

A metabonomic approach to the evaluation of xenobiotic metabolism and toxicity

Imperial College of Science, Technology and Medicine

Kyrillos N. Adesina-Georgiadis

2015

A metabonomic approach to the evaluation of xenobiotic metabolism and toxicity

Kyrillos N. Adesina-Georgiadis

Department of Surgery and Cancer

Imperial College London

Thesis submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy

Primary Supervisor:

Prof Jeremy K. Nicholson, Department of Surgery and Cancer

Secondary Supervisor:

Prof Elaine Holmes, Computational and Systems Medicine

Industrial Supervisor: Dr Peter Webborn, DMPK, AstraZeneca

Head of Group: Prof Ian D. Wilson, Department of Surgery and Cancer

Statement of Originality

This thesis is submitted in partial fulfilment of the requirements for award of the degree of Doctor of Philosophy in Clinical Medicine Research, Imperial College London. I hereby confirm that it is original work, representing my own academic effort and that all sources have been fully acknowledged.

Kyrillos Nikos Adesina-Georgiadis

August 2015

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

© Kyrillos Adesina Georgiadis 2015

Contents

i. Abstract	x
ii. Acknowledgements	xii
iii. List of figures	xiii
iv. List of tables	xix
v. List of abbreviations	xxi
Chapter One:	1
1.1 Introduction to Toxicology and Drug Metabolism	1
1.1.1 Principles of Xenobiotic Metabolism	2
1.2 Introduction to Metabonomics	19
1.2.1 Principles	20
1.2.2 Technologies in Metabonomics	20
1.3 Nuclear Magnetic Resonance	23
1.3.1 Principles of NMR Spectroscopy	23
1.4 Mass Spectrometry	36
1.4.1 Principles of Liquid Chromatography	36
1.4.2 Principles of Mass Spectrometry	49
1.5 Multivariate Statistical Analysis	57
1.5.1 Necessity for use with Metabonomics	57
1.5.2 Principal Component Analysis	57

1.5.3 Partial Least-Square Regression	59
1.6 Structural Assignment	61
1.7 Aims and Objectives	61
1.8 References	64
Chapter Two:	76
2.1 Metabonomics as a tool for the investigation of xenobiotic metabolism	77
Chapter Two Section A	79
2A.1 Introduction	81
2A.1.1 Diclofenac	81
2A.1.2 The hepatic reductase null mouse model	82
2A.1.3 Traditional vs Statistical drug metabolic investigation	84
2A.2 Materials and Methods	85
2A.2.1 Study Design and Sample Collection	85
2A.2.2 Sample Preparation and Experimental	85
2A.2.3 UPLC-MS Data analysis and statistical modelling	86
2A.3 Results and Discussion	87
2A.3.1 Quality control	87
2A.3.2 Chemometric analysis	88
2A.4 Conclusion and future perspectives	93
Chapter Two Section B	94

2B.1 Introduction	96
2B.1.1 Acetaminophen	96
2B.1.2 Metabonomic profiling as a tool for the investigation of acetaminophen metabolism and toxicity	98
2B.2 Materials and Methods	100
2B.2.1 Study Design	100
2B.2.2 Sample Preparation and Experimental Methods	101
2B.2.3 Data Analysis	102
2B.3 Results and Discussion	103
2B.3.1 Quality control	103
2B.3.2 Determination of batch effects	109
2B.3.3 Application of median fold normalisation for batch correction	115
2B.3.4 Effect on acetaminophen metabolites	118
2B.5 Conclusions	125
2B.6 References	126
Chapter Three:	133
3.1 Introduction	135
3.2 Study Details and Sample Preparation	138
3.2.1 Acute 2-Bromophenol Study	138
3.2.2 Chronic Acetaminophen Study	138

3.2.3 Liquid Chromatography -----	139
3.2.4 Mass Spectrometry -----	140
3.2.5 Data Analysis -----	140
3.3 Results and Discussion -----	141
3.3.1 2-Bromophenol -----	141
3.3.1a Quality Control -----	141
3.3.1b Comparison with traditional 2.1mm UPLC-MS profiling -----	146
3.3.1c Metabolite detection -----	150
3.3.2 Acetaminophen -----	151
3.3.2a Quality Control -----	151
3.3.2b Comparison with traditional 2.1mm UPLC-MS profiling -----	153
3.4 Conclusions -----	158
3.5 References -----	159
Chapter Four -----	162
4.1 Introduction -----	164
4.1.1 Toxicity and metabolism of bromobenzene -----	164
4.1.2 Toxicity and metabolism of 2-bromophenol -----	165
4.2 Materials and Methods -----	166
4.2.1 Study Design and Sample Collection -----	166
4.2.2 Sample Preparation and Experimental -----	167

4.2.3 ^1H NMR Spectroscopic Data analysis	167
4.3.4 Sample Preparation and Experimental (LC-MS Analysis)	168
4.3.5 UPLC-MS Data analysis	169
4.3.6 Statistical Modelling	169
4.3 Results and Discussion	171
4.3.1 Comparison with original data	171
4.3.2 Hippurate and, aminohippurate as potential biomarkers of 2-bromophenol nephrotoxicity	173
4.3.3 Insights from LC-MS analysis	183
4.3.4 Extraction of “isotopobromatograms” from the LC-MS data	190
4.3.5 Endogenous Metabolic Changes	192
4.3.6 Histopathology and Clinical Chemistry	197
4.4 Summary and Conclusions	199
4.5 References	201
Chapter Five	205
5.1 Introduction	207
5.1.1 Summary	207
5.1.2 Cumulative Toxicity	207
5.2 Materials and Methods	208
5.2.1 Study Design	208

5.2.2 Sample Preparation and Experimental for ^1H NMR spectroscopy	209
5.2.3 ^1H NMR Spectroscopic Data analysis	210
5.2.4 Sample Preparation and Experimental for UPLC-MS	210
5.2.5 UPLC-MS Data analysis	211
5.2.6 Statistical Modelling	212
5.2.7 ICP-MS	212
5.3 Results and Discussion	213
5.3.1 Cryogenically cooled ^1H NMR spectroscopy	213
5.3.2 UPLC-MS Quality Control	223
5.3.3 UPLC-MS Chemometric analysis	225
5.3.4 Model validation and significance testing	229
5.3.4 Inductively Coupled Plasma Mass Spectrometry	232
5.3.5 Histopathology and Clinical Chemistry	233
5.4 Summary and Conclusions	236
5.5 References	237
Chapter Six	241
6.1 Overall Conclusions	242
6.2 Future Work	248
Appendices	249

Abstract

This thesis comprises of studies which collectively describe the development and application of novel metabonomic approaches to the investigation of xenobiotic metabolism and toxicity. Both computational and UPLC-MS technologies are critically evaluated to facilitate the development of rapid, robust and high throughput analytical approaches for the analysis of compounds in biological matrices, particularly applicable in the area of drug discovery.

Pattern recognition statistical modelling approaches were employed on both a small scale metabolism study of diclofenac and a large toxicology of acetaminophen. This methodology highlighted large differences between the test groups in both studies; however these discriminations were only compound related in the latter study. In a study with low statistical power, the probability of biologically unrelated processes, overwhelming the statistical significance of the compound related effects are quite high. Application of statistical normalisation for the correction of batch effects was also performed in an attempt to emoliate the issue of limited continuous LC-MS instrumental capacity. Batch correction was attempted using variance stabilisation with a *glog* function prior to median fold change normalisation. However, this approach was unable to correct the data to the extent where it was suitable for combined downstream analysis.

A rapid gradient microbore UPLC-MS method is then described to provide a high-throughput solution for metabolic phenotyping of large sample collections. This method increases the flow rate, decreases the column diameter and the gradient elution time, reducing the length of the analysis of each individual while retaining the diagnostic capability of the experiment.

The combined insights gained through this method development process are then applied to two separate toxicity studies of 2-bromophenol, one of acute and one of chronic exposure. UPLC-MS analysis coupled with ^1H NMR spectroscopy identified the major metabolites as the sulphate and glucuronide conjugates, and other changes in endogenous metabolites.

However the lack of reactive intermediates or other toxic metabolites indicate that 2-bromophenol is not a nephrotoxin. While these studies failed to discover any compound related toxicity, they have exemplified the utility of the metabonomic approach in this area.

Acknowledgements

First and foremost I would like to thank God for giving me this opportunity and giving me the strength to proceed successfully through this PhD. It has been my privilege to be supervised and work alongside Prof Jeremy Nicholson and Prof Elaine Holmes, experts and pioneers in the field of metabonomics, who are a true inspiration for young researchers such as myself. Furthermore I am also grateful to my industrial Supervisor Dr Peter Webborn and AstraZeneca, UK, for their support and funding of this PhD.

After joining Imperial Prof Ian Wilson acted as not only as head of our small research group but also as my mentor, personal tutor and guidance councillor, and was undoubtedly one of the major factors which have lead me to completing my PhD. Dr Nicola Gray, part of the Waters team, has taught me everything I know about mass spectrometry, and I would like to thank her for her patience with both me and the SYNAPT. I couldn't have done it without her.

I must extend my thanks to many of our collaborators, for making this, and other work during my PhD possible. Dr Hannah Lees, Dr Marc Dumas, Dr Jia Li, Dr Panagiotis Vorkas, Dr Konstantina Spagou, Mr Muzaffar Anwar, Prof Alun Davies, Dr Leanne Nye and Dr Kirill Vaselkov, all from the Department of Surgery and Cancer, Imperial College, people who I have worked with and have learnt a lot from. I would like to thank Dr Dave Thompson (Keele) for his help with the ICP-MS. I would also like to thank Both Dr Eva Lenz and Dr Richard Goodwin from AstraZeneca for their contributions to this work.

I would like to thank all of my friends and other colleagues both within the department, but also generally at Imperial College, who have made my time at the university bearable. A special shout out to (the now) doctors Edward Lawton, Nan Zang, Jorrit Schafer, but also Tsz, Kazim and the rest of the ICLAC, as well as T&T's finest, Zoe Pierre. Another mention also goes out to the rest of 'Ian's students' who have made my final year hilarious. Finally I would like to thank my family, especially my mother, for their support and encouragement.

List of Figures

1.1 P450 mediated hydroxylation -----	6
1.2 P450 mediated Aromatic hydroxylation -----	7
1.3 P450 mediated Aliphatic hydroxylation -----	7
1.4 P450 mediated epoxidation -----	7
1.5 P450 mediated N-oxidation-----	8
1.6 P450 mediated N- dealkylation, S- dealkylation and O- dealkylation -----	8
1.7 P450 mediated oxidative deamination -----	9
1.8 Non-P450 mediated oxidation -----	10
1.9 Reductive reactions -----	11
1.10 Hydrolysis of esters and amides -----	12
1.11 Hydration of an epoxide -----	12
1.12 Molecular structure of glucuronide -----	13
1.13 Glucuronide conjugation of acetaminophen -----	13
1.14 Sulphoconjugation of bromophenol -----	15
1.15 Structure of glutathione -----	15
1.16 The glutathione conjugative process -----	16
1.17 Stages of the three-phase metabolism model -----	18
1.18 Rotation of positively charged nuclei -----	23
1.19 The two energy states of nuclei -----	24

1.20 Magnetic precession -----	27
1.21 Effect of field on magnetic precession -----	27
1.22 Net magnetisation of nuclei -----	28
1.23 Localised magnetic field -----	31
1.24 Effect of the local field on B_0 -----	32
1.25 NMR spectrum of 1,1,2 trichloroethane -----	34
1.26 Combinations of energy level states -----	35
1.27 Separation of a two part sample through a LC column -----	38
1.28 Components of a typical HPLC system -----	39
1.29 Resolution comparison -----	42
1.30 Theoretical plate model -----	43
1.31 General components of a mass spectrometer -----	50
1.32 The ionisation process by electrospray -----	51
1.33 Plasma induction -----	53
1.34 The quadrupole mass analyser -----	55
2A.1 The structure of diclofenac -----	81
2A.2 PCA scores plot of unfiltered, pareto scaled diclofenac data -----	88
2A.3 PCA scores plot of the diclofenac complete dataset -----	89
2A.4 PCA scores plot comparing the pre dose and 12-24h samples of the WT mice. -----	90
2A.5 S-plot of the PCA model -----	91

2B.1 The molecular structure of paracetamol -----	96
2B.2 The hepatic metabolism and fate of acetaminophen -----	98
2B.3 Overlaid chromatograms of the QC injections for the first batch -----	103
2B.4 PCA scores plot of the unfiltered data for batch one -----	104
2B.5 Overlaid chromatograms of the QC injections for the second batch -----	105
2B.6 PCA scores plot of the unfiltered data for batch two -----	105
2B.7 Stacked comparison of the chromatograms of the QCs -----	106
2B.8 PCA scores plot of the unfiltered data for the third batch -----	106
2B.9 Overlaid chromatograms of the QC injections for the fourth batch -----	107
2B.10 PCA scores plot of the unfiltered data for batch four -----	108
2B.11 Line plot plotting the retention times of detected features -----	111
2B.12 Line plot plotting the mass to charge ratios of detected features between batches -	112
2B.13 Bar plot plotting the peak area of the five most intense features -----	113
2B.14 Comparison of peak area and height for nine randomly selected metabolites -----	114
2B.15 Box plot of sample variation of the non-normalised peak intensities -----	115
2B.16 Dot plots of the standard deviation as a function of the rank of the mean of peak intensity of the QC samples -----	117
2B.17 Box plot of sample variation of the normalised peak intensities -----	118
2B.18 Scores plot from the OPLS-DA models for each batch -----	120

2B.19 Composite s-plot derived from the OPLS-DA models comparing high dose (800mg/kg) and control animals -----	121
2B.20 Composite plot of post batch-correction QC data -----	122
2B.21 Composite plot of post batch correction sample data -----	123
3.1 Overlaid chromatograms of the last QC injection of each run -----	142
3.2 PCA scores plot for the un-filtered data -----	143
3.3 Contribution plot showing the deviation of the QC samples for each ion -----	144
3.4 PCA scores plot depicting the five replicates -----	144
3.5 Bar chart comparing CV% -----	146
3.6 PCA scores plot comparison of methods -----	148
3.7 PCA scores plot comparison of the complete datasets -----	149
3.8. OPLS-DA S plot comparison of methods -----	151
3.9 PCA scores plots showing the initial unfiltered acetaminophen data -----	153
3.10 Overlaid chromatograms of ten selected QCs -----	154
3.11 PCA scores plot for the positive ESI acetaminophen data -----	155
3.12 Major metabolic signatures -----	156
3.13 S-plot of from the OPLS-DA acetaminophen modelling -----	157
4.1. The metabolism of bromobenzene -----	166
4.2 Spectrum comparison between original and repeat analyses -----	172
4.3 Contribution plot of NMR spectral QCs -----	175

4.4 The scores plots NMR data obtained 1H for the 2-bromophenol study -----	176
4.5 S-plot of the OPLS model comparing high dose and controls -----	177
4.6 Xvar plot of the levels of the target metabolites -----	178
4.7 Permutation plot for the NMR data model -----	180
4.8 Chromatographic peaks of the first, 10 th and QC -----	184
4.9 PCA scores plot showing the positioning of the QC data within statistical space -----	185
4.10 PCA scores plot of the model comparing the UPLC-MS data -----	187
4.11 PCA scores plot of the model comparing dose groups at 8 hours post dose -----	187
4.12. PCA scores plot of the model comparing dose groups -----	188
4.13 Composite plot highlighting two major 2-bromophenol metabolites -----	189
4.14. Isotopobromatogram and related mass spectra -----	191
4.15 Line charts of the average relative intensities for each dose group -----	194
4.16 Intensity plot of the contribution of the 2-bromophenol sulphate and 2-bromophenol glucuronide -----	195
4.17 Permutation plot for the UPLC-MS model -----	196
5.1 A comparison of typical control and high dose plasma NMR spectra -----	214
5.2. Scores plots of the PCA analysis of the plasma NMR spectral data -----	216
5.3. Loadings plot from the plasma OPLS-DA analysis with pseudospectrum -----	217
5.4. PCA scores plot of the urine NMR spectroscopy data -----	218
5.5 OPLS and S-plots urine samples comparing high dose and controls -----	219

5.6 Loadings plot of the Urine OPLS-DA analysis	220
5.7 A comparison of typical control and high dose urine NMR spectra	221
5.8 Aromatic region of a typical high dose sample	222
5.9 Overlaid chromatographic peaks of the QCs	223
5.10 PCA scores plot showing the positioning of the QC data within statistical space	224
5.11 Comparison of a typical control and a high dose chromatograms	225
5.12 PCA scores plots of urine UPLC-MS data	226
5.13. Scores and S-plot of the OPLS-DA model comparing the high dose and controls ---	228
5.14 Mass spectra of 2-bromophenol sulphate and glucuronide	229
5.15 Permutation plot for the urine UPLC-MS model	230
5.16. Isotopobromatogram and related mass spectra	231
5.17. Results from the bromine ICP-MS	233

List of Tables

Table 2A.1 Description of the major metabolic changes in the wild type mouse after diclofenac administration -----	91
Table 2A.2 Summary of the metabolites of diclofenac which were identified using conventional mass spectrometry methods -----	92
Table 2B.1 Summary of the statistics for the OPLS-DA models for each batch -----	119
Table 3.1 Table showing the number and percentage of ions satisfying the various CV threshold levels -----	145
Table 4.1 Results of the CV-ANOVA -----	180
Table 4.2, CV percentage scores for the QC data -----	186
Table 4.3 2-Bromophenol metabolites detected: Putative assignments -----	192
Table 4.4, Endogenous metabolites altered on 2-Bromophenol administration -----	193
Table 4.5 Results of the CV ANOVA significance test -----	196
Table 4.6, Clinical chemistry data from the COMET report -----	199
Table 5.1 CV percentage scores for the QC data -----	224
Table 5.2 CV ANOVA scores for the significance test -----	230
Table 5.3 Assignments of the analytes identified by bromine isotope pattern recognition -	232
Table 5.4 Histopathology observations -----	234
Table 5.5 Urine clinical chemistry parameters in the rats on Days 1, 2 and 6. Data are expressed as Mean $\pm$ Standard Deviation -----	235

Table 5.6 Summary of kidney injury biomarkers in the rats on Days 1, 2 and 6. Data are expressed as Mean $\pm$ Standard Deviation -----236

List of Abbreviations

^1H	Proton
<i>ADME</i>	Absorption, distribution, biotransformation and excretion
<i>APAP</i>	Acetaminophen (paracetamol)
<i>APCI</i>	Atmospheric pressure chemical ionisation
<i>ATP</i>	Adenosine triphosphate
<i>COMET</i>	Consortium for Metabonomic Toxicology
<i>CYP</i>	Cytochrome P450
<i>DA</i>	Discriminant analysis
<i>DIMS</i>	Direct infusion mass spectrometry
<i>EI</i>	Electron Impact
<i>ESI</i>	Electrospray ionisation
<i>FAD</i>	flavin adenine dinucleotide
<i>FID</i>	Free Induction Decay
<i>FMN</i>	flavin mononucleotide
<i>FT-ICR</i>	Fourier transform – ion cyclotron resonance
<i>GA</i>	Glucuronic acid
<i>GC</i>	Gas chromatography
<i>GSH</i>	Glutathione
<i>HETP</i>	Height Equivalent to a Theoretical Plate

<i>HILIC</i>	Hydrophilic Interaction Liquid Chromatography
<i>HPLC</i>	High pressure liquid chromatography
<i>ICP</i>	Inductively coupled plasma
<i>LC</i>	Liquid Chromatography
<i>MDR1</i>	multi-drug resistance protein 1 (glycoprotein)
<i>MRP1</i>	multi-drug resistance-associated protein 1
<i>MAS</i>	Magic Angle Spinning
<i>MS</i>	Mass spectrometry
<i>m/z</i>	Mass to charge ratio
<i>NAD</i>	nicotinamide adenine dinucleotide
<i>NADPH</i>	nicotinamide adenine dinucleotide phosphate
<i>NMR</i>	Nuclear magnetic resonance
<i>NOE</i>	Nuclear Overhauser Effect
<i>NPC</i>	Normal phase chromatography
<i>OPLS</i>	Orthogonal partial least square regression
<i>PAPS</i>	3'–phosphoadenosine 5'– phosphosulfate
<i>PCA</i>	Principal Component Analysis
<i>QC</i>	Quality control
<i>RPLC</i>	Reversed-phase liquid chromatography
<i>RT</i>	Retention time

<i>UDP</i>	Uridine diphosphate
<i>UGT</i>	Uridine diphosphate glucuronosyltransferases
<i>UPLC</i>	ultra-performance liquid chromatography
<i>STOCSY</i>	Statistical Total Correlation Spectroscopy
<i>SULT</i>	sulfotransferase
<i>TOF</i>	Time-of-flight

CHAPTER ONE

Introduction

This chapter provides an introduction into the background of xenobiotic metabolism, as well as the use of metabonomics in its study. The theory behind the analytical technologies employed within metabonomics will also be discussed.

1.1 Introduction to Toxicology and Drug Metabolism

1.1.1 Principles of Xenobiotic Metabolism

a. Xenobiotics and Drugs

The term Xenobiotics, derived from the Greek ξένος (xenos) to mean foreign and βίος (bios) to mean life or entity, describes any chemical substance which is not normally produced or expected to be found in an organism. While xenobiotics may be food additives, chemicals of abuse, pesticides, chemical pollutants or industrial by-products, a major category of concern are drugs. Understanding the behaviour of the drug inside a patient is key for their correct use for the treatment of disease and the prevention of adverse side effects. The routes by which an organism eliminates the xenobiotic is called metabolism, and this term describes the total fate of the xenobiotic, and includes absorption, distribution, biotransformation and excretion (collectively termed ADME); although this term is often used synonymously with biotransformation, especially when applied to drugs, as the products of drug biotransformation are called metabolites. Chemicals which are native to the body are also subject to similar processes and routes of biotransformation, called endogenous or endobiotic metabolism, although these pathways of metabolism are highly specific and carefully regulated. However, due to similarities and cross-over of enzymatic activity xenobiotics can also have effects on endogenous metabolism.

The combinations of biotransformation reactions are central to the behaviour of the foreign chemical entity within the body. Indeed few xenobiotics are actually directly excreted from the body unchanged as the parent compound. In most cases biotransformation of the xenobiotic precedes elimination from the host system, therefore any factor that affects the rate and type of metabolism that occurs is likely to alter the overall pharmacokinetic behaviour of the chemical. In most instances metabolic reactions alter the xenobiotic in a way that renders it inactive. However, in the case of some compounds this transformation can result in an active compound that may be responsible for the pharmacological effect

(drug activation). However, depending on the metabolic route taken and the parent compound, it is also possible that the biotransformation may result in a reactive intermediate, or a toxic metabolite, which may cause toxicity. Covalent binding of reactive intermediates to cellular proteins, inhibition of host cell endogenous metabolism, enzymatic inhibition or induction can all lead to unexpected toxicity.

b. Classification of Metabolism

The routes through which xenobiotics can be metabolised are many and varied; they included reactions of oxidation, reduction, hydrolysis, conjugation with endogenous moieties such as e.g. glucuronic acid, but also condensation. Traditionally these metabolic processes are separated into 'phases'; Phase I reactions, which include all the functionalisation reactions and Phase II reactions which involve conjugation. A third phase has recently been defined, Phase III, which involves the active transport of metabolites, however, this will be discussed separately. Phase I reactions have been described as being the preparatory step for Phase II metabolism; the initial functionalisation of the chemical by creating a chemical 'handle', through which Phase II reaction may occur. As such Phase II reactions are thought to be the main process of detoxification, with conjugates thought to be relatively inactive. These conjugates are usually more polar than the parent compound (although reactions such as e.g. N-acetylation actually increase lipophilicity) and may thus be more easily excreted from the body. However, despite the apparently logical order described above, frequently compounds can be targeted by both phase I and II reactions, and there is no clear order of metabolism; a product of Phase II may also be a suitable substrate for Phase I enzymes which in turn may lead to further conjugative reactions. Furthermore, after entering circulation, metabolites may be further modified by the gut microbiota, as gut microbes are known to de-conjugate many compounds (Ilett *et al.* 1990). The metabolism of a compound is therefore difficult at best to predict. However, it is important to study the main pathways as these are the most common routes through which pharmacological or toxicological activities are mediated.

c. Phase I metabolism – Particularly Cytochrome P450 mediated

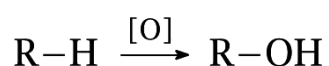
Phase I metabolic reactions, as mentioned before, have been thought of as functionalisation reactions and include oxidation, reduction, hydrolysis, hydration and isomerisation. The main avenue of phase I reactions are through the microsomal mixed-function oxidases, more commonly known as cytochrome P450 or CYP superfamily of enzymes (although flavin monooxidase, or FMO's etc. can also be involved in oxidative metabolism as described below) (Gibson and Skett, 1994).

Cytochrome P450s are heme-containing monooxygenases, with an iron protoporphyrin as the prosthetic group, and are the terminal component of the electron transport system present in the endoplasmic reticulum. These P450s are embedded in the membrane of the endoplasmic reticulum and have a broad range substrate specificity (Rendic and Di Carlo, 1997). CYP enzymes are heavily involved in endobiotic metabolism and are responsible for the biotransformation of steroids and cholesterol as well as bile acid biosynthesis, while they also mediate a large part of xenobiotic metabolism. As there are hundreds of CYP enzymes members of the superfamily are categorized into families and subfamilies and then further into isoforms based on amino acid homology. The system of classification is complex but essentially amounts to enzymes of the same subfamily have 70% or greater sequence homology with each other and members of the same family are at least 40% identical (Nebert *et al.*, 1987). CYP subfamilies one, two three and four, are the major variants involved in the metabolism of xenobiotics. However, adding to the complexity is the fact that many isoforms are differentially expressed and also largely polymorphic, especially in humans, contributing to the differences observed in xenobiotic metabolism from individual to individual (Ingelman-Sundberg *et al.*, 1999).

Some of the main isoenzymes involved in drug metabolism are CYP1A2, CYP2C9, CYP2C19, CYP2D6, and CYP3A4, and are of clinical significance. CYP1A2 has been shown

to metabolise a wide range of substrates, including caffeine and phenacetin, with common inducers of this isoform being compounds with aromatic rings. CYP2C9 is thought to generally metabolise drugs which are weak acids such as phenytoin, ibuprofen and losartan, the efficacy of which are greatly affected by the hypervariable properties of the gene. CYP2C9 is of particular note as it is easily modulated, being inhibited by many drugs (e.g. uconazole, and amiodarone) and activated by many (e.g. rifampicin); consequently it is an avenue of potentially severe drug-drug interactions (Smith *et al.*, 1995). Many important drugs, such as omeprazole, diazepam, and propranolol, are substrates for CYP2C19 (Wrighton and Stevens) whilst CYP2D6 is the main route of the metabolism of many anticancer drugs and CNS agents; both enzymes target hydrophilic weak bases. Finally CYP3A4, the most abundant of the Cytochrome P450s found in both the liver and intestine, is responsible for the metabolism of almost 50% of all known drugs with a wide, structurally diverse, substrate pool including common drugs such as paracetamol, midazolam and cyclosporine (Wrighton *et al.*, 2000).

The underlying metabolic reaction for all CYPs however is the catalysis of an oxidation; this reaction is generally described as:



Where R-H is the target drug and R-OH is the hydroxylated metabolite. During this reaction reduction agents are derived from reduced nicotinamide adenine dinucleotide phosphate (NADPH), and one atom of heterolytically-cleaved molecular oxygen is incorporated into the substrate resulting in C-, N- or S-hydroxylation and simultaneous release of water (Hannemann *et al.*, 2007). Central to this process is the electron transfer from NADPH to the P450, which is mediated by another membrane bound protein, cytochrome P450 reductase. Cytochrome P450 reductase is a flavin-containing protein with a flavin adenine dinucleotide (FAD) and a flavin mononeucleotide (FMN) sub domain. The electrons, initially accepted from NADPH, are shuttled sequentially through the FAD and FMN domains of the

cytochrome P450 reductase and then donated to the ferric protoporphyrin centre of the cytochrome P450 (Guengerich and Macdonald, 1984). The process is visually depicted in figure 1.1 below.

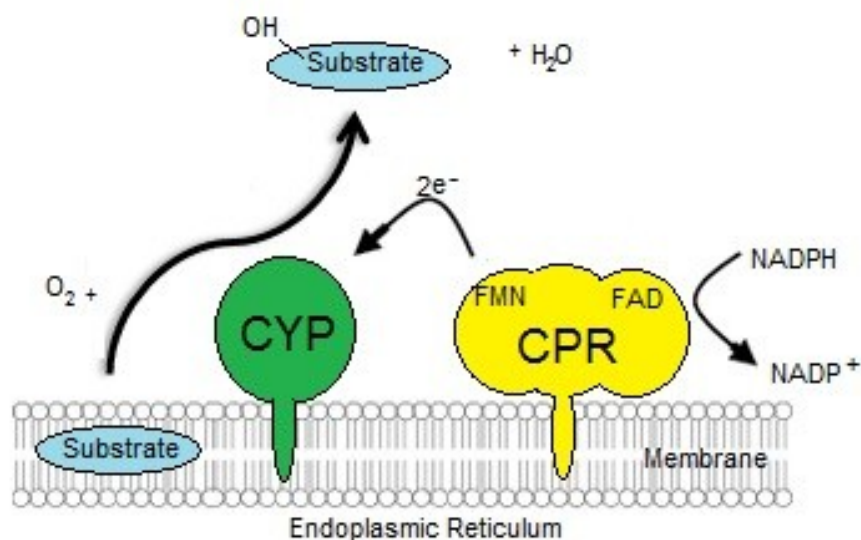


Fig 1.1 Visual representation of the P450 mediated hydroxylation in the endoplasmic reticulum membrane. Initial binding of the substrate to the cytochrome P450 (CYP) activates the cytochrome P450 oxidase (CPR). CPR acts as a transducer shuttling two electrons through its two Flavin domains (FAD and FMN) and subsequently donating these sequentially to CYP. This allows for the heterolysis of molecular oxygen one atom of which is incorporated into the substrate, also resulting in the release of water.

The initial addition of the oxygen during this process will usually be followed by a rearrangement or decomposition in the substrate depending on the type of chemical resulting in the formation of various types of final products. The more common reactions include.

Aromatic hydroxylation. This is the most common reaction for compounds which contain an aromatic ring. The hydroxyl group is attached to the ring resulting in a hydroxylated-derivative, as depicted in figure 1.2

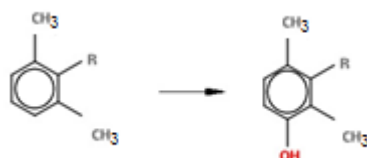


Fig 1.2 Visual representation of P450 mediated Aromatic hydroxylation. The example here shows the formation of a 3-hydroxy derivative, where R may be any prosthetic group or chain

Aliphatic hydroxylation. This is another common reaction, and the oxygen is usually added to terminal (ω) or penultimate ($\omega-1$) carbon atom in the substrate, as shown in figure 1.3.



Fig 1.3 Visual representation of P450 mediated aliphatic hydroxylation. The example here shows the hydroxylation on the terminal carbon, where R may be any prosthetic group or chain.

Epoxidation. Epoxides are generally cyclic ethers with a highly strained three atom ring, which makes them highly reactive. Generally epoxides are unstable intermediates, and can be highly toxic; they are able to bind DNA and cause genetic toxicity through point mutations and deletions are considered potent mutagens (Ehrenberg and Hussain, 1981). Most epoxides isolated are formed from polycyclic compounds and it has been suggested that they are the first step of aromatic hydroxylation (Gibson and Skett, 1994).



Fig 1.4 Visual representation of P450 mediated epoxidation. The example here shows the creation of the three atom ring, where R 1 to 4 may be any prosthetic group or chain.

N- and S- oxidation/hydroxylation. Both N- and S- oxide formation are possible reactions that may be mediated by P450s, although N- oxidation is also possible through other cellular N-oxidases. Many different substrates may be oxidised including amines, amides, hydrazines

and heterocyclic compounds. The oxides can then obtain a hydrogen from the environment to become a hydroxyl group.

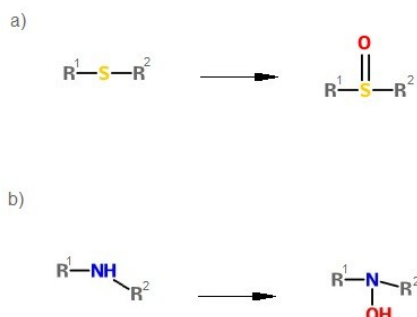


Fig 1.5 Visual representation of P450 mediated S-oxidation (a) and N-hydroxylation (b). The examples here show the creation an S-oxide in the a) and formation of a N-hydroxyl in b) which would be subsequent to a N-oxidation, and where R 1 and 2 may be any prosthetic groups or chains.

Dealkylation. This is a common reaction with compounds containing amines an alkoxy group or an alkyl; it is characterised by a loss of an alkyl group and is classified as N-, O- or S-dealkylation depending on what type of atom the alkyl group is attached to. The reaction is a two-step process where initial hydroxylation of the methyl group followed by a decomposition of the intermediate and the release of formaldehyde.

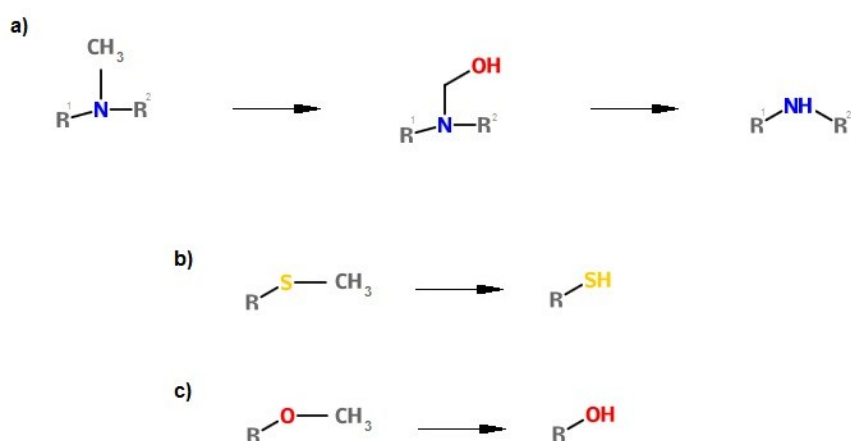


Fig 1.6 Visual representation of P450 mediated N- dealkylation (a), S- dealkylation (b) and O-dealkylation (c). The examples here show the two step process for the removal of the methyl in a), but omit the intermediate for b) and c); R 1 and 2 may be any prosthetic group or chain.

Oxidative deamination. The final P450 mediated reaction to be discussed involves amines and results in the release of ammonia. It is an aliphatic hydroxylation next to the heteroatom which forms a reactive intermediate which spontaneously decomposes to the deaminated form. P450s specifically target amines with the $-\text{CH}(\text{CH}_3)-\text{NH}_2$ structure, making this process distinct to other deaminations such as those of monoamine oxidase or in amino acid metabolism (Gordon and Skett, 1994)

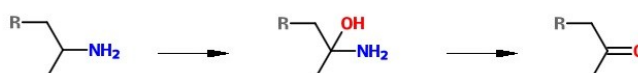


Fig 1.7 Visual representation of P450 mediated oxidative deamination. The example shown here includes the unstable intermediate which decomposes to release the ammonia; R may be any prosthetic group or chain.

d. Phase I – Other reactions

This section has primarily dealt with P450-mediated oxidative metabolism but there are many other enzymes which contribute to Phase I xenobiotic metabolism. These reactions are non-P450 oxidations, reductions, hydrolysis and hydration and will be summarised below.

Non-P450 Oxidation. Oxidation by other enzymes is also possible and is mediated mainly by alcohol dehydrogenase, aldehyde dehydrogenase or the previously mentioned amine oxidases (although many other enzymes are involved). Alcohol dehydrogenase is the primary source of ethanol clearance, and is involved in the metabolism of many other alcohols. Alcohol dehydrogenase forms a dimer and utilises coenzyme NAD to deprotonate two zinc ions within its active site to bind to the hydroxyl group of the alcohol, subsequently releasing the aldehyde product. Aldehyde dehydrogenase is one of a small group of enzymes which are able to oxidise aldehydes (the others being xanthine oxidase and aldehyde oxidase – they all operate in a similar manner). Through the reduction of

nicotinamide adenine dinucleotide (NAD) to NADH an initial hydrate is produced, resulting in the formation of the related carboxylic acid (Gordon and Skett, 1994).. Finally amine oxidases, either monoamine or diamine oxidase, are flavin-containing enzymes which degrade a wide range of both xenobiotic and endobiotic amines; the former is responsible for the endogenous metabolism of many catecholamines and important neurotransmitters such as norepinephrine, dopamine and serotonin (Weyler *et al.*, 1990) and the latter is mostly involved in the breakdown of endogenous polyamines. While amine oxidases primarily metabolise endogenous compounds, monoamine oxidases have been shown to participate in the metabolism of certain dietary components and drugs. The main avenue of reaction is the N-oxidation of their target. Generic examples of the discussed reactions are given below in figure 1.8.

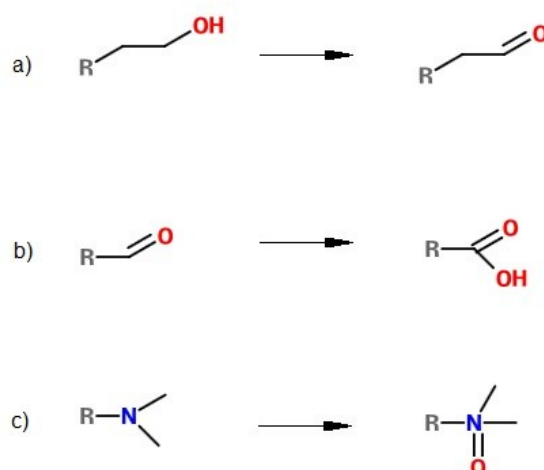


Fig 1.8 Visual representation of non-P450 mediated oxidation. The examples shown here are a) oxidation by alcohol dehydrogenase of an alcohol, b) oxidation of an aldehyde to its corresponding carboxylic acid by aldehyde dehydrogenase and c) monoamine oxidase mediated N-oxidation where the $N=O$ is a dative bond; R may be any prosthetic group or chain.

Reduction. Hepatic microsomes are known to be sites of reduction for a variety of chemicals; these include epoxides, heterocyclic compounds, halogenated hydrocarbons and azo- and nitro- compounds. Epoxides, which may have been the product of earlier epoxidation, are able to be converted back to parent chemical, and heterocyclic compounds may have their

ring structure cleaved. Halogenated hydrocarbons such as fluorocarbons can be dehalogenated. Azo- and nitro-reductions however are thought to proceed in two ways; either the double bond is severed to produce two derivatives or a NADPH mediated step-wise process which requires two molecules of the parent compound. Examples of each case are given below (figure 1.9) (Gordon and Skett, 1994).

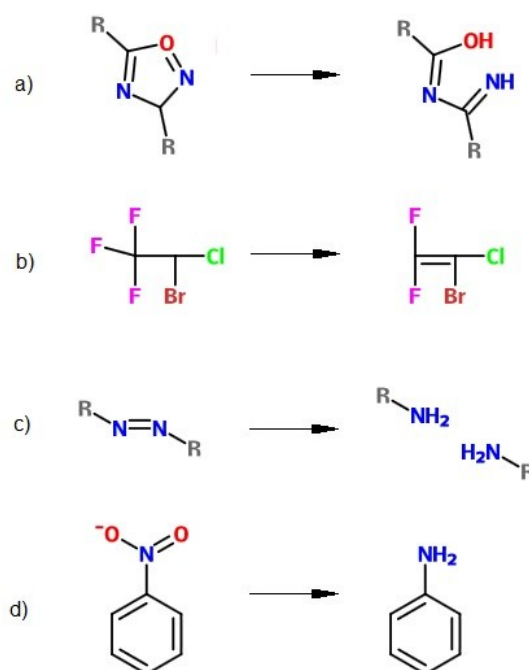


Fig 1.9 Visual representation of reductive reactions. A) heterocyclic reduction, b) the defluorination of haloethane, c) nitro-compound reduction and d) aromatic nitro-group reduction; where R 1 and 2 may be any prosthetic group or chain.

Hydrolysis/Hydration. Hydrolysis is a common reaction that occurs with many classes of compounds, but most commonly in esters and amides, catalysed by plasma esterases and amidases respectively, although esterases have been seen to hydrolyse the equivalent amide albeit with lower specificity. The hydrolytic reaction involves the breaking of bonds with a water molecule. Other compounds which can be hydrolysed include hydrazides and carbamates, and other compounds which already contain a functional moiety; more importantly hydrolysis of proteins and peptides is observed in the gut, which contributes to

the secondary metabolism of conjugates (including de-conjugation). Ester and amide hydrolysis is depicted below (figure 1.10).

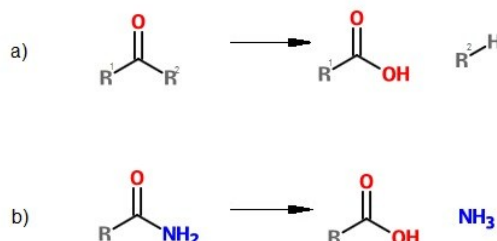


Fig 1.10 Visual representation of hydrolysis of a) esters and b) amides. These reactions are catalysed by the specific plasma enzymes, either esterases or amidases. R 1 and 2 may be any prosthetic group or chain.

Hydration is a very specific case of hydrolysis, and is the catalytic addition of water to a compound without the disassociation of the parent. In the body the most common form of this reaction is the transformation of epoxides into 1,2-diols or glycols by the enzyme epoxide hydrolase (Meynert *et al.*, 1970). This is one of the preferred routes of epoxide removal within the cell. The generic reaction is described in figure 1.11.



Fig 1.11 Visual representation of the hydration of an epoxide. This reaction is generally mediated by epoxide hydrolase, and produces either a glycol or, as in this case, a trans-diol. R1 and 2 may be any prosthetic group or chain.

e. Phase II – Conjugative metabolism

As mentioned previously, conjugative reactions have traditionally been thought to serve as the detoxification endpoint of xenobiotic metabolism. The main conjugation reactions are glucuronidation, sulphation, glutathione conjugation and acetylation.

Glucuronide conjugation. Glucuronidation is a common, and possibly the most important, Phase II reaction. Glucuronidation is part of the metabolism of many endogenous compounds such as hormones, hydroxysteroids, bilirubin and thyroxine (Kiang *et al.*, 2005),

but is also associated with the preparation of drugs for excretion. Uridine diphosphate (UDP)-glucuronosyltransferases (UGTs) are a superfamily of enzymes, and members of the UGT1 and UGT2 families are the key enzymes involved in the conjugative reaction, and due to the high levels of glucuronic acid in the cell, and the relative promiscuity of these enzymes, many xenobiotics are subjects for glucuronidation; in fact up to 70% of known drugs are processed, at least in part, this way (Wells et al., 2004). UGTs are membrane bound enzymes, similar to P450s, located in the endoplasmic reticulum, although their active site is located in the lumen. The UGTs utilise the co-factor UDPGA to transfer glucuronic acid onto a nucleophilic atom on the substrate, forming a conjugate which is highly polar, therefore with increased water solubility which allows easier excretion in the bile or urine. O-glucuronidation is the most common, and is the preferred substrate recognition site. The glucuronic acid in UDPGA is in the alpha configuration and in the drug glucuronide is converted to the beta form both in ether and ester linkages. The figures below depict the glucuronic acid (fig 1.12) and the O-linked glucuronidation of the common drug paracetamol (acetaminophen) (fig 1.13).

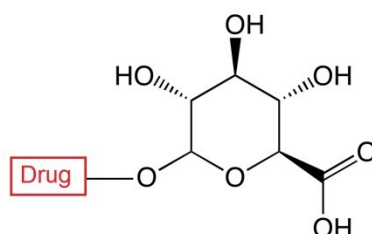


Fig 1.12 The molecular structure of a glucuronide. Glucuronic acid is normally attached to uridine diphosphate (UDP), which is then removed by UGT for the incorporation of the glucuronide into the target compound, at the site labelled 'Drug'.

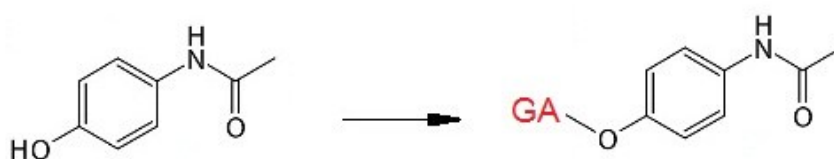


Fig 1.13 The glucuronide conjugation of acetaminophen (paracetamol), shown prior to metabolism on the left. O-linked glucuronidation occurs on the hydroxyl group and is the most abundant acetaminophen metabolite. GA is the added glucuronic acid.

Sulphation. Sulphation, or sulphoconjugation, constitutes another common route of conjugation in the detoxification process but is also heavily involved in endogenous metabolism. Sulphation is mediated by another superfamily of enzymes called sulfotransferases (SULTs), and generally involves the incorporation of a sulphonate to a hydroxyl or amino group on the target compound (O-/N- linked sulphation). The sulphonate is transferred from 3'-phosphoadenosine 5'- phosphosulfate (PAPS) and results in the formation of a sulfuric acid ester (Klaassen and Boles, 1997). Sulfotransferases are generally separated into two classes: membrane associated sulfotransferases and cytosolic sulfotransferases. The first class are embedded in the Golgi body and are concerned with the sulphation of large endogenous biomolecules such as carbohydrates, proteins and other peptides. For endogenous compounds sulphation has also been shown to increase target compound activity and is able to activate or modulate the function of many steroid hormones by altering binding site configuration (Gordon and Skett, 1994). The second class are cytosolic SULTs which while also involved in the metabolism of endogenous compounds such as hormones, catecholamines and vitamin D, are central to the sulphoconjugation of many xenobiotics (Glatt and Mehl, 2004). Similarly to glucuronidation, the addition of the sulphonyl group to the target leads to the formation of a more polar, water soluble, chemical which is then usually more readily excreted from the body. However, sulphation has also been implicated in the mutagenicity of many pro-mutagens such as 2-aminofluorene and 2-acetylaminofluorene (Chou et al., 1995). The SULT iso-enzyme SULT1A1 is perhaps the most important for drug metabolism as it is responsible for the sulphation of acetaminophen, minoxidil, dopamine and also most monocyclic phenols (such as 4-bromophenol) (Glatt and Mehl, 2004). The figure below describes the sulphoconjugation of 4-bromophenol.

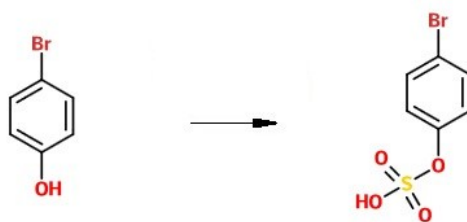


Fig 1.14 The sulfoconjugation of 4-bromophenol, shown prior to metabolism on the left. O-linked sulphation occurs on the hydroxyl group, although if prior hydroxylation on another carbon occurred prior then the location of the sulphonate may differ.

Glutathione conjugation. Glutathione (GSH) is a tripeptide comprised of an unusual γ -glutamyl, cysteinyl and glycine linkage (shown in figure 1.15).

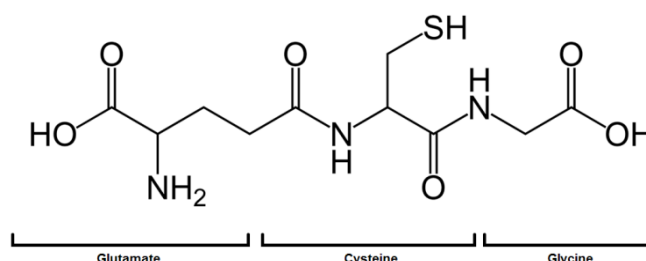


Fig 1.15 The molecular structure of glutathione. The three amino acid components are indicated.

GSH is an important factor in the cellular protection against reactive oxygen species and the general regulation of the redox state of the cell; indeed depletion of GSH is usually indicative of oxidative stress and lack of GSH is thought to instigate apoptosis (Hutter and Greene, 2000). Such thiol-containing compounds are reducing agents and GSH is also involved in the metabolism of both endogenous and exogenous compounds. GSH is also important as a means of detoxifying reactive intermediates formed through other metabolic reactions such as the oxidative Phase I metabolism of xenobiotics such as paracetamol. This detoxification process for reactive metabolites can occur directly via a simple chemical reaction or can be mediated by the glutathione S-transferase group of enzymes. These enzymes catalyse the nucleophilic attack of GSH on lipophilic compounds which contain an electrophilic component (usually a S-), but are also capable of mediating the hydroperoxide reduction of oxygen species (Gordon and Skett, 1994). Glutathione S-

transferases are classified into sub families' α , μ , π , θ and ζ with a similar system to other metabolism enzyme superfamilies based on sequence homology. These enzymes exist as dimers and one sub unit contains a glutathione binding site while the other contains the binding site for the electrophilic component of the target compound. As with most conjugative reactions the product is more polar and is quickly excreted into the bile (not the urine however) or may be transported to the kidney to undergo further transformation. The γ -glutamyl transpeptidase enzyme in nephrocytes is able to remove the γ -glutamyl unit of the GSH conjugate; the exposed cysteine will then usually be acetylated to a mercapturic acid, which is excreted in the urine (Commandeur *et al.*, 1995). The process of glutathione conjugation is shown below in figure 1.16.

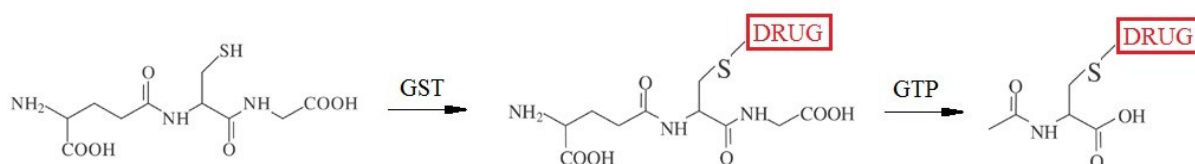


Fig 1.16 Visual representation of glutathione conjugation. Glutathione is added to the target drug through the action of glutathione S- transferase (GST), on an electrophilic atom. This conjugate is the either excreted via the bile or transported to the kidney where the second reaction takes place. Glytamyl transpeptidase (GTP) calatyses the removal of the glutamyl unit, leaving the cysteine residue exposed. This intermediate will usually be acetylated to form the mercapturic acid. The red box labelled 'drug' is the target compound.

Other phase II reactions. There are many other phase II reactions. These include:

- **Methylation:** The addition of a methyl group (CH_3) to the target compound, mediated by methyltransferases (Robertson, 2001).
- **Acetylation:** The addition of an acetyl group (COCH_3) to the target compound, usually mediated by either N-acetyltransferase or acetyl coenzyme A (Sim *et al.*, 2008).

- Amino acid conjugation: The addition of an amino acid, usually glycine, to the target compound. It is thought that this process occurs in the mitochondria on previously carboxylated substrates (Vessey and Kelley, 2001).

In general methylation and acetylation are primarily involved in the metabolism of endogenous compounds, but xenobiotic metabolism through these mechanisms has been seen. Amino acid conjugation, has been known for many years to be involved in xenobiotic metabolism and as there has been a recent resurgence of interest in this topic, it is being recognised as an important factor.

f. Transport

The traditional classification of Phase I and Phase II metabolism, discussed thus far, was introduced by Williams in 1959, and has generally become the 'go to' model for the metabolism of chemicals introduced into an organism. However, there exist issues with this metabolic workflow; studies observed that that Phase I and II reactions do not necessarily occur sequentially, as touched upon previously, but also this model ignores how the compounds involved move in and out of the cell. Metabolism studies had previously assumed that the compounds 'passively' diffuse in, and then diffuse out once modified, or altogether avoided the subject. It was not until the cloning of the drug transporter multi-drug resistance protein 1 (MDR1, also known as permeability glycoprotein) by Roninson *et al.* (1986) that transport was recognised as an important part of distribution and therefore relevant to drug metabolism. Following on from this first study Ishikawa (1992) and Zimniak *et al.* (1993) proceeded to characterise the carrier-mediated efflux of glutathione conjugates by multi-drug resistance-associated protein 1 (MRP1 – not to be confused with MDR1). It was then understood that these transporters are part of the ABC class of transmembrane carriers and are crucial for the elimination of the drugs and their metabolites. This process of

cellular efflux was termed 'phase III detoxification' which was then expanded to also include carrier mediated influx of parent compound (although this is sometimes separately termed Phase 0) (Petzinger and Geyer, 2006). Note that it is incorrect to call this phase 'phase III metabolism' as during the transport of the compounds there are no chemical alterations of the substrates. The importance of phase III is becoming ever more apparent as it has been recognised that in toxic situations, aberrations of normal transporter function, perhaps due to the saturation of carriers or the competitive binding of xenobiotic metabolites may be the mechanism of toxicity. Intracellular transport between sites of metabolism and transporters is carried out by cytosolic binding proteins and the cytoskeleton (Petzinger and Geyer, 2006). A simplified overview of the three phase system is included in the diagram bellow (figure 1.17).

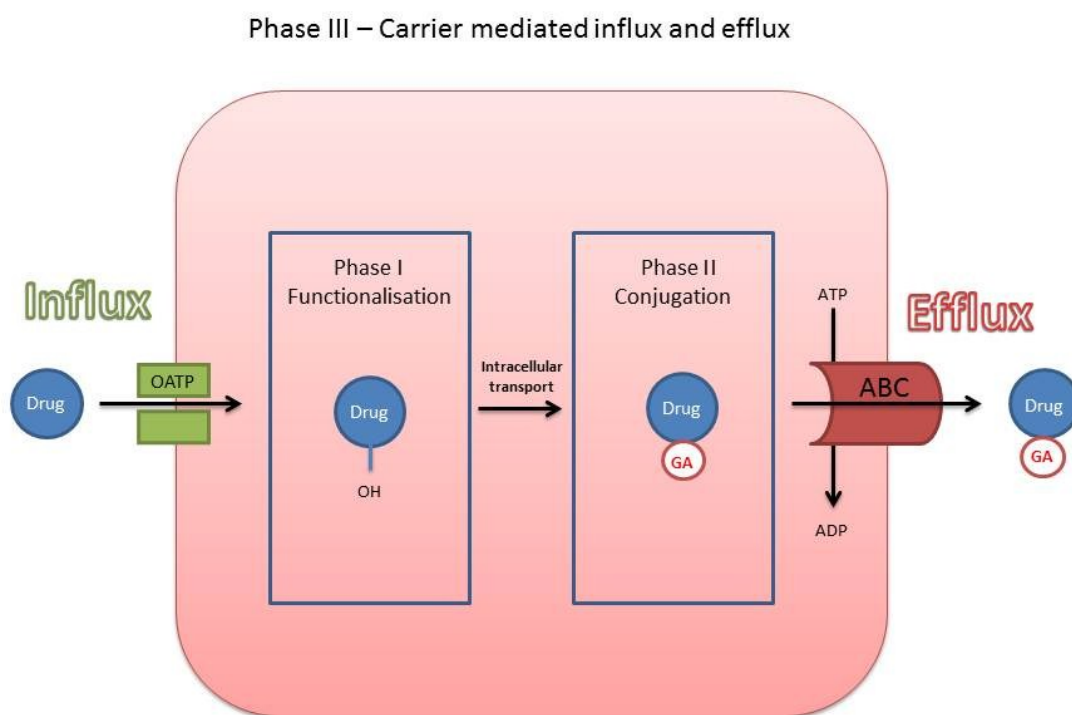


Fig 1.17 Diagram showing the various stages of the three-phase metabolism model as proposed by Petzinger and Geyer. Initial influx of parent compound into the cell is mediated by transporters mainly belonging to the solute carrier family of organic anion transporter polypeptides (OATPs) (Phase III/0). In this example the drug is initially subjected to Phase I functionalisation (hydroxylation in this case) and then Phase II conjugation (glucuronidation in the example). The conjugate is then secreted from the cell via the ABC antiporter transmembrane protein with the use of ATP during phase III detoxification.

1.2 Introduction to Metabonomics

1.2.1 Principles

Within the pharmaceutical industry the process of drug development has become riskier and more expensive. Traditional methods of compound screening frequently fail to identify unsuitable compounds early enough in the pipeline, resulting in much time and money being wasted in the validation steps. Indeed only around 1 in 50 validated lead compounds reach Phase I and only one in six drugs progresses to Phase IV clinical trials (Bains, 2004). There has therefore been a drive for the development of new techniques for the improvement of our understanding of the molecular and cellular alterations occurring and how these relate to pathophysiological and toxic events.

The role that ‘-omic’ technologies can fulfil in improving our understanding of the mechanisms of toxicity and in adverse drug reaction monitoring have been extensively discussed previously (Cockerell *et al.*, 2002). Metabonomics has been identified as one of the emerging, and now possibly established, components of the systems approach to toxicology. Indeed metabonomics has been applied extensively in the investigation of xenobiotic toxicity, and represented one of the first serious applications of the approach (see e.g. Nicholson *et al.*, 2002). Metabolomics was defined as ‘the quantitative measurement of time-related multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification’ (Nicholson *et al.*, 1999). As such Metabonomics can provide a direct picture of cellular activity, health status and the effects of drugs or the environment (Martis *et al.*, 2011) and therefore is an ideal tool for the identification of biomarkers and pathways of toxicity (Craig *et al.*, 2006), the construction of models for the prediction of toxicity (Schnackenberg *et al.*, 2009) and for regulatory purposes.

1.2.2 Technologies in Metabonomics

The experimental process a metabonomic study requires involves the combination of sensitive analytical instruments for the profiling of metabolites and small molecules in the target biological sample, such as serum, CSF, urine, saliva, and faeces (Kaddurah-Daouk *et al.*, 2011). The two main analytical platforms utilised are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). Here these technologies are quickly introduced, with the theory underlying their function discussed in more detail later in this chapter.

a. NMR-Based Metabonomics

In the case of the NMR platform, both ^1H and ^{13}C NMR spectroscopy have been used extensively, and the technique has played a pivotal role in the establishment of metabonomics. The advantages of using NMR spectroscopy include the quantitative nature of the output and the relatively simple sample preparation required. The technology is non-destructive and can also be used for non-invasive *in vivo* investigations (Serkova *et al.*, 2008). NMR spectroscopy is able to provide robust, reproducible quantitative and structural information, unlike other platforms where there are complications in both reproduction and quantification of results. A further advantage of the platform is through the utilisation of Magic Angle Spinning NMR spectroscopy (MAS) it is possible to obtain very high quality proton NMR spectra on small (10-mg) samples of whole tissue. MAS involves spinning the sample at 54.7° to the magnetic field direction, and requires no sample preparation, such as the arduous extraction methods required for MS tissue analysis. The drawbacks however, of NMR spectroscopy, include the lengthy instrument set up and run times, and most notably relatively low sensitivity resulting in the inability of the platform to detect low concentration molecules. These disadvantages however, are being addressed by the implementation of

automated workflows to reduce experimental time; the use of cyroprobes and higher strength magnetic fields has also improved the limits of detection.

b. MS Based Metabonomics

Recently the number of metabonomic studies using MS based technology has increased dramatically. MS is inherently a very diverse platform; there are both a wide range of ionisation methods and the biological sample under investigation can be delivered to the ion source either by direct infusion/injection (DIMS) or following a chromatographic step (so called “hyphenation”) e.g. liquid chromatography (LC-MS). While DIMS is simpler and can be faster, due to the simultaneous ionization of many analytes, of widely different ionisation efficiency, the phenomenon of ion suppression renders the detection of low ionisation efficiency molecules difficult or impossible. The introduction of a chromatographic front end allows molecules of different polarities to elute from the chromatographic column and be analysed separately, allowing for distinction of isobaric compounds. The two main types of chromatography employed are GC (gas chromatography) and LC (liquid chromatography), and their variants. These are usually paired with either Electron Impact (EI) ionization or Electrospray Ionization (ESI) respectively to deliver the ions to a mass analyser. There are in turn various types of mass analysers including Time-Of-Flight, Orbitrap and Fourier transform – ion cyclotron resonance (FT-ICR) mass spectrometers.

The individual properties of the various MS methods that are generally employed will be further discussed later in this chapter however, in general, the obvious advantage of MS is its, structure-dependent, high sensitivity and fast run times; MS is able to detect many subtle differences in all biofluids and tissue extracts. Furthermore, with the recent scientific interest in Mass Spectrometric Imaging (MSI) (Li *et al.*, 2008; Sugiura and Setou, 2010) it is now possible to explore the distribution of small molecules on tissue surfaces or other complex structures. However, all of these techniques require, to some extent, arduous sample preparation and complex instrumental set up. While data rich, the output is not directly

quantitative, as signal intensity is dependent on both concentration and on the ability of the molecule to ionize; a targeted quantitative assay is normally required after the identification of a compound of interest if concentration data are required. Furthermore analytical reproducibility is low, even within a run where detection variability is normalised using quality control samples; comparisons between runs are difficult to perform. Despite the drawbacks MS is a most sensitive analytical method and will continue to provide the backbone of hypothesis free metabonomic studies.

1.3 Nuclear Magnetic Resonance

Nuclear magnetic resonance spectroscopy was first developed shortly after the Second World War by American scientists working in Harvard and Stanford, and since then its applications within science have expanded much further than its original use for structure determination following chemical synthesis. In the subsequent 60 years NMR evolved into a major tool for the investigation of organic substances, from purity testing of drug compounds to medical radiology with Magnetic Resonance Imaging (MRI) technology routinely used in hospitals throughout the world. Although this technology is widely used, the underlying phenomena on which the spectroscopy technology is based on can be quite complex and conceptually hard to grasp. This is due to the fact these are grounded on physical concepts of electricity, nuclear magnetism and quantum mechanics; areas which, while familiar to the physical or organic chemist, may be beyond the expertise of the biologist or medical practitioner. It is possible for the technology to be used empirically, as a 'black box', and the output considered as just another experimental variable; however, an understanding of the underlying principles will surely enrich the appreciation of the system under study. Thus here an overview of the fundamental aspects of NMR spectroscopy is presented, aimed at the biomedical researcher or clinician. For an in-depth description of the chemical properties of nuclei for those more mathematically inclined the publication of Ernst *et al.*, (1987) is recommended; for more information on the many other applications of NMR spectroscopy

beyond those discussed here texts including Cavanagh *et al.* (1996) and Reid (1997) should be consulted.

1.3.1 Principles of NMR spectroscopy

Nuclei have a positive charge and some nuclei also possess a quantised property known as spin. The movement of any charged particle creates a magnetic field (see figure 1.18).

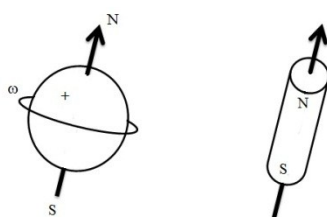


Figure 1.18. Rotation of positively charged nuclei creates a magnetic field similar to a small magnetic bar orientated along the axis of rotation. The created magnetic field is defined by the magnetic moment μ .

Thus the spinning nucleus acts as a small magnet orientated along the axis of rotation, and the magnetic field created corresponds to a magnetic moment termed μ . This moment is defined by:

$$\mu = \gamma_{\chi} \cdot I$$

Where I is the angular momentum, usually called nuclear spin and γ_{χ} is the gyromagnetic ratio constant. The magnetic moment μ is therefore proportional to I (i.e. higher momentum means a stronger magnetic moment). Any nucleus with a non-zero nuclear spin can be observed by NMR spectroscopy (if the spin is zero then there is no magnetic moment). While it is known that empirically all nuclei with odd numbers of protons do possess nuclear spin, there is no currently known law governing the level of spin. What is known however, is that most common isotopes such as ^{12}C and ^{16}O do not, and that biologically relevant isotopes ^1H , ^{13}C and ^{15}N all have nuclear spin quantum number of $I = 1/2$.

The initial orientation of the nuclear spin is random if no external forces exist. In nuclei with a nuclear spin of $1/2$ if an external strong magnetic field is applied to the system, then this will force the nucleus into one of two states: either with the orientation being parallel to the external field (low energy state) or anti parallel (high energy state). There is a significant difference in energy of the system between those two states and this is defined as ΔE , and this is shown in figure 1.19. It is important to note that in nuclei with different nuclear spin there may be more than two possible states, and the number of states can be calculated by $2I + 1$; e.g deuterium ^2H which has a spin of $I = 1$ has 3 possible energy states with magnetic quantum numbers of 1, 0 and -1.

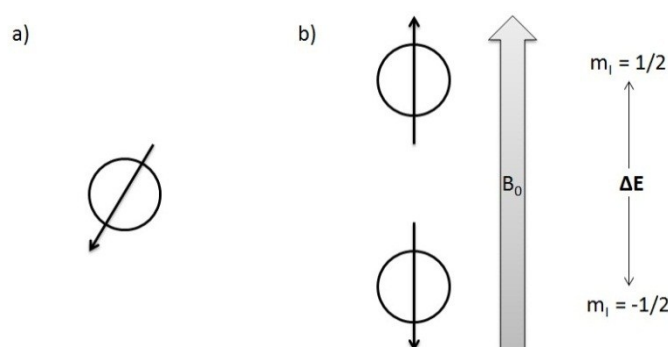


Figure 1.19. The axis of rotation of nuclei with spin is initially randomly orientated (a). With the introduction of an external magnetic field B_0 a nucleus with spin $I = \frac{1}{2}$ can either orientate its axis parallel to B_0 (low energy state) or anti parallel to B_0 (high energy state), where m_1 is the directional quantum number. ΔE is difference in energy between the two states.

It is much more physically favourable for the nucleus to be in the low energy state. The difference in energy between the two states is given by:

$$\Delta E = \gamma_X \cdot h \cdot B_0$$

This can also be found using this equation:

$$\Delta E = \frac{\gamma_X \cdot h \cdot B_0}{2\pi}$$

Where h is the Planck constant, B_0 the strength of the external magnetic field and γ_x the gyromagnetic ratio constant.

Of course in practice there are many more nuclei present within the system, each of which will orientate itself within the magnetic field in one of the two states. As it is physically favourable to be in the low energy state one might wonder why the nuclei do not all flip to the low energy state. However this is a consequence of competition between thermal motion and the Boltzmann distribution to achieve thermomagnetic equilibrium within the system. By utilising the Boltzmann law it is possible to calculate the distribution, and respectively the difference in number between the nuclei in the two states:

$$\frac{n_u}{n_l} = e^{\left(-\frac{\Delta E}{kT}\right)}$$

Where n_u and n_l are the number of nuclei in the higher and lower energy states respectively, k is the Boltzmann constant, T is the absolute temperature and ΔE the difference in energy between the two states; e is the exponential. As the difference between the energy states (ΔE) is very small when compared to the thermal energy of the system, using the principle of $e^{-x} \approx 1 - x$ (when x is small) it is possible to simplify the distribution equation to:

$$\frac{n_l - n_u}{n_l + n_u} = \frac{\Delta E}{2kT}$$

where the terms are as previously defined .

What can be surmised from this equation is that the difference in number between the nuclei in each of the states is dependent on thermal energy of the system and the strength of the magnetic field. This distribution gives the population an overall magnetisation which is termed M_0 , and this magnetisation is aligned along the axis of the external field. One thing that must be noted is that the difference in population between the two states is relatively

small, although dependant on the strength of the magnetic field; for example a field of 19 Tesla would only produce a difference of 64 nuclei per million. The overall magnetisation is effectively what we are able to detect as the signals from the rest of the nuclei cancel each other out, thus the ability to observe the nuclei is quite low, which is the reason that NMR spectroscopy suffers from relatively low sensitivity compared to some other spectroscopic techniques.

Nuclei within the external magnetic field are able to flip between the two energy states spontaneously, although this happens relatively rarely. It is however possible to induce this flipping by applying energy equal to ΔE of the system. If this energy is provided to the system in the form of a radiofrequency pulse (short bursts of electromagnetic energy which are able to produce magnetic fields) the absorption of this energy by the nuclei will cause the transition of both high energy states to low energy states and *via versa*. This palindromic movement produces a voltage which can be detected by coils of wire and can be depicted as a signal called free induction decay (FID), and this resonance phenomenon is the quantum mechanical basis of NMR. After the end of the radiofrequency pulse the nuclei will return to thermomagnetic equilibrium in a process called relaxation, which will be discussed later.

Mechanistically it is possible to consider that the nucleus found within the B_0 magnetic field will have its magnetic moment initially revolve around the axis of the B_0 magnetic field in a manner similar to a spinning top which has been spun at an angle on a friction free surface. This cone shaped movement is called the *Larmor precession*, which is shown below in figure 1.20.

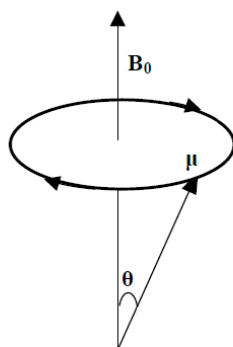


Figure 1.20. The magnetic moment μ of a charged nucleus within a strong external magnetic field B_0 will move in precession around the axis of direction of B_0 . The angle between the magnetic field and the magnetic moment is termed θ .

Now if it is assumed that there is a second, but smaller, external magnetic field applied to the system which is orthogonal (at 90°) to the orientation of B_0 , which is called B_1 . What would occur is that the smaller field B_1 would exert torque on the precession of μ pulling the rotation of the magnetic moment towards its orientation; essentially increasing the angle θ ; returning to the spinning top analogy, it could be imagined that this effect is as the top wobbling as it spun. This is depicted in figure 1.21.

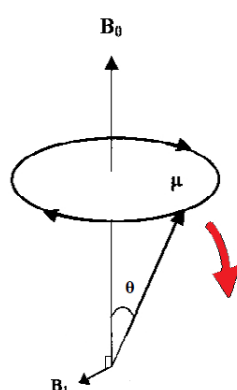


Figure 1.21. The precession of magnetic moment μ around the axis of B_0 is affected by the introduction of a second smaller magnetic field, B_1 which is at an orthogonal orientation to B_0 . This new field exerts torque upon the rotation of μ as shown by the red arrow. Thus the angle of precession θ is increased causing a 'wobble'.

If however the B1 field rotates at a speed equal to the rotation of the magnetic spin μ then there will be an additive and continuous effect on θ .

From what has already been discussed, it can be concluded that in a real situation there will be many more nuclei within the sample and that for nuclei with nuclear spin of $I = \frac{1}{2}$ the nuclear spins of these nuclei will align either parallel or anti parallel to B_0 , and that furthermore the differences of number between the two states will cause an overall magnetisation of the system, M_0 which is parallel to B_0 . If the orientation of B_0 and M_0 are defined as the 'z axis' then it is possible to visualise the system as nuclei in the low energy state precessing around the positive z and those in the high energy state precessing around negative z, as shown in figure 1.22.

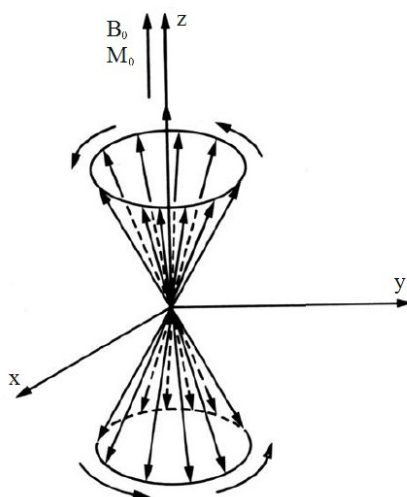


Figure 1.22. Each one of the arrows represents the magnetic spin of a nucleus. In a sample with many nuclei the segregation of nuclei between the high and low energy states in the presence of an external magnetic field B_0 will cause a net magnetisation M_0 which is parallel to B_0 . If we term this axis as Z then low and high energy state nuclei will precess around z and $-z$ respectively.

The frequency of the precession ν_0 can be described by extending the equation discussed previously:

$$\Delta E = \frac{\gamma_h \cdot h \cdot B_0}{2\pi}$$

However using the Bohr condition, $\Delta E = h \cdot \nu$, therefore:

$$h \nu_0 = \frac{\gamma_L \cdot h \cdot B_0}{2\pi}$$

$$\nu_0 = \frac{\gamma_L \cdot B_0}{2\pi}$$

This is called the Larmor equation. Where ν_0 is the Larmor precession frequency and the other terms as previously defined, and the angular frequency is expressed as:

$$\omega_0 = \gamma \cdot B_0$$

Where ω_0 is the angular frequency. While all these equations may seem convoluted and their significance vague, the essential concept equates to the following; to achieve nuclear magnetic absorption, the frequency of the radiation (the radiofrequency pulse) must be adjusted, and applied until the Larmor equation holds.

The application of a B_1 radiofrequency pulse, with a frequency appropriate for the type of nucleus that is being targeted, will result in nuclear magnetic absorption, and perturbation of the thermomagnetic equilibrium of the system. If this B_1 pulse is orientated along the x axis then the torque applied to the nuclear spins will be on the yz plane, pulling the spins in a manner that rotates the magnetisation of the system M_0 along that yz plane (Ernst *et al.* 1987) In one type of nuclear magnetic resonance experiment the B_1 pulse is maintained until the orientation of M_0 has rotated 90° and is then stopped (also called a $\pi/2$ pulse), leading to the aforementioned relaxation and return to equilibrium, and the production of the FID signal.

While it is beyond the scope of this thesis to fully discuss relaxation quantum physics, one note which will be briefly touched upon is signal broadening, and spin interactions. The

energy state switching from application of the B_1 pulse will eventually lead to saturation where the number of nuclei in each state is equal. From this situation the time it takes relaxation to return the system to equilibrium depends on two factors, spin-lattice interactions and spin-spin interactions, each is a first order rate process with its own time constant T_1 and T_2 . Spin-lattice interactions occur because there is energy transfer between the target nuclei and the heteronuclei in the environment, such as the medium, and spin-spin interactions occur when energy is transferred between the different target nuclei either intra- or inter-molecularly. These factors alter relaxation times and influence the FID signal received. Thus, it is thought that T_2 affects signal width under certain circumstances, and is a major problem in the study of macromolecules which have large numbers of identical nuclei interacting with each other reducing relaxation times, broadening peaks and thus reducing spectral resolution as broad peaks may overlap or completely obscure other signals.

The FID signal is initially represented dependent on the time domain; therefore a Fourier Transform is applied to the original signal. The Fourier Transform is a mathematical transformation which transforms a function into oscillatory functions, effectively decomposing a signal into its constituent frequencies. In the case of the signal-averaged FID signal, the output is converted into a frequency domain signal which is represented as the NMR spectrum.

While the above is theoretically true, in reality atoms are not comprised solely of the nucleus, but of the nucleus surrounded by electrons. These electrons are in the form of a negatively charged electronic cloud, and as moving charged particles are also subject to electromagnetic laws, and therefore also affected by the application of the B_0 magnetic field. Indeed B_0 induces circulation within the electron cloud which as predicted by Lenz's law will produce a small local magnetic field which is opposed to B_0 . This field is named B' and is shown in the below figure.

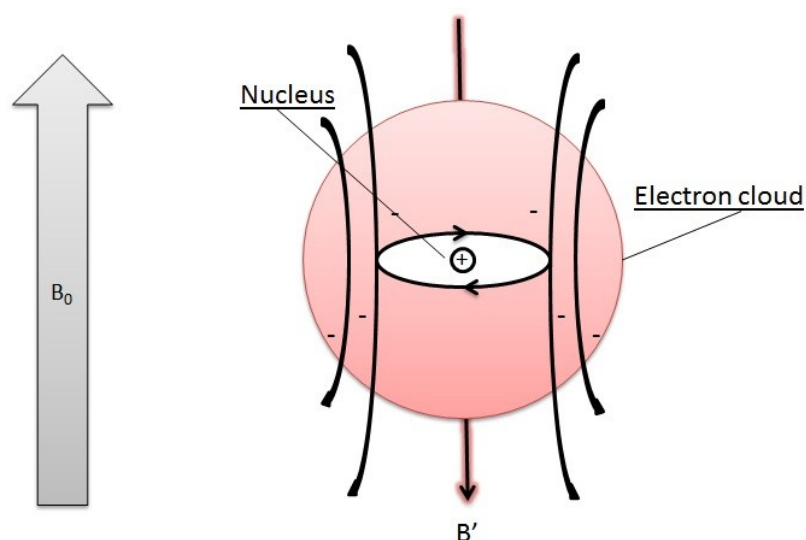


Figure 1.23. The application of external magnetic field B_0 causes rotation within the electron cloud, as shown by the curved arrows. This movement of electric charge produces a small localised magnetic field termed B' , which is opposed to B_0 .

