NTNU – Trondheim
Norwegian University of
Science and Technology

Evaluation of Different MS-Based Methods for Urinary Metabolomic

Agnete Sion Evensen

Chemical Engineering and Biotechnology
Submission date: June 2012
Supervisor: Per Bruheim, IBT
Co-supervisor: Hans Fredrik N. Kvitvang, IBT

Norwegian University of Science and Technology
Department of Biotechnology

Problem description

Background

Mild forms of chronic kidney disease affects a large proportion of the population (10-13%). These patients are at risk of developing severe renal failure requiring dialysis or transplantation. Because there are medical treatments that can reduce or prevent progression to end stage renal disease, population screening is discussed increasingly. Development of new and more specific biomarkers of severe renal failure development will be important for early diagnosis. Mass spectrometric methods due to the high sensitivity and selectivity suitable for the detection of new biomarkers and screening of large sample numbers.

Method

The purpose of this master project is to: 1) establish and test the different mass spectrometric methods, and 2) perform a simple proof-of-concept study on selected samples of urine from patients with different degrees of renal failure.

Abstract

The diagnosis of chronic kidney disease (CKD) by examination of the urine has the potential to improve patients outcome by means of earlier detection. Due to the fact that the urine contains metabolic signatures for many biochemical pathways, this biofluid is ideal for metabolomics. A feature unique to diseases of the kidney is that the components of the kidney excrete urine. On the basis of this, analysis of urine have great potential for discovering new biomarkers for renal failure.

The aim of this study was therefore to compare urine samples obtained from CKD patients with healthy volunteers, in order to observe differences in metabolite concentration. Four different methods were applied for metabolite analysis. The three first methods used targeted analysis with gas chromatography coupled with single and triple quadrupole mass spectrometry and two different derivatization techniques were evaluated, alkylation and silylation respectively. The fourth method used untargeted analysis with hydrophilic interaction liquid chromatography coupled to a time-of-flight mass spectrometer. The combination of these techniques covers a large part of the urine metabolome by enabling detection of amino- and nonamino acids, sugars, sugar alcohols, purines, pyrimidines etc.

The first method identified 36 amino- and nonamino acids in the in-house library as well as finding one unidentified compound present in the samples. The second method identified 59 metabolites using silyaltion as derivatization techniques and identified metabolites which are not amino- and nonamino acids, hypoxanthane and uracil respectively. The third method identified 46 amino- and nonamino acid with absolute quantification. The fourth method using mass profiler professional for feature selection algorithm found 6 accurate masses higher represented in the CKD group, however later it was found that these masses were present in both groups.

The results from this study showed differences in metabolite concentration between the CKD group and the control group, where the excretion of almost all components into urine was decreased for the chronic kidney disease subjects. However, some compounds such as benzoate and proline were observed to be at higher concentration. Finally, the results were comparable with previous studies as well as observing metabolite variations between the two groups. However, there is still a long way to go before this can be applied in clinical settings. Future work needs to be performed on a larger group where the patients are with same diagnosis and off medications.

Sammendrag

Diagnosering av kronisk nyresykdom ved undersøkelse av urin har potensial til å bedre pasientenes prognose ved hjelp av tidligere deteksjon. Urin er ideelt for analyse av metabolitter, da det innholder mye informasjon om de biokjemise mekansimene i kroppen. En egenskap unik for sykdommer i nyrene er at komponentene i nyrene skiller ut urin. På bakgrunn av dette har derfor analyse av urin et stort potensial for å oppdage nye biomarkører under nyresvikt.

Målet med denne master oppgaven var derfor å sammenligne urinprøver fra nyresvikt pasienter med friske personer og observere eventuelle forskjeller i metabolitt konsentrasjon. Fire forskjellige metoder ble anvendt for analyse av metabolitter. De tre første metodene anvendte målrettet (target) analyse med gasskromatografi kombinert med enkel og trippel kvadrupol massespektrometer og to derivatisering teknikkene alkylering og silylering ble evaluert . Den fjerde metoden tok for seg vilkårlig (untarget) analyse med hydrofilisk interaksjons væskekromatografi koplet til et massespektrometer. Kombinasjonen av disse teknikkene dekker dermed en stor del av metabolittene i urinen ved å muliggjøre påvisning av amino-og organiske syrer, sukker, sukker alkoholer, puriner, pyrimidiner osv.

Den første metoden identifiserte 36 amino-og organsike syrer samt å detektere en uidentifisert metabolitt som var tilstede i alle prøvene. Den andre metoden anvendte silylering og identifiserte 59 metabolitter, hvorav noen ikke var amino-og organiske syrer som hypoxanthane og uracil. Den tredje metoden identifiserte 46 amino-og organiske syre ved absolutt kvantifisering. Den fjerde metoden anvendte et filtreringsprogram (Mass Profiler Proferssional) som oppdaget 6 nøyaktige masser som var høyere representert i nyresvikt pasientene. Senere ble det funnet at disse massene var til stede i begge gruppene.

Resultatene fra denne studien viser forskjeller i metabolitt konsentrasjon mellom nyresvikt pasienter og friske. Utskillelsen av nesten alle komponenter i urin ble observert til å redusere under nyresvikt, men det ble også observert at noen metabolitter som benzoate og prolin, hadde høyere konsentrasjon i pasientene. Resultatene ble funnet til å være sammenlignbare med tidligere studier, noe som indikerer gode analyse metoder. Det er fortsatt en lang vei å gå før dette kan brukes i kliniske settinger, da fremtidig arbeid må utføres på en større gruppe der pasientene er med samme diagnose og ikke tar medisiner under urine prøve innsamling.

Preface

This project is my master thesis and represents the final work in the master course Chemical Engineering and Biotechnology. This work was carried out at the Department of Biotechnology, Faculty of Natural Sciences and Technology at the Norwegian University of Science and Technology spring 2012.

A special thanks to my supervisor Associate Professor Per Bruheim for allowing me to work on this project, and for excellent guidance and encouragement. Especially I am thankful for the enthusiasm and care you show everyone of your students.

A warm thanks also to my co-supervisor PhD student Hans Fredrik Nyvold Kvitvang for assisting with the GC/MS part of this thesis, as well as advice when writing and revising. Your good humour and encouragement has been greatly appreciated.

I would also like to thank Stein Hallan and Marius Altern Øvrehus at St.Olavs Hospital for their participating in this project and for supplying necessary samples and information.

Finally I would like to thank my fellow students and members of staff with PhD student Stina Katrine Lien and Kåre Andre Kristiansen in particular.

Contents

Abbreviations

1D	One dimensional
2D	Two dimensional
5MS	5%-phenyl-methylpolysiloxane
ACN	Acetonitrile
AMDIS	Automated mass spectral deconvolution and identification software
AMP	Adenosine monophosphate
ANOVA	Analysis of variance
BP	Blood pressure
BPC	Base peak chromatography
CE	Charge exchange
CEF	Compound exchange forma
CHl	Chloroform
CI	Chemical ionization
CKD	Chronic kidney disease
CRF	Chronic renal failure
DB-5-ms + DG	5%-phenyl-methylpolysiloxane stationary phase + DuraGuard front column without stationary phase
eGFR	Estimated glomerular filtration rate
EI	Electro ionization
ESI	Electrospray ionization
ESRF	End stage renal failure
eV	Electron volt
FAME	Fatty acid methyl esters
GC	Gas chromatography
GC/MS	Gas chromatography coupled with mass spectrometry
HCl	Hydrochloride acid
HH-COSY	Hydrogen-hydrogen correlation spectroscopy
HILIC	Hydrophilic interaction chromatography
HMBC	Heteronuclear multiple bond correlation
HUNT	Helseundersøkelse i Nord-Trøndelag
HSQC	Heteronuclear single quantum coherence
ISTD	Internal standard
KEGG	Kyoto encyclopedia of genes and genomes
LC	Liquid chromatography
LC/MS	Liquid chromatography coupled with mass spectrometry
MeOH	Methanol

MCF	Methyl chloroformate
mmHg	Millimetres of mercury
MPP	Mass profiler professional
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSTFA	N-Methyl-N-(trimethylsilyl)-trifluoroacetamide
m/z	mass-to-charge ratio
NaOH	Sodium hydroxide
NCI	Negative chemical ionization
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
NP	Normal phase
PCA	Principal component analysis
PCI	Positive chemical ionization
PLS-DA	Partial least squares discriminant analysis
pmole	picomole (10^{-12}mole)
PP	Poly propylene
ppm	parts per million
RP	Reverse phase
QC	Quality control
QqQ	triple quadrupole
SEM	Standard error of means
SIM	Single ion monitoring
STD	Standard
SQ	Single quadrupole
TIC	Total ion chromatography
TMCS	Trimethyl chlorosilane
TMS	Trimethylsilylation
TOF	Time of flight
UI	International unit

1 Theory

Research within the human physiology and biology have achieved enhanced knowledge and understanding of diseases. Today treatment of diseases is able to save more lives than before, however several areas in human physiology is still a mystery and needs more research. One of these areas is the kidney and renal diseases, which is a major health problem all over the world [1].

1.1 The function of kidney

The main function of the kidneys is blood filtration and removal of waste products, toxins and excess water from the blood. Every day the kidney clears approximately 180 liters of fluids with a glomerular filtration rate (GFR) of 90 - 120 mL/min/1.73m^2 [2], whereas 1.5 liter turns into urine. This process is important for humans in order to consume different types of food, medications, vitamins and fluids without being concerned about accumulation of waste products. The kidney is crucial for extraction and regulating the amount of calcium, potassium and sodium in blood [1]. It also provides metabolic functions such as protein and peptide catabolism and gluconeogenesis [3].

Both kidneys contain around 106 nephrons, each consisting of the malpighian body and the tubule (Figure 1.1). The malpighian body consists of a tuft of capillaries (glomerulus) surrounded by a double-walled capsule known as Bowman's capsule. The blood enters the glomerulus by an arteriole (renal artery) and the primary urine accumulates in the area between the glomerulus and Bowman's capsule [3].

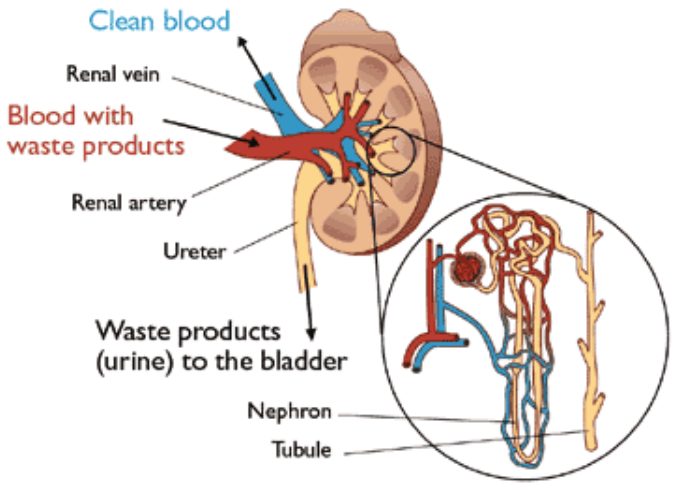


Figure 1.1: *The kidney and its main components [1]*

1.2 Chronic kidney disease and future diagnostic determination

Chronic kidney disease (CKD) is a progressive loss in renal function over a period of months or years [4]. The etiology of CKD is diverse and the diagnosis of CKD can be associated with diabetes, hypertension, glomerulonephritis, chronic interstitial nephritis, cystic disease and vasculities [4]. The function of the nephrons decreases and the filtration decreases, resulting in an increased concentration of waste products in the blood. Waste products and foreign compounds that are toxic are therefore not detoxified and not eliminated from the blood [3]. CKD is divided into 5 stages with decreasing renal function, where Stage 5 is the most serious level. Glomerular filtration rate (GFR) describes the flow rate of filtered fluid through the kidney. Stage 1 CKD is defined with normal or increased GFR (> 90mL/min/1.73m^2), Stage 2 is defined with a mild decrease (60-89mL/min/1.73m^2), Stage 3 CKD is defined as moderate (30-59 mL/min/1.73m^2), Stage 4 is defined as severe (15-29mL/min/1.73m^2) and at Stage 5 the GFR is less then 15 mL/min/1.73m^2 [2]. At this level kidney failure will occur and the patient is dependent on dialysis or kidney transplantation for survival [5].

The most common method to determine renal function and glomerular filtration rate is to determine of creatinine clearance. Creatinine is an end-product of muscle metabolism and is eliminated from the body by the kidneys. Creatinine is usually filtered out of the blood and into the urine. The normal range of serum creatinine concentration is between 1 and 2 mg per 100 mL and a clearance rate of approximately 125 mLmin^{-1} [6]. During kidney failure the plasma concentration of urea and creatinine increases as the filtration rate decreases. Determination of the plasma concentration of creatinine will therefore indicate the degree of kidney failure [1]. As the kidney function declines, the creatinine clearance will fall to 50-60 mL per min depending on age and other factors. The urea concentration is less trustworthy as the production of urea correlates with the intake of proteins [1].

Today patients with kidney disease are diagnosed after renal failure have occurred. The illness is unfortunately irreversible and degree of failure will increase during the following years. New and preventive diagnostic determination techniques are therefore an area that need more investigation. In this respect biofluid such as urine have gained a lot of interest the last 10 years. A feature unique to diseases of the kidney and urinary tract is that most of the components of this organ system are exposed to urine. Utilizing urine for biomarker discovery is therefore believed to have high potential [7]. Identification of patients at an early disease stage by early diagnosis of the specific renal disease enables improved therapeutic treatment [8]. Biomarker discovery with analysis of metabolites in urine from CKD patients is therefore receiving more attention. The recent development in metabolite analysis techniques and instrument development enables more sensitive and precise analysis [9]. Metabolomics has already been reported to be effective for the

discovery of biomarkers for cancer [10] and cardiovascular diseases [11], but the area of renal physiology and kidney disease needs more attention and investigation.

1.3 The increasing usage of metabolomics in nephrology

Figure 1.2 shows a general schematic of the "-omics" organization. The "-omics" organization consists of genomic, transcriptomic, proteomic and metabolomic. Genomics is a study of the genome of an organism. The genome is encoded either in DNA or, for many types of virus, in RNA. The study of transcriptomics examines the expression level of mRNAs in a given cell population, often using high-throughput techniques based on DNA microarray technology [12]. Proteomics is the large-scale study of proteins, particularly their structures and functions [13, 14]. Metabolomics monitors all metabolites and provides information about what is happening in the body, representing the downstream changes in the genome, transcriptome and proteome [15]. Measuring changes in metabolite concentration is a powerful approach for assessing gene function [16].

Metabolomics can be defined as the identification and quantification of all metabolites in a biological system [17]. Metabolites are low molecular weight compounds (less than 1500 Da) as sugars, organic acids and amino acids [18]. These are end-products of cellular processes, for example creatinine is an end-product of muscle metabolism.

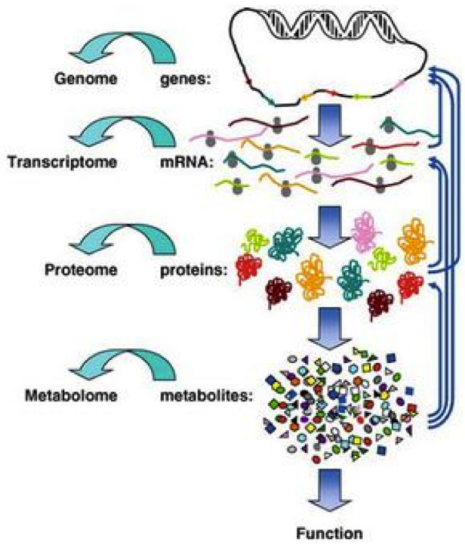


Figure 1.2: *The general flow of information goes from genes to transcripts of proteins and formation of metabolites and how metabolites can provide information about function (phenotype) [19]*

Urine is the biofluid which has received a lot of attention in metabolomics [8, 9, 20, 21]. It consists of by-products excreted from the body and these by-products provide information about cellular processes and changes in these processes. Urine has the additional

advantages of being easy to collect in large volumes and often without the need of excessive ethical/regulatory approval [22]. Previous as well as present studies, focus on the use of proteins and peptides in urine as promising biomarkers [14, 23, 24] . The last few years there has been rising numbers of publications in metabolomics. This demonstrates that metabolomics is an emerging tool to study phenotype and changes in phenotype caused by environmental influences, disease, or changes in genotype [17].

1.4 Metabolome measurement techniques

Because of the varying and diverse composition of urine (proteins, sugars, organic acids, amino acids, etc), metabolomic measurement techniques is very challenging and requires the simultaneous measurement of a large number of metabolites. A number of analytical techniques have been used in metabolomic studies such as Nuclear Magnetic Resonance spectroscopy (NMR) and Mass Spectrometry (MS) [25]. MS coupled with chromatography such as liquid chromatography (LC) or gas chromatography (GC) plays a leading role in this field. MS-based metabolomics offers quantitate analyses with high selectivity and sensitivity [17] hence rapidly and accurately measures the molecular weights and quantities wide variety of substances [16]. The LC/GC separate substances in a sample before injecting the sample into the MS. LC provides separation in liquid state of non-volatile compounds, while GC provides separation of volatile compounds. These measurement techniques are further enhanced by multiple methods of data analysis, ranging from metabolite quantification to multivariate statistical analysis [25].

Targeted and untargeted analysis

The MS-based approach are often divided into targeted and untargeted. As the name suggest, targeted methods are designed to detect and often quantify specific and interesting metabolites in a sample. This approach has the advantage of being highly specific and sensitive [26]. The gas chromatography mass spectrometry (GC/MS) offers targeted analysis of metabolites and the liquid chromatography mass spectrometry (LC/MS) offers untargeted analysis.

Untargeted analysis on the other hand aims to maximize coverage of metabolites, often compromising the sensitivity and specificity for any metabolites. These metabolomic approaches gives high mass accuracy data, but require more data analysis compared to the targeted approach [26]. The most common steps are presented in Figure 1.3. Metabolites are separated and analyzed using MS. The features of interest are selected from raw data using multivariate statistical approaches, and identification are performed with database searches. More details is represented later in Section 1.5.

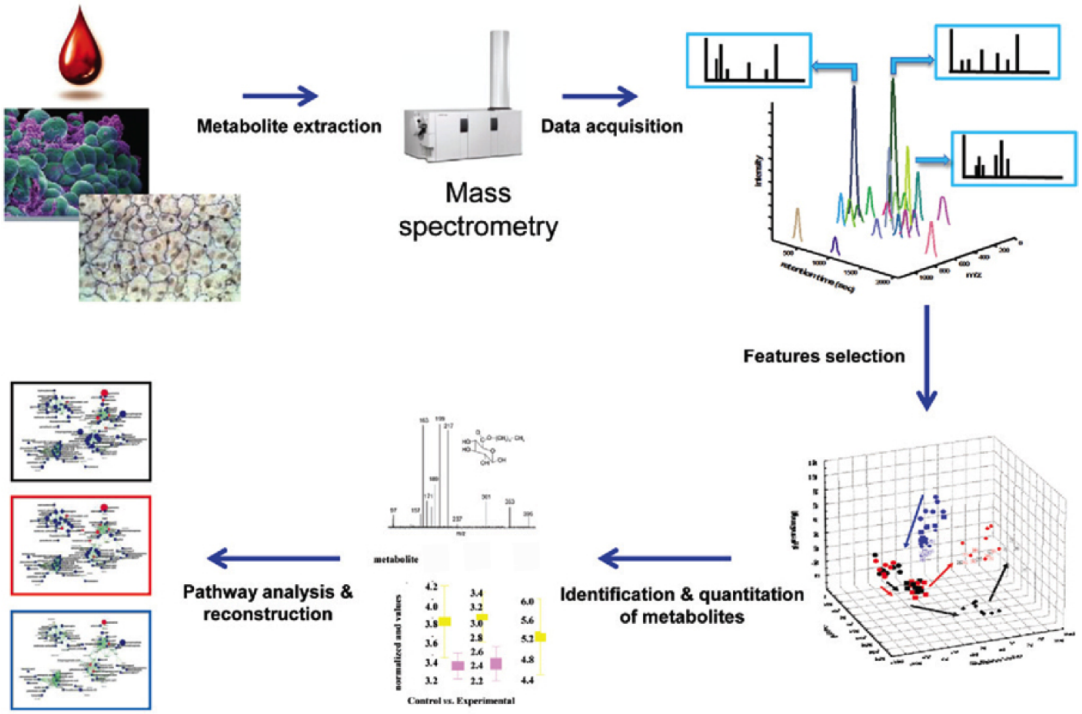


Figure 1.3: *Conventional untargeted metabolomic workflow [26]*

1.4.1 The basic of a gas chromatography - mass spectrometry system

Metabolomic application of GC/MS is considered as the golden standard and so far a common method in metabolomics. Figure 1.4 illustrates the setup for a GC/MS system from injection (with a syringe) to results (chromatograms and mass spectrum). The combination of GC and MS provides high-chromatographic metabolite resolution and analyte-specific detection [17]. The components are presented in the following sections.

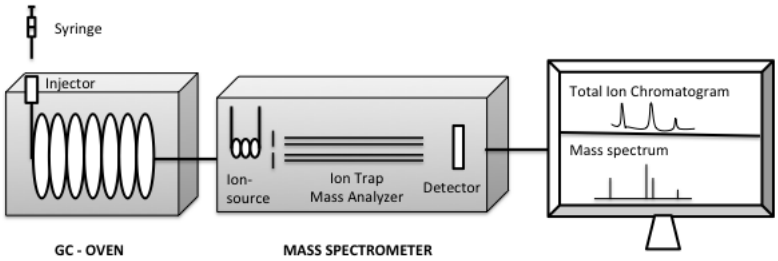


Figure 1.4: *The single quadrupole GC/MS system*

The gas chromatograph

The main purpose of chromatography is to separate a complex mixture of compounds into to discrete chromatographic peaks (Figure 1.5). In the GC/MS system, a sample (often liquid) of metabolites is introduced in the injector where it is evaporated and transferred

to the column by the carrier gas (Helium). The column is a capillary column with fused-silica that has a stationary phase coated on the inner surface. The stationary phase is consisting of a phenyl arylene polymer equivalent to a 5%-phenyl-metylpolysiloxane (i.e DB-5MS+DG). The analytes are separated in the column due to the temperature and compound interactions with the stationary phase of the column (known as partition chromatography). Each compound is characterized with a time called the retention time, which can be used for identification, since all compounds elute with a known retention time.

The compounds elute from the GC and enter the ion source where it is ionized before reaching the MS. The different ions produced in the ion source fragmentize as a consequence of the ionization, and are then separated due to their mass-to-charge (m/z) ratio in the mass analyzer and detected. Quantitative analysis is based on the correlation between the amount of substance that reaches the detector, and the response of the detector for the substance. Components passing the detector are being registered as peaks on the chromatogram. The peak height and peak area reflects the response of the detector [27].

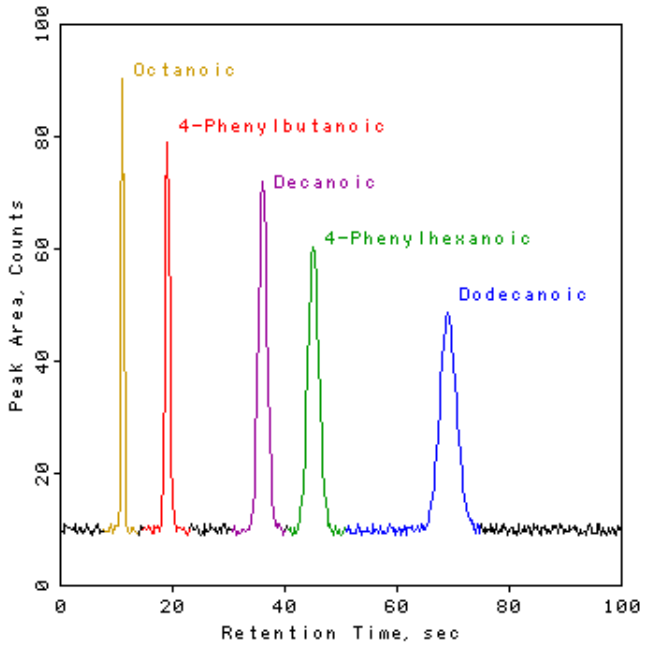


Figure 1.5: *Chromatogram for separation of compounds*

The MS and its three main constituents; ion source, mass analyzer and detector

The mass spectrometer (MS) contains three main consituents; the ion source, the mass analyzer and the detector. The molecules are ionized in the ion source and then fragmented into smaller pieces. The degree of fragmentation corresponds to the structure of

the molecules and the amount of energy during ionization. The masses of the ionized molecules can be determined by separation by their mass to charge values and detection on the detector [27].

The ione source

The ion source is the part of the MS where the molecules are being ionized. The ionization will give rise to both positive and negative fragments. Examples of ionization techniques are electro- and chemical ionization (EI,CI), charge exchange (CE), thermospray and electronspray. The two techniques used in this thesis is EI and CI. EI is the most common method to use in mass spectroscopy. The electrons arise from a filament and are projected as a radiant with an energy of 70 eV through the ion source. The molecules that are present in gas phase at a low pressure are ionized and form positive molecules (M^+). CI is a softer ionization technique. Here the molecules are being ionized by ion/neutral reactions with molecules and ions from a reagent gas. CI can be either be run in positive (PCI) or negative (NCI) mode. Only PCI will be used in this study. The most common reagent gases for PCI is methane and three different ionization processes can occur during PCI. These are proton transfer, hydride abstraction and addition. The proton transfer give *m/z* [M+1], hydride abstraction give *m/z* [M-1], while addition give [M+29] and [M+41] when methane is the reagent [27].

The Mass analyzer

As the name suggests, the mass analyzer analysis the masses in the MS after ionization. There are many different types of mass analyzers such as ion trap, time of flight (TOF), the quadropole- and the magnetic sector instrument. The main difference is how the masses of the components per charge (m/z) values after ionization, are being separated. However, this study will only focus on both the single quadrupole (sQ) and triple quadrupole (QqQ) MS, hence the presentation of the other mass analyzers is omitted.

Quadrupole MS

The single quadrupole is made by four cylindrical rods mounted parallel to each other at the corners of a square. The rods are imposed with direct current voltage U and an alternating voltage V=Vo(cos(ωt)). By varying the voltages and keeping the U/Vo relationship constant, ions with different *m/z* values will be led in-between the four rods and hit the detector one by one. The direct current voltage with a given modified voltage applied allows only ions with a certain *m/z* value to possess a stable trajectory and therefore are able to pass all the way to the detector. All other ions with different m/z values travel unstable, and will pass outside the quadrupole (i.e mass filtration) [27]. If

all ions are allowed to pass (i.e 50-550 m/z), the MS is run in scan mode, whereas if only certain ion masses are allowed to pass it is called single ion monitoring (SIM) [28].

