Towards a Better Understanding of Uraemic Molecules

by

Mitra Nouri Koupaei

A thesis submitted in fulfilment of the requirements for the degree of

Master of Science (MSc.)

UNIVERSITY OF TASMANIA

School of Chemistry

Hobart campus

Submitted April 2013

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other diploma or degree in any tertiary institution, and to the best of my knowledge contains no copy or paraphrase of material previously published or written by another person, except where due reference is made in the text of the thesis.

Mitra Nouri Koupaei

This thesis may be available for loan and limited copying in accordance with the *Copyright Act* 1968.

Mitra Nouri Koupaei

University of Tasmania, Hobart

April 2013

Ethical conduct

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University. This work was undertaken with ethics approval from the Tasmania Health and Medical Human Research Ethics Committee, Reference numbers H9693 and H11758.

Acknowledgements

I could not be able to finish this Master's thesis without the help and support of my supervisors, family and friends. I want to thank and show my gratitude to the following people who helped me to make this happen:

- My supervisory team: Prof. Emily Hilder, A/Prof. Robert Shellie and Prof. Matthew Jose for their encouragement, comments, and support.
- I would like to express my deepest appreciation to Emily Hilder. This thesis would not have been possible without the guidance, valuable suggestions, patience, and support of her.
- Dr. Andras Gaspar for his valuable help and support during his stay in Hobart and for his true friendship.
- A/Prof Michael Breadmore for his suggestions and sharing his scientific experience. Also Dr. Rosanne Guijt for her kind support.
- Thanks to people from CSL, A/Prof Noel Davies and Dr. Richard Wilson for their advice and helpful suggestions.
- Dr. Cameron Johns for his useful comments and experience.
- Special thanks to the special person and my husband, Mohammad, who is my all-time support in the absence of my family who are thousands of kilometres away from me. Thanks for his advices, suggestions and proof reading of my thesis.
- My supportive friends in ACROSS, Kriszta, Jessica, Éadaoin-, Lea, Aliaa, Esme, and Benjamin for their true friendship and making the lab environment friendly and energetic. Thank you all for listening to me and allowing me to vent in the time of disappointment. I already miss

some of you who left Hobart but you will be in my mind forever. Also thanks to lab manager Murray Frith and computer technician Anthony Malone for their kind efforts.

- Manish for being a supportive friend during my studying. Many thanks for listening to me and understanding me.
- Thanks to my family, Mum and Dad for their love, support and encouragement. Also thanks to my uncle and my aunty for supporting me as the only family members that I have in Australia.
- Thanks to my best friend Hoda for her love and support. She definitely has had a great role in whatever I achieved in my life.
- And very special thanks to Dr. Joan Isherwood for being a great friend, advisor and guide. I consider myself a very lucky person to have such a nice friend here.

Abbreviations:

ACN: Acetonitrile

ADMA: Asymmetric dimethylarginine

ANOVA: Analysis of variance

APCI: Atmospheric pressure chemical ionisation

API: Atmospheric pressure ionisation

APPI: Atmospheric pressure photo-ionisation

BGE: Background electrolyte

BSTFA: Bis(trimethylsilyl)trifluoroacetamide

CE-MS: Capillary electrophoresis-mass spectrometry

CGE: Capillary gel electrophoresis

cIEF: Capillary isoelectric focusing

CKD: Chronic kidney disease

CLD: Chemiluminescence detector

CMPF: 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid

CNS: Central nervous system

CRF: Chronic renal failure

CSF: Cerebrospinal fluid

CZE: Capillary zone electrophoresis

DESI: Desorption electrospray-ionisation

DI: Direct immersion

DS: Dextran sulfate

ECD: Electron capture detector

EI: Electron Impact

EOF: Electroosmotic flow

ESI: Electro spray ionisation

FID: Flame inonisation detector

FT-ICR-MS: Fourier transform (FT)-ion cyclotron resonance MS

GC-MS: Gas chromatography- mass spectrometry

GFR: Glomerular filtration rate

HD: Hemodialysis

HDF: Hemodiafilteration

HILIC: Hydrophilic interaction chromatography

HMBD: Human metabolome database

HPLC: High performance liquid chromatography

HS: Headspace

IA: Indoleacetic acid

IS: Indoxyl sulfate

IT: Ion trap

K: Partition coefficient

KNN: K-nearest neighbour

LC-MS: Liquid chromatography- mass spectrometry

LLE: Liquid-liquid extraction

MALDI: Matrix-assisted laser desorption-inoisation

MEKC: Micellar electrokinetic capillary chromatography

MS: Mass spectrometry

MSTFA: N-methyl-N-(trimethylsilyl)trifluoroacetamide)

MVA: Multivariate analysis

NPLC: Normal phase liquid chromatography

PB: Polybrene

PCA: Principle component analysis

PCS: *p*-cresyl sulfate

PDADMAC: Poly(diallyldimethylammonium) chloride

PDMS: Polydimethylsiloxane

PEG: Polyethylene glycols

PLS-DA: Partial least squares discriminant analysis

PP: Protein precipitation

PSS: Poly (sodium 4-styrene-sulfonate)

PVS: Poly (vinyl sulfate)

Q-Ion trap-MS: Quadruple-ion trap-MS

Q-MS: Quadruple-MS

RI: Retention index

RP: Reversed phase

SDMA: Symmetric dimethylarginine

SIMCA: Soft independent modelling by class analogy

SPE: Solid phase extraction

SPME: Solid-phase micro-extraction

TID: Thermionic inonisation

TMS: Trimethylsilyl

TOF-MS: Time-of-flight-MS

TQ: Triple quadruple

Triple-Q-MS: Triple quadruple-MS

UAE: Ultrasonic-assisted extraction

UPLC: Ultra performance liquid chromatography

List of presentations

- Mitra Nouri Koupaei, Emily F. Hilder, Robert A. Shellie, Matthew D. Jose, **Development of a gas chromatography-mass spectrometry method for the analysis of uraemic molecules**. APCE conference, 27-30 Nov. 2011, Hobart. (poster)
- Mitra Nouri Koupaei, Emily F. Hilder, Robert Shellie, Matthew D. Jose, **Towards a better understanding of uraemic molecules: Development of a capillary electrophoresis-mass spectrometry procedure**, 27-30 June 2011, Cairns, Australia. (poster)
- Christopher T. Desire, Mitra Nouri Koupaei, Jianfeng Li, Emily F. Hilder, Robert A. Shellie, Matthew D. Jose, **Towards a better understanding of uraemic molecules**. Research Evening, School of Chemistry, UTAS, 29 April 2010. (poster)

Abstract

Currently, there is no-cure or pre-diagnosis test for patients suffering from Chronic Kidney Disease (CKD). Uremic toxins which normally excreted into the urine by healthy kidneys accumulate in the body due to kidney malfunction causing complications such as higher risk of cardiovascular disease. These complications associated with CKD leading to the high mortality rate among the patients. Although the removal of the uraemic solutes such as urea and creatinine can be achieved by dialysis but it does not affect the survival rate. Therefore, it is suggested that there are some other toxins that cannot be removed by hemodialysis causing the symptoms of CKD. Analysis of serum samples of patients suffering from CKD before and after dialysis compared to samples collected from healthy volunteers can thus give more insight into the effect of uraemic solutes as those remaining after dialysis can lead to the death. In addition, monitoring the changes in metabolites after dialysis can be very helpful to evaluate dialysis membrane efficiency in removal of uraemic toxins.

Analyses of uraemic molecules in serum sample of a group of patients suffering from CKD were performed with three analytical methods including capillary electrophoresis-mass spectrometry (CE-MS), gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS), and the results compared with a healthy (control) group. Both targeted and non-targeted (global profiling) studies were aimed. Unlike GC-MS and LC-MS which proved to be suitable for the analysis of respectively polar and non-polar metabolites, CE-MS practically

failed in providing reproducible results to be compared to the other two techniques.

Targeted study of some of the known uraemic toxins using GC-MS was performed employing the available individual standards following suitable derivatisation procedures. On the other hand, higher numbers of known uraemic molecules were identified in LC-MS upon careful monitoring of MS features such as exact masses of molecular ions and adducts in both positive and negative ion-modes using available data bases such as METLIN library and Human Metabolome Database. Comparison of the results shows that most of the compounds had significant reduction post dialysis, which also reflects the quality of the dialysis treatment. Also unlike water soluble metabolites, the analysis of protein bound solutes was found less conclusive due to the complications associated with them, for example, in finding a suitable sample preparation to be able to efficiently cleave their bounds with proteins.

Global metabolic profiling of the results was also performed employing XCMS platform by visualising the processed data (p-value < 0.05) as distribution profiles. These profiles generally show significant reduction of detected metabolites after dialysis with metabolites had been more efficiently removed from m/z range of 100 to 500. Accordingly, 50% of metabolites were distributed in m/z range of 250-550, 25% in the relatively narrower range of 100-250 and the rest were between m/z of 550 to 950. Also, the majority (90%) of detected metabolites showed relative fold changes less than 70 pre-dialysis which reduced to about 25 after treatment. Moreover, fold change distributions for 50% of metabolites before dialysis was

approximately between 3 and 15 which reduced to less than 5 post dialysis. On the other hand, treatment seems to have insignificant effect on fold changes of about 10% of metabolites. As one application, such information might be beneficial in better understanding of the performance of different dialysis treatments.

Table of contents

Chapter 1

Literature Review

1.1 Introduction

Chronic Kidney Disease (CKD) which is related to renal dysfunction is associated with the dysfunction of multiple organs and a variety of metabolic and enzymatic complications known as Uraemic Syndrome [1]. Due to low kidney function, uraemic solutes that are normally excreted into the urine by healthy kidneys can accumulate in the patient's body and cause many complications. Since these compounds interact negatively with biological functions and result in adverse biological impact in the body, they are known as uraemic solutes or uraemic toxins [2].

Since chronic kidney disease is usually asymptomatic it can be undetected until late stages of the disease [3]. Early diagnosis and early treatments can prevent further damage to the organs. It is believed that uraemic toxins are involved in a number of symptoms that patients with stage-V chronic kidney disease suffer such as: anemia, bone disease, cardiovascular disease, calcification of soft tissue, poor immune response, encephalopathy, and glucose intolerance [4]. Cardiovascular damage is one of the main complications associated with CKD leading to morbidity and mortality in patients [5].

Renal damage is generally detected by proteinuria, albuminuria on urine analysis or measurement of changes in serum creatinine concentration

to calculate the estimated glomerular filtration rate (GFR) or a combination of these. Although these methods can provide good information, they show renal damage at very late stages and they are nonspecific as well. Since biomarkers can usually give us information in the early stages of the disease, metabolomics profiling may have the potential to predict the disease before symptoms occur and can be used as disease assessment techniques [6].

Uraemic toxins identified so far can be classified into three broad categories [4] as summarised in **Table 1.1**. Although toxins like indoxyl sulfate, CMPF, *p*-cresyl sufate, indoleacetic acid and hippuric acid are low MW uraemic toxins in the free state, they can be considered as high MW toxins because they are partially or totally bounded to plasma proteins, mostly albumin. Because of that, their removal with hemodialysis membranes is very difficult [7]. It is also suggested that these group of toxins are responsible for some of the complications associated with CKD like vascular lesions, bone disorders, and the central nervous system (CNS) [7].

It is important to know which of the suspected molecules can cause cardiovascular disease and premature death. Although the level of some of these toxins is reported to be lower in peritoneal dialysis compared to hemodialysis, it does not guarantee that peritoneal dialysis patients are in lower risk of consequences associated with uraemic toxins in the body [8]. As it is currently impossible to cure these diseases at advanced stages or remove all of them, identification of these uraemic toxins offers the best opportunity for early stage diagnosis so that treatment can be used to reduce or even stop progression of the disease.

Table 1.1: Known uraemic toxins.

Free water-soluble, low molecular weight (MW) solutes (MW < 500 Da)	**Protein-bound low-molecular- weight solutes (MW < 500 Da)**	**Middle molecules (12,000 Da > MW > 500 Da)**
asymmetrical dimethylarginine (202), α-keto-δ-guanidinovaleric acid (151), α-N-acetylarginine (216), arab(in)itol (152), argininic acid (175), benzylalcohol (108), β-guanidinobutyric acid (131), β-lipotropin (461), creatine (131), creatinine (113), cytidine (234), dimethylglycine (103), erythritol (122), γ-guanidinobutyric acid (145), guanidine (59), guanidinoacetic acid (117), guanidinosuccinic acid (175), hypoxanthine (136), malondialdehyde (71), mannitol (182), 1-methyladenosine (281), methylguanidine (73), 1-methylguanosine	3-carboxy-4-methyl-5-propyl-2-furanpropionic acid (240), 3-deoxyglucosone (162), fructoselysine (308), glyoxal (58), hippuric acid (179), homocysteine (135), hydroquinone (110), indole-3-3acetic acid (175), indoxyl sulfate (213), kynurenine (208), kynurenic acid (189), melatonin (126), 2-methoxyresorcinol (140), methylglyoxal (72), N-(carboxymethyl)lysine (204), *p*-cresol (108), *p*-cresyl sulfate (188), pentosidine (342), phenol (94), phenyl sulfate (174), phenylacetic acid (136), *p*-hydroxyhippuric acid (195), putrescine	adrenomedullin (5729), atrial natriuretic peptide (3080), cholecystokinin (3866), clara cell protein (CC16) (15800), complement factor D (23750), cystatin C (13300), degranulation inhibiting protein I (angiogenin) (14100), delta-sleep inducing peptide (848), Dinucleoside polyphosphates, *e.g*. diadenosine tetraphosphate (836), β-endorphin (3465), endothelin (4283), hyaluronic acid (25000), interleukin-1β (32000), interleukin-6 (24500), binterleukin-18 (1800), X-Ig light chain (25000),λ-Ig light chain (2500), leptin

(297), 1-methylinosine (282), myoinositol (180), N^2,N^2-dimethylguanosine (311), N^4-acetylcytidine (285), N^6-methyladenosine (281), N-methyl-2-pyridone-5-carboxamide (152), N-methyla-4-pyridone-3-ca rboxamide (152), N^6-threonylcarbamoyladenosine (378), orotic acid (174), orotidine (288), oxalate (90), phenylacetylglutamine (264), pseudouridine (244), symmetrical dimethyl arginine (202), sorbitol (182), taurocyamine (174), threitol (122), thymine (126), uracil (112), urea (60), uric acid (168), uridine (244), xanthine (152), xanthosine (284)	(88), quinolinic acid (167), spermidine (145), spermine (202)	(16000), methionine-enkephalin (555), $\beta 2$-microglobulin (11818), neuropeptide Y (4272), prathyroid hormone (9225), retinol-binding protein (21200), tumor neorosis factor-α (26000)

The identification of these molecules not only helps to evaluate the existing therapeutic treatments but also can make it possible to develop new removal strategies to decrease or normalise the uraemic toxins from plasma [1]. Therefore, further studies are needed for the identification, characterisation, and analytical determination of toxins, causing adverse biological effects with the hope of identifying preventive measures and new therapies for patients suffering from CKD [1].

1.2 Metabolomics

Final products and by-products of the complex biosynthetic and catabolism pathways which are in all live biological systems are called metabolites [9]. Many human diseases impose specific changes in the chemical and biochemical profile of biological fluids and tissues [10]. Direct measurements of metabolites and proteins are therefore necessary steps for the clarification of biological processes in both healthy and diseased states, helping to find the cause of the disease, toxicological progression and identification of biomarkers associated with the disease [10]. The comprehensive study of endogenous low MW metabolites in biological systems is known as metabolomics [11, 12]. On the other hand, tracking the changes in the level of endogenous metabolites in a biological sample, which is resulted from a disease or medical treatment, is usually called metabonomics [13]. More precisely, the quantitative measurement of dynamic multi-parametric response of a living system to pathophysiological stimuli or genetic modification is called metabonomics [9]. The terms metabolomics or metabonomics, are often used interchangeably, stand for global profiling and targeted analysis of endogenous metabolites in a

biological sample [13]. In a simple explanation, metabolomics is a comprehensive and quantitative analysis of the whole metabolome under a given set of condition [14]. Hence, analysis of endogenous low MW metabolites in biological fluids is the main target of metabolomics [11, 15].

There are two approaches that can be used for a metabolomics study: 1) metabolite target analysis and, 2) metabolic profiling. In the first approach, the main focus of the study is on targeted metabolites whose presence and quantification in certain biological samples is considered [16]. Therefore in targeted metabolomics, the quantification of selected metabolites among known compounds is the goal of study [17]. The focus of the second approach is more on the comparison of metabolite profiles of different biological samples in various stages of the disease without a prior knowledge about the nature and the identity of measured metabolites leading to lists of molecular features instead of specific metabolites. Organic acids, amino acids, amines, sugars, steroids, nucleic acids bases and other small molecules, which are intermediate in intracellular metabolism, are the most commonly identified endogenous metabolites in global profiling [9]. Since metabolic profiling provides a wide range of information about the screening of potential diagnostic biomarkers of different diseases; there has been an increasing interest to replace the traditional methods being used in clinical laboratories, which are usually based on the analysis of a few specific metabolites, with profiling approaches [15]. It is worth noting that the consequence of factors like diet, stress, medications, age, exercise, fasting, alcohol consumption, etc. on the metabolic composition of urine and blood should not be underestimated. Thus, minimising possible biological

variations due to these factors before performing sample collection is highly desirable [10].

Although the main purpose of a metabolomics study is to measure the highest possible number of endogenous metabolites in a biological sample preferably in a single analysis, there is no analytical technique capable of giving a comprehensive metabolite profile of a biological sample. In addition, it is impossible to find an extraction method to provide a full recovery for the whole metabolome from tissue or body fluids [15]. As a result of great diversity in the chemical properties of the metabolites and their wide concentration ranges, finding a robust, reproducible and generic method to do metabolic profiling is often problematic and challenging [9, 18].

Each metabolomics study contains these following steps [19]: 1) sample collection and storage 2) sample preparation including metabolite extraction and pre-concentration, etc., 3) analysis of samples, 4) data processing including identification by the aid of metabolomic data bases, and interpretation of the results. These steps are discussed in more detail below.

1.2.1 Sample collection and storage

Sample collection and storage is as important as the other steps in any metabolomics study. Knowing about the history of the samples such as collection procedures, type of containers used, preservatives added to the samples, their stability, processing, storage conditions, and shipment protocols are factors affecting quality of the samples and validity of the

results [10]. Blood, for example, should be collected in clean tubes and allowed to clot or centrifuged in order to separate serum from it. To collect plasma, anticoagulants are added to blood to prevent clotting and cellular components are removed by ultracentrifugation. Clotting time, temperature at collection time, type of vial, anti-clotting agents, centrifugation time and speed, shipping and storage are all important parameters that should be taken into account.

Stability and composition of the sample is greatly affected by storage and the process of freezing and thawing of the samples. Avoiding repeated freezing and drying is recommended to avoid errors in the study and prevent metabolite losses. In addition, to minimise the deterioration of the samples, thawing procedure of frozen samples should be performed on ice, in a cold water bath or at room temperature, not by heating [10]. Serum and plasma samples should be stored in the freezer at -80 °C until analysis time. Aliquoting of samples should be done prior to freezing and storage.

1.2.2 Sample preparation

Limitations such as matrix effects compromise selectivity and decrease detection sensitivity of the analyte of interest during the analysis of a biological sample, even in the most selective and sensitive analytical methods [20], necessitating sample pre-treatment. A good and efficient sample preparation reduces the effect of matrix with selective isolation of analytes of interest from the matrix without losing them. In addition to minimising or elimination of matrix effect, it can also pre-concentrate the sample. Extraction procedures (liquid-liquid, SPE, filtration, etc.) should be

simple yet be able to extract a wide range of metabolites with the highest possible recovery [10].

Nevertheless, sample preparation is the most labour intensive and error-prone step accounting for about 80% of the whole bio-analysis time [20]. Factors such as the type of sample being analysed, method of analysis and whether the study is targeted or profiling all affect the sample preparation procedure chosen for the study [10, 15]. Even the content of the sample matrix is influential; for example, serum and plasma contain proteins, glycoproteins, and lipoproteins, whereas urine has a high concentration of salt and urea; and proteins and salt don't exist in plant extracts [10]. For serum and plasma samples a protein precipitation step is usually applied by organic solvent, acid, and heat. Urine sample preparation is often limited to centrifuging the samples at, for example, 3000 rpm to remove small particles followed by metabolite extraction with solid-phase extraction (SPE) or liquid-liquid extraction, depending on the compound of interest. It is also suggested to do urease treatment prior to mass spectrometry (MS) analysis to reduce the risk of such phenomena as column overloading, peak distortions, matrix effects, and ion suppression. However, the cost of urea elimination by urease is sacrificing some other metabolites such as acotinic acid, hypoxanthine, and the tricarboxylic acid intermediates [10]. Generally speaking, the sample preparation varies from one sample to another and follows the same order of complexity as the biological sample, that is tissue > whole blood > plasma/serum > urine > saliva [20].

Minimal sample pre-treatment is suggested in non-targeted metabolomics study to avoid the loss of metabolites. In the case of analysing

low MW metabolites, it is necessary to separate them from large molecules (like proteins, lipids, and large peptides) and salts to perform proper analysis [21]. In samples like plasma and serum, deproteinisation with an organic solvent is an essential step to prevent further complications during analysis like protein adsorption to the capillary wall in capillary electrophoresis [22]. Denaturation of proteins is generally performed by using an external shock (like strong acid/base, heat or using an organic solvent) [23]. In the most common and simplest protein precipitation method in bio-analysis, three parts of organic solvent is added to one part of sample followed by vortexing and centrifugation. Cold methanol, for example, provided the highest recovery for polar metabolites [21]. Centrifugation can separate supernatant from proteins by making a protein pellet [20]. Since denaturation causes an active change in protein structure, small molecules like metabolites, drugs, and biomarkers usually cleave from proteins and become freely soluble in supernatant and ready for quantitative analysis [20].

For targeted metabolomics studies, there is usually another step after deproteinisation like off-line SPE which not only can desalt the sample but also is capable of eliminating interfering matrix from biological samples [15, 20]. However, due to the highly polar nature of most metabolites, they don't show retention in most RP SPE cartridges and are removed with the salt [15]. A limited number of SPE columns like Oasis HLB from Waters have the capability to extract acidic, basic, and neutral species [24]. The potential of using SPE columns was investigated in selective preconcentration of nucleosides from human urine and nucleoside fingerprinting by capillary electrophoresis [15]. However, there are some studies suggesting that SPE itself can contribute to ion suppression during LC-MS [20]. Due to the use of

organic solvents in the final elution step of SPE, chemical impurities in the cartridge like polyethylene glycols (PEGs) and phthalates can be extracted and therefore interfere with the analysis [20]. Although there are some aqueous-organic solvents with certain pH and ionic strength recommended by the manufacturers to reduce the elution of these impurities, these conditions still favour such impurities causing matrix effects [20].

Another approach for metabolite extraction is solid-phase micro-extraction (SPME) which is based on the extraction of the metabolites into a SPME fibre [26]. Unlike SPE, which is an exhaustive extraction technique, SPME is an equilibrium extraction technique. In a typical SPME protocol, there are two steps: extraction and desorption [20]. The extraction of analytes on the fibre is dependent on several parameters such as pH, salt concentration, sample volume, agitation speed, extraction temperature, and extraction time; and all these parameters should be optimised [20]. There are two types of extraction: direct immersion (DI) into liquid sample and headspace (HS) in which the liquid sample matrix is heated to facilitate the release of volatile compounds into the gas phase until volatile compounds in the sample reach equilibrium [20]. The distribution of the analytes between the two phases at equilibrium is expressed by a thermodynamic equilibrium constant, the partition coefficient (K). This value depends on the solubility of the compounds in the liquid phase, compounds with high solubility have high concentration in the liquid phase; compared to the gas phase resulting in high K values [25]. However, there are some ways to lower the K value (and increase the volatility) like increasing the temperature of the vial, using agitation, or adding inorganic salt to the sample matrix. By performing the latter, the solubility of polar organic compounds in the sample matrix can be

reduced leading to their greater transfer into the headspace by decreasing the *K* values. The effect of salting out on *K* value is different for different compounds [26]. Post volatilisation, analytes are adsorbed to the fibre exposed in the headspace [26]. The success of SPME in extraction is related to physiochemical properties and thickness of the fibre coating. The ultimate step, i.e., sample desorption is by introducing the fibre into the GC injector where the fibre is thermally heated and releases the analytes [20]. SPME can also be used with LC. In this case, desorption of analytes enriched on the fibre is performed via extraction into a solvent.

1.2.3 Analytical techniques

Due to the great differences in physiochemical properties and acid/base characteristics of metabolites, various analytical techniques can be used to separate them prior to mass spectrometric detection [9]. In order to provide reliable results for a metabolomics study an ideal analytical method should have some characteristics [27]:

1) It should cover a wide range in different concentrations.
2) The minimum sample preparation is desired as it prevents losing analytes and altering their compositions so metabolites can be measured in their native states. It reduces the analysis time as well.
3) Reproducibility is also one of the most important factors in analysis which makes the method more reliable and with less experimental variations.
4) It is preferable for a method to be used that is capable of both qualification and quantification.

5) The method should be affordable and easy to handle.

Since there is no single analytical tool that meets all these requirements, a combination of analytical tools and methods is usually used to increase the amount of information collected from a biological sample as each method can cover different types of metabolites [9]. There is a variety of analytical instrumentations that can be used for the analysis of metabolites like gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), nuclear magnetic resonance (NMR), and capillary electrophoresis-mass spectrometry (CE-MS) and each of them has its own advantages and disadvantages [16]. The main requirement to analyse a biological sample is to differentiate the targeted analytes from the interfering compounds and it can be achieved based on chemical shift in a nuclear magnetic resonance (NMR) spectrum, mass to charge ratio on a mass spectrometer (MS), retention time in chromatography or combination of them [28]. NMR spectroscopy has also been used for the metabolic profiling of body fluids, such as serum and urine [29]. Although it is rapid, non-destructive and requires minimal sample preparation, the sensitivity of NMR is limited and only the most abundant species can be detected. Sample amounts of several micrograms are often required [15, 30].

The recent progress in MS provides researchers with a variety of choices for chromatographic separation, ionisation, and mass spectrometric analysis [17]. MS can provide information about the structure of unknown metabolites and it has enough sensitivity for quantification of the metabolites as well [27]. The mass detection of the molecules by MS cannot be achieved unless the analytes in the samples are ionised. Factors that affect the

ionisation efficiency, apart from the chemical properties of the analyte itself, are eluent flow and composition, sample matrix and ionisation type and parameters. Since ion suppression from the sample matrix can cause major problems, especially in semi-quantitative measurements, the use of ionisation enhancers, suitable sample preparation methods, and various ionisation sources can significantly enhance ionisation efficiency [31].

Different ionisation techniques can be used in mass spectrometers hyphenated with chromatographic instruments, like electron impact (EI), electrospray ionisation (ESI), atmospheric pressure chemical ionisation (APCI), matrix-assisted laser desorption-ionisation (MALDI), desorption electrospray-ionisation (DESI), and atmospheric pressure photo-ionisation (APPI) [19]. Based on the ion acceleration and detection technique there are different varieties of mass analyser and configurations, including quadruple-MS (Q-MS), triple quadruple-MS (Triple-Q-MS), quadruple-ion trap-MS (Q-Ion trap-MS),Time-of-flight-MS (TOF-MS), Fourier transform (FT)-ion cyclotron resonance MS (FT-ICR-MS), and, FT-Orbitrap-MS [19]. Each of these ionisation techniques and mass analysers offers its own advantages and disadvantages.