Despite the B' field being very small and in practice perhaps a millionth of the size of field B_0 , it does serve to partially reduce the strength of the actual field experienced by the nucleus. This means that the frequency required for resonance will be slightly increased. This phenomenon is called screening or nuclear shielding and is expressed by a non-directional constant termed σ . This is important because this means that while the nuclei of different atoms will have different resonance frequencies due to the different spins, nuclei of the same atom can also have different absorption frequencies dependant on their local electromagnetic environment.

Initially one might think that this phenomenon is a disadvantage, as the presence of nuclear shielding results in uncertainty of the precise value of the field experienced by the nucleus and therefore makes the measurement of the gyromagnetic ratios difficult. However, in conventional NMR experiments the determination of the value of σ allows for a very robust way of elucidating molecular structure, as the screening constant is further affected by the electronic current of other nearby nuclei. As mentioned before the difference in resonance frequencies can be a million times smaller than the applied field, in the order of a few Hertz,

and furthermore this difference is dependent on the strength of this applied field. To overcome this problem σ is determined relatively to a reference signal, and it is possible to calculate the relative positions of two signals with high precision. This relative position is expressed in parts per million (ppm) or chemical shift (δ) and is calculated by dividing the difference between the target signal and the reference signal in Hertz by the strength of the magnetic field (usually MHz).

A further complication to the spectral signal produced is the fact that there are interactions between the spin states of the nuclei involved. This interaction is able to further alter local electromagnetic environments and the effective magnetic field. If a nucleus with spin $I = \frac{1}{2}$ is considered, i.e. only two possible energy states, such as hydrogen. When two of these nuclei are close enough to each other to exert a magnetic effect and the external B_0 field is applied, for each of the nuclei one of two possible scenarios exist. First that the orientation of the spin of the adjacent nucleus is such that its local magnetic field aligns with the applied field or second that it is opposed. This means that the B_0 field is either reinforced or diminished. This concept is more clearly demonstrated in the below figure.

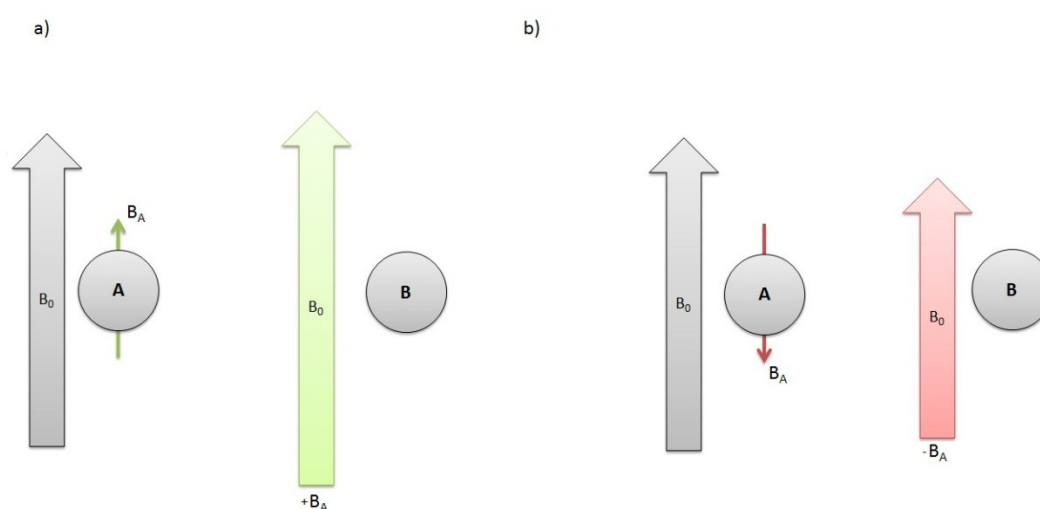


Figure 1.24. Let us assume two homonuclei labelled **A** and **B**, which are close enough to each other to exert an observable magnetic effect on each other. The application of an external magnetic field B_0 will result in nucleus **A** aligning its spin, and therefore its local magnetic field (B_A), either parallel or anti-parallel to B_0 . In case a) B_A is parallel to B_0 and therefore the effective magnetic field applied to **B** is $B_0 + B_A$ (green block arrow). In case b) B_A is opposed to B_0 and therefore the net magnetic force on **B** is $B_0 - B_A$ (red block arrow). In either case the effective field exerted on **B** will be either greater or lower than expected and will produce a slightly altered NMR signal. **B** will also affect **A** in the same way (not shown here). Note that the depicted green and red arrows greatly exaggerate the net effect on B_0 for demonstration purposes.

In practice, there will be a much larger number than two nuclei in a sample, and as each of the two spin orientations are equally as likely then the NMR signal of the above example would be split into two peaks, which are called doublets. The number of peaks a signal can be split into depends on the spin and number of nuclei which are in the close environment. This is called indirect spin-spin coupling. We will now work through a real example.

1,1,2 trichloroethane ($\text{C}_2\text{H}_3\text{Cl}_3$) is a small molecule which contains only three hydrogen atoms (see image below). The two hydrogens on the first carbon (circled red) are chemically equivalent as they share the same environment, whereas the hydrogen on the second carbon (circled blue) differs. If proton (H^1) NMR spectroscopy was performed on this molecule, two signals would be expected, one for the blue hydrogen and one larger signal from the two red hydrogens. These signals would be at $\delta 3.95$ and $\delta 5.77$. However, on examination of the NMR spectrum it is obvious that there are multiple peaks in both positions. In the case of the first doublet, using previous discussions it is possible to deduce that the splitting here occurs because blue hydrogen has aligned either parallel or anti-parallel to the applied field leading to two strengths of field being felt by the red hydrogens. Why, however, is the other signal a triple peak (triplet). This is because the blue hydrogen is affected by both of the red hydrogens. There are four possible combinations possible of spin orientation for the red hydrogens: 1) both can oppose the applied field reducing the magnetic force applied; 2) both could reinforce the applied field increasing the magnetic force applied; 3) the first hydrogen could be parallel and the second anti-parallel 4) the first can be anti-parallel and the second parallel. In options 3) and 4) as the nuclei are identical the net effect is zero. This means that the signal will be split into three peaks, with the middle peak representing 3) and 4) being twice the height as the other two.

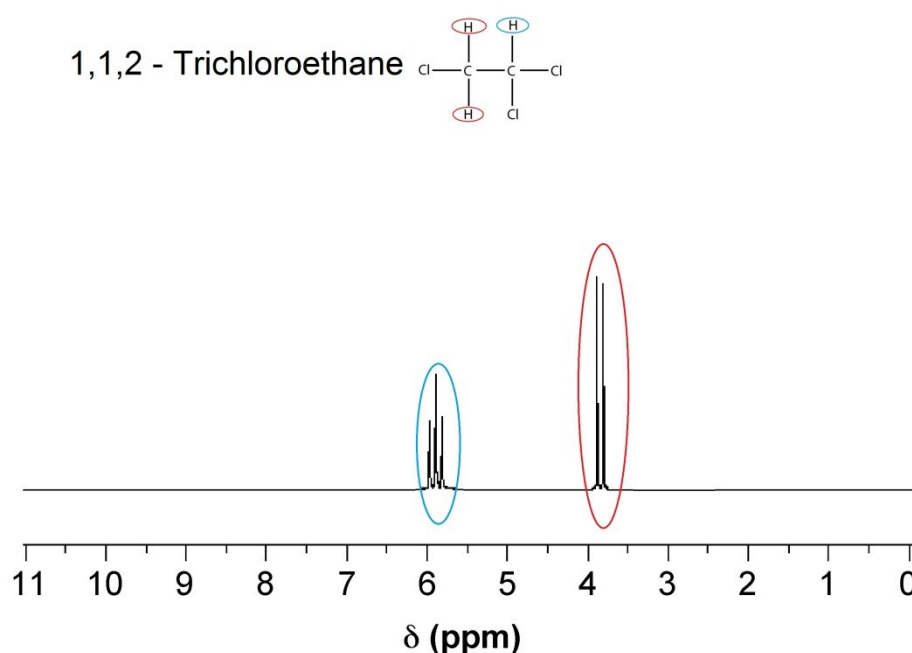


Figure 1.25. The NMR spectrum of 1,1,2 trichloroethane, with the chemical structure shown above. The red hydrogens produce the doublet while the blue hydrogen produces the triplet in the process explained above.

The final consideration for NMR experiments involving nuclei of $I = \frac{1}{2}$ is the nuclear Overhauser effect (NOE), which was first theorised by Albert Overhauser in the early nineteen fifties but wasn't experimentally proven until 1962 by Anderson and Freeman. The previous examples have considered nuclei which are close enough to magnetically interact due to scalar coupling; the interaction is through chemical bonds. However, it is possible that nuclei can interact as a result of the fact that they are spatially close to each other, because of the 3-dimensional configuration of the molecule rather, than being close due to chemical bonds. This is called dipolar coupling. In the situation where there are two dipole-dipole coupled nuclei during the relaxation process it is possible for the spins to affect each other.

If two nuclei are considered, again called **A** and **B** the figure below (fig 1.26) shows the four possible combinations of nuclear spins within an external field are shown, numbered one to four. If irradiated at the absorption frequency of **A** for long enough there will be a net movement from the low energy state of **A** to the high energy state leading to saturation, with

an equal number of nuclei in each energy level. After saturation has occurred during relaxation some high energy state **A** nuclei will flip to the lower state to achieve initial equilibrium. While for two uncoupled nuclei the process by which this would occur would be by the relaxing nuclei to flip independently, due to the dipolar interaction between the two nuclei in the example the flipping of **A** will 'drag' **B** along with it causing it to flip also. This means that the pair will either move from scenario 1) in the figure to scenario 2) in what is called double quantum shift (W_2) or from 3) to 4) in a zero order transition (W_0). If a double quantum shift happens then the effect on the population levels of **B** is that a nucleus in the higher state is forced into the lower energy state. If previous discussions are recalled about equilibrium and how there are more nuclei in the lower energy state normally this means that the difference in populations of **B** increases. As it is the difference in population that NMR detects what this means is that the signal received from **B** is bigger, thus a larger peak. Inversely when zero order transition occurs then a **B** lower energy state nucleus is forced into the higher energy level therefore decreasing the population difference and resulting in a smaller signal.

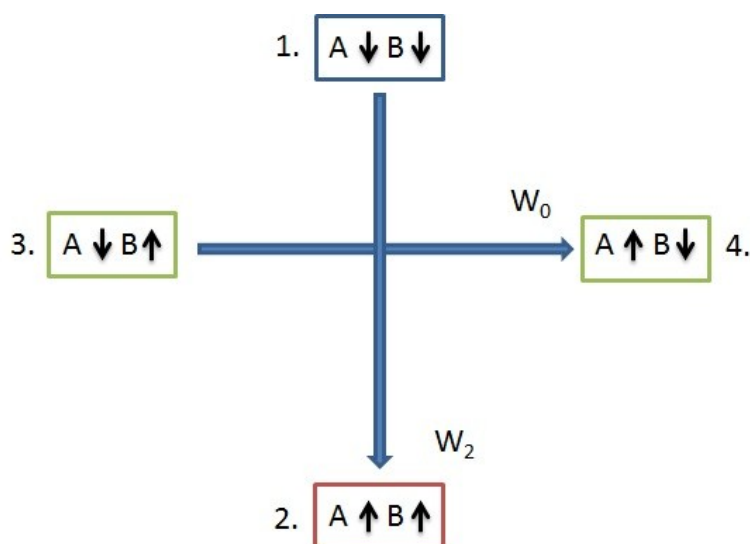


Figure 1.26. The various combinations of energy level states between the two nuclei **A** and **B**. The arrows pointing up or down next to the nucleus name indicate the low or high energy states respectively. During relaxation of two dipolar coupled nuclei the switching of state of **A** will also drag **B** into a forced switch by either the double quantum shift (W_2) pathway by a zero order transition (W_0). These transitions are the basis of the nuclear Overhauser effect.

1.4 Mass Spectrometry

The other major technology employed within metabonomics, and also many other fields of study, is mass spectrometry. Mass spectrometry is an analytical and structural chemistry technique that measures the mass to charge (m/z) ratio of ions. In contrast to spectroscopic techniques, such as NMR spectroscopy, which use some form of electromagnetic signatures to reveal chemical structure, i.e. to 'scope', mass spectrometry relies on chemical and physical reactions and therefore is a type of experiment. Coupled with this is the fact that mass spectrometry is a destructive technique that consumes the sample, and this is the reason for the term 'spectrometry'. While stand-alone use of mass spectrometry (direct infusion) is possible, within metabolic profiling mass spectrometry is generally employed with a front end hyphenated technology, most commonly liquid chromatography, but other technologies such as gas-chromatography or an inductively coupled plasma are also used. In these sections we will first explore the various methods of sample introduction into the ion source, before moving on to mass spectrometry itself.

1.4.1 Principles of Liquid Chromatography

Liquid chromatography is a technique used to separate different components of mixture, with each component being separated based on its interaction between two phases: a stationary phase, and a mobile phase which moves in a definite direction. Due to the fact that there are many different combinations of stationary and mobile phases that can be used for the separation, these various types of chromatography are categorised by the physical state of these phases. Chromatography was first developed at the beginning of the 20th century by Russian botanist Mikhail Tsvet, who separated plant chlorophylls by eluting the sample with ethanol through calcium carbonate. This rudimentary methodology was subsequently refined

in the 50s and 60s into what is today recognised as high performance liquid chromatography.

While this section will discuss the basic principles and underlying science behind liquid chromatography and its implementation, it is not an exhaustive description and a more detailed description of LC kinetics and technology can be found in e.g., 'Liquid Chromatography: Fundamentals and Instrumentation' by Fanali *et al.* (2013).

The general scheme of liquid chromatography is as follows. The components of the sample are separated by distribution between the two phases. The chemical and physical properties of each component determine what affinity that particular compound has for the mobile and stationary phases. Therefore, as the sample passes through the separation column sample components with high affinity for the stationary phase will progress through the column at a slower pace than those with a lower affinity, and thus will separate into bands and exit the column, or 'elute', at different times. These 'bands' can then be analysed with another analytical techniques, such as mass spectrometry. This concept is summarised in figure 1.27 below. The basic set up for liquid chromatography is also summarised in figure 1.28

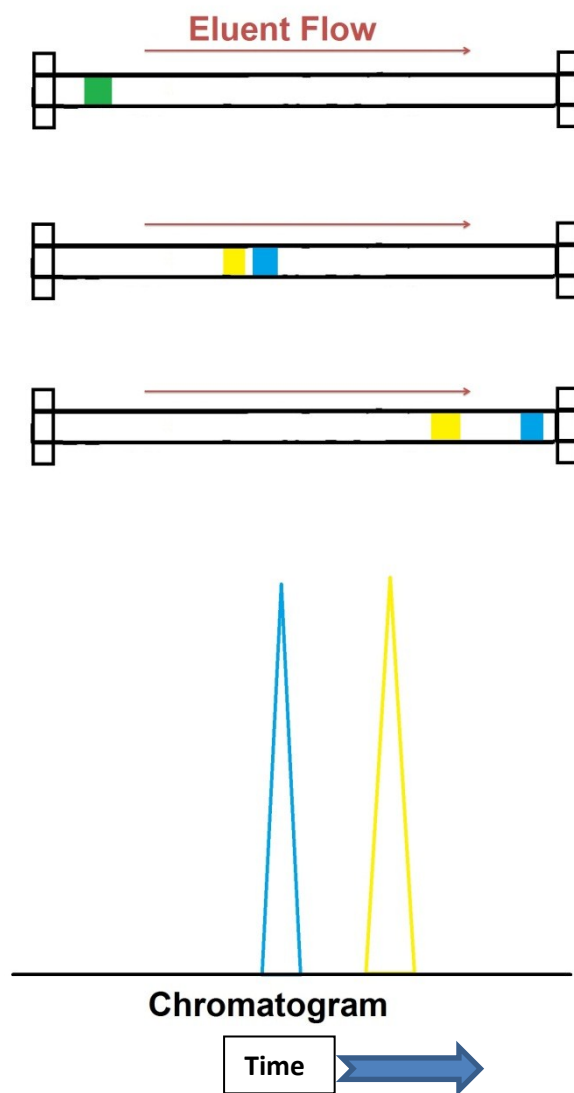


Figure 1.27. The separation of a two part sample through a LC column. The green area in the column represents the entry of the mixture into the column. As elution proceeds, the mixture separates into its two components, represented by the blue and yellow bands. The yellow component has a higher affinity for the stationary phase (column) therefore moves through the column at a slower pace than the blue component which has a lower affinity for the stationary phase. As the blue component elutes earlier its chromatographic peak is detected earlier in the chromatogram.

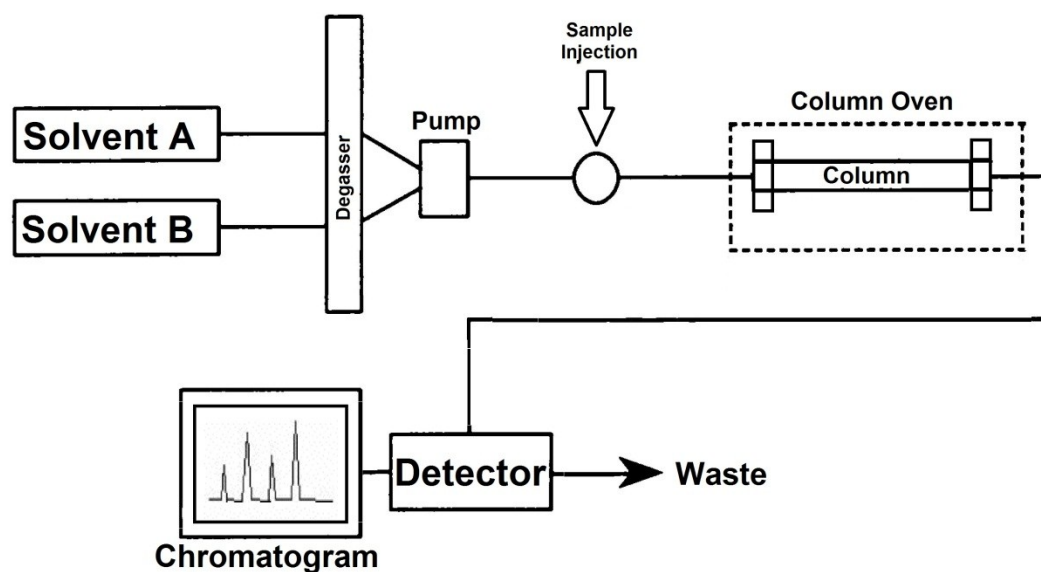


Figure 1.28. The components of a typical high performance liquid chromatography system. There are two solvents, possibly an organic and an aqueous, which are pumped through a degasser before the sample is injected. These then flow through the separation column, and the eluting compounds produce a chromatogram from a detector which converts the digital information acquired from the separation in to a visual representation.

The chromatogram produced by the detector at the end of every LC run provides several important parameters which are crucial in the analysis of the acquired data and are necessary for the quantification and identification of the components of the sample.

a. Void time

After the start of every chromatographic run, it takes some time for the mobile phase to reach the column from the point of injection, which is called the void time (or void volume). This will normally produce a distorted peak, or a baseline fluctuation, where any compounds which are not retained at all by the column pass through in the solvent front, and therefore also represents the amount of time any compound spends within the mobile phase.

b. Retention time

Chromatography is a time dependant separation; therefore the amount of time a particular compound spends within the system is the key characteristic of interest. This is measured by retention time t_R , which is defined as the time taken from injection to peak maximum for an individual component. This effectively calculated as:

$$t_R = t_S + t_M$$

Where t_S is time spent by the solute in the stationary phase and t_M is time spent in the mobile phase. As mentioned before, all components spend time equal to the void time, so retention time is often adjusted to take into account only time spent within the stationary phase.

c. Retention factor

Sometimes during chromatographic analysis it is necessary to identify whether a shift in retention time of a compound is due to actual retention within the column or if the shift is due to some change within the system. To differentiate between the two the retention factor k is used which is defined as the degree of retention of the compound of interest relative to another un-retrained analyte. This is calculated by:

$$k = \frac{t_R - t_0}{t_0}$$

Where t_R is the retention time of the peak in question, and t_0 is the retention time of the unretained peak, also equal to the void time. If k is consistent then the shift in retention is due to the system.

d. Selectivity

For two different compounds to be detected their retention times must be different, i.e., they must have different affinities for the stationary phase, or else they will elute at the same time, and it would be impossible to distinguish between them. As such the ability for a given LC system to separate two analytes is measured by the selectivity factor α which is the ratio of retention factors between the two compounds expressed as:

$$\alpha = \frac{k_B}{k_A}$$

Where k_A and k_B are the retention factors of the two analytes in question. For the system to be able to separate the two compounds the selectivity factor must be greater than one ($\alpha > 1$). Selectivity is directly related to the chemistry of the system, and depends on factors such as the type and strength of solvent used, column packing, temperature (an oft overlooked factor) and possibly mobile phase pH.

e. Resolution

While the selectivity factor is an indicator of the ability of the system to separate two chemicals it does not however, take into account peak shape or width. As retention time is measured from the peak maximum it is possible for two compounds to elute separately but have such broad peak base width that they are not resolved well. To quantify how well two (normally adjacent peaks) are resolved the resolution factor R_s is used and can be calculated as follows:

$$R_s = \frac{t_{R1} - t_{R2}}{0.5 (W_{b1} + W_{b2})}$$

Where T_{R1} and T_{R2} are the retention times of the two peaks, and W_{b1} and W_{b2} are the peak width at the peak base of the same two peaks.

The figure below (figure 1.29) illustrates how two peaks with the same selectivity (α) can have different resolutions.

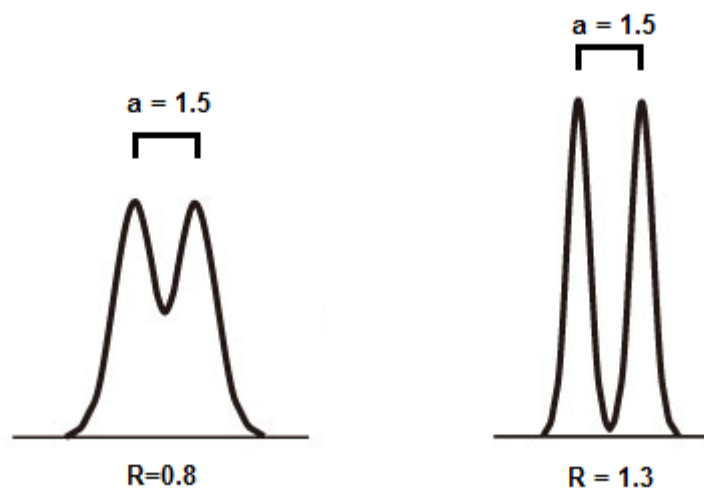


Figure 1.29. The two peaks shown here share the same selectivity value, 1.5, which is over 1 therefore of an acceptable range. However the two peaks on the left have wide peak bases and therefore overlap, leading to poor resolution, hence the low R_s value. The two peaks on the right however, have narrow peak bases and therefore have a better R_s . The situation on the right is therefore more favourable.

f. Peak symmetry

In the ideal scenario, chromatographic peaks would be symmetrical. However, most peaks exhibit a small degree of 'peak tailing' where the latter half of the peak is significantly wider than the first half; 'peak fronting', the reverse of peak tailing, although less common is also possible. This usually occurs due to the fact that more than one retention mechanism is present during separation and one of these mechanisms is being saturated. This phenomenon renders the acquisition of chromatographic signal data quite problematic, as it obfuscates calculations of retention time, peak area and peak width. Furthermore if there is a smaller peak eluting a short time after, it may be obscured in the former peak's tailing, especially if background noise is high.

To quantify this skewedness the asymmetry factor is used which is calculated using the peak width at 10% height. The asymmetry factor A_s is defined as:

$$A_s = \frac{B}{A}$$

Where B is the width of the latter half of the peak at 10% height and A is the width of the first half of the peak at 10% height. A value of 1 ($A_s = 1$) indicates that the peak is perfectly symmetrical whereas if $A_s > 1$ then the peak is tailing and if $A_s < 1$ then the peak is fronting.

g. Plate number

In traditional chromatography theory, the theoretical plate model is used to describe the chemical processes that are occurring in the separation column. In this theoretical model the column is composed of a series of stacked plates, a concept derived from distillation theory. As the mobile phase passes through the column it must equilibrate with each of the plates before it may pass to the next plate. Therefore the analytes proceed through the column through equilibrated mobile phase from plate to plate. Figure 1.30 depicts the column made up of the theoretical plates. It is important to note that these plates do not actually exist but are a convenient way of representing the elution process within the column.

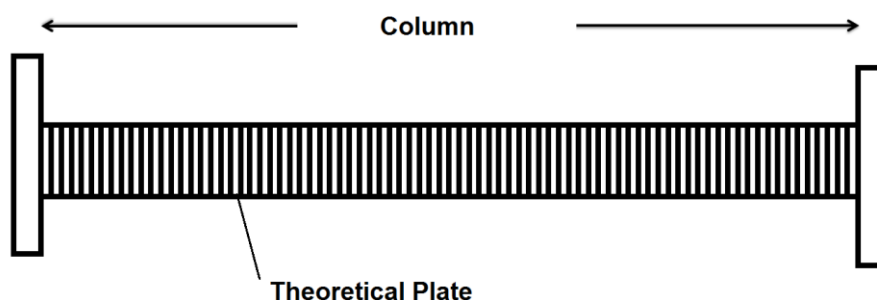


Figure 1.30. The theoretical plate model of chromatography. The separation column is made out of many theoretical plates, and a separate equilibration between the stationary and mobile phase occurs at each plate. The eluting analytes pass through the column by riding the equilibrated mobile phase through each plate.

The number of plates a real column possesses can be determined by utilising a single peak.

The number of plates, N , is calculated thus:

$$N = 5.55 \left(\frac{t_R}{w_{1/2}} \right)^2$$

Where t_R is the retention time of the selected peak, and $w_{1/2}$ is the peak width at half height and the 5.55 coefficient comes from $8\ln 2$ (peak width). The number of theoretical plates can also be used as a measure of column efficiency, either by directly quoting the number of plates, N , (where a higher number of plates means higher column efficiency) or by calculating the height of the plates. This plate height measurement is called the 'Height Equivalent to a Theoretical Plate' or *HETP* and is calculated as follows:

$$HETP = \frac{L}{N}$$

Where L is the column length and N is the number of plates. *HETP* is inversely proportionate to N and therefore a lower *HETP* is desirable. However this measurement is dependent on many parameters of the selected peak, and can change with different solutions. *HETP* is therefore not a measure of overall column quality but rather an indication of the relative suitability of the column for the specific experiment.

h. Band broadening

Band broadening is a phenomenon that occurs after a sample has entered a column and results from different molecules of the same band traveling through the column at different speeds. Generally this reduces the efficiency and quality of column separation, and can complicate both analyte identification and quantification. Furthermore the longer a particular analyte takes to pass through the column, the better it is retained, the higher the likelihood that it will be affected by band broadening. This band broadening can originate from various sources, which will be discussed below, and can include both characteristics of the column

but also may originate from other sources such as the sample injector or the detector. To measure the degree of band broadening the van Deemter equation is used (Gritti and Cuiochon, 2013).

The van Deemter equation calculates the relationship between the *HETP* and the three sources of band broadening, eddy Diffusion (*A*), longitudinal diffusion (*B*) and Mass transfer within the stationary and mobile phase (*C*). The van Deemter equation is as follows:

$$HETP = A + \frac{B}{u} + Cu$$

Where *u* is the linear flow velocity and the other terms are as before. The van Deemter equation also demonstrates the dependence of plate height (*HETP*) to the flow velocity.

The three factors contributing to band broadening are as mentioned previously eddy diffusion, longitudinal diffusion and mass transfer:

- *Eddy diffusion.* Eddy diffusion is the generic term used to describe variations in the flow rate of mobile phase through the column, or the analyte flow path. This relates to the fact that there are small inconsistencies in the homogeneity of the column packaging and the fact that some molecules of the same compound will take different paths through the column packing particles, resulting in some paths being longer than others. This phenomenon is directly related to column particle size, therefore columns with smaller particle size will reduce the effect of eddy diffusion, as there will be less space between particles and a more uniform trajectory for mobile phase flow. Eddy diffusion is not effected by linear velocity.
- *Longitudinal diffusion.* After injection the components of a chromatographic band will disperse outwards due to the naturally occurring concentration gradient present at the outer edges of the band. This is called longitudinal diffusion and results in the widening of the chromatographic band. While this diffusion is possible during the

whole time a compound is within the LC system, it is most prominent within the separation column where the diffusion is orientated along the axis of separation. As this is natural diffusion, the extent of diffusion is dependent on the time spent within the column, and as time spent is itself dependant on linear velocity, it follows that the dispersion will be inversely proportional to this velocity (the faster the mobile phase is flowing, the less time the compound has time to disperse within the column).

- *Mass transfer.* Mass transfer, or more specifically the resistance to mass transfer, is possibly the most important of the contributing factors to band broadening. As molecules are proceeding through the separation column they are constantly diffusing between the mobile and stationary phases. However, each molecule does not attain equilibrium between these stages with the same efficiency, and therefore disparities arise in the length of time taken for the migration of each molecule. The origins of this disparity are twofold. Firstly the depth of diffusion into the stationary phase particles varies, and therefore the deeper a molecule has diffused the longer it takes to diffuse back out into the mobile phase. On the other hand the microarchitecture of the column particles include depressions within the particle surface to maximise interface area. These however, may create localised 'stagnant' pools of mobile phase which are harder for the analytes to traverse, thereby reducing the distribution rate of the solutes between the phases. Both these phenomena result in variability in the time it takes each molecule to pass through the column, creating band broadening.

Besides what has been described above there are also other factors that could affect chromatographic performance and produce band broadening effects. Extra column factors include the sample injector, system tubing and the detector itself (Gritti and Guiochon, 2010). Furthermore these factors become increasingly important as their effect on columns with smaller diameter is ($ID < 2.1\text{mm}$) or columns of small particle size is much greater (Fountain

et al., 2009). This is important as the use of as special modifications to the system are necessary to use column of these types.

1.4.2 Modes of High pressure liquid chromatography

There are a wide range of chemicals and sample mixtures which can be analysed by liquid chromatography. However, there are different modes of operation for LC and each type is more suitable for a different set of target analytes. Indeed there are different combinations of mobile and stationary phase which make the system more specific to certain classes of compounds. The most commonly used modes of operation are the traditional normal phase chromatography (NPC) and reversed-phase chromatography (RPLC). Another widely used type of chromatography, geared towards separation of polar compounds, is hydrophilic interaction chromatography (HILIC). There are many other types of chromatography such as ion-exchange (IEC) (Light *et al.*, 1957) and size exclusion chromatography (Hong *et al.*, 2012) as well as many way of combining the chromatographic modes (mixed mode chromatography) (Oehme and Peters, 2010). However, within the realm of metabonomics reversed-phase and hydrophobic interaction chromatography are the two most widely used so only these two will be described in detail.

a. Reversed-phase liquid chromatography (RPLC)

The traditional method of performing chromatography, normal phase, revolved around a polar stationary phase, typically silica, and a non-polar mobile phase such as hexane. This interaction between analyte and the phases, mediated by the mobile phase, relied mostly on dipole-dipole and hydrogen bonding. While this is a very powerful separation technique for a very large number of chemicals, due to the emergence of RPLC and HILIC NPC has fallen out of favour.

The currently most widely employed form of LC is reversed-phase chromatography. RPLC is preferred due to the fact that it is highly reproducible and is suitable for a wide range of chemical separations including charged and moderately polar molecules. RPLC separation is dependent on the interaction between the analyte and a non-polar stationary phase and a more polar mobile phase. This is the opposite of traditional normal phase chromatography, hence the name 'reversed-phase'. The stationary phase is usually composed of a silica base upon which hydrophobic alkyl chains are attached, usually C₁₈ chains. The mobile phase as mentioned before is polar and is usually comprised of a mixture of an organic solvent such as methanol and an aqueous buffer. The separation of the analyses is based upon the hydrophobic interactions between the lipophilic areas of the analyte and the stationary phase. This results in non-polar compounds being better retained as the hydrophobic interactions with the stationary phase will be stronger. Another benefit of RPLC is the possibility to regulate the composition of the mobile phase by varying the respective levels of organic and aqueous solvents in the mobile phase composition. This allows the polarity of the mobile phase to be modulated also affecting the level of retention within the column; the more polar the mobile phase, the faster elution will occur.

b. Hydrophilic Interaction Liquid Chromatography

Compounds of weak or moderate polarity can be generally well separated by both normal and reversed-phase chromatography, albeit with opposite orders of elution. However, analytes with high polarity tend to be very weakly retained in RPLC, eluting very close to or within the solvent front. Conversely these same polar analytes tend to be very insoluble within the mobile phases used for NPLC, resulting in excessive retention within the column. In both scenarios it can be very difficult to acquire useful data for these compounds.

While it is possible to separate highly polar compounds with normal phase chromatography by using highly aqueous mobile phases it is not easy to do this in a repeatable way. This type of separation was extended by Alpert in 1990 into what is today known as HILIC. HILIC

can be simply described as a chromatographic technique that utilises a polar 'normal phase' stationary phase, with a reversed-phase type mobile phase with a very high percentage of organic solvent. In HILIC retention increases as polarity increases, similarly to NPLC, however due to the high organic content the mobile phase is less viscous, and therefore shows better separation for polar compounds. HILIC stationary phases are generally based either on silica or are hybrid materials including amides, amino and diol chains which offer a wide range of retention specificities.

HILIC has become ever more popular within the scientific community, and has been employed in the investigation of a wide range of biologically important polar compounds. HILIC has also proven to be very suitable for hyphenation with mass spectrometry, and is therefore widely used within metabolomic experiments. While HILIC has been studied extensively, the underlying mechanism of the interactions occurring within the column are still being debated, with hypothesised interactions including dipole-dipole interactions, hydrogen bonding and ion exchange (McAlley, 2010).

1.4.2 Principles of Mass spectrometry

Mass spectrometry is an analytical chemistry technique which is used to quantify known compounds, identify unknown chemicals within a sample and for structural elucidation of unknown analytes. The spectrometry process involves the conversion of the sample into an ion in the gas phase and then analysis by an ion detector which will sort the ions according to mass to charge (m/z) ratios and provide data on the relative abundances of each ion. The general concept is summarised in the below figure (figure 1.31) which details the typical components of a mass spectrometer. The ions generally must be in the gas phase for the application of either an electric or magnetic field to allow for ion separation.

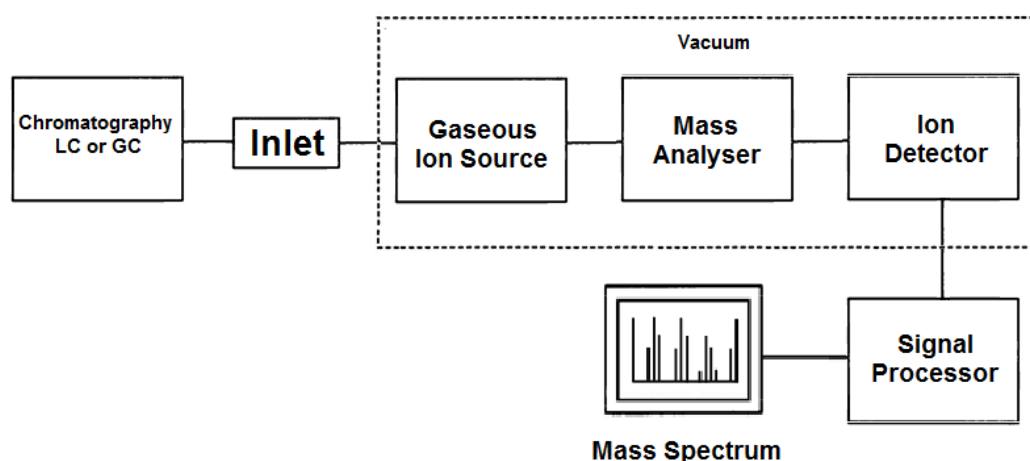


Figure 1.31. A block diagram depicting the general components of a mass spectrometer. After sample introduction gas phase ions are created in the ion source (ionisation), and these ions are then separated according to mass to charge ratio in the mass analyser. The relative abundance of these ions is then calculated in the ion detector and a mass spectrum produced by the signal processing data system.

While the general components for every mass spectrometer are consistent as shown in the above diagram, there are a variety of different ionisation techniques which have been employed, together with different types of mass analysers.

a. Ionisation

Mass spectrometry depends on the ionization of the target compounds to be able to detect and quantify them. Furthermore, for charge-to-mass ratio separation to occur the ions must be introduced in the gas phase into the vacuum of the detector. While this is relatively simple for samples already in the gas phase, such as after gas chromatography, or samples with heat volatile chemicals, it is more complicated when dealing with a LC interface. Traditional ionisation techniques include electron impact (where a gas phase sample is bombarded with electrons) and fast atom bombardment (where ionisation of non-volatile compounds is achieved by bombardment with a stream of either Ar or Xe atoms). These ionisation

methods are however, not very suitable to combination with liquid chromatography. The most commonly used methods are discussed below.

Electrospray Ionisation. Electrospray ionisation (ESI) is the ionisation method that has become the most widely used technique, especially in combined LC-MS systems (Fenn *et al.*, 1989). In this form of ionisation the sample solution is passed through a thin capillary tube, over the opening of which a large electrical charge of a few thousand volts is applied, at atmospheric pressure. This produces a spray of small charged droplets. The electrical current creates a potential difference between the capillary and a counter electrode which is situated near an opening in the vacuum system of the mass analyser. Depending on the voltage used this difference may be either positive or negative, which would attract ions of different charge. This results in two modes of operation for ESI, negative or positive ionisation, and each mode is efficient at ionising a wide range of molecules. As the droplets move through the ionisation chamber a desolvation gas is applied causing the solvent to evaporate increasing both the ion concentration but also the internal charge within the droplet. As the droplet gets smaller the increase in charge causes a 'coulombic explosion' releasing the ions. There are two chemical models by which this explosion is hypothesised to happen, the 'charge residue model' and the ion evaporation model. A detailed description of these models can be found in either Dole *et al.* (1968) or Thompson and Iribarne (1976). An overview of the electrospray process is summarised in the below figure.

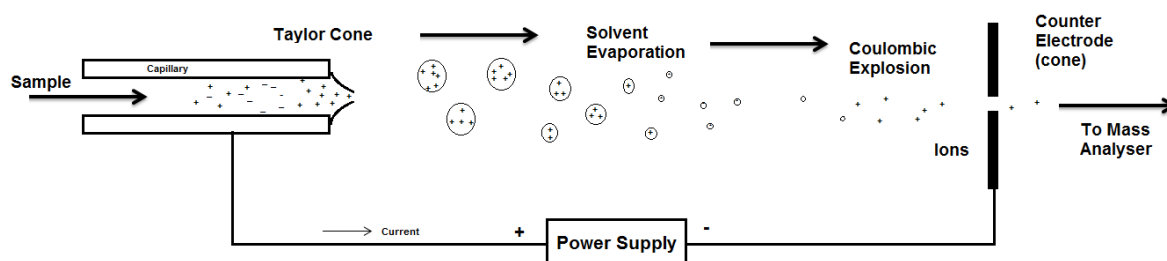


Figure 1.32. A diagram illustrating the ionisation process by electrospray. The application of an electrical current across a capillary opening naturally produces a fine mist spray. The electric potential differences between the capillary and the counter electrode at the other end of the ionisation chamber causes the droplets to move across the chamber, where the application of a desolvating gas, such as nitrogen, will cause solvent evaporation, increasing the internal concentration of ions. Once a charge threshold is reached, a coulombic explosion is caused releasing the ions which then enter the vacuum system of the mass analyser.

Electrospray ionisation has a wide range of suitable target compounds, and is especially useful for the detection of charged polar or basic compounds. Its utility is extended by the simple control of fragmentation by changing the interface potentials, and is very suitable for tandem MS/MS experiments, although this form of ionisation can be easily contaminated and requires quite complicated hardware. Extensions of ESI have been developed such as nano-spray variants which use a very low flow rate, or desorption electrospray ionisation, which extracts ions from a surface sample using the ESI spray.

Atmospheric pressure chemical ionisation (APCI). This method of ionisation uses similar hardware to ESI, and is another form of soft ionisation (Straube *et al.*, 2004). However, instead of the voltage being applied to the spray itself, it is applied to a needle size nebulizer which produces a corona discharge to produce ions (mostly H_3O^+). The sample solution is then introduced into the spray at atmospheric pressure, together with a heated gas that causes the mixture to become volatile. Ionisation occurs in the gas phase through proton transfer from the spray ions resulting in the production of analyte ions. These ions then collide with each other resulting in high ionisation efficiency and the production of both positive and negative analyte ions. APCI is predominately used for less polar compounds, for which ESI is unsuitable, and while the ionisation efficiency is very high, it is generally produces very little fragmentation.

Inductively coupled plasma (ICP) mass spectrometry. ICP is an ionisation method that is used for the quantification of specific elements of interest within a sample (Profrock and Prange, 2012). It is therefore fundamentally different from other types of mass spectrometry as it is not used for structural identification, but it is however innately quantitative, and will give an accurate measurement of the number of atoms within a sample of the target element. ICP is based on the use of a plasma torch to ablate the components of a sample into its constituent atoms. Initially the sample solution is introduced into a nebulizer where small droplets are formed on the injection needle through the action of argon gas. The sample aerosol then passes through a second needle into the torch.

The torch consists of three concentric quartz tubes around which a copper induction coil is wrapped. Argon gas is pumped into the space between the tubes and a spark is introduced to it from a Tesla unit, causing some of the gas to ionize. An oscillating radio frequency electric current is passed through the copper coil which causes the argon ions to interact with the magnetic field and accelerate in alternating directions colliding with other argon atoms creating very high temperatures. With sufficient argon and a high enough oscillation frequency, the plasma will reach equilibrium at temperatures over 6000°C. The sample spray is then introduced to the plasma where it will absorb high amounts of energy and be atomised. These atoms then pass through a vacuum chamber through two cones for rapid cooling of the ions, after which they will be introduced to the mass analyser (Skoog *et al.*, 2007). The general schematic is described in figure 1.33 below.

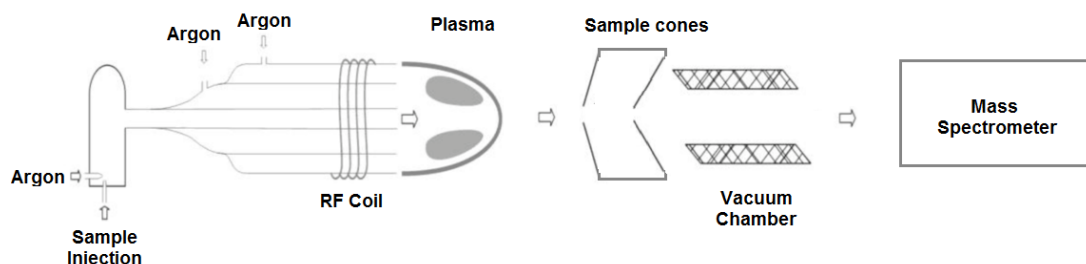


Figure 1.33. A diagram illustrating the plasma induction process. The sample is first nebulized by the introduction of argon gas. In the quartz tubing, argon gas is ignited by a spark and a radio-frequency oscillation causes the ionised gas to superheat. The sample is introduced into this plasma where it is atomised. The ions are then cool in a vacuum chamber before they are introduced into the mass analyser.

Modified from Elkadi *et al.*, 2014.

b. Mass analysers

Once the ions have been produced, they are accelerated through a vacuum into a mass analyser that will separate them according to their mass to charge ratio (M/Z). Various types

of mass analysers have been developed and each type has its own advantages and limitations. However, all mass analysers utilise electromagnetic fields to manipulate the ions and the net force applied on the ions can be summarised by Newton's second law from basic physics theory:

$$\mathbf{F} = m \mathbf{a}$$

Where F is the applied force, m the mass of the ion and a the acceleration. As this force is electromagnetic the above law can be extended by the inclusion of the Lorentz equation:

$$\mathbf{F} = e (\mathbf{E} + \mathbf{vB})$$

Where e is the charge of the ion, E is the strength of the electric field and vB is the product of ion velocity and the applied magnetic field. With the above two mathematical relationships we can surmise that the force applied to the ion is dependent on both the mass of the ion (m) but also the charge of the ion (e), and hence the mass analysers ability to separate by mass to charge ratio.

Quadrupole mass analyser. The quadrupole mass analyser is comprised of four parallel rods, as depicted in figure 1.34, which operates as a mass filter. Two of the rods are applied with a direct current, while the other two with a radio frequency current. Ions from the source are focused and then enter the space between the rods. The oscillations produced by the electromagnetic fields of the rods are such that the ions will experience a trajectory dependant on the mass to charge ratio. Ions with the correct trajectory for a specified m/z will reach the other side of the mass analyser and are able to reach the detector. Ions with the incorrect m/z ratio will have an unstable trajectory and collide with the rods and never reach the detector. The radio-frequency can be adjusted to allow only ions with a specific charge to enter the detector, and as such full mass range can be achieved by slowly scanning through the voltages. Further details for the underlying physics of ion behaviour, and how the force applied to the ion can be derived from what is called the Mathieu equation, the review by March (1997) is highly recommended.

Quadrupoles are also sometimes utilised in tandem, and combined in sequence to provide a system that enables fragmentation of the ions. This is very useful for structural elucidation of unknown compounds. The most common form of tandem quadrupole mass analyser is a triple quadrupole, where three quadrupoles are combined. The first set is used to select the ions of interest, the second set uses guide rods with only the radio-frequency current, and collision energy in combination with an inert gas such as nitrogen to fragment the ion and the third set of quadrupoles separates the fragments thereby allowing them to be detected.

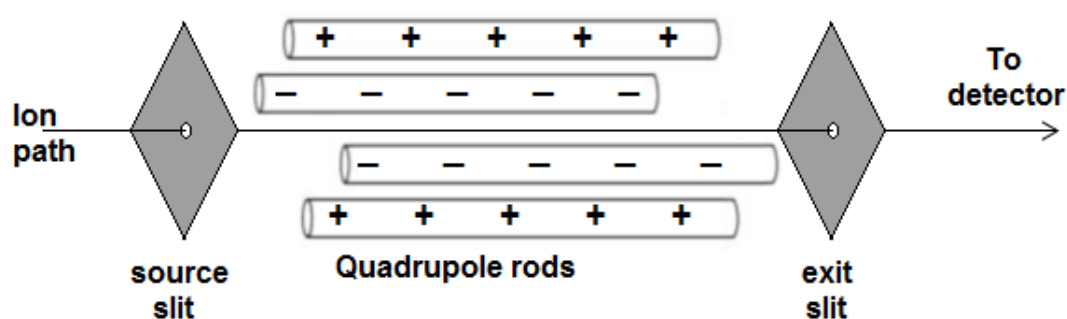


Figure 1.34. A diagram showing a rudimentary quadrupole mass analyser. Ions with the correct m/z matching the radio-frequency current applied to the rods will pass through the mass analyser and reach the detector.

Time of flight analyser. Time of flight (TOF) instruments are some of the most extensively utilised with mass spectrometry. TOF analysers operate on the basic principle of the time needed for a charged particle to travel through a homogenous electric field being dependant on the particle mass to charge ratio. After acceleration of ions by a regulated electric field their velocity will be dependent on the ionic charge. As the length of the mass analyser is known, it is possible to calculate the mass of the ion by using the time taken to travel across the flight tube. The kinetic energy (T) of a charged particle can be calculated by:

$$T = eV = \frac{mv^2}{2}$$

Where e is the ion charge, v is the velocity of the ion and m is the mass. However, the velocity of the ion is the length of the flight tube divided by the time taken to traverse it. By substituting this into the above equation and then solving for time:

$$t = L \sqrt{\frac{m}{2eV}}$$

Where t is the flight time, L is the length of the flight tube and the other terms are as before, it is possible to calculate the mass of the ion.

While the TOF analyser has the advantage of being able to analyse all ions, the length of the flight tube is of central importance as the separation depends on flight time. A longer flight tube will increase the resolution of the separation, and as such length of flight tube has always been the limiting factor for high resolution analysis. Recently the implementation of reflection technology has gained popularity. By introducing a linear field reflection and reversing the flight of the ions the effective length of the flight tube is extended. Furthermore this provides better resolution by eliminating some of the kinetic drift that arises from small variations in the initial kinetic energy of the ions; ions with larger kinetic momentum will penetrate the reflection field more, therefore will take a longer time to be reflected back.

Quadrupole time of flight analysers. As it is possible to combine several quadrupole analysers, it is also possible to combine quadrupole with TOF analysers to produce hybrid Q-TOF instruments. The combination of a quadrupole, a collision chamber and a TOF analyser has become the choice hybrid. The selectivity of the quadrupole for fragmentation coupled with the high resolution and accuracy of the TOF for fragment detection is the reason for this popularity in MS/MS experiments.

1.5 Multivariate Statistical Analysis

a. Necessity for use with Metabonomics

As previously mentioned metabonomics produces a very data rich output, that many times will include complicated or noisy data, which is multicollinear and can contain variation from multiple sources, including electronic noise. For the adequate interpretation and comprehension of the signals, visualisation and identification of the main sources of variation, the need for the use of more variables to characterise observations is apparent. The data analysis techniques routinely utilised within metabonomics include projection based statistics which can be unsupervised, such as Principal Component Analysis, or supervised, such as Partial Least Square Regression and its variants (Eriksson *et al*, 2004). Once again the underlying mathematical theory of the statistical techniques employed is beyond the scope of this thesis, thus only a basic overview and possible applications will be discussed.

b. Principal Component Analysis (PCA)

Principal Component Analysis is a very useful analytical technique because it is hypothesis free, and requires no previous information on the data. As with all projection based methods, a basic assumption of PCA is that the system under study is controlled by an unknown, yet limited, number of Latent variables. Latent variables are hypothetical variables that are inferred, therefore not directly observed. Thus, after obtaining our observed variables, PCA is used to construct further artificial variables which explain the variation in the observations. These artificial variables are called principal components, and can then be used for prediction in further models. PCA is in essence an eigenvector based analysis, the function of which is to identify underlying structures within data, revealing trends and outliers (Trygg *et al*, 2007).

While PCA is a powerful tool for pre-processing, most metabonomic studies, as indeed in most biomedical research, is concerned with the need to find the underlying stratification

which distinguishes between samples of different groupings. This of course includes classes of pathological, disease and metabolic status, different treatment groups and traditional case-control studies. In such cases sample classification can be incorporated to drive the analysis. This can be achieved by the implementation of supervised methods as described below.

Since in metabonomics the output includes a very large amount of variables (around 20,000 on an average resolution setting) PCA is used to describe the data in a lower dimension theoretical space. Each model may have a number of components equal or less than the number of variables in observation. The first model component is constructed using least square regression in a way that the component has the highest variance possible, thus explaining the largest variation within the observations. The second component will have the second largest variation, and so forth (Eriksson, 2006). Initially the n observations are projected onto a hyperspace with p dimensions, where p is the number of observed variables. In this theoretical hyperspace, a hyperplane with dimensions K is constructed from the principal components ($K_1, K_2, K_3 \dots$), each component being linearly orthogonal to each other. The position of every of the n observations within the plane is measured by a value called the score (T), and the position within the space by the loading (P). Each variable therefore has two co-ordinates. Any variation that is not explained by the model is called the residual, and is noted by E . Thus in summary the whole PCA model of a data set 'X' with observed variable number n can be given by:

$$X = T_1P_1T + T_2P_2T + \dots T_nP_nT + E$$

c. Partial Least Square regression (PLS)

Partial least square regression is a method used to project data on latent structures. Normal PLS is used when a quantitative relationship between two matrices X and Y is being investigated; the goal is to predict Y given X . While traditionally this could be achieved using

traditional multiple regression, due to the multicollinear nature of metabonomic datasets this is impossible. The X matrix could be spectral data, NMR signals or similar and Y would normally contain quantitative measurements of metabolite concentrations, a situation where the number of predictors is large compared to the number of observations. Even if Multiple Linear Regression was used, due to the large number of factors, the resultant model would be over-fitted, i.e. while the observed data is well explained, the model fails to predict new data (Trygg *et al*, 2007). However PLS avoids this problem, thus allowing for accurate interpretation of the data. Indeed by using linear regression multiple PLS components are created which can explain much of the sample variation.

Orthogonal PLS (O-PLS) and Orthogonal PLS Discrimination Analysis (O-PLS-DA) Are extensions to the PLS procedure. O-PLS uses a orthogonal signal correction to create an orthogonal component to remove unrelated sources of variation, and by using discriminate Y data can be further extended to O-PLS-DA; Y data can also be other outputs such as clinical parameters or toxicity score. Even though O-PLS-DA is very suitable for metabonomic studies, cross validation is routinely applied to minimise any potential over fitting. This cross validation involves removing a certain portion of data (usually one seventh) and testing the model to see how accurately it can infer the removed part. This process is repeated until every variable has been omitted once, and then the average percentage of successful imputation is acquired, which comprises the often quoted Q2 value. If the Q2 value is above 0.5 (i.e. more than 50% successful prediction) then the model is considered fit for data interpretation, though there are exceptions to this threshold, due to its arbitrary nature (Eriksson, 2006). A further step in model quality control can include creating random Y data and comparing Q2 values between the real and random models. If the Q2 values are significantly different this is a further indication that the model is appropriate.

In PLS the X and Y scores of the components are computed in a fashion so that consecutive scores can have the strongest possible correlation, making the model much more robust than in other redundancy analyses. PLS can also be extended in several ways; in many cases the Y matrix could comprise of qualitative “dummy” data used to represent

discrimination between classes. In this case PLS can be modified to a discriminate analysis, called PLS-DA. Here the model of X can be given by:

$$X = T \cdot P^T + E$$

Where the symbols are as before in PCA.

However, as mentioned before metabonomic data, especially ^1H NMR spectroscopic data, can be noisy, and can include variation which has no relation to the process under study. To negate this problem orthogonal-partial least square regression has been employed. Here the regression is modified so as to split the variation within the observations into two parts, one which is related to the Y data in a linear fashion and depicts the variation caused by it, and one part which is orthogonal to Y and includes all variation which is unrelated to Y (Cloarec *et al*, 2005). Thus in this situation the X model is given by:

$$X = T_p P_p^T + T_o P_o^T + E$$

Where T_p and P_p are the scores and loadings of the predictive variation and T_o and P_o are the scores and loadings of the Y-orthogonal components.

1.5 Structural Assignment

The final step, and ultimate goal, of metabonomic studies is the identification the metabolites that drive the discrimination between groups to obtain “biomarkers” which relate to specific phenotypes. These biomarkers may offer a direct link to the underlying biological pathways which result in the phenotype under study. To be able to identify which metabolite(s) are the ones detected by the statistics various techniques are used. Basic assignment involves observation of e.g., the NMR signal peak, chemical shift and comparing this to a list of standards available on central databases. Furthermore the J-coupling values of the peaks can be compared for a more confident assignment. However, sometimes due to the complex nature of NMR spectroscopic signals, peaks may overlap or be obscured making assignment difficult. In this situation 2D-NMR spectroscopy can be utilised, for separation of peaks. Statistical Total Correlation Spectroscopy (STOCSY) is an alternative approach which computes the correlation between variables in the spectrum, by taking advantage of the multicollinear nature of NMR data. This allows for the identification of related peaks and produces a pseudo- two-dimensional spectrum aiding the metabolite identification process (Cloarec *et al*, 2005). For the identification of compound via MS the M/Z is first used to construct a list of potential chemicals which fit the observations, and also any isotopic patterns present. These are then supplemented by fragmentation data and tandem MS/MS experiments depending on the instrumentation used. This will then lead to the final step which would be the development of a targeted method for the most confident assignments.

1.6 Aim and Objectives

As described earlier, drug discovery, is a complex process, with a very high failure rate. These failures can result from many causes but one major source of these has been toxicity and as part of a strategy to address this, the Consortium for Metabonomic Toxicology (COMET) was established in the early 2000s as international collaboration between Imperial College London and five major pharmaceutical companies: Bristol-Myers-Squibb, Eli Lilly & Co., Hoffmann-La Roche, NovoNordisk, and Pfizer, Inc. The objectives of COMET were a) to

provide a detailed multivariate description and biochemical characterisation of selected rat and mouse strains; b) to assess the ability of metabonomics to differentiate between toxic and non-toxic compounds and identify target organs of toxicity and c) the development of expert systems for the detection of toxicity of xenobiotic molecules (Lindon et al., 2003). To this end the project performed metabonomic studies deploying ^1H NMR spectroscopy as the analytical platform, and using urine and serum samples from rodents exposed to 147 selected compounds, in combination with histological and chemical data. The project was very successful and the results were used to generate a spectral database describing the model compounds (Lindon et al., 2005), which is still widely used for the drug discovery and development process within these companies (Spagou et al., 2011).

One other potential reason for late stage failure in drug development is that the full characterisation of drug metabolism in both preclinical species and man is often incomplete until the definitive radiolabelled studies have been performed. Until these studies are undertaken the detection of drug metabolites in the samples available from e.g., preliminary studies *in vitro*, or in animals (particularly in regulatory toxicology investigations) and early human clinical investigations, relies on the expertise of the DMPK scientist. Metabolite identification in this instance is heavily dependent on the skill of the operator, and this may result in some metabolites not being detected. However, the failure to fully characterise the metabolic fate of compounds at an early stage can result in delayed compound progression through development when they are finally identified in later investigations. This is especially the case where it is found that the species used for toxicity do not replicate the metabolic fate of the compound in humans requiring further toxicology studies in different species. However, in many metabonomic studies where drugs have been administered it has been the case that, on multivariate statistical analysis, both the parent compound and metabolites themselves have been found to be, at least in part, responsible for discriminating between dosed and control animals. The potential therefore exists to use the metabonomic approach as a method for the unbiased detection of the metabolic fate of compounds to supplement

the existing, expert-based, approaches to metabolite discovery. Given that the COMET project resulted in some 35.000 or so biological samples, an investigation of the potential for metabonomic detection of xenobiotic metabolism, based in the main on LC-MS analysis of samples, but also employing other analytical approaches where informative, was considered an appropriate use of this rich legacy. In addition, the re-profiling of COMET samples using LC-MS provides an opportunity to compare the results obtained with those previously acquired using ^1H NMR spectroscopy as a means of detecting biomarkers of toxicity. The aim of the research presented here was therefore as follows:

1. Evaluate the use of LC-MS based strategies as a means for the automatic and unbiased detection of xenobiotic metabolites in biofluids such as urine following the administration of drugs at therapeutic or toxic doses to rodents.
2. Examine, in addition to the detection of drug metabolites, a subset of the COMET samples to see if LC-MS-based methodologies could be applied used to detect signatures specific for mechanism-based toxicity.
3. Compare these LC-MS methodologies, where appropriate to the data that could be acquired using current methods based on ^1H NMR spectroscopy including cryoprobe technology.
4. Determine whether it would be practical to devise a workflow that allowed both biomarkers of toxicity and drug metabolites to be performed.

1.8 References

Alpert, A.J., (1990). Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *Journal of Chromatography*, **499**, 177-196.

Anderson, W.A. and Freeman, R. (1962) Influence of a second radiofrequency field on high-resolution nuclear magnetic resonance spectra. *The Journal of Chemical Physics*, **37**, 85-103.

Bains, W., (2004). Failure rates in drug discovery and development: will we ever get any better? *Drug Discovery World*, **F**, 9–17.

Banwell, C.N. (1994). Fundamentals of molecular spectroscopy (4th edn.). *Mc Graw Hill*, New York.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton., N.J., 1996. Protein NMR Spectroscopy: Principles and Practice. *Academic Press, Inc.*, San Diego

Chou, H.C., Lang, N.P., Kadluba, R.F.F., (1995). Metabolic activation of N-hydroxy arylamines and N-hydroxy heterocyclic amines by human sulfotransferase(s), *Cancer Research*, **55**, 525-531.

Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., Nicholson, J. K., and Holmes, E. (2005). Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in H-1 NMR spectroscopic metabonomic studies. *Analytical Chemistry* **77**, 517-526.

Cockerell, G.L., McKim, J.M., Vonderfecht, S.L., (2002). Strategic importance of research support through pathology. *Toxicology and Pathology*, **30**, 4–7.

Commandeur, J.N., Stijntjes, G.J., Vermeulen, N.P., (1995). Enzymes and transport systems involved in the formation and disposition of glutathione S–conjugates. Role in bioactivation and detoxication mechanisms of xenobiotics. *Pharmacological Reviews*, **47**, 271–330.

Craig, A., Sidaway, J., Holmes, E., Orton, T., Jackson, D., Rowlinson, R., Nickson, J., Tonge, R., Wilson, I., Nicholson, J., (2006). Systems toxicology: integrated genomic, proteomic and metabonomic analysis of methapyrilene induced hepatotoxicity in the rat. *Journal of Proteome Research*, **5**, 1586–1601.

Dole M., Mack L.L., Hines R.L., (1968). Molecular beams of macroions. *Journal of Chemical Physics*, **49**, 2240-2247.

Ehrenberg, L., Hussain, S., (1981). Genetic Toxicity of Some Important Epoxides. *Mutation Research*, **86**, 1–113.

Elkadi, M., Pillay, A., Manuel, J., Khan, M.Z., Stephen, S., Molki, A., (2014). Sustainability Study on Heavy Metal Uptake in Neem Biodiesel Using Selective Catalytic Preparation and Hyphenated Mass Spectrometry. *Sustainability*, **6**, 2413-2423.

Eriksson, L., Antti, H., Gottfries, J., Holmes, E., Johansson, E., Lindgren, F., Long, I., Lundstedt, T., Trygg, J., and Wold, S., (2004). Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Analytical and Bioanalytical Chemistry* **380**, 419-429.

Eriksson, L., Johansson, E., Antti, H., and Holmes, E. (2005). Multi- and megavariable analysis: Finding and data using regularities in metabonomics data. In: *Metabonomics in Toxicity Assessment*, Taylor and Francis Group Press, Boca Raton, pp 263-336.

Ernst, R.R., Bodenhausen, G., Wokaun, A., (1987). Principles of Nuclear Magnetic Resonance in One and Two Dimensions. *Clarendon Press*, Oxford.

Fanali, S., Haddad, P. R., Poole, C., Schoenmakers, P., Lloyd, D.K., (2013). Liquid Chromatography: Fundamentals and Instrumentation, *Elsevier press*, first edition.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science*, **246**, 64-71.

Fountain, K.J., Neue, U.D., Grumbach, E.S., Diehl D.M., (2009). Effects of extracolumn band spreading, liquid chromatography system operating pressure, and column temperature on the performance of sub-2- μ m porous particles. *Journal of Chromatography A*, **1216**, 5979-5988.

Gibson, G.G., Skett, P., (1994). Introduction to Drug Metabolism, *Blackie Academic and Professional Press*, Second Edition.

Glatt, H., Meini, W., (2004). Pharmacogenetics of soluble sulfotransferases (SULTs). *Naunyn-Schmiedeberg's Archives of Pharmacology*, **369**, 55–68.

Gritti, F., Guiochon, G., (2010). Extra column band broadening contributions of modern, very high pressure liquid chromatographs using 2.1 mm ID columns packed with sub 2 μ m particles. *Journal of Chromatography, Science A*, **49**, 7677-7689.

Gritti, F., Guiochon, G., (2013). The van Deemter equation: Assumptions, limits, and adjustment to modern high performance liquid chromatography. *Journal of Chromatography, Science A*, **1302**, 1-13.

Guengerich, F.P., Macdonald, T.L., (1984). Chemical mechanisms of catalysis by cytochromes P-450: A unified view. *Accounts of Chemical Research*, **17**, 9–16.

Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., (2007). Cytochrome P450 systems – biological variations of electron transport chains. *Biochimica et Biophysica Acta*, **1770**, 330–344.

Hong, P., Koza, S., Bouvier, E.S.P., (2012). Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *Journal of Liquid Chromatography & Related Technologies*, **35**, 2923–2950.

Horvath, C., and Melander, W., (1977). Liquid Chromatography with Hydrocarbonaceous Bonded Phases; Theory and Practice of Reversed Phase Chromatography. *Journal of Chromatographic Science*, **15**, 393-404.

Hutter, D., Greene, J.J., (2000). Influence of the cellular redox state on NF-kappaB-regulated gene expression. *Journal of Cell Physiology*, **183**, 45-52.

Ilett, K.F., Tee, L.B., Reeves, P.T., Minchin, R.F., (1990). Metabolism of drugs and other xenobiotics in the gut lumen and wall. *Pharmacology Therapies*, **46**, 67-93.

Ingelman-Sundberg, M., Oscarson, M., McLellan, R.A., (1999). Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *Trends in Pharmacology Science*, **20**, 342-348.

Ishikawa, T., (1992). The ATP-dependent glutathione S-conjugate export pump. *Trends Biochemistry Science*, **17**, 463–468.

Kaddurah-Daouk, R., Kristal, B.S., Weinshilboum, R.M., (2008). Metabolomics: a global biochemical approach to drug response and disease. *Annual Reviews on Pharmacology and Toxicology*, **48**, 653–683.

Kiang, T.K., Ensom, M.H., Chang, T.K., (2005). UDP–glucuronosyltransferases and clinical drug–drug interactions. *Pharmacology and Therapeutics*, **106**, 97–132.

Klaassen, C.D., Boles, J.W., (1997). Sulfation and sulfotransferases 5: the importance of 3'-phosphoadenosine 5'-phosphosulfate (PAPS) in the regulation of sulfation, *FASEB. Journal*, **11**, 404-418.

Li, Y., Shrestha, B., Vertes, A., (2008). Atmospheric pressure infrared MALDI imaging mass spectrometry for plant metabolomics. *Analytical Chemistry*, **80**, 407–420.

Light, A., Acher, R., du Vigneaud, V., (1957). Ion exchange chromatography of purified posterior pituitary preparations. *Journal of Biological Chemistry*, **228**, 633-641.

Lindon, J.C., Nicholson, J.K., Holmes, E., Antti, H., Bollard, M.E., Keun, H., Beckonert, O., Ebbels, T.M., Reilly, M.D., Robertson, D., Stevens, G.J., Luke, P., Breau, A.P., Cantor, G.H., Bible, R.H., Niederhauser, U., Senn, H., Schlotterbeck, G., Sidemann, U.G., Laursen, S.M.,

Tymiak, A., Car, B.D., Lehman-McKeeman, L., Colet, J.M., Loukaci, A., Thomas, C.. Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. *Toxicology and Applied Pharmacology*, **187**, 137-146.

Lindon, J.C., Keun, H.C., Ebbels, T.M., Pearce, J.M., Holmes, E., Nicholson, J.K., (2005). The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. *Pharmacogenomics*, **6**, 691–699.

March, R.E., (1997). An Introduction to Quadrupole Ion Trap Mass Spectrometry. *Journal of Mass Spectrometry*, **32**, 351-369.

Martis, E., Ahire, D., Singh, R., (2011). Metabolomics in drug discovery: a review. *International Journal of Pharmacology*, **1**, 67–74.

McCalley D.V., (2010). Study of the selectivity, retention mechanisms and performance of alternative silica-based stationary phases for separation of ionised solutes in hydrophilic interaction chromatography. *Journal of Chromatography A*, **1217**, 3408-3417.

Meynert, E.W., Foreman, R.L., Watabe, T., (1970). Epoxides as Obligatory Intermediates in the Metabolism of Olefins to Glycols. *The Journal of Biological Chemistry*, **245**, 5234-5238.

Nebert, D.W., Adesnik, M., Coon, M.J., Estabrook, R.W., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., Kemper, B., Levin, W., (1987). The P450 gene superfamily: recommended nomenclature. *DNA*, **6**, 1-11.

Nicholson, J.K., Lindon, J.C., Holmes, E., (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, **29**, 1181–1189.

Nicholson, J.K., Connelly, J., Lindon, J.C., Holmes, E., (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nature, Reviews on Drug Discovery*, **1**, 153–161.

Oehme, F., Peters, J.. (2010). Mixed-mode chromatography in downstream process development. *BioPharm International*, 12–19

Pröfrock, D., Prange, A., (2012). Inductively coupled plasma-mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends. *Applied Spectroscopy*, **66**, 843-868.

Reid, D. G. (1997). Protein NMR Techniques. Methods in Molecular Biology. *Humana Press*, Totowa, New Jersey. 419.

Rendic, S., Di Carlo, F.J., (1997). Human Cytochrome P450 enzymes: a status report summarizing their reactions. Substrates, inducers and inhibitors. *Drug Metabolism Reviews*, **29**, 413-580.

Robertson, K.D., (2001). DNA methylation, methyltransferases, and cancer. *Oncogene*, **28**, 3139-3155.

Roninson, I.B., Chin, J.E., Choi, K.G., Gros, P., Housman, D.E., Fojo, A., Shen, G.W., Gottesman, M.M., Pastan, I., (1986). Isolation of human *mdr* DNA sequences amplified in multidrug-resistant KB carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, **83**, 4538–4542.

Schnackenberg, L.K., Sun, J., Beger, R.D., (2011). Metabolomics in Systems Toxicology: Towards Personalized Medicine. General, Applied and Systems Toxicology, *John Wiley & Sons*, Ltd.

Serkova, N.J., Van Rheen, Z., Tobias, M., Pitzer, J.E., Wilkinson, J.E., Stringer, K.A., (2008). Utility of magnetic resonance imaging and nuclear magnetic resonance-based metabolomics for quantification of inflammatory lung injury. *American Journal of Physiology, Lung Cellular Molecular Physiology*, **295**, 152–161.

Sim, E., Walters, K., Boukouvala, S., (2008). Arylamine N-acetyltransferases: from structure to function. *Drug Metabolism Reviews*, **40**, 479-510.

Skoog, D.A., Holler, F.J., Crouch, S.R., (2007). Principles of Instrumental Analysis. *Thomson Higher Education*, first edition.

Smith, D.A., Ackland, M.J., Jones, B.C., (1997). Properties of cytochrome P450 isoenzymes and their substrates part 1: active site characteristics. *DDT*, **2**, 406-414.

Spagou, K., Wilson, I.D., Masson, P., Theodoridis, G., Raikos, N., Coen, M., Holmes, E., Lindon, J.C., Plumb, R.S., Nicholson, J.K., Want, E.J., (2011). HILIC-UPLC-MS for exploratory urinary metabolic profiling in toxicological studies. *Analytical chemistry*, **83**, 382-390.

Straube, E.A., Dekant, W., Völkel, W., (2004). Comparison of electrospray ionization, atmospheric pressure chemical ionization, and atmospheric pressure photoionization for the analysis of dinitropyrene and aminonitropyrene LC-MS/MS. *Journal of the American Society for Mass Spectrometry*, **15**, 1853-1862.

Sugiura, Y., Setou, M., (2010). Imaging mass spectrometry for visualization of drug and endogenous metabolite distribution: toward in situ pharmacometabolomes. *Journal of Neuroimmune Pharmacology*, **5**, 31-43.

Trygg, J., Holmes, E., and Lundstedt, T. (2007). Chemometrics in metabonomics. *Journal of Proteome Research* **6**, 469-479.

Tswett, M., (1906). Physikalisch-Chemische Studien über das Chlorophyll. Die Adsorption. *Berichte der Deutschen botanischen Gesellschaft*, **24**, 316–326.

Thomson B.A., and Iribarne J.V., (1979). Field- induced evaporation from liquid surfaces at atmospheric pressure. *Journal of Chemical Physics*, **71**, 4451-4463.

Vessey, D.A., Kelley, M., (2001). Characterization of the reaction mechanism for the XL-1 form of bovine liver xenobiotic/medium-chain fatty acid:CoA ligase. *Biochemistry Journal*, **357**, 283-288.

Wells, P.G., Mackenzie, P.I., Chowdhury, J.R., Guillemette, C., Gregory, P.A., Ishii, Y., Hansen, A.J., Kessler, F.K., Kim, P.M., Chowdhury, N.R., Ritter, J.K., (2004). Glucuronidation and the UDP–glucuronosyltransferases in health and disease. *Drug Metabolism and Disposition*, **32**, 281–290.

Weyler, W., Hsu, Y.P.P., Breakefield, X.O., (1990). Biochemistry and genetics of monoamine oxidase. *Journal of Pharmacology Therapies*, **47**, 391-417.

Williams, R.T., (1959). Detoxication Mechanisms. The Metabolism and Detoxification of Drugs, Toxic Substances and Other Organic Compounds. 2nd edition. *Chapman and Hall Ltd*, London.

Wrighton, S.A., Stevens, J.C., (1992). The human hepatic cytochrome P450 involved in drug metabolism, *Critical Reviews in Toxicology*, **22**, 1-21.

Wrighton, S.A., Schuetz, E.G., Thummel, K.E., Shen, D.D., Korzekwa, K.R., Watkins, P.B., (2000). The human CYP3A subfamily practical considerations. *Drug Metabolism Reviews*, **32**, 339-361.

Zimniak, P., Awasthi, S., Awasthi, Y.C., (1993). Phase III detoxification system. *Trends in Biochemistry Science*, **18**, 164–166.

Computational technologies as a tool for
the investigation of metabolism

Chapter Two

The utility of computational approaches for metabolism studies has long since been discussed. In this chapter the utility of such methods are evaluated in both a small scale metabolism study but also in a large scale toxicity study. Furthermore the inherent limitations of the LC-MS platform are explored, with statistical techniques for the correction of batch effect applied.

2.1 Metabonomics as a tool for the investigation of xenobiotic metabolism

The introduction (chapter one) has already touched upon the utility of metabonomics in the investigation of both the metabolism of xenobiotics, but also xenobiotic induced toxicity.

All living organisms are continuously exposed to a variety of exogenous chemicals, through diet, environmental exposure and drug administration. These compounds are actively dealt with by the biological system in the ADME process (absorption, distribution, metabolism and excretion). While there is certainly biological machinery dedicated to the elimination of foreign chemicals, the metabolic fate of the xenobiotic is not as easy to predict as one might initially think. Endogenous metabolism has evolved in a manner that has produced tightly regulated and constrained metabolic pathways. Enzymes are highly specific and the probability of the correct reaction occurring is maximised to ensure maximal chemical efficiency.

This is not the case with metabolism of xenobiotics where it is not possible for cellular biology to have adapted specific metabolic protocols for every possible compound in existence. In this case the foreign compound is altered in a manner purely reliant on the statistical probability that it will come into contact with an enzyme that will be able to interact with. As the enzymes which are able to catabolise xenobiotic reactions tend to have broader specificities than endogenous metabolism, both the type and order of metabolic alteration a xenobiotic may encounter is not determined absolutely (Wilson and Nicholson, 2003). It is therefore apparent that specific combinations of genetic variation and environmental exposures may result in widely different toxic outcomes. Furthermore, the xenobiotic, or metabolites of the xenobiotic, have the potential to alter the innate metabolic processes, and this in itself can be a possible origin of the toxic effects of certain compounds (Klaassen, *et al.*, 2002).

While the metabolism of xenobiotics has been long since studied (Williams, 1947) traditional metabolism techniques are generally hypothesis driven and require specific target

compounds (Lenz and Wilson, 2007). These approaches are useful, but with metabolic outcomes being so hard to predict, there are quite obvious limitations, and can result in incorrect conclusions being reached, which in the realm of drug development can have serious consequences. It is apparent therefore that it is necessary for a hypothesis free approach for the investigation of the effects of exogenous compounds on the metabolic equilibrium. Metabonomics, based on high throughput and robust analytical technologies, is able to detect thousands of metabolites simultaneously within any given biological sample and therefore take a 'snapshot' of the metabolic status present.

