was a large number of metabolites successfully normalised by ^{13}C labelled internal standards (^{13}C paired or co-eluted ^{13}C peaks). Most of those successfully normalised metabolites were present in all groups, as they were normalised by the same internal standard present in all of the samples. A proportion of metabolites in each group could not have been normalised successfully (their percentage RSD was increased compared to non-normalised), because of the lack of suitable internal standards. –These were found to be mainly lipid metabolites.

In the plasma samples, the same calculations were performed. Out of a total number of detected metabolites in each group (526), the number of successfully normalised metabolites were: "control 1st time point" – 388, 73.7%; "control 2nd time point" – 416, 79.0%; "control 3rd time point" – 427, 81.2%; "treatment 1st time point" – 424, 80.6%; "treatment 2nd time point" – 403, 76.6%; "treatment 3rd time point" – 415, 78.8%). Once again, the normalisation in the majority of detected metabolites in each group was successful with a large number of normalised metabolites having a much lower percentage RSD than before. Those metabolites which were not successfully normalised, did not have a suitable internal standard that could have lowered their percentage RSD.

It was noticed that the plasma metabolome in all of the groups (control and treatment) was not normalised as effectively as the urine metabolome – on average, around 70%–80% of plasma metabolites had their percentage RSD decreased, compared to an average of 80%–90% of urine metabolites. The main reason for this was the lack of suitable internal standards found for plasma metabolites, as a large proportion of detected plasma metabolites belonged to the lipid biochemical class, for which a limited number of internal standards were present in (U)-^{13}C labelled *E. coli* (the same source of internal standards has been used in both experiments).

Multivariate analysis

The normalised metabolomes of plasma and urine samples ($^{13}C/^{12}C$ ratios) were arranged in matrix tables (see section 4.2.3) and passed through multivariate analysis to obtain PCA and OPLS-DA plots, for further statistical analysis. For PCA and OPLS-DA score plots, it was decided that "treatment" group (affected by high fatty acid meal) would be called "intervention".

For the urine normalised metabolome, the PCA and OPLS-DA score plots are shown in figures 4.7, 4.8 and 4.9. Figure 4.7 shows the PCA score plot of all the urine control and intervention groups as well as QC samples. The QC samples are shown to be clustered in the middle of the score plot, clearly demonstrating once again the high reproducibility of metabolite data in all of the samples and low variation between replicates. The control and intervention groups are separated from each other, with each group's biological replicates shown to be clustered well, demonstrating good normalisation and highly accurate data with reduced data variation.

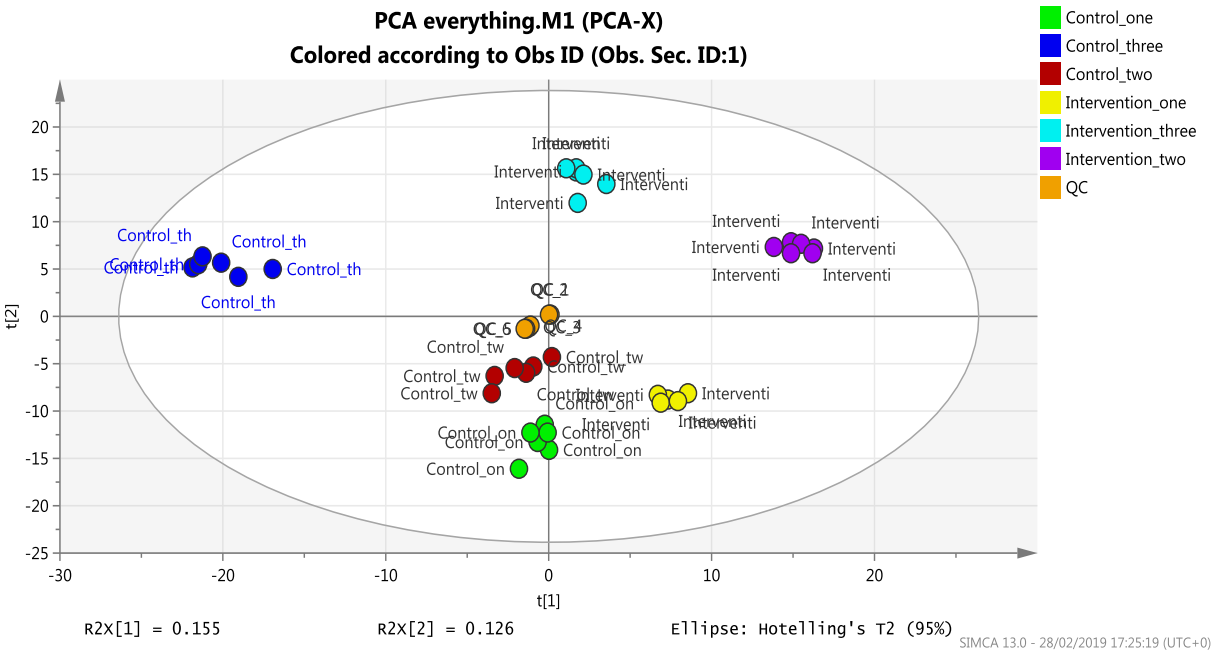


Figure 4.7 PCA score plot of all the urine groups and QC samples

Figure 4.8 shows three PCA score plots of "control vs intervention" groups at three different time points – the time points at which urine samples were collected. There is a clear separation between the data groups in all three plots, as expected, due to the difference in metabolite concentration between control and intervention state. The clustering effect is observed but not to such an extent as in the OPLS-DA models. In figure 4.9, three OPLS-DA models show exactly the same effect of clear separation between control and intervention groups in all three time points groups. There is a good cluster of replicate points in all three OPLS-DA plots, especially in the "third time point" plot, with high Q2 scores in each plot – 0.948 for the first time point plot, 0.963 for the second and 0.972 for the third. This demonstrated good normalisation of metabolites, reduced data variation and improved data accuracy, as well as a clear difference in metabolite activity between control and intervention groups. Urine data was promising for the identification of statistically significant metabolites.

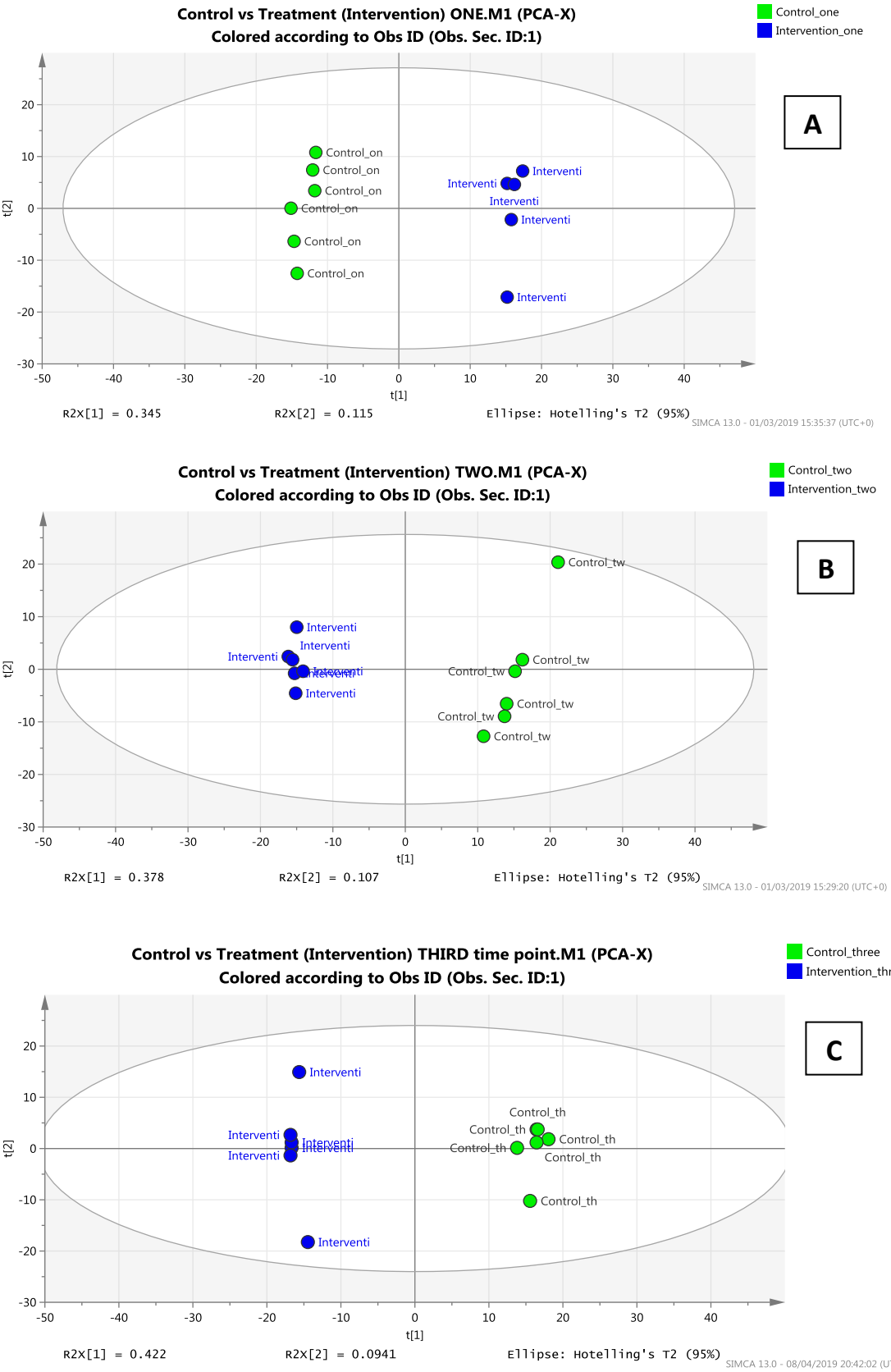


Figure 4.8 PCA score plots of control vs intervention groups at first (A), second (B) and third (C) time points.

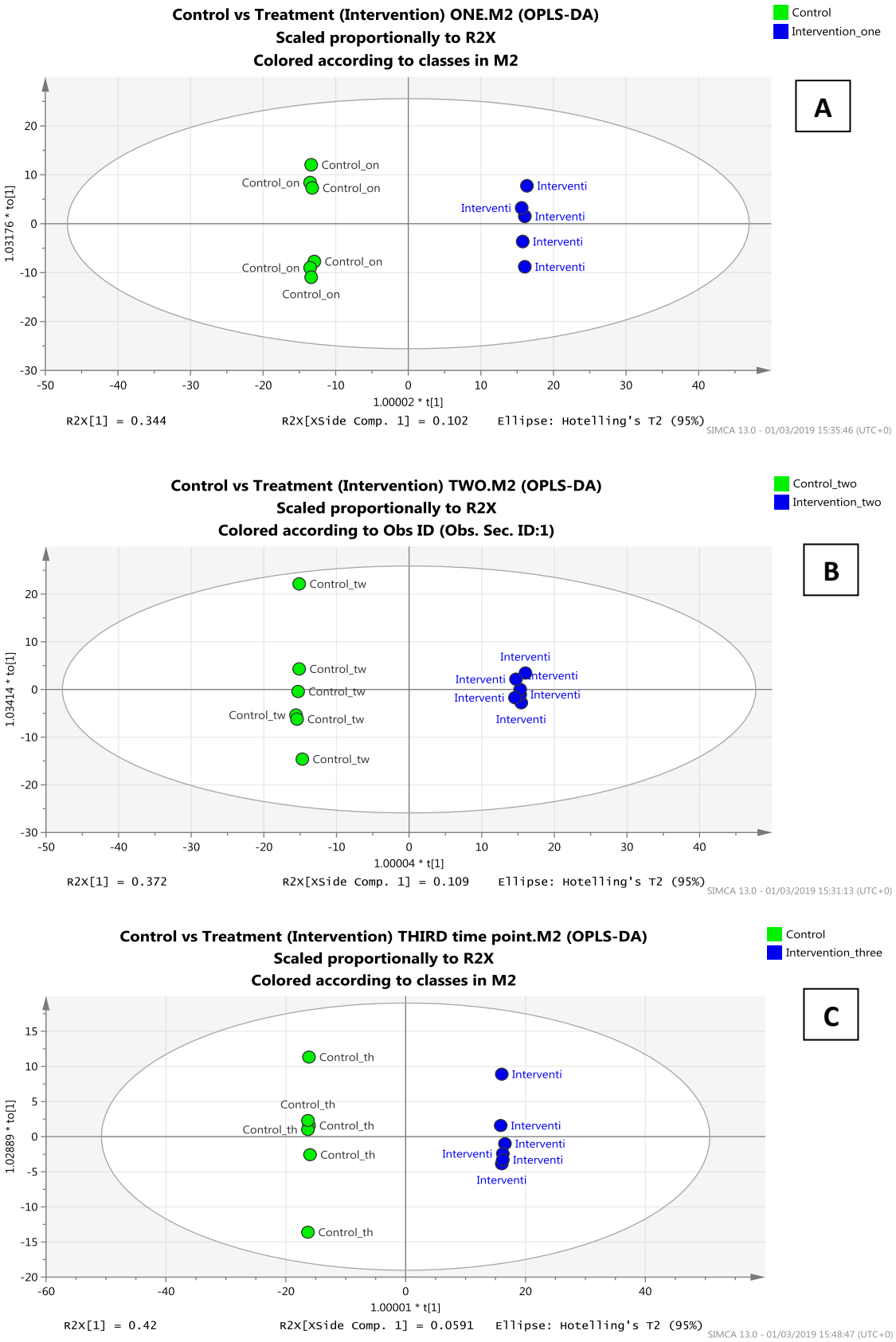


Figure 4.9 OPLS-DA score plots of control vs intervention groups at first (A), second (B) and third (C) time points. Q2 scores were 0.948, 0.963, 0.972 for A, B and C score plots, respectively.

For the plasma normalised metabolome, the PCA and OPLS-DA score plots are shown in figures 4.10, 4.11 and 4.12. Figure 4.10 shows the PCA score plot of all the plasma control and intervention groups as well as QC samples. This time, there is a very poor separation between groups with some replicate points from different groups found to be all over the plot. Even though the data was normalised, an unsupervised PCA model containing all the data groups does not show clear separation all of the time. QC samples are observed to be in the centre of the plot, once again showing the good reproducibility of sample preparation and analytical performance during the experiment.

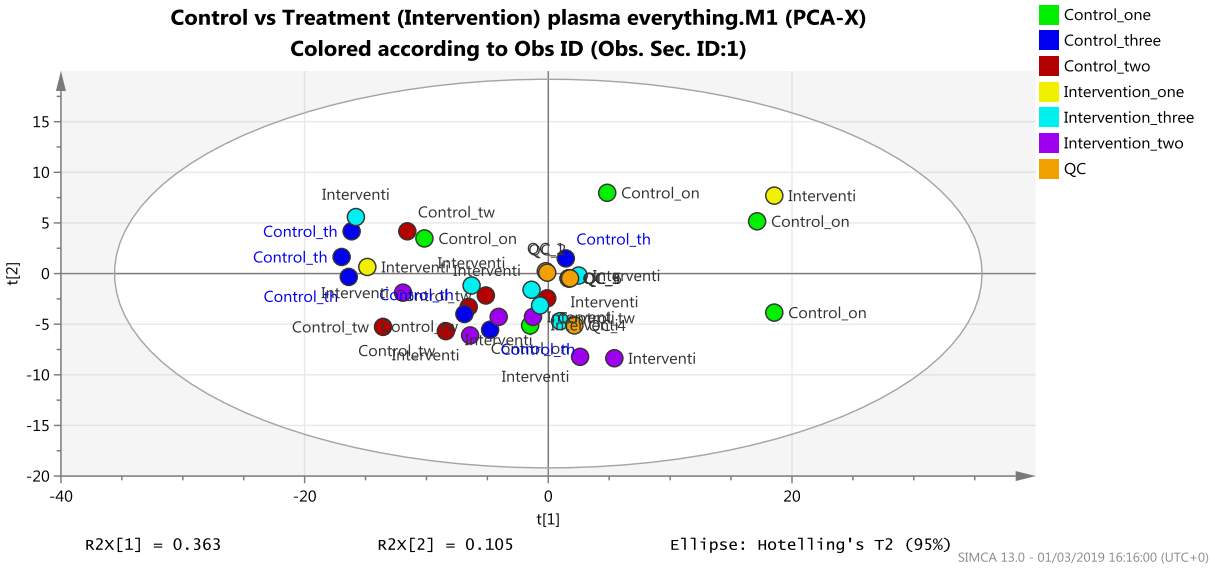
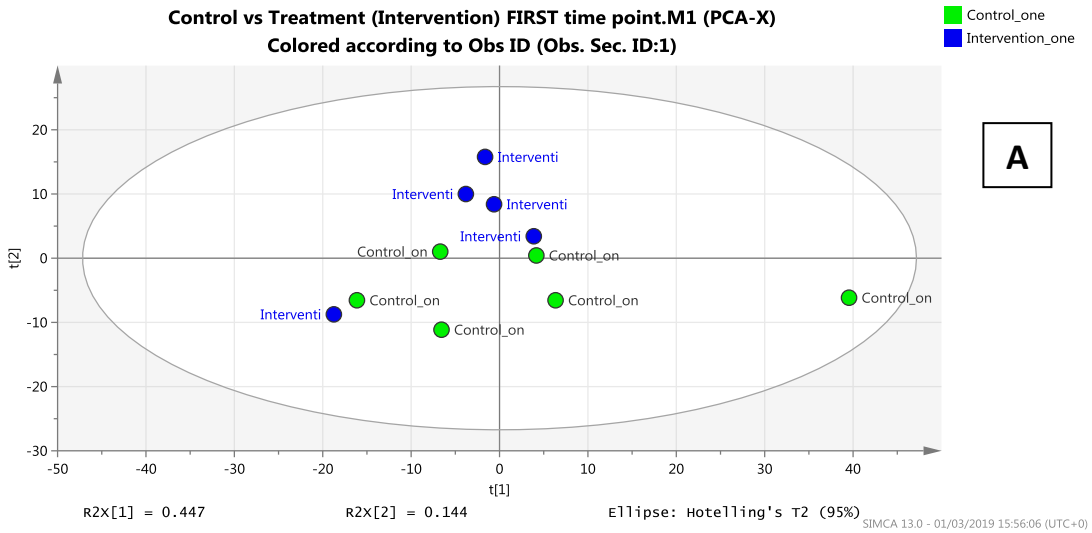
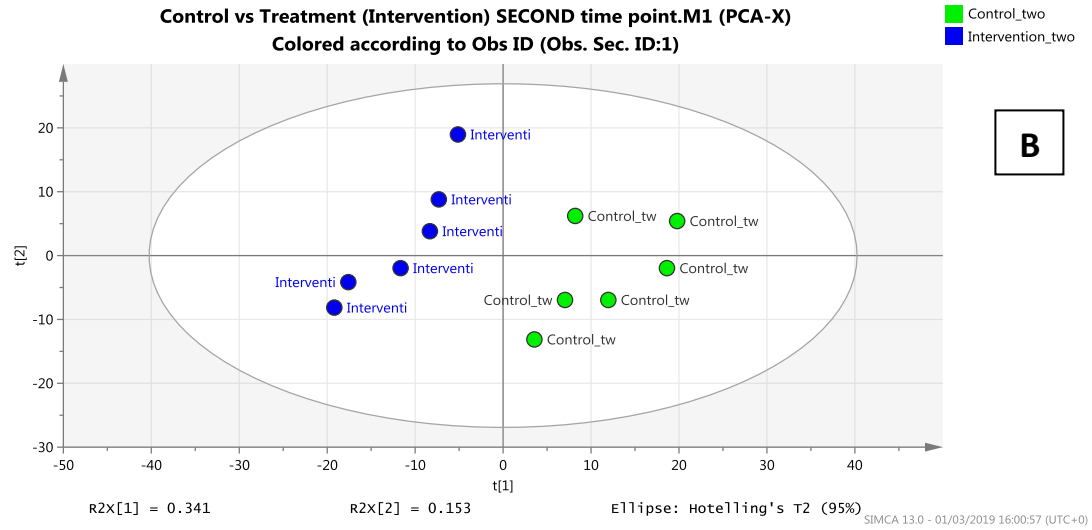


Figure 4.10. PCA score plot of all the plasma groups and QC samples

Three PCA score plots (figure 4.11) and three OPLS-DA plots (figure 4.12) are shown for different time points, in the same manner as for the urine data. PCA plots demonstrate separation between the control and intervention data points but not very clearly (in the first and third time point plots, some of the replicate points are seen to be located on the opposite side close to the other group replicates). However, all three OPLS-DA plots show a clear separation between control and treatment groups, even though the clustering of replicate points is not as tight as in the urine OPLS-DA plots. The Q2 scores for plasma OPLS-DA plots for first, second and third time points are 0.352, 0.807 and 0.746, respectively. These values are lower than the ones in the urine OPLS-DA score plots. Overall, it can be seen that plasma normalisation was not as successful as for urine metabolome, and therefore could not improve the accuracy of data replicate points as much as in the urine data. Nevertheless, the data for the plasma metabolome was found to be accurate and reliable enough for identification of significant metabolites.





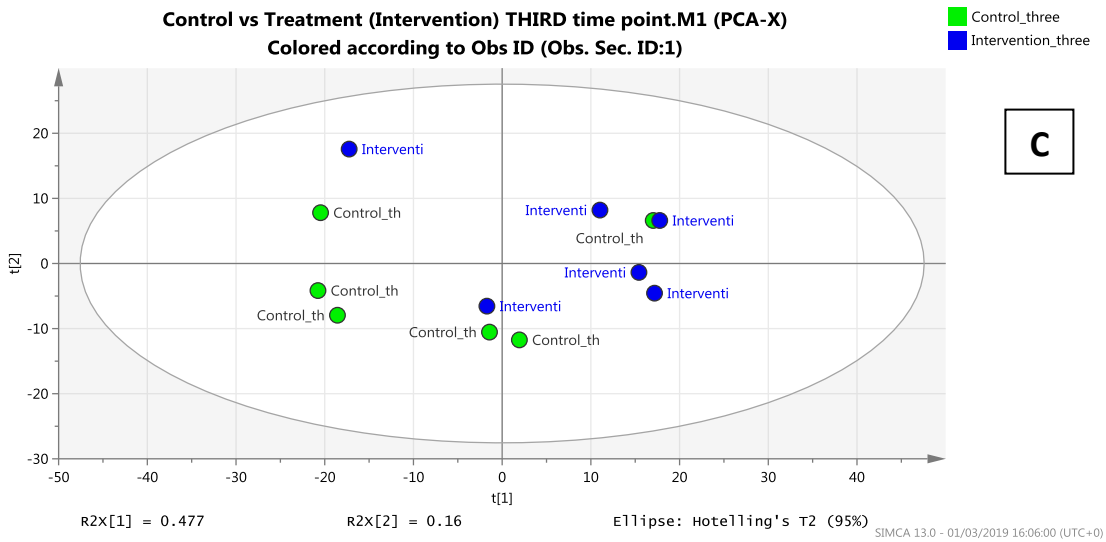


Figure 4.11 PCA score plots of control vs intervention groups at first (A), second (B) and third (C) time points.

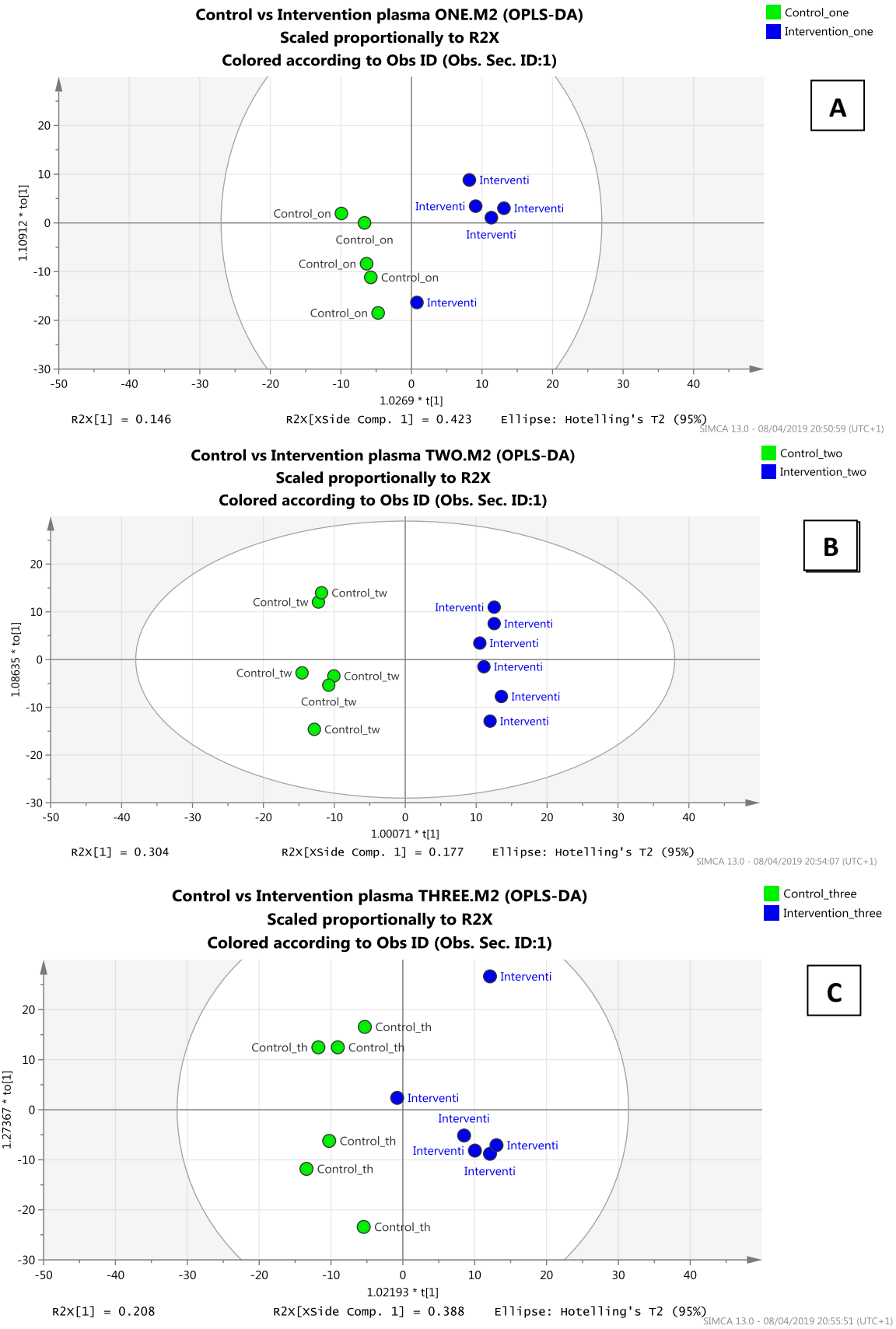


Figure 4.12 OPLS-DA score plots of control vs intervention groups at first (A), second (B) and third (C) time points. Q2 scores were 0.352, 0.807, 0.746 for A, B and C score plots, respectively.

4.3.3 Tentative identification of statistically significant metabolites in urine and plasma samples

Urine key metabolites

During multivariate analysis using OPLS-DA score plots (see figure 4.9), a list of metabolites with a VIP score of 1.0 or above for every time point condition has been generated. The “first time point” OPLS-DA score plot has shown 319 metabolites with high VIP scores of 1.0 or above, in the “second time point” plot there were 278 and in the “third time point” plot there were 240 metabolites.

Univariate analysis involving T-test and FDR correction has expressed a large number of statistically significant metabolites with different fold changes between control and intervention groups, for each time point condition. For the “first time point” data set, there were 194 statistically significant metabolites identified. For the second and third time point data sets, there were 206 and 178 statistically significant metabolites, respectively. These metabolites had a p value of less than 0.05, making them significant between the compared two groups. These metabolites are displayed through volcano plots in figure 4.13, showing a distribution of positive and negative fold changes of metabolites – either being elevated or decreased in concentration during treatment.

Based on the above results from both multivariate and univariate analysis, a list of statistically significant metabolites which passed T-test (p value < 0.05), FDR test (5% significance level) and had a VIP score of 1.0 or above, was generated for each “time point”. Overall, there were 10 significant metabolites found for the “first time point”, 22 for the “second time point” and 30 for the “third time point”. At this stage it was decided to choose the third time point as the most important condition for discovery of statistically significant metabolites, as it was the final collection point of urine throughout the whole experiment – the urine metabolome would contain a wide range of “end of the day” important physiological metabolites. Tables 4.1 to 4.3 show a list of those significant metabolites at each time point. These metabolites were shown with their P values, VIP scores and fold changes, indicating whether they were elevated or decreased in concentration after the treatment was applied. ID confidence levels (IDEOM annotated score levels 1-10, as discussed in chapter 2, section 2.2.3), were assigned to the metabolites. Metabolite annotation was performed in the same way, by matching accurate masses and retention times of observed peaks with metabolites in IDEOM database, assigning

levels 1 (least confident) to 10 (most one). Each metabolite was analysed further for key metabolite activity (refer to the Discussion section later).

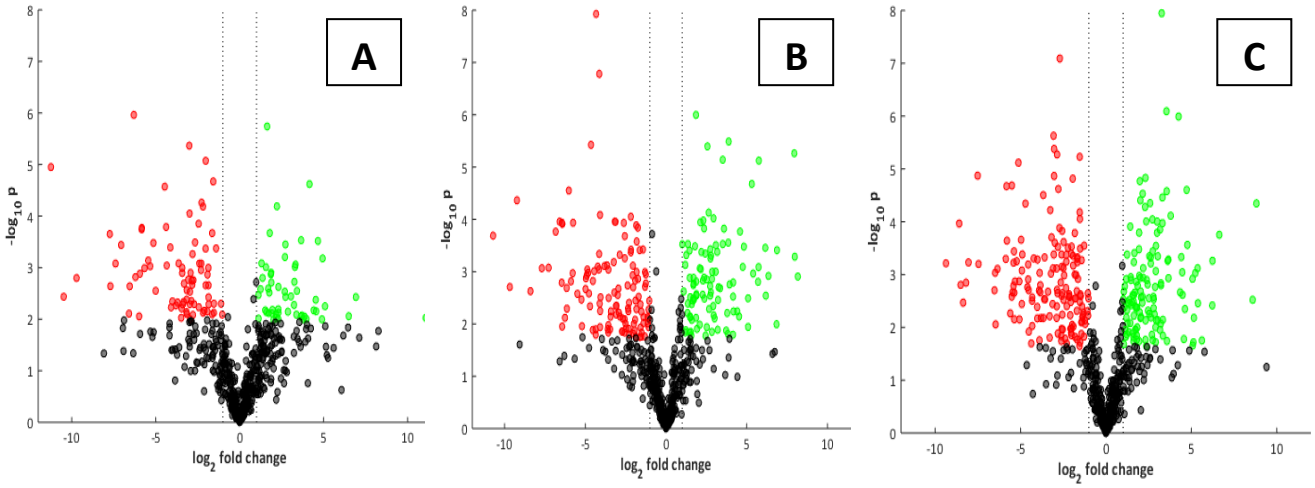


Figure 4.13 Volcano plots showing distribution of statistically significant metabolites detected between control and intervention groups in first (A), second (B) and third (C) time points. Data is analysed by univariate analysis only

Metabolite	Biochemical class	Retention time (mins)	Mass	VIP score	P value	Fold change	ID Confidence
R-Malate	Carbohydrates	8.85	138.0426	1.4875	1.97 E-05	1.638	8
2-oxoglutarate	Amino acids	7.29	150.0534	1.5734	6.37 E-07	6.419	6
5-methylcytidine	Others	7.63	267.1353	1.7781	1.52 E-04	4.345	8
Allantoin	Others	7.89	180.0777	1.2394	4.86 E-04	3.626	8
Androsterone glucoride	Lipids	4.41	491.3469	1.5949	1.17 E-05	2.778	7
cis-aconitate	Carbohydrates	9.46	180.0405	1.6339	3.25 E-09	3.461	6
Farnesoic acid	Others	4.45	251.2276	1.5472	3.61 E-04	2.279	5
Glu-Val	Peptides	7.44	256.1672	1.8305	7.64 E-11	0.752	7
Metaxyl	Others	4.29	294.1973	1.3944	1.25 E-03	3.882	6
Tributyrin	Others	4.53	317.2234	1.3082	3.27 E-06	3.837	7

Table 4.1 A list of statistically significant metabolites at **first** time point of urine collection between control and treatment groups. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10).

Metabolite	Biochemical class	Retention time (mins)	Mass	VIP score	P value	Fold change	ID Confidence
4-Pyridoxolactone	Metabolism of cofactors	4.89	173.0694	1.5014	1.99 E-05	5.697	8
8-keto-7-amino perlagonate	Others	4.52	194.1358	1.4406	5.14 E-05	3.051	6
Allantoin	Others	7.89	180.0777	1.4549	5.83 E-05	3.329	8
Asn-Pro-Tyr	Peptides	5.17	214.1447	1.5925	7.27 E-06	4.593	8
Bendiocarb	Others	5.05	234.1218	1.5427	3.51 E-05	4.048	5
Dodecanoic acid	Lipids	4.67	242.1912	1.6287	1.52 E-06	1.095	6
Heptanal	Lipids	4.48	121.1689	1.1172	1.98 E-05	2.658	5
Indolacryloyglycine	Others	5.53	257.1284	1.2254	2.67 E-04	2.237	7
Kynurenate	Amino acids	5.45	199.0763	1.5238	1.34 E-04	0.611	7
L-Erythrulose	Carbohydrates	7.14	124.0637	1.6783	1.92 E-04	5.416	7
Leu-Gly-Tyr	Peptides	5.44	368.2366	1.6509	1.23 E-07	1.138	6
N-ethyl-2-pyradinamine	Others	5.01	260.1394	1.3686	1.92 E-05	4.167	5
N-acetyl anthranilate	Others	5.37	188.0883	1.2989	2.77 E-06	5.372	6
N-benzoyl-L-glutamate	Others	5.61	263.1196	1.5067	3.71 E-04	2.604	7
Nicotinurate	Metabolism of cofactors	5.42	188.0897	1.4756	3.67 E-06	5.262	6
Phe-Pro	Peptides	5.23	276.1784	1.4854	3.14 E-05	5.371	8
Perillic acid	Biosynthesis of secondary metabolism	5.54	176.1337	1.1942	1.16 E-03	3.082	6
Triangularine	Others	4.36	353.2322	1.3959	1.66 E-04	4.263	8
Ubiquinol	Others	4.54	266.1827	1.6413	3.32 E-07	1.184	7
Urocortisone	Lipids	5.03	385.2944	1.5379	5.74 E-05	3.645	5
Valclavam	Others	5.46	343.2051	1.4935	2.59 E-07	5.508	8
1,3,5-trimethoxybenzene	Metabolism of cofactors	7.42	176.3284	1.6314	7.63 E-06	1.083	7

Table 4.2 A list of statistically significant metabolites at **second** time point of urine collection between control and treatment groups. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10).

Metabolite	Biochemical class	Ret time (mins)	Mass	VIP score	P value	Fold change	ID Confidence
(R)-Malate	Carbohydrates	8.85	138.0426	1.6520	2.00E-05	1.815	8
(S)(+)-Allantoin	Nucleotides	9.53	162.0681	1.4753	4.73E-04	2.196	6
10-Oxodecanoate	Lipids	4.68	196.1649	1.8652	2.00E-11	0.718	6
2',3'-Cyclic AMP	Nucleotides	7.08	339.0923	1.4147	6.52E-04	4.050	7
2-Hydroxyadenine	Others	9.11	156.0711	1.7034	5.00E-05	2.997	8
2-Oxoglutaramate	Amino Acids	7.29	150.0534	1.6886	8.07E-05	4.390	6
5-Methylcytidine	Others	7.63	267.1353	1.6335	1.00E-04	3.973	8
6-Deoxy-L-galactose	Carbohydrates	7.14	170.0939	1.7911	3.53E-06	1.266	9
Allantoin	Nucleotides	7.89	180.0777	1.4364	5.23E-04	4.595	8
Allopurinol	Others	8.17	141.0658	1.7099	9.65E-05	1.508	9
alpha-D-Mannoheptulopyranose	Others	7.51	217.1334	1.5119	2.55E-04	3.574	6
Ascorbate 2-sulfate	Others	11.38	262.0103	1.6339	1.25E-03	0.212	6
Cerulenin	Others	4.28	235.1653	1.3424	2.84E-05	2.199	7
cis-Aconitate	Carbohydrates	9.46	180.0405	1.8227	4.76E-08	2.724	6
Coniferaldehyde glucoside	Others	4.75	356.1722	1.7228	2.59E-05	0.765	7
Deoxycytidine	Nucleotides	7.14	236.1257	1.8415	2.99E-09	5.142	5
D-Glucuronate	Carbohydrates	9.05	200.0635	1.6609	7.28E-04	0.290	8
Didrovaltratum	Lipids	4.14	234.1844	1.4015	6.21E-03	0.277	8
DL-Methionine sulfone	Others	6.88	186.0698	1.6716	2.73E-05	3.717	5
Glu-Asp	Peptides	8.96	271.1155	1.6514	1.00E-04	6.705	6
Glu-His-Thr	Peptides	7.43	400.2152	1.8004	5.86E-07	6.584	8
Glu-Ile-Ile	Peptides	4.53	390.2883	1.5505	3.45E-05	1.710	7
Glu-Val	Peptides	7.44	256.1672	1.8688	9.00E-15	0.913	7
Imidazol-5-yl-pyruvate	Amino Acids	7.48	160.0646	1.5727	1.73E-03	0.601	6
Indole-3-acetyl-glutamine	Others	6.87	318.1733	1.5484	1.47E-04	1.969	5
Inulicin	Others	4.39	325.2268	1.7019	6.22E-05	1.836	6
Isopyridoxal	Metabolism of cofactors	6.89	175.0942	1.8285	1.82E-07	3.841	5
Kynurenate	Amino Acids	5.46	199.0895	1.2656	1.88E-02	0.018	6
L-Erythrulose	Carbohydrates	7.14	124.0637	1.5529	1.61E-03	3.102	7
Prostaglandin A2	Lipids	4.37	354.2851	1.5858	1.00E-05	5.024	7

Table 4.3 A list of statistically significant metabolites at **third** time point of urine collection between control and treatment groups. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10).

Plasma key metabolites

The same statistical analysis (multivariate and univariate procedures) have been applied on plasma metabolomes of three time point data sets. In multivariate analysis from OPLS-DA score plots (see figure 4.12), there were 153 metabolites identified with a VIP score of 1.0 or above for the "first time point" data set, 103 for the "second time point" and 114 for the "third time point". From univariate analysis, the number of statistically significant metabolites detected was lower than for urine samples, with 53 detected in the "first time point" group, 72 in the "second time point" group and 125 in the "third time point" group. The results are presented in figure 4.14 through volcano plots, showing the number of statistically significant metabolites with positive and negative fold changes.

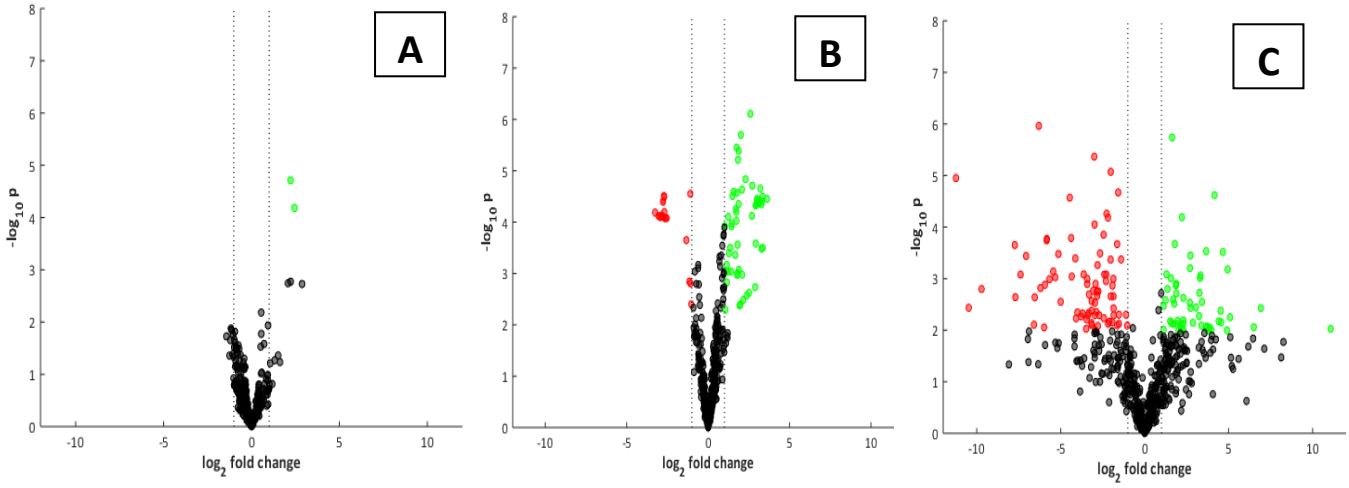


Figure 4.14 Volcano plots showing distribution of statistically significant metabolites detected between control and intervention groups in first (A), second (B) and third (C) time points. Data is analysed by univariate analysis only

The results of multivariate and univariate lists combined together generated the final lists of statistically significant metabolites for each time point condition, between control and intervention groups, based on passed T-test, FDR test and VIP score system. Unfortunately, there were not many statistically significant metabolites detected, compared to the urine results. In the "first time point" data set, there were no statistically significant metabolites passing all of the tests. In the "second time point" data set, there were only two statistically significant metabolites (cis-aconitate and NG,NG-Dimethyl-L-arginine) and in the "third time point" group there were four statistically significant metabolites (10-amino-decanoic acid, allantoin, cis-aconitate and NG,NG-Dimethyl-L-arginine). Their fold changes were all elevated (increased) by 1.5 or 2.0 scale when treatment (liquid milk protein) was applied. Table 4.4

presents the list of statistically significant metabolites at "third time point" collection. Metabolite annotation was performed in the same way as described earlier.

Metabolite	Biochemical class	Retention time (mins)	Mass	VIP score	P value	Fold change	ID Confidence
10-amino-decanoic acid	Lipids	7.33	187.1572	1.746	1.00E-05	4.388	8
Allantoin	Nucleotides	7.89	180.0777	1.331	3.45E-07	5.424	7
cis-aconitate	Carbohydrates	9.46	174.0165	1.282	2.54E-08	2.442	8
NG,NG-Dimethyl-L-arginine	Others	12.17	202.1429	1.384	4.77E-02	3.549	7

Table 4.4. A list of statistically significant metabolites at third time point of plasma collection between control and treatment groups. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10).

Even though in plasma samples, only a few number of statistically significant metabolites were detected, it appears that in common "cis-aconitate" was a key metabolite detected between control and intervention groups, in both urine and plasma samples.

In this study, a statistical paired difference test should have been performed to complete statistical analysis fully, as two sets of measurements (control vs treatment, in both urine and plasma samples) were compared in the same group of subjects (overweight male healthy volunteers). Paired test would have made an assessment whether population means of those measurements significantly differ or not. Such test compares these measurements within subjects, rather than across subjects, and will generally have greater statistical power than an unpaired test. The key issue that motivates the paired difference test is that the subjects will differ substantially from each other before treatment begins – due to biological factors which usually causes biological variance (mentioned earlier). Even with applied normalisation method, biological variance is hard to correct, so running a paired test in this study would have improved an overall accuracy of data by reducing certain degree of inter-subject variability. In OPLS-DA figures it can be seen that inter-subject variability was big, with clear big separation between groups, therefore the paired test would show the groups more clustered.

4.4 Discussion

4.4.1 Quality of data

In this case study, the effect of a meal with a composition of various fatty acids (liquid milk protein) was investigated on a variety of metabolic activity in healthy, slightly overweight volunteers: skeletal muscle protein, fat tissue and biofluid (urine, plasma) metabolism. For the purpose of this thesis, the detailed analyses of human urine and plasma metabolic profiles only were performed, in search of statistically significant metabolites between two conditional groups: control and treatment (high fat dietary intervention). This study did not involve diabetes patients but created the presence of a "pre-diabetes" condition. Therefore, the identification of key metabolites was not a true reflection of diabetes pathology, and the real diabetes patients would need to be involved in the actual study.

From the Results section, one can see that a large number of putatively identified metabolites (mostly polar class) were identified in both the urine and plasma samples of volunteers. This observation was expected, given that both types of samples had undergone hydrophilic extraction with a validated LC-MS analytical method. The samples were collected fresh from volunteers and later prepared without any chemical modifications, apart from extracting them in (U)-^{13}C labelled extract (part of the normalisation process). The rest of the extraction and sample preparation procedures did not hugely affect the nature of the biofluid samples, hence the metabolism content (and metabolite stability) was not altered in any significant way. Further assessment of the quality of LC-MS analytical performance was shown through spectra graphs of samples (clear peak shapes without background noise) and percentage RSD calculations of peak height intensities in QC samples, showing that the majority of metabolites had an RSD value of less than 30%. This indicated that there was no strong biological or analytical variation present in the QC replicates of either urine or plasma, demonstrating good stability, repeatability, robustness and performance of analytical run. This observation was expected, given that urine or plasma pooled QC samples contained an equal amount of every single sample analysed, so it was a good representation of the entire sample set.

Normalisation, using (U)-^{13}C labelled *E. coli* as a source of internal standards, once again caused a significant decrease in data variation for the majority of metabolites – more than 70%–80% of metabolites successfully normalised in plasma, and 80%–90% in urine. This observation was expected, as the previous validation experiments with human urine and plasma samples (see chapter three) had a similar result. There was a large number of ^{13}C labelled metabolites in the

extract, acting as paired or co-eluted internal standards for sample metabolites. A certain proportion of lipids in either plasma or urine samples were not normalised well, because the extract did not contain a large number of lipophilic internal standards appropriate to normalise lipids. In all OPLS-DA plots (in plasma and urine experiments), the normalisation effect was shown clearly through clear separation between control and treatment groups, tight cluster of replicate points and high Q2 scores in most of the OPLS-DA plots. This has demonstrated once again that the proposed normalisation method using stable ^{13}C isotopes as internal standards works effectively, improving data accuracy.

4.4.2 Key metabolites and metabolic pathways

In this study, the effect of a liquid meal full of saturated fat was investigated in skeletal muscles by analysing the metabolome of human urine and plasma samples to observe the effect on metabolic pathways related to secretion and regulation of insulin. There were several statistically significant metabolites detected between control and treatment groups at all three time points (mostly at third one), for both urine and plasma samples. They are presented in table 4.1 to 4.4. An elevation in metabolic activity in most of these metabolites can be observed after the treatment was applied – their fold changes were observed to be positive, while few metabolites were shown to have their activity decreased after the treatment (negative fold changes). As a high fat diet is known to impair the action of insulin and prevents correct utilisation of a glucose in skeletal muscles, it was expected to observe an elevation of glucose and related metabolites in pathways such as glycolysis and Kreb's cycle. Therefore, some of the key detected metabolites were found to relate to those pathways. One research study stated that skeletal muscle insulin resistance is an initial step in the progression towards type 2 diabetes condition and has been linked to increase in dietary fat, so it was expected to observe some of the elevated metabolites usually found in diabetes.[341]

One of the detected metabolites in the urine samples in this study was L-erythrulose, an AGE of L-ascorbate, which is known to be elevated after a high fat diet intake and may cause, for example, the retinopathy complication in diabetes type 1 and 2 patients or atherosclerosis progression.[342,343] From tables 4.1 to 4.3, it can be seen that the metabolite's activity (fold-change) was three times higher in a treatment group, compared to the control state, showing elevation in metabolite activity during treatment. In one dietary study, L-erythrulose levels have been observed to increase dramatically in human urine samples collected from patients who were given a meal with large amount of saturated fat, even though patients were pre-disposed

to diabetes already but not fully diagnosed.[344] Other statistically significant metabolites, which were also found to be upregulated in the treatment group and belonged to AGEs, were cis-aconitate and allantoin. The urine and plasma results showed the elevation of allantoin and cis-aconitate metabolites in treatment samples by almost two and three fold-changes, respectively. Those metabolites have been observed in several high fat dietary studies in activating inflammatory pathways such as RAS, RAGE (receptor for advanced glycation end product) in vascular endothelial cells and skeletal muscles, leading to skin damage and muscle breakdown as well as elevated blood pressure.[345,346] Cis-aconitate is one of the intermediates found in Krebs cycle, which is shown to activate the pathway elevating other intermediates.[347] However, in one study, allantoin was shown to reduce obesity in high fat diet-induced mice by blocking Imidazoline-I1 receptors and improving lipid metabolism.[348] So allantoin remains to be unclear in its role of progressing or preventing disease complications followed by high fat dietary intervention.

R-malate metabolite levels are generally found to be affected following dietary interventions, as it is an intermediate metabolite in the Krebs cycle, producing energy for the cells and elevating glucose levels in the bloodstream.[349] In the urine data results of this study, this metabolite was found to be statistically significant and its fold change was shown to decrease in the treatment samples. These results are supported by another dietary intervention study involving high-protein high-carbohydrate diet, where malate was also shown to be reduced in overweight individuals.[350] In one study, however, levels of malate were found to be elevated in diabetic rat models, significantly impairing the activity of Krebs cycle and concentration of other intermediates in this cycle, for example citrate and pyruvate, causing damage to the local cells and tissues.[349,351] The impairment of Krebs cycle has shown to cause decrease in mitochondrial respiration in skeletal muscles of type 2 diabetic patients.[352] It appears that certain diet influences malate metabolic activity in different ways, depending on whether there is a high or a low protein diet involved.

Nucleotides such as deoxycytidine and hydroxyl-adenine are also normally found to be significantly affected by the effect of high fat and protein content in dietary interventions. They are found to be hydrolysed by enzymes, affecting the DNA and genetic profile of the important cells and tissues regulating fat, glucose utilisation.[353] One study demonstrated an upregulation of gene expression in adipocyte and skeletal muscle tissues followed by a high-fat diet in male mice, by showing an increase and decrease in metabolite activity of certain nucleotides and lipids.[354] From the results in table 4.1, it can be seen that fold changes for those metabolites

were shown to be reduced (downregulated) in the treatment condition – for the deoxycytidine metabolite downregulation was by factor five, for hydroxyl-adenine it was by factor three.

From the plasma data, only a few statistically significant metabolites were detected between control and treatment groups, including cis-aconitate, allantoin and NG,NG-Dimethyl-L-arginine, all of which were shown to be elevated in the treatment group. This once again supported an idea about cis-aconitate being affected by dietary high saturated fat intervention. Metabolite activity of NG,NG-Dimethyl-L-arginine has been found to increase in several dietary studies, including a poly-unsaturated fatty acid diet study where a rise in NG,NG-Dimethyl-L-arginine metabolic activity has shown a beneficial influence on endothelial tissues.[355]

Overall, it appears that several statistically significant metabolites detected in urine and plasma treatment samples, especially cis-aconitate and allantoin, can indeed be influenced by dietary intervention where an increase in saturated fat or protein is observed. In diabetes research, these metabolites can be potential biomarkers, due to their significant differences in fold changes and effect on related metabolic pathways, as demonstrated in other research studies.[356]

4.5 Conclusion

This study has demonstrated a certain effect of saturated fat meal on metabolome activity in plasma and urine samples in human volunteers, showing how treatment using a meal with high fat content may trigger certain metabolic responses in skeletal muscles and be associated with complications such as diabetes, atherosclerosis, cardiovascular and other inflammatory conditions. The main results showed several important key metabolites, such as cis-aconitate, L-erythrulose, R-malate, which were elevated in their metabolite activities during the "treatment" condition. From previous reported studies, it was observed that some of those key metabolites have caused a reduced skeletal muscle ability to secrete proteins to maintain muscle size. Therefore, the study has shown that skeletal muscle is indeed being affected by certain metabolite activities caused by a saturated fat intake.

This study has been investigated through analysis of water-soluble metabolites (via HILIC chromatography conditions), and therefore, few lipid metabolites were detected in the samples overall. As the study involved saturated high fat dietary intervention, the key statistically significant metabolites would be expected to be more of a lipid origin, rather than water-soluble type. However, due to polar analytical conditions implemented, it was not possible to detect those key lipid metabolites and instead few polar ones were found. For this study it would have been better to use lipidomics approach with different column suitable for lipids analysis to obtain more lipophilic compounds. Perhaps, more statistically significant metabolites of lipid origin would have been detected. Key lipid metabolites that would be expected to observe in this study could be triglycerides, LDL-C, HDL-C, ketone bodies, cholesterol, all of which would be expected to increase in their metabolic activity after ingestion of high fat meal.[357]

The proposed analytical normalisation method improved the accuracy of the final data by successfully reducing data variation between biological replicates – as can be seen in section 4.4.1. Improved data reproducibility has helped to distinguish control and treatment groups more clearly, helping to detect key statistically significant biomarkers more accurately and their related metabolic pathways. To obtain even better results on the affected skeletal muscle metabolome, it would have been better to analyse tissue samples directly from the skeletal muscle or take some sort of biopsy to see the morphology of the tissue samples. That would have helped to better understand the significance of high fat meal on human metabolome in skeletal muscles and their ability to generate important proteins for muscle performance.

Chapter Five

Metabolic signatures of low-grade glioma in human brain tissue

5.1 Introduction

5.1.1 Glioma: definition and classification

Glioma is defined as a type of tumour that originates in glial cells of mammalian brain and spine, which make up the central nervous system (CNS). In brain physiology, the main function of glial cells is to insulate the neurons, which are important in the regulation of the CNS, in order to maintain their homeostasis and protect them from external damage.[358] Glial cells, particularly astrocytes, also play an important chemical signalling role in the CNS, controlling the release and action of chemical mediators involved in a variety of CNS mechanisms (GABA, acetylcholine, arachidonic acid metabolites), although neuronal signalling is a lot faster and more important than that of glial cells. Although astrocytes do not conduct action potentials or send signals to other parts of the body, they are very similar to neurons and play an important communication role within the brain.[359] When tumours are detected in glial cells, causing uncontrolled cell proliferation, they may be benign or malignant. Unfortunately, gliomas comprise about 80% of all the malignant brain tumours, and therefore occur more often than the benign type.[360] They are classified by the type of cell from which they have developed, and are graded by their pathogenicity and their location inside the brain.[361]

The type of glial cells which are mainly affected by glioma are astrocytes, causing astrocytoma and glioblastoma multiforme (GBM). Other glial cells affected are oligodendrocytes (oligodendroglioma) and ependymal cells (ependymoma). Gliomas are further categorised by grade: low grade gliomas are well differentiated, where the cell proliferation is slightly uncontrolled and usually leads to benign tendencies. For that reason, low grade gliomas are easier to treat but can increase in grade over time. High grade gliomas, on the other hand, have very poor cellular differentiation (anaplastic) and are referred to as malignant brain tumours, also known as cancers. High grade gliomas are very difficult to treat and there is a bad prognosis for sufferers. In GBM, for example, the grading system is split into four types: grade 1 (least advanced disease where tumour cells appear close to normal), grade 2 (low grade) and grades 3 and 4 (high grade), according to WHO research.[362] Finally, classification of a glioma depends on the brain region (location) that is being affected – below or above the tentorium membrane that separates the cerebellum from the cerebrum inside the brain. Most common cancers of the brain are found to be GBM tumours that are usually defined as an aggressive, high-grade (level 4 according to WHO research), invasive and heterogeneous tumour. Figure 5.1 shows the magnetic resonance imaging (MRI) scans of a human brain affected by GBM at an early stage

and its progression one month later (left image), as well as MRI scans of low grade astrocytoma progression over the same period of time but in a different patient (right image). It can be seen that the rapid growth of a cancer is noticeable in GBM MRI scans, in the central location of the brain, but not so much in low grade MRI scans, showing that low grade glioma is harder to detect and further analysis might be required to confirm it.

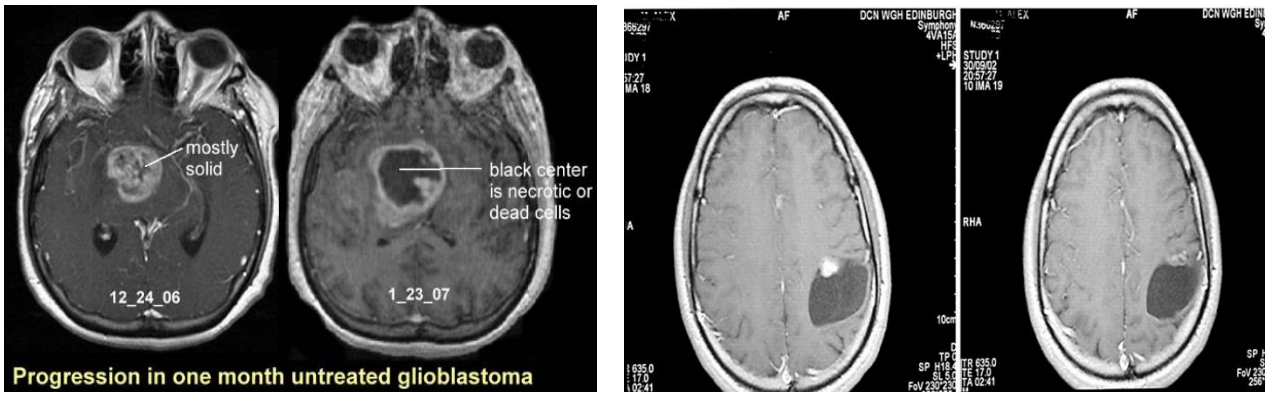


Figure 5.1 Progression of GBM glioma (grade 4) in an adult male brain over one month (left image) and progression of low grade 2 astrocytoma in an adult female brain over one month (right image). Advances in glioblastoma multiforme treatment. E. Ozdemir-Kaynak et al 2018

5.1.2 Low grade glioma: epidemiology, aetiology and symptoms

Epidemiology

Low grade gliomas (LGGs), as defined earlier, are the slowest growing type of tumour (grades 1 to 2), which often arise in young, otherwise healthy adults and children. They are usually developed in astrocytes, for example pilocytic astrocytoma (grade 1), and in oligodendrocytes – grade 2 oligodendroglioma. Despite their initial low, uncontrolled cell proliferation and sometimes asymptomatic state, LGGs may cause considerable morbidity leading to the death of a sufferer, if left untreated, as they can develop into more serious, high grade gliomas over time.[363] Over the last 15 to 20 years, the incidence of LGGs in the UK among adults has gradually increased from 2000 to 3000 cases a year, according to Cancer Research UK.[364] They account for almost 15% of all diagnosed primary brain tumours each year, with increased prevalence among white young men. However, the proportion of adults suffering from low grade glioma among the whole UK population is not so large, only around 3% which is 30 in every 1000 people. In the US for example, between the years 2010 and 2014, the average annual incidence of LGGs appeared to be significantly lower than for high grade glioma, in both males and females, as shown in figure 5.2.[365] There were 11,432 deaths reported in the UK from brain

tumours in 2016 (all types), with a 10-year survival rate of 14% per year according to Cancer Research UK. This shows that even though LGGs are not such serious tumours as high grade (e.g. GBM), they can easily develop into the cancerous type, and therefore need to be treated immediately.

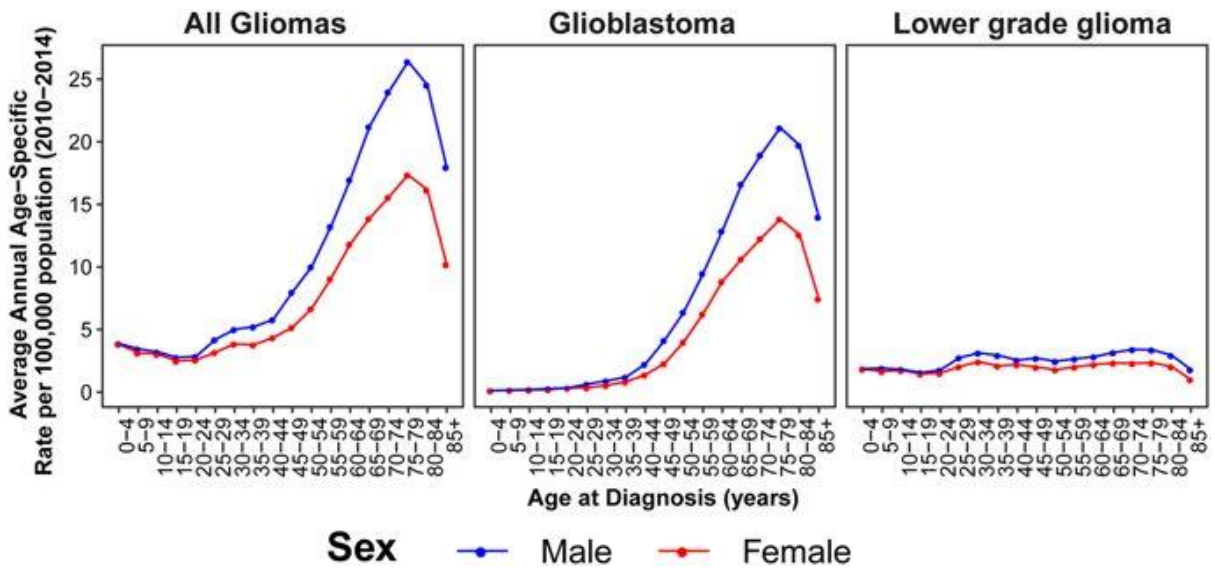


Figure 5.2 Average annual incidence rate of different types of gliomas, including GBM and low grade, over a period of four years (2010–2014) in males and females in the United States. Q. Ostram et al, 2018, Sex-specific glioma genome-wide association study identifies new risk locus at 3p21.31 in females, and finds sex-differences in risk at 8q24.21

Aetiology and symptoms

As for most of brain tumours, it is still uncertain what the main causes of LGGs could be. Scientists are constantly looking into possible origins of the pathogenic mechanisms involved in glioma. There is a theory that electromagnetic radiation from mobile phones[366], power lines[367], microwave radiofrequency radiation[368] and certain food additives such as aspartame[369] (sweetener) may trigger the pathogenesis of a disease, although those factors are still being investigated. It is known that LGGs may occur at any point in life, most commonly developing in children and young adults. As an individual becomes older, the risk of getting any type of glioma, including LGG, increases.[370] A family history of any type of glioma or certain other diseases can also provoke the development of LGG in an individual, where several genes are found to be involved with the progression of a disease, such as isocitrate dehydrogenase (IDH1) gene mutations, neurofibromin (NF1 and NF2) gene mutations associated with neurofibromatosis, tuberous sclerosis (TSC1 and TSC2) gene mutations associated with tuberous sclerosis disease,

and so on. It is thought that IDH1 and IDH2 mutations play an important role in the pathogenesis of a disease. However, there is currently no strong association of genetic influence on the development of LGG overall, so more research needs to be undertaken in this area.[371]

LGG originates inside the brain and normally does not spread outside it but instead grows into the normal brain tissue. As the tumour grows locally and expands, it disrupts connections between brain cells causing severe brain problems, even with a slow growing rate of tumour development.[372] The most common symptom is usually the presence of seizures, occurring in almost 80% of patients diagnosed with LGG. The seizure is usually caused by disorganised electrical activity in the brain, followed by involuntary movements, loss of muscle control and, in severe cases of seizures, loss of consciousness and speech.[373] Other clinical signs and symptoms associated with LGG are headaches, mood swings, memory and vision loss.[374] Sometimes, individuals with diagnosed LGG may be asymptomatic for a long time, not presenting any neurologic or cognitive changes, however they still have to be constantly monitored through imaging and spectroscopic analysis which can show the tumour presence.

Diagnosis of LGG

The diagnosis of any tumour, whether it is a low or high grade, is done via a series of physical, laboratory and X-ray examinations, as well as MRI or computed tomography (CT) scans of brain tissue[375-377], also known as imaging studies. If the presence of a brain tumour is suspected, MRI and CT scans are usually performed to confirm the presence, showing the growth pattern and damage of tissues and cells. The imaging can reveal the full structure of the brain locations affected by the tumour. Biopsy tests are only performed if the suspected tumour has not been fully confirmed by imaging studies, establishing the diagnosis and subtype of tumour. This process involves removing a small amount of tissue from the brain, examining it under a microscope for histological diagnosis – biopsy is not performed unless the other diagnostic tools fail to confirm the presence of a tumour. Biopsy with histological evaluation is generally a very accurate method. For example, low grade astrocytomas have well-differentiated fibrillary astrocytes which are easily identified under a microscope when the tissue is stained with chemicals such as hematoxylin.[378] However, because of the heterogeneity of the tumours, biopsy may not always reflect the highest grade, e.g. GBM, with reported accuracies of 51% to 83%.[379] The decision to undertake biopsy as a diagnostic method in a patient with suspected LGG, requires the clinician to carefully weigh all the risks and benefits associated with this strategy.