Tandem MS

Tandem mass spectrometry, also known as MS/MS, involves two steps of mass spectrometry to examine selectively of the fragmentation of particular ions in a mixture of ions from first quadrupole. Two mass-analyzing quadrupoles are assembled in a tandem with a collision chamber in between them. When two quadrupoles are coupled in tandem in beeing used, it is called a triple quad (qQq) (Figure 1.6) [29].

The first quadrupole selects the precursor ion of interest. The precursor ion is then fragmented in the collision chamber (also referred to as a quadrupole), allowing all fragment ions (product ions) to be transferred into the third quadrupole. The third quadrupole performs mass analysis on the product ions that compose the tandem mass spectra and are rationalized to a structure [30].

One can lock the QqQ on one single ion created in the third quadrupole and only analyze this, and this is known as Multiple Reaction Monitoring (MRM) [29].

The Detector

The last and final component of the MS is the detector. Ions filtered from the MS are amplified in the electron multiplier and recorded on the ion detector. It records the charge induces as an ion hits a surface. In a scanning instrument, as the single quadrupole, the signal produced in the detector during scanning will produce a mass spectrum [27].

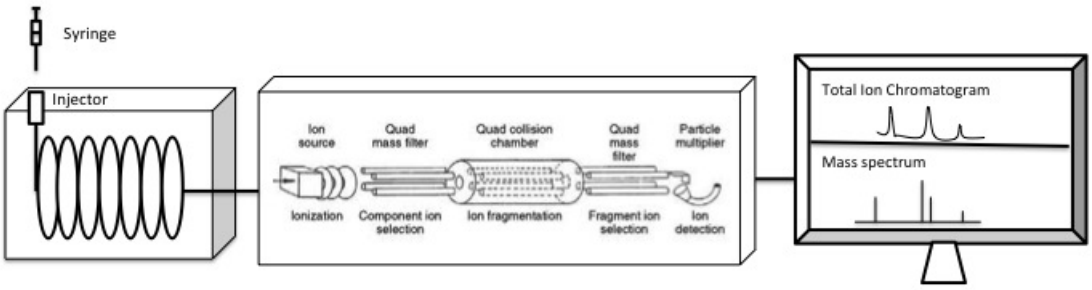


Figure 1.6: *The triple quadrupole GC/MC*

GC derivatization

To be able to analyze the metabolites in an urine sample with a GC/MS-system, it is important that the analytes are stable at the high temperatures [17]. The analysis of polar metabolites requires therefore a preparation technique called derivatization yielding volatile analytes. The derivatization step allows the functional group to reduce polarity and increase thermal stability and volatility. Active hydrogens in functional groups, such as -COOH, -OH, -NH and -SH can be derivatized by alkylation, acylation or silylation [17]. Silylation of organic compounds is the classical and most widely used derivatization procedure for metabolome analysis by GC-MS. N-methyl-N-trimethylsilyl-trifluoracetamide (MSTFA) is a reagent that can be used for trimethylsilylation (TMS). It reacts to replace labile hydrogens on a wide range of polar compounds with a -Si(CH3)$_3$-group [31], as shown in Figure 1.7-1.9. Often 1% trimethylchlorosilane (TMCS) is added to the reagent as a catalyst. Urea is the major organic constituent in human urine and will undergo derivatization. Removing urea with urease, renders the minor metabolites in urine samples accessible to derivatizing agents [16].

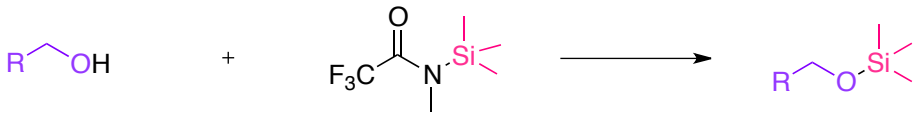


Figure 1.7: *Chemical derivatization of alcohol by silylation using N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA)[31]*

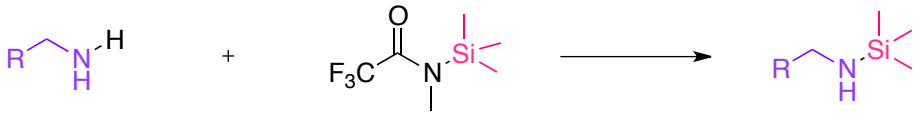


Figure 1.8: *Chemical derivatization of amine by silylation using N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA)[31]*

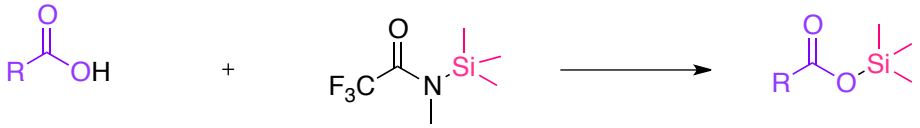


Figure 1.9: *Chemical derivatization of carboxylic acid by silylation using N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA)[31]*

Sugars and their derivatives are the class of metabolites most efficiently derivatized by silylation [32]. Some primary cell metabolites such as some organic- and amino acids produce relatively unstable silylated derivatives [33]. An alternative derivatization agent is methylchloroforomate (MCF). This method is primarily used for derivatization of poly-functional amines and organic acids as shown in Figure 1.10 and 1.11. The alkylation derivatization offers instantaneous reaction without heating or water exclusion, lower reagent cost, and easy separation of the derivatives from the reaction mixture, which causes less damage to the GC-capillary column [31].

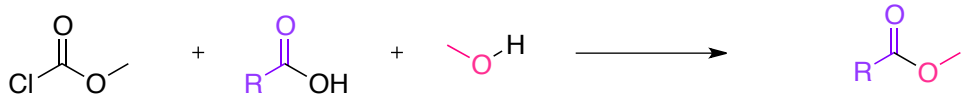


Figure 1.10: *Chemical derivatization of carboxylic acid by alkylation using methyl chloroformate (MCF) with methanol solvent [31]*

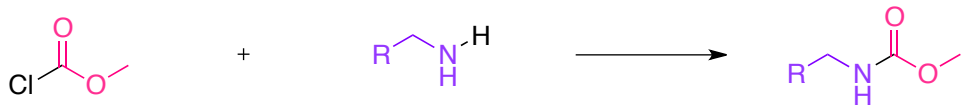


Figure 1.11: *Chemical derivatization of amine by alkylation using methyl chloroformate (MCF) with methanol solvent [31]*

1.4.2 The basic of a liquid chromatography - mass spectrometry system

A different approach in quantitative and qualitative analyses of complex samples can be performed with LC/MS. Similar to GC/MS, LC/MS separates compounds before they are introduced to the ion source and mass spectrometer. LC is used for non-volatile and thermo-labile components, however it requires a good ionization method [27].

Many LC/MS studies only use reversed phase (RP) chromatography, which instantly discriminates against polar compounds. Since urine is aqueous and the content is highly polar, silica-based hydrophilic interaction chromatography (HILIC) columns have been receiving more attention [34]. HILIC is analogous to normal-phase (NP) chromatography in that it utilizes a polar stationary phase [35]. However, unlike NP, HILIC allows the use of aqueous solvents. In direct contrast to RP, gradient elution HILIC begins with a low-polarity organic solvent and elutes polar analytes by increasing the polar aqueous content. Compound are retained by partitioning into a water rich layer, which is partially immobilized on the stationary phase [34]

The HILIC contains a mobile phase which has high content of water-miscible organic solvent. This will promote hydrophilic interactions between the analytes and a water-enriched hydrophilic stationary phase. This technique is also well suited for on-line coupling with electro spray ionization (ESI) [22]. ESI it the most common ionization method

and is performed at atmospheric pressure, generating either positive or negatively charged ions [27].

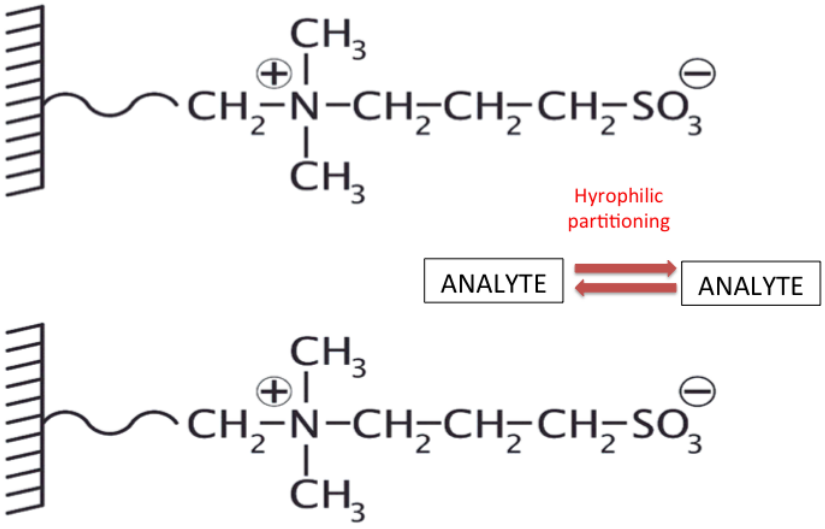


Figure 1.12: *Schematic representation of the retention process on sulfaolkylbetaine zwitterionic stationary phase [22]*

1.5 MS data analysis and identification of metabolites

As mentioned earlier MS data analysis and identification can be performed by targeted and untargeted methods.

1.5.1 Targeted analysis

Target analysis is based on identification of the metabolites retention time and mass spectrum of known metabolites incorporated in a in house library.

GC-SQ-MS data acquisition and identification of the compounds

GC-SQ-MS employs semi-quantitative analysis, yielding results in respons (peak area). Identification of metabolites can be performed by their respective retention time, and the mass spectra they produce when ionized and fragmented. Each compound fragments in a specific way, and the mass spectrums can be found in metabolite libraries. When a sample contains many compounds, the MS data obtained is compared to the mass spectrum in a library.

In some cases where two compounds co-elute, the mass spectrum do not match any compound within the library. This is because of the ratio of unfamiliar ions. In such

cases deconvolution can separate signals from different components present in total ion chromatogram (TIC) (Figure 1.13). One of the tools performing spectra deconvolution and application of retention indices is the Automated Mass Spectral Deconvolution and Identification System (AMDIS) available from National Institute of Standards and Technology (NIST) [17]. A compound is identified when four of its target and qualifier ions are recognized and matched with ions in the library. The match factor is then reported if it is over 60% [36]. Different retention times designate different compounds, and the abundance of a compound is correlated with the area below the peak at the responding time [37].

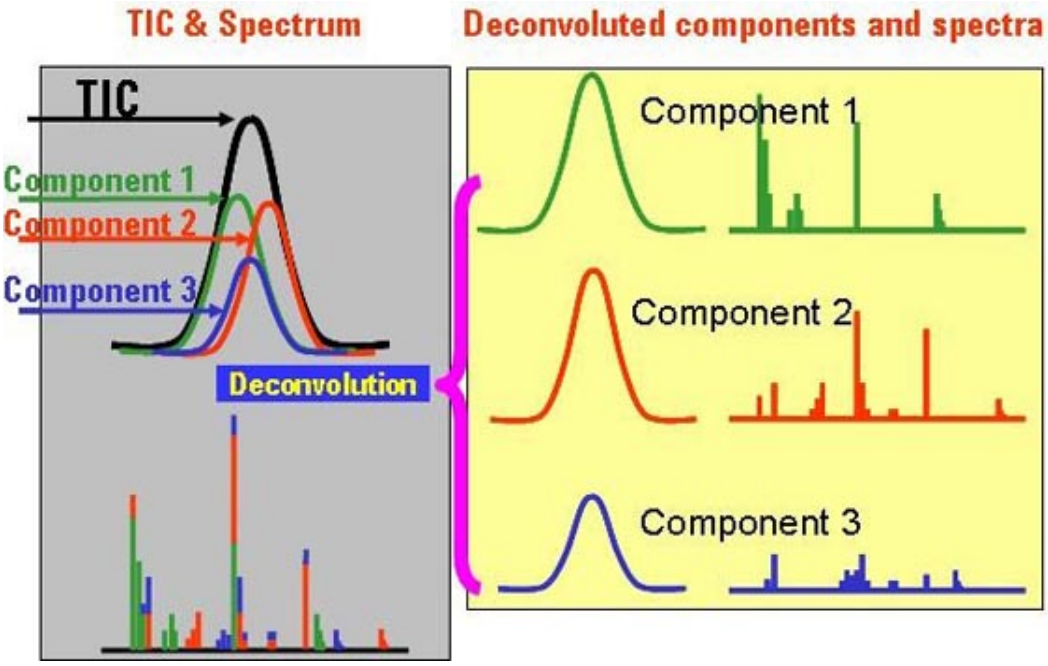


Figure 1.13: *Deconvolution of the TIC by AMDIS. The TIC is the sum of the intensities in all spectrums. AMDIS uses an algorithm to separate which signals originate from which compound, known as deconvolution [36]*

GC-QqQ-MS data acquisition and identification of the compounds

Urine samples analysis by GC/MS/MS is performed with absolute quantification, yielding results as concentrations (μmole/μL injected). During sample preparation the samples are spiked with a deuterated standard mixture. As the MS is operating in MRM mode, metabolites are detected based on the co-elution with the deuterated metabolites with same retention time. The novel method comprises approximately 80 common primary metabolites which mostly belong to the groups of amino and non-amino organic acids [30].

Statistical analysis

During this study, urine samples from CKD patients and controls are evaluated. Urine contains a wealth of information of the physiological condition in a subject. When evaluating the urine samples by metabolite analysis, it is expected to get large data sets as there are many compounds present in urine. To achieve better understanding and visualization of the data results, statistical method such as t-testing, principal component analysis (PCA), analysis of variance (ANOVA) and partial least-square-discriminant analysis (PLS-DA) can be applied.

This thesis aims at comparing a control group and a CKD group and the use of statistical methods will contribute in to evaluating statistical significant differences between these groups.

Student t-distribution

The most common method for comparing quantitative data between two groups is to perform a t-test. Students t-distribution is a statistical method applied for assessing the statistical significance of difference between two sample means, when the sample size is small and population standard deviation is unknown. The best estimate is calculated by the standard deviation (1) of the two groups, and estimation of the uncertainties in the average values with the standard-error-of-means (SEM), is given by equation (3). Student-t value is the significant criteria with a degree of freedom value of 2n-2 [38].

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{1}$$

$$SEM = \bar{\sigma_x} = \frac{\sigma_x}{\sqrt{N}} \tag{2}$$

$$t = \frac{\text{Difference between group average}}{\text{Standard deviation in the groups}} = \frac{\bar{x}_{patient} - \bar{x}_{control}}{\sqrt{\bar{\sigma}^2_{patient} - \bar{\sigma}^2_{control}}} \tag{3}$$

Analysis of variance (ANOVA) can be applied when testing several groups and is a generalization of the t-test. It is a typical statistical approach for comparing variance within each group and variance between each group [38].

$$F = \frac{\text{Variations between the groups}}{\text{Variations within the groups}} \tag{4}$$

As this study aims at comparing only two groups Student's t-distribution is the method of choice.

Principal component analysis (PCA)

The resulting data can be analyzed using the software Unscrambler, which will present the data as principal component analysis plot (PCA). PCA is an orthogonal transformation of multivariate data and is mostly used for exploratory analysis by extracting and displaying systemic variations. PCA is a projection method and attempt to uncover structures by building principal components and describing the maximal variance of the data [39]. The loading plot and score plot provides the relationships among the samples and reveals groupings and trends. For example, Figure 1.14 shows grouping of the triangles to the left and the circles to the right, and the loading plot suggests that the variable X_1 is contributing to clustering of the triangles.

The software also provides one plot know as influence plot (Figure 1.15). This plot shows all of samples and is mainly used for distinguishing samples that might affect the PCA-model. Samples that are high in residual variance and leverage can be considered as outliers and can be eliminated for further study.

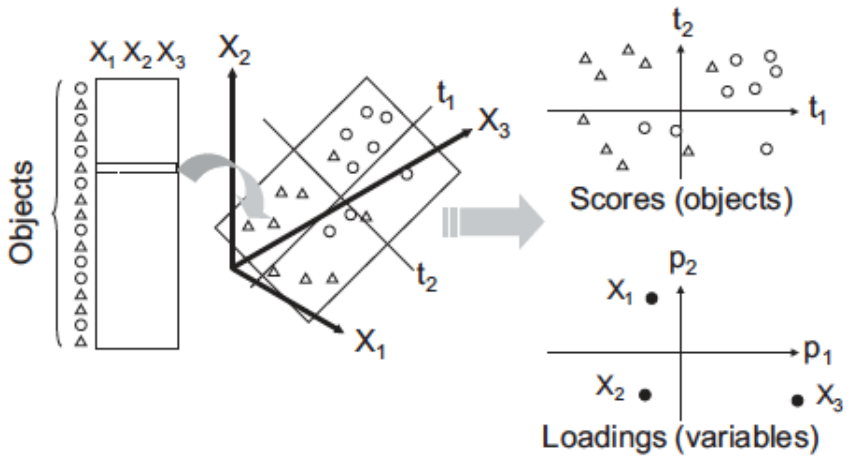


Figure 1.14: *Principal component analysis plot [39]*

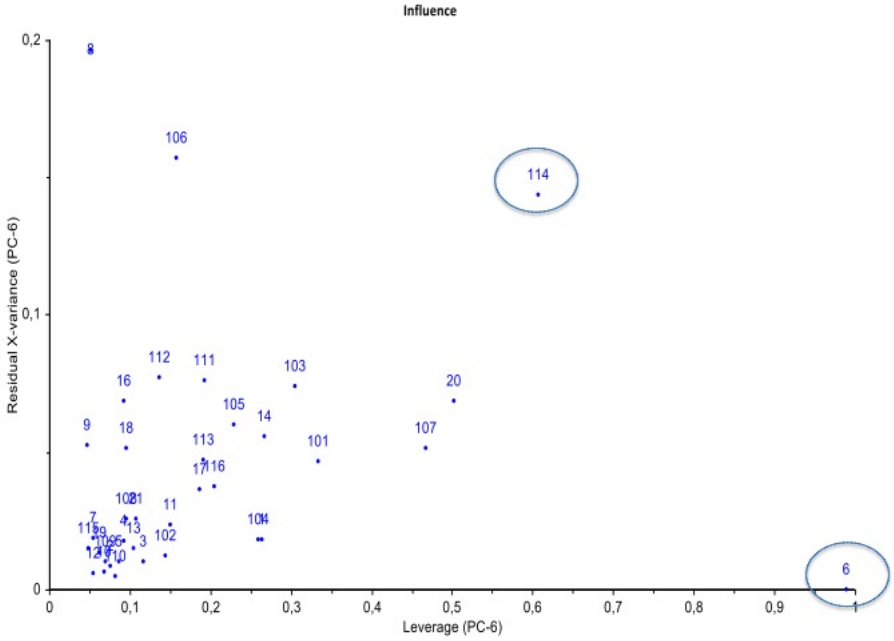


Figure 1.15: *Influence plot showing possible outliers, 114 and 6 respectively*

1.5.2 Untargeted analysis

LC/MS data handling and identification of compounds

Currently, handling of LC/MS data are not performed by automatic peak detection and retention time index as it is for GC/MS. The data handling is mainly performed based on mass accuracy, molecular feature, statistical analysis and identification using databases such as Metlin and KEGG.

Mass Profiler Professional

Mass Profiler Professional (MPP) is a statistical analysis software designed for mass spectrometry data in metabolomics and proteomics. MPP imports CEF-files created by MassHunter Qualitative Analysis and performs a step-by-step filtration of large data sets. The filtration steps are based on steps such as flags, frequency and abundance. Filter Flags allows filtration of input data based upon present-, marginal- and absent-flags, declaring that compounds must appear a minimum of times across all samples. Filter Frequency allows filtration based upon the frequency with which any compound appears in each sample in the experiment. Filter Abundance allows rejection of low intensity masses. The overall procedure is that a total masses of for example 5000 masses can be filtered down to 9 masses, and allows identification of compounds based on the masses. The program is well suited for establishment of relationships between different samples, and provides a great tool for metabolites investigation. An overall workflow is presented in Figure 1.16 below.

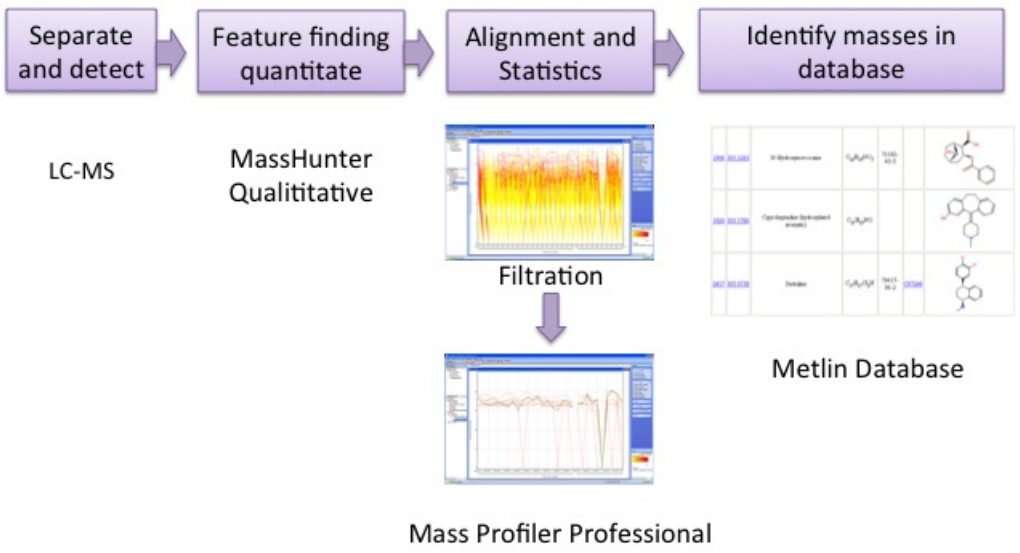


Figure 1.16: *Workflow from separation to identification of metabolites*

Database; Metlin, KEGG

Present, a retention index system for LC is yet to be established so the metabolomic libraries aims primarily at the molecular feature identification of metabolites [17]. Although a number of genomics and proteomics databases are available, metabolomics databases are still limited. Biochemical databases can be used for identification of the structure or biological function and are available i Kyoto Encyclopedia of Genes and Genomes (KEGG) databases [17]. The Metlin database catalogues metabolites, MS/MS spectra and LC/MS profiles of human plasma and urine samples [40] and provides a database with 55 000 metabolites [41].

1.6 Scope of this study

The purpose of this master thesis was to evaluate and test different MS-based methods (targeted and untargeted) for urinary metabolomics and to perform a simple proof-of-concept study on urine samples collected from CKD patients and healthy subjects.

The first method performed targeted analysis on ten samples with GC-SQ-MS and MCF-derivatization was employed. The method also aimed at searching for unidentified metabolites which could be added to the in-house library. The second method employed MSTFA-derivatization on all samples and analysis with GC-SQ-MS. The third method was performed with absolute quantitative analysis with GC-QQQ-MS and MCF-derivatization was employed. The fourth method performed untargeted analysis with LC-HILIC-MS with subsequent filtration in Mass Profiler Professional (MPP). Finally, the data results were visualized in principal component analysis plot (PCA plot).

2 Material and Method

Urine was collected in the morning 12.02.10 from colleagues regarded as "healthy" and samples from CKD patients labeled HUNT (Helse undersøkelse i Nord Trøndelag), giving a total of 37 urine samples. Figure 2.1 shows all urine samples, where 21 represents the patient group and 16 representing the "healthy" group. The patients were labeled with number 1-21 and the control group is labeled with number from 101 to 116. The CKD group contains subjects of different age and gender. Each person are diagnosed with different kidney disease (Nephrosclerosis, Interstitial nephritis, Alport syndrome etc) as well as other conditions as obesity, diabetes, heart failure and hypertension to mention some. The urine samples collected from colleagues are also of different age, gender and were regarded as healthy at the urine collection date two years ago.

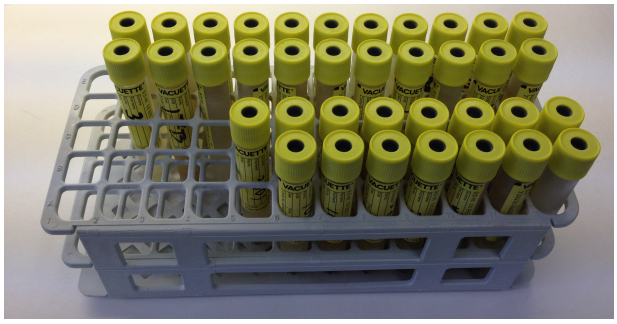


Figure 2.1: *Urine samples from 21 CKD patients and 16 controls*

A simplified workflow is presented in Figure 2.2. The first step includes sample collection and storage of the sample between each study,-32°C. The next step is sample preparation which includes urease degradation, precipitation of proteins with methanol and derivatization. GC-SQ-MS data analysis is performed with semi-quantitive analysis by integration in Chemstation and mass spectrum deconvolution in AMDIS. GC-QqQ-MS data analysis is performed with absolute quantification. The final step is exporting the data and analyze the results with statistical programs such as Unscrambler and MPP, yielding visualization of results in PCA-plot.

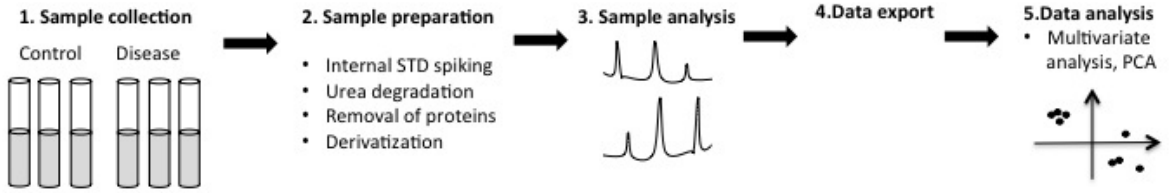


Figure 2.2: *Simplified workflow*

2.1 Alkylation with MCF for metabolite analysis on GC/MS

2.1.1 Preparation of urine samples

100 μL of matrix fluid was mixed with 400 μL methanol in poly propylene (PP) tubes (75 mm x 11mm) and incubated on ice for 30 minutes. The sample was then centrifuged at max speed (5000 rpm) for 10 minutes. The supernatant was transferred to a new PP tube and concentrated with a vacuum concentrator (Speedvac-Savant SPD20110).

2.1.2 MCF derivatization of urine samples for analysis on GC-SQ-MS

The MCF derivatization protocal was developed by Villas Bôas and co-workers [42]. The check sequence used consisted of chloroform, FAME, blank (1 M NaOH), 10 mM d5-glutamate (20 μL), std1 and std2 (40 μL), where as the latter three was derivatized. The steps during the MCF derivatization are presented in Figure 2.3. d5-glutamate was used as an internal standard during derivatization. For more details about the mixtures see Appendix A.2 and A.3.