By using multivariate statistical comparisons between large datasets acquired from analysis of the samples it is possible to detect subtle changes and identify both markers of xenobiotic exposure and also changes within the endogenous metabolism of the organism simultaneously. Through statistical modelling it is possible to position every sample within a theoretical hyperspace. By constructing a model containing the targets of comparison it is possible to consider the relative positions of each subject in relation to each other. By comparing the statuses of the various individuals it is possible to infer directions of influence, and also identify which variables contribute to this influence. Given enough statistical power it is therefore possible to 'blindly' identify which are the most influential metabolites, on a purely statistical level. Further targeted work is possible to then identify these metabolites, and also pave the way for the elucidation of metabolic pathways and mechanisms of toxicity. Metabonomics is therefore potentially a valuable tool in xenobiotic metabolism studies.

Two examples of the use of a statistical approach are described here. The first uses a small "traditional" drug metabolism protocol, with a limited number of animals and a pharmacological dose of the drug diclofenac (Section 2A). The second adopts a toxicological study design with larger doses and numbers of animals to look at the fate of the drug acetaminophen (paracetamol) (Section 2B).

Computational technologies as a tool for
the investigation of the metabolism of
diclofenac in the HRN mouse

Chapter Two [A]

Previous work by Sarda et al. had utilized HPLC-MS to investigate the metabolism of Diclofenac in the mouse. The major routes of metabolism were thought to be glucuronic acid and taurine conjugation. This study was repeated here in the HRN and wild type mouse, with UPLC-MS metabolic profiling. Multivariate statistical analysis was used to identify the metabolic differences. Computational analysis failed to identify the previously known metabolites of diclofenac, as diurnal variation predominated the statistics.

SUMMARY:

- Diclofenac, a non-steroidal anti-inflammatory drug, is a known hepatotoxin.
- Diclofenac was dosed orally by gavage to three male HRN null and three male wild type mice (strain: C57BL/6J). Urine was collected at regular intervals, and after two days the animals were sacrificed and livers retrieved.
- Urine samples were analysed by LC-MS (UPLC-MS) to determine both the effects of diclofenac on urinary endogenous metabolite profile and the metabolic fate of the compound itself.
- Chemometric analysis showed that the main discriminators between the pre dose and 24 hour time points were unassigned analytes 1 (m/z 203.14, 7.93RT) and 2 (m/z 314.02, 7.65RT). These did not match the previously identified metabolites of diclofenac.
- Manual interrogation of the spectrometry data showed that both diclofenac and major metabolites were present in the urine samples.
- The conclusion from this investigation was that classic manual investigation of the urinary profiles showed presence of diclofenac metabolites. However, the combination of low dose, small sample size and low statistical power rendered the computational approach unable to identify these same metabolites.

2A.1 Introduction

2A.1.1 A Metabolism study in Mice following administration of Diclofenac

Diclofenac (2-[2-(2,6-dichlorophenylamino)phenyl] acetic acid), molecular weight 294.01, the chemical structure of which is shown below in figure 2A.1, is a non-steroidal anti-inflammatory drug (NSAID), widely used to reduce inflammation or for pain relief in certain conditions. It is prescribed for the treatment of many inflammatory conditions including rheumatoid arthritis, osteoarthritis and acute muscle pain, but is also used for the pain management of severe migraines and menstrual cramps (Small, 1989). Diclofenac is generally administered as a sodium salt, although in some countries a potassium salt is used.

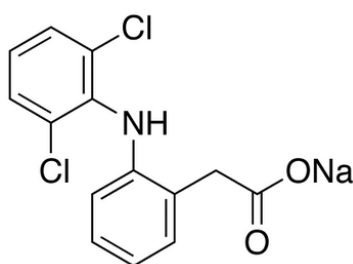


Figure 2A.1 The molecular structure of diclofenac as a sodium salt.

Diclofenac has been associated with a wide range of adverse drug reactions, some of which include increased risk of heart disease, serious gastrointestinal events, and renal toxicity characterised by increased serum creatinine (Novartis report, 2014). Administration of diclofenac has also been linked to rare, but severe, hepatotoxicity in around 4% of patients (Boelsterli *et al.*, 1995). Curiously as these toxic episodes typical present with delayed onset of symptoms, and a lack of a clear relationship between dose and severity, they have been defined as idiosyncratic toxicity rather than individual susceptibility (Banks *et al.*, 1995).

Due to these characteristics diclofenac has been the subject of intense scientific interest, and many studies have attempted to elucidate the mechanisms underlying the toxicity. Studies of diclofenac toxicity have been performed both *in vivo* and *in vitro*, with studies in animal models (Kallio *et al.*, 2010) and in human patients (Poon *et al.*, 2001). This research effort has provided some insight into the metabolism that diclofenac is subject to. Highly reactive quinone-imine metabolites of diclofenac are thought to be produced by the microsomal mixed function oxidase mediated oxidation of initial hydroxylation products (Tang *et al.*, 1999). Ester glucuronides of diclofenac itself have been detected as have both ether and ester glucuronide conjugates of the drugs downstream hydroxy-metabolites. Generally ether glucuronides have been shown to be quite inert but acyl glucuronides, on the other hand, are chemically reactive and have been proposed as a key component in many toxic events (Faed, 1984). Diclofenac acyl glucuronides have been widely observed in studies in animals and humans (Kretz-Rommel and Boelsterli, 1994). In addition glutathione conjugates of diclofenac have also been observed in rat bile (Waldon *et al.*, 2010). It is generally accepted that the presence of glutathione conjugates indicates the presence of reactive intermediates of the parent compound (Sims and Grover, 1974) as one of the primary protective mechanisms of the cell against such chemicals is inactivation through the action of glutathione S-transferase. These conjugates are then generally excreted in the urine as the mecapturate (Wood, 1970); indeed Poon *et al.* (2001) identified the mecapturic acid metabolite of diclofenac in both human and rodent urine.

2A.1.2 The hepatic reductase null mouse model

The potential toxic metabolites of diclofenac includes both the quinone-imine and acyl glucuronides however, as these will be produced at the same time in normal mice it is hard to pin point which, if not both, are responsible for the observed toxicity or, if both are toxic, to what degree each contributes. Traditional metabolism experiments to try to deconvolute this

problem would involve the use of chemical inhibitors of the enzymes responsible for the production of these two groups of metabolites, and then look for their effects in modifying toxicity. However, while this is generally quite easy *in vitro* or in synthetic systems, such as those used in the above mentioned work by Kretz-Rommel and Boelsterli (1994), this approach is harder *in vivo* as these inhibitors can have unpredictable effects on the organism under study, which can be clinically significant (Lynch and Price, 2007).

Within biomedical research the use of genetically modified animals for the investigation of genes which have been sequenced but whose function is unknown has been extensive. These 'knock out' animals have the gene of interest either replaced or deactivated through insertion of foreign DNA. This technology is highly applicable to studies focusing on the contribution a particular enzyme to the metabolism of a xenobiotic. Mice that have a particular P450 isoenzyme deactivated have been used to identify the role that particular CYP has in the toxicity of various compounds (Pineau *et al.*, 1995). More recently Le and Sauer (2001) described a knock out mouse model, created using the Cre/LoxP system, which has only the hepatic P450 enzymes silenced, called the Hepatic Reductase Null (HRN) mouse model. Cre recombinase, the enzyme upon which the knock out is dependent, catalyses recombination at very specific 34-base pair sequence (LoxP). Furthermore this gene ablation can be tightly controlled and targeted to certain cell types and can therefore be used to create such conditional knock outs (Gu *et al.*, 1994). The HRN mouse, whilst have no activity with respect to liver P450's, will contain maintain normal functional P450s in all other tissues, thereby avoiding the embryonic lethality that an animal with no P450 would suffer. The HRN mouse is therefore the ideal animal model for the investigation of the role of diclofenac acyl glucuronide vs oxidative metabolism in liver toxicity, as it is thought the majority of its metabolism occurs within the hepatocytes. If coupled with metabonomic global metabolic profiling, the comparison of the HRN metabolic fingerprint with wild type animals could potentially be used to identify which non-P450-derived diclofenac metabolites are produced by hepatic metabolism.

2A.1.3 Traditional vs Statistical drug metabolic investigation

Previous work by Sarda *et al.*, (2011) in the mouse, has already identified the key urinary metabolites of diclofenac using radiolabelled diclofenac and radio-HPLC and MS detection using an ion trap mass spectrometer. This work identified that the main routes of metabolism in the mouse involved direct taurine conjugation, or 4 and 5 hydroxylation followed by taurine or glucose conjugation. Furthermore the authors reported on how the main route of conjugation was via formation of the ether, not the acyl, glucuronide. This was somewhat different to previous results, as there no mecapturic acid or quinone-imine derivatives of diclofenac were detected, although most of previous the work in rodents had been performed in the rat.. Other minor metabolites of diclofenac were also described including sulfate, glucose and ribose conjugates of the Phase I diclofenac metabolites.

This previous work utilised the skills of experienced spectrometrists, combined with radiotracer experiments to detect the diclofenac metabolites amongst the mass of endogenous compounds. This process is both time intensive and demands highly skilled individuals for good results. Furthermore without multivariate statistical comparisons between the groups this approach may fail to identify other subtle effects on endogenous metabolism. The aim of this initial pilot study described below was to ascertain whether it is possible to perform an investigation into the metabolism of diclofenac utilising purely computational based methods. As this was a 'worked example', where the metabolism has already been described using conventional methods, the aim was to see if mining the data using multivariate statistical methods could derive the same results. Global metabolic profiling was performed on the samples using UPLC-MS and then chemometric techniques were employed to discern what differences were present between the metabolic landscapes of the HRN and wild type mice.

2A.2 Materials and Methods

2A.2.1 Study Design and Sample Collection

The analytical chemistry in this study was performed on urine samples previously collected in a diclofenac metabolism study (courtesy of Dr H Lees) at AstraZeneca (Alderley Park). Diclofenac was dosed orally by gavage to three male HRN null and three male wild type mice (strain: C57BL/6J). Male wild type mice aged between 8 and 10 weeks were supplied by the AstraZeneca Rodent Breeding Unit, Alderley Park. The HRN (strain: B6.Cg-Port^{m1Wolff}Tg(Alb-cre)21Mgn N7) mice aged between 7 and 8 weeks were purchased from Taconic (Taconic Farms Inc., Germantown, NY, USA). Mice were treated with 10mg/kg of diclofenac as the sodium salt, with a total dose volume of 5mL/kg. Mice were housed in metabolic cages according to strain and urine was collected for metabonomic analysis. Spot urine was collected at 16 hours pre dose and then total urines were collected for the time periods of 0-12 hours and 12-24 hours. Clinical observations and body weight were recorded daily. The animals were sacrificed 24 hours after dosing by fluothane inhalation. Kidneys and livers were removed and were examined histopathologically. Samples were stored at -80 °C prior to analysis.

2A.2.2 Sample Preparation and Experimental

Prior to analysis the sample order was randomised. Of the 18 samples, one sample (animal three HRN, 0-12 hours) had insufficient volume and was therefore omitted from the study. A 40 µL aliquot of urine was mixed with 40 µL water to dilute the salt concentration for each sample. These samples were vortexed and left at -20 °C overnight before centrifugation for 5 min at 12 000 G at 4 °C. For reversed-phase chromatography, 50 µL of sample was taken and added to 150 µL water in 350 µL 96-well plates (Waters Corporation, Milford, USA). A pooled quality control (QC) sample was prepared by combining 10 µL of each sample and diluting 1:4 with water. Five injections of the pooled QC sample were used to condition the

column with a QC sample injected prior to analysis of the individual urines and then after every 10 injections. Each sample was injected twice.

Reversed-phase liquid chromatography was performed on an Acquity UPLC system (Waters Ltd, Elstree, UK) coupled to a Xevo G2 triple quadrupole time-of-flight mass spectrometer (Waters MS technologies, Ltd., Manchester). UPLC-MS conditions for urine analysis were those previously described for the reversed-phase analysis of urine (Want *et al.*, 2010) with the separation was performed using a Acquity BEH C18 (1.7 μ m, 2.1 x 100 mm) column and 2 μ L per sample were injected. Mass spectrometry was performed with electrospray ionisation operated in positive (ESI+) ion mode. The instrument was operated in sensitivity mode and set to acquire data over the m/z range 50 – 1200 with a scan time of 0.1 s. All mass spectral data were collected in centroid mode using the MS^e data acquisition function. For mass accuracy, leucine enkephalin (MW = 555.62) was used as a lock mass at a concentration of 200 pg/ μ L (in 50/50 CH₃CN:H₂O, 0.1 % formic acid) infused at a flow rate of 20 μ L/min via a lock spray interface. The data were collected using MassLynx V 4.1 software (Waters Corporation, Milford, USA).

2A.2.3 UPLC-MS Data analysis and statistical modelling

The UPLC-MS data set was pre-processed using XCMS Bentobox (Tautenhahn *et al.* 2012) software in order to convert the three-dimensional LC-MS raw data into time-aligned detected features of retention time, m/z ratio and intensity (peak area). The ion intensities for each peak detected were then normalised, within each sample, to the median peak intensity in that sample. The resulting peak marker tables, comprising m/z , RT and normalised intensity values for each of the 1732 variables in every sample, were exported to SIMCA-P+ 11.0.4 software (Umetrics, Umea, Sweden) to perform multi-variate statistical analysis.

Prior to any statistical modelling the QCs were evaluated, through the use of positioning within the PCA plots. If there was significant variation within the quality control data, or if there was significant drift within the position of the QCs within PCA scores plots the data

would have been unfit for analysis. SIMCA-P+ 11.0.4 was used for the statistical modelling; Principal component analysis (PCA) was utilised to explore any underlying stratification within the data. OPLS S-plots were used to investigate any differences within the sample groups and the loadings were used to identify which (if any) variables were driving the separation.

2A.3 Results and Discussion

2A.3.1 Quality control

As previously discussed LC-MS techniques are subject to instrumental instability resulting in possibly unreliable data. For quality assurance purposes aliquots from every sample were combined in a pooled sample which was then analysed at regular intervals throughout the experiment. Initial PCA scores plots used to scrutinise analytical drift showed that the QCs together with the conditioning injections grouped well together in the centre of statistical space, as can be seen in figure 2A.2 below. As this experiment contained very few injections (44 including both the duplicates and the QCs), it was not expected that there would be any major analytical inconsistencies due to excessive drift during the run due to column/detector contamination, and this seems to be borne out in the results.

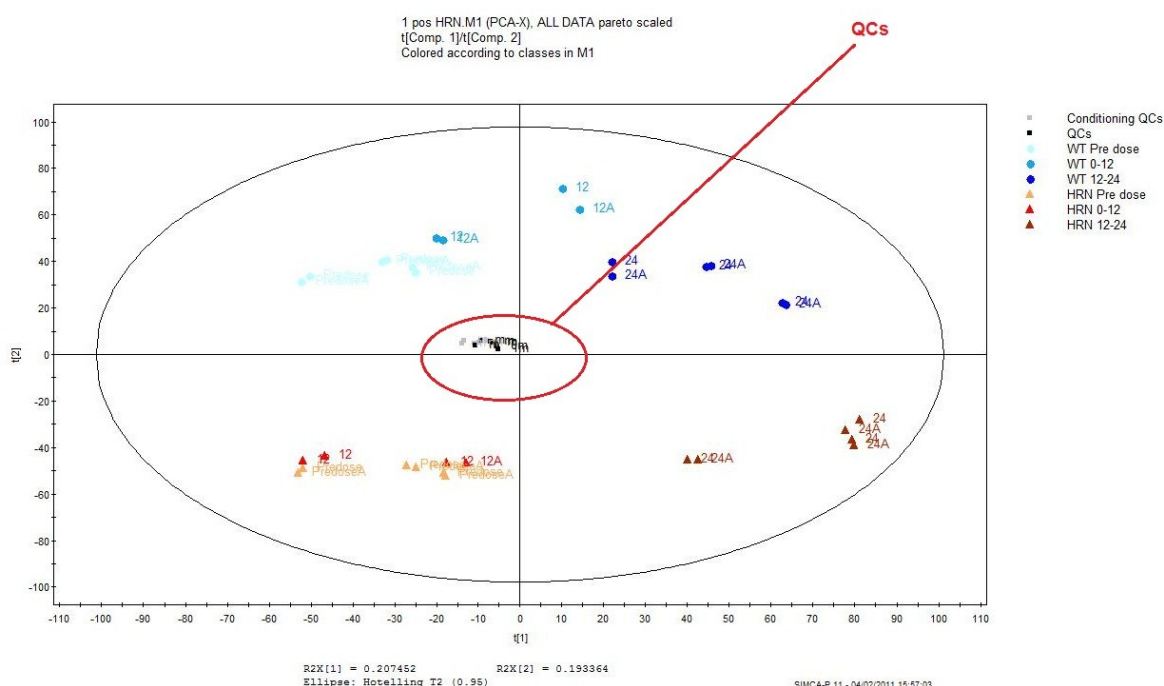


Figure 2A.2 A PCA scores plot of unfiltered, pareto scaled data. The red highlighted area indicates where the pooled sample QCs have clustered together in the centre of statistical space, indicating that there was good analytical stability throughout the run.

2A.3.2 Chemometric analysis

As the aim of this pilot study was to ascertain the suitability of computational methods as a means of detecting drug metabolites in metabolic studies, a 'blind' hypothesis free approach was utilised in the statistical analysis. Initial PCA modelling of the complete dataset revealed significant differences between the metabolic profiles of the wild type and HRN mice. It is therefore apparent that there are significant metabolic differences in the two strains, both initially and after dosing, as knocking out the hepatic p450 would undoubtedly have significant effects. Figure 2A.3 illustrates the separation between wild type and HRN mouse and in addition shows a time-dependant progression within each strain. Pre-dose and 0-12 hour samples from the HRN mice seem to be metabolically very similar, whereas there is significant deviation from pre-dose metabolotype of the wild type animals over the same time period. Maximal differences in metabolic phenotypes for the HRN mouse were found

between 12 and 24 hours. In both strains it seems that the maximal response occurred within these first 24 hours. This matches previous results from Sarda *et al.*, (2012) as they identified that the majority (49.7%) of excretion of drug-related material occurred in this time period.

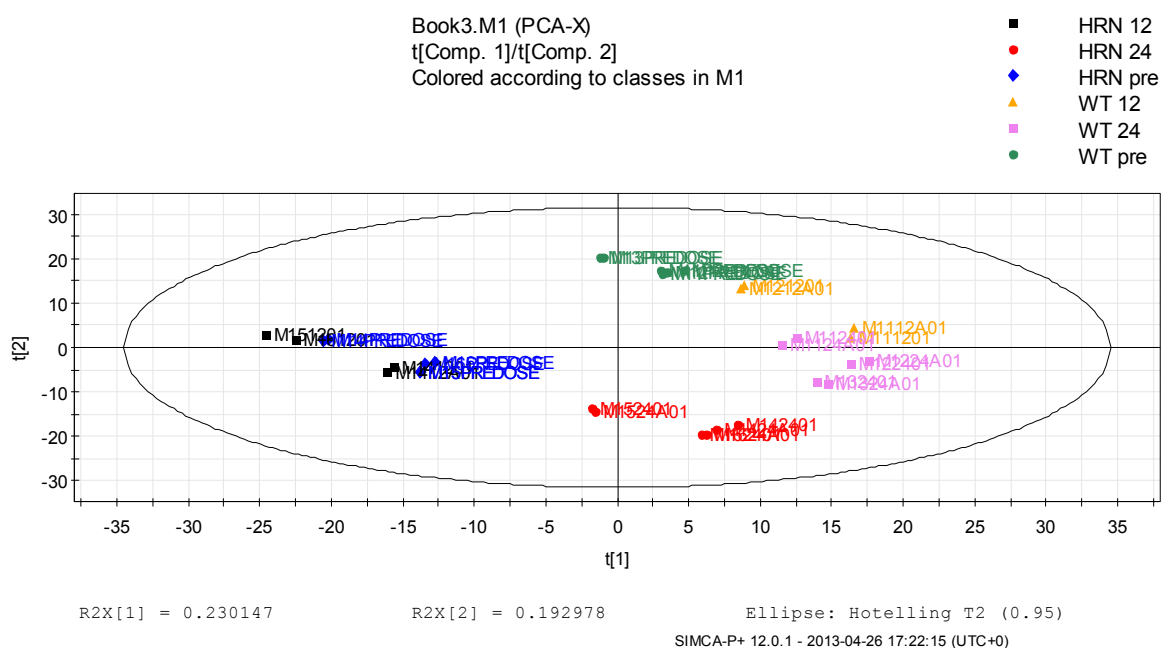


Figure 2A.3 A PCA scores plot of the complete dataset, coloured according to strain and time point.

For the identification of the key metabolic differences after administration of diclofenac, the pre-dose wild type samples were compared to the wild type samples at 24 hours post dose. PCA modelling showed a clear difference between the metabolic states of the mice. In this comparison the pre-dose samples act as a control group. Figure 2A.4 shows the PCA scores

plot for this comparison. To identify the key metabolites driving this separation S-plots were constructed. The S-plot highlights which variables are contributing the most to the differences between the groups (figure 2A.5). The signals identified by the scores plot are detailed in table 2A.1.

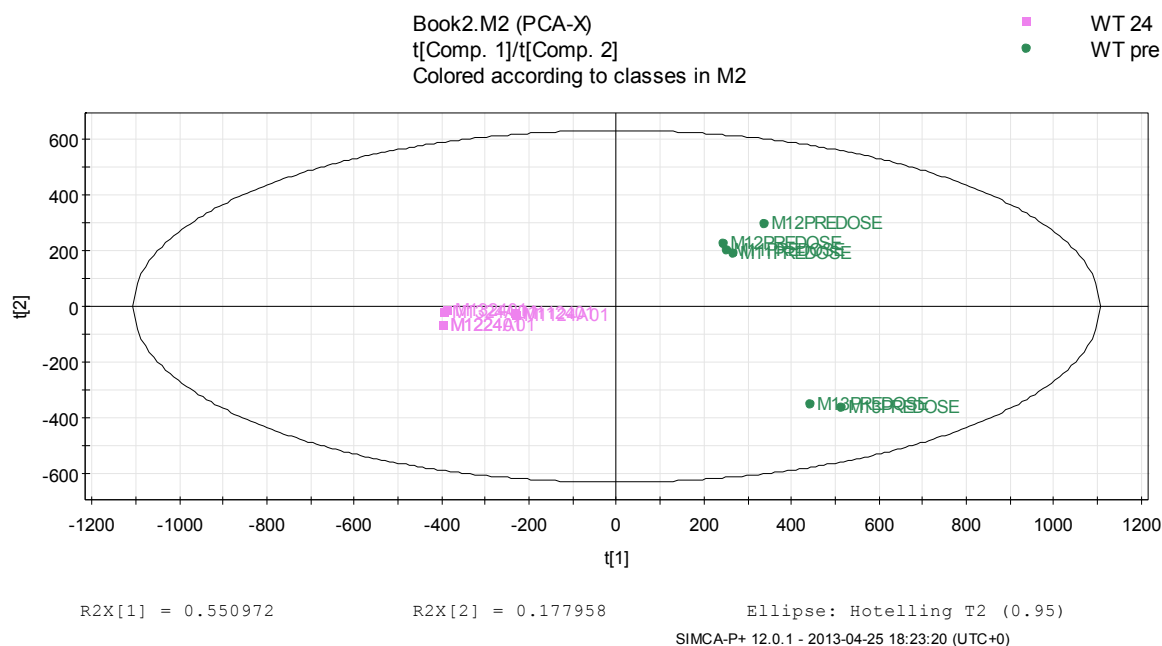


Figure 2A.4 A PCA scores plot comparing the pre dose and 12-24 hour samples of the wild type mice. There are clear differences in the metabolic phenotypes of the mice with time after diclofenac dosing, the influence of which is described along the first principal component.

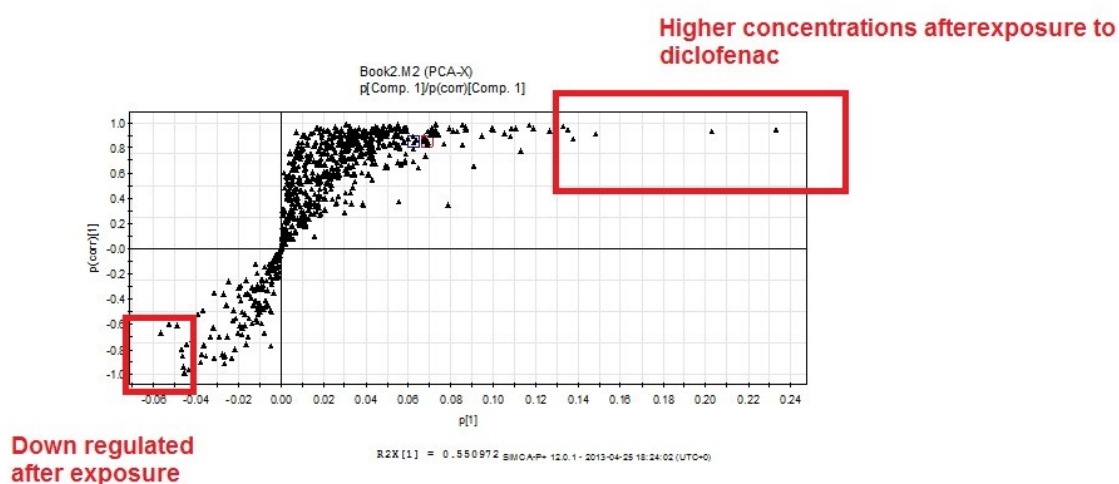


Figure 2A.4 S-plot of the PCA model comparing the pre dose and 12-24 hour samples of the wild type mouse. The highlighted points at the top represent the metabolites which are up regulated after diclofenac dosing, while the ones at the bottom left are down regulated.

	Accurate Mass	Retention Time(min)	Up/Down regulated
Unassigned 1	203.14	7.93	Up
Unassigned 2	203.11	6.65	Up
Unassigned 3	314.02	7.65	Up
Unassigned 4	132.08	0.71	Up
Unassigned 5	297.16	3.50	down

Table 2A.1 Description of the major metabolic changes in the wild type mouse after diclofenac administration

The identified ions described in the table above do not match either the parent compound (diclofenac has a molecular mass of 294.91) or indeed any of the previously identified major metabolites, nor any of the product ions or fragments of those metabolites. In addition, none of these compounds displayed the diagnostic isotope pattern that would be expected for a molecule containing 2 chlorine atoms (see Sarda *et al*, 2011). This suggests that the metabolites detected are not from diclofenac but are of endogenous origin. As such they might reflect the pharmacological effects of the drug on metabolic processes in the animals or simply factors such as diurnal variation. As the aim of this pilot study was to determine how well an unbiased statistical approach to data mining for detecting xenobiotic metabolites

could replicate results found by a human expert these endogenous markers were not investigated further.

There was however, the possibility that the failure of the analysis to detect the metabolites that had been characterised in the previous study could have been due to their not being present in these samples. To ascertain whether our analytical method was successful, standard mass spectrometry practice was used to manually search for and identify these known metabolites. As indicated above diclofenac contains two chlorine atoms, and enabling the identification of drug related products more easily through the use of the well-known chlorine-isotope pattern. The same MS data that had been analysed using statistical methodology was therefore analysed using this targeted, expert-based approach and, as indicated in table 2A.2, a large number of diclofenac metabolites were found in this fashion.

Compound	Mass	Formula	Key Ions
Diclofenac	294.0086	C ₁₄ H ₁₀ O ₂ N ₁ Cl ₂	250, 214, 178
Hydroxy	310.0033	C ₁₄ H ₁₀ O ₃ N ₁ Cl ₂	266, 230, 194, 166
Di-hydroxy	325.9983	C ₁₄ H ₁₀ O ₄ N ₁ Cl ₂	290, 282, 246, 210
Hydroxy, sulfate	389.9599	C ₁₄ H ₁₀ O ₆ N ₁ Cl ₂ S ₁	310, 266, 230, 194
Taurine	401.0123	C ₁₆ H ₁₅ O ₄ N ₂ Cl ₂ S ₁	124, 107, 80
Hydroxy, taurine	417.0070	C ₁₆ H ₁₅ O ₅ N ₂ Cl ₂ S ₁	381, 345, 317, 292, 237, 124, 80
Demethylated, Acyl glucuronide	456.0245	C ₁₉ H ₁₆ O ₈ N ₁ Cl ₂	280, 276, 244, 200, 175
1-β-O-Acyl glucuronide	470.0040	C ₂₀ H ₁₈ O ₈ N ₁ Cl ₂	193
Hydroxy, glucose	472.0558	C ₂₀ H ₂₀ O ₈ N ₁ Cl ₂	428, 266, 230, 161
Hydroxy, glucuronide	486.0350	C ₂₀ H ₁₈ O ₉ N ₁ Cl ₂	310, 266, 193, 175, 131, 113

Table 2A.2 Summary of the metabolites of diclofenac which were identified using conventional mass spectrometry methods.

2A.4 Conclusion and future perspectives

The discovery of the known metabolites within the LC-MS data signified that the statistical approach lacked the sensitivity to identify drug related compounds within our data set. This could be attributed either to the small sample size, or to the low dose administered. In either case the presence of diclofenac and its metabolites in the sample was masked by underlying diurnal variation and changes in endogenous metabolism. The untargeted analysis simply lacked the statistical power to unearth the presence of diclofenac and its phase I and II metabolites hidden in the endogenous metabolic profile, because these changes were subtle compared to other fluctuations. This highlights a potential limitation of the computational approach, which generally is biased towards the greatest numerical difference, irrespective of whether this difference is biologically relevant or not. For future work, it was decided that a much greater sample size and higher dosing levels could offer a more robust approach to statistical metabolic evaluation.

Application of computational analysis and batch correction normalisation to the investigation of the metabolism of acetaminophen

Chapter Two [B]

After the work described in the previous sub chapter (2A), application of computational analysis was performed on an acetaminophen toxicity study. UPLC-MS was performed on a large number of urine samples from a COMET study of chronic acetaminophen exposure, which were split into four batches. Data analysis of the individual batch datasets identified clear differences between the metabolic profiles of the different dose groups. These differences were driven by the presence of both high concentrations of parent compound and the major metabolites. These major metabolites were identified to be the glucuronide and sulphate conjugates. Attempts to merge the data from the different batches were made using median fold change normalization after variance stabilization using a glog function. However, this was largely unsuccessful as the resultant data had lost much biologically relevant information.

SUMMARY:

- Acetaminophen, also known as paracetamol, a commonly used analgesic and antipyretic drug, is a human and rodent hepatotoxin at high doses..
- 40 Sprague Dawley rats were split into four equal groups and were administered with a single IP dose of either vehicle (carboxymethylcellulose), 200 mg/kg, 400 mg/kg or 800 mg/kg of acetaminophen. Urine was collected at regular intervals for seven days resulting in a total of 680 samples.
- Urine samples were split into four batches and then analysed by LC-MS (UPLC-MS) to determine both the effects of acetaminophen on urinary endogenous metabolite profiles and the metabolic fate of the compound itself.
- Chemometric analysis of each batch showed that the main discriminators between dosed and control animals were associated with metabolites of acetaminophen.
- A number of acetaminophen metabolites were identified in urine, including with the glucuronide and the sulphate, and the detection of the N-acetylcysteiny conjugate, itself a metabolite of the glutathione conjugate of acetaminophen, evidence for the production of reactive metabolites.
- Batch correction was attempted by using median fold change normalisation after variance stabilisation with a glog function, with the QCs as a measure of normalisation. While this approach did normalise the data, when the complete data set was analysed much of the biologically interesting variation had been masked.

2B.1 Introduction

2B.1.1 Acetaminophen

Acetaminophen (APAP, N-acetyl-p-aminophenol), molecular weight 151.16, widely known in Europe as paracetamol, the structure of which is shown below in figure 2B.1, is a commonly used, over the counter analgesic and antipyretic. As a mild analgesic is it used in the treatment of headaches and other minor pains in addition to fever management however, it exhibits limited anti-inflammatory action. The mechanism of action of acetaminophen appears to involve the inhibition of prostaglandin synthesis, a group of hormone-like glycoproteins which are involved in pain perception and temperature control (Kanada *et al.*, 2013).

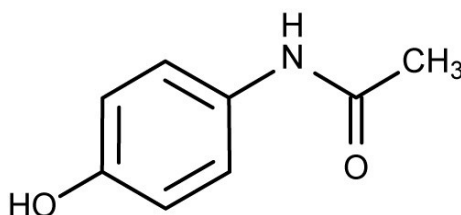


Figure 2B.1 The molecular structure of paracetamol.

Despite being readily available, acetaminophen has been linked to a wide range of side effects and toxicity; in fact acetaminophen overdose is the most common cause of acute hepatic failure in western populations (Bernal *et al.*, 2010). This situation may be in part due to unintentional overdosing due to concurrent use of multiple acetaminophen-containing medications. Nonetheless concerns over the rising incidence of mortality due to acetaminophen triggered multi-organ failure lead the Food and Drug Administration (FDA) in the USA to introduce new legislation in February 2014 limiting acetaminophen content to 325 mg per unit dose (FDA guidelines, 2011).

After oral dosing acetaminophen is rapidly absorbed through the gastrointestinal tract after which it undergoes extensive hepatic metabolism. The majority of the drug is subject to phase II metabolism, with approximately 50-70% of acetaminophen glucuronidated by UDP-glucuronosyltransferases and a further 20-30% sulphated by sulphotransferase enzymes (Prescott, 1980). These conjugates are then readily excreted into the urine. A further 10% of the dose of acetaminophen is oxidised by microsomal mixed function oxidases, primarily CYP2E1 in humans, to the reactive intermediate N-acetyl-p-benzoquinone imine (NAPQI) (Dahlin *et al.*, 1984). NAPQI is then conjugated with glutathione, either spontaneously or catalysed by glutathione-S-transferase enzymes (GST), into acetaminophen-GSH, which is then subject to further biotransformation into cysteinyl and related compounds, the acetaminophen N-acetyl-L-cysteinyl (mercapturate) and acetaminophen cysteinyl-glycine conjugates. These metabolites together with the glutathione conjugate are excreted in the bile, or excreted as the cysteinyl and or the mercapturic acid derivatives in the urine. Urine is also seen to contain the parent compound and catechol derivatives. The hepatic metabolism of acetaminophen is summarised in figure 2B.2 below.

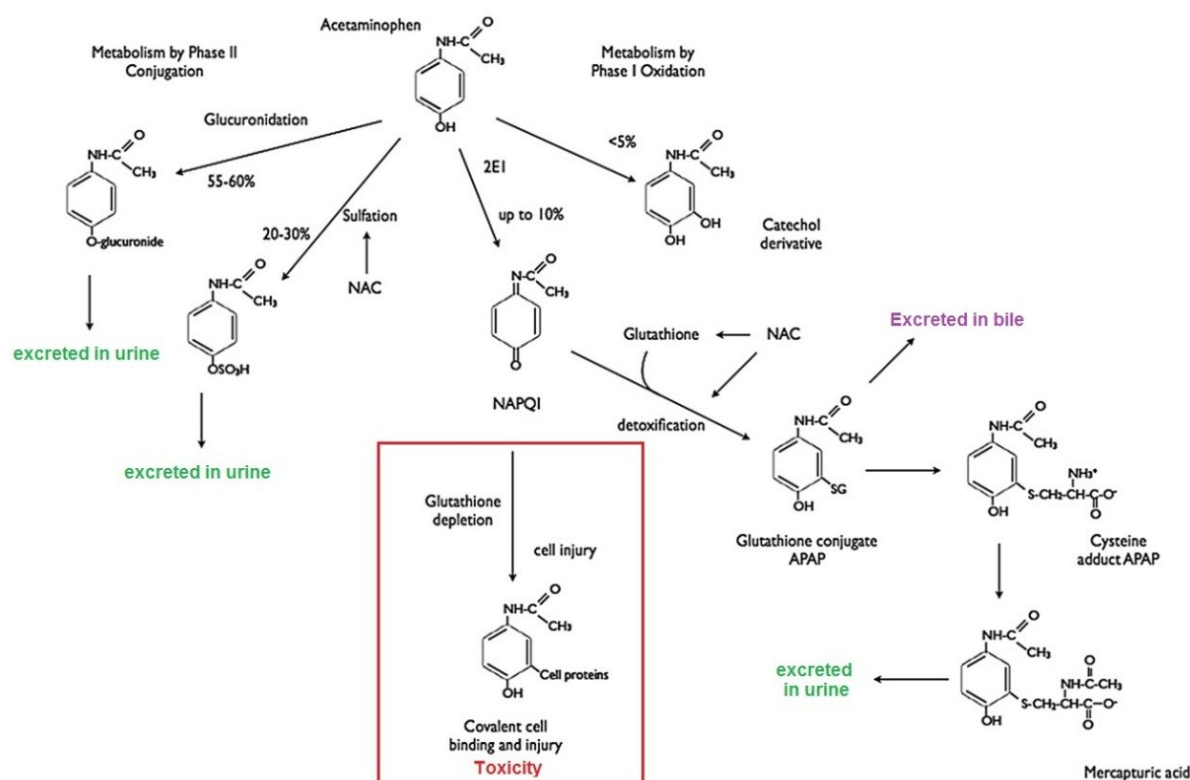


Figure 2B.2 Diagram illustrating the hepatic metabolism and fate of acetaminophen. At normal doses major metabolites APAP-glucuronide, APAP-sulphate and mercapturic acid derivatives are excreted in the urine. At toxic levels however, high levels of NAPQI result in glutathione depletion and a cascade of further hepatotoxic events.

Modified from Hodgman and Garrard (2012)

At toxic doses continued production of NAPQI results in eventual glutathione depletion. The reactive electrophilic NAPQI will then cause cell damage by protein binding (Mitchell *et al.*, 1973). This glutathione depletion and covalent binding to proteins will trigger a plethora of downstream events, the generation of reactive oxygen species oxidative stress leading to mitochondrial failure and eventual cell death (Jones *et al.*, 2010), although the precise cellular pathways by which this occurs are still under investigation.

2B.1.2 Metabonomic profiling as a tool for the investigation of acetaminophen metabolism and toxicity

As mentioned previously metabonomics comprises a suitable platform for the comprehensive investigation of the metabolism and toxicity of xenobiotics, with the added

benefit of enabling the assessment of any perturbation on endogenous metabolism. As one of the most well studied, but also most clinically relevant drugs, acetaminophen has been the subject of many metabolic profiling studies. Indeed one of the first studies utilising ^1H NMR spectroscopy for profiling urine (Bales *et al.*, 1984) succeeded in identifying acetaminophen and its major metabolites, the glucuronide, sulphate and cysteine conjugates, in human urine samples after dosing at therapeutic levels. Furthermore the authors were able to describe alterations in the excretion of the endogenous metabolites creatinine, citrate, hippurate and sarcosine, although the latter signal is now known to be dimethylamine (Coen *et al.*, 2003). Further work by the same group (Bales *et al.*, 1988) involved a pioneering toxicology study which identified markers of toxicity of acetaminophen using the same metabolic profiling approach. Evidence of the saturation of sulphaconjugation, increased production of NAPQI, in addition to altered quantities of amino acids, were seen as indicative of hepatic cell damage.

^1H NMR spectroscopy-based metabolic profiling has also been used to extensively study the metabolism of acetaminophen in many animal models. The introduction of a ^{13}C stable label into the acetyl group of the compound, coupled with NMR spectroscopy allowed the demonstration of so-called “futile de-acetylation” i.e., the cyclic deacetylation and re-actylation of the compound in the rat and human, a process which had previously been largely overlooked (Nicholls *et al.*, 1995). Animal models have also been used in metabolome-wide mapping for the comprehensive evaluation of the effect acetaminophen has on endogenous metabolism. In combination with traditional clinical chemistry, ^1H NMR spectroscopy was used to identify elevated amounts of triglyceride and monounsaturated fatty acids and a decrease in the amounts of polyunsaturated fatty acids in the mouse (Coen *et al.*, 2003). Monounsaturated fatty acids are linked to mitochondrial dysfunction while polyunsaturated fatty acids are indicative of beta-oxidation inhibition (Osmundsen and Bjørnstad, 1985), which suggests that acetaminophen may cause dysregulation of cellular energy metabolism.

Building upon these previous results, the COMET consortium undertook two large scale animal toxicology studies to study both single and chronic exposure to acetaminophen. These ^1H NMR-based studies identified further treatment associated biomarkers, with increases in the relative amounts of taurine and creatinine, with a decrease in allantoin and succinate (internal COMET report, 2003).

A small proportion of studies have employed the UPLC-MS platform for metabolic profiling of acute exposure (Sun *et al.*, 2009), screening for reactive APAP intermediates (Li *et al.*, 2011) and even utilising hyphenated combined HPLC-NMR-MS techniques for the detection of the minor APAP metabolite ether glucuronide of 3-methoxy-acetaminophen which was undetectable by direct ^1H NMR of urine (Godejohann *et al.*, 2004). However, there have been few comprehensive studies on the long term metabolic perturbations of chronic acetaminophen administration with analysis by LC-MS. As such, the aims of this study were to utilise previously collected urine samples, from the chronic COMET acetaminophen study, in an untargeted UPLC-MS metabolic profiling approach for the combined elucidation of acetaminophen metabolites and biomarkers of toxicity and treatment. A further secondary objective was evaluation of computational techniques (discussed in the previous section) in a larger cohort, with a relatively high dose.

2B.2 Materials and Methods

2B.2.1 Study details

The reanalysis performed in this study was performed on the samples collected in the COMET study S09 where urine and serum samples were collected after the administration of acetaminophen by gastric intubation, to male Sprague-Dawley rats for seven days. The rats (10 animals/group) were dosed with 200 mg/kg, 400 mg/kg or 800 mg/kg of acetaminophen in 0.2% carboxymethylcellulose, (control rats were dosed with 0.2% carboxymethylcellulose only). Rats were housed in metabolic cages and urine was collected

for metabonomic and clinical analysis. Urine samples were collected at 16 hours pre-dose, at time of dose and at 8, 24, 32, 48, 56, 72, 80, 96, 104, 120, 144, 168, 192, 216, 240, 264, 288 and 312 hours post dose, with a total of 680 samples collected. Clinical observations and body weight were recorded daily. The rats were split into two equal sacrifice groups (Groups A and B). Clinical chemistry was performed on study days 8, 24 hours after last treatment (Group A) and on study day 14, 168 hours after last the treatment (group B). Necropsy examinations were performed on the same days. Half of the animals from each group were sacrificed on study day 8 and the other half were sacrificed on study day 14. Kidneys were removed and were examined histopathologically. Samples were stored at -40 °C prior to analysis.

2B.2.2 Sample Preparation and Experimental Methods

Prior to analysis the sample order was randomised using the Microsoft excel random number generator function. The urine samples were then separated into four equal batches of 170 samples each, which were prepared simultaneously but analysed separately. A 150 µL aliquot of urine was mixed with 300 µL water to dilute the salt concentration before protein removal via precipitation with acetonitrile (1:2 v/v). These samples were vortexed and left at -20 °C overnight before centrifugation for 5 min at 12 000 G at 4 °C. For reversed-phase chromatography, 50 µL of sample was taken and added to 150 µL water in 350 µL 96-well plates (Waters Corportation, Milford, USA). A pooled quality control (QC) sample was prepared (Want *et al.*, 2010) by combining 15 µL of each sample (across all batches, i.e. the QCs contained urine from all 680 samples) and diluting 1:4 with water. Dilutions of 1 in 2, 1 in 4 and 1 in 8 of this QC sample in water were made. The 96-well plates were stored at -20 °C until the analysis by LC-MS as described below. The plates were centrifuged at 1700 g before being placed into the autosampler at 4 °C.

Reversed-phase liquid chromatography was performed on an Acquity UPLC system (Waters Ltd, Elstree, UK) coupled to a Xevo G2 triple quadrupole time-of-flight mass spectrometer

(Waters MS technologies, Ltd., Manchester). UPLC-MS conditions were those previously described for the reversed-phase analysis of urine (Want *et al.*, 2010) with the separation performed using a Waters HSS T3 (1.8 μm , 2.1 x 100 mm) column. For UPLC-MS analysis 2 μL per sample were injected. The chromatographic mobile phase was composed of 0.1 % formic acid in water (A) and 0.1 % formic acid in acetonitrile (B). The column temperature was maintained at 40 °C and linear gradient elution was performed at either 0.5 mL/min. The starting composition was 1 % B, held for 1.0 min before increasing to 15 % at 3.0 min, 50 % at 6.0 min, 95 % at 9.0 min for a 1.0 min wash and returning to 1 % B for a 2 min re-equilibration step (total cycle time 12 min). Mass spectrometry was performed with electrospray ionisation operated in positive ESI. The instrument was operated in sensitivity mode and set to acquire data over the m/z range 50 – 1200 with a scan time of 0.1 s. All mass spectral data were collected in centroid mode using the MS^e data acquisition function. For mass accuracy, leucine encephalin (MW = 555.62) was used as a lock mass at a concentration of 200 pg/ μL (in 50/50 $\text{CH}_3\text{CN}:\text{H}_2\text{O}$, 0.1 % formic acid) infused at a flow rate of 20 $\mu\text{L}/\text{min}$ via a lock spray interface. The data were collected using MassLynx V 4.1 software (Waters Corporation, Milford, USA). In between batches the inlet of the mass spectrometer and the ESI source were cleaned, and the instrument and detector recalibrated.

2B.2.3 Data Analysis

The raw data from each of the four individual batches were processed by Markerlynx XS data analysis software (Waters Corporation, Milford, USA) for peak picking, alignment and normalization, to produce peak intensities for retention time (RT) and m/z data pairs. Further statistical analysis was performed on the resulting normalized peak intensities using SIMCA P 13.0.2 (Umetrics, Umea, Sweden) as described previously.

2B.3 Results and Discussion

2B.3.1 Quality control

As mentioned in the previous chapter it is common practice that aliquots from every sample are combined in a pooled sample and then analysed at regular intervals throughout the experiment for the purpose of quality assurance/control (Want et al., 2010). As each batch was analysed separately each data set was assessed individually.

- *Batch one.* Initial visual inspection of the QC chromatograms showed that there was significant variation in intensity and some fluctuations of retention times in the major chromatographic peaks. Figure 2B.3 shows the overlaid chromatograms of the QCs of batch one. These results were mirrored in the initial PCA scores plot of the unfiltered data (figure 2B.4), with a time-related drift across the second component (PC2). Taken together these data would generally be seen as in need of further examination before further analysis.

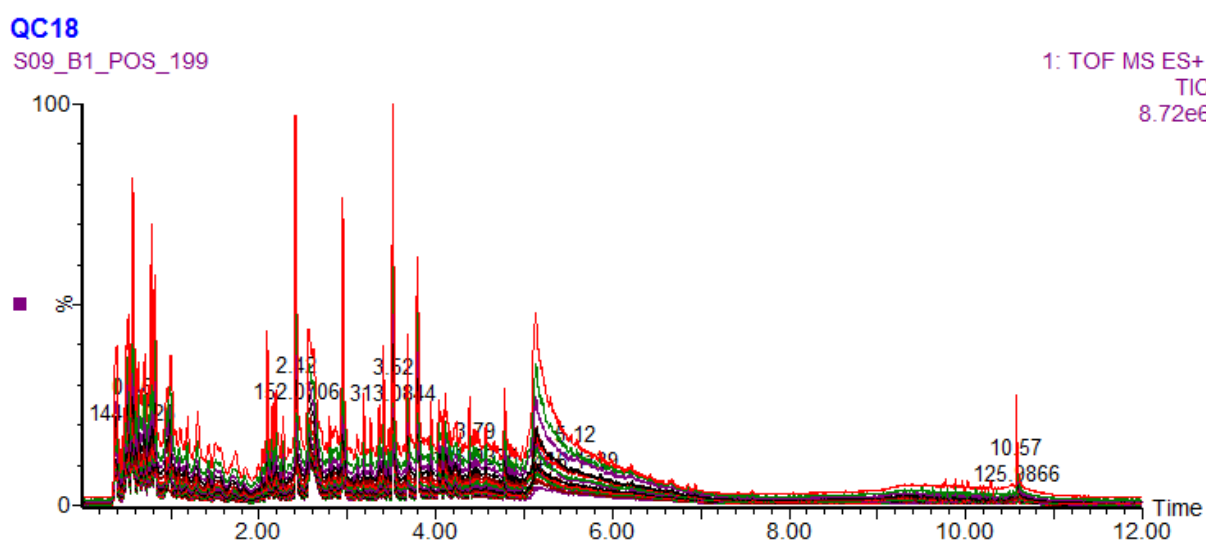


Figure 2B.3 Overlaid chromatograms of the QC injections for the first batch. There is considerable intensity variation coupled with small retention time shifts, although the majority of the major peaks seem to elute at a consistent time.

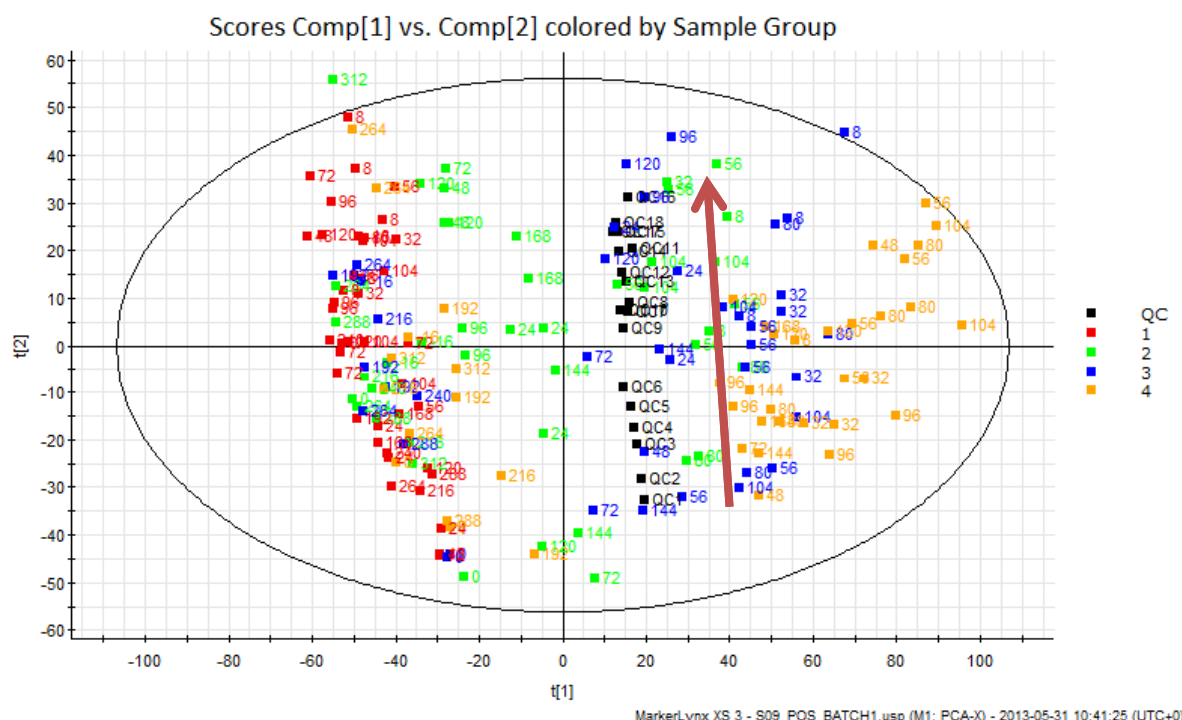


Figure 2B.4 PCA scores plot of the unfiltered data for batch one, coloured by dose group. The black points are the QCs and exhibit a time dependant drift across the second component. The red arrow shows the direction of the drift across statistical space.

- *Batch two.* Initial inspection of the QC chromatograms showed that there was much better analytical stability in the second batch. Figure 2B.5 shows the overlaid chromatograms of the QCs of batch two, illustrating the more consistent results this time. These results were also mirrored in the initial PCA scores plot of the unfiltered data (figure 2B.6), which show the QCs well grouped in the centre of the plot. The data acquired was considered to be of good quality, and suitable to be taken forward for statistical analysis without further investigation.

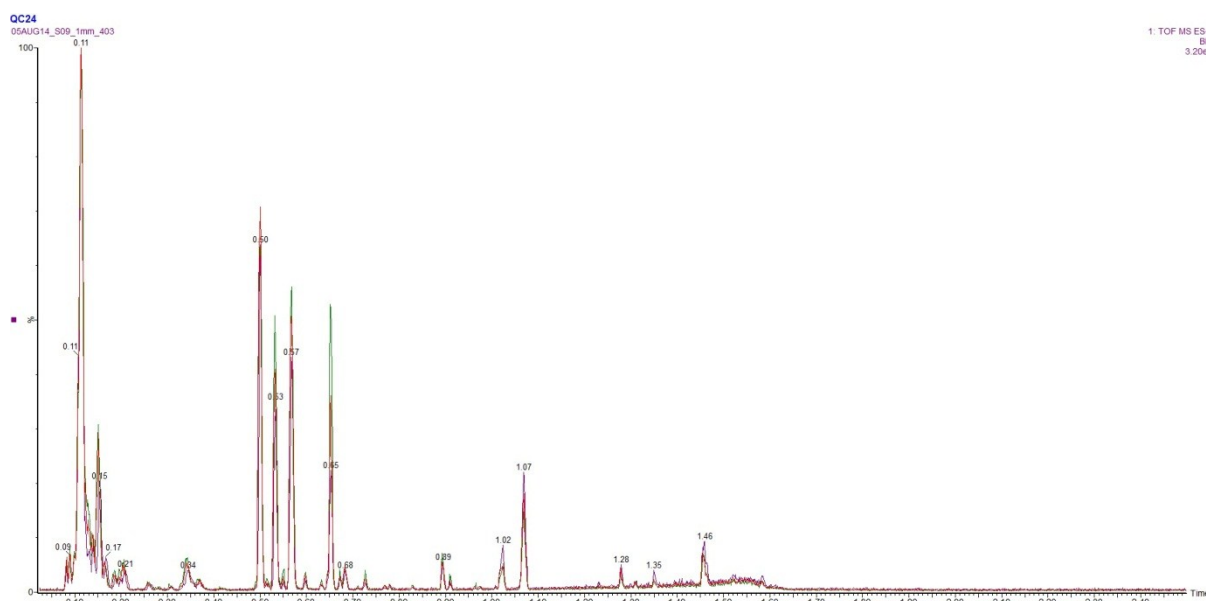


Figure 2B.5 Overlaid chromatograms of the QC injections for the second batch. There chromatographic peaks match up well, with little evidence of chromatographic drift.

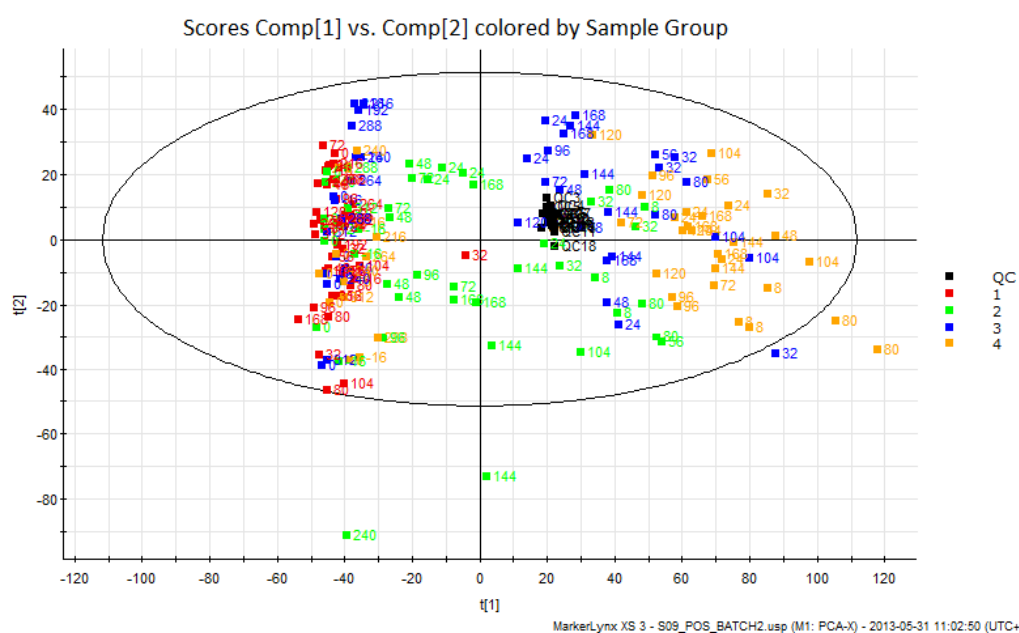


Figure 2B.6 PCA scores plot of the unfiltered data for the second batch, coloured by dose group. The black points are the QCs and cluster well in the centre of statistical space. This suggests a high level of analytical stability throughout the second run.

- *Batch three.* Similar evaluation of batch three revealed that this batch showed subject to both retention time shifts, and a general reduction of intensity as the run

progressed, resulting in data with a time-related bias. These effects are illustrated in figure 2B.7 and 8 chromatograms (figure 2B.7) and the scores plot (figure 2B.8) for the QC data. This data of this batch was considered, like batch 1 to be of lower quality than the data acquired on batch 2, requiring investigation before further statistical analysis.

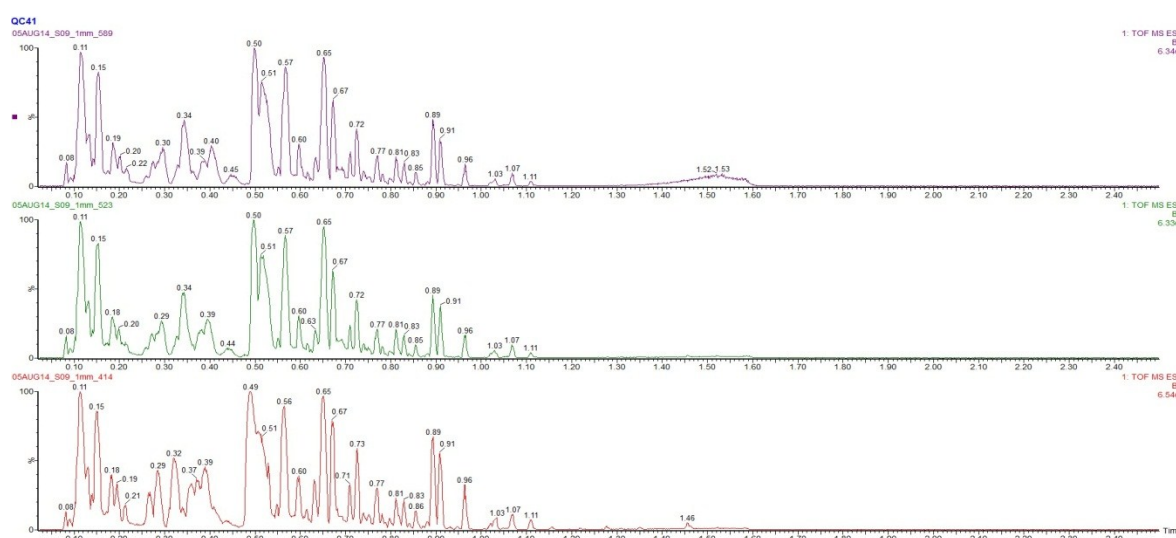


Figure 2B.7 Stacked comparison of the chromatograms of the QCs 1 (top), 10 (middle) and 18 (bottom) for batch three. There are small retention time shifts in addition to a general loss of intensity throughout the run.

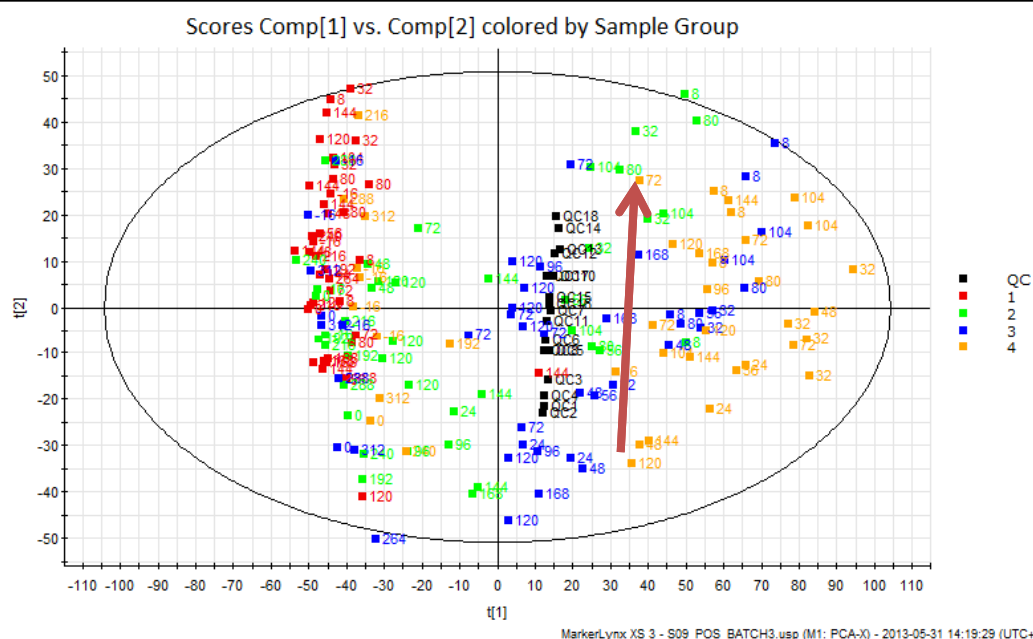


Figure 2B.8 PCA scores plot of the unfiltered data for the third batch, coloured by dose group. The black points are the QCs and show (similarly to batch one) a significant time related drift, the influence of which is across the second component. The red arrow show the direction of the time related bias.

- Batch four.** The QC chromatograms of this batch showed that there was much better analytical stability in the fourth batch, similar to that seen for batch two than seen for batches 1 and 3. Figure 2B.9 shows the overlaid chromatograms of the QCs of batch four, and the PCA scores plot (figure 2B.10) which also shows the QCs well grouped in the centre, albeit with a small time related effect. The data acquired from this run were considered to be of acceptable quality, and suitable for statistical analysis.

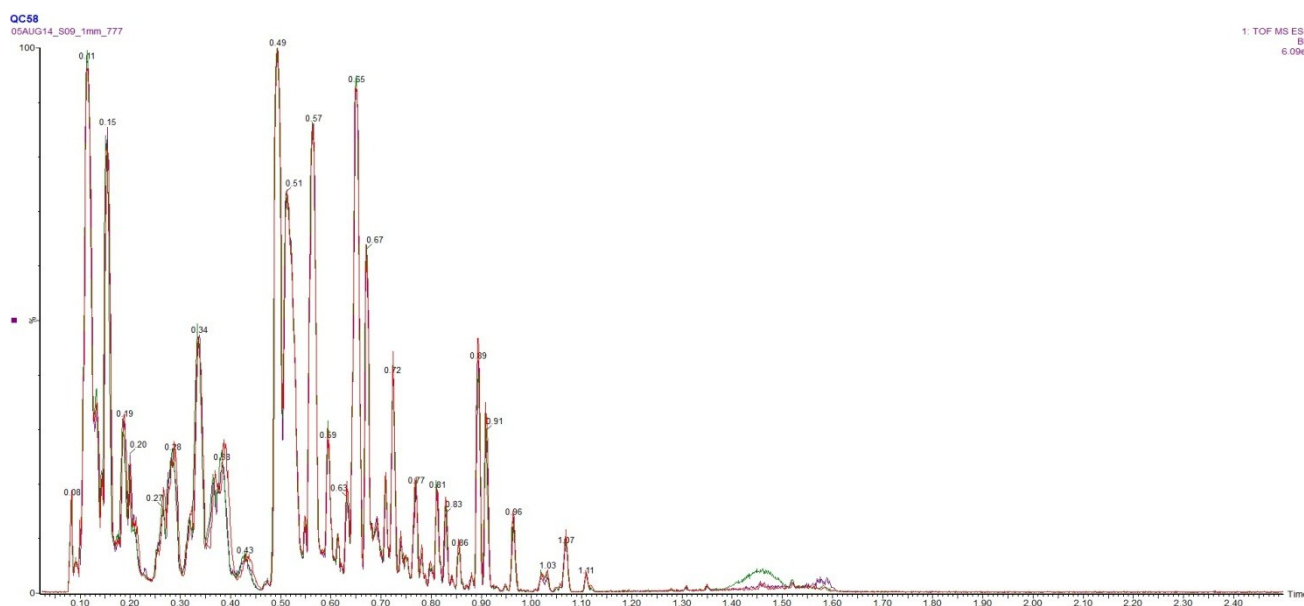


Figure 2B.9 Overlaid chromatograms of the QC injections for the fourth batch. There chromatographic peaks match up well, again with little evidence of chromatographic drift.

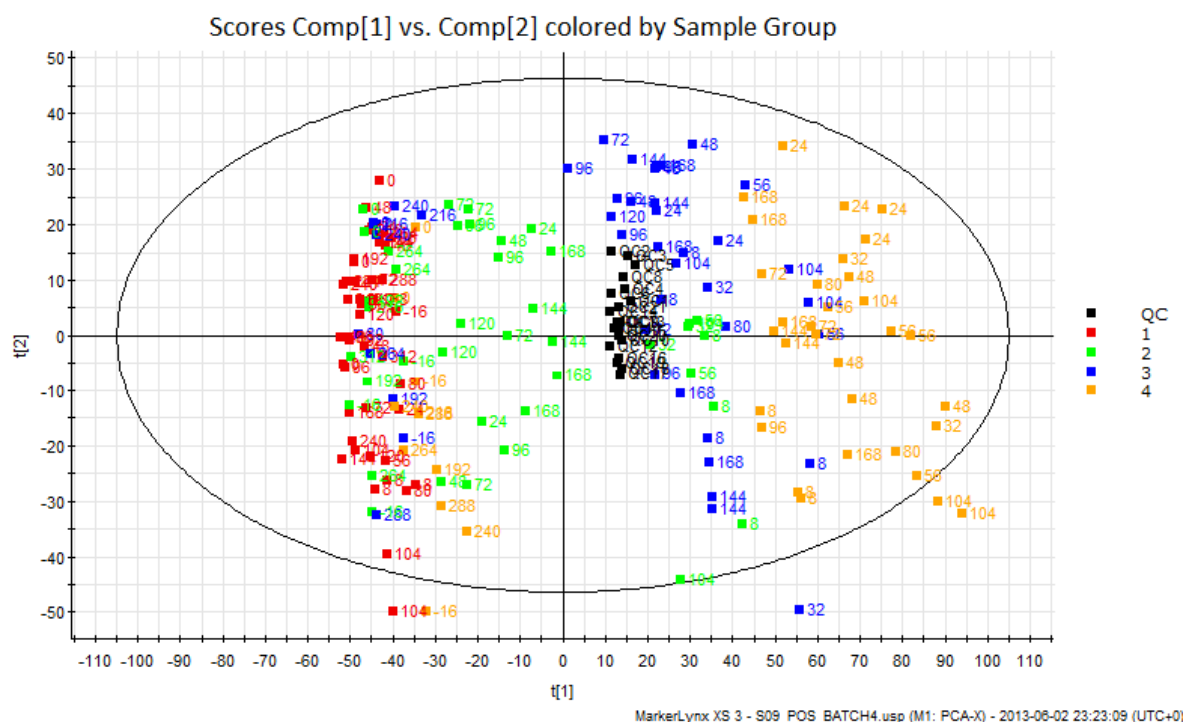


Figure 2B.10 PCA scores plot of the unfiltered data for the fourth batch, coloured by dose group. The black points are the QCs and cluster well in statistical space, although there is a slight time related influence along the second component. This suggests a reasonable level of analytical stability throughout the fourth run.

After the quality control, batches one and three were deemed to be of lower quality than batches two and four, which were considered to be more suitable for chemometric analysis. Batch one detected 27558 ions, while batches two, three and four detected 16365, 35934 and 40824 respectively. While there was a considerable amount of variation in detected signals, there seems to be no correlation between quality of the data and number of ions detected. This high level of variability between the data from each batch, is despite using identical methods on the same instrument.

However, whilst it is clear that there are a number of problems with the data generated from these analysis that require further investigation with respect to normalisation etc., it is also evident that for all four batches similar grouping of the dose groups is seen when PCA is performed. From this it is apparent that there is a clear distribution of the data that relates to acetaminophen administration and dose level. What is not possible with these results, in

their present state, is the direct comparison of the results from one batch with another, or their combination into a single data set for further statistical analysis. This is clearly a major limitation and, for this reason, an investigation of the nature of the differences in the batches was undertaken.

2B.3.2 Determination of batch effects.

The analytical drift observed here produced data which exhibited both within and between batch variations. It was therefore impossible to concurrently analyse the data from all the batches before extensive corrections were applied to remove batch bias. Batch effects are an existential problem for large scale metabolic profiling when performed using MS for detection, and as such have attracted much scientific research attention. Indeed, it was the realization that changes, in particular to sensitivity, through the course of an analytical run could lead to variability that led to the adoption of the use of the QC approach to monitor for such effects (Sangster *et al.*, 2007). Assessments of batch to batch variations have been previously performed, and the need for robust methods for removing batch bias has been extensively discussed.

Analytically, variation in detector response etc., has typically been compensated for by the application of an internal standard. The inclusion of a chemical with similar properties to the compound(s) under study makes it possible to calibrate and curve fit with the internal standard eliminating systematic sources of bias. This approach, however, whilst stable isotope labelled internal standards are widely used for target compound analysis it is difficult to apply this approach for hundreds if not thousands of analytes performed in multicomponent metabolic phenotyping analysis using MS-based methods. This is due to the large differences in seen ionisation for different metabolites, even when these may seem to be structurally similar. Furthermore, in a hypothesis free metabolic profiling experiment it is impossible to pre-emptively select a single, or indeed a representative set, of internal standards as many the metabolites present in the sample are unknown before the analysis has been undertaken.

Previously a range of computational and statistical techniques have also been employed to correct both within- and between-batch variations. When there is heteroscedastic disparity between the batches, i.e. when the variation of each chromatographic peak between each batch is unequal, scaling can be employed for correction, as is routinely used in microarray gene expression data (Luo *et al.*, 2010). Scaling typically is applied by the use of a scaling factor (which is usually the standard deviation or square root of the standard deviation); this factor is then subtracted from each feature and then the result is divided by the same scaling factor. This approach is limited by the fact that the scaling factor and the relative abundance for each peak must be calculated for each sample which may not always be possible (i.e. peaks are not present in all samples) and must also be recalculated each time a new batch data set is acquired.

Other methodologies have been based upon normalisation of the data using the QC samples. As the QCs should be representative of the entirety of the analytes within the samples, assessment of the QCs by calculating the relative standard deviation of peak intensity for each ion across the QC samples gives a good indication of reproducibility. Such QC based normalisation is either based upon quantile normalisation where the peak intensities of all samples are adjusted to account for time- related effects in a sample wide manner (Draisma *et al.*, 2010) or linear regression normalisation which corrects each ion individually using a (typically weighted) regression curve (Kamleh *et al.*, 2012). A wide variety of statistical packages and workflow strategies have been developed, for LC-MS (Wang *et al.*, 2012), DIMS (Kirwan *et al.*, 2012) and even for comparisons of data originating from heterogeneous sources (Kirwan *et al.*, 2014). While all these various methods exist for data correction, each approach based upon algorithms of ever increasing complexity, their use is still not wide spread, and there seems to be no consensus on the 'best' method, nor does there seem to be an approach which is universally applicable to every dataset.

For the determination of the extent of batch related bias within the data derived from the LC-MS analysis of these rat urine samples, the retention times, peak areas (before normalisation) and mass accuracy was compared across the four batches using the QCs.

Any feature which was not present in the QCs of all four batches was removed and remaining features were matched by mass to charge ratio, with a window of 0.01 Da. Plots comparing the retention time and m/z of the features were constructed, and are shown below (figures 2B.11 and 2B.12). The comparisons showed that there was little disparity in the RT and m/z between batches, showing a consistency of detected features across the batches.

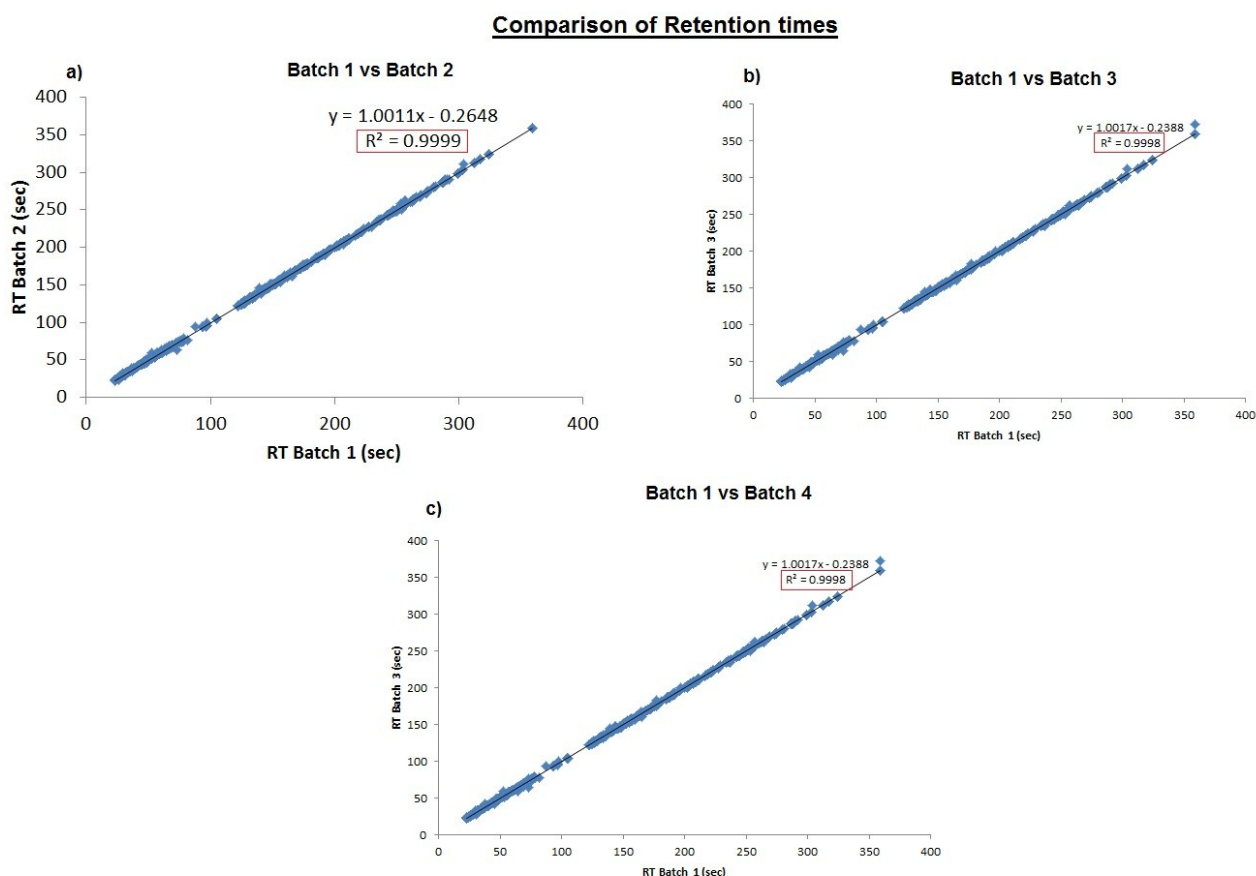


Figure 2B.11 Line plot, plotting the retention times of detected features between batch one and two (a), one and three (b) and one and four (c). R2 values (highlighted red) are an indication of the match between the two batches (expressed as a goodness of fit). The closer the value to 1, the more consistent the retention times are between the batches.