5.1.3 Low grade glioma: current treatment strategies

Currently, it is a great challenge to develop a suitable treatment for any type of LGG or even design and evaluate clinical trials showing the effectiveness of proposed therapy. Each therapeutic agent or technique requires a complicated procedure of treatment, high cost, safety, success of removing the cancer without any post adverse effects and improving the quality of life for a patient overall.[380] At the moment, surgical removal of a LGG tumour is one of the prime methods in therapy, especially as it has shown to improve seizure control in patients.[381] Several studies demonstrated that overall the surgical intervention was found to be safe and effective in LGG patients, with high survival rates compared to non-surgical strategies.[379,382,383] They concluded that there was a trend in survival rate improvement with more extensive surgical resection. A number of advanced techniques in surgical procedures has allowed surgeons to maximise the degree of tumour resection without affecting good healthy tissue, significantly improving the life expectancy of an individual.[384] Biopsy methods, mentioned earlier, are part of such surgical operations, where the removal of potentially cancerous brain tissue and an immediate diagnosis outweigh the risks of the procedure. In most LGG patients, this sort of operation is done regularly and sometimes does not require full anaesthesia with sedation – depending on the location of the tumour within the brain.[385]

Other treatment procedures include radiotherapy and chemotherapy. The role of radiotherapy in the management of patients with LGG still remains very unclear. Several clinical studies investigated the overall effectiveness of high and low doses of radiation response in LGG adult patients. In one study, 379 patients received radiotherapy post-operatively (after biopsy surgery), showing no overall significant difference in survival rates, compared to a single surgery treatment. There was also no significant difference in survival rates between high and low dose radiotherapy (58% and 59% survival rates, respectively).[386] Moreover, in another randomised clinical study, there was a high incidence of radiation neurotoxicity (adverse effect on nervous system) coming from a high and low dose therapies – 5% and 2.5% incidence rates, respectively.[387] Another study showed that long term radiation therapy caused significant cognitive deficits in almost half of the study subjects (patients), affecting their memory.[388] Therefore, there is still uncertainty whether to use radiotherapy in LGG patients or not, with higher chances of risks and the development of complications, outweighing benefits. Chemotherapy, on the other hand, has been regularly used in all types of brain cancer, more often in high grade gliomas, and certain drugs with lower doses in LGGs. Oncologists are mainly focusing on a drug called temozolomide, an alkylating agent used as an anticancer drug in brain

gliomas, mainly in high grade GBM and low grade astrocytoma.[389,390] The drug has shown a distinctive advantage over other alkylating agents such as procarbazine or lomustine, due to its ability to pass the blood–brain barrier more efficiently, being more active against tumour cells and showing less toxic adverse reactions.[391] In one clinical phase 2 study of astrocytoma grade 2 cancer, this drug has managed to demonstrate excellent pharmacological anticancer properties – out of 46 patients tested with the drug, 61% achieved complete or partial drug response with a large reduction in the number of tumour cells, and 78% of patients showed prolonged survival rate for an additional 22 months.[392] Clinical phase 2 studies with temozolomide applied in LGG patients (astrocytoma) after the surgical procedure showed even better results, as expected, with significant improvement in survival rate among 250 participants for almost 3 years before LGG turned into a high grade type where the drug was no longer effective.[393] However, the combination of radiotherapy with temozolomide or other alkylating drug (vincristine) in high risk LGGs studies could not show an overall improvement in survival rates, compared to the studies where only the drug was used, once again showing that chemotherapy did not reduce cancer cells significantly.[394,395]

Treatment management of LGGs is constantly changing, with new strategies being proposed, balancing the risks and benefits in reducing tumours. Considering the evidence currently available from the research in LGG treatment, the optimal treatment strategy may be a combination of surgical resection followed by chemotherapy (with a drug such as temozolomide), delaying radiotherapy for as long as possible.

5.1.4. Low grade glioma: molecular pathology, metabolic pathways and potential biomarkers

Molecular pathology studies of LGG help scientists in predicting tumour response to treatment, by developing a thorough understanding of the pathogenesis of those tumours. Most pathological mechanisms are related to genetic mutations and alterations in important metabolic pathways. The presence of LGG related genetic and metabolic alterations in tumours affect the patient's response to specific cancer therapies, and so an accurate screening of those alterations can assist scientists in deciding an optimal treatment based on molecular characteristics of the tumour.[396] These monitored changes can lead to the discovery of potential biomarkers, which could be specific genes, proteins, cell receptors or metabolites in signalling pathways of the CNS. This information also shows how signalling pathways are generally affected and which of them may be targeted with specific therapies. With current standard

treatment strategies in place involving surgery, chemotherapy and radiotherapy, large variations may exist in clinical patient data (survival rates of patients, number of reduced tumour cells) with various types of glioma, including LGG. Therefore, there is a need for specific biomarkers and metabolic pathways identification for making an accurate prognosis of a cancer, sub classification of LGGs based on molecular parameters and searching for optimised treatments targeting specific areas of the LGG.[397] Research into molecular pathology of LGG has revealed several affected metabolic pathways which are responsible for controlling cell growth and death. Some of these key metabolic pathways are phosphoinositide-3-kinase (PI3K), MAPK and tricarboxylic acid (TCA or Krebs) cycle, all of which are crucial in regulating cell proliferation and preventing cancer.[398] These pathways can be greatly affected by genetic mutations caused by LGG, and contribute to energy production in glioma cells, increasing their number and the size of tumour.

As mentioned earlier, IDH 1 and 2 gene mutations occur in the vast majority of LGGs. These gene mutations were found to have a strong effect on the metabolic flux of isocitrate, NADP, glutamate and glutamine, which play a major role in the TCA cycle.[399] The TCA cycle is known to generate a pool of chemical energy (in the form of ATP) for cell growth and proliferation, by a series of oxidation processes of certain metabolites such as acetyl-CoA, NADP and isocitrate. Under normal circumstances, IDH 1 and 2 genes encode catalytic enzymes promoting these oxidative reactions: decarboxylation of the metabolite isocitrate to alpha-ketoglutarate, reduction of NADP to NADPH, and so on. The oxidation of alpha-ketoglutarate to succinate helps to control cell proliferation and suppress tumour cell growth.[400] However, when mutated, the IDH genes become deactivated and change the function of catalytic enzymes, altering cellular metabolism and causing unwanted angiogenesis (proliferation) of tumour cells – the oxidation of isocitrate to 2-hydroxyglutarate, an oncogenic metabolite.[401] Such mutations are considered oncogenic, although their relevance to LGG is still not fully understood and is currently under investigation.[402] Recent genetic sequencing studies demonstrated high IDH mutations present in 59–90% of grade 2 astrocytomas and in 68–85% of grade 2 oligodendrogliomas, indicating that these mutations play an important role and definitely can be used as potential biomarkers.[403] IDH-mutant LGG astrocytomas are found to be less aggressive and less cancerous than IDH-mutant GBM, and therefore may require less aggressive treatment such as chemotherapy.[404] Several studies have demonstrated a great impact on IDH mutations in response to the temozolomide drug in low grade astrocytomas, reducing the number of tumour cells significantly[405,406], while other studies have shown no improved response to temozolomide chemotherapy at all.[407] Therefore, more research is required into IDH gene mutations in order

to identify potential biomarkers. Figure 5.3 shows how the TCA cycle is thought to be affected by such mutations.

Another potential pathogenic mechanism for LGG appears to be methylation of a DNA repair protein O6-methylguanine-methyltransferase (MGMT). MGMT is a DNA repair enzyme that functions by removing alkyl groups from guanine nucleotides and repairing DNA adducts. However, methylation of the promoter region of the MGMT gene leads to deactivation of the protein function, reducing DNA repair activity and initiating cancer.[408] MGMT methylation has also been observed to affect the sensitivity of gliomas to antineoplastic and alkylating drugs, increasing the efficiency of temozolomide drug response, for example. Although the exact mechanism of the drug efficiency increase is unknown, increased MGMT activity and DNA repair deficiency are suggested as the reasons.[409] MGMT methylation usually occurs in about 50% of LGG astrocytomas and in about 40% of GBM high grade cancers, making it the most common cause of gliomas. Therefore, MGMT methylation status is currently investigated as a potential biomarker. It was shown to have a massive impact on the TCA cycle, as MGMT methylation is strongly associated with IDH gene mutations.[410]

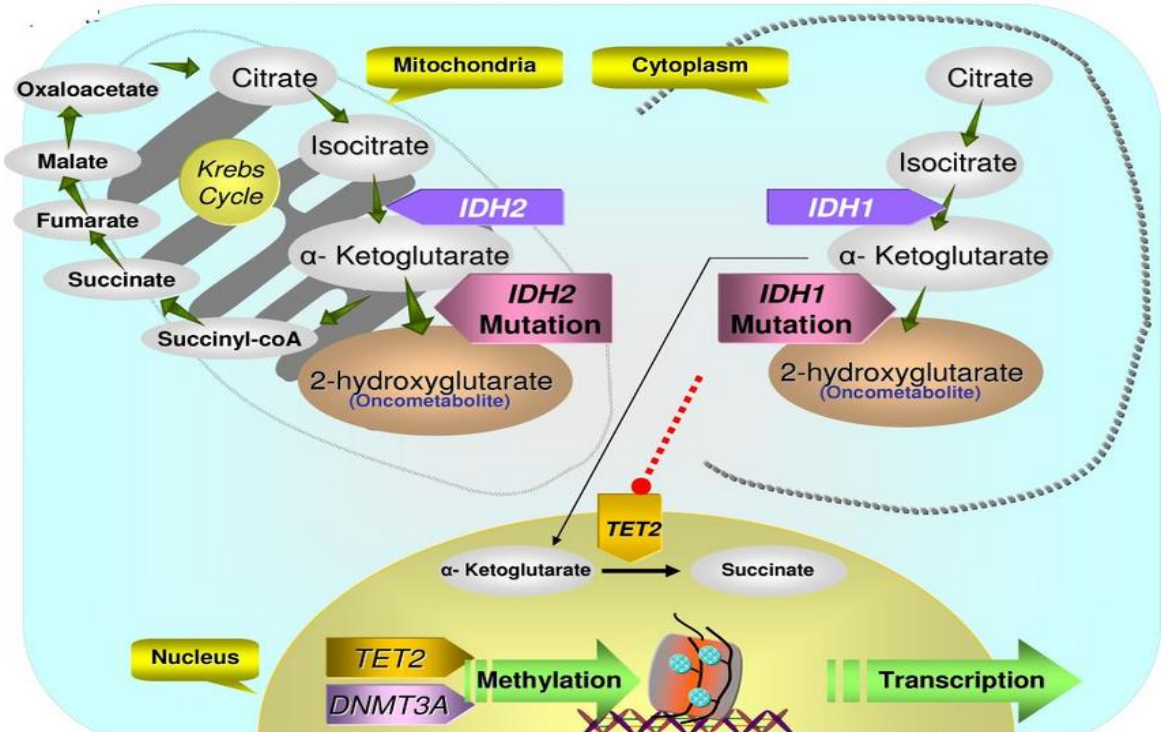


Figure 5.3 Effect of IDH 1 and 2 gene mutations on the metabolic activity in the TCA cycle. H.Yang et al, *IDH1* and *IDH2* Mutations in Tumorigenesis: Mechanistic Insights and Clinical Perspectives. 2012

Other gene mutations also associated with activation of LGG cancer-related signalling pathways, occur in a variety of proteins such as “tumour protein 53” (TP53), BRAF kinase protein, Ras protein, and phosphatase and tensin homolog (PTEN) tumour suppressor protein, although their full pathogenicity has not yet been fully understood.[411,412] For example, gene mutations in PTEN tumour suppressor protein are known to over activate the PI3K pathway.[413] PTEN gene mutation involves methylation of a protein and its deactivation, which in turn increases phosphorylation of protein kinase B found in the PI3K metabolic pathway. This triggers the angiogenesis of tumour cells.[414] Although PTEN mutations are generally rare in LGGs, these tumours may exhibit lower levels of PTEN protein and high levels of phosphorylated protein kinase B, but mostly it happens in high grade GBM.[413,415] The Ras signalling pathway (also known as MAPK pathway) is usually activated in pilocytic astrocytoma, an example of LGG, increasing tumour cell growth and proliferation.[416] The MAPK pathway is activated by a variety of methods including increased activity of Ras protein caused by mutations in the NF1 gene[417], TP53 protein caused by mutations in the TP53 gene[418], and BRAF kinase protein by mutations in the BRAF gene[419]. These mutations are more common in LGGs than in high grade gliomas. It has been found that TP53 mutations are more effectively targeted by radiotherapy than chemotherapy or surgery, while NF1 and BRAF gene mutations are investigated by proposed drugs such as BRAF and Ras inhibitors, which are currently being designed and tested.[420]

In low grade oligodendroglioma and astrocytoma, a common alteration is usually observed to be codeletion of chromosomes 1p and 19q[421], which also affects the production of important protein suppressing tumours, such as capicua transcriptional repressor (CTR) and FUSE binding protein (FUBP).[422] The gene coding for the CTR protein is located in chromosome 19q, while the gene for FUBP protein is on chromosome 1p. The deletion of these genes during LGG disease will reduce levels of their suppressor proteins produced, elevating tumour growth. In one study it has been observed that long term chemotherapy with temozolomide, targeting the LLGs caused by 1p and 19q codeletions, leads to a progressive decrease in “mean tumour diameter” after the onset of a drug in almost all of the 36 patients suffering from LGG (94.4% success). In some patients (27%), the mean tumour diameter increased again after the drug was discontinued for several months. This study demonstrated that 1p and 19q chromosome codeletions can be an effective biomarker for LGG, being successfully targeted by the temozolomide drug.[423]

All of the above mentioned gene mutations and related proteins affecting metabolites’ activity and related metabolic pathways, are currently under investigation as potential biomarkers for LGG.[424] One way to assist the monitoring of metabolic pathways and search for potential

biomarkers for LGG is the use of the untargeted metabolomics technique. Brain tissue samples from sufferers affected by LGG can be analysed using LC-MS metabolomics to help scientists in the search of biomarkers and affected biological pathways, and to observe closely how LGGs can progress to high grade gliomas.

5.1.5 The importance of intra-tumour heterogeneity and central core

A variety of metabolomics, lipidomics, genomics, proteomics, cellular morphology studies and other relevant biochemical tests have been performed on LGG brain tissues in the past, in search of potential biomarkers and related metabolic, lipidomic pathways affected by the disease.[425-427] However, accumulating evidence suggests that intra-tumour heterogeneity (ITH), defined as an observation of distinct morphological and phenotypic profiles between biological samples taken from the **same** tumour, is very poorly understood.[428] For example, in brain tumours, ITH may show the morphological and phenotypic differences between individuals with the **same** brain tumour and even between tissues taken from the same brain location within just one individual. This phenomenon occurs due to a simple consequence of DNA gene mutations. This opposes inter-tumour heterogeneity, where the differences are observed between different types of tumours. Investigating ITH in brain tumours can help to obtain better data on tumour pathology, design more effective treatment strategies and reduce current treatment failures.[429]

So far, only a limited amount of research has been done with respect to ITH in brain tissues affected by LGGs or high grade gliomas. In the past, a variety of metabolomics, lipidomics, morphological and gene expression studies have only been performed on LGG brain tissue to observe the general differences between control and disease states in search of biomarkers and inter-tumour heterogeneity.[430] Recently, a GBM high grade glioma study in human volunteers has demonstrated clear differences in metabolite activity not only between control and disease samples as expected, but also a presence of ITH between brain tissues taken from specific parts of brain, for **each** individual. In that study, several anatomical surgical resections of brain tissues (also known as plane of views) from each volunteer were shown to express differences in metabolic activities, fully demonstrating ITH and difference in metabolism within one individual's brain tissues.[431] Those surgical resections are normally applied to study the human brain physiology in more details when brain tissues are affected by certain disease, giving a physician more information on what parts of the brain affected and how they are different. Therefore, it seems that even for each individual affected by tumour, a certain degree of

metabolite difference can be expected when comparing different brain tissues taken from the same individual – it can help to detect potential biomarkers more accurately. Figure 5.4 shows the typical brain anatomical plane of views, usually applied in sectional anatomy to examine the brain structure in physiology, including “superficial”, “anterior”, “central core”, and “posterior” views.

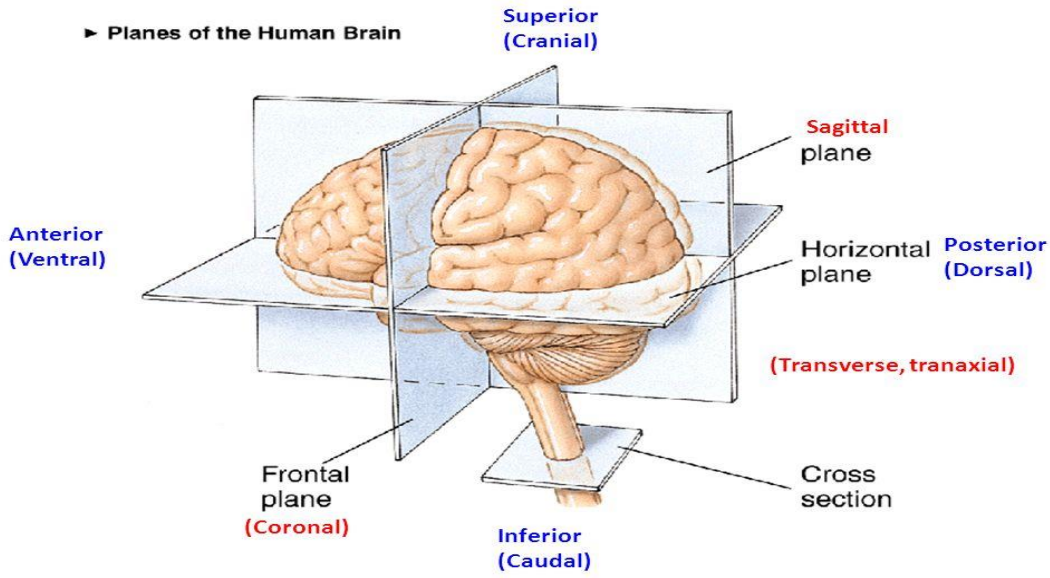


Figure 5.4. Human brain tissue anatomical surgical resections (or planes of view). **J. Park et al, A proposal of new reference system for the standard axial, sagittal, coronal planes of brain based on serially-sectioned images, 2010.**

The central core is one of the most important functional parts of human brain. The central core consists of five main regions helping to regulate basic life processes including breathing, movement and sleep. It stands deep inside each hemisphere of the brain, as a solid block between brainstem and cerebral lobes. When a tumour such as LGG affects brain tissues, the large presence of cancerous cells are normally detected in the central core. The non-central core brain tissues include the margins outside the central core and are normally close to residual cells. Residual cells are malignant cells which are left untouched and stay behind after surgical treatments, having a strong resistance to surgical therapy (and chemotherapy), creating major driving force of mortality. They are found difficult to remove during surgery and may cause further tumour progression. This is why they are one of the unique cellular targets in brain tumours and are worth investigating.[432] As they are found close to the non-central core, physicians like to study morphological and metabolomics differences between the central and non-central core, in order to further understand the role of residual cells in cancer development.

Therefore, understanding the differences between central and non-central core metabolism, as well as detecting ITH presence between brain tissues, help the scientists to fully observe the effect of glioma on different parts of human brain and contribution in pathogenesis – this may help scientists to discover personalised medicine for each individual against cancer.

This chapter investigates the effect of LGG grade 2 IDH-1 mutated astrocytoma on metabolic and cellular morphology of human brain tissues, collected via biopsy from five adult patients suffering from the same condition. The ultimate goal of the study was not only to identify potential biomarkers (metabolites, lipids) between control and disease states and the effect on related pathways, but also to observe presence of ITH between brain tissues collected from each individual, as well as to compare central to non-central core metabolism affected by glioma.

For the purposes of this chapter, two hypotheses have been set up:

1. There is a significant difference in metabolic activity between different brain tissues (from different **anatomical plane of views**) for **each individual**, confirming ITH presence.
2. There is a significant difference in metabolic activity between **central and non-central** brain core tissues, for each individual.

These hypotheses were designed to observe if there is indeed an ITH between those individuals suffering from LGG astrocytoma by looking into their metabolic profile of brain tissue.

5.1.6 Aims and objectives

With the above hypotheses in mind, the following aims and objectives of this study, for the purposes of this chapter, are:

1. To perform an LC-MS data analysis on human LGG astrocytoma brain tissue to obtain metabolome information of brain tissues affected by LGG

- Extract and prepare brain tissue samples
- Run the samples with the LC-MS instrument (same methodology as in previous chapters) with applied normalisation technique
- Process the data with IDEOM, mzMatchIso software for obtaining metabolome and assessing the efficiency of labelled metabolites
- Test normalisation of the data as performed in chapter three

2. To address the two hypotheses by:

- Designing and running multivariate analysis on the data to create specific PCA and OPLS-DA models (discussed later)
- Running univariate analysis to obtain the list of statistically significant metabolites for each hypothesis
- Linking those statistically significant metabolites to metabolic pathways in order to obtain potential biomarkers
- Using the obtained information to answer the hypothesis questions set by the study

5.2 Materials and methods

5.2.1 Tissue sample collection and information

In this study, five different volunteers (adults, over 25 years of age), with confirmed astrocytoma IDH-1 mutant grade 2 cancer, were selected to provide their brain tissue. The volunteers were recruited within the UK NHS Trust centre at Nottingham University Hospital NHS Trusts, Queen's Medical Centre. The study has been granted ethical approval from the National Research Ethics Service Committee from East Midlands, UK (ethics approval reference 11/EM/0076). Briefly, a small amount (close to 150–200mg weight) of brain tissue from various brain locations was surgically removed from each participant by biopsy. The procedure was done over a period of several days. The brain locations from which tissue samples were removed are summarised in table 5.1. Each participant was assigned a code for confidentiality purposes, with the letters LGG followed by a number for participant and another number for brain location. For example, one patient had codes LGG 2.1 for one location, LGG 2.2 for another and so on. Brain tissue, from each participant and each brain location, was further cut into five smaller pieces (20mg–40mg) for the purpose of creating biological replicates. The tissue was cut on a glass slide over dry ice using a disinfected scalpel, to prevent it from thawing, and each small tissue fragment was transferred into a 2mL Eppendorf tube. There were no analytical replicates created in this study.

The control sample for this study was brain tissue from the temporal lobe region of a volunteer who did not have LGG astrocytoma, but suffered from epilepsy. The ideal control brain tissue sample should have been cancer free (no presence of tumour) and collected from a participant with LGG. However, as it is almost impossible to retrieve any non-diseased brain tissue from a cancerous brain, it was decided in this study to obtain the control non-cancerous tissue from a volunteer with epilepsy, who suffers a similar adverse effect (e.g. seizure) found in LGG. This way it made the control sample closer to the physiological nature of the experimental samples.

Overall, there were 24 LGG tissues (with five biological replicates of each) and five biological replicates of control sample, giving in total 125 tissue samples. The tissue samples were put into Eppendorf tubes, labelled appropriately and stored in -80°C freezer conditions. Table 5.1 summarises the replicate brain tissue samples collected from each participant.

Patient and code	Code number for tissue sample and location	Brain anatomical plane of views	Approximate weight of collected tissue (mg) used for each biological replicate (5 replicates)
Patient ONE (LGG 2)	LGG 2.1	Superficial	20, 20, 40, 40, 40
	LGG 2.2	Central core	30, 40, 40, 30, 40
	LGG 2.3	Superior	20, 30, 30, 30, 40
	LGG 2.4	Posterior	20, 40, 30, 30, 40
Patient TWO (LGG 5)	LGG 5.1	Superficial	30, 30, 40, 20, 20
	LGG 5.2	Anterior	20, 20, 30, 30, 30
	LGG 5.3	Posterior	30, 40, 30, 30, 20
	LGG 5.4	Central core	30, 40, 30, 20, 30
	LGG 5.5	Inferior	30, 40, 30, 30, 30
Patient THREE (LGG 7)	LGG 7.1	Superficial	40, 30, 40, 30, 30
	LGG 7.2	Central core	30, 30, 40, 20, 20
	LGG 7.3	Posterior	20, 20, 20, 20, 20
	LGG 7.4	Anterior	30, 30, 30, 30, 30
	LGG 7.5	Medial	20, 20, 20, 20, 30
Patient FOUR (LGG 11)	LGG 11.1	Rim	30, 20, 20, 40, 30
	LGG 11.2	Central core	40, 30, 40, 40, 30
	LGG 11.3	Inferior	40, 40, 40, 40, 40
	LGG 11.4	Lateral	30, 30, 30, 30, 40
	LGG 11.5	Invasive margin	20, 30, 20, 20, 20
Patient FIVE (LGG 12)	LGG 12.1	Superficial	30, 30, 30, 30, 30
	LGG 12.2	Anterior	40, 30, 30, 40, 40
	LGG 12.3	Posterior	30, 30, 30, 40, 40
	LGG 12.4	Central core	20, 30, 40, 20, 20
	LGG 12.5	Deep	30, 40, 20, 40, 30
Control (Epilepsy patient)	Control	Temporal lobe	30, 30, 20, 20, 30

Table 5.1 Brain tissue samples from each volunteer taken from different brain plane of views

5.2.2 Sample preparation and extraction

Each of the LGG tissue sample and control sample replicates were prepared in the same way, following the same extraction protocol, in randomised order. Brain tissue was precooled at 4°C and mixed with (U)-^{13}C labelled *E. coli* extract solution, proportional to the weight (100µL of solution per 10mg of tissue). Disruption of the tissue in extract solution was performed using a hand-held homogeniser (Bibby Scientific Stuart, SHM1) until the disappearance of visible clumps. Prior to using the homogeniser for each tissue, the device was cleaned with 70% IMS and Chem G disinfectant by dipping the rotor into the solutions. 300µL (proportional to 10mg tissue) of ice cold (4°C) chloroform (HPLC grade) was then added to each sample and vortexed for 30 seconds, followed by administration of 100µL (proportional to 10mg tissue) of 4°C deionised (purified) water (HPLC grade) and further vortexing for 30 seconds. These steps were carried out in 2mL Eppendorf tubes with rounded bottoms to make it easier for homogenisation. Samples were then centrifuged for 10 minutes at 13000 x g at 4°C (Harrier Centrifuge) to separate the two phases (polar and lipid) for further extraction. The upper phase (polar) was collected first using a pipette and was transferred into fresh pre-cooled Eppendorf tubes, labelled and placed in the -80°C freezer until LC-MS analysis. By carefully displacing the protein sediment layer between the two separated phases, the lower phase (lipid) was collected and transferred in fresh Eppendorf tubes, labelled and stored in the -80°C freezer. Note that the lipid fractions were stored for use in a separate, completely different project not related to this thesis. Prior to LC-MS analysis, brain tissue polar extracts were transferred into LC-MS vials. The QC sample was prepared by pooling 5µL from each sample, including the control, into one Eppendorf tube.

5.2.3 LC-MS analysis

The LC-MS analysis was carried out exactly as previously described in chapter two (section 2.2.1 and 2.2.2) in a randomised sequence order. Ten injections of pooled QC sample were analysed at the beginning of the run to condition (stabilise) the column before the analysis of randomised clinical samples. Pooled QC sample was then injected once after each randomised 10 samples for the purpose of monitoring the stability, repeatability, robustness and performance of the analytical run. A reagent blank (methanol) was injected alongside every QC sample to minimise any possible carryover effect.

5.2.4 Data processing and statistical analysis

An identical data processing method as described in chapter four (see section 4.2.3) has been applied to the clinical raw data from brain tissue samples. The same software programmes have been applied to a raw data set: XCMS for untargeted peak-picking[144], mzMatch for alignment and annotation of peaks[142], IDEOM for noise filtering and metabolite putative identification.[147] During processing with IDEOM, the raw data was grouped according to brain region and patients. Once the list of putatively identified metabolites was generated, showing a number of detected metabolites in each group (patient, brain location), further analysis was performed on metabolite peak height intensities for normalisation. The same approach was performed as in chapter three (see section 3.2.4), generating ratios of $^{13}C/^{12}C$ peak height intensities for each metabolite in each biological replicate for each group. ^{13}C peaks were identified using the proposed Python code, used in the validation experiments. The quality of the metabolomics analysis was assessed by calculating percentage RSD values (normalisation), but they were not compared to "non-normalised" or other normalisation techniques, as it is beyond the scope of this chapter.

For multivariate analysis, the decision had to be made on how to compare the brain tissue data groups effectively, in order to answer the set hypotheses. As the hypotheses were testing intra-tumour heterogeneity within patients, as well as comparing central and non-central core metabolism, various matrix tables were created – see the results section 5.3. Based on these matrix tables containing normalised data (ratios), PCA and OPLS-DA plots were constructed in the same manner as in previous chapters three and four, comparing control to tumour (treatment). VIP score analysis was done in every OPLS-DA plot, choosing metabolites with a score of 1.0 or above only. Univariate analysis was performed on the same grouped data as used in multivariate analysis, creating a list of statistically significant metabolites between control and treatment groups. The parameters for T-test and FDR statistical correction were chosen to be the same as in chapter three and chapter four (p value of less than 0.05, FDR rate 5%). These lists were then compared to their corresponding multivariate VIP score lists to produce final lists of statistically significant metabolites. These metabolites were then compared to research papers, in order to confirm the identification of potential biomarkers and to observe which metabolic pathways were affected. The fold changes obtained from univariate analysis were used to decide whether a potential biomarker was elevated or decreased during the "disease" condition compared to the control.

5.3 Results

5.3.1 Demographic data on chosen participants

The following table 5.2 summarises the demographics of participants involved in the study. There were six participants involved, five of which had a presence of LGG grade 2 astrocytoma and one who had epilepsy only (set as a control). The brain tissue was collected from a certain anatomical part of the brain (lobe) of a participant. Inside the particular lobe, the brain tissues were taken from various anatomical resections (plane of views), for example superior, anterior, posterior and superficial (described in table 5.1 and shown in figure 5.4). For the purposes of this study, the ethnicity, age and gender of an individual was not particularly relevant, as the aim of the study investigated metabolic differences between brain locations and individuals, rather than specific ethnicity, age or gender differences – it was difficult to recruit participants with the same genetic characteristics, due to limited choice of available patients at the time.

Patient code	Age	Gender	Ethnicity	Anatomical brain lobes	Condition
LGG 2	25	Male	White British	Parietal lobe	LGG astrocytoma grade 2
LGG 5	37	Female	White British	Left frontal lobe	LGG astrocytoma grade 2
LGG 7	27	Female	Asian British	Right frontal lobe	LGG astrocytoma grade 2
LGG 11	19	Female	White British	Left frontal lobe	LGG astrocytoma grade 2
LGG 12	32	Male	White British	Left frontal lobe	LGG astrocytoma grade 2
Control	18	Male	Asian British	Temporal lobe	Epilepsy

Table 5.2 Demographics data of each volunteer participating in the study (names are omitted for confidentiality purposes).

5.3.2 Analytical performance and metabolomics data quality

Metabolite profiling of brain tissue samples

IDEOM data processing revealed 497 putatively identified metabolites overall (407 in common between control and treatment groups). The majority of detected metabolites in all of the samples belonged to biochemical classes such as amino acids, carbohydrates, nucleotides and other polar compounds, with a certain proportion of lipids (non-polar class). The composition of biochemical classes is shown in figure 5.5. It was expected that a large number of hydrophilic

compounds would be observed, as the extraction of brain tissue samples was performed in hydrophilic conditions with polar extraction solvent (methanol).

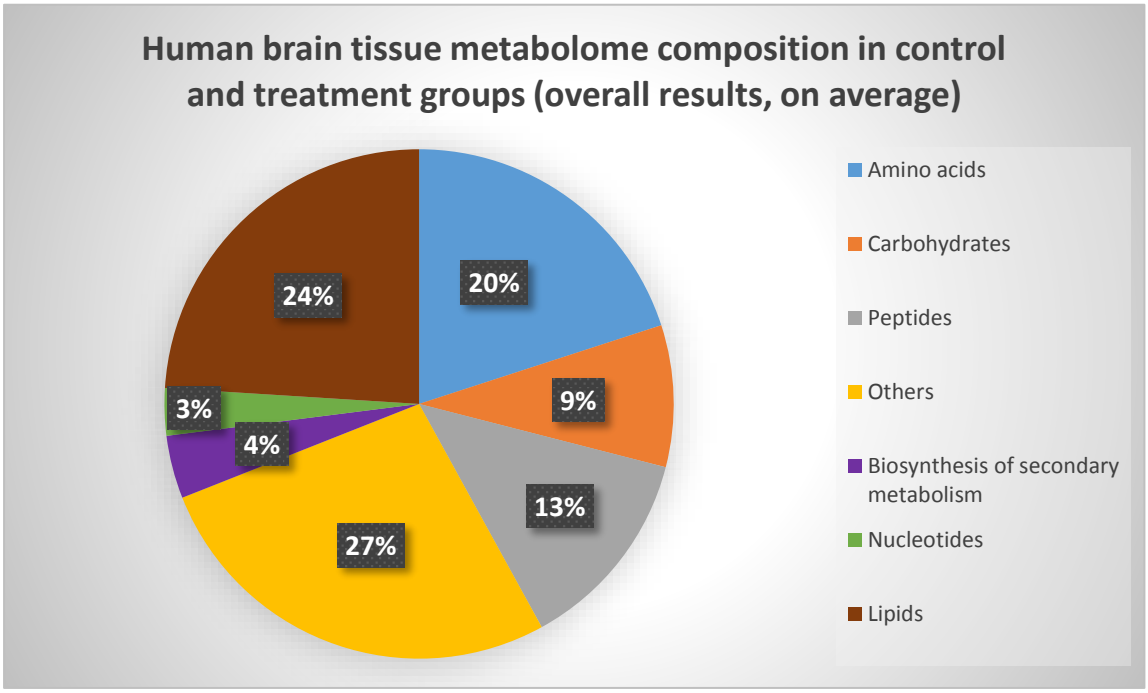


Figure 5.5. Composition of putative metabolites in human brain tissue samples (control and treatment groups) according to biochemical class

LC-MS analytical performance

Analytical runs of brain tissue samples (treatment groups and control) have generated clear spectra with distinctive metabolite peaks, high TIC values (≥ E^8) and no background noise or carryover effects. The spectra of QC samples in both experiments also showed clear peaks with high TIC values, demonstrating a good analytical stability of sample metabolites and reproducibility. The shape of the spectra peaks of QC samples was found to be similar to the control and treatment spectra, with some peaks identical, showing that QC samples contained the common metabolite peaks as the actual samples. Examples of some spectra graphs are shown on the next page in figure 5.6. It can be seen that the spectra of the selected samples (Patient one with four different brain locations, control and QC samples) all look very similar to each other in the number of peaks and their shapes.

During data processing, QC samples were also analysed to check for analytical performance of the experiment. The first 10 QC samples in the analytical run were not included in data processing, as they were used to stabilise the system. The other 10 QC samples were analysed by IDEOM software. The peak height intensity of every metabolite in each of the 10 QC samples

was calculated and percentage RSD values were determined to observe the degree of data variation of metabolites between QC samples. It has been observed that the majority of detected metabolites (374 out of 497 metabolites) expressed an RSD of less than 30% between QC samples – that is before normalisation calculations. This demonstrated good analytical performance and instrument suitability throughout the experiment.

Normalisation data analysis

Once the normalisation analysis had been applied to datasets of brain tissue samples, percentage RSD values of peak height intensities of every metabolite between five biological replicates were calculated in each control and treatment group, in the same way that the data was normalised in validation experiments in chapter three. Out of 497 metabolites, on average, around 70% of metabolites were normalised successfully in each group (from 340 to 375). The majority of normalised metabolites were from amino acids, carbohydrates and other polar classes. Metabolites which were not successfully normalised, did not have a suitable internal standard that could have lowered their percentage RSD – mainly they were of lipid nature. There was no significant difference in the number or nature of successfully normalised metabolites between the data groups (patients and brain locations). Further quality of normalised data can be observed through multivariate PCA and OPLS-DA plots (see the next section), where the data points were observed to be clustered and data groups were well separated, showing reduced data variation.

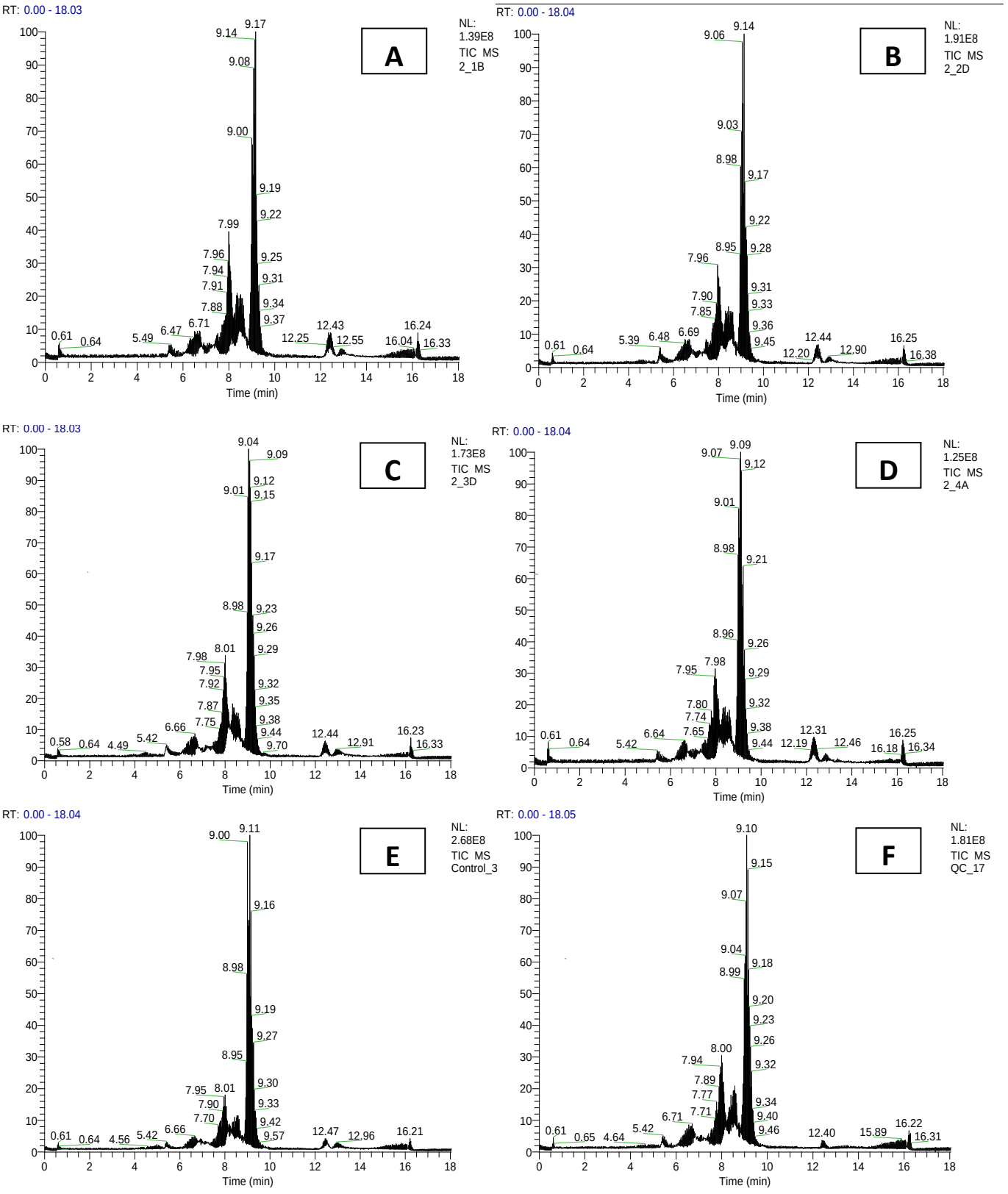


Figure 5.6. Examples of LC-MS spectra of human brain tissue samples from patient ONE (LGG 2) with four brain locations: superficial, central core, superior, posterior (A-D). Also, from Control sample (E) and QC sample (F).

5.3.3 Metabolomics analysis of LGG brain tissue

Normalised metabolomes of brain tissue samples ($^{13}C/^{12}C$ ratios) were arranged in matrix tables and passed through multivariate analysis to obtain PCA and OPLS-DA plots, for further statistical analysis. The data was arranged according to the tested hypotheses. For the first hypothesis, where metabolic activity between different brain plane of views was compared for each patient, five matrix tables were created (one for each patient). For the second hypothesis, where metabolic activity was compared between central and non-central core brain tissues, one single matrix table was designed separating central and non-central core data points.

Figures 5.7a and 5.7b, on the next two pages, demonstrate multivariate data analysis based on the first hypothesis. In those figures, five OPLS-DA score plots show separation between brain tumour tissue (from various anatomical resections) and control data points, for each of the LGG patients – patients datasets (patient ONE to FIVE, A–E, respectively). The separation between the data groups (treatment and control) in each of the OPLS-DA plots is very clear with observed cluster of data points. High Q2 scores are also observed on the models (0.646, 0.829, 0.872, 0.747 and 0.821 for Patients ONE to FIVE, respectively). This shows that there is a difference in metabolic activity between control and disease states but also a presence of ITH, observed in every patient.

In figure 5.8, the OPLS-DA scores plot shows the relationship in metabolic activity between central and non-central core brain tissue samples (all of the other types of brain locations grouped together), for all of the LGG patients, supporting hypothesis 2. The OPLS-DA showed a clear separation between central core versus non-central core data groups, with clustered data points and high Q2 score (0.897), demonstrating an expected biological difference between the groups.

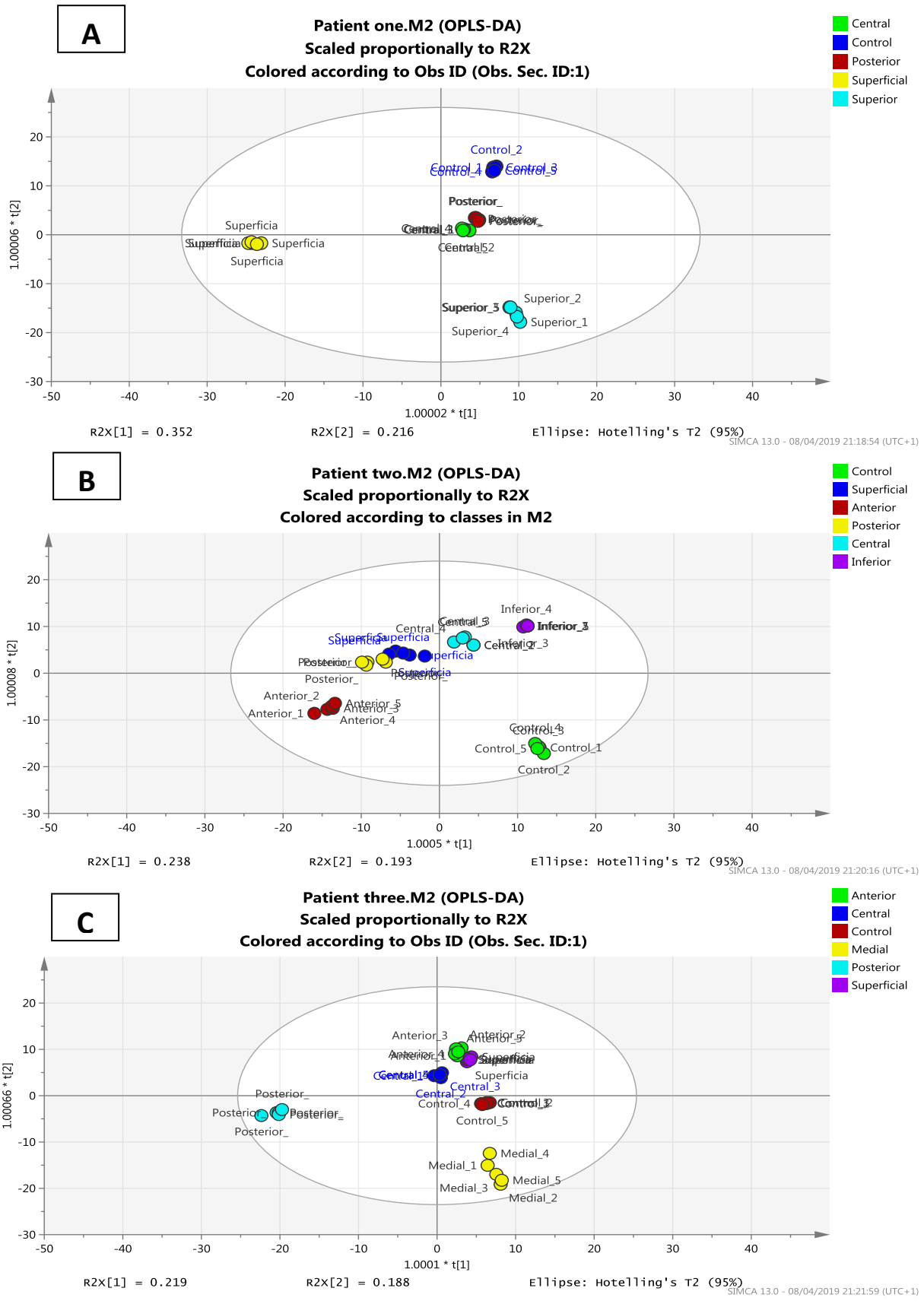
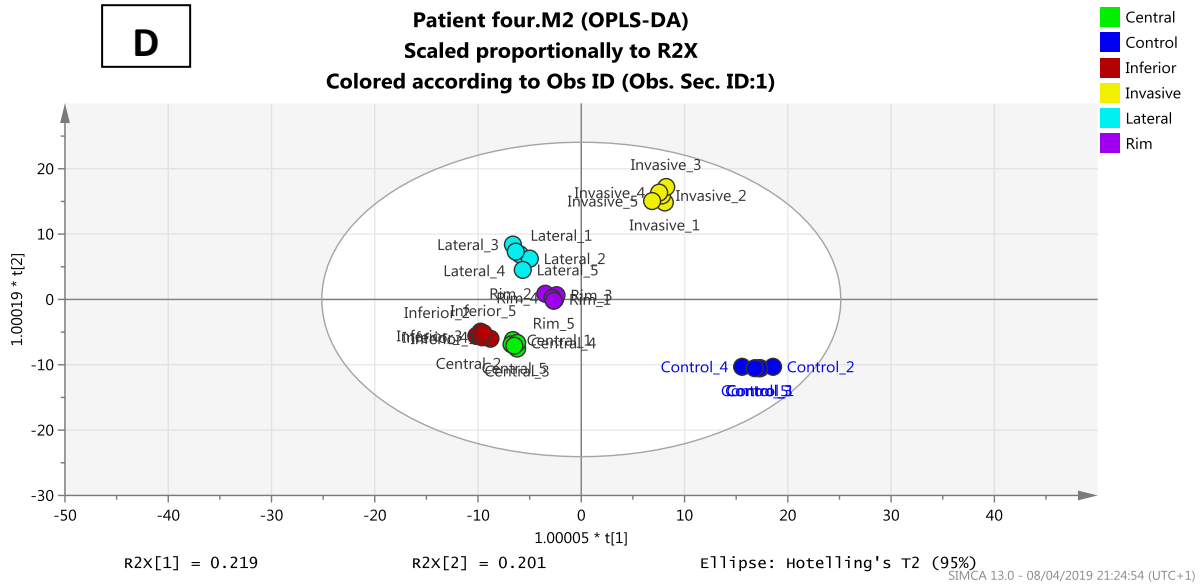


Figure 5.7a. OPLS-DA score plots of control and LGG brain tissue from various locations (treatment) data groups for **Patients ONE (A), TWO (B) and THREE (C)**. The Q2 score values for Patients ONE, TWO and THREE were calculated as 0.646, 0.829 and 0.872, respectively.



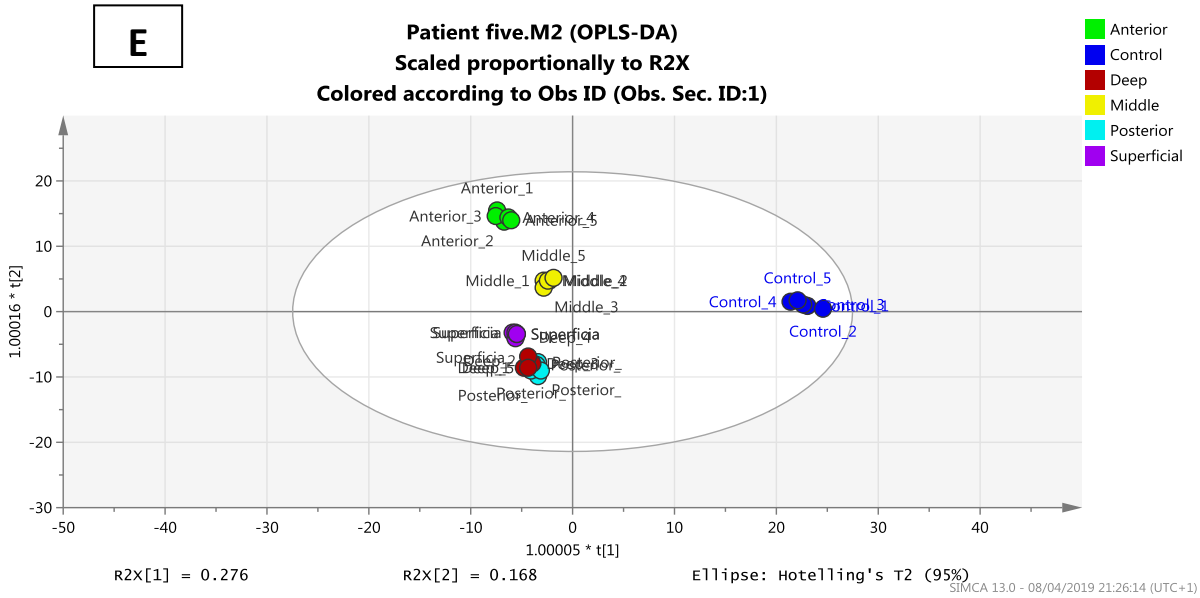


Figure 5.7b. OPLS-DA score plots of control and LGG brain tissue from various locations (treatment) data groups for **Patients FOUR (D) and FIVE (E)**. The Q2 score values for Patients FOUR and FIVE were calculated as 0.747 and 0.821, respectively.

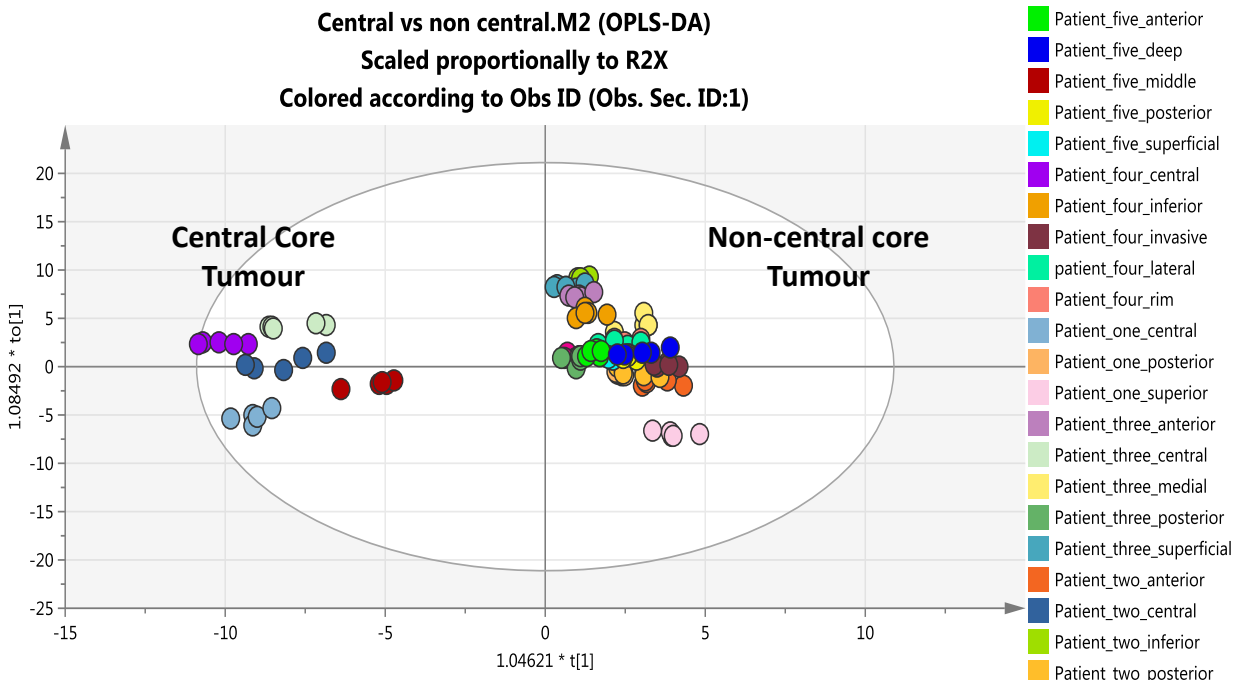


Figure 5.8. OPLS-DA score plot of central core versus non-central core data groups, for all patients. The R2 score value was calculated as 0.576; Q2 score value was calculated as 0.897.

Overall, in all of the shown OPLS-DA models (for both hypotheses), it can be seen that there is a difference in metabolism between control and treatment groups, as well as expected difference between central and non-central core (figure 5.8). Further statistical analysis, with VIP score (multivariate) and univariate T-test and FDR statistical corrections, was performed to determine the significant differences in metabolic activity between these parameters in more detail, to assist in finding potential biomarkers.

5.3.4 Tentative identification of key significant metabolites (potential biomarkers) in brain tissue samples affected by LGG

First hypothesis (presence of ITH between brain tissues for each patient)

During multivariate analysis using OPLS-DA scores plots (see figures 5.7a and b), a list of metabolites with a VIP score of 1 or above has been generated for each patient dataset. The "Patient ONE" OPLS-DA scores plot has shown 274 metabolites with a high VIP score of 1 or above, in "Patient TWO" plot there were 256, in "Patient THREE" plot there were 281, in "Patient FOUR" plot there were 264 and in "Patient FIVE" plot there were 260 metabolites.

Univariate analysis involving T-test and FDR correction has expressed a certain number of statistically significant metabolites with different fold changes between control and treatment groups, for each patient. For "Patient ONE" data set, there were 110 statistically significant metabolites identified by T-test. For "Patient TWO", "Patient THREE", "Patient FOUR" and "Patient FIVE" data sets, there were 154, 104, 129 and 143 statistically significant metabolites, respectively. These metabolites had a p value of less than 0.05, making them significant between the compared control and treatment groups. These metabolites are displayed through volcano plots in figure 5.9, showing a distribution of positive and negative fold changes of metabolites – either being elevated or decreased in concentration, in each patient dataset (A–E for patient ONE to FIVE, respectively).

Based on the above results from both multivariate and univariate analysis, a list of statistically significant metabolites which passed T-test (p value < 0.05), FDR statistical correction (5% significance level) and had a VIP score of 1 and above, was generated for each of Patient ONE to FIVE datasets (see tables 5.3, 5.4, 5.5, 5.6 and 5.7). There were several statistically significant metabolites in common between patient datasets such as D-tryptophan, D-fructose-1,6-bisphosphate and 2-phospho-D-glycerate, discussed in more details in the next section. Most of these metabolites were identified as level 2 by metabolomics standard initiative, or with high putative confidence levels as presented in the table.[138]

As for chapter four**, statistical paired test** should have been performed in this chapter as well, as the study compared control and treatment groups in the same group of individuals (LGG astrocytoma patients). Details of the test are described in previous chapter (see 4.3.2). The overall results might have shown more clustering of replicate points on OPLS-DA plots with better separation of groups.

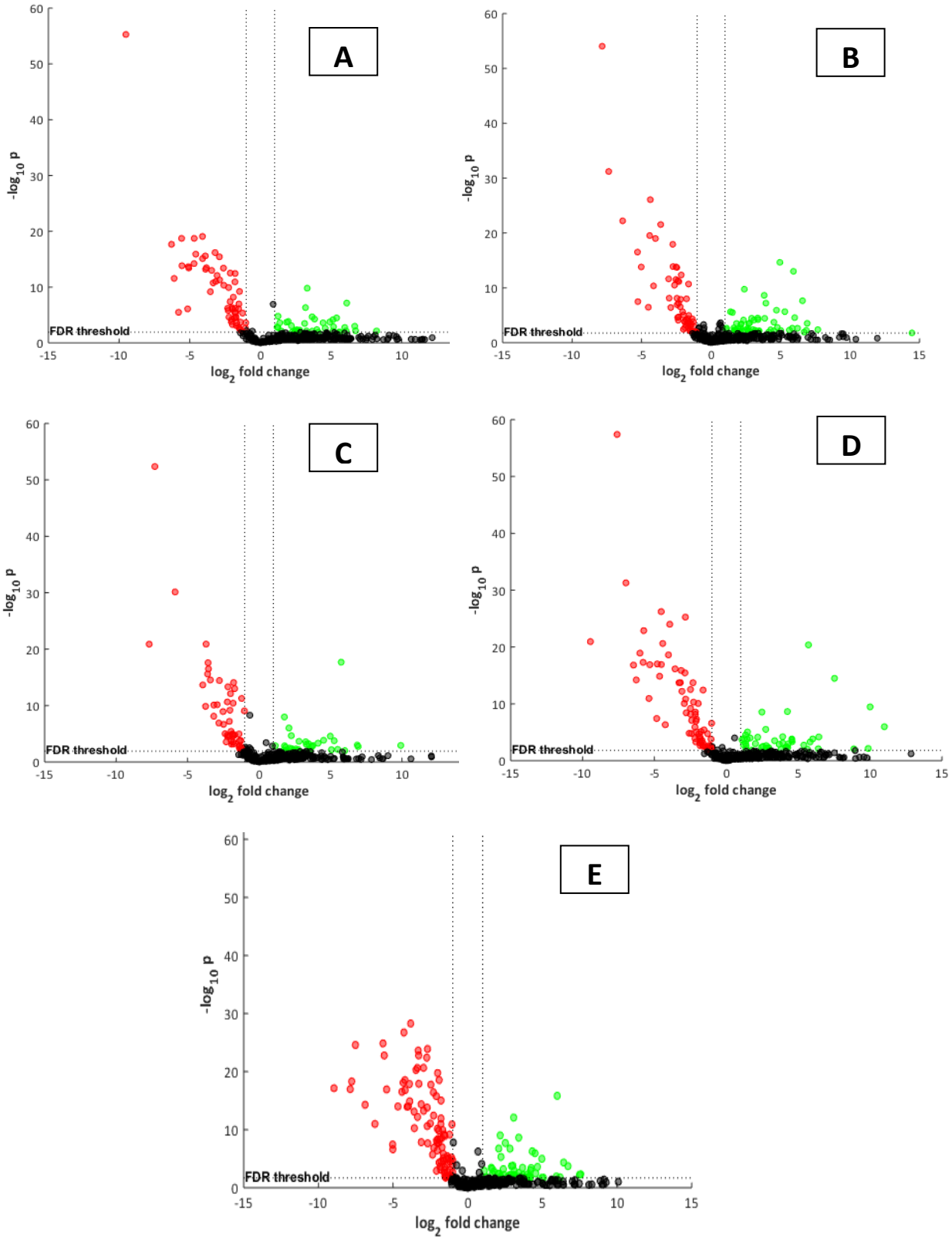


Figure 5.9. Volcano plots showing distribution of statistically significant metabolites detected between control and treatment groups in Patients ONE (A), TWO (B), THREE (C), FOUR (D) and FIVE (E). Green points represent positive fold change in significant metabolites, Red points represent negative fold change. Data is analysed by univariate analysis only.

Metabolite	Biochemical class	Retention time (mins)	Molecular mass	ID confidence	VIP score	P value	Fold change
D-Fructose 1,6-bisphosphate	Energy metabolism	9.07581	346.015	8	1.00695	<1.00E-17	1.356
2-Phospho-D-glycerate	Carbohydrates	8.98775	185.99	8	1.02937	1.89E-05	0.624
D-Tryptophan	Amino acids	8.41097	215.126	6	1.03453	<1.00E-19	2.471
Ecgonine	Biosynthesis of secondary metabolism	7.96432	194.135	5	1.02927	<1.00E-15	3.746
Isocitrate	Carbohydrates	7.79821	192.03	8	1.04528	3.42E-06	0.937
L-Rhamnulose	Carbohydrates	8.38324	170.088	6	1.04292	<1.00E-16	2.336
L-Histidine	Amino acids	9.23138	155.07	5	1.01435	2.98E-06	2.283
Glutathione	Amino acids	8.09843	307.08	8	1.13633	<1.00E-15	2.163
Carnosine	Amino acids	9.07357	235.136	8	1.12911	<1.00E-15	1.393

Table 5.3 List of statistically significant metabolites between control and treatment (brain locations grouped) in Patient **ONE**. The positive fold changes represent an upregulated metabolic activity and fold changes less than 1 represent downregulated activity affected by LGG. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Metabolite	Biochemical class	Retention time (mins)	Molecular mass	ID confidence	VIP score	P value	Fold change
2-Phospho-D-glycerate	Carbohydrates	8.98775	185.99	8	1.02937	1.75E-03	0.537
D-Fructose 1,6-bisphosphate	Energy metabolism	9.07581	346.015	8	1.00695	<1.00E-12	2.849
L-Hypoglycin	Others	7.03673	148.102	7	1.04528	2.48E-02	1.343
Glutathione	Amino acids	8.09843	307.08	8	1.13633	<1.00E-14	1.792
Isocitrate	Carbohydrates	7.79821	192.03	8	1.12911	<1.00E-15	0.884
Calystegin B2	Others	7.79402	182.108	7	1.02093	1.39E-04	4.246
dXMP	Others	8.44035	358.08	7	1.15394	<1.00E-15	1.831
L-Cysteine	Amino acids	9.07892	124.03	10	1.00387	<1.00E-15	1.884

Table 5.4 List of statistically significant metabolites between control and treatment (brain locations grouped) in Patient **TWO**. The positive fold changes represent an upregulated metabolic activity and fold changes less than 1 represent downregulated activity affected by LGG. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Metabolite	Biochemical class	Retention time (mins)	Molecular mass	ID confidence	VIP score	P value	Fold change
D-Tryptophan	Amino acids	8.41097	215.126	6	1.48489	<1.00E-20	1.673
6-diazo-5-oxonorleucine	Others	7.96148	177.084	5	1.14785	<1.00E-18	3.514
Isocitrate	Carbohydrates	7.79821	192.03	8	1.03723	3.75E-04	0.283
Homostachydrine	Others	7.61303	165.137	5	1.05549	<1.00E-18	5.428
D-Fructose 1,6-bisphosphate	Energy metabolism	9.07581	346.015	8	1.04953	2.49E-03	1.474
2-Phospho-D-glycerate	Carbohydrates	8.98775	185.99	8	1.05752	<1.00E-20	0.736
D-Proline	Amino acids	8.52481	115.06	8	1.01748	4.58E-06	1.628
L-Cysteine	Amino acids	9.07892	124.03	10	1.04841	<1.00E-18	2.384

Table 5.5 List of statistically significant metabolites between control and treatment (brain locations grouped) in Patient **THREE**. The positive fold changes represent an upregulated metabolic activity and fold changes less than 1 represent downregulated activity affected by LGG. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Metabolite	Biochemical class	Retention time (mins)	Molecular mass	Confidence ID	VIP score	P value	Fold change
L-Cysteine	Amino acids	9.07892	124.03	10	1.04841	<1.00E-20	1.479
6-diazo-5-oxonorleucine	Others	7.96148	177.084	5	1.38492	<1.00E-20	2.799
D-Fructose 1,6-bisphosphate	Energy metabolism	9.07581	346.015	8	1.57583	<1.00E-20	2.663
D-phenylalanine	Amino acids	7.54839	165.08	8	1.15935	<1.00E-20	2.539
3-Phosphoenolpyruvate	Amino acids	9.19384	167.98	8	1.12637	<1.00E-20	2.654
Glutathione	Amino acids	8.09843	307.08	8	1.99345	<1.00E-20	4.741
D-Tryptophan	Amino acids	8.41097	215.126	6	1.04959	<1.00E-20	2.445

Table 5.6 List of statistically significant metabolites between control and treatment (brain locations grouped) in Patient **FOUR**. The positive fold changes represent an upregulated metabolic activity and fold changes less than 1 represent downregulated activity affected by LGG. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Metabolite	Biochemical class	Retention time (mins)	Molecular mass	Confidence ID	VIP score	P value	Fold change
3-Deoxy-D-manno-octulosonate 8-phosphate	Biosynthesis of secondary metabolism	8.44545	326.061	5	1.16991	<1.00E-14	3.475
D-Tryptophan	Amino acids	8.41097	215.126	6	1.45885	3.57E-08	4.284
Ecgonine	Biosynthesis of secondary metabolism	7.96432	194.135	5	1.27894	<1.00E-15	2.758
O-Succinyl-L-homoserine	Amino acids	7.96313	227.101	8	1.39435	1.36E-04	4.322
L-Cysteine	Amino acids	9.07892	124.03	10	1.99483	<1.00E-20	2.637
Ala-Met-Gln-Gln	Peptides	7.44822	494.265	7	1.02023	1.83E-12	2.849

Table 5.7 List of statistically significant metabolites between control and treatment (brain locations grouped) in Patient **FIVE**. The positive fold changes represent an upregulated metabolic activity and fold changes less than 1 represent downregulated activity affected by LGG. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Second hypothesis (Central versus non-central core brain tissue metabolomics)

In this type of analysis, central core brain tissue metabolome (from all patients) was compared to non-central core metabolome (other brain locations from all patients grouped together), both affected by tumour. From the OPLS-DA scores plot (figure 5.8), there were 207 metabolites identified with a VIP score 1 or above. After univariate analysis was performed on the metabolome data, there were 123 statistically changed metabolites between central and non-central core (with p value less than 0.05 and FDR 5% corrected) – shown in the volcano plot below (figure 5.10). More than half of those metabolites expressed positive fold changes (elevation in metabolite concentration in central core tumour, as opposed to non-central), although there was a large proportion of metabolites expressing negative fold change too.

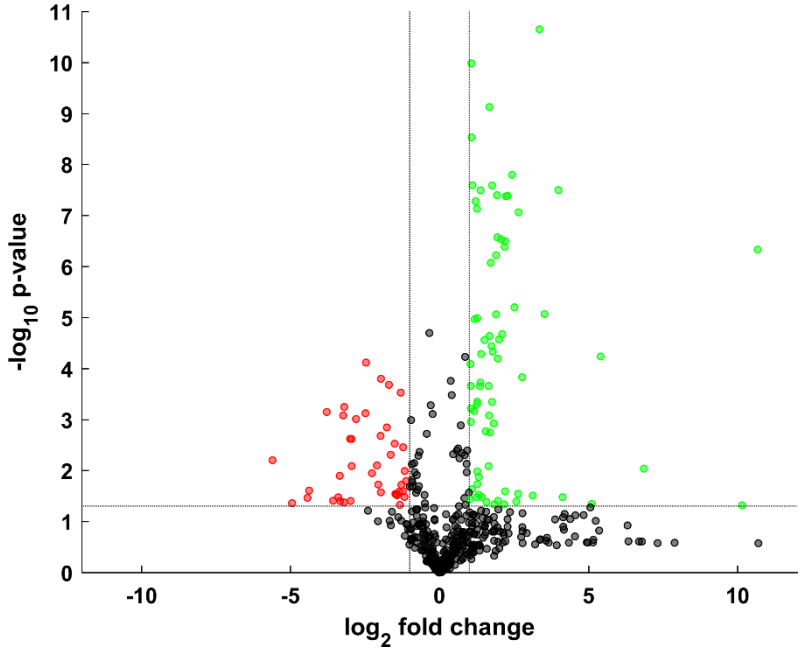


Figure 5.10. Volcano plot showing distribution of statistically significant metabolites detected between central core tumour and non-central core tumour. Green points represent positive fold change in significant metabolites; red points represent negative fold change. Data is analysed by univariate analysis only.