The performance of a mass spectrometer can be determined by some parameters like mass resolving power (resolution), mass accuracy, linear dynamic range, and sensitivity [19, 32]. Single quadruple is the most commonly used mass analyser; however its resolution is 4-times lower than that of TOF-MS. On the other hand, Fourier transform (FT)-ion cyclotron (ICR)-MS has the resolution of over 1,000,000 but it is very expensive equipment and costly for maintenance. A mass spectrometer with high mass

accuracy is capable of differentiating between very close m/z signals which can significantly affect the quality and quantity of interpreting the mass results [19, 33]. Due to the high sensitivity and high resolving power (10,000) of hybrid TOF-MS instruments (like Q-TOF-MS), they are becoming very popular in the field of metabolomics. However, they also have some limitations such as in the operating linear range, which is limited by the properties of the time-to-digital converter detector, being only capable of recording one ion per dead time. Hence, very intense mass signals become overloaded which badly affects real intensity and causes distortions on peak shape. These issues can result in deviations in the mass accuracy especially for low *m*/*z* signals. However, this can be partly tackled by using an on-line lock mass spray which acts as an internal standard and corrects these deviations in the *m*/*z* axis [19, 34].

Among all the mass spectrometers, FT-ICR-MS and FT-Orbitrap-MS have the ability to measure a very high mass accuracy in a very wide dynamic range. Moreover, FT-ICR-MS has the highest resolution that is currently accessible (> 1,000,000) compared to all other mass spectrometers and mass accuracy around 1ppm [35]. Although the recently developed FT-Orbitrap-MS has lower performance and mass accuracy (~2ppm, with internal standard) in comparison to FT-ICR-MS, due to the high speed of ion transmission, as a result of shorter accumulation times, coupling this instrument to chromatographic techniques like LC is much more beneficial [19, 36]. The opportunity of gaining high mass accuracy with a very wide dynamic range can improve the capability of identification of metabolites in very complex biological samples. So, as far as coupling LC with MS is

concerned for metabolomics study, LC-FT-Orbitrap-MS can be considered as a golden technique.

Hyphenation of MS with separation techniques, like LC, GC, or CE is very common in metabolomics studies. However, GC-MS and LC-MS are more popular since they are more convenient to operate than CE-MS and can cover a wider range of compounds [19]. GC-MS is suitable for the analysis of stable, volatile metabolites, and derivatives such as fatty acids, steroids, flavonoids, etc. On the other hand, LC-MS is commonly used for non-volatile analytes such as peptides, lipids, nucleotides, ionic species, etc. [10]. Using each of these techniques has some advantages and disadvantages, so for a comprehensive study one can use them together to take advantage of their orthogonality. **Figure 1.1** schematically shows a general flow-chart for the analysis of metabolites.

1.2.3.1 Capillary electrophoresis-mass spectrometry (CE-MS)

Capillary electrophoresis is a separation technique capable of analysing highly polar and charged compounds. As most of the metabolites are ionic and highly polar compounds CE is now recognised as one of the best options for metabolomics. Since the separation mechanism is different than LC, CE can provide complementary details about the content of the biological sample [16].

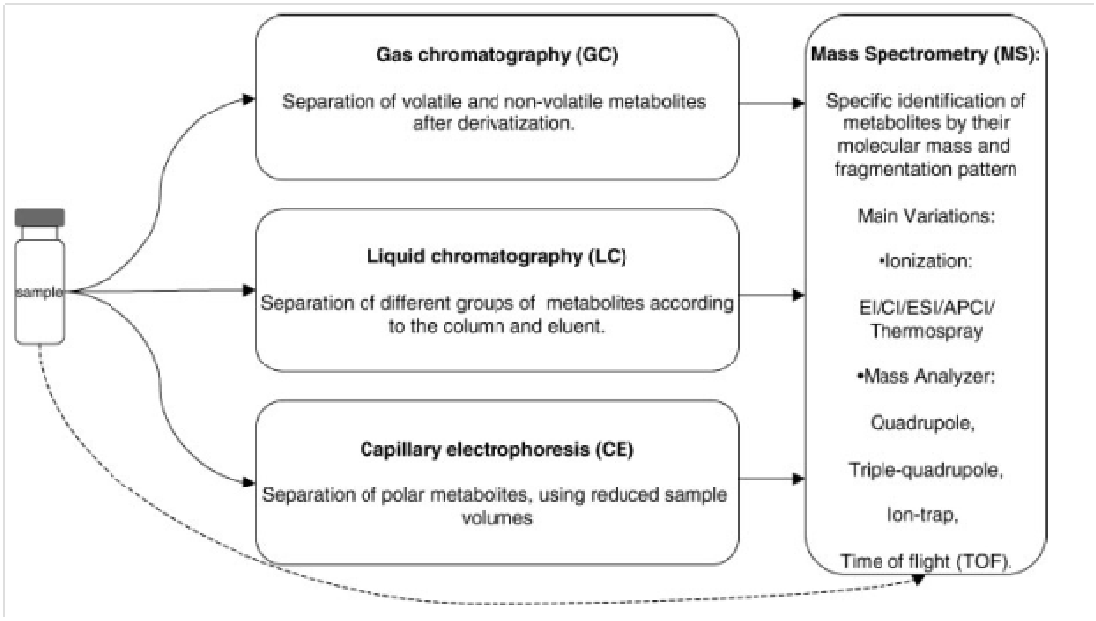

Figure 1.1: Different analytical options for the analysis of metabolites. The sample can be directly analysed with MS or it can be analysed by coupling any of chromatographic techniques to MS [28].

Like any other analytical technique, CE has its own advantages and disadvantages. The amount of sample needed for CE is in the nanoliter range so it is the best option for volume limited samples. Other advantages of using CE are the low running cost, and small amount of reagents needed. Also, separation is performed using fused-silica capillaries instead of expensive LC or GC columns.

On the other hand, poor concentration sensitivity is one of the drawbacks for CE which is due to the short optical path length when UV absorbance detection is employed. While this issue can be tackled when CE is hyphenated to a mass spectrometer, new complications will arise as a result of this hyphenation. For example, one inherent limitation of CE-MS is that the small amount of sample that can be injected into the capillary due to

the limited volume of the capillary, compromises the detection sensitivity in MS [15].

There are variety of CE modes used to separate compounds including, capillary zone electrophoresis (CZE), capillary isoelectric focusing (cIEF), micellar electrokinetic capillary chromatography (MEKC), and capillary gel electrophoresis (CGE) [37]. CZE is the most common CE technique used in CE-MS metabolomics [38], which has highest compatibility with MS detection due to the possibility of using volatile BGEs [15]. On the other hand, although MEKC offers better selectivity and better quality of separation compared to CZE, hyphenating with MS is more difficult and reduces the sensitivity [15]. In addition, due to the presence of polymers and surfactants in MEKC, its connection to the MS can be problematic [37] . Nevertheless, CZE with polyacrylamide coated capillary and reversed polarity was recently employed for anionic metabolites, and MEKC was used for fingerprinting of cationic and neutral metabolites in urine samples [39]. Capillary electrochromatography (CEC) is another CE technique capable of measuring more metabolites and high resolution than other CE modes. It also shows a good potential for coupling with MS for the analysis of proteins and peptides in plasma samples from patients suffering from gangrenous and phlegmonous appendicitis [40].

Capillary electrophoresis can be coupled to ESI-MS via a sheath-liquid interface or (more recently) a sheath-less interface [41]. So far, the sheath-liquid interface is the most commonly used in metabolomics study. The coupling of CE to MS is very challenging since it should be built in a way that CE separation column and spray tip to be in a single continuous unit to

prevent any dead volume minimising the efficiency of the separation [41]. Besides, this unit should preserve the electric circuits between the CE system and the spray tip [41]. To do so, a stainless steel capillary (ESI needle) is located around CE capillary which delivers sheath liquid [41]. Although hyphenation of MS with CE can improve the sensitivity (in comparison to UV detection), the sheath liquid interface dilutes the sample and reduces the sensitivity. This can be improved by employing some online or offline pre-concentration methods with loading more samples into the column [42].

To avoid dilution of sample with sheath liquid, CE can be coupled to MS via a more recent configuration called sheathless interface in which the outlet end of the capillary is covered with a metal coating and then a voltage is applied to the CE buffer at the capillary outlet which provides the necessary ESI potential [41]. The suitability of a sheathless interface for the metabolic profiling of urine was investigated by Ramautar and co-workers [43]. According to their results, 900 molecular features were detected in urine using a sheathless interface compared to only 300 detected compounds with sheath liquid interface, which suggests significant improvements in sensitivity can be gained by employing sheathless interface.

Although ESI is very well suited for the analysis of polar compounds, it has some limitations from the low tolerance to salts (CE buffers) and susceptibility to matrix effect [38]. As a matter of fact, ESI is compatible with volatile solvents, which often cannot compete in performance to standard CE buffers for the quality of separation [44]. The reason that standard buffers like phosphate, borate, and sulfate are not suitable for ESI is because they are non-volatile and form ion pairs in solution. Using formic acid and

ammonium acetate, due to the high volatility of them, is very common in CE-MS [41]. It is recommended to work from neutral pH to pH 2 in positive ion mode and neutral pH to pH 10 at negative ion mode [45]. Moreover, the concentration of these BGEs should be low to prevent minimising ion production [44]. Higher concentrations of formic acid as the BGE can improve the separation, but impose a high current, which is also an issue leading to Joule heating [46]. Alternatively, organic modifiers like methanol can be added to high concentration BGEs in order to improve the separation. The potential of this modification on separation efficiency has been shown in the CE-MS analysis of amino acids, in which separation of the isomers leucine and isoleucine was performed by applying 2M formic acid (pH 1.8) containing 20% methanol [46].

There are different mass analysers being used for CE-MS metabolomics. Triple quadruple (TQ) and ion trap (IT) are the most common analysers for analysis of low-molecular-weight metabolites [47]. Although they both provide good sensitivity, especially TQ, their scanning process is lower than the speed of CE separation, making them unable to get data points across very narrow CE peaks. In this regard, TOF analysers are suitable alternatives being widely used in non-targeted approaches. Moreover, TOF can provide more resolution and higher scanning speed and accuracy. Despite these advantages and due to the high sensitivity of the instrument, interferences from solvent ions, contaminants in the lab and compounds, etc. with the same masses as the metabolites can be problematic in the analysis [48]. This problem can be tackled by employing FT-ICR analyser which can give very accurate mass measurements with minimal errors. Despite this high accuracy, it has a disadvantage of low speed of mass

measurements which again can compromise the analysis of narrow peaks [15, 48]. Targeted and non targeted metabolomics studies have been demonstrated on different biological samples like, blood, urine, serum, plasma, saliva, sweat, tissue, bacteria.

1.2.3.1.1 Recent Developments in CE-MS of metabolites

No discussion on CE-MS analysis of metabolites is complete without mentioning the significant contributions of Soga and co-workers. They developed several methods capable of quantitative measurement of metabolites in different matrices such as plasma, serum, bacteria, tissue, plant extraction [47]. **Table 1.2** summarises one of the most frequently used methods developed by this group.

Soga and co-workers used CE-MS for the large scale metabolite analysis of *Bacillus subtilis* to monitor the changes during sporulation. A total of 150 metabolites have been identified from 1692 measured metabolites. Results suggest that many metabolites in the glycolytic, pentose phosphate, and tricarboxylic acid pathway were reduced in the early stage of sporulation [49]. Similarly, pH of the BGE was modified to 7.5 by substitution of 1M formic acid with 50mM ammonium acetate solution for the analysis of nucleotides in *E.coli* [47]. Interestingly, phosphate buffer was used to mask the silanol groups of the fused-silica capillary before the start of separation to prevent the interaction of nucleotides with the capillary wall.

Table 1.2: Typical procedures for CE-MS analysis of metabolites [47].

BGE	1M Formic acid	50mM ammonium acetate (pH=8.5)
Injection time	50 mbar, 3 sec (3nL)	50 mbar, 30 sec (30 nL)
Temperature	20 °C	20 °C
Capillary length	100 cm fused silica	100 cm cationic polymer coated
Voltage	+30 kV	-30 kV
Polarity	Positive	Negative
Sheath Liquid	5mM ammonium acetate in 50% methanol and water	5mM ammonium acetate in 50% methanol and water
Capillary Voltage	4000 V	3500 V
Sheath Flow	10 μL/min	10μL/min
Dry gas temp.	300 °C	300 °C

The usefulness of this procedure was demonstrated by the analysis of phosphorylated species in *E.coli* wild type providing valuable information about the enzyme activity. In another study, the same method has been applied for the metabolite profiling of colon and stomach cancer. Metabolites were extracted from tissue by using methanol and chloroform followed by ultrafilteration with 5-kDa cut-off filters. This resulted in quantification of 94 metabolites in colon and 95 metabolites in stomach contributing in glycolysis, the pentose phoshphate pathways, the tricarboxylic acid cycle (TCA) and urea cycles, and amino acid and nucleotide metabolism [50].

These findings enabled them to identify some cancer specific metabolite traits.

Saliva in another biological sample which is readily accessible and informative and is ideal for early detection of a wide range of diseases [51]. Except for the sheath liquid which is methanol/water (50% v/v) and 0.5 μM reserpine, the rest of the method is the same as that described in **Table 1.2**. Saliva samples of individuals with oral, pancreatic and breast cancer were analysed and the results compared to a control group revealing 57 metabolites that can possibly contribute to the disease. The concentration of metabolites in patient samples was significantly higher compared to the control group demonstrating the correlation between patient's characteristics and the quantified metabolites [51].

The main metabolites in rice leaves was also analysed using the same method (**Table 1.2**). Metabolites were extracted by using methanol and chloroform followed by ultrafilteration with cut-off filters. This study resulted in a successful measurement of 88 main metabolites involved in glycolysis, the TCA, the pentose phosphate pathway, photorespiration, and amino acid biosynthesis [52].

In 2010 Toyohara and his colleagues employed CE-TOF-MS for metabolomics profiling of uraemic solutes in CKD patients. They employed Soga's method (**Table 1.2**) for the analysis of cations and anions. Sample preparation was also done with methanol and chloroform extraction with 5-kDa centrifugal filters and then reconstitution in water. They successfully found 22 cations and 30 anions, with their accumulation suggested as an

indication of decrease in estimated glomerular filtration rate (eGFR). 9 cations and 27 anions were successfully identified and also 7 cations (2 new) and 5 anions (all new) were found to contribute in eGFR reduction [53].

CE-MS analyses of metabolites using coated capillaries have been investigated by other groups. The usefulness of non-covalently coated capillaries for the metabolic profiling of urine was demonstrated by Ramautar and co-workers [54]. Under low BGE pH, polybrene-dextran sulfate-polybrene (PB-DS-PB) coating was employed for the analysis of cationic compounds, whereas polybrene- poly (vinyl sulfonate) (PB-PVS) was used in the case of anionic compounds. Approximately 600 molecular features were measured in rat urine by employing PB-DS-PB, while half of this amount was detected with PB-PVS capillary. In general, the response for the analysis of anionic compounds by CE-MS was found to be much lower than that of cationic compounds. The possible explanation could be lower ionisation efficiency and ion suppression caused by the BGE. Nevertheless, in terms of throughput analysis, the PB-PVS/CE-MS showed superiority over PB-DS-PB/CE-MS by providing higher number of metabolites, yet at the expense of longer analysis time.

Ramautar *et al.* also used non-covalently coated capillaries with a bilayer of PB and PVS for the analysis of amino acids in cerebrospinal fluid (CSF) [10]. The coating provides a considerable EOF at low pH employing 1M formic acid as the BGE which facilitated the fast separation of amino acids. The retention time repeatability was related to the stable and very consistent PB-PVS coating. A small plug of 12.5% ammonium hydroxide was injected before injecting the sample to perform pH mediated stacking which

increases the loading capacity of the capillary. This method has been applied for metabolic profiling of CSF samples collected from a patient suffering from complex regional pain syndrome in comparison to a healthy sample [54, 55].

Coated capillaries were also employed for a CE-TOF-MS separation and detection of tryptophan metabolites of the kynurenic pathway. To deactivate the fused silica wall and to generate a stable reversed EOF for a reproducible migration time, the capillary inner was modified with 1-(4-iodobutyl) 4-aza-1-azoniabicyclo [2, 2, 2] octane iodide, also called M7C4I. This coating enabled analysis of tryptophan, kynurenine, kynurenic acid, and metabolites spiked into 5-fold diluted CSF in less than 5 min [56].

1.2.3.2 Gas chromatography-mass spectrometry (GC-MS)

Coupling GC to MS has resulted in a very useful analytical technique mainly due to the high sensitivity, peak resolution, and reproducibility involved. In addition, access to GC-MS electron impact (EI) libraries makes identification of diagnostic biomarkers possible [57]. The only requirement for a compound to be analysed by GC-MS is for it to be volatile [9]. Due to the polar and non-volatile nature of most of metabolites, derivatisation is a necessary step to increase their volatility, make them thermally stable and to decrease their polarity [9]. Derivatisation requires the compound to have necessary functional groups including, amino-, organic-, and aromatic-acids, amines, sugars up to trisaccharides, alcohols, polyols and some mono-phosphorylated functionalities [58]. However, extensive sample preparation and long analysis time makes GC-MS a relatively low throughput technique for metabolomics studies. This might be circumvented by using

comprehensive two dimensional GC (GC×GC) with TOF-MS which enhances the separation efficiency, as a result of increasing peak capacity, and detection sensitivity [9].

Although MS can provide us with more details about the compounds in the sample, there are other detectors which can be used with GC such as flame ionisation detector (FID), thermionic ionisation (TID), chemiluminescence detector (CLD), and electron capture detector (ECD). Among them FID has been repeatedly used for targeted analysis and metabolic profiling, as a general detector for the carbon containing compounds [10]. TID, CLD, and ECD are less suitable for metabolomics because of the limitations in the detection of some compounds, but they can be successfully used for the detection of targeted metabolites only for certain compounds of interest where detection of them with these detectors is possible. Regardless of their ability to give us qualitative and quantitative data, these detectors are unable to identify the chromatographic peaks with the confidence level of MS detection [10].

Although GC-MS is one of the commonly used techniques in metabolomics investigation, it still has some issues such as variations associated with the derivatisation step or the complexity of biological samples. In addition, interpreting the data collected from the GC-MS and using a suitable library database is one of the critical steps in this method.

A typical GC-MS metabolomics study contains four steps [59]:

1) The first step is metabolite extraction through which metabolites are isolated from the rest of a sample [60]. Protein precipitation (PP),

liquid-liquid extraction (LLE), solid phase extraction (SPE) and ultrasonic-assisted extraction (UAE) are some of the extraction methods commonly used for metabolites. Protein precipitation can simply be achieved by the addition of organic solvent to the sample followed by centrifugation and supernatant removal. Metabolite extraction can improve sensitivity as it pre-concentrates the sample [20].

2) As most of the metabolites are naturally polar and non volatile, making them volatile while thermally stable is also another crucial step to be performed before GC-MS analysis [59, 60].
3) Analysis: separation of the metabolites is the goal of this step.
4) Identification and quantification are performed through de-convolution of mass spectra followed by identification of the metabolites based on retention time and mass spectra libraries.

Although low volatility might be related to the size of the molecule, sometimes smaller molecules might have low volatility as a result of very strong intermolecular interactions among polar groups. In this case, an increase in volatility can be achieved by masking the polar groups with derivatisation reagents. Derivatisation can also be used to decrease volatility to allow analysis of low molecular weight compounds in order to prevent losses in manipulation and to separate these from the solvent peak. In addition to increasing the sample volatility, derivatisation has some more advantages including increasing stability, improving sensitivity, enhancing detectability, and minimising the irreversible interaction between the sample and GC column (polar analytes have a tendency to adsorb to the active surfaces of the columns) as mentioned below [61, 62]. Derivatisation also

helps better accentuation of the differences between sample components to improve chromatographic separation. Silylation, acylation and alkylation are the main types of derivatisation used particularly in metabolomics profiling of biological samples.

Silylation is the most prevalent method which alters the active hydrogen with a trimethylsilyl (TMS) group driven by the fact that the better leaving group has the better siliylation efficiency. Pyridine is usually drives the reaction forward (as it is an acid scavenger) and is the most common solvent for silylation, although it can cause some peak tailing. The reactivity order of functional groups toward silylation is: Alcohol > Phenol > Carboxyl > Amine > Amide.

The use of acylation is limited to addition of fluorinated groups. Highly polar and multifunctional compounds such as carbohydrates and amino acids are targets of acylation reagents. Acyl derivatives are formed with acyl anhydrides, acyl halides and activated acyl amide reagents. Fluorinated acyl groups, going from trifluoroacetyl to heptafluorobutyryl, can be used to increase retention times. Also, acyl derivatives tend to direct the fragmentation patterns of compounds in MS, and so provide useful information on the structure of analytes.

Alkylation is commonly employed to protect certain active hydrogens or as a first step for further derivatisation. During the alkylation, active hydrogens are replaced by alkyl group resulting in molecular polarity reduction.

Methoxamine in pyridine can also be used for the oxamine reaction before silylation in order to stabilise carbonyl groups in metabolites therefore minimises the conversion reactions during silylation. By using a TMS reagents, usually BSTFA (bis(trimethylsilyl)trifluoroacetamide) or MSTFA (N-methyl-N-(trimethylsilyl)trifluoroacetamide), functional groups are altered to TSM ethers, TMS esters, TMS sulfides or TMS amines. Although TMS derivatisation is very efficient, it can cause column damage, formation of unwanted derivatives and multiple derivatives, signal overlapping, and being very labour intense and time consuming [59, 62].

1.2.3.2.1 Recent developments in GC-MS metabolomics

Nishiumi and co-workers developed a derivatisation based GC-MS method to analyse the metabolites in different stages of the pancreatic cancer [63]. In this study, 60 metabolites have been detected in serum and levels of 18 of the metabolites were found significantly higher in patients in comparison to a healthy group (p-value < 0.05). They could also successfully differentiate the different stages of the pancreatic cancer, stage III, stage IVa, and stage IVb by the aid of partial least squares discriminant analysis (PLS-DA) assessment.

Ikeda and colleagues used a derivatisation based GC-MS method for the metabolic profiling of gastrointestinal cancer [64]. Serum samples were collected from 15 esophageal, 11 gastric, and 12 colorectal cancer patients compared to 12 healthy samples. After processing the data, 58 metabolites were identified and the level of 9, 5, and 12 metabolites (p-value < 0.05) were found to dramatically change in esophageal, gastric and colorectal cancer

patients, respectively. According to their results, changes in the levels of malonic acid and L-serine is related to esophageal cancer, variations in the level of 3-hdroxypropionic acid and pyruvic acid is contributed to the gastric cancer, and finally colorectal cancer patients can be characterized by changes in the level of L-alanine, glucuronic lactone and L-glutamine. This study shows the potential of such a metabolomics study as an early diagnostic test [64].

In a study published by Welthagen *et al*. two-dimentional GC×GC connected to MS was shown to increase the number of identified metabolites by an increase in peak capacity. Using two different columns with different polarities in GC×GC-MS can increase the number of peaks identified to 1200, whereas only 500 compounds were detected with conventional GC-MS [65].

Tao and co-workers developed a GC-MS method for the analysis of endogenous metabolites in serum samples of uraemic patients [66]. Derivatisation was done by ethyl chloroformate which has some benefits over TMS. The disadvantage of this method is that derivatives should be extracted with hexane before GC-MS analysis. The metabolic profiling of the samples showed significant differences between patients and healthy. Moreover, levels of valine, leucine, and isoleucine were much lower in patients than control group, whereas levels of myristic acid and linoleic acid were higher in uraemic patient compared to healthy.

SPME combined with GC-MS has been used for the determination of volatile metabolites in human colon cancer cell line by Zimmermann and colleagues [67]. Compounds like undecan-2-ol and pentadecan-2-one were

introduced for the first time in this study to associate with human metabolism. Also this study resulted in some findings that indicated significant differences between the metabolism of colon cancer cells and healthy volunteers. Moreover, comparison of different stages of the disease SW 480 (grade IV, Duke B) and SW 1116 (grade II, Duke A) showed a change in metabolites related to the progress of disease.

1.2.3.3 Liquid chromatography-mass spectrometry (LC-MS)

Unlike GC-MS which usually restricted to the hard ionisation method, electron impact (EI) ionisation, LC-MS uses soft ionisation techniques such as atmospheric pressure ionisation (API), electrospray ionisation (ESI), and atmospheric pressure chemical ionisation (APCI). Liquid chromatography has more versatility compared to other chromatography techniques as it facilitates separation of compounds over a wide range of polarity with less sample preparation step compared to GC-MS [19]. Although sample preparation is much simpler than GC-MS, samples still need to undergo some pre-treatments to be analysed by LC-MS. For example, since most of the salts precipitate in organic solvent, sample desalting might be required before injection [68].

LC coupled to MS can provide very useful information about the structure and quantity of metabolites without the need for derivatisation [4]. Due to advantages such as high throughput in analysis time and the amount of sample needed for injection, recently developed monolithic-based capillary LC is gaining a lot of attention in metabolomics studies [27]. In addition, recently developed UHPLC (ultra high pressure liquid chromatography) systems can not only enhance the speed of separation but

also can provide better resolution compared to conventional HPLC [19]. Nevertheless, because many of the metabolites in body fluids are very polar and ionic, separation of them with RPLC-MS is challenging [69]. Recently, there has been success with some new RPLC stationary phases such as Atlantis dC18 and Alltima HP C18, which show an increase in the retention of polar compounds [70], [15]. One way to increase the retention of ionic compounds in RPLC is to employ ion-pair agents which was applied for the quantitative analysis of polar metabolites, like nucleotides, coenzyme A esters, sugar nucleotides, and sugar phosphates in bacterial extracts [71]. Although using ion-pair agents in RPLC can be a useful tool for metabolomics study, the use of ion-pair agents can cause ionisation suppression in mass spectrometry [71].

Another approach to separate highly polar compounds is hydrophilic interaction chromatography (HILIC). Because of the high content of organic solvent used in this technique (typically > 70% ACN), it is highly suited for coupling to MS. RP columns are capable of analysing semi-polar compounds such as phenolic acids, flavonoids, glycosylated steroids, alkaloids, and other glycosylated species, whereas HILIC columns can separate polar compounds such as sugars, amino sugars, amino acids, vitamins, carboxylic acids and nucleotides. In a plant metabolomics study the efficiency of two-micro bore HILIC columns, Polyhydroxyethyl A and TSK Gel Amide 80 were compared to normal phase LC (NPLC) silica columns. HILIC columns showed better selectivity and separation efficiency compared to NPLC facilitating the detection of positively and negatively charged metabolites [72].

1.2.3.3.1 Recent developments in LC-MS metabolomics

Metabolomics analysis of uraemic toxins in the serum sample of rats with chronic renal failure (CRF) was carried out using RPLC-ESI-MS/MS [73]. Comparison of the results with normal rats showed a significant increase in indoxyl sulfate, phenyl sulfate, hippuric acid and *p*-cresyl sulfate in the samples of CRF rats.

HPLC-q-TOF MS followed by multivariate data analysis was used for the analysis of cocoa phytochemicals in human urine in order to investigate the health benefits of cocoa consumption [74]. Samples were collected from volunteers before and after random intake of either cocoa with milk or water, or milk without cocoa. Results showed the effect of cocoa powder on urinary metabolism 24 h after cocoa consumption. In addition, no difference was found between cocoa intake with milk or water. These findings resulted in the identification of 27 metabolites related to cocoa phytochemicals such as alkaloid derivatives, polyphenol metabolites and processing-derived products like diketopiperazines, all contributing to the urinary modifications.