20 μL 10 mM d5-glutamate, 380 μL 1 M NaOH, 333 μL MeOH and 67 μL pyridine was added to the samples and vigorously mixed. 80 μL of MCF was then added and the sample was mixed for 60 seconds. 400 μL chloroform was added and mixed for 10 seconds. 400 μL of 50 mM $NaHCO_3$ was then added and mixed for 10 seconds.

The chloroform-phase (lower phase) was transferred into a new clean PP tubes and 2-3 spatula of dried $NaSO_4$(s) and mixed for 5 seconds. The water-free chloroform phase was then transferred to GC/MS vials with insert and analyzed on the GC/MS.

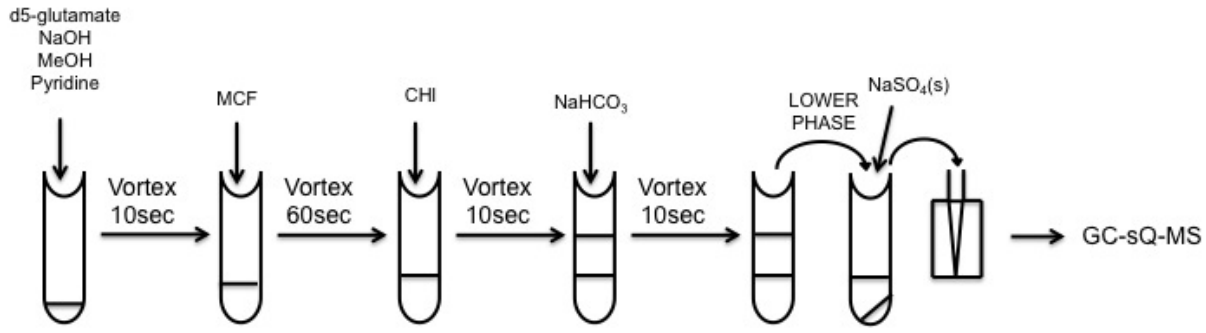


Figure 2.3: *MCF derivatization before metabolomic application of single quadrupole*

2.1.3 GC/MS instrumentation and data analysis

Samples were run on a GC/MS system (Agilent 7890A series GC system coupled with Agilent 5975 Mass Selective Detector). The carrier gas used was Helium gas. 2 μL was injected in pulsed splitless mode (split ratio 50:1). The GC was equipped with a J&W Scientific DB-5MS+DG column (30 m, 250 μm inner diameter, 0.25 μm film thickness). The GC oven was kept at 45 °C for 2 minutes before raised to 300 °C using a linear gradient of 10 °C/ min for 7.5 minutes resulting in a total run time of 35 minutes. The MS operated in scan mode (start after 6 min, mass range 50-550 m/z at 2.5scans/s) and utilizing EI source operating at 70 eV. d5-glutamate was used as standard for retention time locking of the method.

The data collected, was obtained using Agilent MSD Productivity ChemStation. The data set was loaded to ChemStations QEdit function, both automatic and manually integrated ion chromatogram up against a MCF-library. The spectra obtained were also compared to library spectrums available in NIST Automated Mass Spectral Deconvolution and Identification Software (AMDIS). AMDIS separates ion chromatogram signals in a complex TIC's, creating pure compound peak free of overlapping signals (Deconvolution) [36].

After identification, the results were normalized against d5-glutamate to correct for variations in the run.

2.1.4 MCF derivatization of urine samples for analysis on GC-QqQ-MS system

Metabolite quantification on the GC-QqQ-MS is performed by first preparing a standard solutions with increasing concentration for each metabolite (75 metabolites). The absolute quantification is then based on a standard (STD)-curve (made by the solutions) with samples spiked with deuterated standard mix (ISTD). The absolute quantification provides results in concentration.

400 μL 1M NaOH, 333 μL MeOH and 67 μL pyridine was added to the dried metabolites. 80 μL of MCF was then added and the sample was vigorously mixed for 60 seconds. 400 μL chloroform was added and mixed for 10 seconds. 400 μL of 50 mM $NaHCO_3$ was then added and mixed for 10 seconds.

The chloroform-phase (lower phase) was transferred into a new clean PP tube and 2-3 spatula of dried $NaSO_4$(s) was added prior to mixing for 5 seconds. The water-free chloroform phase was then added to GC/MS vials without insert. 30 μL internal standard (ISTD) was added to inserts and every inserts was inspected ensuring that the pipette volume was correct. 170 μL of the chloroform phase was then carefully mixed with the ISTD. The process in known as spiking. The steps are shown in Figure 2.4

2.1.5 Preparation of STD-curve samples

The standard solution (1.49 mM) was brought to room temperature prior to derivatization. A dilution series of standard mix (1.49mM) was made with the following dilution 1:1, 1:3, 1:27 and 1:243, by dilution with 1M NaOH(aq). For more detail of the standard mix, see Appendix A.4.

The 1:3 dilution was prepared by diluting 66.67 μL of the STD-mix with 133.33 μL NaOH (1M) gaining a total volume of 200 μL with a concentration of 0.498 mM. The 1:27 dilution was prepared by diluting 22.22μL of the 0.498 mM STD-mix with 177.8 μL 1M NaOH gaining a total volume of 200 μL with a concentration of 5.533 μM. The 1:243 dilution was prepared by diluting 22.22 μL of the 5.533 μM STD-mix with 177.8 μL 1M NaOH gaining a total volume of 200 μL with a concentration of 6.142 μM.

2.1.6 MCF derivatization of STD-curve samples

100 μL of the 1:1, 1:3, 1:27 and 1:243 dilution was added in four PP tubes together with 300 μL 1M NaOH, 333 μL MeOH and 67 μL pyridine. The STD-mix was vigorously mixed for 10 seconds to ensure homogenous solution. 80 μL of MCF was then added and the sample was mixed for 60 seconds. 400 μL chloroform was added and mixed for 10 seconds. 400 μL of 50 mM $NaHCO_3$ was added and mixed for 10 seconds. The chloroform-phase (lower phase) was transferred into a new clean PP tube with 2-3 spatula of dried $NaSO_4$(s) and mixed for 5 seconds. The water-free chloroform phase was then transferred to GC/MS vials without insert.

2.1.7 d_3-MCF / d_4-MeOH derivatization of internal standard

The identification of metabolites in the urine samples is quantified by spiking the samples with a derivatized metabolite standards. The metabolite standards was separately derivatized with deuterated methyl chloroformate (d_3-MCF) and methanol (d_4-MeOH).

100 μL of 1:3 diluted STD mix was added in four PP tubes together with 300 μL 1M NaOH, 333 μL heavy labelled methanol (d_4-MeOH) and 67 μL pyridine. The ISTD-mix was vigorously mixed for 10 seconds. 80 μL of heavy labelled MCF (d_3-MCF) was then added and vigorously mixed for 60 seconds. 400 μL chloroform was added and vigorously mixed for 10 seconds and then 400 μL of 50mM $NaHCO_3$ was added and vigorously mixed for 10 seconds. The chloroform-phase was transferred into a new clean PP tube together with 2-3 spatula of dried $NaSO_4$(s) and vigorously mixed for 5seconds. The water-free chloroform phase was the added to GC/MS vial without insert.

The same procedure was performed for the remaining three tubes. The four tubes were at the end mixed together to give a homogenous concentration of the internal standards.

2.1.8 Spiking of urine samples and spiking of STD-curve samples

30 μL of the d_3-MCF/d_4-MeOH derivatized ISTD-mix was added into GC-MS vial with insert. 170 μL of the MCF/MeOH derivatized STD-curve samples was added to the insert and mixed carefully with the ISTD-mix. 30 μ of the d_4-MeOH/d_3-MCF derivatized ISTD-mix was also mixed with 170 μL of the MeOH/MCF derivatized urine samples (36 samples). Figure 2.4 presents the derivatization steps of urine samples and the last step presents spiking with heavy labelled compounds

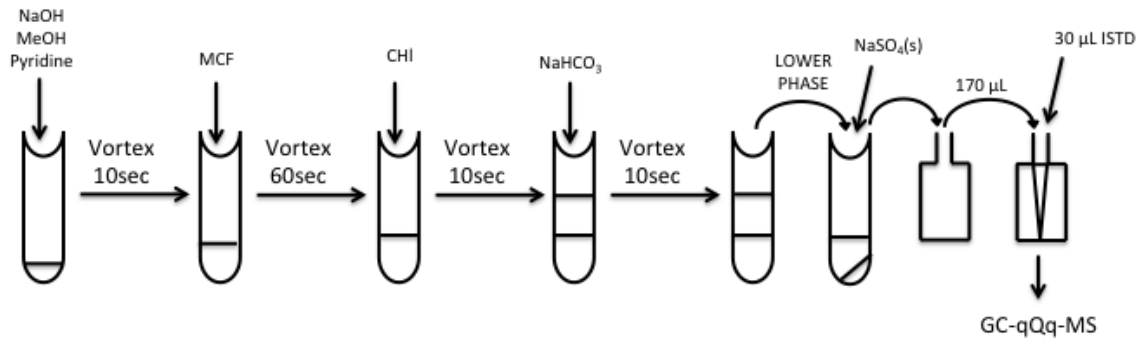


Figure 2.4: *MCF derivatization and spiking with deuterized internal standard*

2.1.9 GC-QqQ-MS instrumentation and data analysis

The samples were run on a GC-QqQ-MS system (Agilent 7890A series GC system coupled with Agilent 7000B triple quad MS). The carrier gas used was Helium 6.0 with a pressure on 14.5 psi. 1 μL was injected operating in pulsed splitless mode. The GC oven was equipped with a J&W Scientific DB-5MS+DG column (30 m + 10 m DG, 250μm inner

diameter, 0.25μm film thickness). The GC oven was kept at 40°C for 3 minutes and then raised to 320°C using a linear gradient of 20°C/min for 5 minutes resulting in a total run time of 20 minutes. The interface temperature was set to 280°C and the MS was operated in MRM (multiple reaction monitoring) mode for absolute quantification of metabolites (75). Methane reagent gas was used for PCI.

The absolute quantification of the 75 metabolites was performed by first making a linear standard curve from the highest concentration to the lowest concentration which is the limit of quantification (LOQ). The standard curve was forced to go through zero, and the best fit of either linear or power are used (R^2=0.98).

The metabolites was analyzed and detected based on the co-elution with the deuterated metabolites (same retention time) as actual metabolites and concentration was automatically calculated from the standard curve and ISTD-peak area.

2.2 Silylation with MSTFA+ 1% TMCS for metabolite analysis on GC-SQ-MS system

The aim of this method was to identify metabolites such as sugars, sugar alcohols, purines, pyrimidenes, amino acids and organic acids.

2.2.1 Preliminary study for urease degradation

Using biofluid such as urine, which contains a great amount of urea, degradation of urea must be performed prior to silylation. Unlike alkylation, urea will undergo derivatization with MSTFA + 1% TMCS yielding high symmetrical peaks in the chromatograms. Enzymatic treatment with urease was performed in a preliminary study. The preliminary study also determined the urine concentration needed for highest response. The method and strategy used was recommended by Huang [20]. Fresh human urine was collected from 3 young healthy volunteers (aged 22-25) early in the morning. Two replicates were taken from each sample. The response was best for the samples where 250 μL of the supernatant was used (data not shown).

2.2.2 Urease degradation and preparation of urine samples

The preparation of the samples and degradation of urease was based on the preliminary study using the method recommended by Huang [20]. A total of 37 urine samples,16 healthy urine samples and 21 urine samples from patients was used. Figure 2.5 presents the steps for urease degradation. For chemical details, see Appendix A.1.

10.28 μL (40UI) of urease and 189.7 μL ion-free water was mixed with 200 μL of urine and incubated at 37°C for 45 minutes. 600 μL methanol was then added and the samples

were vigorously mixed for 10 seconds. The samples were ultrasonicated for 5 minutes and centrifuged at max speed (5000 rpm) for 5 minutes. 250 μL of the supernatant, giving a volume of 50 μL urine in each samples was mixed with 25 μL d27-myristic acid solution. The samples were then concentrated for 4 hours at 65°C with a vacuum concentrator (Speedvac-Savant SPD20110).

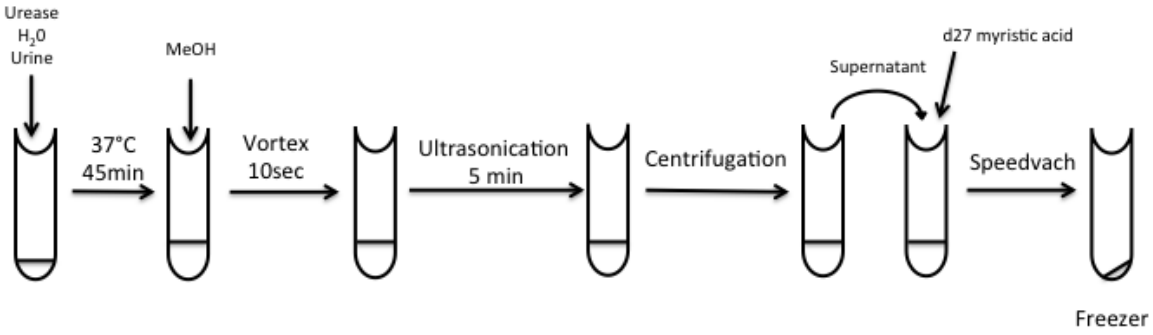


Figure 2.5: *Flowsheet for urease degradation and precipitation of proteins*

2.2.3 Siyilation of urine samples and check sequence

The silylation protocol is a modified method obtained from Fiehn [43]. The silylation was peformed on urine samples and for the check sequence (d27-myristic acid, fatty acid methyl esters mix (FAME), standard mix (STD1/STD2) and blank). For more detail about the mixtures see Appendix A.2 and A.3. Figure 2.6 presents the the steps during silylation and Table 2.1 presents the volumes of reagents added.

25 μL d27-myristic acid solution was added to empty vials, labeled d27-myristic acid and fame. 75 μL of 2.0mM standard mix was added to two empty vials labeled std1 and std2. The check sequence samples were then evaporated to dryness. 10 μL FAME was then added to the vial labeled fame. 50 μL 4% w/v methoxyamine hydroxychloride (HCl) in pyridine was added to the urine samples, d27-myristic acid vial, fame and to an empty vial labeled blank. 100 μL 4% methoxyamine HCl in pyridine was added to STD1 and STD2. The vials were then shaken carefully at 30 °C for 90 minutes. 20 μL of the samples, d27-myristic acid, fame and blank and 50μL of the std1 and std2 was transferred to new vials. 180 μL MSTFA + 1% TMCS was then added to the samples, d27-myristic acid and the two standards, while 150 μL MSTFA + 1% TMCS was added to FAME. The vials were shaken carefully at 37 °C for 30 minutes and then cooled to room temperature. The solutions were transferred to insert and placed back into the vials and analyzed on GC-SQ-MS.

Table 2.1: *Volumes and reagents used during silylation*

	Urine samples	Blank	d-27 myristic acid	FAME	STD1/STD2
2 mM sugar mix					75 μL
4% methoxyamine HCl in pyridine	50 μL	50 μL	50 μL	50 μL	100 μL
Volume transferred	20 μL	20 μL	20 μL	20 μL	50 μL
MSTFA + 1% TMCS	180 μL	180 μL	180 μL	150 μL	180 μL

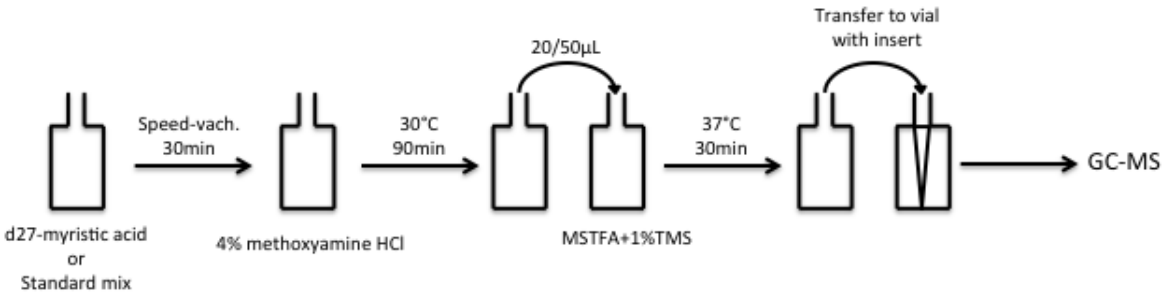


Figure 2.6: *Flowsheet of the steps during silylation*

2.2.4 GC/MS instrumentation and data analysis

Samples were run on a GC-SQ-MS (Agilent 7890A series GC system coupled with Agilent 5975 Mass Selective Detector). 1 μL was injected in split mode (split ratio 10:1). The carrier gas used was helium. The GC was operated at a constant flow equipped with a J&W Scientific DB-5MS column + DG (30 m, 250 μm inner diameter, 0.25 μm film thickness). The GC oven was kept at 60°C for 1 minute then raised to 325°C using a linear gradient of 10°C/min for 10 minutes resulting in a total run time of 37.5 minutes. The MS operated in scan mode (start after 5.9 min, mass range 50-600 m/z at 2.66 scan/s) and utilized an EI source.

The data collected, was obtained using Agilent MSD Productivity ChemStation. The data set was loaded to ChemStations QEdit function, both automatic and manually integrated ion chromatogram up against a MCF-library consisting 1034 metabolites (i.e qualitative). The spectra obtained were also compared to library spectrums available in NIST Automated Mass Spectral Deconvolution and Identification Software (AMDIS). AMDIS separates ion chromatogram signals in a complex TIC's, creating pure compound peak free of overlapping signals (i.e. deconvolution) [36].

After identification, the results were normalized against d-27 Myristic acid to correct for variations in each run.

2.3 LC-TOF-MS instrumentation

No sample preparations was performed in this study, as the LC/MS results were obtained from a previous run. The method was provided by Bajad et al [44]. The samples were run on a LC-TOF-MS (Agilent) equipped with a HILIC column (2.0 x 250 mm). The temperature of the column was 25°C. The injection volume was 10 μL. The mobil phase was at pH 9.45 where mobile phase A consisted of 20 mM ammonium acetate + 20 mM ammonium hydroxide (95:5, water:acetonitrile), while mobile phase B consisted of acetonitrile (ACN). The gradient started with 85 % ACN and decreasing to 0 % ACN within 15 minutes. The mobile phase was then kept at 0 % 15 minutes. The total analysis was 30 minutes with a 10 minutes post run for re-equilibrium. The MS used ESI as ion source utilizing both positive and negative mode, ESI^+ and ESI^- respectively. The TOF-MS was operating in scan mode (mass range 50-1500 m/z at 2.5 scans s^{-1}).

3 Results and Discussion

3.1 Clinical characteristics

Urine samples from patients with kidney disease were compared with samples from healthy voluenteers. Information about the subjects was provided by St. Olavs Hospital. The clinical characteristics of the patients and control group are shown in Table 3.1. The two groups were different in age, gender and diagnosis. Estimated Glomerular Filtration Rate (eGFR) in patients showed a significant decrease (17.4 ± 7.0ml/min/1.73m^2), which is compatible with stage 4 CKD patients. Stage 4 is defined with renal function of between 15-29 ml/min/1.73m^2 [2].

Table 3.1: *Clinical characteristics of patients included in the study*

Variable	Control (No.=16)	CKD patients (No.=21)
Age	48,8 ± 10,2	64,3 ±16,2
Sex (male/female)	6 / 21	15 / 21
eGFR (ml/min/1.73m^2)	>90-78	17,4 ±7,0
Systolic BP (mmHg)	131,8 ±14,3	144,0 ± 24,3
Diastolic BP (mmHg)	83,4 ±9,4	80,4 ±10,8

3.2 Evaluation of MS methods for analysis of urine

Four different MS methods were applied to evaluate urine samples from CKD patients and the control group. The three first methods uses GC/MS and performs targeted analysis. The fourth method uses the LC/MS and performs untargeted analysis:

- Method 1: Targeted analysis using alkylation and GC-SQ-MS on 10 samples, identifying amino and nonamino acids.
- Method 2: Targeted analysis using silylation and CG-SQ-MS on all samples, identifying amino- and nonamino acids, purines, pyrimidines, sugars, sugar alcohols etc.
- Method 3: Targeted analysis with absolute quantification using alkylation and GC-QqQ-MS on all samples, identifying amino and nonamino acids.
- Method 4: Untargeted analysis of data from LC-HILIC-MS identifying metabolites with high mass accuracy.

Results from each method will be presented in the following sections 3.3-3.6.

3.3 Results from method 1: Alkylation and metabolomic application of GC-SQ-MS

The aim of this approach was to first identify metabolites using the in-house build library in ChemStation and AMDIS. Since the MS was operated in scan mode, there are also unknown (i.e metabolites not in the library) which also could be detected. The aim of the second approach was therefore to investigate unidentified peaks with high abundance and observe how this peaks were represented in the sick and healthy group.

As this was a preliminary study, further studies was performed using absolute quantification (Section 3.6), only ten samples were analyzed. The samples were derivatized by alkylation and analyzed on GC-SQ-MS system.

3.3.1 Identification of organic acids and amino acids

A total number of 36 metabolites was identified. The raw-data was normalized against d5-glutamate to correct for variation in each run through the GC-MS and the result is presented as average peak area per 100 μL urine in Figure 3.1-3.3 (for additional data, see Appendix A.5). The bar plots shows differences between patient and healthy for all of the metabolites and the error bars illustrates significant differences for malanoic acid, succinate, threonine, citraconate, benzoate, aspargine, aspartate, serine-2, glutamine, isocitrate, tryptophane, citrate and histidine. During sample preparation no replicates were made, so the average peak area is the sum of 5 samples. As the plot presents, the standard deviation are observed to be very large for many of the metabolites. In the case of phenylalanine sample number 105 (the control groups) is 10 times higher in peak area compared to the others, which contributes a lot in the calculation when only five samples are evaluated. However, this is expected since the urine samples were collected from subjects with individual differences which will contribute to variations in the data.

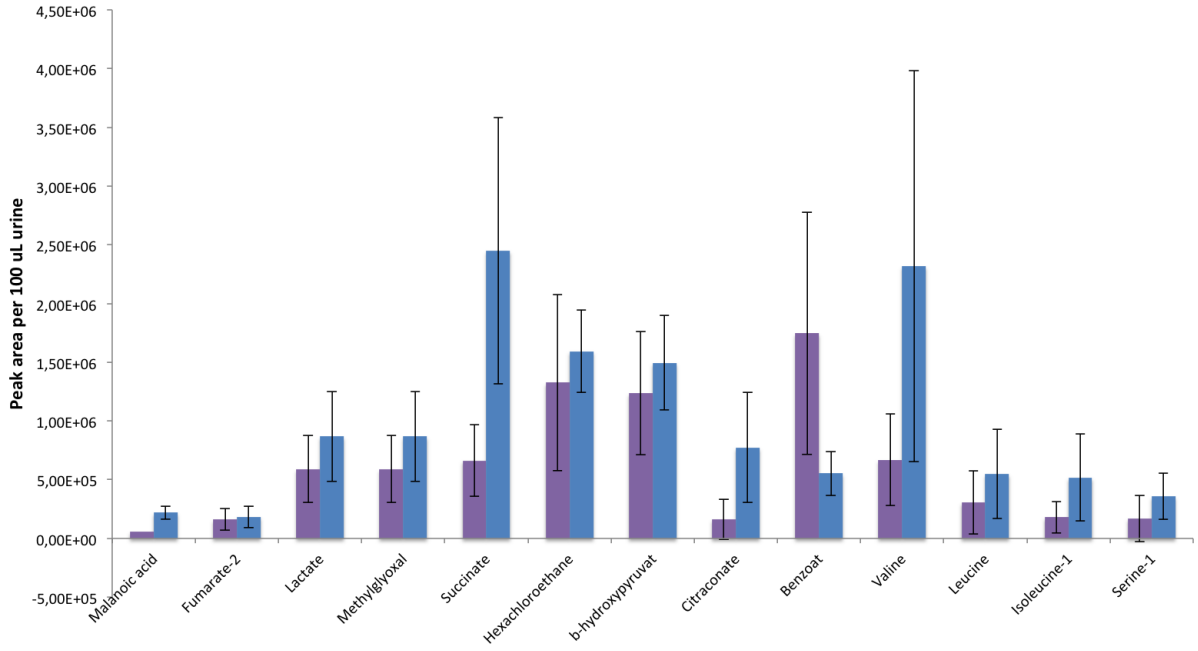


Figure 3.1: *Bar plot of average peak area per 100 µL urine of metabolites in sick (purple) and healthy (blue)*

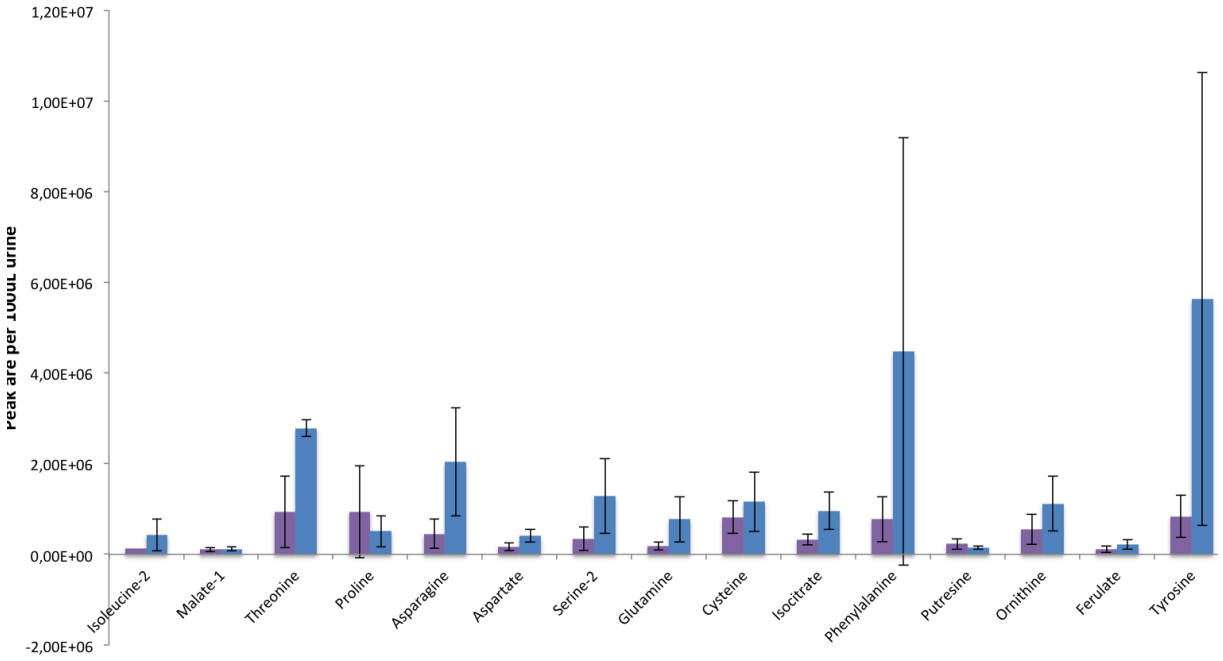


Figure 3.2: *Bar plot of average peak area per 100 µL urine metabolites in sick (purple) and healthy (blue)*

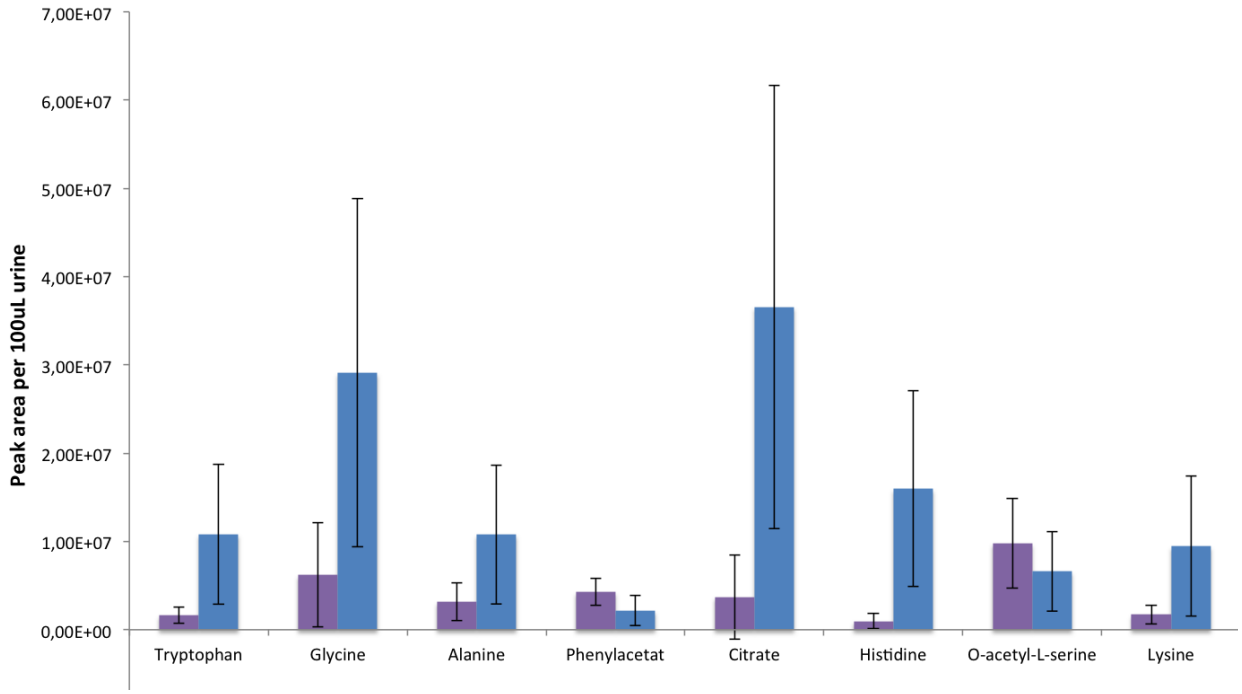


Figure 3.3: *Bar plot of average peak area per 100 μL urine of metabolites in sick (purple) and healthy (blue)*

Several metabolites found in the healthy group have higher concentration compared to the patient group; Valine, tryptophan, glycine, citrate, histidine, tyrosine, phenylalanine, threonine, glutamine and succinate to mention some. Benzoate and proline, on the other hand were found to be higher in urine samples from the patient group. The metabolites were further evaluated statistically with T-test. Metabolites found to be statistical different was O-acetyl-L-serine, glycine, tryptophane, ferulate, isocitrate, serine-2, aspartate, aspargine, citraconate, succinate, methyl glyoxal, lactate, benzoate, glutamine, citrate and histidine. For more calculation details see Appendix A.6.