The list of statistically significant metabolites from univariate analysis was compared to the VIP score list, to generate the list of statistically significant metabolites which passed T-test, were FDR corrected and had a VIP score of 1.0 or above. There were 31 significantly changed metabolites (shown in table 5.8). Most of them were found to be elevated in concentration (central tumour over non-central tumour).

Metabolite	Biochemical class	Ret time (mins)	Molecular mass	ID confidence	VIP score	P value	Fold change
1,8-Diazacyclotetradecane-2,9-dione	Others	5.60265	238.208	7	1.42423	1.93E-06	0.478
D-Fructose 1,6-bisphosphate	Energy metabolism	9.07581	346.015	8	1.38493	<1.00E-13	2.578
2-Phospho-D-glycerate	Carbohydrates	8.98775	185.99	8	1.04532	1.88E-05	0.853
D-Tryptophan	Amino acids	8.41097	215.126	6	1.40692	<1.00E-15	2.487
Ecgonine	Biosynthesis of secondary metabolism	7.96432	194.135	5	1.57557	<1.00E-15	1.773
Citrate	Carbohydrates	7.79821	192.03	10	1.04758	<1.00E-20	1.684
Homostachydrine	Others	7.61303	165.137	5	1.00455	<1.00E-14	1.322
L-Rhamnulose	Carbohydrates	8.38324	170.088	6	1.64369	<1.00E-20	3.886
L-Histidine	Amino acids	9.23138	155.07	5	1.75563	2.97E-06	2.342
Carnosine	Amino acids	9.07357	235.136	8	1.56694	<1.00E-20	3.854
L-Cysteine	Amino acids	9.07892	124.03	10	1.73576	<1.00E-20	2.847
6-diazo-5-oxonorleucine	Others	7.96148	177.084	5	1.48593	<1.00E-20	1.492
3-Phosphoenolpyruvate	Amino acids	9.19384	167.98	8	1.47583	<1.00E-14	2.426
Glutathione	Amino acids	8.09843	307.08	8	1.22942	<1.00E-20	2.445
D-Proline	Amino acids	8.52481	115.06	8	1.45583	<1.00E-12	1.277
D-phenylalanine	Amino acids	7.54839	165.08	8	1.45635	<1.00E-20	2.631
Ala-Met-Gln-Gln	Peptides	7.44822	494.265	7	1.38485	8.76E-14	0.573
D-Apiitol	Others	8.05021	141.09	7	1.28531	9.19E-05	3.573
DL-2-Aminooctanoicacid	Others	8.38775	167.152	5	1.39996	5.97E-06	1.236
gamma-L-Glutamyl-L-cysteinyl-beta-alanine	Others	7.60821	332.135	7	1.40984	<1.00E-20	0.037
D-Serine	Amino acids	9.30846	105.042	9	1.64832	<1.00E-20	1.573
O-Acetyl-L-homoserine	Amino acids	7.76138	167.089	8	1.36243	7.65E-07	0.479
O-Succinyl-L-homoserine	Amino acids	7.96313	227.101	8	1.26072	<1.00E-20	2.403
D-methionine	Amino acids	8.04627	149.05	8	1.37483	<1.00E-20	4.332
Hypotaurine	Amino acids	9.13925	109.02	8	1.64953	5.24E-03	2.853

2-Furoate	Others	9.37829	112.02	5	1.48495	3.32E-05	0.858
Tremetone	Others	8.92748	202.10	5	1.43874	<1.00E-20	0.048
5'-Phosphoribosylglycinamide	Nucleotide	8.61633	286.06	8	1.76843	<1.00E-20	2.644
1-Methyladenosine	Others	10.17484	281.11	7	1.33924	1.73E-02	3.284
Phenol sulphate	Others	4.83522	174.84	7	1.41573	<1.00E-20	3.274
D-Xylulose 5-phosphate	Carbohydrates	8.55547	230.02	6	1.82934	3.46E-03	4.284

Table 5.8 List of statistically significant metabolites between central core tumour and non-central core tumour among all patients. The positive fold changes represent an upregulated metabolic activity (increased concentration) of metabolites in central core over non-central core tumour, while fold changes less than 1 represent downregulated activity. ID confidence: metabolite identification level according to the putative identification confidence levels by IDEOM (1-10). Red highlighted metabolites were found to be related to metabolic pathways.

Metabolic pathway analysis

The list of potential biomarkers (key statistically significant metabolites) have been passed through metabolic pathway enrichment analysis (MetaboAnalyst) to observe any important metabolic pathways affected during intervention (glioma effect). The results below show the affected pathways with calculated fold enrichment based on metabolites' fold changes. Figure 5.11 shows the map of the main affected pathways and related metabolites' fold enrichment (based on metabolite concentrations/fold changes).



Figure 5.11. Metabolic pathway enrichment analysis overall results showing which pathways were mostly affected by the intervention (LGG disease). Data taken from MetaboAnalyst data processing.

5.4 Discussion

In this clinical application, an investigation was performed in metabolic profiling of human brain tissue affected by a single LGG tumour (astrocytoma grade 2), answering hypothesis questions on intra-tumour heterogeneity (ITH), difference between central and non-central core metabolome activity, and trying to find significant key metabolites as potential biomarkers, which can strongly affect metabolic pathways developing LGG pathogenesis. In the past, metabolic heterogeneity has been investigated across several brain tumours[433], but not in the context of a single brain tumour. To date, the ITH concept has been used in other types of organ cancers, such as lung[434] and kidney[435]. Multi-region sampling in five LGG patients has been performed in this study to detect metabolic ITH.

Quality of metabolome data

A large number of putative metabolites, with a wide range of biochemical classes, were detected in human brain tissue. As was expected, the majority of metabolites belonged to a polar class, due to polar extraction of tissues in hydrophilic conditions. Several untargeted and targeted metabolomics studies have previously been performed on mammalian brain tissue, for a variety of reasons, showing a wide range of detected polar metabolites, high quality spectra data and good analytical performance throughout the experiments (good QC results).[68,436,437] However, these studies showed a wide biological variation in their samples. Most of them did not use any normalisation or advanced normalised method, for example ^{13}C isotopes, to minimise the vast data variation that developed during sample extraction, preparation and analytical analysis of brain tissue samples. Therefore, an advanced normalisation method proposed earlier in this thesis has been applied to brain tissue samples in this study to help improve data accuracy and to interpret complicated metabolic queries relating to the ITH of brain tissue and metabolic pathways analysis.

In this study, the main results of normalised metabolomics data of LGG-affected human brain tissue have shown a clear difference between control (epilepsy-affected human brain tissue) and treatment groups (brain tissues taken from various anatomical resections in each individual), through OPLS-DA models – shown in figure 5.7. In those models, a clear separation between the groups with high Q2 scores (greater than 0.5) and clustered data points (reduced variation), have demonstrated that there might be a presence of ITH in metabolic activity between brain tissues (taken by various resections) in each individual. Further statistical analysis using the univariate approach helped to detect a range of key statistically significant metabolites

(shown in tables 5.3 to 5.7) confirming the presence of ITH. Among those key metabolites, there were some which were found to be linked to major metabolic pathways, such as the TCA cycle, glycolysis and MAPK (mentioned in the introduction). Examples of common, significantly different metabolites found among patients ONE to FIVE in the experimental data, were D-fructose-1,6-bisphosphate, 2-phospho-D-glycerate, isocitrate, D-tryptophan, L-cysteine and glutathione. All of these metabolites are hugely involved in physiological metabolic pathways, closely related to cell proliferation and growth, so a change in their metabolic activity influenced by genetic mutations can massively affect cell proliferation, and hence cause cancer.[438]

Statistically significant metabolites and metabolic pathways

D-fructose-1,6-bisphosphate and 2-phospho-D-glycerate metabolites both play an essential role in the glycolysis and glutamine-glutamate pathways which are important for cell growth, converting glucose into the metabolite pyruvate and releasing energy in the form of ATP, which is used in further important metabolic reactions.[439] Pyruvate then enters pyruvate metabolism, generating acetyl-CoA, which then activates the TCA cycle metabolic reactions.[440] Therefore, an inhibitory or elevated effect of D-fructose-1,6-bisphosphate and 2-phospho-D-glycerate can have huge consequences on glycolysis and related pathways, significantly affecting cell proliferation. From the experimental data in this study, the levels of both metabolites were elevated (increased concentration levels) in LGG brain tumour tissue compared to the control. For D-fructose-1,6-bisphosphate, the metabolic activity was increased by a fold change of 1.4 (patient ONE), 2.8 (patient TWO), 1.5 (patient THREE) and 2.6 (patient FOUR) – patient FIVE tumour samples did not display the significance of this metabolite. For 2-phospho-D-glycerate, the metabolic activity was also increased, but the fold change was smaller – 0.624 (patient ONE), 0.573 (patient TWO), 0.736 (patient THREE), with the tumour samples of patients FOUR and FIVE not displaying any significant reaction to this metabolite. The elevation of these metabolites has already been observed in metabolomics studies in the past. For example, in one study with GMB brain samples, an increased rate of phosphorylation of fructose-6-phosphate into fructose-1,6-bisphosphate was observed in tumour samples, showing highly elevated concentrations of fructose-1,6-biphosphate in cancer cells.[441] In another study, the metabolite levels of 2-phospho-glycerate was observed to be elevated in human astrocytoma LGG tissues. That study showed overexpressed activity of the enzyme phosphoglycerate mutase 1 that normally catalyses 3-phosphoglycerate into 2-phosphoglycerate. With its high activity, the production of 2-phosphoglycerate was very rapid in the tumour cells.[442] In epilepsy, these metabolites have never been shown to be affected in their metabolic activity and have never been investigated as potential biomarkers in epilepsy, even though the rate of glycolysis was

found to increase during seizures.[443] Therefore, the control sample did not influence the metabolic levels of D-fructose-1,6-bisphosphate and 2-phospho-D-glycerate in any way that could affect interpretation of the data. Figure 5.11 shows how the two major metabolic pathways (glutathione, glutamate have highest amount of key metabolites fold change).

Another statistically significant metabolite observed between control and treatment samples in this study was isocitrate, the key metabolite in the TCA cycle pathway. In TCA metabolism, isocitrate is normally reduced to the metabolite alpha-ketoglutarate by IDH-1 and IDH-2 enzymes (see figure 5.3). In LGG there is a high chance of IDH gene mutations, mentioned earlier, which lead to deactivation of these enzymes, metabolising isocitrate into 2-hydroxyglutarate, interrupting and downregulating the TCA cycle. Therefore, the levels of isocitrate may be reduced. In this study, it can be seen that metabolic fold changes between treatment and control for isocitrate were less than 1 – in patient ONE (0.9) and in patient TWO (0.8), showing the decrease in metabolic activity of isocitrate in LGG tumour tissue. The downregulation of isocitrate, citrate and TCA cycle related metabolites has been observed in glioma studies before.[444,445] Figure 5.11 shows that in TCA cycle pathway there were low number of fold changes in key metabolites. Unfortunately, the expected elevation in metabolic activity of 2-hydroxyglutarate or decrease in alpha-ketoglutarate were not observed in this study, as the metabolites were not even detected in this metabolic LC-MS analysis.

D-Tryptophan has also been observed as a statistically significant metabolite between control and LGG tumour, in almost all patients in this study. Its metabolic activity has been elevated in LGG tissue, with increased fold changes in patient ONE (by 2.5), patient THREE (by 1.6), patient FOUR (by 2.4) and patient FIVE (by 4.2). Generally, tryptophan is found to activate oncogenes and shield glioma cells from detection by immune cells. Several metabolic studies in glioma have demonstrated an increase in tryptophan metabolite levels, showing it to play an important role in pathogenicity, as it is linked to the TCA cycle – tryptophan activates the metabolite acetyl CoA, increasing the activity of TCA metabolism and leading to tumour development.[446-448] Tryptophan was also found to aid glioma resistance to oxidative stress (radicals that may destroy tumour cells).[449]

L-cysteine and glutathione also play an important pathogenic role in gliomas. In this study, their metabolic activity has been increased in LGG tumour brain tissue, compared to the control – fold changes were elevated in a few patients. This observation was expected, because in gliomas, the cellular uptake of cysteine is generally increased. As cysteine is a precursor of glutathione biosynthesis, glutathione levels would also be expected to increase in glioma

cells.[450] Increased metabolic levels of glutathione in glioma cells has a huge effect on the TCA cycle, as well as the MAPK pathway. Elevated levels of glutathione were shown to inhibit the MAPK pathway, preventing antioxidant protection of healthy brain cells.[451]

Central versus non-central core metabolism

Metabolism of central and non-central core tissues was compared to each other in order to understand the effect of LGG tumour development throughout the brain tissues and the role of residual cells in LGG pathogenesis. OPLS-DA model (figure 5.8) showed a clear separation between the central core and non-central core groups, demonstrating metabolic differences in tumour activity between those regions. Further statistical analysis helped to reveal some key statistically significant metabolites (table 5.8), some of which were in common with key metabolites detected for individual patient datasets (tables 5.3 to 5.7). It appears that in central core LGG brain tissues, some metabolites were involved in important metabolic pathways such as TCA cycle and glycolysis, and had higher metabolic activity than in non-central core tissues. This supports the theory that metabolism in central core brain tissues is hugely affected during cancer development, especially in the early stages of the disease. The tumour then progresses and affects other brain tissues surrounding central core (non-central core tissues), further spreading the disease throughout the brain. Presence of tumour activity in non-central core may lead to the development of residual cells, causing recurrent tumour. Therefore, further statistical analysis could be performed on non-central core metabolism data, in order to see how the tumour is progressing through residual cells. However, this is outside of the scope of this chapter.

5.5 Conclusion

In this study, several statistically significant metabolites involved in important metabolic pathways have been identified between LGG brain glioma astrocytoma and control brain tissue, as well as between central core and non-central core brain glioma tissue. It appears that there is a significant difference in metabolic activity between various brain tissue regions for each patient – as shown in the OPLS-DA models (figure 5.7 and 5.8) and the list of statistically significant metabolites. Therefore, according to the acquired data from this study, the first hypothesis can be fully accepted, showing the extent of heterogeneity between patients. For the second hypothesis, stating that there is ITH between patients for each brain location, the data was very hard to interpret, as some patients did not provide their tissue from all of the

brain locations which were tested in this study – for example, patient ONE did not provide LGG tissue from the anterior location, while patient FIVE did not provide tissue from the superficial location. Therefore, it was difficult to interpret this comparison. The reason for not obtaining brain tissues from each location for every single patient, which would make data more consistent, might have been due to risks of affecting healthy tissue during biopsy technique or further damaging brain of a patient – in some situations, it is easy to remove certain brain tissue from particular location, especially if it is found deep inside the brain.

The key statistically significant metabolites detected between the LGG tumour and control tumour were D-fructose-1,6-bisphosphate, 2-phospho-D-glycerate, isocitrate, D-tryptophan, L-cysteine and glutathione. These metabolites can act as potential biomarkers for the LGG disease, as targeting their metabolism or related metabolic pathways may help to suppress LGG tumour cell proliferation.

LC-MS analysis of LGG brain tissue samples produced clear and distinctive results. With the advanced normalisation technique involved in this analysis, the highly accurate acquired data helped to detect important metabolite changes in metabolic pathways and detect potential biomarkers. This information can help researchers to understand more about low grade glioma cell proliferation and how can it develop into high grade glioma by monitoring changes in metabolic activity. Some of the key metabolites detected in this study may also be of use as diagnostic parameters, rather than fully relying on MRI scans and other types of imaging. For example, tryptophan, cysteine and alpha-ketoglutarate have already been used as diagnostic parameters of gliomas.[452]

Now that ITH in LGG brain tissue samples between individuals and brain regions has been observed, it would be interesting to see the same observation in high grade gliomas, such as GBM. For this case study, GBM would be a perfect comparison, as grade 2 astrocytoma LGG normally develops into GBM as the disease progresses over the time. The data from both LGG astrocytoma and GBM can provide better understanding on potential biomarkers, metabolic pathways affected during those conditions and designing effective treatment.

Chapter Six

Conclusions and Future Work

6.1 Method development and validation

In this thesis, a proposed LC-MS method with an advanced normalisation technique involving (U)-^{13}C labelled *E. coli* bacterium as a source of multiple internal standards (^{13}C labelled metabolites) has been applied in variety of clinical untargeted studies. The main aim of this project was to demonstrate the effectiveness of this normalisation technique in improving accuracy and quality of metabolome data for a wide range of polar and non-polar metabolites in the mammalian samples, which were involved in untargeted metabolite profiling experiments. Great attention to detail was applied in the development of the method, validation using a small number of biological samples from human and animal biofluids and tissues, and application of the method to a range of clinical mammalian samples with different diseases to obtain highly accurate data on potential biomarkers involved in specific metabolic pathways. As mentioned earlier, currently there is very limited research performed in ^{13}C normalisation of mammalian metabolomes in untargeted studies, possibly due to a lack of resources (availability of commercial ^{13}C labelled internal standards for a wide range of metabolites) and their high cost, availability of clinical samples on demand or other related reasons. Therefore, in this thesis it was decided to address this issue by designing a new method of ^{13}C normalisation using a uniformly labelled microorganism and applying it to a set of untargeted metabolomics studies with mammalian samples. The thesis was split into two main parts: chapters two and three described method development and its validation, while chapters four and five involved the application of the developed method in clinical studies.

Chapter two has looked at the method development. The results have demonstrated that the *E. coli* bacterial organism was a suitable source of internal standards because of its large detected metabolome (over 700 putative metabolites on average), the majority of which (80%) were found to be in common with the metabolome of human biofluids and tissues, based on information from HMDB and ECDB online databases.[53,214] One of the challenging steps in method development was finding the most suitable and optimised method of efficient extraction for *E. coli*. A number of already existing metabolite extraction techniques have been used in the past, applying different methods to obtain a good quality extract, each having their own advantages and limitations.[190,269,285,291] The experimental data in chapter two concluded that for *E. coli,* using ice-cold methanol extraction solvent followed by freeze-thaw cycles and centrifugation generated a very good quality of extract with a larger number of detected metabolites, compared to other extraction methods such as drying extraction or using another solvent (acetonitrile). In one study, using methanol and freeze-thaw cycle extraction has also

shown the best extraction results for *E. coli*.[269] In fact, some of the techniques such as drying extraction involving speed vacuum centrifugation can be associated with possible metabolite degradation during the drying stage, as shown in another extraction study.[453] *E. coli* and *Spirulina* were chosen for comparative analysis, as both microorganisms are extensively involved in clinical and pharmaceutical areas of research.[454-456] *E. coli* metabolome results, on average, have demonstrated a clear advantage over *Spirulina* extract in terms of number of detected metabolomes, easy extraction method, cost and good metabolite peak signal response.

When it came to labelling the bacterial organisms with ^{13}C isotopes, the main aim was to generate an extract with a wide range of ^{13}C labelled metabolites. The chosen (U)-^{13}C-labelled *E. coli*, on average, has shown to have a large proportion of metabolites labelled with ^{13}C isotope (80%), with the remaining 20% of the metabolome being unlabelled, possibly due to introduction of carbon dioxide during bacterial growth or the contribution of ^{12}C isotope from LB media. This was perhaps the most challenging task, to generate a fully labelled *E. coli* extract without any presence of ^{12}C isotope. It was for this reason that the M9 minimal medium was used during the bacterial growth procedure, as it has a minimal amount of ^{12}C isotopes and enough ingredients to generate a good bacterial growth. However, because *E. coli* needed an LB medium as well, for an initial "boost" of growth and to help it reach a higher OD, this might have introduced a certain amount of ^{12}C isotopes. Several studies in the past have been growing *E. coli* extract in a ^{13}C labelled environment for a variety of purposes – mainly for tracking activity and function of microorganisms in a biological system, measuring metabolite fluxes or for monitoring the metabolic activity of a peptidoglycan polymer in bacterial cell membranes. These studies have successfully managed to label a large number of metabolites with ^{13}C isotope, showing high isotopic purity of ^{13}C, but there was also a presence of ^{12}C isotopes. This showed that it was almost impossible to have an entire metabolome labelled successfully and therefore demonstrated that the technique has its limitations.[189,457,458]

The next step was to validate the proposed normalisation method using appropriate biological samples of interest. (U)-^{13}C labelled *E. coli* was applied in human urine, human brain tissue and mouse plasma samples, all of which did not contain any traces of serious clinical disease, apart from the brain tissue which was taken from volunteers suffering from seizures. In all three studies, a wide range of metabolites have shown a significant reduction in metabolite data analytical variation through percentage RSD calculations – in human urine and brain studies (approximately 70% had their RSD reduced); in the mouse plasma study (90% of metabolites) – in contrast to non-normalised metabolomes. Multivariate analysis (PCA and OPLS-DA score

plots) also showed effective normalisation results by demonstrating clear separation between data groups with sample replicates clustered and high recorded Q2 scores. Finally, the normalisation technique has helped to detect more statistically significant metabolites in all three studies, compared to non-normalised datasets. This method has been shown to pass a validation stage, reducing the metabolite variation overall and improving the accuracy of data by detecting a high number of statistically significant metabolites. As mentioned earlier in the introduction chapter and chapter three, the ^{13}C normalisation technique has already been applied in a variety of metabolomics studies (mainly targeted analysis) to assist in an accurate identification and quantification of intracellular metabolites, with limited research in untargeted studies.[190,195,206] Despite the use of ^{13}C labelled internal standards in normalising metabolite data, research scientists still face issues such as matrix effects and ion suppression problems in metabolome data, even after normalisation. Some studies have demonstrated that internal standards had an effect on analyte structure and concentration, and flow rate in chromatography, all of which could cause ion suppression and lead to development of matrix effects. However, the biggest limitation which was found even with an improved ^{13}C normalisation technique, was the issue around the "co-elution" principle – where an exact internal standard could not be found for a particular metabolite, so a different internal standard with similar chemical properties had to be used instead. Some studies did not agree with this principle, stating retention time does not fully reflect all matrix and chemical properties of a metabolite, and it might not be very accurate to use one internal standard on several metabolites which happen to co-elute with the same retention time.[194] Regardless of this, the main advantages of using ^{13}C internal standards in normalising metabolome data outweigh these limitations. The experiments in chapter three have confirmed that the technique successfully reduced metabolome data variation among analytical replicates and improved the accuracy of data. However, such a great difference in %RSD was not observed between ^{13}C normalisation and other normalisation methods. Other normalisation methods such as MSTUS, TIC and Creatinine demonstrated similar degree of reduction in %RSD for the majority of metabolites in all three case studies compared to ^{13}C normalisation. ^{13}C normalisation still demonstrated the best results and lower %RSD variance but not as effectively as was expected in comparison to other techniques. ^{13}C normalisation is still a relatively new method, and therefore the labelling efficiency of all metabolites covering all carbon atoms with ^{13}C in the structure is proving to be a tricky step, affecting the efficiency of normalisation overall, hence %RSD variance reduction is not huge in contrast to other normalisation techniques.

6.2 The clinical applications of the developed LC-MS normalisation method

In chapters four and five, a fully validated LC-MS method with a ^{13}C normalisation technique has been applied in two clinical investigation studies, in order to obtain highly accurate information on potential biomarkers and their effect on metabolic pathways linked with disease progression. In both clinical studies, ^{13}C normalised metabolome data has helped to detect a number of statistically significant metabolites which were observed to affect certain metabolic pathways involved in the investigated diseases, showing potential to be biomarkers. Some of these statistically significant metabolites and their associated metabolic pathways were linked to previous studies, confirming their importance in the pathogenesis of a disease. For example, in chapter five, where low grade glioma pathogenesis was investigated in human brain tissue, a set of potential biomarkers including metabolites D-fructose-1,6-bisphosphate, 2-phospho-D-glycerate were observed. They were shown to be elevated in the glycolysis pathway, which was found to be important in glioma progression, confirming their significance. In chapter four, important, statistically significant metabolites such as L-erythrulose, cis-aconitate and allantoin were shown to be elevated in urine and plasma samples of individuals treated with a high fatty acid meal. Therefore, it can be concluded that an improved normalisation technique has indeed played an important role in minimising data variation and helping to obtain highly accurate metabolome data in clinical samples, showing promise for potential use in future large clinical studies investigating potential biomarkers.

However, one of the clinical studies (in chapter four) did not contain actual disease samples – urine and plasma were taken from individuals who took a high fatty acid meal, but were otherwise healthy and did not have diabetes. This meal might have triggered certain metabolites and metabolic pathways associated with diabetic response (increase in blood glucose levels), as shown in the final results, but these statistically significant metabolites and their associated metabolic pathways cannot be fully interpreted in the context of the diabetes disorder. This was one of the limitations found in this project. To fully confirm the usefulness of a validated normalisation method in clinical samples, actual disease samples have to be used. In chapter five, the samples were taken from individuals with a developed disorder (LGG brain tumour), so the clinical results (a list of potential biomarkers and linked metabolic pathways) were shown to be appropriate for the disease and could be used further in the clinical interpretation of glioma. However, the difficulty in obtaining a consistent number of brain tissue samples from each brain region for all volunteers, may have complicated the interpretation of

the results and therefore could not confirm the second hypothesis – intra-tumour heterogeneity between individuals for each brain location. Also, the control sample in this study was chosen to be brain tissue from a patient with epilepsy, rather than obtaining healthy brain tissue from an individual with low grade glioma. As it is almost impossible to obtain healthy tissue from glioma patients to use as a control, because most of the brain tissue in each individual was affected by the disease, the healthy tissue had to be taken from an individual with a similar brain physiology closely related to the symptoms of glioma (seizures from an epileptic patient). These factors were found to be the limitations in this study.

Generally, for the clinical studies to be more viable, the number of volunteers and samples needs to be increased. In both clinical studies, there were only five or six volunteers who provided their samples. Increasing the number of participants in such studies can provide more metabolite and biomarker information for a particular disease. Due to time and fund constraints, most of the reported potential biomarkers were only tentatively identified. They could be better defined by using targeted approaches with LC-MS/MS instrumentation, as well as using pure authentic standards for statistically significant metabolites, to fully confirm their identity and increase the confidence levels of those biomarkers.

Despite the points outlined above, the clinical studies have clearly demonstrated the differences in metabolic activity between control and disease samples.

6.3 Future Work

Now that it has been established that ^{13}C normalisation can indeed help to achieve useful results for untargeted mammalian metabolomics studies, additional work which is beyond the scope of this thesis could be beneficial to observe what could be done to further improve the normalisation method in LC-MS metabolite profiling experiments of clinical samples.

One of the future work goals should focus on how to avoid unsuccessful normalisation of metabolites after normalisation of metabolome data. From chapter three, it can be seen that metabolites which were unsuccessfully normalised were not paired with their corresponding ^{13}C labelled peaks. This could have happened because the unpaired metabolites simply were not detected in the *E. coli* extract, therefore would not have a paired ^{13}C labelled internal standard in the first place. One way to improve this issue would possibly be to introduce commercially available ^{13}C labelled internal standards for those metabolites which did not have a paired ^{13}C

internal standard. The commercially purchased internal standards would be introduced to the samples of interest, which would be analysed again, followed by the normalisation data processing only for any remaining unsuccessfully normalised metabolites. Even though it is expensive to purchase individual labelled internal standards, it can guarantee that all of the metabolites in the samples would be normalised effectively.

Another improvement to the technique would be to apply the normalisation method in a wider range of mammalian biofluids and tissues. Biofluid samples such as saliva, cerebrospinal fluid and body sweat also contain important metabolites which are significantly changed by diseases, so are worth applying a normalisation method to. Application to a variety of tissues would also be beneficial – liver tissue, heart and kidney, which are constantly affected by a variety of serious diseases. It would be interesting to observe to what degree data variation can be reduced in these types of samples.

Once the proposed normalisation LC-MS method can be applied in a variety of mammalian biofluid and tissue samples in metabolomics, the method can then be tested in other omics studies, including lipidomics and proteomics, to see if normalisation can work for different classes of biochemical compounds. At this stage, the method would have to be optimised again to fit certain requirements of lipid profiling or proteomics. For example, ^{13}C labelling of wide range of lipids, using the identical uniform labelling technique but using slightly different chromatographic and solvent parameters appropriate to lipidomics, can be useful to label hydrophobic/lipophilic metabolites in mammalian samples to fully cover the range of metabolites and lipids for full normalisation effect.

In clinical studies, to fully validate the potential biomarkers, it would be wise to use the targeted metabolomics approach with LC-MS/MS instrumentation for those clinical samples. That would give more qualitative and quantitative information about the reported biomarkers.

In summary, the contents of this thesis demonstrated that in all of the validation and clinical studies, the proposed LC-MS method using ^{13}C normalisation technique has shown a clear advantage in obtaining highly accurate metabolome data by demonstrating a significant reduction in metabolite data variation in biological samples, through percentage RSD calculations, multivariate and univariate statistical analysis. These results have demonstrated a potential use of this method towards normalising a wide range of metabolites in untargeted metabolomics studies using mammalian samples. This can be of benefit by providing further understanding of the biological processes related to the study in question.

Chapter Seven

References

(1) Mathew, A.; Padmanaban, V. Metabolomics: the apogee of the omics trilogy. *Int J Pharm Pharm Sci* **2013**, *5*, 45-48.

(2) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. Systematic functional analysis of the yeast genome. *Trends in biotechnology* **1998**, *16*, 373-378.

(3) Zhou, B.; Xiao, J. F.; Tuli, L.; Ressom, H. W. LC-MS-based metabolomics. *Mol Biosyst* **2012**, *8*, 470-481.

(4) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. *Mass Spectrom Rev* **2007**, *26*, 51-78.

(5) Gowda, G. N.; Zhang, S.; Gu, H.; Asiago, V.; Shanaiah, N.; Raftery, D. Metabolomics-based methods for early disease diagnostics. *Expert review of molecular diagnostics* **2008**, *8*, 617-633.

(6) Gutteck, U.; Rentsch, K. M. Therapeutic drug monitoring of 13 antidepressant and five neuroleptic drugs in serum with liquid chromatography-electrospray ionization mass spectrometry. *Clinical chemistry and laboratory medicine* **2003**, *41*, 1571-1579.

(7) Robertson, D. G.; Watkins, P. B.; Reily, M. D. Metabolomics in toxicology: preclinical and clinical applications. *Toxicological Sciences* **2010**, *120*, S146-S170.

(8) Bundy, J. G.; Davey, M. P.; Viant, M. R. Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **2009**, *5*, 3.

(9) Brown, D. G.; Rao, S.; Weir, T. L.; O'Malia, J.; Bazan, M.; Brown, R. J.; Ryan, E. P. Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. *Cancer & metabolism* **2016**, *4*, 11.

(10) Khamis, M. M.; Adamko, D. J.; El-Aneed, A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrom Rev* **2017**, *36*, 115-134.

(11) Madsen, R.; Lundstedt, T.; Trygg, J. Chemometrics in metabolomics—a review in human disease diagnosis. *Analytica chimica acta* **2010**, *659*, 23-33.

(12) Patti, G. J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nature reviews Molecular cell biology* **2012**, *13*, 263.

(13) Chen, R.; Mias, G. I.; Li-Pook-Than, J.; Jiang, L.; Lam, H. Y.; Chen, R.; Miriami, E.; Karczewski, K. J.; Hariharan, M.; Dewey, F. E. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* **2012**, *148*, 1293-1307.

(14) Gehlenborg, N.; O'donoghue, S. I.; Baliga, N. S.; Goesmann, A.; Hibbs, M. A.; Kitano, H.; Kohlbacher, O.; Neuweger, H.; Schneider, R.; Tenenbaum, D. Visualization of omics data for systems biology. *Nature methods* **2010**, *7*, S56.

(15) Klenø, T.; Kiehr, B.; Baunsgaard, D.; Sidelmann, U. Combination of 'omics' data to investigate the mechanism (s) of hydrazine-induced hepatotoxicity in rats and to identify potential biomarkers. *Biomarkers* **2004**, *9*, 116-138.

(16) Baas, T.; Baskin, C.; Diamond, D. L.; Garcia-Sastre, A.; Bielefeldt-Ohmann, H.; Tumpey, T.; Thomas, M.; Carter, V.; Teal, T.; Van Hoeven, N. Integrated molecular signature of disease: analysis of influenza virus-infected macaques through functional genomics and proteomics. *Journal of virology* **2006**, *80*, 10813-10828.

(17) Luan, H.; Meng, N.; Liu, P.; Fu, J.; Chen, X.; Rao, W.; Jiang, H.; Xu, X.; Cai, Z.; Wang, J. Non-targeted metabolomics and lipidomics LC–MS data from maternal plasma of 180 healthy pregnant women. *GigaScience* **2015**, *4*, 16.

(18) Nikolskiy, I.; Mahieu, N. G.; Chen, Y.-J.; Tautenhahn, R.; Patti, G. J. An untargeted metabolomic workflow to improve structural characterization of metabolites. *Analytical chemistry* **2013**, *85*, 7713-7719.

(19) Zhang, W.; Li, F.; Nie, L. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology* **2010**, *156*, 287-301.

(20) Zhang, A.; Sun, H.; Yan, G.; Wang, P.; Wang, X. Mass spectrometry-based metabolomics: applications to biomarker and metabolic pathway research. *Biomedical Chromatography* **2016**, *30*, 7-12.

(21) Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nature reviews molecular cell biology* **2004**, *5*, 763.

(22) Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature reviews Drug discovery* **2016**, *15*, 473.

(23) Saghatelian, A.; Trauger, S. A.; Want, E. J.; Hawkins, E. G.; Siuzdak, G.; Cravatt, B. F. Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry* **2004**, *43*, 14332-14339.

(24) Cheng, S.; Rhee, E. P.; Larson, M. G.; Lewis, G. D.; McCabe, E. L.; Shen, D.; Palma, M. J.; Roberts, L. D.; Dejam, A.; Souza, A. L. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* **2012**, CIRCULATIONAHA. 111.067827.

(25) Sreekumar, A.; Poisson, L. M.; Rajendiran, T. M.; Khan, A. P.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R. J.; Li, Y. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **2009**, *457*, 910.

(26) De Filippis, F.; Pellegrini, N.; Vannini, L.; Jeffery, I. B.; La Storia, A.; Laghi, L.; Serrazanetti, D. I.; Di Cagno, R.; Ferrocino, I.; Lazzi, C. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. *Gut* **2015**, gutjnl-2015-309957.

(27) Lin, C. Y.; Viant, M. R.; Tjeerdema, R. S. Metabolomics: methodologies and applications in the environmental sciences. *Journal of Pesticide Science* **2006**, *31*, 245-251.

(28) Cappello, T.; Mauceri, A.; Corsaro, C.; Maisano, M.; Parrino, V.; Paro, G. L.; Messina, G.; Fasulo, S. Impact of environmental pollution on caged mussels Mytilus galloprovincialis using NMR-based metabolomics. *Marine pollution bulletin* **2013**, *77*, 132-139.

(29) Kaddurah-Daouk, R.; Weinshilboum, R.; Network, P. R. Pharmacometabolomics: implications for clinical pharmacology and systems pharmacology. *Clinical Pharmacology & Therapeutics* **2014**, *95*, 154-167.

(30) Wikoff, W. R.; Frye, R. F.; Zhu, H.; Gong, Y.; Boyle, S.; Churchill, E.; Cooper-Dehoff, R. M.; Beitelshees, A. L.; Chapman, A. B.; Fiehn, O. Pharmacometabolomics reveals racial differences in response to atenolol treatment. *PloS one* **2013**, *8*, e57639.

(31) Dunn, W. B.; Ellis, D. I. Metabolomics: current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry* **2005**, *24*, 285-294.

(32) Weckwerth, W. Metabolomics in systems biology. *Annual review of plant biology* **2003**, *54*, 669-689.

(33) Asiago, V. M.; Alvarado, L. Z.; Shanaiah, N.; Gowda, G. N.; Owusu-Sarfo, K.; Ballas, R. A.; Raftery, D. Early detection of recurrent breast cancer using metabolite profiling. *Cancer research* **2010**, 0008-5472. CAN-0010-1319.

(34) Bonnet, J.-L.; Bonnemoy, F.; Dusser, M.; Bohatier, J. Toxicity assessment of the herbicides sulcotrione and mesotrione toward two reference environmental microorganisms: Tetrahymena pyriformis and Vibrio fischeri. *Archives of environmental contamination and toxicology* **2008**, *55*, 576-583.

(35) Chiang, Y.-M.; Chuang, D.-Y.; Wang, S.-Y.; Kuo, Y.-H.; Tsai, P.-W.; Shyur, L.-F. Metabolite profiling and chemopreventive bioactivity of plant extracts from Bidens pilosa. *Journal of Ethnopharmacology* **2004**, *95*, 409-419.

(36) Alonso, A.; Marsal, S.; Julia, A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front Bioeng Biotechnol* **2015**, *3*, 23.

(37) Lam, C.-W.; Law, C.-Y. Untargeted mass spectrometry-based metabolomic profiling of pleural effusions: fatty acids as novel cancer biomarkers for malignant pleural effusions. *Journal of proteome research* **2014**, *13*, 4040-4046.

(38) Sawada, Y.; Akiyama, K.; Sakata, A.; Kuwahara, A.; Otsuki, H.; Sakurai, T.; Saito, K.; Hirai, M. Y. Widely targeted metabolomics based on large-scale MS/MS data for elucidating metabolite accumulation patterns in plants. *Plant and Cell Physiology* **2008**, *50*, 37-47.

(39) Ruhaak, L. R.; Koeleman, C. A.; Uh, H.-W.; Stam, J. C.; van Heemst, D.; Maier, A. B.; Houwing-Duistermaat, J. J.; Hensbergen, P. J.; Slagboom, P. E.; Deelder, A. M. Targeted biomarker discovery by high throughput glycosylation profiling of human plasma alpha1-antitrypsin and immunoglobulin A. *PloS one* **2013**, *8*, e73082.

(40) Chi, J.; Jayewardene, A. L.; Stone, J. A.; Motoya, T.; Aweeka, F. T. Simultaneous determination of five HIV protease inhibitors nelfinavir, indinavir, ritonavir, saquinavir and amprenavir in human plasma by LC/MS/MS. *Journal of pharmaceutical and biomedical analysis* **2002**, *30*, 675-684.

(41) Inoue, K.; Tsuchiya, H.; Takayama, T.; Akatsu, H.; Hashizume, Y.; Yamamoto, T.; Matsukawa, N.; Toyo'oka, T. Blood-based diagnosis of Alzheimer's disease using fingerprinting metabolomics based on hydrophilic interaction liquid chromatography with mass spectrometry and multivariate statistical analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* **2015**, *974*, 24-34.

(42) Michell, A. W.; Mosedale, D.; Grainger, D. J.; Barker, R. A. Metabolomic analysis of urine and serum in Parkinson's disease. *Metabolomics* **2008**, *4*, 191.

(43) van Wietmarschen, H. A.; Dai, W.; van der Kooij, A. J.; Reijmers, T. H.; Schroën, Y.; Wang, M.; Xu, Z.; Wang, X.; Kong, H.; Xu, G. Characterization of rheumatoid arthritis subtypes using symptom profiles, clinical chemistry and metabolomics measurements. *PloS one* **2012**, *7*, e44331.

(44) Lewis, G. D.; Wei, R.; Liu, E.; Yang, E.; Shi, X.; Martinovic, M.; Farrell, L.; Asnani, A.; Cyrille, M.; Ramanathan, A. Metabolite profiling of blood from individuals undergoing planned myocardial infarction reveals early markers of myocardial injury. *The Journal of clinical investigation* **2008**, *118*, 3503-3512.

(45) Potts, B. C.; Deese, A. J.; Stevens, G. J.; Reily, M. D.; Robertson, D. G.; Theiss, J. NMR of biofluids and pattern recognition: assessing the impact of NMR parameters on the principal component analysis of urine from rat and mouse. *Journal of pharmaceutical and biomedical analysis* **2001**, *26*, 463-476.

(46) Wei, J.; Xie, G.; Zhou, Z.; Shi, P.; Qiu, Y.; Zheng, X.; Chen, T.; Su, M.; Zhao, A.; Jia, W. Salivary metabolite signatures of oral cancer and leukoplakia. *International Journal of Cancer* **2011**, *129*, 2207-2217.

(47) Czech, C.; Berndt, P.; Busch, K.; Schmitz, O.; Wiemer, J.; Most, V.; Hampel, H.; Kastler, J.; Senn, H. Metabolite profiling of Alzheimer's disease cerebrospinal fluid. *PLoS one* **2012**, *7*, e31501.

(48) Calderón-Santiago, M.; Priego-Capote, F.; Jurado-Gámez, B.; de Castro, M. L. Optimization study for metabolomics analysis of human sweat by liquid chromatography–tandem mass spectrometry in high resolution mode. *Journal of Chromatography A* **2014**, *1333*, 70-78.

(49) Rastrelli, F.; Jha, S.; Mancin, F. Seeing through Macromolecules: T 2-Filtered NMR for the Purity Assay of Functionalized Nanosystems and the Screening of Biofluids. *Journal of the American Chemical Society* **2009**, *131*, 14222-14224.

(50) Li, M.; Song, Y.; Cho, N.; Chang, J. M.; Koo, H. R.; Yi, A.; Kim, H.; Park, S.; Moon, W. K. An HR-MAS MR metabolomics study on breast tissues obtained with core needle biopsy. *PloS one* **2011**, *6*, e25563.

(51) Salek, R. M.; Colebrooke, R. E.; Macintosh, R.; Lynch, P. J.; Sweatman, B. C.; Emson, P. C.; Griffin, J. L. A metabolomic study of brain tissues from aged mice with low expression of the vesicular monoamine transporter 2 (VMAT2) gene. *Neurochemical research* **2008**, *33*, 292-300.

(52) Wishart, D. S. Proteomics and the human metabolome project. *Expert review of proteomics* **2007**, *4*, 333-335.

(53) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S. HMDB: the human metabolome database. *Nucleic acids research* **2007**, *35*, D521-D526.

(54) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B. The human serum metabolome. *PloS one* **2011**, *6*, e16957.

(55) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S. HMDB: a knowledgebase for the human metabolome. *Nucleic acids research* **2008**, *37*, D603-D610.

(56) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nature reviews Molecular cell biology* **2016**, *17*, 451.

(57) Velenosi, T. J.; Hennop, A.; Feere, D. A.; Tieu, A.; Kucey, A. S.; Kyriacou, P.; McCuaig, L. E.; Nevison, S. E.; Kerr, M. A.; Urquhart, B. L. Untargeted plasma and tissue metabolomics in rats with chronic kidney disease given AST-120. *Scientific reports* **2016**, *6*, 22526.

(58) Bogdanov, M.; Matson, W. R.; Wang, L.; Matson, T.; Saunders-Pullman, R.; Bressman, S. S.; Beal, M. F. Metabolomic profiling to develop blood biomarkers for Parkinson's disease. *Brain* **2008**, *131*, 389-396.

(59) Fu, H.; Liu, X.; Yu, W.; Zhu, L.; Zheng, D.; Wang, J. Screening of urinary biomarkers in patients with type 2 diabetes mellitus. *Wei sheng yan jiu= Journal of hygiene research* **2013**, *42*, 907-914.

(60) Montuschi, P.; Paris, D.; Melck, D.; Lucidi, V.; Ciabattoni, G.; Raia, V.; Calabrese, C.; Bush, A.; Barnes, P. J.; Motta, A. NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax* **2011**, thoraxjnl-2011-200072.

(61) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. The human urine metabolome. *PLoS One* **2013**, *8*, e73076.

(62) Rasmussen, L. G.; Savorani, F.; Larsen, T. M.; Dragsted, L. O.; Astrup, A.; Engelsen, S. B. Standardization of factors that influence human urine metabolomics. *Metabolomics* **2011**, *7*, 71-83.

(63) Ryan, D.; Robards, K.; Prenzler, P. D.; Kendall, M. Recent and potential developments in the analysis of urine: a review. *Analytica Chimica Acta* **2011**, *684*, 17-29.

(64) Shen, C.; Sun, Z.; Chen, D.; Su, X.; Jiang, J.; Li, G.; Lin, B.; Yan, J. Developing urinary metabolomic signatures as early bladder cancer diagnostic markers. *OMICS* **2015**, *19*, 1-11.

(65) Klepacki, J.; Klawitter, J.; Klawitter, J.; Thurman, J. M.; Christians, U. A high-performance liquid chromatography–tandem mass spectrometry–based targeted metabolomics kidney dysfunction marker panel in human urine. *Clinica Chimica Acta* **2015**, *446*, 43-53.

(66) Calderón-Santiago, M.; Priego-Capote, F.; Turck, N.; Robin, X.; Jurado-Gámez, B.; Sanchez, J. C.; De Castro, M. D. L. Human sweat metabolomics for lung cancer screening. *Analytical and bioanalytical chemistry* **2015**, *407*, 5381-5392.

(67) Gu, H.; Chen, H.; Pan, Z.; Jackson, A. U.; Talaty, N.; Xi, B.; Kissinger, C.; Duda, C.; Mann, D.; Raftery, D. Monitoring diet effects via biofluids and their implications for metabolomics studies. *Analytical Chemistry* **2007**, *79*, 89-97.

(68) Ni, Y.; Su, M.; Lin, J.; Wang, X.; Qiu, Y.; Zhao, A.; Chen, T.; Jia, W. Metabolic profiling reveals disorder of amino acid metabolism in four brain regions from a rat model of chronic unpredictable mild stress. *FEBS letters* **2008**, *582*, 2627-2636.

(69) Suhre, K. Metabolic profiling in diabetes. *J Endocrinol* **2014**, *221*, R75-R85.

(70) Moco, S.; Vervoort, J.; Bino, R. J.; De Vos, R. C.; Bino, R. Metabolomics technologies and metabolite identification. *TrAC Trends in Analytical Chemistry* **2007**, *26*, 855-866.

(71) Lu, W.; Bennett, B. D.; Rabinowitz, J. D. Analytical strategies for LC–MS-based targeted metabolomics. *Journal of Chromatography B* **2008**, *871*, 236-242.

(72) Boroujerdi, A. F.; Jones, S. S.; Bearden, D. W. NMR analysis of metabolic responses to extreme conditions of the temperature-dependent coral pathogen Vibrio coralliilyticus. *Letters in applied microbiology* **2012**, *54*, 209-216.

(73) Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M. Targeted profiling: quantitative analysis of 1H NMR metabolomics data. *Analytical chemistry* **2006**, *78*, 4430-4442.

(74) Ludwig, C.; Ward, D.; Martin, A.; Viant, M.; Ismail, T.; Johnson, P.; Wakelam, M.; Günther, U. Fast targeted multidimensional NMR metabolomics of colorectal cancer. *Magnetic Resonance in Chemistry* **2009**, *47*, S68-S73.

(75) Law, W. S.; Huang, P. Y.; Ong, E. S.; Ong, C. N.; Li, S. F. Y.; Pasikanti, K. K.; Chan, E. C. Y. Metabonomics investigation of human urine after ingestion of green tea with gas chromatography/mass spectrometry, liquid chromatography/mass spectrometry and 1H NMR spectroscopy. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry* **2008**, *22*, 2436-2446.

(76) Seli, E.; Sakkas, D.; Scott, R.; Kwok, S. C.; Rosendahl, S. M.; Burns, D. H. Noninvasive metabolomic profiling of embryo culture media using Raman and near-infrared spectroscopy correlates with reproductive potential of embryos in women undergoing in vitro fertilization. *Fertility and sterility* **2007**, *88*, 1350-1357.

(77) Cherney, D. P.; Ekman, D. R.; Dix, D. J.; Collette, T. W. Raman spectroscopy-based metabolomics for differentiating exposures to triazole fungicides using rat urine. *Analytical chemistry* **2007**, *79*, 7324-7332.

(78) Kim, D.-H.; Jarvis, R. M.; Allwood, J. W.; Batman, G.; Moore, R. E.; Marsden-Edwards, E.; Hampson, L.; Hampson, I. N.; Goodacre, R. Raman chemical mapping reveals site of action of HIV protease inhibitors in HPV16 E6 expressing cervical carcinoma cells. *Analytical and bioanalytical chemistry* **2010**, *398*, 3051-3061.

(79) Villas-Bôas, S. G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. Mass spectrometry in metabolome analysis. *Mass spectrometry reviews* **2005**, *24*, 613-646.

(80) Lei, Z.; Huhman, D. V.; Sumner, L. W. Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry* **2011**, *286*, 25435-25442.

(81) Fernandes, V. C.; Vera, J. L.; Domingues, V. F.; Silva, L. M.; Mateus, N.; Delerue-Matos, C. Mass spectrometry parameters optimization for the 46 multiclass pesticides determination in strawberries with gas chromatography ion-trap tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 2187-2197.

(82) Fanos, V.; Caboni, P.; Corsello, G.; Stronati, M.; Gazzolo, D.; Noto, A.; Lussu, M.; Dessì, A.; Giuffrè, M.; Lacerenza, S. Urinary 1H-NMR and GC-MS metabolomics predicts early and late onset neonatal sepsis. *Early human development* **2014**, *90*, S78-S83.

(83) Lanza, I. R.; Zhang, S.; Ward, L. E.; Karakelides, H.; Raftery, D.; Nair, K. S. Quantitative metabolomics by 1H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes. *PloS one* **2010**, *5*, e10538.

(84) Liu, Y.; Hong, Z.; Tan, G.; Dong, X.; Yang, G.; Zhao, L.; Chen, X.; Zhu, Z.; Lou, Z.; Qian, B. NMR and LC/MS-based global metabolomics to identify serum biomarkers

differentiating hepatocellular carcinoma from liver cirrhosis. *International journal of cancer* **2014**, *135*, 658-668.

(85) Sangster, T.; Major, H.; Plumb, R.; Wilson, A. J.; Wilson, I. D. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *Analyst* **2006**, *131*, 1075-1078.

(86) Chen, M.; Rao, R. S. P.; Zhang, Y.; Zhong, C. X.; Thelen, J. J. A modified data normalization method for GC-MS-based metabolomics to minimize batch variation. *SpringerPlus* **2014**, *3*, 439.

(87) Zhang, T.; Watson, D. G.; Wang, L.; Abbas, M.; Murdoch, L.; Bashford, L.; Ahmad, I.; Lam, N. Y.; Ng, A. C.; Leung, H. Y. Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry Based Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery. *PLoS One* **2013**, *8*, e65880.

(88) Abdelrazig, S.; Ortori, C. A.; Davey, G.; Deressa, W.; Mulleta, D.; Barrett, D. A.; Amberbir, A.; Fogarty, A. W. A metabolomic analytical approach permits identification of urinary biomarkers for Plasmodium falciparum infection: a case–control study. *Malaria journal* **2017**, *16*, 229.

(89) Theodoridis, G.; Gika, H. G.; Wilson, I. D. LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. *TrAC Trends in Analytical Chemistry* **2008**, *27*, 251-260.

(90) Xiao, J. F.; Zhou, B.; Ressom, H. W. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC Trends in Analytical Chemistry* **2012**, *32*, 1-14.

(91) Becker, S.; Kortz, L.; Helmschrodt, C.; Thiery, J.; Ceglarek, U. LC–MS-based metabolomics in the clinical laboratory. *Journal of Chromatography B* **2012**, *883*, 68-75.

(92) Buszewski, B.; Noga, S. Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique. *Anal Bioanal Chem* **2012**, *402*, 231-247.

(93) Williams, P.; Strauss, C.; Wilson, B.; Massy-Westropp, R. Use of C18 reversed-phase liquid chromatography for the isolation of monoterpene glycosides and nor-isoprenoid precursors from grape juice and wines. *Journal of Chromatography A* **1982**, *235*, 471-480.

(94) Tanaka, T.; Nakashima, T.; Ueda, T.; Tomii, K.; Kouno, I. Facile discrimination of aldose enantiomers by reversed-phase HPLC. *Chemical and pharmaceutical bulletin* **2007**, *55*, 899-901.

(95) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC− MS/MS) for large-scale protein analysis: the yeast proteome. *Journal of proteome research* **2003**, *2*, 43-50.

(96) Churchwell, M. I.; Twaddle, N. C.; Meeker, L. R.; Doerge, D. R. Improving LC–MS sensitivity through increases in chromatographic performance: Comparisons of UPLC–ES/MS/MS to HPLC–ES/MS/MS. *Journal of Chromatography B* **2005**, *825*, 134-143.

(97) Jandera, P. Stationary and mobile phases in hydrophilic interaction chromatography: a review. *Analytica chimica acta* **2011**, *692*, 1-25.

(98) Zhang, R.; Watson, D. G.; Wang, L.; Westrop, G. D.; Coombs, G. H.; Zhang, T. Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites. *J Chromatogr A* **2014**, *1362*, 168-179.

(99) Chen, J.; Wang, W.; Lv, S.; Yin, P.; Zhao, X.; Lu, X.; Zhang, F.; Xu, G. Metabonomics study of liver cancer based on ultra performance liquid chromatography coupled to mass spectrometry with HILIC and RPLC separations. *Analytica Chimica Acta* **2009**, *650*, 3-9.

(100) Gros, M.; Petrović, M.; Barceló, D. Development of a multi-residue analytical methodology based on liquid chromatography–tandem mass spectrometry (LC–MS/MS) for

screening and trace level determination of pharmaceuticals in surface and wastewaters. *Talanta* **2006**, *70*, 678-690.

(101) Wahlen, R.; Catterick, T. Comparison of different liquid chromatography conditions for the separation and analysis of organotin compounds in mussel and oyster tissue by liquid chromatography–inductively coupled plasma mass spectrometry. *Journal of Chromatography B* **2003**, *783*, 221-229.

(102) Han, X.; Gross, R. W. Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry a bridge to lipidomics. *Journal of lipid research* **2003**, *44*, 1071-1079.

(103) Blount, B. C.; Milgram, K. E.; Silva, M. J.; Malek, N. A.; Reidy, J. A.; Needham, L. L.; Brock, J. W. Quantitative detection of eight phthalate metabolites in human urine using HPLC− APCI-MS/MS. *Analytical chemistry* **2000**, *72*, 4127-4134.

(104) Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* **2004**, *306*, 471-473.

(105) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S.: Unraveling the mechanism of electrospray ionization. ACS Publications, 2012.

(106) Kujawinski, E. B.; Freitas, M. A.; Zang, X.; Hatcher, P. G.; Green-Church, K. B.; Jones, R. B. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Organic geochemistry* **2002**, *33*, 171-180.

(107) Awad, H.; Khamis, M. M.; El-Aneed, A. Mass spectrometry, review of the basics: ionization. *Applied Spectroscopy Reviews* **2015**, *50*, 158-175.

(108) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. Metabolomic applications of hilic–lc–ms. *Mass spectrometry reviews* **2010**, *29*, 671-684.

(109) Gika, H. G.; Wilson, I. D.; Theodoridis, G. A. LC–MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. *Journal of Chromatography B* **2014**, *966*, 1-6.

(110) Eng, J. Sample size estimation: how many individuals should be studied? *Radiology* **2003**, *227*, 309-313.

(111) Billoir, E.; Navratil, V.; Blaise, B. J. Sample size calculation in metabolic phenotyping studies. *Briefings in bioinformatics* **2015**, *16*, 813-819.

(112) Nakagawa, S. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral ecology* **2004**, *15*, 1044-1045.

(113) Fitzner, K.; Heckinger, E. Sample size calculation and power analysis: a quick review. *The Diabetes Educator* **2010**, *36*, 701-707.

(114) Blaise, B. J.; Correia, G. a.; Tin, A.; Young, J. H.; Vergnaud, A.-C.; Lewis, M.; Pearce, J. T.; Elliott, P.; Nicholson, J. K.; Holmes, E. Power analysis and sample size determination in metabolic phenotyping. *Analytical chemistry* **2016**, *88*, 5179-5188.

(115) Wu, H.; Southam, A. D.; Hines, A.; Viant, M. R. High-throughput tissue extraction protocol for NMR-and MS-based metabolomics. *Analytical biochemistry* **2008**, *372*, 204-212.

(116) Dettmer, K.; Nürnberger, N.; Kaspar, H.; Gruber, M. A.; Almstetter, M. F.; Oefner, P. J. Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols. *Analytical and bioanalytical chemistry* **2011**, *399*, 1127-1139.

(117) Teng, Q.; Huang, W.; Collette, T. W.; Ekman, D. R.; Tan, C. A direct cell quenching method for cell-culture based metabolomics. *Metabolomics* **2009**, *5*, 199.

(118) Canelas, A. B.; Ras, C.; Ten Pierick, A.; van Dam, J. C.; Heijnen, J. J.; Van Gulik, W. M. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics* **2008**, *4*, 226-239.

(119) Pinu, F. R.; Villas-Boas, S. G.; Aggio, R. Analysis of Intracellular Metabolites from Microorganisms: Quenching and Extraction Protocols. *Metabolites* **2017**, *7*, 53.

(120) Dunn, W. B.; Wilson, I. D.; Nicholls, A. W.; Broadhurst, D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis* **2012**, *4*, 2249-2264.

(121) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007**, *3*, 211-221.

(122) Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. Liquid chromatography and ultra-performance liquid chromatography–mass spectrometry fingerprinting of human urine: sample stability under different handling and storage conditions for metabonomics studies. *Journal of Chromatography A* **2008**, *1189*, 314-322.

(123) Mushtaq, M. Y.; Choi, Y. H.; Verpoorte, R.; Wilson, E. G. Extraction for metabolomics: access to the metabolome. *Phytochemical Analysis* **2014**, *25*, 291-306.

(124) Sellick, C. A.; Knight, D.; Croxford, A. S.; Maqsood, A. R.; Stephens, G. M.; Goodacre, R.; Dickson, A. J. Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets. *Metabolomics* **2010**, *6*, 427-438.

(125) Wei, G.-T.; Yang, Z.; Chen, C.-J. Room temperature ionic liquid as a novel medium for liquid/liquid extraction of metal ions. *Analytica Chimica Acta* **2003**, *488*, 183-192.

(126) Salem, M.; Bernach, M.; Bajdzienko, K.; Giavalisco, P. A simple fractionated extraction method for the comprehensive analysis of metabolites, lipids, and proteins from a single sample. *Journal of visualized experiments: JoVE* **2017**.

(127) Vuckovic, D. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Analytical and bioanalytical chemistry* **2012**, *403*, 1523-1548.

(128) Ju, Z. Y.; Howard, L. R. Effects of solvent and temperature on pressurized liquid extraction of anthocyanins and total phenolics from dried red grape skin. *Journal of Agricultural and food Chemistry* **2003**, *51*, 5207-5213.

(129) Sellick, C. A.; Hansen, R.; Stephens, G. M.; Goodacre, R.; Dickson, A. J. Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling. *nature protocols* **2011**, *6*, 1241.

(130) Beltran, A.; Suarez, M.; Rodríguez, M. A.; Vinaixa, M.; Samino, S.; Arola, L.; Correig, X.; Yanes, O. Assessment of compatibility between extraction methods for NMR-and LC/MS-based metabolomics. *Analytical chemistry* **2012**, *84*, 5838-5844.

(131) Bruce, S. J.; Tavazzi, I.; Parisod, V. r.; Rezzi, S.; Kochhar, S.; Guy, P. A. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Analytical chemistry* **2009**, *81*, 3285-3296.

(132) Vuckovic, D.; Pawliszyn, J. Systematic evaluation of solid-phase microextraction coatings for untargeted metabolomic profiling of biological fluids by liquid chromatography− mass spectrometry. *Analytical chemistry* **2011**, *83*, 1944-1954.

(133) Angerer, J.; Ritter, A. Determination of metabolites of pyrethroids in human urine using solid-phase extraction and gas chromatography–mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications* **1997**, *695*, 217-226.

(134) Maubach, J.; Bracke, M. E.; Heyerick, A.; Depypere, H. T.; Serreyn, R. F.; Mareel, M. M.; De Keukeleire, D. Quantitation of soy-derived phytoestrogens in human breast tissue and biological fluids by high-performance liquid chromatography. *Journal of Chromatography B* **2003**, *784*, 137-144.

(135) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature protocols* **2013**, *8*, 17.

(136) Forgacsova, A.; Galba, J.; Garruto, R. M.; Majerova, P.; Katina, S.; Kovac, A. A novel liquid chromatography/mass spectrometry method for determination of neurotransmitters in brain tissue: Application to human tauopathies. *Journal of Chromatography B* **2018**, *1073*, 154-162.

(137) Römisch-Margl, W.; Prehn, C.; Bogumil, R.; Röhring, C.; Suhre, K.; Adamski, J. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. *Metabolomics* **2012**, *8*, 133-142.

(138) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3*, 211-221.

(139) Hayton, S.; Maker, G. L.; Mullaney, I.; Trengove, R. D. Experimental design and reporting standards for metabolomics studies of mammalian cell lines. *Cellular and Molecular Life Sciences* **2017**, *74*, 4421-4441.

(140) Katajamaa, M.; Orešič, M. Data processing for mass spectrometry-based metabolomics. *Journal of chromatography A* **2007**, *1158*, 318-328.

(141) Sugimoto, M.; Kawakami, M.; Robert, M.; Soga, T.; Tomita, M. Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. *Current bioinformatics* **2012**, *7*, 96-108.

(142) Wei, X.; Shi, X.; Kim, S.; Zhang, L.; Patrick, J. S.; Binkley, J.; McClain, C.; Zhang, X. Data preprocessing method for liquid chromatography–mass spectrometry based metabolomics. *Analytical chemistry* **2012**, *84*, 7963-7971.

(143) Lommen, A.; Kools, H. J. MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. *Metabolomics* **2012**, *8*, 719-726.

(144) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical chemistry* **2006**, *78*, 779-787.

(145) Boccard, J.; Rudaz, S.: Extracting Knowledge from MS Clinical Metabolomic Data: Processing and Analysis Strategies. In *Clinical Metabolomics*; Springer, 2018; pp 371-384.

(146) Hendriks, M. M.; van Eeuwijk, F. A.; Jellema, R. H.; Westerhuis, J. A.; Reijmers, T. H.; Hoefsloot, H. C.; Smilde, A. K. Data-processing strategies for metabolomics studies. *TrAC Trends in Analytical Chemistry* **2011**, *30*, 1685-1698.

(147) Creek, D. J.; Jankevics, A.; Burgess, K. E.; Breitling, R.; Barrett, M. P. IDEOM: an Excel interface for analysis of LC-MS-based metabolomics data. *Bioinformatics* **2012**, *28*, 1048-1049.

(148) Creek, D. J.; Jankevics, A.; Breitling, R.; Watson, D. G.; Barrett, M. P.; Burgess, K. E. Toward global metabolomics analysis with hydrophilic interaction liquid chromatography–mass spectrometry: improved metabolite identification by retention time prediction. *Analytical chemistry* **2011**, *83*, 8703-8710.

(149) Bartel, J.; Krumsiek, J.; Theis, F. J. Statistical methods for the analysis of high-throughput metabolomics data. *Computational and structural biotechnology journal* **2013**, *4*, e201301009.

(150) Matsuda, F.; Shinbo, Y.; Oikawa, A.; Hirai, M. Y.; Fiehn, O.; Kanaya, S.; Saito, K. Assessment of metabolome annotation quality: a method for evaluating the false discovery rate of elemental composition searches. *PLoS One* **2009**, *4*, e7490.

(151) Smilde, A. K.; Jansen, J. J.; Hoefsloot, H. C.; Lamers, R.-J. A.; Van Der Greef, J.; Timmerman, M. E. ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics* **2005**, *21*, 3043-3048.

(152) Broadhurst, D. I.; Kell, D. B. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2006**, *2*, 171-196.

(153) Saccenti, E.; Hoefsloot, H. C.; Smilde, A. K.; Westerhuis, J. A.; Hendriks, M. M. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics* **2014**, *10*, 361-374.

(154) Worley, B.; Powers, R. Multivariate analysis in metabolomics. *Current Metabolomics* **2013**, *1*, 92-107.

(155) Zhang, J.; Yan, L.; Chen, W.; Lin, L.; Song, X.; Yan, X.; Hang, W.; Huang, B. Metabonomics research of diabetic nephropathy and type 2 diabetes mellitus based on UPLC–oaTOF-MS system. *Analytica chimica acta* **2009**, *650*, 16-22.

(156) Jansen, J. J.; Hoefsloot, H. C.; van der Greef, J.; Timmerman, M. E.; Westerhuis, J. A.; Smilde, A. K. ASCA: analysis of multivariate data obtained from an experimental design. *Journal of Chemometrics: A Journal of the Chemometrics Society* **2005**, *19*, 469-481.

(157) Vis, D. J.; Westerhuis, J. A.; Smilde, A. K.; van der Greef, J. Statistical validation of megavariate effects in ASCA. *BMC bioinformatics* **2007**, *8*, 322.

(158) Boccard, J.; Rutledge, D. N. A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Analytica chimica acta* **2013**, *769*, 30-39.

(159) Triba, M. N.; Le Moyec, L.; Amathieu, R.; Goossens, C.; Bouchemal, N.; Nahon, P.; Rutledge, D. N.; Savarin, P. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems* **2015**, *11*, 13-19.

(160) Westerhuis, J. A.; van Velzen, E. J.; Hoefsloot, H. C.; Smilde, A. K. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics* **2010**, *6*, 119-128.

(161) Martínez-Arranz, I.; Mayo, R.; Pérez-Cormenzana, M.; Mincholé, I.; Salazar, L.; Alonso, C.; Mato, J. M. Enhancing metabolomics research through data mining. *Journal of proteomics* **2015**, *127*, 275-288.

(162) Cambiaghi, A.; Ferrario, M.; Masseroli, M. Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Briefings in bioinformatics* **2017**, *18*, 498-510.

(163) Saude, E. J.; Adamko, D.; Rowe, B. H.; Marrie, T.; Sykes, B. D. Variation of metabolites in normal human urine. *Metabolomics* **2007**, *3*, 439-451.

(164) Yuan, L.; Zhang, D.; Jemal, M.; Aubry, A. F. Systematic evaluation of the root cause of non-linearity in liquid chromatography/tandem mass spectrometry bioanalytical assays and strategy to predict and extend the linear standard curve range. *Rapid Communications in Mass Spectrometry* **2012**, *26*, 1465-1474.