Kristensen and co-workers investigated the effect of fresh apple or apple-pectin on the urinary metabolome of rats with UHPLC-QTOF-MS in both positive and negative ionisation modes [75]. The male rats were grouped in three, and two of the groups were fed a standard diet with either 7% apple-pectin or 10 g raw apple for 24 days. The urine samples then analysed in order to find the exposure end effect markers of fruit and fruit fibre. Based on the results, quinic acid, *m*-coumaric acid and (-)epicatechin

were identified as exposure markers of apple intake whereas hippuric acid behaved as an effect marker. Pyrrole-2-carboxylic acid and 2-furoylglycine behaved as pectin exposure markers while 2-piperidinone was recognized as a potential pectin effect marker.

LC-MS was also employed for the analysis of metabolites in one of the epidemic tropical disease called *Onchocerciasis* [76]. Serum and plasma samples were collected from 73 African affected patients. The analysis of these serum and plasma samples revealed 14 biomarkers which can discriminate between *Onchocercavolvulus* positive and negative samples. This approach can be used as a diagnostic tool not only for early detection of *Onchocerciasis* but also for the detection of other tropical disease.

The effect of cannabis consumption on schizophrenia by affecting endocannabinoid system was investigated by analysing serum samples of 74 patients with LC-MS [77]. Based on previous studies, it is known that primary fatty acid amides are elevated in drug induced schizophrenia and affective disorder. Their study revealed the role of the endocannabinoid system in the pathology of schizophrenia.

The urine samples from mice with Crohn's disease and wild-type mice were collected and analysed with LC-MS to find biomarkers responsible for inflammation of colon [78]. The analysis resulted in the identification of xanthurenic acid and glucuronides of xanthurenic acid and α-CEHC (2,5,7,8-tetramethyl-2-(2′-carboxyethyl)-6-hydroxychroman). The study supports the correlation between the degree of inflammation and concentrations of xanthurenic acid and α-CEHC glucuronide.

1.2.4 Data processing and metabolites identification

There are some factors affecting irreproducibility of the results including matrix effect, column temperature fluctuation due to change in lab temperature, variations in mobile phase content and flow rate and pressure fluctuations, etc. Consequently, reproducibility in chromatographic techniques can change from run to run. In order to get meaningful interpretation in global metabolic profiling, retention times should be reproducible. These variations are unavoidable, resulting in errors in peak assignments and intensity. However, there is software developed to minimise these variations with the capability of noise filtration, peak detection, alignment and normalisation. MZmine and XCMS are two automated software packages capable of detecting peaks imported in MS CDF file format from different instruments platforms (for example, LC-IT MS or GC-TOF MS). Other types of software such as, msInspect, open MS, ApecArray, and XAlign are also used for this purpose [10].

XCMS is a unique software package which has the potential to do nonlinear retention time alignment, matched filtration, peak detection, and peak matching without even the need to use an internal standard [10]. This ability can be achieved by using hundreds of endogenous metabolites as standards enabling the calculation of a non linear retention time correction for each sample. Afterwards, relative intensities can be used as a scale for comparison of changes in endogenous metabolites in different sample groups. It is also available in both online and offline version and both are free to use.

1.2.4.1 Statistical approaches for data analysis

The analysis of biological samples results in a huge amount of data which can be very difficult to process and analyse manually. In order to get meaningful interpretation of the data, suitable statistical tools should be employed to manipulate the large raw data sets providing a workable and understandable format [10]. To interpret the complex metabolic pathway information from the measurements, different multi-dimensional and multivariate statistical analyses were developed. A metabolomic profiles which discriminate between two samples has both known and unknown metabolites, therefore the use of multivariate analysis (MVA) may help identify which of these metabolites are important. The required parameters for GC-MS and HPLC-MS statistical data analysis are peak retention time, intensity, and mass/charge (*m*/*z*) ratio. Various statistical methods for data analysis have their advantages and disadvantages. MVA technique was employed by markerview™ software for the supervised and unsupervised approaches for data analysis by utilizing PCA which divides large datasets into new components. This method provides the capability of detecting variations between sample groups in a simplified manner which is very useful as a first step in data analysis. There are also some other statistical software using statistical models such as Principle Component Analysis (PCA), Soft Independent Modelling by Class Analogy (SIMCA), Partial Least Squares Discriminant Analysis (PLS-DA),K-Nearest-Neighbor (KNN), etc.

Issaq *et al.* used an LC-MS method to analyse metabolites from urine samples collected from 41 bladder cancer patients and compared these to 48 health volunteers [79]. For those peaks that showed difference between the

diseased and healthy, *m*/*z*, retention time and intensity of each peak in the chromatogram in each sample set was submitted to PCA and OPLS-DA to perform statistical data analysis. Both procedures can discriminate between the two groups at greater than 90% sensitivity and specificity (**Figure 1.2**).

Some scientists use two different statistical tools for their study. In one study, MZmine and XCMS both were used for processing the data generated with HILIC-LC-MS, RP-UPLC-MS, and GC-TOF-MS from kidney cancer samples [80]. The results were analysed using a feature selection algorithm with subsequent univariate analysis of variance (ANOVA) and a multivariate partial least squares (PLSs) approach. By using two different statistical software packages, more than 2000 mass spectral features were detected which shows a major difference between renal cancer samples compared to the control group.

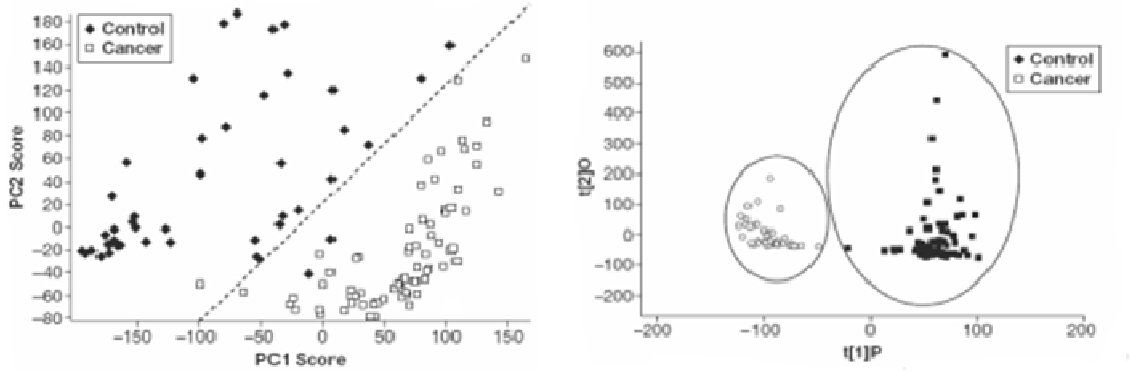

Figure 1.2: PCA (left) and OPLS-DA (right) from LC-MS study of patients with bladder cancer compared to control [81].

1.2.5 Identification

Unlike low molecular weight metabolites, proteins, peptides, DNA and RNA are less challenging to identify by MS mainly because of the linear and repetitive nature of their building blocks. In contrast, the structure of most of the metabolites are not a combination of repeating monomer units, but they appear as combination of C, H, N, O, S ,and P making compound categories including saccharides, simple sugars, lipids, steroids, isoprenoids, porphyrines, purines, pyrimids, amino acids, catecholamines, acids, ketoacids, amines, and many synthetic and natural exogenous compounds. Moreover, a wide range of chemical modifications such as hydroxylation, methylation, epoxidation, esterification, glycosylation, oxidation, reduction, and isomerisation can increase the complexity of the metabolomics study making the estimated number of metabolites from 200,000 to 1000,000 [10, 11, 82].

Identification of metabolites is very dependent on the type of study. For targeted metabolomics studies, identification of known metabolites can simply be achieved by comparing the results with that of pure standards. On the other hand, if metabolites are not known, such as for metabolite profiling for a specific disease, the metabolite identification is more difficult. This kind of study involves employing high resolution MS for accurate mass analysis, statistical analysis software packages for data processing, metabolite databases, and access to hundreds of metabolite standards for confirmation of the results [10]. Using the available metabolomics databases can facilitate the identification process of unknown metabolites.

In these databases hits for known metabolites are based on searching for *m*/*z* values in both positive and negative ion modes that produce protonated $[M+H]^+$ and deprotonated $[M-H]^-$ ions, respectively. Under some circumstances other molecular adducts like Na^+, K^+, $NH4^+$, *etc*. and aggregate ions (*eg*., $[2M+H]^+$, $[3M+H]^+$, *etc*.) are also generated that allow for further confirmation of the observed MW [10].

For each metabolite in the library, the system automatically calculates all the possible adducts and protonated and deprotonated forms which allows the system to look for those possible *m*/*z* values in the real processed result from each sample in each MS mode. Sometimes, searching the library results in multiple metabolites hits and sometimes it results in zero hit. The reason for zero hits is because there are many metabolites in biological fluids which are not known and therefore they do not yet exist in metabolomics databases. If so, MS/MS might be an option to gain structural information about them [10].

1.3 Project aims

Results in the literature so far demonstrate the suitability of separation methods including CE, GC and LC hyphenated with MS for metabolomics studies. Such hyphenations provide powerful tools being capable of resolving complexity associated with monitoring metabolites in biological samples.

The general aim of this project was to systematically investigate the metabolites (uraemic toxins) in patients suffering from CKD via analysis of

serum samples collected from patients pre- and post-dialysis, and healthy volunteers. The specific aims of the project were:

- To develop analysis methods based on CE-MS, GC-MS and LC-MS to be utilised for the analysis of serum samples.
- To implement the developed methods for both the targeted and non-targeted study of uraemic toxins and to compare their performance with the hope of getting complementary information.

Chapter 2

Experimental

This chapter describes chemicals and reagents, instrumentation, and procedures used throughout this study, unless otherwise specified in particular chapters.

2.1 Chemicals and reagents

Unless otherwise specified, the chemicals used were of analytical grade and are listed in the following table.

Table 2.1: Chemicals and reagents

Chemical	Supplier
Acetonitrile ≥ 99.9%	Fluka
Methanol ≥ 99.9%	Fluka
2-propanol ≥ 99.9%	Fluka
n-Hexane	Fluka
Acetone	Merck, reagents grade
Formic acid ≥ 98%	Fluka
Sodium hydroxide ≥ 99%	Sigma-Aldrich
Sodium chloride ≥ 99.5%	Sigma-Aldrich
MOX™ reagent, contains 2% methoxyamine.HCl in pyridine	Pierce
N-methyl-N-(trimethylsilyl)trifluoroacetamide	Sigma-Aldrich (derivatisation grade)
Alkanes (C10-C22)	Fluka

D-Mannitol	Max Planck Ins.
L-(-)-Arabitol 98%	Alfa Aesar
Myo-inositol ≥ 99%	Sigma-Aldrich
D-Sorbitol	Sigma-Aldrich
Erythritol	Max Planck Ins.
L-Threitol 99%	Sigma-Aldrich
Polybrene ≥ 95 %	Sigma-Aldrich
Poly (sodium 4-styrene-sulfonate), 70 kDa	Sigma-Aldrich
Poly (diallyldimethylammonium chloride), (40% wt)	Sigma-Aldrich
Dextran sulfate	Sigma-Aldrich

2.2 Instruments

CE-MS experiments were performed using an Agilent CE (Agilent Technologies, Waldbronn, Germany) connected to an Agilent 6320 Ion Trap LC/MS equipped with a G1607A CE-electrospray ionisation (ESI)-MS sprayer kit and an Agilent 1200 Series pump operated on the split mode with the split ratio of 1:100. Agilent ChemStation software and 6300 Ion Trap Control software were used for system control and data acquisitions in CE and MS, respectively. The spectra were collected from *m*/*z* 70-400.

GC-MS analyses were performed on a Shimadzu QP2010-plus GC-MS (Shimadzu Scientific Instruments, Sydney, Australia) fitted with a BPX-5 capillary column (25 m × 0.22 μm i.d., 0.25 μm film thickness).

LC-MS analyses were conducted using a hybrid Linear Trap Quadrupole / Orbitrap high-resolution mass spectrometer (Thermo Fisher

Scientific, Bremen, Germany). The samples were separated using a Waters 2690 separations module (Rydalmere BC NSW, Australia) employing a Waters Nova-Pak 4.0 μm, 3.9 × 150 mm C18 column.

2.3 General equipments

General equipments used are summarised in **Table 2.2.**

Table 2.2: General equipments

Equipment	Manufacturer
Ultrasonic bath	Technolab
Vortex mixer	Ratek VM1 vortex mixer
Centrifuge	Eppendorf bench top centrifuge 5424
Vacuum oven	SEM oven
Speed vacuum centrifuge	MiVac duo vacuum concentrator from Genevac Ltd (Ipswich, UK)
Syringe pump	Harvard apparatus model PHD 2000
Diamond capillary column cutter	Shortix ™, SGT

2.4 Procedures

2.4.1 Sample collection and storage

The first batch of samples (**Table 2.3**) was collected from 31 patients before dialysis and 8 patients after dialysis (HD and HDF) plus 6 healthy samples. For the second batch of samples, 13 samples were collected from patients before dialysis and 13 samples after dialysis (**Table 2.4**) plus 6

samples collected from healthy volunteers (**Table 2.5**). The samples from patients after dialysis are a combination of HD and HDF procedures.

Table 2.3: Sample information for batch one.

No.	patient code	Age	Gender	No.	Patient code	Age	Gender
1000	Pre-dialysis	54	F	1020	Post HD	66	F
1001	Pre-dialysis	55	M	1021	Post HD	76	F
1002	Pre-dialysis	53	F	1022	Post filtration	53	F
1003	Pre-dialysis	82	F	1023	Post HD	63	F
1004	Pre-dialysis	76	F	1024	Post filtration	55	M
1005	Pre-dialysis	63	F	1025	Post HD	58	M
1006	Pre-dialysis	66	F	1026	Post filtration	82	F
1007	Pre-dialysis	49	M	1027	Post filtration	49	M
1008	Pre-dialysis	60	M	1028	Pre-dialysis	58	M
1009	Pre-dialysis	58	M	1029	Pre-dialysis	60	M
1010	Pre-dialysis	66	F	1030	Pre-dialysis	49	M
1011	Pre-dialysis	76	F	1031	Pre-dialysis	82	F
1012	Pre-dialysis	63	F	1032	Pre-dialysis	74	M
1013	Pre-dialysis	53	F	1033	Pre-dialysis	55	M
1014	Pre-dialysis	49	M	1034	Pre-dialysis	66	F
1015	Pre-dialysis	82	F	1035	Pre-dialysis	54	F
1016	Pre-dialysis	55	M	1036	Pre-dialysis	63	F
1017	Pre-dialysis	74	M	1037	Pre-dialysis	76	F
1018	Pre-dialysis	60	M	1038	Pre-dialysis	53	F
1019	Pre-dialysis	58	M				

Samples 1025, 1032, and 1038 were excluded from the study.

5mL serum specimen was taken, left to sit for 15 min, spun down to make the serum, then transferred the aliquot into at least 2 × 1 mL Eppendorf vials labelled with participant number before storing into the -20 ˚C freezer.

T**able 2.4:** Sample information for batch two[1].

No.	Age	Gender	No.	Age	Gender
3001	66	M	3008	63	M
3002	76	M	3009	77	M
3003	79	M	3010	84	M
3004	74	M	3011	72	M
3005	77	M	3012	63	F
3006	76	M	3013	90	M
3007	56	M			

[1] All samples are end stage renal disease (ESRD) type. Sample number 3001 was excluded from the study.

Table 2.5: Sample information for healthy (control) group.

No.	Age	Gender	No.	Age	Gender
2000	37	F	2003	55	F
2001	39	F	2004	49	F
2002	56	F	2005	43	F

2.4.2 Capillary coatings for CE-MS

2.4.2.1 PB-PSS coating

A fused-silica capillary (50 μm i.d. × 100 cm, Polymicro Technologies (Phoenix, AZ, USA)was rinsed with Milli-Q water for 5 min at 300 μL/h

followed by 15 min flushing with 1M sodium hydroxide with the same flow rate. Then the capillary had been washed with Milli-Q water until it reached neutral pH before it was flushed with 10% (w/v) PB solution for 30 min at 300 μL/h. Afterwards, the capillary was washed with Milli-Q water for 3 min followed by flushing with the last layer solution containing 5% (w/v) PSS for 30 min at 300 μL/h. Finally, the capillary was washed again with Milli-Q water for 5 min. Before the analysis coated capillaries should be equilibrated with BGE of choice which is 1 M formic acid in this study for at least 30 min.

2.4.2.2 PDADMAC-PSS-PDADMAC-PSS coating

Bare fused-silica capillary was rinsed with water followed by activating the capillary inner wall by flushing with 1 M NaOH for 15 min at 300 μL/h. Then it was rinsed with Milli-Q water again and 0.1% PDADMAC solution in 0.5 M NaCl was chosen as the first polymer layer and the solution was flushed for 30 min at 300 μL/h. After flushing each layer of polymer coating the capillary was flushed with water for 3 min. Then the coating procedure continued with flushing the 5% (w/v) PSS solution for 30 min (300 μL/h) followed by rinsing with Milli-Q water for another 3 min. Afterwards, the procedure was repeated and another PDADMAC layer was made by flushing the capillary with 0.1% PDADMAC solution in 0.5 M NaCl for 30 mins and successively with water for 3 min (300 μL/h). The last layer coating was performed by flushing the 5% PSS solution for 30 min and final rinsing with water for 5 mins (300 μL/h). Before starting of the run the coated capillary was equilibrated with 1M formic acid for 30 min.

2.4.2.3 PDADMAC-DS-PDADMAC-DS coating

Fused-silica capillary was rinsed with Milli-Q water and then activated by flushing the 1M NaOH solution for 15 min and followed by rinsing with Milli-Q water again till it reached neutral pH. The coating procedure was performed by flushing 0.1% PDADMAC solution in 0.5 M NaCl for 30 min and rinsing with water for 3 min (300 μL/h).The next layer of the coating is made of 1% DS solution which was flushed for 30 mins at 300μL/h. Then the capillary was washed with Milli-Q water for another 3 min and 0.1% PDADMAC in 0.5 M NaCl was flushed for another 30 min (300 μL/h). The capillary had been rinsed with water before it was flushed with another 1% DS solution for 30 min (300 μL/h). At the end capillary was washed with Milli-Q water and then equilibrated by flushing 1M formic acid for 30 mins.

2.4.3 CE-MS analysis

100 μL of one of the pre-dialysis serum samples was mixed with 300 μL of acetonitrile (ACN) and then vortexed and kept in the fridge for 30 min before being centrifuged at 14000 rpm for 15 min. Then, 350 μL of the aliquot was removed and transferred to a new vial for further drying in the vacuum oven at room temperature. After almost 2 h, the dried samples were reconstituted in 40 μL of Milli-Q water and then they centrifuged again and the top 35 μL of the supernatant was transferred to CE vial for analysis *via* one of the following methods:

2.4.3.1 Method 1

The method reported by Ramautar *et al.* was used with some modifications [55]. Sample injection was performed hydrodynamically for 40 s at 50 mbar into the 100 cm of 50 μm i.d. coated (PB-PSS or PDADMAC-PSS-PDADMAC-PSS) capillaries. During analysis, a constant pressure of 30 mbar was applied. The buffer is 1M formic acid (pH 1.8) and the capillaries were preconditioned with BGE for at least 30 min before use and 10 min between runs. At the end of the day, capillaries were washed with water and dried by air. The capillary voltage was +30 kV and capillary temperature was set at 20 °C. The CE was coupled to MS via sheath liquid interface containing 0.1% (v/v) formic acid in 50% (v/v) methanol and water. Both of the solutions were filtered and degassed to avoid having interfering spikes during the run. The sheath liquid was pumped with the flow rate of 4 μL/min. Mass spectrometer parameters were set at 180 °C, drying gas flow at 4 L/min, 7.3 psi nebulizer gas pressure with the capillary voltage of -4500 V. Data were collected in positive ion mode from 70-400 *m*/*z*.

2.4.3.2 Method 2

Separation was carried out in both coated (PDADMAC-DS-PDADMAC-DS) and bare fused silica capillaries. In the case of using fused-silica, it was first rinsed with water then activated with 1 M sodium hydroxide for 15 min followed by flushing with water until neutral. In addition, a constant pressure of 30 mbar was applied from the beginning till the end of the run. 1M formic acid was chosen as BGE and the capillaries were preconditioned with BGE for at least 30 min before use and 10 min between runs. At the end of the day, capillaries were washed with water and

dried by air. Samples were injected hydro-dynamically for 30 s at 50 mbar. Before sample injection a small plug of water for 15 s was injected to induce stacking. Sample injection was followed by injection of a small plug of BGE for 5 s at 50 mbar. The capillary temperature was maintained at 20 °C and the applied voltage was set to +30 kV. Sheath liquid containing 0.1% (v/v) FA in 50% (v/v) methanol-water was delivered at 4 μL/min. ESI MS was set in the positive ion mode and the capillary voltage at -4000 V. A flow of dry nitrogen gas (temperature 250 °C) was kept at 6 L/min and nebulizer pressure of 5 psi. During the injection time nebulizer gas was shut to prevent unwanted additional suction of sample during the injection [47, 83]. As there is not an automatic programming for this purpose, it was performed manually before the start of the injection and then the gas flow was turned back to operation after the injection.

2.4.4 GC-MS analysis

Frozen samples were thawed to room temperature and 50 μL of each sample was added to 100 μL of acetonitrile and then vortexed for 30 s. Afterwards, they were incubated at 4 °C in the fridge for 30 minutes. Then samples were centrifuged at 14,000 rpm at room temperature for 15 minutes. The top 100 μL of supernatant has been removed and transferred to a glass vial to be dried at room temperature (25 °C) in vacuum oven. Finally, the dried samples in the glass vials were kept in -20 °C until the analysis time when they were derivatised freshly to prevent any unwanted changes in metabolites.

For oximation, 20 µL of 2% methoxyamin.HCl in pyridine were mixed with dried sample and vortexed for 20 s before being left at 30 °C for 90 min in heating block. In the next step, 80 µL of derivatisation reagent, N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA), was added to the sample followed by vortexing and keeping in 37 °C for 30 min in a heating block.

Standards of mannitol, erythritol, sorbitol, arabitol, myo-inositol, and threitol were derivatised by adding 20 µL of MSTFA to a detectable amount of sample followed by keeping in incubator at 37 °C for 30 min.

For SPME extraction, the manual holder of the SPME is transferred to the injection port of GC to be ready for desorption of the compounds. The fibre containing the sample is being left at 230 ˚C for 5 min until all the compounds desorbed completely. The injection mode was splitless. Between the runs the fibre was exposed to the injection port of the GC for at least 10 min followed by injecting a blank run to eliminate any possible carry over. The GC column temperature was programmed to hold at 40 ˚C for 1 min then with the rate of 10 ˚C per min rose to 280 ˚C which then was held for 0.50 min.

For derivatised samples, 1µL was injected in splitless mode at the inlet temperature of 230 ˚C for 0.50 min. For the standards of metabolites, 0.5 µL of the derivatised samples were injected. After each injection, 3 washing steps with acetone were performed in the auto-sampler to thoroughly wash the injection syringe and eliminate any carryover to the next injection. The GC column temperature was programmed to hold at 40 ˚C for 1 min then with the rate of 10 ˚C per min rose to 330 ˚C which then be held for 0.50 min.

The carrier gas was helium with the flow rate of 1 mL/min and the column pressure was 11.1 psi. The MS conditions were as follows: Ion source temperature was set at 200 ˚C and interface temperature at 250 ˚C. The spectra were recorded from min 1 for SPME extraction and from min 5.5 for derivatised method with a *m*/*z* range of 35-500 and scan speed of 5000 amu/s.

The Golm Database as a private metabolite library obtained from collaboration with the Max Planck Institute was searched for identification of metabolites in GC-MS study.

2.4.4.1 Data processing of GC-MS results: Post-run analysis

The data were processed with "GC-MS Post-run Solution" in the software by making a template manually from one of the pre-dialysis samples as a reference. Such sample was chosen because of the presence of the highest number of peaks in it (as a result of the disease). In the template, each metabolite peak was manually added to the table. The peaks were added after min 10 in order to avoid the interference from the solvent peak. 94 peaks were ultimately added to the template including the corresponding reference ions and the retention time of the individual peaks. In the end, peak integration was performed in order to ensure the adjustment of each individual peak. Unidentified peaks were then added manually. Finally the table was transferred to the other sample chromatograms. Using the "Quantitative Parameters" option in the software, the integration and adjustment were double-checked for each sample and peaks were manually to ensure that the right peak had been chosen by the software within the defined retention time shift window (4s).

The peak areas for each metabolite were then exported to a MS Excel spread sheet and the averaged area, and Student's t-test were calculated for the normalised peak areas of each metabolite. Only peaks that showed major differences (p-value < 0.05) between the normalised peak area of healthy and patients (pre & post-dialysis) were picked-up as metabolites of interest and the rest were discarded.

2.4.5 LC-MS analysis

The samples were prepared just before analysis to avoid changes in the structure of metabolites. 100 μL aliquots of serum samples were mixed with 300 μL of methanol, then vortexed for 30 s and left at 4 °C in the fridge for 30 min. Then the samples were centrifuged at 14,000 rpm for 15 min and the supernatant was removed and transferred to 500 μL Eppendorf vials. The supernatant was dried under vacuum in a MiVac duo vacuum concentrator (Genevac Ltd., Ipswich, UK) for about 2 hours. The dried samples were reconstituted in 100 μL of Milli-Q water, mixed well and centrifuged again at 4,000 rpm for 10 min. Then the top 60 μL upper layer was removed and transferred to LC vials for the analysis.

LC separations were performed at a flow rate of 0.8 mL/min using HPLC-grade solvents (0.1% formic acid in water as mobile phase A and 0.1% formic acid in methanol as mobile phase B). A 10 μL aliquot of each sample was injected and compounds were eluted over a step gradient of 10-50% mobile phase B over 3 min, then 50-80% mobile phase B over 8 min and 80-100% mobile phase B over 1 min. After holding at 100% mobile phase B for 2 min the column was re-equilibrated in 10% mobile phase B for 0.5 min. The

column was operated at ambient temperature (20 °C) and the samples maintained at 10 °C.

Centroid mass spectra were acquired in the *m*/*z* range of 50-1000 at a target resolution of 30,000 and according to the following parameters: capillary temperature of 300 °C; sheath gas and auxiliary gas flow rates were set to 30 au and 5 au, respectively. A capillary voltage of 7 V was used for positive ion acquisition and -44 V used for negative ion acquisition.

2.4.5.1 LC-MS data analysis

Raw MS files were exported and analysed using the online version of XCMS software (available at https://xcmsonline.scripps.edu/). This is a versatile tool making it suitable for targeted and untargeted metabolomics and providing solution like feature detection, retention time correction, alignment, annotation, statistical analysis, and data visualization. For identification of the compounds METLIN (http://metlin.scripps.edu/) library provide us with *m*/*z* equivalent to protonated metabolites and metabolites adduct.

Chapter 3

Capillary electrophoresis-mass spectrometry (CE-MS)

3.1 Introduction:

Capillary electrophoresis (CE) is a suitable technique for the analysis of polar compounds. When hyphenated with a mass spectrometer, the technique enables metabolomics studies. Among all CE separation modes, capillary zone electrophoresis (CZE) is the most commonly used for this purpose. The potential of using both bare fused-silica capillary and its non-covalently coated counterpart have been investigated in many publications. Coated capillaries can potentially prevent adsorption of proteins or other matrix components to the capillary wall which results in migration time shifts and irreproducible runs [15].