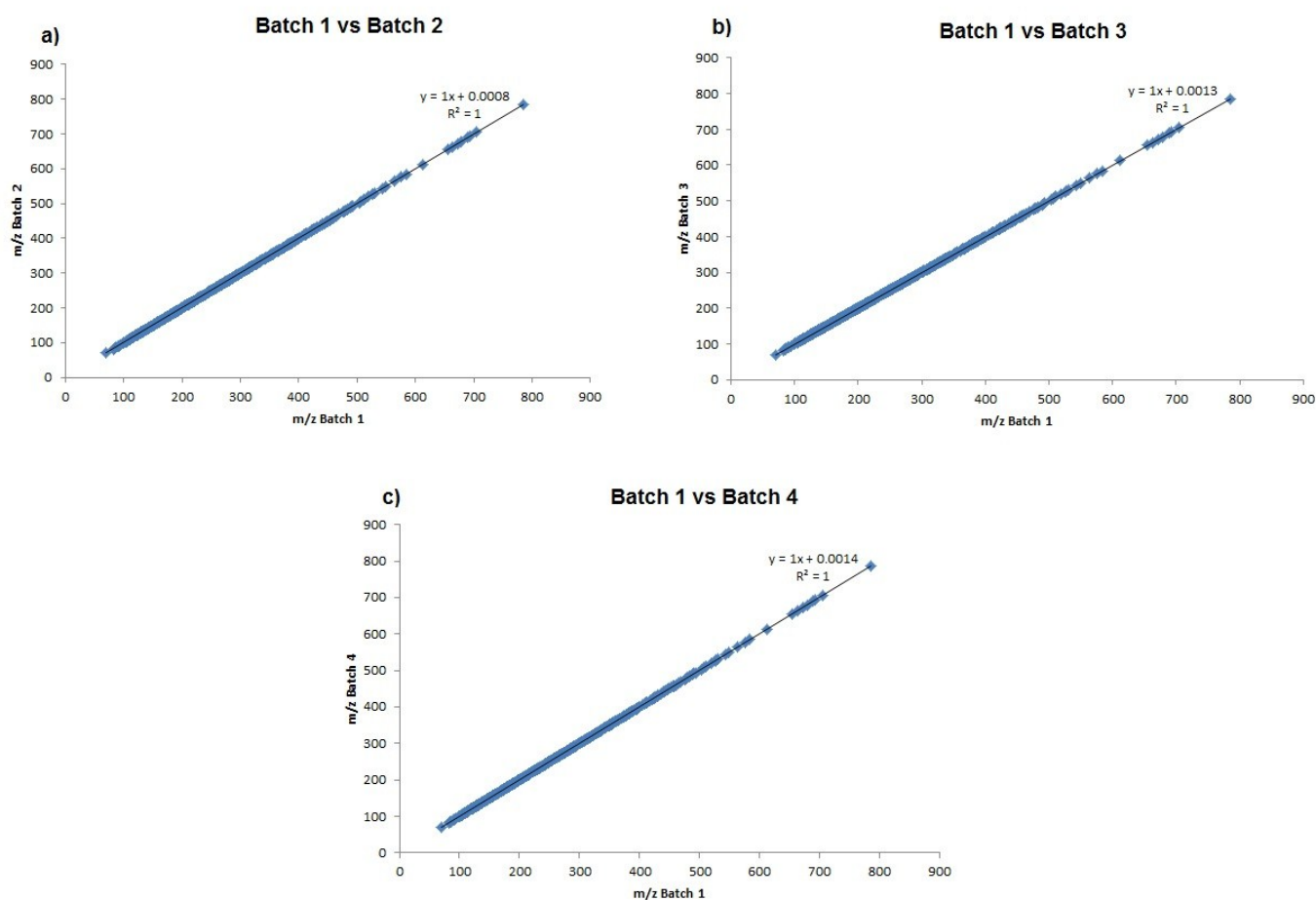
Comparison of m/z 

Figure 2B.12 Line plot, plotting the mass to charge ratios of detected features between batch one and two (a), one and three (b) and one and four (c). R^2 values (highlighted red) are an indication of the match between the two batches (expressed as a goodness of fit). The closer the value to 1, the more consistent the mass to charge ratios are between the batches.

For the assessment of any intensity related effects the chromatographic peak area of the five most intense features of the QCs were compared. There seems to be a distinct yet consistent trend with batches one and three showing a sharp decline in intensity throughout the run, with batches two and four being more stable. This is illustrated in figure 2B.13. To ascertain whether these trends were only limited to a specific types of ions which ionise well we blindly selected nine new features which spanned the intensity range. Hippurate was also evaluated as this is a biologically significant metabolite to a toxicology study to make sure these trends were also mirrored in analytes relevant to the study. These are illustrated in figure 2B.14.

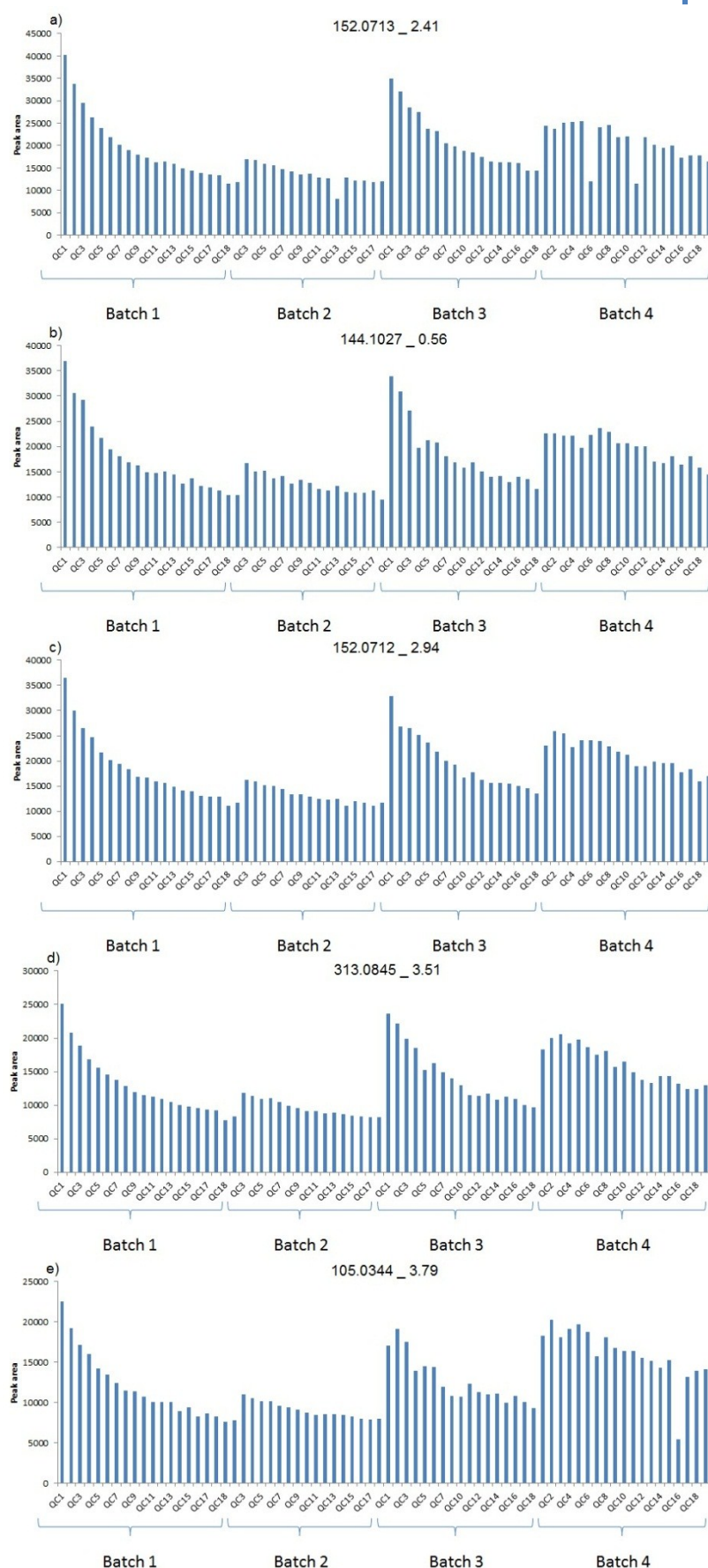


Figure 2B.13 Bar plot, plotting the peak area of (a) the first (m/z 152.07, RT 2.41), (b) second (m/z 144.10, RT 0.56), (c) third (m/z 152.07, RT 2.94), (d) fourth (m/z 313.08, RT 3.51) and (e) fifth (m/z 105.03, RT 3.79) most intense features across the QC of the four batches. The intensities show a consistent trend across all the features.

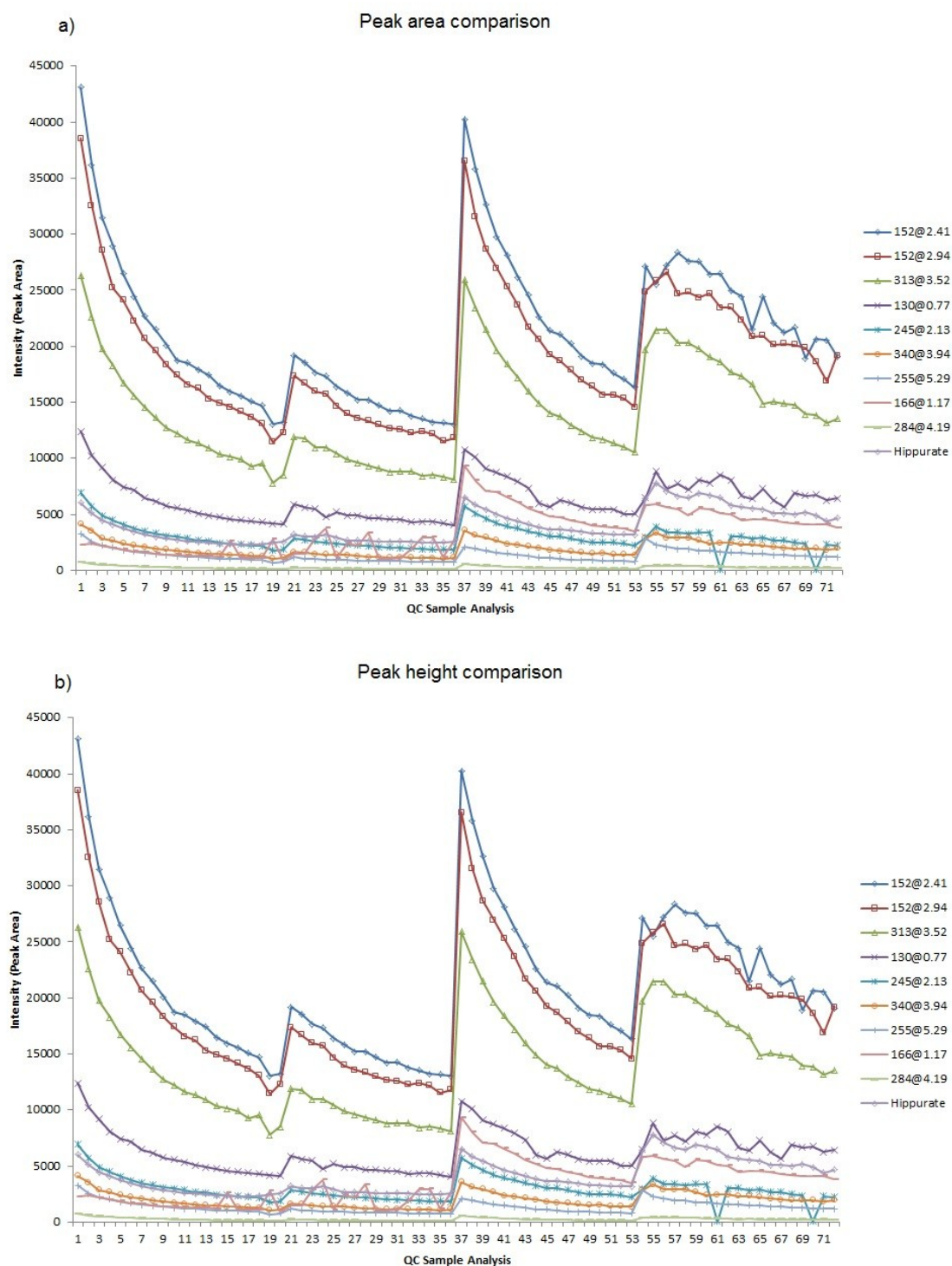


Figure 2B.14 Line plot, plotting the comparison of peak area of (a) and peak height (b) for the nine randomly selected metabolites and hippurate across the QCs of each batch. Each metabolite is represented by a coloured line. The trends observed match those seen previously, showing that the systematic bias is consistent across all ions.

2B.3.3 Application of median fold normalisation for batch correction

As discussed above there are many strategies available for the correction of batch effects. It was therefore important to evaluate which approach would be most suitable for the dataset. Diagnostics of the variation in overall intensity between samples revealed large relative variation between both samples of the different batches but also within the same batch. If it is assumed that up to half of peak intensities are increased or decreased due to biological effects, the log ratios of non-differentially abundant metabolite peaks should have a comparably small spread around zero across samples (as shown by the box in figure 2B.15) and a median of zero. However, the peak intensities of non-normalized samples exhibit large variation (up to a 16 fold difference) in overall intensities. Thus it is apparent that the overall sample intensities are attenuated within the batch.

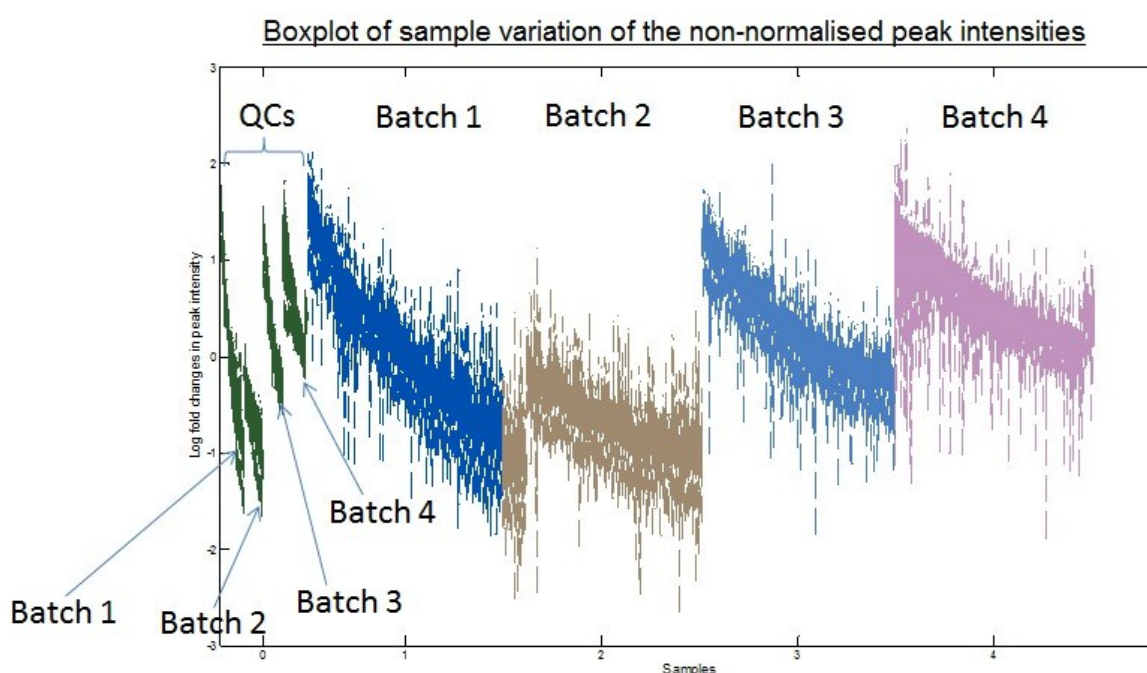


Figure 2B.15 Box plot of sample variation of the non-normalised peak intensities, coloured by batch. Each (very thin) box represents the sample variation in log fold change in peak intensity. It is very apparent that there is a significant level of difference between the variance from sample to sample, with variation decreasing as the run progresses, in line with the assumption that variation decreases as intensity decreases.

After determining that the bias within the data had a multifaceted aspect with clear heteroscedastic noise structure, the application of simple scaling could result in altered downstream statistical analysis and was therefore unsuitable for the datasets. This is due to the fact that non biological sources of variation tend to have a greater effect in features of higher intensity (Systi-Aho *et al.*, 2007), what is called ‘multiplicative noise’. Most multivariate statistical modelling employed (including PCA and OPLS) assumes that the underlying noise structure is independent of signal intensity. Thus it was necessary to apply variance stabilisation before normalisation could be employed.

In collaboration with Dr Kirill Vaselkov, Imperial College London, glog variance stabilisation was applied to the dataset. The algorithm employed has been extensively explained by Vaselkov *et al.* (2011) but in summary aims to remove the intensity dependence of the sample variation by transforming the multiplicative error into additive. Under the assumption that background noise is small glog is calculated as:

$$\text{glog}(x) = \log\left(\frac{x + \sqrt{x + c}}{2}\right)$$

Where x is the feature and x is the intensity of the feature, with c as the fitted value of the feature. This is an appropriate variance stabilising transformation as this approximation is close to linear for low intensity features, close to logarithmic for high intensity peaks and a smooth intermediate for features in between. The function smoothly ‘interpolates’ between linear and logarithmic and is therefore called ‘linear interpolation’. Figure 2B.16a illustrates the influence of the multiplicative noise by presenting the standard deviation as a function of the mean peak intensities of the QC samples; metabolites with greater peak intensities to exhibit larger variability when repeatedly measured. The plot to the right (figure 2B.16b) illustrates how the application of variance stabilisation succeeded in correcting the data.

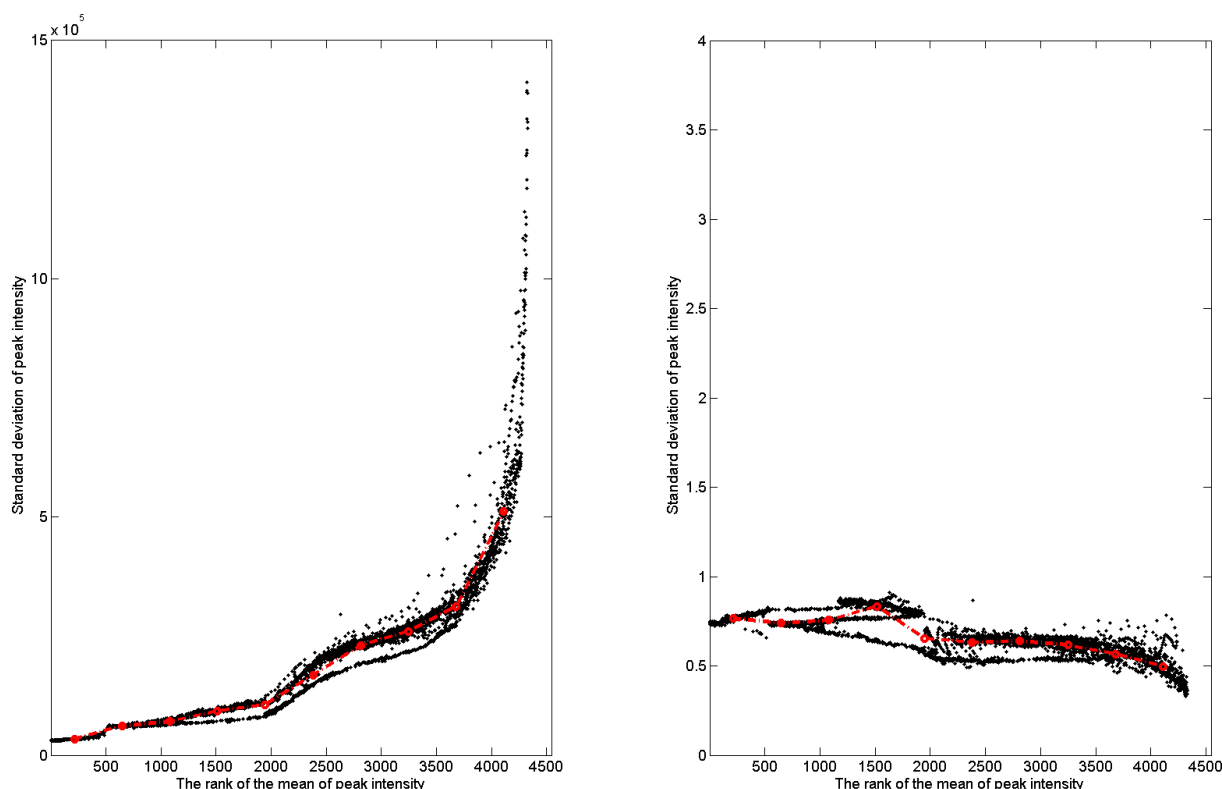


Figure 2B.16 Dot plots of the standard deviation as a function of the rank of the mean of peak intensity of the QC samples of (a) the original data and (b) after glog variance stabilisation. The dotted red line is the running median of the standard deviation. In the absence of multiplicative noise this line should be horizontal.

Median fold change normalisation was subsequently applied to equalize the variation in overall intensity across the samples of all batches. Median fold normalisation has been extensively used in NMR spectroscopy, and removes sources of systematic variation by setting the median of log fold changes across the peak intensities of all samples to zero (Bolstad *et al*, 2003). The underlying assumption of this form of normalisation is that the analyte peaks affected only by dilution will behave in the same manner across all the QC samples. This fold change is calculated by:

$$n_k^{\text{MFC}} = \text{median} \left(\frac{x_{ik}}{x_{ir}} \right)$$

Where n_k^{MFC} is the normalisation factor calculated from the median fold change of peak intensity between the k th sample (x_{ik}) and the target sample (x_{ir}). Figure 2B.17 shows the variance box plots of the samples from each dataset after median fold normalisation. As the box plots align it is apparent that the samples are now directly comparable across batches.

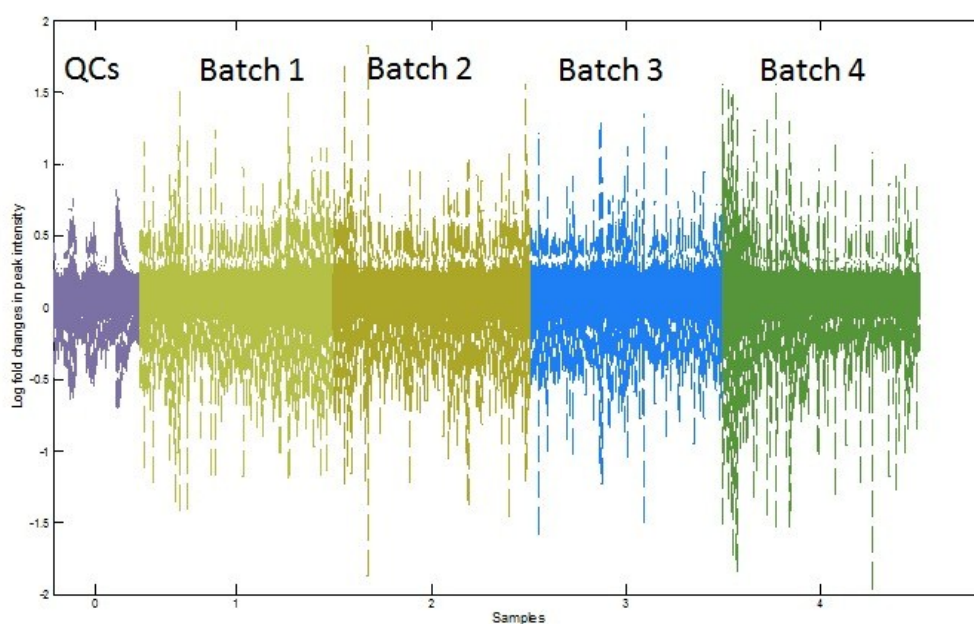


Figure 2B.17 Box plot of sample variation of the normalised peak intensities, coloured by batch. Each (very thin) box represents the sample variation in log fold change in peak intensity. The variation has been normalised across the four batches, suggesting that the datasets are now directly comparable.

2B.3.4 Effect on acetaminophen metabolites

To identify the effect the normalisation had on the biologically relevant acetaminophen metabolites the data was interrogated to find the major expected metabolites, acetaminophen glucuronide and acetaminophen sulphate, in addition to the parent compound.

Initial multivariate assessment of the peak area normalised data for each batch was conducted. The analysis was the same for each batch and involved initial PCA modelling to reveal distribution patterns. All four batches showed discrimination between dose groups,

with maximal response to dosing for the high dose group between 48 and 56 hours, with the metabolic profiles of the recovery animals returning to 'normal' (i.e. similar to control animal profiles) by the 192 hour time point (8 days after initial dose and within 48 hours after dosing had ceased). The response seemed to be dose and time dependant which could indicate that the metabolic strain caused by acetaminophen is accumulative. Subsequent OPLS modelling of the high dose (800mg/kg) group against the control animals provided strong results for each of the batches (summarised in table 2B.1); these models were then used to identify the signals driving the separation. Figure 2B.16 shows the OPLS-DA scores plots for each batch with 2 orthogonal components. The variables with the greatest influence on the separation were consistent throughout the models for each of the batches and represented acetaminophen (m/z 152.07), acetaminophen glucuronide (m/z 328.10 or 350.08 as the sodium adduct), acetaminophen sulphate (m/z 232.02) and N-acetylcysteinyl-acetaminophen (m/z 313.08). This is illustrated in figure 2B.17 which presents the S- plots of the OPLS-DA models for each batch.

Batch	R ²	Q ²	P Value (CV)
One	0.62	0.54	3.15883e ⁻⁰¹²
Two	0.64	0.54	3.34866e ⁻⁰¹³
Three	0.65	0.53	1.06355e ⁻⁰¹⁰
Four	0.75	0.64	1.82352e ⁻⁰¹⁷

Table 2B.1 Summary of the statistics for the OPLS-DA models for each batch

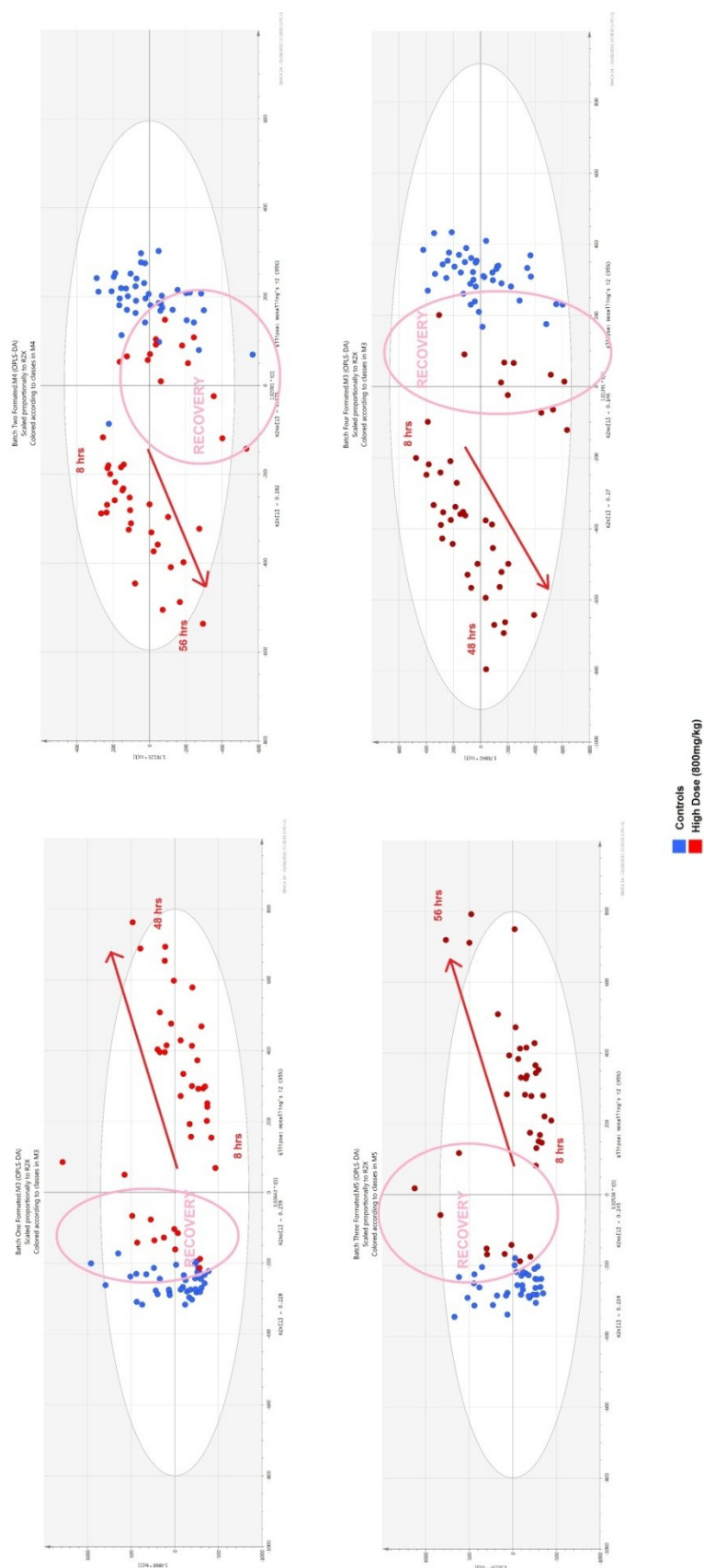


Figure 2B.18 Scores plot from the OPLS-DA models for each batch. There is a clear separation between the high dose samples (red) and the controls (blue), with a time related trend. The red arrow shows the direction of the time influence. The high dose samples within the pink circle are samples after dosing has ceased and therefore exhibit a recovery.

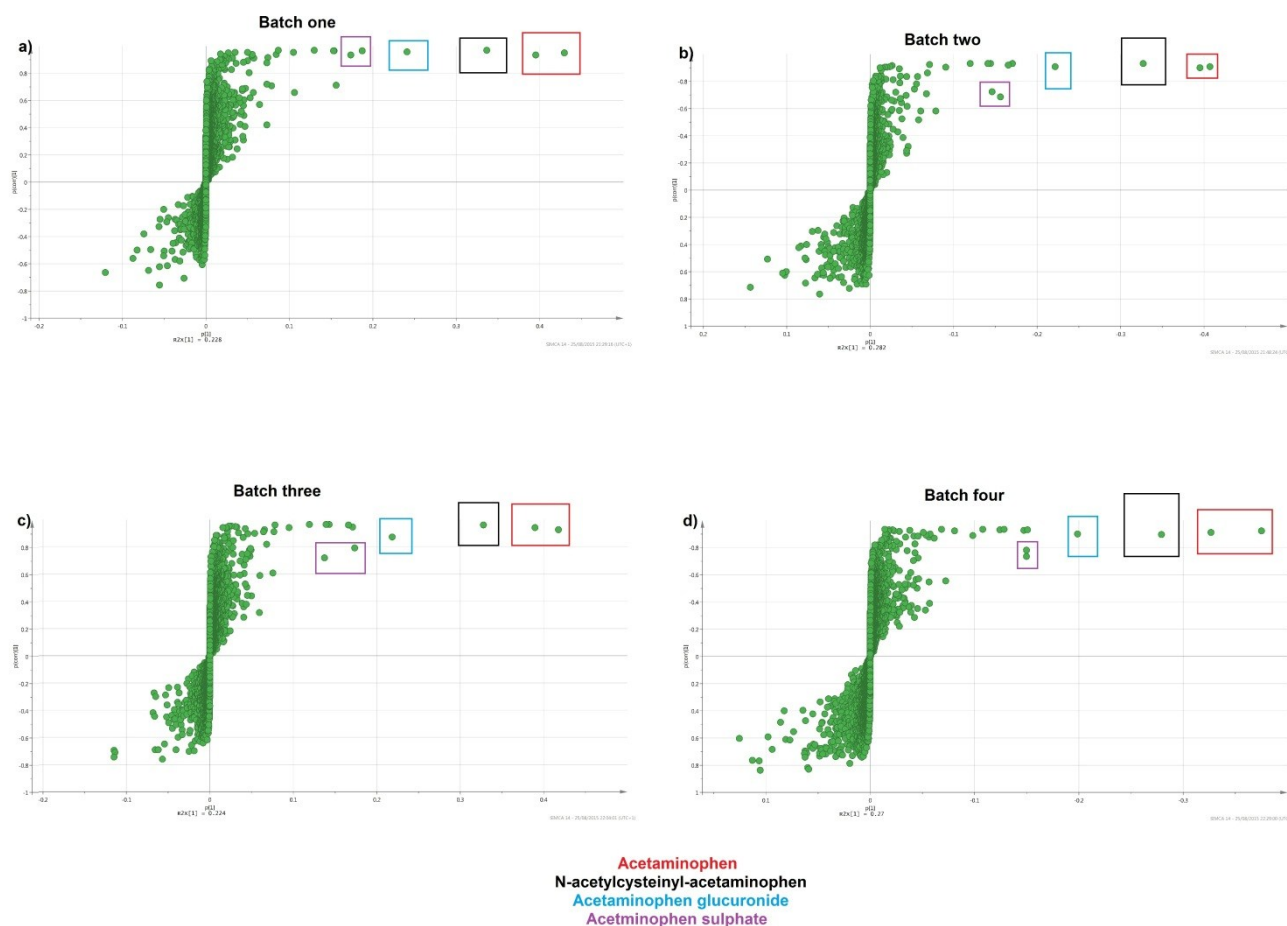


Figure 2B.19 Composite s-plot derived from the OPLS-DA models comparing high dose (800mg/kg) and control animals. The S-plot shows which variables exact the greatest influence on the model, driving separation of the classes. Variables in the top right are more abundant in the dose group, while variables in the lower left are more abundant in controls. The dots highlighted in the red boxes relate to the signals from the parent compound acetaminophen, in the black boxes N-acetylcysteinyl-acetaminophen, acetaminophen glucuronide in the cyan and acetaminophen sulphate in the purple boxes respectively. While there is variation in the quality of the data, and the exact statistics differ slightly, it is apparent that the same metabolites have been detected in all batches, and are in every case the signals influencing the separation

These results showed that for each batch separately, it is possible to differentiate between the dose groups and, crucially, the signals driving this separation are the acetaminophen metabolites (unlike the diclofenac example discussed in the first part of this chapter). After the batch correction was applied the effect this had on these signals for the drug metabolites was investigated. Thus the chromatographic peaks for acetaminophen, and the sulphate and glucuronide conjugates were evaluated. Figure 2B.20 shows the variation that these analytes exhibited in the QC samples even after batch correction. PCA modelling of just the

batch corrected QCs showed that the QC samples separated into two sub groups, which should not be happening if the samples were analytically identical. When the same metabolites were compared in the total dataset, including all 720 test samples, the variation still followed a batch-related pattern, and differentiation between dose groups was obscured as the batch influence was too strong, as seen in figure 2B.19. This seems to suggest that the batch correction was unsuccessful in stitching the QCs from the different batches together, while batch effects are also still present in the test samples themselves.

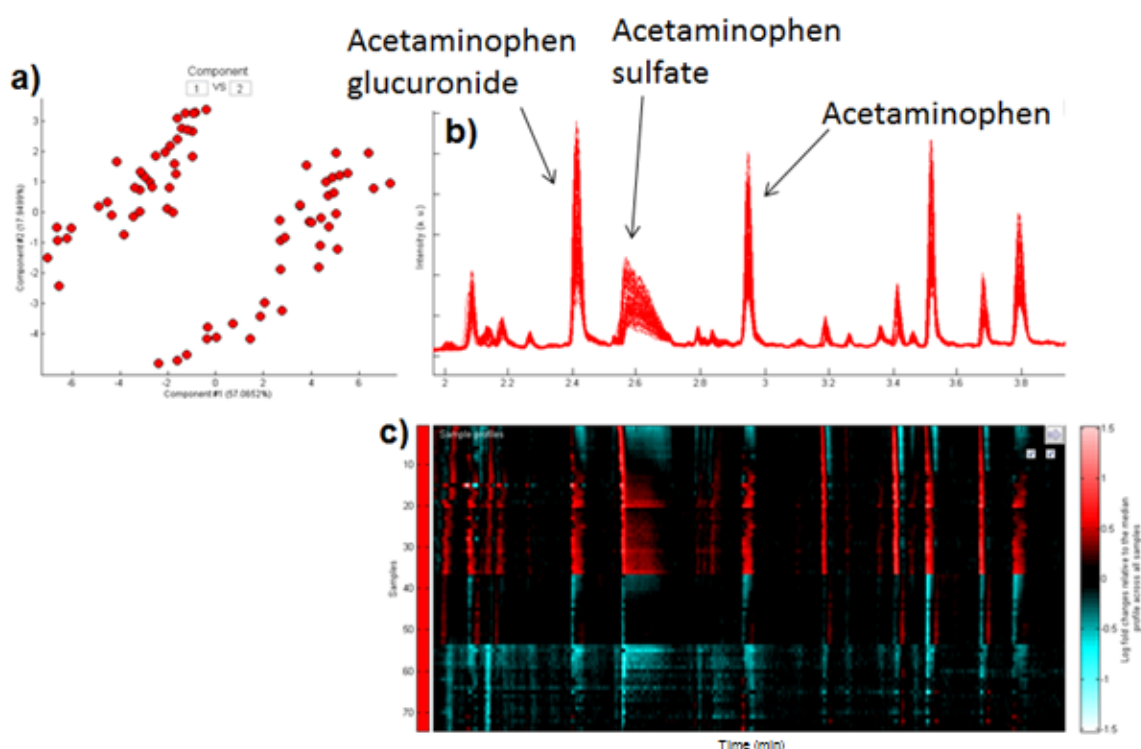


Figure 2B.20 Composite plot of post batch correction data. A) The PCA scores plot of the combined QC samples from all four batches. The QCs segregate into two sub groups, the top left being from batches two and four and the group to the lower right from batches one and three; if the batches were directly comparable there should be no clear distinction between the QCs of different batches. B) The overlaid chromatograms of the QC samples, with the acetaminophen, acetaminophen glucuronide and acetaminophen sulphate peaks highlighted. The peaks of these important metabolites show high variability, despite the fact that as QCs each sample should produce the same signal. C) Heatmap plot indicating the log fold distance for each QC sample for the median. Each row signifies one sample and the colour indicates how far that m/z range deviates from the median. There are clearly two subgroups as indicated by the mostly red appearance of half the QCs while the other half are mostly blue. In the situation where there is little variation between the samples (i.e. all samples are close to the median for the peaks, the

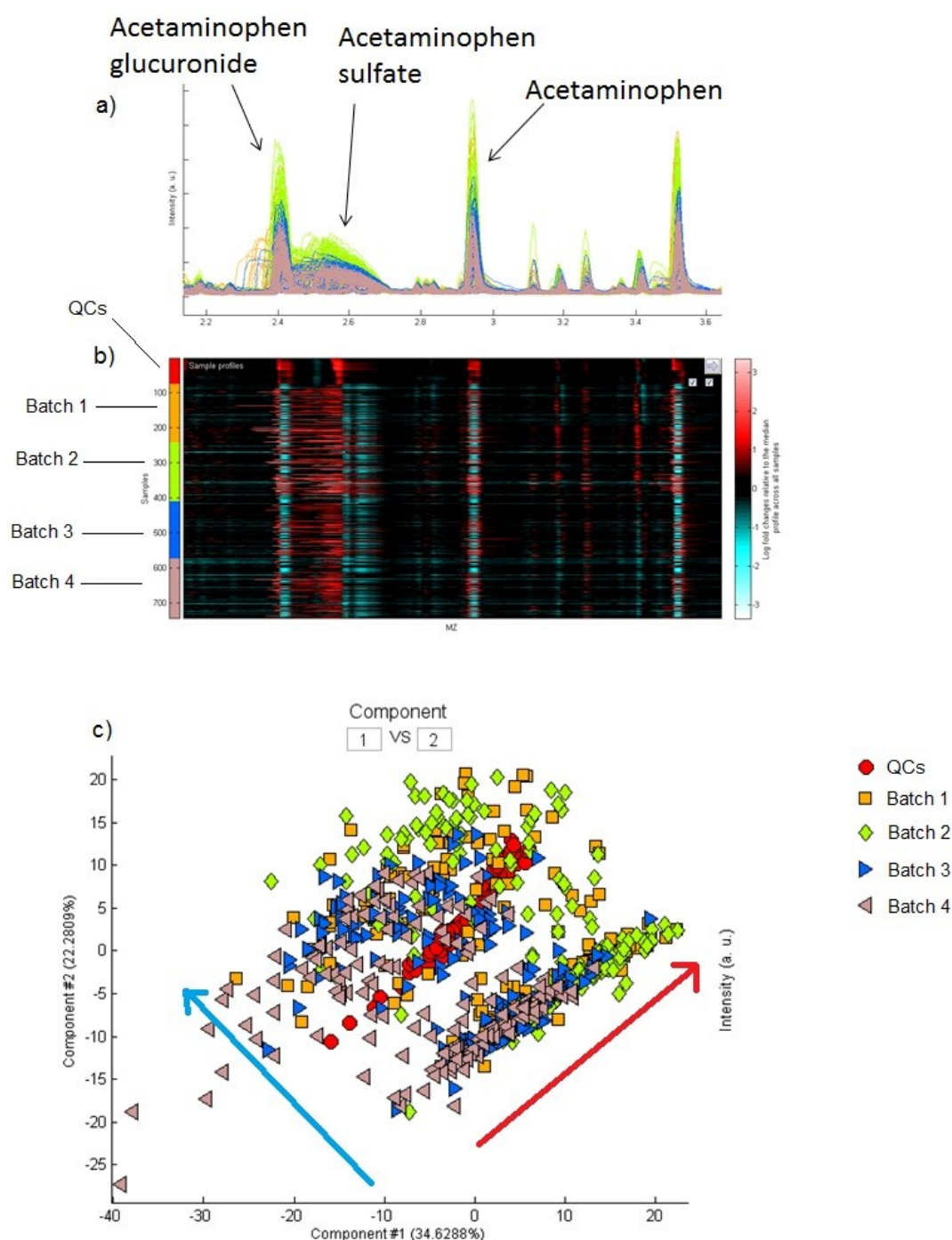


Figure 2B.21 Composite plot of post batch correction data. A) The overlaid chromatograms of the complete data set containing all samples, coloured by batch, with the acetaminophen, acetaminophen glucuronide and acetaminophen sulphate peaks highlighted. The peaks of these important metabolites show batch dependant variability. B) Heatmap plot indicating the log fold distance for each sample from the median. Each row signifies one sample and the colour indicates how far that m/z range deviates from the median. There are clearly batch related effects. The fold change should be dose dependant if batch correction was successful. C) The PCA scores plot of all the combined samples from all four batches, coloured by batch, with the red points being the QCs. While there is some evidence of dose related segregation, the influence of which is shown by the blue arrow, there are significant batch related effects (red arrow) for the statistical modelling to be obscured.

In the previous study (chapter 2A) the utility of computational methodologies for the investigation of the metabolism of xenobiotics, in this case diclofenac, was evaluated. The conclusion was that in a small scale metabolism study, using a small number of animals and a low dose, biological effects such as diurnal variation within the samples will dominate the statistics, rendering detection of both metabolites of the compound under study but also markers of the effect of administration quite difficult. This highlights an unavoidable limitation of the computational approach, as detection is based on statistical power. Indeed most previous studies using statistical methods have been *in vitro* where the system is tightly regulated and the relative dose is higher.

In a larger scale toxicological study, such as the COMET acetaminophen study presented here, both the sample size and the dose levels are much larger than used for the diclofenac work. In such a study, where the parent compound and its metabolites dominate the metabolic landscape, statistical methods were highly applicable. PCA and OPLS-DA modelling were able to easily distinguish between treatment groups, and were able to also identify the expected key metabolites and the presence of the parent drug. However, the large scale nature of the study necessitated the use of multiple batches for the analysis. This brought the severe problems of batch-to-batch variation and within- batch injection order to the forefront. After the determination of the extent of the batch effects, linear interpolation and median fold change normalisation were applied in an attempt to make the data directly comparable and to merge the datasets into one. However, despite mathematically reducing multiplicative noise and equalising variations across the batches, this approach was unable to correct the data to the extent where it was suitable for combined downstream analysis, as not only was the variation causing the batch differences not eradicated but subsequently there was loss of potentially biologically related differences.

There has been much research effort directed at the development of algorithms for the correction of this well-known limitation of MS based metabonomics. While many studies have reported varying levels of success, the method attempted here was ultimately

unsuccessful. There are many other normalisation methods which could have been attempted, including more sophisticated algorithms such as those based on locally weighted scatter plot smoothing (LOESS) normalization (Dunn *et al.*, 2011). However, further tampering with the data could result in loss of biologically important information and was therefore deemed an invalid approach. Ultimately the origin of the problem lies in the analytical instability of the mass-spectrometry platform, and as such is in essence an analytical issue. It is therefore possibly more sensible to attempt to ameliorate the problem at the source, thereby acquiring good data, rather to continue to undertake ever more convoluted processes for the correction of poor data.

2B.5 Conclusions

The work described here are an evaluation of the utility of computational approaches to the investigation of xenobiotic metabolism. The results of the previous sub chapter (2A) showed that in studies where the amount of administered drug is low, the effects of biological processes such as diurnal variation dominate the statistical analysis of the urinary metabolotypes. Furthermore preclinical drug metabolism studies generally involved small samples sizes, which results in reduced statistical power. This lack of ability to detect an effect coupled with the obscuring of any signals originating from the drug render statistical approached ineffective in such situations. Indeed in the presented work the analysis failed to detect the major metabolites of diclofenac, despite the fact that the presence of metabolites was confirmed using manual scanning of the data.

In the case of acetaminophen presented in this sub chapter (2B), however, both the dosing and sample size was much greater. This set up is more common in large scale toxicology studies. In this situation the metabolic signatures from both parent compound and the major metabolites are so intense that they dominate both the metabolic and the statistical landscape. This enabled the successful identification of the expected metabolites through

the computational means. In this case the application of such technologies is highly favourable.

However, the analysis of large numbers of samples in turn exposes one of the fundamental limitations of the mass spectrometry platform. As mentioned previously in more details it is necessary to split the study analysis into several batches, which results in many problems with trying to combine data. The fluctuations within the data, mainly due to changes in detector sensitivity throughout the run, require complicated algorithms to adjust for. The work presented here ultimately was not able to adequately correct the data. This prompted the further work discussed in the next chapter, where development of a rapid gradient was performed in an attempt to reduce analysis time and therefore alleviate the need for batches.

2B.6 References

Banks, A.T., Zimmerman, H.J., Ishak, K.G., Harter, J.G., (1995). Diclofenac associated hepatotoxicity: analysis of 180 cases reported to the Food and Drug Administration as adverse reactions. *Hepatology*, **22**, 820–827.

Bernal, W., Auzinger, G., Dhawan, A., Wendon, J., (2010). Acute liver failure. *Lancet*, **376**, 190–201.

Boelsterli, U.A., Zimmerman, H.J., Kretz-Rommel, A., (1995). Idiosyncratic liver toxicity of nonsteroidal anti-inflammatory drugs: molecular mechanisms and pathology. *Critical Reviews in Toxicology*, **25**, 207–235.

Bolstad, B. M., Irizarry, R. A., Astrand, M., Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**, 185-193.

Coen, M., Lenz, E.M., Nicholson, J.K., Wilson, I.D., Pognan, F., Lindon, J.C., (2003). An integrated metabonomic investigation of acetaminophen toxicity in the mouse using NMR spectroscopy. *Chemical research in Toxicology*, **16**, 295-303.

Dahlin, D.C., Miwa, G.T., Lu, A.Y., Nelson, S.D., (1984). N-acetyl-pbenzoquinone imine: A cytochrome P-450-mediated oxidation product of acetaminophen. *Proceedings of the National Academy of Science USA*, **81**, 1327–1331.

Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles J.D., Halsall, A., Haselden, J.N., Nicholls., A.W., Wilson, I.D., Kell D.B., Goodacre, R., HUSERMET Consortium, (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, **6**, 1060–1083.

Draisma, H.H.M., Reijmers, T.H., van der Kloet, F., Bobeldijk-Pastorova, I., Spies-Faber, E., Vogels, J.T.W.E., Meulman, J.J., Boomsma, D.I., van der Greef, J., Hankemeier, T., (2010) Equating, or correction for between-block effects with application to body fluid LC–MS and NMR metabolomics data sets. *Analytical Chemistry*, **82**, 1039–1046.

Faed, E.M., (1984). Properties of acyl glucuronides: implications for studies of the pharmacokinetics and metabolism of acidic drugs. *Drug Metabolism Reviews*, **15**, 1213–1249.

FDA guidelines (2011). Prescription Drug Products Containing Acetaminophen; Actions To Reduce Liver Injury From Unintentional Overdose. *Federal Register*, **76**, 2691–2697.

Godejohann, M., Tseng, L.H., Braumann, U., Fuchser, J., Spraul, M., (2004). Characterization of a paracetamol metabolite using on-line LC-SPE-NMR-MS and a cryogenic NMR probe. *Journal of Chromatography A*, **1058**, 191–196.

Gu, H., Marth, J. D., Orban, P. C., Mossman, H., Rajewsky, K., (1994) Deletion of a polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science*, **265**, 103–106.

Jones, D.P., Lemasters, J.J., Han, D., Boelsteri, U.A., Kaplowitz, N., (2010). Mechanisms of pathogenesis in drug hepatotoxicity putting the stress on mitochondria. *Molecular Interventions*, **10**, 98–111.

Kallio, J.M., Lahti, M., Oikari, A., Kronberg, L., (2010). Metabolites of the aquatic pollutant diclofenac in fish bile. *Environmental Science Technology*, **44**, 7213–7219.

Kamleh, M.A., Ebbels, T.M.D., Spagou, K., Masson, P., Want, E.J., (2012) Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies. *Analytical Chemistry*, **84**, 2670–2677.

Kanda, H., Kobayashi, K., Yamanaka, H. and Noguchi, K. (2013), COX-1-dependent prostaglandin D2 in microglia contributes to neuropathic pain via DP2 receptor in spinal neurons. *Glia*, **61**, 943–956.

Klaassen, C.D., Watkins, J.B., Casarett, L.J., (2010). Casarett & Doull's essentials of toxicology, 2nd ed., *McGraw-Hill Medical*, New York.

Kirwan, J.A., Broadhurts, D.I., Davidson, R.L., Vlant, M.R., (2013). Characterising and correcting batch variation in an automated direct infusion mass spectrometry (DIMS) metabolomics workflow. *Analytical and Bioanalytical Chemistry*, **406**, 5147-5157.

Kretz-Rommel, A., Boelsterli, U.A., (1994). Mechanism of covalent adduct formation of diclofenac to rat hepatic microsomal proteins. Retention of the glucuronic acid moiety in the adduct. *Drug Metabolism and Disposition*, **22**, 956–961.

Le, Y., Sauer, B., (2001). Conditional gene knockout using Cre recombinase. *Molecular Biotechnology*, **17**, 269–275.

Lenz, E.M., Wilson, I.D., (2007) Analytical strategies in metabonomics. *Journal of Proteome Research*, **6**, 443-458.

Li, F., Lu, J., Ma, X., (2011). Profiling the reactive metabolites of xenobiotics using metabolomic technologies. *Chemical Research in Toxicology*, **24**, 744–751.

Liu, G., Zhou, Y., Herman, D., Li, Y., Deng, Y., Fang, H., Bushel, P., Woods, M., Zhang, J. (2010). A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. *The Pharmacogenomics Journal*, **10**, 278–291.

Luo, J., Schumacher, M., Scherer, A., Sanoudou, D., Megherbi, D., Davison, T., Shi, T., Tong, W., Shi, L., Hong, J., Zhao, C., Elloumi, F., Shi, W., Thomas, R., Lin, S., Tillinghast, Mitchell, J.R., Jollow, D.J., Potter, W.Z., Gillette, J.R., Brodie, B.B., (1973), Acetaminophen-induced hepatic necrosis: Protective role of glutathione. *Journal of Pharmacology and Experimental Therapeutics*, **187**, 211–217.

Nelson, S.D., (1982). Metabolic activation and drug toxicity. *Journal of Medical Chemistry*, **25**, 753–765.

Nicholls, A.W., Farrant, R.D., Shockcor, J.P., Uneger, S.E., Wilson, I.D., Lindon, J.C., Nicholson, J.K., (1997). NMR and HPLC-NMR spectroscopic studies of futile deacetylation in paracetamol metabolites in rat and man. *Journal of Pharmaceutical and Biomedical Analysis*, **15**, 901–910.

Osmundsen, H., Bjørnstad, K. (1985). Inhibitory effects of some long-chain unsaturated fatty acids on mitochondrial beta-oxidation. Effects of streptozotocin-induced diabetes on mitochondrial beta-oxidation of polyunsaturated fatty acids. *Biochemical Journal*, **230**, 329–337.

Pineau, T., Fernandez-Salguero, P., Lee, S.S., McPhail, T., Ward, J.M., Gonzalez, F.J., (1995). Neonatal lethality associated with respiratory distress in mice lacking cytochrome P450 1A2. *Proceedings of the National Academy of Sciences, USA*, **92**, 5134–5138.

Poon, G.K., Chen, Q., Teffera, Y., Ngui, J.S., Griffin, P.R., Braun, M.P., Doss, G.A., Freeden, C., Stearns, R.A., Evans, D.C., Baillie, T.A., Tang, W., (2001). Bioactivation of diclofenac via benzoquinone imine intermediates identification of urinary mercapturic acid derivatives in rats and humans. *Drug Metabolism and Disposition*, **29**, 1608–1613.

Prescott, L.F., (1980). Kinetics and metabolism of paracetamol and phenacetin. *British Journal of Clinical Pharmacology*, **10**, 291–298.

Sangster, T.P., Wingate, J.E., Burton, L., Teichert, F., Wilson, I.D. (2007) Investigation of analytical variation in metabonomic analysis using liquid chromatography/mass spectrometry. *Rapid Communications in Mass Spectrometry*, **21**, 2965–2970.

Sarda, S., Page, C., Pickup, K., Schulz-Utermoehl, T., Wilson, I.D., (2012). Diclofenac metabolism in the mouse: Novel in vivo metabolites identified by high performance liquid chromatography coupled to linear ion trap mass spectrometry. *Xenobiotica*, **42**, 179-194.

Small, R.E., (1989). Diclofenac sodium. *Clinical Pharmacology*, **8**, 545–558.

Sims, P. and Grover, P.L., (1976). Epoxides in polycyclic aromatic hydrocarbon metabolism and carcinogenesis. *Advances in Cancer Research*, **20**, 166-262.

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

Sun, J., Schnackenberg, L.K., Beger, R.D., (2009). Studies of acetaminophen and metabolites in urine and their correlations with toxicity using metabolomics. *Drug Metabolism Letters*, **3**, 130–136.

Tang, W., Stearns, R.A., Bandiera, S.M., Zhang, Y., Raab, C., Braun, M.P., Dean, D.C., Pang, J., Leung, K.H., Doss, G.A., Strauss, J.R., Kwei, G.Y., Rushmore, T.H., Chiu, S.H., Baillie, T.A., (1999). Studies on cytochrome P-450-mediated bioactivation of diclofenac in rats and in human hepatocytes: identification of glutathione conjugated metabolites. *Drug Metabolism and Disposition*, **27**, 365–372.

Tautenhahn, R., Patti, G.J., Rinehart, D. Siuzdak, G. (2012). XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Analytical Chemistry*, **84**, 5035–5039.

Veselkov, K.A., Vingara, L.K., Masson, P., Robinette, S.L., Want, E., Li, J.V., Barton, R.H., Boursier-Neyret, C., Walther, B., Ebbels, T.M., Pelczer, I., Holmes, E., Lindon, J.C., Nicholson, J.K., (2011). Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Analytical chemistry*, **83**, 5864-5872.

Wang, S.Y., Kuo, C.H., Tseng, Y.J., (2012). Batch Normalizer: A Fast Total Abundance Regression Calibration Method to Simultaneously Adjust Batch and Injection Order Effects in Liquid Chromatography/Time-of-Flight Mass Spectrometry-Based Metabolomics Data and Comparison with Current Calibration Methods. *Analytical Chemistry*, **85**, 1037–1046.

Want, E.J., Wilson, I.D., Gika, H., Theodoridis, G., Plumb, R.S., Shockcor, J., Holmes, E., Nicholson, J.K. (2010). Global metabolic profiling procedures for urine using UPLC-MS. *Nature Protocols*, **5**, 1005-1018.

Williams, R.T., (1947). Detoxication mechanisms. The metabolism of drugs and allied organic compounds. *Chapman & Hall, Ltd.*, London, 288-290.

Wilson, I.D., and Nicholson, J.K., (2003). Do metabolic pathways for xenobiotics really exist? *Xenobiotica*, **33**, 887-901.

Wood, J.L., (1970). Biochemistry of mercapturic acid formation In Metabolic conjugation and metabolic hydrolysis. *New York Academic Press*, Volume II, 261-299.

Development of a Rapid Microbore UPLC-MS Method for High-Throughput Metabolic Phenotyping

Chapter Three

Previous large scale UPLC-MS metabolic profiling work has been hindered by the analytical stability of the MS platform. A rapid gradient microbore UPLC-MS method was developed to provide a high-throughput solution for metabolic phenotyping of large sample collections and employed for the analysis of urine. Comparison showed that traditional UPLC-MS and rapid gradient approaches provided peak capacities of 150 and 50 respectively, with a reduction of detected features from 19,000 to 6000 with similar levels of repeatability seen for both methods. Despite the reduced peak capacity and the reduction in ions detected, the fast gradient method was able to achieve similar levels of group discrimination as conventional UPLC-MS when applied to two different sets of rodent urine samples.

SUMMARY:

- A major limitation of the current practice in UPLC-MS metabolic profiling studies is the inherent lack of analytical stability of the platform.
- Previous work detailed in chapter 2 identified the difficulties in running studies in batches, and the inability of satisfactory batch correction to be achieved.
- A rapid gradient microbore UPLC-MS method has been developed to provide a high-throughput analytical platform for metabolic phenotyping of large scale studies.
- The approach was based on miniaturizing the conventional reversed-phase UPLC-MS method for urinary profiling from a 2.1 x 100 mm column to a 1 x 50 mm column and reducing run time from 15 min to 2.5 min per sample.
- Comparison of convention and microbore methods showed that the rapid gradient approach resulted in a reduction of detected features from 19,000 to 6000 with similar levels of repeatability.
- However, despite the reduced peak capacity and the reduction in ions detected, the fast gradient method was able to achieve similar levels of group discrimination as conventional UPLC-MS when applied to both the previously discussed acetaminophen cohort, and a 2-bromophenol study in the rat.
- In conclusion, the reduced analysis time in addition to the retained sensitivity, indicate that the proposed methodology is highly suitable for high-throughput metabonomic analysis of complex mixtures such as urine.

3.1 Introduction

Metabonomics as a field of investigative science has progressed rapidly since its inception. As alluded to in chapter one, metabolic phenotyping is generally performed in either a targeted manner, when a marker of interest is already known, or in an exploratory, hypothesis free, untargeted approach where the metabolic profiles of two groups of interest (such as a traditional case-control) are compared. Untargeted metabolic phenotyping has been frequently applied in studies of toxicology and epidemiology, and in the case of the latter are therefore being applied to increasingly larger cohorts, with thousands of samples. The analytical techniques employed, as discussed previously are mainly ^1H NMR spectroscopy and LC-MS, with ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) having become the most used hyphenated mass spectrometry technique (Want *et al*, 2010).

One of the key goals of untargeted profiling is to have global metabolite coverage coupled with high resolution and sensitivity. However, metabolites from biological samples cover a wide range of polarities, rendering isocratic chromatographic solvent systems (where the composition of the mobile phase remains constant throughout the run) unsuitable for metabolic profiling. A gradient system is therefore used, where the elutropic strength of the solvent is increased as the analysis proceeds, allowing for compounds of different polarities to elute at different times. The chromatographic resolution of a UPLC run is therefore somewhat dependent on the solvent gradient used; in general the longer the gradient the higher the resolution with more ions detected as, during short runs, ion suppression and co-elution become more apparent. While some studies which attempt complete metabolome coverage have used gradient lengths of over 30 minutes (Guy *et al.*, 2008) there is a balance to be considered between resolution and throughput. Generally a widely accepted compromise seems to be a chromatographic run time of around 10 to 12 minutes with a short re-equilibration period of two to three minutes, providing good metabolite separation

combined with moderate throughput, with approximately 100 samples possible to be analysed per day.

In any metabotyping study a major consideration is the analytical stability of the mass spectrometer. Generally during a mass spec run there is, as was shown in the studies reported in Chapter 2, a decrease in sensitivity as contaminants build up on the ESI source, and after a certain point the data retrieved will become unreliable, and difficult to compare to previous data. At this point the column, the mass spectrometer inlet and the ESI source should be carefully cleaned to ensure continued good performance. The level and rate of contamination depend heavily on the type of sample used, as some biofluids, such as serum, which are highly concentrated, can cause significant loss of performance if injected undiluted, and without the prior removal of proteins. For most conventional UPLC-MS experiments, using a 15 min run time/sample, generally approximately 100-250 rodent urine samples may be analysed in a single run, though this number falls to *ca.* 100-120 for serum samples (Michopoulos *et al.*, 2009). Nonetheless for any experimental set up there will be a 'hard' limit after which continued analysis is impractical and must be interrupted for source cleaning. Run length and therefore sample number limitations persist and constitute an existential problem for the continuation of LC-MS based metabonomic experiments for large sample sets. Various techniques have been previously employed to attempt to overcome the problem. The application of nano-electrospray ionization (NESI) has indeed considerably extended the usability of the typical MS setup, as due to the reduced volume of sample injected less material is introduced into the mass spectrometer and therefore less contaminants build up on the source (Karas *et al.*, 2000). Another approach is to attempt to reduce analysis times by completely eliminating the chromatography and directly injecting samples into the mass spectrometer, a process known as direct infusion mass spectrometry (DIMS) (Madalinski *et al.* 2008), and the above two techniques have also been implemented together in nano-electrospray DIMS. The benefits of such methodology include an obvious reduction in analysis time, small sample volume consumption (Draper *et al.*, 2013) and claimed comparable prediction and discrimination capabilities to traditional liquid

chromatography-mass spectrometry (Lin *et al.*, 2010). However, these approaches can still take several minutes per sample, and due to the lack of chromatographic separation suffer from susceptibility to significant ion suppression and an inability to distinguish between isomeric and isobaric compounds (Giavalisco *et al.*, 2009). Furthermore due to the loss of the time dimension provided by a chromatographic separation the acquired data is formed from a single composite and complex spectrum for each sample, with only mass to charge ratio and intensity values, which makes metabolite identification close to impossible (Sumner *et al.*, 2007).

A middle ground between the current 'traditional' UPLC-MS profiling methods and direct infusion would be to use shorter chromatographic analysis times, which alleviate some of the disadvantages of DIMS by retaining the chromatographic separation while accepting the loss of resolution and fewer ions detected. A previous study by Plumb *et al.* (2005) has introduced this concept, and has already showed the utility of such an approach. In their experiment the authors compared a shortened run time of 1.5 minutes with a traditional 10 minute gradient and were able to observe the same class distinction despite the decreased number of features detected.

There is ultimately a need for higher sample throughput within the field of metabolic profiling. In this chapter, the utility of runs with shorter chromatographic analysis times is evaluated as a means of eliminating the need for several batches has been evaluated. In addition the previous methodology was also extended to include microbore liquid chromatography, using a 1mm ID column which has recently been shown to enhance information retention with increased sensitivity but also reduced solvent usage (Gray *et al.*, 2015). This combination of microbore UPLC-MS with a rapid gradient was applied to the both the acetaminophen study described in the previous chapter and to a second COMET study exploring the metabolism and toxicity of 2-bromophenol in the rat. The resulting data were compared with conventional UPLC-MS for both compounds.

3.2 Study Details and Sample Preparation

3.2.1 Acute 2-Bromophenol Study

Urine samples were collected from three groups of male Sprague-Dawley rats (10 rats/group) administered a single intraperitoneal dose of 2-bromophenol at either 100 mg/kg or 200 mg/kg in corn oil. A control group was treated with corn oil alone. Rats were housed in metabolic cages to enable the collection of urine and samples were taken at -16, 0, 8, 24, 48, 72, 96, 120, 144 and 168 hours after administration. The urine samples were stored at -40 °C prior to analysis.

A 120 µL aliquot of urine was mixed with 120 µL water to dilute the salt concentration before protein removal with acetonitrile (1:3 v/v). These samples were vortex mixed and left at -20 °C overnight before centrifugation for 5 min at 15,000 g at 4 °C. For analysis, 50 µL was removed and added to 250 µL water in 350 µL 96-well plates. A pooled quality control (QC) sample was prepared by combining 100 µL of each sample and diluting 1:4 with water. The 96-well plates were stored at -20 °C until analysis and the plates were centrifuged again for 5 min at 700 g before being placed into the autosampler at 4 °C. The plate from this study was analysed five times, as a replicate, to assess the robustness of the approach.

3.2.2 Chronic Acetaminophen Study

Urine samples were collected and stored from four groups of male Sprague-Dawley (10 rats/group) rats dosed orally with acetaminophen (APAP) as described in Chapter 2 (pages 98 to 99)

For the analyses described here a 20 µL aliquot of each urine sample was mixed with 60 µL methanol in order to precipitate proteins. The samples were vortex mixed and left at -20 °C overnight before centrifugation for 5 min at 15,000 g at 4 °C. For analysis, 20 µL was removed and added to 180 µL water in 250 µL 384-well plates. A pooled quality control

(QC) sample, used for system conditioning and assessment of analytical performance was prepared by combining 5 μL of each sample and diluting 1:9 with water. The well plates were centrifuged again for 5 min at 700 g before being placed into the autosampler at 4 °C.

3.2.3 Liquid Chromatography

Liquid chromatographic analysis was performed on an Acquity I-class UPLC system, equipped with a binary solvent manager, sample manager and column heater (Waters Corporation, Milford, MA, USA), interfaced with a Synapt G2-S HDMS mass spectrometer (Waters Corporation, Manchester, UK). The chromatographic separations were performed on a HSS T3 1.8 μm stationary phase of either 2.1 x 100 mm for the conventional analysis or 1 x 50 mm for the fast, microscale analysis. The chromatographic mobile phase was composed of 0.1 % formic acid in water (A) and 0.1 % formic acid in acetonitrile (B). The column temperature was maintained at 40 °C and linear gradient elution was performed at either 0.5 mL/min for the 2.1 x 100 mm column or 0.4 mL/min for the 1 x 50 mm column. For the conventional method the starting composition was 1 % B, held for 1.0 min before increasing to 15 % at 3.0 min, 50 % at 6.0 min, 95 % at 9.0 min for a 1.0 min wash and returning to 1 % B for a 2 min re-equilibration step (total cycle time 12 min). A 2 μL injection of sample was performed using the flow through needle. For the fast, microscale method the starting composition was 1 % B, held for 0.14 min before increasing to 15 % at 0.42 min, 50 % at 0.83 min, 95 % at 1.25 min for a 0.25 min wash and returning to 1 % B for a 1 min re-equilibration step (a total cycle time of 2.5 min). A 1 μL (2-bromophenol study) or 0.2 μL (acetaminophen study) injection of sample was performed using the flow through needle. In both cases the purge solvent was 95:5 (v/v) $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ and sample manager wash was CH_3OH . For each study, prior to the analysis of the samples themselves 50 consecutive injections (2 μL volume) of the pooled samples were made at the start of the chromatographic run to 'condition' the column to ensure that the analytical system was fully equilibrated. The QC sample was injected at the beginning of the analytical run and every 10 injections thereafter to monitor instrument stability.

For use of 1 mm I.D. columns, minor system modifications were necessary to reduce peak dispersion (Gray *et al.*, 2015). For the microbore separations the standard outlet tubing i.d. (0.004") was reduced to 0.0025" of minimum length and the ESI stainless steel capillary (125 μm i.d.) was replaced with a narrow bore variant (50 μm i.d.).

3.2.4 Mass Spectrometry

Mass spectrometry for metabolic profiling was performed on a Synapt G2-S HDMS accurate mass instrument with electrospray ionization (ESI) operated in positive (ESI+) ion mode and a Xevo G2 operated in the same way. For the Synapt G2-S the capillary voltage was 1.0 kV, cone voltage was 25 V, source temperature was set at 120 °C with a cone gas (nitrogen) flow rate of 50 L/h, a desolvation gas temperature of 600 °C and a nebulization gas (nitrogen) flow of 1000 L/h. The Synapt G2-S was operated in resolution (V optics) mode whereas the Xevo G2 was operated in sensitivity (V optics) mode and both detectors were set to acquire data over the m/z range 50-1200 with a scan time of 0.1 s for the conventional method or 0.05 s for the fast method. All mass spectral data were collected in centroid mode using the MS^e data acquisition²¹ function to obtain fragmentation data simultaneously. In function one a low collision energy (4 eV) was used and in the second function a high collision energy (ramp 15-45 eV) was used for fragmentation. For mass accuracy, leucine enkephalin (MW = 555.62) was used as a lock mass at a concentration of 200 pg/ μL (in 50:50 CH₃CN/H₂O, 0.1 % formic acid) infused at a flow rate of 20 $\mu\text{L}/\text{min}$ via a lock spray interface. Lockmass scans were collected every 15 s and averaged over 3 scans to perform mass correction. The instrument was calibrated before analysis with 0.5 mM sodium formate solution. The data were collected using MassLynx V 4.1 software (Waters Corporation, Manchester, UK).

3.2.5 Data Analysis

The raw data were processed by Progenesis QI data analysis software (Non-Linear Dynamics, Newcastle, UK) for peak picking, alignment and normalization, to produce peak intensities for retention time (RT) and m/z data pairs. Further statistical analysis was performed on the resulting normalized peak intensities using SIMCA P 13.0.2 (Umetrics, Umea, Sweden).

3.3 Results and Discussion

3.3.1 2-Bromophenol

a. Quality Control

As discussed in the introduction chapter, LC-MS techniques can be subject to lengthy sample preparation and also fluctuation of instrumental performance leading to acquiring unreliable data. The data for acetaminophen presented in Chapter 2 has highlighted that there is period of analytical instability at the start of every run, characterized by initial changes in chromatographic retention time and both short and longer term effects on signal intensity which, in extreme cases, can lead to unusable data. This is a common observation and to overcome this issue the use of conditioning injections, repeated injections of matrix, to modify the system and therefore stabilize retention times and signal intensities prior to the start of an analysis, has become commonplace (Gika *et al.*, 2007). Other studies have shown that an increase in the volume of the conditioning injections coupled with a rapid gradient is an efficient way of decreasing the time required to achieve instrumental stability and increased reproducibility (Michopoulos *et al.*, 2009). A similar approach was adopted here as the use of such a rapid gradient was well suited to the use of a large number of conditioning injections without resulting in a significant increase in the required analysis time. In this experiment 50 conditioning injections were performed prior to the start of the sample analysis, which required less than two hours of analysis time. Furthermore to enhance conditioning the injection volume was doubled to 2 μ L to give the equivalent of 100 conditioning QC injections.

As well as their use for conditioning the pooled quality control samples were also used to monitor the performance of the method throughout the run, and in order to do this a QC

sample was analysed at the start and then every ten sample injections thereafter. Initial visual inspection of the QC chromatograms showed that there was good consistency throughout the whole run, with the chromatograms of the QC injections overlaying perfectly. Figure 3.1 depicts the last QC injection from each of the five replicate runs a good indication of the instrumental stability throughout the approximately 500 injections made over the course of the analysis.

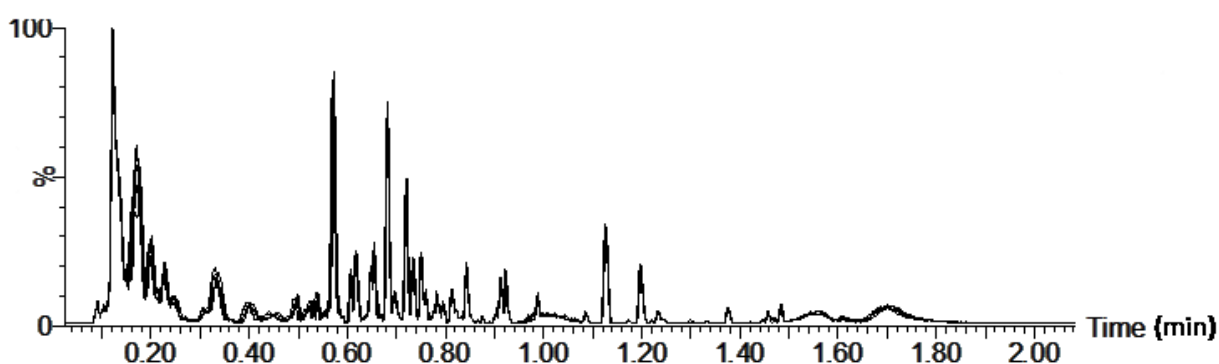


Figure 3.1 Overlaid chromatograms of the last QC injection of each of the 5 runs. The chromatograms align well suggesting that there was a high level of stability throughout the analysis.

The stability of the approach was mirrored in the initial PCA scores plots used to scrutinize analytical drift. Figure 3.2 shows the initial un-filtered data for the negative electrospray ionisation (ESI) runs; the red points in the scores plot represent the QC samples. In the scores plot we can see that the QC samples cluster together in the center of statistical space, with some evidence for a small amount of time-related drift, the direction of which seems to proceed along the first component. This seems to indicate that there was a slight deviation in analytical stability as the experiment progressed, however since the deviation was not significant in the overall variation within the data, and as the sample data were not greatly affected, the data were accepted as suitable for further evaluation.

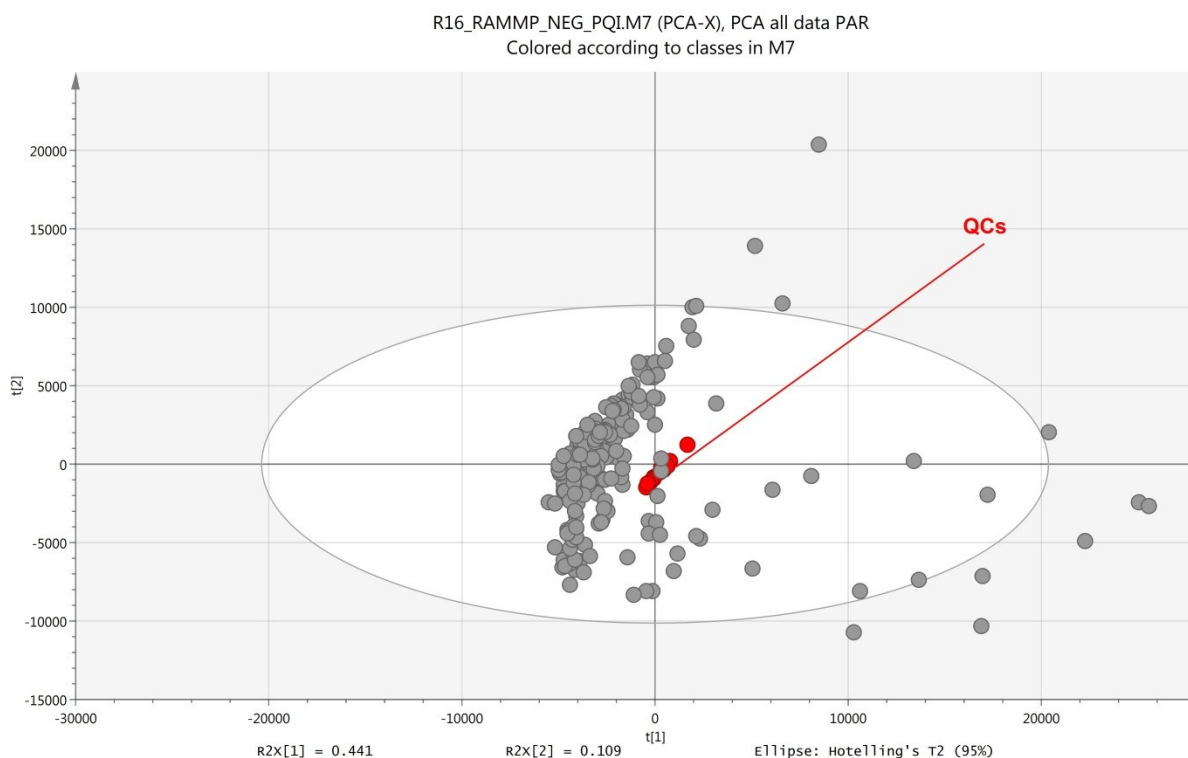


Figure 3.2. PCA scores plot for the un-filtered data for the negative electrospray ionization RAMMP run. The red points in the scores plot represent the QC samples while the gray points indicate the samples. The QCs cluster in the centre of statistical space, with a slight time dependent drift along the first component.