(165) Beach, D. G.; Gabryelski, W. Linear and nonlinear regimes of electrospray signal response in analysis of urine by electrospray ionization-high field asymmetric waveform ion mobility spectrometry-MS and implications for nontarget quantification. *Analytical chemistry* **2013**, *85*, 2127-2134.

(166) Ouerdane, L.; Meija, J.; Bakirdere, S.; Yang, L.; Mester, Z. n. Nonlinear signal response in electrospray mass spectrometry: implications for quantitation of arsenobetaine using stable isotope labeling by liquid chromatography and electrospray Orbitrap mass spectrometry. *Analytical chemistry* **2012**, *84*, 3958-3964.

(167) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. An overview of matrix effects in liquid chromatography–mass spectrometry. *Mass spectrometry reviews* **2011**, *30*, 491-509.

(168) Enke, C. G. A predictive model for matrix and analyte effects in electrospray ionization of singly-charged ionic analytes. *Analytical Chemistry* **1997**, *69*, 4885-4893.

(169) Cappiello, A.; Famiglini, G.; Palma, P.; Pierini, E.; Termopoli, V.; Trufelli, H. Overcoming matrix effects in liquid chromatography− Mass Spectrometry. *Analytical chemistry* **2008**, *80*, 9343-9348.

(170) Smeraglia, J.; Baldrey, S.; Watson, D. Matrix effects and selectivity issues in LC-MS-MS. *Chromatographia* **2002**, *55*, S95-S99.

(171) Xu, Y.-F.; Lu, W.; Rabinowitz, J. D. Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography–mass spectrometry-based metabolomics. *Analytical chemistry* **2015**, *87*, 2273-2281.

(172) Livera, A. M. D.; Sysi-Aho, M.; Jacob, L.; Gagnon-Bartsch, J. A.; Castillo, S.; Simpson, J. A.; Speed, T. P. Statistical methods for handling unwanted variation in metabolomics data. *Analytical chemistry* **2015**, *87*, 3606-3615.

(173) Parsons, H. M.; Ekman, D. R.; Collette, T. W.; Viant, M. R. Spectral relative standard deviation: a practical benchmark in metabolomics. *Analyst* **2009**, *134*, 478-485.

(174) Shurubor, Y. I.; Paolucci, U.; Krasnikov, B. F.; Matson, W. R.; Kristal, B. S. Analytical precision, biological variation, and mathematical normalization in high data density metabolomics. *Metabolomics* **2005**, *1*, 75-85.

(175) Gormley, S. M.; McBride, W. T.; Armstrong, M. A.; Young, I. S.; McClean, E.; MacGowan, S. W.; Campalani, G.; McMurray, T. J. Plasma and urinary cytokine homeostasis and renal dysfunction during cardiac surgery. *Anesthesiology: The Journal of the American Society of Anesthesiologists* **2000**, *93*, 1210-1216.

(176) Lin, C. Y.; Wu, H.; Tjeerdema, R. S.; Viant, M. R. Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics* **2007**, *3*, 55-67.

(177) Mizuno, H.; Ueda, K.; Kobayashi, Y.; Tsuyama, N.; Todoroki, K.; Min, J. Z.; Toyo'oka, T. The great importance of normalization of LC–MS data for highly-accurate non-targeted metabolomics. *Biomedical Chromatography* **2017**, *31*, e3864.

(178) Ejigu, B. A.; Valkenborg, D.; Baggerman, G.; Vanaerschot, M.; Witters, E.; Dujardin, J.-C.; Burzykowski, T.; Berg, M. Evaluation of normalization methods to pave the way towards large-scale LC-MS-based metabolomics profiling experiments. *Omics: a journal of integrative biology* **2013**, *17*, 473-485.

(179) Tang, K. W.; Toh, Q. C.; Teo, B. W. Normalisation of urinary biomarkers to creatinine for clinical practice and research – when and why. *Singapore Medical Journal* **2015**, *56*, 7-10.

(180) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. Use of a pre-analysis osmolality normalisation method to correct for variable urine concentrations and for improved metabolomic analyses. *J Chromatogr A* **2016**, *1431*, 103-110.

(181) Warrack, B. M.; Hnatyshyn, S.; Ott, K. H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M. Normalization strategies for metabonomic analysis of urine samples. *J Chromatogr B Analyt Technol Biomed Life Sci* **2009**, *877*, 547-552.

(182) Deininger, S.-O.; Cornett, D. S.; Paape, R.; Becker, M.; Pineau, C.; Rauser, S.; Walch, A.; Wolski, E. Normalization in MALDI-TOF imaging datasets of proteins: practical considerations. *Analytical and bioanalytical chemistry* **2011**, *401*, 167-181.

(183) Sun, N.; Walch, A. Qualitative and quantitative mass spectrometry imaging of drugs and metabolites in tissue at therapeutic levels. *Histochemistry and cell biology* **2013**, *140*, 93-104.

(184) Brunius, C.; Shi, L.; Landberg, R. Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. *Metabolomics* **2016**, *12*, 173.

(185) Edmands, W. M.; Ferrari, P.; Scalbert, A. Normalization to specific gravity prior to analysis improves information recovery from high resolution mass spectrometry metabolomic profiles of human urine. *Analytical chemistry* **2014**, *86*, 10925-10931.

(186) Wu, Y.; Li, L. Dansylation metabolite assay: a simple and rapid method for sample amount normalization in metabolomics. *Analytical chemistry* **2014**, *86*, 9428-9433.

(187) Silva, L. P.; Lorenzi, P. L.; Purwaha, P.; Yong, V.; Hawke, D. H.; Weinstein, J. N. Measurement of DNA concentration as a normalization strategy for metabolomic data from adherent cell lines. *Analytical chemistry* **2013**, *85*, 9536-9542.

(188) Redestig, H.; Fukushima, A.; Stenlund, H.; Moritz, T.; Arita, M.; Saito, K.; Kusano, M. Compensation for systematic cross-contribution improves normalization of mass spectrometry based metabolomics data. *Analytical chemistry* **2009**, *81*, 7974-7980.

(189) Kim, D. H.; Achcar, F.; Breitling, R.; Burgess, K. E.; Barrett, M. P. LC-MS-based absolute metabolite quantification: application to metabolic flux measurement in trypanosomes. *Metabolomics* **2015**, *11*, 1721-1732.

(190) Schatschneider, S.; Abdelrazig, S.; Safo, L.; Henstra, A. M.; Millat, T.; Kim, D.-H.; Winzer, K.; Minton, N. P.; Barrett, D. A. Quantitative isotope dilution high-resolution mass spectrometry analysis of multiple intracellular metabolites in Clostridium autoethanogenum using uniformly 13C-labelled standards derived from Spirulina. *Analytical chemistry* **2018**.

(191) Gagnebin, Y.; Tonoli, D.; Lescuyer, P.; Ponte, B.; de Seigneux, S.; Martin, P.-Y.; Schappler, J.; Boccard, J.; Rudaz, S. Metabolomic analysis of urine samples by UHPLC-QTOF-MS: Impact of normalization strategies. *Analytica chimica acta* **2017**, *955*, 27-35.

(192) Hughes, G.; Cruickshank-Quinn, C.; Reisdorph, R.; Lutz, S.; Petrache, I.; Reisdorph, N.; Bowler, R.; Kechris, K. MSPrep—Summarization, normalization and diagnostics for processing of mass spectrometry–based metabolomic data. *Bioinformatics* **2013**, *30*, 133-134.

(193) De Livera, A. M.; Dias, D. A.; De Souza, D.; Rupasinghe, T.; Pyke, J.; Tull, D.; Roessner, U.; McConville, M.; Speed, T. P. Normalizing and integrating metabolomics data. *Anal Chem* **2012**, *84*, 10768-10776.

(194) Bueschl, C.; Kluger, B.; Lemmens, M.; Adam, G.; Wiesenberger, G.; Maschietto, V.; Marocco, A.; Strauss, J.; Bodi, S.; Thallinger, G. G.; Krska, R.; Schuhmacher, R. A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research. *Metabolomics* **2014**, *10*, 754-769.

(195) Weindl, D.; Wegner, A.; Jäger, C.; Hiller, K. Isotopologue ratio normalization for non-targeted metabolomics. *Journal of Chromatography A* **2015**, *1389*, 112-119.

(196) Li, P.; Li, Z.; Beck, W. D.; Callahan, P. M.; Terry, A. V.; Bar-Peled, M.; Bartlett, M. G. Bio-generation of stable isotope-labeled internal standards for absolute and relative quantitation of phase II drug metabolites in plasma samples using LC–MS/MS. *Analytical and bioanalytical chemistry* **2015**, *407*, 4053-4063.

(197) Jemal, M.; Schuster, A.; Whigan, D. B. Liquid chromatography/tandem mass spectrometry methods for quantitation of mevalonic acid in human plasma and urine: method validation, demonstration of using a surrogate analyte, and demonstration of unacceptable matrix effect in spite of use of a stable isotope analog internal standard. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 1723-1734.

(198) Stokvis, E.; Rosing, H.; Beijnen, J. H. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry* **2005**, *19*, 401-407.

(199) Wang, S.; Cyronak, M.; Yang, E. Does a stable isotopically labeled internal standard always correct analyte response?: A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *Journal of pharmaceutical and biomedical analysis* **2007**, *43*, 701-707.

(200) Hou, H.; Xiong, W.; Zhang, X.; Song, D.; Tang, G.; Hu, Q. LC-MS-MS measurements of urinary creatinine and the application of creatinine normalization technique on cotinine in smokers' 24 hour urine. *Journal of analytical methods in chemistry* **2012**, *2012*.

(201) Chen, Y.; Shen, G.; Zhang, R.; He, J.; Zhang, Y.; Xu, J.; Yang, W.; Chen, X.; Song, Y.; Abliz, Z. Combination of injection volume calibration by creatinine and MS signals' normalization to overcome urine variability in LC-MS-based metabolomics studies. *Analytical chemistry* **2013**, *85*, 7659-7665.

(202) Kennedy, A. D.; Miller, M. J.; Beebe, K.; Wulff, J. E.; Evans, A. M.; Miller, L. A.; Sutton, V. R.; Sun, Q.; Elsea, S. H. Metabolomic profiling of human urine as a screen for multiple inborn errors of metabolism. *Genetic testing and molecular biomarkers* **2016**, *20*, 485-495.

(203) Vogl, F. C.; Mehrl, S.; Heizinger, L.; Schlecht, I.; Zacharias, H. U.; Ellmann, L.; Nürnberger, N.; Gronwald, W.; Leitzmann, M. F.; Rossert, J. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Analytical and bioanalytical chemistry* **2016**, *408*, 8483-8493.

(204) Pérez-Victoria, I.; Martín, J.; Reyes, F. Combined LC/UV/MS and NMR strategies for the dereplication of marine natural products. *Planta medica* **2016**, *82*, 857-871.

(205) Gyamlani, G. G.; Bergstralh, E. J.; Slezak, J. M.; Larson, T. S. Urinary albumin to osmolality ratio predicts 24-hour urine albumin excretion in diabetes mellitus. *American journal of kidney diseases* **2003**, *42*, 685-692.

(206) Boysen, A. K.; Heal, K. R.; Carlson, L. T.; Ingalls, A. E. Best-Matched Internal Standard Normalization in Liquid Chromatography–Mass Spectrometry Metabolomics Applied to Environmental Samples. *Analytical chemistry* **2018**, *90*, 1363-1369.

(207) Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Orešič, M. Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC bioinformatics* **2007**, *8*, 93.

(208) Bennett, B. D.; Yuan, J.; Kimball, E. H.; Rabinowitz, J. D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat Protoc* **2008**, *3*, 1299-1311.

(209) Katajamaa, M.; Orešič, M. Processing methods for differential analysis of LC/MS profile data. *BMC bioinformatics* **2005**, *6*, 179.

(210) Jonsson, P.; Wuolikainen, A.; Thysell, E.; Chorell, E.; Stattin, P.; Wikstrom, P.; Antti, H. Constrained randomization and multivariate effect projections improve information extraction and biomarker pattern discovery in metabolomics studies involving dependent samples. *Metabolomics* **2015**, *11*, 1667-1678.

(211) Trost, B.; Pajon, R.; Jayaprakash, T.; Kusalik, A. Comparing the similarity of different groups of bacteria to the human proteome. *PloS one* **2012**, *7*, e34007.

(212) Lv, H.; Hung, C. S.; Chaturvedi, K. S.; Hooton, T. M.; Henderson, J. P. Development of an integrated metabolomic profiling approach for infectious diseases research. *Analyst* **2011**, *136*, 4752-4763.

(213) Jewison, T.; Knox, C.; Neveu, V.; Djoumbou, Y.; Guo, A. C.; Lee, J.; Liu, P.; Mandal, R.; Krishnamurthy, R.; Sinelnikov, I. YMDB: the yeast metabolome database. *Nucleic acids research* **2011**, *40*, D815-D820.

(214) Guo, A. C.; Jewison, T.; Wilson, M.; Liu, Y.; Knox, C.; Djoumbou, Y.; Lo, P.; Mandal, R.; Krishnamurthy, R.; Wishart, D. S. ECMDB: the E. coli Metabolome Database. *Nucleic Acids Res* **2013**, *41*, D625-630.

(215) Rampler, E.; Coman, C.; Hermann, G.; Sickmann, A.; Ahrends, R.; Koellensperger, G. LILY-lipidome isotope labeling of yeast: in vivo synthesis of 13 C labeled reference lipids for quantification by mass spectrometry. *Analyst* **2017**, *142*, 1891-1899.

(216) Soufi, B.; Macek, B.: Stable isotope labeling by amino acids applied to bacterial cell culture. In *Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC)*; Springer, 2014; pp 9-22.

(217) Creek, D. J.; Nijagal, B.; Kim, D. H.; Rojas, F.; Matthews, K. R.; Barrett, M. P. Metabolomics guides rational development of a simplified cell culture medium for drug screening against Trypanosoma brucei. *Antimicrob Agents Chemother* **2013**, *57*, 2768-2779.

(218) Kwon, Y.; Park, S.; Shin, J.; Oh, D.-C. Application of 13 C-labeling and 13 C–13 C COSY NMR experiments in the structure determination of a microbial natural product. *Archives of pharmacal research* **2014**, *37*, 967-971.

(219) Yang, C.; Harrison, C.; Jin, E. S.; Chuang, D. T.; Sherry, A. D.; Malloy, C. R.; Merritt, M. E.; DeBerardinis, R. J. Simultaneous steady-state and dynamic 13C NMR can differentiate alternative routes of pyruvate metabolism in living cancer cells. *Journal of Biological Chemistry* **2014**, jbc. M113. 543637.

(220) Moreno, K. X.; Satapati, S.; DeBerardinis, R. J.; Burgess, S. C.; Malloy, C. R.; Merritt, M. E. Real-time detection of hepatic gluconeogenic and glycogenolytic states using hyperpolarized [2-13C] dihydroxyacetone. *Journal of Biological Chemistry* **2014**, jbc. M114. 613265.

(221) Colombo Serra, S.; Karlsson, M.; Giovenzana, G. B.; Cavallotti, C.; Tedoldi, F.; Aime, S. Hyperpolarized 13C-labelled anhydrides as DNP precursors of metabolic MRI agents. *Contrast Media & Molecular Imaging* **2012**, *7*, 469-477.

(222) Vadlapudi, V.; Borah, N.; Yellusani, K. R.; Gade, S.; Reddy, P.; Rajamanikyam, M.; Vempati, L. N. S.; Gubbala, S. P.; Chopra, P.; Upadhyayula, S. M. Aspergillus Secondary Metabolite Database, a resource to understand the Secondary metabolome of Aspergillus genus. *Scientific reports* **2017**, *7*, 7325.

(223) Marcinowska, R.; Trygg, J.; Wolf-Watz, H.; Mortiz, T.; Surowiec, I. Optimization of a sample preparation method for the metabolomic analysis of clinically relevant bacteria. *Journal of microbiological methods* **2011**, *87*, 24-31.

(224) Bose, U.; Hewavitharana, A. K.; Ng, Y. K.; Shaw, P. N.; Fuerst, J. A.; Hodson, M. P. LC-MS-Based metabolomics study of marine bacterial secondary metabolite and antibiotic production in Salinispora arenicola. *Marine drugs* **2015**, *13*, 249-266.

(225) Bharucha, A. E.; Camilleri, M.; Veil, E.; Burton, D.; Zinsmeister, A. R. Comprehensive assessment of gastric emptying with a stable isotope breath test. *Neurogastroenterology & Motility* **2013**, *25*, e60-e69.

(226) Chokkathukalam, A.; Kim, D.-H.; Barrett, M. P.; Breitling, R.; Creek, D. J. Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis* **2014**, *6*, 511-524.

(227) Yan, S.; Zhu, C.; Yu, T.; Huang, W.; Huang, J.; Kong, Q.; Shi, J.; Chen, Z.; Liu, Q.; Wang, S. Studying the Differences of Bacterial Metabolome and Microbiome in the Colon between Landrace and Meihua Piglets. *Frontiers in microbiology* **2017**, *8*, 1812.

(228) Zhong, F.; Xu, M.; Metz, P.; Ghosh-Dastidar, P.; Zhu, J. A quantitative metabolomics study of bacterial metabolites in different domains. *Analytica Chimica Acta* **2018**.

(229) Kapoore, R. V.; Vaidyanathan, S. Towards quantitative mass spectrometry-based metabolomics in microbial and mammalian systems. *Phil. Trans. R. Soc. A* **2016**, *374*, 20150363.

(230) Mossel, D.; Laarhoven, T. M. B. V.; Ligtenberg-Merkus, A. M. T.; Werdler, M. E. Quality assurance of selective culture media for bacteria, moulds and yeasts: an attempt at standardization at the international level. *Journal of Applied Bacteriology* **1983**, *54*, 313-327.

(231) Bryant, M. P.; Burkey, L. A. Cultural methods and some characteristics of some of the more numerous groups of bacteria in the bovine rumen. *Journal of Dairy Science* **1953**, *36*, 205-217.

(232) Widdel, F. Theory and measurement of bacterial growth. *Di dalam Grundpraktikum Mikrobiologie* **2007**, *4*.

(233) Neidhardt, F. C.; Bloch, P. L.; Smith, D. F. Culture medium for enterobacteria. *Journal of bacteriology* **1974**, *119*, 736-747.

(234) Sezonov, G.; Joseleau-Petit, D.; D'Ari, R. Escherichia coli physiology in Luria-Bertani broth. *Journal of bacteriology* **2007**, *189*, 8746-8749.

(235) Ezraty, B.; Henry, C.; Hérisse, M.; Denamur, E.; Barras, F. Commercial Lysogeny Broth culture media and oxidative stress: A cautious tale. *Free Radical Biology and Medicine* **2014**, *74*, 245-251.

(236) Papasotiriou, D. G.; Markoutsa, S.; Meyer, B.; Papadioti, A.; Karas, M.; Tsiotis, G. Comparison of the membrane subproteomes during growth of a new Pseudomonas strain on lysogeny broth medium, glucose, and phenol. *Journal of proteome research* **2008**, *7*, 4278-4288.

(237) Mizunoe, Y.; Wai, S. N.; Takade, A.; Yoshida, S.-i. Restoration of culturability of starvation-stressed and low-temperature-stressed Escherichia coli O157 cells by using H2O2-degrading compounds. *Archives of Microbiology* **1999**, *172*, 63-67.

(238) Arulanantham, R.; Pathmanathan, S.; Ravimannan, N.; Niranjan, K. Alternative culture media for bacterial growth using different formulation of protein sources. *J Nat Prod Plant Resour* **2012**, *2*, 697-700.

(239) Wijtes, T.; McClure, P.; Zwietering, M.; Roberts, T. Modelling bacterial growth of Listeria monocytogenes as a function of water activity, pH and temperature. *International Journal of Food Microbiology* **1993**, *18*, 139-149.

(240) Presser, K.; Ross, T.; Ratkowsky, D. Modelling the growth limits (growth/no growth interface) of Escherichia coli as a function of temperature, pH, lactic acid concentration, and water activity. *Applied and Environmental Microbiology* **1998**, *64*, 1773-1779.

(241) Madkour, F. F.; Kamil, A. E.-W.; Nasr, H. S. Production and nutritive value of Spirulina platensis in reduced cost media. *The egyptian journal of aquatic research* **2012**, *38*, 51-57.

(242) Ismaiel, M. M. S.; El-Ayouty, Y. M.; Piercey-Normore, M. Role of pH on antioxidants production by Spirulina (Arthrospira) platensis. *Brazilian journal of microbiology* **2016**, *47*, 298-304.

(243) Ratkowsky, D.; Olley, J.; McMeekin, T.; Ball, A. Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology* **1982**, *149*, 1-5.

(244) Pietikäinen, J.; Pettersson, M.; Bååth, E. Comparison of temperature effects on soil respiration and bacterial and fungal growth rates. *FEMS Microbiology Ecology* **2005**, *52*, 49-58.

(245) Repaske, R.; Clayton, M. Control of Escherichia coli growth by CO2. *Journal of bacteriology* **1978**, *135*, 1162.

(246) Eklund, T. The effect of carbon dioxide on bacterial growth and on uptake processes in bacterial membrane vesicles. *International Journal of Food Microbiology* **1984**, *1*, 179-185.

(247) Dixon, N. M.; Kell, D. B. The inhibition by CO2 of the growth and metabolism of micro-organisms. *Journal of Applied Bacteriology* **1989**, *67*, 109-136.

(248) Fee, J.; Lees, A.; Bloch, P.; Gilliland, P.; Brown, O. Oxygen stasis of bacterial growth: analogy between the stasis of E. coli by hyperbaric oxygen and by aerobic paraquat. *Biochemistry International* **1980**, *1*, 304-311.

(249) Sanders, E. R. Aseptic laboratory techniques: plating methods. *Journal of visualized experiments: JoVE* **2012**.

(250) Nebe-von-Caron, G.; Stephens, P.; Hewitt, C.; Powell, J.; Badley, R. Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. *Journal of microbiological methods* **2000**, *42*, 97-114.

(251) DeLand, F.; Wagner, H. N. Automated radiometric detection of bacterial growth in blood cultures. *Translational Research* **1970**, *75*, 529-534.

(252) Zhu, C.; Lee, Y. Determination of biomass dry weight of marine microalgae. *Journal of applied phycology* **1997**, *9*, 189-194.

(253) Monod, J. The growth of bacterial cultures. *Annual Reviews in Microbiology* **1949**, *3*, 371-394.

(254) Stevenson, K.; McVey, A. F.; Clark, I. B.; Swain, P. S.; Pilizota, T. General calibration of microbial growth in microplate readers. *Scientific reports* **2016**, *6*, 38828.

(255) Sutton, S. Measurement of microbial cells by optical density. *Journal of Validation technology* **2011**, *17*, 46.

(256) Zwietering, M.; Jongenburger, I.; Rombouts, F.; Van't Riet, K. Modeling of the bacterial growth curve. *Applied and environmental microbiology* **1990**, *56*, 1875-1881.

(257) Begot, C.; Desnier, I.; Daudin, J. D.; Labadie, J. C.; Lebert, A. Recommendations for calculating growth parameters by optical density measurements. *Journal of Microbiological Methods* **1996**, *25*, 225-232.

(258) Korz, D.; Rinas, U.; Hellmuth, K.; Sanders, E.; Deckwer, W.-D. Simple fed-batch technique for high cell density cultivation of Escherichia coli. *Journal of biotechnology* **1995**, *39*, 59-65.

(259) Soni, S.; Agrawal, K.; Srivastava, S.; Gupta, S.; Pankaj, C. K. Growth performance and biochemical analysis of Spirulina platensis under different culture conditions. *J. Algal Biomass Utln* **2012**, *3*, 55-58.

(260) Caputo, J. L. Biosafety procedures in cell culture. *Journal of tissue culture methods* **1988**, *11*, 223-227.

(261) Zaki, A. N. Biosafety and biosecurity measures: management of biosafety level 3 facilities. *International journal of antimicrobial agents* **2010**, *36*, S70-S74.

(262) Bertini, I.; Hu, X.; Luchinat, C. Global metabolomics characterization of bacteria: pre-analytical treatments and profiling. *Metabolomics* **2014**, *10*, 241-249.

(263) Pinu, F.; Villas-Boas, S. Extracellular microbial metabolomics: The state of the art. *Metabolites* **2017**, *7*, 43.

(264) Faijes, M.; Mars, A. E.; Smid, E. J. Comparison of quenching and extraction methodologies for metabolome analysis of Lactobacillus plantarum. *Microbial cell factories* **2007**, *6*, 27.

(265) Canelas, A. B.; ten Pierick, A.; Ras, C.; Seifar, R. M.; van Dam, J. C.; van Gulik, W. M.; Heijnen, J. J. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. *Analytical chemistry* **2009**, *81*, 7379-7389.

(266) Spura, J.; Reimer, L. C.; Wieloch, P.; Schreiber, K.; Buchinger, S.; Schomburg, D. A method for enzyme quenching in microbial metabolome analysis successfully applied to gram-positive and gram-negative bacteria and yeast. *Analytical biochemistry* **2009**, *394*, 192-201.

(267) Minnikin, D.; O'donnell, A.; Goodfellow, M.; Alderson, G.; Athalye, M.; Schaal, A.; Parlett, J. An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *Journal of Microbiological Methods* **1984**, *2*, 233-241.

(268) de Koning, W.; van Dam, K. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Analytical biochemistry* **1992**, *204*, 118-123.

(269) Maharjan, R. P.; Ferenci, T. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of Escherichia coli. *Analytical biochemistry* **2003**, *313*, 145-154.

(270) Cuellar-Bermudez, S. P.; Aguilar-Hernandez, I.; Cardenas-Chavez, D. L.; Ornelas-Soto, N.; Romero-Ogawa, M. A.; Parra-Saldivar, R. Extraction and purification of high-value metabolites from microalgae: essential lipids, astaxanthin and phycobiliproteins. *Microbial biotechnology* **2015**, *8*, 190-209.

(271) Muhamadali, H.; Chisanga, M.; Subaihi, A.; Goodacre, R. Combining Raman and FT-IR spectroscopy with quantitative isotopic labeling for differentiation of E. coli cells at community and single cell levels. *Analytical chemistry* **2015**, *87*, 4578-4586.

(272) Wang, Y.; Song, Y.; Tao, Y.; Muhamadali, H.; Goodacre, R.; Zhou, N.-Y.; Preston, G. M.; Xu, J.; Huang, W. E. Reverse and multiple stable isotope probing to study bacterial metabolism and interactions at the single cell level. *Analytical chemistry* **2016**, *88*, 9443-9450.

(273) Kubryk, P.; Kölschbach, J. S.; Marozava, S.; Lueders, T.; Meckenstock, R. U.; Niessner, R.; Ivleva, N. P. Exploring the potential of stable isotope (resonance) Raman microspectroscopy and surface-enhanced Raman scattering for the analysis of microorganisms at single cell level. *Analytical chemistry* **2015**, *87*, 6622-6630.

(274) Tang, Y. J.; Meadows, A. L.; Kirby, J.; Keasling, J. D. Anaerobic central metabolic pathways in Shewanella oneidensis MR-1 reinterpreted in the light of isotopic metabolite labeling. *Journal of Bacteriology* **2007**, *189*, 894-901.

(275) Marley, J.; Lu, M.; Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *Journal of biomolecular NMR* **2001**, *20*, 71-75.

(276) Panetta, J.; Gajjar, A.; Hijiya, N.; Hak, L.; Cheng, C.; Liu, W.; Pui, C.; Relling, M. Comparison of native E. coli and PEG asparaginase pharmacokinetics and pharmacodynamics in pediatric acute lymphoblastic leukemia. *Clinical Pharmacology & Therapeutics* **2009**, *86*, 651-658.

(277) Zhao, M.; Lepak, A. J.; Marchillo, K.; VanHecker, J.; Andes, D. R. In Vivo Pharmacodynamic Characterization of a Novel Odilorhabdin Antibiotic, NOSO-502, against Escherichia coli and Klebsiella pneumoniae in a Murine Thigh Infection Model. *Antimicrobial agents and chemotherapy* **2018**, *62*, e01067-01018.

(278) Castanheira, M.; Griffin, M. A.; Deshpande, L. M.; Mendes, R. E.; Jones, R. N.; Flamm, R. K. Detection of mcr-1 among Escherichia coli clinical isolates collected worldwide as part of the SENTRY antimicrobial surveillance program during 2014-2015. *Antimicrobial agents and chemotherapy* **2016**, AAC. 01267-01216.

(279) Lou Ver Leigh, A.; Andaya, A. G. The effects of Spirulina (Arthrospira platensis) dietary supplement as an adjunct therapy for children aged 7-14 years old with asthma: A randomized-double blind placebo controlled clinical trial. *Journal of Allergy and Clinical Immunology* **2018**, *141*, AB208.

(280) Xiong, J.; Liu, S.; Pan, Y.; Zhang, B.; Chen, X.; Fan, L. Combination of fish oil and ethanol extracts from Spirulina platensis inhibits the airway inflammation induced by ovalbumin in mice. *Journal of Functional Foods* **2018**, *40*, 707-714.

(281) Yousefi, R.; Mottaghi, A.; Saidpour, A. Spirulina platensis effectively ameliorates anthropometric measurements and obesity-related metabolic disorders in obese or overweight healthy individuals: A randomized controlled trial. *Complementary therapies in medicine* **2018**, *40*, 106-112.

(282) Cai, M.; Huang, Y.; Sakaguchi, K.; Clore, G. M.; Gronenborn, A. M.; Craigie, R. An efficient and cost-effective isotope labeling protocol for proteins expressed in shape Escherichia coli. *Journal of biomolecular NMR* **1998**, *11*, 97-102.

(283) Marles, R. J.; Barrett, M. L.; Barnes, J.; Chavez, M. L.; Gardiner, P.; Ko, R.; Mahady, G. B.; Dog, T. L.; Sarma, N. D.; Giancaspro, G. I. United States pharmacopeia safety evaluation of Spirulina. *Critical reviews in food science and nutrition* **2011**, *51*, 593-604.

(284) Surrati, A.; Linforth, R.; Fisk, I. D.; Sottile, V.; Kim, D.-H. Non-destructive characterisation of mesenchymal stem cell differentiation using LC-MS-based metabolite footprinting. *Analyst* **2016**, *141*, 3776-3787.

(285) Rabinowitz, J. D.; Kimball, E. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Analytical Chemistry* **2007**, *79*, 6167-6173.

(286) BaniMustafa, A. H.; Hardy, N. W.: A strategy for selecting data mining techniques in metabolomics. In *Plant Metabolomics*; Springer, 2011; pp 317-333.

(287) Chokkathukalam, A.; Jankevics, A.; Creek, D. J.; Achcar, F.; Barrett, M. P.; Breitling, R. mzMatch-ISO: an R tool for the annotation and relative quantification of isotope-labelled mass spectrometry data. *Bioinformatics* **2013**, *29*, 281-283.

(288) Gromski, P. S.; Xu, Y.; Hollywood, K. A.; Turner, M. L.; Goodacre, R. The influence of scaling metabolomics data on model classification accuracy. *Metabolomics* **2015**, *11*, 684-695.

(289) Dallmann, R.; Viola, A. U.; Tarokh, L.; Cajochen, C.; Brown, S. A. The human circadian metabolome. *Proc Natl Acad Sci U S A* **2012**, *109*, 2625-2629.

(290) Danesi, E. D. G.; Rangel-Yagui, C. O.; Sato, S.; Carvalho, J. C. M. d. Growth and content of Spirulina platensis biomass chlorophyll cultivated at different values of light intensity and temperature using different nitrogen sources. *Brazilian Journal of Microbiology* **2011**, *42*, 362-373.

(291) Hasunuma, T.; Kikuyama, F.; Matsuda, M.; Aikawa, S.; Izumi, Y.; Kondo, A. Dynamic metabolic profiling of cyanobacterial glycogen biosynthesis under conditions of nitrate depletion. *Journal of experimental botany* **2013**, *64*, 2943-2954.

(292) Falquet, J.; Hurni, J. The nutritional aspects of Spirulina. *Antenna Foundation. Available online at: https://www. antenna. ch/wp-content/uploads/2017/03/AspectNut_UK. pdf (Accessed July 25, 2017)* **1997**.

(293) Mendiola, J. A.; Marín, F. R.; Hernandez, S.; Arredondo, B. O.; Señoráns, F. J.; Ibañez, E.; Reglero, G. Characterization via liquid chromatography coupled to diode array detector and tandem mass spectrometry of supercritical fluid antioxidant extracts of Spirulina platensis microalga. *Journal of separation science* **2005**, *28*, 1031-1038.

(294) Wu, Y.; Li, L. Sample normalization methods in quantitative metabolomics. *Journal of Chromatography A* **2016**, *1430*, 80-95.

(295) Clendinen, C. S.; Stupp, G. S.; Ajredini, R.; Lee-McMullen, B.; Beecher, C.; Edison, A. S. An overview of methods using 13C for improved compound identification in metabolomics and natural products. *Frontiers in plant science* **2015**, *6*, 611.

(296) Agalloco, J. Validation: an unconventional review and reinvention. *PDA journal of pharmaceutical science and technology* **1995**, *49*, 175-179.

(297) Food, U.; Administration, D. Guideline on general principles of process validation. *US FDA: Rockville, MD* **1987**.

(298) Health, U. D. o.; Services, H. Guidance for industry, bioanalytical method validation. *http://www. fda. gov/cder/guidance/index. htm* **2001**.

(299) Naz, S.; Vallejo, M.; García, A.; Barbas, C. Method validation strategies involved in non-targeted metabolomics. *Journal of Chromatography A* **2014**, *1353*, 99-105.

(300) López-López, Á.; López-Gonzálvez, Á.; Clive-Baker, T.; Barbas, C. A review of validated biomarkers obtained through metabolomics. *Expert review of molecular diagnostics* **2018**.

(301) t'Kindt, R.; Scheltema, R. A.; Jankevics, A.; Brunker, K.; Rijal, S.; Dujardin, J.-C.; Breitling, R.; Watson, D. G.; Coombs, G. H.; Decuypere, S. Metabolomics to unveil and understand phenotypic diversity between pathogen populations. *PLoS neglected tropical diseases* **2010**, *4*, e904.

(302) Chadha, V.; Garg, U.; Alon, U. S. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatric Nephrology* **2001**, *16*, 374-382.

(303) Mwenechanya, R.; Kovářová, J.; Dickens, N. J.; Mudaliar, M.; Herzyk, P.; Vincent, I. M.; Weidt, S. K.; Burgess, K. E.; Burchmore, R. J.; Pountain, A. W. Sterol 14α-demethylase mutation leads to amphotericin B resistance in Leishmania mexicana. *PLoS neglected tropical diseases* **2017**, *11*, e0005649.

(304) Pouwels, P. J.; Frahm, J. Regional metabolite concentrations in human brain as determined by quantitative localized proton MRS. *Magnetic resonance in medicine* **1998**, *39*, 53-60.

(305) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012**, *2*, 775-795.

(306) Ward, J. L.; Harris, C.; Lewis, J.; Beale, M. H. Assessment of 1H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of Arabidopsis thaliana. *Phytochemistry* **2003**, *62*, 949-957.

(307) Gromski, P. S.; Xu, Y.; Kotze, H. L.; Correa, E.; Ellis, D. I.; Armitage, E. G.; Turner, M. L.; Goodacre, R. Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites* **2014**, *4*, 433-452.

(308) Salek, R. M.; Maguire, M. L.; Bentley, E.; Rubtsov, D. V.; Hough, T.; Cheeseman, M.; Nunez, D.; Sweatman, B. C.; Haselden, J. N.; Cox, R. A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. *Physiological genomics* **2007**, *29*, 99-108.

(309) Rhee, E. P. Metabolomics and renal disease. *Current opinion in nephrology and hypertension* **2015**, *24*, 371.

(310) Kiefer, P.; Schmitt, U.; Vorholt, J. A. eMZed: an open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. *Bioinformatics* **2013**, *29*, 963-964.

(311) Maughan, R. J.; Griffin, J. Caffeine ingestion and fluid balance: a review. *Journal of human nutrition and dietetics* **2003**, *16*, 411-420.

(312) Gika, H. G.; Theodoridis, G. A.; Wingate, J. E.; Wilson, I. D. Within-day reproducibility of an HPLC− MS-based method for metabonomic analysis: application to human urine. *Journal of proteome research* **2007**, *6*, 3291-3303.

(313) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc* **2010**, *5*, 1005-1018.

(314) Zhang, A.; Sun, H.; Wu, X.; Wang, X. Urine metabolomics. *Clinica Chimica Acta* **2012**, *414*, 65-69.

(315) Mulder, T. P.; Rietveld, A. G.; van Amelsvoort, J. M. Consumption of both black tea and green tea results in an increase in the excretion of hippuric acid into urine–. *The American journal of clinical nutrition* **2005**, *81*, 256S-260S.

(316) Huang, H.; Zhang, J.; Wakana, S.; Zhang, W.; Ren, T.; Richards, L. J.; Yarowsky, P.; Donohue, P.; Graham, E.; van Zijl, P. C. White and gray matter development in human fetal, newborn and pediatric brains. *Neuroimage* **2006**, *33*, 27-38.

(317) Whiley, L.; Godzien, J.; Ruperez, F. J.; Legido-Quigley, C.; Barbas, C. In-vial dual extraction for direct LC-MS analysis of plasma for comprehensive and highly reproducible metabolic fingerprinting. *Analytical chemistry* **2012**, *84*, 5992-5999.

(318) Huang, X.; Chen, Y.-J.; Cho, K.; Nikolskiy, I.; Crawford, P. A.; Patti, G. J. X13CMS: global tracking of isotopic labels in untargeted metabolomics. *Analytical chemistry* **2014**, *86*, 1632-1639.

(319) Shrimpton, R.; Victora, C. G.; de Onis, M.; Lima, R. C.; Blössner, M.; Clugston, G. Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics* **2001**, *107*, e75-e75.

(320) Prentice, A. M.; Ward, K. A.; Goldberg, G. R.; Jarjou, L. M.; Moore, S. E.; Fulford, A. J.; Prentice, A. Critical windows for nutritional interventions against stunting–. *The American of Clinical Nutrition* **2013**, *97*, 911-918.

(321) Ammerman, A. S.; Lindquist, C. H.; Lohr, K. N.; Hersey, J. The efficacy of behavioral interventions to modify dietary fat and fruit and vegetable intake: a review of the evidence. *Preventive medicine* **2002**, *35*, 25-41.

(322) Baldwin, C.; Parsons, T.; Logan, S. Dietary advice for illness-related malnutrition in adults. *Cochrane Database of Systematic Reviews* **2007**.

(323) Lozano, I.; Van der Werf, R.; Bietiger, W.; Seyfritz, E.; Peronet, C.; Pinget, M.; Jeandidier, N.; Maillard, E.; Marchioni, E.; Sigrist, S. High-fructose and high-fat diet-induced disorders in rats: impact on diabetes risk, hepatic and vascular complications. *Nutrition & metabolism* **2016**, *13*, 15.

(324) Wong, S. K.; Chin, K.-Y.; Suhaimi, F. H.; Ahmad, F.; Jamil, N. A.; Ima-Nirwana, S. Osteoporosis is associated with metabolic syndrome induced by high-carbohydrate high-fat diet in a rat model. *Biomedicine & pharmacotherapy* **2018**, *98*, 191-200.

(325) Yamamoto, T.; Takabatake, Y.; Takahashi, A.; Kimura, T.; Namba, T.; Matsuda, J.; Minami, S.; Kaimori, J.-y.; Matsui, I.; Matsusaka, T. High-Fat Diet–Induced Lysosomal Dysfunction and Impaired Autophagic Flux Contribute to Lipotoxicity in the Kidney. *Journal of the American Society of Nephrology* **2017**, *28*, 1534-1551.

(326) Chaar, L. J.; Coelho, A.; Silva, N. M.; Festuccia, W. L.; Antunes, V. R. High-fat diet-induced hypertension and autonomic imbalance are associated with an upregulation of CART in the dorsomedial hypothalamus of mice. *Physiological reports* **2016**, *4*.

(327) O'Neill, A. M.; Burrington, C. M.; Gillaspie, E. A.; Lynch, D. T.; Horsman, M. J.; Greene, M. W. High-fat Western diet–induced obesity contributes to increased tumor growth in mouse models of human colon cancer. *Nutrition Research* **2016**, *36*, 1325-1334.

(328) Moon, H.; Ruelcke, J. E.; Choi, E.; Sharpe, L. J.; Nassar, Z. D.; Bielefeldt-Ohmann, H.; Parat, M.-O.; Shah, A.; Francois, M.; Inder, K. L. Diet-induced hypercholesterolemia promotes androgen-independent prostate cancer metastasis via IQGAP1 and caveolin-1. *Oncotarget* **2015**, *6*, 7438.

(329) Organization, W. H.: *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*; World Health Organization, 2013.

(330) Olazarán, J.; Reisberg, B.; Clare, L.; Cruz, I.; Peña-Casanova, J.; Del Ser, T.; Woods, B.; Beck, C.; Auer, S.; Lai, C. Nonpharmacological therapies in Alzheimer's disease: a systematic review of efficacy. *Dementia and geriatric cognitive disorders* **2010**, *30*, 161-178.

(331) Newman, S.; Steed, L.; Mulligan, K. Self-management interventions for chronic illness. *The Lancet* **2004**, *364*, 1523-1537.

(332) Forouhi, N. G.; Misra, A.; Mohan, V.; Taylor, R.; Yancy, W. Dietary and nutritional approaches for prevention and management of type 2 diabetes. *BMJ* **2018**, *361*, k2234.

(333) Dutta, T.; Chai, H. S.; Ward, L. E.; Ghosh, A.; Persson, X.-M. T.; Ford, G. C.; Kudva, Y. C.; Sun, Z.; Asmann, Y. W.; Kocher, J.-P. A. Concordance of changes in metabolic pathways based on plasma metabolomics and skeletal muscle transcriptomics in type 1 diabetes. *Diabetes* **2012**, DB_110874.

(334) Suthahar, N.; Meijers, W. C.; Brouwers, F. P.; Heerspink, H. J.; Gansevoort, R. T.; van der Harst, P.; Bakker, S. J.; de Boer, R. A. Heart failure and inflammation-related biomarkers as predictors of new-onset diabetes in the general population. *International journal of cardiology* **2018**, *250*, 188-194.

(335) Gedela, S.; Rao, A. A.; Medicherla, N. R. Identification of Biomarkers for Type 2 Diabetes and Its Complications: A Bioinformatic Approach. *International journal of biomedical science: IJBS* **2007**, *3*, 229.

(336) Maranhão, P. A.; de Souza, M. d. G. C.; Panazzolo, D. G.; Nogueira Neto, J. F.; Bouskela, E.; Kraemer-Aguiar, L. G. Metabolic Changes Induced by High-Fat Meal Evoke Different Microvascular Responses in Accordance with Adiposity Status. *BioMed research international* **2018**, *2018*.

(337) Stitt, A. W.; Jenkins, A. J.; Cooper, M. E. Advanced glycation end products and diabetic complications. *Expert opinion on investigational drugs* **2002**, *11*, 1205-1223.

(338) Moreno-Fernández, S.; Garcés-Rimón, M.; Vera, G.; Astier, J.; Landrier, J.; Miguel, M. High Fat/High Glucose Diet Induces Metabolic Syndrome in an Experimental Rat Model. *Nutrients* **2018**, *10*, 1502.

(339) DeFronzo, R. A.; Tripathy, D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. *Diabetes care* **2009**, *32*, S157-S163.

(340) Weir, J. M.; Wong, G.; Barlow, C. K.; Greeve, M. A.; Kowalczyk, A.; Almasy, L.; Comuzzie, A. G.; Mahaney, M. C.; Jowett, J. B.; Shaw, J. Plasma lipid profiling in a large population based cohort. *Journal of lipid research* **2013**, jlr. P035808.

(341) Naples, S. P.; Borengasser, S. J.; Rector, R. S.; Uptergrove, G. M.; Morris, E. M.; Mikus, C. R.; Koch, L. G.; Britton, S. L.; Ibdah, J. A.; Thyfault, J. P. Skeletal muscle mitochondrial and metabolic responses to a high-fat diet in female rats bred for high and low aerobic capacity. *Applied physiology, nutrition, and metabolism* **2010**, *35*, 151-162.

(342) Simpson, G. L.; Ortwerth, B. The non-oxidative degradation of ascorbic acid at physiological conditions. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2000**, *1501*, 12-24.

(343) Uribarri, J.; Woodruff, S.; Goodman, S.; Cai, W.; Chen, X.; Pyzik, R.; Yong, A.; Striker, G. E.; Vlassara, H. Advanced glycation end products in foods and a practical guide to their reduction in the diet. *Journal of the American Dietetic Association* **2010**, *110*, 911-916. e912.

(344) Singh, V. P.; Bali, A.; Singh, N.; Jaggi, A. S. Advanced glycation end products and diabetic complications. *The Korean Journal of Physiology & Pharmacology* **2014**, *18*, 1-14.

(345) Ramasamy, R.; Yan, S. F.; D'Agati, V.; Schmidt, A. M. Receptor for Advanced Glycation Endproducts (RAGE): a formidable force in the pathogenesis of the cardiovascular complications of diabetes & aging. *Current molecular medicine* **2007**, *7*, 699-710.

(346) Li, C.; Culver, S. A.; Quadri, S.; Ledford, K. J.; Al-Share, Q. Y.; Ghadieh, H. E.; Najjar, S. M.; Siragy, H. M. High-Fat Diet Amplifies Renal Renin Angiotensin System Expression, Blood Pressure Elevation and Renal Dysfunction Caused by Ceacam1 Null Deletion. *American Journal of Physiology-Heart and Circulatory Physiology* **2015**.

(347) Gaster, M.; Nehlin, J. O.; Minet, A. D. Impaired TCA cycle flux in mitochondria in skeletal muscle from type 2 diabetic subjects: marker or maker of the diabetic phenotype? *Archives of physiology and biochemistry* **2012**, *118*, 156-189.

(348) Chung, H.-H.; Lee, K. S.; Cheng, J.-T. Decrease of obesity by allantoin via imidazoline I1-receptor activation in high fat diet-fed mice. *Evidence-Based Complementary and Alternative Medicine* **2013**, *2013*.

(349) Bowman, R. H. Effects of diabetes, fatty acids, and ketone bodies on tricarboxylic acid cycle metabolism in the perfused rat heart. *Journal of biological Chemistry* **1966**, *241*, 3041-3048.

(350) Mu, C.; Yang, Y.; Luo, Z.; Zhu, W. Metabolomic analysis reveals distinct profiles in the plasma and urine of rats fed a high-protein diet. *Amino acids* **2015**, *47*, 1225-1238.

(351) Schrauwen, P.; Hesselink, M. Reduced tricarboxylic acid cycle flux in type 2 diabetes mellitus? *Diabetologia* **2008**, *51*, 1694.

(352) Mogensen, M.; Sahlin, K.; Fernström, M.; Glintborg, D.; Vind, B. F.; Beck-Nielsen, H.; Højlund, K. Mitochondrial respiration is decreased in skeletal muscle of patients with type 2 diabetes. *Diabetes* **2007**, *56*, 1592-1599.

(353) Lunkes, G. I.; Lunkes, D.; Stefanello, F.; Morsch, A.; Morsch, V. M.; Mazzanti, C. M.; Schetinger, M. R. C. Enzymes that hydrolyze adenine nucleotides in diabetes and associated pathologies. *Thrombosis research* **2003**, *109*, 189-194.

(354) Okada, Y.; Sakaue, H.; Nagare, T.; Kasuga, M. Diet-induced up-regulation of gene expression in adipocytes without changes in DNA methylation. *Kobe J Med Sci* **2009**, *54*, E241-E249.

(355) Eid, H. M.; Arnesen, H.; Hjerkinn, E. M.; Lyberg, T.; Ellingsen, I.; Seljeflot, I. Effect of diet and omega-3 fatty acid intervention on asymmetric dimethylarginine. *Nutrition & metabolism* **2006**, *3*, 4.

(356) Tam, Z. Y.; Ng, S. P.; Tan, L. Q.; Lin, C.-H.; Rothenbacher, D.; Klenk, J.; Boehm, B. O. Metabolite profiling in identifying metabolic biomarkers in older people with late-onset type 2 diabetes mellitus. *Scientific reports* **2017**, *7*, 4392.

(357) Li, Z.-Y.; Ding, L.-L.; Li, J.-M.; Xu, B.-L.; Yang, L.; Bi, K.-S.; Wang, Z.-T. 1H-NMR and MS based metabolomics study of the intervention effect of curcumin on hyperlipidemia mice induced by high-fat diet. *Plos one* **2015**, *10*, e0120950.

(358) Schwartzbaum, J. A.; Fisher, J. L.; Aldape, K. D.; Wrensch, M. Epidemiology and molecular pathology of glioma. *Nature Reviews Neurology* **2006**, *2*, 494.

(359) Reynolds, B. A.; Weiss, S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *science* **1992**, *255*, 1707-1710.

(360) Ostrom, Q. T.; Gittleman, H.; Liao, P.; Vecchione-Koval, T.; Wolinsky, Y.; Kruchko, C.; Barnholtz-Sloan, J. S. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014. *Neuro-oncology* **2017**, *19*, v1-v88.

(361) Louis, D. N.; Holland, E. C.; Cairncross, J. G. Glioma classification: a molecular reappraisal. *The American journal of pathology* **2001**, *159*, 779-786.

(362) Wesseling, P.: Classification of Gliomas. In *Emerging Concepts in Neuro-Oncology*; Springer, 2013; pp 3-20.

(363) Kortmann, R.-D.; Paulino, A. C.: Low-Grade Glioma. In *Radiation Oncology for Pediatric CNS Tumors*; Springer, 2018; pp 189-211.

(364) Network, C. G. A. R. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *New England Journal of Medicine* **2015**, *372*, 2481-2498.

(365) Ostrom, Q. T.; Kinnersley, B.; Wrensch, M. R.; Eckel-Passow, J. E.; Armstrong, G.; Rice, T.; Chen, Y.; Wiencke, J. K.; McCoy, L. S.; Hansen, H. M. Sex-specific glioma genome-wide association study identifies new risk locus at 3p21. 31 in females, and finds sex-differences in risk at 8q24. 21. *Scientific reports* **2018**, *8*, 7352.

(366) Hepworth, S. J.; Schoemaker, M. J.; Muir, K. R.; Swerdlow, A. J.; van Tongeren, M. J.; McKinney, P. A. Mobile phone use and risk of glioma in adults: case-control study. *Bmj* **2006**, *332*, 883-887.

(367) Wrensch, M.; Yost, M.; Miike, R.; Lee, G.; Touchstone, J. Adult glioma in relation to residential power frequency electromagnetic field exposures in the San Francisco Bay area. *Epidemiology* **1999**, 523-527.

(368) Hardell, L.; Carlberg, M.; Söderqvist, F.; Hansson Mild, K. Meta-analysis of long-term mobile phone use and the association with brain tumours. *International journal of oncology* **2008**, *32*, 1097-1103.

(369) Marinovich, M.; Galli, C. L.; Bosetti, C.; Gallus, S.; La Vecchia, C. Aspartame, low-calorie sweeteners and disease: regulatory safety and epidemiological issues. *Food and chemical toxicology* **2013**, *60*, 109-115.

(370) Corell, A.; Carstam, L.; Smits, A.; Henriksson, R.; Jakola, A. Age and surgical outcome of low-grade glioma in Sweden. *Acta Neurologica Scandinavica* **2018**.

(371) Zhang, J.; Wu, G.; Miller, C. P.; Tatevossian, R. G.; Dalton, J. D.; Tang, B.; Orisme, W.; Punchihewa, C.; Parker, M.; Qaddoumi, I. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. *Nature genetics* **2013**, *45*, 602.

(372) van Nieuwenhuizen, D.; Douw, L.; Klein, M.; Peerdeman, S. M.; Heimans, J. J.; Reijneveld, J. C.; Stam, C. J.; Hillebrand, A. Cognitive functioning and functional brain networks in postoperative WHO grade I meningioma patients. *Journal of neuro-oncology* **2018**, 1-9.

(373) Santos-Pinheiro, F.; Park, M.; Liu, D.; Kwong, L. N.; Cruz, S.; Levine, N. B.; O'Brien, B. J.; Chen, M. Seizure burden pre-and postresection of low-grade gliomas as a predictor of tumor progression in low-grade gliomas. *Neuro-Oncology Practice* **2018**.

(374) Lahti, J.; Saunamäki, T.; Salo, J.; Niemelä, A.; Jehkonen, M. Cognitive Impairment and Recovery in Meningiomas and Low-Grade Gliomas. *Journal of Behavioral and Brain Science* **2018**, *8*, 473.

(375) Borja, M. J.; Plaza, M. J.; Altman, N.; Saigal, G. Conventional and advanced MRI features of pediatric intracranial tumors: supratentorial tumors. *American Journal of Roentgenology* **2013**, *200*, W483-W503.

(376) Fan, G.; Deng, Q.; Wu, Z.; Guo, Q. Usefulness of diffusion/perfusion-weighted MRI in patients with non-enhancing supratentorial brain gliomas: a valuable tool to predict tumour grading? *The British journal of radiology* **2006**, *79*, 652-658.

(377) Huang, W.; Muo, C.; Lin, C.; Jen, Y.; Yang, M.; Lin, J.; Sung, F.; Kao, C. Paediatric head CT scan and subsequent risk of malignancy and benign brain tumour: a nation-wide population-based cohort study. *British journal of cancer* **2014**, *110*, 2354.

(378) Brat, D. J.; Scheithauer, B. W.; Fuller, G. N.; Tihan, T. Newly codified glial neoplasms of the 2007 WHO Classification of Tumours of the Central Nervous System: angiocentric glioma, pilomyxoid astrocytoma and pituicytoma. *Brain pathology* **2007**, *17*, 319-324.

(379) Pouratian, N.; Asthagiri, A.; Jagannathan, J.; Shaffrey, M. E.; Schiff, D. Surgery Insight: the role of surgery in the management of low-grade gliomas. *Nature Reviews Neurology* **2007**, *3*, 628.

(380) Aarsen, F. K.; Van Dongen, H.; Paquier, P.; Van Mourik, M.; Catsman-Berrevoets, C. E. Long-term sequelae in children after cerebellar astrocytoma surgery. *Neurology* **2004**, *62*, 1311-1316.

(381) Jakola, A.; Skjulsvik, A.; Myrmel, K.; Sjåvik, K.; Unsgård, G.; Torp, S.; Aaberg, K.; Berg, T.; Dai, H.; Johnsen, K. Surgical resection versus watchful waiting in low-grade gliomas. *Annals of Oncology* **2017**, *28*, 1942-1948.

(382) McGirt, M. J.; Chaichana, K. L.; Attenello, F. J.; Weingart, J. D.; Than, K.; Burger, P. C.; Olivi, A.; Brem, H.; Quinoñes-Hinojosa, A. Extent of surgical resection is independently associated with survival in patients with hemispheric infiltrating low-grade gliomas. *Neurosurgery* **2008**, *63*, 700-708.

(383) McGirt, M. J.; Goldstein, I. M.; Chaichana, K. L.; Tobias, M. E.; Kothbauer, K. F.; Jallo, G. I. Extent of surgical resection of malignant astrocytomas of the spinal cord: outcome analysis of 35 patients. *Neurosurgery* **2008**, *63*, 55-61.

(384) Sanai, N.; Berger, M. S. Glioma extent of resection and its impact on patient outcome. *Neurosurgery* **2008**, *62*, 753-766.

(385) Duffau, H. Is non-awake surgery for supratentorial adult low-grade glioma treatment still feasible? *Neurosurgical review* **2018**, *41*, 133-139.

(386) Karim, A. B.; Maat, B.; Hatlevoll, R.; Menten, J.; Rutten, E. H.; Thomas, D. G.; Mascarenhas, F.; Horiot, J. C.; Parvinen, L. M.; van Reijn, M. A randomized trial on dose-response in radiation therapy of low-grade cerebral glioma: European Organization for Research and Treatment of Cancer (EORTC) Study 22844. *International Journal of Radiation Oncology* Biology* Physics* **1996**, *36*, 549-556.

(387) Shaw, E.; Arusell, R.; Scheithauer, B.; O'fallon, J.; O'neill, B.; Dinapoli, R.; Nelson, D.; Earle, J.; Jones, C.; Cascino, T. Prospective randomized trial of low-versus high-dose radiation therapy in adults with supratentorial low-grade glioma: initial report of a North Central Cancer Treatment Group/Radiation Therapy Oncology Group/Eastern Cooperative Oncology Group study. *Journal of Clinical Oncology* **2002**, *20*, 2267-2276.

(388) Douw, L.; Klein, M.; Fagel, S. S.; van den Heuvel, J.; Taphoorn, M. J.; Aaronson, N. K.; Postma, T. J.; Vandertop, W. P.; Mooij, J. J.; Boerman, R. H. Cognitive and radiological effects of radiotherapy in patients with low-grade glioma: long-term follow-up. *The Lancet Neurology* **2009**, *8*, 810-818.

(389) Houillier, C.; Wang, X.; Kaloshi, G.; Mokhtari, K.; Guillevin, R.; Laffaire, J.; Paris, S.; Boisselier, B.; Idbaih, A.; Laigle-Donadey, F. IDH1 or IDH2 mutations predict longer survival and response to temozolomide in low-grade gliomas. *Neurology* **2010**, *75*, 1560-1566.

(390) Hoang-Xuan, K.; Capelle, L.; Kujas, M.; Taillibert, S.; Duffau, H.; Lejeune, J.; Polivka, M.; Criniere, E.; Marie, Y.; Mokhtari, K. Temozolomide as initial treatment for adults with low-grade oligodendrogliomas or oligoastrocytomas and correlation with chromosome 1p deletions. *Journal of clinical oncology* **2004**, *22*, 3133-3138.

(391) Sherman, J. H.; Moldovan, K.; Yeoh, H. K.; Starke, R. M.; Pouratian, N.; Shaffrey, M. E.; Schiff, D. Impact of temozolomide chemotherapy on seizure frequency in patients with low-grade gliomas. *Journal of neurosurgery* **2011**, *114*, 1617-1621.

(392) Quinn, J. A.; Reardon, D. A.; Friedman, A. H.; Rich, J. N.; Sampson, J. H.; Provenzale, J. M.; McLendon, R. E.; Gururangan, S.; Bigner, D. D.; Herndon, J. E. Phase II trial of temozolomide in patients with progressive low-grade glioma. *Journal of clinical oncology* **2003**, *21*, 646-651.

(393) Kaloshi, G.; Benouaich-Amiel, A.; Diakite, F.; Taillibert, S.; Lejeune, J.; Laigle-Donadey, F.; Renard, M.-A.; Iraqi, W.; Idbaih, A.; Paris, S. Temozolomide for low-grade gliomas Predictive impact of 1p/19q loss on response and outcome. *Neurology* **2007**, *68*, 1831-1836.

(394) Shaw, E. G.; Wang, M.; Coons, S. W.; Brachman, D. G.; Buckner, J. C.; Stelzer, K. J.; Barger, G. R.; Brown, P. D.; Gilbert, M. R.; Mehta, M. P. Randomized trial of radiation therapy plus procarbazine, lomustine, and vincristine chemotherapy for supratentorial adult low-grade glioma: initial results of RTOG 9802. *Journal of Clinical Oncology* **2012**, *30*, 3065.

(395) Stupp, R.; Mason, W. P.; Van Den Bent, M. J.; Weller, M.; Fisher, B.; Taphoorn, M. J.; Belanger, K.; Brandes, A. A.; Marosi, C.; Bogdahn, U. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine* **2005**, *352*, 987-996.

(396) Li, Z.; Yin, Y.; Liu, F. Recent developments in predictive biomarkers of pediatric glioma. *Oncology letters* **2017**, *14*, 497-500.

(397) Hochberg, F. H.; Atai, N. A.; Gonda, D.; Hughes, M. S.; Mawejje, B.; Balaj, L.; Carter, R. S. Glioma diagnostics and biomarkers: an ongoing challenge in the field of medicine and science. *Expert review of molecular diagnostics* **2014**, *14*, 439-452.

(398) Libby, C. J.; Tran, A. N.; Scott, S. E.; Griguer, C.; Hjelmeland, A. B. The pro-tumorigenic effects of metabolic alterations in glioblastoma including brain tumor initiating cells. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **2018**.

(399) Molenaar, R. J.; Maciejewski, J. P.; Wilmink, J. W.; Noorden, C. J. Wild-type and mutated IDH1/2 enzymes and therapy responses. *Oncogene* **2018**, 1.

(400) Reisch, A. S.; Elpeleg, O. Biochemical assays for mitochondrial activity: assays of TCA cycle enzymes and PDHc. *Methods in cell biology* **2007**, *80*, 199-222.

(401) Mustafa, D. A.; Swagemakers, S. M.; Buise, L.; van der Spek, P. J.; Kros, J. M. Metabolic alterations due to IDH1 mutation in glioma: opening for therapeutic opportunities? *Acta neuropathologica communications* **2014**, *2*, 6.

(402) Cohen, A. L.; Holmen, S. L.; Colman, H. IDH1 and IDH2 mutations in gliomas. *Current neurology and neuroscience reports* **2013**, *13*, 345.

(403) Chang, E. F.; Smith, J. S.; Chang, S. M.; Lamborn, K. R.; Prados, M. D.; Butowski, N.; Barbaro, N. M.; Parsa, A. T.; Berger, M. S.; Mcdermott, M. M. Preoperative prognostic classification system for hemispheric low-grade gliomas in adults. **2008**.

(404) Andronesi, O. C.: Precision oncology in the era of radiogenomics: the case of D-2HG as an imaging biomarker for mutant IDH gliomas. Oxford University Press US, 2018.

(405) Saito, T.; Sugiyama, K.; Hama, S.; Yamasaki, F.; Takayasu, T.; Nosaka, R.; Muragaki, Y.; Kawamata, T.; Kurisu, K. Prognostic importance of temozolomide-induced neutropenia in glioblastoma, IDH-wildtype patients. *Neurosurgical review* **2018**, *41*, 621-628.

(406) Grimmer, M. R.: Abstract LB-376: Clinico-genomic analysis of temozolomide-induced hypermutation in both molecular subtypes of IDH-mutant glioma. AACR, 2018.

(407) Abe, H.; Natsumeda, M.; Kanemaru, Y.; Watanabe, J.; Tsukamoto, Y.; Okada, M.; Yoshimura, J.; Oishi, M.; Fujii, Y. MGMT Expression Contributes to Temozolomide Resistance in H3K27M-Mutant Diffuse Midline Gliomas and MGMT Silencing to Temozolomide Sensitivity in IDH-Mutant Gliomas. *Neurologia medico-chirurgica* **2018**, *58*, 290-295.

(408) Nakamura, M.; Watanabe, T.; Yonekawa, Y.; Kleihues, P.; Ohgaki, H. Promoter methylation of the DNA repair gene MGMT in astrocytomas is frequently associated with G: C→ A: T mutations of the TP53 tumor suppressor gene. *Carcinogenesis* **2001**, *22*, 1715-1719.

(409) Gerson, S. L. MGMT: its role in cancer aetiology and cancer therapeutics. *Nature Reviews Cancer* **2004**, *4*, 296.

(410) Wick, W.; Meisner, C.; Hentschel, B.; Platten, M.; Schilling, A.; Wiestler, B.; Sabel, M. C.; Koeppen, S.; Ketter, R.; Weiler, M. Prognostic or predictive value of MGMT promoter methylation in gliomas depends on IDH1 mutation. *Neurology* **2013**, *81*, 1515-1522.

(411) Kyritsis, A. P.; Bondy, M. L.; Xiao, M.; Berman, E. L.; Cunningham, J. E.; Lee, P. S.; Levin, V. A.; Saya, H. Germline p53 gene mutations in subsets of glioma patients. *JNCI: Journal of the National Cancer Institute* **1994**, *86*, 344-349.

(412) Pfister, S.; Janzarik, W. G.; Remke, M.; Ernst, A.; Werft, W.; Becker, N.; Toedt, G.; Wittmann, A.; Kratz, C.; Olbrich, H. BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. *The Journal of clinical investigation* **2008**, *118*, 1739-1749.

(413) Wiencke, J. K.; Zheng, S.; Jelluma, N.; Tihan, T.; Vandenberg, S.; Tamgüney, T.; Baumber, R.; Parsons, R.; Lamborn, K. R.; Berger, M. S. Methylation of the PTEN promoter defines low-grade gliomas and secondary glioblastoma. *Neuro-oncology* **2007**, *9*, 271-279.

(414) Zysman, M. A.; Chapman, W. B.; Bapat, B. Considerations when analyzing the methylation status of PTEN tumor suppressor gene. *The American journal of pathology* **2002**, *160*, 795-800.

(415) Lavanya, C.; Chetan, G.; Sibin, M.; Jeru, M. M.; Srinivas, B. M. Promoter methylation of PTEN gene in high grade gliomas. *Molecular cytogenetics* **2014**, *7*, P10.

(416) Hatzivassiliou, G.; Song, K.; Yen, I.; Brandhuber, B. J.; Anderson, D. J.; Alvarado, R.; Ludlam, M. J.; Stokoe, D.; Gloor, S. L.; Vigers, G. RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* **2010**, *464*, 431.

(417) Lo, H.-W. Targeting Ras-RAF-ERK and its interactive pathways as a novel therapy for malignant gliomas. *Current cancer drug targets* **2010**, *10*, 840-848.

(418) Horbinski, C.; Nikiforova, M. N.; Hagenkord, J. M.; Hamilton, R. L.; Pollack, I. F. Interplay among BRAF, p16, p53, and MIB1 in pediatric low-grade gliomas. *Neuro-oncology* **2012**, *14*, 777-789.

(419) Dougherty, M. J.; Santi, M.; Brose, M. S.; Ma, C.; Resnick, A. C.; Sievert, A. J.; Storm, P. B.; Biegel, J. A. Activating mutations in BRAF characterize a spectrum of pediatric low-grade gliomas. *Neuro-oncology* **2010**, *12*, 621-630.

(420) Schreck, K. C.; Guajardo, A.; Lin, D. D.; Eberhart, C. G.; Grossman, S. A. Concurrent BRAF/MEK Inhibitors in BRAF V600–Mutant High-Grade Primary Brain Tumors. *Journal of the National Comprehensive Cancer Network* **2018**, *16*, 343-347.