In this chapter, optimisation of a CE-MS method for the analysis of metabolites in serum samples of patients suffering from CKD is demonstrated. The effect of different coating procedures on the quality of separation was investigated by monitoring the migration time shifts and the results compared with those obtained with bare fused-silica capillaries.

3.2 Experimental

The general experimental details including chemicals and instrumentation are presented in Chapter 2. Specific experimental conditions are given in each of the figure captions.

3.3 Results and Discussion

Like any metabolomics study, sample preparation is the first step followed by analysis and data processing. For sample analysis, several procedures were followed for coating the capillary and the results compared with bare fused-silica. The work flow-chart has been summarised in **Figure 3.1**.

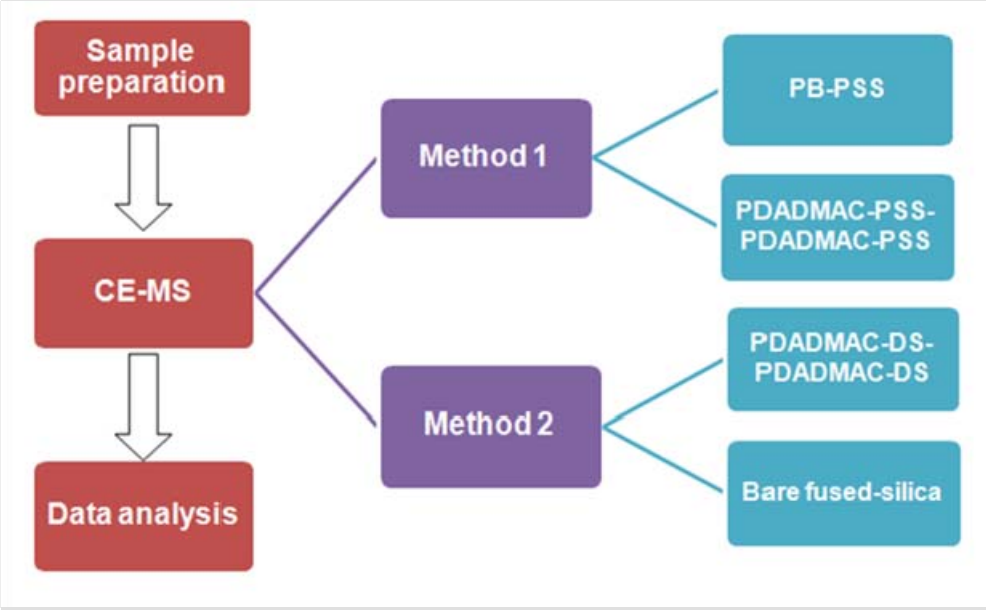

Figure 3.1: Work flow-chart for CE-MS analysis of metabolites.

3.3.1 Sample preparation

Aside from the analysis method employed, sample preparation is an integral part of every bioanalytical method [20]. Selective isolation of analyte(s) of interest from the matrix, minimisation or elimination of matrix components, and pre-concentration of the analytes are the target in an appropriate sample preparation method. The quality of bioseparation is therefore directly affected by the sample preparation step. In CE-MS analysis, for instance, while a larger sample plug (more than 1-5% of the total capillary length [41]) can be injected in order to address the low sensitivity arisen from the low concentration of metabolites, this frequently caused a current drop. One reason can be because of the conductivity difference between the BGE and sample matrix. In the meantime, the current drop can be also due to the irreversible adsorption of sample matrix components (like lipids, amino acids, proteins, etc.) on the capillary wall leading to irreproducible EOF [55].

Protein precipitation can generally be done by adding some organic solvents (usually methanol or ACN). By increasing the ratio of ACN added to the serum one can increase the efficiency of protein elimination preventing them from attachment to the capillary wall, but dilution of the sample also needs to be taken into account. The amount of ACN added to the serum was increased sequentially from 1:1 to 3:1 followed by pre-concentration via drying in vacuum oven at room temperature. Also reconstitution of the dried samples in water was used to induce the stacking effect and increase the amount of sample introduced to the capillary [41]. The ACN to serum sample ratio of 1: 3 (v/v) was found appropriate for sample preparation.

3.3.2 Bioanalysis

Using a low pH BGE together with bare fused silica capillary can cause a significant decrease in EOF [55]. Also as mentioned above, matrix components mostly in biological samples can adsorb to the capillary wall causing irreproducible EOFs and resulting in migration time shift. The suggested mechanism is based on the multi-point attachment between biomolecules hydrophobic and charged regions, and ionised silanols, affecting the whole surface charge and causes EOF changes [84]. It consequently results in samples loss, deterioration of separation performance and poor repeatability.

Covering the silanol groups is often employed to resolve the problems, provided that the coating is homogenous and stable [85]. In some studies, non-covalently coated capillaries with bi-layer of oppositely charged polymers PB-PVS, were successfully used for the rapid and reproducible analysis of peptides [86]. The PB-PVS coating facilitates the rapid separation by providing a considerable EOF at low pH BGEs [86]. In this study the procedure offered by Ramautar and colleagues [55] was slightly modified by replacing the PVS with PSS as negatively charged final layer. Before the start of coating, 100 cm of 50 μm i.d. capillary was cut with a diamond capillary column cutter. As the precision of spraying in CE-MS is highly dependent on the quality of the cut capillary ends should be checked under the magnifying glass to make sure the cuts are flat. Clearly, jagged edges as a result of poor cut not only reduce the quality of the spraying but also can act as adsorptive sites for sample components affecting repeatability and sensitivity [87]. To prevent damage to the upcoming coating, the polyamide coating of bare

fused-silica was removed from both capillary ends by using a flame before the start of coating procedure and then wiped with isopropanol. Unfortunately despite all these considerations, the method worked only for a few runs and then current drop occurred. The same happened for a freshly prepared capillary as well. The reason might be because of the attachment of amino acids or other matrix compounds to the capillary wall as a result of inappropriate covering or bleeding of the coating.

The suitability of multi-layer coating, which is known to be more stable, was therefore examined following the procedure reported by Nehme and colleagues [85, 88], with some modifications for creating a four-layered coating consists of PDADMAC-PSS-PDADMAC-PSS. Instead of PB in the original procedure, 0.1% PDADMAC solution in 0.5 M NaCl was used which may give more stability compared to PB [89]. Also PDADMAC is known as a polycationic polyelectrolyte and it owes its popularity to its high charge density, availability and suitability for coating of silicon and glass surfaces [85]. Additionally, it is a "strong polycation" with the ionisation state independent of pH making it more suitable for the coating purposes [85]. This negatively charged capillary wall coating can also bring higher EOF and increase the separation speed [86].

A sample was prepared and analysed according to CE-MS method 1 to evaluate the performance and quality of the capillary coated with PDADMAC-PSS- PDADMAC-PSS. To regenerate the coating, the capillary was flushed with the last polymer layer after a few runs. To avoid contamination of the MS, the capillary was removed from the sprayer before flushing with the last layer coating solution for a few minutes. Although this

multiple layer coating was more efficient and provided more stable current up to 20 runs, instability in current emerged again thereafter, which can be again related to bleeding or attachment of some biological samples. Unlike bare fused-silica, it is not possible to wash the capillary surface with sodium hydroxide in order to refresh the capillary wall since it easily removes the coating. In both of the experiments the sensitivity was very poor and frequent current drop prevented method optimisation (data not shown). Hence, efforts were focused primarily on addressing the current drop issue and then to develop a method with the highest possible sensitivity for the analysis of metabolites.

Based on the literature, DS is another monomer capable of providing a stable coating featuring pH-independent EOF from the anode to cathode in pH range of 2-11 [90]. Therefore, a procedure [85, 88] was followed to make a four-layered PDADMAC-DS-PDADMAC-DS coated capillary. Analysis was performed based on CE-MS method 2. A plug of water was also injected prior to the sample to induce the stacking effect and increase the volume of sample loaded into the capillary. Nevertheless, the reproducibility was very poor represented by a marked shift in migration times. It is likely again that the coating was not stable enough and bled from the capillary causing irreproducible EOF. While using coated capillaries for CE-MS analysis is frequently seen in the literature, there are also many instances of metabolomics studies supporting the use of bare fused-silica capillaries (see for eg. [91, 92]). These findings therefore led us to try bare fused-silica capillaries for further studies.

1 M formic acid (pH 1.8) is the most commonly used BGE for the CE-MS analysis of cationic metabolites using bare fused silica capillary [49, 50, 55, 91]. While due to the higher number of active silanol groups on capillary wall at higher pH (pH > 4) EOF is sufficient, such a low pH necessitates applying additional pressure after injection to push forward the sample plug since in this case the electrophoretic mobility is also not adequately high to bring about a fast run. A sample was injected a few times with the same method but unfortunately the reproducibility was still poor. **Figure 3.2** shows the base peak chromatograms of four repeated runs of the same sample. A gradual shift in migration times is observed, which is likely due to the attachment of matrix components to the capillary wall leading to a change the EOF and consequently in migration time. Reduced intensity of the peaks from run-to-run also shows lack of repeatability and poor sensitivity (see **Figure 3.2**(c)). These issues were addressed by flushing the capillary with 1M sodium hydroxide solution between the runs followed by washing with water and preconditioning with 1M formic acid again (**Figure 3.2** (d)). Although this procedure greatly increased the reproducibility and repeatability of the method, it is very time consuming. For each flush the capillary should be removed from the sprayer chamber to avoid contamination of the MS. Considering the re-adjustment of the capillary inside the sprayer, it is very tedious and needs a skilled operator. Also the capillary should be flushed with BGE thereafter for at least 30 min. Despite these difficulties, a big advantage of using bare fused-silica capillary instead of a coated one is the absence of current drop, which enhances reproducibility.

Overall, the sensitivity issue, repeated current drop in addition to poor reproducibility are the main factors that hindered progress in using this technique. Although current drop can be circumvented by using a bare fused silica capillary instead of a coated capillary, the sensitivity was not still enough to enable a suitable metabolomics study, evident from a very low number of the metabolites detected in comparison to the other metabolomics studies reported before.

3.3.3 Troubleshooting

Along with the current drop, generation of spikes and having an unstable baseline is another issue in CE-MS bioanalysis. Our experience with this issue is presented here together with some recommendations to tackle some of the operational difficulties associated with CE-MS in bioanalysis.

As the primary source of spikes is air bubbles, extensive degassing of freshly prepared sheath liquid and BGE was conducted before analysis, but spikes were still appearing during analysis. Logically, the presence of air bubbles even after degassing can be addressed in the content of the sheath liquid.

The choice of organic solvent for the sheath liquid is typically limited to some specific organic solvents, mostly methanol and isopropanol. While ACN is a popular solvent in LC-MS applications, its usage as a solvent in sheath liquid is usually avoided due to the tendency of ACN to remove the polyamide coating from the capillary, leading to MS contamination.

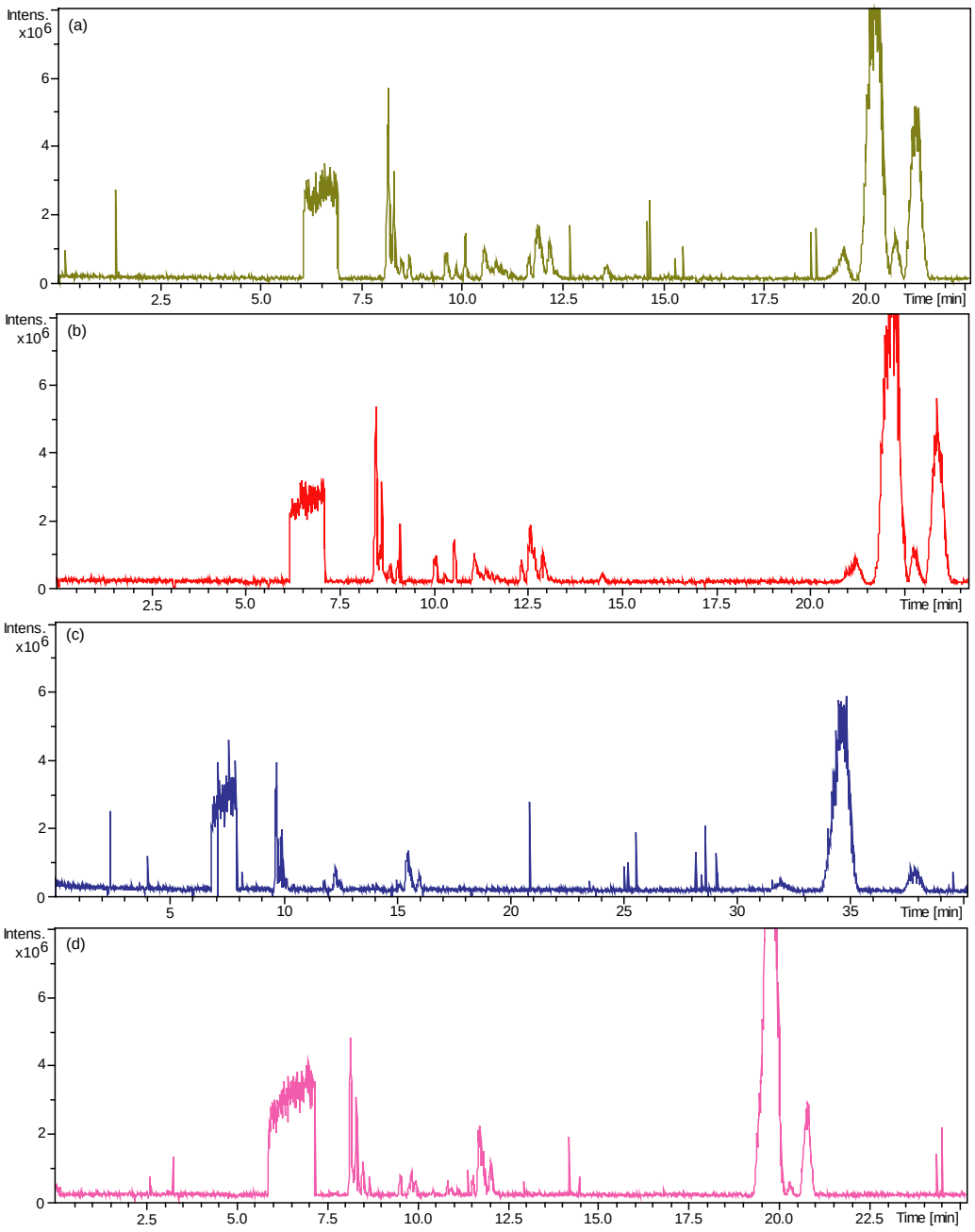

Figure 3.2: BPC of CE-MS analysis of a serum sample (following Method 2 in Experimental section). (a) first injection in a bare fused-silica capillary activated with 1M NaOH followed by preconditioning with BGE, (b) second injection in the same capillary, (c) third injection, (d) injection of the sample in the same capillary re-activated with 1M NaOH followed by preconditioning with BGE. (Note: time scales are different).

Therefore, two sheath liquids one containing 0.1% (v/v) formic acid in 50% (v/v) methanol/water and the other containing 0.1% (v/v) formic acid in 50% (v/v) isopropanol/water were freshly prepared and their performance were compared. Sheath liquid containing isopropanol was prone to produce less air bubbles resulting in fewer spikes in the baseline, most likely due to the higher viscosity. Nevertheless, the use of methanol in the sheath liquid is preferable as it provides better ionisation efficiency than isopropanol. It was experimentally found that increasing the dry gas flow can significantly reduce the spikes, possibly by refracting the air bubbles from the aerosol before entering the MS analyser. The effect of dry gas flow on the presence of spikes is shown in **Figure 3.3**. Although increasing the dry gas flow can reduce the spikes and helps to have more stable baseline, it can reduce the sensitivity as well. Hence, a compromise between these two parameters needs to be considered during method optimisation.

Although in this study the CE-MS technique failed to provide useful results, some of the major issues that a user is likely to face with in bioanalysis are presented here (**Table 3.1**). Suggestions are also made which might help to tackle these issues.

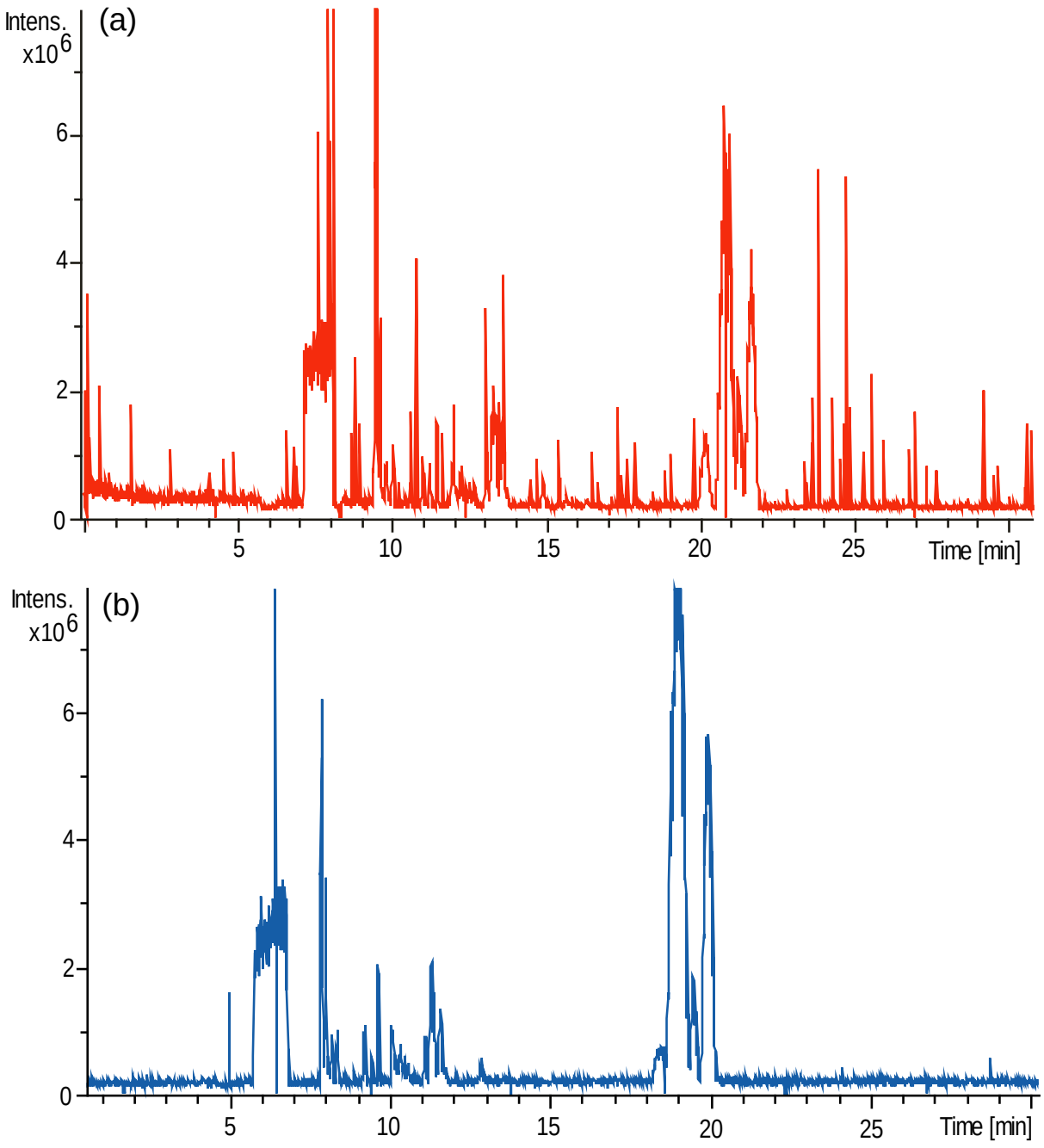

Figure 3.3: The effect of dry gas flow on spikes generation with sheath liquid containing methanol. Method 2 in experimental section was employed with different dry gas flows. Dry gas flow is 4 L/min (a) and 6 L/min (b).

Table 3.1: Troubleshooting for CE-MS bioanalysis.

Issue	Suggestions
Low sensitivity	Pre-concentration of the sample or using one of the stacking techniques/ decreasing the flow rate of sheath liquid to avoid sample dilution/ making sure about the position of capillary in the sprayer needle/ checking if capillary ends are smooth and flat [87]
Poor reproducibility and current drop	In case of using coated capillary flushing with the solution of last layer coating after every few runs/ In case of fused silica capillary flushing with 1M NaOH is suggested to have better reproducibility
Spikes in the TIC	Extensive degassing of the sheath liquid and BGE can minimise the spikes/ sheath liquid made of isopropanol has less sipkes compared to that of methanol/increasing the dry gas flow can reduce the spikes however it compromises the sensitivity
Unstable sensitivity	Turning the nebulizer gas off till the end of the injection [83],[47]
Frequent capillary blockage	Can be resolved by some changes in sample preparation, after the first centrifuge when supernatant has been removed it should be centrifuged again

	to precipitate the tiny residues that can block the capillary/Filtering the BGE to get rid of the small particles
Mass spectrometer contamination	Ion source should be wiped with a soft cloth using isopropanol/ All the chemicals for making the solution should be mass grade to avoid contamination in the instrument and getting signals for the impurities
Unstable signal	Check to see if the capillary is properly cut/ Capillary should be readjusted/ or sprayer tip has some erosion and needs to be replaced [87]

3.4 Conclusions

Despite some advantages reported for the applicability of coated capillaries such as increasing the EOF in lower pH ranges [86], our experience with CE-MS analysis of metabolites suggested the superiority of bare fused-silica capillary over coated counterparts. Coated capillaries also suffered from the time consuming (and irreproducible) procedures for coating, and bleeding after few runs. More importantly, coated capillaries investigated in this study failed to fulfil their mission, that is the permanent prevention of adsorption of biomolecules to the capillary wall during analysis [85]. While adsorbed species in bare fused-silica capillaries can be removed using concentrated sodium hydroxide, this reagent can easily remove the coatings from coated capillaries. On the other hand, the slow

EOF associated with the use of bare fused silica capillary in low pH ranges can be resolved by adding some pressures during the run.

In spite of all the optimisation and method developments employed by using different capillaries and samples preparations, unfortunately our CE-MS study failed to provide the required sensitivity for a reliable metabolomics study. Even after improving the sensitivity, the ion-trap MS is in fact technically limited in providing the accuracy needed for such a metabolomics study. Using MS analysers with higher resolving power such as TOF or Orbitrap together with employing a recently introduced sheath-less interface for CE-MS has been demonstrated to significantly improve the sensitivity and accuracy of metabolomics studies with CE-MS [43], which can be the focus of our future studies.

Chapter 4

Gas chromatography-mass spectrometry (GC-MS)

4.1 Introduction:

Capillary GC hyphenated with MS has proved to be very useful for metabolomics studies mainly due to its sensitivity, high resolution, and reproducibility. This technique owes its popularity to the availability of GC-MS (EI) spectral library facilitating the identification of biomarkers and aiding the subsequent mechanistic elucidation of the biological or pathological variations [9]. Regardless of all the benefits that analysis with GC provides, the prerequisite for GC-MS analysis is to make the compounds volatile yet thermally stable. Due to highly polar and non-volatile nature of the metabolites, they cannot be directly analysed by GC-MS unless undergoing chemical derivatisation step [9].

Generally there are 6 steps to do GC-MS analysis [59]:

1) Extraction is usually the first step of metabolomics study with any analytical technique. This step should be performed very carefully to avoid degradation, modification or lose of metabolites.
2) Derivatisation is the necessary step to make the compounds volatile for GC-MS analysis.
3) Separation which involves method development and optimisation to separate as many metabolites as possible.

4) Ionisation of the compounds which elute from the GC. This step is required for subsequent mass spectrometry. Electron impact (EI) ionisation which is the most widely used technique, produces reproducible fragmentation patterns and molecular ions.
5) Detection of molecular ions that can be achieved with different mass spectrometry analysers such as, single quadruple, ion trap, and time of flight.
6) Data processing is the final, and perhaps the most important, step that needs special softwares or tools. Since shift in migration times is usually happened, particularly when a large number of samples analysed, a suitable software is required to align the retention times and mass-spectral fragmentation patterns. Failure in doing this step carefully can easily result in false reports. More information can be achieved by matching the aligned signals to metabolomics library hits and interpret the patterns in different sample groups.

The sample preparation step can be facilitated through developing a suitable solid-phase micro-extraction (SPME) method, by which metabolite extraction can be accomplished within an SPME fibre in a single step. SPME as a sample preparation that offers fast and minimal sample handling can minimise the potential for metabolite loss during the sampling and sample preparation. This technique has also the potential to eliminate the artifacts producing from sample preparation, extraction and storage. In the mean

time, the desorption step can be combined to GC-MS analysis by leaving the fibre in the injection port of GC for thermal desorption [93].

Headspace and direct extraction are two different modes that have been reported in the literature for the extraction process and the choice depends on the volatility of analytes of interest. *In vivo* SPME has been successfully applied for the extraction of metabolites from different biological systems, for example, human emissions [93]. Due to the popularity of this technique for metabolomics study as a result of its availability and straightforward analysis, in this study we initially explored the feasibility of *in vitro* SPME for extraction of volatile metabolites, if any, from the headspace of serum samples. Classic derivatisation method was also performed and the results were compared.

4.2 Experimental

The general experimental details including chemicals and instrumentation are presented in Chapter 2. Specific experimental conditions are given in figure captions.

4.3 Results and discussion

4.3.1 SPME-based GC-MS analysis

Initially, SPME fibres were exposed to the headspace of a serum sample without doing any sample pre-treatment. The main purpose was to optimise a suitable condition in which the maximum amount of volatile compounds in the sample can be adsorbed to the fibre. The time required for

sample headspace to reach to equilibrium with the bulk solution is one of the important parameters that should be optimised for getting reproducible results [93].

When the volume of the sample is small (100 μL), the system reaches to equilibrium faster if the smallest possible sample vial is used. Therefore, 100 μL of the serum sample without doing protein precipitation or any other pre-treatment has been transferred to a 2.5 ml glass vial sealed with a septum. Then, the manual SPME holder was used for fibre (100 μm PDMS, Supelco, Kit 1, 24 GA) exposure to the headspace of a sample for different time frames from 30 min to overnight at room temperature. The extracted samples were then analysed immediately after completion of extraction by inserting the manual holder of the device in the injection port of the GC and leaving for 5 min until the volatile compounds were desorbed at 230 °C. Obviously, the greater the exposure time in the headspace, the better extraction was achieved. The longest extraction time tried was 12 hours which resulted in the highest number of peaks seen in the spectrum. After each injection the fibre was thermally cleaned by leaving it in the injection port of the GC for about 10 min to allow the release of any carryover from the previous sample or lab environment. Because of the thermolabile nature of the biological samples, it is better to avoid leaving them out of fridge for a long time. Therefore, different possibilities were tried to increase the speed of extraction in the shortest possible time. Salting-out is one possibility that may expedite the extraction process.