The reproducibility of the methodology is further demonstrated by the consistency of the acquired data between the replicate runs, as shown in figures 3.3 and 3.4. Figure 3.3 which depict the scale adjusted variation exhibited by the five replicates for each of a randomly selected subset of variables (detected features) shown as a line graph. As the variation does not intersect the first standard deviation threshold it can be surmised that there was little variation between replicate analyses during the run. Figure 3.4 is a PCA scores plot showing the statistical positioning of a random subset of control samples from the urine analysis. As each one of the five replicates for each sample clusters closely together, this adds further confidence to the repeatability of the experiment.

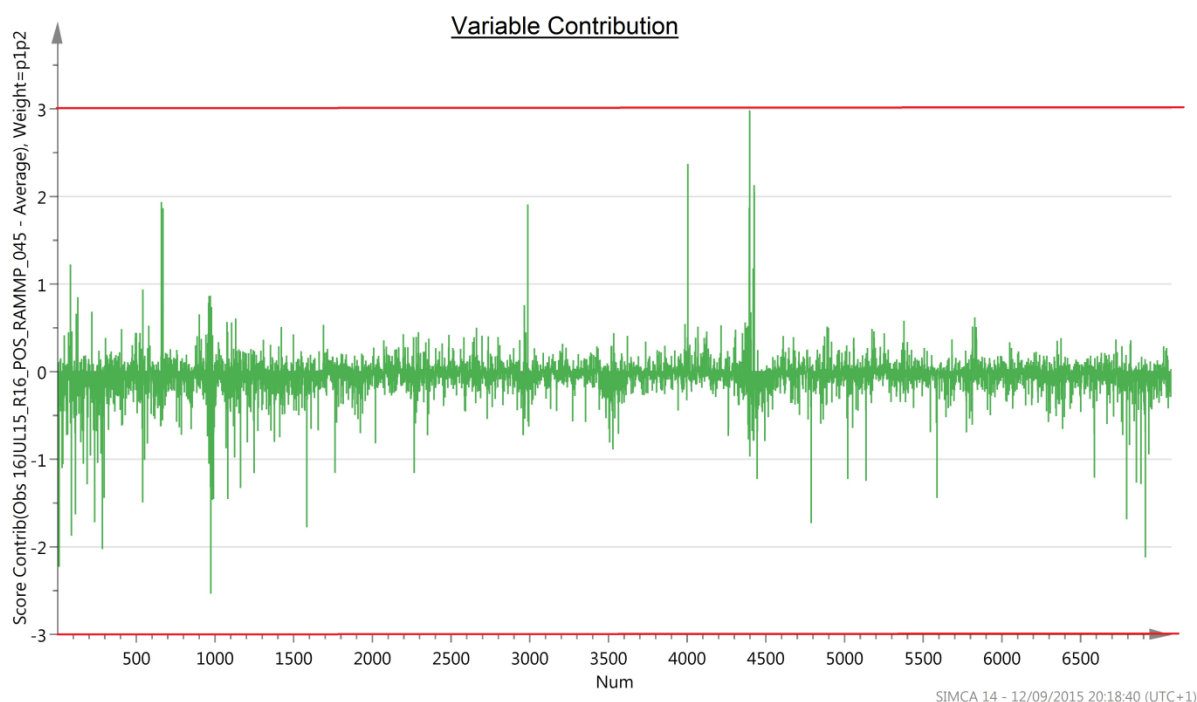


Figure 3.3. Contribution plot showing the deviation of the QC samples for each ion from the statistical mean of the principal component analysis expressed in standard deviations. The plot displays the vector $v_k(x_{ik} - avg_{ik})$ for each variable as a bar, the orientation of which indicates the direction of the deviation. As none of the bars intersect the standard deviation threshold we can be confident in the analytical conditions were stable throughout the run.

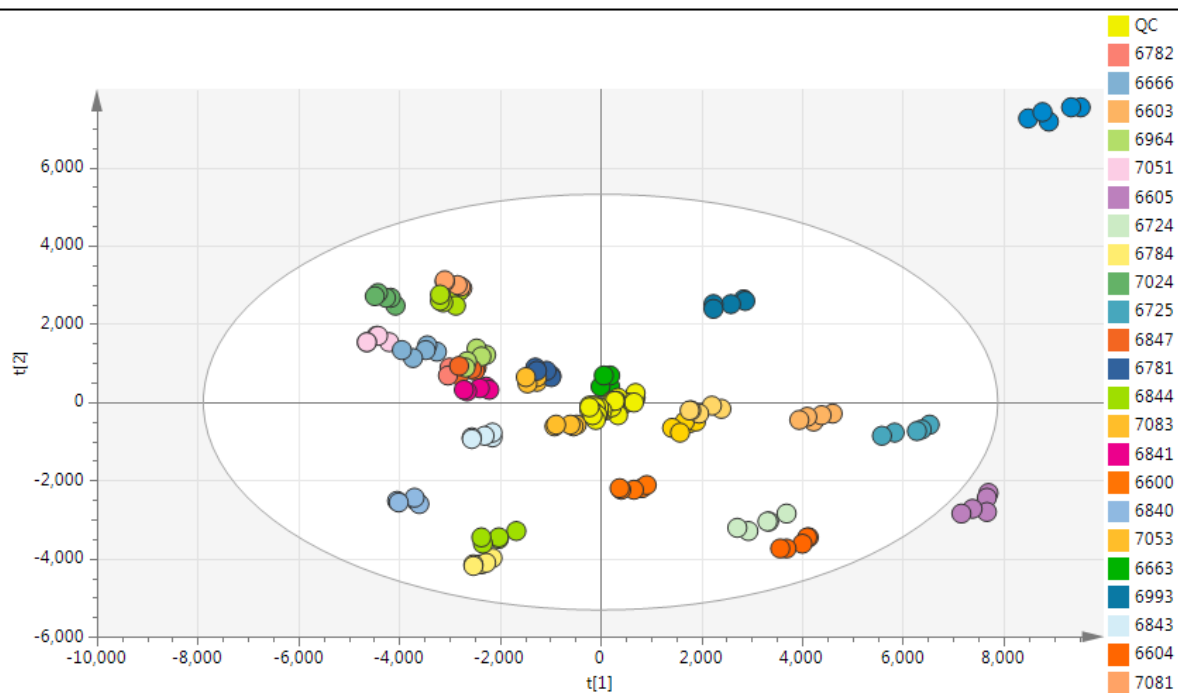


Figure 3.4. PCA scores plot depicting the five replicates of a random subset of control samples from the 2-bromophenol urine rapid gradient study, together with Quality control samples (yellow, in the centre). The clustering of each of the replicates within statistical space suggests high levels of analytical reproducibility.

Another method of determining data reliability is through the use of the coefficient of variance. By computing this value for the peak height for each of the features detected of in the QC samples (which should contain all compounds/features present in the sample set as they are pooled samples) throughout the analysis should give an indication of how robust detection of that particular metabolite/feature is within the analysis, and therefore how confident we can be of that metabolite as a biomarker. The coefficient of variation (CV) expressed as a percentage is defined as:

$$CV = \frac{\sigma}{\mu} * 100$$

Where σ is the standard deviation and μ is the expected return, in this case the mean. There is no particular value which would constitute a threshold for acceptance, as the criteria for acceptance within hypothesis free metabolome mapping have yet to be standardized, however values of a CV filter ranging between 15% and 30% are generally used depending on how rigorous an analysis needs to be. Table 3.1 shows the number and percentage of ions passing the various cut off CV thresholds for the negative ESI subset study.

Negative	CV less than		
	<15%	<20%	<30%
Number of ions	3059	3703	5517
Percentage	55.2%	67.2%	82%

Table 3.1 Table showing the number and percentage of ions satisfying the various CV threshold levels.

Generally ion intensity stability is greater when the peak intensity is greater, therefore low intensity peaks or those suffering from ion suppression will have higher levels of variation. As the rapid microbore methodology has a steeper and faster gradient, there is an increased

likelihood of co-elution of metabolites in addition to overlap and lower resolution between peaks, which should result in more pronounced ion suppression effects. Thus it was expected that there would be smaller percentage of ions at each CV cut off in this fast analysis methodology than in conventional UPLC-MS. In figure 3.5 the percentage of ions passing each CV threshold is compared using a bar chart. Examination of the figure shows that there is a small reduction in repeatability, with 89% of total features detected by traditional 2.1mm ID UPLC-MS passing the 30% CV threshold as opposed to 82% from the rapid analysis method, and similarly with the lower CV criteria. However, this reduction in stability is small compared to the massive throughput gain achieved.

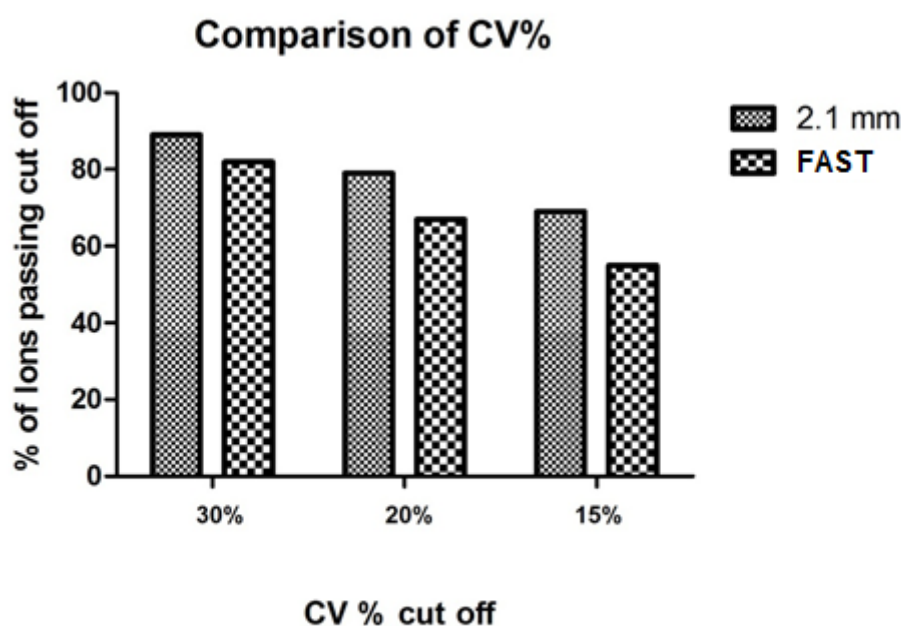


Figure 3.5. Bar chart comparing the percentage of ions passing the acceptance criteria for each of the cut off levels for both the 2.1 ID UPLC-MS and the rapid gradient method. There is a consistent increase of variability in the rapid profiling method (less ions pass the cut off) throughout all the cut off thresholds.

b. Comparison with traditional 2.1mm UPLC-MS profiling method

As we have previously alluded to, the reduction of chromatographic separation time is expected to result in the loss of some detected features, and therefore a lower metabolic

coverage scope. To ascertain the level of information loss, the a subset of the 2-bromophenol sample set was also analysed with traditional UPLC-MS methodology with a 2.1mm ID column and a gradient time of 12 minutes. Initial comparison of the negative ESI data showed that there was indeed a lower number of ions detected, with the rapid analysis only detecting 6,188 features, whereas the conventional mass spectrometry detected 18,823; a nominally significant loss. The positive ESI mode runs showed a similar loss of information, with 6,129 ions detected with the rapid method, compared to the 17,816 features detected by traditional mass spectrometry. Note that comparisons of the full negative and positive ESI 2-bromophenol datasets for the conventional UPLC-MS method will be included in the main 2-bromophenol metabolism chapter, Chapter 4.

However, despite the loss of detected ions, when PCA analysis was performed on the negative ESI data acquired for the subset of samples a similar degree of group separation and statistical prediction was observed between the rapid microbore analysis and the conventional analysis. Figure 3.5 shows the comparison between the PCA scores plots of the conventional UPLC-MS method and the rapid gradient method for the initial pilot study, while figure 3.6 shows the sample comparison for the full R16 2-bromophenol data sets, in negative ESI.

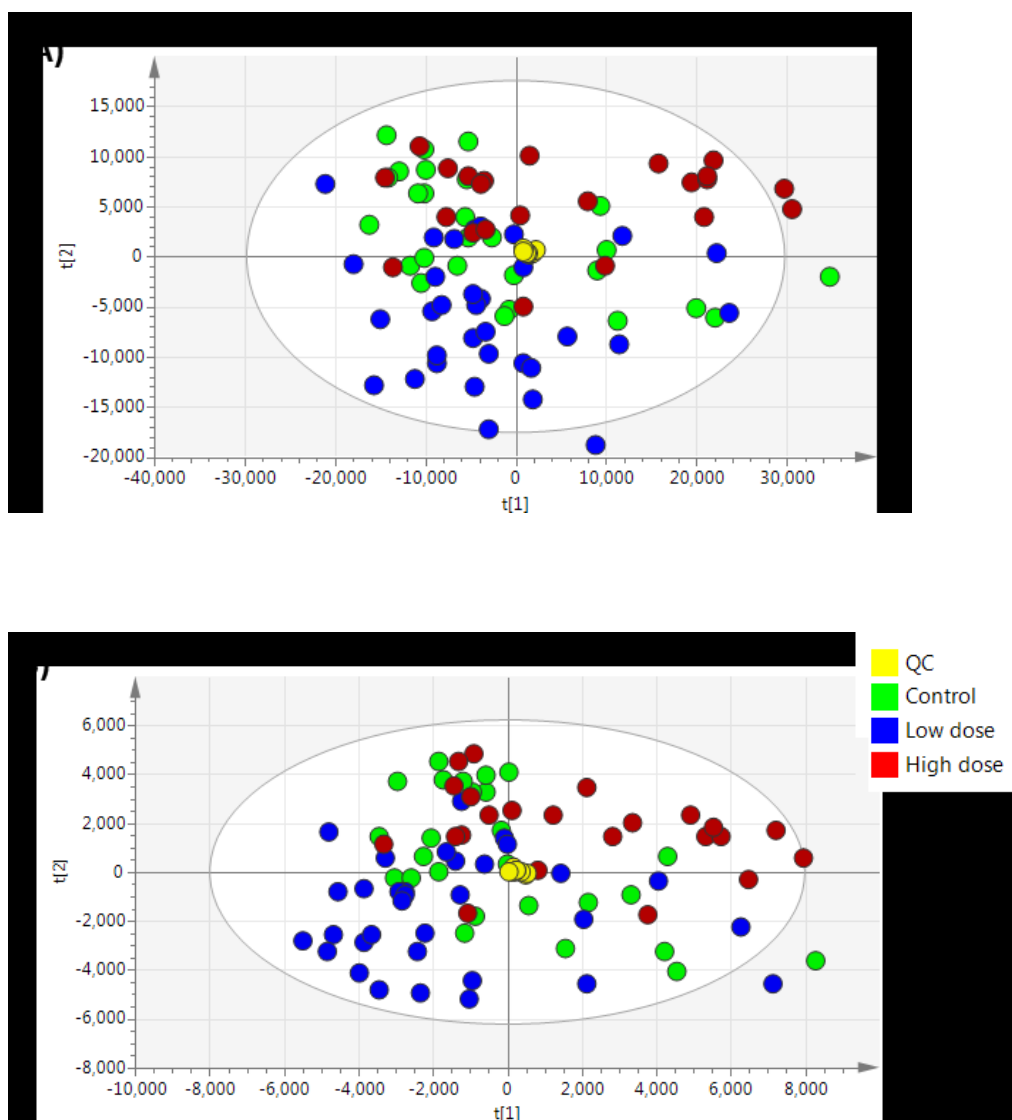


Figure 3.6. PCA scores plot depicting the same sub-set of rat urine samples taken from the R16 2-bromophenol toxicity study analysed either via the conventional 2.1mm ID method A) or the rapid gradient 1 mm ID. method B). Similar trends are apparent in both PCA models, with similar data distribution along statistical space, despite the rapid gradient data containing less than half the number of variables.

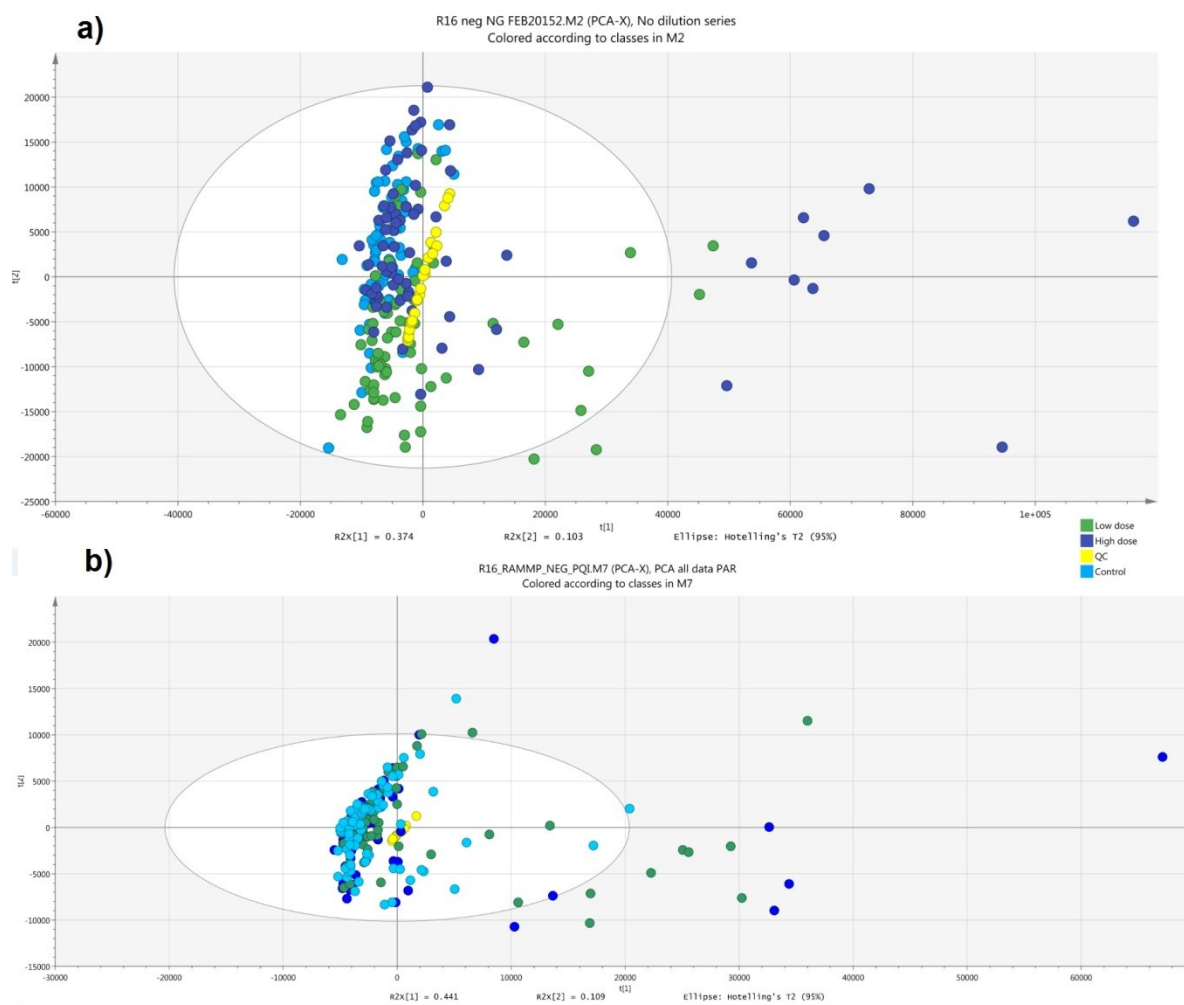


Figure 3.7. PCA scores plot depicting the complete set of rat urine samples taken from the R16 2-bromophenol toxicity study analysed either via the conventional 2.1mm ID method a) or the rapid gradient 1 mm ID. method b) in negative ionisation. Similar trends are apparent in both PCA models, with similar data distribution along statistical space, despite the rapid analysis method data containing less than half the number of variables. Note that there seems to be a reduced time related drift (as shown by the QCs) in the rapid gradient method data indicating a further potential advantage of the methodology.

In both figures almost identical trends for the data can be observed within the model, giving similar levels of discrimination between the test groups. It can therefore be concluded that, despite the lower number of variables within the model, the variables contributing most to the separation are still being observed. This allows similar conclusions to be drawn from either methodology with a similar patterns in the distribution of the three groups of animals (control, low dose (100 mg/kg) and high dose (200 mg/kg)) and the quality control (QC) observed.

The QCs in all analyses cluster around the center of the scores plots, indicating a good level of analytical consistency. However, in figure 3.6 we can see that there is a reduced contribution in the fast method of the time related drift observed for the conventional 2.1 ID UPLC-MS method. This should be self-evident seeing as the total time of analysis in the fast method is greatly reduced, but would also suggest that the rapid analysis approach could be less susceptible to time dependent degradation of instrumental sensitivity.

c. Metabolite detection

In the previous chapter, the potential for statistical group discrimination to occur due to the influence of variables not related to the administered drug has been explored. As such, while our fast method had produced similar statistical distributions between dose groups in the PCA models, it is necessary to investigate whether these results are driven by the same metabolic signals. It is, for example, possible that the rapid analysis methodology had an inherent bias towards certain ions, meaning that the nature of the metabolites highlighted as markers would differ. To ascertain whether this was the case, orthogonal partial least square regression models were constructed, for both conventional and rapid gradient analysis data sets. By utilizing the model residuals it is possible to determine the contribution that each variable makes towards the discrimination. The OPLS S-plot shown (from a model comparing the high dose with the control group) below highlights the fact that the major contributing ions are the same in both cases, lending further support for the equivalence of the conventional and rapid gradient methodology. It must be noted, that the predictive residual value ($p[1]$) of the fast 1mm ID method were slightly lower than conventional LC-MS. However, while producing a slightly less predictive model, the same ions were identified, and the individual p-values were still significant.

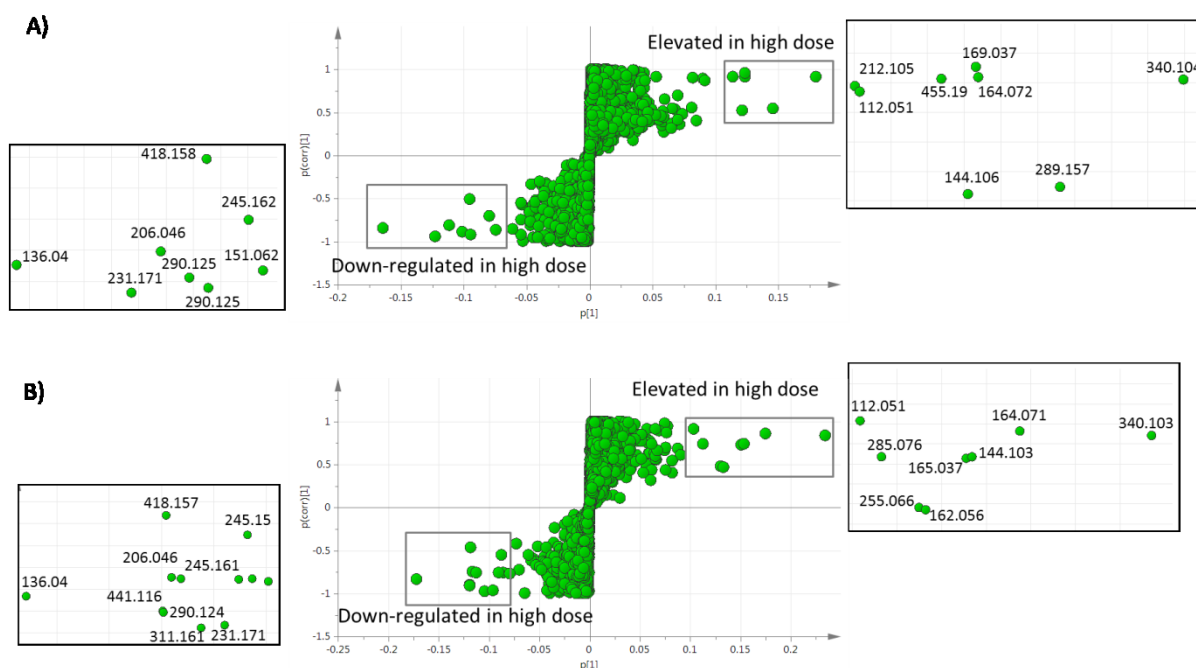


Figure 3.8. OPLS-DA S plot of models constructed from data of samples taken from the R16 2-bromophenol toxicity study analysed either via the conventional 2.1mm ID method A) or the rapid gradient 1 mm ID method B). The same discriminating variables are identified by both methods, highlighted in the zoomed in regions.

3.3.2 Acetaminophen

a. Quality control

The increasingly large cohorts being employed in metabonomic studies have shown how analytical time for LC-MS methods can be very long, and can easily reach several days if not weeks. We have already discussed the inherent limitations of the analytical platform, and how there is a critical threshold of instrument stability. The interruption of runs to allow for cleaning and other maintenance procedures inevitably requires the separation of experimental samples into batches. However, the LC-MS platform is characteristically unstable, and is susceptible to time-related analytical drift and ion contamination. This is exacerbated by the process of source cleaning and the replacement of eluent with new batches, which have small micro- variations in composition. These can culminate in

significant fluctuations in retention times, mass accuracy and signal intensity. The data obtained for each batch, while valid within itself as any bias is uniform, cannot be integrated or compared with the data from other batches.

In the previous chapter an attempt was made to analyse an acetaminophen toxicology study in the rat where urine samples had been collected over seven days (20 time points) from 40 rats, in 3 dose and one control group, was described. This study had resulted in 680 samples which were randomized and divided into 4 batches. The four batches were analysed using the convention UPLC-MS method and run by traditional LC-MS on consecutive days, interspaced by source cleaning and instrument maintenance. The resultant data was highly inconsistent across batches and was characterized by large variations in instrument response within and between runs, as well as some differences in retention times. Attempts to use statistical and computational methods to compensate for the observed batch effects using the QCs as a point of normalization were made, but met with limited success, although other groups have reported greater success using this approach (e.g., Kirwan *et al.*, 2013). The extent of the drift is illustrated in the PCA scores plots for QC quality control (figure 3.9).

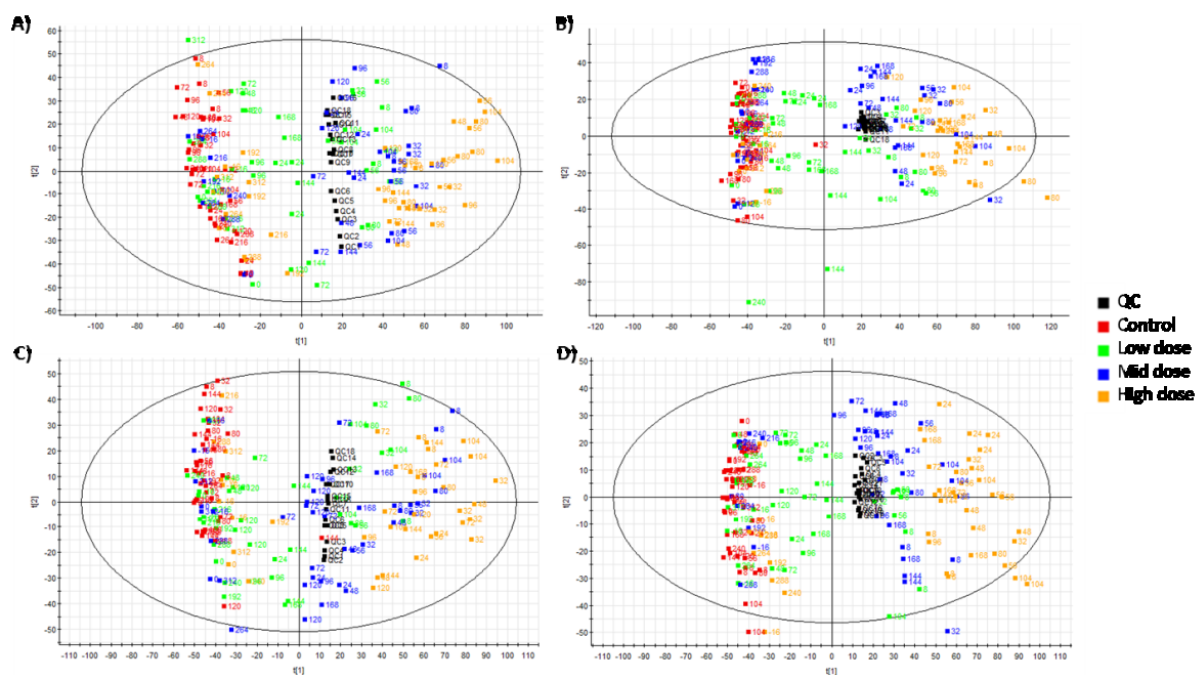


Figure 3.9. PCA scores plots showing the initial unfiltered data colored by dose group for each batch (a) batch one, (b) batch two, (c) batch three and (d) batch four. This data is from conventional 2.1mm ID LC-MS. The black points are the QC samples for that particular batch. It can be seen that there was significant time-related drift in batches one and three, with batch four being more acceptable and batch two being the most consistent. However, computational methods were not able to correct for batch differences even between the two most consistent batches, two and four.

One of the aims of the application of the fast 1mm ID column method was to evaluate if the increase in sample throughput could eliminate the need for batches in LC-MS experiments. Indeed the 680 samples of this acetaminophen study were prepared in two 385 micro-well plates, and were analyzed as a single batch, with a run time for each sample of 2.5 minutes, resulting in a total analysis time of just over 28 hours, a fraction of the time required for conventional methods. The same column matrix conditioning strategy was employed as in the 2-bromophenol experiment. Initial visual inspection of the QC chromatograms showed that there was a consistency throughout the whole run, with the chromatograms of the QC injections overlaying perfectly. Figure 3.10 depicts every 5th QC injection spaced approximately 90 samples apart throughout the run, a good indication of the instrumental stability throughout the approximately 700 injections made over the course of the analysis. The chromatograms overlay almost perfectly, suggesting that the analytical conditions were

stable throughout the run. The final QC does, however, show some deviation in one of the major peaks, however, potentially indicating the limit of stability of the method.

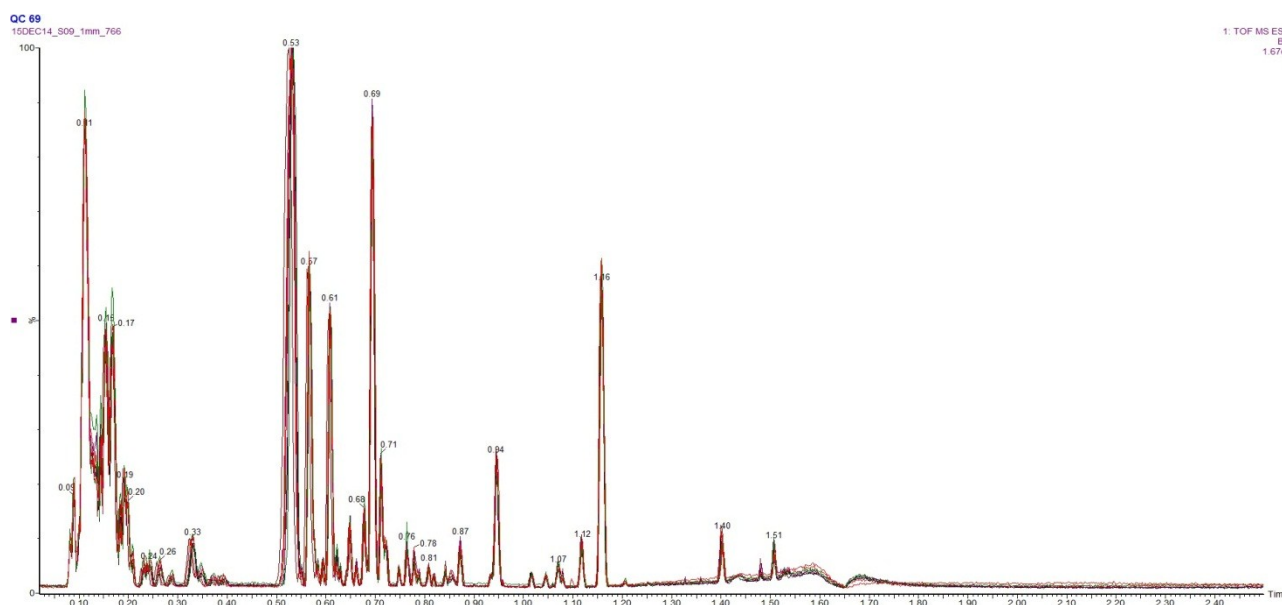


Figure 3.10. Overlaid chromatograms of ten selected QCs, spaced approximately 90 injections apart. The chromatographic peaks show remarkable consistency in retention times, suggesting low levels of drift. The final QC (injection 766) does show a slight deviation in the retention of the major peak at approximately 0.55mins, however this drift is small and seems to not have affected the rest of the run.

Inspection of the unfiltered PCA scores plot for the dataset (figure 3.11) also revealed that the QC samples exhibited low levels of variation, and grouped together tightly in the center of the model hyperspace. Generally the QC samples suggested a high level of analytical quality and consistency with our employed methods. It is possible that the rigorous column conditioning has significantly reduced analytical variability, and is therefore highly recommended for any rapid gradient liquid chromatography method.

Initial inspection of the PCA scores plots revealed that the data showed clear separation between the dose groups, with an obvious correlation between response and dose level. Maximum response for all groups was at 5 to 6 days post initial dose. A subset of animals in the initial study were sacrificed early, while another subset were retained for 7 days after dosing had ceased to identify any biomarkers of recovery. Indeed the metabolic profiles of the recovery animals gradually returned to normal and, by the end of the study, seemed to be very similar to control samples and pre dose status. Figure 3.12 illustrates the good statistical separation achieved by PCA modeling of the data. These PCA scores plots match the trends seen within the conventional LC-MS data.

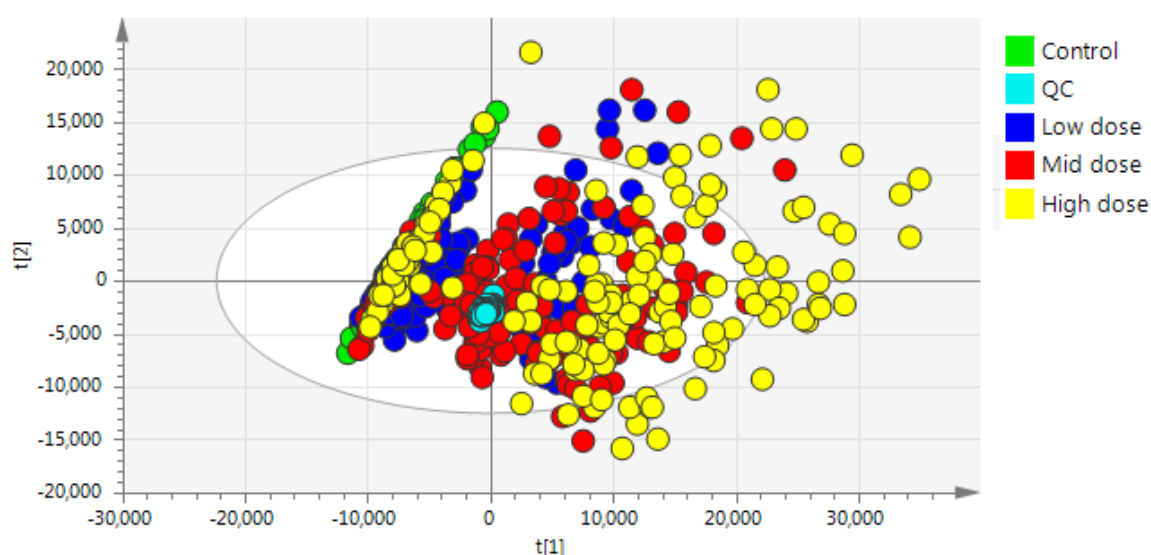


Figure 3.11 PCA scores plot for the positive electrospray ionization rapid analysis run of the acetaminophen study with all time points included, colored according to dose level. Maximal response is in the high dose (800 mg/kg) group at 5 days post first dose. The metabolic profiles of animals in the recovery group return to metabolic equilibrium after the end of dosing.

With the reduction in chromatographic separation time, a loss of resolution is expected with the rapid analysis approach. The major metabolites of acetaminophen have already been discussed in chapter 2, but will be briefly summarized again. The majority of acetaminophen is metabolized by phase-II conjugative reactions, with acetaminophen-glucuronide and

acetaminophen-sulphate the two most prevalent metabolites. Hepatic glutathione conjugation of other reactive intermediates is also apparent, and the products of these conjugations are further metabolized downstream into N-acetyl-L-cysteinyl- and cysteinylacetaminophen conjugates (Nelson, 1982). For the employed methodology to be robust enough to be routinely used in metabolic and toxicological studies, the above well-known metabolites of acetaminophen should be detected in addition to the drug itself. Extraction of the known molecular masses from the chromatograms revealed that the parent compound (m/z 152.0712), and major metabolites APAP-glucuronide (m/z 328.1032), APAP-sulfate (m/z 232.0280) and APAP-N-acetylcysteinyl (m/z 313.0858) were all well resolved. The figure below shows the peaks for acetaminophen and the major sulphate and glucuronide metabolites detected along with their corresponding mass spectra (figure 3.12).

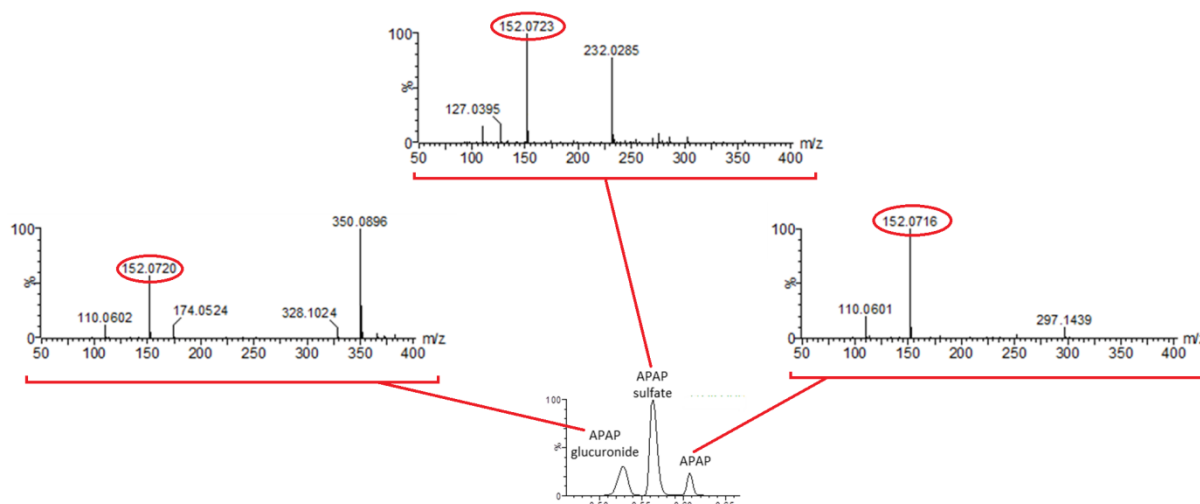


Figure 3.12 The chromatographic peaks (bottom center) and related mass spectra of the major metabolites of acetaminophen detected by use of the HSS T3 1 × 50 mm 1.8 μ m column. The data are from an animal in the high dose (800 mg/kg) group. The major metabolites identified were APAP glucuronide and APAP sulfate. APAP glucuronide was present mostly as the sodium adduct, as seen in the mass spectrum (m/z 350.08). Parent compound was detected, and contributed the common ion (m/z 152.07) observed in all metabolites (circled red).

Construction of OPLS models comparing high dose and control groups at 24 hours post initial dose time point provided very strong models ($R^2 = 0.89$, $Q^2 = 0.63$), which were used to identify the signals driving the separation, and determine if these signals were acetaminophen-related metabolites. The S-plot produced from the model loadings shows that the most significant ions influencing the group discrimination are indeed the major acetaminophen metabolites APAP-glucuronide, APAP-sulfate and APAP-NAC, in combination with the presence of the parent compound. This is highlighted in the below figure 3.13. It is interesting to note that the statistics are dominated by the major metabolites, and while there may be effects on the endogenous metabolism it is difficult to disentangle those signals from the background noise. This relates to the earlier diclofenac study (chapter 2A) where, in contrast, the endogenous metabolites effectively masked any drug related products.

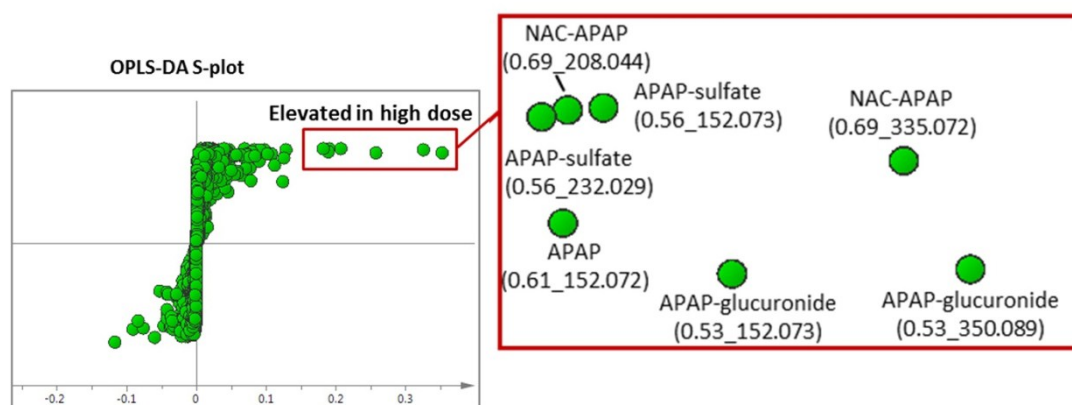


Figure 3.13 The S-plot of the loadings from the OPLS-DA model comparing high dose (800 mg/kg) and control groups at 24 hours post first dose. The zoomed in region at the top is to illustrate the important variables. The ions relate to the expected major acetaminophen metabolites and the parent compound.

3.4 Conclusions

Metabonomics is now an established field of scientific research and as such studies will continue to grow in scale and scope, with large scale biomarker discovery, patient stratification and epidemiological applications. Methodologies employed within metabolic fingerprinting will need to evolve to be able to cope with the increased sample numbers and the need for higher throughput. As one of the primary analytical technologies, the reproducibility, affordability and efficiency of the UPLC-MS platform must also be addressed. The novel approach described here using a rapid chromatographic gradient and a lower ID separation column, combined with fast scanning high resolution accurate mass spectrometry, attempts to address some of the future needs within metabolic profiling. The rapid method achieved analysis times comparable to direct infusion, without the loss of the chromatographic dimension, allowing for separation of analytes with the same molecular weight, and with reduced risk of ion suppression effects. Indeed, without the separation provided by the rapid gradient chromatography it would not have been possible to relate the key common ion 152.07 to the presence of unchanged acetaminophen in the samples using DIMS. The use of smaller diameter UPLC columns has the added benefit of requiring lower volumes of both sample and eluents, which both lowers environmental impact but also enables comprehensive analysis even with limited sample availability.

The most obvious limitation of the approach is the reduction of resolution and incomplete metabolome coverage, which could result in the loss of critical endogenous metabolic signatures, missing potential biomarkers or other more subtle effects. However, this approach is a good for quick interrogation of the samples, and is sensitive enough to detect stratification within the cohort. This should give an indication of the suitability of the study and could pave the way for further, more in depth, possibly targeted, profiling. Nonetheless the key consideration is the ability for the rapid analysis approach to reduce analysis time to the extent that batch analysis is no longer necessary, which precludes the need for copious data manipulation for batch effect correction. In conclusion the results of the study presented

here suggest the rapid analysis methodology is well suited to the analysis of large numbers of urine samples. This approach is expected to be implemented in future work.

3.5 References

Draper, J., Lloyd, A. J., Goodacre, R., Beckmann, M., (2013). *Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: A review. Metabolomics*, **9**, 4–29.

Giavalisco, P., Kohl, K., Hummel, J., Seiwert, B., Willmitzer, L., (2009). ^{13}C isotope-labelled metabolomes allowing for improved compound annotation and relative quantification in liquid chromatography-mass spectrometry-based metabolomic research. *Analytical Chemistry*, **81**, 6546–6551.

Gika, H.G., Theodoridis, G.A., Wingate, J.E., Wilson, I.D., (2007). Within day reproducibility of an HPLC-MS-based method for metabonomic analysis: application to human urine. *Journal of Proteome Research*, **6**, 3291–3303.

Gray, N., Lewis, M.R., Plumb, R.S., Wilson, I.D., Nicholson, J.K., (2015). High-Throughput Microbore UPLC–MS Metabolic Phenotyping of Urine for Large-Scale Epidemiology Studies. *Journal of Proteome Research*, **14**, 2714–2721.

Guy, P.A., Tavazzi, I., Bruce, S.J., Ramadan, Z., Kochhar, S., (2008). Global metabolic profiling analysis on human urine by UPLC-TOFMS: issues and method validation in nutritional metabolomics. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences*. **871**, 253–260.

Karas, M., Bahr, U., Dulcks, T., (2000). Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius' Journal of Analytical Chemistry*, **366**, 669-676.

Kirwan, J.A., Broadhurst, D.I., Davidson, R.L., Viant, M.R., (2013). Characterising and correcting batch variation in an automated direct infusion mass spectrometry (DIMS) metabolomics workflow. *Analytical and Bioanalytical Chemistry*, **406**, 6075-6086.

Lin, L., Quan, Y., Xiaomei Y., Wei H., Jiaxin Z., Jinchun X., Benli H., (2010). *Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer. Analyst*, **135**, 2970–2978.

Madalinski, G., Godat, E., Alves, S., Lesage, D., Genin, E., Levi, P., Labarre, J., Tabet, J.C., Ezan, E., Junot, C., (2008). Direct introduction of biological samples into a LTQ-Orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Analytical Chemistry*, **80**, 3291-3303.

Michopoulos, F., Lai, L., Gika, H., Theodoridis, G., Wilson, I.D., (2009). UPLC-MS-based analysis of human plasma for metabonomics using solvent precipitation or solid phase extraction. *Journal of Proteome Research*, **8**, 2114-2121.

Plumb, R.S., Granger, J.H., Stumpf, L.C., Johnson, K.A., Smith, B.W., Gaultitz, S., Wilson, I.D., Castro-Perez, J., (2005). A rapid screening approach to metabonomics using UPLC and oa-TOF mass spectrometry: application to age, gender and diurnal variation in normal/Zucker obese rats and black, white and nude mice. *Analyst*, **130**, 844–849.

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., 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., (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, **3**, 211–221.

Want, E.J., Wilson, I.D., Gika, H., Theodoridis, G., Plumb, R.S., Shockcor, J., Holmes, J., Nicholson, J.K., (2010). Global metabolic profiling procedures for urine using UPLC–MS. *Nature Protocols*, **5**, 1005-1018.

Investigation of the metabolism and potential nephrotoxicity of 2-Bromophenol after a single dose

Chapter Four

Previous work by Lau et al. had identified 2-bromophenol, a phenolic metabolite of bromobenzene, as a potential nephrotoxin. Due to these findings 2-bromophenol was chosen as one of the test compounds of the COMET studies. Here continuing on from those studies, a repeat of the original NMR experiments was performed, with further information UPLC-MS data acquired from the same samples. Application of the previously described rapid microbore methodology was also performed. Data analysis identified 2-bromophenol glucuronide and 2-bromophenol sulphate as the major metabolites of 2-bromophenol. However no toxic metabolites were detected and after analysis of the histopathology and clinical chemistry, there is little evidence for any toxic action of 2-bromophenol after a single dose.

SUMMARY:

- 2-bromophenol, a metabolite of the hepato- and nephrotoxic compound bromobenzene, is a suspected nephrotoxin.
- 30 Sprague Dawley rats were split into three equal groups and were administered a single IP dose of either vehicle (corn oil), 100mg/kg or 200mg/kg of 2-bromophenol. Urine was collected at regular intervals, and after seven days the animals were sacrificed and kidneys retrieved.
- Urine samples were analysed by ^1H NMR spectroscopy and UPLC-MS to determine both the effects of 2-bromophenol on urinary endogenous metabolite profile and the metabolic fate of the compound itself.
- Chemometric analysis showed that the main discriminators between dosed and control animals were associated with metabolites of 2-bromophenol. Little to no evidence of toxicity was seen in terms of altered endogenous metabolite profiles, histopathology or clinical chemistry.
- A number of 2-bromophenol metabolites were identified in urine, including with 2-bromophenol glucuronide and 2-bromophenol sulphate but there was no evidence for the production of reactive metabolites.
- The conclusion from this investigation was that 2-bromophenol caused little or no nephrotoxicity at the dose administered after a single administration of either 100 or 200 mg/kg of 2-bromophenol to the rat.

4.1 Introduction

As discussed in Chapter 1, the COMET project resulted in 147 studies, including many on toxins, being performed on rodents (rats and mice) with the specific aim of generating diagnostic metabolic profiles which might further illuminate the various mechanisms of the toxicity exhibited by the dosed compounds (Lindon *et al.*, 2005). One of the original compounds selected for the COMET program is 2-Bromophenol (COMET study R16). 2-bromophenol is a metabolite of bromobenzene, which, as described below, is a well-known hepato- and nephrotoxin.

4.1.1 Toxicity and metabolism of bromobenzene

Halogenated hydrocarbons have long been considered to be potent kidney toxins, and bromobenzene has been one of the compounds of choice for the investigation of the mechanism(s) of toxicity (Rush, *et al.*, 1984). Administration of bromobenzene to rodents has been shown to induce proximal renal tubule cell necrosis in addition to covalent binding of bromobenzene and its metabolites to renal cell proteins (Read, 1973). Lau *et al.* in a series of studies (1984a; 1984b) have investigated the mechanisms responsible for the toxicity and the metabolism of bromobenzene and these are thought to have been well characterised by these authors. Cytochrome P450 monooxygenase (CYP) mediated biotransformation of bromobenzene occurs in the liver; CYP enzymes are known to metabolize some chemicals to highly reactive forms, such as epoxides, and two such products of the metabolism of bromobenzene have been observed: bromobenzene 2,3 epoxide and bromobenzene 3,4 epoxide (Madhu *et al.*, 1992). Bromobenzene 2,3 epoxide is relatively stable, although it has been shown to bind to only infrequently to blood proteins such as haemoglobin, and is then further metabolised into 2,3 bromo-dihydrodiol by microsomal epoxide hydrolase (Lau and Zannoni, 1981) after which it will then be removed by the traditional detoxication routes, which were discussed in chapter 1.

Bromobenzene 3,4 epoxide however, unlike bromobenzene 2,3 epoxide, is electrophilic and therefore highly reactive. It has been shown to directly bind to glutathione S-transferase resulting in conjugation to glutathione. This conjugation to glutathione detoxifies the reactive metabolites of bromobenzene but results in glutathione depletion in hepatocytes (Madhu and Klaassen, 1992). The epoxide can then be further hydrolysed by CYPs into 2- and 4-bromophenol. 4-Bromophenol, and another potential metabolite 4-bromocatechol, have been shown to not induce toxicity (Madhu and Klaassen, 1992), but 2-bromophenol can be oxidised to bromohydroxyquinone as discussed below.

4.1.2 Toxicity and Metabolism of 2-Bromophenol

2-Bromophenol and bromohydroxyquinone are readily transported to the kidneys from the liver, an organ rich in glutathione S-transferase and the combination with glutathione produces various mono- and di- substituted molecules (Parke and Piotrowski, 1996). The accumulation of these conjugates, in addition to the depletion of the local glutathione pool, is thought to be the cause of the renal toxicity (Lau et al. 1984b). Figure 4.1 describes schematically the biotransformation of bromobenzene to 2-bromophenol and bromohydroxyquinone. The loss of intracellular protection against oxidative stress and other hazardous substances can also lead to indirect events further exacerbating cellular damage. However, there is little information about the mechanism of toxicity of 2-bromophenol and there is no information on the bioavailability of the compound. Bruchajzer et al. (2002) however, concluded that evaluation of the nephrotoxicity of 2-bromophenol should be based on urine analysis, as they observed changes in urine which included increased protein excretion and elevated epithelial cell content, despite variability in the levels of the 'classic markers of nephrotoxicity' such as creatinine.

This study aims to utilize the original COMET samples, combining new technologies with previous analyses to provide a broader view on the acute nephrotoxicity of 2-bromophenol

after a single administration, investigate the reversibility of the effects and identify both biomarkers of toxicity and perturbations of endogenous metabolism.

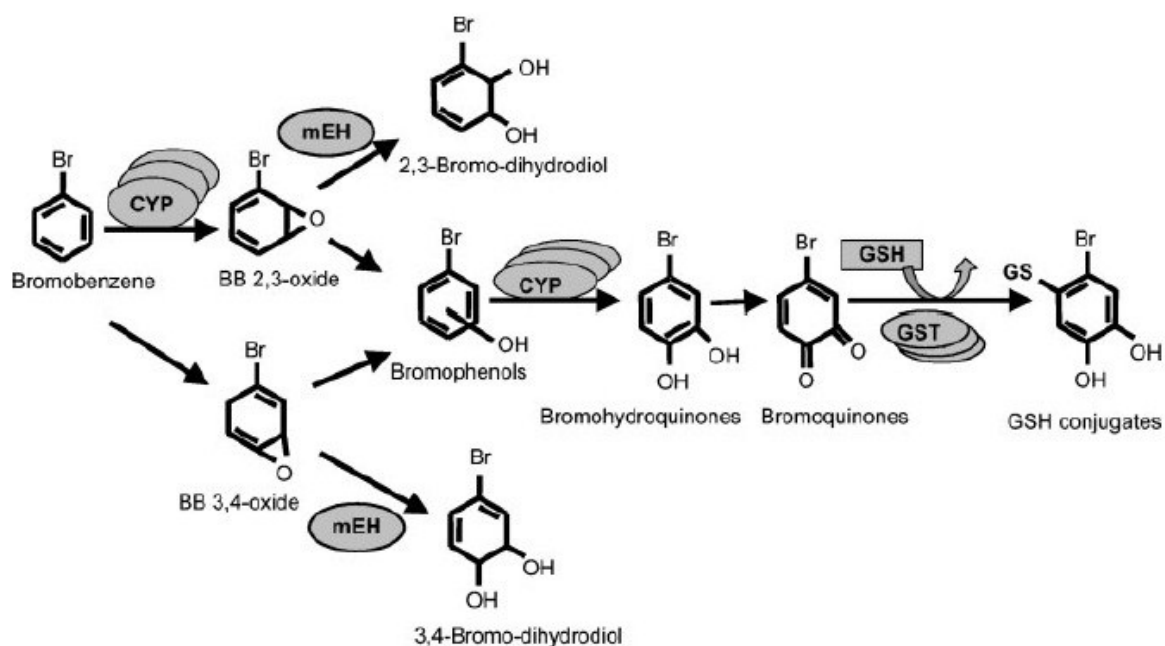


Figure 4.1. A visual presentation of the metabolism of bromobenzene in the rat liver and kidney.

4.2 Materials and Methods

4.2.1 Study Design and Sample Collection

The reanalysis performed in this study was performed on the samples collected in the COMET study R16 where urine and serum samples were collected after the intraperitoneal (IP) administration of 2-bromophenol as a single dose to male Sprague-Dawley rats. The rats (10 animals/group) were dosed with 100 mg/kg or 200 mg/kg of 2-bromophenol in corn oil (control rats were dosed with corn oil). Rats were housed in metabolic cages and urine

was collected for metabonomic and clinical analysis. Urine samples were collected at 16 hours pre dose, at time of dose and at 8, 24, 48 72, 120, 144 and 168 hours post dose, with a total of 230 samples collected. Clinical observations and body weight were recorded daily. Clinical chemistry was performed on study day 2, 24 hours after treatment (all rats); study day 3, 48 hours after treatment (group A); and on study day 8, 168 hours after treatment (group B). Five animals from each group were sacrificed on study day 3 and five were sacrificed on study day 8. Kidneys were removed and were examined histopathologically. Samples were stored at -40 °C prior to analysis.

4.2.2 Sample Preparation and Experimental (NMR spectroscopy)

The collected urine was spun down in a centrifuge and 250 μ L of sample aliquoted into 2 mL cryo-vials and diluted 1:2 with phosphate buffer. A quality control mixture containing 10 μ L of each sample was also prepared. Each urine sample was centrifuged at 12000g at 4 °C for 5 min and subsequently 500 μ L was transferred into 96-well plates using a Bruker Sample Track system and a Liquid Handler 215 preparation robot. ^1H NMR spectroscopy was performed at a field strength of 14.1 T (^1H 600.29 MHz) using a 5 mm broadband inverse configuration probe with a z axis magnetic field-gradient capability. The spectrometer was equipped with a Bruker Sample Jet system set to 5mm shuttle mode with a cooling rack of refrigerated tubes at 6 °C and controlled via a Bruker Avance III console. The ^1H NMR spectra of the urine were measured using one-dimensional (1D) nuclear Overhauser enhancement spectroscopy (NOESY)- presat (noesypr1d) pulse sequence (Stonehouse and Keeler, 1995). The data were acquired using Topspin 3.2 and run under automation by IconNMR.

4.2.3 ^1H NMR Spectroscopic Data analysis

The same computational methodology was employed for the chemometric analysis of all of the various NMR datasets. After phasing to correct for first and zero-order errors, and baseline correction to account for distortions of the base values, the data was imported to

MATLAB R2012b (Natick, Massachusetts, U.S.A.) using the Metaspectra program script, using a resolution of 0.00055. The data were then aligned using Recursive Segment-Wise Peak Alignment (RSPA) (Veselkov et al., 2008) and normalised by Probabilistic Quotient to account for, and exclude, systematic sources of bias within samples which were not due to biological processes or environment. Statistical Recouping of Variables (SRV) (Blaise et al., 2009) was then used for peak detection, and the integrated dataset was exported to SIMCA-P+ 13.0.2 software for multivariate analysis.

4.2.4 Sample Preparation and Experimental (LC-MS Analysis)

Prior to analysis the sample order was randomised using the Microsoft Excel random number generator function. All 230 samples were included in the analysis. A 150 μL aliquot of urine was mixed with 300 μL water to dilute the salt concentration before protein removal via precipitation with acetonitrile (1:2 v/v). These samples were vortexed and left at $-20\text{ }^{\circ}\text{C}$ overnight before centrifugation for 5 min at 12 000 g at $4\text{ }^{\circ}\text{C}$. For reversed-phase chromatography, 50 μL of sample was taken and added to 150 μL water in 350 μL 96-well plates (Waters Corporation, Milford, USA). A pooled quality control (QC) sample was prepared (Want *et al.*, 2010) by combining 100 μL of each sample and diluting 1:4. Dilutions of 1 in 2, 1 in 4 and 1 in 8 of this QC sample in water. The 96-well plates were stored at $-20\text{ }^{\circ}\text{C}$ until the analysis by LC-MS as described below. Immediately before being placed in the autosampler the plates were centrifuged at 1700 g before being placed into the autosampler at $4\text{ }^{\circ}\text{C}$.

Reversed-phase liquid chromatography was performed on an Acquity UPLC system (Waters Ltd, Elstree, UK) coupled to a Synapt G2 triple quadrupole dual time-of-flight mass spectrometer (Waters MS technologies, Ltd., Manchester). UPLC-MS conditions for urine analysis were those previously described for the reversed-phase analysis of urine (Want *et al.*, 2010) where the separation was performed using a Waters HSS T3 ($1.8\text{ }\mu\text{m}$, $2.1 \times 100\text{ mm}$) column and $2\mu\text{L}$ per sample were injected. The chromatographic mobile phase was

composed of 0.1 % formic acid in water (A) and 0.1 % formic acid in acetonitrile (B). The column temperature was maintained at 40 °C and linear gradient elution was performed at 0.5 mL/min. The starting composition was 1 % B, held for 1.0 min before increasing to 15 % at 3.0 min, 50 % at 6.0 min, 95 % at 9.0 min for a 1.0 min wash and returning to 1 % B for a 2 min re-equilibration step (total cycle time 12 min). Mass spectrometry was performed with electrospray ionisation operated in both positive (ESI+) and negative ion modes (ESI-) The instrument was operated in sensitivity mode and set to acquire data over the m/z range 50 – 1200 with a scan time of 0.1 s. All mass spectral data were collected in centroid mode using the MSe data acquisition function. For mass accuracy, leucine enkephalin (MW = 555.62) was used as a lock mass at a concentration of 200 pg/ μ L (in 50/50 CH₃CN:H₂O, 0.1 % formic acid) infused at a flow rate of 20 μ L/min via a lock spray interface. The data were collected using MassLynx V 4.1 software (Waters Corporation, Milford, USA).

4.2.5 UPLC-MS Data analysis

Each UPLC-MS data set was pre-processed using both Transomics (Non Linear Dynamics, UK) and XCMS Bentobox (Tautenhahn *et al.* 2012) software in order to convert the three-dimensional LC-MS raw data into time-aligned detected features of retention time, m/z ratio and intensity (peak area). The ion intensities for each peak detected were then normalised, within each sample, to the median peak intensity in that sample. The resulting peak marker tables, comprising m/z , RT and normalised intensity values for each variable in every sample, were exported to SIMCA-P+ 13.0.2 software (Umetrics, Umea, Sweden) to perform multi-variate statistical analysis.

4.2.6 Statistical Modelling

Prior to any statistical modelling the QC's were evaluated, through both the use of CV and positioning within PCA plots. If there was gross variation within the quality control data, or if there was significant drift within the position of the QCs within PCA scores plots the data

would have been unfit for analysis. SIMCA-P+ 13.0.2 was used for the statistical modelling; Principal component analysis (PCA) (unsupervised) and Partial least square regression discriminant analysis (OPLS-DA) were utilised to explore any underlying stratification within the data. Scores plots were used to investigate any differences within the sample groups and the loadings were used to identify which (if any) variables were driving the separation. Validation by permutation testing was also performed to test the predictive capabilities of the models created. Each peak or feature of interest was then tested for significance using a modified Multiple Analysis of Variance (MANOVA) and the p values corrected for multiple testing using Beynjamini-Yekutielli correction. Metabolite assignment was then performed in the manner previously discussed in the introduction.

4.3 Results and Discussion

4.3.1 Comparison with original data

The samples used in this study were collected in 2003 and subsequent to the initial COMET NMR spectroscopy have been kept in storage at -40°C . In order to confirm the integrity of the samples and determine if there had been any time related changes in their composition, the original ^1H NMR data, were compared with the new data from the analysis of the samples. When compared sample-to-sample, after accounting for the increased resolution of the newer probe used, relatively few differences were noted between the spectra, as illustrated for the spectra obtained for sample R006600 in 2003 and in the re-analysis in 2014 depicted in figure 4.2. Using Statistical Recouping of Variables (SRV) for automatic peak detection with identical parameters, the total number of features detected in each sample in the reanalysis was found to be increased by approximately 10% (17 more features per sample) compared with the original data (discussed below). However, every feature detected in the original data was present on reanalysis in the present study, and the PCA models produced from each dataset showed the same trends. This evidence suggests that the storage conditions used for the samples were appropriate for maintaining their integrity, and supports re-analysis using more modern NMR instrumentation and UPLC-MS technology. From this point, all the ^1H NMR results discussed were generated from the reanalysis of the samples using data acquisition at 600 MHz.

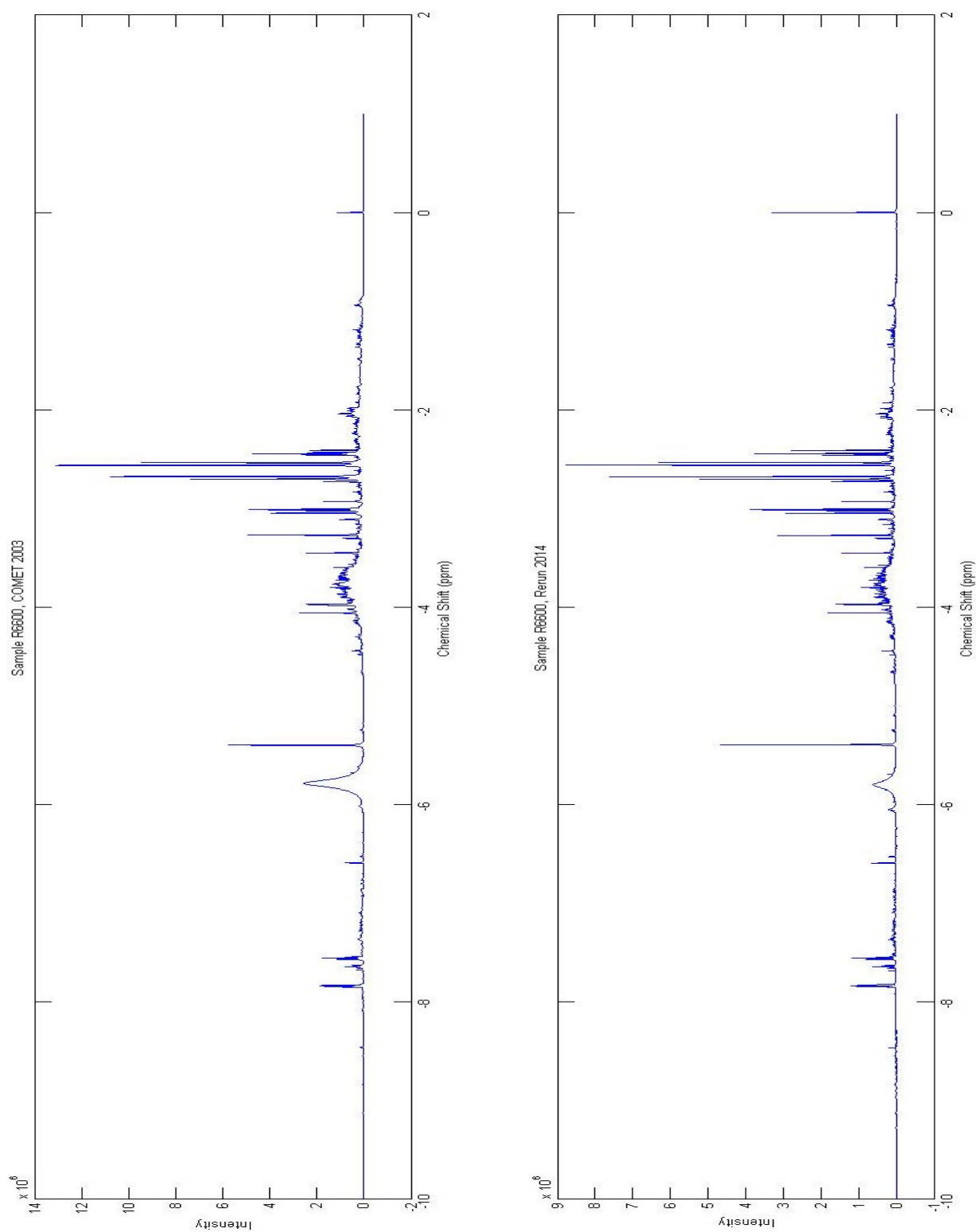


Figure 4.2 A comparison between the spectral data for sample R006600 acquired in the original NMR run in 2003 (top) and the data from the same sample obtained in 2014 to highlight the chemical stability of the samples.

4.3.2 Hippurate and, aminohippurate as potential biomarkers of 2-bromophenol nephrotoxicity

a. Original COMET results

The results of the original COMET project have been summarised by Lindon et al. (2005) but are briefly reiterated here. In the case of 2-bromophenol the data produced were analysed using the “traditional” chemometric methods employed for metabonomics at that time and the metabolic signatures were compared between treatment groups. Multivariate analysis of the data using PCA showed that there was a general separation of the high dose group from the other low dose and control groups within statistical space, whereas there seems to have been little response to the low dose administration of 2-bromophenol. The original findings stated that for the high dose group the maximal response was detected to be at 48hrs post treatment. This response was indicated by an increase in the amounts of creatinine in the urine, with other smaller elevations in the concentrations of citrate and hippurate and a decrease in the quantity of aminohippurate.

b. Results for the ^1H NMR Reanalysis of the R16 Urine Samples

Quality Control Samples

To assess the analytical stability of the ^1H NMR spectroscopic analysis a pooled quality control sample was included from which spectra were spectra at the beginning, middle and end of the run. These QCs can be seen tightly grouped together in the PCA scores plots (e.g. see Figure 4.4) indicating a good degree of consistency in the analytical process. If these data had showed that the QCs were not similar, i.e. showing a drift along statistical space then this would indicate that there had been time-related analytical changes (e.g., resulting from degradation whilst samples were queued for analysis). These data were further investigated via the calculation of the variance for each integrated feature for the QCs

to further examine consistency, and pinpoint any sources of variability. Variance is defined as:

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

Where $f(x)$ is the probability density function of a continuous variable and simplified in this case to:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

Here variance is the sum of the squared distances of each feature in the spectrum from the mean (μ), divided by the number of features detected (N). Total cumulative intensity variance for the QCs was 4.63×10^{-10} , which for the deviation of this data set is very low, therefore suggesting that the QCs were highly consistent. As the QCs are a pool, formed by mixing equal volumes from every sample analysed, they should represent an analytical ‘average’ when components are constructed in a PCA. They can also be interrogated by comparing the deviation of the QC samples from this theoretical average to reveal the difference between the weighted difference for the QC data and the predicted average of the model, which is presented in figure 4.3. The bars within the plot show how far each variable deviates from the theoretical mean, and the smaller the bars are the more consistent that area of the data set is. As none of the bars intersect the one standard deviation threshold we can be confident in the analytical comparability of the data. This procedure cannot be used to assess the quality of the original COMET study data as QC samples were not used).

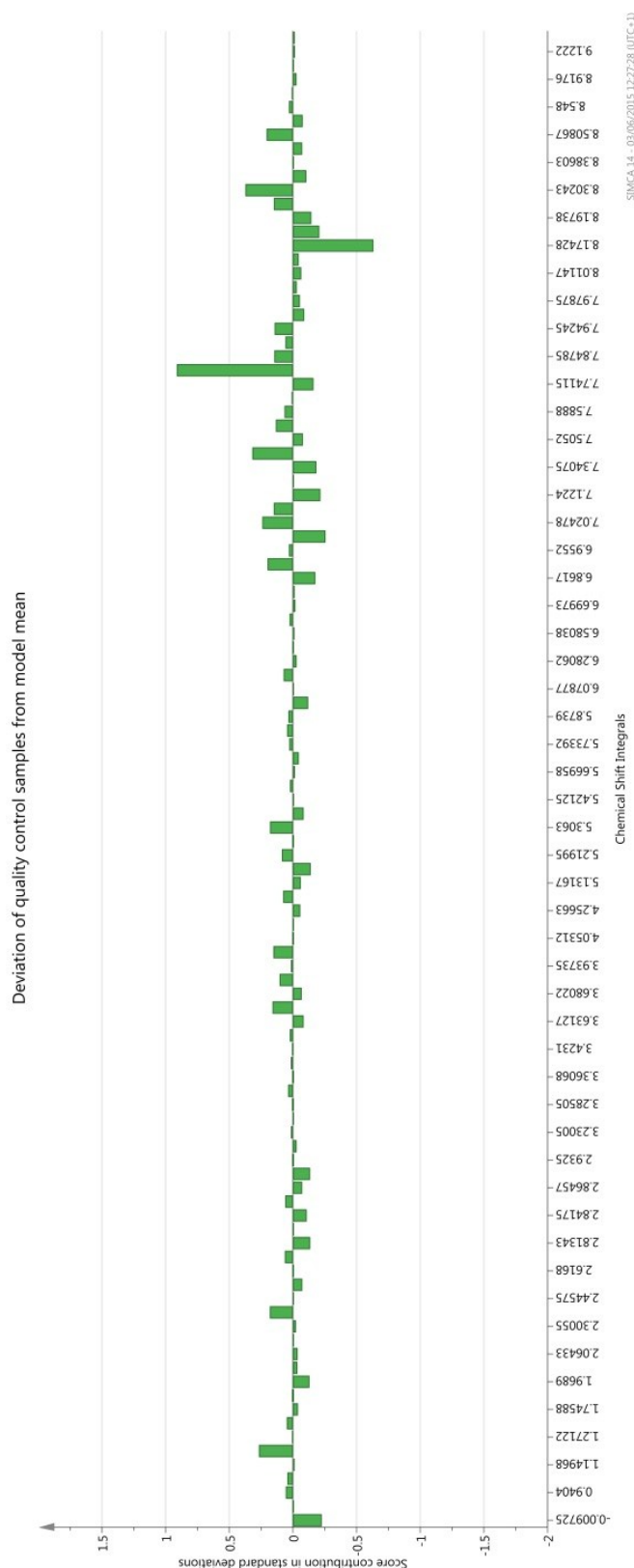


Figure 4.3 Contribution plot showing the deviation of the QC samples in each integrated spectral region from the statistical mean of the principal component analysis expressed in standard deviations. The plot displays the vector $v_k(x_{ik} - \text{avg}_{ik})$ for each variable as a bar, the orientation of which indicates the direction of the deviation. As none of the bars intersect the one standard deviation threshold we can be confident in the analytical conditions were stable throughout the run.

c. Chemometric Analysis

Chemometric analysis of the newly acquired ^1H NMR data did however, highlight the same trends as the previous COMET analysis. Construction of PCA models indicated a general separation of the high dose group (200 mg/kg) from the low dose (100 mg/kg) and control groups within statistical space. However, when the complete dataset was used, there was little difference between the data acquired from the controls and for samples obtained following low dose administration, a trend which was echoed by supervised O-PLS discriminant analysis (Fig 3.4a and 3.4b). By removal of data for the low dose group from the analysis we attempted to identify the discriminating signals driving the separation of the control and high dose groups. By investigating each time point separately it was found that the point of maximum metabolic difference following administration of 2-bromophenol at 200 mg/kg was seen in the samples acquired at 8 hours post dose, with a reduced response at 24 hours where the samples at this time point are situated closer in the statistical space to the controls (figure 4.4). There also seems to be a separate group within the low dose animals which responded differently to the administration of 2-bromophenol.