Literature search was performed to evaluate the results from this method and to see if the results were comparable with previous studies. Ying-Yong Zao et al. investigated biochemical changes due to chronic renal failure (CRF) in rats and found 12 potential biomarkers. Phenylalanine was found to be lower in the CRF rat group compared to the control group [45], showing the same results as for this experiment.

Hayashi K et al. investigated serum and urine metabolites in patients with stage 1-2 chronic kidney disease [9]. They discovered that aspartate in urine increased in the patient group, while histidine and glutamine increased in the control group. Similar increase of histidine and glutamine was also observed in this experiment, however aspartate was observed to be opposite. This might be due to the fact that their patients are in the first stages of the kidney disease, while patient investigated in this thesis are in the stage 4 CKD and it is assumed that this will yield some different results. Hayashi K et al.

also discovered higher concentration of citrate in urine samples for the healthy group, which was also observed in this study. The higher concentration of citrate, might suggest changes in the glucose metabolism and may orginate from the disease [9].

The observations in this study is therefore found to be comparable with previous literature and studies [9, 45], as well as presenting interesting features for other metabolites such as benzoate, proline and aspartate.

3.3.2 Identification of metabolites not included in the in-house MCF GC-MS metabolite library

This data-method was applied in order to find unidentified peaks with high abundance. For each sample AMDIS would present a number of target compounds and compound present in the sample (e.g 43 targets and 500 components). In order to find the additional metabolites, peaks with response over 2000 (peaks under this limit is considered as disturbance in the chromatogram) were chosen for further study. The mass spectrum for each was used to search through the in-house library obtained from NIST.

Identified metabolites in addition to the in-house MCF GC-MS library

Figure 3.4 presents the number of additional metabolites found in each sample. A higher number of metabolites was identified for urine samples from the control (101-106), compared to the CKD group (1-6). The names of the metabolites and response in peak hight from this data-method are presented in Table 3.2-3.4.

For each hit in NIST, a probability match was represented in precent. The identification of one component was therefore based on this percentage which could be between 20%-98% and only components with a matching percentage over 80% were chosen.

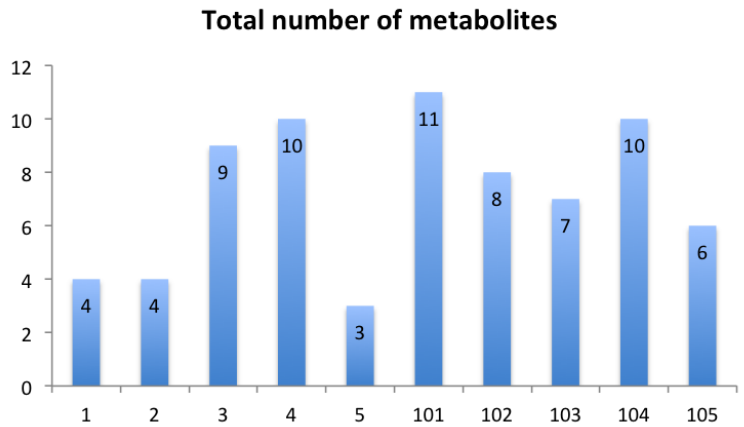


Figure 3.4: *Total number of metabolites identified in each sample*

Table 3.2: *Response in peak height of polar compounds with retention time from RT = 6.248 - 9.814 minutes*

	Butanoic acid	Propanoic acid	Propanedioic acid	Tricyclo(2,4) -hex-3-ene-3 -carbonitrile	Propanedioic acid	Heptanoic-/ Pentanoic acid	Butanedioic acid	Acetic- / 3-Octenoic acid
RT [min]	6.248	6.458	7.398	7.797	8.562	9.177	9.427	9.814
1						320000		48000
2						340000		
3					75000	280000		
4	< 5000		24000	55000	125000			105000
5						320000		
101	140000	150000	115000					340000
102	40000		32000					
103	85000		75000					
104	80000					350000	340000	
105	260000	170000				360000	380000	

Table 3.3: *Response in peak height of polar compounds with retention time from RT = 12.378-20.25 minutes*

	Hexanedioic acid	Acetic acid	1-Propene-1,2,3-tricarboxylic acid	Motrin methyl ester	Trimethyl 2-methoxy propane-1,2,3-tri carboxylate	1H-Indole-3-acetic acid	2-(4-methoxyphenyl)-2-methyl-1,3-Dioxolane
RT[min]	12,378	12,680	14,485	16.008	17,416	19,514	20.25
1				550000			
2						1250000	
3		1300000	350000				
4			600000			380000	
5						260000	
101	220000	480000	3000000			800000	
102	180000	145000	650000			350000	
103			1150000			800000	1600000
104	240000		1900000		1000000	410000	
105			2200000				

Table 3.4: *Response in peak height of polar compounds with retention time from RT = 20.541 - 24.36 minutes*

	Pentadecanoic acid or Hexadecanoic acid	Methyl 3-methoxy -4,5-methylenedioxybenzoate	Octadecanoic acid	L-Glutamine
RT [min]	20.541	21.068	22.611	24.36
1				480000
2	280000			450000
3	320000		170000	125000
4	320000			
5	210000			
101	500000	380000		650000
102	400000		210000	
103	480000			
104	390000			520000
105				600000

Peaks which were not identified in the in-house library of NIST

Identification of unknown peaks by NIST was not so successful as there was a number of metabolites (Figure 3.5) which had no match (<15%). However, as this study also had the intention of finding new metabolites which could be added to the in-house library and since the peak height were interestingly high it was decided to present the results.

Table 3.5 and 3.6 present the unidentified compounds and response in height which was not successfully identified.

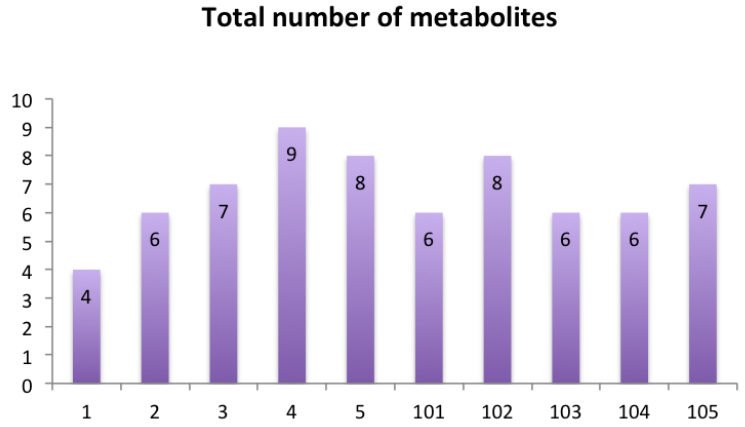


Figure 3.5: *Total number of unidentified compounds found in each sample*

Table 3.5: *Response in peak height of unidentified metabolites with retention time between RT= 8.004 -17.531 minutes*

	Unidentified compounds									
RT [min]	8.004	10.122	**10.421**	10.577	11.891	13.046	16.1256	17.283	17.423	17.531
1			**800000**							
2			**900000**			240000				2200000
3			**800000**	220000						
4		250000	**900000**	200000	160000	300000				700000
5			**800000**		900000		300000			
101	700000		**1400000**		400000	400000			1600000	3000000
102			**800000**			160000	400000		400000	800000
103		450000	**900000**			350000		500000		2200000
104			**900000**					400000		1200000
105			**800000**		700000	625000			1200000	1600000

Table 3.6: *Response in peak height of unidentified metabolites with retention time between RT= 19.523 - 25.316 minutes*

	Unidentified compounds								
RT	19.523	20.252	20.614	22.3	22.988	23.014	23.232	24.669	25.316
1		650000				1300000	780000		
2					700000		400000		450000
3	600000	600000	450000		1100000				180000
4		1400000			700000		500000		
5				700000	350000		4800000	625000	850000
101									
102			900000		200000		160000		
103		1600000							
104		1000000				1200000			
105	900000				1600000				

As Table 3.5 shows one metabolite was observed to be present in all of the samples (RT:10.421). Except from sample 101, the concentration of the compound is observed to be the same in the urine samples (Figure 3.6) and the differences was not found to be statistically different on a 95 % significance level. For calculation details, see Appendix

A.6. No further study was performed on this compound. Future study in identifying the compound (scan is presented in Appendix A.7) could be done by purification and structure characterization with one- and two dimensional (1D/2D) NMR. 1D provides information about every single atom, while 2D provides information about how they correlate to each other. Recommended analysis method is ^{1}H-NMR and ^{13}C-NMR with 2D analysis such as HMBC (Heteronuclear Multiple Bond Correlation), HSQC (Heteronuclear Single Quantum Coherence) and H-H-COSY (Hydrogen-Hydrogen Correlation Spectroscopy) [46]. For more information of the techniques, the reader is referred to the literature (e.g [46]).

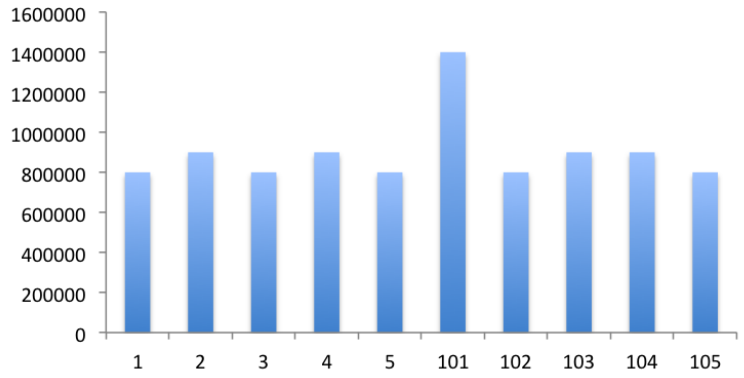


Figure 3.6: *Response in peak height for an unidentified metabolite at retention time 10.421 minutes in 10 different samples*

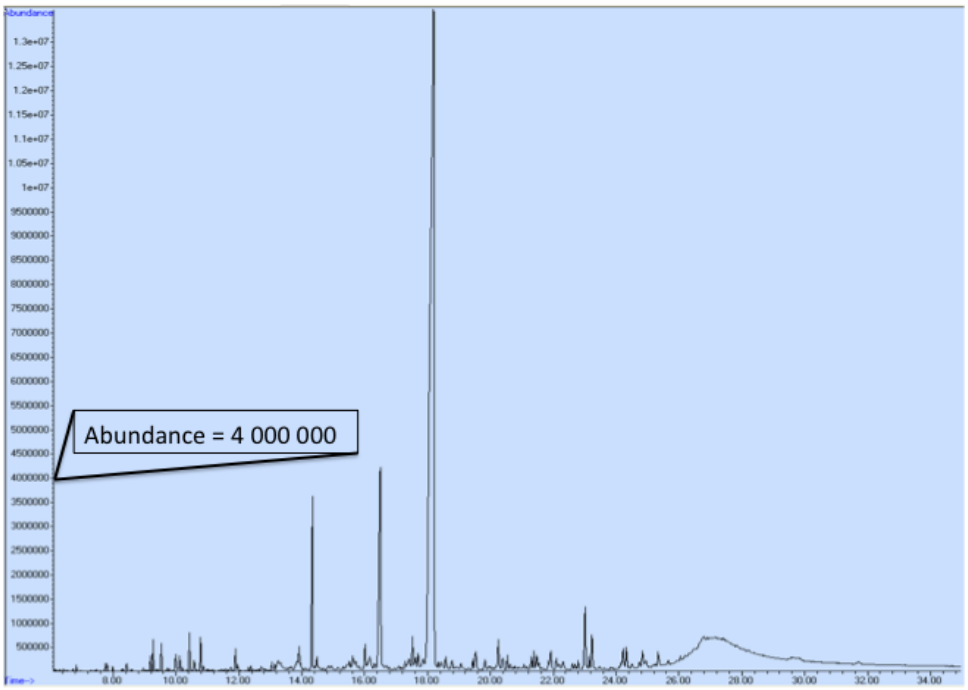
The overall results from this study

Table 3.7 presents the names of the compounds and how they occur in sick and healthy. As the table shows some metabolites was found in both sick and healthy samples, while some was found in only one of the groups. Even though some metabolites only occur in one of the groups, the metabolites were not present in all of the samples within the group. Some of the metabolites such as hexanedioic was found to be present in three out of five samples in the healthy group and absent in the urine sample from patients. Other metabolites such as motrin methyl ester was found in only one of the samples in the sick group. Only one component, butanoic acid was observed to be more represented in the controls. However, butanoic acid was detected very early in the chromatogram. Each sequence was run through the GC/MS with a check sequence and blank samples. Butanoic acid was also observed in the start of the chromatograms for the blank sample (data not shown). Therefore, any conclusions of butanoic acid as a possible biomarker cannot be made.

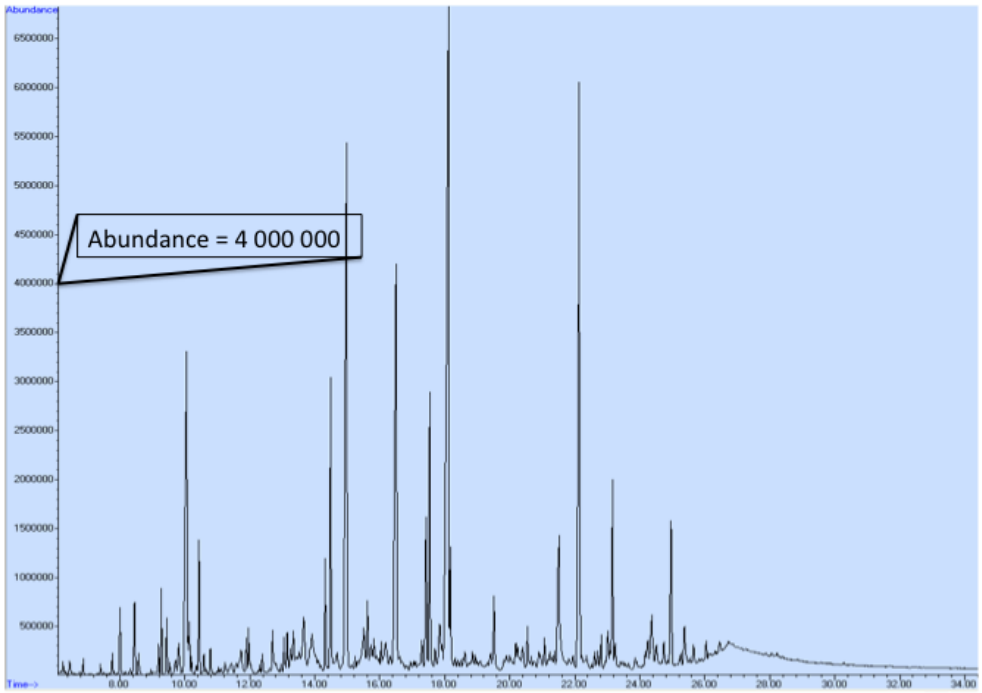
Table 3.7: *The table shows the untargeted metabolites from 10 samples, identified in in NIST. Some of the metabolites were only present in sick or healthy*

Metabolites identified in NIST	Retention Time	Found in	
		sick	healthy
Butanoic acid	6,648	✓	✓
Propanoic acid	6,458		✓
Propanedioic acid	7,398	✓	✓
Tricyclo [3.1.0.0(2,4)]hex-3-ene-3-carbonitrile	7,797	✓	
Propanedioic acid	8,562	✓	
Heptanoic / Pentanoic acid	9,177	✓	✓
Butanedioic acid	9,427		✓
Acetic acid / 3-Octenoic acid	9,814	✓	✓
Hexanedioic acid	12,378		✓
Acetic acid	12,68	✓	✓
1-Propene-1,2,3-tricarboxylic acid	14,485	✓	✓
Motrin methyl ester	16,008	✓	
Trimethyl 2-methoxypropane-1,2,3-tricarboxylate	17,416		✓
1-H-Indole-3-acetic acid	19,514	✓	✓
2-(4-methoxyphenyl)-2-methyl-1,3-Dioxolane	19,514		✓
Pentadecanoic acid / Hexadecanoic acid	20,541	✓	✓
Methyl-3-methoxy-4,5-methylenedioxybenzoate	21,068		✓
Octadecanoic acid	22,611	✓	✓
L-Glutamine	24,36	✓	✓

The TIC for a CKD patient and a control is presented in Figure 3.7. The TIC presents the abundance of the metabolites in each sample. So far, the study has found that the metabolites are higher in concentration for the controls compared to the CKD subject and that variations caused by individual differences might contribute to differences. The peak with the highest abundance is d5-glutamate (which was added for normalization due to variations caused by errors in sample preparation as well as variation in the detector). The chromatograms in Figure 3.7 shows a mark for the abundance value of 4×10^6. By elevating the mark in (a) to the same point as in (b), more peaks will become visible and the chromatograms would not be that different. However, some peaks after 20 minutes are observed to be more abundant in the control sample, concluding previous observation of decreases in concentration of metabolites in chronic kidney disease patients.



(a)



(b)

Figure 3.7: *Total ion chromatogram for (a) patient and (b) control. The x-axis shows the retention time (0- 34 minutes) and the y-axis shows the abundance*

3.4 Results from method 2: Silylation and metabolomic application of GC-SQ-MS

The aim of using this method was to identify new metabolites such as sugars, sugar alcohols, purines and pyrimidines, in addition to amino-and nonamino acids. The in-house library contained approximately 1000 compound's, and 59 metabolites were successfully identified.

In each identified metabolite the urine samples within one group (CKD or control) appeared with varying concentration and the metabolites was not found in some of the urine samples. For example fructose was observed in 7 out of 21 samples of the CKD samples and in 5 out of 16 of the controls. It could therefore not be concluded if some metabolites are more present in CKD subjects or healthy groups (For additional information, see Appendix A.8) However, some metabolites shows some differences worth noticing. 6-deoxy-D-glucose, glycolic acid, hypoxanthine, uracil and citric acid are observed to be more present in the healthy group, (Figure 3.8 - 3.12).

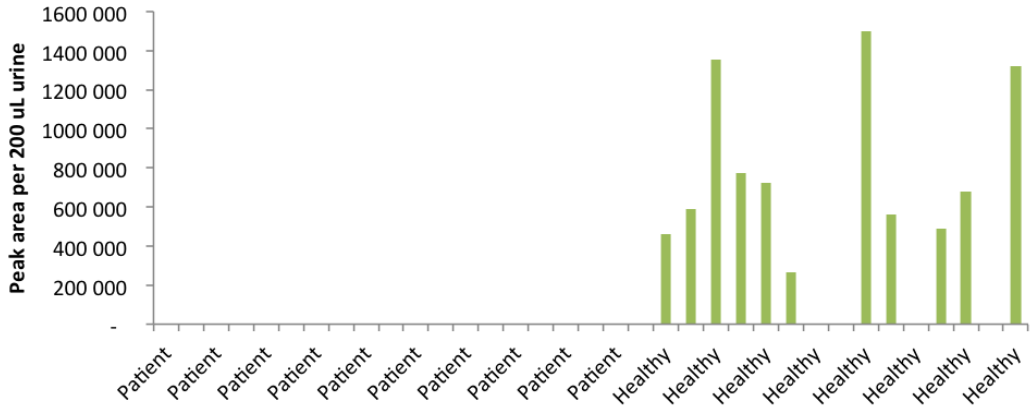


Figure 3.8: *6-deoxy-D-glucose*

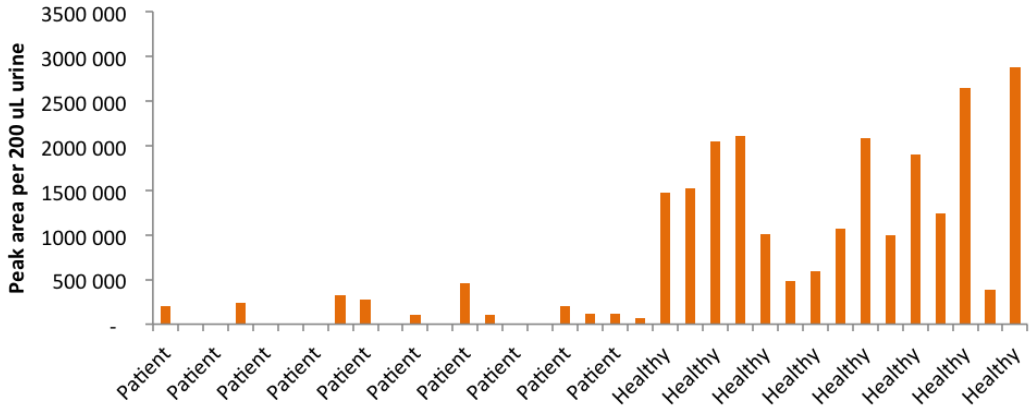


Figure 3.9: *Glycolic acid*

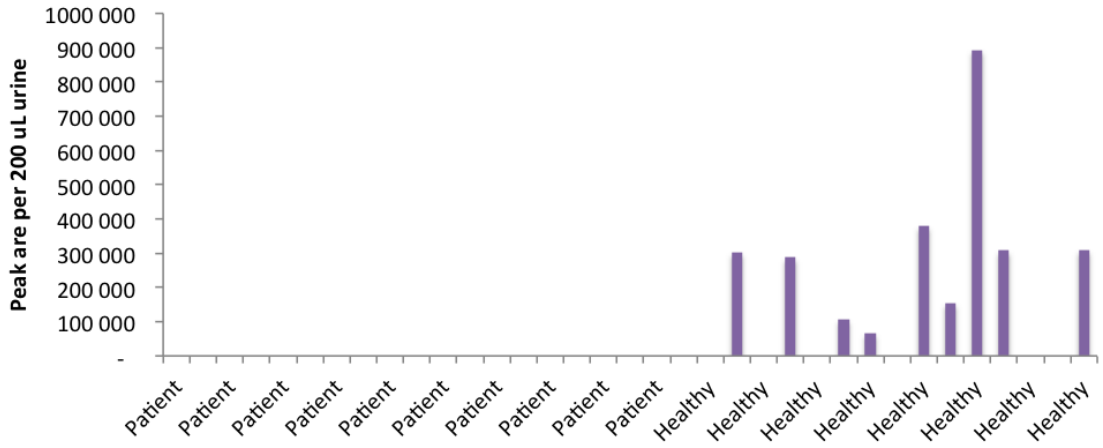


Figure 3.10: *Hypoxanthine*

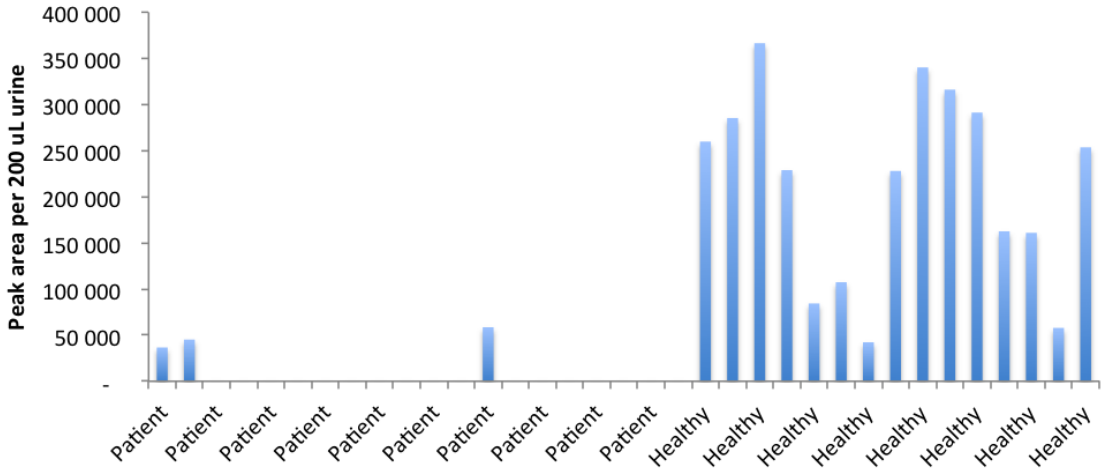


Figure 3.11: *Uracil*

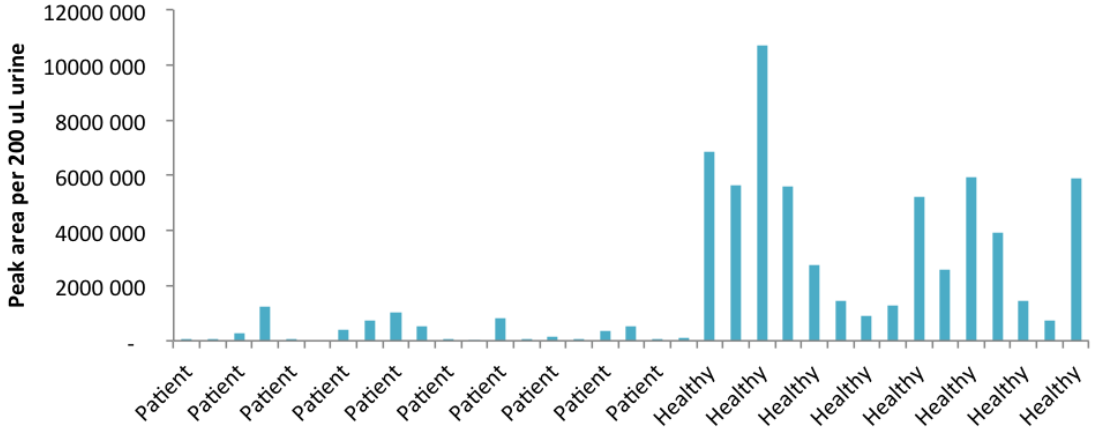


Figure 3.12: *Citric acid*

Literature search was performed to evaluate the results and to see if the data were comparable from previous studies. As presented earlier, Hayashi K et al. investigated metabolic changes i stage 1-2 CKD patients [9]. They observed that the levels of hypoxanthine was increased in the serum of CKD patients. The degradation of the purine adenosine monophosphate (AMP) yields hypoxanthine with uric acid as end product [47]. In this thesis hypoxanthine was found to be at higher levels in urine for healthy people and by comparing this results with Hayashi K et al, filtration of hypoxanthine might be decreased causing elevated levels in the serum. It might also indicating that kidney disease might interfere with the catabolism of purine nucleotides [9]. There has also been a study in measuring plasma concentration of hypoxanthine and uric acid [48] showing that the concentration of hypoxanthine increased in the plasma in patients with end stage renal failure (ESRF). Some literatures have focused on cardiovascular disease in relation to kidney disease [49] and if hypoxanthine in plasma may cause increased risk of cardiovascular disease [50].