To stimulate more compounds to be transferred to the gas phase, a mixture of saturated NaCl was prepared and 10 μL of that was added to 100

μL of serum sample and vortexed for 30 s. Afterwards the PDMS fibre was exposed to the headspace of the sample and left for different time frames to evaluate the effect of salting-out on the extraction time. As can be seen in **Figure 4.1**, sample extracted in the presence of saturated salt demonstrated slightly higher intensity for some of the volatile compounds in time periods, for example, 10.5-12 min but reduction in the intensity in some other areas, eg. 6-8 min, compared to the sample without salt. Therefore, adding salt to achieve more effective extraction was not conclusive. On the other hand, NIST library was unable to identify any relevant peak where most of the compounds were recognised as hydrocarbons. To get more confidence about the reliability of the results, 100 μm PDMS fibre was exposed to the headspace of an empty vial and a vial with 10 μL of saturated salt as controls. Surprisingly, analysis of both control samples revealed a significant overlap with the spectra of real samples which reduced the level of certainty and resulted in the conclusion that perhaps some impurities in the saturated salt solution or vial septum itself can significantly contribute to the release of volatile compounds and interfere with our compounds of interest (see **Figure 4.1**). Given that some of these peaks could also be seen even when the fibre was exposed to an empty vial, it is more likely that the source of contamination was from the vial septum or the lid. Besides, the laboratory environment has some volatile compounds which might be adsorbed to the fibre in the time between removal from extraction vial and transfer to the injection port of the GC.

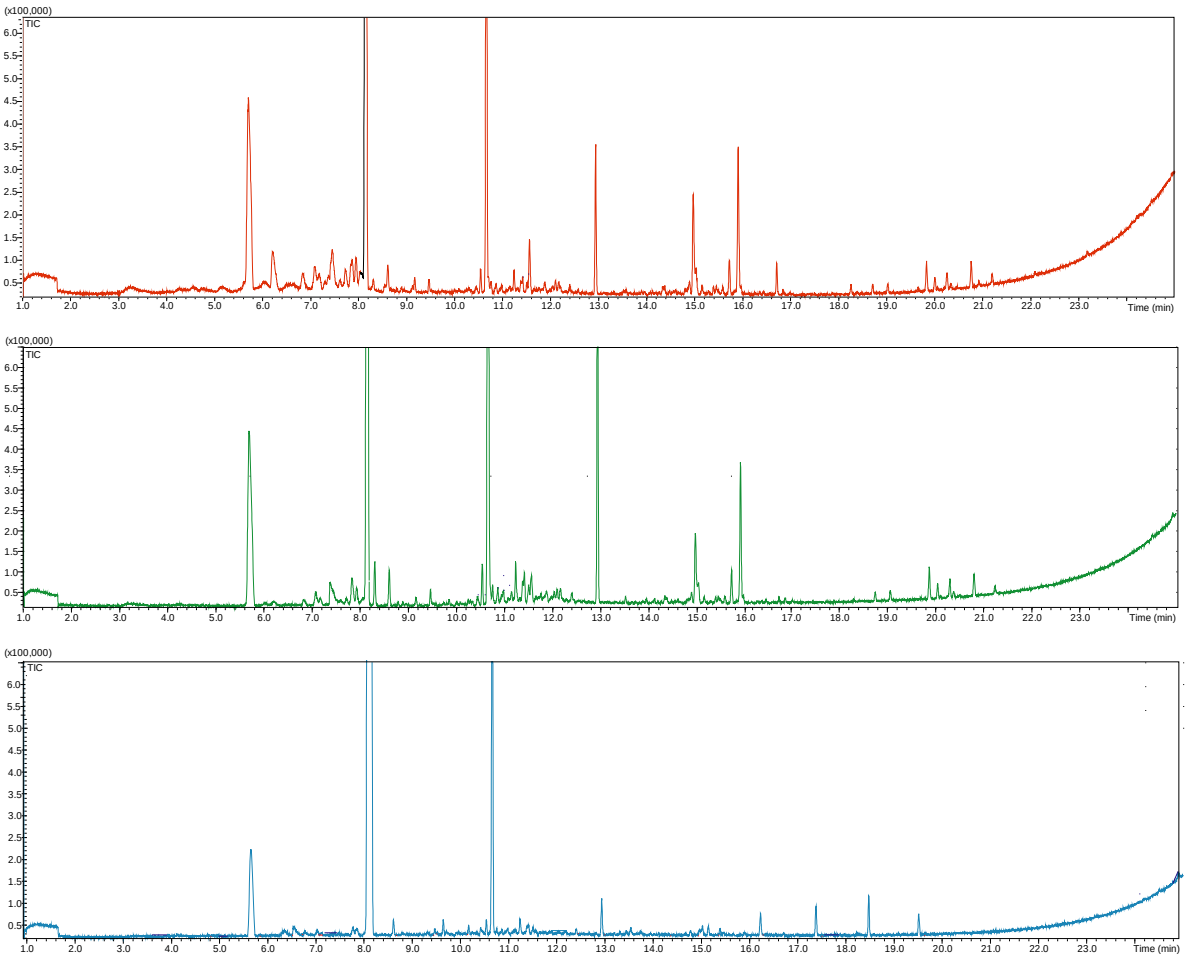

Figure 4.1: TICs for GC-MS analyses of samples extracted with SPME (100 µm PDMS fibre). SPME was exposed for 90 mins in the headspace of vial containing 100 µL of serum sample (red), SPME was exposed for 90 mins in the headspace of vial containing 100 µL of serum plus 10 µL of saturated salt (green), SPME left in the headspace of empty vial for 90 mins (blue).The fibre containing the sample is being left at 230 ˚C for 5 min. The injection mode was splitless. The GC column temperature was hold at 40 ˚C for 1 min then rose to 280 ˚C (10 ˚C/min) and held for 0.50 min. The carrier gas was helium with the flow rate of 1 mL/min and the column pressure was 11.1 psi. MS conditions: Ion source temperature, 200 ˚C and interface temperature, 250 ˚C. The spectra were recorded from min 1 with *m/z* range of 35-500 and scan speed of 5000 amu/s.

Due to the lack of enough evidence for the source of these peaks and also the low possibility of the presence of volatile compounds contributing to the kidney disease in the serum samples, SPME-based extraction was not explored further in favour of classic derivatisation-based procedures. This conclusion is also supported by the fact that SPME-based extraction for biological samples has been mostly focused on direct exposure of extraction fibre to the liquid samples followed by LC-MS analysis, while headspace has been employed for the volatile emissions in, for example, plants, cell cultures [94], human breath [95], and insects [96].

4.3.2 Derivatisation-based GC-MS analysis

Three groups of sample, pre-dialysis, post-dialysis and control group were analysed after derivatisation. For two samples from each category, the replicate analysis was done to ensure the reproducibility and consistency for the results. The replicates of the same samples were compared together and once the samples showed good reproducibility between the runs then replicate analysis stopped and each sample has been analysed only once.

Before derivatisation the samples should be completely dried as silylation reagents are extremely sensitive to moisture otherwise it can cause degradation of derivatives [97]. One of the disadvantages of silylation is its contribution in conversion reactions , for example, Halket *et al*. described how arginine is converted into ornithine by reaction with BSTFA or MSTFA [98].

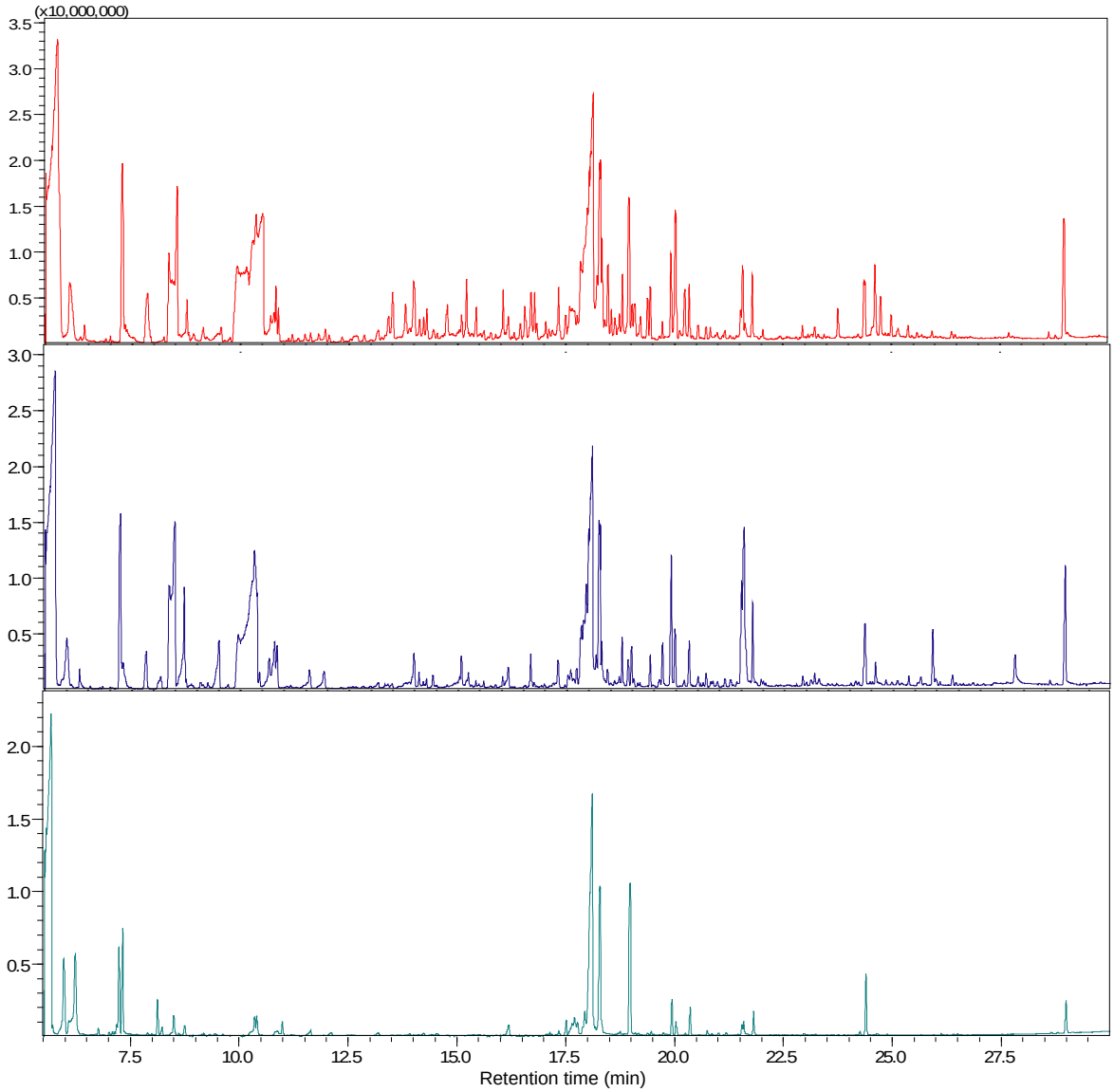

Figure 4.2: Typical GC-MS TICs for pre-dialysis (red), post-dialysis (blue) and healthy (green) samples (Note: scales are different). GC conditions: The GC column temperature was held at 40 ˚C for 1 min then rose to 330 ˚C (10 ˚C/min) and was held for 0.50 min. Other conditions as Figure 4.1.

Examples of GC-MS TICs for samples before and after dialysis, and a healthy control were demonstrated in **Figure 4.2**. It can be clearly seen that the level of most of the metabolites are higher in pre-dialysis compared to

post-dialysis and healthy. Even after dialysis, the level of metabolites is generally higher than healthy.

Targeted metabolites were then identified and cross-checked throughout 3 steps: 1) the results were searched against Golm metabolomics library, 2) the authenticity of each metabolite was cross-checked by comparing the calculated *versus* reference retention indices, and 3) the available standards of identified metabolites were injected as a final confirmation.

4.3.3 Targeted metabolomics study

Identification of metabolites in the library based on the similarity percentage is the most convenient, yet uncertain option since for most of the peaks there is more than one hit (and sometime with similar percentage) suggested by the library. In addition, like many other metabolomics libraries, the Golm library also is not complete and for many metabolites there is no hit in the library (considered as unknown). Also, standards of many of the metabolites are not available or they are very expensive.

Although mass spectral matching is shown to be very useful for identification of the metabolites, it is not sufficient for the non-ambiguous identification. It is mainly due to the presence of different isomers of a compound in very complex biological samples. Therefore, retention indices (RI) based on n-alkenes can provide additional information about the compound of interest and gives more certainty to the library search. There are some spectral libraries can facilitate the prediction process for RI [99].

Table 4.1: The retention times of alkenes from C_{10} to C_{22} in the BPX5 column.

Number of Carbons	t_R (min)	Number of Carbons	t_R (min)
C_{10}	6.205	C_{17}	16.233
C_{11}	7.867	C_{18}	17.375
C_{12}	9.462	C_{19}	18.468
C_{13}	10.975	C_{20}	19.507
C_{14}	12.403	C_{21}	20.520
C_{15}	13.752	C_{22}	21.468
C_{16}	15.027	-	-

The potential of this method was examined and proven by the injection of available standards (only for those metabolites in the list of known uraemic toxins). Then the method was used for the identification of those hits in the library that there was no standard available for them.

Although RI can give some information about the expected order of targeted metabolite, sometimes for some metabolites and their isomers RIs are very close to each other and therefore are very hard to differentiate. There are a lot of other factors like column type, length, and diameter and sample preparation that affect RI, therefore RI can give a relative order for metabolites retention instead of giving a very accurate retention time. As a slight retention time shift is expected in different sample runs, for those metabolites that have RI very close to their isomers, the possibility of confirming them with calculated RI is highly uncertain.

Table 4.2: Identified metabolites in targeted study.

Compound	t_R (min) in the sample (based on library search)	Similarity search	RI Ref.[1]	RI calc.
Myo-inositol	20.022	90	2091.9	2051
Threitol	13.412	95	1501.7	1475
Erythritol	13.523	96	1510.2	1483
Mannitol	18.323	-	1928.8	1887
Sorbitol	18.387	90	1935.8	1893

[1]Retention indices extracted from Ref. [100].

RIs for all detected metabolites were calculated from the data in **Table 4.1** and compared with those extracted from the literature [100] (RI Reference). **Table 4.2** summarised metabolites identified based on minimum differences between their calculated and referenced RIs and also got confirmed by the library search. While there is a systematic shift (average= -35±7) between referenced and calculated RIs, the authenticity of the identified compounds was further confirmed by injecting the individual standards of these metabolites. Due to the absence of carbonyl groups in these compounds, methoximation, which prevents cyclisation and stabilises carbonyl groups in β –position of reducing sugars [101] by protecting ketone and aldehyde groups [102], is not required and derivatisation can be simply performed only by silylation. As can be seen in **Table 4.3**, there is a very good agreement between the expected (in the sample) and observed (in the standard) retention times of all compounds demonstrating the suitability of this study.

Table 4.3: Retention times for individual standards and the corresponding retention times in samples.

Standard name	t_R (min)	t_R (min) of the metabolite in the sample	Deviation
Myo-inositol	20.018	20.022	+0.004
Threitol	13.415	13.413	-0.002
Erythritol	13.519	13.523	+0.004
Mannitol	18.316	18.323	+0.007
Sorbitol	18.384	18.387	+0.003
Arabitol	16.055	16.053	-0.002

As mentioned before, for many metabolites a standard is not available. Therefore, a RI study can be used to predict their retention time or their elution order leading to their recognition for further studies.

The ultimate goal of this study is to monitor the fold change of metabolites of interest (targeted) in samples before and after dialysis (in comparison to healthy). **Table 4.4** summarised such results for the metabolites identified so far. These results show the course of changes in the level of these sugar alcohols in different groups. As can be seen, although dialysis treatment significantly reduced the amount of these compounds, still the level of them are higher than in the control group, with myo-inositol very close to healthy (fold change 1.4) in contrast to mannitol, which is still 20 time higher.

Table 4.4: this table shows the changes in the level of targeted metabolites in patients before dialysis and after dialysis compared to that of healthy.

Metabolite	Calc. RI	Fold change pre-dialysis *vs*. control	Fold change post-dialysis *vs*. control
Myo-inositol	20.022	14	1.4
Threitol	13.413	138	8
Erythritol	13.523	70	3
Mannitol	18.323	371	20
Arabitol	16.053	571	4

Based on these findings, one can conclude that by increasing the time of dialysis treatment, the level of these metabolites is likely to reach to the normal level. However, it might be also the case that the rate of reduction is not linear so spending more time on dialysis doesn't lead to any further elimination of these metabolites. To gain more insight needs more investigation.

4.3.4 Non-targeted metabolomics study

Unlike known water-soluble uraemic toxins, there are some other metabolites that don't exist in the list of known uraemic toxins but level of them is significantly higher in patients compared to healthy volunteers (unknown). Study of these compounds is important since their extraordinary levels can potentially cause toxicity associated with CKD, thus permits discovering new uraemic molecules. Examples of such compounds have been displayed in **Table 4.5**. The RIs, obtained via predictions from the

Table 4.5: Identified non-targeted metabolites.

Compound	Reference RI[1]	Calculated RI	Fold change pre *vs.* control	Fold change post *vs.* control
Proline, DL	1303.4	1289	35.7	5.7
Serine, DL	1369.3	1345	25	6
Erythronic acid	1548.7	1512	6.8	1.4
Threonic acid	1568.2	1519	9.6	7.5
Amino acid	-	1529	86.2	1.6
Phenylalanine	1635.4	1620	16.6	8.3
carboxylic acid	-	1649	19.5	6.1
Unknown	-	1730	667	4.3
Citric acid	1827.8	1796	120	10.7
Unknown	-	1971	43	4.6

[1] Obtained from Ref. [100].

literature [100]. For each of the metabolites in the table, there was more than one possibility based on the library search. RI for each of them was calculated and the one being closer to the reference RI was chosen (p-value < 0.05).

Metabolomics profiling can give us a picture of the whole metabolome changes after any treatment without the need for individual identification of each metabolite. This picture can be very useful in terms of monitoring the effects of different treatments (for example, therapeutic medications, dialysis, etc.) on a group of patients. In this case, it can provide us with invaluable information about the level of small molecules in the patients before and after a treatment compared to control.

Figure 4.3 demonstrates global metabolomics profiling study (*p*-value < 0.05) for samples from patients before and after dialysis compared to healthy volunteers. It can be seen that dialysis treatment was quite effective in (partial) removal of most of the detected metabolites.

4.3.5 Principal component analysis (PCA)

Principal Component Analysis (PCA) is used to identify the correlations within a set of observed variables and samples. Samples from patients before and after dialysis, together with those from healthy volunteers were fed in Multibase as an Excel Add-Ins program, which enables direct processing of Excel data. Consequently, three principal components were extracted as three sets of scatter plots. PCA plots summarised the relationships among the samples and normalised peak areas of common metabolites. **Figure 4.4** shows the PCs with the greatest contribution in explaining the original data, which is 74% for the first component and 8% for the second component. Therefore, the accumulated contribution of component 1 and 2 goes up 82%, reflecting the high reliability of the results. As can be seen, there is a significant difference between the distributions of pre-dialysis group and the other two groups making the recognition very easy. Furthermore, there is a significant overlap seen between post-dialysis and healthy groups, suggesting the effectiveness of the dialysis treatments. This analysis can offer a very quick and useful way to differentiate healthy and pre-dialysis patients.

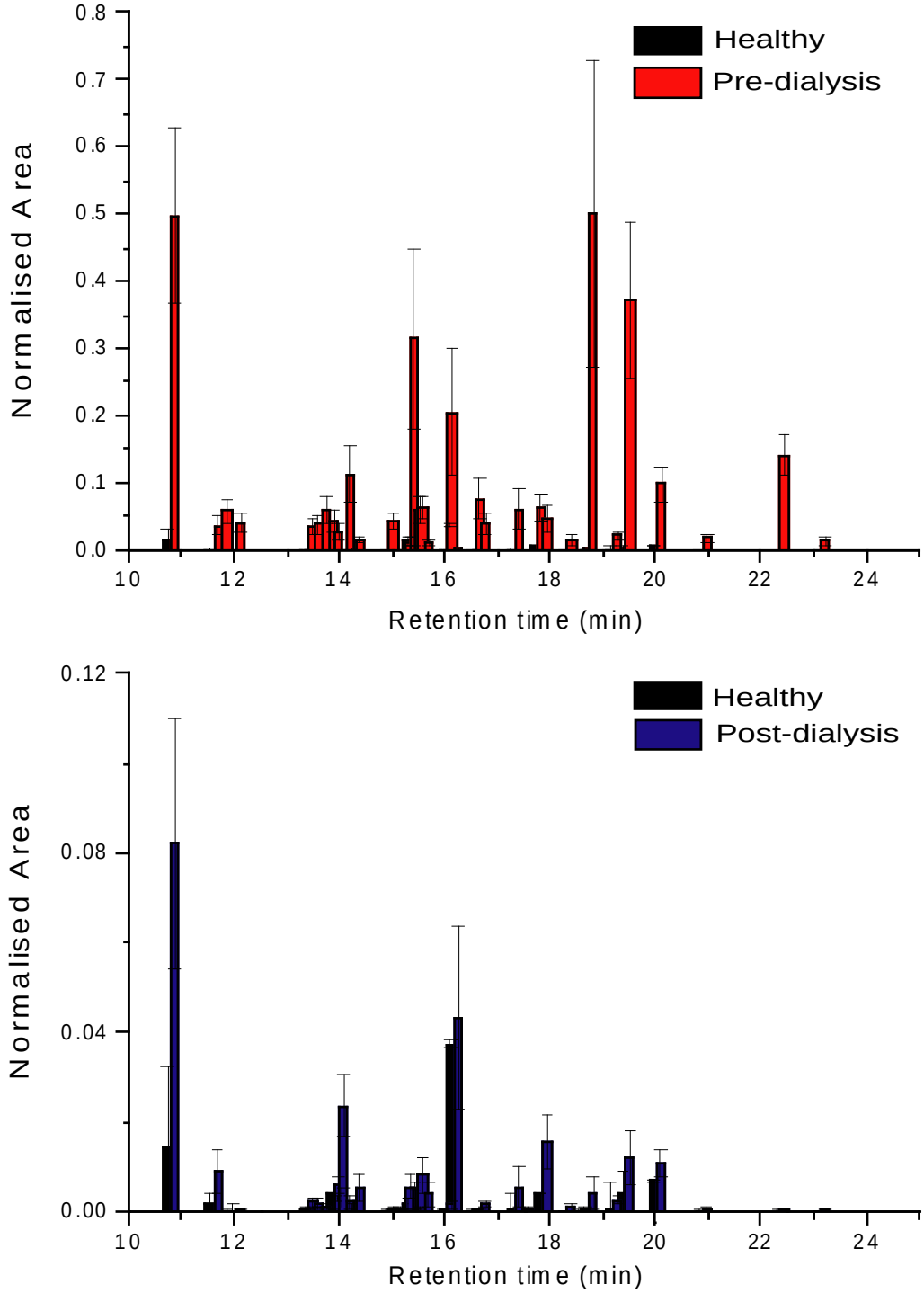

Figure 4.3: Global metabolome profiling. Comparison between pre-dialysis treatment and healthy (top) and post-dialysis and healthy (bottom). (NOTE: different scale was used to make visible the results for healthy samples).

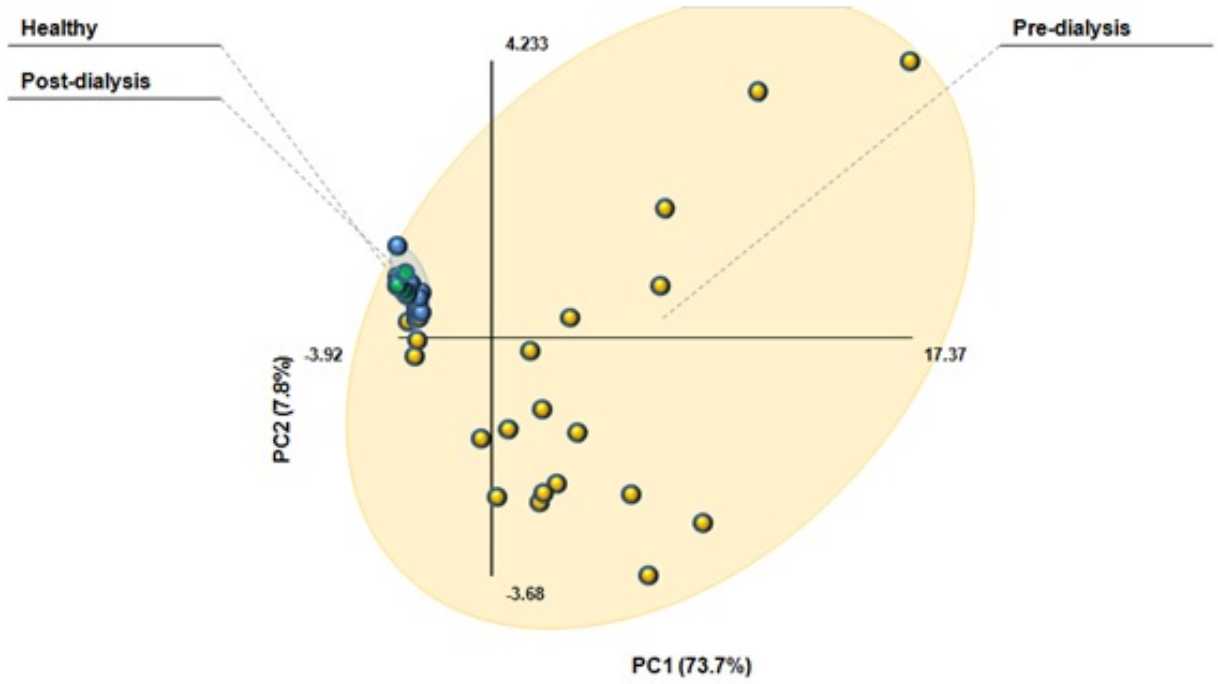

Figure 4.4: Pre-dialysis (yellow), post dialysis (blue), and healthy (green).

4.3.6 Repeatability of the results

Aside from the type of method chosen for metabolomics study (targeted or non-targeted), it is always desirable to evaluate to what extent that study can be generalised. It can be achieved, for example, by repeating the study on a different batch of sample.

A new batch of samples were therefore analysed to see if the results from different patient suffering from CKD can confirm our previous study. Moreover, since the LC-MS analysis was already performed on this batch, it enables to compare the GC-MS results with the LC-MS results, as a complementary technique. The comparison was performed for the known uraemic toxins with available standards.

Table 4.6: In this table the level of metabolites that changed after dialysis is presented.

Metabolite	t_R (min) of the metabolite in sample	Fold change between pre-dialysis and control	Fold change between post-dialysis and control
Myo-inositol	20.022	4.3	2
Threitol	13.413	10	6.1
Erythritol	13.523	15	5
Mannitol[1]	18.323	444.4	333
Arabitol	16.053	15	4.2

[1]*p*-value > 0.05

As can be seen in **Table 4.6**, the fold changes for these metabolites follow the same trend as the previous batch (see **Table 4.4**), i.e. significant reduction after dialysis. However in this study mannitol has *p*-value higher than 0.05 which makes any conclusion about the level in the body uncertain. The similarity of fold changes between post-dialysis and control in the two batches seems to be independent of the initial level of metabolites before dialysis (depicted as a fold change between pre-dialysis and healthy). Actually, it is not surprising by considering this fact that the content of metabolites in the body is in direct correlation with a lot of factors such as, diet, medication, age, stress, exercise, gender, etc.

Both of the studies confirm that this class of toxins (sugar alcohols) are removable by the dialysis membrane. The main issue is that their elimination from serum is not complete and it might be related to the number of hours a patient is on dialysis. Based on the fact that the kidney is the organ that continuously takes part in removal of the toxins in the body, probably the

longer hours on dialysis or increasing this procedure to overnight procedures done on daily plan can help to reach the goal of complete elimination of these water soluble uraemic toxins from patients with CKD.