(421) Hu, N.; Richards, R.; Jensen, R. Role of chromosomal 1p/19q co-deletion on the prognosis of oligodendrogliomas: A systematic review and meta-analysis. *Interdisciplinary Neurosurgery* **2016**, *5*, 58-63.

(422) Boots-Sprenger, S. H.; Sijben, A.; Rijntjes, J.; Tops, B. B.; Idema, A. J.; Rivera, A. L.; Bleeker, F. E.; Gijtenbeek, A. M.; Diefes, K.; Heathcock, L. Significance of complete 1p/19q co-deletion, IDH1 mutation and MGMT promoter methylation in gliomas: use with caution. *Modern Pathology* **2013**, *26*, 922.

(423) Izquierdo, C.; Alentorn, A.; Idbaih, A.; Simó, M.; Kaloshi, G.; Ricard, D.; Barritault, M.; Meyronet, D.; Bruna, J.; Honnorat, J. Long-term impact of temozolomide on 1p/19q-codeleted low-grade glioma growth kinetics. *Journal of neuro-oncology* **2018**, *136*, 533-539.

(424) Mörén, L.; Bergenheim, A. T.; Ghasimi, S.; Brännström, T.; Johansson, M.; Antti, H. Metabolomic screening of tumor tissue and serum in glioma patients reveals diagnostic and prognostic information. *Metabolites* **2015**, *5*, 502-520.

(425) Görke, R.; Meyer-Bäse, A.; Wagner, D.; He, H.; Emmett, M. R.; Conrad, C. A. Determining and interpreting correlations in lipidomic networks found in glioblastoma cells. *BMC systems biology* **2010**, *4*, 126.

(426) Hiratsuka, M.; Inoue, T.; Toda, T.; Kimura, N.; Shirayoshi, Y.; Kamitani, H.; Watanabe, T.; Ohama, E.; Tahimic, C. G.; Kurimasa, A. Proteomics-based identification of differentially expressed genes in human gliomas: down-regulation of SIRT2 gene. *Biochemical and biophysical research communications* **2003**, *309*, 558-566.

(427) Frieboes, H. B.; Lowengrub, J. S.; Wise, S.; Zheng, X.; Macklin, P.; Bearer, E. L.; Cristini, V. Computer simulation of glioma growth and morphology. *Neuroimage* **2007**, *37*, S59-S70.

(428) Gerlinger, M.; Rowan, A. J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England journal of medicine* **2012**, *366*, 883-892.

(429) Sottoriva, A.; Spiteri, I.; Piccirillo, S. G.; Touloumis, A.; Collins, V. P.; Marioni, J. C.; Curtis, C.; Watts, C.; Tavaré, S. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences* **2013**, *110*, 4009-4014.

(430) Cusnir, M.; Cavalcante, L. Inter-tumor heterogeneity. *Human vaccines & immunotherapeutics* **2012**, *8*, 1143-1145.

(431) Smith, S. J.; Diksin, M.; Chhaya, S.; Sairam, S.; Estevez-Cebrero, M. A.; Rahman, R. The invasive region of glioblastoma defined by 5ALA guided surgery has an altered cancer stem cell marker profile compared to central tumour. *International journal of molecular sciences* **2017**, *18*, 2452.

(432) Glas, M.; Rath, B. H.; Simon, M.; Reinartz, R.; Schramme, A.; Trageser, D.; Eisenreich, R.; Leinhaas, A.; Keller, M.; Schildhaus, H. U. Residual tumor cells are unique cellular targets in glioblastoma. *Annals of neurology* **2010**, *68*, 264-269.

(433) Chinnaiyan, P.; Kensicki, E.; Bloom, G.; Prabhu, A.; Sarcar, B.; Kahali, S.; Eschrich, S.; Qu, X.; Forsyth, P.; Gillies, R. The metabolomic signature of malignant glioma reflects accelerated anabolic metabolism. *Cancer research* **2012**.

(434) Hensley, C. T.; Faubert, B.; Yuan, Q.; Lev-Cohain, N.; Jin, E.; Kim, J.; Jiang, L.; Ko, B.; Skelton, R.; Loudat, L. Metabolic heterogeneity in human lung tumors. *Cell* **2016**, *164*, 681-694.

(435) Okegawa, T.; Morimoto, M.; Nishizawa, S.; Kitazawa, S.; Honda, K.; Araki, H.; Tamura, T.; Ando, A.; Satomi, Y.; Nutahara, K. Intratumor heterogeneity in primary kidney cancer revealed by metabolic profiling of multiple spatially separated samples within tumors. *EBioMedicine* **2017**, *19*, 31-38.

(436) Li, P.; Beck, W. D.; Callahan, P. M.; Terry Jr, A. V.; Bartlett, M. G. Quantitation of cotinine and its metabolites in rat plasma and brain tissue by hydrophilic interaction chromatography tandem mass spectrometry (HILIC–MS/MS). *Journal of Chromatography B* **2012**, *907*, 117-125.

(437) Inoue, K.; Miyazaki, Y.; Unno, K.; Min, J. Z.; Todoroki, K.; Toyo'oka, T. Stable isotope dilution HILIC-MS/MS method for accurate quantification of glutamic acid, glutamine, pyroglutamic acid, GABA and theanine in mouse brain tissues. *Biomedical Chromatography* **2016**, *30*, 55-61.

(438) Marie, S. K. N.; Shinjo, S. M. O. Metabolism and brain cancer. *Clinics* **2011**, *66*, 33-43.

(439) Bélanger, M.; Allaman, I.; Magistretti, P. J. Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. *Cell metabolism* **2011**, *14*, 724-738.

(440) Malthankar, G. V.; White, B. K.; Bhushan, A.; Daniels, C. K.; Rodnick, K. J.; Lai, J. C. Differential lowering by manganese treatment of activities of glycolytic and tricarboxylic acid (TCA) cycle enzymes investigated in neuroblastoma and astrocytoma cells is associated with manganese-induced cell death. *Neurochemical research* **2004**, *29*, 709-717.

(441) Sanzey, M.; Rahim, S. A. A.; Oudin, A.; Dirkse, A.; Kaoma, T.; Vallar, L.; Herold-Mende, C.; Bjerkvig, R.; Golebiewska, A.; Niclou, S. P. Comprehensive analysis of glycolytic enzymes as therapeutic targets in the treatment of glioblastoma. *PloS one* **2015**, *10*, e0123544.

(442) Liu, Z. G.; Ding, J.; Du, C.; Xu, N.; Wang, E. L.; Li, J. Y.; Wang, Y. Y.; Yu, J. M. Phosphoglycerate mutase 1 is highly expressed in C6 glioma cells and human astrocytoma. *Oncology letters* **2018**, *15*, 8935-8940.

(443) Yang, H.; Wu, J.; Guo, R.; Peng, Y.; Zheng, W.; Liu, D.; Song, Z. Glycolysis in energy metabolism during seizures. *Neural regeneration research* **2013**, *8*, 1316.

(444) Lu, V. M.; McDonald, K. L. Isocitrate dehydrogenase 1 mutation subtypes at site 132 and their translational potential in glioma. *CNS oncology* **2018**, *7*, 41-50.

(445) Ren, J.-G.; Seth, P.; Ye, H.; Guo, K.; Hanai, J.-i.; Husain, Z.; Sukhatme, V. P. Citrate suppresses tumor growth in multiple models through inhibition of glycolysis, the tricarboxylic acid cycle and the IGF-1R pathway. *Scientific reports* **2017**, *7*, 4537.

(446) Bosnyák, E.; Kamson, D. O.; Behen, M. E.; Barger, G. R.; Mittal, S.; Juhász, C. Imaging cerebral tryptophan metabolism in brain tumor-associated depression. *EJNMMI research* **2015**, *5*, 56.

(447) Kamson, D. O.; Juhász, C.; Buth, A.; Kupsky, W. J.; Barger, G. R.; Chakraborty, P. K.; Muzik, O.; Mittal, S. Tryptophan PET in pretreatment delineation of newly-diagnosed gliomas: MRI and histopathologic correlates. *Journal of neuro-oncology* **2013**, *112*, 121-132.

(448) Kamson, D. O.; Mittal, S.; Robinette, N. L.; Muzik, O.; Kupsky, W. J.; Barger, G. R.; Juhász, C. Increased tryptophan uptake on PET has strong independent prognostic value in patients with a previously treated high-grade glioma. *Neuro-oncology* **2014**, *16*, 1373-1383.

(449) Sahm, F.; Oezen, I.; Opitz, C. A.; Radlwimmer, B.; von Deimling, A.; Ahrendt, T.; Adams, S.; Bode, H. B.; Guillemin, G. J.; Wick, W. The endogenous tryptophan metabolite and NAD+ precursor quinolinic acid confers resistance of gliomas to oxidative stress. *Cancer research* **2013**.

(450) Chung, W. J.; Lyons, S. A.; Nelson, G. M.; Hamza, H.; Gladson, C. L.; Gillespie, G. Y.; Sontheimer, H. Inhibition of cystine uptake disrupts the growth of primary brain tumors. *Journal of Neuroscience* **2005**, *25*, 7101-7110.

(451) Limón-Pacheco, J. H.; Hernández, N. A.; Fanjul-Moles, M. L.; Gonsebatt, M. E. Glutathione depletion activates mitogen-activated protein kinase (MAPK) pathways that display organ-specific responses and brain protection in mice. *Free Radical Biology and Medicine* **2007**, *43*, 1335-1347.

(452) Huang, J.; Weinstein, S. J.; Kitahara, C. M.; Karoly, E. D.; Sampson, J. N.; Albanes, D. A prospective study of serum metabolites and glioma risk. *Oncotarget* **2017**, *8*, 70366.

(453) Taymaz-Nikerel, H.; De Mey, M.; Ras, C.; ten Pierick, A.; Seifar, R. M.; Van Dam, J. C.; Heijnen, J. J.; van Gulik, W. M. Development and application of a differential method for reliable metabolome analysis in Escherichia coli. *Analytical biochemistry* **2009**, *386*, 9-19.

(454) Karkos, P.; Leong, S.; Karkos, C.; Sivaji, N.; Assimakopoulos, D. Spirulina in clinical practice: evidence-based human applications. *Evidence-based complementary and alternative medicine* **2011**, *2011*.

(455) Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature chemical biology* **2009**, *5*, 593.

(456) Ong, I. M.; Glasner, J. D.; Page, D. Modelling regulatory pathways in E. coli from time series expression profiles. *Bioinformatics* **2002**, *18*, S241-S248.

(457) Wegener, G.; Kellermann, M. Y.; Elvert, M. Tracking activity and function of microorganisms by stable isotope probing of membrane lipids. *Current opinion in biotechnology* **2016**, *41*, 43-52.

(458) Rattray, J. E.; Geenevasen, J. A.; Van Niftrik, L.; Rijpstra, W. I. C.; Hopmans, E. C.; Strous, M.; Schouten, S.; Jetten, M. S.; Sinninghe Damsté, J. S. Carbon isotope-labelling experiments indicate that ladderane lipids of anammox bacteria are synthesized by a previously undescribed, novel pathway. *FEMS microbiology letters* **2009**, *292*, 115-122.

Appendices

Appendix A1. Authentic standard metabolites

No	Authentic standard name	Formula	Monoisotopic mass	pHilic Ret time (mins)	Concentration (mM)
1	(R)-2-Hydroxyglutarate	C5H8O5	148.0372	8.11	100
2	(R)-Acetoin	C4H8O2	88.0524	12.49	10
3	(R)-Lactate	C3H6O3	90.0317	7.39	100
4	(S)-Malate	C4H6O5	134.0215	8.46	80
5	1,3-Diaminopropane	C3H10N2	74.0844	13.34	100
6	1-Aminocyclopropane-1-carboxylate	C4H7NO2	101.0477	13.96	100
7	1-Aminopropan-2-ol	C3H9NO	75.0684	15.51	100
8	1-Butanol	C4H10O	74.0732	15.37	10
9	1H-Imidazole-4-ethanamine	C5H9N3	111.0796	11.07	10
10	1-Naphthylacetic acid	C12H10O2	186.0681	6.06	5
11	2,3-Bisphospho-D-glycerate	C3H8O10P2	265.9593	15.27	10
12	2-Aminobutan-4-olide	C4H7NO2	101.0477	15.99	10
13	2-Deoxy-D-glucose	C6H12O5	164.0685	9.79	10
14	2-Ethylhexyl phthalate	C16H22O4	278.1518	5.74	10
15	2-hydroxyethyldisulfide	C4H10O2S2	154.0122	6.62	100
16	2-Methylcitrate	C7H10O7	206.0427	9.69	100
17	2-Oxoadipate	C6H8O5	160.0372	7.39	100
18	2-Oxobutanoate	C4H6O3	102.0317	6.53	100
19	2-Oxoglutarate	C5H6O5	146.0215	8.11	100
20	3-(3-Hydroxy-phenyl)-propanoic acid	C9H10O3	166.0630	7.14	10
21	3-(4-Hydroxyphenyl)pyruvate	C9H8O4	180.0423	6.64	10
22	3,4-Dihydroxymandelate	C8H8O5	184.0372	9.61	100
23	3,4-Dihydroxyphenylacetate	C8H8O4	168.0423	7.39	100
24	3',5'-Cyclic AMP	C10H12N5O6P	329.0525	11.39	10
25	3',5'-Cyclic GMP	C10H12N5O7P	345.0474	10.72	10
26	3alpha,12alpha-Dihydroxy-5beta-cholan-24-oic Acid	C24H40O4	392.2927	5.97	10
27	3alpha,7alpha-Dihydroxy-5beta-cholan-24-oic Acid	C24H40O4	392.2927	5.99	10
28	3-Hydroxyphenylacetate	C8H8O3	152.0473	6.53	10
29	3-Methoxy-4-hydroxymandelate	C9H10O5	198.0528	7.16	10
30	3-Methoxy-4-Hydroxyphenylglycolsulfate	C9H12O7S	264.0304	7.58	10
31	3-Methoxytyramine	C9H13NO2	167.0946	7.86	10
32	3-Methyl-2-oxobutanoic acid	C5H8O3	116.0474	6.19	100
33	3-Phospho-D-glycerate	C3H7O7P	185.9929	9.14	10
34	4-Aminobenzoate	C7H7NO2	165.0790	6.65	10
35	4-Aminobutanoate	C4H9NO2	103.0633	23.47	100
36	4-Aminohippuricacid	C9H10N2O3	194.0691	7.51	100
37	4-Coumarate	C9H8O3	164.0473	6.30	100
38	4-Hydroxybutanoic acid	C4H8O3	104.0473	14.22	10
39	4-Hydroxyphenylacetate	C8H8O3	152.0473	7.51	100
40	4-Trimethylammoniobutanoate	C7H15NO2	145.1103	14.31	100
41	5-Aminolevulinate	C5H9NO3	131.0582	15.44	100
42	5-Hydroxyindoleacetate	C10H9NO3	191.0582	8.73	10

43	5-Methoxytryptamine	C11H14N2O	190.1106	12.53	100
44	5-Methoxytryptamine	C11H14N2O	190.1106	12.68	50
45	5'-Methylthioadenosine	C11H15N5O3S	297.0896	9.75	10
46	5-Oxoproline	C5H7NO3	129.0426	7.82	100
47	6-Phospho-D-gluconate	C6H13O10P	276.0246	11.24	10
48	Acetaldehyde	C2H4O	44.0262	8.47	100
49	Acetyl-CoA	C23H38N7O17P3S	809.1258	12.48	10
50	Acetylcysteine	C5H9NO3S	163.0303	6.79	100
51	Adenine	C5H5N5	135.0545	13.32	10
52	Adenosine	C10H13N5O4	267.0968	11.07	10
53	ADP	C10H15N5O10P2	427.0294	14.94	10
54	Agmatine	C5H14N4	130.1218	15.75	100
55	Allantoin	C4H6N4O3	158.0440	11.12	10
56	Alpha-Hydroxyisobutyricacid	C4H8O3	104.0473	6.74	100
57	AMP	C10H14N5O7P	347.0631	12.70	10
58	Ascorbate	C6H8O6	176.0321	10.30	10
59	ATP	C10H16N5O13P3	506.9958	15.33	10
60	Benzoate	C7H6O2	122.0368	6.17	10
61	Betaine	C5H11NO2	117.0790	12.69	100
62	Biopterin	C9H11N5O3	237.0862	10.64	100
63	Biotin	C10H16N2O3S	244.0882	22.48	10
64	Cadaverine	C5H14N2	102.1157	16.42	100
65	CDP	C9H15N3O11P2	403.0182	6.24	10
66	Choline phosphate	C5H14NO4P	183.0660	15.27	100
67	citrate	C6H8O7	192.0270	11.34	10
68	CMP	C9H14N3O8P	323.0519	14.11	10
69	CoA	C21H36N7O16P3S	767.1152	6.72	10
70	Creatine	C4H9N3O2	131.0695	14.32	100
71	Creatinine	C4H7N3O	113.0589	15.09	100
72	CTP	C9H16N3O14P3	482.9845	13.57	10
73	Cytidine	C9H13N3O5	243.0855	15.36	100
74	Cytosine	C4H5N3O	111.0433	15.67	10
75	dAMP	C10H14N5O6P	331.0682	12.49	10
76	D-Arabinose	C5H10O5	150.0528	10.84	100
77	Deoxyadenosine	C10H13N5O3	251.1018	10.86	10
78	Deoxycarnitine	C7H16NO2	146.1181	6.38	10
79	Deoxyguanosine	C10H13N5O4	267.0968	10.24	10
80	Deoxyribose	C5H10O4	134.0579	12.14	100
81	Deoxyuridine	C9H12N2O5	228.0746	7.99	10
82	D-Erythrose	C4H8O4	120.0423	10.14	100
83	D-Erythrose 4-phosphate	C4H9O7P	200.0086	11.79	100
84	D-Fructose	C6H12O6	180.0634	10.82	100
85	D-Fructose 1,6-bisphosphate	C6H14O12P2	339.9961	13.92	10
86	D-Fructose 6-phosphate	C6H13O9P	260.0297	12.53	100

87	D-Galactarate	C6H10O8	210.0376	12.46	10
88	D-Galactose	C6H12O6	180.0634	11.79	100
89	D-Galacturonate	C6H10O7	194.0427	11.95	100
90	D-Gluconic acid	C6H12O7	196.0583	11.66	100
91	D-Glucono-1,4-lactone	C12H20O2	178.0477	19.34	100
92	D-Glucosamine	C6H13NO5	179.0794	17.23	100
93	D-Glucosamine 6-phosphate	C6H14NO8P	259.0457	15.29	100
94	D-Glucose	C6H12O6	180.0634	11.55	100
95	D-glucose 6-phosphate	C6H13O9P	260.0297	12.70	100
96	D-Glycerate	C3H6O4	106.0266	9.22	100
97	dGMP	C10H14N5O7P	347.0631	12.16	100
98	dIMP	C10H13N4O7P	332.0522	11.21	10
99	DL-Glyceraldehyde 3-phosphate	C3H7O6P	169.9980	12.70	10
100	Dopamine	C9H11NO4	197.0688	10.12	100
101	D-Ribose	C5H10O5	150.0528	9.96	100
102	D-Ribose 5-phosphate	C5H11O8P	230.0192	11.91	100
103	D-Threose	C4H8O4	120.0423	9.99	100
104	dUMP	C9H13N2O8P	308.0410	10.32	100
105	Ethanolamine phosphate	C2H8NO4P	141.0191	14.02	100
106	FAD	C27H33N9O15P2	785.1571	12.37	10
107	FMN	C17H21N4O9P	456.1046	9.76	10
108	Folate	C19H19N7O6	441.1397	9.56	5
109	Fumarate	C4H4O4	116.0110	6.84	100
110	Gallate	C7H6O5	170.0215	9.36	10
111	GDP	C10H15N5O11P2	443.0243	5.38	10
112	Glutarate	C5H8O4	132.0423	6.76	100
113	Glutathione	C10H17N3O6S	307.0838	12.59	100
114	Glutathione disulfide	C20H32N6O12S2	612.1520	14.99	10
115	Glyceraldehyde	C3H6O3	90.0317	5.38	100
116	Glycerol	C3H8O3	92.0473	7.73	100
117	Glycine	C2H5NO2	75.0320	14.40	100
118	Glycolate	C2H4O3	76.0160	8.16	100
119	Glycylglycine	C4H8N2O3	132.0535	15.89	100
120	Glyoxylate	C2H2O3	74.0004	10.14	100
121	GMP	C10H14N5O8P	363.0580	12.59	10
122	GTP	C10H16N5O14P3	522.9907	10.47	10
123	Guanidine	CH5N3	59.0484	13.57	10
124	Guanidinoacetate	C3H7N3O2	117.0538	15.15	10
125	Guanine	C5H5N5O	151.0494	12.50	10
126	Guanosine	C10H13N5O5	283.0917	10.86	10
127	HEPES	C8H18N2O4S	238.0987	15.98	10
128	Hexanoylcarnitine	C13H25NO4	259.1784	11.03	100
129	Hippurate	C9H9NO3	179.0582	6.48	100
130	Homocystine	C8H16N2O4S2	268.0551	15.61	100

131	homovanillate	C9H10O4	182.0579	6.40	10
132	Hypoxanthine	C5H4N4O	136.0385	9.58	10
133	IDP	C10H14N4O11P2	428.0134	6.23	10
134	Imidazole	C3H4N2	68.0375	11.27	100
135	Imidazole-4-acetate	C5H6N2O2	126.0429	14.38	100
136	IMP	C10H13N4O8P	348.0471	11.53	10
137	Indoxylsulfate	C8H7NO4S	213.0096	17.38	100
138	Inosine	C10H12N4O5	268.0808	9.59	10
139	isobutanal	C4H8O	72.0575	8.23	50
140	Isonicotinic acid	C6H5NO2	123.0320	8.64	10
141	Itaconate	C5H6O4	130.0266	6.94	100
142	L-2,3-Diaminopropanoate	C3H8N2O2	104.0586	18.57	100
143	L-2,4-Diaminobutanoate	C4H10N2O2	118.0742	18.46	100
144	L-2-Aminoadipate	C6H11NO4	161.0688	17.21	100
145	L-2-Aminoadipate	C6H11NO4	161.0688	13.27	100
146	L-Adrenaline	C9H13NO3	183.0895	15.89	100
147	L-Alanine	C3H7NO2	89.0477	13.94	100
148	L-Arginine	C6H14N4O2	174.1117	18.98	100
149	L-Asparagine	C4H8N2O3	132.0535	14.05	100
150	L-Aspartate	C4H7NO4	133.0375	13.63	100
151	L-Carnitine	C7H15NO3	161.1052	14.77	100
152	L-Citrulline	C6H13N3O3	175.0957	14.39	100
153	L-Cystathionine	C7H14N2O4S	222.0674	15.83	10
154	L-Cysteate	C3H7NO5S	169.0045	12.06	100
155	L-Cysteine	C3H7NO2S	121.0198	13.00	100
156	L-Cystine	C6H12N2O4S2	240.0239	15.45	10
157	L-Dehydroascorbate	C6H6O6	174.0164	10.20	100
158	Leucinicacid	C6H12O3	132.0786	6.27	100
159	L-Glutamate	C5H9NO4	147.0532	13.53	100
160	L-Glutamine	C5H10N2O3	146.0691	13.91	100
161	L-Gulono-1,4-lactone	C6H10O6	178.0477	7.03	100
162	L-Histidine	C6H9N3O2	155.0695	18.35	100
163	L-Homocysteine	C4H9NO2S	135.0354	12.86	10
164	L-Homoserine	C4H9NO3	119.0582	13.90	100
165	Lipoamide	C8H15NOS2	205.0595	6.26	10
166	Lipoate	C8H14O2S2	206.0435	5.97	10
167	L-Isoleucine	C6H13NO2	131.0946	12.18	100
168	L-Kynurenine	C10H12N2O3	208.0848	11.62	10
169	L-Leucine	C6H13NO2	131.0946	11.84	100
170	L-Lysine	C6H15N2O2	146.1055	13.36	100
171	L-Metanephrine	C10H15NO3	197.1052	13.90	10
172	L-Methionine	C5H11NO2S	149.0510	12.10	100
173	L-Methionine sulfone	C4H9NO5S	181.0409	13.37	50
174	L-Noradrenaline	C8H11NO3	169.0739	16.97	100

175	L-Norleucine	C6H13NO2	131.0946	11.70	100
176	L-Normetanephrine	C9H13NO3	183.0895	14.60	100
177	L-Ornithine	C5H12N2O2	132.0899	18.93	100
178	L-Phenylalanine	C9H11NO2	165.0790	14.60	100
179	L-Proline	C5H9NO2	115.0633	12.67	100
180	L-Rhamnose	C6H12O5	164.0685	10.92	100
181	L-Serine	C3H7NO3	105.0426	14.22	100
182	L-Threonine	C4H9NO3	119.0582	18.64	100
183	L-Tryptophan	C11H12N2O2	204.0899	12.18	10
184	L-Tyrosine	C9H11NO3	181.0739	12.74	10
185	L-Valine	C5H11NO2	117.0790	12.77	10
186	Maleic acid	C4H4O4	116.0110	6.20	100
187	Malonate	C3H4O4	104.0110	7.78	100
188	Mannitol	C6H14O6	182.0790	11.26	100
189	Melatonin	C13H16N2O2	232.1212	6.28	5
190	Menadione	C11H8O2	172.0524	6.02	10
191	Mercaptoethanol	C2H6OS	78.0139	8.84	100
192	Mesaconate	C5H6O4	130.0266	6.54	100
193	meso-2,6-Diaminoheptanedioate	C7H14N2O4	190.0954	16.50	10
194	methylglyoxal	C3H4O2	72.0211	11.37	100
195	Methylguanidine	C2H7N3	73.0640	15.13	100
196	Methylmalonate	C4H6O4	118.0266	6.86	100
197	MOPS	C7H15NO4S	209.0722	11.31	100
198	myo-Inositol	C6H12O6	180.0634	5.35	100
199	N(pi)-Methyl-L-histidine	C7H11N3O2	169.0851	18.57	100
200	N6-Acetyl-L-lysine	C8H16N2O3	188.1161	13.07	100
201	N6-Methyl-L-lysine	C7H16N2O2	160.1212	13.13	100
202	N-Acetyl-D-glucosamine	C8H15NO6	221.0899	10.09	100
203	N-Acetylglutamine	C7H12N2O4	188.0797	8.51	10
204	N-Acetyl-L-glutamate	C7H11NO5	189.0637	7.58	10
205	N-Acetyl-L-phenylalanine	C11H13NO3	207.0895	6.31	100
206	N-Acetylneuraminate	C11H19NO9	309.1060	10.96	100
207	N-Acetylornithine	C7H14N2O3	174.1004	15.04	10
208	N-Acetylputrescine	C6H14N2O	130.1106	7.48	100
209	NAD+	C21H27N7O14P2	663.1091	4.28	50
210	NADH	C21H29N7O14P2	665.1248	12.73	10
211	NADP+	C21H28N7O17P3	743.0755	7.51	10
212	NADPH	C21H30N7O17P3	745.0911	5.33	20
213	Ne,Ne dimethyllysine	C8H18N2O2	174.1368	19.00	10
214	Nicotinamide	C6H6N2O	122.0480	8.75	100
215	O-Acetylcarnitine	C9H17NO4	203.1158	13.29	100
216	O-Acetyl-L-serine	C5H9NO4	147.0532	11.96	10
217	O-Butanoylcarnitine	C11H21NO4	231.1471	9.27	100
218	Octadecanoic acid	C18H36O2	284.2715	5.74	100

219	O-decanoyl-R-carnitine	C17H33NO4	315.2410	9.70	10
220	Orotate	C5H4N2O4	156.0171	8.73	10
221	Orotidine	C10H12N2O8	288.0594	9.68	10
222	Oxalate	C2H2O4	89.9953	7.75	50
223	Oxaloacetate	C4H4O5	132.0059	7.70	100
224	Pantothenate	C9H17NO5	219.1107	7.04	10
225	Phenolsulfonphthalein	C19H14O5S	354.0562	5.44	100
226	Phenylacetylglycine	C10H11NO3	193.0739	8.34	10
227	Phenylhydrazine	C6H8N2	108.0687	8.83	100
228	Phenylpyruvate	C9H8O3	164.0473	6.08	100
229	Phosphate	O4P	97.9769	17.42	100
230	Phosphoenolpyruvate	C3H5O6P	167.9824	12.46	100
231	Picolinic acid	C6H5NO2	123.0320	8.65	100
232	Putrescine	C4H12N2	88.1000	16.72	100
233	pyrazinoate	C5H4N2O2	124.0273	8.84	10
234	Pyridoxal	C8H9NO3	167.0582	13.77	100
235	Pyridoxal phosphate	C8H10NO6P	247.0246	11.63	10
236	Pyridoxine	C8H11NO3	169.0739	13.76	100
237	Pyrophosphate	O7P2	177.9432	6.32	100
238	Pyruvate	C3H4O3	88.0161	7.00	100
239	S-Adenosyl-L-homocysteine	C14H20N6O5S	384.1216	15.50	10
240	S-Adenosylmethioninamine	C14H22N6O3S	354.1474	17.39	10
241	Salicyluric acid	C9H9NO4	195.0532	6.38	50
242	Sarcosine	C3H7NO2	89.0477	4.52	100
243	Selenomethionine	C5H11NO2Se	196.9955	12.12	100
244	Serotonin	C10H12N2O	176.0950	14.97	10
245	sn-Glycero-3-Phosphocholine	C8H20NO6P	257.1028	13.93	10
246	sn-Glycerol 3-phosphate	C3H9O6P	172.0137	11.49	10
247	Sorbitol 6-phosphate	C6H15O9P	262.0454	6.37	100
248	Spermidine	C7H19N3	145.1579	14.11	100
249	Spermine	C10H26N4	202.2158	17.47	100
250	Succinate	C4H6O4	118.0266	7.26	100
251	Sucrose	C12H22O11	342.1162	11.71	100
252	Taurine	C2H7NO3S	125.0147	12.77	10
253	Taurocholate	C26H45NO7S	515.2917	5.25	10
254	Thiamin diphosphate	C12H18N4O7P2S	424.0371	21.19	100
255	Thiamine	C12H17N4OS	265.1123	6.37	100
256	Thymidine	C10H14N2O5	242.0903	7.39	100
257	Thymine	C5H6N2O2	126.0429	7.39	10
258	trans-Aconitate	C6H6O6	174.0164	8.15	100
259	UDP	C9H14N2O12P2	404.0022	16.20	100
260	UDP-Glucose	C15H24N2O17P2	566.0550	4.24	100
261	UMP	C9H13N2O9P	324.0359	10.97	10
262	Uracil	C4H4N2O2	112.0273	9.68	5

263	Urate	C5H4N4O3	168.0283	10.72	100
264	Urea	CH4N2O	60.0324	8.82	100
265	Uridine	C9H12N2O6	244.0695	9.01	10
266	UTP	C9H15N2O15P3	483.9685	13.42	100
267	Xanthine	C5H4N4O2	152.0334	8.99	10
268	Xanthosine	C10H12N4O6	284.0757	7.37	50

Appendix A2. Comparison between Human Biofluid (Blood and Urine) metabolome with Bacterial metabolome (E. coli and Spirulina)

Metabolite	Molecular Mass	Retention time (minutes)	Presence in Human Blood	Presence in Human Urine	Presence in Spirulina	Presence in E. coli
(S)-Malate	134.0215	10.8400	+	+	+	+
1-Methyladenosine	281.1122	11.2613	+	+	-	+
2-Aminoadipate	161.0688	16.7158	+	+	+	+
2-Ethylhexyl phthalate	278.1517	3.6959	-	+	+	+
2-hydroxy-4-methylthiobutanoate	150.0350	6.2140	-	+	+	+
2-Methylcitrate	206.0426	19.1039	+	+	-	+
2-Oxoadipate	160.0370	16.5016	+	+	-	+
2-Oxobutanoic acid	102.0317	17.3729	+	+	+	+
2-Oxoglutaric acid	146.0215	17.2320	+	+	+	+
3-(3-Hydroxy-phenyl)-propanoic acid	166.0630	6.8257	+	+	+	+
3-(4-Hydroxyphenyl)lactate	182.0580	9.7322	+	+	+	+
3-(4-Hydroxyphenyl)pyruvate	180.0423	7.9286	+	+	-	+
3-Hydroxy-9-hexadecenoylcarnitine	413.3143	3.8674	+	-	-	+
3-Hydroxyanthranilate	153.0425	9.0597	+	+	-	+
3-Methoxytyramine	167.0946	8.8173	+	+	-	+
3-Methyl-2-oxobutanoic acid	116.0473	16.5395	+	+	-	+
3-Methyladenine	149.0701	10.4729	+	+	-	+
3-Phospho-D-glycerate	185.9929	18.5314	+	-	+	+
4-Aminobutanoate	103.0633	16.8776	+	+	-	+
4-Guanidinobutanoate	145.0853	16.5743	+	+	+	+
4-Hydroxybenzoate	138.0317	11.6575	+	+	-	+
4-Hydroxybutanoic acid	104.0474	8.4710	+	+	-	+
4-Hydroxyphenylacetylglycine	209.0688	9.9614	+	+	-	+
5,6-Dihydrouracil	114.0430	7.8103	+	+	+	+
5-6-Dihydrouridine	246.0851	9.7186	+	+	+	+
5-Methylcytidine	257.1011	9.2099	+	+	+	+
5-Methylthioadenosine	297.0897	8.3786	+	+	+	+
8-Hydroxy-7-methylguanine	181.0601	13.0409	+	+	-	+
Aconitate	174.0164	11.7800	+	+	+	-
Adenine	135.0545	9.7888	+	+	+	+
Adenosine	267.0968	9.2139	+	+	+	+
ADP	427.0294	17.1089	+	+	-	+
Aminoacetone	73.0528	10.2767	+	+	-	+

AMP	347.0631	15.5966	+	+	+	+
Arginino-succinate	290.1226	21.4819	+	+	-	+
ATP	506.9957	18.7103	+	+	-	+
Betaine	117.0790	12.0710	+	+	+	+
Cadaverine	102.1156	26.0065	+	+	-	+
Choline	104.1075	19.9449	+	+	+	+
Citrate	192.0270	19.7514	+	+	+	+
Creatine	131.0695	10.7900	+	+	+	-
Creatinine	113.0589	8.4300	+	+	+	-
Cytidine	243.0855	12.7438	+	+	+	+
Cytosine	111.0432	14.7501	+	+	+	+
D-Aspartate	133.0375	15.6252	+	+	-	+
Deoxyadenosine	251.1017	7.9441	+	+	+	+
Deoxyguanosine	267.0967	12.3685	+	+	-	+
D-Gluconic acid	196.0583	15.8439	+	+	+	+
D-Glucose	180.0634	14.6522	+	+	-	+
Dihydrothymine	128.0585	16.2892	+	+	-	+
Dimethylglycine	103.0633	13.9787	+	+	-	+
D-Ribose	150.0527	10.5210	+	+	+	+
D-Sorbitol	182.0791	10.5605	+	+	+	+
D-Xylulose	150.0527	10.5210	+	+	-	+
Ergothioneine	229.0885	10.8600	+	+	+	-
FAD	785.1580	13.6304	+	-	-	+
FMN	456.1050	13.6527	+	-	-	+
Fumaric acid	116.0110	17.6885	+	+	-	+
gamma-L-Glutamyl-L-cysteine	250.0625	15.9484	+	+	-	+
Gluconolactone	178.0477	14.4712	+	-	+	+
Glutathione	307.0839	16.1554	+	+	-	+
Glucose 6-phosphate	260.0297	11.1900	+	-	+	-
Glutathione disulfide	612.1520	11.0600	+	-	+	-
Glycerate	106.0266	13.7847	+	+	+	+
Glycine	75.0320	17.1318	+	+	+	+
Guanidinoacetate	117.0539	12.0181	+	+	-	+
Guanine	151.0493	13.1950	-	+	+	+
Guanosine	283.0917	12.4492	+	+	+	+
Hexanoyl-carnitine	259.1784	11.1480	+	+	-	+

Hydroxypyruvic acid	104.0110	17.4567	+	+	-	+
Hypoxanthine	136.0385	10.6418	+	+	+	+
IMP	348.0471	17.1549	+	-	+	+
Inosine	268.0808	10.5413	+	+	-	+
L-2,4-Diaminobutanoate	118.0742	22.7170	+	+	-	+
L-Alanine	89.0477	16.2945	+	+	+	+
L-Arabinose	150.0529	11.6319	+	+	-	+
L-Arginine	174.1117	27.5430	+	+	+	+
L-Asparagine	132.0535	17.7527	+	+	+	+
L-Aspartate	133.0375	16.7606	+	+	+	+
L-Carnitine	161.1051	13.6965	+	+	-	+
L-Citrulline	175.0957	16.8890	+	+	+	+
L-Cysteine	121.0197	15.6688	+	+	-	+
L-Dehydroascorbic acid	174.0163	15.6955	+	+	-	+
L-Glutamate	147.0532	16.4727	+	+	+	+
L-Glutamine	146.0691	16.5357	+	+	+	+
L-Histidine	155.0695	16.5145	+	+	+	+
L-Homocitrulline	189.1113	18.7257	+	+	-	+
L-Homocysteine	135.0354	14.4098	+	+	-	+
L-Leucine	131.0946	11.6630	+	+	+	+
L-Lysine	146.1055	25.9237	+	+	+	+
L-Methionine	149.0511	12.6955	+	+	+	+
L-Ornithine	132.0899	24.3443	+	+	-	+
L-Phenylalanine	165.0790	11.1397	+	+	+	+
L-Proline	115.0633	13.8927	+	+	+	+
L-Serine	105.0426	17.3035	+	+	+	+
L-Threonine	119.0583	15.9681	+	+	+	+
L-Tryptophan	204.0899	12.6388	+	+	+	+
L-Tyrosine	181.0739	14.3075	+	+	+	+
L-Valine	117.0790	13.2160	+	+	-	+
Malate	398.2118	10.8800	+	+	+	-
Maleic acid	116.0110	13.2559	+	+	-	+
Methylimidazoleacetic acid	140.0586	8.0557	+	+	+	+
Methyl-lysine	160.1212	23.6252	+	+	-	+
Methylmalonic acid	118.0266	7.4756	+	+	-	+
myo-Inositol	180.0633	9.2600	+	+	+	-

N1-Acetylspermidine	187.1685	30.0027	+	+	-	+
N-Acetyl-arginine	216.1222	16.2584	+	+	-	+
N-Acetyl-aspartate	175.0481	16.3764	+	+	-	+
N-Acetylcadaverine	144.1262	13.4295	+	+	-	+
N-Acetyl-glutamate	189.0637	15.9383	+	+	-	+
N-Acetylglutamine	188.0797	9.7948	-	+	-	+
N-Acetyl-L-leucine	173.1052	6.9487	-	+	-	+
N-Acetyl-L-phenylalanine	207.0895	7.5848	-	+	-	+
N-Acetyl-ornithine	174.1004	17.2354	+	+	-	+
N-Acetylputrescine	130.1106	21.4011	+	+	+	+
NAD+	664.1169	15.9890	+	-	-	+
NADP+	744.0833	18.3386	+	-	-	+
Neopterin	253.0811	16.2152	+	+	+	+
Nicotinamide	122.0480	7.0582	+	+	+	+
Nicotinate	123.0321	7.7920	+	+	+	+
N-Ribosylnicotinamide	254.0901	25.0572	-	+	-	+
O-Acetyl-L-serine	147.0532	12.1477	+	+	-	+
Orotic acid	156.0171	11.4306	+	+	-	+
Pantothenic acid	219.1107	9.5039	+	+	-	+
Phosphoenolpyruvate	167.9824	18.9824	+	-	+	+
Piperidine	85.0891	11.4086	-	+	-	+
Putrescine	88.1000	27.1472	+	+	-	+
Pyridoxal	167.0582	7.9609	+	-	-	+
Pyridoxamine	168.0897	12.4449	+	-	-	+
Pyridoxine	169.0738	7.8442	+	+	-	+
Pyruvic acid	88.0162	8.6060	+	+	-	+
Quinic acid	192.0634	8.9600	+	-	+	-
Quinolinic acid	167.0219	16.3597	+	+	-	+
S-Adenosyl-methionine	399.1451	17.8846	+	+	-	+
Shikimic acid	174.0528	15.6252	+	+	-	+
Sorbitol 6-phosphate	262.0454	17.6510	+	-	-	+
Spermidine	145.1579	37.4335	+	+	-	+
Succinate	118.0266	16.9126	+	+	+	+
Sulfate	97.9674	19.3844	+	+	+	+
Tetradecanoyl-carnitine	371.3036	3.8460	+	+	-	+
Thymine	126.0430	7.2705	+	+	+	+

Trimethyl-lysine	188.1525	23.4346	-	+	-	+
Tyramine	137.0841	8.6116	+	+	-	+
UMP	324.0359	24.6076	+	-	+	+
Uracil	112.0273	8.4913	+	+	-	+
Urate	168.0283	12.1415	+	+	-	+
Uridine	244.0695	12.8318	+	+	+	+
Xanthine	152.0334	12.4425	+	+	+	+

Appendix A3. Human Urine Study Ethical Documentation

Health Questionnaire

Project:

"**Global metabolite profiling of human urine samples treated with caffeine-rich tea and the application of normalisation methodology on the metabolome data."**

Researcher: Sergey Evseev

Supervisors: Dr Dong-Hyun Kim, Prof Dave Barrett

1. Your sex gender. Please cirlce: Male Female
2. Are you a smoker? Please cirle: YES NO
3. Do you have any medical conditions such as Diabetes, Hypertension or Renal diseases? If yes, please state below:
4. Are you taking any medications (including prescription and OTCs)? If yes, please state below:
5. Have you read the Participant Information Sheet? Please circle: YES NO
6. Do you have any questions? If yes, please state below

Participant Information Sheet

Project:

"**Global metabolite profiling of human urine samples treated with caffeine-rich tea and the application of normalisation methodology on the metabolome data."**

Researcher: Sergey Evseev

Supervisors: Dr Dong-Hyun Kim, Prof Dave Barrett

Study: This study will investigate the effect of caffeine-rich tea ("Earl Grey") on human urine metabolome, collected from healthy individuals by the researcher, and the application of a normalisation method using Uniformly-^{13}C labelled *E.coli* on those samples to improve the accuracy of the data and biological relevance. Participants will be asked to drink "Earl Grey" tea (1 cup) in the early afternoon, their urine samples will then be taken twice during the day, after the consumption of the tea. The samples will be analysed with LC-MS instrument, and the data will be processed as a part of the project for PhD studies. This study will be one of the first to apply this normalisation method on human biofluid samples. The overall aim is to normalise the metabolite data i.e. improve the data accuracy – reduce the Coefficient Variation (CV) values for metabolite peak areas and retention times, and to improve the separation between data groups in Principal Component Analysis (PCA).

Choosing participants: The volunteers will be asked to answer a series of healthcare questions before taking the experiment (health questionnaire), in order to determine whether they are suitable for the study or not. The ones who are successful will be notified. 10 participants are expected in this study (all males).

On the day of the experiment: The experiment will take place in Boots Science Building, common room area (canteen). Before attending the experiment, the chosen participants will be asked to avoid drinking any strong caffeine drinks in the morning and drink glass of water instead, as well as going to the toilet to remove the urine that was affected by overnight. Once the participants arrive, they will be asked to take their urine samples into falcon 50ml tubes, **before** drinking "Earl Grey" tea, around 9am. Few minutes after that, participants will be asked to drink one cup of tea and their urine samples will be collected **twice** into the new tubes. All the samples (tubes) will be passed to researcher who will aliquot them into Eppendorf tubes and put them in a safe location inside the -80 degrees freezer.

Afterwards: Once the study will be over and participant's data will be processed, it will be anonymised and kept secured by password protection in online research folders (University secured Research Drive and "Centre for Analytical Bioscience data back-up website platform "Kingfisher Homes") for seven years after which they will be destroyed. The participant anonymised data may also be used for publishing a research paper. Participants will be allowed to access their own data if they want to, provided the approval of supervisors and researcher.

If you are interested in this study and would like to participate, could you please answer the questionnaire (provided on separate document).

PARTICIPANT CONSENT FORM

Project title: Global metabolite profiling of human urine samples treated with caffeine-rich tea and the application of normalisation methodology on the metabolome data.

Researcher's name: Sergey Evseev

Supervisors: Dr Dong-Hyun Kim and Prof Dave Barrett

- I have read the Participant Information Sheet and the nature and purpose of the research project has been explained to me. I understand and agree to take part.

- I understand the purpose of the research project and my involvement in it.

- I understand that I may withdraw from the research project at any stage

- I understand that while information gained during the study may be published, I will not be identified, and my personal research data will remain confidential and anonymised. I may ask for a copy of my personal data.

- I understand that my personal anonymised research data will be retained for at least seven years from the date of any publication which is based upon them. Data will be stored on electronic Research Drive, as well as on an online data back-up website platform "Kingfisher Homes", both accessible by Researcher and Supervisor (named above). I will also be able to access my own personal research data if I wish so, at any time, by contacting Researcher or Supervisors (named above). After seven years my personal data will be destroyed accordingly (deleted from Research drive and "Kingfisher").

- I understand that I may contact the researcher or supervisor if I require further information about the research, and that I may contact the Chair of School of Pharmacy Research Ethics Committee, if I wish to make a complaint relating to my involvement in the research.

- I understand that if I learn anything unsettling about my health while participating in the project or after accessing my data, it is advisable to discuss this with my GP (but not mandatory).

Signed …… (research participant)

Print name …… **Date** ……

Contact details

Researcher: *Sergey Evseev, Centre for Analytical Biosciences, Division of Advanced Materials and Healthcare Technologies, School of Pharmacy, Boots Science Building, University of Nottingham, NG7 2RD. Email: paxse3@nottingham.ac.uk*

Project Supervisors:

*Dr Dong-Hyun Kim (*Advanced Materials & Healthcare Technologies, School of Pharmacy, University of Nottingham, Tel: +44 (0)115 748 4697, Email: dong-hyun.kim@nottingham.ac.uk); *Prof Dave Barrett (*Advanced Materials & Healthcare Technologies, School of Pharmacy, University of Nottingham, Tel: +44 (0)115 951 5062, Email: David.Barrett@nottingham.ac.uk).

Chair of School of Pharmacy Research Ethics Committee:

Dr Franco Falcone, Associate Professor in Allergy and Infectious Diseases, School of Pharmacy, University of Nottingham, Tel. +44(0)115 84 66073, Email: Franco.Falcone@nottingham.ac.uk

School of Pharmacy Undergraduate and Postgraduate Research Projects Ethical Approval

University Policy requires that all research involving human participants, their data, or their tissues undergoes ethical review by a research ethics committee. The type of review required depends on the associated level of risk (as defined by the ESRC), and whether approved standard procedures are available.

For all projects, please select as applicable (use a separate form for each project). Your answers to the tick list overleaf will determine whether the project meets the ESCR criteria for "minimal risk".

☐ a. This project is minimal risk (all "no" answers in ethics tick list) and standard procedures for this project have been approved by a research ethics committee in the last three years. I have considered the ethical implications of this research, and approve this project.

☒ **b. This project is minimal risk (all "no" answers in ethics tick list) but approved standard procedures for this project are not available. An outline of the project is enclosed for review by the Research Ethics Committee including a description of the procedures to be used.**

☐ c. This project is more than minimal risk (at least one "yes" answer in ethics tick list). An outline of the project is enclosed for review by the Research Ethics Committee, detailing the procedures to be used and addressing the ethical issues raised in the tick list.

☐ d. This project will take place outside the UK.

Supervisor's Name: Dr Dong-Hyun Kim

Project title: Global metabolite profiling of human urine samples treated with caffeine-rich tea and the application of normalisation methodology on the metabolome data.

Supervisor's Signature:  **Date: 03/11/2016**

Please return this form, and additional documents where required, by email to: franco.falcone@nottingham.ac.uk, or via internal mail to Franco Falcone, School of Pharmacy, BSB, University Park.

ETHICS TICK LIST (please answer all questions):

1. Will the study involve participants who are particularly vulnerable or unable to give informed consent (*e.g.* children, people with learning disabilities, individuals in a dependent or unequal relationship, individuals who lack capacity as defined by the Mental Capacity act, 2005)?

   No

2. Will the study require the co-operation of a gatekeeper for initial access to the groups or individuals to be recruited (*e.g.* students at school, members of self-help group, residents of nursing home)?

   No

3. Will it be necessary for participants to take part in the study without their knowledge and consent at the time (*e.g.* covert observation of people in non-public places)?

   No

   Will the study involve discussion of sensitive topics (*e.g.* sexual activity, drug use)?

   No

4. Are drugs, placebos or other substances (*e.g.* food substances, vitamins) to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?

   No

5. Will blood or tissue samples be obtained from participants?

   No (urine samples)

6. Is pain or more than mild discomfort likely to result from the study?

   No

7. Could the study induce psychological stress or anxiety or cause harm or negative consequences beyond the risks encountered in normal life?

   No

8. Will the study involve prolonged or repetitive testing?

   No (one day only)

9. Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?

   No

10. Will the study involve recruitment of patients or staff through the NHS?

    No

11. Will the study involve deceased persons, body parts or other human elements?

    No

12. Will the study involve administrative data or secure data for which ethical approval covering this study has not previously been obtained?

    No

13. Will the study involve access to records of personal or sensitive confidential information, including genetic or other biological information, concerning identifiable individuals?

    No

14. Will the study involve intrusive interventions or data collection methods – for example, the administration of substances, vigorous physical exercise, or techniques such as hypnotism?

    No

15. Will the safety of the researcher be in question?

    No

16. Will the project involve members of the public in a research capacity in research data collection *e.g.* participatory research?

    No

17. Will the project involve respondents through the internet, in particular where visual images are used, and where sensitive issues are discussed?

    No

18. Does the research involve data sharing of confidential information beyond the initial consent given - for example where the research topic or data gathering

involves a risk of information being disclosed that would require the researchers to breach confidentiality conditions agreed with participants?

No

If you have answered “no” to all the questions in the ethics tick list, the project meets the ESRC criteria for “minimal risk”, and can be approved by the project supervisor provided that standard procedures are available that have been approved by an ethics committee.

If approved standard procedures are not available, ethical review by a research ethics committee is required.

If you have answered “yes” to any of the questions in the ethics tick list, the project does not meet the ESRC criteria for minimal risk, and requires ethical approval by a research ethics committee.

Guidance can be found in the University’s Policy for Ethical Review, the ESRC’s framework for research ethics, and the University’s Code of Research Conduct and Research Ethics, all accessible via the research tab on the Pharmacy workspace. For more details, please contact the School Research Ethics Officer (Dr Franco Falcone, x66073, franco.falcone@nottingham.ac.uk).

Appendix A4. Python code generated to run a search for ^{13}C peaks in metabolome data for normalisation

```
#This program will try to use a series of predifined specifications to generate a better C13 model
#!/usr/bin/env python
#Last update 30/08/17
import xlrd
import numpy as np
import re
"""------------EQUATIONS-------------"""
def CV(L): #oefficient of variation for a list
    return(np.std(L)/np.average(L)*100)
def checkzeros(L): #checks if the list contains the value 0, when it appears will return True
    for a in L:
        if a==0:
            return True
    return False
def PPM(mass1,mass2,ppm):     #check if 2 masses are in a PPM range
    nppm=abs((mass1-mass2)/mass1*1000000)
    if nppm<ppm:
        return True
    else:
        return False
"""------------IMPORTING--------------"""

correctingDATA=open("mzMATCHoutput_Control1.txt",'r').readlines()     #data name
correctingDATA.pop(0)
book = xlrd.open_workbook("CV1.xlsx")                                 #excel name
print ("The number of worksheets is", book.nsheets)
sh = book.sheet_by_index(0)
print (sh.name, sh.nrows, sh.ncols)

colums=6                                                              #colum number
headers =[]
core=[]
D={}
out=open("sergey-analisis-important-peak-hight.txt",'w')                          #Output name

"""------------PROCESING--------------"""
for a in range(sh.nrows):
    if a==0:
        #header
        for i in range(sh.ncols-1):
            headers.append(sh.cell_value(rowx=a, colx=i))
    else:
        D[sh.cell_value(rowx=a, colx=0)]=[]
        for i in range(1,colums*2+6):
            D[sh.cell_value(rowx=a, colx=0)].append(float(sh.cell_value(rowx=a, colx=i)))
            #print ("Then we are in ", sh.cell_value(rowx=a, colx=i))

##Check if a metabolite needs repair
good=0
bad=0
improved=0
for key, value in D.items():

    #calculate the label/unlabel this set the start and finish line
    L=value[2:colums+3] #change to change the number of colums
    uL=value[colums+5:colums*2+6]   #change to change the number of colums
    LuL=[]
    #print(key,L,uL,checkzeros(uL))
    for i in range(len(L)-1):
        LuL.append(uL[i]/L[i])
        #print(key,CV(L),CV(LuL))
```

```
if CV(LuL)<CV(L):
    print("good>>>>>>>", key)
    out.write(key+'\t')
    for k in value:
        out.write(str(k)+'\t')
    for k in LuL:
        out.write(str(k)+'\t')
    out.write(str(np.average(LuL))+'\t')
    out.write(str(np.std(LuL))+'\t')
    out.write(str(CV(LuL))+'\t')
    out.write(str(100*(1-CV(LuL)/CV(L)))+'\t')
    out.write('\n')
    good+=1
else:
    bad+=1
    print("bad >>>>>>>", key)
    MASS=value[1]
    RT = value[0]
    Improved=False
    for l in correctingDATA:
        valuesl=[]
        for j in l.split('\t'):
            valuesl.append(float(j))
        MASSl=valuesl.pop(0)
        RTl=valuesl.pop(0)
        if PPM(MASSl,MASS,20) and checkzeros(valuesl)==False:            #MASS LIMIT
            if RTl> RT-9.001 and RTl<RT+9.001:                             #RT LIMIT
                #print("possible match..")
                uL=valuesl
                LuL=[]
                for i in range(len(L)-1):
                    LuL.append(uL[i]/L[i])
                #print(value)
                #print(key,CV(L),CV(LuL))
                if CV(LuL)<CV(L):
                    print("improved", int(100-(CV(LuL)/CV(L))*100), "%")
                    out.write(key+'\t')
                    for k in value[:colums*2]:
                        out.write(str(k)+'\t')
                    for k in uL:
                        out.write(str(k)+'\t')
                    for k in LuL:
                        out.write(str(k)+'\t')
                    out.write(str(np.average(LuL))+'\t')
                    out.write(str(np.std(LuL))+'\t')
                    out.write(str(CV(LuL))+'\t')
                    out.write(str(100*(1-CV(LuL)/CV(L)))+'\t')
                    out.write('\n')
                    Improved=True
                    improved+=1
                    break
                else:
                    pass
                    #print("not improved")

        if Improved==False:
            out.write(key+'\t')
            for k in value:
                out.write(str(k)+'\t')
            for k in LuL:
                out.write(str(k)+'\t')
            #out.write("BAD"+'\t')
            out.write(str(np.average(LuL))+'\t')
            out.write(str(np.std(LuL))+'\t')
            out.write(str(CV(LuL))+'\t')
            out.write(str(100*(1-CV(LuL)/CV(L)))+'\t')
            out.write('\n')

out.close()

print("\n\n\n\n\n>>>The final results show that:")
print("A total of ",good," metabolites were already correct")
print("A total of ",bad," metabolites were bad")
print("From the bad ones a total of ",improved," metabolites have been properly normalized")
print("Leaving a final ", str(bad-improved), " non improved metabolites and ", good+improved,"improved")
print("This means that ", str(int(float(good+improved)/float(bad + good)*100)) ,
    "% of the metabolites have been improved by the methodology")
```

Appendix A5. Comparison between %RSD of non-normalised and normalised metabolites in Human Urine study

Metabolite	Retention time (minutes)	Mass	CONTROL URINE		TREATMENT (TEA) URINE	
			Non-normalised RSD (%)	^{13}C-normalised RSD (%)	Non-normalised RSD (%)	C13 Normalised RSD (%)
N1-Methyl-2-pyridone-5-carboxamide	7.96	152.0585	90.91	20.41	87	30.47
Butanal	5.46	72.0575	44.22	25.37	122.95	132.38
Phe-Tyr	10.5	328.1418	77.23	25.8	81.62	18.14
3-Methylguanine	8.87	165.0651	46.65	27.53	90.74	37.53
L-2,3-Dihydrodipicolinate	6.24	169.0376	100.73	27.72	71.23	32.43
N-Acetyl-D-quinovosamine	10.12	205.0951	57.04	28.67	78.3	35.02
GW 9662	9.59	276.0303	57.62	30.61	143.89	17.47
N2-Acetyl-L-lysine	9.14	188.1161	60.65	31.05	59.56	25.59
N-Feruloylglycine	6.23	251.0793	83.61	31.23	87.1	30.14
3-Sulfolactate	10.34	169.9885	33.13	32.11	109.05	22.72
Glycylproline	8.7	172.0848	46.06	33.54	70.99	16.56
N3-fumaramoyl-L-2,3-diaminopropanoate	9.51	202.0589	54.87	33.99	94.36	22.08
Betaine	8.9	117.0789	68.14	36.82	54.76	42.18
L-Homoserine	9.77	119.0583	33.18	36.93	83.41	45.7
Propanoyl phosphate	8.48	154.003	41.24	37.02	87.48	52.75
trans-4-Hydroxycyclohexanecarboxylate	8.06	144.0786	194.93	37.54	75.03	36.91
2,3,6-Trihydroxypyridine	8.14	127.0269	41.56	38.39	97.88	58.64
Valpromide	12.16	143.131	33.93	38.43	110.52	24.18
[FA methyl,hydroxy,oxo(5:2/4:0)] methyl 4-[2-(2-formyl-vinyl)-3-hydroxy-5-oxo-cyclopentyl]-butanoate	6.23	254.1156	91.09	38.59	108.49	33.49
3''-Deamino-3''-oxonicotianamine	9.66	302.1114	55.62	38.65	94.47	42.64
Diethyl adipate	5.01	202.1206	101.46	38.76	140.48	31.51
(S)-ATPA	9.24	228.111	37.74	38.79	50.83	16.78
N-Acetyl-L-phenylalanine	5.39	207.0896	62.96	38.89	64.66	36.05
Met-Met-Thr-Tyr	8.44	272.1008	90.21	39.11	67.85	48.41
N-D-Glucosylarylamine	8.23	255.1106	87.43	39.44	96.83	26.18
Ferulate	7.69	194.058	90.84	40.31	119.54	50.53
3-Mercaptolactate	8.09	122.0038	66.72	40.86	106.62	70.78
Precocene 2	5.15	220.11	35.45	41.05	92.63	61.36
Pendimethalin	8.11	281.1376	52.77	41.15	157.04	51.14
2-Aminomalonate semialdehyde	8.56	103.027	161.13	41.79	48.3	59.69
Ala-Ser	8.38	176.0797	51.05	42.07	75.4	38.92

Laserpitin	4.79	450.2617	99.99	42.14	118.51	24.11
2,6-Dihydroxypseudooxynicotine	7.71	210.1005	108.37	42.42	67.13	42.2
4-Heptyloxyphenol	5.4	208.1464	34.93	42.46	47.67	53.56
[6]-Gingerol	4.54	294.183	59.96	42.47	95.55	65.19
(S)-2-Aceto-2-hydroxybutanoate	9.13	146.0578	62.07	42.95	72.67	38.9
Europine	7.84	329.1839	55.53	43.63	72.89	49.75
(R)-2,3-Dihydroxy-3-methylbutanoate	6.33	134.0579	68.85	43.65	69.31	23.6
3-mercapto-1,2-propanediol	8.07	108.0245	178.08	44.21	40.86	54.16
N2-(D-1-Carboxyethyl)-L-lysine	9.17	218.1267	93.28	44.34	49.62	42.49
(1-Ribosylimidazole)-4-acetate	8.51	258.0852	138.19	44.98	91.69	38.74
L-Rhamnulose	9.35	164.0685	61	45.05	53.88	25.84
Formyl-N-acetyl-5-methoxykynurenamine	7.61	264.111	73.85	45.09	141.94	51.71
5-Hydroxyferulate	6.23	210.0529	31.75	45.74	64.51	42.95
gamma-L-Glutamylputrescine	12.19	217.1427	44.17	46.11	67.06	28.32
N-Formyl-L-methionine	6.44	177.046	107.98	46.41	91.09	46.28
Cortolone-3-glucuronide	6.19	542.2729	75.12	46.79	78.78	81.66
2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate	8.46	191.0795	70.68	47.29	126.96	55.52
sulfoquinovose	9.96	244.0251	44.37	47.59	57.21	42.87
L-Histidinal	8.12	139.0745	155.77	47.61	62.63	35.6
4-Hydroxy-2-butynal	6.27	84.0211	91.67	47.67	85.75	82.29
5-Guanidino-2-oxopentanoate	10.13	173.0801	52.73	48	262.43	70.01
[FA methyl,hydroxy(5:0)] 3R-methyl-3,5-dihydroxy-pentanoic acid	6.25	148.0736	79.24	48.08	38.59	46.71
Otonecine	5.47	185.1052	88.66	48.25	72.51	42.83
2-Aminomuconate	8.3	157.0375	62.34	48.3	186.83	52.2
Indole-3-acetate	7.88	175.0633	47.73	48.6	71.53	68.86
Glutarylcarnitine	8.41	275.1369	58.83	48.76	133.71	43.27
Deoxycytidine	8	227.0906	27.51	48.83	75.7	72.2
Tetranor-PGF1alpha	5.12	300.1936	73.8	49.22	91.51	35.83
N2-Acetyl-L-aminoadipate	9.23	203.0794	44.05	49.62	98.52	61.62
N-Acetyl-L-histidine	7.88	197.0801	34.74	49.82	66.82	65.99
L-Erythrulose	7.95	120.0423	70.68	49.83	98.46	44.24
Carnosine	9.98	226.1066	30.72	49.99	99.34	56.86
7-Deoxyloganate	6.18	360.142	61.32	50.12	53.85	47.84
Diacetyl	7.68	86.0367	28.89	50.19	38.76	62.28
4-Guanidinobutanoate	10.19	145.0851	51.41	50.66	42.46	55.85
L-N2-(2-Carboxyethyl)arginine	12.63	246.1328	54.91	50.71	65.05	39.97

(S)-Malate	9.71	134.0215	78.89	50.83	42.47	73.56
3-(Dimethylamino)propyl benzoate	7.97	207.1259	62.31	50.96	56.42	49.12
Arene oxide	8.57	94.0419	29.02	51.02	174.87	73.37
Phenylhydrazine	9.32	108.0688	28.67	51.89	62.66	79.6
Hippurate	6.15	179.0583	50.59	52.13	82.3	50.88
DL-2-Aminooctanoicacid	9.27	159.126	49.78	52.19	64.2	67.27
4-Hydroxyphenylacetate	8.01	152.0473	42.05	52.22	148.51	36.61
2,5-Dioxopentanoate	10.19	130.0267	72.03	52.29	42.42	47.97
Xanthine	8.91	152.0334	68.48	52.38	107.13	51.91
trans-3-Hydroxycotinineglucuronide	7.91	368.1222	38.51	52.43	37.54	45.8
Mercaptopyruvate	10.05	119.9881	51.07	52.55	108.33	66.95
Gamma-Aminobutyryl-lysine	12.58	231.1583	63.84	53.19	87.09	38.44
L-Tyrosine	9.69	181.0739	46.88	53.35	63.72	50.96
1-Methylhypoxanthine	8.19	150.0542	44.82	53.49	131.97	61.17
7-Ethoxycoumarin	5.27	190.063	70.05	53.85	93.19	51.59
L-Serine	10.2	105.0426	55.17	53.88	78.76	94.87
Monodehydroascorbate	8.24	175.0242	48.69	54.43	70.45	28.62
2-Oxoglutaramate	8.09	145.0375	216.24	54.64	85.04	53.1
3,4-Dihydroxyphenylacetate	8.47	168.0422	58.98	54.75	84.35	63.82
betaine aldehyde hydrate	12.57	119.0946	167.28	54.76	64.84	40.25
2-Aminoacrylate	8.09	87.032	61.6	54.85	41.79	82.25
Lys-Asp	10.18	261.1325	66.3	55	116.3	26.14
Propanethial S-oxide	6.24	90.014	70.8	55.07	37.02	43.1
L-Carnitine	9.3	161.1052	60.11	55.93	66.42	67.46
Linamarin	8.41	247.1055	87.53	56.01	132.38	40.92
Gln-Pro	13.73	243.122	47.03	56.11	82.96	53.29
Indolylacryloylglycine	6.22	244.0847	66.31	56.26	119.22	90.5
N-Acetyl-D-fucosamine	8.34	205.0951	79.69	56.32	76.15	62.57
Paeonolide	6.27	460.1587	58.07	56.34	103.95	45.78
[PR] (+)-exo-5-Hydroxycamphor	6.23	168.1149	121.23	56.38	65.42	38.65
3,4-Dihydro-7-methoxy-2-methylene-3-oxo-2H-1,4-benzoxazine-5- carboxylic acid	5.53	235.0481	63.16	56.42	54.75	46.57
Ala-Asp-Pro	10.17	301.1275	36.86	56.8	42.07	84.72
methyl pyruvate	5.05	102.0317	42.67	57.06	61.54	54.4
Dihydroclavaminic acid	8.1	200.0798	81.69	57.14	52.19	33.69
Tarennoside	5.6	358.1263	63.73	57.21	60.31	78.39