Regarding the observation of 6-deoxy-d-glucose, glycolic acid, uracil and citric acid, no literature related to kidney disease was found to verify the results. However, uracil have been measured in plasma and urine for monitoring drug delivery in patients with cancer, receiving anti-metabolic chemotherapy drug such as 5-Fluorouracil [51],[52]. Due to limiting knowledge about pharmacokinetics and drug delivery, it is not possible to conclude if uracil might be a degradation product from the medications.

Principal component analysis plot of the data

Unscrambler was applied to perform PCA plot and to detect any structure in relationships between the metabolites. PCA is a tool to illustrate possible trends and grouping of the patient group and the healthy group as described in the Theory. Figure 3.13 presents the score plot and Figure 3.14 presents the loading plot, representing 93% of the variation of the data. The loading plot shows that glycerol and phosphoric acid dominates the plot, which can be explained by the metabolites having higher values. Different approaches were performed in order to normalize the data to achieve a clearer view in the loading plot. It was thereby decided to divide the data of each metabolites by their standard deviation. The new score- and loading plot is represented in Figure 3.15 and 3.16. The score plot (Figure 3.15) presents clustering of urine samples obtained from CKD patients (pink circle). Comparing the score plot with the loading plot (Figure 3.16) it is observed that D-mannitol, trans-4-hydroxy-L-proline, L-proline and L-glutamine contribute to the clustering of the CKD samples.

No literature was found regarding D-mannitol and trans-4-hydroxy-L-proline and their connection to kidney disease. In the previous study (method 1) the metabolite proline was found to be at higher concentration in the CKD samples, however glutamine was

observed to be opposite. Glutamine is quantitatively the most important donor of -NH_3 in the kidney. The -NH_3 is cleaved from glutamine by the action of phosphate-dependent glutaminase, which is subjected to the acid-bace balance in the kidney [53]. Hayashi K et al. measured a lower concentration of glutamine in urine for 1-2 stage CKD patients [9]. This indicates that the results with focus on glutamine from this study are not so reliable. This method used MSTFA + 1% TMCS as derivatization reagent, while method 2 used MCF as derivatization reagent. The silylation- and alkylation method will therefore be discussed later in this thesis.

The overall results from this method is that the silylation was successful in finding metabolites which was not observed in the results from the alkylation, hypoxanthine and uracil respectively. The principal component analysis also showed clustering of the CKD subjects due to metabolite similarities.

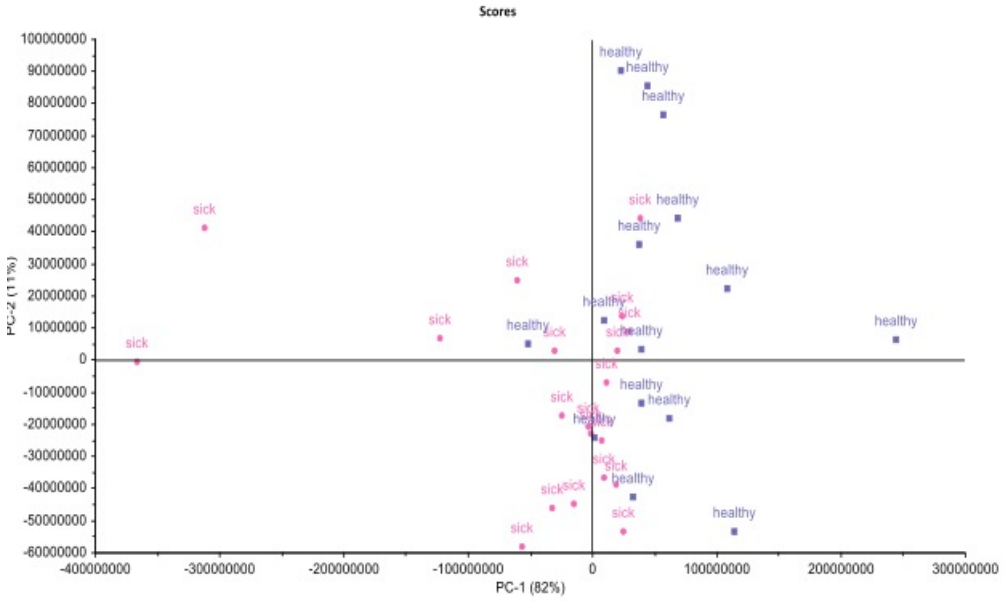


Figure 3.13: *Score plot of the urine samples explaining 93% of the variations of the data*

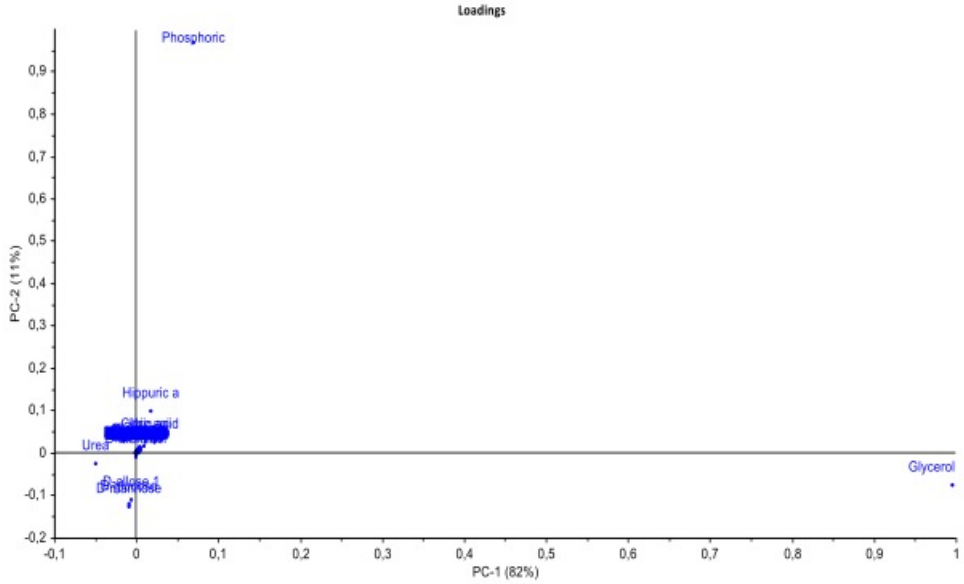


Figure 3.14: *Loading plot of the urine samples where phosphoric and glycerol dominates the plot*

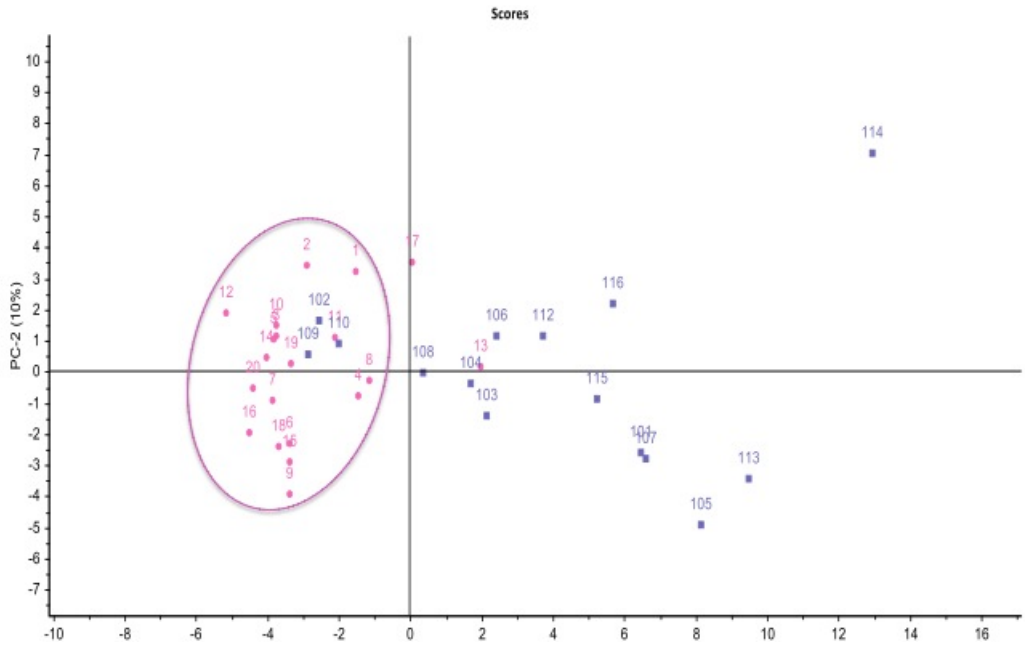


Figure 3.15: *Score plot of the urine samples where the pink circle marks the CKD subjects*

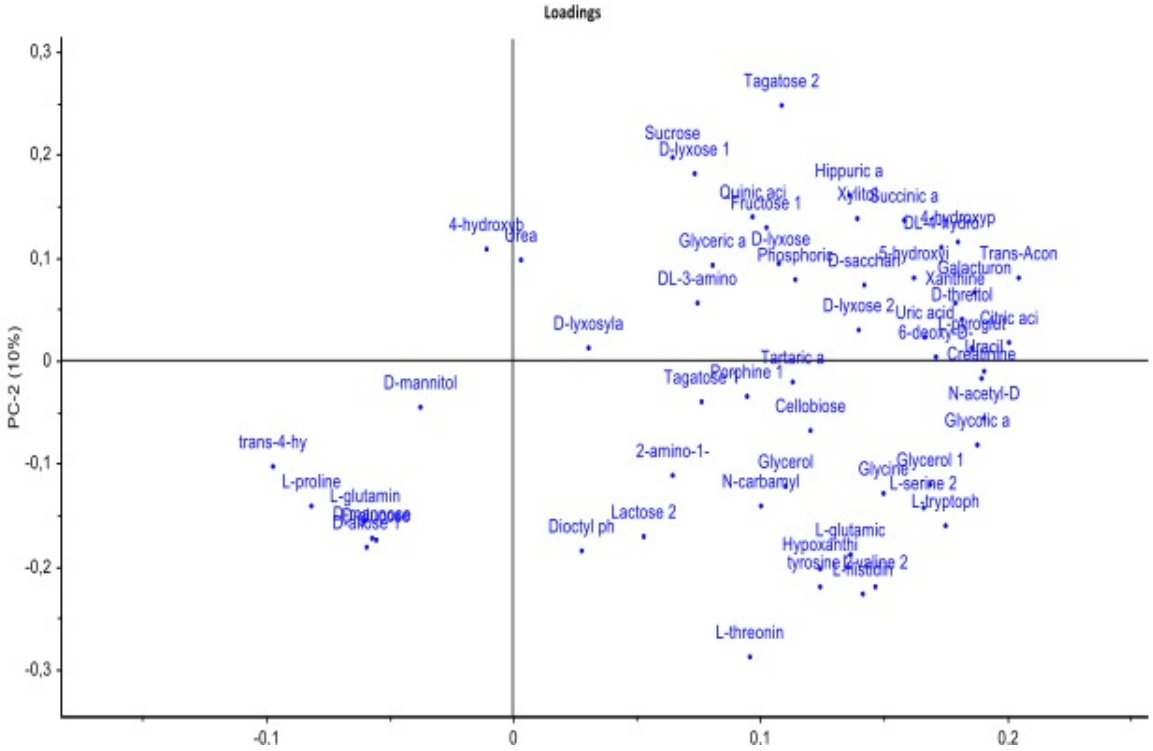


Figure 3.16: *Loading plot of the urine samples where D-mannitol, L-proline, L-glutamine and trans-4-hydroxyprolin contributes in clustering of CKD subjects in Figure 3.15*

3.5 Results from method 3: Alkyation and application of GC-QqQ-MS in metabolomics

The aim of this data analysis was to further study the results from the GC-SQ-MS scan data analysis, by performing analysis with tandem mass spectroscopy on all samples. GC-QqQ-MS operates in MRM (multiple reaction monitoring) mode and higher sensitivity and selectivity of the analytes are achieved.

A total of 46 metabolites were identified. The results (Figure 3.17-3.24) shows that almost all of the metabolites have a higher concentration in the control group compared to the CKD group which is compatible with the results in the two previous studies (for additional data, see A.9). The error bars shows that the metabolites malonate, citraconate, 2-aminobutyrate, citrate and 4-imidazoleacrylate are significantly different. Statistical calculations with student-t distribution was not performed as the number of subjects in each group was not the same, 16 and 21 respectively.

In method 1, the standard deviation were observed to be very large and it was believed that the number of subjects evaluated (only 5 in each group) was one factor contributing to the large error bars. However, as these plots present, the standard deviation are also observed to be very large for many of the metabolites in this method (no parallels were made). One can therefore conclude that individual variations will contribute to the data results. For future studies it is recommended to collect more samples. This will be discussed later.