On the other hand, regardless of the anomalous results for mannitol, one may come to this conclusion that the capacity of the dialysis membrane is limited to certain degrees, since, for example, the relative levels of erythritol were reduced by dialysis from 70 and 15 to 3 and 5 (see **Table 4.4** and **4.6**). So it is likely that further reduction cannot be achieved by extending the dialysis period(s). Further investigations are therefore required to get a firm conclusion.

4.4 Conclusions

Headspace SPME was found to be unsuitable for metabolic profiling of serum samples, due to the lack of volatility of metabolites in serum matrix. Analysis of non-volatile metabolites was successfully accomplished on samples treated with classic derivatisation procedures. Recognition of some of the metabolites was performed using RI and confirmed by injecting individual standards. Results revealed a general descent trends for studied metabolites in post-dialysis samples. Also a similar trend was observed for a new batch of samples with comparable levels for known uraemic toxins, which were also independent of the initial concentration of them in patients.

To get a picture of the level of whole metabolites, global metabolic profiling of the samples was also performed which revealed a significant reduction in the level of metabolites by dialysis treatment, although the level

of many of them is still higher than healthy. To what extent the level of metabolites is allowed to be higher than healthy without causing symptoms associated with CKD may be a focus of another study.

Chapter 5

Liquid chromatography-mass spectrometry (LC-MS) of uraemic metabolites

5.1 Introduction

LC-MS and related techniques, such as capillary LC and ultra-high performance LC (UHPLC), are currently recognised as the most widely used separation technique for metabolomics studies [10]. High compatibility of LC with biological samples, widespread availability together with the continuously emerging technologies such as capillary monolithic columns coupled to the high resolution and sensitivity of MS make this combination highly suitable for both targeted and global profiling of metabolites. While the technique is ideal for direct injection of some relatively simple biological samples such as urine, the required sample preparation for more complex samples such as serum is also well-established and can be performed with less difficulties than other separation techniques such as GC-MS, which requires, for example, a pre-derivatisation step, etc. [20, 103]. Also in comparison to GC, LC is not limited to the separation of thermally stable, volatile or underivatised compounds, enabling analysis of a much wider range of analytes than GC, and therefore is well suited for metabolomics studies. However, GC-MS offers the advantage of lower cost for instrumentation than LC-MS, and is preferred for targeted metabolomics [10].

To date, reversed-phase LC (RPLC) in gradient mode hyphenated with ESI in both positive and negative ion modes represents the most popular platform for LC-MS analysis of metabolites [10, 103]. For LC separation, conventional column formats (typically 2.1-4.6 mm i.d., 5-25 cm L, and 3-5 μm particle sizes) are typically used at flow-rates of 0.1-1.0 mL/min which usually needs to be splitted before introduction to MS. While very well suited for separation of compounds with low to medium polarity, RPLC of highly polar compounds (like some of the metabolites) is challenging. For such polar/ionic compounds separation under hydrophilic interaction (HILIC) mode is a suitable option [104, 105].

It is also shown that the scope of a metabolomics study can be widened by taking advantage of a combination of different separation techniques. In one study, RP-UHPLC-MS, HILIC-MS and GC-TOF-MS were employed for a comprehensive metabolomics study of urine samples from patients suffering from kidney cancer [106]. Both MZmine and XCMS platforms were employed for peak alignment and data processing resulted in more than 2000 peaks. Several metabolites with significant variations were detected which enabled discrimination between renal cell carcinoma patients and healthy controls.

In this study, RPLC-MS analysis of serum samples collected before and after dialysis from patients suffering from CKD were performed in both positive and negative ionisation modes and compared with a control group. Both targeted and global metabolic profiling studies were conducted. Monitoring fold changes in both water soluble and protein-bound known uraemic toxins in the former study, and establish a general profile of whole

metabolites in the latter study were aimed, employing XCMS platform for data processing.

5.2 Experimental

The general experimental details including chemicals and instrumentation are presented in Chapter 2. Specific experimental conditions are given in each of the figure captions.

5.3 Results and discussion

In order to get maximum resolution and accuracy, LC-Orbitrap-MS was used for sample analysis. The LC-MS method developed by Edwards, *et al*. for the metabolomics study in human plasma was used in this study [107]. Moreover, serum samples were pre-concentrated by using speed vacuum to get the highest possible sensitivity. Since only a few file formats including "netCDF", "mzXML", "mzData", and Agilent ".d" folders [108] are readable for XCMS, collected raw MS data were converted to "mzXML" employing "MS convertor" and then uploaded in online XCMS server. Once all data is uploaded in the stored datasets part, a job can be created. Since it is still not possible to do multi-class comparison, each group of samples (pre-dialysis and post-dialysis) were compared to control (healthy) group to avoid variations. XCMS (like other data-mining softwares) apply special algorithms that filter and bin the raw data and then detect the relevant features, i.e., ionic species assigned as a pair of m/z and retention time. The software then aligns and normalises the features found in the sample set. As an example, **Figure 5.1** demonstrates the overlaid TICs of 12 pre-dialysis and

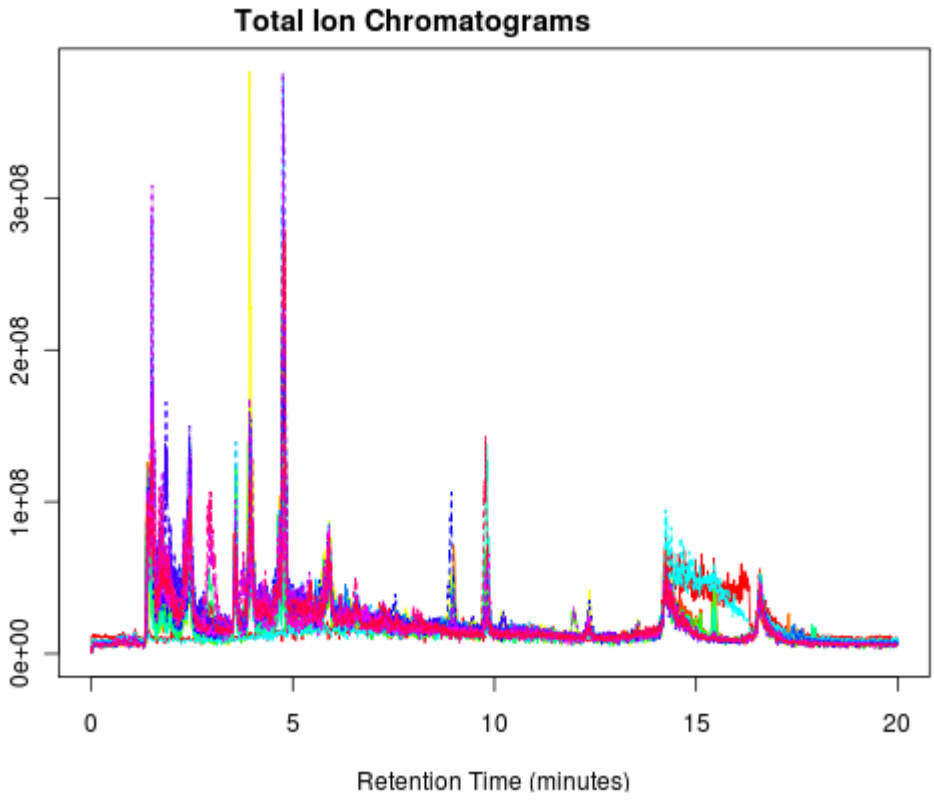

Figure 5.1: Overlaid TICs of pre-dialysis and post-dialysis samples aligned with XCMS (positive ion mode).

12 post-dialysis samples after peak alignment. Finally, it creates a peak table, which is actually a large data matrix, that is subjected to PCA and other statistical features [103].

The job can be submitted by just choosing one of the pre set parameters which is specified for each instrument. In this case HPLC/Orbitrap I was chosen. For this study the library error was set to 5 ppm to have a broader window of compounds. The processing time is very dependent to the size of the file varying from several minutes up to several hours. The out-of-range values were eliminated manually.

As a rule of thumb, any metabolite which the level of that is high in pre-dialysis compared to healthy can be a potential uraemic toxin. In addition, monitoring these toxins after dialysis can be a good indicator for the efficiency of dialysis membranes. Based on this hypothesis there are three possibilities:

Compounds that have very high intensity in patients compared to healthy and based on the library search they also exist in the list of known uraemic toxins.

Compounds that have high concentration in the patients compared to healthy, and there is hit for them in the METLIN library but do not exist in the list of known uraemic toxins; hence they hold the possibility of being new uraemic toxins.

For many compounds with significant fold changes (*eg.*, > 10) there is no hit in the METLIN or other metabolomics databases, and they do not exist in the list of known uraemic toxins. These metabolites are recognised as unknown. MS/MS is recommended to get some useful information about their structure and chemistry.

In this study we focus on the molecules which are in the list of known uraemic toxins, and they are either water soluble or protein-bounds in both positive and negative ion modes. In addition, metabolic profiling of small molecules (p-value < 0.05) was performed for both pre- and post-dialysis compared to healthy. The latter is very useful in getting a whole picture of metabolome changes in patients, for example, before and after getting a treatment.

5.3.1 Targeted study

5.3.1.1 Water soluble molecules

Processing of raw data produced a huge number of peaks; *ca*. 10,000 in positive ion mode; 7000 in negative ion mode. Filtering peaks with *p*-values higher than 0.05 using the "mirror plot" tool in XCMS reduced the number of peaks to around 6,000 and 4,400 respectively. The *p*-value is calculated from the t-statistics (Welch t-test, unequal variances) [108]. The targeted study was accomplished by searching for the known uraemic toxins in the XCMS results. To do so, the ionic species including protonated and deprotonated as well as sodium adduct forms of these metabolites were required. Also in this study, the less common ionic species, i.e., M+Na-2H was taken into account. Such features together with more detailed information like chemical structure and exact MW are accessible in the METLIN library for most known uraemic toxins. However, some of the toxins in the list of known uraemic toxins do not exist in the METLIN library or are known with their less common names, which makes looking for them even harder. Therefore, other metabolomics databases like Human Metabolome Database (HMDB; http://www.hmdb.ca/) were also searched for the rest of metabolites. Finally, for those known uraemic toxins that could not be found in any common databases the exact MW of the compounds and molecular ions were manually calculated based on the relevant chemical structures found in SciFinder (http://www.cas.org/products/scifinder).

Table 5.1: MS features of water-soluble, known uraemic toxins.

Compound	Chemical structure	MW	Ionic species			
			$[M+Na]^+$	$[M+H]^+$	$[M-H]^-$	$[M+Na-2H]^-$
1-methyladenosine	$C_{11}H_{15}N_5O_4$	281.1124	304.1016	282.1197	280.1051	302.0871
1-methylguanosine	$C_{11}H_{15}N_5O_5$	297.1073	320.0965	298.1146	296.1000	318.0820
1-methylinosine	$C_{11}H_{14}N_4O_5$	282.0964	305.0856	283.1037	281.0891	303.0711
ADMA	$C_8H_{18}N_4O_2$	202.1429	225.1322	203.1503	201.1357	223.1176
α-keto-δ-guanidinovaleric acid	$C_6H_{11}N_3O_3$	173.0800	196.693	174.0873	172.0728	194.0547
α-N-acetylarginine	$C_8H_{16}N_4O_3$	216.1222	239.1115	217.1295	215.1150	237.0969
Arab(in)itol	$C_5H_{12}O_5$	152.0684	175.0577	153.0758	151.0612	173.0431
Argininic acid	$C_6H_{13}N_3O_3$	175.0956	198.0849	176.1030	174.0884	196.0704
Benzylalcohol	C_7H_8O	108.0575	131.0467	109.0648	107.0502	129.0322
β-guanidinopropionic acid	$C_4H_9N_3O_2$	131.0694	154.0587	132.0768	130.0622	152.0441
Creatine	$C_4H_9N_3O_2$	131.0694	154.0587	132.0768	130.0622	152.0441
Creatinine	$C_4H_7N_3O$	113.0589	136.0481	114.0662	112.0516	134.0336
Cytidine	$C_9H_{13}N_3O_5$	243.0855	266.0747	244.0928	242.0782	264.0602

Dimethylglycine	$C_4H_9NO_2$	103.0633	126.0525	104.0706	102.0561	124.0380
Erythritol	$C_4H_{10}O_4$	122.0579	145.0471	123.0652	121.0506	143.0326
γ-guanidinobutyric acid	$C_5H_{11}N_3O_2$	145.0851	168.0743	146.0924	144.0779	166.0598
Guanidine	CH_5N_3	59.0483	82.0376	60.0556	58.0411	80.0230
Guanidinoacetic acid	$C_3H_7N_3O_2$	117.0538	140.0430	118.0611	116.0466	138.0285
Guanidonosuccinic acid	$C_5H_9N_3O_4$	175.0593	198.0485	176.0666	174.0520	196.0340
Hypoxanthine	$C_5H_4N_4O$	136.0385	159.0277	137.0458	135.0312	157.0132
Malondialdehyde	$C_3H_4O_2$	72.0211	95.0103	73.0284	71.0139	92.9958
Mannitol	$C_6H_{14}O_6$	182.0790	205.0683	183.0863	181.0718	203.0537
Methylguanidine	$C_2H_7N_3$	73.0639	96.0532	74.0713	72.0567	94.0387
Myoinositol	$C_6H_{12}O_6$	180.0633	203.0526	181.0707	179.0561	201.0381
N^2,N^2-dimethylguanosine	$C_{12}H_{17}N_5O_5$	311.1229	334.1122	312.1302	310.1157	332.0976
N^4-acetylcytidine	$C_{11}H_{15}N_3O_6$	285.0960	308.0853	286.1034	284.0888	306.0708
N^6-methyladenosine	$C_{11}H_{15}N_5O_4$	281.1124	304.1016	282.1197	280.1051	302.0871
N^6-threonylcarbamoyladenosine	$C_{15}H_{19}N_6O_8$	411.1265	434.1162	412.1343	410.1186	432.1006
Orotic acid	$C_5H_4N_2O_4$	156.0171	179.0063	157.0244	155.0098	176.9918

Orotidine	$C_{10}H_{12}N_2O_8$	288.0593	311.0486	289.0666	287.0521	309.0340
Oxalate	$C_2H_2O_4$	89.9953	112.9845	91.0026	88.9880	110.9700
Phenylacetylglutamine	$C_{13}H_{16}N_2O_4$	264.1110	287.1002	265.1183	263.1037	285.0857
Pseudouridine	$C_9H_{12}N_2O_6$	244.0695	267.0588	245.0768	243.0623	265.0442
SDMA	$C_8H_{18}N_4O_2$	202.1430	225.1322	203.1502	201.1357	223.1176
Sorbitol	$C_6H_{14}O_6$	182.0790	205.0683	183.0863	181.0718	203.0537
Taurocyamine	$C_3H_9N_3O_3S$	167.0364	190.0257	168.0437	166.0292	188.0111
Threitol	$C_4H_{10}O_4$	122.0579	145.0471	123.0652	121.0506	143.0326
Thymine	$C_5H_6N_2O_2$	126.0429	149.0321	127.0502	125.0357	147.0176
Uracil	$C_4H_4N_2O_2$	112.0272	135.0165	113.0346	111.0200	133.0019
Urea	CH_4N_2O	60.0323	83.0216	61.0396	59.0251	81.0070
Uric acid	$C_5H_4N_4O_3$	168.0283	191.0176	169.0356	167.0211	189.0030
Uridine	$C_9H_{12}N_2O_6$	244.0695	267.0588	245.0768	243.0623	265.0442
Xanthine	$C_5H_4N_4O_2$	152.0334	175.0226	153.0407	151.0262	173.0081
Xanthosine	$C_{10}H_{12}N_4O_6$	284.0756	307.0649	285.0830	283.0684	305.0504

Exact MW and calculated masses for some ionic species of all water-soluble known uraemic toxins have been summarised in **Table 5.1**. Due to the high mass resolution of Orbitrap-MS, the exact masses need to be calculated with four digits after decimal point to minimise the errors.

In the next step, all possible molecular ions listed in **Table 5.1** were manually searched in the processed results in XCMS (p-value < 0.05) in both positive and negative ion modes for both groups (pre-dialysis and healthy and post dialysis and healthy) by considering the m/z tolerance (mass error) of less than 5 ppm.

It is worth mentioning that for some of the features (ionic species) searched, the METLIN library revealed more than one hit with mass error < 5 ppm. Although unambiguous identification of a metabolite is only possible by MS/MS study of both sample and standard, only those hits existing in the list of known uraemic toxins were chosen as our targeted metabolites. In the meantime, the other hits might be considered as potentially new uraemic toxins, particularly when fold changes for their representative features are significant.

Table 5.2 summarised the identified water soluble, known uraemic toxins in positive ion mode with their corresponding fold changes and mass errors in two studied groups. As can be seen, for some compounds both the protonated form and sodium adducts are observed with different fold changes, which reflects the different tendency of the metabolite to the sodium ion or proton; however it should be noted that both ion forms should be retained at the same retention time.

Table 5.2: Water soluble, known uraemic toxins identified with LC-MS in positive ion mode.

No	Metabolite (protonated and sodium adduct)	Pre-dialysis & Healthy			Post-dialysis & Healthy		
		Fold change	Mass[1] error (ppm)	t_R (min)	Fold change	Mass[1] error (ppm)	t_R (min)
1	Urea, $[M+H]^+$	4.2	0	1.54	1.9*	0	1.49
2	Urea, $[M+Na]^+$	2.3	-4.8	1.51	1.4*	-3.6	1.44
3	Creatinine, $[M+H]^+$	8.0	-2.6	1.58	2.4	-2.6	1.55
4	Creatinine, $[M+Na]^+$	5.8	-2.2	1.58	2.2	-2.2	1.55
5	Guanidinoacetic acid, $[M+H]^+$	3.9	+4.2	16.51	3.0	+4.2	16.44
6	Benzylalcohol or *p*-cresol,$[M+Na]^+$	3.6	-1.5	7.2	3.4	-1.5	7.12
7	Phenylacetylglutamine, $[M+H]^+$	37.1	-0.3	4.83	10.8*	-1.1	4.74
8	Phenylacetylglutamine, $[M+Na]^+$	11.9	-0.3	4.83	5.6	-0.6	4.74
9	1-Methyladenosine or N^6-methyladenosine, $[M+Na]^+$	220.8*	+1.9	2.08	32.6*	+2.6	2.03
10	1-Methyladenosine or N^6-methyladenosine, $[M+H]^+$	2.4	-1.0	1.97	1.6*	-1.0	1.92
11	Xanthosine, $[M+Na]^+$	212.2	+0.9	1.58	23.6	+0.9	1.54
12	N^4-Acetylcytidine, $[M+H]^+$	5.4	-1.0	3.83	1.8*	-0.6	3.75

13	N^4-Acetylcytidine, $[M+Na]^+$	4.5	-0.6	3.83	1.8*	-0.6	3.75
14	N^6-Threonylcarbamyladenosine, $[M+H]^+$	14.1	0	4.79	3.4	-0.4	4.76
15	N^2,N^2 -Dimethylguanosine or 1,7-Dimethylguanosine, $[M+H]^+$	13.3	-0.9	4.06	6.7	-0.9	4.02
16	N^2,N^2-Dimethylguanosine or 1,7-Dimethylguanosine, $[M+Na]^+$	8.7	-1.1	4.06	6.4	-1.1	4.01
17	N^2,N^2-Dimethylguanosine or 1,7-Dimethylguanosine, $[M+Na]^+$	67.1	+2.0	2.98	24.5	2.3	2.93
18	1-Methylinosine,$[M+H]^+$	30.9	-0.7	3.57	12	-0.3	3.50
19	1-Methylinosine,$[M+Na]^+$	21.9	-0.3	3.57	12.3	0	3.57
20	Uridine or Pseudouridine, $[M+H]^+$	10	-1.6	1.67	3.3	-1.2	1.67
21	Uridine or Pseudouridine, $[M+Na]^+$	9.8	-1.8	1.68	N.D.	N.D.	N.D.
22	1-Methylguanosine, $[M+H]^+$	2.0*	+0.3	3.62	1.8*	-0.3	3.57

23	Creatine or β-Guanidinopropionic acid, $[M+H]^+$	0.4	-3	1.54	0.2*	-3.0	1.50
24	Creatine or β-Guanidinopropionic acid, $[M+Na]^+$	0.2*	-2.5	1.55	0.2*	-1.9	1.51
25	Erythritol or Threitol	9.6	-1.3	1.48	4	-1.3	1.43

* p-value < 0.05; for the other peaks p-value < 0.001 was obtained

[1] Mass errors were calculated manually as: $(\frac{measuredmass-theoreticalmass}{theoreticalmass}) \times 10^6$

Even in the negative ion mode, the molecular ions (deprotonated, etc.) for each metabolite should be observed at the same retention time as their molecular ions in the positive ion mode. This can be used as a very simple and useful hint to confirm the identified peak and to eliminate fake discovered hits in the library. As seen, a slight shift in the retention times also exist between two groups which is mostly less than 6 s.

Comparison of the results shows that most of the compounds had significant reduction post dialysis, which also reflects the quality of the dialysis treatment. Therefore, it may be concluded that increasing the dialysis time can reduce these toxins to the normal level. However, at least for metabolites 6 and 22 it seems to be unlikely, given that they had slight reduction after dialysis so probably the dialysis membrane is not efficient enough to remove them properly. Some unusual trends are also observed, for example, adducts for compounds 23 and 24 (which can be either creatine or β-guanidinopropionic acid) both had lower levels in patients compared to healthy. Also for compound 5, the error with guanidinoacetic acid is around

4 ppm which is high considering the accuracy of Orbitrap-MS and therefore, it is unlikely that this compound is guanidinoacetic acid. Metabolite 9 (which can be sodium adduct of either 1-methyladenosine or N^6-methyladenosine) shows the highest intensity before dialysis which is around 220.8 (sodium adduct) and even after dialysis it is still 32.6 times higher than the normal level. Whether this level can cause toxicity or not might be considered as the focus of another study. The same can be said for compound 22 either which shows almost unaltered change before and after dialysis. The rest of the compounds shown in **Table 5.2** had significant reduction after dialysis and their level in pre-dialysis is higher than post-dialysis with various level of removal by dialysis treatment.

Some metabolites have higher tendency to lose a proton rather than gaining one, so they have higher sensitivity in negative ion mode. As can be seen in **Table 5.3**, all the identified compounds showed a downward trend post dialysis. Comparison with positive ion mode (**Table 5.2**) reveals some common compounds, i.e., 1, 5, 7, 8, 9, and 10, that appeared at the same retention times. Metabolite 6 has a very high concentration before dialysis and almost two-thirds of it had been removed by dialysis and can be either uridine or pseudouridine.

The same possibilities also appeared for metabolite 5 but at a different retention time. On the other hand, these two compounds were also suggested in positive ion mode as numbers 20 and 21 (**Table 5.2**) but at the same retention time with metabolite number 5 (negligible shift).

Table 5.3: Water soluble, known uraemic toxins identified with LC-MS in negative ion mode.

No.	Metabolite (deprotonated and sodium adduct)	Pre-dialysis & Healthy			Post-dialysis & Healthy		
		Fold change	Mass error[1] (ppm)	t_R (min)	Fold change	Mass error[1] (ppm)	t_R (min)
1	Benzylalcohol, $[M-H]^-$	5.3	+0.9	7.19	3.9	+0.9	7.13
2	Uric acid, $[M-H]^-$	1.4	0	2.01	0.6	0	1.93
3	Orotidine, $[M-H]^-$	32.5	-3.1	6.09	25.2	-3.4	6.08
4	Argininic acid, $[M-H]^-$	3.3	-0.5	1.53	1.3*	-0.5	1.49
5	Uridine or Pseudouridine, $[M-H]^-$	9.5	0	1.71	3.7	0	1.67
6	Uridine or Pseudouridine $[M+Na-2H]^-$	157.6	-0.7	3.0	54.7	-0.3	2.90
7	Phenylacetylglutamin, $[M-H]^-$	22.3	-0.7	4.79	8.2	-0.7	4.75
8	N4-acetylcytidine, $[M-H]^-$	18	-0.3	3.82	3.6	0	3.74
9	N2-N2 dimethylguanosine, $[M-H]^-$	12.7	-0.3	4.07	7.0	-0.3	3.99
10	Creatine or β-Guanidinopropionic acid, $[M-H]^-$	0.4*	-0.7	1.59	0.3*	-0.7	1.55

* *p*-value < 0.05; for the other peaks *p*-value < 0.001 was obtained

[1] Mass errors were calculated manually similar to Table 5.2.

It simply means that only one of these two compounds was detected in positive ion mode but both in negative ion mode however, it is not possible to confidently differentiate between these two compounds without injecting standards for them. Compound number 10 in **Table 5.3** which also appeared in positive ion mode had lower concentration in both patients before and after dialysis compared to the healthy state.

Since XCMS is a batch processing platform, it might be interesting to see that manual measurement of fold changes for individual compounds also follows the same trend as those obtained by batch processing. **Figure 5.2** shows the mass spectra in positive ion mode for one of the identified known uraemic toxins, phenylacetylglutamine, in the same sample before and after dialysis. The protonated form of this compound is the base peak (relative abundance, 100%) in the pre-dialysis sample, whereas the same peak shows about 37% relative abundance post dialysis, which corresponds to approximately 63% reduction. This example shows a very good agreement with the batch processing study, i.e., ~ 70% (see **Table 5.2**). Also, the theoretical mass for the protonated form of this compound is 265.1183 and for its sodium adduct is 287.1002, which is only slightly different to the observed *m*/*z* in both samples (mass error < 1 ppm).