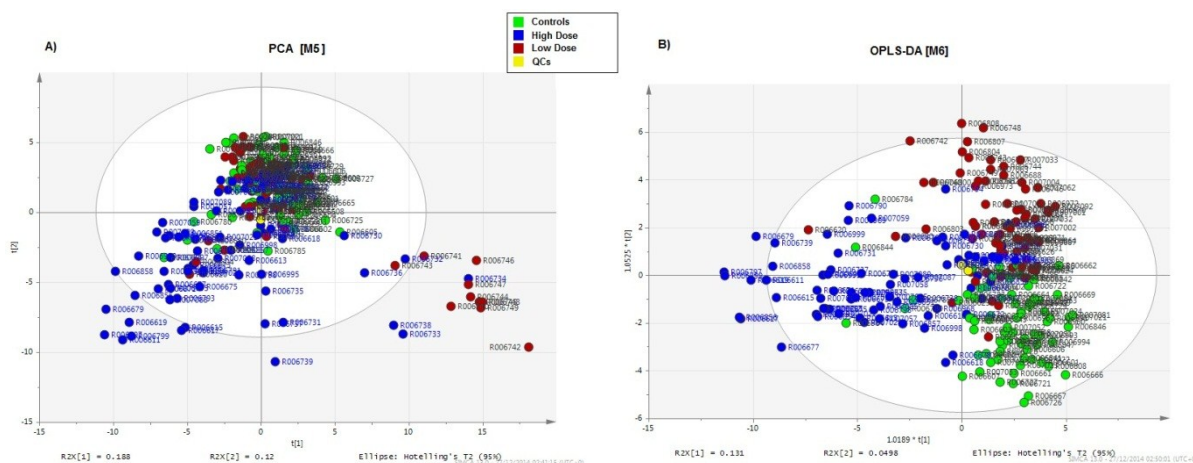


Figure 4.4 The scores plots from the multivariate analysis of the ^1H NMR data obtained for the 2-bromophenol study A) and B) are the scores of PCA and OPLS-DA, respectively, of the whole data set. There seems to be a distinct trend in statistical space pushing the high dose group away from the other two. The samples coloured in Yellow are QC samples and these are present consistently in the middle of statistical space indicating good stability for the analytical platform.

Using the S-plot (figure 4.5) from the supervised O-PLS discriminant analysis of the data it is possible to identify the signals driving group differences. From examination of the data acquired for the samples collected at 8 hours post dose it is possible to identify the major changes in endogenous metabolites. These were an increase in the amount of hippurate ($\delta 3.96/7.56/7.84$, $p = 7 \times 10^{-17}/3.78 \times 10^{-18}/6.16 \times 10^{-17}$) and aminohippurate ($\delta 6.87$, $p = 1.49 \times 10^{-17}$) and a decrease in 3-methylhistamine ($\delta 2.84$, $p = 2.29 \times 10^{-20}$) (these identifications are based on database searches and match what was previously reported by the original COMET report, provided in the appendices). The p-values quoted are derived from the ANOVA significance testing, after Benjamini-Yekutieli multiple testing correction. A variance plot depicting the intensities of the integrated region of the three metabolites mentioned above for each of the samples has been constructed for each metabolite signal and also presented below (figure 4.6). Through interrogation of the 600 MHz ^1H NMR spectra no parent compound/metabolites was seen, possibly because the signals from the 2-bromophenol/metabolites were obscured by other larger peaks. While the PCA trends identified here might also be expected to occur due to the excretion of dosed compound and its metabolites in the urine, which are known to be present as discussed later, statistically there seem to be other areas driving the differences.

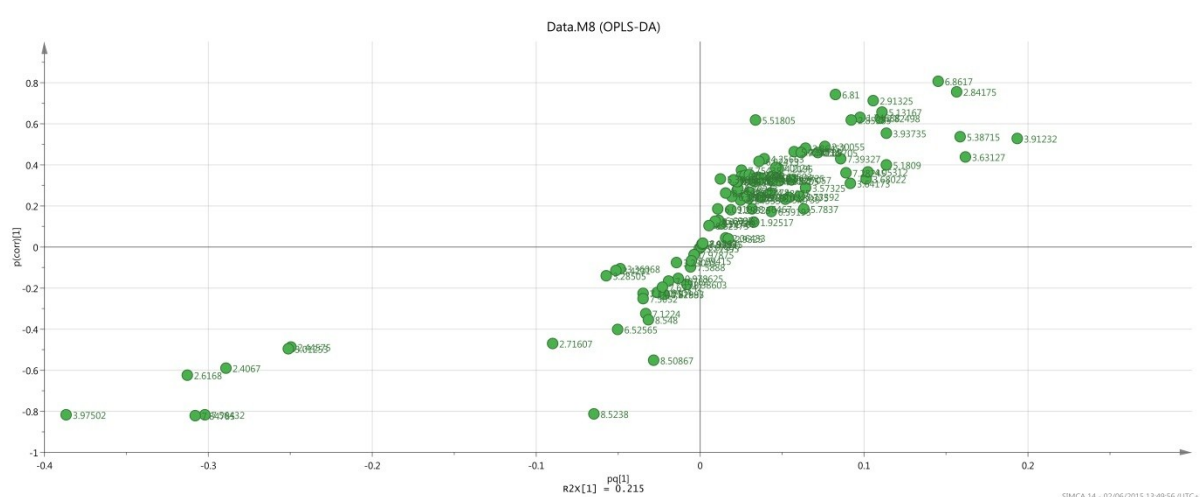


Figure 4.5 S plot derived from the OPLS model comparing high dose and control groups at 8 hours post dose. Variables clustered towards the bottom left are variables which represent metabolites more abundant in the high dose group, whereas variables in the top right are down regulated in the treatment group.

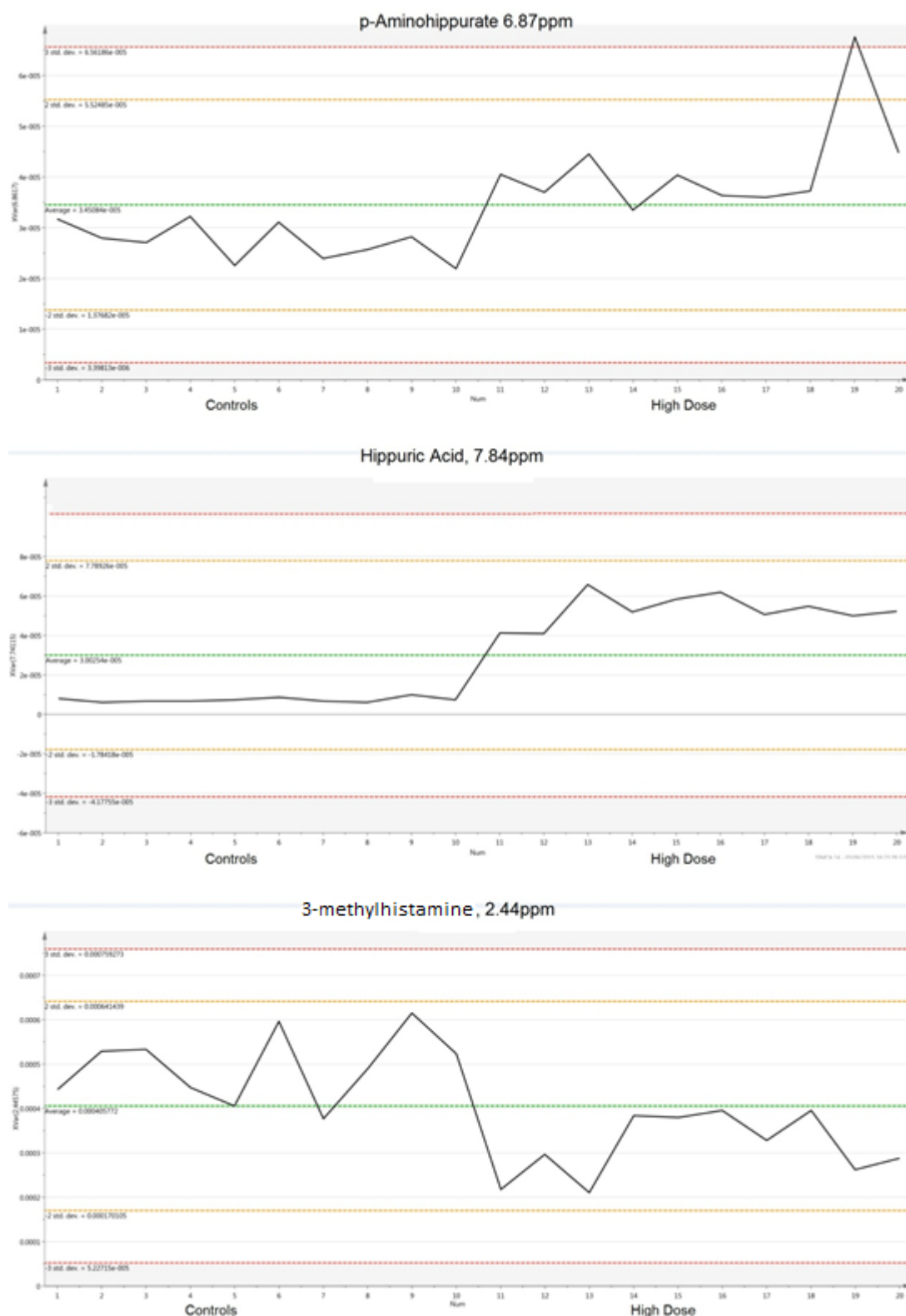


Figure 4.6 Xvar plot of the levels of the indicated metabolites in the various animals. The yellow threshold in the graph is 1 standard deviation whereas the second red threshold is the second deviation from the theoretical mean. There are clear trends in the amounts of hippurate and *p*-aminohippurate, whereas there seems to be greater within group variability in 3-methylhistamine concentrations.

The OPLS model was able to provide convincing test statistics with an R^2 of 0.95 and a Q^2 of 0.91). The R^2 shows the amount of Y variables explained by the model after cross validation and gives the 'goodness-of-fit' of the model, the Q^2 gives information about the predictive quality of the model.

To test for the validity of the model and assess the possibility that the results were spurious permutation testing was performed, with 1000 iterations recalculated every 100 permutations (due to technical constraints). Permutation testing compares the model generated from the data with random chance, which is to say that the performance of the test statistics of the model against randomly generated Y-data is compared. Figure 4.7 illustrates the results of the permutation testing. In this figure the green and blue points on the far right are the original R^2 and Q^2 values, and the points to the left of these are the values of the randomly generated data. The two dotted lines are the regression lines for each of the two statistics. The criteria for validity of the model are that all the Q^2 values of the random data are lower than the original and that the Q^2 regression line intersects the vertical axis at, or below 0; and in both cases the model passed these acceptance criteria. While R^2 values always show a degree of optimisation, and will not always follow the same trend as the Q^2 values, the fact that in the test permutation all of random data R^2 values were also lower than the original which further adds validity to the model.

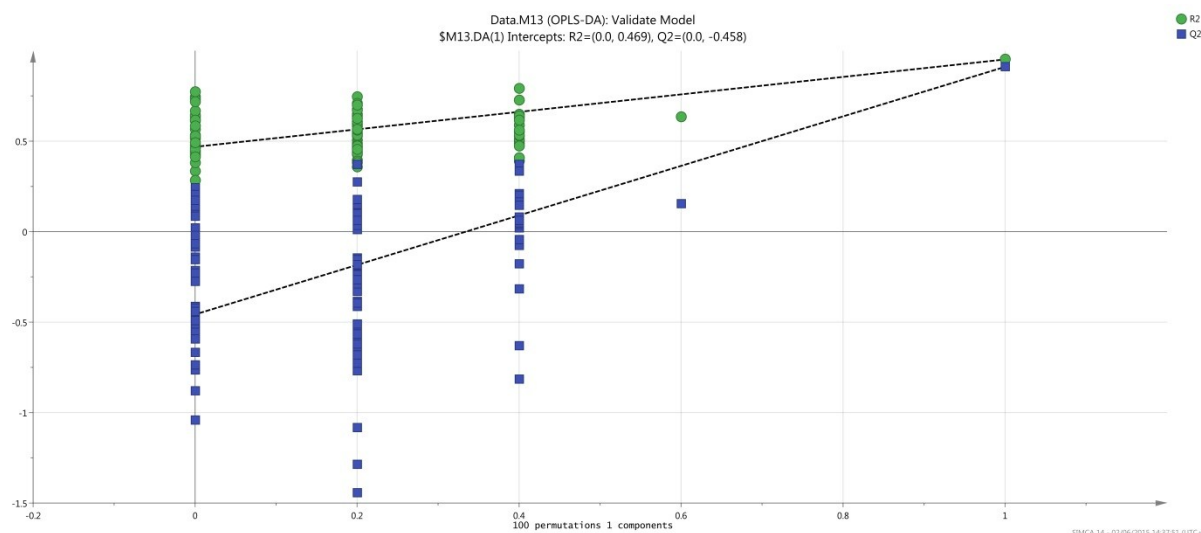


Figure 4.7 Permutation plot for the model constructed. The green and blue dots represent the R2 and Q2 of the dummy data respectively. The permutation values are lower than the actual values suggesting that the model is highly robust.

As a final step the significance of the model was investigated using ANOVA of the cross-validated residuals (CV-ANOVA). This is achieved by comparing the significance level of the OPLS model in the variation around the mean. Table 4.1 summarises the results of the CV-ANOVA. In the table the important number is the p value in the final column which, which is highly significant, meaning we can be highly confident in the validity of the models which were constructed.

Controls vs High Dose	Sum of Squares	Degrees of Freedom	Mean Squares	F value	P value
Total corr.	19	19	1		
Regression	17.3042	4	4.32605	38.2659	1.05543e-7
Residual	1.69579	15	0.113052		

Table 4.1 Results of the CV-ANOVA

We should note that investigation of the data for pre-dose time point revealed basal stratification of the animals into two distinct groups, with a minor subgroup of animals showing inherently elevated urinary concentrations of hippurate and lower concentrations of

other endogenous metabolites including 4-aminohippurate and N-acetylhistamine. While members of this minor group were all part of the low dose treatment (100mg/kg), and were therefore omitted from the majority of the analyses (due to their potential to bias the analysis), they could have potentially skewed the results of the PCA. Furthermore they could indicate inconsistencies in metabolic state of the animals prior to the start of the study, or differences in microbiomes, housing or feeding.

d. Hippurate as a potential biomarker of the effects of 2-bromophenol administration

The metabolite that was seen to be consistently detected as significantly changed as a result of administration of 2-bromophenol at 200 mg/kg, throughout the data interrogation, was hippuric acid or hippurate. Hippurate is the glycine conjugate of benzoic acid and has been shown to be readily detectable by both ^1H NMR spectroscopy (Bales et al., 1984) and LC-MS (Sakai et al., 1983) analytical platforms. In the rat the kidneys (together with the liver) serve as a major site of glycine conjugation as well as the main route of hippurate excretion via urine. During impaired renal function hippurate has been seen to accumulate in the serum (Niwa, 1996) coupled with a reduction of urinary concentrations (Bairaktari et al., 2001). As such it has been identified as an indicator of active secretion competence and kidney amino acylase activity, as well as expression of organic anion transporters (Christians et al., 2008). As the main location of filtration, glomerular nephrotoxicity specifically causes perturbation of hippuric acid elimination; indeed reduction of urinary hippurate concentrations have been shown to directly correlate with lesion severity in human patients with glomerulonephritis, an acute nephrotic syndrome characterised by damage to the glomeruli (Psihogios et al., 2007). The utility of hippurate as a biomarker in metabonomic studies has been reviewed by Lees *et al.* (2013) and as such should be considered for further investigation. Our findings are in agreement with previous studies, and may indicate that 2-bromophenol may localise to the glomerulus.

e. *p*-Aminohippurate

p-Aminohippurate is another well-established marker of nephrotoxicity and has been used to specifically measure renal plasma flow. Transport of organic anions, including *p*-aminohippurate occurs almost exclusively in the proximal tubule cell through active anion exchange. This is a two part process. Initially *p*-aminohippurate is exchanged for intracellular 2-oxoglutarate by the PAH/2-oxoglutarate membrane bound exchange protein PAHT (Lu et al., 1999); subsequently the metabolite is exported to the tubular lumen through the action of the multidrug resistance protein 2 (MRP2) an ATP-dependent exporter protein (Leier et al., 2000). While PAHT has been shown to have broad substrate specificity, and is able to transport a variety of organic substrates (Sekine et al., 1997), MRP2 is specific to amphiphilic anions, such as *p*-aminohippurate and other endogenously formed glutathione and sulfate conjugates. A decrease in *p*-aminohippurate secretion has long been thought to result from the inhibition of these transport mechanisms through competitive binding to the transporter protein of a toxin (Berndt and Grote, 1978).

4.3.3 Insights from LC-MS analysis

While with the ^1H NMR reanalysis confirmed sample integrity and, as in the original COMET analysis, enabled the 2-bromophenol dose groups to be separated via e.g. PCA as well as highlighting a number of endogenous metabolites perturbed by 2-bromophenol administration it revealed little new information. Therefore UPLC-MS was undertaken to further characterise the effects of 2-bromobenzene administration by providing an orthogonal, and potentially, more sensitive analytical platform.

a. Quality control

As discussed in previous chapters LC-MS techniques often require more significant sample preparation and are also subject to the degradation of instrumental performance during analysis leading to the acquisition of unreliable data. For quality assurance/control aliquots from every sample were combined in a pooled sample and then analysed at regular intervals throughout the experiment (Want *et al.*, 2010). Initial visual inspection of the QC chromatograms showed good consistency throughout the run. Figure 4.8 illustrates how the chromatograms of QCs throughout the run overlay well, with little evidence of retention time drift. These trends were mirrored in the initial PCA scores plots used to scrutinise analytical drift. Figure 4.9 shows the initial un-filtered data for the negative electrospray ionisation (ESI) runs; the red points in both scores plots represent the QC samples. In the positive ESI run it can be seen that all the QCs cluster together in the centre of statistical space, until QC 18 (injection 203). From this point the QCs seem to drift off in a time dependant manner with the final injection, QC 23 being the furthest. This seems to indicate that there was a drop in analytical stability after approximately 200 injections however, since the deviation was not statistically significant in the overall variation within the data the data was accepted as usable. Interestingly a similar trend was observed for the negative ESI data, with the QCs initially being well clustered with the same deviation occurring at QC 19 (injection 206), again this drift was not significant and the data were also accepted. Nonetheless the data suggests

that there is a typical loss of consistency after around 200 injections, perhaps indicating the limit of the methodology employed.

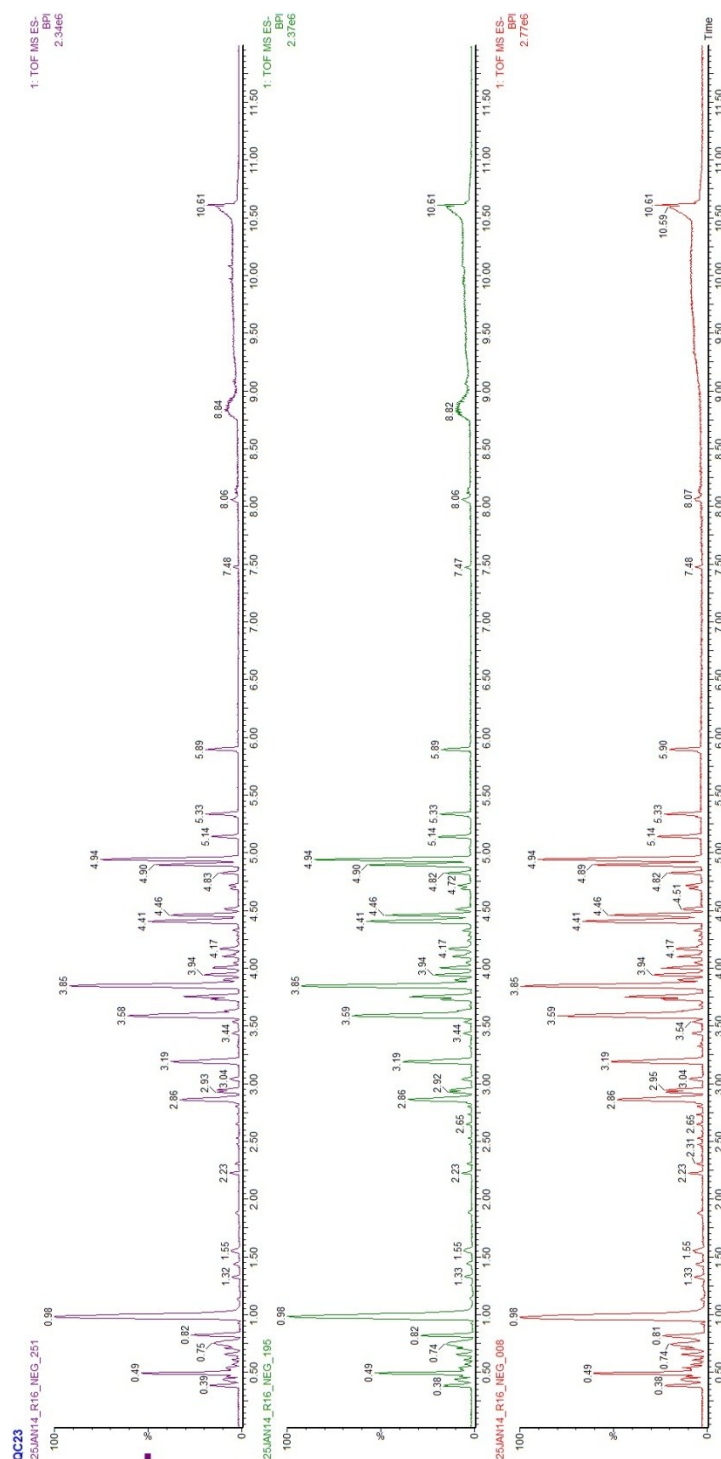


Figure 4.8, chromatographic peaks of the first QC (red trace, injection 8) the 10th QC (green trace, injection 195) and the last QC (purple trace, injection 251). The chromatograms are highly consistent with no visible retention time drift, suggesting that the analysis achieved good analytical stability

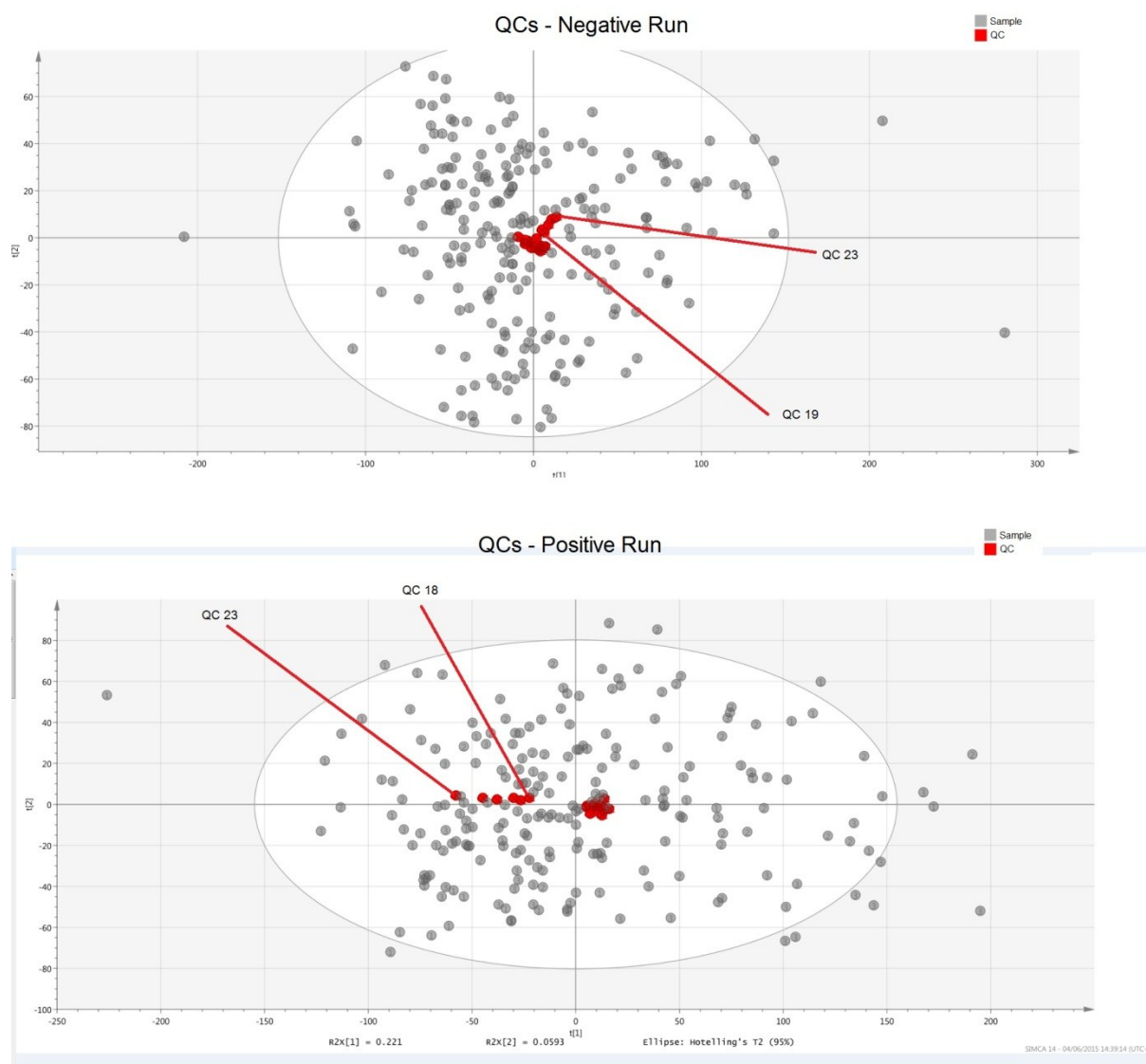


Figure 4.9 PCA scores plot showing the positioning of the QC data within statistical space, in both the positive and negative ionisation modes. The red dots are the QCs while the grey dots are the samples. The QCs in both modes tend to cluster in the centre of the plot, with a time dependant deviation after around 200 injections.

To ensure that the QCs were consistent throughout, and that there were no regions of the acquisition where there was instrumental instability masked the overall stability of the metabolic profile the coefficient of variation for each of the features detected in the QCs was calculated. Ion intensity stability is typically poor for low intensity ions and greater the more

intense the ion (providing it is not saturating the detector) and those features that are subject to ion suppression/enhancement will show greater variation than those that are resolved from ion suppressing matrix components. The coefficient of variation (CV) expressed as a percentage is defined as:

$$CV = \frac{\sigma}{\mu} * 100$$

Where σ is the standard deviation and μ is the expected return, in this case the mean. The below table (table 3.2) shows the results of the CV calculations. From this analysis it is clear the that negative ESI results were less stable than those from the positive ESI analysis, although the results could be skewed due to the fact that the negative run detected more features (18589 features) than the positive run (17816 features). Typically the features detected in negative ESI are less intense than those in positive mode, and therefore more subject to background noise, which can result in greater variability.

	CV less than		
	<15%	<20%	<30%
Negative	29.2%	43.2%	64%
Positive	50.7%	64%	79.2%

Table 4.2, CV percentage scores for the QC data.

b. Chemometric analysis

After computational analysis with Transomics software, the extracted metabolic signatures, defined by retention time and m/z ratio, both the positive and negative ESI datasets were investigated using multivariate statistics in a similar way to the ^1H NMR spectroscopic data.

The positive ESI data was analysed first, but did not show any trends in the PCA modelling scores in either the overall dataset model, or when high dose and controls were compared at 8, 24 and 48 hours (figure 4.10), although this does match expectations, as it was predicted that bromophenol and its metabolites would ionize better in negative mode. However, there

were some time related trends between the pre dose time point and the 8 hour time point in the high dose group. These trends appeared to be due to changes in the endogenous metabolite profile.

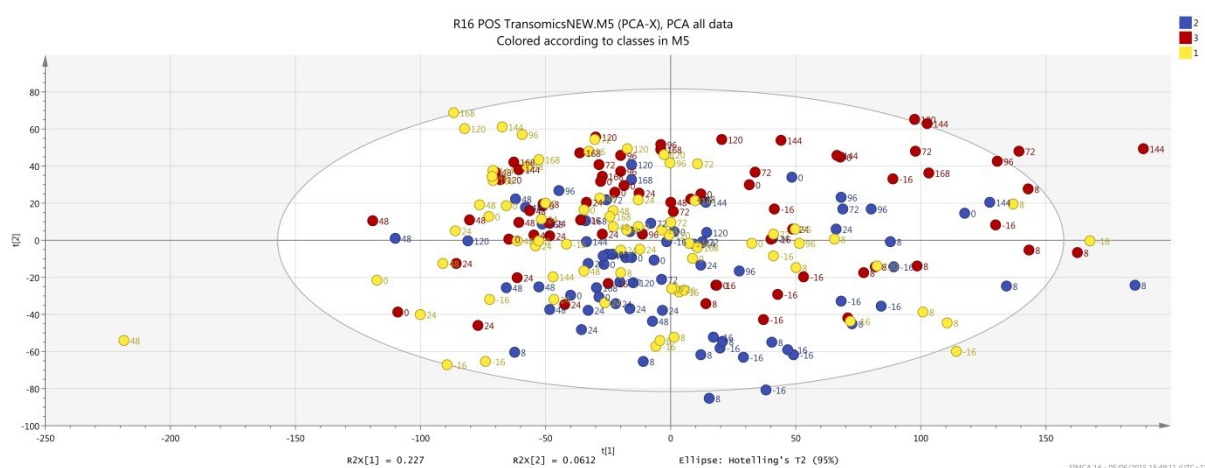


Figure 4.10 PCA scores plot of the model comparing dose groups, acquired by positive mode ionisation. Coloured by dose group, yellow being controls, blue low dose and red high dose. There does not seem to be overall good discrimination between the groups.

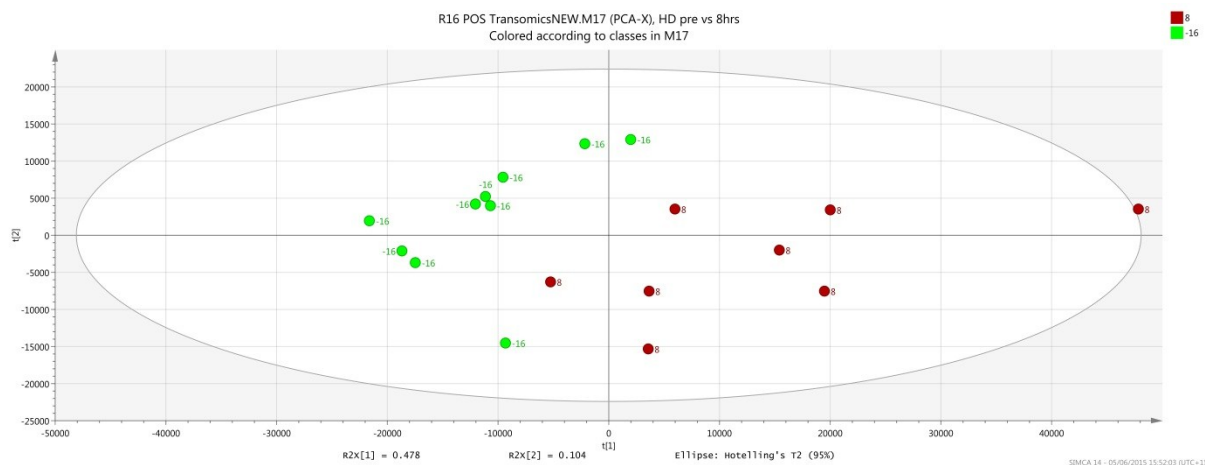


Figure 4.11 PCA scores plot of the model comparing dose groups (controls green and high dose red), at 8 hours post dose, acquired by positive mode ionisation. Maximal response seems to be at 8 hours, with reasonable separation between the groups.

The negative ESI data proved to be more interesting. Initial PCA modelling of the complete dataset revealed distribution patterns within the models which were remarkably consistent and presented an identical separation to those resulting from the ^1H NMR spectral data. Maximal “response” for the high dose group was observed at 8 hours post dose with a diminishing effect at 24 hours after which the metabolic profiles of the animals returned to ‘normal’ (Figure 4.12). Furthermore, due to the greater sensitivity of MS analysis, more subtle differences were detected leading to some differentiation between the low dose and control classes which were previously undetected by ^1H NMR spectroscopy. Subsequent OPLS modelling of the 8 hour time point provided very strong models ($R^2 = 0.99$, $Q^2 = 0.71$), which we used to identify the signals driving the separation.

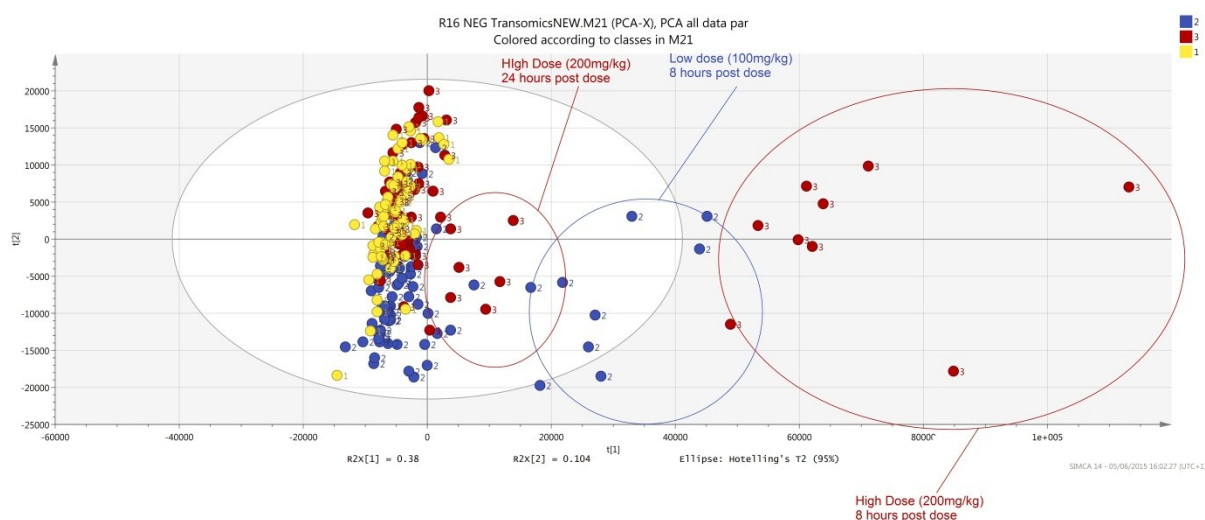


Figure 4.12. PCA scores plot of the model comparing dose groups, acquired by negative mode ionisation. There is good separation between the groups and the distance the samples deviate from the controls seems to be dose dependant.

Examination of the MS data revealed both changes in endogenous compounds and also showed the presence of brominated metabolites of the dosed 2-bromophenol which were readily detected as a result of the characteristic bromine isotope pattern. These data were further examined with OPLS modelling to determine the nature of the excreted metabolites.

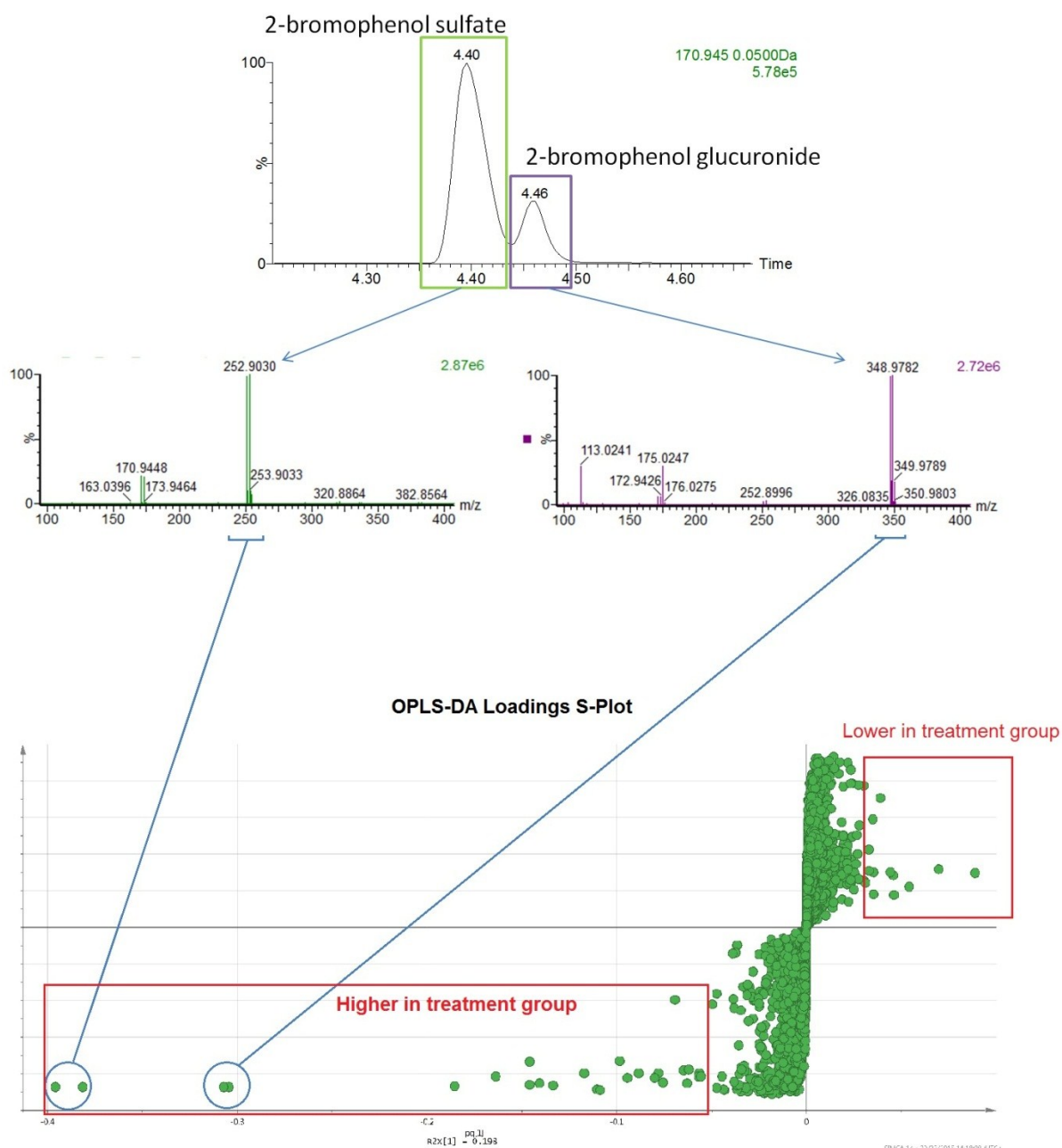


Figure 4.13 Composite plot highlighting the two most prominent 2-bromophenol metabolites. Inspection of the s-plot (bottom) derived from the OPLS model comparing low dose and controls, the ions influencing discrimination relate to 2-bromophenol sulphate (m/z 252.9) and 2-bromophenol glucuronide (m/z 348.97). The chromatographic peak extraction (top) and the related mass spectra are also provided.

The results showed that the two main avenues of phase two xenobiotic metabolism for 2-bromophenol in these animals, as seen via the metabolites excreted in the urine, involved sulfation and glucuronidation. Thus, from both the mass chromatograms and the statistically

discriminating variables within the OPLS models the ions at m/z 252.93 and 348.97 were identified as the key differentiating masses. These ions corresponded to 2-bromophenol sulfate and 2-bromophenol glucuronide respectively (fig 4.13) with both showing the distinctive bromine 1:1 isotope pattern and appropriate losses of 80 and 176 Da, diagnostic for the sulphate and glucuronic acid moieties respectively, due to fragmentation in the ion source. A 2-bromophenol glucuronide dimer, pertaining to m/z 718.94, with a similar retention time to the glucuronide was also seen. However, this was most likely a result of in source dimerization as this reaction is highly unlikely in biology (and elution of both the glucuronide and the dimer, if real, at the same retention time is also highly unlikely). In order to see how well the statistically discriminating variables within the OPLS models had extracted the bromophenol metabolites the LC-MS data were also examined using more traditional mass spectrometric approaches as described below.

4.3.4 Extraction of “isotopobromatograms” from the LC-MS data, application to metabolite profiling and identification

Compounds containing bromine have a very distinctive isotope pattern in mass spectra, with the ^{79}Br and ^{81}Br isotopes present at a ratio of approximately 1:1. This isotopic signature enables the chromatographic data to be searched for this characteristic mass difference, providing a powerful tool for detecting and identifying bromine-containing metabolites. An example is illustrated in Figure 4.14, where an 0-8 hr high dose urine sample (negative ESI data) has been used to generate the chromatogram (Figure 4.14A) by extracting all ions containing an isotopic pattern with a mass difference of 2 Da at a ratio of 1:1, thereby resulting in a “isotopobromatogram” showing only ions containing the bromine isotopic pattern. This selective trace highlights 6 bromine-containing peaks, for which the spectra are detailed in Figure 4.14B.

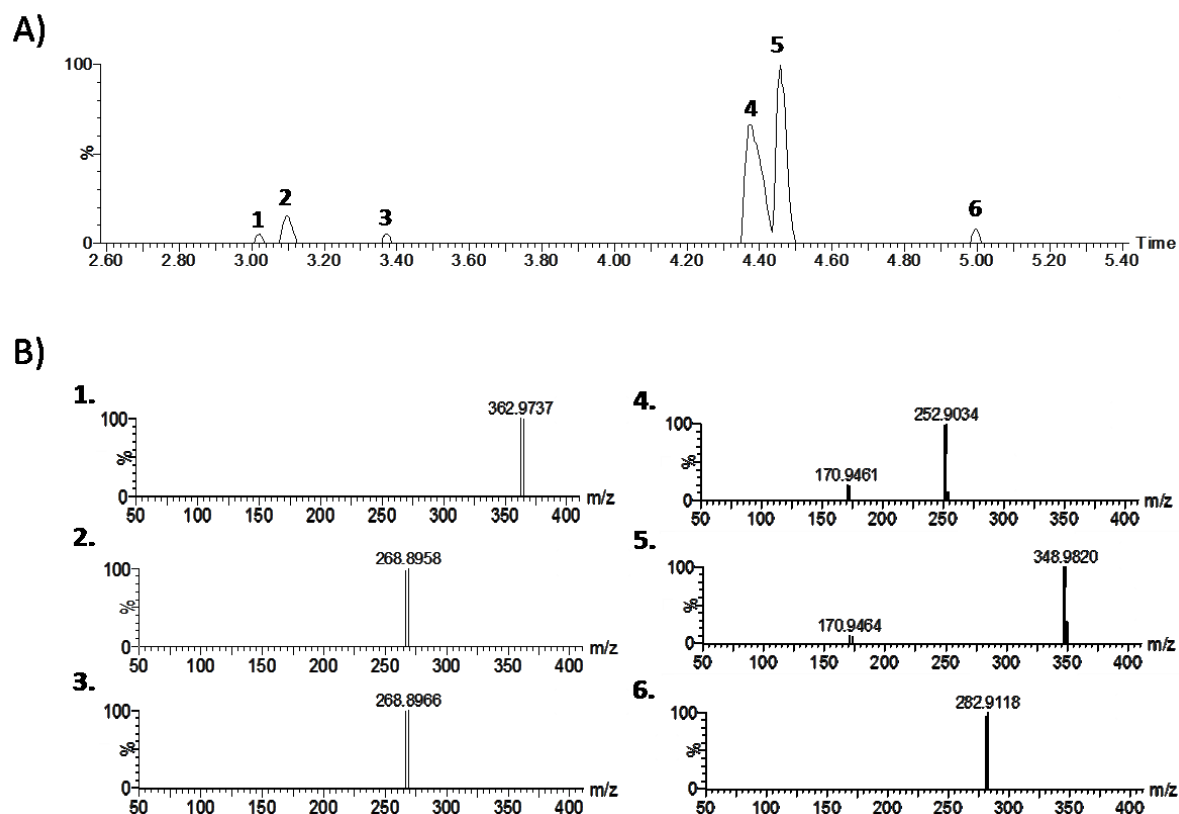


Figure 4.14. Isotopobromatogram (top) and the related mass spectra (bottom). The chromatographic peaks are numbered and match up to the numbered mass spectra.

The peaks detected with the isotopobromatogram and their masses are listed below in table 4.5. The masses and the presence of the bromine, plus the MS/MS data, have led us to postulate some putative assignments but further work would be necessary to confirm the identity of some of these minor 2-bromophenol metabolites (e.g. isolation and ^1H NMR spectroscopy), which appear to be further hydroxylated versions of 2-bromophenol that have also been glucuronidated and sulphated. Peaks 4 and 5 from the above figure corresponded to the 2-bromophenol sulphate and glucuronide metabolites respectively, which is in concordance with the rest of our findings. The MS/MS of the sixth peak contained both the characteristic sulphate 80Da loss; this in addition to the m/z ratio led to the assignment of this compound as 2-hydroxy(3-methoxy)bromophenol sulphate. The methylation of vicinal aromatic diols is quite common within the cell, by enzyme such as

catechol-O-methyltransferase (Huotari *et al.*, 2002). Thus if the parent 2-bromophenol undergoes initial hydroxylation in the 3 position this would produce a vicinal diol which can be readily methylated, and then undergo sulphoconjugation. This would also fit the longer retention observed as the methyl group would increase the lipophilicity of the compound. The MS/MS spectra for the above ions are provided in the appendix.

The standard for 2-bromophenol was also run on the system to ascertain whether we could detect the parent compound however, the compound was not detected under the conditions used. It should be noted that there was no evidence of the presence of glutathione conjugates, nor were any mercapturic acid derivatives detected in the urine. Glutathione conjugates are one of the major indications of potential toxicity via reactive metabolites as they are a direct result of GHS-transferase attack on electrophilic intermediate metabolites (Boyland and Chasseaud, 1969). This suggests that 2-bromophenol, at these doses and in this strain of rat, does not produce reactive intermediates following I.P. administration.

Metabolite	Measured Mass	Retention Time	Formula	IBG
Hydroxy 2-bromophenol glucuronide	362.97	3.01	C ₁₂ H ₁₅ O ₉	1
Hydroxy 2-bromophenol sulphate	268.89	3.09	C ₆ H ₇ O ₅ BrS	2
Hydroxy 2-bromophenol sulphate	268.89	3.39	C ₆ H ₇ O ₅ BrS	3
2-bromophenol sulphate	252.90	4.47	C ₆ H ₆ O ₄ BrS	4
2-bromophenol glucuronide	348.98	4.43	C ₁₂ H ₁₄ O ₈	5
2-bromophenol glucuronide dimer	718.94	4.46	--	5
Hydroxy (3-methoxy) 2-bromophenol sulphate	282.91	5.14	C ₇ H ₈ O ₅ BrS	6

Table 4.3 2-Bromophenol metabolites detected: Putative assignments. IBG: Isotopobromatogram peak number.

4.3.5 Endogenous Metabolic Changes

To investigate any potential changes in endogenous metabolites that had been masked by the presence of the 2-bromophenol metabolites the ions for the sulphate and glucuronide conjugate ions were removed from the MS data and the statistical analysis was rerun. This reanalysis showed that, by both by PCA and OPLS, it was still possible to discriminate the groups, and the signals detected as being responsible are listed in the table below (table 4.4). MS/MS analysis of the ions detected this way allowed for the identification of hippurate, the main discriminating endogenous metabolite detected by ^1H NMR spectroscopy, and citrate to be elevated in the dose group.

Endogenous Metabolites Lower in dose group			
	m/z	Retention time	P(cor)
Unassigned 1	418.12	4.83	0.49
Unassigned 2	253.05	4.44	0.87
Higher in dose group			
	m/z	Retention time	P(cor)
Citric Acid	191.02	0.99	0.68
Hippurate	178.05	3.84	0.65

Table 4.4, Endogenous metabolites altered on 2-Bromophenol administration

From the Analysis of the LC-MS data the clear discriminating factor seems to be the presence of the 2-bromophenol sulphate and glucuronide. The graph below, figure 4.15, depicts the time course of excretion of both metabolites over time in each of the two dose groups. The control samples were blank. It is interesting to note that while these animals were only dosed once, there seems to have been an initial rapid phase of excretion shortly after dose administration, and a second spike after about 48 hours.

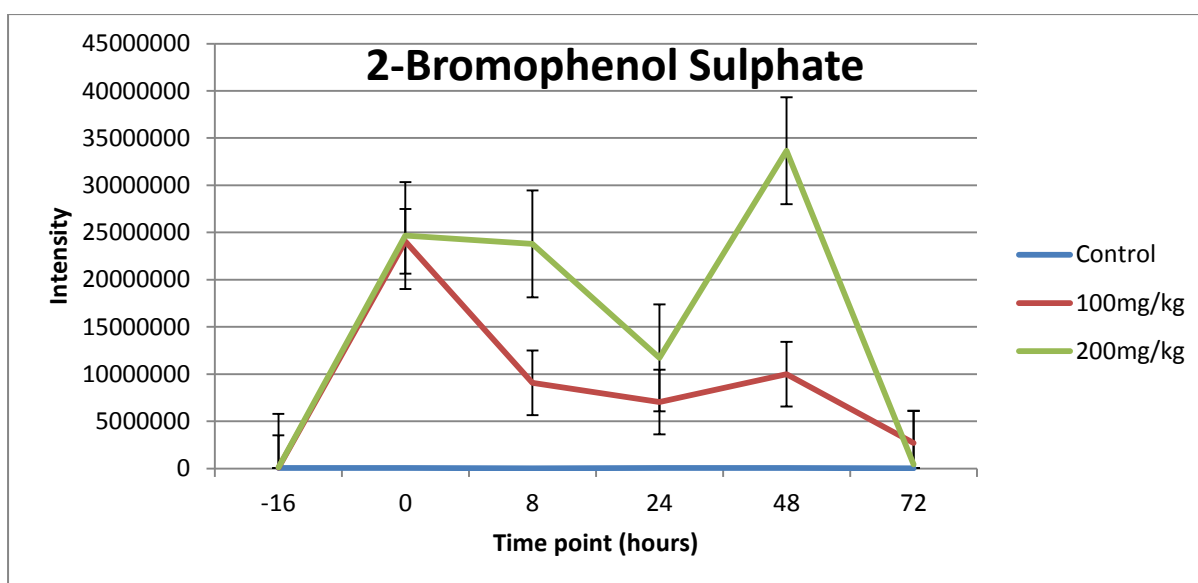
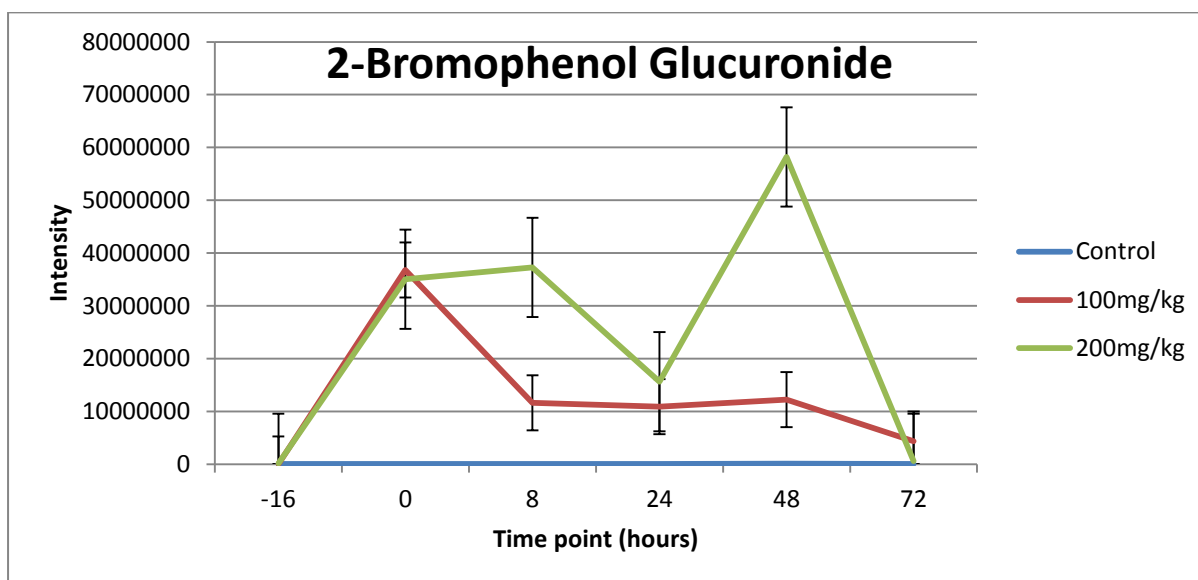


Figure 4.15 Line charts showing the average relative, total volume corrected, intensities for each dose group, across the various time points.

It may also be of worth to note that there seems to be significant inter animal variation in terms of the metabolic fate of 2-bromophenol within the high dose group. Indeed when scrutinising the contribution plots of the high dose samples we can see that there are significant differences from rodent to rodent. The main difference seems to be the ratio

between the sulphate and glucuronide, with some profiles indicating a much higher relative signal from the sulphate conjugate and other animals with more similar signals of the two (figure 4.16). While, due to the varying levels of ionisation being the determinant of peak height, it is not possible to deduce the absolute quantitative ratio, although it is possible to surmise that some animals are saturating their sulphate conjugation pathway capacity much more than others, causing this 'high responder' metabotype. This between-animal variation may due to differential metabolism, due to variations in exposure as a result of the route of administration, genetic, or microbiome differences etc.

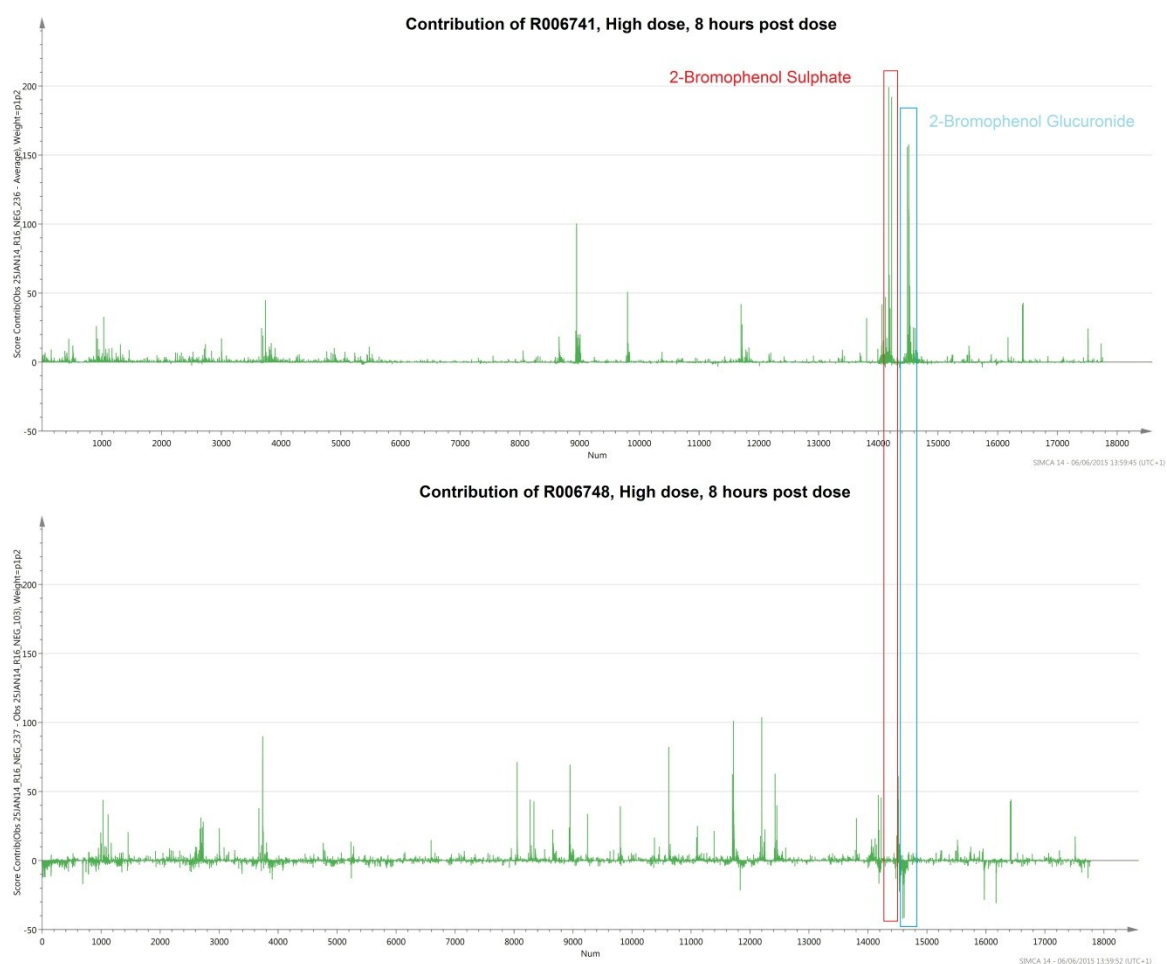


Figure 4.16 Intensity plot showing the contribution of the 2-bromophenol sulphate (highlighted red) and 2-bromophenol glucuronide (highlighted light blue) to the between group discrimination for two different high dose animals at 8 hours post dose from the negative ESI data. The y axis is the scores contribution value for each variable. The fact that the contribution differs greatly between samples at the same time point suggests that the response to bromophenol administration differs between animals.

To validate the model the same permutation testing was employed as in the ^1H NMR spectroscopic data. This is of critical importance for mass spectrometry data as the platform detects many more features than samples and therefore the data is highly multi-collinear. The same 1000 iterations recalculated every 100 permutations were used for each of the models and comparisons performed; the below figure (figure 4.16) presents the permutation plot for the comparison of the high dose and controls at 8 hours post dose (permutation plots for the other models available in the appendices). The criteria for validity are once again that all the Q2 values of the random data are lower than the original and that the Q2 regression line intersects the vertical axis at or below 0. Indeed once again our model fulfils these criteria. Also included are the CV-anova statistics in the table below for completeness (table 4.5).

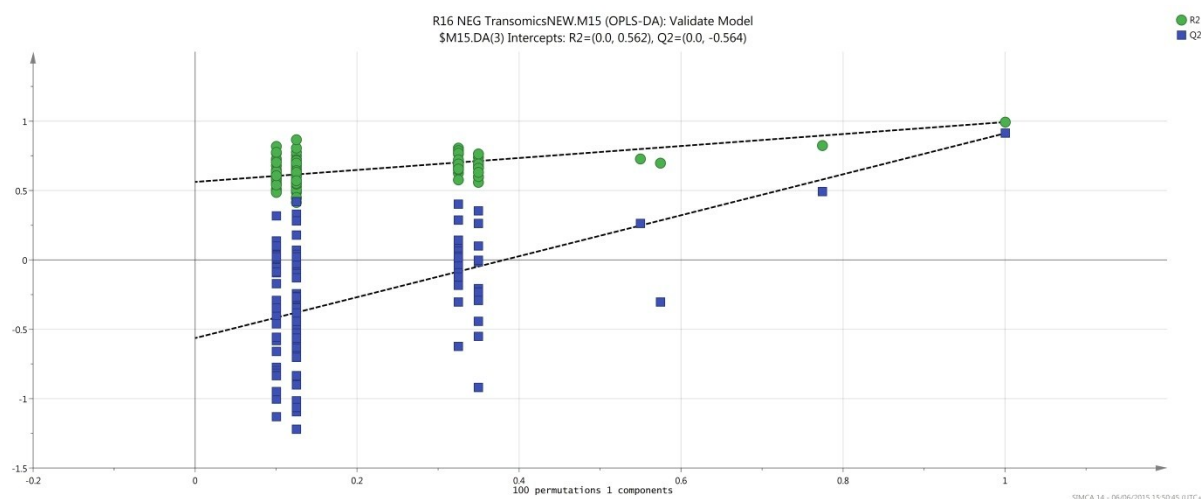


Figure 4.17 Permutation plot for the model constructed. The green and blue dots represent the R2 and Q2 of the dummy data respectively. The permutation values are lower than the actual values suggesting that the model is highly robust.

Control vs High Dose (8hours)	Sum of Squares	Degrees of Freedom	Mean Squares	F value	P value	Standard Deviation
Total corr.	17	17	1			1
Regression	15.5075	4	3.87686	33.7672	9.40394e-007	1.96898
Residual	1.49255	13	0.114812			0.338839

Table 4.5 Results of the CV ANOVA significance test

4.3.6 Histopathology and Clinical Chemistry

The main objective of the original COMET study was to identify the main metabolic changes related to 2-bromophenol induced nephrotoxicity and determining this was also assessed using classical toxicological methods. Both kidney and liver samples were taken from the animals in the study for histopathological assessment, as were clinical chemistry tests undertaken on both the urine and blood samples. These data were provided by the companies in the COMET group who performed the live phase of the study.

a. Histopathology

The administration of bromophenol has been previously found to cause nephrotoxicity characterised by necrotic damage to the proximal tubule cells. Information on whether toxicity was achieved by the current study was provided by the histopathology report (see appendix). It was expected that the high dose would cause changes in the kidney but no mortality. The high dose group showed mild clinical signs, including transient tremors, slight laboured respiration and inactivity; however these effects were minor and the animals had recovered by days 2-3 post dose.

Inspection of the kidneys revealed that there was minimal general non-specific nephropathy present in the kidneys of all animals (dosed and control), which was however, not severe enough to cause any symptoms, although two low dose group animals did show signs of minimal localised inflammation. Two animals out of ten from the High dose group showed evidence of slight tubular regeneration, characterized by pale basophilic epithelial cells and necrotic cellular debris within the lumen of the proximal tubules. However, this is a low number and the rest of the animals in the group showed no kidney damage whatsoever. Furthermore one control animal and one high dose animal also showed signs of hyaline

droplet formation. Hyaline droplets are cytoplasmic vesicles containing amorphous materials surrounded by a membrane and are thought to originate from a loss of glycocalyx in normal apical vesicles. Hyaline droplets have been linked to tubular necrosis and are thought to be indicative of degenerating proximal tubular epithelium (Sato *et al.* 2005). This finding would indicate a low level background of tubular damage in the test animals. Overall, there seemed to be no difference between the kidney tissues of any of the groups.

While bromophenol is thought to be a kidney toxin, for completeness the livers were also collected and inspected. Surprisingly the livers of every animal in the study, except one low dose animal, showed the presence of cytoplasmic vacuolation of hepatocytes and inflammation. This cytoplasmic accumulation of glycogen in liver cells is seen consistently after mild acute liver injury or the introduction of hepatotoxins. This alteration is believed to protect the cells from further damage by increasing toxin resistance and blocking cellular degeneration (Nayak *et al.* 1996). This observation raises questions about the overall health of the animals in the study, as these signs cannot be related to the administration of the bromophenol.

b. Clinical Chemistry

Clinical chemistry showed that there were negligible pathology changes in the low dose group, and only one of the animals in the high dose group. This particular animal showed increased blood urea nitrogen and glucose, in addition elevated total bilirubin, serum potassium, serum phosphorus and decreased serum sodium, chloride, and altered albumin/globulin ratio. The clinical chemistry of the urine samples showed no trends and all animals were within the normal ranges for each test. The overall results of the clinical chemistry are summarised in table 4.6.

Treatment	Dose mg/kg/day	Protein	glucose	Creatinine	Potassium
8 Hours					
Vehicle (corn oil)	0	0.23 ± 0.11	1.19 ± 0.38	8.55 ± 1.33	236.4 ± 37.6
2-Bromophenol	100	0.18 ± 0.04	1.05 ± 0.27	6.92 ± 1.43	217 ± 43.9
2-Bromophenol	200	0.27 ± 0.12	0.85 ± 0.27	5.21 ± 1.77	164.7 ± 51.6
DAY 1					
Vehicle (corn oil)	0	0.31 ± 0.2	1.52 ± 0.28	9.85 ± 1.23	389.5 ± 52
2-Bromophenol	100	0.239 ± 0.06	1.43 ± 0.35	8.18 ± 1.4	340.8 ± 71.2
2-Bromophenol	200	0.44 ± 0.13	1.45 ± 0.32	8.57 ± 2.03	317 ± 81.6

Table 4.6, Clinical chemistry data from the COMET report

4.4 Summary and Conclusions

Previous work had identified 2-bromophenol, a phenolic metabolite of bromobenzene, as a potential nephrotoxin (Lau *et al*, 1984b). This study in Sprague Dawley rats, was part of the COMET 1 program and was performed to assess the nephrotoxicity of the compound. After combined ^1H NMR spectroscopic and UPLC-MS metabolic profiling, statistical analysis showed that there was a significant metabolic perturbation due to administration of the compound, but clinical chemistry and histopathological examination of the kidneys showed that the toxic effects of the compound were minimal. Effects were noted on endogenous metabolites such as hippuric and *p*-aminohippuric acids by ^1H NMR spectroscopy and statistical analysis of the LC-MS data also showed effects of 2-bromophenol administration. However, in the case of LC-MS the major discriminating ions were due not to changes in endogenous compounds but to the detection of the excretion of the sulphate and glucuronide conjugates of 2-bromophenol. More detailed analysis of the urine via bromine-isotope specific detection revealed the presence of a number of minor, hydroxylated conjugates of 2-bromophenol but there was no evidence of glutathione cysteinyl or metcapturic acid conjugates. The absence of these “markers” of the formation of reactive metabolites is consistent with apparent absence of kidney damage indicated by

metabonomic phenotyping, clinical chemistry and histopathology. Statistical interrogation of the data obtained in this study shows that, whilst the major xenobiotic metabolites was revealed reasonably efficiently, the detection of minor metabolites required a more targeted approach.

The overall conclusion from this study with respect to nephrotoxicity was therefore that, after single doses of either 100 or 200 mg/kg, 2-bromophenol is not a nephrotoxin. While these results do agree with the conclusions reached during the original COMET project (see internal report, appendix) they are in contrast to previous work by Lau *et al.* It is, however, possible that achieving toxicity requires greater exposure to 2-bromophenol, either through higher single doses or repeated administration. To investigate the potential for repeat dosing of 2-bromophenol to cause nephrotoxicity it was therefore decided to do a repeat dose study as described in chapter 5.

4.5 References

- Bairaktari, E., Liamis, G., Tsolas, O., Elisaf, M., (2001). Partially reversible renal tubular damage in patients with obstructive jaundice. *Hepatology*, **33**, 1365 – 1369.
- Bales, J. R., Higham, D. P., Howe, I., Nicholson, J. K., Sadler, P. J. (1984). Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clinical Chemistry*, **30**, 426–32.
- Berndt W.O. and Grote D., (1968) The accumulation of ¹⁴C-dinitrophenol by slices of rabbit kidney cortex. . *Journal of Pharmacology, Experimental Therapeutics*, **164**, 223–231.
- Blaise, B.J., Shintu, L., Elena, B.N.D., Emsley, L., Dumas, M.-E., Toulhoat, P. (2009). Statistical Recoupling Prior to Significance Testing in Nuclear Magnetic Resonance Based Metabonomics. *Analytical Chemistry*, **81**, 6242-6251.
- Boyland, E., Chasseaud, L.F., (1969). Role of glutathione and glutathione S-transferases in mercapturic acid biosynthesis. *Advances in Enzymology*, **32**, 173–177.
- Bruchajzer, E., Szymanska, J.A., Piotrowski, J.K., 2002. Acute and subacute nephrotoxicity of 2-bromophenol in rats. *Toxicology Letters*, **5**, 245–252.
- Christians, U. , Schmitz, V. , Schöning W., (2008). Toxicodynamic therapeutic drug monitoring of immunosuppressants: promises, reality, and challenges. *Therapeutic Drug Monitoring*, **30**, 151–158.

Huotari, M., Gogos, J.A., Karayiorgou, M., Koponen, O., Forsberg, M., Raasmaja, A., Hyttinen, J., Männistö, P.T., (2002). Brain catecholamine metabolism in catechol-O-methyltransferase (COMT)-deficient mice. *European Journal of Neuroscience*, **15**, 246-256.

Lau S.S., Monks T.J., Greene K.E., Gillette J.R., (1984a). The role of *ortho*-bromophenol in the nephrotoxicity of bromobenzene in rats. *Toxicology and Applied Pharmacology*, **72**, 539-549.

Lau S.S., Monks T.J., Gillette J.R., (1984b). Identification of 2-bromohydroquinone as a metabolite of bromobenzene and *o*-bromophenol: Implications for bromobenzene-induced nephrotoxicity. *Journal of Pharmacology, Experimental Therapeutics*, **230**, 360-366.

Lau, S.S. and Zannoni, V.G. (1981). Bromobenzene epoxidation leading to binding on macromolecular protein sites. *Journal of Pharmacology, Experimental Therapeutics*, **219**, 563–572.

Lees, H. J., Swann, J. R., Wilson, I. D., Nicholson, J. K., Holmes, E., (2013). Hippurate: The Natural History of a Mammalian-Microbial Cometabolite. *Journal of Proteome Research*, **12**, 1527–1546.

Lindon, J.C., Keun, H.C., Ebbels, T.M., Pearce, J.M., Holmes, E., Nicholson, J.K., (2005). The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. *Pharmacogenomics*, **6**, 691–699.

Lu, R., Chan, B.S., Schuster, V.L., (1999). Cloning of the human kidney PAH transporter: Narrow substrate specificity and regulation by protein kinase C. *American Journal of Physiology*, **276**, 295-303.

Madhu, C. and Klaassen, C.D., (1992). Bromobenzene-glutathione excretion into bile reflects toxic activation of bromobenzene in rats. *Toxicology Letters*, **60**, 227–236.

Nayak, N.C., Sathar, S.A., Mughal, S., Duttagupta, S., Mathur, M., Chopra, P., (1996). The nature and significance of liver cell vacuolation following hepatocellular injury — an analysis based on observations on rats rendered tolerant to hepatotoxic damage. *Virchows Archives*, **428**, 353–365.

Niwa, T., (1996). Organic acids and the uremic syndrome: protein metabolite hypothesis in the progression of chronic renal failure. *Seminars on Nephrology*, **16**, 167 –182.

Parke, D.V. and Piotrowski, J.K., (1996). Glutathione: its role in the detoxication of reactive oxygen and environmental chemicals. *Acta Toxicologica*, **4**, 1–13.

Psihogios, N. G., Kalaitzidis, R. G., Dimou, S., Seferiadis, K. I., Siamopoulos, K. C., Bairaktari, E. T., (2007). Evaluation of tubulointerstitial lesions ' severity in patients with glomerulonephritides: an NMR-based metabonomic study. *Journal of Proteome Research*, **6**, 3760 – 3770.

Reid, W.D., (1973). Mechanism of renal necrosis induced by bromobenzene or chlorobenzene. *Experimental Molecular Pathology*, **19**, 197-214.

Rush, G.F., Kuo, C., Hook, J.B., (1984). Nephrotoxicity of bromobenzene in mice. *Toxicology Letters*, **20**, 23-32.

Sakai, T., Ninuma, Y., Yanagihara, S., Ushio, K., (1983). Simultaneous determination of hippuric acid and o-, m- and p-methylhippuric acids in urine by high-performance liquid chromatography. *Journal of Chromatography*, **276**, 182 – 188.

Sato, S., Kitamura, H., Ghazizadeh, M., Adachi, A., Sasaki, Y., Ishizaki, M., Inoue, K., Wakamatsu, K., Sugisaki, Y.. (2005). Occurrence of hyaline droplets in renal biopsy specimens: an ultrastructural study. *Medical Molecular Morphology*, **38**, 63-71.

Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H. (1997) Expression cloning and characterization of a novel multispecific organic anion transporter. *Journal of Biological Chemistry*, **272**, 18526–18529.

Stonehouse, J., and Keeler, J., (1995). A Convenient and Accurate Method for the Measurement of the Values of Spin-Spin Coupling-Constants. *Journal of Magnetic Resonance A*, **112**, 43.

Tautenhahn, R., Patti, G.J., Rinehart, D. Siuzdak, G. (2012). XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Analytical Chemistry*, **84**, 5035–5039.

Veselkov, K.A., Lindon, J.C., Ebbels, T.M.D., Crockford, D., Volynkin, V.V., Holmes, E., Davies, D.B., Nicholson, J.K. (2008). Recursive Segment-Wise Peak Alignment of Biological ¹H NMR Spectra for Improved Metabolic Biomarker Recovery. *Analytical Chemistry*, **81**, 56-66.

Want, E.J., Wilson, I.D., Gika, H., Theodoridis, G., Plumb, R.S., Shockcor, J., Holmes, E., Nicholson, J.K. (2010). Global metabolic profiling procedures for urine using UPLC-MS. *Nature Protocols*, **5**, 1005-1018.

Investigation of the metabolism and potential nephrotoxicity of 2-bromophenol after chronic exposure

Chapter Five

Previous work described in chapter 4 found that there were no toxic effects of 2-bromophenol after a single dose in Sprague Dawley rats. A repeat study was therefore performed to identify the effects of 2-bromophenol after chronic exposure. UPLC-MS and ¹H NMR spectroscopy was performed on urine and plasma samples acquired from Han Wistar rats administered with 2-bromophenol for 3 days. ICP-MS was also performed to identify total bromine concentrations. Data analysis identified 2-bromophenol glucuronide and 2-bromophenol sulphate as the major metabolites of 2-bromophenol. However no toxic metabolites were detected and after analysis of the histopathology and clinical chemistry, there was still little evidence for any toxic action of 2-bromophenol even after repeated dosing.

SUMMARY:

- 2-bromophenol, a metabolite of the hepato- and nephrotoxic compound bromobenzene, was also suspected of being a nephrotoxin. However, previous work showed only limited effects after a single dose.
- In a further study 20 Han Wistar rats were split into five equal groups and were administered three IP doses of vehicle (corn oil), 100mg/kg or 200mg/kg of 2-bromophenol at 24 hour intervals. Urine and blood were collected at regular intervals. Three groups (12 rats) were sacrificed after four days and the remainder after seven days and the kidneys retrieved.
- Urine samples were analysed by ^1H NMR spectroscopy and UPLC-MS to determine both the effects of 2-bromophenol on urinary endogenous metabolite profiles and the metabolic fate of the compound itself.
- Chemometric analysis showed that the main discriminators between dosed and control animals were associated with metabolites of 2-bromophenol. Little evidence of toxicity was seen in terms of altered endogenous metabolite profiles, histopathology or clinical chemistry.
- The same 2-bromophenol metabolites were identified in urine as shown in for the single dose study reported in Chapter 4, and included 2-bromophenol glucuronide and 2-bromophenol sulphate. Again there was no evidence for the production of reactive metabolites.
- ICP-MS was performed to determine total bromine concentrations in urine. The resulting data showed that there was the expected elevation of bromine containing compounds in the dose groups. Bromine concentrations remained elevated in the recovery group suggesting that 2-bromophenol related material is excreted slowly.

5.1 Introduction

As discussed in Chapter 4 Lau *et al.*, (1984a; 1984b), in a series of studies in Sprague Dawley rats have previously investigated the mechanisms responsible for the toxicity and the metabolism of bromobenzene. These studies had concluded that 2-bromophenol was one of the toxic downstream metabolites of bromobenzene (the metabolism of bromobenzene and 2-bromophenol has already been discussed in Chapter 4, and will not be repeated here).

However, as described, analysis of samples also obtained from Sprague Dawley rats dosed with 2-bromophenol after administration of a single dose of 2-bromophenol at either 100 or 200 mg/kg, as part of the COMET program, failed to reveal any discernible toxic effects. Thus, histopathological examination of the kidneys of these animals did not indicate any tissue damage nor were there any indications of raised amounts of traditional clinical chemistry markers of kidney damage. Both ^1H NMR spectroscopy and LC-MS were employed to metabolically phenotype the urine samples collected from control and dosed animals but, while LC-MS (but not ^1H NMR spectroscopy) was able to detect metabolites of 2-bromophenol, there were no elevations of the typical biomarkers of toxicity, such as glucose or lactate. The major metabolites of 2-bromophenol were identified by LC-MS to be the glucuronide and sulphate conjugates and, with respect to endogenous compounds, there were notable effects on the excretion of hippurate, *p*-aminohippurate, 3-methylhistamine and citric acid.

5.1.2 Cumulative Toxicity

Clearly these results do not align with the findings of Lau *et al.*, who reported “mild toxicity” on administration of 2-bromophenol. One potential limitation of the COMET study that could be advanced to explain this difference is that the exposures achieved failed in some way to replicate those obtained by Lau *et al.*, and that either an increased dose, or multiple

administration of 2-bromophenol might be required to produce a toxic effect. Indeed, the toxic effects of a wide variety of compounds are cumulative. Many chemicals thought to be either innocuous, or of low order toxicity, such as the agricultural compounds ethylene oxide aerosols (Klonne *et al.*, 1987) and methomyl (Tanaka *et al.*, 1988) have been shown to be toxic after repeated exposure. Furthermore many well studied drugs, such as the cancer drugs cisplatin (Goren *et al.*, 1986) and ifosfamide (Loebstein and Koren, 1998) have been shown to exhibit cumulative toxicity. In fact, the pre-clinical toxicity evaluation of ifosfamide failed to detect any significant toxic effects, and its potential for nephrotoxicity was not discovered until much later (Barnen, 1982). Therefore, it followed logically from the previous single dose study that there was a need to assess the potential for cumulative 2-bromophenol toxicity following multiple doses. As such the aims of the work presented here were to investigate the effects of the repeated administration of 2-bromophenol on both on the metabolism of the compound itself and its effects on the kidney.