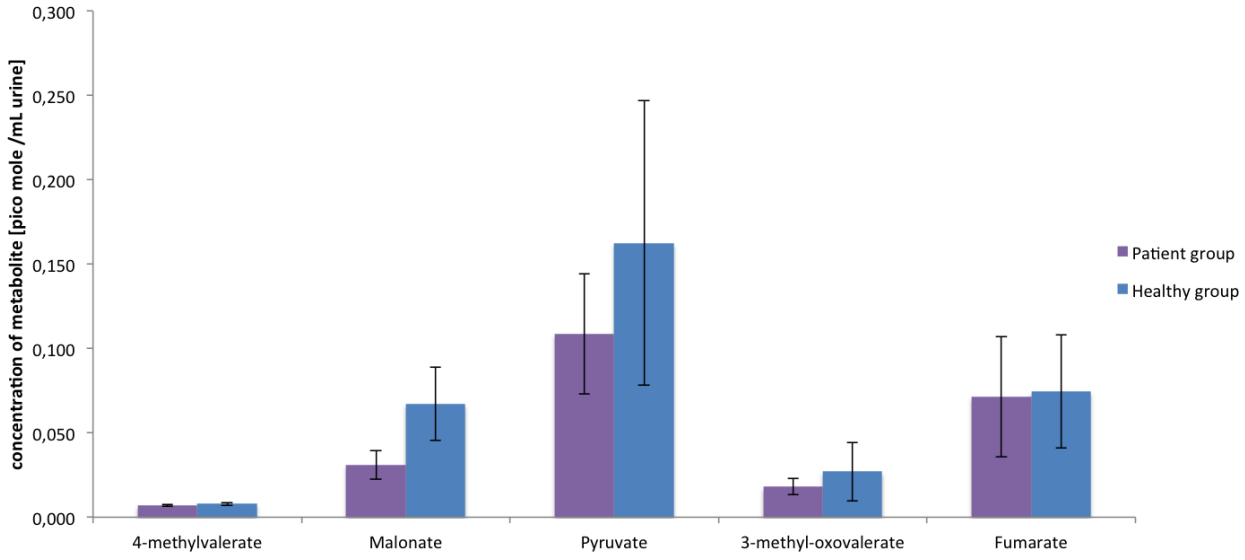


Figure 3.17: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis.*

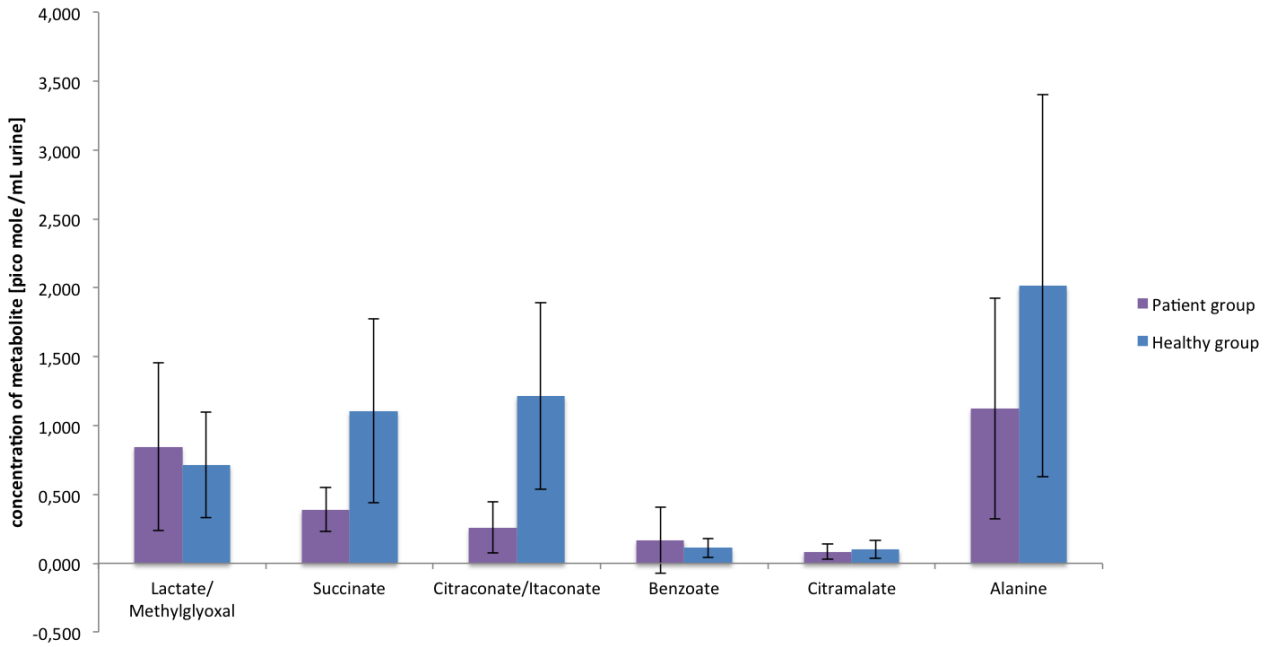


Figure 3.18: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

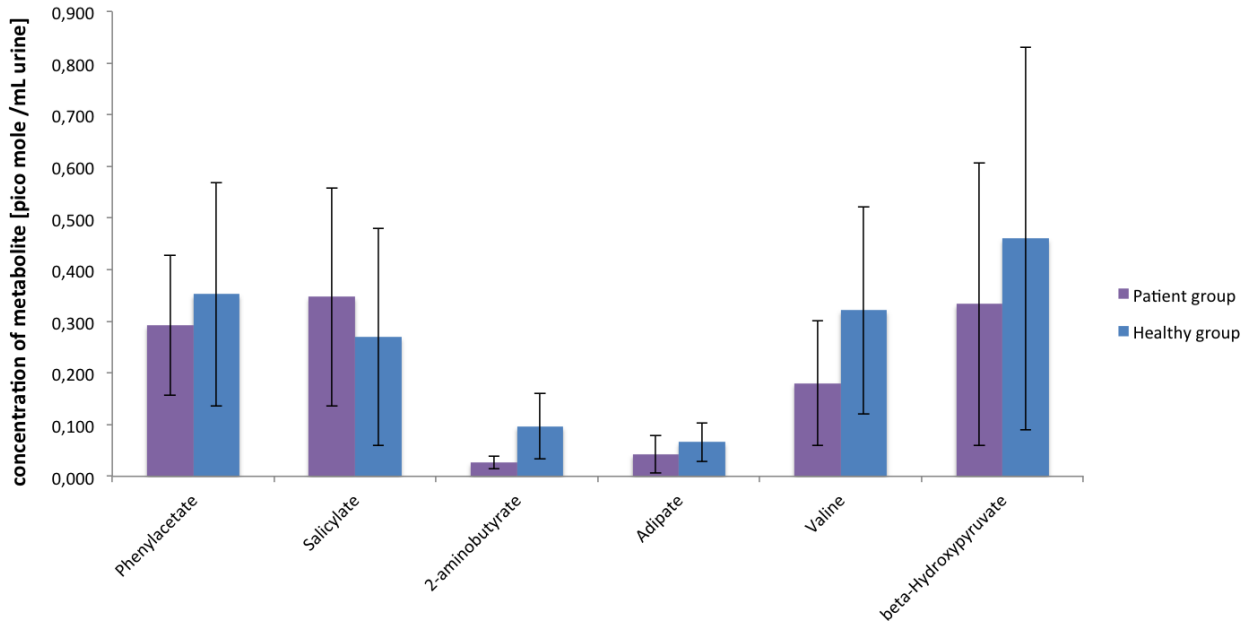


Figure 3.19: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

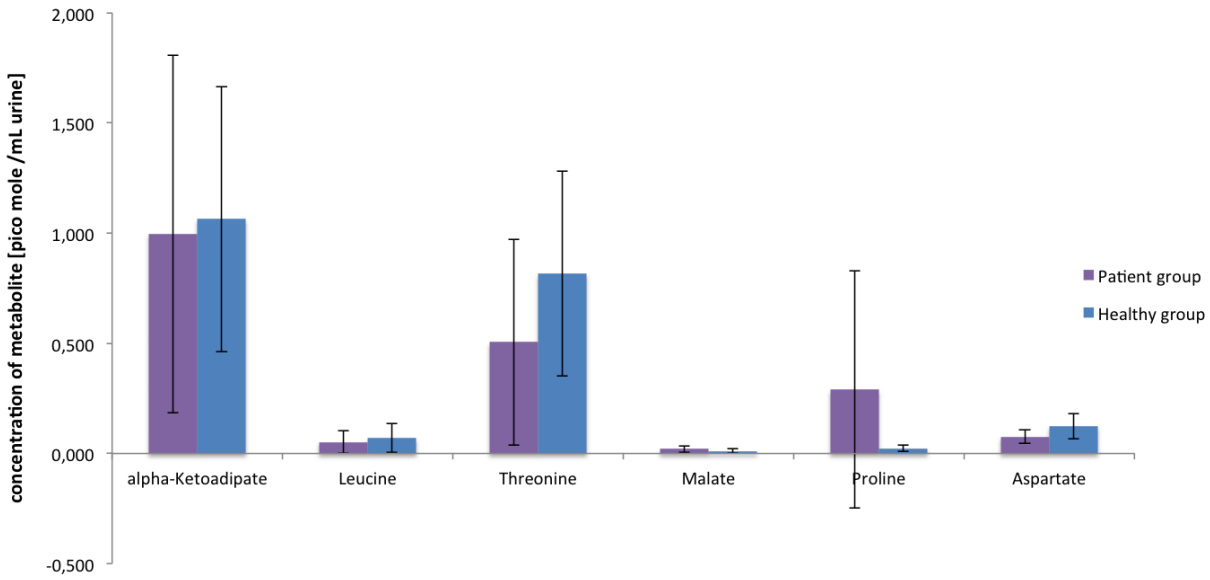


Figure 3.20: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

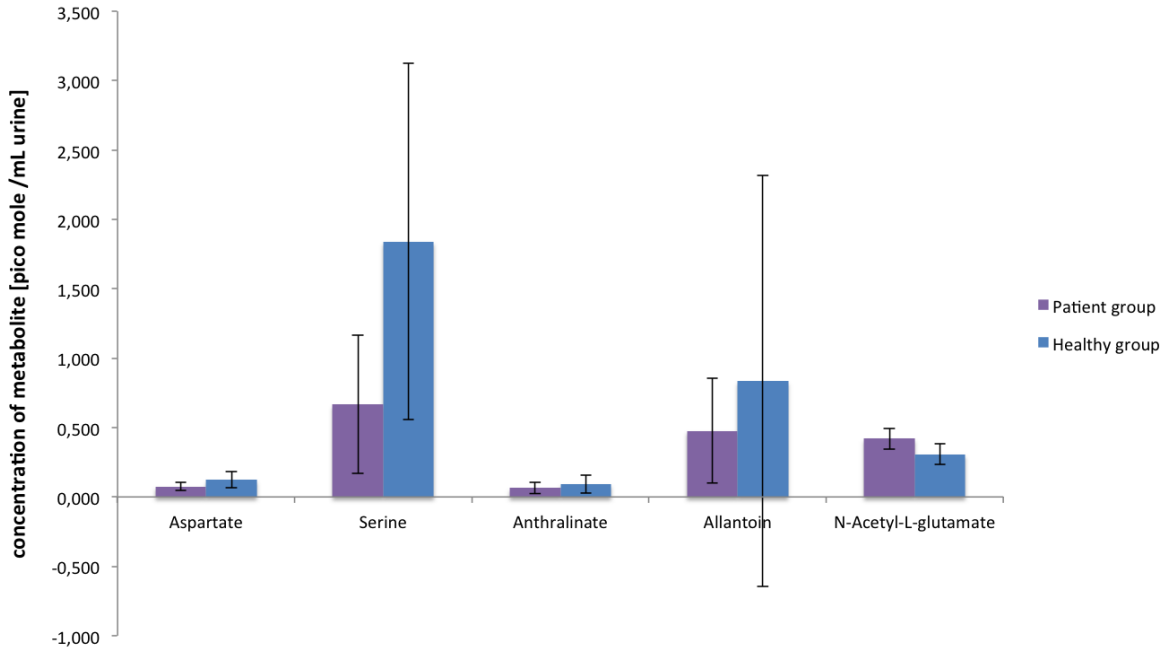


Figure 3.21: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

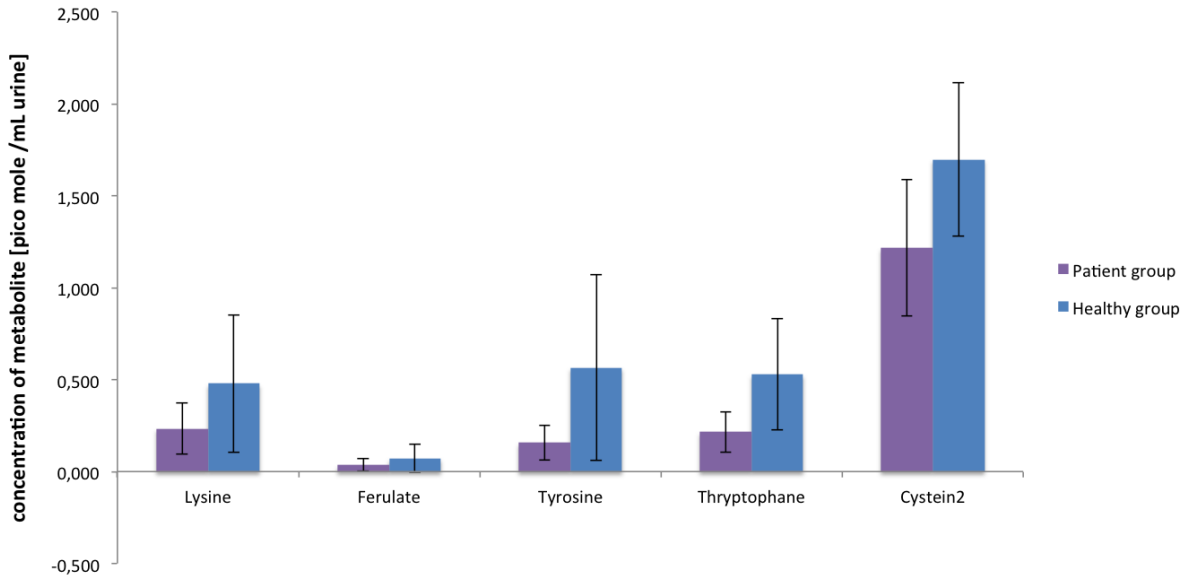


Figure 3.22: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

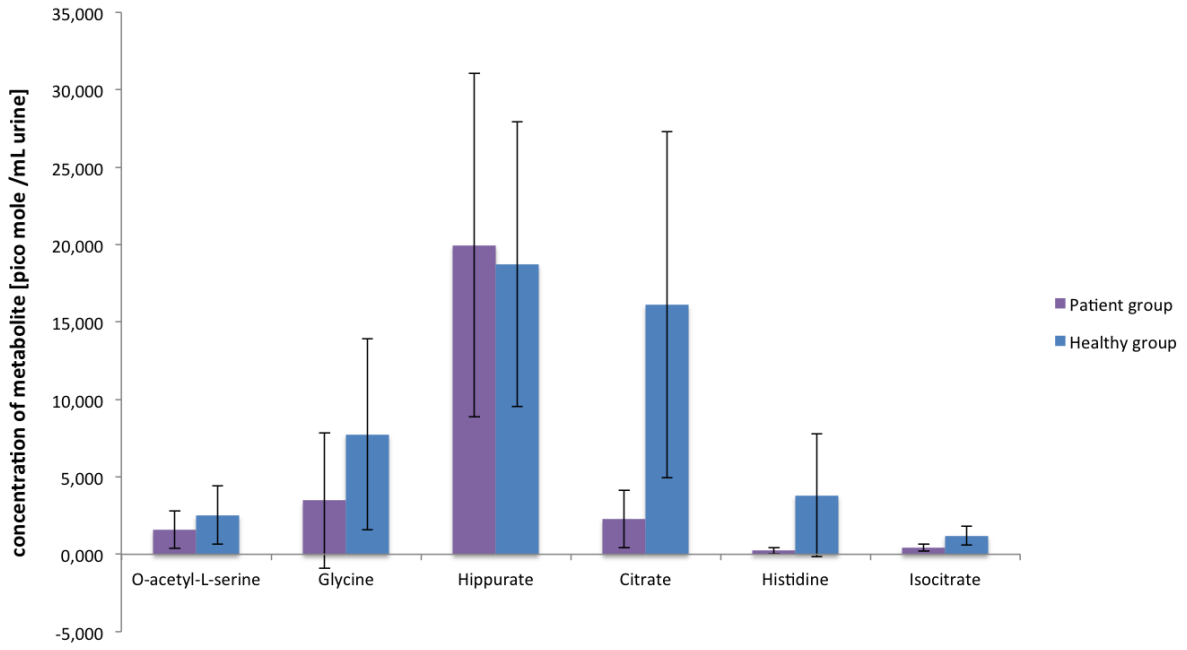


Figure 3.23: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

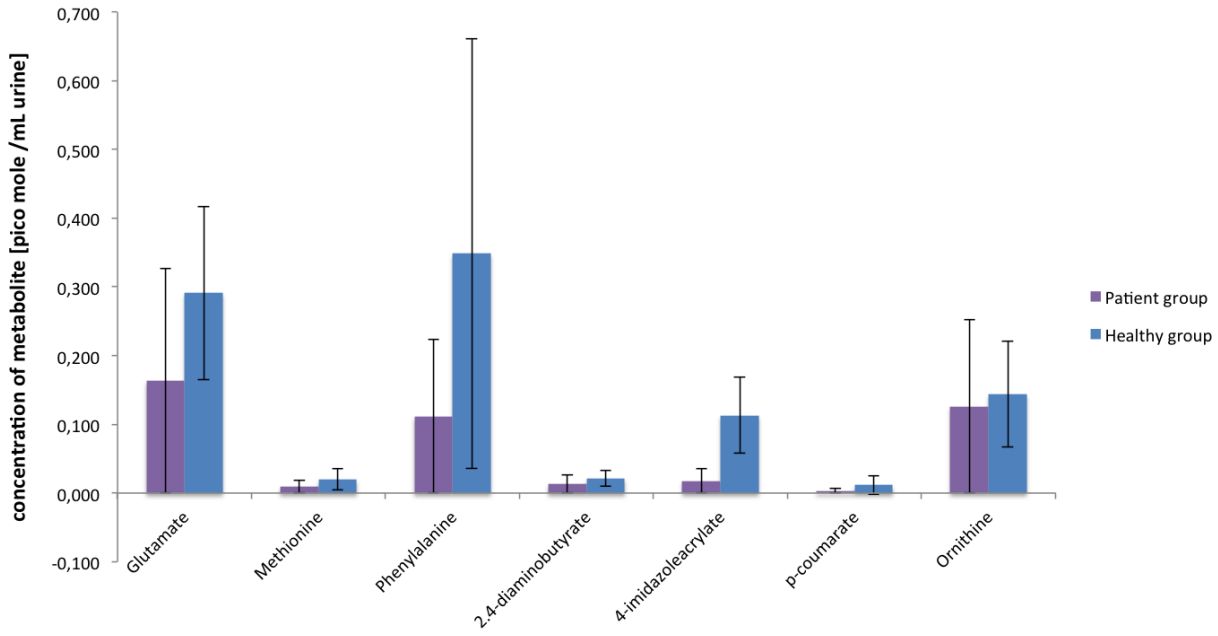


Figure 3.24: *Bar plot of average concentration (y-axis) of metabolites in sick (purple) and healthy (blue) on the x-axis*

Evaluating the bar plots visually without considering the error bars, proline, lactate, benzoate, salicylate and N-acetyl-L-glutamate was observed to have higher concentration in the CKD group. The amino acid proline was observed to be at a higher level in CKD patient, for all the three evaluated methods. Further literature search was therefore proceeded on proline. Tizianello et al. studied the net amino acid and ammonia metabolism across the kidney in subject with normal renal function and in patients with chronic renal insufficiency [54]. The amino acid analysis was performed by an automated ion-exchange chromatography (not coupled to any MS). They discovered significant differences in urinary excretion only for 3-methylhistidine and proline. Instrumental development has been achieved since their study, however it is very interesting that proline was also found to be excreted in higher amounts.

PCA-data analysis

PCA was applied to detect structure in relationships between the variables (metabolites). Figure 3.25 shows the score plot containing all samples, representing 83 % of the variations in the data. It is observed some grouping of the samples from CKD patients (pink color), while the control group (purple color) are spread over the plane. The loading plot is dominated by citrate, glycine and hippurate. It was observed that citrate is negatively correlated to the sick samples. This means that the concentration of citrate is lower in the sick group compared to the healthy group, which was also observed in the bar plot (Figure 3.23). Citrate, glycine and hippurate dominates the loading plot due to their

magnitude. Since it is already observed that these three metabolites are different in the two groups, it was desirable to investigate new trends. It was therefore decided to remove this three metabolites for further analysis. The influence plot, see Appendix A.9, shows residual variance and leverage. It is observed that sample 114 and sample 6 shows high residual and leverage. Comparing their values in Table A.11 and A.12 in Appendix it is observed that the result in samples 6 was different from the other values, which might due to insufficient sample preparation. Sample 6 was therefore removed and new PCA plot was made.

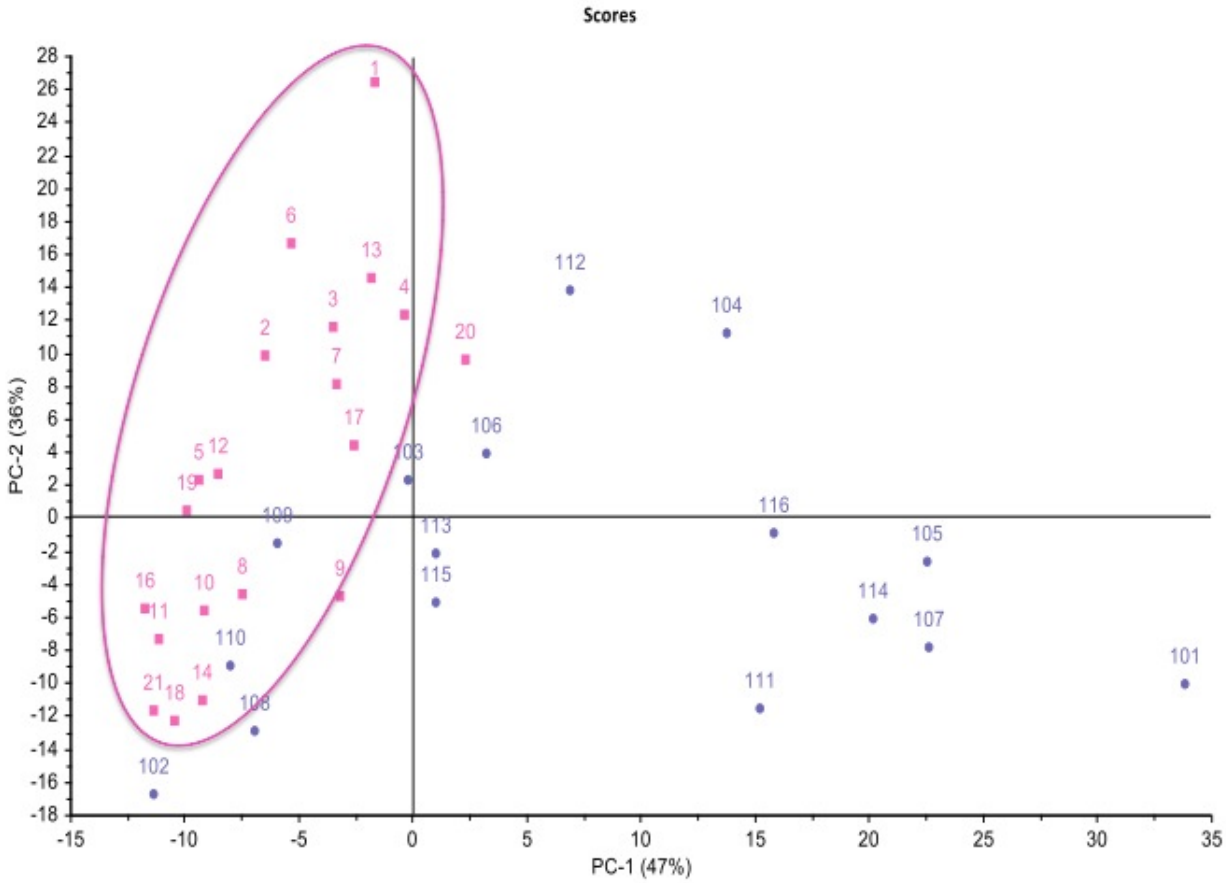


Figure 3.25: *Score plot of the urine samples. Sick are represented as pink and healthy are represented as purple*

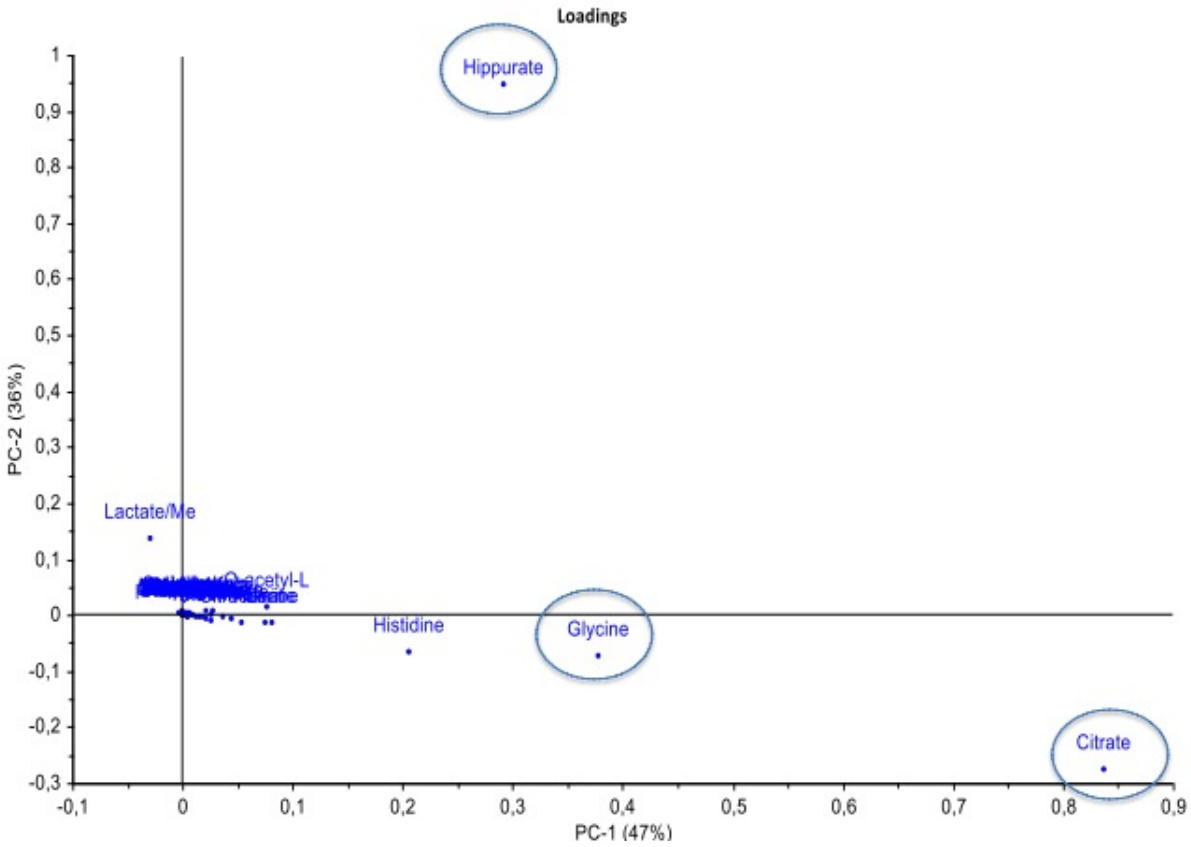


Figure 3.26: *Loading plot of the urine samples. Sick are represented as pink and healthy are represented as purple*

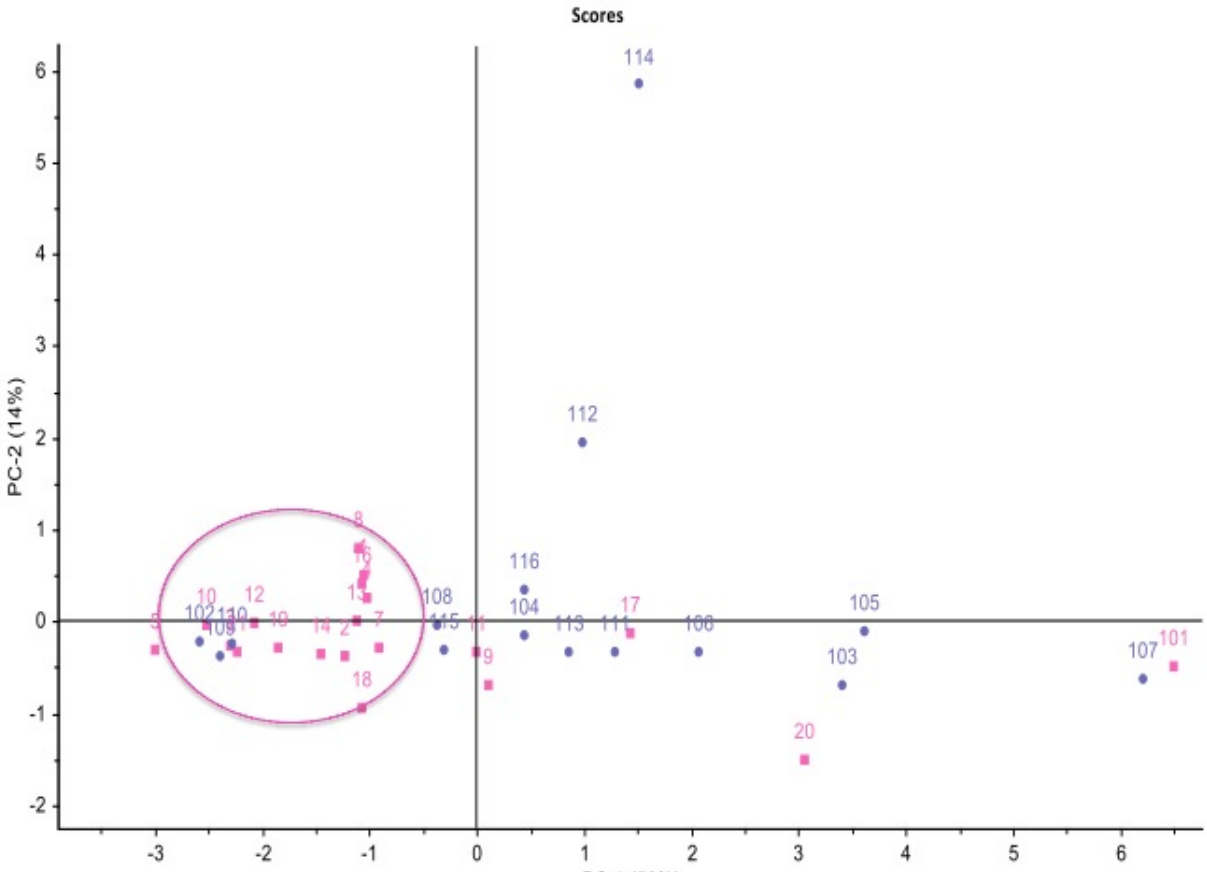


Figure 3.27: *Score plot of the urine samples. Sick are represented as pink and healthy are represented as purple*

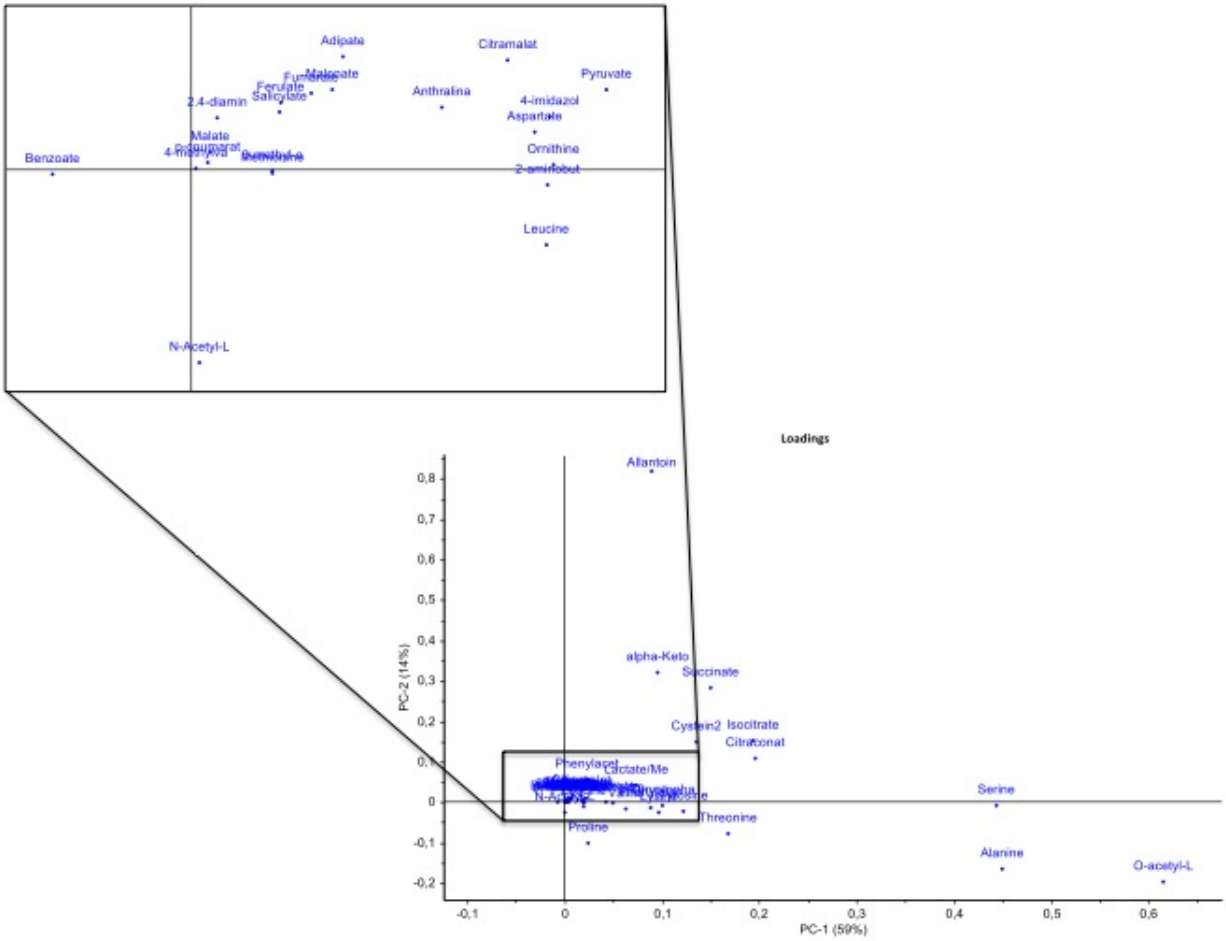


Figure 3.28: *Loading plot of the urine samples. Benzoate is contributing to the clustering of the CKD patients in Figure 3.27*

Figure 3.27 shows the new score plot after removal of sample 6 and the metabolites glycine, citrate and hippurate. The plot presents the same trend as the first score plot. The CKD subjects are slightly clustering together while the healthy group are spread over the plane. The loading plot now shows a more clear distribution of the metabolites, but the high number of metabolites makes it difficult to distinguish each metabolite.

By again comparing the score plot with the urine samples and loading plot with the metabolites, benzoate was found to be higher correlated to the sick group. This is the same result as found in using during the analysis on GC-SQ-MS. An explanation of the appearance of benzoate in the sick group, was not found in any literatures. Benzoate can be used in treatments for individual with urea cycle defects. Careful administration of benzoate in the diet would lower the level of ammonia in the blood [47]. During the urine sample collection, the patients were taking a great amount of medications which might effect the results. Due to limiting knowledge about pharmocokinetics and drugdelivery, it is not possible to conclude if benzoate might be a degradation product from the medications.

The overall results from the three GC/MS methods

The overall results from the three GC/MS method are presented in Table 3.8. The metabolites are found to be less excreted in the urine from chronic kidney disease patients, however some metabolites such as proline and benzoate were observed at higher concentration in chronic kidney disease subjects. The derivatization methods silylation and alkylation gave as expected different results. The GC-SQ-MS and GC-QqQ-MS analysis both showed the same appearance of the metabolites in CKD subjects and controls, except from lactate and O-acetyl-L-serine which were opposite. The error bars of lactate/methylglyoxal and O-acetyl-L-sering are however not observed to be significantly different. Further evaluation of the differences is therefore omitted. Absolute quantitative GC-QqQ-MS analysis has higher sensitivity compared to semi-quantitative analysis with GC-SQ-MS. However, GC-QqQ-MS do not perfoms search for unidentified peaks, as the GC-SQ-MS does. A combination of the two methods is therefore recommended for further studies. On the basis of this, if a perfect biomarker for renal failure is discovered, it is recommended to use absolute quantitative analysis.