Calystegin B2	8.51	175.0844	39.27	57.25	139.29	42.87
2-Dehydro-3-deoxy-L-rhamnonate	8.94	162.0528	36.7	57.26	189.29	69.32
D-Glycerate	8.57	106.0266	60.38	57.29	50.19	56.41
beta-D-Digitalopyranose	7.83	178.0841	105.99	57.51	36.82	56.73
[FA (7:1)] 2-heptenal	5.56	112.0888	55.15	57.54	84.92	64.96
4-Hydroxybenzoate	11.49	138.0316	71.67	57.56	52.22	57.24
[PR] Alliacol A	4.56	264.1361	41.52	57.73	120.84	74.86
5,6-Dihydrouracil	8.43	114.0429	97.67	57.88	248.15	40.69
[FA (13:0/2:0)] Tridecanedioic acid	5.53	244.1673	37.02	58.06	64.05	64.05
1-Deoxy-D-xylulose	7.48	134.0579	223.04	58.4	107.82	91.6
3-Phospho-D-glycerate	9.81	185.9929	93.99	58.41	76.32	55.23
Amabiline	7.71	283.1783	48.09	58.5	95.78	38.63
Phenol	5.42	94.0419	46.26	58.73	79.64	54.81
2,3,4-trihydroxybenzoate	7.67	170.0215	95.48	58.82	38.39	67.41
8-Oxocoformycin	8.22	282.0964	46.21	58.88	72.55	49.34
[FA oxo(5:0)] 3-oxo-4-pentenoic acid	6.2	114.0318	141.04	58.94	90.05	64.81
5-Methoxyindoleacetate	7.67	205.0739	54.64	58.96	94.13	41.43
4-Hydroxy-4-methylglutamate	7.96	177.0637	44.02	59.06	122.66	57.31
N-Acetyl-L-aspartate	9.08	175.048	41.31	59.14	103.25	74.07
(4E)-2-Oxohexenoic acid	7.7	128.0473	37.24	59.35	114.28	85.49
5'-Methylthioadenosine	7.76	297.0896	49.19	59.42	60.72	81.09
N3-(4-methoxyfumaroyl)-L-2,3-diaminopropanoate	9.21	216.0746	47.97	59.56	33.99	55.93
N-acetyl -D- glucosaminitol	9.28	223.1055	75.97	59.7	94.92	41.75
(+)-7-Isojasmonic acid	5.52	210.1256	65.13	59.77	60.41	79.84
N(alpha)-Benzyloxycarbonyl-L-leucine	8	265.1314	59.76	59.77	63.33	40.51
coniferaldehyde glucoside	5.52	340.1157	82.42	59.92	140.64	92.88
Taurine	10.12	125.0147	56.46	60.31	64.78	72.82
3-Oxododecanoic acid	4.91	214.1569	65.1	60.32	72.1	45.77
(+/-)-6-Hydroxy-3-oxo-alpha-ionone	4.56	222.1256	23	60.41	44.98	92.75
Val-His	11.65	254.1379	101.32	60.57	38.43	84.97
2-trans,4-cis-Decadienoylcarnitine	7.66	311.2096	61.18	60.63	84.76	74.96
Sulfonylbismethane	7.77	94.0088	83.31	60.7	47.59	58.5
5-Methylthio-D-ribose	6.27	180.0456	34	60.72	62.42	70.37
VAI-2	4.73	227.1522	47.17	61.07	109.93	43.29
Caffeine	6.38	194.0804	48.53	61.48	57.25	52.84
Ala-Ala-Gln	12.81	288.1434	49.22	61.5	56.8	43.85

Methylimidazole acetaldehyde	7.74	124.0637	49.65	61.54	72.21	46.8
Thr-Tyr	5.36	282.1215	63.23	61.77	75.9	85.79
Senecionine	4.91	335.1733	74.98	61.88	66.51	51.63
formyl-isoglutamine	8.28	174.0641	63.29	61.93	45.09	54.82
N6-Methyl-L-lysine	12.87	160.1212	66.55	62.17	59.7	51.2
1-O-Methyl-myo-inositol	8.11	194.0789	54.03	62.36	84.13	58.98
5'-methylthioformycin	10.15	297.0896	45.67	62.42	93.72	56.41
1-(beta-D-Ribofuranosyl)-1,4-dihydronicotinamide	8.35	256.1059	71.83	62.51	111.92	82.53
1-butanesulfonate	8.48	138.0351	65.88	62.61	58.4	36.91
L-Histidine	9.91	155.0695	49.7	62.63	151.42	21.74
Phenylthioacetohydroximate	5.29	167.0405	41.86	62.66	263.5	92.05
Gln-Ser-His	7.95	185.08	75.4	62.76	81.52	58.71
Glycine	10.45	75.032	27.91	62.77	33.54	86.33
[FA (7:0/2:0)] Heptanedioic acid	8.37	160.0735	34.5	63.02	57.54	78.83
(R)-1-Aminopropan-2-ol	8.9	75.0684	47.11	63.03	43.65	61.29
Ptilosteroid B	4.8	430.2023	42.18	63.06	67.03	37.55
D-Glucuronolactone	12.13	176.032	48.94	63.09	88.55	74.78
8-Hydroxy-7-methylguanine	9.13	181.06	129.89	63.3	81.91	58.8
N-(L-Arginino)succinate	9.73	290.1227	49.63	63.33	94.42	30.33
[FA (6:0)] O-hexanoyl-R-carnitine	7.74	259.1783	41.84	63.57	63.02	67.41
L-Tyrosine methyl ester	7.7	195.0896	38.55	63.72	75.08	46.11
[FA oxo,hydroxy(5:1/5:0)] (1S,2R)-3-oxo-2-(4'-hydroxy-2'Z-pentenyl)-cyclopentaneacetic acid	7.7	226.1205	45.32	64	113.85	74.81
[FA (14:0/2:0)] 10E,12E-tetradecadiene-4,6-diynoic acid	5.23	216.1151	125.2	64.05	212.39	57.63
D-Glucuronate	9.86	194.0426	245.29	64.13	63.09	47.11
DL-Methionine sulfone	6.26	181.0411	75.38	64.2	78.12	56.33
3-(4-Hydroxyphenyl)pyruvate	7.7	180.0424	107.95	64.27	50.96	78.98
5-Hydroxyindoleacetate	9.82	191.0582	105.04	64.51	70.25	64.39
Nalpha,Nalpha-Dimethyl-L-histidine	10.07	183.1008	90.51	64.66	73.58	39.3
Indole-3-ethanol	5.23	161.084	116.96	64.72	110	91.5
Taurochenodeoxycholate-3-sulfate	4.66	289.6268	56.89	64.78	66.14	73.48
Bisphenol A dimethacrylate	8.51	364.167	102.43	64.84	25.37	78.59
N-Methyl-L-glutamate	11.1	161.0688	82.91	64.9	128.81	46.72
S-Adenosyl-L-methionine	10.19	398.1376	55.28	64.92	80.57	63.71
7-Aminomethyl-7-carbaguanine	8.16	179.0807	33.4	64.97	50.12	63.03
L-Octanoylcarnitine	6.22	287.2096	54.7	65.05	89.59	114.14

1-Pyrroline-4-hydroxy-2-carboxylate	9.87	129.0426	80.06	65.3	68.7	80.97
(2S)-2-Isopropyl-3-oxosuccinate	6.25	174.0529	44.61	65.31	79.03	54.18
[FA methyl(7:1)] 2-methyl-2E-heptenoic acid	6.25	142.0994	70.08	65.33	48.08	48.68
[PR] (1S,4R)-1-Hydroxy-2-oxolimonene	7.67	168.115	37.52	65.42	57.73	77.83
L-proline amide	9.32	114.0793	43.01	66.04	149.4	84.62
Glycerol	9.52	92.0474	46.59	66.06	92.96	65.4
Taurocyamine	10.65	167.0365	124.23	66.14	104.03	45.64
L-Citrulline	10.2	175.0957	58.99	66.42	68.57	40.56
LL-2,6-Diaminoheptanedioate	9.5	190.0953	83.41	66.42	77.09	63.59
tetrahydrothiophene 1-oxide	7.67	104.0296	84.79	66.5	93.45	47.08
serinol	12.16	91.0633	44.99	66.51	115.57	51.34
sn-Glycerol 3-phosphate	9.23	172.0136	59.96	66.66	68.68	62.98
Lunacrine	4.58	273.1363	138.87	66.71	55	54.31
beta-Cymaropyranose	7.65	162.0891	155.16	66.72	57.51	93.72
D-Sorbitol	9.8	182.0791	71.7	66.74	105.16	52.42
N-Acetyl-L-leucine	5.61	173.1053	54.9	66.82	38.89	34.11
Pyridine-2,3-dicarboxylate	9.5	167.0218	83.88	67.03	76.65	61.18
Gentiopicrin	6.22	356.1108	69.11	67.06	124.54	53.76
[FA oxo,hydroxy(2:0)] 9-oxo-11R,15S-dihydroxy-5Z,13E-prostadienoic acid	4.78	352.2251	61.12	67.13	64	82.47
2,7-Anhydro-alpha-N-acetylneuraminic acid	8.87	291.0955	86.11	67.13	80.45	51.66
Tyr-OEt	6.23	209.1052	46.36	67.42	115.58	90.26
Dopaquinone	7.81	195.0531	39.48	67.53	107.35	78.03
N6,N6,N6-Trimethyl-L-lysine	12.34	188.1525	50.78	67.83	163.82	63.51
Monocrotaline	8.18	325.1526	53.15	67.85	54.43	54.91
Gabaculine	8	139.0633	177.25	67.91	53.19	61.1
Ascorbate 2-sulfate	10.8	255.989	32.3	68.16	109.49	49.13
[FA methyl(12:1)] 2-methyl-2-dodecenoic acid	5.41	212.1776	39.37	68.33	87.6	99.45
D-Proline	8.09	115.0633	143.7	68.38	66.74	58.03
4-Trimethylammoniobutanoate	9.28	145.1102	44.14	68.39	191.38	116.57
3-Butyn-1-ol	8.07	70.0419	73.4	68.42	116.08	66.69
L-Cystathionine	10.04	222.0674	56.48	68.57	71.58	81.1
Stachydrine	8.71	143.0946	133.24	68.68	76.08	88.18
1-thio-β-D-glucose	7.67	196.0406	54.72	68.7	81.94	51.39
Alizarin	7.86	120.0211	39.37	68.86	99.14	47.69
1-7-Dimethylguanosine	8.17	311.123	187.34	68.93	146.68	65.9

1-(3,4-dimethoxylphenyl)ethane-1,2-diol	5.19	198.0892	118.66	69.22	102.46	68.59
(R)-2,3-Dihydroxy-3-methylpentanoate	7.73	148.0736	34.58	69.31	89.93	84.26
5-Methylcytosine	10.38	125.0589	83.03	69.37	59.42	48.36
Leu-Lys-Asp	10.5	187.1082	39.37	69.48	85.06	64.2
Ankorine	6.21	335.2098	54.79	69.55	51.02	49.51
N-Acetyl-L-glutamate 5-semialdehyde	6.28	173.0689	69.24	69.66	49.82	93.56
4-Amino-4-deoxychorismate	7.69	225.0638	54.06	70.08	97.11	65.52
5-Hydroxy-L-tryptophan	8.05	220.0848	129.16	70.25	82.71	64.72
Lys-Val	12.62	245.1739	53.6	70.4	70.51	119.21
myo-Inositol	8.61	180.0634	43.65	70.45	99.03	64.25
Maleimide	7.68	97.0164	41.76	70.51	83.74	54.33
Dodecanedioic acid	5.13	230.1518	104.41	70.58	95.61	103.14
N-Acetyl-D-glucosamine	9.23	221.0901	48.74	70.76	102.34	113.23
N-(3-acetamidopropyl)-4-aminobutanal	11.69	186.1368	47.39	70.79	87.98	96.71
sinapaldehyde glucoside	5.37	370.1263	157.52	70.83	259.88	39.89
Tiglylcarnitine	7.94	243.1471	46.23	70.84	52.43	82.71
Glycylprolylhydroxyproline	6.21	285.1327	52.71	70.99	104.36	92.69
L-2-Amino-3-oxobutanoic acid	7.78	117.0426	50.65	71.23	89.13	40.57
6-Hydroxynicotinate	8.96	139.0269	63.02	71.32	79.03	95.29
indole-3-acetyl-glutamine	6.28	303.1219	63.25	71.53	64.72	55.15
L-Cysteine	8.51	121.0198	34.55	71.58	169.22	42.76
Urate	8.92	168.0283	34.47	71.65	61.07	89.81
N-Carbamoylsarcosine	8.42	132.0535	94	71.71	141.91	89.14
melibionate	9.44	358.111	63.92	71.83	52.55	51.16
Indole	7.71	117.0579	45.27	71.98	48.6	53.32
3-Methoxy-4-hydroxyphenylethyleneglycolsulfate	8.07	264.0303	42.88	72.02	94.87	46.53
3-Oxopropanoate	7.97	88.016	116.72	72.1	58.41	88.19
Methylimidazoleacetic acid	7.84	140.0585	150.8	72.21	85.83	47.39
[PR] (+)-Bornane-2,5-dione	7.66	166.0994	57.69	72.24	56.38	70.25
3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	11.41	204.1474	43.83	72.25	153.03	62.68
Oxalosuccinate	9.96	190.0113	66.71	72.51	56.34	66.88
9-Decenoylcarnitine	6.2	313.2253	52.11	72.55	87.89	69.25
(S)-3-Methyl-2-oxopentanoic acid	5.29	130.063	73.01	72.67	38.79	58.06
Loratadine	4.54	382.1449	112.98	72.73	137.15	52.57
3-Hydroxyanthranilate	7.85	153.0425	44.93	72.77	95.16	71.07
Asp-Asp	9.8	248.0644	79.87	72.83	101.3	43.38

Thymine	8.47	126.0429	50.9	72.87	104.99	89.07
Fagomine	9.17	147.0895	193.08	72.89	40.31	99.75
N1,N8-diacetylspermidine	11.98	229.179	43.58	72.95	20.41	71.3
2-Hexenoylcarnitine	7.79	257.1627	130.38	73.02	94.65	63.23
D-Galactosamine	11.1	179.0793	82.14	73.06	89.78	84.08
L-gamma-Glutamyl-L-hypoglycin	9.12	270.1215	62.6	73.11	90.11	64.61
2,3-dihydroxy-5-methyl-1,4-benzoquinone	5.09	154.0266	72.14	73.38	87.26	64.84
3-Isopropylbut-3-enoic acid	8.15	128.0837	66.33	73.45	44.21	46.02
N-Amidino-L-aspartate	9.67	175.0593	46.27	73.58	92.15	105.86
Nitrilotriacetic acid	9.99	191.043	46.37	73.65	87.66	87.17
Hypoxanthine	8.99	136.0385	141.49	73.79	117.21	66.71
Choline sulfate	9.2	183.0564	52.58	73.85	75.75	88.54
Gly-Ser-His	10.54	299.123	38.88	74.12	74.63	118.37
L-Lysine 1,6-lactam	10.63	128.0949	53.14	74.58	113.8	90.83
Graphinone	4.81	296.1626	107.81	74.63	96.69	41.6
L-Lysine	13.1	146.1055	98.87	74.66	74.58	68.91
2-Octenoylcarnitine	7.68	285.194	47.07	74.71	136.11	69.26
1-5-diazabicyclononane	12.95	126.1157	52.23	74.72	75.64	105.95
trans-Cinnamate	5.08	148.0525	102.23	75.03	81.57	90.25
L-Tyrosine methyl ester 4-sulfate	5.54	275.0464	75.01	75.08	66.71	63.59
2-oxosuberate	8.4	188.0685	46.25	75.15	149.53	90.83
Glu-Ser	9.37	234.0851	153.84	75.16	48.76	69.2
Ala-Ser-Ser	9.79	263.1117	105.37	75.4	68.86	86.72
15-Keto-prostaglandinE2	5.09	350.2092	30.37	75.64	68.93	61.52
Deoxyguanidinoproclavaminic acid	9.7	228.1223	67.84	75.7	148.12	63.82
Cinnamaldehyde	5.3	132.0575	55.23	75.75	105.94	55.84
N-hexanoyl-D,L-homoserine lactone	4.96	199.1209	169.6	75.89	161.79	75.47
THTA	9.25	146.0401	69.72	75.9	96.33	86.18
3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate	5.41	406.2717	38.71	75.97	124.13	62.69
Suberic acid	7.96	174.0891	89.24	76.08	84.93	104.27
N-Methyl-2-pyrrolidinone	7.83	99.0684	137.5	76.14	64.9	53.99
N-Acetyl-D-glucosaminate	10	237.0848	76.09	76.15	70.76	111.86
7-Methylguanine	9.95	165.065	79.01	76.24	93.02	70.98
delta-Guanidinovaleric acid	10.14	159.1007	40.95	76.31	48.83	50.05
3-phosphoglucarate	9.7	290.004	75.39	76.32	32.11	51.77
4-Methylene-L-glutamate	7.69	159.0531	59.26	76.42	129.9	110.93

[FA (8:1)] 2Z-octenoic acid	7.66	142.0994	140.01	76.43	76.75	83.89
Pyridoxine	8.54	169.0738	46.19	76.65	107.93	54.65
N-Butyryl-L-homoserine lactone	5.35	171.0895	70.46	76.71	71.71	93.88
[FA (8:2)] 2,4-octadienal	5.15	124.0888	107.73	76.75	78.83	70.09
L-Leucine	9.04	131.0946	52.34	77.09	74.66	72.72
L-Pipecolate	9.04	129.079	111.43	77.54	66.04	77.2
3-4-DihydroxyphenylglycolO-sulfate	9.21	250.0147	52.54	77.81	75.97	63.85
D-Lysine	9.89	146.1056	68.49	78.12	95.61	74.47
N-Acetyl-D-tryptophan	7.5	246.1005	67.55	78.3	100.66	67.22
[FA dioxo(8:0)] 4,7-dioxo-octanoic acid	8.05	172.0736	34.03	78.35	92.16	53.23
[FA hydroxy(8:0)] 6,8-dihydroxy-octanoic acid	5.31	176.1048	61.46	78.62	122.87	53.62
L-Threonine	8.03	119.0583	97.38	78.76	53.35	41.76
Creatine	10.02	131.0695	148.08	78.78	164.43	44.21
[FA (9:1)] 3-nonenal	5.32	140.1201	44.42	78.83	138.77	76.09
alpha-Curcumene	4.94	202.1722	181.32	79.02	122.85	117.92
(2S,3S)-2-Hydroxytridecane-1,2,3-tricarboxylate	5.05	332.1835	80.52	79.03	59.35	44.26
6-Keto-decanoylcarnitine	6.22	329.2203	45.51	79.03	64.97	93.99
Sulfoacetate	10.56	139.9779	35.67	79.08	60.7	147.87
12-trans-Hydroxy juvenile hormone III	5.04	282.1831	47.87	79.22	74.72	61.53
2,3-dioctanoylglyceramide	6.2	357.2516	50.06	79.39	147.03	50.85
Phenol sulfate	5.41	173.9987	52.84	79.64	105.3	64.7
Tuberonic acid glucoside	5.02	388.1732	96.61	79.77	103.53	62.78
Homocarnosine	9.38	240.1222	87.16	79.8	132.17	70.01
Imidazol-5-yl-pyruvate	8.3	154.0378	72.19	79.97	71.98	21.59
5-Acetylamino-6-amino-3-methyluracil	8.63	198.0753	90.34	80.17	116.81	48.74
Asp-Arg	10.41	289.1387	63.54	80.21	72.83	57.72
2-6dimethylheptanoylcarnitine	6.2	301.2253	60.68	80.45	141.54	75.27
Orotate	8.18	156.0171	76.07	80.51	89.46	54.18
(-)-11-hydroxy-9,10-dihydrojasmonic acid 11-beta-D-glucoside	5.1	390.1889	40.7	80.52	59.77	37.65
Sedoheptulose	8.33	210.0739	52.92	80.57	61.88	61.15
Vinylacetylglycine	6.34	143.0582	98.84	80.82	52.38	37.08
Leu-Pro	8.49	228.1474	48.13	80.97	86.79	40.17
beta-Alanyl-N(pi)-methyl-L-histidine	11.51	240.1222	53.04	81.2	86.73	80.64
Glu-Asp	9.68	262.0801	35.44	81.52	86.29	48.35
Triacetate lactone	5.05	126.0317	125.78	81.57	130.77	79.26
Isocitrate	7.79	192.027	198.5	81.62	86.87	77.78

Pimelylcarnitine	8.09	303.1682	51.59	81.62	149.4	65.85
Phe-Ala	8.56	236.1162	81.58	81.64	112.27	70.97
8-Methoxykynurenate	5.42	219.0532	47.61	81.91	98.49	120.28
2-(Acetamidomethylene)succinate	7.68	187.0481	69.96	81.94	103.58	71.33
Indolelactate	5.49	205.0739	76.06	81.98	56.26	71.36
O-Acetylcarnitine	8.58	203.1158	29.81	82.23	114.91	41.64
Homoarginine	13.32	188.1274	66.4	82.3	79.8	108.56
sn-glycero-3-Phosphoethanolamine	9.81	215.0559	130.45	82.44	66.66	99.93
5-L-Glutamyl-taurine	9.53	254.0573	59.93	82.71	58.96	77.46
Deoxyloganin	5.1	374.1577	71.23	82.75	96.59	73.94
Fructoselysine	11.37	308.1584	54.97	82.87	103.6	22.18
Gln-Pro-His	8.05	380.1805	124.12	82.96	62.76	96.11
D-Glucosamine 6-phosphate	9.4	259.0457	59.68	83.02	102.72	79.33
hexanesulfonate	7.91	166.0664	27.79	83.38	52.13	99.46
L-Hypoglycin	7.99	141.079	98.52	83.41	56.01	100.77
Nonanoylcarnitine	7.66	301.2252	71.67	83.41	98.83	74.94
(S)-1-Pyrroline-5-carboxylate	8.18	113.0477	55.89	83.48	42.95	75.8
6beta-Hydroxy-D-homo-17a-oxaandrost-4-ene-3,17-dione	4.79	318.1831	50.09	83.48	71.32	69.87
5-6-Dihydrouridine	7.99	246.0852	100.57	83.62	221.94	98.85
Malonate	9.51	104.011	43.8	83.74	71.83	80.82
1-Pyrroline	7.68	69.0579	32.59	84.13	65.3	97.06
Theophylline	7.57	180.0647	62.82	84.19	106.98	81.65
3,5/4-Trihydroxycyclohexa-1,2-dione	7.73	160.0371	46.68	84.35	119.16	78.35
N(pi)-Methyl-L-histidine	9.03	169.0851	29.16	84.61	72.95	121.13
2-Amino-4-hydroxy-6-(D-erythro-1,2,3-trihydroxypropyl)-7,8-dihydropteridine	10.26	255.0967	27.98	84.73	54.85	35.1
2-trans,6-trans-Farnesal	5.22	220.1828	44.42	84.76	149.21	68.39
4-(beta-Acetylaminoethyl)imidazole	7.86	153.0901	79.48	84.81	147.95	93.37
[FA (7:2)] 2,4-heptadienal	5.4	110.0732	104.31	84.92	76.43	138.13
Succinate	9.36	118.0266	84.48	84.93	177.65	51.48
2-Oxoglutarate	9.62	146.0215	115.37	85.04	75.15	122.78
Leu-Lys-Pro	12.03	356.2425	48.52	85.06	80.97	67.28
N4-Acetylcytidine	8.12	285.0961	43.2	85.45	67.83	98.54
2-Hydroxybutane-1,2,4-tricarboxylate	10.01	206.0425	58.1	85.62	121.78	123.81
2-Hydroxy-3-(4-hydroxyphenyl)propenoate	11.74	180.0423	45.07	85.72	85.62	72.76
4'-Hydroxy-3'-prenylacetophenone	4.96	204.1151	165.84	85.75	59.06	105.16

(R)-AMAA	8.32	172.0484	74.58	85.83	101.1	113.29
Methylitaconate	6.23	144.0422	41.28	85.83	39.11	78.97
N-Succinyl-L-glutamate 5-semialdehyde	8.51	231.0742	73.85	86.11	82.23	75.55
Glucosylgalactosylhydroxylysine	11.78	486.2063	64.83	86.29	90.42	67.82
[FA oxo,amino(6:0)] 3-oxo-5S-amino-hexanoic acid	9.31	145.0739	43.94	86.67	67.13	21.57
beta-Citryl-L-glutamic acid	10.57	321.0697	50.69	86.73	66.72	82.13
pentanesulfonate	8.18	152.0507	70.61	86.77	81.64	82.64
Leu-Thr	10.1	232.1423	43.54	86.79	90.11	85.32
Kynurenate	6.24	189.0426	191.2	86.87	89.6	125.34
Sucrose	10.15	342.1163	67.24	86.88	137.61	75.72
N2-(D-1-Carboxyethyl)-L-arginine	9.96	246.1328	72.84	87	44.34	37.82
gamma-Amino-gamma-cyanobutanoate	6.19	128.0586	51.97	87.09	87.7	125.26
N-Formimino-L-glutamate	9.34	174.0641	46.74	87.1	89.95	50.09
2,3-Dimethylmaleate	9.82	144.0422	72.88	87.26	79.39	81.48
3-Methylglutarylcarnitine	8.22	289.1525	69.19	87.39	27.53	81.67
Pseudouridine	9.38	244.0694	51.04	87.48	63.06	56.4
[FA methyl(18:0)] methyl 9-hydroperoxy-10,12,13,15-bisepidioxy-16E-octadecenoate	5.45	388.2097	64.91	87.6	65.33	123.28
N-methyl glucamine	10.41	195.1108	47.16	87.66	76.14	86.67
gamma-Glutamyl-gamma-aminobutyraldehyde	6.26	216.1111	51.62	87.7	46.11	68.65
a Cysteine adduct	9.51	191.0252	123.79	87.89	168.73	82.02
corresponding carbamoyl amino acid	6.27	194.0691	40.68	87.95	46.79	75.52
Pancratistatin	8.17	325.0797	58.19	87.96	107.09	79.26
N-(3-oxooctanoyl)-L-homoserine	8.24	259.142	91.87	87.98	59.77	68.8
8-Epiiridotrial glucoside	5.14	344.147	113.98	88.23	63.3	51.39
L-Glutamine	9.88	146.0691	29.34	88.48	47.61	64.25
D-Glutamate	9.85	147.0531	106.1	88.55	57.29	68.34
Hydroxybutyrylcarnitine	8.6	247.1419	68.53	88.64	103.47	86.51
Quassin	5.33	194.0944	135.33	88.86	190.16	76.63
L-Alanine	10.01	89.0477	114.13	89.13	89.42	99.13
1-Methyladenosine	10.95	281.1125	48.45	89.21	53.49	78.22
L-Albizziine	10.56	147.0643	126.27	89.42	151.6	78.1
O-Succinyl-L-homoserine	8.88	219.0743	45.74	89.46	48.25	171.85
2-Hydroxyethanesulfonate	8.43	125.9987	37.44	89.54	128.05	69.46
Val-Val	8.97	216.1474	69.3	89.57	80.82	69.77
Lophophorine	5	235.1209	85.2	89.59	72.73	84.54

L-1-Pyrroline-3-hydroxy-5-carboxylate	6.2	129.0426	48	89.6	27.72	42.78
D-Galactosyl-3-(N-acetyl-beta-D-galactosaminyl)-L-serine	6.16	235.0878	91.61	89.78	89.92	77.29
D-Glucarate	9.98	210.0375	59.86	89.92	96.53	35.94
(R)-2-Hydroxyglutarate	9.26	148.0372	49.22	89.93	106.7	104.89
N-Formyl-L-aspartate	9.48	161.0323	79.92	89.95	46.41	104.03
[FA oxo(5:1/5:0/4:0)] (1R,2R)-3-oxo-2-(2'Z-pentenyl)-cyclopentanebutanoic acid	5.21	238.1567	48.88	90.05	95.04	65.74
Leu-Tyr	7.7	294.1579	57.89	90.11	181.34	110.89
L-Glutamate 5-semialdehyde	7.85	131.0582	40.78	90.11	88.48	79.15
Ectoine	9.4	142.0742	66.49	90.12	43.63	53.27
Lys-Pro	13.01	243.1583	64.36	90.19	70.4	47.25
Glu-Leu	8.15	260.1372	54.54	90.42	75.16	65.52
2-Deoxy-D-ribose 5-phosphate	9.67	214.0244	70.4	90.46	73.02	88.61
carbidopa	7.69	226.0954	92.6	90.59	49.99	81.4
3-Oxo-delta4-steroid	4.97	272.214	113.56	90.74	60.32	88.67
NG,NG-Dimethyl-L-arginine	12.63	202.143	52.1	91.09	133.41	66.51
N-Acetyl-aspartyl-glutamate	9.61	304.0906	62.71	91.17	56.32	154.51
4-(4-Deoxy-alpha-D-gluc-4-enuronosyl)-D-galacturonate	10.27	352.0641	83.39	91.26	84.81	52.56
Theogallin	8.24	344.0741	49.2	91.51	84.19	51.35
3b-16a-Dihydroxyandrostenonesulfate	4.6	384.1605	54.86	91.61	212.89	84.78
(25S)-11alpha,20,26-trihydroxyecdysone	5.6	512.2985	32.85	91.69	65.31	133.01
Homocysteinethiolactone	6.23	117.0249	116.21	92.06	150.41	89.01
N-benzoyl-L-glutamate	7.65	251.0792	37.24	92.15	76.71	62.82
[FA dioxo,hydroxy(4:0/2:0)] 9,15-dioxo-11R-hydroxy-2,3,4,5-tetranor-prostan-1,20-dioic acid	5.3	328.1521	48.7	92.16	110.29	81.69
Procollagen 5-(D-galactosyloxy)-L-lysine	12.11	324.1533	72.66	92.63	55.07	79.62
Pyruvate	10.04	88.016	45.39	92.84	88.86	83.88
Glycerophosphoglycerol	9.91	246.0505	61.55	92.96	62.77	78.12
7-Methylxanthine	8.03	166.0491	102.37	93.02	93.46	77.21
7-Methyladenine	10.94	149.0701	48.49	93.19	76.24	86.97
Tetranor-PGE1	4.82	298.1782	90.65	93.45	49.22	79.13
8-Epiiridodial glucoside	5.22	330.1679	66.76	93.46	88.23	41.55
5-Oxoproline	7.74	129.0426	41.92	93.72	94.43	84.55
S-Acetyldihydrolipoamide	7.99	249.086	120	93.75	64.92	32.78
5-Methylcytidine	10.7	257.1012	56.64	94.13	69.37	47.13
N4-(Acetyl-beta-D-glucosaminyl)asparagine	9.72	335.1329	45.13	94.36	109.64	93.82

N-(octanoyl)-L-homoserine	7.94	245.1627	52.28	94.42	84.61	90.98
5-Valerolactone	9.96	100.0524	49.2	94.43	98.52	119.31
3-hydroxy-15-dihydrolubimin	5.1	254.1882	49.42	94.47	107.95	60.04
2-Hydroxy-2,4-pentadienoate	7.71	114.0317	37.31	94.65	85.72	110.04
Aspidinol	4.78	224.1049	96.01	94.75	101.35	90.39
N-Ethylmaleimide	9.29	125.0477	40.81	94.8	96.34	133.09
3-Methyl-2-oxobutanoic acid	6.16	116.0474	51.02	94.87	87.39	73.71
N-acetyl-(L)-arginine	9.84	216.1223	52.67	94.92	91.17	141.8
[FA oxo(5:2/5:0/4:0)] (1S,2S)-3-oxo-2-pentyl-cyclopentanebutanoic acid	5.19	240.1725	37.27	95.04	86.67	75.32
[FA hydroxy,oxo(7:0/2:0)] 4-hydroxy-2-oxo-Heptanedioic acid	8.25	190.0477	35.6	95.1	68.33	88.31
3-Hydroxydodecanedioicacid	7.68	246.1467	76.74	95.16	72.25	91.17
[ST trihydroxy(2:0)] 11beta,17,21-trihydroxypregn-4-ene-3,20-dione, 21-sulfate	4.8	442.166	69.51	95.35	69.22	70.3
2-Isopropylmaleate	8.88	158.0579	66.9	95.35	212.38	112.13
Homovanillicacidsulfate	10.82	262.0146	134.2	95.41	123.12	66.95
[FA (10:0)] O-decanoyl-R-carnitine	6.18	315.241	114.79	95.55	123.95	118.44
D-Methionine	7.73	149.051	42.09	95.61	70.58	122.04
Dodecanedioylcarnitine	6.23	373.2466	87.64	95.61	178.35	63.55
Androsterone glucuronide	4.97	466.2566	73.32	95.78	69.55	74.8
THTC	8.51	132.0245	36.69	96.33	72.87	35.04
Cys-Ser	9.86	208.0518	46.41	96.34	113.95	104.29
Neuraminic acid	10.15	267.0953	63.92	96.34	31.23	90.48
D-Gluconic acid	9.19	196.0582	31.57	96.53	83.02	95.15
Dethiobiotin	7.93	214.1318	43.73	96.59	73.06	65.33
p-Cresolglucuronide	6.24	284.0896	117.4	96.68	107.5	30.05
Guanidinoacetate	10.47	117.0538	38.16	96.69	30.61	83.61
Ne,Ne dimethyllysine	12.25	174.1369	61.9	96.83	94.8	91.33
2-C-Methyl-D-erythritol 4-phosphate	9.48	216.0401	86.61	96.86	57.26	119.14
4-Aminohippuricacid	7.66	194.0691	135.95	97.11	50.66	73.58
2,3-Diaminosalicylic acid	6.42	168.0535	113.63	97.88	73.38	138.84
4-Trimethylammoniobutanal	11.99	129.1154	61.74	98.06	68.39	92.32
3-ureidoacrylate	8.04	130.0379	77.78	98.35	117.14	133.99
4-O-beta-D-Glucosyl-4-hydroxycinnamate	6.24	326.1001	112.42	98.42	116.94	123.91
Leu-Ala	9.12	202.1318	48.03	98.46	108.09	71.97
8-methylthiooctylhydroximate	4.77	221.0906	47.07	98.49	58.88	92.52

6-Amino-2-oxohexanoate	7.85	145.0738	105.75	98.52	83.48	60.91
N2-Acetyl-L-aminoadipate semialdehyde	7.68	187.0844	37.54	98.52	31.05	106.73
Norepinephrinesulfate	11.42	249.0306	101.81	98.83	86.11	81.25
S-Methyl-1-thio-D-glycerate	7.92	136.0194	56.47	99.02	156.08	84.8
N-(2,3-Dihydroxybenzoyl)-L-serine	10.57	241.0586	48.05	99.03	70.79	52.85
Citrate	10.26	192.027	44.33	99.13	59.92	70.51
Allantoin	10.4	158.0439	49.86	99.14	127.43	98.06
Cerulenin	4.8	223.1209	57.58	99.34	73.85	89
α-oxo-1-carboxy-5,6-dihydrothiopyranpropanoate S-oxide	9.87	246.0197	36.14	100.59	80.52	132.43
N-Acetylglutamine	8.05	188.0797	141.49	100.66	141.1	70.45
(R,R)-Tartaric acid	10.06	150.0163	180.97	101.1	83.48	67.35
Asp-Asp-Pro	9.45	345.1174	106.73	101.3	94.75	74.12
Asp-Leu-Met-Cys	8.34	240.0859	74.88	101.35	114.45	91.59
Acetone cyanohydrin	7.82	85.0527	251.89	102.34	250.85	128.77
N-Acetyl-D-glucosamine 6-sulfate	9.3	301.0466	57.83	102.34	28.67	72.35
3-(4-Hydroxyphenyl)lactate	7.73	182.0579	86.68	102.41	64.27	33.1
1-(3-aminopropyl)-4-aminobutanal	13.52	144.1262	75.44	102.46	62.51	222.95
D-Glucose 6-sulfate	10.81	260.02	32.57	102.72	64.13	92.8
1-O-(8R-hydroxy-8-methyl-3Z,9-decadienoyl)-beta-D-glucopyranose	4.72	360.1785	93.31	103.1	62.36	96.07
N-Acetyl-L-citrulline	8	217.1061	79.86	103.25	69.66	94.07
Hymexazol O-glucoside	9.69	261.0848	64.96	103.47	119.92	136.68
Tuliposide B	9.84	294.0951	37.3	103.53	67.42	87.16
2(alpha-D-Mannosyl)-D-glycerate	9.43	268.0794	52.01	103.58	263.62	98.52
fumarylacetic acid	5.04	158.0215	73.38	103.6	67.91	60.76
p-aminobenzoate-β-D-glucopyranosyl ester	8.4	299.1005	54.94	103.95	87.96	172.96
Tauropine	7.75	197.0358	82.47	104.03	66.5	118.2
[FA trihydroxy(2:0)] 9S,11R,15S-trihydroxy-5Z,13E-prostadienoic acid	4.83	354.2407	27.2	104.35	137.96	58.99
Gly-His	7.16	212.0909	117.15	104.36	74.12	53.1
3-Aminopropane-1,2-diol	13.73	91.0633	44.09	104.61	91.61	75.56
Tiglic acid	8.7	100.0525	44.77	104.99	70.84	92.34
11-beta-hydroxyandrosterone-3-glucuronide	5.36	482.2516	81.07	105.01	79.22	69.88
D-Xylulose	8.11	150.0528	55.84	105.16	90.12	62.67
Phenylacetaldehyde	5.75	120.0575	221.74	105.3	51.89	77.55
citraconate	7.81	130.0266	133.36	105.94	99.13	91.23

4-(2-Aminophenyl)-2,4-dioxobutanoate	8.22	207.0531	71.88	106.36	91.26	52.95
3-Mercaptolactate-cysteinedisulfide	9.31	241.0078	48.8	106.62	262.27	66.92
(R)-3-Hydroxybutanoate	7.81	104.0474	48.48	106.7	85.83	82.96
L-Azetidine 2-carboxylic acid	8.23	101.0477	45.82	106.71	55.93	82.63
Thiophanate-methyl	9.57	342.0462	96.66	106.98	107.3	81.24
Pantothenate	6.3	219.1107	46.03	107.09	96.68	139.06
Xanthurenic acid	8.43	205.0376	104.81	107.13	258.34	92.05
Thr-Asn-His	6.94	185.0801	81.27	107.3	164.24	63.61
D-Phenylalanine	8.53	165.079	40.32	107.35	68.38	106.86
P-DPD	9.01	212.0086	56.18	107.5	41.15	156.04
1-deoxyxylonojirimycin	10.13	133.0738	71.08	107.82	89.21	97.44
Pyrrole-2-carboxylate	8.06	111.032	150.63	107.93	92.84	71.33
3-Hydroxy-2-methylpyridine-4,5-dicarboxylate	9.08	197.0327	64.96	107.95	72.77	125.89
Leu-Lys	11.71	259.1896	49.28	108.09	69.48	141.53
methiin	8.39	151.0304	127.83	108.33	57.06	94.64
[FA oxo(12:1)] 12-oxo-10E-dodecenoic acid	5.05	212.1413	44.45	108.49	58.94	78.93
2-Pyridyl hydroxymethane sulfonic acid	5.58	189.0096	82.46	108.76	60.63	74.54
3-sulfopropanoate	10.02	153.9935	71.58	109.05	98.35	90.97
D-Cathinone	8.27	149.0841	71.97	109.3	111.13	137.41
Asn-Asp-Ser	8.66	334.1125	75.37	109.49	80.21	60.09
N4-Acetylaminobutanal	12.3	129.079	39.24	109.64	85.45	85.04
Val-Asp-Pro	6.18	329.1587	48.3	109.93	60.57	79.63
[PR] Limonene-1,2-diol	5.5	170.1306	56.05	109.98	123.71	95.61
Indoleacrylicacid	8.28	187.0634	47	110	129.25	87.22
[FA hydroxy(10:0)] N-(3S-hydroxydecanoyl)-L-serine	8.11	275.1732	64.36	110.29	78.62	83.63
[FA trihydroxy(2:0)] 9S,11R,15S-trihydroxy-2,3-dinor-5Z,13E-prostadienoic acid-cyclo[8S,12R]	4.96	326.2092	68.37	110.45	104.35	79.33
Val-Tyr	5.02	280.1423	69.43	110.52	89.57	107.03
Dehydrojuvabione	5.01	264.1725	82.46	111.13	76.31	90.4
10-nitro-9E-octadecenoic acid	6.19	327.241	44.38	111.92	105.01	146.15
O-Butanoylcarnitine	8.06	231.1471	63.65	112.02	119.54	106.57
Phe-Asp	7.77	280.1059	68.69	112.27	58.73	91.4
L-Methionine S-oxide	6.26	165.0461	53.61	113.8	50.71	43.25
[FA oxo,hydroxy(5:1/5:0)] (1S,2R)-3-oxo-2-(5'-hydroxy-2'Z-pentenyl)-cyclopentaneacetic acid	4.68	226.1205	47.01	113.85	110.45	82.47
D-Alanyl-D-alanine	7.8	160.0848	167.32	113.95	109.3	161.56

(L-Seryl)adenylate	9.81	218.0556	74.85	114.28	63.03	127.36
Asp-Ser	9.51	220.0696	200.71	114.45	220.7	110.84
O-Acetyl-L-homoserine	8.39	161.0688	95.63	114.91	112.02	166.5
Serotonin	6.28	176.0951	251.67	115.57	122.85	100.95
ubiquinol-1	6.13	252.1362	58.74	115.58	116.51	106.98
Lotaustralin	8.29	261.1212	97.88	115.83	77.54	88.71
3-Butynoate	8.57	84.0211	67.69	116.08	159.08	109.21
Lys-Gly	9.03	203.127	45.33	116.3	90.19	69.14
Uracil	8.46	112.0273	51.34	116.51	71.65	40.04
5-Amino-4-chloro-2-(2,3-dihydroxyphenyl)-3(2H)-pyridazinone	9.44	253.0256	253.96	116.81	48	156.88
4-Pyridoxate	7.66	183.0532	129.98	116.94	98.06	92.97
3-Ureidoisobutyrate	6.18	146.0692	67.9	117.14	106.36	109.48
Ile-Thr	12.35	232.1423	42.56	117.21	79.97	104.32
L-Asparagine	9.93	132.0534	67.52	118.51	106.71	113.23
3,5-Dihydroxyanthranilate	7.67	169.0375	113.7	119.16	133.64	99.17
Indoxyl	7.71	133.0527	65.97	119.22	119.51	218.34
Indoxylsulfate	7.71	213.0096	65.44	119.51	81.62	162.91
Formylanthranilate	5.55	165.0426	66.29	119.54	61.93	52
O-Propanoylcarnitine	5.13	217.1314	40.2	119.54	80.51	49.6
Hypaphorine	8.27	246.1368	81.96	119.92	73.79	108.52
Cyclohex-2-enone	5.39	96.0575	52.86	120.08	96.34	150.67
[PR] Flakinin A	4.8	280.131	53.58	120.84	109.98	47.98
2-Hydroxydecanedioicacid	7.8	218.1154	38.12	121.78	89.54	90.98
4-Hydroxyaniline	5.37	109.0528	66.33	122.66	57.56	82.18
alpha-N-Phenylacetyl-L-glutamine	6.17	264.111	66.92	122.85	58.5	149.66
Ser-Ser	10.97	192.0747	67.05	122.85	70.83	89.78
[FA hydroxy(9:1)] 4-hydroxy-2-nonenal	6.22	156.1151	29.14	122.87	95.1	41.38
Butanoic acid	8.64	88.0524	85.39	122.95	61.48	101.93
Hydantoin-5-propionate	9.44	172.0484	48.2	123.12	88.64	49.78
[PR] Tretinoin/All-Trans Retinoic Acid	4.82	300.2092	35.64	123.71	141.82	115.94
[FA (10:1/2:0)] 2E-Decenedioic acid	4.83	200.1049	45.67	123.95	58.06	73.6
3-Amino-3-(4-hydroxyphenyl)propanoate	7.94	181.074	129.52	124.13	104.61	115.25
Geranylhy droquinone	4.99	246.162	46.86	124.54	56.11	64.8
2-Acetolactate	8.63	132.0423	86.09	125.98	47.29	103.36
trihomomethionine	8.41	191.098	72.79	126.01	79.77	66.07
2-amino-4-carboxypyrimidine	8.05	139.0381	116.43	126.96	84.73	48.55

Allylisothiocyanate	9.49	99.0143	59.19	127.43	79.02	75.55
Glu-Val	8.38	246.1215	44.65	127.56	66.06	80.12
2-Hydroxypyridine	7.99	95.0371	87.54	128.05	95.35	66.8
Nonane-4,6-dione	7.66	156.115	94.06	128.81	83.41	106.38
Indolebutyric acid	5.36	203.0947	89.1	129.25	81.98	94.54
4-Nitrophenylsulfate	10.71	218.9837	90.85	129.9	98.42	103.56
Triadimefon	8.35	293.0932	57.75	130.77	126.01	119.97
1-Methyluric acid	8.03	182.044	102.37	131.97	103.1	56.59
Homocysteinesulfinicacid	7.72	167.0252	54.48	132.17	92.06	115.74
L-Indospicine	8.94	173.1164	54.4	132.38	66.42	126.06
N-Heptanoylhomoserine lactone	4.83	213.1365	143.18	133.41	75.89	74.03
3-17-Androstanediolglucuronide	5.37	468.2724	100.84	133.64	77.81	249.5
Glutathionylaminopropylcadaverine	4.86	448.246	107.15	133.71	127.56	172.23
2-Oxazolidinone	6.28	87.032	135.34	136.11	54.64	123.05
L-Ornithine	12.61	132.0899	125.53	137.15	115.83	130.86
Sulfoacetaldehyde	9.87	123.983	71.12	137.61	79.08	118.98
[FA trihydroxy(4:0)] 2,3,4-trihydroxy-butanoic acid	8.67	136.0372	42.58	137.96	72.24	76.1
[FA amino(4:0)] 2-amino-4-cyano-butanoic acid	7.65	128.0585	43.57	138.77	149.67	95.04
Capryloylglycine	4.87	201.1365	93.81	139.29	90.59	90.04
Dihydrobiopterin	9.7	239.1018	44.44	140.48	57.14	50.18
Coniferin	5.11	342.1315	84.28	140.64	87.95	54.09
N-Acetylisatin	7.26	189.0426	33.66	141.1	59.14	163.31
2-8-Dihydroxyquinoline-beta-D-glucuronide	8.14	337.0798	104.51	141.54	125.98	133.49
[ST hydrox] 3alpha-hydroxy-5alpha-androstan-17-one 3-sulfate	4.48	370.1812	26.17	141.82	95.35	81.06
N-Carbamyl-L-glutamate	9.44	190.059	56.86	141.91	39.44	170.92
Formylpyruvate	9.22	116.011	96.51	141.94	82.87	123.83
Heptanoylcarnitine	7.67	273.1939	52.58	143.89	83.38	71.43
3-Indolepropionicacid	5.58	189.079	82.36	146.24	73.45	122.5
17-Hydroxypregnenolonesulfate	4.52	412.1918	115.58	146.68	62.61	78.63
2,5-Dihydroxybenzoate	10.17	154.0266	81.6	147.03	52.29	106.58
4,6-Dihydroxyquinoline	8.12	161.0477	49.13	147.95	70.08	131.88
Deoxyinosine	9.24	252.0858	57.84	148.12	82.75	82.79
4-Imidazolone-5-propanoate	8.31	156.0535	36.75	148.51	76.42	84.16
3-(3,4-Dihydroxyphenyl)lactate	7.81	198.0528	207.4	149.21	102.41	116.17
L-rhamnitol	9.03	166.084	62.52	149.4	45.05	33.24
p-nitroso-N,N-dimethylaniline	8.45	150.0793	92.43	149.4	41.05	131.24

2-Phenylacetamide	7.49	135.0684	166.45	149.53	108.76	137.73
[FA amino(9:0)] 9-amino-nonanoic acid	7.9	173.1415	61.51	149.67	78.35	80.16
Homogentisate	10.93	168.0422	53.89	150.41	95.41	113.12
L-Homocitrulline	9.98	189.1114	49.94	151.42	36.93	73.83
L-Arginine	15.8	174.1117	54.23	151.6	42.14	123.68
3-Hydroxytetradecanedioicacid	5.26	274.1781	177.68	153.03	146.24	122.1
S-Methyl-L-methionine	9.25	163.0667	52.46	156.08	82.44	169.67
pentahomomethionine	7.85	219.1293	51.46	157.04	86.77	186.76
S-(3-oxo-3-carboxy-n-propyl)cysteine	9.26	221.0357	39.95	157.59	93.75	145.9
3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid	5.23	240.0998	144.42	159.08	38.65	74.53
Nicotinate	9.48	123.0321	85.7	161.79	73.65	99.48
N6-Acetyl-N6-hydroxy-L-lysine	8.54	204.1107	102.11	163.82	62.17	137.94
Thr-Pro	7.65	216.111	70.17	164.24	61.77	84.99
Crotanecine	8.51	171.0895	83.01	164.43	120.08	150.47
Acetaminophenglucuronide	7.75	327.0955	117.18	168.73	102.34	70.38
L-Cystine	9.68	240.0238	73.93	169.22	49.83	167.27
Ascorbate	9.61	176.032	159.41	174.87	68.16	81.26
Succinyladenosine	9.05	383.1076	35.82	177.65	86.88	28.8
Dopamine 3-O-sulfate	10.88	233.0357	123.88	178.35	67.53	81.49
Leu-Val	9.05	230.1631	93.54	181.34	73.11	143.84
2-Aminophenol	8.57	109.0529	157.32	186.83	96.86	52.43
2-Dehydro-D-xylonate	9.54	164.0321	71.8	189.29	90.46	89.19
Quinate	8.75	192.0634	125.57	190.16	225.84	59.88
5,6-Dihydrothymine	10.03	128.0586	123.83	191.38	57.88	182.18
2-O-alpha-L-Rhamnopyranosyl-D-glucopyranose	10.03	326.1211	127.08	212.38	74.71	146.9
[FA (5:2)] 2,4-pentadienoic acid	7.83	98.0368	73.28	212.39	63.57	139.05
3beta-Hydroxyandrost-5-en-17-one 3-sulfate	4.46	368.1656	56.97	212.89	68.42	74.98
Benzofuran	5.54	118.0419	193.92	220.7	81.2	53.81
5-Acetamidopentanoate	9.19	159.0896	157.1	221.94	80.17	26.44
Retronecine	9.49	155.0948	134.57	225.84	157.59	130.62
5'-Hydroxycotinine	7.19	192.09	178.39	246.38	45.74	71.53
5,6-Dihydroxyindole	6.12	149.0477	48.83	248.15	83.62	120.82
Acronycidine	5.06	289.095	141.24	250.85	61.5	218.69
Xylitol	9.58	152.0683	39.59	258.34	100.59	160.51
Slaframine	8.56	240.1473	45.61	259.88	99.02	93.5
3-Methoxy-4-hydroxyphenylethyleneglycol	5.78	184.0737	230.49	262.27	72.02	152.01

5-guanidino-3-methyl-2-oxo-pentanoate	10.05	187.0957	89.52	262.43	246.38	160.31
Phe-Thr	5.18	266.1267	118.23	263.5	25.8	101.93
2,3,4,5-Tetrahydropyridine-2-carboxylate	9.9	127.0634	262.1	263.62	58.82	67.73

Appendix A6. Comparison between %RSD of non-normalised and normalised metabolites in Human Brain Tissues

Metabolite	Retention time (minutes)	Mass	GREY MATTER		WHITE MATTER	
			Non-normalised RSD (%)	C13 Normalised RSD (%)	Non-normalised RSD (%)	C13 Normalised RSD (%)
Creatine	10.15	135.0827	9.58	6.48	5.07	3.3
Leu-Ser	5.17	227.1564	16.21	7.48	13.35	7.17
L-Histidinal	7.35	145.0943	9.24	8.17	24.95	21.12
Imidazole-4-acetaldehyde	9.24	115.0646	25.88	9.28	23.79	21.39
Val-Pro-His	8.21	367.2433	17.87	10.75	13.37	3.63
1-Methylnicotinamide	16.29	143.0867	36.08	10.97	25.83	23.03
Glu-Asp-Cys	7.94	377.1289	37.89	11.02	34.85	29.33
Dihydrobiopterin	8.5	248.1316	17.23	11.03	31.52	24.92
L-Hypoglycin	7.28	148.1021	23.65	11.59	44.53	125.47
Pro-Ser-Ser	8.55	300.1637	14.44	11.77	20.8	16.79
7-Methyladenine	7.82	155.09	24.88	11.82	19.26	16.17
Aspartyl-L-proline	9.44	239.1198	21.85	11.88	21.92	19.04
N6-Methyl-L-lysine	13.86	167.1442	14.1	11.9	22.27	20.26
Phe-Thr	5.13	279.1696	18.27	12.09	24.43	20.81
5,6-Dihydrothymine	10.19	133.0752	14.23	12.34	14.49	11.65
L-Serine	10.45	108.0525	31.46	12.72	24.3	21.56
Leu-Pro-Pro	5.16	341.253	20	13.32	18.37	13.77
1,2-Didecanoylglycerol	3.84	423.3952	20	13.4	36.74	25.67
2-Oxoglutaramate	7.7	150.054	18.2	13.66	23.3	21.24
Maltotriose	10.51	522.2289	20.76	13.74	21.31	15.86
D-Gluconic acid	8.92	202.0782	17.77	13.77	14.13	12.64
Thr-Asp-His	8.64	385.1905	18.95	13.83	27.53	23.82
Hydroxybutyrylcarnitine	8.43	258.1783	43.72	14.01	29.45	22.53
Pimelylcarnitine	4.99	317.2143	18.66	14.17	39.55	34.04

4-Methylene-L-glutamine	9.25	164.0889	24.82	14.19	16.31	14.53
Asp-Gln-Gln-Pro	13.1	505.2692	41.62	14.21	19.08	14.31
Suberylglycine	5.14	241.1437	22.22	14.36	18.22	16.58
allylcysteine	5.13	167.0708	16.48	14.38	11.14	9.04
5-Hydroxypentanoate	5.16	123.0796	25.51	14.49	16.37	6.09
Leu-Ala-Pro	5.14	313.2307	19.97	14.5	13.09	10.24
Cardiospermin	5.18	286.1369	16.85	14.62	11.41	10.52
Ethyl (R)-3-hydroxyhexanoate	4.51	168.1364	27.23	14.74	25.25	22.64
N,N-Dimethylglycine	10.35	107.0766	27.99	14.87	37.9	33.49
N5-Ethyl-L-glutamine	9.48	181.1235	16.99	14.97	20.65	16.18
Met-Pro	5.01	256.1367	18.09	15.01	13.04	10.36
Acetylcysteine	5.23	168.0468	27.07	15.18	15.76	14.21
Methylimidazoleacetic acid	7.26	146.0783	25.3	15.3	25.86	23.83
Ile-Thr	5.16	242.1754	18.04	15.39	17.89	15.78
Gly-His	8.82	220.1176	18.66	15.51	14.51	18.45
Asn-Met-Gln-Ser	7.86	495.2409	33.03	15.61	49.52	36.41
Ala-Asn-Pro	5.27	312.1829	25.68	15.63	25.64	17.93
Uracil	8.09	116.0405	20.5	15.77	20.2	14.93
3-Carboxy-4-methoxy-N-methyl-2-pyridone	4.71	191.0797	20.94	15.79	55.97	29.71
Orotate	7.57	161.0337	23.72	15.92	32.79	28.37
Glutarylcarnitine	7.92	287.1764	39.58	16.13	36.96	27.77
(2S)-2-Isopropyl-3-oxosuccinate	5.33	181.0759	21.15	16.14	15.28	10.81
Leu-Thr-Pro	5.16	344.2445	20.4	16.19	18	14.16
Guanidinoacetate	10.78	120.0637	20.05	16.3	46.85	29.26
4-Guanidinobutanoate	10.38	150.1016	27.73	16.4	23.97	19.26
Leu-Gly-Pro	5.16	298.2119	19.4	16.4	14.15	11.09
1-(3-aminopropyl)-4-aminobutanal	14.08	151.1494	19.81	16.54	27.28	34.27

S-Adenosyl-L-methionine	10.51	413.1869	27.43	16.54	41.29	35.7
Glaucarubin	4.16	521.3125	18.62	16.58	12.35	10.42
Atherospermidine	8.58	323.1277	22.43	16.67	33.77	30.39
(S)-Malate	9.85	138.0348	20.88	16.72	21.04	19.12
4-Acetamidobutanoate	5.23	151.0937	20.55	16.74	14.01	12.03
Dodecanedioylcarnitine	4.46	392.3091	19.31	16.77	52.71	29.41
Glycyl-leucine	5.17	196.1425	19.5	16.79	18.85	10.01
2-Furoate	10.65	117.0326	29.1	16.98	44.78	36.68
6-Deoxyerythronolide B	4.02	407.3365	27.51	17.16	18.21	13.03
Leu-Gln-Pro	5.16	372.2587	19.12	17.27	15.25	10.76
Specionin	7.43	225.1447	23.78	17.36	26.33	18.79
Glu-Pro-Thr	8.29	359.1997	24.46	17.38	32.98	24.58
2,3-Dimethylmaleate	10.17	150.062	29.81	17.59	43.66	111.08
Isocitrate	8.26	198.0468	21.85	17.62	31.81	13.79
Cys-Pro	5.13	226.099	28.87	17.63	16.82	14.54
Deoxyadenosine	7.38	261.1348	30.04	17.67	19.17	17.61
Val-Pro-Pro	7	326.234	26.48	17.72	37.39	32.14
Homoarginine	15.17	195.1504	35.8	17.79	22.38	27.37
4-Aminobutanoate	7.01	107.0765	23.31	17.85	28.88	35.48
Ala-Gly-Pro	5.46	253.1552	31.67	18.03	17.52	12.24
Leu-Pro-Arg	8.78	401.3046	30.15	18.07	28.55	17.41
benzenesulfonate	5.25	164.0235	23.09	18.18	25.08	27.34
Ibuprofenacylglucuronide	4.16	401.2259	26.43	18.26	21.5	16.11
γ-thiomethyl glutamate	7.01	199.0607	20.74	18.31	17.07	19.88
3-N4-ethenocytosine	6.96	141.0632	20.73	18.32	41.36	101.46
Ala-Leu-Pro-Ser	5.17	403.2726	20.59	18.36	21.8	19.38
D-glycero-D-manno-Heptose 1,7-bisphosphate	10.46	377.0299	34.55	18.41	47.99	27.18

Ile-Ala	5.16	211.1615	29.31	18.42	17.98	13.81
Leu-Asn-Pro	5.17	357.2399	28.16	18.47	17.12	12.41
Choline	14.37	108.1162	36.79	18.6	21.58	17.71
N-(2,3-Dihydroxybenzoyl)-L-serine	8.74	251.0917	23.67	18.62	39.53	35.98
Slaframine	8.19	252.187	34.13	18.65	22.98	20.43
L-Proline	9.33	120.0798	24.36	18.67	25.58	20.4
Thr-Ala-Pro	5.22	299.1881	24.67	18.68	21.88	10.96
N-Acetyl-D-glucosamine	8.87	229.1162	25.45	18.72	13.35	11.97
Val-Val-Ala	8.58	300.2274	31.16	18.86	45.57	15.56
L-Arginine	14.9	180.1315	32.09	18.88	27.96	23.2
Thiamin	14.98	276.1441	32.91	19.06	29.41	42.59
Creatinine	8.22	117.0721	27.76	19.11	13.5	10.7
2-Hydroxyethanesulfonate	7.92	128.0053	29.25	19.2	26.4	22.79
GDP-L-fucose	10.43	605.1353	24.59	19.28	29.11	23.86
Nitrilotriacetic acid	9.36	197.0627	32.13	19.3	15.92	13.1
Imidazol-5-yl-pyruvate	7.83	160.0575	32.61	19.32	33.21	24.36
1-Methylhypoxanthine	7.62	156.0739	30.56	19.39	21.66	14.99
Guanine	9.61	156.0659	25.27	19.42	21.91	28.03
CMP-2-aminoethylphosphonate	9.47	441.1023	26.47	19.45	29.09	37.55
Leu-Pro-Ser	5.17	329.2258	29.33	19.51	13.46	12.34
S-Ribosyl-L-homocysteine	10.18	276.1074	33.25	19.54	19.2	14.98
(R)-2-Hydroxyglutarate	9.53	153.0537	33.18	19.6	47.65	124.44
Oleandolide	4.34	406.2968	28.61	19.62	20.98	10.98
neuromedin N	3.98	649.4839	31.9	19.63	39.75	44.14
dTDP-6-deoxy-L-talose	5.89	564.1338	39.34	19.75	14.74	19.26
7-O-Acetylsalutaridinol	7.55	392.2422	23.15	19.82	20.13	13.77
Cimifugin	4.22	322.1633	23.7	19.85	32.11	35.66

6-Methylpretetramide	7.94	385.1557	44.29	19.92	30.89	26.99
1'-Acetoxyeugenol acetate	4.27	278.1459	25.44	19.95	28.13	33.05
L-isoleucyl-L-proline	8.09	239.1837	32.52	19.95	25	21.57
Ala-Cys-Cys-Ser	10.12	394.1378	27.9	19.97	53.01	136.85
N(pi)-Methyl-L-histidine	9.2	176.1082	28.49	20	48.7	19.46
3,4-Dihydroxymandelaldehyde	4.85	176.0688	29.15	20.09	20.04	16.87
L-Glutamate methylester	5.34	167.0885	25.5	20.18	44.36	131.2
sn-Glycerol 3-phosphate	9.46	175.0236	35.71	20.21	34.1	25.37
D-Glucarate	9.76	216.0574	28.7	20.35	39.42	17.4
N-hydroxyl-tryptamine	5.26	186.128	20.37	20.37	11.38	7.63
5-6-Dihydrouridine	7.61	255.1149	32.46	20.4	29.02	24.22
N-Acetyl-D-quinovosamine	5.76	213.1214	30.02	20.41	68.22	93
N-Succinyl-L-glutamate 5-semialdehyde	8.34	240.1039	35.88	20.42	31.99	26.93
Dethiobiotin	7.01	224.1648	25.87	20.5	37.28	31.3
N-Succinyl-LL-2,6-diaminoheptanedioate	9.73	301.1476	25.08	20.5	24.74	21.31
Urate	8.69	173.0447	35.12	20.51	24.92	21.92
gamma-Glutamyl-gamma-aminobutyraldehyde	7.59	225.1407	32.32	20.53	24.9	19.14
N-(3-acetamidopropyl)-4-aminobutanal	8.16	195.1665	32.38	20.53	25.48	22.96
D-Glutamine	7.08	151.0856	58.49	20.57	75.92	23.88
Glucosyloxyanthraquinone	4.48	406.1667	27.92	20.63	20.84	9.6
Scoparone	4.63	217.0941	26.53	20.65	26.54	21.82
Asn-Gln-Pro	8.25	371.2111	30.14	20.71	40	32.45
N-Carbamoylputrescine	15.62	136.1223	25.96	20.75	50.47	136.78
Diethanolamine	12.94	109.0922	19.53	20.83	21.67	18.8
L-Valine	8.6	122.0955	30.16	20.88	14.8	10.61
Asp-Val-Gly-Pro	5.34	402.233	38.88	20.95	11.77	15.78
Pederin	4.08	528.3926	31.34	21.07	21.53	18.85

Proacacipetalin	5.16	270.142	26.15	21.13	22.75	17.97
3-Hydroxysebacicacid	5.11	228.1484	31.37	21.17	21.49	17.94
Thr-Pro-Pro	7.06	327.21	24.33	21.23	73.04	124.68
Glu-Ile-Pro	5.18	373.2428	26.18	21.24	25.49	22.66
Iminoaspartate	7.07	135.035	31.13	21.29	43.18	39.03
D-Glucuronate 1-phosphate	9.44	280.0289	34.73	21.34	43.76	32.25
Deoxyinosine	7.94	262.1187	25.19	21.35	32.23	28.18
Thr-Thr-Thr	7.48	333.1931	37.16	21.47	16.1	12.53
IMP	9.51	358.0804	24.02	21.5	29.24	22.22
S-8-methylthiooctylhydroximoyl-L-cysteine	8.5	320.1628	29.26	21.5	23.81	25.82
[FA methyl(18:2)] methyl 9-hydroperoxy-10,12-epidioxy-13,15-octadecadienoate	4.5	375.2828	17.6	21.54	21.72	22.35
N-Ribosylnicotinamide	15.59	265.1265	34.06	21.54	52.36	133.67
Glutarylglycine	5.29	196.0869	24.22	21.59	13.17	11.72
N-Acetyl-D-fucosamine	7.99	213.1214	36.59	21.61	42.65	36.14
1-nitrobutane	5.16	107.0765	38.9	21.68	21.31	13.28
His-Gly-Ser-Ser	7.93	207.1235	25.23	21.68	26.62	20.74
O-Succinyl-L-homoserine	8.79	227.1007	32.21	21.72	21.3	17.19
(R,R)-Tartaric acid	10.07	154.0295	24.52	21.73	27.29	34.89
L-Olivosyl-oleandolide	4.1	542.3798	26.36	21.73	16.15	11.53
Ala-Val-Pro-Ser	5.18	388.254	28.72	21.76	23.31	15.4
L-Glutamine	9.88	151.0856	24.99	21.77	87.34	38.37
myo-Inositol	11.07	186.0833	33.84	21.79	17.46	15.49
8-Oxocoformycin	7.61	293.1326	34.4	21.8	24.56	20.55
D-Mannose	7.13	186.0832	25.94	21.8	23.79	19.3
2,3-dioctanoylglyceramide	4.14	376.3142	25.56	21.86	20.92	17.33
L-Arabinonate	8.72	171.0644	26.69	21.86	22.77	20.17
D-Mycinose	5.18	200.1262	36.22	21.87	35.35	25.53