5.3.1.2 Protein-bound known uraemic toxins

Among uraemic toxins, protein bound toxins are very difficult to remove by conventional dialysis because they bind strongly to serum albumin and therefore they cannot pass through the dialysis membrane [7]. Organic anions, such as indoxyl sulfate (IS), 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF), *p*-cresyl sulfate (PCS), indoleacetic acid (IA),

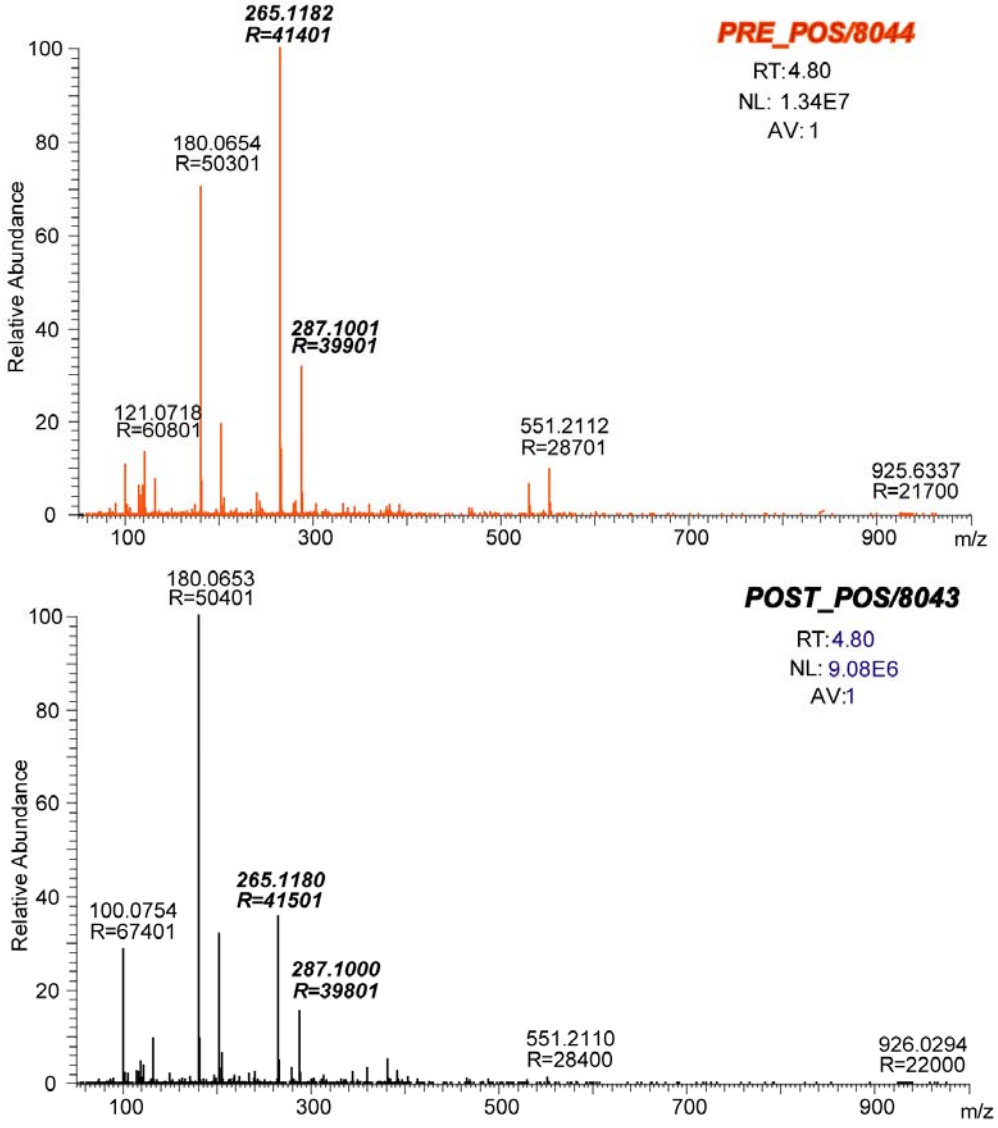

Figure 5.2: Positive ion mode MS spectra for phenylacetylglutamine in the same sample before (top) and after dialysis.

and hippuric acid are examples of protein bound uraemic toxins accumulating in kidneys and other tissues. Some of the cationic guanidino compounds are also linked to produce proinflammatory effect [109]. IS, PCS, HA, and IA can be partially removed by hemodialysis (HD) but removal of CMPF has been reported to be very difficult by hemodialysis [7]. In the case

of phenylacetic acid, the level of affinity to serum albumin and the efficiency of its removal by hemodialysis is still unknown [7]. Since these compounds are not fully bound to serum albumin (eg., 94% PCS and 93% IS [110]), hemodilaysis is only capable of removing the free fraction of them which is able to pass through the dialysis membrane [7, 111]. It is suggested that daily HD can minimise pre-HD concentration of some of protein bound toxins like indole-3-aceticacid, IS, *p*-cresol, and homocysteine [112, 113]. However, *p*-cresol can be removed more efficiently with hemodiafilteration when compared to high-flux HD [114].

It is known that during deproteinisation with organic solvents, biomarkers which are bound to protein are cleaved due to denaturation of the protein [20]. Also in previous studies, the role of employing different organic solvents, as a protein precipitation technique, in the efficiency of cleaving bounds from proteins have been investigated. An increase in the release of protein-bound solutes from their binding sites, which also results in higher number of detected peaks, was reported [115].

Due to the use of organic solvent in protein precipitation in this study, it is expected that some of the toxins bound to proteins become available in the supernatant together with the small percentage of free ones. The exact MW for this group of uraemic toxins has been searched from databases (mostly METLIN). Also for those toxins that METLIN or HMDB were unable to give enough information, the exact MW of the metabolites and related adducts have been calculated manually based on their chemical structure. Exact MW and calculated masses for some ionic species of protein-bound known uraemic toxins have been summarised in **Table 5.4**.

Table 5.4: MS features for protein-bound, known uraemic toxins.

Compound	Chemical structure	MW	Ionic species			
			$[M+H]^+$	$[M+Na]^+$	$[M-H]^-$	$[M+Na-2H]^-$
2-Methoxyresorcinol	$C_7H_8O_3$	140.0473	141.0546	163.0366	139.0401	161.0220
3-Deoxyglucosone	$C_6H_{10}O_5$	162.0528	163.0601	185.0420	161.0455	183.0275
CMPF	$C_{12}H_{16}O_5$	240.0998	241.1070	263.0890	239.0925	261.0744
Fructoselysine	$C_{12}H_{24}N_2O_7$	308.1584	309.1656	331.1476	307.1511	329.1330
Glyoxal	$C_2H_2O_2$	58.0054	59.0128	80.9947	56.9982	78.9801
Hippuric acid	$C_9H_9NO_3$	179.0582	180.0655	202.0475	178.0510	200.0329
Homocystein	$C_4H_9NO_2S$	135.0354	136.0427	158.0246	134.0281	156.0101
Hydroquinone	$C_6H_6O_2$	110.0368	111.0441	133.0260	109.0295	131.0114
Indole-3-acetic acid	$C_{10}H_9NO_2$	175.0633	176.0706	198.0525	174.0561	196.0380
Indoxyl sulfate	$C_8H_7NO_4S$	213.0095	214.0169	235.9988	212.0023	233.9842
Kynurenine	$C_{10}H_{12}N_2O_3$	208.0848	209.0921	231.0740	207.0775	229.0595
Kynurenic acid	$C_{10}H_7NO_3$	189.0426	190.0499	212.0318	188.0353	210.0173
Melatonin	$C_{13}H_{16}N_2O_2$	232.1212	233.1285	255.1104	231.1139	253.0958

Methylglyoxal	$C_3H_4O_2$	72.0211	73.0284	95.0103	71.0139	92.9958
N-(Carboxymethyl)lysine	$C_8H_{16}N_2O_4$	204.1110	205.1183	227.1002	203.1037	225.0857
p-Cresol	C_7H_8O	108.0575	109.0648	131.0467	107.0502	129.0322
p-Cresyl sulfate	$C_7H_8O_3S$	172.0194	173.0267	195.0086	171.0121	192.9941
Pentosidine	$C_{17}H_{26}N_6O_6$	378.2015	379.2088	401.1908	377.1943	399.1762
Phenol	C_6H_6O	94.0418	95.0491	117.0311	93.0346	115.0165
Phenylacetic acid	$C_8H_8O_2$	136.0524	137.0597	159.0416	135.0452	157.0271
p-Hydroxyhippuric acid	$C_9H_9NO_4$	195.0531	196.0604	218.0424	194.0459	216.0278
Putrescine	$C_4H_{12}N_2$	88.1000	89.1073	111.0893	87.0928	109.0747
Quinolinic acid	$C_7H_5NO_4$	167.0218	168.0291	190.0111	166.0146	187.9965
Spermidine	$C_7H_{19}N_3$	145.1578	146.1652	168.1471	144.1506	166.1326
Spermine	$C_{10}H_{26}N_4$	202.2157	203.2230	225.2050	201.2085	223.1904

In analysing the results, once again the *m*/*z* with error range up to 5 ppm has been searched and the fold changes for the components with *p*-value < 0.05 have been reported. For example, searching for hippuric acid in positive ion mode resulted in finding both of protonated and sodium adduct forms at two different retention times 4.80 and 16.44 (see **Table 5.5**). It is assumed that they are two different isomers of the same compound that have the same MW, yet different retention. While unambiguous identification is possible by using standards, this metabolite can also be differentiated from its isomer by considering the rationale that hippuric acid is a protein bound uraemic toxin and its removal is hampered by dialysis [7], therefore one can come to the conclusion that the one which shows minimal removal after dialysis should be hippuric acid (t_R = 16.44 min) and the other one (t_R = 4.80 min) to be its isomer.

For indoxyl sulfate (IS) both protonated and sodium adduct were found with very low error range, but they did not elute at the same retention time (one appeared at 1.22 and the other at 6.59 min) which means just one of them is possibly IS. Results show that the sodium adduct form, which is eluted at 6.59 min, demonstrates very limited removal post dialysis (only 0.4 reduction in fold change) and can be assigned as IS. Such a minor fold change can also be related to the free form of IS. Consequently, the other compound with the same MW to IS, which has been suggested as the protonated form of IS, should be another compound, possibly a new uraemic

Table 5.5: Protein-bound, known uraemic toxins identified with LC-MS in positive ion mode.

No	Metabolite (protonated and sodium adduct)	Pre-dialysis & Healthy			Post-dialysis & Healthy		
		Fold change	Mass error[1] (ppm)	t_R (min)	Fold change	Mass error[1] (ppm)	t_R (min)
1	Hippuric acid H^+	20.1	-0.5	4.80	11.0*	-0.5	4.72
2	Hippuric acid Na^+	5.9	-0.4	4.80	4.8*	-0.4	4.72
3	Hippuric acid H^+	15.6	-1	16.44	11.6	-1	16.38
4	Hippuric acid Na^+	2.8	-0.4	16.62	2.0	-0.4	16.57
5	Indoxyl sulfate H^+	108.7	+2.0	1.22	42.3	+2.0	1.01
6	Indoxyl sulfate Na^+	20.7	-1	6.59	20.3	-0.8	6.60
7	Hydroxy hippuric acid H^+	123.2	-1	4.05	35.2	-1	3.98
8	Hydroxy hippuric acid Na^+	63.7	-1	4.16	26.2	-0.9	4.11
9	Hydroxy hippuric acid H^+	46.7	-0.5	5.46	24.4	-1	5.41
10	Kynurenic acid H^+	16.9*	-1	5.39	11.5*	-1	5.31
11	Kynurenic acid Na^+	67.4	-0.4	5.42	42.8*	-0.9	5.34
12	Kynurenine H^+	1.5	-1	3.40	1.5	-0.9	3.34
13	CMPF H^+	5.2	-0.8	9.87	5.7*	-0.8	9.75
14	CMPF Na^+	3.3*	-1	9.87	3.6*	-1	9.75
15	CMPF H^+	8.9	-0.4	7.65	7.1	-0.4	7.54
16	CMPF Na^+	6.6	-1	7.56	4.7	-1	7.54

* p-value < 0.05; for the other peaks p-value < 0.001 was obtained

[1] Mass errors were calculated manually similar to Table 5.2.

molecule. This assumption is supported by the fact that the fold change for this compound significantly reduced from 108.7 to 42.3 times by dialysis, which is the characteristic of uraemic toxins or biomarkers.

There are two possibilities for hydroxy hippuric acid (HHA), one with both protonated form and sodium adduct which elutes at 4.05 min and, the other relates to just a protonated form which elutes at 5.46 min. Obviously, one of these two compounds can be HHA. Unlike aforementioned examples, results show that dialysis could eliminate a big percentage of both of them, so identification of HHA based on the previous rationale looks impossible. Meanwhile, significant fold changes once again suggest the possibility of introducing two new uraemic molecules. MS/MS study is suggested for such investigation since they appear at different retention times.

Kynurenic acid was eluted at 5.4 min with its both forms showing some reduction post dialysis. Kynurenine, on the other hand, showed no change after dialysis with the level of it in both pre and post-dialysis to be 1.5. Such a stable fold change may reflect the possibly of a very high percentage being protein-bound for this compound.

A somewhat more complicated situation can be observed for CMPF. Although the difference in the fold change for the compound eluted at 9.7 min is trivial, the difference for the other compound (t_R = 7.6 min) is not significantly high either, making it difficult to exclude this compound. As an inherent complication in metabolomics study, it should be kept in mind that minor differences between fold changes can be because of the contribution of analysis-related factors such as baseline fluctuation, variation in peak

Table 5.6: Protein-bound, known uraemic toxins identified with LC-MS in negative ion mode.

No.	Metabolite (deprotonated and sodium adduct)	Pre-dialysis & Healthy			Post-dialysis & Healthy		
		Fold change	Mass error[1] (ppm)	t_R (min)	Fold change	Mass error[1] (ppm)	t_R (min)
1	Hydroxy hippuric acid H	133.9	0	3.99	38.7	0	3.99
2	Kynurenic acid H	18.0	0	5.36	17.5	+0.5	5.34
3	Kynurenine H	2.8	0	3.44	2.1	0	3.33
4	CMPF H	4.1*	-0.4	9.72	4.4*	-0.4	9.73
5	CMPF +Na-2H	2.8*	-0.3	9.72	2.9*	-0.3	9.73
6	CMPF H	9.2	-0.4	7.62	7.5	0	7.53

* p-value < 0.05; for the other peaks p-value < 0.001 was obtained

[1] Mass errors were calculated manually similar to Table 5.2.

alignment, ion suppression etc. [103]. Therefore, the (minor) differences between test samples and control groups may be artificial rather than real.

As expected, fewer compounds were detected in negative ion mode; nevertheless, the identified compounds demonstrated more or less the same pattern as positive ion mode (see **Table 5.6**). In the case of hydroxyl hippuric acid there were two possibilities at two different retention times, 3.99 and 5.41 mins. However, the corresponding fold change for the latter was unorthodoxly 40% higher post dialysis (data not shown) which suggests exclusion of this hit. The other option eluting at 3.99 min in both positive and negative ion mode showed similar fold change patterns in both ion modes and reduced significantly by the treatment. Like positive ion mode, CMPF

suggested for two different retention times: 9.7 and 7.6 mins. Again while the peak at 9.7 min is more likely to be CMPF, the other option showed about 20% reduction by dialysis, a bit higher than in positive ion mode.

5.3.2 Global metabolic profiling in positive ion mode

While getting some information about the level of known uraemic toxins in the patients before and after dialysis treatment is highly important, as it enables, for example, to design more efficient dialysis membranes for better removal of "targeted" toxins, global metabolite profiling, on the other hand, provides a full picture of the general status of all metabolites. Such a picture when compared to healthy can be very useful in better understanding of a disease, for example, in monitoring the course of a treatment or even predicting the stage of the disease.

As mentioned before, processing the data with XCMS results in a massive load of data, which is partially reduced post statistical filtration. One way for better understanding of the results is by visualising them as some distribution profiles. Such profiles can be drawn after exporting the data in software packages like Origin (Northampton, MA, USA).

XCMS analysis of the data in positive ion mode resulted in about 6,000 peaks (p-value < 0.05), for which more information such as retention time, fold change, m/z, etc. was provided. Such information was exploited for redrawing graphical views. For example, **Figure 5.3** compares the distribution of fold changes against m/z between two studied groups (pre-dialysis *vs.* healthy and post-dialysis *vs.* healthy). This figure generally shows significant reduction of detected metabolites after dialysis.

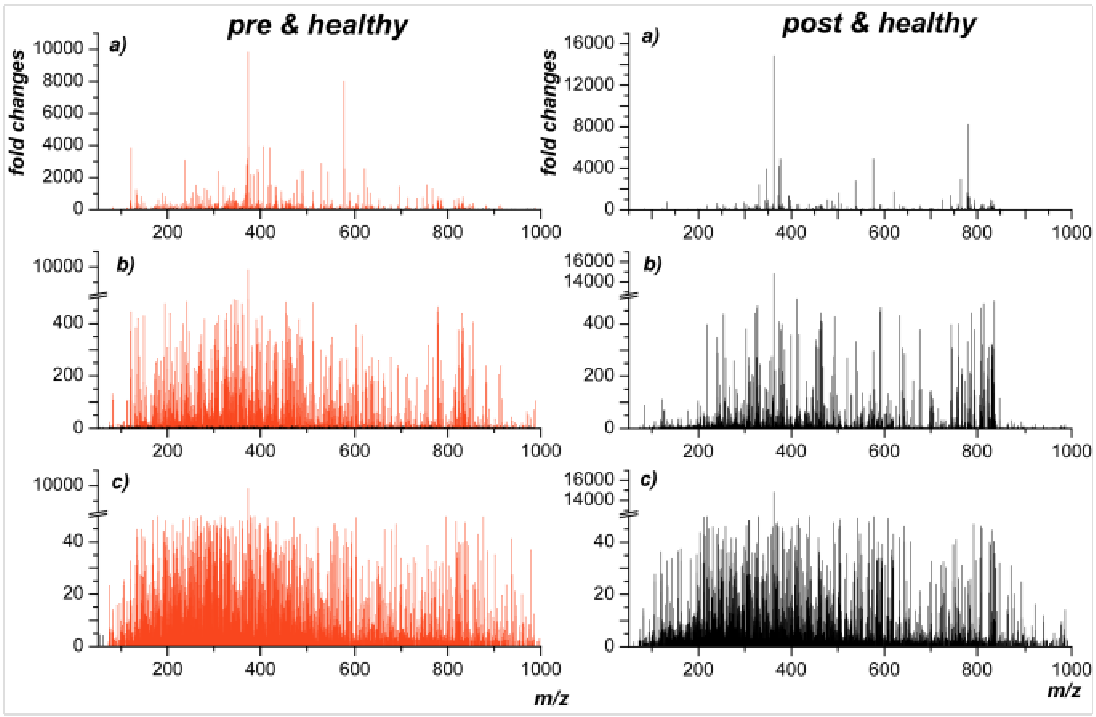

Figure 5.3: *m/z vs.* fold change plots in different scales comparing the metabolite distribution profiles pre- and post-dialysis.

Zooming in (b and c graphs) shows that metabolites had been more efficiently removed from *m/z* range of 100 to 500.

More precise estimation of the dialysis treatment can be obtained by viewing the percentage of metabolites before and after dialysis. The box-charts in **Figure 5.4** demonstrate such statistics for profiled metabolites. The right panel in **Figure 5.4** is actually a different viewing of **Figure 5.3** representing the mass distribution of profiled metabolites. Accordingly, 50% of metabolites were distributed in *m/z* range of 250-550, 25% in the relatively narrower range of 100-250 and the rest were between 550 to 950.

As can be seen in left panel, the majority (90%) of detected metabolites had relative fold changes less than 70 pre-dialysis which reduced to about 25

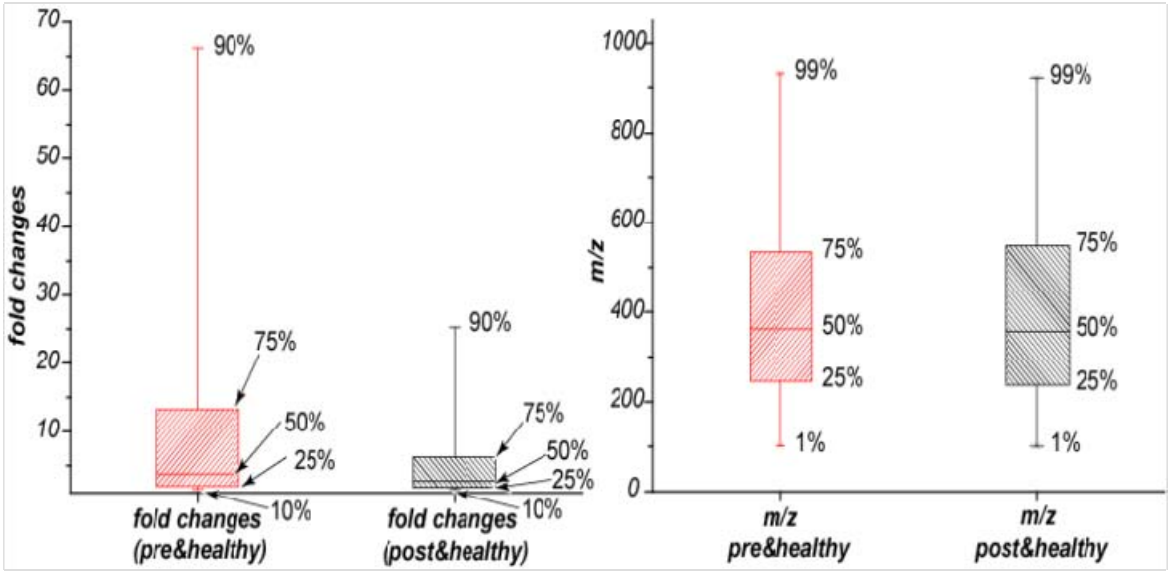

Figure 5.4: Box-charts representing fold change (left panel) and *m*/*z* (right panel) distribution percentages for profiled metabolites.

after treatment. Also, fold change distribution for 50% of metabolites before dialysis was approximately between 3 and 15 which was reduced to less than 5 post dialysis. On the other hand, treatment seems to have insignificant effect on fold changes of about 10% of metabolites. As one application, such information might be beneficial in better understanding of the performance of different dialysis treatments.

3-D viewing of the analysed results may also provide some more insights into the profiled population of metabolites**. Figure 5.5** shows density views for the distribution of metabolites before and after dialysis based on both fold changes and *m*/*z*. Although created from the same data set, these profiles can provide complementary information. For example, histograms on the bottom of each panel demonstrate the abundance of metabolites over the entire mass range as 50 evenly distributed bins, each covers 20 Da.

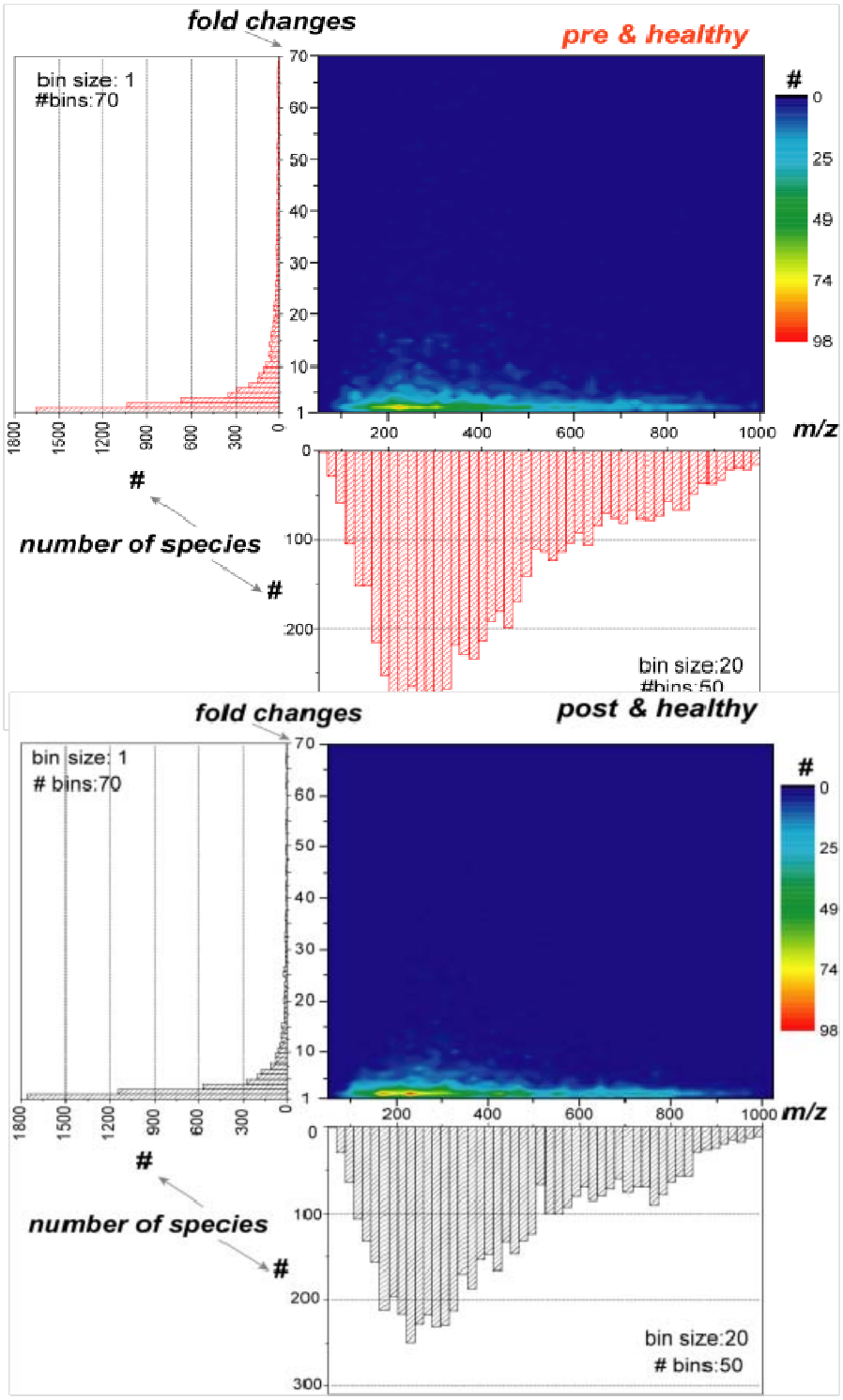

Figure 5.5: Density views and histograms of globally profiled metabolites before and after dialysis in comparison to healthy.

While identical distribution profiles were obtained for metabolites before and after from the corresponding box-charts (**Figure 5.4**, right panel), histograms revealed that along with differences in fold changes before and after dialysis the treatment also altered the mass distribution of metabolites as well. In other words, it suggests complete elimination of some species by the dialysis treatment. Such changes will be more visible if mass bins form one group are compared one-by-one to their counterparts in the other group. For example, comparison between bin 9 in two groups shows entire elimination of 50 species after the treatment (number of species 300 *vs.* 250). Also, such a significant elimination is more visible in *m*/*z* range of 200 to 400.

An overall picture for the performance of the treatment can be illustrated via 3-D, density views. **Figure 5.6** shows such views with emphasise on fold changes less than 40 for better viewing. As can be seen, the dialysis treatment has been more effective for the elimination of those metabolites in the *m*/*z* of 200-400 whose levels before dialysis had been 10 to 30 times higher than healthy group.