5.2 Materials and Methods

5.2.1 Study Design

Urine and serum samples were collected from a 2-bromophenol toxicology study where the compound was administered intraperitoneally as a repeated dose, every 24 hours for three administrations, to male Han Wistar rats. The rats were treated with 100 mg/kg (4 rats) or 200 mg/kg (8 rats) in corn oil. Control rats were treated with corn oil only (8 rats). Rats were housed in metabolic cages for 6 hours daily (between 08:00 and 14:00) and urine was collected for metabonomic analysis. 0.3ml of whole blood was collected from the tail vein at 0.1, 6, 9, 24 and 48hrs post dose for the sacrifice group A (three dose groups) and at 24, 48, 72, 96, 120 and 144hrs post dose for sacrifice group B (two dose groups). Blood samples were taken in lithium heparin tubes and centrifuged for 3 minutes at 1500 g under refrigeration (2 to 8°C). Plasma was transferred into 0.5 mL polypropylene tubes and kept

frozen at -80°C until analysis. There were four animals for each of the 5 dose groups. Selected organs were examined histopathologically, and the kidneys removed and stored for analysis.

A more detailed description of the study design is given in the Protocol in the appendix.

5.2.2. Sample Preparation and Experimental for ^1H NMR spectroscopy

^1H NMR (750 MHz) of the plasma and urine was performed at AstraZeneca, Alderley Park, using a flow-injection process.

a. Plasma

The collected plasma samples were centrifuged for 10 min at 10621 g at 4°C . 400 μl of each sample was aliquoted and 100 μl of D_2O added. The samples were then manually transferred into 5mm NMR tubes.

b. Urine

The collected urine samples were centrifuged for 10 min at 10621 g at 4°C in Eppendorf tubes. Then 400 μl of each sample was aliquoted into an Eppendorf tube and 100 μl of 0.1% TSP/ D_2O and phosphate buffer at pH added. The samples were then manually transferred into 5mm NMR tubes.

A subset of samples ($n = 18$) for both plasma and urine samples was then selected which it was hoped would be representative of the whole study. These samples included controls (pre-dose and 24 hours), high dose (pre-dose and 24 hours) and recovery animals (24 and 144 hours). Both plasma and urine samples were then run on a 5 mm Prodigy TCI cryogenically cooled probe in a Bruker Avance III Spectrometer operating at 700 MHz. The ^1H NMR spectra of the urine were measured using one-dimensional (1D) nuclear Overhauser enhancement spectroscopy (NOESY) presat (noesypr1d) pulse sequence

(Stonehouse and Keeler, 1995); NOESY was preferred over CPMG as this is not routinely used for water suppression in urine and plasma. The data were acquired using Topspin 2.1.

5.2.3 ^1H NMR Spectroscopic Data analysis

The same computational methodology was employed for the chemometric analysis of both the urine and plasma ^1H NMR datasets. After phasing to correct for first and zero-order errors, and baseline correction to account for distortions of the base values, the data was imported to MATLAB R2012b (Natick, Massachusetts, U.S.A.) using the Metaspectra program script, using a resolution of 0.00055. The data were then aligned using Recursive Segment-Wise Peak Alignment (RSPA) (Veselkov et al., 2008) and normalised by Probabilistic Quotient to account for, and exclude, systematic sources of bias within samples which were not due to biological processes or environment. Statistical Recouping of Variables (SRV) (Blaise et al., 2009) was then used for peak detection, and the integrated dataset was exported to SIMCA-P+ 13.0.2 software for multivariate analysis.

5.2.4 Sample Preparation and Experimental for UPLC-MS

UPLC-MS analysis was performed at Imperial College London.

Prior to analysis the sample order was randomised using the Microsoft excel random number generator function. Only urine samples were included in the analysis. A 150 μL aliquot of urine was mixed with 300 μL water to dilute the salt concentration before protein removal via precipitation with acetonitrile (1:2 v/v). These samples were vortexed and left at $-20\text{ }^\circ\text{C}$ overnight before centrifugation for 5 min at 12 000 g at $4\text{ }^\circ\text{C}$. For reversed-phase chromatography 50 μL of sample was taken and added to 150 μL water in 350 μL 96-well plates (Waters Corporation, Milford, USA). A pooled quality control (QC) sample was prepared (Want et al., 2010) by combining 100 μL of each sample and diluting 1:4 with water. The 96-well plate was stored at $-20\text{ }^\circ\text{C}$ until the analysis by LC-MS as described below. Immediately before being placed in the autosampler the plate was centrifuged at 1700 G before being placed into the autosampler at $4\text{ }^\circ\text{C}$.

Reversed-phase liquid chromatography was performed on an Acquity UPLC system (Waters Ltd, Elstree, UK) coupled to a Synapt G2 triple quadrupole dual time-of-flight mass spectrometer (Waters MS technologies, Ltd., Manchester). UPLC-MS conditions for urine analysis were those previously described for the reversed-phase analysis of urine (Want et al., 2010) with the separation performed using a Waters HSS T3 (1.8 μm , 2.1 x 100 mm) column. For sample analysis aliquots of 2 μL per sample were injected. The chromatographic mobile phase was composed of 0.1 % formic acid in water (A) and 0.1 % formic acid in acetonitrile (B). The column temperature was maintained at 40 °C and linear gradient elution was performed at either 0.5 mL/min for the 2.1 x 100 mm column or 0.4 mL/min for the 1 x 50 mm column. The starting composition was 1 % B, held for 1.0 min before increasing to 15 % at 3.0 min, 50 % at 6.0 min, 95 % at 9.0 min for a 1.0 min wash and returning to 1 % B for a 2 min re-equilibration step (total cycle time 12 min). Mass spectrometry was performed with electrospray ionisation operated in negative ion mode (ESI⁻). The instrument was operated in sensitivity mode and set to acquire data over the m/z range 50 – 1200 with a scan time of 0.1 s. All mass spectral data were collected in centroid mode using the MS^e data acquisition function. For mass accuracy, leucine enkephalin (MW = 555.62) was used as a lock mass at a concentration of 200 pg/ μL (in 50/50 CH₃CN:H₂O, 0.1 % formic acid) infused at a flow rate of 20 $\mu\text{L}/\text{min}$ via a lock spray interface. The data were collected using MassLynx V 4.1 software (Waters Corporation, Milford, USA).

5.2.5 UPLC-MS Data analysis

The UPLC-MS data set was pre-processed using Transomics (Non Linear Dynamics, UK) software in order to convert the three-dimensional LC-MS raw data into time-aligned detected features of retention time, m/z ratio and intensity (peak area). The ion intensities for each peak detected were then normalised, within each sample, to the median peak intensity in that sample. The resulting peak marker tables, comprising m/z , RT and normalised intensity values for each variable in every sample, were exported to SIMCA-P+ 13.0.2 software (Umetrics, Umea, Sweden) to perform multi-variate statistical analysis.

5.2.6 Statistical Modelling

Prior to any statistical modelling the QCs were evaluated, through both the use of CV and positioning within PCA plots. If there was significant variation within the quality control data, or if there was significant drift within the position of the QCs within PCA scores plots the data would have been unfit for analysis. SIMCA-P+ 13.0.2 was used for the statistical modelling; Principal component analysis (PCA) (unsupervised) and Partial least square regression discriminant analysis (OPLS-DA) were utilised to explore any underlying stratification within the data. Scores plots were used to investigate any differences within the sample groups and the loadings were used to identify which (if any) variables were driving the separation. Validation was also performed to test the predictive capabilities of the models created. Each peak or feature of interest was then tested for significance using a modified Multiple Analysis of Variance (MANOVA) and the p values corrected for multiple testing using Benjamini-Yekutielli correction. Metabolite assignment was then performed.

5.2.7 ICP-MS

ICP-MS data was provided by Dr David Thompson, Keele university. Concentrations of Br containing material in the urine were calculated in reference to a standard curve prepared using *o*-bromophenol. Ten calibration standards were prepared in control urine covering the range 0.00004 to 4 mg mL⁻¹. 50 µL of each standard was directly injected, into a NexION 350 ICP mass spectrometer (Perkin Elmer UK Ltd.), set to monitor ⁷⁹Br and ⁸¹Br, and the response measured over 1.5 minutes. The NexION ICP-MS was operating at the conditions as described in Duckett *et al.*, (2015). The measured peaks were then used to construct a standard curve. Urine samples were then directly injected, in duplicate, into the ICP-MS, set to detect both ⁷⁹Br and ⁸¹Br in addition to ¹⁹²Ir as an internal standard. The responses were measured over 1.5 minutes and then total bromine concentrations were determined by

utilising the reference curve. The data were collected using MassLynx V 4.1 software (Waters Corporation, Milford, USA)

5.3 Results and Discussion

5.3.1 Cryogenically cooled ^1H NMR spectroscopy

The ^1H NMR spectroscopic analysis of 2-bromophenol described for the COMET sample set in Chapter 4 was performed on conventional 600 Mhz instruments. As there were few discriminating signals detected from this methodology the application of a cryoflow probe aimed to improve coverage as it has been shown to have the ability to improve the limit of detection (Spraul *et al.*, 2003). By supercooling the radio-frequency coils of the instrument, the signal to noise ratio is substantially increased. Furthermore by cooling the tuning and matching circuits the random thermal motion of electrons in the conductors is reduced and the resistivity of the pure metals also decreases with lower temperature, all of which additionally reduces the noise. The preamplifier, filters and transmit-receive-switch are also cooled in order to improve the noise figure of the electronic components. This also allows for the use of smaller volumes of sample (100 μl as opposed to the 500 μl that was necessary for conventional ^1H NMR spectroscopy) which was crucial as there were limited sample volumes available and a wide range of analyses were planned. It was also hoped that the application of cryoprobe technology would result in the detection of 2-bromophenol and its metabolites which, whilst readily detected by UPLC-MS, were not observed using ^1H NMR spectroscopy for analysis of the COMET samples, and were assumed to be previously below the limit of detection.

a. Plasma

Analysis of plasma samples via the cryoprobe study showed that the detection of low concentration endogenous metabolites was substantially increased and sharper peaks were produced when compared to the ^1H NMR spectral data from the COMET study.

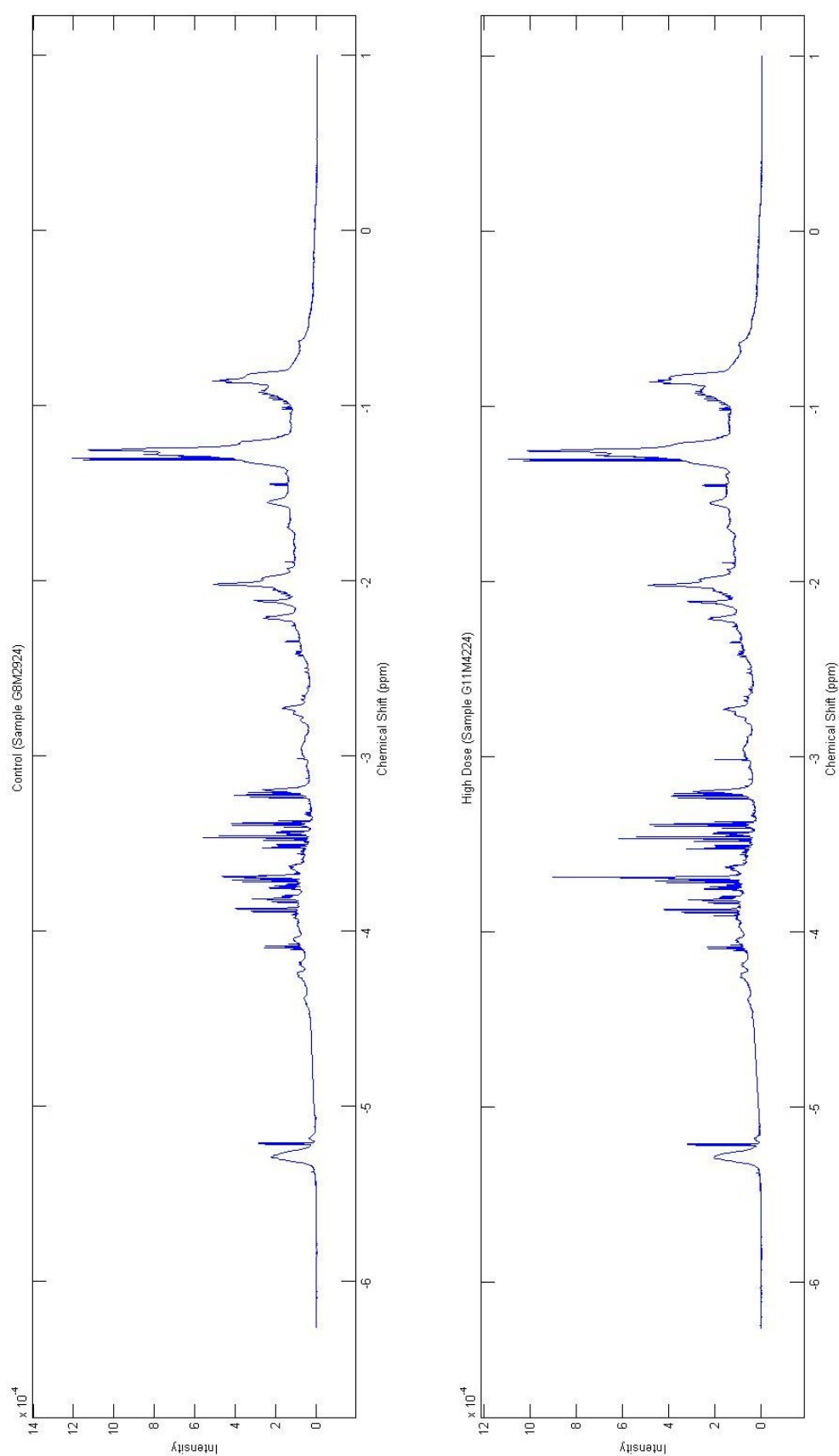


Figure 5.1 A comparison of the typical control (top) and high dose (bottom) NMR spectra of plasma obtained at 24 hr post dose with either corn oil or 200 mg/kg 2-bromophenol.

Chemometric analysis of the plasma data identified a clear separation between the high dose treatment groups (200 mg/kg) and the controls in statistical space when PCA modelling was employed (figure 5.2a). Maximal response was identified to be 24 hours post dose, as seen in figure 5.2b, which is in agreement to the results presented for urine in Chapter 4 for the single dose study.

This trend was also mirrored in the OPLS-DA modelling performed, and comparisons of the control and high dose group at 24 hours (figure 5.2c) showed a clear metabolic difference between the animals. However, when investigating any time related effects it became apparent that there was little metabolic difference between high dose and recovery group (96hrs after dosing termination/144hrs post initial treatment) (figure 5.2d) This would suggest either that the systemic metabolic changes from the xenobiotic administration are irreversible or insufficient time had passed for significant metabolic reversal. The OPLS-DA statistical models produced achieved high correlation values and predictive power ($R^2 = 0.78$, $Q^2 = 0.61$) (figure 5.2) and allowed for the identification of the signals influencing the separation most. However, due to low sample size ($n=18$), despite the clear metabolic difference, individual variables failed to pass the multiple testing correction, with the exception of $\delta 3.54$, Glycine ($p = 0.036$) which was increased in the high dose group (figure 5.3c).

The role glycine may have in the response to 2-bromophenol administration is not clear. Glycine is a non-essential amino acid and has various roles in metabolism, including bile acid conjugation and other endogenous metabolic pathways (Robert, 1986). Biologically it has a role in inflammation and protects against renal and hepatic toxicity. Glycine possess vasodilatory activity in renal vessels and can increase renal blood flow and glomerular filtration rates (Brezis *et al.*, 1984). Glycine also is involved in the protection of hepatic cells from oxidative stress and attenuates the glutathione depletion caused by many hepatotoxins (Deters *et al.*, 1998).

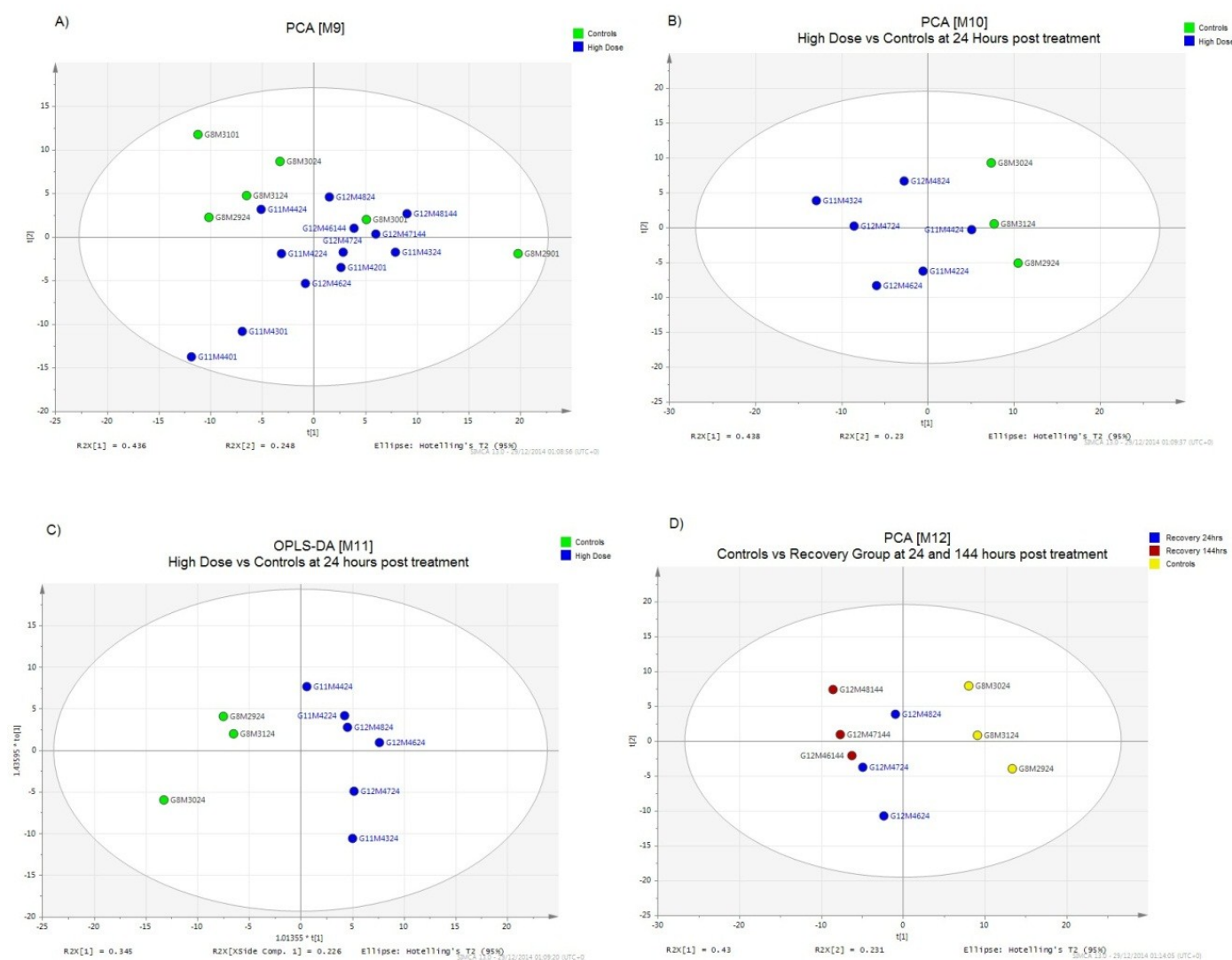


Figure 5.2. The scores plots from the multivariate analysis of the ^1H NMR spectral data obtained from the plasma of the multiple dose 2-bromophenol study, coloured by dose group. A) Is the PCA scores plot comparing the high dose and control groups. B) Is the PCA scores plot comparing the high dose and control animals at 24 hours post initial dose, which is the time point of maximal response. There is a clear metabolic difference between the two dose groups. C) Is the OPLS-DA scores plot of the same comparison, further highlighting the dose related trend. D) Is the PCA scores plot comparing the controls with the two different time points of the recovery group. There is little difference between 24 hours (maximal response) and 144 hours (3 days after dosing has ended) indicating that the metabolic perturbation has not resolved.

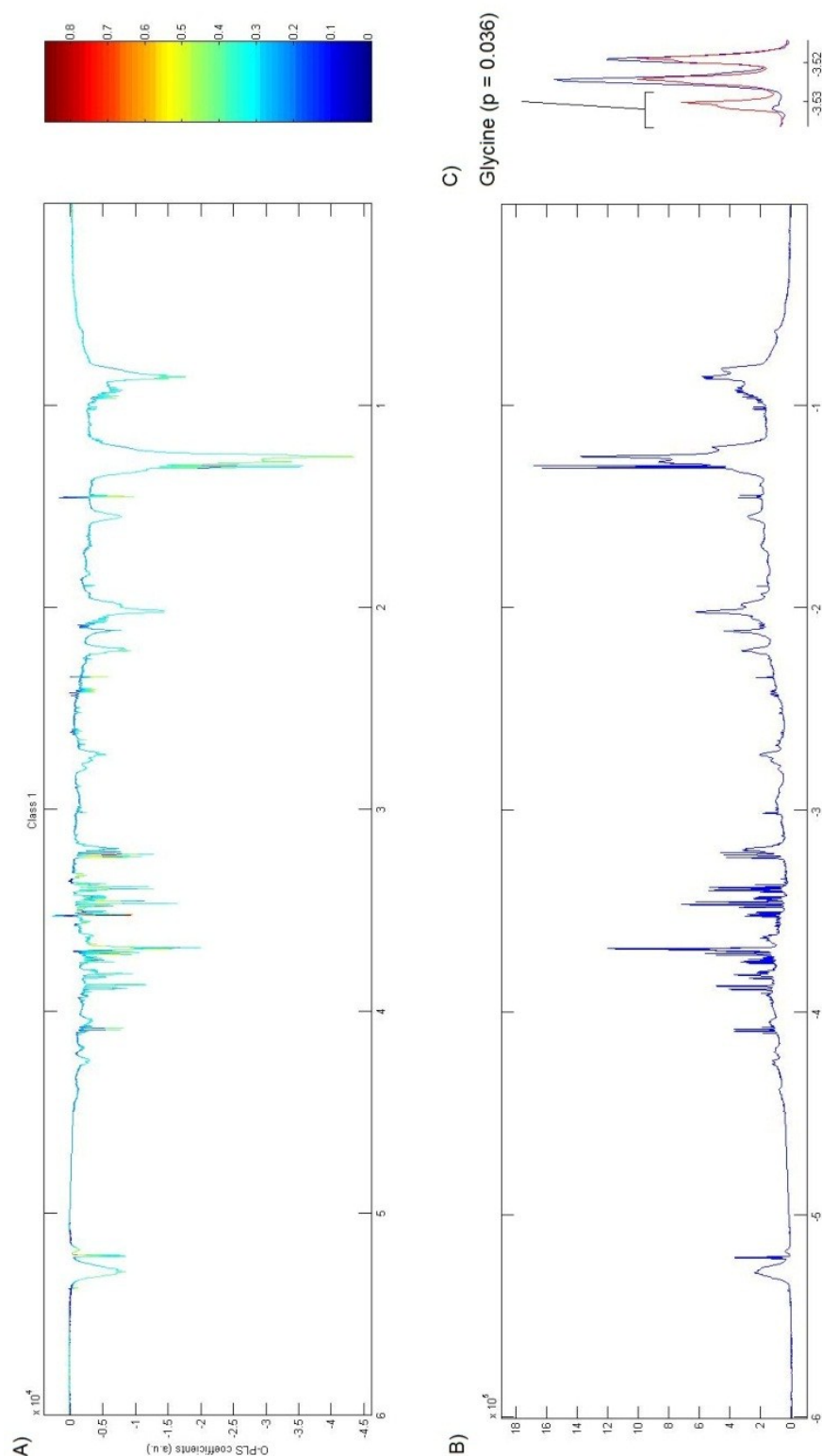


Figure 5.3. The loadings plot (a) from the plasma OPLS-DA analysis displaying a pseudospectrum overlaid with a heat map representation of the correlation scores for each variable and the spectrum of a high dose sample (b) with the x axis (ppm) aligned for comparison. There are few areas of high correlation, and the majority of the variables of influencing peaks failed the multiple testing correction. C) Is the enlarged glycine peak (red) overlaid with a control sample of the region (blue). Glycine is the only metabolite which was of statistical significance.

Chemometric analysis of the urine ^1H NMR spectroscopic data showed a strong difference in metabolic profiles between control and treatment groups with these differences more pronounced than those seen for plasma. Construction of PCA models indicated a strong separation of the high dose group (200 mg/kg) from the low dose (100 mg/kg) and control groups within statistical space. On initial observation it may seem as if there are a number of high dose samples mixed with the controls in statistical space however, when coloured by time point to show the pre-dose and recovery times points, as well as the post dose samples, it is possible to visualise the ‘metabolic journey’ through time from a “normal” pre-dose profile to metabolic events at 24 hours post dose and back to equilibrium after the end of the treatment (figure 5.4).

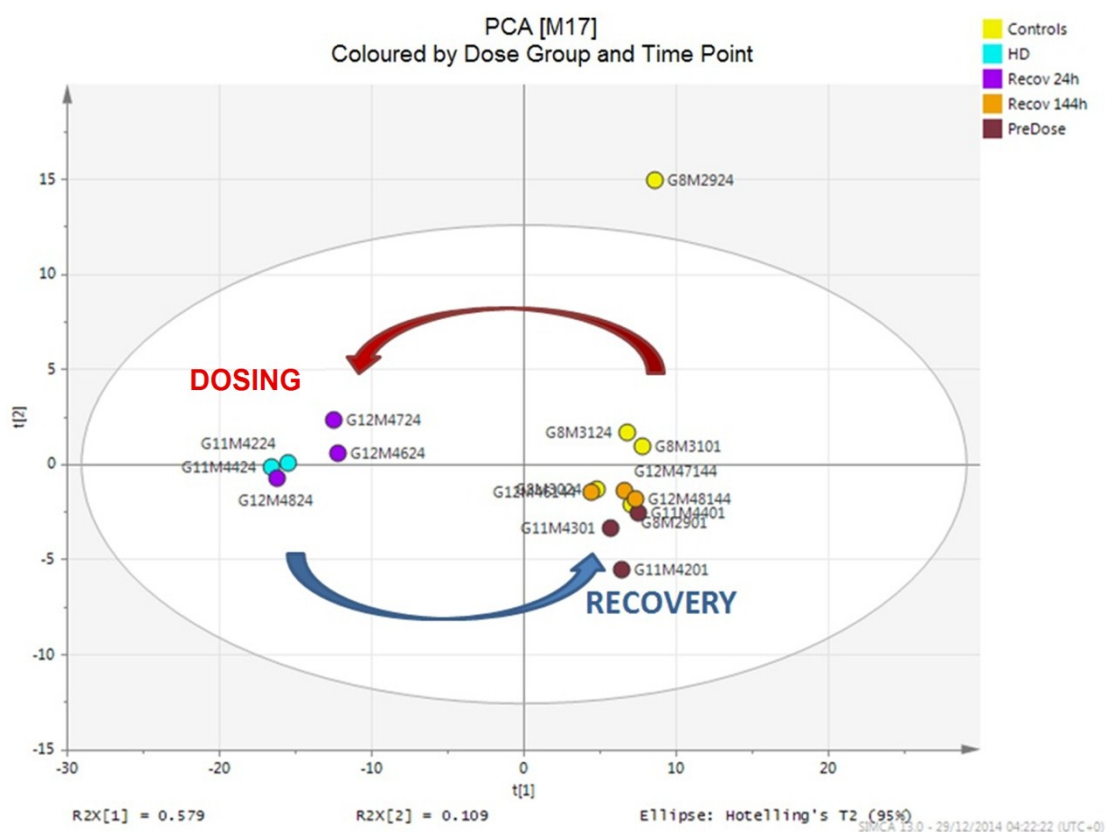


Figure 5.4. PCA scores plot of the urine samples, coloured by both dose group and time point. Using this colouring scheme it is possible to visualise the ‘metabolic journey’ journey of the animals, as indicated by the coloured arrows. Initially all samples are similar to the controls (pre dose). Compound dosing alters the metabolic profiles. After dosing has ended the effects gradually diminish until the metabolic profiles have returned to equilibrium.

OPLS-DA modelling was also performed, and comparisons of the control and high dose group at 24 hours (figure 5.6a) illustrate the significant metabolic differences between the animals. The statistical models produced achieved very high correlation values and predictive power ($R^2 = 0.99$, $Q^2 = 0.98$) and allowed for the identification of the signals driving the separation. Figure 5.6b illustrates the OPLS-DA s-plot, with the enlarged region indicating which spectral signals are most influential. Despite the low sample size the differences in variables were large enough for the identified peaks to pass multiple testing correction and were highly significant.

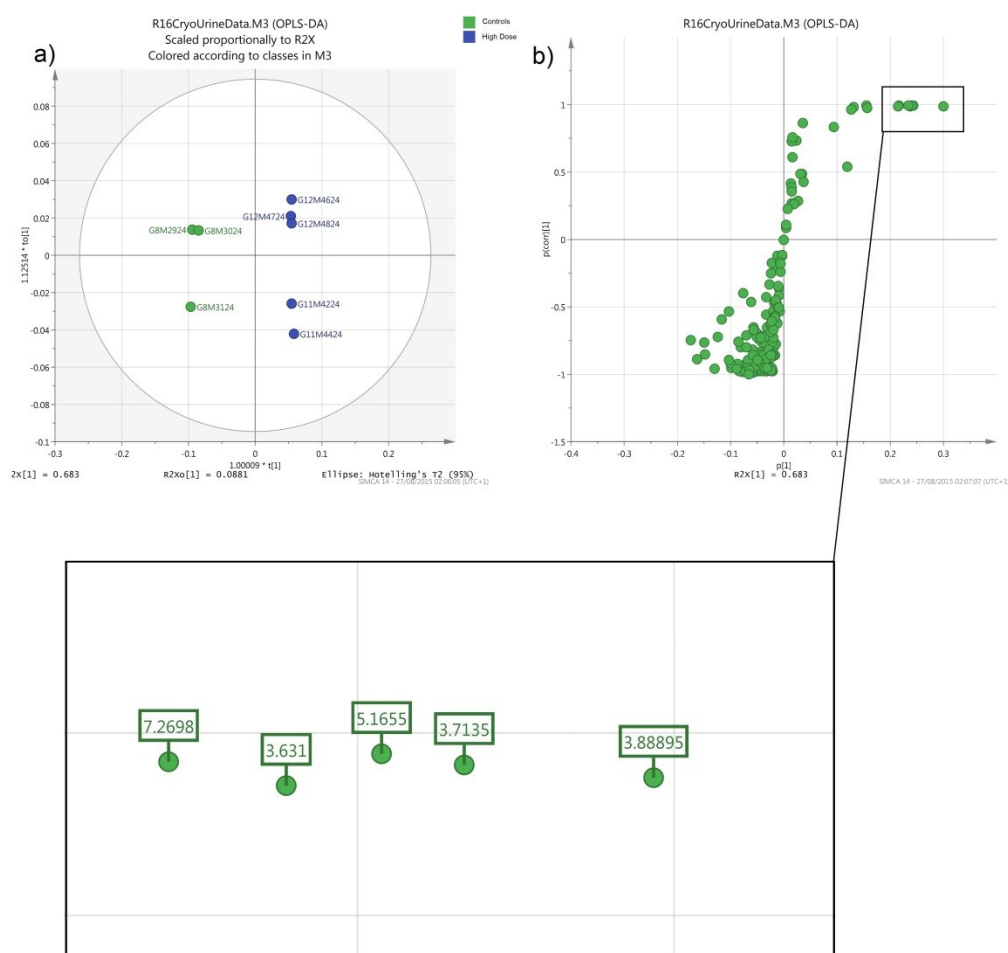


Figure 5.5. A) is the OPLS scores plot of the urine samples comparing the high dose and control animals, coloured by dose group. There is a clear distinction between the treatment groups. B) Is the OPLS-DA s-plot indicating which are the variables contributing most to the separation. The integrated peak regions in the top right of the plot are those which are elevated in the high dose. This area is zoomed in at the bottom and the chemical shift is also indicated.

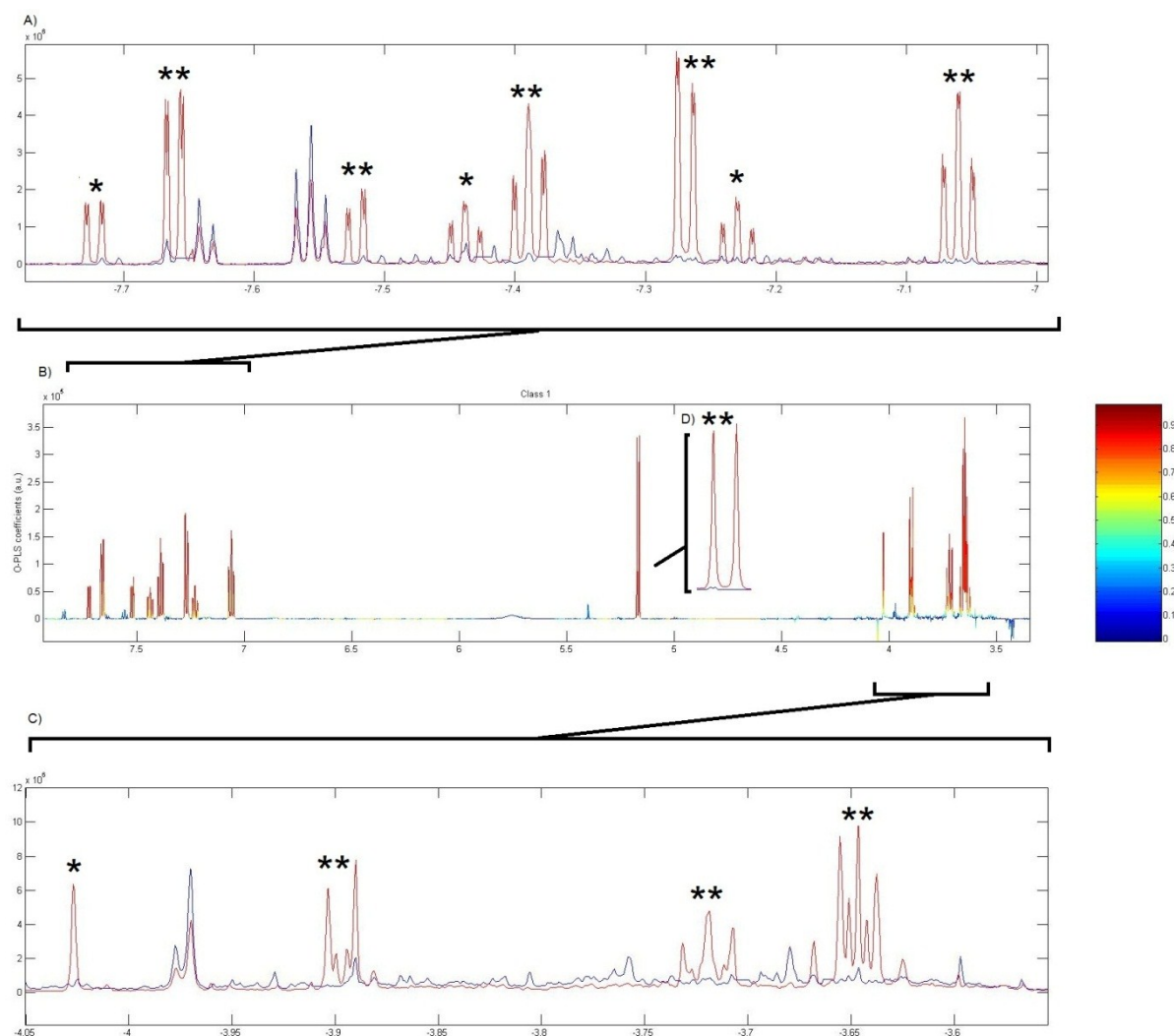


Figure 5.6 The loadings plot (B) from the Urine OPLS-DA analysis displaying a pseudospectrum overlaid with a heat map representation of the correlation scores for each variable. A), C) and D) are zoomed in representations of the real peaks in the regions indicated (red = High dose, blue = control). Asterisks above peaks indicate MANOVA significance levels: ** $p < 10^{-10}$, * $p < 10^{-5}$

After identifying the regions of interest putative assignments could be made using both results from the previous studies but also database information. Key analytes were the glucuronide and sulphate conjugates (δ 5.16/ δ 7.16/ δ 7.27/ δ 7.43/ δ 7.66) of 2-bromophenol and the endogenous metabolite hippuric acid (δ 3.91/ δ 7.52/ δ 7.54). A comparison of the high dose and control ^1H NMR spectra is included below (Figure 5.7) together with a zoomed in region of the putative 2-bromophenol/glucuronide/sulphate assignment (figure 5.8).

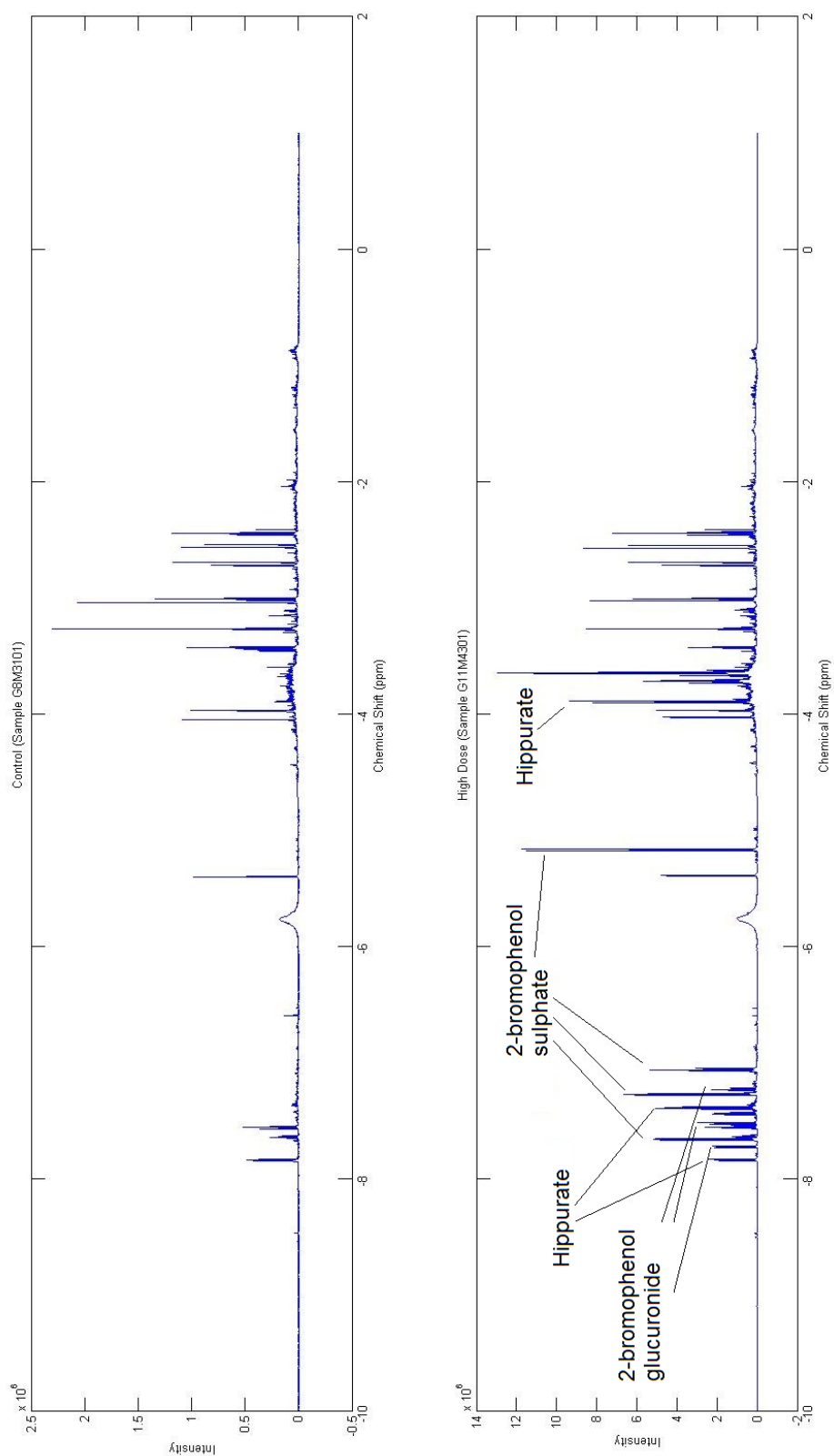


Figure 5.7 A comparison of the typical control (top) and high dose (bottom) ^1H NMR spectra, with putative assignments included on the high dose spectrum obtained at 1 hr post initial dose.

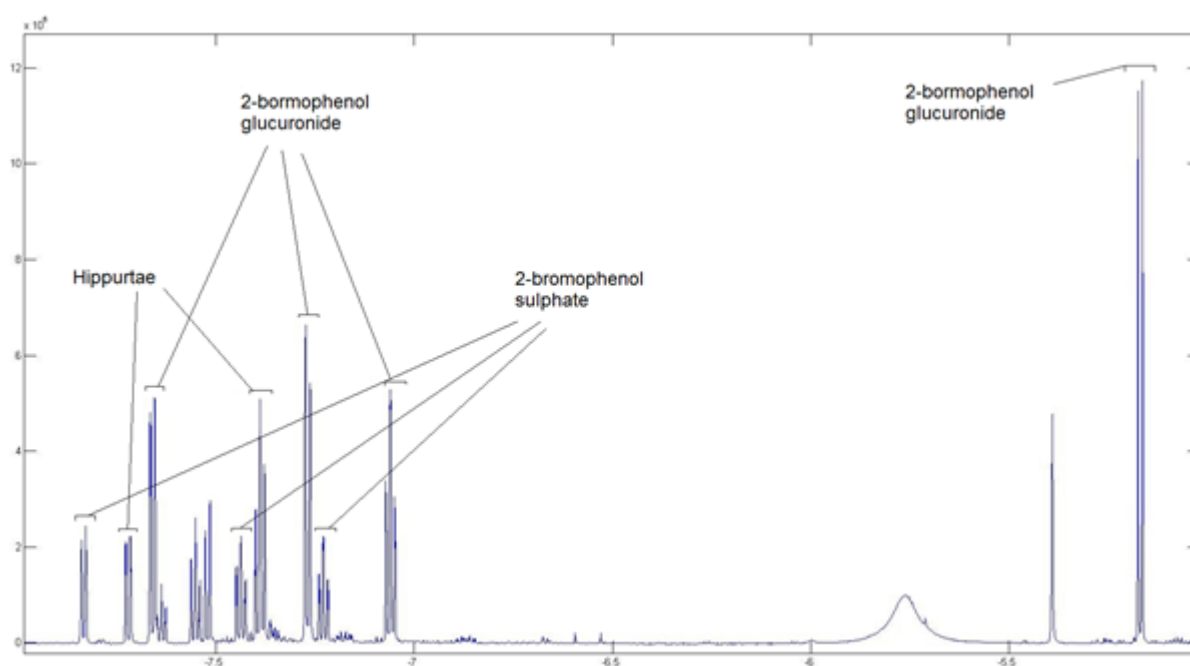


Figure 5.8 A zoomed in representation of the aromatic region of a typical high dose sample, at 1 hour post first dose. Putative assignment of the signals from 2-bromophenol sulphate and 2-bromophenol glucuronide.

These results, which show very high concentrations of 2-bromophenol-related compounds in the urine of these Han Wistar rats at should be contrasted with the low (undetectable) amounts seen in the COMET study samples obtained after dosing Sprague Dawley animals. There is an analytical difference, as these data were obtained using a 700 MHz NMR spectrometer with a cryoprobe as to an instrument operating at 600 MHz with a conventional probe. However the reason for the differences observed is unclear considering dose, dose formulation and dose route was the same, but may reflect strain differences in the handling of the compound. Irrespective of the reason for these differences it is evident that kidney exposure to the potential toxicity of 2-bromophenol, as measured by urinary excretion, is much higher here than in the COMET study (see also the UPLC-MS urinary data discussed below).

5.3.2 UPLC-MS Quality Control

Standard QC based quality control of the analytical stability of the experiment was performed on the dataset. The reasons for the need of such quality assurance have been extensively discussed in previous chapters. Initial visual inspection of the QC chromatograms showed good consistency throughout the run. Figure 5.9 illustrates how the chromatograms of the QCs throughout the run overlay, with no evidence of retention time drift. These trends were mirrored in the initial PCA scores plots used to scrutinise analytical drift. Figure 5.10 shows the initial un-filtered data for the negative ESI MS run; the red points in the scores plot represent the QC samples. It can be seen that all the QCs cluster together almost perfectly in the centre of statistical space. It was concluded that the data acquired was of very good quality and suitable for further statistical analysis.

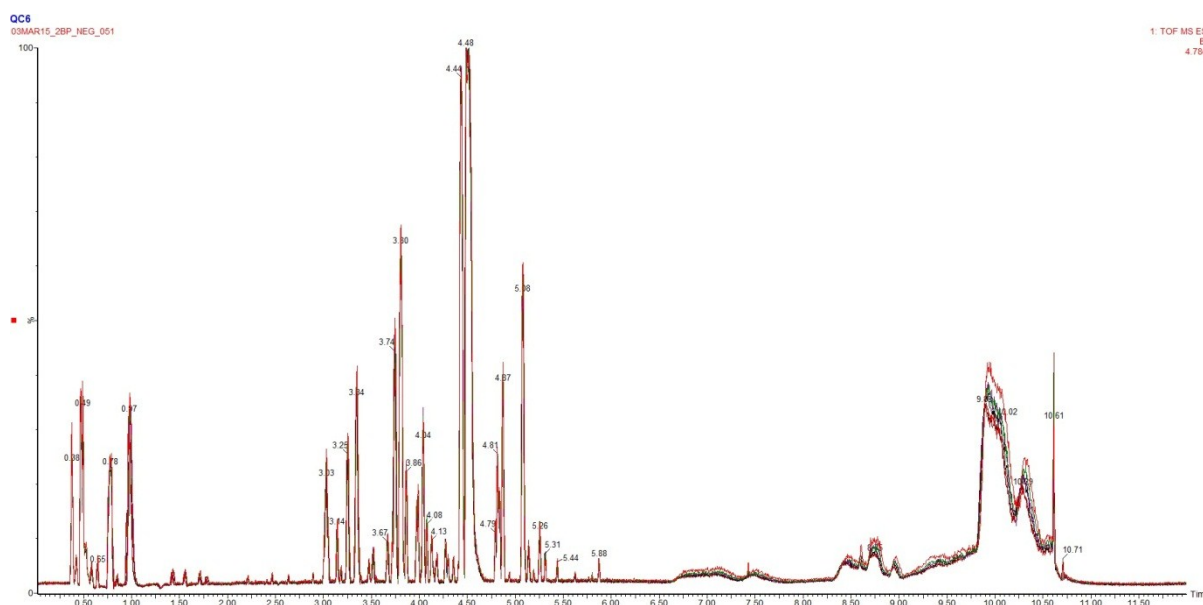


Figure 5.9 chromatographic peaks of the QCs throughout the run. The chromatograms are highly consistent and overlay well, with no visible retention time drift, suggesting that the analysis achieved good analytical stability.

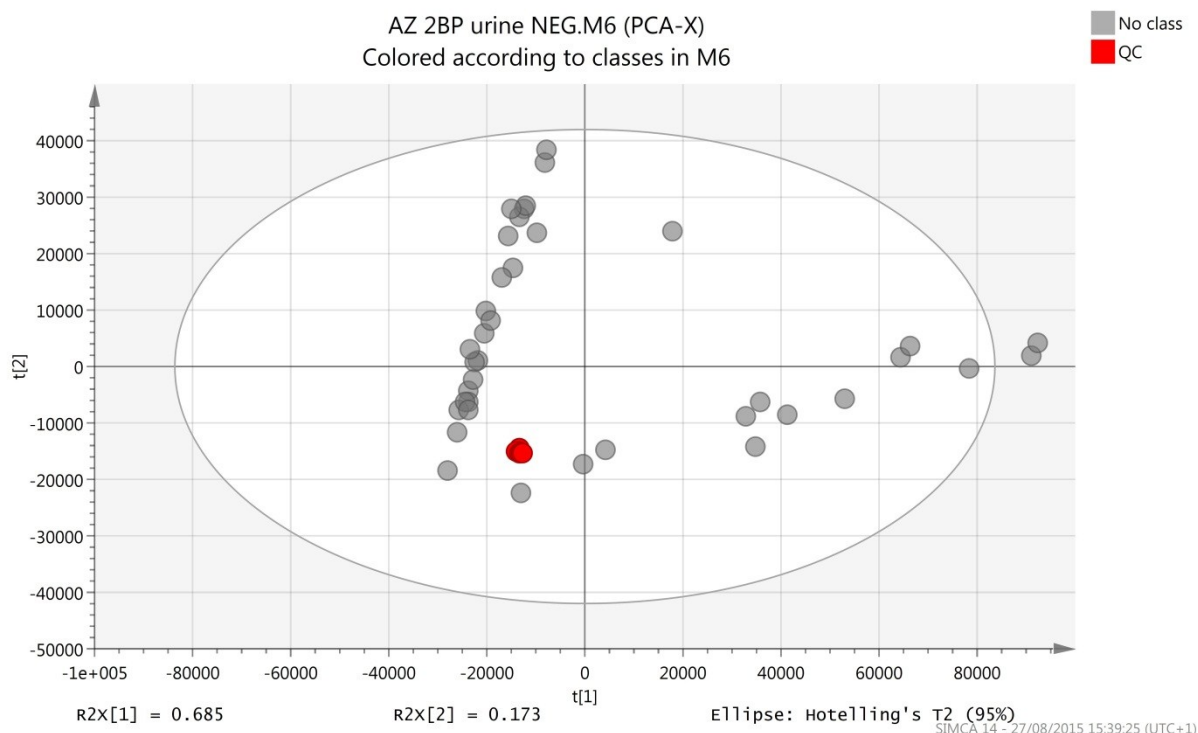


Figure 5.10 PCA scores plot showing the positioning of the QC data within statistical space, in the negative ionisation mode. The red dots are the QCs while the grey dots are the samples. The QCs cluster well towards the centre of the plot.

To ensure that the QCs were consistent throughout, the coefficient of variation for each of the features detected in the QCs was calculated. Ion intensity stability is typically poor for low intensity ions and greater the more intense the ion (providing it is not saturating the detector) and those features that are subject to ion suppression/enhancement will show greater variation than those that are resolved from ion suppressing matrix components. The CV was calculated in the same manner as previous chapters, and the results are in the table below (Table 5.1). The high percentage of ions passing even the stringent 15% CV threshold gave a high level of confidence in the analytical stability of the run.

	CV less than		
	<15%	<20%	<30%
Negative Mode	71.7%	80.8%	90.2%

Table 5.1 CV percentage scores for the QC data.

5.3.3 UPLC-MS Chemometric analysis

After computational analysis with Transomics software, the extracted metabolic signatures, defined by retention time and m/z ratio, were investigated using multivariate statistics in a similar way to the previous chapter. Initial observation of the chromatograms showed that there were peak differences between the high dose and low dose groups. Of special note was a huge peak, of an intensity suggesting probable detector saturation, eluting after 4.5 minutes which seemed to dominate the peak profile as shown in figure 5.11.

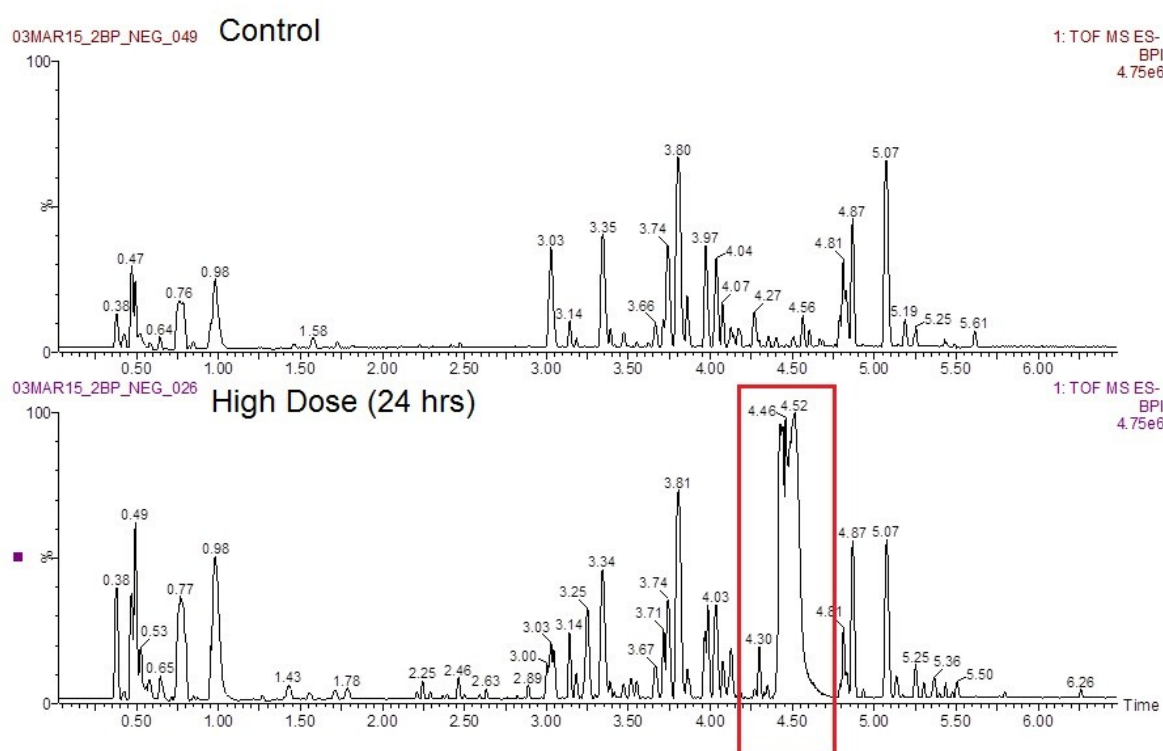


Figure 5.11 A comparison between the chromatograms of a typical control (top) and a high dose urine (obtained from animal 10M39) at 24 hours post administration of 2-bromophenol. The large peak highlighted in red is the major difference between the profiles.

This large peak represents bromophenol metabolites (see below) and should also be compared to the results obtained on the samples from the COMET study undertaken in Sprague Dawley rats. These results confirm, if confirmation were necessary given the ^1H

NMR spectroscopic data on these samples, that much higher kidney exposure to 2-bromophenol-related material was attained in this study in the Han Wistar rat.

Initial PCA modelling produced scores plots which showed a convoluted distribution of the samples across metabolic space. There was no immediately discernible trend, and there seemed to be complicated responses from both the high dose and low dose animals. However, when time points were taken into account it became apparent that maximum response to bromophenol administration was at 24 hours post dose, and when the profiles of the high dose group at this time were compared to the controls there was a significant discrimination as seen in figure 5.12.

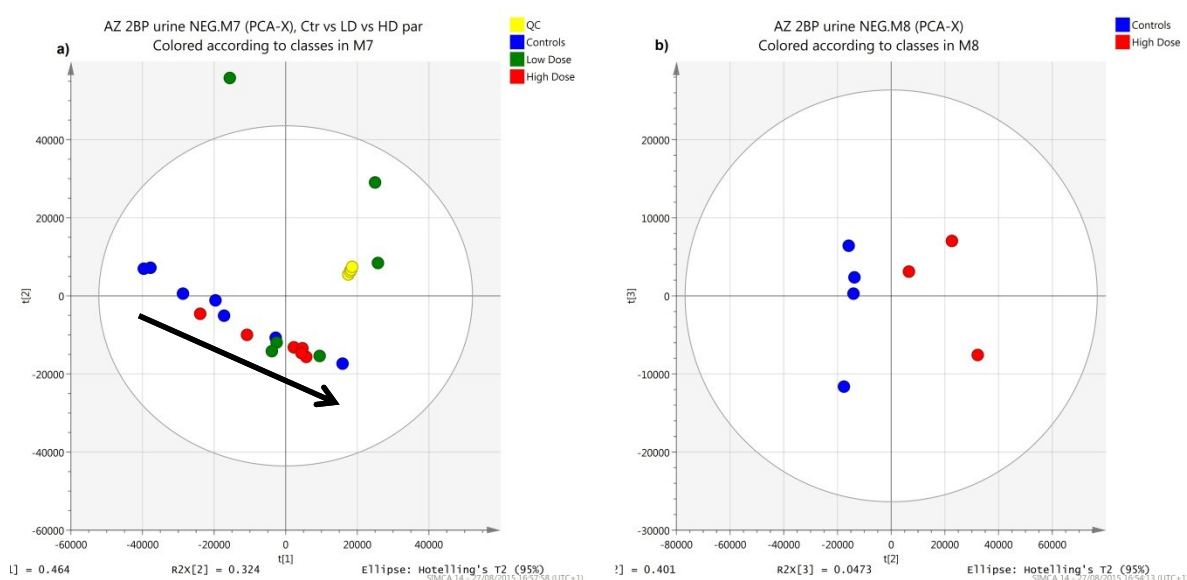


Figure 5.12 PCA scores plots. A) Is the model which includes all the data, coloured by dose group. There is no obvious separation although there is a time related trend as shown by the black arrow. B) is the comparison of just the high dose group and controls at 24 hours post dose. There is now a clear separation between the treatment groups, and this proves to be the timepoint of the maximum statistical difference between the groups.

Subsequent OPLS modelling of the 24 hour time point provided very strong models ($R^2 = 0.95$, $Q^2 = 0.86$), which were used to identify the signals driving the separation. Figure 5.13 shows the OPLS-DA scores plot and the S-plot used to identify the critical ions. The ions highlighted are m/z 252.9, m/z 175.02 and m/z 348.97 which match the mass to charge

ratios of 2-bromophenol sulphate, 2-bromophenol (parent) and 2-bromophenol glucuronide respectively, and also match previous findings (see chapter 4.4.3). The retention times of the signals also showed that they were all found to be the peaks eluting in the large distinctive peak highlighted previously. Given the likely large differences between the physico chemical properties of the 2-bromophenol the presence of the 175.02 ion probably resulted from in source decomposition of the sulphate and glucuronide metabolites (see below) rather than the presence of any of the free aglycone itself. This pattern of metabolites was similar to that seen for the Sprague Dawley rats used for the COMET study and described in Chapter 4. To ensure that these assignments were consistent with the previous, evaluation of the mass spectra of the ions identified the distinctive bromine 1:1 isotope pattern and appropriate losses of 80 and 176 Da, diagnostic for the sulphate and glucuronic acid moieties respectively, due to fragmentation in the ion source, and are highlighted in figure 5.14 together with the chromatographic peak.

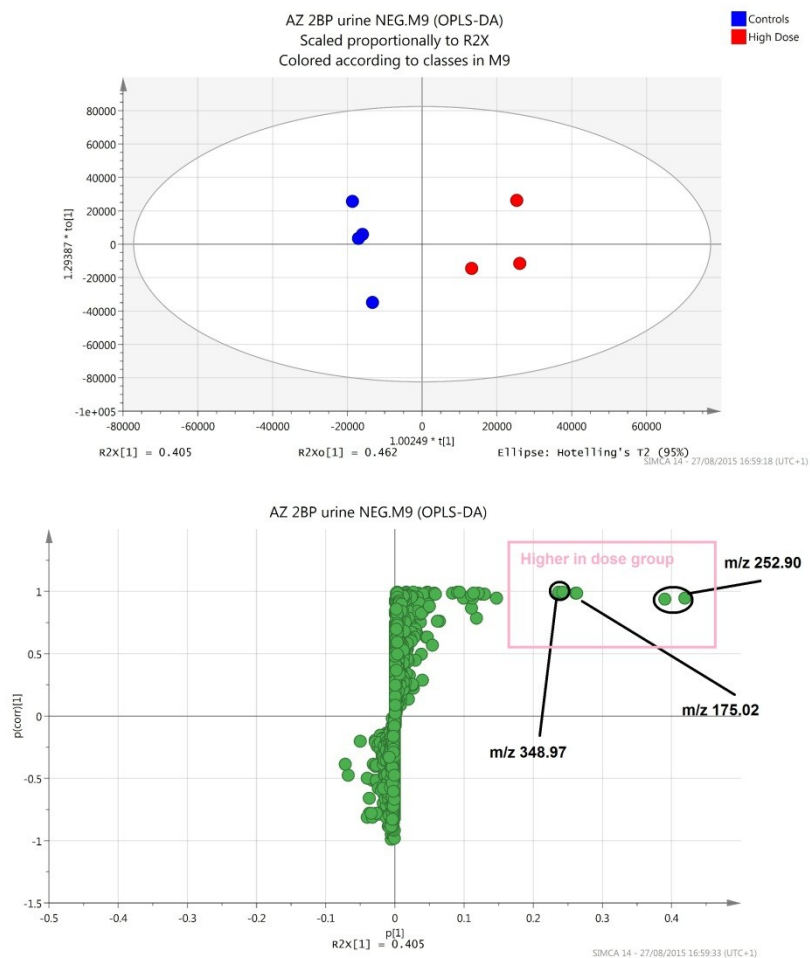


Figure 5.13. a) scores plot of the OPLS-DA model comparing the high dose and control group animals at 24 hours post first dose, showing clear separation between the classes. Inspection of the s-plot (b) derived from the OPLS model comparing low dose and controls, the ions influencing discrimination are m/z 252.9, m/z 175.02 and m/z 348.97.

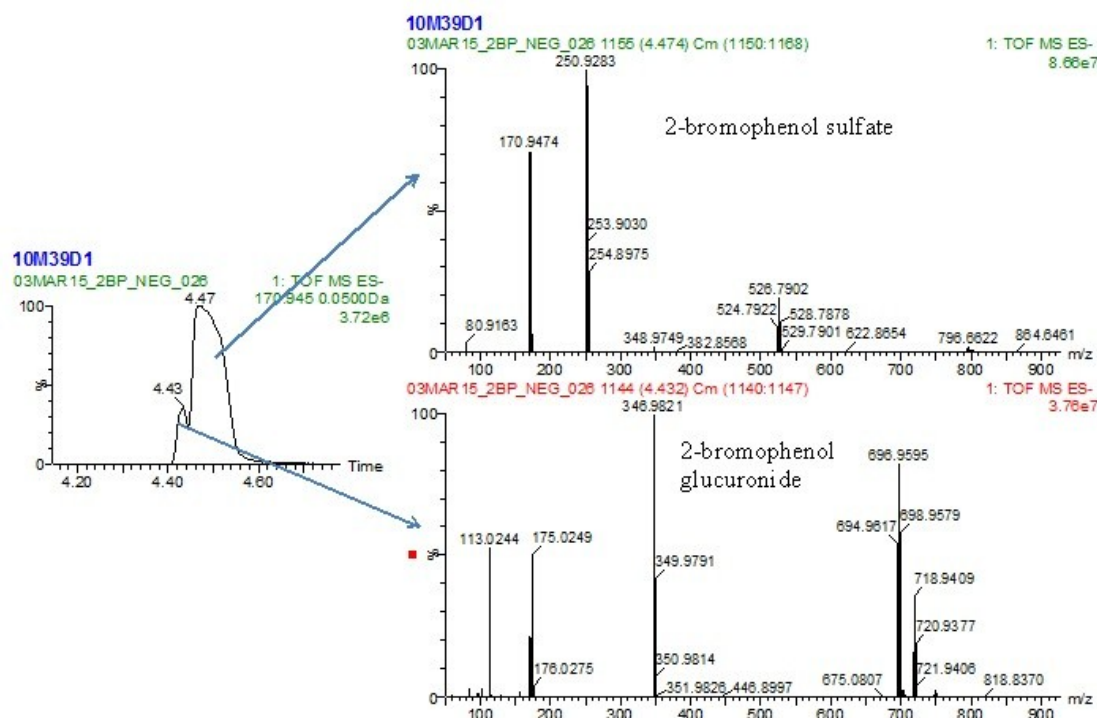


Figure 5.14 The mass spectra of 2-bromophenol sulphate (top right) and 2-bromophenol glucuronide (lower right) together with their respective chromatographic peaks (left). The mass spectra revealed the appropriate losses of 80 and 176 Da, for the sulphate and glucuronide moieties respectively, supporting these assignments. The higher mass ions seem to be the same source related dimer as described in Chapter 4.

5.3.4 Model validation and significance testing

To validate our model the same permutation testing employed previously was again used. This is of critical importance for mass spectral data as the platform detects many more features and therefore the data is highly multi-collinear. Once again 1000 iterations recalculated every 100 permutations were used for each of the models and comparisons performed and presented below in figure 5.15. The permutation plot for the comparison of the high dose and controls at 24 hours post dose is presented. Also included are the CV-anova statistics in the table below for completeness (table 5.2).

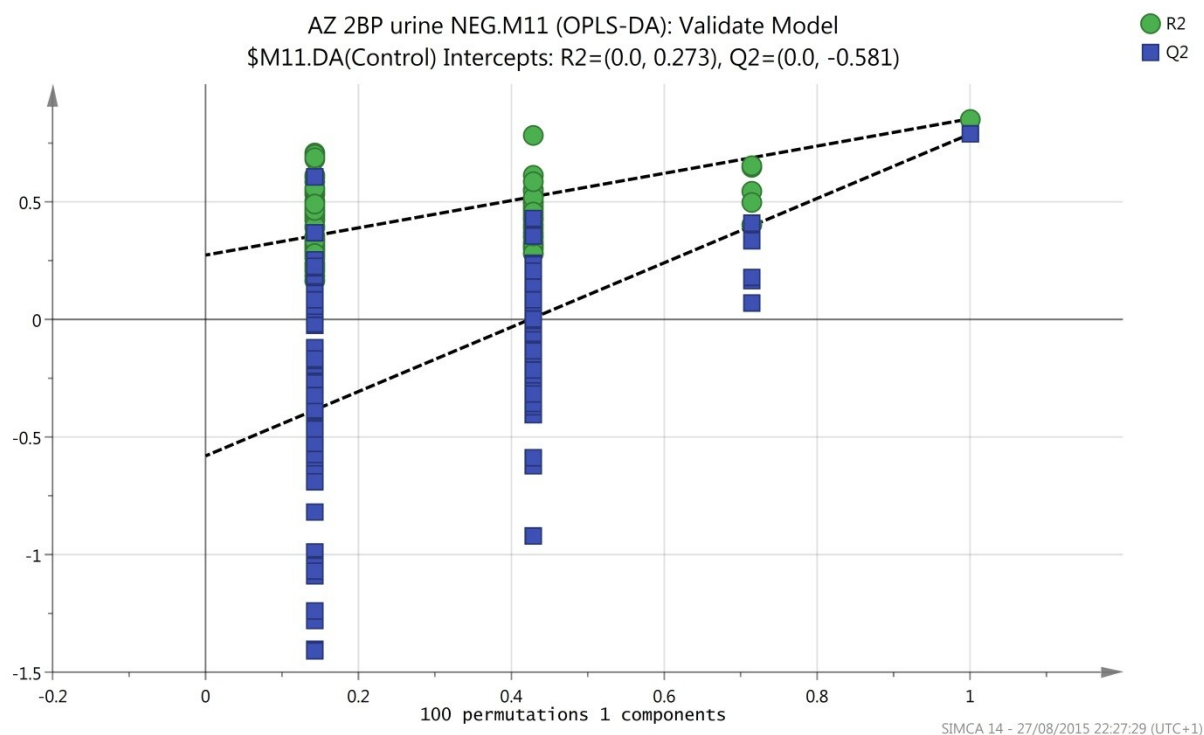


Figure 5.15 Permutation plot for the model constructed. The green and blue dots represent the R2 and Q2 of the dummy data respectively, and the dashed line is the median value trend for the dummy values. The permutation values are lower than the actual values suggesting that the model is highly robust.

OPLS-DA	Sum of Squares	Degrees of Freedom	Mean of Squares	F value	P Value	Standard Deviation
Total corr.	13	13	1			1
Regression	10.2453	4	2.56132	8.3681	0.0042196	1.60041
Residual	2.75473	9	0.306081	1		0.553246

Table 5.2 CV ANOVA scores for the significance test.

5.3.5 Extraction of “isotopobromatograms” from the LC-MS data

As discussed in the previous chapter, compounds containing bromine have a very distinctive isotope pattern in their mass spectra, with the ^{79}Br and ^{81}Br isotopes present at 2 Da apart and at a ratio of approximately 1:1. This isotopic signature enables the chromatographic data to be searched for this characteristic mass difference and ratio. This selective tracing highlighted 6 bromine-containing peaks, for which the spectra are detailed in figure 5.16.

These are the same bromine containing-compounds which were identified in the previous study, and their (repeated) assignments are summarised in table 5.3. As the ions, and their fragments, matched those examined in the previous chapter further MS/MS was not performed.

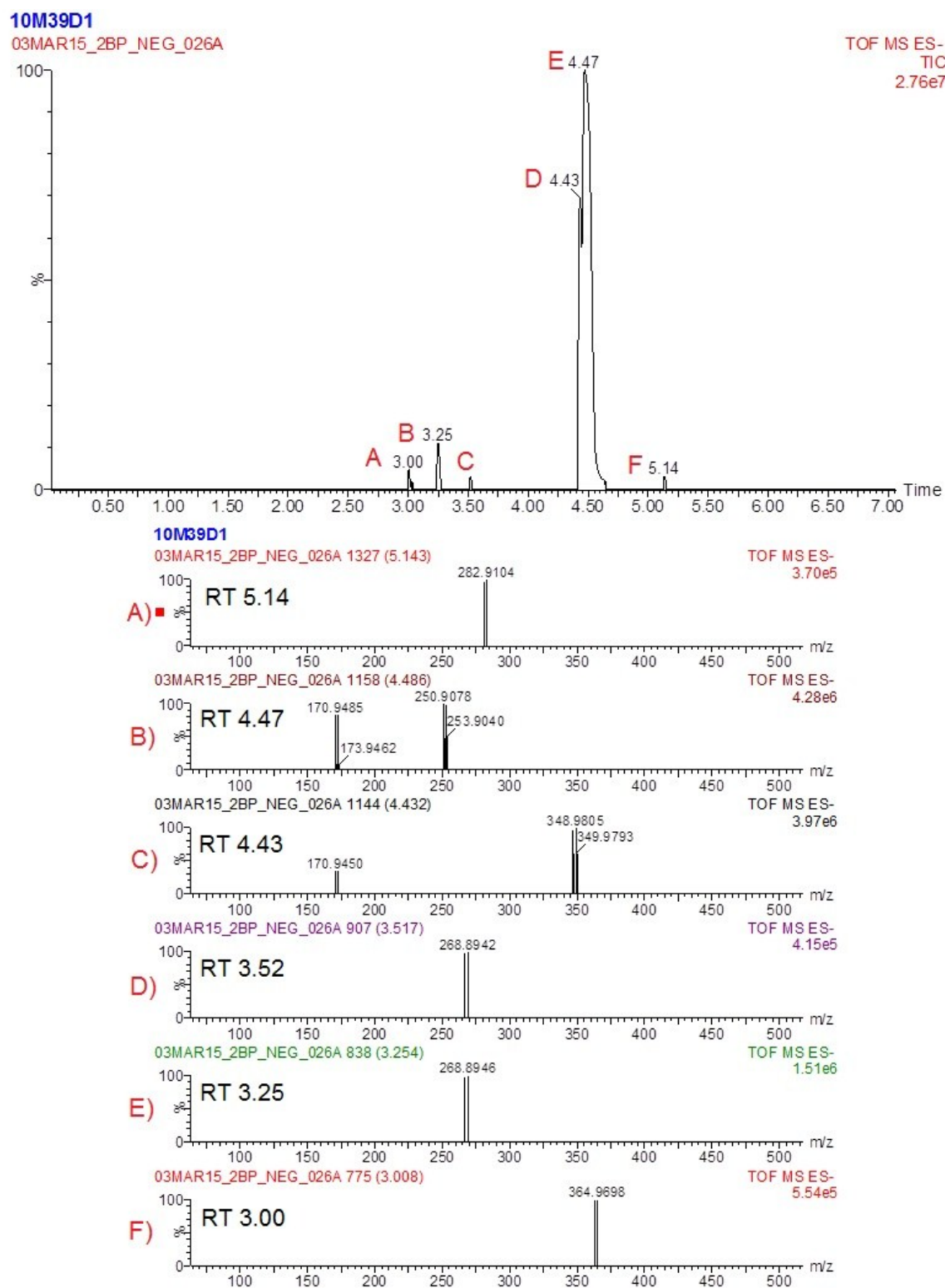


Figure 5.16. Isotopobromatogram (top) and the related mass spectra (bottom). The chromatographic peaks are labelled and match up to the mass spectra. All the mass spectra show the distinctive 1:1 bromine isotope pattern.

Metabolite	Measured Mass	Retention Time	Formula	IBG
2- hydroxy bromophenol glucuronide	364.96	3.00	C ₁₂ H ₁₅ O ₉	F
2- hydroxy bromophenol sulphate	268.89	3.25/3.52	C ₆ H ₇ O ₅ BrS	D/E
2-bromophenol sulphate	253.90	4.47	C ₆ H ₆ O ₄ BrS	B
2- bromophenol Glucuronide	349.97	4.43	C ₁₂ H ₁₄ O ₈	C
2- hydroxy (3-methoxy) bromophenol sulphate	282.91	5.14	C ₇ H ₈ O ₅ BrS	A

Table 5.3 Assignments of the analytes identified by the bromine isotope pattern recognition.

5.3.4 Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

As discussed in the introduction, ICP-MS is a powerful elemental detection that allows for the direct quantification of the target element. ICP-MS allows studies on the metabolic fate of compounds, without the need for the copious and expensive synthesis of radiolabelled standards which are otherwise necessary for this type of experiment. By using 79/81 Br-ICPMS for quantification in this study, it was possible to determine concentrations of total bromine in the urine samples. Figure 5.17b below presents the progression of bromine concentrations in a typical high dose animal, with a clear decrease in 2-bromophenol-related material present in the urine following the cessation of dosing. A comparison between the control and high dose (and recovery) groups show that the baseline bromine urinary concentrations are 0.55 ± 0.3 mg/mL whereas at 24 hours, the time of maximum response, bromine concentrations reached 3.01 ± 0.8 mg/mL in the high dose group. In the recovery group, 4 days after dosing had stopped the concentrations of bromine were only slightly reduced, to 2.57 ± 0.8 mg/mL, suggesting that perhaps 2-bromophenol and its related metabolites are excreted relatively slowly. Figure 5.17a illustrates the average urinary bromine concentrations for the three groups.