Table 3.8: *Metabolites found in the analysis approach using alkylation (MCF) and silylation (MSTFA)*

	MCF GC-QqQ-MS		MCF GC-SQ-MS		MSTFA GC/MS	
	sick	healthy	sick	healthy	sick	healthy
Amino acid						
Proline	✓		✓		✓	
Phenylalanine		✓		✓		
Serine		✓		✓		
Tryptophane		✓		✓		
Valine		✓		✓		
Glutamine				✓	✓	
Aspartate				✓		
Glycine				✓		
Histidine		✓		✓		
Tyrosine		✓		✓		
Nonamino acid						
Citric acid						✓
Glycolic acid						✓
Other:						
Glycolic acid						✓
Lactate	✓			✓		
Benzoate	✓		✓			
Salicylate	✓					
N-acteyl-L-glutamate	✓					
O-acetyl-L-serine		✓	✓			
Citrate		✓		✓		
Citraconate		✓		✓		
Pyruvate		✓				
trans-4-hydroxy-L-proline					✓	
Uracil						✓
Hypoxanthine						✓

Other remarks and future perspectives of GC/MS

The quality of the urine samples

After collection of urine samples from patients and healthy volunteers the samples were kept in the freezer at -32°. Before each study samples were brought to room temperature. The samples were collected two years ago (2010) and how the quality of the samples have changed over the last two years was not investigated in this study. For future work storage time and measurements of metabolite change over time should be investigated. Saude et al. investigated urine stability with focus on metabolites and how storage time and temperature would effect the quality of the urine [55]. They concluded that urine kept at -80° over a period of four weeks best reflected the original metabolite concentration. The metabolite change was not found to be significant different. However, this was a four week study and the samples were kept at a much lower temperature and any conclusion of metabolites stability can not be drawn. It would therefore be interesting to investigate the concentration of the metabolites at the collection day and how it changes over a greater time range when kept at -32°.

Evaluation of the subjects and clinical characterization

The urine samples were collected from 36 different subjects, see clinical characteristics in Table 3.1. As presented in the table subjects are different in age, gender and diagnosis.

Due to this it is expected that the subjects will have metabolite variation in the urine. At the collection time the control group was considered healthy with no underlying chronic illness or major medical problem. However, it is now know that sample 102, 106, 108 and 112 have a medical condition. Still, the subjects were not left out from the study. There was also no attempt to normalize diet or other factors as medication. The chronic kidney disease patients are on a number of different medications due to their critical medical conditions at the time of the urine collection. It is therefore uncertain how medications would effect the metabolism in the body and the variation of the metabolites in the urine for each of the CKD patients.

For future work it is recommended to collect a larger number of urine samples for analysis, as 36 samples will only provide some information about metabolite concentration. A greater number of patients and urine samples per patient must therefore be collected for increased knowledge of the metabolite changes during renal failure. The collection of urine should also be performed while the CKD subjects are of any medications. The patient group should also consist of subjects with same diagnosis and other conditions such as obesity and diabetic. Future studies could also be divided into two group based on gender, as it is known that these factors contributes to metabolite variations [25]. Metabolite variations must be taken into account when using urine for analysis. Urine is not in constant circulation through the body like blood [25], and differences in gender, sampling time and individuals must be considered. Partial least square discriminating analysis (PLS-DA) have become a very common method for discrimination variations in age and gender [34]. It is therefore recommended to use PLS-DA for future work.

Sample preparation

The steps during sample preparation includes removal of proteins, urease degradation, centrifugation, vacuum drying, transfer from one tube to another etc. All of the steps will contribute to loss of metabolites.

There is much sediment in urine, which includes proteins, salt, cellular debris and compounds absorbed into the sediment or insoluble in the urine after storage in a freezer [20]. Removal of proteins was performed by adding an organic solvent, MeOH, to the urine sample. Other solvents such as acetone, acetonitrile and ethanol have been investigated, concluding that methanol and water are the best solvents for maximize the peak area of the metabolites [20]. The ratio of methanol to water to urine 3:1:1 is observed to be the best for maximize the peak area of the metabolites [20]. This could be applied for future studies.

The centrifugation and transfer-steps contributes to loss of metabolites. This study did not add any internal standard (IS) during the preparation steps, so the loss of metabolites could not be evaluated on the basis of the IS. For future studies it is recommended to

add internal standard when adding methanol to correct for loss during centrifugation and transfer-steps and other preparation steps

After vacuum drying the tubes with dried metabolites were kept in the freezer prior to derivatization. During the first derivatization step sodium hydroxide (1M NaOH) is added to dissolve the dried metabolites. Loss of metabolites in this step could occur if the tubes were dried for to long, yielding poor dissolving or alkaline hydrolysis could occur.

Urease degradation

Urease degradation was performed prior to silylation and analysis on GC-SQ-MS system. The enzymatic treatment of urea was observed to be successful for almost all of the samples (Figure 3.29). The enzymatic treatment is therefore recommended and must be performed for future investigation of urine samples during the silylation step for eliminating urea. Interestingly, Kind et al discovered by comparing urease treated urine against non-urease treated urine that metabolites such as citrate, succinate, tyrosine and glycerol, were found to be greatly diminished with urease treatment [21]. In this thesis tyrosine was observed to be present in 12 out of 21 in the CKD group and in 6 out of 16 in control group. Citrate and succinate were not observed in the samples, however the metabolites were found in urine samples treated with alkylation. On the basis of this, urea degradation must still be performed prior to silylation to render the minor metabolites in the urine samples accessible to derivatization reagents. Evaluation of the two different derivatization techniques is therefore of interest.

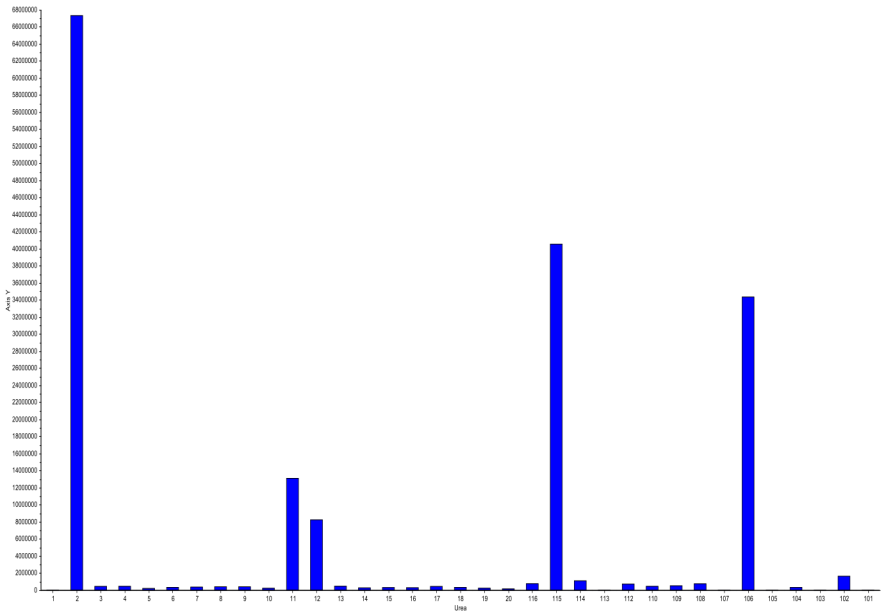


Figure 3.29: *The peak area up to 6.8×10^7 (y-axis) of urea in urine samples (x-axis)*

Evaluation of alkylation and silylation

During this study alkylation with methyl chloroformate (MCF) was found to be successful when evaluating metabolite differences between CKD subjects and controls. Silylation with MSTFA was not observed to be that successful and this might due to poor substitutions of $Si(CH_3)_3$ and that the reaction was not driven to completion [31]. Silylated derivatives are also found to be less stable and the samples must be analyzed immediately after derivatization reaction [31, 56]. According to literatures and studies performed by Villas-Bôas et al the analytical performance of silylation have been found to be unsatisfactory for amino and non-amino organic acids and nucleotides [31]. Koek et al reported that, in general, GC-inlet liners required changing after 20 samples had been injected as well as occasional removal of a small section of the front end of the analytical column to restore performance [33]. This was taken in consideration during this study. The urine samples were injected in the runs over a period of three days, however the column was not cut prior to each run. However, blank samples were run prior to the samples. As these showed no sign of contamination it believed that the run of the samples were sufficient. The samples derivatized with MCF were injected into the GC/MS under pulsed splitless injection while the MSTFA derivatized samples were injected in split mode. This would results in a larger proportion of MCF samples would reached the column compared to MSTFA samples [31] and might explain lower detection limit for the derivatives.

Statistical analysis

Statistical analysis was performed using T-test calculations in order to see if the values are statistical significant different. T-test can only be used when the groups have the same number of samples, so the calculations were only performed on the results from the first method with alkylation and application of GC-SQ-MS where five samples from each group was studied. For calculation details see Table A.6 in Appendix A.6.

On a 95% significance level it was found that glutamine, citrate, histidine, benzoate o-acetyl-L-serine, glycine, tryptophane, ferulate, isocitrate, serine-2, aspartate, aspergine, citraconate, succinat, methyl glyoxal and lactateare statistically significant different. However phenylalanine was not found to be statistically different for the two groups even though the bar-plot shows a difference. This might due to the standard deviation of phenylalanine is very big; $(7.70 \pm 4.92)\times10^4$ and effects the results. Metabolites found not to be statistical different were hexacloroethane, leucine, isoleucine, tyrosine, proline, cystein, lysine and putrecine.

This study did not use analysis of variance for statistical analysis, due to limited knowledge and usage of the method. However, as analysis of variance can be used for groups with different number of subjects [38] this could be performed in further studies.

Variations in the GC/MS-system

Variations in the GC/MS can provide analytical errors. The performance of the GC/MS instrument will vary within a sequence of analysis, during maintenance intervals, contamination of the ion source and after tuning where the detector sensitivity is changed [30]. The check sequence is run prior to the run to evaluate variations in the GC-MS. For further studies the samples could be reanalyzed by injecting the samples several times to observe any variations.

3.6 Results from method 4: LC-HILIC-MS and non-targeted analysis

Data analysis was performed on old data results obtained from a run on LC-HILIC-MS. The results investigated in this section are from the negative ionization (ESI^-). Injection of the samples into the TOF-MS results in generation of raw-chromatograms (wiff-files) that needs to be processed into CEF-files before uploading the results into MPP. This process was already completed and the following sections presents the filtration steps.

The software MPP enables statistical analysis of large datasets and contributes to detection of interesting masses. The data files from negative ESI, were imported into to the MPP program and samples were labeled with the parameters H (healthy) and D (disease). The data files are all independent and represent different subjects and were set to be 'non-average' (Average is usually used on replicates). MPP shows the results as intensity values for visualization and analysis, where each line represent one entity (i.e mass). A total of 5024 entities was detected and the profile plot are shown in Figure 3.30.

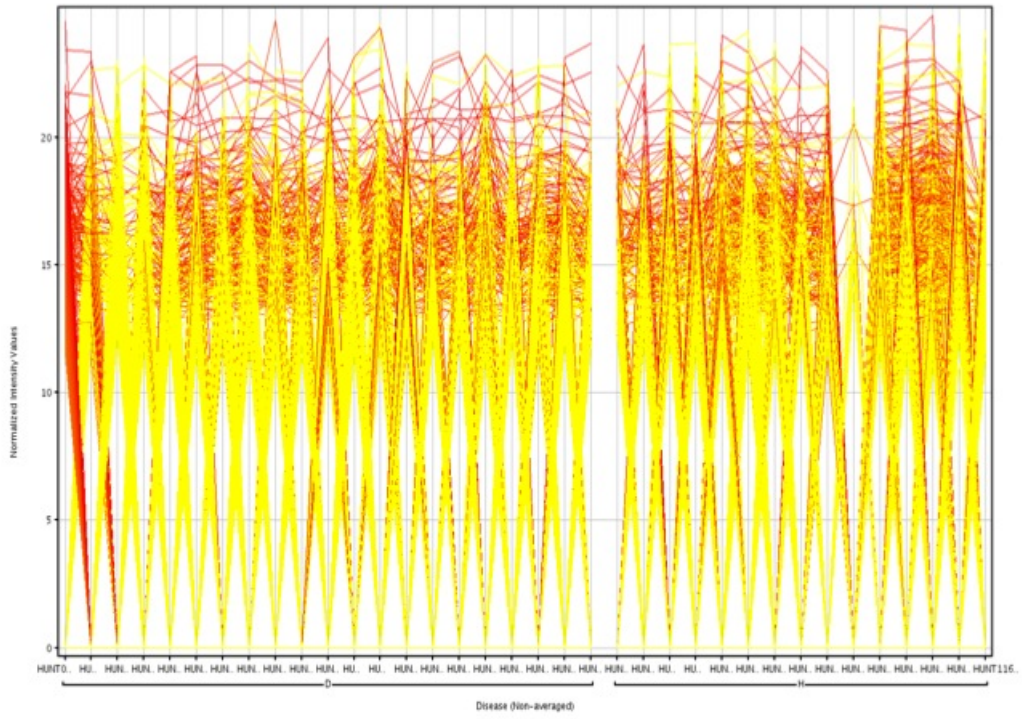


Figure 3.30: *Profile plot of 5024 entities found in all samples from negative ESI. The x-axis shows the samples with the CKD group on the left and control group on the right. The y-axis presents intensity values from 0-20*

3.6.1 Quality control (QC) of the samples

The first step was to evaluate the data samples in a PCA plot. The' QC on samples' displays a PCA plot to help verify the quality of the samples. This plot often shows any samples that might have exceptionally different values from the rest, or if a sample was assigned as an incorrect parameter value. The PCA plot in Figure 3.31 shows that all of the red dots (representing the sick group) except from three samples, appears at the same place in the plot (marked with the red circle). The gathering of these at almost the same place is an very interesting observation and this shows that the masses found in the CKD samples clusters due to similar metabolites.

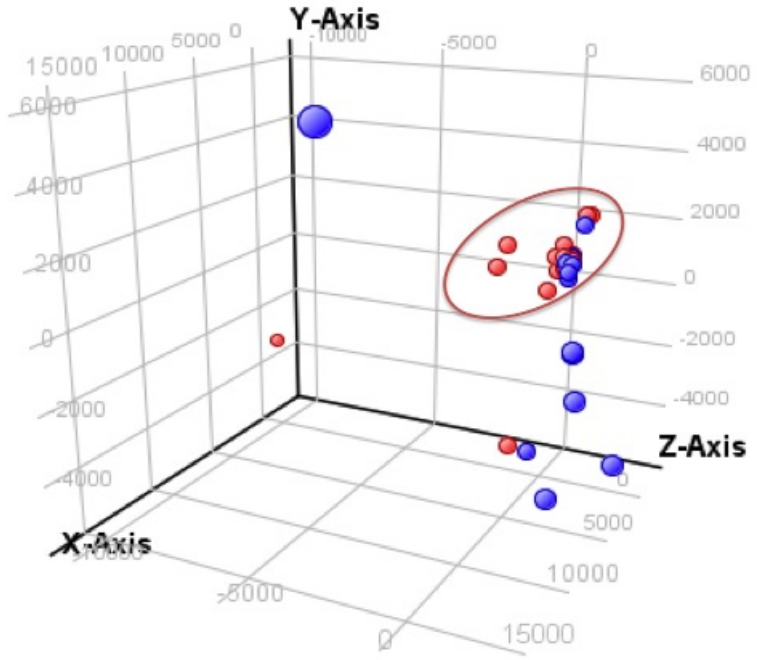


Figure 3.31: *Three dimensional PCA-plot of all of the samples. Red dot represent the sick group and blue represent the healthy group*

3.6.2 Filter parameter: Filter by Flags

'Filter by Flags' allows filtration of data based upon Present, Marginal or Absent flags. The focus in this thesis was to investigate differences between the two groups. Identifying masses only present in one of the groups is therefore the major aim in this step. The parameters was set at absent and to retain entities which at least 100% of the values in any 1 out of 2 conditions have acceptable values. The profile plot of all entities present after the filtration is shown in Figure 3.32. 3556 masses were able to pass the filter whereas 1468 masses were rejected for further analysis.

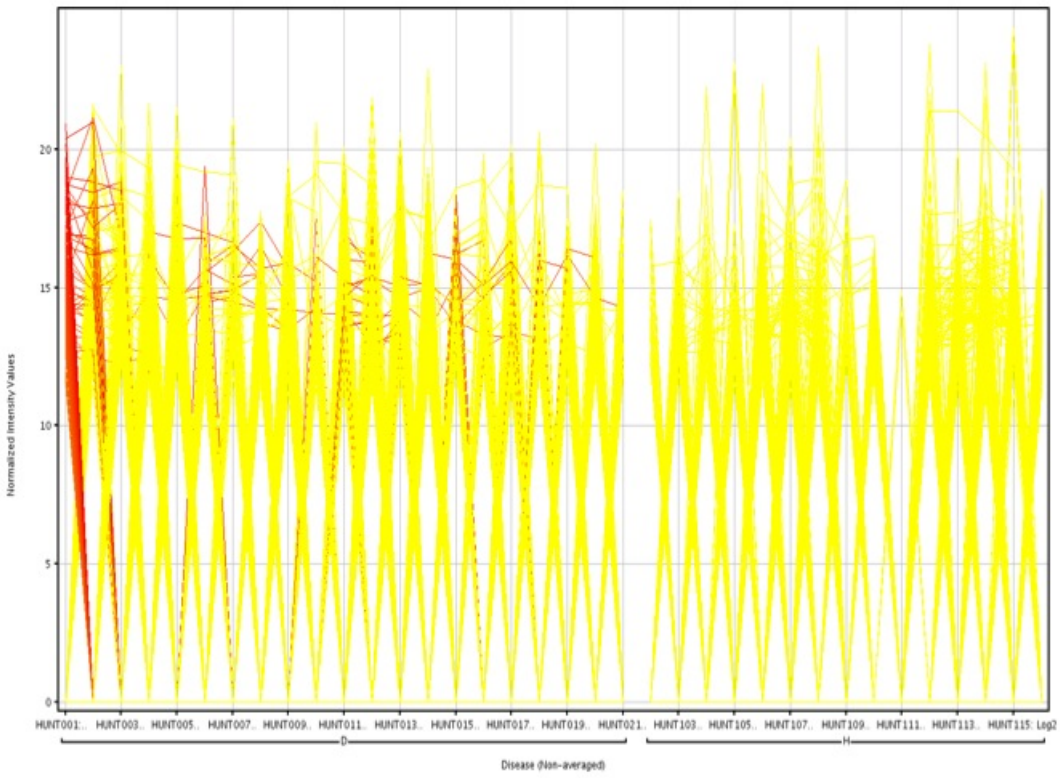


Figure 3.32: *Profile plot after filtration by flags, displaying 3556 out of 5024 entities. The x-axis shows the samples with the CKD group on the left and control group on the right. The y-axis presents intensity values from 0-20*

3.6.3 Fitration by abundance

Filtration based on abundance was performed on the 3556 masses. The filter rejected low-intensity masses. The cut range was set to have a percentile between 20 and 100. The filter passed 1865 masses whereas 1691 masses were rejected. Figure 3.33 shows the profile plot containing 1865 entities.

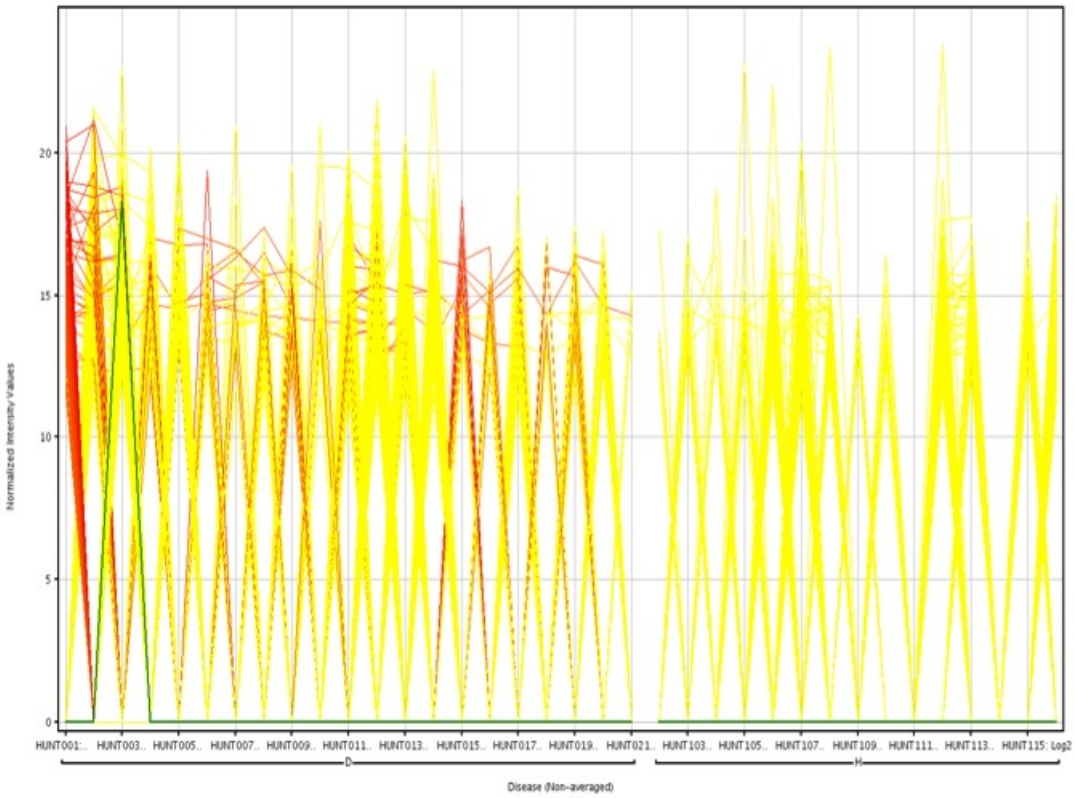


Figure 3.33: *Profile plot after filtration by abundance, displaying 1865 out of 3556 entities. The x-axis shows the samples with the CKD group on the left and control group on the right. The y-axis presents intensity values from 0-20*

3.6.4 Filtration by frequency

Filtration by frequency allows filtration based on frequency with which any compound appears in each sample in the experiment. Figure 3.34 displays 6 of 1865 entities where at least 50% of samples is in exactly 1 out of 2 conditions. The "zig-zag pattern" shows a drop in intensity of each masses, and there are no masses found in the healthy group (H). Table 3.9 presents the six compounds obtained from MPP and identification of the compounds was performed using databases in Metlin. Metlin database provides a library of up to 55,214 metabolites [41] and it is possible to do simple search based on the mass as well advanced search based on molecular feature and mass. However, the six compounds were not successively identified in Metlin as the database search gave different suggestion (based on a ±30 ppm). For example the mass of 206.0346 got 6 metabolite hits and the mass 164.0664 got 17 possible metabolites (Table 3.9)

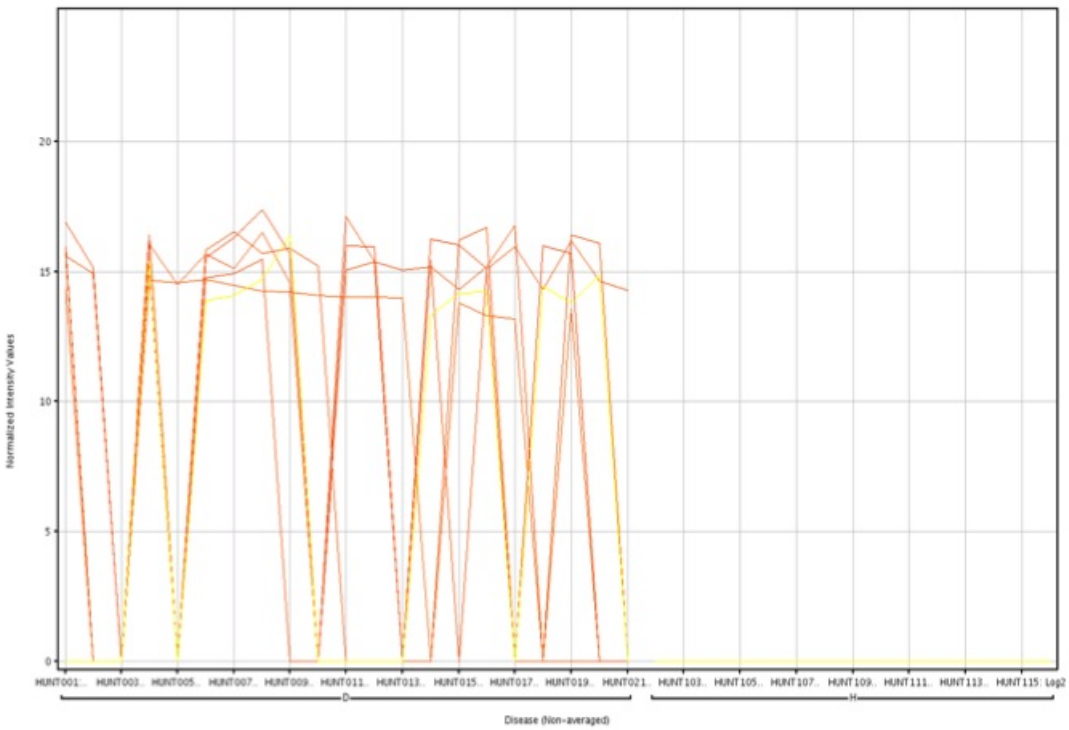


Figure 3.34: *Profile plot after filtration by frequency, displaying 6 of 1865 entities. The x-axis shows the samples with the CKD group on the left and control group on the right. The y-axis presents intensity values from 0-20*

Table 3.9: *Metabolites suggested by Metlin*

g/mol	m/z	Mode	Metabolites (±30 pmm)
138.0293	137.0213	Negative	5-Hydroxypyrazinamide N-Monomethyl-2-aminoethylphosphonate
164.0664	163.0587	Negative	L-rhamnose, Viburnitol, Quercitol, β-D-Fucose, 2-Deoxyglucose, 1,5-Anhydrosorbitol, D-quinovose, 6-Deoxy-L-galactose, L-fucose, 2-deoxy-D-galactose, Alpha-D-Fucose, Digitoxic acis, 6-Deoxy-D-Glucose, L-fucolose, L-Rhamnofuranose, L-Rhamnulose
206.0384	205.0302	Negative	Benzthiazuron, Pterin-6-carboxylic acid, 4-(2-Aminophenyl)-2,4-dioxobutanoic acid, 1-Nitro-5,6-dihydroxy-dihydronaphtalene, 2-Formaminobenzoylacetate, 2-amino-4,7-dihydroxy-8-methylcoumarin
294.0525	293.0587	Negative	Estazolam, Aplysin, Laurinterol, Filiformin, Cyclic de-hypoxanthine futalosine
264.0426	263.0343	Negative	Frutinone A, Ethidimuron, (1R,2R)-3-[(1,2-Dihydro-2-hydroxy-1-naphthalenyl)thio]-2-oxopropanoic acid
504.2725	503.2682	Negative	Ceanothine B, Taxuyunnanin C, Phorbol 12,13-dibutanoate

MassHunter Qualitiative Analysis software was used to see how reliable the results obtained from MPP were. The software provides different functions in search for specific masses (m/z). Extract Chromatograms allows the user to search for masses. One urine sample from each group was randomly picked, 116 and 9 respectively, and $m/z = 263$ was searched for in both samples. As Figure 3.35 and Figure 3.36 shows, a compound with $m/z = 263$ was found present in both samples.The discovery of finding the mass present in both samples, indicates that the filtration steps performed by MPP was not optimal. One limitation with using MPP is that while setting the acceptable range all the values below the range will be set as zero. For example, if one filtration step is set to be from 2000 (e.g. abundance), values below 2000 such as 1900 will be set as zero. This might be the explanation of why $m/z = 263$ is found only in the sick group by using MPP, while the extracted chromatograms presents a different result.