S-Acetylphosphopantetheine	5.17	413.1499	24.69	21.88	17.97	14.75
[PR] 2,3-Didehydrogibberellin A10	10.17	333.2139	24.87	21.93	19.53	20.85
UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-6-carboxy-L-lysyl-D-alanyl- D-alanine	9.4	637.8064	27.47	21.94	28.57	29.61
4-Imidazolone-5-propanoate	7.96	162.0733	20.96	21.95	23.32	20.32
N-Acetyl-L-aspartate	9.11	181.0679	27.28	21.96	23.84	15.06
L-Methionine S-oxide	9.08	170.0624	27.31	22	27.18	23.98
3-Methylguanine	8.33	171.0848	24.3	22.06	47.39	30.52
(S)-ATPA	9.55	238.1442	38.15	22.13	32.05	23.29
(R)-3-Hydroxy-3-methyl-2-oxopentanoate	5.19	152.0777	28.23	22.15	21.74	18.86
1-(3,4-dimethoxylphenyl)ethane-1,2-diol	4.77	208.1223	33.15	22.17	23.65	18.92
Gly-Pro-Pro	8.05	281.1769	30.5	22.17	31.7	29
Rishitin	4.81	236.2082	25.22	22.24	35.13	38.43
Phenylpropanoate	5.04	159.0978	25.57	22.26	23.72	21.35
Inosine	8.66	278.1138	24.83	22.37	16.71	15.15
N3-Metyladenine	9.92	157.1057	26.77	22.52	73.68	57.32
Val-Asn-Pro	5.02	342.2206	26.02	22.55	15.69	12.04
6-Phospho-D-gluconate	10.4	282.0445	35.53	22.56	20.92	12.81
N-Acetyl-beta-alanine	5.28	136.0747	26.16	22.62	16.54	14.21
omega-Carboxy-trinor-leukotriene B4	4.03	354.2169	29.56	22.68	28.33	30.88
Thiomorpholine 3-carboxylate	6.96	152.0519	26.6	22.68	41.01	116.08
Glu-Val	9.75	256.1546	31.43	22.69	21.07	16.91
Ecgonine	8.57	194.1349	27.66	22.8	39.23	34.7
3-Methoxy-4-hydroxyphenylethyleneglycol	5.12	193.1033	31.4	22.81	22.84	13.92
Isoetin 4'-glucuronide	7.4	499.1435	32.33	22.95	25.15	22.41
Coniferin	4.36	358.1846	30.39	22.99	20.04	16.67
Glu-Pro-Val	7.69	358.2237	29.6	23.02	40.77	36.86
Acremoauxin A	4.25	339.1897	25.43	23.03	24.3	21.46

Ergothioneine	10.25	238.1182	27.22	23.03	49.94	40.77
sn-glycero-3-Phosphocholine	9.6	265.1293	18.06	23.03	28.23	20.13
Met-Met-Thr	8.35	395.186	25.73	23.15	21.52	16.67
Lys-Tyr	8.03	324.2182	37.28	23.18	24.77	31.61
Paeonoside	4.88	343.1656	30.15	23.28	23.58	16.56
alpha-N-Phenylacetyl-L-glutamine	5.16	277.1537	30.69	23.31	32.22	22.17
Cytidine	9.15	252.1152	29.4	23.38	20.55	18.94
Proclavaminic acid	8.52	210.1218	31.18	23.38	22.76	19.41
Sorbitol 6-phosphate	9.55	268.0652	33.04	23.38	31.85	28.16
N4-(Acetyl-beta-D-glucosaminyl)asparagine	10.25	347.1728	32.16	23.4	67.98	120.14
GammaGlutamylglutamicacid	9.98	286.1288	31.16	23.43	25.53	32.19
3',5'-Cyclic dGMP	5.42	339.0856	29.73	23.44	18.59	20.91
L-Leucine	8.5	137.1145	29.8	23.47	24.26	16.43
Asn-Pro-Pro	5.23	340.2046	28.28	23.53	30.59	20.99
Ala-Leu-Thr-Pro	7.33	218.1755	29.67	23.54	24.56	18.33
Asp-Gly-His	7.83	339.1576	37.88	23.54	29.83	23.17
D-Glucose 6-phosphate	9.77	266.0495	32.02	23.59	25.57	21.39
Amabiline	4.45	298.2278	33.51	23.65	17.97	14.53
3-phosphoglucarate	9.86	296.0239	26.7	23.69	17.86	13.82
3-O-alpha-L-rhamnopyranosyl-3-hydroxydecanoic acid	5.13	350.2518	28.27	23.89	50.65	125.82
trans-Cinnamoyl beta-D-glucoside	4.36	325.1549	28.31	23.91	19.5	11.78
L-Glutamate	9.36	152.0696	27.9	23.95	15.53	13.87
1'-Acetoxychavicol acetate	4.65	247.1323	30.94	24.02	24.44	17.45
L-Methionine	8.76	154.0675	33.63	24.06	31.62	20.6
Chorismate	9.53	236.0807	35.69	24.07	14.9	16.31
Ile-Pro-Pro	7	341.2529	33.6	24.07	45.65	57.94
Pantetheine 4'-phosphate	7.25	369.133	28.27	24.09	37.33	32.94

5-Oxopentanoate	4.9	121.0639	30.06	24.17	34.19	29.74
2-Hydroxylauroylcarnitine	4.28	378.3298	33.96	24.19	27.73	33.63
Hydroxypropionylcarnitine	7.06	243.1594	31.8	24.2	32.63	28.33
3-Phospho-D-glycerate	10.12	189.0027	35.91	24.22	68.75	51.71
Glycogen	10.57	690.3013	38.94	24.23	48.31	43.31
(1-Ribosylimidazole)-4-acetate	7.32	268.1183	32.11	24.32	16.61	14.5
L-Ornithine	13.52	137.1063	30.27	24.35	42.25	28.04
[FA methyl(20:2)] methyl 5-hydroperoxy-6,8,9,11-bisepidioxy-12,14-eicosadienoate	4.27	435.295	21.81	24.37	35.4	42.96
5,6,7-Trimethoxycoumarin	4.51	248.1082	30.53	24.54	20.12	18.38
Thr-Ala-Asp	9.44	316.1586	36.36	24.59	27.73	34.93
N-Acetyl-L-glutamate	8.82	196.0869	31.86	24.61	20.85	18.63
L-thiazolidine-4-carboxylate	5.43	137.0329	28.64	24.63	26.25	23.03
Thr-Thr-Ala	7.63	302.1792	27.4	24.64	47.18	20.89
N-Acetylglutamine	7.61	195.1028	33.18	24.74	26.22	23.81
14-Dihydroxycornestin	4.22	324.179	27.38	24.89	19.82	18.09
captopril	5.08	226.107	30.07	24.92	45.12	128.72
(2R,3S)-3-Isopropylmalate	5.64	183.0916	28.16	24.93	48.02	64.97
Bendiocarb	4.5	234.1208	34.78	24.98	17.94	15.59
O-Acetylcarnitine	8.35	212.1455	33.99	24.99	32.56	25.11
Ile-Val	7.98	241.1993	34.47	25.06	24.83	21.33
1,2-dioctanoyl-1,2,6-hexanetriol	3.86	408.3763	19.73	25.08	25.73	33.66
Biotin amide	8.23	253.1371	37.88	25.11	24.87	22.01
L-methioninamide	5.29	153.0835	31.89	25.11	22.22	11.53
PC(16:0/P-18:0)	4.43	787.737	23.06	25.17	30.53	31.75
Monodehydroascorbate	9.4	181.044	29.91	25.21	42.76	135.59
Arbutin	5.16	284.1294	31.16	25.22	24	17.87
N-Methylnicotinium	14.2	187.1677	32.62	25.23	46.14	27.71

1alpha,5alpha-Dimercaptoandrostane-3alpha,17beta-diol	4.31	375.2465	28.36	25.25	20.46	17.73
3-Methylglutarylcarnitine	7.99	302.1953	34.88	25.26	52	136.41
D-Glyceraldehyde 3-phosphate	9.46	173.0079	40.55	25.28	25.71	32.79
2-Hydroxybutane-1,2,4-tricarboxylate	10.11	213.0657	30.78	25.33	24.95	18.5
N1,N8-diacetylspermidine	11.87	240.2154	32.69	25.35	21.44	18.03
NADP+	9.82	764.1449	32.01	25.35	23.27	18.52
Glycylproline	8.75	179.1079	29.74	25.39	21.44	17.74
[FA trihydroxy(18:1)] 9S,12S,13S-trihydroxy-10E-octadecenoic acid	4.17	348.3003	23.25	25.52	33.46	37.8
S-Adenosyl-L-homocysteine	9.26	398.1677	30.41	25.6	31.86	21.21
N,N-Dimethylaniline N-oxide	5.77	145.1104	28.22	25.68	18.48	16.4
L-Lysine	14.2	152.1253	31.31	25.74	42.22	33.94
8-Epiiridotrial glucoside	4.45	360.2002	33.09	25.77	21.18	17.15
L-Carnitine	9.29	168.1282	32.79	25.78	18.77	13.25
4-Oxoglutaramate	5.26	150.054	28.32	25.79	19.12	16.54
Asn-Cys-Ser	8.65	332.1279	28.38	25.79	18.71	17.16
Butyl 3-O-beta-D-glucopyranosyl-butanoate	5.15	336.2089	29.36	25.8	29.44	21.23
Ala-Pro-Ser	7.35	284.1688	37.12	25.82	26.31	23
Ethanolamine phosphate	9.95	143.0257	34.2	25.83	46.26	116.03
N-Acetylputrescine	14.24	136.1305	29.01	25.84	50.38	44.71
Leu-Ala	8.76	211.1615	30.58	25.88	48.29	114.52
Thr-Val	7	227.1564	35.22	25.91	33.12	28.74
2-Acetolactate	5.26	137.0588	29.33	25.93	28.23	20.47
8-Epiiridodial glucoside	4.92	346.2209	31.43	25.93	14.96	15.63
Val-Ser	8.5	212.1373	29.42	25.93	41.95	36.68
Ala-Ala-Ala	10.43	240.1517	34.31	26.01	27.61	25.39
3-Deoxy-D-manno-octulosonate	9.01	246.0954	31.33	26.02	30.2	27.57
nicotinamide guanine dinucleotide	9.39	700.1725	38.02	26.04	29.09	23.96

L-Rhamnulose	9.1	170.0883	30.47	26.05	28.66	18
Tetranor-PGF1alpha	4.26	316.2464	33.58	26.06	41.47	118.84
N-(3-oxooctanoyl)-L-homoserine	4.75	271.1816	32.48	26.17	22.21	18.51
trans-4-Hydroxycyclohexanecarboxylate	5.15	151.1017	29.15	26.28	21.36	17.32
Cytosine	9.15	115.0565	29.16	26.29	23.88	20.34
Ecgonine methyl ester	8.21	209.1538	28.93	26.3	23.33	18.33
N-Acetyl-L-leucine	5.08	181.1317	30.12	26.41	23.61	21.74
Xanthine	8.59	157.05	31.66	26.46	22.57	20.3
[ST] androst-16-en-3alpha-ol	8.7	156.1774	18.45	26.59	11.5	11.75
4-Trimethylammoniobutanoate	9.28	152.1333	31.64	26.59	62.53	24.56
N3-methylthymine	5.46	146.0783	30.09	26.62	24.25	29.89
2-Isopropylmaleate	5.32	165.081	29.53	26.64	20.26	18.63
Uplandicine	4.28	374.2349	29.75	26.69	19.11	16.85
3-Oxododecanoic acid	4.28	226.1966	32.92	26.72	18.64	13.97
Ala-Thr-Thr-Ser	7.54	392.2213	26	26.76	45.12	40.11
10-Hydroxydecanoic acid	4.21	198.1743	30.49	26.81	16.49	13.09
N-Formyl-L-methionine	5.16	183.0658	51.67	26.91	14.07	9.06
Artecanin	4.39	293.165	33.95	26.98	19.77	18.2
CMP	9.63	332.0816	34.84	26.99	50.49	115.34
O-beta-D-Glucopyranosyl-cis-zeatin	8.42	397.2179	35.29	27	29.87	21.46
2-C-Methyl-D-erythritol 2,4-cyclodiphosphate	9.63	283.0122	35.98	27.06	26.07	30.67
D-Ribose 5-phosphate	9.38	235.0357	30.48	27.07	41.87	38.34
ADP-D-glycero-D-manno-heptose	9.53	636.1494	42.68	27.12	29.7	26.74
Acetyl-CoA	5.18	832.2016	36.99	27.15	47.63	42.73
carbidopa	5.2	236.1284	30.33	27.18	22.53	15.8
3-Amino-3-(4-hydroxyphenyl)propanoate	9.49	190.1036	30.55	27.19	21.82	20.09
3-(3-Hydroxy-phenyl)-propanoic acid	4.75	175.0926	33.28	27.2	21.04	17.93

Diethyl (2R,3R)-2-methyl-3-hydroxysuccinate	4.92	213.1296	30.53	27.2	20.34	13.71
Leu-Asp-Pro	5.2	358.2238	33.04	27.27	25.21	22.21
Europine	4.31	345.2366	32.32	27.3	23.74	21.11
O-Propanoylcarnitine	7.83	227.1644	37.21	27.31	24.86	27.14
PE(16:0/P-18:1(11Z))	4.28	740.6646	26.12	27.35	16.85	17.74
Heptanoylcarnitine	4.24	287.2401	32.09	27.38	23.47	17.78
L-Cystathionine	10.45	229.0906	33.52	27.38	17.76	14.89
Hypoxanthine	8.53	141.055	39.05	27.43	9.35	8.41
Thiamin diphosphate	9.1	436.0768	33.53	27.45	26.54	23.76
Scorpioidin	4.11	292.1891	30.68	27.52	22.99	17.74
[FA hydroxy(18:0)] 9,13-dihydroxy-10-ethoxy-11-octadecenoic acid	3.97	378.3383	25.28	27.55	43.64	115.14
Tributyrin	5.13	317.2226	31.78	27.59	28.88	24.38
9-Oxononanoic acid	4.72	181.1397	31.25	27.71	23.79	20.1
Glu-Pro	9.16	254.1389	31.11	27.76	26.45	23.07
Ferulate	5.16	204.0908	32.22	27.81	12.13	10.96
Thr-Thr-Gln	7.62	361.2073	33.22	27.84	42.26	30.41
Glaucarubolone	8.99	217.1477	37.08	27.89	35.86	23.04
6-methyltetrahydropterin	8.51	188.1195	39.9	27.9	39.08	30.17
[FA hydroxy,dioxo(4:0/2:0)] 11R-hydroxy-9,15-dioxo-2,3,4,5-tetranor-prostan-1,20-dioic acid	4.32	342.1896	19.44	27.93	28.33	34.25
2-isocapryloyl-3R-hydroxymethyl-γ-butyrolactone	4.31	255.1947	35.86	27.97	21.24	29.01
3-4-5-Trimethoxycinnamicacid	4.91	250.1239	39.57	28	32.1	40.18
Dodecanedioic acid	4.21	242.1914	34.6	28.08	44.46	118.49
2-Dehydro-3-deoxy-L-rhamnonate	8.84	168.0727	34.87	28.11	30.39	24.34
3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	12.22	213.1771	34.43	28.13	47.61	105.33
D-Galactosamine	8.6	185.0991	34.08	28.13	18.73	14.7
2-Methylbutyroylcarnitine	5.49	257.2023	33.38	28.15	23.42	29.8

Ala-Cys-Pro-Tyr	10.13	472.2383	31.44	28.17	30.51	25.33
2,3,4,5-Tetrahydrodipicolinate	5.17	178.0763	46.13	28.19	21.08	18.13
N2-Succinyl-L-ornithine	9.12	241.1356	32.87	28.29	29.68	24.9
L-Albizziine	10.94	151.0775	37.61	28.31	27.19	24.95
L-Pipecolate	9.09	135.0988	31.83	28.31	23.83	14.47
3,4-Dihydroxyphenylacetaldehyde	4.28	160.0737	33.8	28.33	23.59	29.02
Homocarnosine	10.45	250.1553	34.79	28.33	29.62	24.8
1-Methyladenosine	10.58	292.1487	35.49	28.38	21.43	19.65
L-Histidine	10.44	161.0893	34.88	28.48	29.24	21.07
Ala-Pro	7.67	194.1268	31.27	28.49	20.57	18.81
Taurine	10.17	127.0213	35.44	28.6	24.68	19.88
L-Tryptophan	9.1	215.1262	31.78	28.65	44.11	116.8
Urea-1-carboxylate	9.06	106.0288	33.94	28.7	19.06	17.44
GTP	11.13	533.0239	48.43	28.78	59.97	53.84
Leu-Met-Pro	4.7	375.2411	32.33	28.79	34.62	30.55
Hydroxymethylphosphonate	9.03	112.9959	33.02	28.8	36.05	31.77
cis-Aconitate	10.35	180.0362	44.38	28.83	42.05	32.5
Thr-Ala	14.2	197.1185	32.52	28.84	40.86	16.3
3-Dehydroxycarnitine	5.06	152.1333	39.85	28.85	37.76	18.92
N6-Acetyl-N6-hydroxy-L-lysine	7.06	212.1375	64.45	28.85	38.95	34.11
Phosphoenolpyruvate	10.47	170.9919	32.37	28.92	39.66	35.67
N-Acetyl-leucyl-leucine	4.26	300.2351	35.46	28.94	17.52	15.26
Ala-Ser-Ser	9.86	272.1414	32.18	28.95	24.93	21.17
Tiglic acid	4.95	105.0689	35.46	28.95	34.94	27
dihydrothymidine	4.99	254.1388	36.22	28.98	47.02	121.24
UDP-glucose	9.86	581.1049	34.4	28.98	62.46	41.31
D-Fructose 1,6-bisphosphate	10.49	346.0158	34.3	29.04	20.86	14.49

Glu-Leu	9.02	271.1733	32.29	29.05	32.65	27.39
Daminozide	5.82	166.1043	37.13	29.08	23.39	19.84
Leu-Lys-Pro	12.98	373.2984	32.86	29.1	46.19	93.66
N-Acetyl-leu-leu-leu-tyr-amide	3.99	590.448	36.5	29.15	23.21	18.1
Oxotremorine	5.4	218.1814	48.33	29.18	17.22	13.06
N2-(D-1-Carboxyethyl)-L-arginine	9.79	255.1627	36.47	29.19	37.16	18.24
N-acetyl -D- glucosaminitol	9.26	231.1321	28.7	29.22	42.79	117.66
Choline phosphate	9.39	188.0826	32.84	29.27	90.27	47.56
Propanoyl phosphate	8.17	157.013	35.9	29.35	41.08	26.35
[SP (16:0)] N-(hexadecanoyl)-sphing-4-enine-1-phosphocholine	4.56	741.6965	21.25	29.38	22.62	23.5
Gamma-Aminobutyryl-lysine	13.74	241.1912	34.34	29.47	41.95	37.63
Riboflavin	7.58	393.1943	34.19	29.47	23.56	20.43
Succinate	9.45	122.0398	32.93	29.47	24.69	22.61
Psicofuranin	8.22	308.1437	37.87	29.52	21.31	15.55
Glu-Leu-Pro	7.31	373.2427	34.14	29.64	35.57	45.47
Glutathione	9.24	317.1167	35.56	29.7	38.09	26.63
6-Deoxy-L-galactose	6.45	170.0883	22.52	29.71	26.99	24.42
L-Aspartate	9.58	137.0507	33.81	29.75	28.89	18.56
[FA (9:1)] 3-nonenal	4.2	149.1498	22.39	29.79	24.23	29.03
Syringin	4.55	389.1985	33.15	29.84	22.18	18.1
5-O-(1-Carboxyvinyl)-3-phosphoshikimate	9.59	334.0579	39.17	29.86	29.16	25.61
GMP	10.15	373.0914	37.66	29.92	23.98	21.36
beta-Citryl-L-glutamic acid	11.29	332.1061	38.66	29.96	17.93	15.41
Guanosine	9.56	293.1245	33.68	30.01	29.09	24.91
N2-Succinyl-L-arginine	9.64	284.161	51.17	30.04	27.68	15.67
NAD+	8.98	684.1785	35.8	30.08	26.57	16.45
D-Phenylalanine	8.04	174.1087	33.36	30.2	32.67	28.2

PS(36:1)	3.94	831.6907	23.38	30.2	16.45	19.24
5,6-Dihydroxy-3-methyl-2-oxo-1,2-dihydroquinoline	4.5	201.0913	35.91	30.21	43.07	101.81
Nicotinamide	7.09	128.0679	37.49	30.27	14.1	12.6
DL-2-Aminooctanoicacid	9.23	167.1523	36.21	30.38	29.69	22.75
7-oxoheptanoate	7.03	151.1017	33.65	30.4	29.57	22.43
Hymexazol O-glucoside	10.04	271.1179	35.78	30.41	33.85	29.1
N-Succinyl-2-L-amino-6-oxoheptanedioate	10.26	300.1161	33.89	30.41	29.83	24.15
N2-N2-Dimethylguanosine	7.57	323.1625	37.22	30.42	21.47	18.92
L-Threonine	9.79	123.0715	38.05	30.47	76.67	22.53
CoA	5.18	276.7745	38.29	30.49	25.84	20.14
[SP amino,trihydroxy,hydroxy,methyl,oxo(20:0)] 2S-amino-3R,4R,5S-trihydroxy-2-(hydroxymethyl)-14-oxo-eicos-6E-enoic acid	4.52	438.3417	21.08	30.52	18.73	20.03
dTDP-4-acetamido-4,6-dideoxy-D-glucose	5.23	607.1673	35.57	30.52	36.69	31.97
Lotaustralin	8.4	272.1575	36.27	30.57	59.51	20.74
LL-2,6-Diaminoheptanedioate	9.59	197.1184	23.08	30.58	39.56	26.25
3-Hydroxyanthranilate	7.3	160.0657	36.2	30.65	35.44	30.19
SM(d18:1/22:0)	4.5	831.8096	25.64	30.66	27.4	28.75
SM(d19:1/24:1(15Z))	4.49	874.852	29.86	30.7	20.53	22.43
S-propyl-D-cysteine	5.1	169.0864	33.99	30.8	9.83	8.23
N4-Acetylcytidine	7.5	296.1323	42.49	30.86	20.8	25.44
Traumatic acid	4.46	240.1758	33.95	30.94	47.2	150.8
Tecostanine	6.99	194.1987	40.96	31.08	24.02	17.06
2-Deoxy-D-ribose 5-phosphate	8.75	219.0408	36.77	31.12	24.2	21.99
(4E)-2-Oxohexenoic acid	5.19	134.0672	34.53	31.16	28.07	25.66
CDP-ethanolamine	10.01	457.097	36.08	31.21	19.1	17.49
N4-Acetylaminobutanal	7.1	135.0988	39.32	31.25	27.12	24.35
[FA trihydroxy(2:0)] 9S,11R,15S-trihydroxy-2,3-dinor-5Z,13E-prostadienoic acid-cyclo[8S,12R]	4.14	344.2687	29	31.34	35.54	40.66

N-hydroxyvaline	6.99	138.0904	40.77	31.38	18.76	17.1
1-Deoxy-D-altro-heptulose 7-phosphate	9.49	281.0685	38.28	31.39	25.2	22.67
sn-glycero-3-Phospho-1-inositol	9.91	343.0964	37.98	31.51	51.36	121.14
GDP-mannose	10.7	621.1302	42.24	31.63	30.07	22.26
Val-Asp-Pro	5.22	343.2054	37.7	31.68	31.91	27
2S-Hydroxytetradecanoic acid	4.13	258.25	35.5	31.78	21.94	20.04
1-(beta-D-Ribofuranosyl)-1,4-dihydronicotinamide	7.78	267.1422	35.63	31.95	22.87	15.14
Thiamin monophosphate	9.01	356.1103	37.26	31.95	25.14	18.98
2,3-Bisphospho-D-glycerate	10.62	268.9692	40.3	31.96	49.28	43.74
Glucosylgalactosylhydroxylysine	13.1	504.2657	37.4	31.96	69.64	29.84
CDP-choline	9.45	502.1537	35.86	31.99	25.53	23.51
L-Cysteate	10.15	172.0144	35.74	32.01	44.82	21.63
Ile-Lys-Pro	8.78	373.2984	38.62	32.04	30.96	24.01
Gly-Arg	10.85	239.1596	38.99	32.1	22.95	11.95
Carnosine	10.22	235.1363	37.66	32.15	41.59	32.47
Graphinone	4.29	312.2152	35.98	32.2	23.51	31.24
2-oxosuberate	7.04	196.0948	35.88	32.21	19.83	9.02
(10S)-Juvenile hormone III acid diol	4.19	285.2326	28.31	32.23	29.87	31.54
D-Glycerate	8.26	109.0365	35.63	32.24	59.42	28.35
D-Apiitol	8.59	141.09	39.33	32.25	44.18	114.74
Adenosine	7.86	277.1297	36.63	32.3	43.56	122.02
Leu-Val-Pro	4.77	343.269	39.99	32.32	49.51	132.36
Leu-Ala-Asp	5.22	330.202	35.47	32.34	25.97	14.89
Ala-Leu-Val-Pro	8	417.3156	46.24	32.36	41.45	36.28
Succinyladenosine	10.09	397.1541	35.78	32.36	32.61	27.92
N-Acetyl-D-glucosaminyl-N-acetylmuramoyl-L-Ala-D-glutamyl-6-carboxy- L-lysyl-D-alanine	8.9	976.5139	39.87	32.38	37.45	31.61
(-)-Menthyl O-beta-D-glucoside	4.91	334.257	36.26	32.5	29.75	25.86

Pyrimidine nucleoside	5.22	223.1251	38.04	32.5	41.01	30.81
N-Acetylneuraminate	8.7	320.1426	37.71	32.52	19.04	15.15
[FA (7:0/2:0)] Heptanedioic acid	5.15	167.0967	30.61	32.62	38.49	39.45
2,6-Dioxo-6-phenylhexanoate	4.65	232.1133	39.69	32.62	20.87	17.65
gamma-L-Glutamyl-L-cysteinyl-beta-alanine	8.35	332.1358	41.38	32.64	58.19	31.05
L-Ala-L-Glu	7.5	226.1166	38.17	32.65	55.7	120.54
sn-glycero-3-Phosphoethanolamine	10.07	220.0724	38.68	32.65	19.34	24.84
(2S,3R)-3-Hydroxybutane-1,2,3-tricarboxylate	7.06	213.0657	38	32.67	26.63	28.83
Trimethylolpropane triacrylate	4.42	311.1756	40.88	33.1	21.99	20.04
Monomethyl sulfate	7.06	112.9864	38.92	33.21	28.18	23.95
[FA hydroxy(14:0)] 3,11-dihydroxy-tetradecanoic acid	4.45	274.245	32.4	33.28	24.93	27.98
(R)-glycerol 1-acetate	5.18	139.0745	37.91	33.34	29.92	26.15
3-Isopropylbut-3-enoic acid	4.51	135.1068	37.11	33.43	24.51	22.56
3-(4-Hydroxyphenyl)lactate	9.6	191.0876	40.72	33.44	52.27	122.29
(-)-erythro-(2R,3R)-dihydroxybutylamide	7.02	123.0716	24.22	33.65	25.91	16.98
Ascorbate	9.3	182.0519	46.28	33.71	32.88	24.62
Lys-Val	13.67	256.2102	39.43	33.83	26.26	22.28
Asp-Pro-Arg	9.23	401.2409	38.82	33.84	27.84	18.5
N-Hydroxy-2-acetamidofluorene	7.54	254.1441	39.38	33.84	26.96	24.73
3-O-Sulfogalactosylceramide(d18	3.84	939.8056	32.69	33.93	25.1	21.07
[SP] (3'-sulfo)Galbeta-Cer(d18:0/2-OH-24:0)	3.85	955.8003	28.82	33.94	35.94	40.01
N-Formimino-L-glutamate	7.98	180.0841	41.64	33.95	29.32	25.15
CoA-glutathione	9.62	567.1941	39.85	34.01	38.09	28.47
PE(36:1)	4.42	786.6974	27.59	34.04	29.89	30.38
N-Acetyl-aspartyl-glutamate	9.96	315.1271	37.95	34.14	18.85	13.58
2-Octenoylcarnitine	4.3	300.2435	39.2	34.17	17.23	15.42
O-Butanoylcarnitine	7.47	242.1833	38.07	34.28	58.41	110.65

Deoxycytidine	7.43	236.1203	42.3	34.43	25.18	14.7
Met-Ser	10.56	244.1095	40.36	34.72	29.24	37.2
UDP	9.95	413.0321	44.52	34.79	31.97	27.15
Triethanolamine	7.96	155.125	38.32	34.85	33.99	30.54
1,2-Dihydroxy-5-(methylthio)pent-1-en-3-one	7.08	168.0549	38.67	34.94	28.37	20.56
Pinolidoxin	4.27	356.2323	39.87	35.02	19.52	17.02
Val-Val	9.23	226.1804	40.56	35.15	35.4	30.02
[FA hydroxy(16:0/2:0)] 9-hydroxy-hexadecan-1,16-dioic acid	4.52	318.2623	33.06	35.26	20.98	23.46
N-Acetyl-L-histidine	7.51	205.1065	43.8	35.39	24.28	18.6
Phe-Ala-Pro	5	350.2248	39.51	35.4	14.96	13.74
Adenine	7.87	140.071	40.32	35.56	17.24	13.49
UDP-N-acetyl-D-glucosamine	9.2	624.1382	42.52	35.59	56.7	116.69
(R)-3-Hydroxybutanoate	7.11	108.0606	46.16	35.61	44.74	36.56
Leu-Leu-Pro	4.59	358.2877	40.54	35.62	20.49	17.29
Ile-Leu-Gly-Pro	4.82	417.3154	41.62	35.68	28.72	23.4
Calystegin B2	8.52	182.1074	27.54	35.85	30.12	26.42
gamma-L-Glutamylputrescine	13.33	226.1723	35.84	35.88	38.6	32.13
(R)-3-((R)-3-Hydroxybutanoyloxy)butanoate	7.24	198.1105	39.72	35.94	35.51	23.26
Sucrose	10.35	354.156	40.05	35.94	28.24	20.23
Hydroxyhexanoycarnitine	7.59	288.216	46.72	36.02	30.78	22.91
Thr-Pro	5.81	225.1407	40.47	36.05	20.33	27.31
AminoDAHP	9.05	294.0639	40.13	36.12	43.23	32.43
Furfural diethyl acetal	4.56	179.124	42	36.22	32.47	28.75
N-(octanoyl)-L-homoserine	7.2	257.2023	40.5	36.24	55.07	120.41
Glycerophosphoglycerol	8.49	252.0703	40.29	36.31	30.03	24.51
NADH	8.48	686.195	41.63	36.4	21.41	18.71
D-Sedoheptulose 7-phosphate	9.85	297.0633	41.31	36.5	12.84	11.73

[FA (7:1)] 2-heptenal	4.61	119.1119	27.44	36.75	22.98	27.39
Phosphodimethylethanolamine	8.93	173.0635	36.72	37.15	39.87	26.32
Leu-Leu-Ala	8.34	330.2654	44.56	37.18	37.31	33.4
(R)-S-Lactoylglutathione	8.53	392.148	43.11	37.24	47.25	41.37
[FA hydroxy,oxo(7:0/2:0)] 4-hydroxy-2-oxo-Heptanedioic acid	7.82	197.0709	34.51	37.31	22.29	26.72
[FA trihydroxy(16:0)] 2,15,16-trihydroxy-hexadecanoic acid	4.54	320.2779	36.13	37.53	13.43	15.87
[FA oxo(8:0)] 3-oxo-octanoic acid	5.03	166.1207	37.1	37.7	20.75	23.51
Ala-Thr-Ala-Ser	5.74	361.2075	49.43	37.85	45.79	39.56
O-Acetyl-L-homoserine	7.11	167.0885	41.99	37.96	21.87	19.85
L-Fucose 1-phosphate	8.65	250.0547	65.68	38.08	43.73	25.02
[FA dioxo(8:0)] 4,7-dioxo-octanoic acid	5.1	180.1	28.9	38.22	19.46	23.73
Leu-Phe-Asn	5.03	411.2679	53.43	38.23	27.81	22.11
nonulose 9-phosphate	7.04	184.0601	42.29	38.27	32.58	18.37
Phenol sulfate	4.87	180.0184	42.32	38.35	17.95	16.02
PC(30:0)	4.46	743.6564	35.74	38.36	20.88	21.3
(2S)-2-Isopropylmalate	7.79	183.0915	35.36	38.47	26.31	18.44
2-(alpha-Hydroxyethyl)thiamine diphosphate	8.83	482.1097	36.3	38.5	51.18	27.48
L-Glutamate 5-semialdehyde	7.06	136.0747	42.93	38.5	44.94	38.23
Methyl5-(but-3-en-1-yl)amino-1,3,4-oxadiazole-2-carboxylate	5.57	205.1065	48.93	38.64	47.56	120.12
(1S,4S)-4-Hydroxy-3-oxocyclohexane-1-carboxylate	6.35	165.081	28.49	38.7	45	130.04
NG,NG-Dimethyl-L-arginine	13.29	210.1694	35.18	38.7	31	39.37
Ala-Val-Val-Pro	4.9	402.2967	43.9	38.85	45.37	133.23
D-Xylulose	7.53	155.0692	43.65	39.01	40.17	36.53
Glu-Gly	7.06	211.0978	55.67	39.18	34.61	31.3
Sulcatone	4.11	134.1309	36.1	39.25	26.73	24.43
S-Acetyldihydrolipoamide	7.28	259.1192	36.33	39.39	39.76	29.27

Asp-Cys-Ser-Tyr	17.35	505.2042	47.86	39.63	40.92	32.96
Kinobscurinone	8.17	324.1119	45.14	39.7	64.32	40.31
Mycosporine	5.49	272.1575	49.22	39.72	60.75	54.73
[FA trihydroxy(2:0)] 9S,11R,15S-trihydroxy-5Z,13E-prostadienoic acid 1(3)-glyceryl ester	4	451.3535	31.97	39.79	25.42	28.08
Phenolsulfonphthalein	5.05	373.1192	50.6	39.79	28.18	21.58
N-Methylpelletierine	5.27	164.1607	48.83	39.88	25.6	22.51
[SP (24:0)] N-(tetracosanoyl)-1-beta-glucosyl-sphing-4-enine	4.38	859.8487	39.66	39.93	43.59	52.27
Acetylenedicarboxylate	9.43	118.0086	37.74	40.25	20.38	18.42
L-Erythrulose	7.35	124.0555	46.24	40.45	26.37	21.29
dAMP	8.27	341.1011	38.33	40.51	34.53	29.64
[FA amino(7:0)] 2R-aminoheptanoic acid	12.78	152.1333	37.31	40.66	32.72	37.31
Berberastine	10.08	371.1763	45.36	40.74	42.08	52.32
Uridine	8.3	253.0992	46.49	40.8	40.67	54.09
[FA (8:1)] 2Z-octenoic acid	4.42	150.1258	28.58	40.92	49.02	110.87
Leu-Thr	10.51	242.1753	49.34	41.01	34.76	42.86
(+)-pinoresinol	9.19	378.207	60.75	41.15	37.54	23.8
CMP-N-acetylneuraminate	9.97	634.2121	38.85	41.41	21.36	27.74
PC(32:1)	4.44	771.678	33.12	41.49	24.28	24.89
Dephospho-CoA	7.13	708.2186	46.5	41.82	50.36	44.04
2-C-Methyl-D-erythritol 4-phosphate	10.27	221.0565	41.14	41.98	47.06	57.7
Met-Ala-Ser	10.83	318.1564	57.07	42.22	35.15	25.87
Shikimate	7.09	181.0759	47.85	42.59	42.95	35.52
2-Dehydro-D-xylonate	9.59	169.0486	52.56	42.79	22.4	16.25
Tridecanoylglycine	4.03	286.2643	56.61	42.93	18.37	13.31
[FA hydroxy(18:1)] 9,10-dihydroxy-12Z-octadecenoic acid	3.97	332.3054	32.39	43.15	22.23	24.15
2-Benzothiazolesulfonamide	4.04	221.0098	54.63	43.43	21.37	19.61
Ser-Val-Ser-Ser	5.67	392.2214	55.18	43.74	27.61	22.93

Botrydial	4.04	327.234	82.02	43.76	39.24	28.64
alpha-D-Glucosamine 1-phosphate	9.13	265.0657	51.71	43.79	52.02	41.25
[SP amino,trihydroxy(20:0)] 2S-amino-5S-acetoxy-3R,4R,14R-trihydroxyeicos-6E-enoic acid	4.68	453.3606	36.88	43.84	20.22	21.11
[SP (26:0)] N-(17Z-hexacosenoyl)-sphing-4-enine-1-phosphocholine	4.48	889.8703	38.66	43.95	56.75	63.24
7,8-H2pterin-6-ylmethyl-l-(4-aminophenyl)-1-deoxy-D-ribitol 5'-phosphate	9.14	500.1917	31.15	44.13	34.39	45.37
N-Acetyl-ala-ala-ala-methylester	7.09	299.1878	55.8	44.25	63.53	81.25
bacimethrin	5.84	161.0894	54.91	44.61	67.33	55.49
SM(d18:0/18:1(9Z))	4.54	771.7343	34.63	44.63	36.88	37.16
[FA oxo,hydroxy(18:3)] 12-oxo-14,18-dihydroxy-9Z,13E,15Z-octadecatrienoic acid	4.25	342.253	37.24	44.94	42.51	103.57
Leu-Asn	7.05	255.1706	51.08	45.26	20.85	19.07
[FA methyl(5:1/5:2/8:0)] methyl 8-[3,5-epidioxy-2-(3-hydroperoxy-1-pentenyl)-cyclopentyl]-octanoate	4.12	418.3238	35.22	45.37	27.31	30.86
[SL (6:0/6:0)] 2,3-di-0-hexanoyl-alpha-glucopyranose	4.87	394.2694	37.38	45.39	23.76	27.54
[FA (8:0)] octanoic acid	4.08	152.1414	36.56	45.42	26.17	29.96
PC(34:1)	4.43	801.7163	34.17	45.69	23.93	24.27
SM(d18:1/23:0)	4.5	846.8297	38.38	45.9	21	24.6
Porphobilinogen	7.04	236.1284	52.77	46.12	48.25	42.44
[FA (14:5)] 5,7,9,11,13-tetradecapentaenoic acid	4.44	232.177	42.85	46.17	46.09	49.26
[FA trihydroxy(4:0)] 2,3,4-trihydroxy-butanoic acid	8.41	140.0504	36.93	46.2	30.07	39.04
[FA hydroxy,oxo(16:0)] 9-hydroxy-16-oxo-hexadecanoic acid	4.44	302.2674	43.14	46.39	24.69	27.12
2-Amino-5-formylamino-6-(5-phosphoribosylamino)pyrimidin-4(3H)-one	9.57	391.1019	65.78	46.54	36.59	31.09
[FA (12:4/2:0)] 2E,4E,8E,10E-Dodecatetraenedioic acid	4.67	234.1289	35.85	46.79	21.14	21.72
[FA amino(11:0)] 11-amino-undecanoic acid	7.39	212.2092	34.57	47.06	11.92	15.71
MG(0:0/16:1(9Z)/0:0)	3.96	347.3241	36.7	47.8	20.01	21.16
27-nor-campestan-3beta,4beta,5alpha,6alpha,7beta,8beta,14alpha,15alpha,24-nonol	3.98	543.4195	41.68	47.89	37.82	43.58

Tetranor-PGE1	4.18	314.231	41.72	48.19	31.65	34.02
ADT	7.75	120.0539	36.5	48.45	35.44	31.78
UDP-glucuronate	10.87	595.0841	72.34	48.5	32.63	19.8
Val-Gly-Pro	5.18	283.1929	72.12	48.67	14.84	8.63
Asp-Leu-Ser-Ser	10.33	436.2388	44.22	48.83	47.42	33.56
PS(40:4)	3.92	885.7197	41.65	48.87	22.28	23.72
Leu-Val	4.71	241.1993	58.36	49.21	45.32	37.36
Val-Gly	7.05	181.1236	54.41	49.32	39.45	50.27
Queuine	12.54	289.157	65.35	49.83	62.56	45.9
N6,N6-Dimethyladenosine	13.32	307.1676	79.19	50.83	58.88	52.46
L-Xylulose	6.38	155.0691	67.51	51	29.07	18.89
3,4-Dehydrothiomorpholine-3-carboxylate	7	150.0362	65.23	51.78	63.42	35.11
(8S,9S,10S)-8,10-dimethyl-1-octalin	4.15	176.1961	34.34	51.83	45.6	52.74
PC(36:1)	4.42	831.7541	45.5	52.1	19.97	20.36
[FA hydroxy(7:0)] 2-hydroxy-heptanoic acid	4.73	153.1174	38.02	52.27	23.1	25.2
campest-22E-en-3beta,4beta,5alpha,6alpha,8beta,14alpha,15alpha,25,28-nonol	3.94	556.4228	41.68	53	27.1	28.19
(S)-Dihydroorotate	8.01	163.0492	59.08	53.26	30	27.01
8,8a-Deoxyoleandolide	4.01	392.3175	41.63	53.71	19.72	14.22
Leu-Val-Gly	12.31	300.2273	87.74	53.93	42.54	27.17
[FA hydroxy(9:0)] 2-hydroxy-nonanoic acid	4.97	183.1553	40.21	55.23	26.92	30.31
Galbeta1-3GalNAcbeta1-4(NeuAcalpha2-8NeuAcalpha2-3)Galbeta1-4Glcbeta-Cer(d18:1/20:0)	4.99	1018.7871	47.6	55.24	31.34	31.67
GDP	10.61	453.0576	40.66	55.25	37.37	30.24
[FA trihydrox] 9S,11R,15S-trihydroxy-2,3-dinor-13E-prostaenoic acid-cyclo[8S,12R]	4.42	346.2846	51.38	55.26	21.64	22.17
1,8-Diazacyclotetradecane-2,9-dione	5.41	238.2077	51.97	55.34	64.42	82.81
[FA hydroxy(11:2/11:2)] 2R,9R-dihydroxy-3S,4S,7S,8S-diepoxy-5E,10-undecadien-1-ol	4.86	239.1362	42.79	55.97	48.52	158.45

[FA oxo(5:2/5:0/6:0)] (1R,2R)-3-oxo-2-pentyl-cyclopentanehexanoic acid	4	284.2567	44.88	56.29	25.37	30.66
N-Carbamoyl-L-aspartate	10.16	181.0597	62.52	56.32	34.88	31.67
5'-methylthioformycin	7.05	308.1259	63.58	56.37	27.01	22.46
Gamma-Glutamylglutamine	9.38	285.1448	70.98	56.71	57.31	17.71
2-Hydroxyethylphosphonate	9.27	128.0149	58.36	57.19	33.87	30.49
D-Alanyl-D-alanine	7.25	166.1046	70.65	57.24	24.18	21.09
N-Acetylmethionine	6.99	198.0848	77.08	57.83	20.97	12.05
3'-AMP	6.14	357.0961	46.68	58.16	70.94	92.52
[FA (11:0/2:0)] Undecanedioic acid	7.04	227.1725	44.44	58.55	39.1	49.12
Hydantoin-5-propionate	7.05	178.0682	80.9	58.84	34.72	31.73
D-Glucuronate	9.4	200.0624	80.65	59.13	25.34	18.48
[FA hydroxy(10:0)] N-(3S-hydroxydecanoyl)-L-serine	4.64	288.2161	41.92	59.28	20.84	24.33
Acetylcholine	7.98	152.1333	70.87	59.29	20.52	18.86
beta-Cymaropyranose	7.52	169.1122	53.91	59.44	48.13	110
D-Sorbitol	9.79	188.0989	69.7	59.46	21.4	17.54
Oxalosuccinate	7.08	196.0311	66.81	60.08	38.58	30.56
L-Histidinol	15.24	147.11	80.72	60.33	25.13	21.89
Nona-2,6-dienal	4.93	147.1341	44.62	60.54	27.56	24.65
Leu-Lys	12.43	271.2292	65.35	60.66	17.6	15.17
FAD	5.19	812.2468	42.9	60.69	49.43	105.28
Leu-Pro	7.06	239.1837	72.77	61.28	39.46	51.1
1,6,6-Trimethyl-2,7-dioxabicyclo[3.2.2]nonan-3-one	4.36	194.143	43.13	62.52	26.52	31.62
N-acetyl-(L)-arginine	9.95	224.1486	88.31	62.63	14.79	6
PC(36:2)	4.43	829.7383	43.31	62.74	19.35	19.48
[FA (9:2)] 2,6-nonadienoic acid	4.89	163.129	57.39	63.03	37.03	48.47
[SP amino,hydroxy(4:0/20:0)] 2S-amino-3R,4R,5S,14R-tetrahydroxyeicos-6E-enoic acid	4.21	409.3434	48.14	63.2	34.32	40

3-(4-Hydroxyphenyl)pyruvate	7.03	189.0722	57.57	63.89	26.64	24.55
(S)-2-Aceto-2-hydroxybutanoate	7.06	152.0777	64.81	63.94	44.49	40.47
Pantothenate	7.04	228.1404	74.27	64.33	78.91	30.63
7-Deoxyloganate	5.16	376.1952	83.01	64.34	23.37	18.02
[FA (17:1/2:0)] 8E-Heptadecenedioic acid	4.11	315.2705	44.11	64.45	25.34	28.54
Sinalbin	9.11	438.0839	45.5	64.57	24.68	19.72
[FA trihydroxy(2:0)] 9S,11,15S-trihydroxy-2,3-dinor-thromboxa-5Z,13E-dien-1-oic acid	4.21	360.2638	48.2	64.97	47.47	120.32
L-Citrulline	10.44	181.1155	119.8	68.37	15.03	10.81
[FA oxo(12:1)] 12-oxo-10E-dodecenoic acid	4.36	224.1809	51.53	69.02	29.67	31.89
PC(32:0)	4.44	773.6943	58.78	69.56	29.2	29.31
[FA oxo(8:1)] 5-oxo-7-octenoic acid	4.85	164.105	59.56	70.8	25.2	27.13
a Cysteine adduct	7.06	197.0452	97.3	73.52	29.14	24.5
Thiamine aldehyde	7.15	274.1285	103.41	74.12	28.57	24.77
[6]-Gingerol	4.05	311.2393	62.15	75.19	29.14	31.17
Citrate	10.72	198.0468	86.17	75.34	29.5	26.18
SM(d18:1/20:0)	4.52	801.772	55.92	75.45	22	22.01
Acalyphin	10.1	374.1624	92.34	80.41	40.84	52.17
ATP	10	517.029	62.38	80.79	54.04	116.4
[SP (18:0)] N-(octadecanoyl)-sphing-4-enine-1-phosphate	4.56	681.6291	63.39	81.15	8.74	9.18
SM(d18:1/24:1(15Z))	4.48	859.8318	57.32	85.89	27.01	28.7
Stealthin C	9.24	325.1435	61.79	86.67	25.54	26.7
[PK] 5-O-Methylvisamminol	4.29	306.1682	72.68	88.48	26.47	31.57
[FA hydroxy(13:0)] 2-hydroxy-tridecanoic acid	4.34	243.2311	52.06	90.54	42.49	107.95
SM(d18:1/18:1(9Z))	4.54	769.7184	53.98	90.72	33.99	34.05
1alpha,5alpha-Epidithio-17a-oxa-D-homoandrostan-3,17-dione	4.2	385.1944	54.32	91.71	21.26	25.07
[FA methyl(18:0)] methyl 9-hydroperoxy-10,12,13,15-bisepidioxy-16E-octadecenoate	4.39	407.2728	50.51	91.94	30.44	31.7

ADP	9.32	437.0629	55.36	92.27	25.03	22.11
[FA methyl(14:0/2:0)] 3-methyl-tetradecanedioic acid	4.01	287.2485	52.2	93.62	23.87	27.29
[FA methyl,hydroxy,oxo(18:0)] methyl 9,12-dihydroxy-13-oxo-10-octadecenoate	4.27	361.3035	45.72	95.27	39.13	42.08
16-bromo-7E,15E-hexadecadien-5,13-diynoic acid	5.61	338.1096	56.47	95.95	38.53	50.6
[FA hydroxy(15:0)] 2-hydroxy-pentadecanoic acid	4.06	273.2692	75.41	97.19	34	38.23
Sinapyl alcohol	4.87	221.1255	51.75	98.9	24.71	22.21
N6,N6,N6-Trimethyl-L-lysine	13.26	197.1822	53.11	99.66	26.05	19.99
[FA (8:2)] 2,4-octadienal	4.38	132.1152	53.48	99.72	44.88	53.86
adenosine thiamine triphosphate	8.63	775.1624	56.3	99.76	22.4	15.25
10,16-Dihydroxyhexadecanoic acid	4.34	304.283	45.69	101.06	20.09	18.34
12-Hydroxydodecanoic acid	4.36	228.2122	54.3	102.52	27.8	21.49
[Fv hydrox] 4,2'-Dihydroxychalcone 4-glucoside	5.2	423.2005	52.57	103.36	25.96	14.4
[FA hydroxy(9:1)] 4-hydroxy-2-nonenal	4.88	165.1447	46.16	103.49	51.99	55.27
PS(36:2)	3.93	829.675	78.2	103.81	23.77	26.6
[ST hydrox] N-(3alpha-hydroxy-5beta-cholan-24-oyl)-taurine 3-sulfate	4.07	589.3438	57.44	105.44	28.22	31.9
SM(d18:0/24:1(15Z))	4.49	861.8493	64.1	106.05	80.16	89.27
[FA dioxo,hydroxy(4:0/2:0)] 9,15-dioxo-11R-hydroxy-2,3,4,5-tetranor-prostan-1,20-dioic acid	4.38	344.2053	49.54	107.98	27.75	29.81
[FA (10:0/2:0)] Decanedioic acid	7.12	212.1536	46.26	108.49	20.94	21.91
(-)-11-hydroxy-9,10-dihydrojasmonic acid 11-beta-D-glucoside	5.12	408.249	39.43	108.54	24.35	27.3
PS(40:6)	3.92	881.6883	87.68	108.91	21.02	21.83
[FA (10:1/2:0)] 2E-Decenedioic acid	4.76	210.138	59.63	109.19	22.59	23.95
[FA hydroxy,methoxy(18:0)] 12,13-dihydroxy-11-methoxy-9-octadecenoic acid	4.04	363.3193	52.84	109.6	20.99	25.48
UMP	9.31	333.0657	53.24	109.68	50.04	101.17
[ST dimethyl(3:0/3:0)] (5Z,7E)-(1S,3R)-24,24-difluoro-26,27-dimethyl-24a-homo-9,10-seco-5,7,10(19)-cholestatriene-1,3,25-triol	4.2	277.2773	55.02	110.21	22.54	24.02

N6-(1,2-Dicarboxyethyl)-AMP	10.46	245.5831	61.66	110.42	24.03	21.79
[FA (13:0/2:0)] Tridecanedioic acid	5.14	257.2104	51.39	110.51	27.92	30.01
3-Hydroxytetradecanedioicacid	5.14	288.2243	52.18	111.57	28.27	25.35
Homocysteinesulfinicacid	7.06	171.0384	59.68	113.64	55.2	46.35
(3S,4S)-3-Hydroxytetradecane-1,3,4-tricarboxylate	4.37	363.2555	65.28	114.47	27.19	32.19
isopropyl β-D-galactopyranoside	7.05	231.1404	44.74	114.6	72.43	60.86
[FA (14:0/2:0)] Tetradecanedioic acid	5.04	272.2294	50.84	114.88	12.05	13.76
[FA oxo,amino(6:0)] 3-oxo-5S-amino-hexanoic acid	7.08	151.0937	49.67	115.06	26.25	29.05
[FA oxo,hydroxy(5:1/5:0)] (1S,2R)-3-oxo-2-(5'-hydroxy-2'Z-pentenyl)-cyclopentaneacetic acid	4.76	238.1601	53.09	115.28	31.85	33.6
Deidaclin	4.94	283.1453	49.84	117.25	24.24	16.09
Prostaglandin F1alpha	4.13	376.3222	43.39	117.56	43.63	109.28
Ginkgolide A	4.07	428.2087	43.75	117.76	24.67	29.57
N(alpha)-t-Butoxycarbonyl-L-leucine	4.63	242.1834	56.4	118.56	21.55	19.42
5-6-DihydroxyprostaglandinF1a	4.15	408.3125	61.45	119.1	54.17	65.64
Ala-Pro-Pro	8.07	296.1961	54.59	119.27	28.72	25.62
AMP	8.67	357.0964	52.98	120.8	22.08	16.77
UDP-L-Ara4FN	9.52	578.1052	62.62	121.42	54	120.32
Genipin	4.94	237.1204	58.39	124.3	27.26	23.16
Prostaglandin G2	4.09	388.2862	53.35	125.87	21.24	18.99
2,3,6-Trihydroxypyridine	7.54	132.0435	43.29	125.88	27.34	24.45
[SP] Fumonisin C4	3.98	708.4929	118	126.56	27.93	33.43
Ala-Pro-Tyr	5.17	366.22	45.31	129.34	12.36	9.54
Camptothecin	8.71	368.1764	54.9	129.35	29.85	21.4
[SP] (3'-sulfo)Galbeta-Cer(d18:0/2-OH-22:0)	3.86	925.7623	57.3	129.58	26.54	27.14
6-(2-Hydroxyethyl)-5,6-dihydrosanguinarine	7.9	399.1983	46.24	130.66	41.25	32.15
(10S)-Juvenile hormone III diol	4.12	300.2515	48.41	131.73	15.23	18.29
Norajmaline	5.24	331.2463	44.26	132.16	12.87	11.24

3''-Deamino-3''-oxonicotianamine	7.48	314.1509	50.1	132.78	16.15	12.46
N1-Methyl-2-pyridone-5-carboxamide	7.23	159.0816	49.62	137.05	28.73	23.3
hexane-1,3,4,6-tetracarboxylate	7.31	272.1019	53.65	137.33	24.68	21.3
dTMP	8.19	332.0896	62.37	139.62	21.14	18.69
Alpha-CEHC	4.03	294.2047	56.65	142.51	27.16	32.93
[FA methyl(6:0)] 2-methyl-hexanoic acid	4.74	137.1225	91.94	144.22	18.95	19.65

Appendix A7. Comparison between %RSD of non-normalised and normalised metabolites in Mouse Plasma study

Metabolite	Retention time (minutes)	Mass	THREE MONTHS AGE		EIGHTEEN MONTHS AGE	
			Non-normalised RSD (%)	13C-normalised RSD (%)	Non-normalised RSD (%)	C13 Normalised RSD (%)
[PC (18:1/20:4)] 1-(1Z-octadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphocholine	4.39	839.7498	20.36	5.43	5.43	3.65
N-Hydroxy-2-acetamidofluorene	7.61	254.1442	20.14	5.93	5.93	3.39
Isocitrate	8.63	198.0468	14.37	6.22	6.22	5.38
[FA trihydroxy(4:0)] 8S,10R,14S-trihydroxy-4Z,12E,16Z,19Z-neuroprostatetraenoic acid-cyclo[7R,11R]	4.7	400.3138	29.55	6.25	6.25	5
1-Oleoylglycerophosphocholine	4.94	547.4342	8.09	6.36	6.36	5.41
PS(19:0/0:0)	4.83	564.4051	66.11	6.68	6.68	5.95
PA(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/20:5(5Z,8Z,11Z,14Z,17Z))	4.68	811.6048	20.06	6.81	6.81	5.19
PA(20:2(11Z,14Z)/0:0)	4.93	485.3512	19.35	7.12	7.12	9.26
[PC (22:6)] 1-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-sn-glycero-3-phosphocholine	4.89	597.4311	18.97	7.17	7.17	5.68
SM(d18:2/24:1)	4.52	857.818	20.98	7.35	7.35	5.68
[FA (24:0)] 15Z-tetracosenoic acid	3.99	390.4292	15.37	7.61	7.61	4.25
Phenethylamineglucuronide	5.43	311.1673	19.44	7.74	7.74	9.41
[SP (16:0)] N-(hexadecanoyl)-sphing-4-enine-1-phosphocholine	4.65	741.6959	9.6	7.81	7.81	6.04
D-Sorbitol	9.86	188.0989	16.18	7.89	7.89	4.43
Presqualene diphosphate	4.82	616.4183	22.71	7.93	7.93	6.6
LPA(0:0/18:2(9Z,12Z))	4.98	455.3126	28.49	8.14	8.14	5.46
PE(40:4)	4.4	840.7258	7.86	8.15	8.15	6.71
PC(38:6)	4.4	851.7133	8.83	8.25	8.25	7.23
Linoleamide	4.55	297.3155	13.46	8.28	8.28	10.11
18'-hydroxy-χ,χ-caroten-18-oate	4.59	614.4756	48.54	8.36	8.36	6.76
[PE (16:1/22:6)] 1-O-(1Z-hexadecenyl)-2-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-sn-glycero-3-phosphoethanolamine	4.2	790.6616	23.19	8.47	8.47	7.28
Cholesterolsulfate	3.92	493.4008	14.57	8.48	8.48	7.13
[PE (18:1)] 1-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine	4.79	502.3769	17.76	8.77	8.77	8.06
octyl β-1,6-D-galactofuranosyl-α-D-glucopyranoside	4.99	474.3073	19.33	8.79	8.79	6.14
PC(40:6)	4.37	881.751	18.93	8.85	8.85	7.36
[FA (16:4)] 4,7,10,13-hexadecatetraenoic acid	4.11	264.2305	27.8	8.97	8.97	5.78
[FA oxo(18:3)] 4-oxo-9Z,11E,13E-octadecatrienoic acid	4.12	310.2633	29.52	9.01	9.01	7.09
L-cysteine sulfinate	7.41	156.0194	17.46	9.01	9.01	10.31
[FA (8:2)] 2,4-octadienal	4.5	132.1152	45.85	9.13	9.13	7.93

[FA (17:0)] heptadecanoic acid	4.05	287.312	21.44	9.16	9.16	8.2
[PC (18:1)] 1-(11Z-octadecenyl)-sn-glycero-3-phosphocholine	4.99	533.4548	11.09	9.21	9.21	6.72
N-Acetyl-D-tryptophan	5.33	259.1433	20.41	9.36	9.36	8
Butyl 3-O-beta-D-glucopyranosyl-butanoate	5.26	336.2089	32.23	9.4	9.4	8.42
LPA(0:0/18:1(9Z))	4.97	457.3283	28.01	9.48	9.48	7.07
Hexadecanoic acid	4.06	272.2931	19.28	9.49	9.49	8.33
[PC (16:2)] 1-hexadecyl-sn-glycero-3-phosphocholine	5.07	505.4326	11.44	9.51	9.51	6.87
gamma-Amino-gamma-cyanobutanoate	10.24	133.075	35.68	9.56	9.56	8.52
Dihydroclavaminic acid	7.69	208.1062	37.1	9.66	9.66	8.32
L-Glutamine	10.11	151.0856	13.93	9.69	9.69	6.12
[PG (18:0)] 1-octadecanoyl-sn-glycero-3-phospho-(1'-sn-glycerol)	4.67	536.3916	6.84	9.75	9.75	7.79
Lithocholate3-O-glucuronide	4.6	582.4294	32.04	9.75	9.75	8.7
[FA hydroxy(20:2)] 11R-hydroxy-12E,14Z-eicosadienoic acid	4.06	344.3325	19.91	9.82	9.82	7.76
2-Methylbutyroylcarnitine	5.68	257.2023	38.34	9.82	9.82	8.75
[GP (16:0)] 1-hexadecanoyl-2-sn-glycero-3-phosphate	5.01	429.3062	15.22	9.84	9.84	7.06
[FA trihydroxy(22:0/6:0)] 7S,16,17S-trihydroxy-docosa-4Z,8E,10Z,12E,14E,19Z-hexaenoic acid	4.75	398.2982	20.77	9.85	9.85	8.59
D-Glucarate	10.09	216.0573	18.77	9.88	9.88	8.98
N-Undecanoylglycine	4.55	256.2264	27.64	10.12	10.12	11.68
[FA oxo(5:1/5:0/8:0)] (1S,2S)-3-oxo-2-(2'Z-pentenyl)-cyclopentaneoctanoic acid	4.1	312.2789	25.73	10.17	10.17	8.25
PA(34:2)	4.46	709.5953	11.86	10.17	10.17	6.49
PC(34:2)	4.43	799.7002	15.88	10.17	10.17	5.98
L-Histidine	10.72	161.0893	17.46	10.21	10.21	8.58
Xanthine	8.67	157.0499	35.1	10.22	10.22	8.33
PA(36:2)	4.44	739.6333	20.23	10.28	10.28	7.49
Nonadecanoicacid	4.03	317.3493	26.2	10.29	10.29	9.29
[SP (24:0)] N-(tetracosanoyl)-sphing-4-enine	4.25	691.7765	23.08	10.31	10.31	8.72
L-Rhamnulose	9.12	170.0883	18.83	10.39	10.39	7.25
[GP (20:4)] 1-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphate	4.92	481.3194	12.11	10.5	10.5	8.73
PC(36:3)	4.42	827.7223	15.09	10.77	10.77	7.82
[FA (11:2)] 2E,4E-undecadienoic acid	4.05	193.1671	32.34	10.83	10.83	7.61
[GP (21:4)] 1-(6Z,9Z,12Z,15Z-heneicosatetraenoyl)-sn-glycero-3-phosphate (Ammonium salt)	4.98	513.3646	15.99	10.85	10.85	8.37
[GP (18:0)] 1-octadecanoyl-2-sn-glycero-3-phosphate	4.9	459.3439	33.37	10.88	10.88	9.25
[ST] (5Z,7E)-9,10-seco-5,7,10(19)-cholestatriene	5.96	395.4333	42.51	10.88	10.88	11.79

N-Decanoylglycine	4.51	241.2073	31.16	11	11	9.36
(S)-3-Methyl-2-oxopentanoic acid	4.73	136.0829	31.68	11.08	11.08	9.73
Urate	8.89	173.0448	30.69	11.1	11.1	9.86
[GL (16:1/17:1)] 1-(9Z-hexadecenoyl)-2-(9Z-heptadecenoyl)-sn-glycerol	3.87	614.6103	23.72	11.16	11.16	9.67
Glucosylceramide (d18:1/24:1(15Z))	4.38	857.8333	25.51	11.17	11.17	8.45
PI(38:5)	3.98	931.6973	13.95	11.22	11.22	13.07
hydrogen iodide	10.56	127.9123	6.01	11.23	11.23	13.71
[ST methy] 24-methylene-cholest-5-en-3beta-ol-3-sulfate	3.91	506.404	15.07	11.4	11.4	8.65
N-Lauroylglycine	4.44	271.2456	28.91	11.44	11.44	7.31
[PE (16:1)] 1-(1Z-hexadecenyl)-sn-glycero-3-phosphoethanolamine	4.84	458.3603	18.48	11.48	11.48	9.93
PS(20:1(11Z)/0:0)	4.98	577.4087	21.72	11.48	11.48	7.86
Taurine	10.23	127.0213	13.15	11.53	11.53	10.05
Pyrimidodiazepine	0.21	230.1209	29.11	11.57	11.57	7.95
[SP (16:0)] N-(hexadecanoyl)-sphing-4-enine	4.34	571.6242	19.9	11.61	11.61	9.34
4-8dimethylnonanoylcarnitine	4.51	347.3159	33.65	11.63	11.63	10.17
PA(20:3(8Z,11Z,14Z)/0:0)	4.97	483.3351	14.73	11.72	11.72	8.56
L-Carnitine	9.34	168.1282	16.86	11.75	11.75	10.82
4-Imidazolone-5-propanoate	8.07	162.0733	26.57	11.81	11.81	10.7
[PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phosphocholine	5.03	519.4115	23.55	11.83	11.83	7.32
[SP (24:0)] N-(15Z-tetracosenoyl)-sphing-4-enine	4.25	689.7602	24.12	11.83	11.83	8.75
[PR] 2-methylbacteriohopane-32,33,34,35-tetrol	3.84	596.5996	24.18	11.89	11.89	10.02
FAD	7.8	812.2463	29.03	11.89	11.89	10.76
LysoPC(20:1(11Z))	4.87	577.4717	17.72	11.89	11.89	10.92
[PC (14:1)] 1-(9Z-tetradecenoyl)-sn-glycero-3-phosphocholine	5.02	487.3576	12.94	11.92	11.92	5.82
PA(38:4)	4.44	765.64	18.75	11.92	11.92	9.1
Asp-Asp-Gly	7.26	315.119	12.65	11.98	11.98	9.88
[FA (18:1)] 9Z-octadecenoic acid	4.04	300.3152	12.29	12.02	12.02	8.34
PC(20:2(11Z,14Z)/P-18:1(11Z))	4.39	841.7648	8.74	12.11	12.11	8.53
N-Acetyl-L-citrulline	7.72	225.1326	35.85	12.26	12.26	10.06
[FA (20:0)] eicosanoic acid	4.02	332.3689	13.58	12.35	12.35	10.51
[PC (18:0)] 1-octadecanoyl-sn-glycero-3-phosphocholine	4.9	549.4494	11.22	12.38	12.38	10.81
[FA (8:0)] octanoic acid	4.45	152.1414	22.87	12.45	12.45	7.26
5-Aminopentanoate	17.67	122.0955	11.83	12.49	12.49	8.65
[PR] (1S,4R)-1-Hydroxy-2-oxolimonene	4.42	178.148	20.18	12.53	12.53	10.14
3-Carboxy-4-methoxy-N-methyl-2-pyridone	4.74	191.0795	15.11	12.53	12.53	8.93