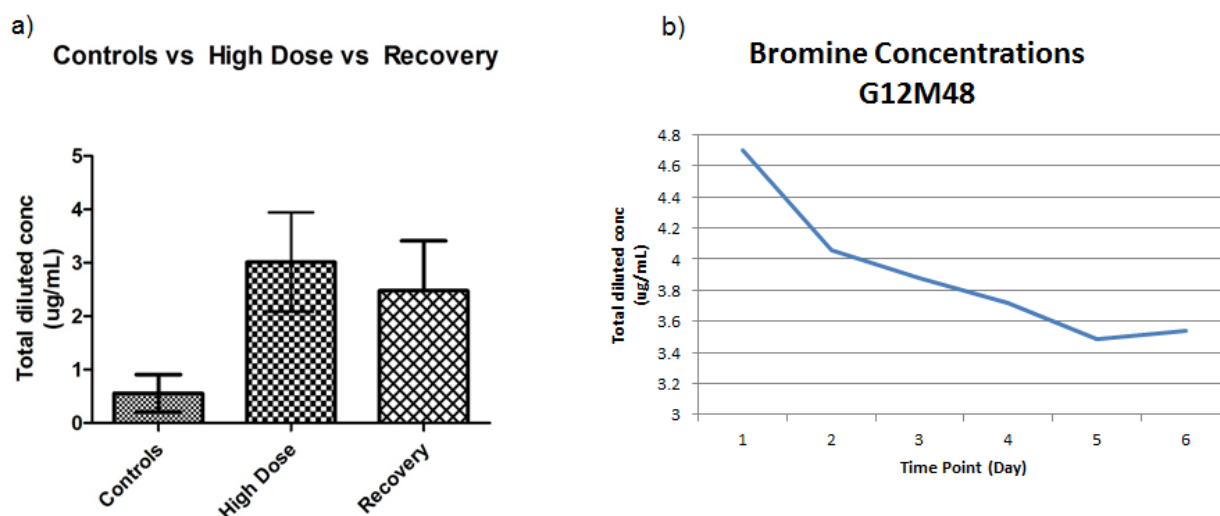


Figure 5.17. Results from the bromine ICP-MS. A) is the adjusted averages for the total bromine concentrations for each treatment group. There is a baseline bromine level indicated by the controls, however, administration of 2-bromophenol greatly increases urinary bromine levels as expected. There is however little difference between the concentrations at 24 hours (high dose) and recovery (144 hours, 3 days after dosing has stopped), suggesting that the compound and its major metabolites are excreted slowly. B) The typical bromine excretion curve for a high dose animal over the 6 day study.

5.3.5 Histopathology and Clinical Chemistry

As for the single dose study described in Chapter 4 “traditional” toxicology assessments, including histopathology on the kidneys removed from the animals at termination and clinical chemistry tests were undertaken on both the urine and blood samples. These data were provided by Kim Maratea, from DSM Boston, Astrazeneca.

a. Histopathology

Lau *et al.* (1984) reported that the administration of 2-bromophenol resulted in nephrotoxicity characterised by necrotic damage to the proximal tubule cells. Information on whether toxicity was achieved by the current study was provided by the histopathology report (see appendix). Again it was hoped that the high dose would cause changes in the kidney but no mortality. However, neither the low nor the high dose groups showed any clinical signs, and

the rats appeared comfortable and in good health throughout the experiment. One low dose rodent exhibited signs of very minor tubular degeneration, and one high dose recovery animal showed mild mononuclear cell damage coupled with minor tubular degeneration. However, this is a low number and the rest of the animals in the group showed no kidney damage whatsoever. An overview of the total histopathology findings are detailed below (table 5.4).

Compound Dose (mg/kg/day)	Incidence (Median score)		
	Vehicle (corn oil) 0	2-bromophenol 100	2-bromophenol 200
Day 3 (Main)			
Cyst, cortex	0/4	0/4	0/4
Degeneration, tubular	0/4	1/4 (1)	0/4
Fibrosis, interstitial	0/4	0/4	0/4
Infiltrate, mononuclear cell	0/4	0/4	0/4
Mineralization, tubular	0/4	0/4	0/4
Day 7 (Recovery)		NA	
Degeneration, tubular	0/4	-	1/4 (1)
Fibrosis, interstitial	0/4	-	0
Infiltrate, mononuclear cell	0/4	-	1/4 (1)
Day 24 (Recovery)		NA	
Degeneration, tubular	-	-	-
Fibrosis, interstitial	-	-	-
Infiltrate, mononuclear cell	-	-	-

Incidence = number affected/number examined

N/A = not applicable

Table 5.4 Histopathology observations

b. Clinical chemistry

Haematology clinical chemistry data (provided in the appendix) showed no changes in the haematology parameters after 2-bromophenol administration and all animals were within the normal ranges for each test. Urinary clinical chemistry, shown in table 5.5, showed a slight overall elevation in urinary protein concentrations, although creatinine, glucose and urea levels remained unaffected.

Treatment	Dose mg/kg/day	UVOL L	UCRE mmol/L	UTP ratio [§]	UGLU ratio [§]	UUREA ratio [§]
DAY 1						
Vehicle (corn oil)	0	0.0017 ± 0.0014	10.33 ± 3.93	105.08 ± 16.40	0.080 ± 0.011	66.33 ± 10.96
2-Bromophenol	100	0.0026 ± 0.0008	7.67 ± 4.70	133.04 ± 31.95	0.082 ± 0.007	64.96 ± 6.79
2-Bromophenol	200	0.0029 ± 0.0013	8.13 ± 3.46	217.25 ± 62.27	0.082 ± 0.009	67.15 ± 8.29
DAY 2						
Vehicle (corn oil)	0	0.0016 ± 0.0008	10.91 ± 5.48	89.46 ± 20.32	0.083 ± 0.018	64.90 ± 10.72
2-Bromophenol	100	0.0014 ± 0.0005	8.35 ± 5.35	163.84 ± 26.91	0.0811 ± 0.011	62.70 ± 4.86
2-Bromophenol	200	0.0025 ± 0.0020	10.27 ± 5.24	164.19 ± 41.23	0.085 ± 0.006	70.49 ± 6.82
DAY 6						
Vehicle (corn oil)	0	0.0024 ± 0.0007	N/A	87.84 ± 36.90	N/A	N/A
2-Bromophenol	200	0.0038 ± 0.0013	N/A	69.82 ± 21.52	N/A	N/A

§Normalized to urine creatinine (UCRE)

Table 5.5 Urine clinical chemistry parameters in the rats on Days 1, 2 and 6. Data are expressed as Mean ± Standard Deviation.

Testing for the detection of any salient kidney injury biomarkers was also performed, to ensure that any kidney injury which was not detectable by histopathological inspection was not missed. Key biomarkers investigated were α -glutathione-S-transferase (α GST), indicative of glutathione depletion, rat glutathione-S-transferase Yb1 (GSTYb1) a glutathione transferase which is localised to the renal tubule cells, renal papillary antigen (Rpa-1) specific for collecting duct cell damage and also a biomarker of 2-bromoethanamine nephrotoxicity (Price *et al.*, 2010), clusterin, indicative of toxicant induced acute kidney injury in the rat (Vaidya *et al.*, 2009), neutrophil gelatinase-associated lipocalin (NGAL, also known as lipocalin-2) an early indicator of nephrotoxicity (Devarajan, 2008), osteopontin, shown to be indicative of glomerular damage and inflammation (Panzer *et al.*, 2000) and kidney injury molecule-1 (Kim-1) one of the major diagnostic markers for renal proximal tubule injury (Han *et al.*, 2002). While the biomarker testing was extensive there were no overall trends within the data, suggesting that there had been no kidney damage. The results of the analysis are presented in table 5.6 below.

Treatment	Dose mg/kg/day	aGST	GSTYb1	RPA	Clusterin	NGAL	Osteopontin	Kim-1
DAY 1								
Vehicle (corn oil)	0	255.33 ± 122.86	7.34 ± 7.80	2.32 ± 1.64	5.77 ± 1.75	323.99 ± 62.63	0.93 ± 0.49	0.85 ± 0.11
2-Bromophenol	100	230.11 ± 179.16	5.59 ± 4.31	1.99 ± 1.44	3.08 ± 1.01	394.14 ± 110.40	0.91 ± 0.71	0.60 ± 0.08
2-Bromophenol	200	434.13 ± 266.27	17.33 ± 7.00	5.51 ± 5.28	6.10 ± 2.73	390.15 ± 91.72	3.12 ± 1.42	0.94 ± 0.12
DAY 2								
Vehicle (corn oil)	0	265.18 ± 99.13	7.10 ± 5.88	2.84 ± 2.33	6.46 ± 7.21	345.18 ± 72.17	0.93 ± 0.98	0.87 ± 0.12
2-Bromophenol	100	363.91 ± 265.18	2.01 ± 2.54	3.75 ± 2.46	2.31 ± 1.97	62.72 ± 2599.78	1.16 ± 0.97	0.66 ± 0.97
2-Bromophenol	200	261.87 ± 173.88	10.14 ± 6.89	5.63 ± 8.27	3.26 ± 1.61	473.73 ± 177.95	4.16 ± 3.03	0.88 ± 0.07
DAY 6								
Vehicle (corn oil)	0	135.04 ± 62.71	2.14 ± 3.22	0.88 ± 0.21	4.84 ± 0.96	N/A	N/A	N/A
2-Bromophenol	200	150.90 ± 38.40	8.78 ± 12.24	1.17 ± 0.64	9.22 ± 3.54	N/A	N/A	N/A

Table 5.6 Summary of kidney injury biomarkers in the rats on Days 1, 2 and 6. Data are expressed as Mean ± Standard Deviation.

5.4 Summary and Conclusions

Previous work had identified 2-bromophenol, a phenolic metabolite of bromobenzene, as a potential nephrotoxin (Lau et al, 1984b). The previous chapter described a 2-bromophenol toxicity study in Sprague Dawley rats, as part of the COMET program. That study assessed the nephrotoxicity of the compound by combining ^1H NMR spectroscopic and UPLC-MS metabolic profiling. While there was evidence of extensive metabolism of 2-bromophenol, with key 2-bromophenol metabolites and changes in endogenous metabolism identified, the overall conclusion was that there was no nephrotoxicity at the doses administered.

This chapter describes the continuation of this research avenue with the investigation of the effects of chronic 2-bromophenol exposure, this time in Han Wistar rats. After combined ^1H NMR spectroscopic and UPLC-MS metabolic profiling, statistical analysis showed that there was a significant metabolic perturbation due to administration of the compound. However, once again, even with a longer dosing schedule, the histopathology and clinical chemistry analysis still showed no evidence of kidney damage. Use of cryoprobe technology with the ^1H NMR spectroscopy, combined with the urinary excretion of significantly higher concentrations of 2-bromophenol-related material, did allow for the identification of the major

2-bromophenol metabolites, the glucuronide and sulphate conjugates, which were also the key discriminating signals together with hippuric acid. The analysis of the urine-derived LC-MS data revealed that the major discriminating ions were not the result of changes in endogenous compounds but to the detection of the sulphate and glucuronide conjugates of 2-bromophenol. Further detailed analysis of the urine via bromine-isotope specific detection revealed the presence of a number of minor, hydroxylated conjugates of 2-bromophenol, replicating the previous results. There was once again no evidence of glutathione conjugates, generally agreeing with the theme of no toxic value.

The overall conclusion from both this and the previous study, with respect to nephrotoxicity, was that after administration of 100 or 200 mg/kg, either as a single dose or repeated, 2-bromophenol is not a kidney toxin. There is overwhelming evidence supporting this conclusion as three separate studies (including the original COMET), in different animals in different animal laboratories achieved the same results. It is therefore apparent that all current data is not in agreement with the previous work by Lau *et al.*; further work would be needed to ascertain why this discrepancy is so prominent.

4.7 References

Barnes, D., (1982). Pre-clinical toxicology of ifosfamide. *Seminars in Oncology*, **9**, 8 – 13.

Blaise, B.J., Shintu, L., Elena, B.N.D., Emsley, L., Dumas, M.-E., Toulhoat, P. (2009). Statistical Recoupling Prior to Significance Testing in Nuclear Magnetic Resonance Based Metabonomics. *Analytical Chemistry*, **81**, 6242-6251.

Brezis, M., Silva, P., Epstein, F.H., (1984). Amino acids induce renal vasodilatation in isolated perfused kidney: Coupling to oxidative metabolism. *American Journal of Physiology*, **247**, 999-1004.

Deters, M., Siegers, C.P., Strubelt, O., (1998). Influence of glycine on the damage induced in isolated perfused rat liver by five hepatotoxic agents. *Toxicology*, **128**, 63-72.

Devarajan, P., (2008). Neutrophil gelatinase-associated lipocalin (NGAL): A new marker of kidney disease. *Scandinavian Journal of Clinical and Laboratory Investigation*. (Supplementum), **241**, 89–94.

Duckett, C., McCullagh, M., Smith, C., Wilson, I.D., (2015). The metabolism of 4-bromoaniline in the bile-cannulated rat: application of ICPMS ((79/81)Br), HPLC-ICPMS & HPLC-*oa*TOFMS. *Xenobiotica*, **45**, 672-680.

Goren, M.P., Wright, R.K., Horowitz, M.E., (1986). Cumulative renal tubular damage associated with cisplatin nephrotoxicity. *Cancer Chemotherapy and Pharmacology*, **18**, 69-73.

Han, W.K., Bailly, V., Abichandani, R., Thadhani, R., Bonventre, J.V., (2002). Kidney Injury Molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury. *Kidney International*, **62**, 237-244.

Klonne, D.R., Nachreiner, D.J., Dodd, D.E., Losco, P.E., Tyler, T.R., (1987). Acute and 2-week inhalation toxicity studies on aerosols of selected ethylene oxide/propylene oxide polymers in rats. *Fundamental and Applied Toxicology*, **9**, 773-784.

Loebstein, R., Koren, G., (1998). Ifosfamide-induced nephrotoxicity in children: critical review of predictive risk factors. *Pediatrics*. **101**, E8.

Panzer, U., Thaiss, F., Zahner, G., Barth, P., Reszka, M., Reinking, R.R., Wolf, G., Helmchen, U., Stahl, R.A., (2001). Monocyte chemoattractant protein-1 and osteopontin differentially regulate monocytes recruitment in experimental glomerulonephritis. *Kidney International*, **59**, 1762–1769

Price, S.A., Davies, D., Rowlinson, R., Copley, C.G., Roche, A., Falkenberg, F.W., Riccardi, D., Betton, G.R., (2010). Characterization of renal papillary antigen 1 (RPA-1), a biomarker of renal papillary necrosis. *Toxicological pathology*, **38**, 346-358.

Robert, H., (1986). Carbohydrate Metabolism I: Major Metabolic Pathways and their Control. New Delhi, *John Wiley and Sons*, 432-440.

Stonehouse, J., and Keeler, J., (1995). A Convenient and Accurate Method for the Measurement of the Values of Spin-Spin Coupling-Constants. *Journal of Magnetic Resonance A*, **112**, 43.

Tanaka, I., Igisu, H., Haratake, J., Cho, S., Mori, K., Fujishiro, K., Inoue, N., Horie, A., Akiyama, T., (1987). Cumulative toxicity potential of methomyl aerosol by repeated inhalation. *Journal of the American Industrial Hygiene Association*, **48**, 330-334.

Vaidya, V.S., Ferguson, M.A., & Bonventre, J.V., (2008). Biomarkers of Acute Kidney Injury. *Annual Review of Pharmacology and Toxicology*, **48**, 463–493.

Veselkov, K.A., Lindon, J.C., Ebbels, T.M.D., Crockford, D., Volynkin, V.V., Holmes, E., Davies, D.B., Nicholson, J.K. (2008). Recursive Segment-Wise Peak Alignment of Biological ¹H NMR Spectra for Improved Metabolic Biomarker Recovery. *Analytical Chemistry*, **81**, 56-66.

Want, E.J., Wilson, I.D., Gika, H., Theodoridis, G., Plumb, R.S., Shockcor, J., Holmes, E., Nicholson, J.K. (2010). Global metabolic profiling procedures for urine using UPLC-MS. *Nature Protocols*, **5**, 1005-1018.

Conclusion

Chapter Six

This chapter provides a summary of the work undertaken during the research leading to this Thesis, and also some general conclusions. Based on the outcomes garnered from the collective evaluation of the work presented in this thesis suggestions for future research are presented.

6.1 Summary and Conclusions

The utility of metabonomics as a platform for studying xenobiotic metabolism and toxicity has been extensively discussed both in this thesis but also in the available literature. What separates metabonomics from other types of investigations is the fact that metabolic profiling enables the hypothesis free observation of the effects of exogenous compounds on whole organisms by utilising multiparametric measurements of the response to the exposure over time. In the cellular environment, every level of biomolecular organisation is mutually reliant, and thereby equally susceptible to external events. Downstream phenomena, be it the direct effect of the introduction of the foreign compound, or a metabolite thereof, or the alteration of the abundance of endogenous substances, can all be mapped upon a metabolic landscape. This landscape, constructed by the metabolic profiling of biofluids or tissues, when incorporated into a systems biology framework, allows an understanding of the relationships between the experimental inputs and real world endpoints; which is not always the case with other 'omic' fields.

Integral to the 'metabonomic' process is the characterisation of these metabolic changes through appropriate analytical methods, with NMR spectroscopy and LC-MS predominating in this regard. The acquired data is then analysed using multivariate statistics by various chemometric and bioinformatic tools. However, as time progresses both the analytical platforms and the computational resources available change. Furthermore, recent trends within epidemiological and toxicological studies (amongst others) indicate that there is a distinct focus on large-scale population level studies, requiring high-throughput technologies. This means that for the continuing evolution of metabonomics as a field, and indeed its survival, new analytical methods and computational approaches must be developed to satisfy this increase in demand. The work presented in this thesis has, in addition to its application to xenobiotic metabolism and toxicity, revolved around the expansion of the capacity of current experimental practice, with the onus being on surmounting the intrinsic limitations of the technologies employed.

The first experimental chapter of this thesis (Chapter Two) is centred on the application of ‘computational’ spectrometry, utilising multivariate statistical significance, primarily PCA modelling, to identify which metabolic signals are relevant to the xenobiotic metabolism and toxicity processes under study. Traditional mass spectrometry generally requires an analytical chemist to manually scan through the data using both metabolic and analytical expertise to highlight ions of interest (although software programs exist where “expected metabolites” can be automatically searched for). This process is both time intensive and demands highly skilled individuals for good results. Furthermore, results between different specialists may be different as both experience and observer bias may influence areas of focus. Therefore using multivariate statistics to circumvent or supplement the human element is a highly attractive prospect.

However, in the pilot experiment investigating the effect of diclofenac administration, at a relatively low “therapeutic” dose to a limited number of animals, by urinary UPLC-MS, the statistical approach was unable to detect diclofenac metabolites, despite a clear difference in overall metabolic profiles. Major diclofenac metabolites were, nonetheless, present, which was retrospectively confirmed by manual evaluation of the data. The signatures from the metabolites were being masked by unrelated biological processes which lowered the significance of the drug administration effects. This study was, however, a small scale metabolism study, typical of the type of study performed by conventional drug metabolism groups in early drug development, which lacked both statistical power and sufficient cohort size.

When the same statistical modelling approach was employed in a large scale COMET study in the rat using the drug acetaminophen, with over 600 urine samples and a high “toxicological” dose, the results were markedly different. OPLS modelling highlighted large differences between the test groups, and readily highlighted the key metabolites. Using model validation, the major metabolites were identified as the expected Phase II acetaminophen conjugates, as well as the unchanged parent and the “minor” mercapturates.

Manual evaluation revealed that the major metabolite peaks dominated the chromatography, which in turn made the variable a key contributor to the model statistics.

In principle the statistical approach is mathematically holistic; while this does mean that it is able to detect subtle effects on endogenous metabolism this does not necessarily equate to a superior exploratory tool. The statistics are mechanically detached from any biological process and therefore may occlude discovery of important information because equal weight is given to all variables. In the diclofenac study, the major metabolic differences were due to unavoidable diurnal variation. The trained specialist would be able to identify this and then effectively exclude any signals pertaining to this process, leading to the discovery of the target metabolites; the computational algorithms unfortunately cannot. Ultimately it is this lack of adaptability that may render such methodologies inviable. While there is of course a place for computational metabonomics, especially in the discovery of novel biomarkers and the validation of animal models of disease, its usage should still typically be supplemented by relevant analytical and biological expertise.

One of the analytical problems encountered, and subsequently explored, was the issue of limited continuous instrumental capacity. Thus, if the sample size of the study exceeds the time available on the instrument, current practice involves splitting of the samples into separate runs. Also, as has been frequently mentioned LC-MS is an analytically unstable platform, with a noticeable decrease in sensitivity as contaminants build up on the ESI source. After a certain point the data retrieved can become unreliable, and difficult to compare with previous data. However, one of the further complications is that data from different runs, either required by instrument cleaning or “time-constrained” batch splitting are impossible to merge as subtle differences in the analytical environment between runs renders the data incomparable.

As described in the Thesis, many attempts to utilise statistical batch-correction algorithms have been published, and generally revolve around using the QC samples to normalise the

data. In the acetaminophen study described here, the samples were separated into four batches. Batch correction was attempted using variance stabilisation with a *glog* function prior to median fold change normalisation. While mathematically reducing multiplicative noise and equalising variations across the batches, this approach was unable to correct the data to the extent where it was suitable for combined downstream analysis. Further tampering with the data could result in loss of biologically important information and was therefore deemed an invalid approach. Indeed what could be technically suitable for one study or for one type of data might not be for another, which currently precludes the synthesis of a universally applicable method. Ultimately attempting to correct “bad” data, when the potential for acquiring “good” data exists, is a counter intuitive approach to research. As the origin of the problem lies in the analytical instability of the mass-spectrometry platform, and as such is in essence an analytical issue, an analytical solution was considered to be required.

In chapter three, a rapid gradient microbore UPLC-MS method was described to provide a high-throughput solution for metabolic phenotyping of large sample collections enabling all samples to be run as a single batch. Whilst this does not remove the need for instrument cleaning at some point it does, potentially, remove the need to attempt to “stitch” different batches together. Typically in UPLC-MS a gradient separation, rather than isocratic chromatography, is used. The use of a solvent gradient allows for compounds of very different polarities to be eluted in the same analytical separation. The length, and therefore the steepness, of the gradient is therefore generally proportionate to the resolution of the analysis. By decreasing the gradient elution time the length of the analysis of each individual sample can be reduced from 12 minutes to 1.5 minutes. In addition, by reducing the column diameter, as described here from 2.1 to 1 mm i.d., sample size can be reduced to maximise the use of limited samples such as those obtained from rodent toxicity studies. Whilst employing a more rapid gradient does inevitably reduce chromatographic resolution, equating to around a 70% decrease in ions detected, if there are major metabolic changes

present these will still be represented. Indeed when this rapid gradient method was applied to the previous acetaminophen study, and to a subset of the COMET 2-bromophenol study samples the same discrimination as conventional UPLC-MS was observed. Furthermore, in both sample sets, the same major metabolites were identified.

This approach does indeed have some limitations, as many subtle yet possibly biologically important changes could be missed. However, the advantages of speed and the elimination of the need for separate batches far outweigh the disadvantages, and the rapid gradient approach could be employed as an exploratory investigation to ascertain whether more detailed profiling is required or indeed warranted. As the limitations of the LC-MS platform constitute an existential problem for the future of large scale metabonomic studies it is imperative that high throughput technologies are developed. While the rapid gradient described here is not a panacea of all analytical ills, it is a significant advance without being reliant on complex statistical manipulation of the data.

The collective knowledge and experience gained during the metabonomic method development process described was then applied to a further, smaller scale, toxicology study concerning acute 2-bromophenol, a suspected toxic metabolite of the nephrotoxin bromobenzene. 2-Bromophenol was one of the test compounds of the COMET project and in the present investigation ^1H NMR spectroscopy was combined with UPLC-MS analysis to investigate the toxicity and metabolic fate of this compound. Major Phase II metabolites identified by the UPLC-MS platform included the sulphate and glucuronide, with a number of minor hydroxylated metabolites also observed, and perturbations in the endogenous metabolism (for e.g., hippurate) were also detected. These metabolites of 2-bromophenol were readily detected by multivariate statistical analysis, and indeed represented major discriminating ions. However, glutathione-derived metabolites, indicative of the production of toxic reactive species, such as mecapturates were not detected as being present even using Br-specific isotopic mining of the MS or manual interrogation of the data. This coupled with the histopathology and clinical chemistry results indicated that 2-bromophenol did not elicit a

toxic response when administered, which was in contrast to the results of the literature. However, the results of the study showed that, using a single metabolic profiling approach could be used for the investigation of both xenobiotic and endogenous metabolite profiling.

It was, however, considered possible that achieving toxicity required greater exposure to 2-bromophenol than achieved in the COMET study, either through higher single doses or repeated administration. To investigate the potential for repeat dosing of 2-bromophenol to cause nephrotoxicity it was therefore decided to do a chronic exposure study. After combined ^1H NMR spectroscopic and UPLC-MS metabolic profiling, statistical analysis showed that there was a significant metabolic perturbation due to administration of the compound, with the same major metabolites of 2-bromophenol identified as before. Interestingly, in this study, the 2-bromophenol metabolites formed a major part of the ^1H NMR profile of the urine, which had not been the case for the COMET samples indicating a potentially much higher kidney exposure to the compound. Nevertheless, once again, even with a longer dosing schedule, and apparently higher exposure of the putative toxin the target organ, the histopathology and clinical chemistry analysis still showed no evidence of kidney damage, nor were any reactive metabolites detected.

The lack of toxicity exhibited by 2-bromophenol in these studies was puzzling in the context of the published results. The nephrotoxicity of bromobenzene, as previously mentioned has been well described. Furthermore the work by Lau *et al.* in a series of papers indicated that there was nephrotoxicity associated with the administration of bromophenol to the rat (Lau *et al.*, 1984a). Lau clearly stated that administration of bromophenol the same doses as the work presented in this thesis caused nephrotic damage as well as an increase in blood urea nitrogen concentrations and glutathione depletion in the rat (Lau *et al.*, 1984b). The presented results were also supplemented by histopathological investigation of the kidneys. While the techniques employed by Lau *et al.* for the measure of both BUN and glutathione levels (a Sigma Kit 535A) are outdated by current standards, and the fact that BUN levels have been shown to not be an absolute metric of kidney damage, there seems to be no

explanation for the differences in results especially as the Sprague Dawley rat was used for both Lau's work and ours. However, it should be noted that some of the rats in the Lau study were pretreated with the cytochrome P450 enzyme inducer phenobarbital prior to administration of 2-bromophenol which has been shown to cause some level of hepatotoxicity and enhance the toxicity of other concurrently administered drugs (Dayrell-Hart *et al.*, 1991). If this could be contributing to the toxicity is hard to determine, however, based on the current studies reported in this thesis the conclusion is that 2-bromophenol is not a nephrotoxin in the rat at the doses administered.

In overall conclusion the work presented in this thesis highlight the utility of metabonomics in the field of xenobiotic metabolism and toxicology. It offers information on organism level functional integrity over time after xenobiotic exposure, and allows for the identification of characteristic changes which directly relate to the target tissues or internal processes. Method development effort has allowed for the evaluation of the computational pattern recognition methods, and the identification of limitations of current practice. Insights into potential solutions, both analytical and computational have been presented, with practical real world examples. These advances pave the way for future advances which will enable the continued transition of the field into the "big data" era of high throughput experiments.

6.2 Future Work

A number of future directions for the application of LC-MS for the evaluation of xenobiotic metabolism and toxicity follow from the research described in this Thesis.

- 1) The development of the rapid UPLC-MS methodology offers the potential to re-analyse the remaining COMET samples for biomarker discovery with respect to mechanisms of toxicity and also for xenobiotic metabolites. Combination of this new data with that generated via ^1H NMR spectroscopy will provide a more holistic view of both the changes in endogenous metabolism and its relation to exposure to the various toxins used in COMET.

- 2) Whilst the rapid UPLC-MS method does offer advantages of speed and analytical stability over batch-driven analysis further work on improving the analytical methodology to enable statistical methods of normalisation for batch correction would undoubtedly be of value.
- 3) The rapid UPLC-MS methodology used here worked well for urine samples, but it would also be of value to validate the approach for other sample types such as tissue extracts and blood-derived samples such as serum and plasma.
- 4) In the specific case of 2- bromophenol it would be interesting to expand upon the use of the presence of bromine and obtain further information by ICP-MS on the plasma and urine Br concentrations (and metabolite profiles). Such information might enable the pharmacokinetics, and therefore relative exposures to 2-bromophenol and its metabolites, in the single and multiple dose studies, to be performed. Such information might provide further insights into the disposition and effects of 2-bromophenol in the rat.

6.3 References

Dayrell-Hart, B, Steinberg, S.A., VanWinkle, T.J., Farnbach, G.C., (1991). Hepatotoxicity of phenobarbital in dogs: 18 cases (1985-1989). *Journal of the American Veterinary Medical Association*, **199**, 1060-1066.

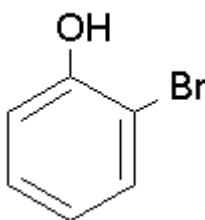
Lau S.S., Monks T.J., Greene K.E., Gillette J.R., (1984a). The role of *ortho*-bromophenol in the nephrotoxicity of bromobenzene in rats. *Toxicology and Applied Pharmacology*, **72**, 539-549.

Lau S.S., Monks T.J., Gillette J.R., (1984b). Identification of 2-bromohydroquinone as a metabolite of bromobenzene and *o*-bromophenol: Implications for bromobenzene-induced nephrotoxicity. *Journal of Pharmacology, Experimental Therapeutics*, **230**, 360-366.

Appendix

COMET consortium study with o-bromophenol: Justification for dose selection**1. Chemical Name and Chemical Properties**

- o-Bromophenol or 2-Bromophenol
- MW: 173.01
- Liquid
- Solubility in aqueous solution: Not applicable
- Storage temperature: Room Temp
- Available from Sigma-Aldrich at 98% purity

2. Toxicity: Kidney toxicant

o-Bromophenol, a toxic phenolic metabolite of bromobenzene, undergoes further metabolic transformation to form a toxic metabolite, 2-bromohydroxyquinone. Both o-bromophenol and 2-bromohydroxyquinone are known to induce nephrotoxicity (Lau et. al 1984a). They cause renal necrosis in rats and mice. The o-bromophenol is converted to 2-bromohydroquinone in the liver. 2-bromohydroquinone is then carried to the kidney where it induces renal toxicity. Both o-Bromophenol and 2-bromohydroquinone deplete glutathione and may be converted to various toxic glutathione conjugates in the kidney (Lau et al.1984b and Monks TJ 1984).

3. Proposed Dosing regimen:

Route of administration: IP injection

Dose groups: Vehicle: Sesame oil

Low-dose: 100 mg/kg (0.58 mmol/kg)

High-dose: 200 mg/kg (1.16 mmol/kg)

Dose Selection Justification:

While the mechanism of o-bromophenol toxicity is well understood, the acute toxicity of o-bromophenol in rats is not very well defined. Lau et. al, 1984a, reported that single dose of >1.28 mmol/kg (221.5 mg/kg), IP, in Sprague Dawley rats, caused increase in serum BUN levels and severe renal necrosis, 48 hrs after treatment. In the same study some mortality was noted at a dose of 1.75 mmol/kg (302.8 mg/kg) dose. Renal lesions at the 1.75 mmol/kg dose consisted of a striking coagulative necrosis of the renal tubular cells in corticomedullary region within 24 to 48 hours after treatment. At lower doses apparent regeneration of damaged tubules was noted by 48 hours post dose. At a higher dose of 1.92 mmol/kg (332.2 mg/kg), o-bromophenol caused significant decrease in renal glutathione levels within 90 minutes post dose. The highest dose of 2.03 mmol/kg (351.2 mg/kg), IP has been tested in rats. There were no reports on effect of single or repeated doses of o-bromophenol beyond 48 hours of treatment in rats or mice. In another study (Rush GF, et al., 1983), single dose of bromophenol isomers greater than 0.2 mmol/kg (34.6 mg/kg), IV, caused severe convulsions and death. Convulsions were not observed at 0.14 mmol/kg (24.2 mg/kg), IV. No data on plasma exposure was available to correlate the IV to the IP dose.

After personal communication with Dr. Lau regarding the acute effects of o-bromophenol following IP injection, we decided to conduct an in-house pilot study to give us a better understanding of the maximum tolerated dose and histopathological changes in the kidney on Day 7 after a single IP dose.

Six male Sprague Dawley rats were randomly assigned to the following dose groups listed in the table below. Three rats were sacrificed at 48 hrs post dose (Day 2) and the remaining three rats were sacrificed on Day 7 following a single intraperitoneal injection of o-bromophenol. The vehicle was sesame oil. Clinical observations were conducted daily. Blood samples for serum chemistry were collected on Day 2 and 7 prior to necropsy. Only kidneys were evaluated for microscopic changes.

Group	Number of Animals/ group (M)	Concentration of o-bromophenol (mg/mL)	Dose Volume (ml/kg)	Dose of o-bromophenol (mg/kg)	Dose of o-bromophenol (mmol/kg)
1	6	0	1	0 (vehicle)	0
2	6	25	1	25	0.15
3	6	50	1	50	0.29
4	6	100	1	100	0.58
5	6	200	1	200	1.16

The dose selection for the pilot study was based on the study conducted by Lau et al. (1984a) and personal communication with her, where she indicated that convulsions are noted immediately following dosing at doses ~0.57 mmol/kg (98.6 mg/kg). For the metabonomic study the high dose should cause changes in the kidney but no mortality. The doses for this study were chosen to evaluate the maximum tolerated dose for o-bromophenol following single i.p. administration.

One animal in the 200 mg/kg dose group that was scheduled to be sacrificed on Day 7 was sacrificed early on Day 2 due to deteriorating clinical condition. No convulsions were noted at any dose, but slight transient tremors were noted in 4/6 animals at the 200 mg/kg dose immediately post dose, however. Most animals in the 200 mg/kg dose group had slight labored respiration and inactivity. These effects were minor and recovered by Day 2-3 post dose. In lower dose groups, slight transient labored respiration was noted in fewer animals, mainly on the day of dosing.

Clinical pathology changes were noted in one 200 mg/kg rat that was sacrificed early on study Day 2. The changes included increased blood urea nitrogen, glucose, aspartate aminotransferase, alanine aminotransferase, total bilirubin, serum potassium, serum phosphorus, and globulin. Decreased serum sodium, chloride, and albumin/globulin ratio were also present in that animal. There were no compound-related clinical pathology changes at 25, 50, or 100 mg/kg.

Gross pathology changes attributed to dosing with o-bromophenol included white discolored or pale livers in one of three animals at 200 mg/kg on Day 2 and 7, and one of three 100 mg/kg animals on Day 2. Foreign material (dosing solution) was present in the peritoneal cavity or peritoneal skeletal muscle in several animals on Day 7.

Microscopically, Day 2 treatment-related changes in the kidney consisted of slight to moderate tubular regeneration, with necrotic debris in tubular lumens, in two of three animals at 200 mg/kg. Two of three 200 mg/kg animals at Day 7 had similar minimal to slight changes. At 100 mg/kg, minimal tubular regeneration, with necrotic debris in tubular lumens, was present in one of three animals at Day 2 and two of three animals at Day 7. There were no compound-related anatomic pathology changes at 25 or 50 mg/kg.

The clinical and anatomic pathology changes present in this study are much milder than those reported by Lau et al. One high dose animal in this study had an elevated BUN, but the kidney changes in that animal appeared no different than the other high dose animals that did not have an elevated BUN. Increase in BUN in that one animal could be due to pre-renal reasons. The many other clinical pathology changes in that animal can not be explained by examination of the kidney sections. Slight hemolysis present in the blood sample may have contributed to the

elevated AST, but not the other changes. The mild tubular changes present in the 100 and 200 mg/kg kidneys were not severe enough to affect the clinical health of the animals.

At 200 mg/kg (1.16 mmol/kg), the kidney changes are associated with some (but not excessive) clinical signs of toxicity. Therefore a high dose of 200 mg/kg (1.16 mmol/kg) is recommended. At the 100 mg/kg dose the histopathological changes were very subtle and no overt signs of toxicity were noted and therefore a low dose of 100 mg/kg (0.58 mmol/kg) is recommended for the definitive metabonomic study.

References:

1. Lau S.S., Monks T.J., Greene K.E., Gillette J.R., 1984a. The role of *ortho*-bromophenol in the nephrotoxicity of bromobenzene in rats. *Toxicol. Appl. Pharmacol.* **72**:539-549.
2. Lau S.S., Monks T.J., Gillette J.R., 1984b. Identification of 2-bromohydroquinone as a metabolite of bromobenzene and *o*-bromophenol: Implications for bromobenzene-induced nephrotoxicity. *J. Pharmacol. Expt. Ther.* **230**(2): 360-366.
3. Monks T.J., Lau S.S., Pohl L.R., Gillette J.R. 1984. The mechanism of formation of *o*-bromophenol from bromobenzene. *Drug Met. Disp.* **12**(2):193-198.
4. Rush G.F., Newton J.F., Maita K., Kuo C-H, Hook J.B. 1984. Nephrotoxicity of phenolic bromobenzene metabolites in the mouse. *Toxicology.* **30**:259-272.

COMET CONSORTIUM SUMMARY

O-BROMOPHENOL SINGLE DOSE METABONOMICS STUDY IN MALE SPRAGUE DAWLEY RATS

INTRODUCTION

The objective of this study was to characterize the toxicity of o-bromophenol and to provide serum, urine and tissue samples for subsequent evaluation of the profile of endogenous metabolites using NMR spectroscopy.

In this study, o-bromophenol was administered as a single intraperitoneal injection to male Sprague-Dawley rats. The following dose levels were selected:

Group	Dose	No. of Rats	Necropsy
1	Control (corn oil)	10	48 h (Group 1A) or 168 h (Group 1B) after dose
2	100 mg/kg	10	48 h (Group 2A) or 168 h (Group 2B) after dose
3	200 mg/kg	10	48 h (Group 3A) or 168 h (Group 3B) after dose

VIABILITY/MORTALITY

There were no deaths.

CLINICAL SIGNS

There were no clinical signs attributable to treatment with o-bromophenol.

BODY WEIGHTS

Most animals in all groups, but particularly those treated at 200 mg/kg, lost weight during the first 24 hours after dosing. This was attributed, in part, to the dosing procedure. Recovery was apparent in all groups from Day 2 although at a slightly slower rate in the high dose animals compared with the control and low dose animals. By Day 8, body weights were similar for all three groups.

CLINICAL BIOCHEMISTRY

There were no consistent patterns of variation in the clinical biochemistry parameters to suggest an effect of treatment with o-bromophenol.

URINALYSIS

Statistically significant decreases in urinary osmolality, creatinine, potassium and glucose were observed in high dose animals between 0 and 8 hours after dosing on Day 1.

ORGAN WEIGHTS

There was no consistent pattern of variation in organ weight data to indicate an effect of treatment with o-bromophenol.

MACROSCOPIC FINDINGS

No macroscopic changes attributable to the administration of o-bromophenol were observed in animals killed at 48 or 168 hours after dosing.

MICROSCOPIC FINDINGS

For sub-group A animals (Day 3), a minor multifocal tubular degeneration/regeneration was observed in the kidney of two high dose animals. This finding was characterized by pale basophilic epithelial cells and necrotic cellular debris within the lumen of the proximal tubules.

For sub-group B animals (Day 8), the spectrum of microscopic findings in the liver and kidney of treated animals was generally comparable with that in controls, suggesting a complete reversal of the findings seen in the kidney of the high dose Subgroup A animals.

ASSESSMENT AND CONCLUSION

o-Bromophenol is a toxic phenolic metabolite of bromobenzene which undergoes further metabolic transformation to form a toxic metabolite, 2-bromohydroxyquinone. Both o-bromophenol and 2-bromohydroxyquinone are known to induce nephrotoxicity in rats and mice, evident as renal necrosis. The o-bromophenol is converted to 2-bromohydroxyquinone in the liver. 2-bromohydroxyquinone is then carried to the kidney where it induces renal toxicity. Both o-bromophenol and 2-bromohydroxyquinone deplete glutathione and may be converted to various toxic glutathione conjugates in the kidney.

In this study, the decreases in urinary osmolality, potassium and glucose recorded at the high dose between 0 and 8 hours after dosing on Day 1 and the minor multifocal tubular degeneration/regeneration seen in the kidney of two high dose animals are considered to be representative of the expected toxic effects of the test article at sub-lethal doses.

The study successfully completed the objectives of characterizing the toxicity of the test article and generating serum, urine and tissue samples for the determination of the metabolic profile of the test article by NMR spectroscopy.



Study Plan (Discovery)

Study number	3844DR
GEL ID	INF.000-156-829
Test compound(s)	AZ13719017, AZ13647935 & 2-Bromophenol
Project identifier	PMB
Document status	DRAFT
GLP compliance	No
Date approved	<< >>

Polymyxin B, Polymyxin B nonapeptide & 2-Bromophenol: 1-Day Subcutaneous Investigative Kidney Toxicity Study in Han Wistar Rats

This is not a GLP study but all critical operations and methods will be performed according to current written local Standard Operating Procedures (SOPs) unless otherwise noted.

Study Director/Study Manager	Location of study
-------------------------------------	--------------------------

Nakpangi Johnson	DSM Boston
------------------	------------

Responsible Technician/Scientists	Location
--	-----------------

Crystal Brown, Study Coordinator	DSM Boston
Letitia Cheatham, Study Co-coordinator	DSM Boston
Natalie Kierstead, Pathologist	DSM Boston
Debra Snow, Histopathology	DSM Boston
Marie Blais, Clinical Pathology	DSM Boston
Aixiang Xue, Formulations	DSM Boston
Richard Goodwin, DMPK	DMPK Alderley Park
Peter Webborn, DMPK	DSM Alderley Park
John Swales, DMPK	DSM Alderley Park
Nick Adesina-Georgiadis, DMPK	DSM Alderley Park

Study Timings	Date(s)
Start of pre-study period	<<dd Month yyyy>>
Proposed experiment start date	<<dd Month yyyy>>
Proposed necropsy start and finish date	<<dd Month yyyy>>
<<additional timings if needed>>	<<dd Month yyyy>>
Final Study Report	<<dd Month yyyy>>
Data Storage and Archiving	Archives of DSM Boston

OBJECTIVE

The purpose of the study is to investigate the disposition, retention and effects of polymyxin B, polymyxin nonapeptide and 2-bromophenol exposure on the kidney. This study will also explore the endogenous metabolic profiles of the rat using metabonomic and tissue imaging methods with the aim of evaluating these as a means of obtaining suitable and specific biomarkers for polymyxin nephrotoxicity.

STUDY DESIGN

Species and strain	Rat, Han Wistar (Charles River)
	Animal identification and conditions of housing, acclimatisation, environment, diet and water will be in accordance with the current Standard Operating Procedures
Number and gender	48 males, housed 2/cage
Age and weight at start of dosing	Approximately 10 weeks old 300-325 g

Table 1 **Dose Groups and Animals**

Group	Compound / Vehicle	Study Duration	Animal reference numbers	Dose levels mg/kg/dose	Dose Volume mL/kg	Concentration of dosing formulation mg/mL
1	Vehicle (Saline)	Main D3	1-4	0	5	0
2	Vehicle (Saline)	Recov D7	5-8	0	5	0
3	AZ13719017	Main D3	9-12	2	5	0.4

Group	Compound / Vehicle	Study Duration	Animal reference numbers	Dose levels mg/kg/dose	Dose Volume mL/kg	Concentration of dosing formulation mg/mL
4	AZ13719017	Main D3	13-16	6.25	5	1.25
5	AZ13719017	Recov D7	17-20	6.25	5	1.25
6	AZ13647935	Main D3	21-24	10	5	2
7	AZ13647935	Recov D7	25-28	10	5	2
8	Vehicle (corn oil)	Main D3	29-32	0	5	0
9	Vehicle (corn oil)	Recov D7	33-36	0	5	0
10	2-bromophenol	Main D3	37-40	100	5	20
11	2-bromophenol	Main D3	41-44	200	5	40
12	2-bromophenol	Recov D7	45-48	200	5	40

Dosing information

Days 1 to 2: The animals in Groups 1 to 7 will be dosed 4 times daily for 2 days by subcutaneous injection in the intrascapular region, approximately every 6 hours; groups 8 to 12 will be dosed once per day orally. Control animals in Groups 1 and 2 will be dosed with sterile saline for injection. Groups 3 to 5 will be dosed with Polymyxin B (2 or 6.25 mg/kg/dose), groups 6 and 7 with Polymyxin nonapeptide (6mg/kg/dose). Groups 8 and 9 will be dosed with corn oil (vehicle for 2-bromophenol) and animals in groups 10 through 12 will be dosed with 2-Bromophenol (100 or 200 mg/kg/day).

Day 3: All groups will be dosed once with vehicle, Polymyxin B (2 or 6.25 mg/kg/dose), Polymyxin nonapeptide (6mg/kg/dose) or 2-Bromophenol (100 or 200 mg/kg/day), as previously. All groups, with the exclusion of recovery groups 2, 5, 7, 9 and 12, will be sacrificed 4 hours post final dose.

Please see table in appendix for overview of dosing, urine collection and bleeding.

Dose justification

Doses for both Polymyxins and 2-Bromophenol are identical to the dose levels of several other polymyxin analogues investigated in previous studies (in house and external). Clinical signs such as decreased motor activity, swollen limbs, feet and muzzles, cold extremities, cyanosis, excessive scratching and red discoloration of the limbs, feet and muzzles have been observed in Polymyxin dosing. However, it is not anticipated that signs will develop in the low dose groups and those in the high dose should not persist.

Test compound, formulation and hazard information

Test Compound	Batch	Molecular weight (g/mol)	Base factor	Purity	Formulated in
AZ13719017 Polymyxin B	001	1203.0	1	100	Sterile saline for injection
AZ13647935 Polymyxin B nonpeptide	002	963.10	1	95	Sterile saline for injection

2-Bromophenol

Sigma Aldrich

Lot#/Exp Date:

Storage

The formulations need to be protected from light.

Containers will be labelled with study number, ID of compound, concentration, and manufacturing date.

Remaining formulations

All remaining formulations will be sent to Aixiang Xue in LAS for disposal.

Hazards

All test substances undergoing safety assessment and evaluation must be regarded as potent pharmacologically active agents with unknown toxic properties. Precautions for high hazard materials must be taken since there is insufficient data for assigning to a specific health and hazard group. Handling and use of the test substances are restricted to designated areas. Use these materials in laboratory fume hoods or glove boxes when preparing formulations or other solutions. Wear suitable protective clothing, gloves, and safety glasses. If any specific hazard is identified, appropriate precautions must be taken.

Samples for analysis

Appropriate samples of the dosing formulations will be taken after preparation and frozen. Samples may be used for future concentration verification analysis.

Information to be provided with samples: Sample list. Labels on the vials should be marked with study number, substance code or control, nominal concentration (parent form), sample number and sampling date.

GOOD STATISTICAL PRACTICE

This study is confirmed as being a non-standard study design. A statistician has performed a statistical review of the study design. A copy of the statistical review will be archived with the study data. If there are any planned changes to the study that affect GSP principles, a revised statistical review will be produced and archived with the study data.

PARAMETERS TO BE MEASURED

In-life observations

Clinical observations	8 am and 2pm doses: Animals will be observed pre-dose and at least 1 to 3 hours post-dose. 8 pm, 2 am and weekend/holiday doses: Animals will be observed pre-dose and at approximately 1 hour post-dose. Necropsy day dose: The animals will also be observed the morning of necropsy on Days 3 or 7 prior to dosing and 1-3 hours post-dose.
-----------------------	--

Body weight	Dosing phase: Will be taken at randomization; then recorded daily prior to dosing and/or until necropsy.
-------------	--

Food and water consumption	Will not be recorded.
----------------------------	------------------------------

Laboratory

Toxicokinetics

Groups:	1-12
Animals per time point:	4/group

Time points and site/route/volume:	Groups 1-3 and 5, Polymyxin B/Polymyxin Nonapeptide: (following the 2pm dose) 6 (immediately prior to dosing), 7, 9, 12 (immediately prior to dosing), 30 (immediately prior to dosing), 31, 33 35 & 51 (terminal) hrs post dose 1 on Day 1. Group 4, 6 and 7, Recovery: (following the 2pm dose) 6 (immediately prior to dosing), 7, 9, 12 (immediately prior to dosing), 30 (immediately prior to dosing), 31, 33 36. At hours 6, 7, 9, 12 hrs daily thereafter once dosing has concluded. Group 8-9, 2-Bromophenol (following the single dose) 6, 7, 9, 12, 30, 31, 33 35 & 51 (terminal). Groups 10-11, Recovery: (following the single dose) 6, 7, 9, 12, 30, 31, 33 36. At hours 6, 7, 9, 12 hrs daily thereafter once dosing has concluded.
	Tail vein*, 0.2 mL whole blood *Saphenous veins and/or retro-orbital sinus will be used if tail vein attempt is unsuccessful. If an alternate route is used or partial sample collected, the TK timesheet and platemap will contain a note detailing the collection problem.
Sample preparation/storage and anticoagulant:	Samples will be taken in lithium heparin tubes and kept on ice. Samples will be centrifuged for 3 minutes at 1500 g under refrigeration (2 to 8°C) within approximately 0.5h of collection. Plasma will be transferred into 0.5 mL polypropylene tubes with snap caps and immediately frozen. The plasma samples will be kept frozen at approximately -70°C until analysis.
Samples will be analysed for:	AZ13709017, AZ13649735, 2-Bromophenol
Parameters calculated:	C_{max} , $AUC_{(0-24)}$, $t_{1/2}$
Estimated sample ship date:	

Samples sent to: Nick Adesina-Georgiadis c/o
Peter Webborn
Mereside
35/F/11
Alderley Park, Macclesfield,
Cheshire, England
SK10 4TF
+44 1625 233101

Clinical pathology

Groups / Animals	Days	Sampling route, volume, anti-coagulant and Time points
1-12/1-48	-1,2,3	Urine chemistry and DIKI Biomarkers.
1, 3, 4, 6, 8, 10, 11/ 1-4, 9-16, 21-24, 29-32, 37-44	3	Vena cava, isoflurane, plasma chemistry up to 4.0 mL (lithium heparin) hematology up to 2.0 mL (EDTA)
2, 5, 7, 9, 12/5-8, 17-20, 25-28, 33-36, 45-48	4-7	Urine chemistry and DIKI Biomarkers.
2, 5, 7, 9, 12/5-8, 17-20, 25-28, 33-36, 45-48	7	Vena cava, isoflurane, plasma chemistry up to 4.0 mL (lithium heparin) hematology up to 2.0 mL (EDTA)

Hematology

Erythrocytes (RBC)	Platelets (PLT)
Hemoglobin (HGB)	Leukocytes(WBC)
Hematocrit (HCT)	Absolute Neutrophils(ANEUT)
Mean red cell volume (MCV)	Absolute Lymphocytes(ALYM)
Mean red cell hemoglobin (MCH)	Absolute Monocytes(AMONO)
Mean red cell haemoglobin concentration (MCHC)	Absolute Eosinophils(AEOS)
Red cell distribution width (RDW)	Absolute Basophils(ABASO)
Reticulocytes (RET%)	Absolute Large unstained cells (ALUC)
Absolute Reticulocytes (ARET)	

A peripheral blood smear will be made on all hematology samples but only examined where automated analysis indicates this is necessary or at the discretion of the Pathologist or Study Director.

Plasma chemistry

Creatinine: CREA	Blood Urea Nitrogen: BUN
Total Protein: TP	Albumin: ALB
Cholesterol: CHOL	Triglycerides: TRIGS
Total Bilirubin: TBIL	Glucose: GLU
Sodium: NA	Potassium: K

Alanine aminotransferase: ALT

Globulin: GLOB¹

Albumin/Globulin ratio: AG¹

Inorganic phosphorus: IPHOS

¹ Calculated parameter

Alkaline phosphatase: ALP

Aspartate transferase: AST

Total Calcium: CA

Abnormal color and lipemia to be noted

Urine analysis

Baseline urines will be collected on Day –1 for all animals. All animals will also have individual urine samples collected on ice, daily on Days 1 to 3. The rats will be placed in metabolic cages immediately after dosing for 6 hours.

For the recovery groups 2, 5, 7, 9 and 12, individual urine samples will be collected daily on Days 4 through 7, every 24hr by placing rats in metabolic cages for 6 hours.

Water will be available during the collection period but food will be withdrawn.

Urine Chemistry

Volume¹

Glucose

Urea

Protein

Creatinine

¹Measured by graduated conical tube

Urinary Biomarkers

AGST (alpha glutathione-s-transferase)

Albumin

GSTYb1 (glutathione-s-transferase Yb1)

Creatinine

TIM-1 (KIM-1, kidney injury molecule)

Clusterin

Urine aliquots for analysis (All Groups)

If sufficient volumes are available, remaining unprocessed urine will be transferred to labeled microcentrifuge tubes and stored at -70°C or below until transfer to DMPK for analysis.

Pathology

Necropsy and histopathology: Day 3 (Groups 1, 3, 4, 6, 8, 10, 11); Day 7 (Groups 2, 5, 7, 9, 12)

Study animals in Groups 1 to 12 will be subjected to necropsy. Animals will have free access to food and water overnight before the necropsy. The body weight will be recorded prior to scheduled necropsy. Animals will be euthanized by exsanguination from the caudal vena cava under isoflurane anesthesia. External features of the animal, the thoracic and abdominal cavities and contents will be examined. Macroscopic abnormalities will be recorded and collected.

Moribund and Decedent animals: The study director and study pathologist must be contacted in the event of a moribund or decedent animal. Moribund and decedent animals from Groups 1 to 12 will be necropsied when possible and study plan required tissues collected. Additional tissues will be collected at the discretion of the study pathologist. In both cases, collected specimens will be processed and examined at the discretion of the study pathologist. Organ weights will not be taken from moribund or early decedent animals. Blood samples will only be drawn on moribund animals, not decedents and will be stored on ice until the plasma is separated. In the event of moribund animals on the weekend/holiday, hematology and plasma chemistry samples will be collected. The hematology (EDTA) whole blood samples will be refrigerated at 2-8°C and the plasma chemistry samples (Lithium heparin) will be spun down and the plasma frozen at -70°C by LAS.

The following tissues will be sampled, processed and examined:

Kidney

Tissues processed and examined by light microscopy: All preserved tissues will be processed to paraffin blocks. Blocks will be sectioned, stained with hematoxylin and eosin and examined. Additional sections and stains may be requested by the study pathologist to determine the significance of any observed findings.

Organ Weights (Groups 1-12)

The following organs will be weighed intact and unfixed at scheduled necropsy:

Kidneys

Paired organs will be weighed together and combined weight recorded. Organ weights will not be taken from decedents or moribund animals.

DMPK Sampling (Groups 1-12)

The following tissue will be collected for MALDI and exposure. Upon collection, the tissue should be weighed and recorded. Tissue will be kept frozen until transfer to DMPK for analysis.

MALDI Groups 1-12	Animals in groups 1-12 will have half of the right kidney (transverse section) collected. Organs will be snap frozen slowly by free floating in chilled isopentane (2-methybutane) in disposable weigh boats on dry ice. Once frozen they should be placed in aluminium foil, transferred on dry ice and stored at -80°C Samples will be sent to: Richard Goodwin 35F53/6 Mereside, Alderley Park, Macclesfield, Cheshire, England SK10 4TG +44 1625 518751
Bioanalysis Groups 1-12	The other half of the right kidney will be snap frozen in liquid nitrogen. Samples will be sent to: Nick Adesina-Georgiadis c/o Peter Webborn Mereside 35/F/11 Alderley Park, Macclesfield, Cheshire, England SK10 4TF +44 1625 233101

REPORTING AND STATISTICAL ANALYSIS

A summary study report will be prepared according to the current Discovery Reporting guidelines. Statistical analysis will be performed at the discretion of the Study Director following review of the data. Any analysis done will be recorded in the study file.

REFERENCES

3579DR: AZ11983219: Investigative subcutaneous tolerability study

3587DR: AZ13584324: Investigative subcutaneous tolerability study

3545KR: AZ11983219: Colistin Sulphate 7-Day repeat Intravenous Infusion in Han Wistar Rats

3611DR: AZ119983219 & AZ13647935: Investigative 2-Day Tolerability Study in Han Wistar Rats

3693DR : AZ13695479 & AZ13719017: Investigative 2-Day Subcutaneous Tolerability Study in Han Wistar Rats

3607DR: AZ13695479 & AZ13719017: 7 Day Subcutaneous Investigative Kidney Toxicity Study

3716DR: AZ13708229 & AZ13712953: Investigative 2-Day Subcutaneous Tolerability Study in Han Wistar Rats

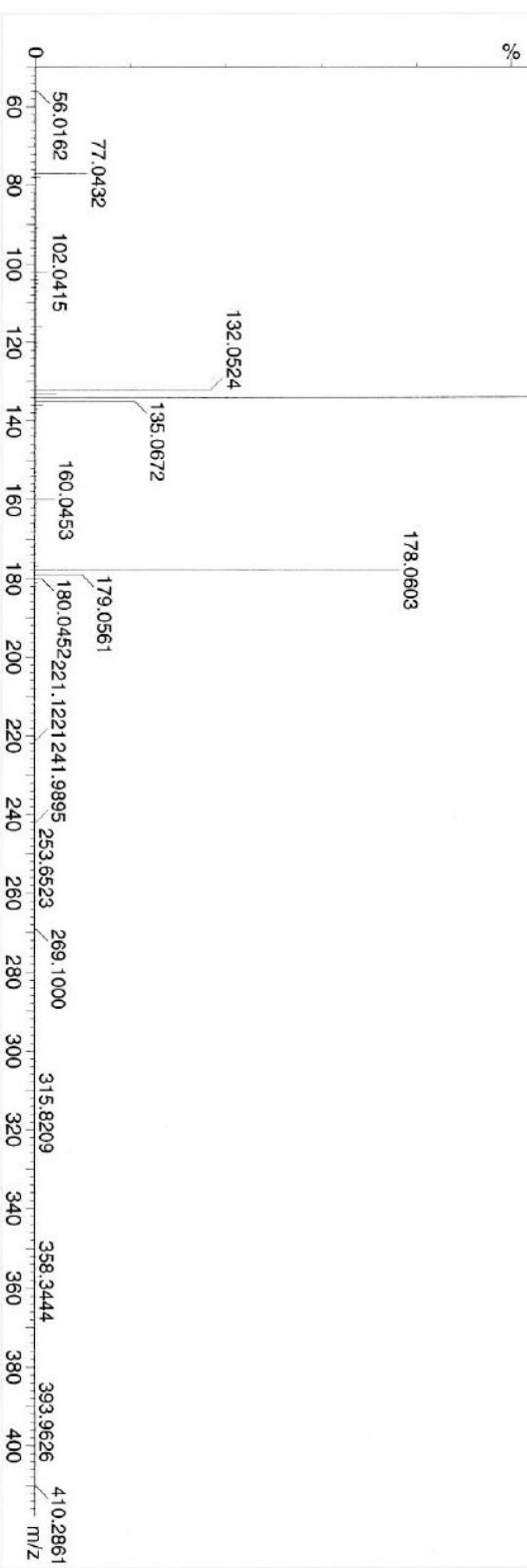
Abdelraouf K, Braggs KH, Yin T, Truong LD, Hu M, Tam VH (2012). Characterization of polymyxin B-induced nephrotoxicity: implications for dosing regimen. *Antimicrob. Agents Chemother.* 59 (9): 4625-4629.

R16 QC MSMS 1 in 5 m/z 600

13JUL15_R16_NEG_MSMS_019 165 (3.876) Cm (155:178)
134.0681

3: TOF MSMS 178.06ES-
2.07e5

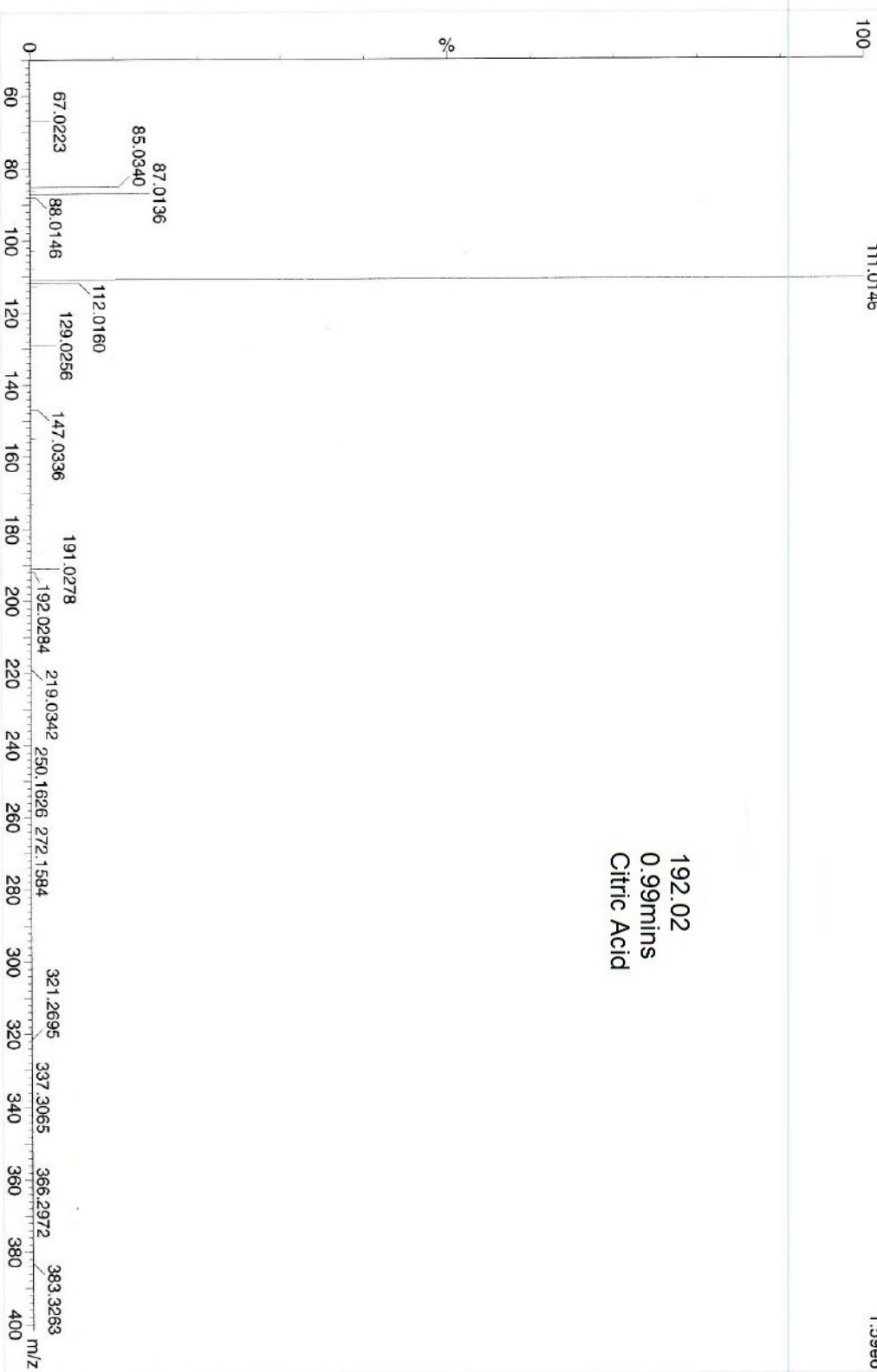
178.05
3.84mins
Hippurate



R16 QC MSMS 1 in 5 m/z 600

13JUL15_R16_NEG_MSMS_019 514 (1.019) Cm (491:565)
111.0146

1: TOF MSMS 191.02ES-
1.59e6



192.02
0.99mins
Citric Acid

R16 HD 8 hr R006742 MSMS m/z 600

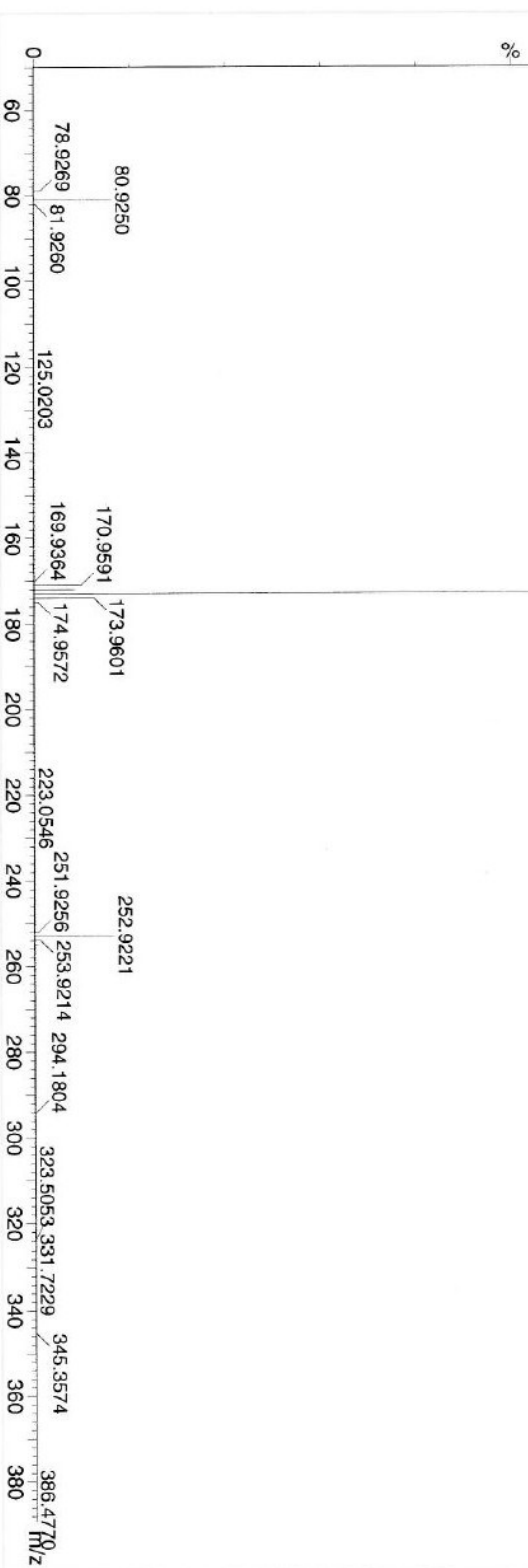
13JUL15_R16_NEG_MSMS_020 99 (4.376) Cm (93:125)

4: TOF MSMS 252.90ES-

100 172.9664

5.70e7

252.90
4.41mins
2-bromophenol Sulphate



R16 QC MSMS 1 in 5 mz 600

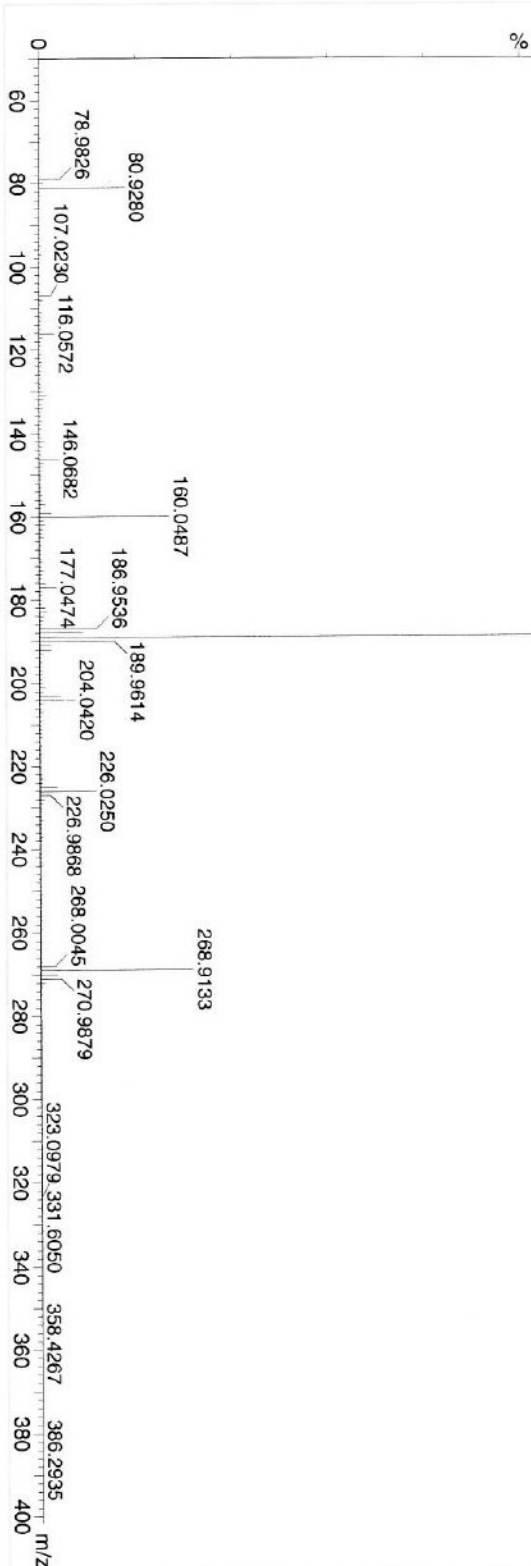
13JUL15_R16_NEG_MSMS_019 216 (3.223) Cm (204:240)

100

188.9483

2: TOF MSMS 268.89ES-
1.31e5

268.89
3.09mins
Hydroxy 2-Bromophenol Sulphate

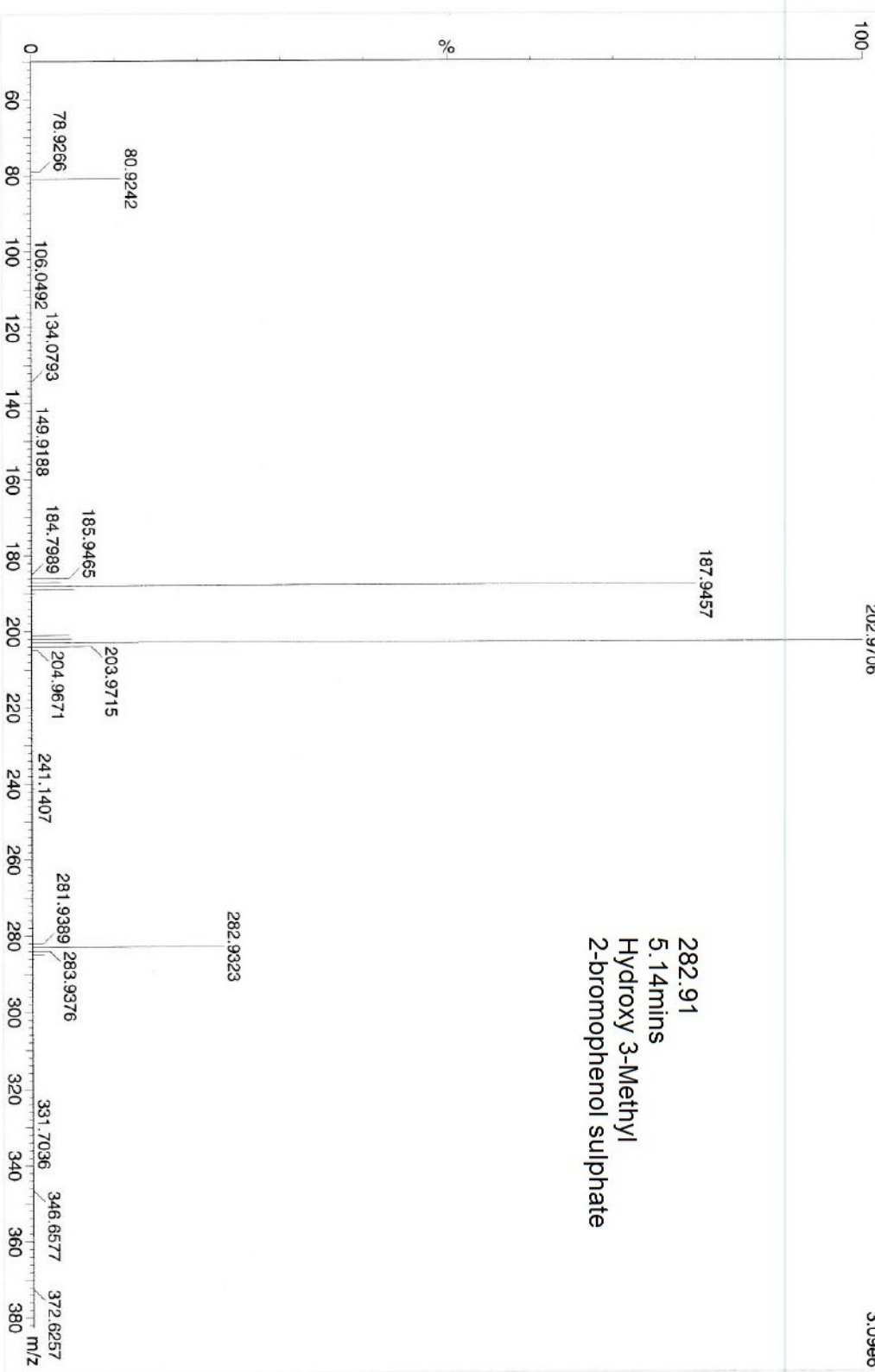


R16 HD 8 hr R006742 MSMS m/z 600

13JUL15_R16_NEG_MSMS_020 95 (5.018) Cm (85:110)

202.9706

5: TOF MSMS 282.91ES-
3.09e6



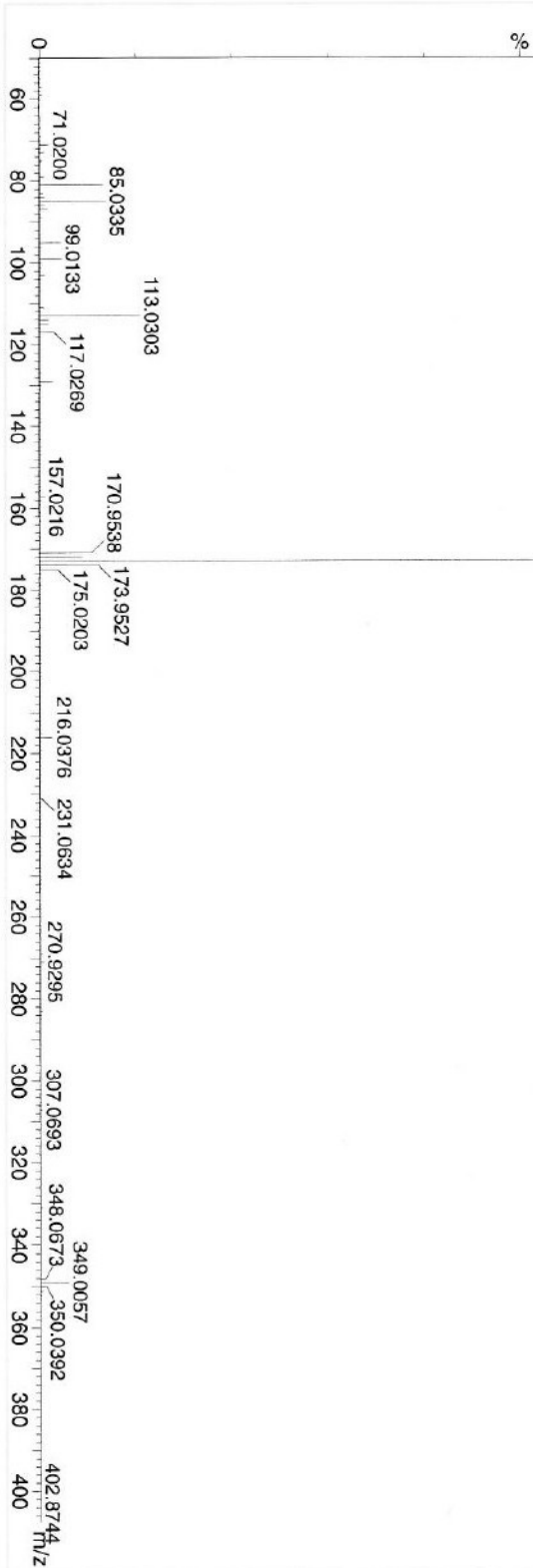
282.91
5.14mins
Hydroxy 3-Methyl
2-bromophenol sulphate

R16 QC MSMS 1 in 5 m/z 600

13JUL15_R16_NEG_MSMS_019_42 (4.471) Cm (40:45)
172.9524

4: TOF MSMS 348.98ES-
2.36e5

348.98
4.46mins
2-bromophenol glucuronide



R16 QC MSMS 1 in 5 higher ce
13JUL15_R16_NEG_MSMS_021.42 (4.465) Cm (41.44)

100

6: TOF MSMS 718.94ES-
718.9789 5.2864