5.4 Conclusions

RPLC-MS analysis of serum samples collected from patients suffering from CKD was performed in both positive and negative modes. Water soluble and protein-bound known uraemic toxins were targeted in order to link their fold changes before and after dialysis to the symptom of the disease. While processing of the huge number of peaks generated (eg., 6000 in positive ion mode) using XCMS platform enabled monitoring the fold changes of several known uraemic toxins identified based on MS features

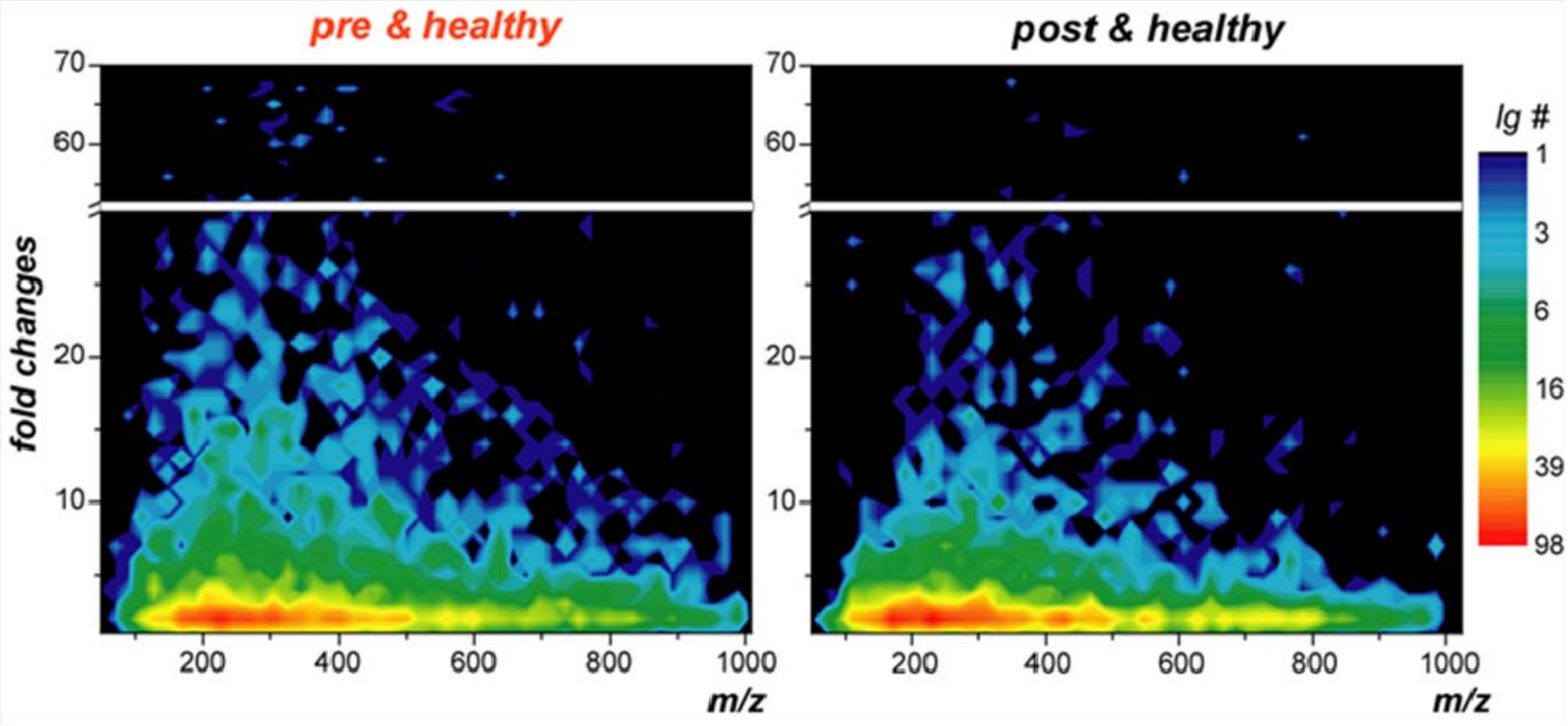

Figure 5.6: Magnified density views in Figure 5.5.

(such as ion species, mass error, etc.), absolute identification might be only possible by the analysis of authentic standards (if available) and study of the MS/MS features. Global metabolic profiling, on the other hand, provided a general picture of distribution and level of all detected metabolites. Such pictures clearly demonstrated partial or total removal of some of the metabolites upon dialysis treatment, which might be of value in monitoring the course of a treatment, comparison between, for example, different type of dialysis treatments or even in designing new membranes with complementary qualifications, in order to enhance the quality of dialysis treatments.

Chapter 6

General conclusion and future studies

Analysis of pre-dialysis and post-dialysis samples from patients suffering from CKD was successfully performed employing analytical techniques including CE-MS, GC-MS and LC-MS and the results were subsequently compared with a control (healthy) group. While good orthogonality between GC-MS and LC-MS for the analysis of respectively polar and non-polar metabolites was observed, CE-MS practically failed in providing reproducible results to be compared with the other two techniques.

Regardless of the method of analysis used, the level of many of the metabolites was found higher in patients especially before dialysis treatment. Post-dialysis the level of some of the metabolites significantly decreased in comparison to some with moderate or no change by dialysis. Although based on the list of known uraemic toxins, with only less than 100 compounds are known as uraemic toxins, it does not necessarily mean that the other compounds that have higher level in patients do not contribute to the effect of uraemic disease. From this point of view, metabolic profiling is a very useful study enabling a global view over three groups of samples for detecting changes, differences as well as similarities. Such a view is very useful for example in monitoring the course of a treatment or comparing the effect of different treatment in addressing the disease.

On the other hand, a targeted study enables monitoring known uraemic molecules and their possible contribution in uraemic symptoms. Therefore, targeted and non-targeted studies can provide complementary results helping to find the mechanisms of metabolite toxicity in the body and eventually curing of the disease by, for example, designing dialysis membranes for complete removal of the toxins from the body perhaps on a daily basis. Certainly, water soluble metabolites are much easier to remove with dialysis as also confirmed by these results. On the other hand, protein bound toxins are more sophisticated to eliminate as they are bound to protein and their removal can cause protein removal from the blood. Also their unambiguous monitoring is challenging and required consideration of effectors such as the efficiency of protein cleavage during sample preparation and also understanding about the ratio of protein bound to the free species.

In the following there are some recommendations that might be considered for future investigations:

- Using a higher number of samples in this kind of study can minimise the chance of variations and helps to get more certainty in statistical tests. Moreover, because of accumulation of the toxins due to kidney failure, many of these toxins can contribute in other diseases. Therefore, each of these diseases can produce specific metabolites and biomarkers interfering with the study. Based on clinical reports some of the samples were collected from the patients that have cancer in addition to their kidney disease. This will add other compounds and

causes significant variation in the results. Also, the different medications and treatments that patients received can possibly affect the metabolic pathways and since the kidney is not functioning properly to excrete them into urine they can show toxicity as well. Therefore for future studies, samples are recommended to be collected from a bigger population of each group preferably with much more similarities among the people in each group. The medications used should be similar and preferably it would be better to just focus on one disease rather than collecting samples from patients with different medical conditions who have the common kidney disease as well.

- In this study samples were collected before and after dialysis from patients treated with two types of dialysis procedure, i.e., HD and HDF. Possibly one might be more efficient in removing of toxins compared to the other one. If samples are collected from two groups of patients each treated individually with only one procedure, it is likely to find differences in the performance of each procedure when compared to control group and possibly the best dialysis treatment in removing more compounds.
- The analysis of dialysate wastes can be informative as well. A correlation between the reduction of metabolites post-dialysis and the accumulation of them in dialysate wastes is expected, and might be used to evaluate the reliability of this study.

Unfortunately, such a study was not possible to fit into the time frame of the current project.

- This study was aimed to monitor the changes in known uraemic toxins plus general metabolic profiling. It would be of value if identification of the metabolites that are not in the list of known uraemic toxins is also taken into account in future studies. This can be achieved by doing MS/MS and/or accessing individual standard of metabolites.
- Another part that can be a focus for future study is the analysis of protein bound uraemic toxins which are the most complicated metabolites. Since the focus of this study was not primarily on protein bound uraemic toxins and their measurement was not part of the plan, the sample preparation perhaps was not able to efficiently cleave most of them from the proteins. ACN is one of the organic solvent employed for cleaving protein bound uraemic toxins [115]. In some studies activated charcoal was employed for this purpose [73]. Hence, finding a suitable sample preparation which specifically targets cleaving proteins might be of value for future investigation as well.
- The analysis of middle molecules was one of the missing parts of the puzzle in this study and worth working on in the future. All the MS parameters in our study have been set to be suitable for small molecules. Different MS settings are better to be optimised for middle size molecules (see **Table 1.1**).

- Some compounds were seen to have higher levels in the control group in comparison to patients. This possibly means that these are the metabolites that normally should not be excreted into the urine but in patients with CKD the kidney fails to keep these metabolites in the body so they are released into the urine. Since the urine of these patients has been collected already, it is possible to monitor the trace of these metabolites to evaluate this assumption. However, if the level of them is not high in urine of patients it can be concluded that the level of these metabolites was originally low; probably due to the fact that the metabolic pathways lead to their production had been already affected and produced less of those metabolites in patients. Again, such information would also assist in gaining a complete picture of the underlying biological changes occurring.
- Last but not least, more comprehensive separation techniques such as two-dimensional LC (2D-LC), which has been proved before to be highly suitable for the analysis of complex matrices such as in proteomics, might be considered as an independent study for the current purpose. Due to the unsuitability of RPLC in separating highly polar solutes, the orthogonality of RPLC and hydrophilic interaction chromatography (HILIC) modes can be exploited in a 2D-LC set-up for achieving the highest possible peak capacity and separation efficiency that also facilitates the detectability of metabolites with enough confidence and sensitivity.

References

[1] Cohen, G., Glorieux, G., Thornalley, P., Schepers, E., Meert, N., Jankowski, J., Jankowski, V., Argiles, A., Anderstam, B., Brunet, P., Cerini, C., Dou, L., Deppisch, R., Marescau, B., Massy, Z., Perna, A., Raupachova, J., Rodriguez, M., Stegmayr, B., Vanholder, R., Horl, W. H., *Nephrol Dial Transplant* 2007, *22*, 3381-3390.

[2] Liabeuf, S., Barreto, D. V., Barreto, F. C., Meert, N., Glorieux, G., Schepers, E., Temmar, M., Choukroun, G., Vanholder, R., Massy, Z. A., *Nephrol Dial Transplant* 2010, *25*, 1183-1191.

[3] Chadban, S. J., Briganti, E. M., Kerr, P. G., Dunstan, D. W., Welborn, T. A., Zimmet, P. Z., Atkins, R. C., *J Am Soc Nephrol* 2003, *14*, S131-138.

[4] Niwa, T., *Mass Spectrom Rev* 2011, *30*, 510-521.

[5] Vanholder, R., Baurmeister, U., Brunet, P., Cohen, G., Glorieux, G., Jankowski, J., *J Am Soc Nephrol* 2008, *19*, 863-870.

[6] Good, D. M., Zurbig, P., Argiles, A., Bauer, H. W., Behrens, G., Coon, J. J., Dakna, M., Decramer, S., Delles, C., Dominiczak, A. F., Ehrich, J. H., Eitner, F., Fliser, D., Frommberger, M., Ganser, A., Girolami, M. A., Golovko, I., Gwinner, W., Haubitz, M., Herget-Rosenthal, S., Jankowski, J., Jahn, H., Jerums, G., Julian, B. A., Kellmann, M., Kliem, V., Kolch, W., Krolewski, A. S., Luppi, M., Massy, Z., Melter, M., Neususs, C., Novak, J., Peter, K., Rossing, K., Rupprecht, H., Schanstra, J. P., Schiffer, E., Stolzenburg, J. U., Tarnow, L., Theodorescu, D., Thongboonkerd, V., Vanholder, R., Weissinger, E. M., Mischak, H., Schmitt-Kopplin, P., *Mol Cell Proteomics* 2010, *9*, 2424-2437.

[7] Watanabe, H., Miyamoto, Y., Otagiri, M., Maruyama, T., *J Pharm Sci* 2011, *100*, 3682-3695.

[8] Lameire, N., Vanholder, R., De Smet, R., *Kidney Int Suppl* 2001, *78*, S292-297.

[9] Pasikanti, K. K., Ho, P. C., Chan, E. C., *J Chromatogr B Analyt Technol Biomed Life Sci* 2008, *871*, 202-211.

[10] Issaq, H. J., Van, Q. N., Waybright, T. J., Muschik, G. M., Veenstra, T. D., *J Sep Sci* 2009, *32*, 2183-2199.

[11] Fiehn, O., *Plant Mol Biol* 2002, *48*, 155-171.

[12] van der Greef, J., Stroobant, P., van der Heijden, R., *Curr Opin Chem Biol* 2004, *8*, 559-565.

[13] Nicholson, J. K., Wilson, I. D., *Nat Rev Drug Discov* 2003, *2*, 668-676.

[14] Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N., Willmitzer, L., *Nat Biotechnol* 2000, *18*, 1157-1161.

[15] Ramautar, R., Somsen, G. W., de Jong, G. J., *Electrophoresis* 2009, *30*, 276-291.

[16] Ramautar, R., Mayboroda, O. A., Somsen, G. W., de Jong, G. J., *Electrophoresis* 2011, *32*, 52-65.

[17] Lu, W., Bennett, B. D., Rabinowitz, J. D., *J Chromatogr B Analyt Technol Biomed Life Sci* 2008, *871*, 236-242.

[18] Glassbrook, N., Beecher, C., Ryals, J., *Nat Biotechnol* 2000, *18*, 1142-1143.

[19] Moco, S., Vervoort, J., Bino, R. J., De Vos, R. C. H., Bino, R., *TrAC Trends in Analytical Chemistry* 2007, *26*, 855-866.

[20] Kole, P. L., Venkatesh, G., Kotecha, J., Sheshala, R., *Biomed Chromatogr* 2011, *25*, 199-217.

[21] Maharjan, R. P., Ferenci, T., *Anal Biochem* 2003, *313*, 145-154.

[22] Desiderio, C., De Rossi, A., Inzitari, R., Mancinelli, A., Rossetti, D. V., Castagnola, M., Messana, I., *Anal Bioanal Chem* 2008, *390*, 1637-1644.

[23] Englard, S., Seifter, S., *Methods Enzymol* 1990, *182*, 285-300.

[24] Weigel, S., Kallenborn, R., Huhnerfuss, H., *J Chromatogr A* 2004, *1023*, 183-195.

[25] Bruno Kolb, L. S. E., *Static Head-space Gas Chromatography*, 2006.

[26] RESTEK, *A Technical Guide for Static Headspace Analysis Using GC*, obtained from: www.restek.com/pdfs/59895B.pdf.

[27] Lenz, E. M., Wilson, I. D., *J Proteome Res* 2007, *6*, 443-458.

[28] Villas-Boas, S. G., Mas, S., Akesson, M., Smedsgaard, J., Nielsen, J., *Mass spectrometry reviews* 2005, *24*, 613-646.

[29] Beckonert, O., Keun, H. C., Ebbels, T. M., Bundy, J., Holmes, E., Lindon, J. C., Nicholson, J. K., *Nat Protoc* 2007, *2*, 2692-2703.

[30] Wilson, I. D., Plumb, R., Granger, J., Major, H., Williams, R., Lenz, E. M., *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, *817*, 67-76.

[31] Mallet, C. R., Lu, Z., Mazzeo, J. R., *Rapid Commun Mass Spectrom* 2004, *18*, 49-58.

[32] McLuckey, S. A., Wells, J. M., *Chem Rev* 2001, *101*, 571-606.

[33] Chernushevich, I. V., Loboda, A. V., Thomson, B. A., *J Mass Spectrom* 2001, *36*, 849-865.

[34] Moco, S., Bino, R. J., Vorst, O., Verhoeven, H. A., de Groot, J., van Beek, T. A., Vervoort, J., de Vos, C. H., *Plant Physiol* 2006, *141*, 1205-1218.

[35] Brown, S. C., Kruppa, G., Dasseux, J. L., *Mass Spectrom Rev* 2005, *24*, 223-231.

[36] Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., Horning, S., *Anal Chem* 2006, *78*, 2113-2120.

[37] Haselberg, R., de Jong, G. J., Somsen, G. W., *J Chromatogr A* 2007, *1159*, 81-109.

[38] Monton, M. R., Soga, T., *J Chromatogr A* 2007, *1168*, 237-246; discussion 236.

[39] Guillo, C., Barlow, D., Perrett, D., Hanna-Brown, M., *J Chromatogr A* 2004, *1027*, 203-212.

[40] Johannesson, N., Olsson, L., Backstrom, D., Wetterhall, M., Danielsson, R., Bergquist, J., *Electrophoresis* 2007, *28*, 1435-1443.

[41] Landers, J. P., *Handbook of Capillary And Microchip Electrophoresis And Associated Microtechniques*, CRC Press 2008.

[42] Janini, G. M., Zhou, M., Yu, L. R., Blonder, J., Gignac, M., Conrads, T. P., Issaq, H. J., Veenstra, T. D., *Anal Chem* 2003, *75*, 5984-5993.

[43] Ramautar, R., Busnel, J. M., Deelder, A. M., Mayboroda, O. A., *Anal Chem* 2012, *84*, 885-892.

[44] Nilsson, S. L., Andersson, C., Sjoberg, P. J., Bylund, D., Petersson, P., Jornten-Karlsson, M., Markides, K. E., *Rapid Commun Mass Spectrom* 2003, *17*, 2267-2272.

[45] BrukerDaltonic, *micrOTOFQ User Manual*, obtained from: http://www.instrument.com.cn/show/literature/c111524.pdf.

[46] Mayboroda, O. A., Neususs, C., Pelzing, M., Zurek, G., Derks, R., Meulenbelt, I., Kloppenburg, M., Slagboom, E. P., Deelder, A. M., *J Chromatogr A* 2007, *1159*, 149-153.

[47] Soga, T., Ishikawa, T., Igarashi, S., Sugawara, K., Kakazu, Y., Tomita, M., *J Chromatogr A* 2007, *1159*, 125-133.

[48] Baidoo, E. E., Benke, P. I., Neususs, C., Pelzing, M., Kruppa, G., Leary, J. A., Keasling, J. D., *Anal Chem* 2008, *80*, 3112-3122.

[49] Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M., Nishioka, T., *J Proteome Res* 2003, *2*, 488-494.

[50] Hirayama, A., Kami, K., Sugimoto, M., Sugawara, M., Toki, N., Onozuka, H., Kinoshita, T., Saito, N., Ochiai, A., Tomita, M., Esumi, H., Soga, T., *Cancer Res* 2009, *69*, 4918-4925.

[51] Sugimoto, M., Wong, D. T., Hirayama, A., Soga, T., Tomita, M., *Metabolomics* 2010, *6*, 78-95.

[52] Sato, S., Soga, T., Nishioka, T., Tomita, M., *Plant J* 2004, *40*, 151-163.

[53] Toyohara, T., Akiyama, Y., Suzuki, T., Takeuchi, Y., Mishima, E., Tanemoto, M., Momose, A., Toki, N., Sato, H., Nakayama, M., Hozawa, A., Tsuji, I., Ito, S., Soga, T., Abe, T., *Hypertens Res* 2010, *33*, 944-952.

[54] Ramautar, R., Torano, J. S., Somsen, G. W., de Jong, G. J., *Electrophoresis* 2010, *31*, 2319-2327.

[55] Ramautar, R., Mayboroda, O. A., Deelder, A. M., Somsen, G. W., de Jong, G. J., *J Chromatogr B Analyt Technol Biomed Life Sci* 2008, *871*, 370-374.

[56] Arvidsson, B., Johannesson, N., Citterio, A., Righetti, P. G., Bergquist, J., *Journal of Chromatography A* 2007, *1159*, 154-158.

[57] Want, E. J., Nordstrom, A., Morita, H., Siuzdak, G., *J Proteome Res* 2007, *6*, 459-468.

[58] Márquez, A. J., Stougaard, J., *Lotus japonicus handbook*, Springer, Dordrecht 2005.

[59] Nikolau, B. J., Wurtele, E. S., *Concepts in Plant Metabolomics*, Springer 2007.

[60] Kopka, J., Fernie, A., Weckwerth, W., Gibon, Y., Stitt, M., *Genome Biol* 2004, *5*, 109.

[61] Karl Blau, J. M. H., *Handbook of Derivatives for Chromatography*, Wiley 1993 (2nd Ed.).

[62] Knapp, D. R., *Handbook of Analytical Derivatization Reactions*, Wiley 1979.

[63] Nishiumi, S., Shinohara, M., Ikeda, A., Yoshie, T., Hatano, N., Kakuyama, S., Mizuno, S., Sanuki, T., Kutsumi, H., Fukusaki, E., Azuma, T., Takenawa, T., Yoshida, M., *Metabolomics* 2010, *6*, 518-528.

[64] Ikeda, A., Nishiumi, S., Shinohara, M., Yoshie, T., Hatano, N., Okuno, T., Bamba, T., Fukusaki, E., Takenawa, T., Azuma, T., Yoshida, M., *Biomed Chromatogr* 2012, *26*, 548-558.

[65] Welthagen, W., Shellie, R., Spranger, J., Ristow, M., Zimmermann, R., Fiehn, O., *Metabolomics* 2005, *1*, 65-73.

[66] Tao, X., Liu, Y., Wang, Y., Qiu, Y., Lin, J., Zhao, A., Su, M., Jia, W., *Anal Bioanal Chem* 2008, *391*, 2881-2889.

[67] Zimmermann, D., Hartmann, M., Moyer, M., Nolte, J., Baumbach, J., *Metabolomics* 2007, *3*, 13-17.

[68] Visser, N. F. C., Lingeman, H., Irth, H., *Analytical and Bioanalytical Chemistry* 2005, *382*, 535-558.

[69] Ullsten, S., Danielsson, R., Backstrom, D., Sjoberg, P., Bergquist, J., *J Chromatogr A* 2006, *1117*, 87-93.

[70] t'Kindt, R., Alaerts, G., Vander Heyden, Y., Deforce, D., Van Bocxlaer, J., *J Sep Sci* 2007, *30*, 2002-2011.

[71] Coulier, L., Bas, R., Jespersen, S., Verheij, E., van der Werf, M. J., Hankemeier, T., *Anal Chem* 2006, *78*, 6573-6582.

[72] Tolstikov, V. V., Fiehn, O., *Anal Biochem* 2002, *301*, 298-307.

[73] Kikuchi, K., Itoh, Y., Tateoka, R., Ezawa, A., Murakami, K., Niwa, T., *J Chromatogr B Analyt Technol Biomed Life Sci* 2010, *878*, 1662-1668.

[74] Llorach, R., Urpi-Sarda, M., Jauregui, O., Monagas, M., Andres-Lacueva, C., *J Proteome Res* 2009, *8*, 5060-5068.

[75] Kristensen, M., Engelsen, S., Dragsted, L., *Metabolomics* 2012, *8*, 64-73.

[76] Denery, J. R., Nunes, A. A., Hixon, M. S., Dickerson, T. J., Janda, K. D., *PLoS Negl Trop Dis* 2010, *4*.

[77] Schwarz, E., Whitfield, P., Nahnsen, S., Wang, L., Major, H., Leweke, F. M., Koethe, D., Lio, P., Bahn, S., *Front Biosci (Elite Ed)* 2011, *3*, 308-314.

[78] Otter, D., Cao, M., Lin, H. M., Fraser, K., Edmunds, S., Lane, G., Rowan, D., *J Biomed Biotechnol* 2011, *2011*, 974701.

[79] Issaq, H. J., Nativ, O., Waybright, T., Luke, B., Veenstra, T. D., Issaq, E. J., Kravstov, A., Mullerad, M., *J Urol* 2008, *179*, 2422-2426.

[80] Kind, T., Tolstikov, V., Fiehn, O., Weiss, R. H., *Anal Biochem* 2007, *363*, 185-195.

[81] Issaq, H. J., Nativ, O., Waybright, T., Luke, B., Veenstra, T. D., Issaq, E. J., Kravstov, A., Mullerad, M., *J Urol* 2008, *179*, 2422-2426.

[82] Ott, M. A., Vriend, G., *BMC Bioinformatics* 2006, *7*, 517.

[83] Jason A. Day, J. A. C., *a J. Sabine Beckerb and Hans-Joachim Dietze, *Journal of Analytical Atomic Spectrometry* 2000, *15*, 1343-1348.

[84] Ghosal, S., *Electrophoresis* 2004, *25*, 214-228.

[85] Nehme, R., Perrin, C., Cottet, H., Blanchin, M. D., Fabre, H., *Electrophoresis* 2008, *29*, 3013-3023.

[86] Catai, J. R., Torano, J. S., de Jong, G. J., Somsen, G. W., *Electrophoresis* 2006, *27*, 2091-2099.

[87] AgilentTechnologies, *Installing the Agilent CE ESI-MS Sprayer Kit manual*

obtained from:
http://www.chem.agilent.com/Library/usermanuals/Public/G1600-90117.pdf.

[88] Nehme, R., Perrin, C., Cottet, H., Blanchin, M. D., Fabre, H., *Electrophoresis* 2009, *30*, 1888-1898.

[89] Kazarian, A. A., Hilder, E. F., Breadmore, M. C., *Analyst* 2010, *135*, 1970-1978.

[90] Katayama, H., Ishihama, Y., Asakawa, N., *Anal Chem* 1998, *70*, 2254-2260.

[91] Soga, T., *Methods Mol Biol* 2007, *358*, 129-137.

[92] Soga, T., Ueno, Y., Naraoka, H., Ohashi, Y., Tomita, M., Nishioka, T., *Anal Chem* 2002, *74*, 2233-2239.

[93] Vuckovic, D., Risticevic, S., Pawliszyn, J., *Angew Chem Int Ed Engl* 2011, *50*, 5618-5628.

[94] Chen, X., Xu, F., Wang, Y., Pan, Y., Lu, D., Wang, P., Ying, K., Chen, E., Zhang, W., *Cancer* 2007, *110*, 835-844.

[95] Miekisch, W., Fuchs, P., Kamysek, S., Neumann, C., Schubert, J. K., *Clin Chim Acta* 2008, *395*, 32-37.

[96] Cai, L., Koziel, J. A., O'Neal, M. E., *J Chromatogr A* 2007, *1147*, 66-78.

[97] Dettmer, K., Aronov, P. A., Hammock, B. D., *Mass Spectrom Rev* 2007, *26*, 51-78.

[98] Halket, J. M., Waterman, D., Przyborowska, A. M., Patel, R. K., Fraser, P. D., Bramley, P. M., *J Exp Bot* 2005, *56*, 219-243.

[99] Strehmel, N., Hummel, J., Erban, A., Strassburg, K., Kopka, J., *J Chromatogr B Analyt Technol Biomed Life Sci* 2008, *871*, 182-190.

[100] Birkemeyer, C. S., Max-Planck-Institute of Molecular Plant Physiology 2005.

[101] Roessner, U., Wagner, C., Kopka, J., Trethewey, R. N., Willmitzer, L., *Plant J* 2000, *23*, 131-142.

[102] Abate, S., Ahn, Y. G., Kind, T., Cataldi, T. R., Fiehn, O., *Rapid Commun Mass Spectrom* 2010, *24*, 1172-1180.

[103] Theodoridis, G., Gika, H. G., Wilson, I. D., *TrAC Trends in Analytical Chemistry* 2008, *27*, 251-260.

[104] Cubbon, S., Bradbury, T., Wilson, J., Thomas-Oates, J., *Analytical chemistry* 2007, *79*, 8911-8918.

[105] Horie, K., Ikegami, T., Hosoya, K., Saad, N., Fiehn, O., Tanaka, N., *Journal of chromatography. A* 2007, *1164*, 198-205.

[106] Garcia-Perez, I., Vallejo, M., Garcia, A., Legido-Quigley, C., Barbas, C., *Journal of chromatography. A* 2008, *1204*, 130-139.

[107] Registech, *GC Derivatization*, obtained from: http://www.registech.com/Library/gcderrev.pdf.

[108] Tautenhahn, R., Patti, G. J., Rinehart, D., Siuzdak, G., *Anal Chem* 2012.

[109] Schepers, E., Glorieux, G., Dou, L., Cerini, C., Gayrard, N., Louvet, L., Maugard, C., Preus, P., Rodriguez-Ortiz, M., Argiles, A., Brunet, P., Cohen, G., Jankowski, J., Jankowski, V., Massy, Z., Rodriguez, M., Vanholder, R., *Blood Purif* 2010, *30*, 277-287.

[110] Niwa, T., *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, *877*, 2600-2606.

[111] Jourde-Chiche, N., Dou, L., Cerini, C., Dignat-George, F., Vanholder, R., Brunet, P., *Semin Dial* 2009, *22*, 334-339.

[112] Nesrallah, G., Suri, R., Moist, L., Kortas, C., Lindsay, R. M., *Am J Kidney Dis* 2003, *42*, 13-17.

[113] Fagugli, R. M., De Smet, R., Buoncristiani, U., Lameire, N., Vanholder, R., *Am J Kidney Dis* 2002, *40*, 339-347.

[114] Bammens, B., Evenepoel, P., Verbeke, K., Vanrenterghem, Y., *Am J Kidney Dis* 2004, *44*, 278-285.

[115] Vanholder, R., Hoefliger, N., De Smet, R., Ringoir, S., *Kidney Int* 1992, *41*, 1707-1712.