PE(20:2(11Z,14Z)/P-18:1(11Z))	4.25	796.7095	16.2	12.58	12.58	8.52
Queuine	12.56	289.157	27.51	12.58	12.58	8.46
[FA (17:2)] 10E,16-heptadecadien-8-ynoic acid	4.12	279.2494	17.99	12.67	12.67	11.19
Tetrapentylammonium	8.16	317.4055	62.11	12.67	12.67	14.68
PA(40:5)	4.41	793.6626	19.8	12.7	12.7	11.44
Cer(d18:1/23:0)	4.32	676.7571	17.2	12.74	12.74	11.6
PG(34:1)	4.02	788.6573	18.46	12.85	12.85	9.6
Phenyllactate	5.08	175.0926	19.06	12.85	12.85	6.37
Propanoyl phosphate	8.24	157.013	17.06	13.05	13.05	12.44
[SP (2:0)] sphinga-4E,14Z-dienine	4.54	315.326	29	13.06	13.06	6.46
PC(32:0)	4.45	773.694	25.81	13.06	13.06	8.51
3,7-Dimethyl-8,11-dioxo-2E,6E,9E-dodecatrienal	4.42	248.1718	78.31	13.15	13.15	5.44
10,16-Dihydroxyhexadecanoic acid	4.28	304.2827	20.72	13.18	13.18	14.66
Tetradecanoic acid	4.1	242.2552	18.43	13.19	13.19	11.63
2,4-Diaminotoluene	5.27	129.1074	15.62	13.23	13.23	9.64
Inosine	8.73	278.1137	50.01	13.27	13.27	11.33
PC(36:2)	4.42	829.7379	10.32	13.33	13.33	11.11
27-Nor-5b-cholestane-3a-7a-12a-24-25-pentol	3.94	464.4205	14.92	13.39	13.39	11.47
[PE (18:1/18:1)] 1-(1Z-octadecenyl)-2-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine	4.24	770.7018	17.99	13.4	13.4	11.54
2-Dehydro-D-xylonate	8.74	169.0486	16.98	13.44	13.44	7.64
[PC (16:1)] 1-(9Z-hexadecenoyl)-sn-glycero-3-phosphocholine	5.07	517.396	12.22	13.46	13.46	10.41
4-Heptyloxybenzoic acid	4.36	250.1874	82.8	13.48	13.48	6.74
Creatine	10.18	135.0827	25.1	13.51	13.51	8.03
PI(36:4)	3.98	903.6745	18.94	13.53	13.53	15.18
[FA (12:4/2:0)] 2E,4E,8E,10E-Dodecatetraenedioic acid	4.36	234.1288	14.83	13.58	13.58	10.04
LysoPE(0:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	4.73	552.3746	15.82	13.6	13.6	11.39
(-)-erythro-(2R,3R)-dihydroxybutylamide	5.67	123.0714	22.17	13.66	13.66	11.61
[PC (15:0)] 1-pentadecanoyl-sn-glycero-3-phosphocholine	5.08	504.3926	9.45	13.69	13.69	7.9
PE(36:4)	4.23	780.6506	35.48	13.72	13.72	10.81
D-Glycerate	8.35	109.0365	16.13	13.76	13.76	8.65
SM(d16:1/18:1)	4.65	739.68	9.49	13.77	13.77	9.44
[FA methyl(6:0)] 2-methyl-hexanoic acid	4.61	137.1225	25.64	13.8	13.8	12.87
SM(d18:0/18:1(9Z))	4.2	771.7343	19.79	13.84	13.84	12.64
Leucyl-leucine	11.83	256.2182	19.03	13.87	13.87	9.68
PE(42:6)	4.4	866.7318	14.58	13.89	13.89	17.34

[PR] Juvenile hormone III	4.14	282.241	20.77	13.92	13.92	12.06
3-Amino-3-(4-hydroxyphenyl)propanoate	7.2	190.1036	14.62	13.96	13.96	10.15
Uracil	7.57	116.0405	19.66	14	14	11.6
[FA oxo,hydroxy,dimethyl(2:0)] 9-oxo-15R-hydroxy-16,16-dimethyl-5Z,13E-prostadienoic acid	3.99	386.3341	28.9	14.01	14.01	10.34
[FA (11:0/2:0)] Undecanedioic acid	7.18	227.1724	12.04	14.08	14.08	7.34
[FA hydroxy(18:2)] 9S-hydroxy-10E,12Z-octadecadienoic acid	4.08	314.2946	28.5	14.08	14.08	8.39
Glucosylceramide (d18:1/12:0)	3.85	679.6217	32.76	14.1	14.1	11.24
PC(42:10)	4.38	903.7265	12	14.13	14.13	11.01
Glu-Gly	9.61	211.0977	21.4	14.18	14.18	10.93
PS(14:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	4.74	791.5754	27.89	14.2	14.2	13.11
[FA methoxy,trihydrox] 1-methoxy-9S,11R,15S-trihydroxy-5Z,13E-prostadiene	4.46	375.3464	45.29	14.31	14.31	12.47
N-Acetyl-L-histidine	5.85	205.1065	26.3	14.32	14.32	17.16
LysoPC(20:4(5Z,8Z,11Z,14Z))	4.92	571.4248	14.16	14.34	14.34	12.57
1-18:2-lysophosphatidylethanolamine	5.1	515.3795	19.29	14.39	14.39	12.88
Cytidine	9.23	252.1152	22.72	14.4	14.4	15.43
Linoleate	4.05	298.2997	21.34	14.4	14.4	7.74
O-Butanoylcarnitine	7.54	242.1835	16.94	14.44	14.44	9.13
PC(40:4)	4.38	885.7824	8.51	14.44	14.44	12.58
[FA hydroxy(20:3)] 8,9-dihydroxy-5Z,11Z,14Z-eicosatrienoic acid	4.12	358.3116	22.04	14.46	14.46	11.53
1-Hexadecanoyl-2-(9Z-octadecenoyl)-sn-glycero-3-phosphonocholine	4.43	785.7219	10.79	14.46	14.46	8.55
[FA (14:0/2:0)] Tetradecanedioic acid	4.7	272.2293	16.2	14.51	14.51	9.37
12-Hydroxy-12-octadecanoylcarnitine	5.1	468.4427	12.97	14.59	14.59	6.81
PA(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	4.87	507.3258	13.1	14.61	14.61	7.14
SM(d18:1/24:1(15Z))	4.51	859.8314	11.75	14.62	14.62	12.67
PC(34:1)	4.43	801.716	16.65	14.64	14.64	10.93
Proclavaminic acid	7.69	210.1219	25.43	14.68	14.68	8.26
(3R)-beta-Leucine	17.69	137.1144	10.25	14.69	14.69	12.8
3-Methyl-2-oxobutanoic acid	5.15	121.0639	20.56	14.71	14.71	13
Kaempferol 3-(6'''-rhamnosyl-2'''-(6-malyl-glucosyl)-glucoside)	7.46	909.3438	14.07	14.78	14.78	11.05
[FA hydroxy(18:1)] 9,10-dihydroxy-12Z-octadecenoic acid	4.18	332.3051	25.03	14.79	14.79	11.05
L-5-benzyl-hydantoin	8.13	200.1073	24.96	14.88	14.88	13.56
[FA (18:2)] 9S-hydroperoxy-10E,12Z-octadecadienoic acid	4.13	330.2896	32.37	14.95	14.95	13.73
[PE (18:1/22:6)] 1-(1Z-octadecenyl)-2-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-sn-glycero-3-phosphoethanolamine	4.2	820.6996	22.83	14.97	14.97	12.82

Cytosine	9.23	115.0564	24.18	15.01	15.01	13.45
PE(38:4)	4.27	810.6881	24.09	15.06	15.06	11.62
Glu-Arg	11.12	314.1904	12.13	15.1	15.1	12.67
[SP (3:0)] sphinga-4E,8E,10E-trienine	4.55	313.3104	16.29	15.18	15.18	11.61
Betaine	8.66	122.0955	33.21	15.18	15.18	12.94
[FA (20:5)] N-(5Z,8Z,11Z,14Z,17Z-eicosapentaenoyl)-ethanolamine	4	367.3394	35.07	15.2	15.2	11.84
N-Acetyl-beta-alanine	5.4	136.0747	19.37	15.32	15.32	11.41
Scymnol	3.81	495.4343	22.97	15.33	15.33	14.13
Hypoxanthine	8.64	141.055	37.07	15.36	15.36	13.27
N-Acetylglutamine	7.6	195.1029	10.73	15.36	15.36	12.56
PA(19:0/0:0)	4.85	474.363	21.27	15.36	15.36	11.32
PC(32:2)	4.47	769.6618	14.84	15.38	15.38	13.36
SM(d18:1/23:0)	4.51	846.8277	12.36	15.38	15.38	13.81
[FA oxo(12:1)] 12-oxo-10E-dodecenoic acid	4.39	224.1808	42.04	15.4	15.4	10.93
4-Trimethylammoniobutanoate	9.33	152.1333	25.78	15.4	15.4	9.2
7alpha-Hydroxycholest-4-en-3-one	3.8	427.4233	25.49	15.4	15.4	13.36
PE(36:2)	4.41	784.6825	14.44	15.4	15.4	12.6
[PR] Limonene-1,2-diol	4.39	180.1637	21.88	15.46	15.46	9.08
α -ethynylagmatine	15.43	161.1449	19.46	15.51	15.51	8.81
L-Pipecolate	9.15	135.0988	20.28	15.69	15.69	7.33
N-acetyl -D- glucosaminitol	9.33	231.132	15.61	15.77	15.77	3.89
[FA (18:3)] 9Z,12Z,15Z-octadecatrienoic acid	4.06	296.284	18.59	15.8	15.8	7.57
Monomethylglutaricacid	4.39	152.0777	25.08	15.87	15.87	13.93
Methyl oxalate	7.2	122.0399	12.44	15.88	15.88	12.66
(2R,4S)-2,4-Diaminopentanoate	10.46	137.1063	30.08	16.06	16.06	14.16
[FA (11:1)] 10-undecenoic acid	4.23	195.1826	31.94	16.11	16.11	14.9
Glucosylceramide (d18:1/16:0)	4.44	739.6976	20.04	16.11	16.11	13.94
7E,9E,11-Dodecatrienyl acetate	3.98	236.2083	36.51	16.14	16.14	13.48
[FA hydroxy(11:1)] 2-hydroxy-10-undecenoic acid	5.16	211.1776	33.41	16.15	16.15	14.08
Palmitoylglycerone phosphate	5.07	427.2904	15.8	16.24	16.24	10.89
Dodecanoic acid	4.16	212.2172	17.07	16.31	16.31	11.4
[FA (12:4)] 2E,4E,8Z,10E-dodecatetraenoic acid	4.4	204.1546	25.84	16.48	16.48	14.54
27-Norcholestanehexol	3.91	480.4152	16.84	16.48	16.48	14.19
Guanidinoacetate	10.84	120.0637	42.59	16.48	16.48	13.95
PE(18:3(6Z,9Z,12Z)/P-18:1(11Z))	4.24	764.656	20.56	16.48	16.48	13.48
Thymine	7.25	131.0594	16.42	16.51	16.51	11.83

N(pi)-Methyl-L-histidine	9.06	176.1082	69.29	16.59	16.59	11.07
[FA methyl(14:0)] 12-methyl-tetradecanoic acid	4.08	257.2741	16.6	16.6	16.6	13.99
2,3,2',3'-Tetrakis(3-hydroxytetradecanoyl)-D-glucosaminyl-1,6-beta-D-glucosamine 1-phosphate	4.9	1393.1117	23.56	16.8	16.8	13.99
PC(16:0/P-18:0)	4.43	787.7367	20.89	16.81	16.81	13.48
[FA (11:2)] 4,10-undecadiynal	4.37	173.1407	50.96	16.84	16.84	12.79
N-Acetylmethionine	5.28	198.0847	23.21	16.86	16.86	7.14
PI(40:7)	5.1	503.4316	22.04	16.87	16.87	13.44
[FA trihydroxy(2:0/2:0)] N,N-dimehyl-9S,11R,15S-trihydroxy-5Z,13E-prostadien-1-amide	4	403.3605	16.62	16.9	16.9	15.18
Choline	14.26	108.1162	22.88	16.93	16.93	10.92
[PE (20:4)] 1-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphoethanolamine	4.78	526.3678	12.24	16.94	16.94	14.86
3-Acetyl-5alpha-androstane-3beta,17beta-diol 3-acetate	4.53	399.3376	19.71	16.97	16.97	12.85
[PC (18:3)] 1-(9Z,12Z,15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	5.04	543.4029	17.62	16.98	16.98	12.95
N-Methylpelletierine	4.69	164.1607	6.88	16.98	16.98	20.26
PA(17:0/0:0)	4.97	444.3251	25.74	16.98	16.98	14.23
(10S)-Juvenile hormone III diol	4.43	300.2517	26.67	17.06	17.06	11.15
PE(22:4(7Z,10Z,13Z,16Z)/P-18:0)	4.21	824.7304	12.35	17.07	17.07	11.76
L-Kynurenine	8.68	218.1179	33.6	17.09	17.09	11.14
N-Acetyl-D-fucosamine	8.39	213.1218	28.97	17.14	17.14	8.96
2-Oxooctadecanoic acid	4.07	316.31	20.83	17.15	17.15	13.35
N-Formimino-L-glutamate	8	180.0839	27.17	17.15	17.15	14.63
2-Furoate	10.8	117.0325	17.66	17.17	17.17	13.29
VAI-2	4.25	239.1918	40.46	17.18	17.18	10.82
NG,NG-Dimethyl-L-arginine	13.34	210.1694	27.84	17.28	17.28	13.68
Suberylglycine	5.33	241.1436	16.02	17.28	17.28	15.55
Diethanolamine	12.96	109.0922	30.15	17.31	17.31	6.85
PA(38:3)	4.43	767.6553	25.08	17.32	17.32	15.54
Glycerophosphoglycerol	8.64	252.0702	23.7	17.39	17.39	17.4
Indan-1-ol	4.49	143.1028	47.56	17.42	17.42	21.79
[FA (18:3)] 13S-hydroperoxy-9Z,11E,14Z-octadecatrienoic acid	4.15	328.274	26.97	17.5	17.5	14.65
Anibine	5.43	214.0946	23.18	17.5	17.5	19.97
2-Arachidonoylglycerol	4.92	401.3531	19.51	17.57	17.57	15.93
[GP (14:0)] 1-tetradecanoyl-2-sn-glycero-3-phosphate	5.15	399.268	23.55	17.59	17.59	15.25
2-Oxoglutarate	9.74	151.038	33.42	17.63	17.63	12.33

3-Isopropylcatechol	4.77	161.1134	48.45	17.65	17.65	20.13
[FA (20:4)] 5Z,8Z,11Z,14Z-eicosatetraenoic acid	4.03	324.3062	23.92	17.66	17.66	16.07
Skatole	9.19	140.1031	20.66	17.77	17.77	16.07
10-Oxodecanoate	5.23	196.1586	27.64	17.78	17.78	14.97
N6-Methyl-L-lysine	13.86	167.1443	18.85	17.8	17.8	15.76
1-thio-β-D-glucose	7.17	202.0604	25.51	17.81	17.81	15.09
Glu-Ser	9.35	242.1115	23.78	17.81	17.81	8.55
PE(22:2(13Z,16Z)/P-18:1(11Z))	4.23	826.7477	13.8	17.88	17.88	8.92
PS(O-20:0/0:0)	5.1	565.4439	19.76	17.9	17.9	16.5
[FA hydroxy(20:3)] 15S-hydroxy-8Z,11Z,13E-eicosatrienoic acid	4.03	342.3166	28.4	17.92	17.92	13.3
Decanoic acid	4.28	182.1793	22.66	17.92	17.92	11.46
Lotaustralin	8.52	272.1575	17.89	17.92	17.92	15.14
Kynurenate	5.33	199.0756	26.92	18	18	15.87
L-Glutamate methylester	5.52	167.0885	20.9	18.01	18.01	15.42
Indole-3-acetate	7.23	185.0963	37.13	18.02	18.02	13.72
Ethyl (R)-3-hydroxyhexanoate	4.61	168.1363	31.54	18.04	18.04	13.51
PI(36:3)	3.98	905.6891	25.22	18.13	18.13	15.25
PA(20:1(11Z)/0:0)	4.9	487.3662	15.37	18.14	18.14	14.43
Costunolide	4.01	247.1959	22.17	18.17	18.17	14.99
Icosatrienoic acid	4.04	326.3219	20.01	18.19	18.19	14.05
Thiamin	14.78	276.144	32.42	18.2	18.2	17.8
alpha-Ribazole	5.32	292.1727	18.04	18.3	18.3	12.13
PE(34:2)	4.24	754.6443	35.38	18.3	18.3	13.95
myristic amide	4.68	241.2711	20.1	18.42	18.42	16.31
Ile-Glu	9.12	271.1733	13.75	18.45	18.45	13.66
1-O-(1Z-Tetradecenyl)-2-(9Z-octadecenoyl)-sn-glycerol	3.92	585.6114	20.6	18.48	18.48	14.14
PA(36:3)	4.46	737.6175	20.02	18.48	18.48	12.62
3-Butylidene-7-hydroxyphthalide	4.47	216.1182	13.71	18.53	18.53	15.35
[PE (16:0)] 1-hexadecanoyl-sn-glycero-3-phosphoethanolamine	4.88	474.3548	29.31	18.56	18.56	20.13
PE(38:2)	4.42	814.719	8.49	18.6	18.6	6.32
Glycylproline	8.76	179.1079	51.97	18.64	18.64	11.25
(9Z)-Hexadecenoic acid	4.07	270.2775	18.09	18.71	18.71	8.88
Stearoylcarnitine	4.87	452.4483	19.83	18.76	18.76	12.32
N-Methylhistamine	15.72	131.1151	28.27	18.84	18.84	20.34
[PC (20:0)] 1-eicosanoyl-sn-glycero-3-phosphocholine	4.84	579.4878	13.72	18.86	18.86	16.21
N5-Ethyl-L-glutamine	9.37	181.1235	49.64	18.91	18.91	7.19

Pantothenate	7.19	228.1404	31.17	18.94	18.94	16.29
[FA (7:2)] 2,4-heptadienal	4.5	117.0962	53.75	19	19	12.87
Indicine	5.05	314.2228	46.4	19.24	19.24	13.78
Met-Pro-Pro-Tyr	7.34	277.1895	19.99	19.24	19.24	14.23
Perillic acid	4.39	176.1323	21.91	19.26	19.26	12.79
Thymidine	7.23	252.1233	16.53	19.26	19.26	16.41
(L-Seryl)adenylate	9.81	231.0985	22.97	19.27	19.27	14.16
3-Hydroxysebacicacid	5.42	228.1484	23.08	19.3	19.3	16.03
(1S,2R,4S)-(-)-Bornyl acetate	5.17	208.186	21.11	19.37	19.37	18.5
L-Glutamate 5-semialdehyde	7.22	136.0747	28.65	19.39	19.39	8.49
Prostaglandin A2	4.81	354.2804	42.36	19.48	19.48	12.82
Vinylacetylglycine	9.37	149.078	31.4	19.55	19.55	14.37
L-Arginine	14.92	180.1315	28.16	19.65	19.65	16.58
[FA (14:2)] 5,8-tetradecadienoic acid	4.06	238.224	16.02	19.69	19.69	17.14
Retinol	4.35	306.2953	22.15	19.7	19.7	17.76
sn-glycero-3-Phosphocholine	9.64	265.1292	22.59	19.71	19.71	14.02
[PC (18:1/18:0)] 1-(1Z-octadecenyl)-2-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine	4.41	815.7579	8.1	19.74	19.74	16.65
[FA hydroxy(7:0)] 2-hydroxy-heptanoic acid	4.57	153.1174	23.86	19.77	19.77	14.35
ubiquinone-1	4.26	264.1668	30.83	19.77	19.77	18.2
[FA (13:0/2:0)] Tridecanedioic acid	5.08	257.2103	33.61	19.78	19.78	14.69
[ST hydroxy(4:0)] 2beta,3alpha,7alpha,12alpha-Tetrahydroxy-5beta-cholestan-26-oic acid	3.88	493.4189	19.18	19.86	19.86	17.57
[GL (17:2/20:3)] 1-(9Z,12Z-heptadecadienoyl)-2-(8Z,11Z,14Z-eicosatrienoyl)-sn-glycerol	3.99	668.6386	24.53	19.88	19.88	13.85
Leu-Pro	8.07	239.1837	31.52	19.89	19.89	11.48
5-Guanidino-2-oxopentanoate	10.3	179.0998	15.42	19.9	19.9	15.51
[FA hydroxy(18:0)] 2S-hydroxy-octadecanoic acid	4.03	318.326	19.31	19.95	19.95	15.9
4-Hexyloxyphenol	4.26	206.1702	23.27	20.07	20.07	15.4
D-Histidine	14.21	161.0893	22.39	20.08	20.08	18.52
Nonanoic acid	4.36	167.1603	22.95	20.13	20.13	14.19
carbidopa	5.46	236.1283	14.12	20.2	20.2	14.33
[ST hydrox] 3alpha,12alpha-Dihydroxy-5beta-chol-6-en-24-oic Acid	3.93	414.3563	18.83	20.21	20.21	15.19
Ovothiol A-cysteine disulfide	8.7	330.0947	41.59	20.22	20.22	18.12
omega-Cyclohexylundecanoic acid	4.05	285.2963	17.89	20.24	20.24	11.77
N-Butyryl-L-homoserine lactone	5.19	179.1159	34.03	20.29	20.29	12.95

[FA (20:0)] 11Z-eicosenoic acid	4.02	330.3532	19.07	20.31	20.31	12.19
[FA hydroxy(13:0)] 2-hydroxy-tridecanoic acid	4.18	243.2311	17.68	20.37	20.37	15.23
3-Dehydrocarnitine	7.22	166.1127	22.8	20.38	20.38	9.86
[FA methyl(18:0)] 11R,12S-methylene-octadecanoic acid	4.04	315.3345	23.03	20.42	20.42	12.59
[PE (18:1/20:4)] 1-(1Z-octadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphoethanolamine	4.23	794.6945	23.96	20.43	20.43	15.05
Tiglic acid	9.27	105.0689	25.53	20.52	20.52	13.77
Indolepyruvate	7.11	214.0946	41.46	20.53	20.53	17.73
dihydroartemisinic acid hydroperoxide	4.78	283.2168	47.48	20.55	20.55	8.94
Tridecanoylglycine	4.35	286.2642	27.55	20.59	20.59	18.31
[FA (10:0/2:0)] Decanedioic acid	5.32	212.1535	23.88	20.62	20.62	16.72
2-Pyridyl hydroxymethane sulfonic acid	6.1	195.0294	41.59	20.69	20.69	14.25
4-Guanidinobutanoate	10.43	150.1016	47.32	20.7	20.7	12.61
Urea-1-carboxylate	9.09	106.0288	19.63	20.7	20.7	8.43
[SP (22:0)] N-(docosanoyl)-sphing-4-enine	4.32	661.7379	20.71	20.75	20.75	17.1
[FA oxo(16:0)] 3-oxo-hexadecanoic acid	4.12	286.2721	26.6	20.78	20.78	16.18
L-Cysteate	10.23	172.0143	20.45	20.78	20.78	17.12
[ST hydroxy(2:0)] (5Z,7E)-(1R,2R,3R)-2-(3-hydroxypropoxy)-9,10-seco-5,7,10(19)-cholestatriene-1,3-diol	3.84	504.47	24.82	20.82	20.82	18.97
PC(38:4)	4.4	855.7443	7.68	20.84	20.84	21.8
O-Acetylcarnitine	8.39	212.1455	47.17	20.85	20.85	19.27
Bendiocarb	5.37	234.1208	33.46	20.93	20.93	10.06
[FA methyl,hydroxy(5:0)] 3R-methyl-3,5-dihydroxy-pentanoic acid	7.2	154.0934	26.63	20.94	20.94	15.64
Norajmaline	7.23	331.2463	19.66	20.95	20.95	19.21
2,6-dichloro-4-nitrophenol	10.32	212.969	13.41	21.02	21.02	19.45
Capsidiol	4.04	251.2272	16.63	21.14	21.14	19.44
[FA (22:0)] 13Z-docosenoic acid	4	360.3911	32.66	21.24	21.24	11.19
L-Citrulline	10.46	181.1155	28.91	21.24	21.24	18
Teasterone	3.86	476.4477	20.92	21.35	21.35	14.23
10-Hydroxydecanoic acid	4.28	198.1742	43.07	21.37	21.37	14.91
ubiquinol-1	4.18	266.1824	30.28	21.39	21.39	17.89
[FA (18:0)] N-(9Z-octadecenoyl)-ethanolamine	4.53	345.364	13.29	21.41	21.41	14.7
Diisopropyl adipate	7.15	242.1914	23.96	21.48	21.48	18.35
Diethyl adipate	7.23	212.1536	39.85	21.52	21.52	19.84
Ne,Ne dimethyllysine	13.12	182.1633	21.75	21.53	21.53	11.55
[PR] Iridotrial	7.16	192.1273	33.16	21.55	21.55	19.08

PE(38:5)	4.23	808.6728	28.15	21.56	21.56	17.72
9-Oxononanoic acid	5.21	181.1396	48.9	21.6	21.6	17.37
Homocysteinesulfinicacid	7.2	171.0384	35.4	21.64	21.64	13.5
N-(2-hydroxyoctadecanoyl)-4,8-sphingadienine	3.83	615.6426	21.78	21.68	21.68	20.06
N-Acetyl-L-phenylalanine	4.96	218.1258	38.92	21.68	21.68	16.46
Trimetaphosphate	10.92	15	11.45	21.69	21.69	19.37
L-Arabinose	10.28	155.0692	16.49	21.72	21.72	21.14
16-hydroxypalmitate	4.02	288.2881	33.82	21.77	21.77	18.29
Coniferyl alcohol	4.75	190.1115	12.1	21.78	21.78	16.22
Arg-Cys-Gly-Ser	15.87	435.2207	30.04	21.84	21.84	19.17
Nitrilotriacetic acid	9.45	197.0628	19.77	21.85	21.85	15.89
[PC (14:0)] 1-tetradecanoyl-sn-glycero-3-phosphocholine	5.15	489.3737	17.06	21.87	21.87	18.95
[FA hydroxy(9:0)] 2-hydroxy-nonanoic acid	4.36	183.1552	22.28	21.89	21.89	15.17
Citrate	10.83	198.0467	20.08	21.91	21.91	20.2
N-Ribosylnicotinamide	15.57	265.1265	21.28	21.91	21.91	19.96
15-Keto-prostaglandinE2	5.35	370.2754	32.6	21.93	21.93	20.06
2,3-dihydroxy-5-methyl-1,4-benzoquinone	4.66	161.0497	57.35	22.11	22.11	16.52
1,2,4-Trimethylbenzene	4.5	129.1236	38.11	22.13	22.13	19.64
PE(40:6)	4.26	836.6941	22.89	22.17	22.17	20.08
Mayolene-16	3.82	566.5616	18.85	22.19	22.19	17.68
Isoquinoline	7.14	138.0875	23.72	22.22	22.22	18.3
1-O-Methyl-myo-inositol	8.48	201.102	57.13	22.29	22.29	22.08
S-Adenosyl-L-methionine	10.57	413.1866	19.76	22.3	22.3	13.3
PI(34:2)	3.99	877.6676	22.15	22.31	22.31	16.67
PE(22:4(7Z,10Z,13Z,16Z)/P-18:1(11Z))	4.22	822.7154	13.33	22.36	22.36	18.31
3-Hydroxydodecanedioicacid	5.32	258.1861	41.97	22.49	22.49	19.05
L-Allothreonine	7.24	123.0716	16.03	22.49	22.49	11.56
(S)-1-Pyrroline-5-carboxylate	7.23	118.0642	35.13	22.51	22.51	18.82
[FA oxo(19:0)] 10-oxo-nonadecanoic acid	4.15	331.3291	14.75	22.51	22.51	11.58
[ST hydroxy(3:0)] (5Z,7E)-(1S,3R)-1,3-dihydroxy-9,10-seco-5,7,10(19)-cholestatrien-22-one	4.34	441.4028	28.84	22.64	22.64	20.26
Fagomine	9.34	153.1094	22.31	22.66	22.66	19.49
Palmitoleamide	4.56	269.2932	12.48	22.81	22.81	27.47
2-Oxoglutaramate	7.61	150.054	15.69	22.83	22.83	22.68
CL(1'-[18:2(9Z,12Z)/0:0],3'-[18:2(9Z,12Z)/0:0])	3.73	969.6622	17.71	22.86	22.86	20.14
Diethyl 2-methyl-3-oxosuccinate	8.25	211.1138	77.61	22.9	22.9	19.67

1-(3-aminopropyl)-4-aminobutanal	15.52	151.1493	31.27	22.98	22.98	18.3
8-Methoxykynurenate	7.88	230.0895	48.09	23.04	23.04	12.87
3-Dehydroteasterone	3.86	474.4321	22.85	23.13	23.13	15.2
2-Hydroxy-2,4-pentadienoate	10.29	119.0483	12.66	23.23	23.23	19.37
L-Ala-L-Glu	9.19	226.1167	36.78	23.23	23.23	20.73
meso-2,6-Diaminoheptanedioate	9.66	197.1182	20.9	23.24	23.24	16.75
6-Deoxocastasterone	3.86	478.4634	25.55	23.36	23.36	16.94
L-Tryptophan	9.18	215.1262	13.23	23.38	23.38	18.43
Ergothioneine	10.31	238.1182	37.77	23.39	23.39	12.06
PE(38:6)	4.22	806.6571	32.63	23.39	23.39	20.76
12-Nitro-9Z,12Z-octadecadienoic acid	4.72	343.2847	16.64	23.47	23.47	21.45
[6]-Gingerol	4.1	311.2391	52.56	23.53	23.53	21.15
[FA (20:5)] 5Z,8Z,11Z,14Z,17Z-eicosapentaenoic acid	4.06	322.2906	26.97	23.71	23.71	21.9
Tylactone	5.31	417.348	32.25	23.74	23.74	20.32
[FA amino(11:0)] 11-amino-undecanoic acid	7.29	212.2093	27.2	23.77	23.77	21.58
Succinate	9.57	122.0399	28.68	23.77	23.77	19.61
5-Hydroxy-L-tryptophan	7.62	231.1212	29.77	23.86	23.86	21.04
Thromboxane B2	4.44	390.3016	14.89	23.92	23.92	21.89
D-Mycinose	5.33	200.1261	31.8	23.93	23.93	20.98
Sulfamycin B	7.38	1298.4971	17.03	23.98	23.98	20.8
Ecgonine methyl ester	7.17	209.154	17.15	23.99	23.99	22.18
[FA dimethyl(13:0)] 2,5-dimethyl-2E-tridecenoic acid	4.09	255.2587	25.45	24.05	24.05	18.07
[PR] Tretinoin/All-Trans Retinoic Acid	4.07	320.2749	29.11	24.12	24.12	19.98
Glu-Leu	7.75	271.1735	26.63	24.2	24.2	20.63
gamma-Glutamyl-gamma-aminobutyraldehyde	5.46	225.1407	20.35	24.26	24.26	19.47
O-Acetyl-L-homoserine	7.23	167.0885	30.16	24.32	24.32	21.79
[ST hydroxy(2:0)] (5Z,7E)-(1R,2R,3R)-2-(5-hydroxypentoxy)-9,10-seco-5,7,10(19)-cholestatriene-1,3-diol	3.83	534.5078	19.88	24.64	24.64	18.69
1-Methyladenosine	10.62	292.1486	22.64	24.68	24.68	16.5
D-Galactosamine	8.67	185.0991	34.89	24.73	24.73	18.56
2-Oxophytanate	4.04	346.3482	16.92	24.97	24.97	26.16
L-Tyrosine	9.58	190.1037	26.39	25.01	25.01	18.35
[FA oxo(8:0)] 3-oxo-octanoic acid	5.37	166.1207	40.98	25.07	25.07	22.28
Acetoacetate	10.3	106.0449	18.1	25.13	25.13	18.24
D-Glucose	10.32	186.0832	13.76	25.19	25.19	21.59
PI(40:6)	3.97	959.72	21.94	25.22	25.22	18.49

PE(40:2)	4.41	844.7568	13.51	25.27	25.27	19.8
PS(40:6)	3.94	881.6879	19.14	25.28	25.28	21.05
Methyl5-(but-3-en-1-yl)amino-1,3,4-oxadiazole-2-carboxylate	7.5	205.1064	10.51	25.33	25.33	26.42
Orotidine	8.63	298.0924	19.43	25.43	25.43	7.88
Magnoshinin	4.54	438.2831	12.56	25.46	25.46	20.65
N-Acetyl-D-glucosamine 6-phosphate	9.37	309.0828	24.5	25.49	25.49	20.93
Docosanoic acid	4.02	362.4068	20.89	25.8	25.8	21.8
Octadecanamide	4.53	301.3467	20.28	25.83	25.83	32.35
DG(15:0/18:4(6Z,9Z,12Z,15Z)/0:0)	3.81	610.5788	37.71	25.91	25.91	21.87
DL-2-Aminooctanoicacid	9.28	167.1523	22.63	25.91	25.91	18.8
2-Oxoadipate	9.38	166.0569	25.01	26.07	26.07	14.6
L-Hypoglycin	7.24	148.1021	29.39	26.13	26.13	19.03
1-(beta-D-Ribofuranosyl)-1,4-dihydronicotinamide	7.81	267.1423	23.96	26.16	26.16	21.73
DL-Methionine sulfone	14.19	186.0573	17.67	26.17	26.17	33.03
L-Ornithine	13.58	137.1064	33.01	26.18	26.18	31.55
(10S)-Juvenile hormone III acid diol	4.67	285.2328	34.9	26.19	26.19	23.63
Glu-Tyr	9.09	324.1627	29.05	26.27	26.27	22.89
Glu-Val	8.38	256.1545	38.44	26.29	26.29	13.78
[PI (18:0)] 1-octadecanoyl-sn-glycero-3-phospho-(1'-myo-inositol)	4.23	627.4167	30.61	26.36	26.36	15.27
[FA trihydroxy(18:1)] 9S,12S,13S-trihydroxy-10E-octadecenoic acid	4.35	348.2999	30.15	26.49	26.49	29.64
[ST (4:0)] (7E)-(1R,2S,3R)-19-nor-9,10-seco-5,7-cholestadiene-1,2,3,25-tetrol	3.89	446.4098	19.74	26.55	26.55	18.83
Imidazol-5-yl-pyruvate	7.85	160.0575	16.19	26.59	26.59	20.6
Octadecanoic acid	4.03	302.3308	22.31	26.64	26.64	17.05
L-Lysine	14.22	152.1253	16.05	26.65	26.65	32.16
[PR] Citronellyl acetate	4.22	210.2015	39.72	26.67	26.67	20.6
[PC acetyl(17:2)] 1-heptadecyl-2-acetyl-sn-glycero-3-phosphocholine	4.86	564.4684	19.11	26.7	26.7	24.56
3-Indolepropionicacid	5.34	200.1153	11.69	26.71	26.71	18.6
N-Methyl-L-glutamate	9.51	167.0886	30.03	26.72	26.72	21.53
[FA hydroxy(18:0)] 9,10-dihydroxy-octadecanoic acid	4.18	334.3207	17.13	26.74	26.74	19.89
Jasmone	5.18	175.1564	31.07	26.81	26.81	23.82
PI(38:4)	3.98	933.7128	16.85	26.86	26.86	24.89
[FA oxo(5:1/5:0/4:0)] (1R,2R)-3-oxo-2-(2'Z-pentenyl)-cyclopentanebutanoic acid	4.14	252.2031	12.38	26.91	26.91	17.31
LysoPE(0:0/18:2(9Z,12Z))	4.84	500.3615	15.18	27.09	27.09	24.67

Myristoylglycine	4.74	301.2832	17.13	27.2	27.2	13.28
Acetylpseudotropine	5.23	193.159	28.08	27.27	27.27	15.94
(S)-Malate	9.99	138.0348	24.34	27.34	27.34	11.81
10-nitro-9E-octadecenoic acid	4.1	345.3004	41.81	27.37	27.37	15.02
Dodecanedioic acid	5.31	242.1913	32.38	27.44	27.44	24.95
[SP (18:0)] N-(octadecanoyl)-sphing-4-enine	4.34	601.6625	14.13	27.48	27.48	21.95
all-trans-Retinoyl-beta-glucuronide	4.54	264.2062	25.16	27.63	27.63	23.94
2-Hydroxydecanedioicacid	7.36	228.1484	14.56	27.65	27.65	25.08
Gamma-Glutamylglutamine	9.53	285.1445	29.09	27.65	27.65	13.22
Carnosine	10.31	235.1364	15.02	27.66	27.66	18.18
3,4-Dihydro-2H-Pyrrole-2-carboxylate	10.51	118.0642	33.39	27.75	27.75	20.84
PC(32:1)	4.45	771.6783	10.7	27.94	27.94	25.09
[FA (18:1)] 9Z-octadecenamide	4.54	299.3311	15.83	27.95	27.95	35.35
2-Acetolactate	10.17	137.0587	16.21	27.98	27.98	34.73
Leu-Lys-Pro	12.97	373.2983	21.43	28.2	28.2	36.3
[FA (24:6)] 4,8,12,15,19,21-tetracosahexaenoic acid	3.99	380.3509	23.98	28.36	28.36	24.78
N3-methylthymine	5.63	146.0783	32.82	28.39	28.39	29.24
Lys-Val	13.63	256.2102	18.69	28.42	28.42	25.37
Phe-Pro	5.29	276.1779	10.09	28.65	28.65	23.77
N-(octanoyl)-L-homoserine	7.27	257.2023	51.35	28.69	28.69	26.29
[SP hydroxy,hydroxy,methyl(10:2/2:0)] 6R-(8-hydroxydecyl)-2R-(hydroxymethyl)-piperidin-3R-ol	4.84	303.2988	26.31	28.7	28.7	23.72
[FA (16:0)] N-hexadecanoyl-glycine	4.73	331.3212	15.41	28.79	28.79	16.56
L-Aspartate	9.67	137.0507	28.64	28.81	28.81	24.44
2S-Hydroxytetradecanoic acid	4.11	258.25	25.67	28.87	28.87	23.04
2,3,4,5-Tetrahydropyridine-2-carboxylate	7.47	133.0831	29.19	28.88	28.88	22.36
4-Aminobutanoate	9.02	107.0765	10.17	28.92	28.92	28.81
L-Glutamate	9.46	152.0696	17.64	28.94	28.94	20.67
[FA amino(8:0)] 3-amino-octanoic acid	5.35	167.1523	58.18	29.01	29.01	25.11
Gibberellin A12	4.82	352.2649	24.55	29.07	29.07	11.25
2-(7'-methylthio)heptylmalate	4.97	290.1584	31.98	29.08	29.08	25.49
L-Octanoylcarnitine	5.35	302.2591	11.89	29.11	29.11	25.64
9-nitro-9E-octadecenoic acid	5.36	345.3001	38.21	29.29	29.29	26.56
Oblongolide	3.89	234.1927	22.63	29.3	29.3	15.98
[SP amino(18:0/2:0)] 2S-aminooctadecane-1,3R-diol	7.16	319.3574	18.15	29.31	29.31	26.62

[PI (20:4)] 1-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phospho-(1'-myo-inositol)	4.25	649.3918	31.66	29.34	29.34	20.27
[SP] Sphing-4-enine-1-phosphate	5.26	397.3083	32.67	29.42	29.42	24.38
[ST (3:0)] (5Z,7E)-(1S,3R,24R)-22-oxa-9,10-seco-5,7,10(19)-cholestatriene-1,3,24-triol	3.89	444.3943	26.97	29.63	29.63	21.47
N-stearoyl tyrosine	4.88	474.4238	23.91	29.75	29.75	23.23
5-Methylthio-D-ribose	8.11	186.0654	30.32	29.76	29.76	17.41
PI(36:2)	4	907.7063	24.75	29.94	29.94	19.9
Quinoline	5.44	138.0875	15.02	29.95	29.95	20.42
5-Methylcytosine	8.03	130.0754	39.26	30	30	22.35
[ST (4:0/4:0)] (5Z,7E,22E)-(1S,3R,25S)-9,10-seco-5,7,10(19),22-ergostatetraene-1,3,25,26-tetrol	3.86	472.4164	23.39	30.04	30.04	22.59
L-Threonine	9.87	123.0714	24.49	30.04	30.04	19.36
[ST hydroxy(4:0)] 2alpha,3alpha,22R,23R-tetrahydroxy-6,7-seco-5-campestano-6,7-lactone	3.88	508.4379	24.56	30.08	30.08	27.61
L-Indospicine	14.31	180.1395	32.88	30.12	30.12	27.6
6-Acetyl-D-glucose	7.8	230.1003	51.43	30.15	30.15	14.56
Choline phosphate	9.48	188.0826	23.71	30.19	30.19	25.86
8-Amino-7-oxononanoate	5.31	196.1506	46.48	30.23	30.23	23.02
[FA (20:2)] 15S-hydroperoxy-11Z,13E-eicosadienoic acid	3.96	360.3272	18.87	30.28	30.28	27.5
6alpha-Hydroxy-castasterone	3.88	494.4587	25.98	30.31	30.31	27.81
Creatinine	8.29	117.0721	34.32	30.36	30.36	24.51
PS(O-18:0/0:0)	5.24	535.4067	22.14	30.41	30.41	27.37
3-(4-Hydroxyphenyl)lactate	7.25	191.0877	50.19	30.56	30.56	26.99
N-(3-oxooctanoyl)-L-homoserine	5.32	271.1816	24.05	30.58	30.58	24.24
[FA (26:0/2:0)] Hexacosanedioic acid	3.91	452.4569	21.72	30.62	30.62	28.04
Ala-Ser	13.58	182.0995	31.08	30.78	30.78	37.18
[ST (5:0)] 5beta-Cholestane-3alpha,7alpha,12alpha,25,26-pentol	3.87	479.4393	30.4	30.98	30.98	20.78
[ST hydrox] N-(3alpha,7alpha-dihydroxy-5beta-cholan-24-oyl)-taurine	4.32	525.3826	79.03	30.99	30.99	21.41
Mephobarbital	7.12	259.1433	40.42	31.01	31.01	18.12
Capryloylglycine	5.24	211.1695	36.43	31.2	31.2	18
L-Serine	10.51	108.0525	32.57	31.22	31.22	23.07
[PR] Vitamin E	4.36	459.4767	9.38	31.31	31.31	27.17
Docosahexaenoicacid	4	350.3128	31.89	31.32	31.32	27.43
L-Homocitrulline	10.19	196.1344	25.43	31.37	31.37	21.73
LL-2,6-Diaminoheptanedioate	14.2	197.1185	19.8	31.53	31.53	28.67

2-Hydroxybutane-1,2,4-tricarboxylate	7.23	213.0656	29.79	31.55	31.55	19.29
Tiglylcarnitine	5.29	255.1866	12.83	31.55	31.55	25.41
D-Ribose 5-phosphate	9.47	235.0356	34.85	31.61	31.61	26.17
[FA trihydroxy(4:0)] 2,3,4-trihydroxy-butanoic acid	8.64	140.0505	26.23	31.64	31.64	27.67
Gibberellin A12 aldehyde	4.17	336.2696	14.59	31.7	31.7	25.45
Arecoline	7.18	163.1211	42.99	31.71	31.71	22.85
[PC (18:1/22:6)] 1-(1Z-octadecenyl)-2-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-sn-glycero-3-phosphocholine	4.37	865.7561	13.04	31.74	31.74	19.19
5beta-Pregnane-3alpha,17alpha,20alpha-triol	4.14	357.3357	16.16	31.85	31.85	23.42
cis-Aconitate	10.64	180.0361	28.25	32.01	32.01	26.68
3-Methylguanine	8.39	171.0847	45.64	32.02	32.02	27.67
Leu-Gln-Pro	9.14	372.2584	21.38	32.07	32.07	24.96
1-Pyrroline-2-carboxylate	5.51	118.0642	28.1	32.12	32.12	19.13
[GL (16:1/17:2)] 1-(9Z-hexadecenoyl)-2-(9Z,12Z-heptadecadienoyl)-sn-glycerol	3.81	612.5953	29.33	32.24	32.24	27.9
3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	12.26	213.1771	21.68	32.39	32.39	28.68
Cholate	5.19	432.3669	68.88	32.4	32.4	23.28
Retronecine	5.31	163.121	38.35	32.4	32.4	16.36
[SP hydrox] 6-hydroxysphing-4E-enine	4.71	333.3368	21.5	32.42	32.42	23.62
5-Oxopentanoate	7.51	121.0639	22.38	32.42	32.42	28.39
[FA (22:5)] 4,7,10,13,16-Docosapentaynoic acid	3.95	182.1615	30.85	32.54	32.54	25
Ethanolamine phosphate	10.03	143.0257	14.7	32.54	32.54	20.21
Antibiotic TA	3.97	658.5551	16.6	32.64	32.64	23.22
Hippurate	5.3	188.088	19.65	32.69	32.69	37.11
2-Hexenoylcarnitine	7.22	270.2056	22.34	32.78	32.78	28.98
PI(18:2(9Z,12Z)/0:0)	4.29	623.3854	22.89	32.81	32.81	16.84
N-Acetyl-D-glucosamine 6-sulfate	8.86	309.0731	12.12	32.9	32.9	20.3
3-Oxohexobarbital	5.32	262.1349	29.51	32.96	32.96	28.91
N-Formyl-L-methionine	5.34	183.0658	10.71	33.1	33.1	24.9
[FA (17:1)] 7-heptadecynoic acid	4.05	283.2805	12.2	33.11	33.11	30.6
[FA hydroxy(20:4/2:0)] 5S,12R-dihydroxy-6Z,8E,10E,14Z-eicosatetraene-1,20-dioic acid	4.5	386.2703	66.86	33.14	33.14	24.73
[FA (18:2)] N-(9Z,12Z-octadecadienoyl)-ethanolamine	4.51	343.3483	25.31	33.26	33.26	25.22
N-Heptanoylhomoserine lactone	7.15	224.1728	19.91	33.26	33.26	25.43
D-Glucono-1,5-lactone	9.43	184.0675	17.81	33.37	33.37	25.03
Indolelactate	7.17	216.1102	36.37	33.43	33.43	27.72

L-1-Pyrroline-3-hydroxy-5-carboxylate	10.06	134.0592	13.18	33.56	33.56	29.22
Leu-Lys	12.46	271.2293	25.09	33.75	33.75	21.5
PC(36:4)	6.81	825.7068	26.06	33.75	33.75	41.93
Eugenol methyl ether	4.06	189.1357	32.29	33.85	33.85	27.7
N6,N6,N6-Trimethyl-L-lysine	13.27	197.1823	18.12	33.92	33.92	24.53
[FA hydroxy(20:4)] 15S-hydroxy-5Z,8Z,11Z,13E-eicosatetraenoic acid	4.06	340.3011	27.73	33.98	33.98	31.09
Ascorbate 2-sulfate	10.89	262.0087	16.41	34.08	34.08	22.5
[FA (21:0/2:0)] Heneicosanedioic acid	4.46	377.362	43.76	34.1	34.1	25.71
Tetradecanoylcarnitine	5.1	392.3727	26.57	34.14	34.14	15.4
[FA amino(9:0)] 3R-aminononanoic acid	5.28	182.1713	47.38	34.16	34.16	25.35
N-Nonanoylglycine	7.59	226.1885	34.01	34.23	34.23	25.88
D-Gluconic acid	9.32	202.0781	20.52	34.24	34.24	30.41
2-C-Methyl-D-erythritol 4-phosphate	10.06	221.0566	26.47	34.3	34.3	29.66
6beta-acetoxy-24-methylcholestan-3beta,5alpha,22R,24-tetrol	3.88	538.476	13.19	34.38	34.38	30.35
4-Chlorotoluene	4.99	133.0468	6.56	34.51	34.51	21.79
Leu-Ala	9.9	211.1614	38.97	34.51	34.51	26.34
Halstoctacosanolide A	4.66	470.4258	20.74	34.68	34.68	25.27
Homoarginine	14.1	195.1505	29.6	34.68	34.68	31.87
[SP] Sphinganine-1-phosphate	7.16	399.3243	20.84	34.71	34.71	19.24
L-Urobilinogen	4.58	629.4656	36.15	34.72	34.72	30.93
D-Proline	13.58	120.0798	21.54	34.73	34.73	34.02
L-Arabinonate	9.08	171.0642	15.54	34.75	34.75	25.13
(R)-3-Hydroxybutanoate	7.22	108.0606	69.84	34.96	34.96	25.78
Spongipregnoloside A	3.93	345.2849	28.08	35.03	35.03	18.9
Urocanate	8.07	144.0627	21.22	35.11	35.11	15.42
Indoxyl	7.17	141.0791	28.94	35.24	35.24	23.55
Palmiticamide	4.59	271.3089	25.98	35.25	35.25	27.56
myo-Inositol	8.25	186.0831	23.15	35.33	35.33	21.73
O-Propanoylcarnitine	7.92	227.1644	27.69	35.4	35.4	32.47
[PR trihydroxy(2:0)] (-)-1alpha,7beta,12-trihydroxy-2,9-illudadien-8-one	5.39	279.1856	18.42	35.55	35.55	45.97
1,2-dioctanoyl-1-amino-2,3-propanediol	5.24	362.3347	35.35	35.58	35.58	21
N-Acetyl-L-glutamate	8.95	196.0868	26.02	35.7	35.7	29.17
Chenodeoxycholicaciddisulfate	5.32	300.1824	39.34	35.83	35.83	31.38
N-acetyl-(L)-arginine	10.02	224.1487	20.87	35.83	35.83	32.58
Icosadienoic acid	4.02	328.3376	24.2	35.87	35.87	24.34

PA(38:5)	4.44	763.6242	16.22	35.91	35.91	28.1
[GL methyl(15:0/8:0)] 1-(14-methyl-pentadecanoyl)-2-(8-[3]-ladderane-octanyl)-sn-glycerol	3.97	641.6561	27.34	35.95	35.95	25.07
[FA methyl(12:1)] 2-methyl-2-dodecenoic acid	4.13	225.2206	46.01	36.28	36.28	19.25
[FA (16:2)] 9,12-hexadecadienoic acid	4.08	268.2619	31.57	36.31	36.31	30.91
Bradykinin	11.27	579.9458	62.07	36.42	36.42	31.96
3beta-Hydroxy-4beta-methyl-5alpha-cholest-7-ene-4alpha-carboxylate	4.28	473.4559	36.06	36.49	36.49	32.77
PE(40:7)	4.2	834.6797	31.96	36.51	36.51	30.94
[ST trihydroxy(3:0)] (5Z,7E)-(1S,3R,23S,25R)-1,3,25-trihydroxy-9,10-seco- 5,7,10,(19)-cholestatrieno-26,23-lactol	4.2	473.3924	39.03	36.53	36.53	33.06
N-Acetylneuraminate	8.84	320.1423	25.6	36.78	36.78	22.62
1-Pyrroline-4-hydroxy-2-carboxylate	6.02	134.0592	46.63	37.11	37.11	34.86
Tauropine	7.22	202.0522	32.41	37.14	37.14	33.64
2-Dehydro-3-deoxy-L-rhamnonate	9.14	168.0726	24.41	37.34	37.34	28.72
[FA dioxo(8:0)] 4,7-dioxo-octanoic acid	7.61	180.0999	25.68	37.4	37.4	27.65
Hexanoylglycine	7.18	181.1317	42.36	37.43	37.43	27.27
Phensuximide	7.14	200.1153	27.05	37.49	37.49	24.48
Pro-Ser-Ser	8.7	300.1638	58.43	37.5	37.5	22.7
Nicotinamide	7.25	128.0678	38.71	37.64	37.64	32.34
13'-hydroxy-α-tocopherol	4.37	475.4717	9.43	37.71	37.71	33.92
Taxa-4(20),11(12)-dien-5alpha-yl acetate	4	352.3285	28.19	37.76	37.76	29.97
Tetranor-PGF1alpha	4.24	316.2461	26.41	37.79	37.79	27.22
[PR] Coenzyme Q4	4.01	483.4041	17.23	37.82	37.82	22.02
3-(4-Hydroxyphenyl)pyruvate	7.21	189.072	35.48	37.86	37.86	34.52
Ngamma-Monomethyl-L-arginine	15.16	195.1505	18.59	37.88	37.88	33.49
2-Quinolinecarboxylic acid	5.46	183.0807	20.84	37.99	37.99	27.54
cis-5-Tetradecenoylcarnitine	5.14	390.357	35.06	38.08	38.08	25.77
5-Methoxytryptophol	8	202.131	43.97	38.14	38.14	32.84
Dehydrojuvabione	4.18	280.2254	19.99	38.22	38.22	28.35
N-Acetyl-L-aspartate	9.22	181.0679	48.94	38.28	38.28	29.53
3-(3-Hydroxy-phenyl)-propanoic acid	7.2	175.0926	54.77	38.33	38.33	34.52
Anacardic acid	3.99	370.3392	15.49	38.61	38.61	23.47
Ala-Pro	7.68	194.1269	21.5	38.74	38.74	33.7
(R)-2-Hydroxyglutarate	9.53	153.0537	26.69	39.01	39.01	30.2
[FA oxo(5:2/5:0/6:0)] (1R,2R)-3-oxo-2-pentyl-cyclopentanehexanoic acid	4.07	284.2561	40.2	39.05	39.05	28.44

3alpha,7alpha,12alpha-Trihydroxy-5beta-cholestan-26-al	3.88	461.4287	27.95	39.07	39.07	31.73
(2S)-2-{[1-(R)-Carboxyethyl]amino}pentanoate	5.33	197.1264	16.09	39.25	39.25	33.39
Nicotine	5.55	172.1487	38.3	39.31	39.31	34.09
[PI (18:0/18:0)] 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phospho-(1'-myo-inositol)(ammonium salt)	3.98	924.732	31.44	39.62	39.62	26.4
N4-(Acetyl-beta-D-glucosaminyl)asparagine	10.42	347.1727	52.59	39.71	39.71	32.75
Neoquassin	5.12	412.2771	31.08	39.8	39.8	34.8
N-Acetyl-L-leucine	5.23	181.1316	34.88	39.94	39.94	23.53
creatinine phosphate	11.33	197.0387	19.5	39.99	39.99	22.27
Homogentisate	7.33	176.0687	32.29	40.24	40.24	33.26
[FA amino(10:0)] 10-amino-decanoic acid	5.19	197.1902	44.97	40.27	40.27	31.44
4-Oxocyclohexanecarboxylate	5.53	149.0861	17.45	40.28	40.28	33.93
3,4-Dihydroxyphenylacetaldehyde	5.06	160.0737	19.33	40.48	40.48	36.54
3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate	4.66	430.3512	76.16	40.61	40.61	37.01
Delphinidin 3-O-(2''-O-galloyl-6''-O-acetyl-beta-galactopyranoside)	4.21	689.2243	39.12	40.61	40.61	36.77
Homocarnosine	9.58	250.1552	29.85	40.82	40.82	36.28
Gly-Arg	10.88	239.1595	86.76	40.93	40.93	37.44
Tributyrin	5.39	317.2224	22.64	40.94	40.94	31.91
PS(40:8)	3.96	877.6564	28.74	40.97	40.97	41.73
sn-Glycerol 3-phosphate	9.46	175.0236	28.09	40.99	40.99	30.8
3, 5-Tetradecadiencarnitine	5.19	388.3414	29.83	41.06	41.06	31.74
5-Acetamidopentanoate	5.35	166.1126	18.98	41.22	41.22	32.61
[FA hydroxy,oxo(7:0/2:0)] 4-hydroxy-2-oxo-Heptanedioic acid	7.81	197.0708	17.68	41.33	41.33	27.55
D-Phenylalanine	8.15	174.1087	32.15	41.51	41.51	29.86
[FA] O-Palmitoyl-R-carnitine	4.97	422.4105	18.96	41.84	41.84	29.99
4-(2-Aminophenyl)-2,4-dioxobutanoate	7.86	217.0862	16.03	41.93	41.93	48.35
[ST hydroxy(2:0)] 17,21-dihydroxypregn-4-ene-3,20-dione	5.17	367.2835	12.21	41.99	41.99	35.38
Methylimidazoleacetic acid	7.27	146.0784	53.84	42.45	42.45	30.57
2-Amino-9,10-epoxy-8-oxodecanoic acid	5.31	225.1488	27.45	42.58	42.58	38.03
Glu-Trp	8.66	349.1855	26.53	42.62	42.62	26.79
9,12-Hexadecadienoylcarnitine	5.06	418.3792	33.5	42.64	42.64	35.53
2-Methylcitrate	8.67	213.0657	18.83	42.86	42.86	21.43
PS(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	4.2	597.3678	49.02	43.02	43.02	31.87
[FA (22:4)] 7Z,10Z,13Z,16Z-docosatetraenoic acid	4.01	354.3442	24.6	43.04	43.04	31.43
Formylanthranilate	5.07	173.069	21.78	43.07	43.07	37.94

Cyanidin 3-O-(2''-O-galloyl-6''-O-alpha-rhamnopyranosyl-beta-galactopyranoside	4.24	781.2899	37.08	43.21	43.21	39.19
Phenylpyruvate	4.8	173.077	16.7	43.65	43.65	47.12
N-Glycoloyl-neuraminate	9.38	336.1371	30.81	43.68	43.68	40.2
4-Acetamidobutanoate	5.32	151.0937	26.43	43.95	43.95	31.15
Elaidiccarnitine	4.9	450.4327	24.9	43.95	43.95	28.67
(24R,24'R)-Fucosterol epoxide	4.31	457.4611	8.69	43.99	43.99	39.45
2-8-Dihydroxyquinoline-beta-D-glucuronide	7.65	352.1294	26.73	44.24	44.24	29.12
1(2H)-Isoquinolinone	6.99	154.0824	16.28	44.37	44.37	33.44
[FA hydroxy(22:5)] (+/-)-19,20-dihydroxy-4Z,7Z,10Z,13Z,16Z-docosapentaenoic acid	4.06	384.3181	35.41	44.4	44.4	28.64
S-Adenosyl-L-homocysteine	9.32	398.1676	20.43	44.53	44.53	38.48
[Fv Hydroxy,methoxy,methyl(9:1)] 5-Hydroxy-6-methoxy-3',4'-methylenedioxyfurano[2'',3'':7,8]flavanone	7.2	373.136	41.25	44.63	44.63	39.21
3-Hydroxy-9-hexadecenoylcarnitine	5.26	436.3897	35.48	44.71	44.71	41.35
PS(34:1)	4.23	801.6517	13.27	44.84	44.84	41.2
Butenylcarnitine	5.12	240.1677	32.04	45.27	45.27	34.42
Trp-Asp-Gly	7.66	393.1944	13.7	45.3	45.3	55.4
3-Hydroxytetradecanedioicacid	4.78	288.224	28.18	45.32	45.32	51.06
[FA (11:2)] 2,5-undecadienal	4.37	177.1721	22.22	45.94	45.94	28.89
3-Oxododecanoic acid	4.19	226.1965	48.68	45.97	45.97	35.14
Kolanone	4.06	535.4167	41.85	45.97	45.97	39.18
5-6-Dihydrouridine	8.37	255.1148	19.64	46.06	46.06	27.44
[FA (12:0)] 12-[5]-ladderane-dodecanoic acid	4.01	382.3663	26.03	46.36	46.36	38.49
cis-1,2-Dihydroxy-1,2-dihydrodibenzothiophene	9.22	230.0796	11.93	46.51	46.51	33.18
12-Hydroxydodecanoic acid	4.16	228.2121	46.42	46.83	46.83	40.85
1-Palmitoylglycerophosphocholine	6.99	519.4117	50.32	46.86	46.86	36.21
Pseudouridine	8.89	253.0992	23.06	46.9	46.9	30.02
3-Hydroxyhexadecanoylcarnitine	5.23	438.4053	28.83	46.93	46.93	40.52
4-Hydroxybenzoate	4.98	145.0547	17.62	46.93	46.93	42.17
Linoelaidylcarnitine	4.92	448.4171	26.73	47.02	47.02	41.89
L-Methionine	8.82	154.0675	26.44	47.28	47.28	37.85
(E)-2-Butenyl-4-methyl-threonine	7.15	196.1506	48.64	47.35	47.35	32.48
L-Cystine	10	246.0436	51.17	48.15	48.15	40.46
Biliverdin	4.19	615.357	28.79	48.29	48.29	43.81
PS(38:4)	3.94	855.6812	32.39	48.3	48.3	43.45

(9Z)-Tetradecenoic acid	4.11	240.2396	26.55	48.36	48.36	40.68
Fructoselysine	12.43	320.1978	19.66	48.46	48.46	56.81
Cerulenin	4.34	235.1604	10.94	48.53	48.53	42.31
N-hexanoyl-D,L-homoserine lactone	4.5	209.1538	15.75	48.57	48.57	34.76
[FA hydroxy(10:0)] N-(3S-hydroxydecanoyl)-L-serine	7.57	288.2162	22.91	48.91	48.91	43.47
L-Leucine	8.53	137.1144	36.84	49.15	49.15	54.38
Linamarin	8.64	257.1385	21.82	49.17	49.17	39.24
trans-Hexadec-2-enoylcarnitine	5.01	420.3948	22.32	49.3	49.3	36.45
L-2-Amino-3-oxobutanoic acid	7.37	121.0558	40.83	49.45	49.45	42.4
[FA hydroxy,oxo(16:0)] 9-hydroxy-16-oxo-hexadecanoic acid	4.27	302.267	31.63	49.48	49.48	37.17
L-Olivosyl-oleandolide	5.08	284.2325	25.76	49.73	49.73	37.32
2-Hydroxyethanesulfonate	7.95	128.0053	20.67	50.29	50.29	39.95
2-Naphthylamine	8.46	153.1065	17.56	50.4	50.4	35.04
N-Acetyl-aspartyl-glutamate	10.04	315.127	47.82	51.18	51.18	47.33
[FA oxo(14:0)] 3-oxo-tetradecanoic acid	4.1	256.2345	54.53	51.22	51.22	45.58
3-4-DihydroxyphenylglycolO-sulfate	8.98	258.0411	15.26	51.62	51.62	41.79
3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid	5.42	252.1395	23.76	52.37	52.37	47.55
3-Methyleneoxindole	5.35	154.0825	38.31	52.5	52.5	47.5
S-Sulfo-L-cysteine	10.82	203.9863	44.05	52.83	52.83	47.49
[SP (14:0)] 1-deoxy-tetradecasphinganine	5.1	243.2867	56.71	53.18	53.18	42.2
Leukotriene A4	4.07	338.2856	29.47	53.65	53.65	32.03
PG(34:2)	3.89	786.6418	15.5	53.72	53.72	47.73
L-Proline	9.4	120.0798	25.21	53.86	53.86	66.74
α-oxo-1-carboxy-5,6-dihydrothiopyranpropanoate S-oxide	9.49	255.0495	37.05	54.68	54.68	31.58
5,6-Dihydrothymine	7.87	133.075	44.87	55.09	55.09	44.83
Allantoin	10.42	162.0571	15.95	56.61	56.61	48.74
D-Glyceraldehyde 3-phosphate	9.54	173.0082	46.45	57.09	57.09	63.6
Taurocholate	4.7	541.3773	66.71	57.12	57.12	50.99
3-Hydroxy-cis-5-tetradecenoylcarnitine	5.32	406.352	42.28	57.49	57.49	52.57
2-Octenoylcarnitine	7.19	300.2435	26.13	57.51	57.51	50.99
L-Methionine S-oxide	9.2	170.0623	22.66	57.56	57.56	71.17
L-Asparagine	10.15	136.0666	32.39	57.78	57.78	48.18
Calystegin A3	9.13	166.1128	15.96	57.94	57.94	51.94
[ST hydrox] N-(3alpha-hydroxy-5beta-cholan-24-oyl)-glycine 3-sulfate	4.33	539.3617	53.89	58.44	58.44	47.15

[FA oxo(5:2/5:0/4:0)] (1S,2S)-3-oxo-2-pentyl-cyclopentanebutanoic acid	4.1	254.2188	53.13	58.9	58.9	54.26
3-Methylglutarylcarnitine	8.18	302.1954	17.75	58.91	58.91	45.69
1-Methylnicotinamide	16.09	143.0868	19.93	59.06	59.06	63.3
Monomethyl sulfate	7.2	112.9864	49.72	59.38	59.38	51.9
L-Histidinal	5.23	145.0944	21.2	59.43	59.43	45.65
Phenylacetylglycine	5.23	203.1071	26.53	59.76	59.76	38.59
3-Hydroxyanthranilate	7.25	160.0657	23.89	59.89	59.89	54.26
Thr-Pro	7.18	225.1407	26	60.1	60.1	32.43
[FA oxo,amino(6:0)] 3-oxo-5S-amino-hexanoic acid	9.41	151.0937	20.72	60.32	60.32	50.05
Deoxyuridine	7.59	237.1044	21.5	61.2	61.2	78.37
N-Carbamoylsarcosine	8.46	136.0666	52.76	61.21	61.21	52.92
L-Erythrulose	7.31	124.0555	26.94	61.34	61.34	50.04
Hydroxybutyrylcarnitine	8.48	258.1782	47.53	62.56	62.56	37.7
N1-Methyl-2-pyridone-5-carboxamide	7.2	159.0817	23.62	63.09	63.09	56.89
4-Hydroxyaminoquinoline N-oxide	7.24	185.0883	25.28	64.06	64.06	51.57
Benzoate	5.31	129.0599	22.16	65.16	65.16	44.72
Phenol sulfate	5.52	180.0185	19.26	65.18	65.18	58.12
(S)-ATPA	9.45	238.144	19.8	65.34	65.34	60.35
Glu-Met-Met-Met	11.25	560.241	32.57	65.37	65.37	86.31
5,6-Dihydrouracil	8.45	118.0561	46.93	65.77	65.77	53.52
Urocortisone	4.5	385.2945	31.74	65.88	65.88	50.63
[ST trihydrox] 3alpha,11beta,21-5alpha-trihydroxy-pregnane-20-one	4.41	371.315	17.11	66.06	66.06	43.4
Cotinineglucuronide	11.37	368.1797	21.34	67.35	67.35	57.66
Phosphocreatine	9.42	215.049	41.46	67.76	67.76	44.15
3-Hydroxy-5, 8-tetradecadiencarnitine	5.35	404.3363	34.06	68.74	68.74	62.79
2-Octenedioicacid	9.81	180.1001	42.09	70.48	70.48	36.09
D-Glucuronolactone	9.6	182.0519	100.95	70.68	70.68	65
[ST trihydroxy(2:0)] 11beta,17,21-trihydroxypregn-4-ene-3,20-dione	4.53	383.2788	14.75	72.39	72.39	64.92
N(alpha)-t-Butoxycarbonyl-L-leucine	6.19	242.1833	33.19	73.8	73.8	53.01
Deoxyinosine	7.73	262.119	24.93	76.12	76.12	70.15
UDP-N-acetyl-D-glucosamine	9.3	624.1375	23.26	77.71	77.71	40.65
Lys-Gly	8.85	211.1535	22.63	78.18	78.18	68.39
Glu-Thr	9.29	257.1305	19.11	78.68	78.68	74.41
indole-3-acetonitrile oxide	5.33	182.0967	56.8	80.34	80.34	68.32
Deoxycytidine	8.56	236.1203	45.42	90.3	90.3	50.24

N2-Acetyl-L-aminoadipate	8.92	211.1058	14.89	96.36	96.36	48.14
D-Xylulose	7.35	155.0693	44.26	98.15	98.15	96.41
2 -(Butylamido)-4-hydroxybutanoic acid	9.54	197.1267	72.72	106.61	106.61	54.02
Pro-Pro	8.47	222.149	24.53	113.55	113.55	61.25
Val-Arg	8.66	284.2163	59.08	115.9	115.9	41.22
LysoPC(18:0)	6.71	549.4497	24.33	138.35	138.35	60.9