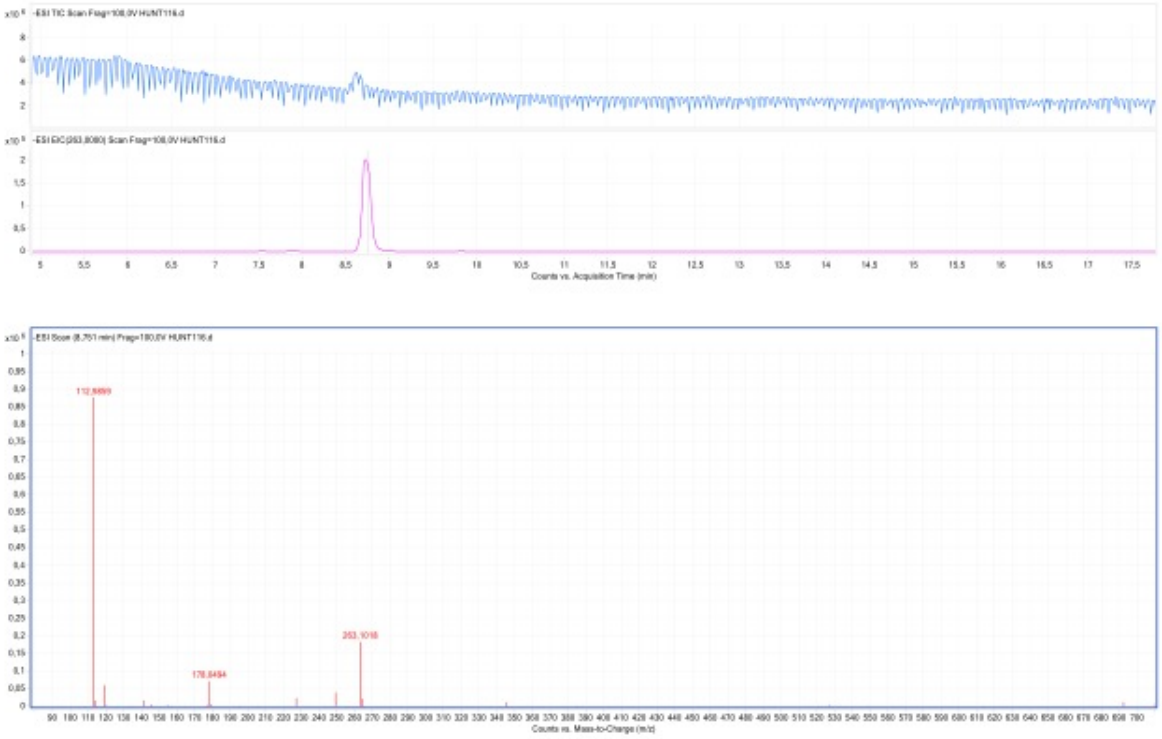


Figure 3.35: *TIC (upper) and extracted chromatogram of* m/z = *263 in sample 116. The plot at the bottom shows is the scan*

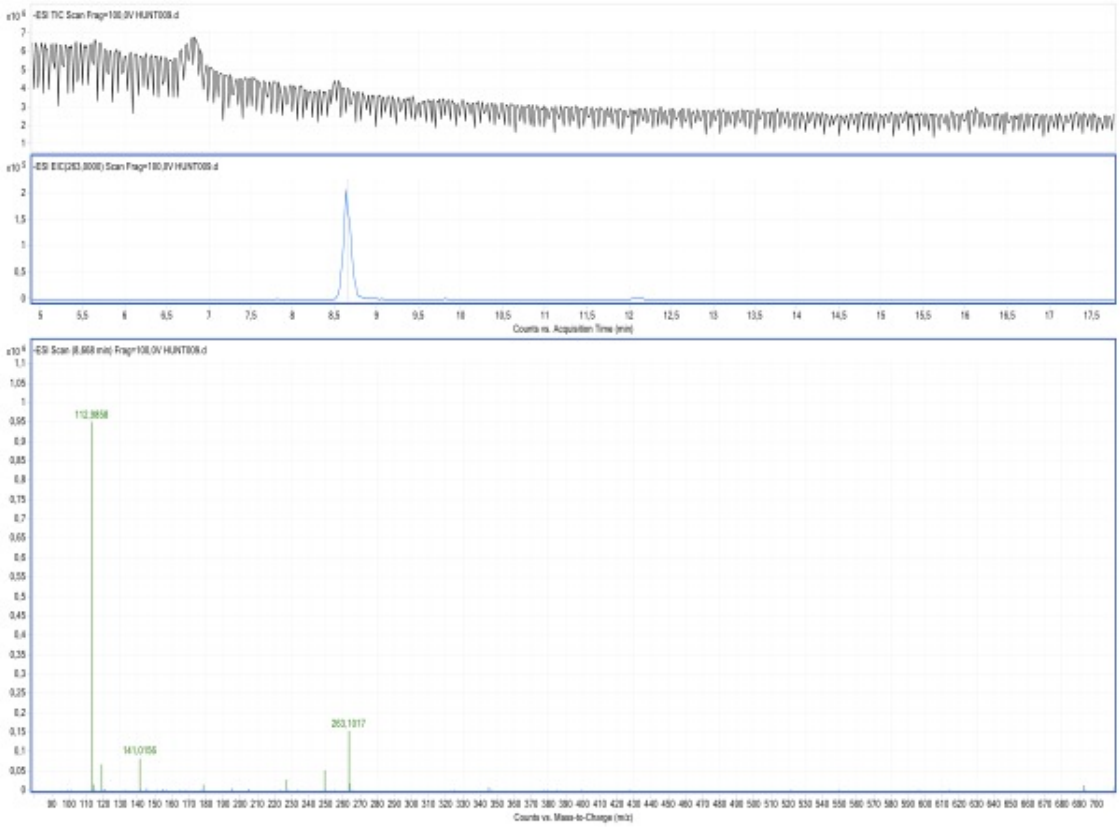


Figure 3.36: *TIC (upper) and extracted chromatogram of* m/z = *263 sample 9. The plot at the bottom shows is the scan*

As the MPP filtration was not found to be successful in this study, due to limited experience with the method, a different approach was performed. The aim of this method was to observe peak differences between the groups as well as observing if there were some peaks that would appear at higher intensity (i.e peak height) in only one of the groups. Figure 3.37-3.40 presents the total ion chromatography (TIC) and the base peak chromatography (BPC) of eight samples with masses between 130 and 200. The TIC represents the intensity across the range of masses being detected and as the figures show, provides limited information when multiple analytes elute simultaneously. The BPC shows only the most intense peaks in each spectrum, resulting in a clearer view of the analytes. As the BPC shows there are no peaks that are found to more present in one of the groups. The BPC also shows individually differences in the groups, which is expected.

The overall results from this method is that the MPP filtration did not manage to provide any new features of the metabolite in the urine samples. Further studies should be proceeded with LC/MS as it is a promising approach for urinary metabolomic.

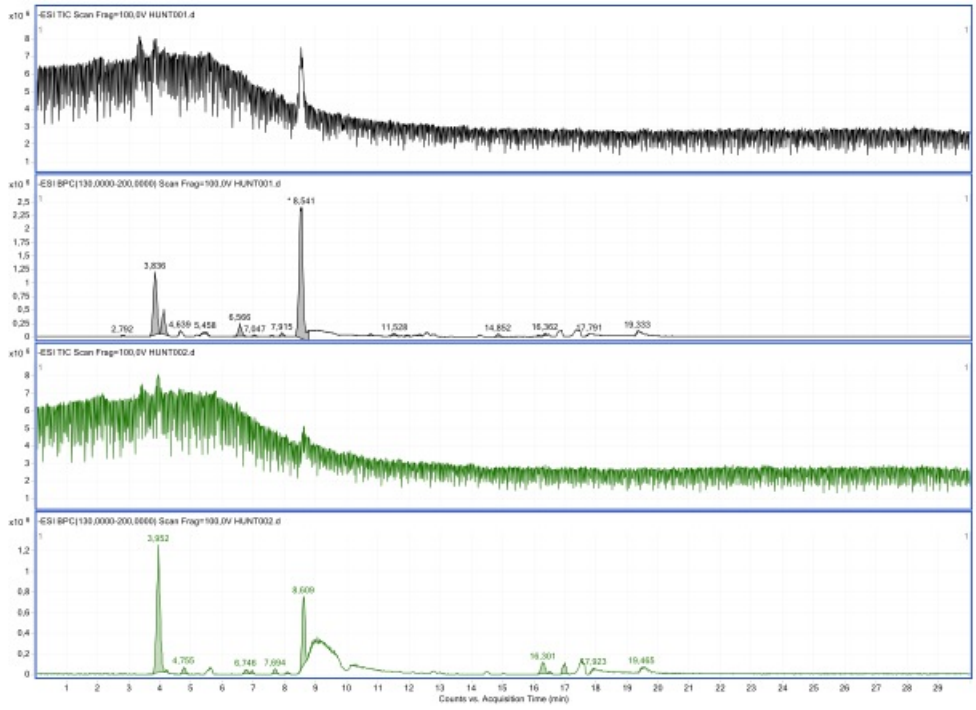


Figure 3.37: *TIC and BPC for urine sample 1 and 2*

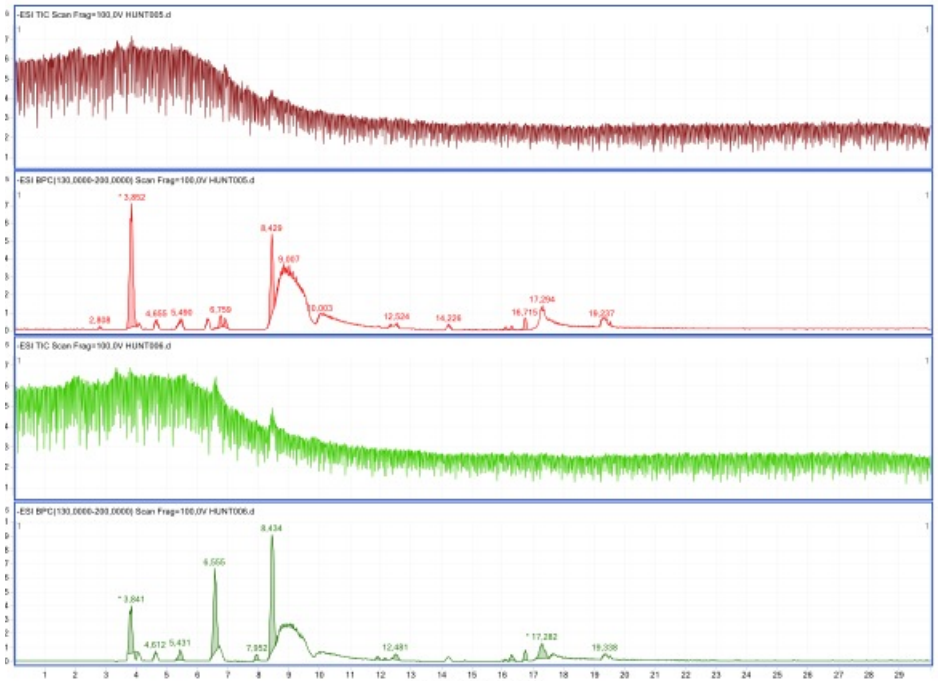


Figure 3.38: *TIC and BPC for urine sample 5 and 6*

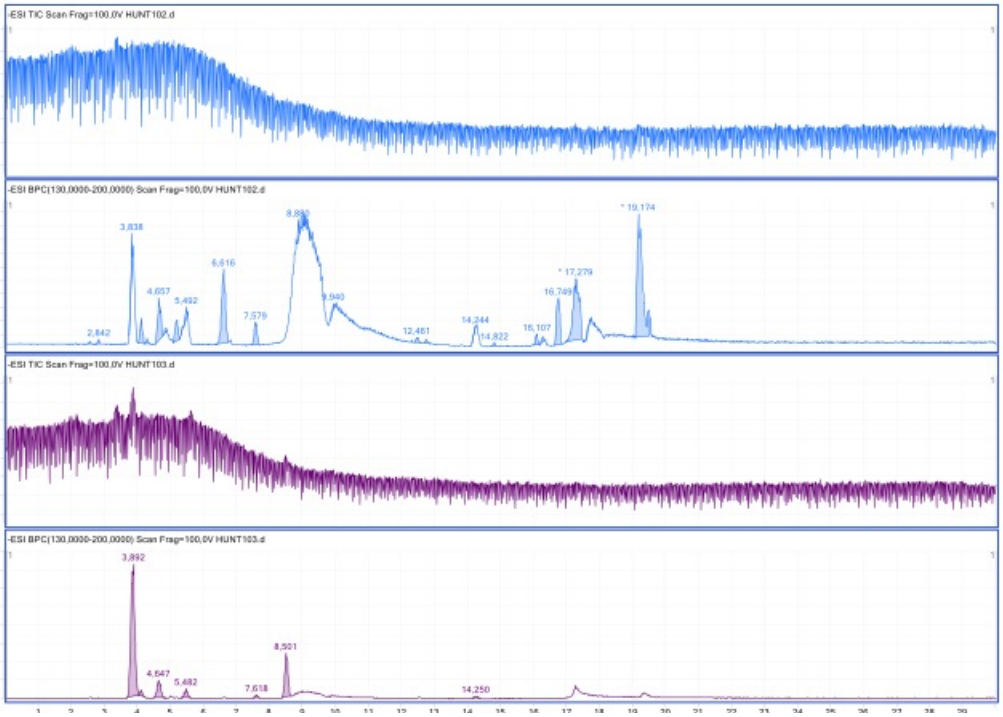


Figure 3.39: *TIC and BPC for urine sample 102 and 103*

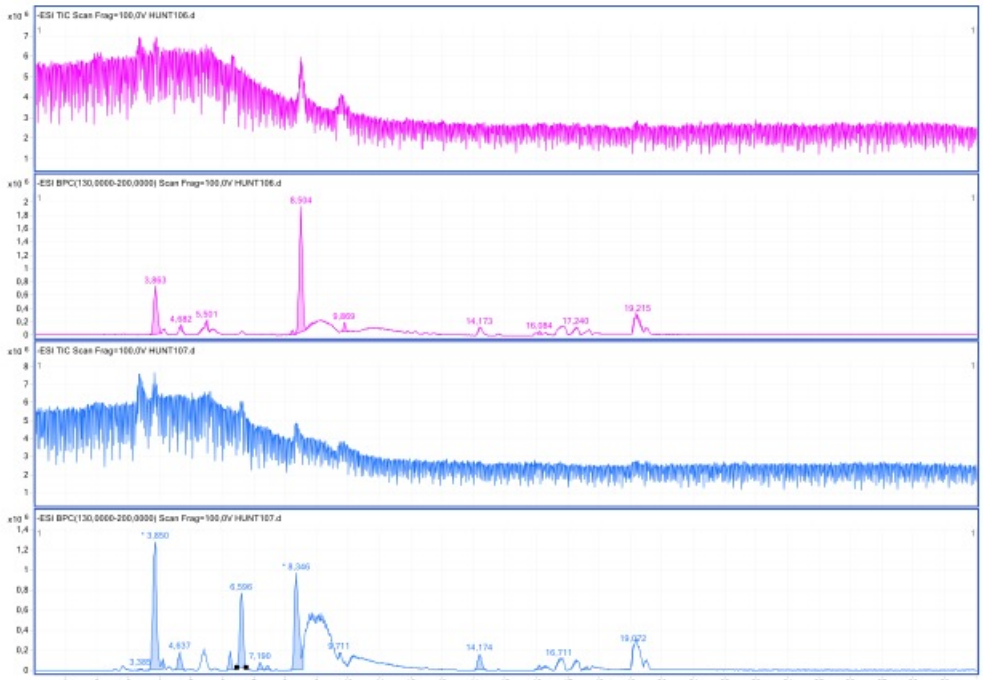


Figure 3.40: *TIC and BPC for urine sample 106 and 107*

Other remarks and future perspectives using LC/MS

Sample preparation

As the data results from the LC/MS were obtained from a previous run, no sample preparation and run on the LC-HILIC-MS was proceeded during this study. The application of LC/MS on the urine samples was found to be difficult but it is a promising approach in urine metabolomics.

The base curve chromatograms (Figure 3.37-3.40) illustrates variation in the samples. The urine samples obtained from humans exhibit a large amount of variation caused by physiological factors such as sex, state of health, age, diet, stress to mention some. This requires normalization to account for the variation and to give each sample equal importance in analysis. The most common normalization step for analysis of urine involves scaling each sample according to the creatinine content. However, this thesis uses urine samples from patients with kidney disease and from healthy persons and the creatinine levels might perturbed by the disease [34]. For further data analysis, statistical analysis such as PCA and PLS can be applied. Since the urine samples are obtained from different individuals, PLS-DA is a nice tool for discriminating for example sex, time of collection and age [34].

Variations in the instrument

As for the GC/MS, LC/MS can provide analytical errors. The performance of the instrument will vary within a sequence of analysis [30] which might yield variation in the results. The use of internal and external calibration method have been a common approach for analysis with LC/MS. External calibration method is performed by analysis of known metabolites of analyte standard. Generally, at least five-point standard curve are established for each standard analyte by plotting integrated peak areas versus concentration [22]. In contrast, the internal standard (IS) calibration is performed by addition of an IS at a constant concentration in all calibration standards and samples. Is is assumed that variations in the IS peak area are represented over the whole analysis, and therefore, the IS can be used to account for instrumental variations as well as sample preparation variations. For future usage of HILIC MS, it is recommended to use internal and external standards.

Evaluation of the LC-HILIC-MS method

The evaluation of LC-HILIC-MS analysis of metabolites did not receive the most attention in this study, due to time limit. The method is a promising approach in metabolomics. Because LC/MS is broadly applicable, it is the technique of choice for untargeted metabolomics studies and further study should be performed.

LC/MS is generally favored for discovering unknown metabolites, based on the wider range of compounds it can analyze as well as it performs injection of samples in "natural state" which is very optimal for analysis of urine and other biofluids. Analysis on LC/MS has the advantage of less preparation steps, compared to GC/MS where time consuming derivatization steps must be performed. The hydrophilic interaction chromatography is ideal for urine analysis of polar molecules and the high organic mobile phase leads yielding the separation technique compatible with ESI-MS [34].

Drawbacks in GC-MS methods is that they only target a specific class of metabolites and might only detect a small number of metabolites. In LC-HILIC-MS the samples can be run in positive- and negative mode, providing between 2500-4000 metabolites in any sample in about 30 minutes with accurate masses. However, some disadvantages are observed of HILIC. The mechanism is not entirely understood and column overloading and ion suppression can be a problem [36].

4 Conclusion

The aim of the study was to evaluate different MS-based methods in order to find new and promising biomarkers for patients diagnosed with chronic kidney diseases (CKD). Urine analysis of CKD patients and healthy volunteers were performed with targeted and untargeted methods. Targeted analysis was performed on GC-SQ-MS and GC-QqQ-MS and two different derivatization methods were evaluated, alkylation and silylation respectively. Run on GC-SQ-MS provided semi-quantitative analysis on GC-SQ-MS and absolute quantification was provided on GC-QqQ-MS (Tabel 4.1). Untargeted analysis was performed with LC-HILIC-MS results obtained from a previous run.

Table 4.1: *Metabolites found in the analysis approach using alkylation (MCF) and silylation (MSTFA)*

	MCF GC-QqQ-MS		MCF GC-SQ-MS		MSTFA GC/MS	
	sick	healthy	sick	healthy	sick	healthy
Amino acid						
Proline	✓		✓		✓	
Phenylalanine		✓		✓		
Serine		✓		✓		
Tryptophane		✓		✓		
Valine		✓		✓		
Glutamine				✓	✓	
Aspartate				✓		
Glycine				✓		
Histidine		✓		✓		
Tyrosine		✓		✓		
Nonamino acid						
Citric acid						✓
Glycolic acid						✓
Other:						
Glycolic acid						✓
Lactate	✓			✓		
Benzoate	✓		✓			
Salicylate	✓					
N-acteyl-L-glutamate	✓					
O-acetyl-L-serine		✓	✓			
Citrate		✓		✓		
Citraconate		✓		✓		
Pyruvate		✓				
trans-4-hydroxy-L-proline					✓	
Uracil						✓
Hypoxanthine						✓

To sum up, the first method with analysis on GC-SQ-MS identified 36 metabolites where proline and benzoate were found to be higher concentrated for CKD patients. The method also had the intension of identifying metabolites which were not found in the in-house library. Additional 19 metabolites were identified, but the metabolites were not found to be more abundant in either the control or the CKD group. However, the method also focused on discovering new metabolites which could be added to the in-house library. One compound was found to be present at retention time 10.42 minutes and identification methods with NMR were suggested.

The second method with silylation identified 59 metabolites. It was observed that 6-deoxy-D-glucose, glycolic acid, hypoxanthine, uracil and citric acid were found at higher concentration in the control group and trans-4-hydroxy-L-proline was observed at higher

concentration in the CKD subjects.

The third method was a absolute quantitative GC-QqQ-MS analysis where 46 metabolites were identified. Proline, lactate, benzoate, salicylate and N-acetyl-L-glutamate were observed to be at higher concentration for the CKD patients.

The fourth method used untargeted analysis on LC-HILIC-MS data obtained from a previous study. The data results were analyzed with Mass Profiler Professional yielding six unidentified accurate masses only occurring in CKD patients. However, the MPP results were unreliable as the masses were observed in both groups in the chromatograms. Given this, utilizing LC/MS for urine analysis is still a promising method that needs more investigation

The overall observation of the MS methods (Table 4.1) is that the concentration of metabolites is found to be lower in CKD patients, which indicates that only minimal amounts of amino acids are excreted into the urine during renal failure. Finally it can be said, that the results from the three GC-MS methods were comparable with previous studies as well as observing some new trends. Utilizing urine for biomarker discovery in chronic kidney disease patients can therefore be said to have good potential, however more investigations and further work must be proceeded.

References

[1] E. Haug, O. Sand, and Ø.V. Sjaastad. *Menneskets fysiologi*. Universitetsforlaget, 1996.

[2] AS. Levey, KU. Eckardt, Y. Tsukamoto, A. Levin, J. Coresh, J. Rossert, et al. Definition and classification of chronic kidney disease: a position statement from kidney disease: Improving global outcomes (kdigo). *Kidney international*, 67(6):2089–2100, 2005.

[3] S. Silbernagl and A. Despopoulos. *Color atlas of physiology*. George Thieme Verlag, 2009.

[4] G. Eknoyan and N.W. Levin. K/doqi clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis*, 39(2 Suppl 1):S1–266, 2002.

[5] F. Kronenberg. Emerging risk factors and markers of chronic kidney disease progression. *Nature Reviews Nephrology*, 5(12):677–689, 2009.

[6] S. Jambhekar, P.J. Breen, and Royal Pharmaceutical Society of Great Britain. *Basic pharmacokinetics*. Pharmaceutical Press London, UK:, 2009.

[7] R.H. Weiss and K. Kim. Metabolomics in the study of kidney diseases. *Nature Reviews Nephrology*, 8(1):22–33, 2011.

[8] J. Jantos-Siwy, E. Schiffer, K. Brand, G. Schumann, K. Rossing, C. Delles, H. Mischak, and J. Metzger. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *Journal of proteome research*, 8(1):268–281, 2008.

[9] K. Hayashi, H. Sasamura, T. Hishiki, M. Suematsu, S. Ikeda, T. Soga, H. Itoh, et al. Use of serum and urine metabolome analysis for the detection of metabolic changes in patients with stage 1-2 chronic kidney disease. *Nephro-Urology Monthly*, 3(03):164–171, 2011.

[10] J.L. Spratlin, N.J. Serkova, and S.G. Eckhardt. Clinical applications of metabolomics in oncology: a review. *Clinical Cancer Research*, 15(2):431–440, 2009.

[11] C.L. Waterman, C. Kian-Kai, and J.L. Griffin. Metabolomic strategies to study lipotoxicity in cardiovascular disease. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1801(3):230–234, 2010.

[12] Z. Wang, M. Gerstein, and M. Snyder. Rna-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63, 2009.

[13] R.E. Banks, M.J. Dunn, D.F. Hochstrasser, J.C. Sanchez, W. Blackstock, D.J. Pappin, and P.J. Selby. Proteomics: new perspectives, new biomedical opportunities. *The Lancet*, 356(9243):1749–1756, 2000.

[14] P. Perco, C. Pleban, A. Kainz, A. Lukas, G. Mayer, B. Mayer, and R. Oberbauer. Protein biomarkers associated with acute renal failure and chronic kidney disease. *European journal of clinical investigation*, 36(11):753–763, 2006.

[15] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, et al. Hmdb: the human metabolome database. *Nucleic acids research*, 35(suppl 1):D521–D526, 2007.

[16] M. Tomita and T. Nishioka. *Metabolomics: the frontier of systems biology*. Springer Verlag, 2005.

[17] K. Dettmer, P.A. Aronov, and B.D. Hammock. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26(1):51–78, 2007.

[18] P.A. Guy, I. Tavazzi, S.J. Bruce, Z. Ramadan, and S. Kochhar. Global metabolic profiling analysis on human urine by uplc–tofms: issues and method validation in nutritional metabolomics. *Journal of Chromatography B*, 871(2):253–260, 2008.

[19] R. Goodacre. Metabolomics–the way forward. *Metabolomics*, 1(1):1–2, 2005.

[20] Q. Huang, G. Wang, W. Zha, B. Yan, H. Ren, S. Gu, Y. Zhang, Q. Zhang, F. Shao, L. Sheng, et al. Global analysis of metabolites in rat and human urine based on gas chromatography/time-of-flight mass spectrometry. *Analytical biochemistry*, 379(1):20–26, 2008.

[21] T. Kind, V. Tolstikov, O. Fiehn, and R.H. Weiss. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Analytical biochemistry*, 363(2):185–195, 2007.

[22] S. Cubbon, C. Antonio, J. Wilson, and J. Thomas-Oates. Metabolomic applications of hilic–lc–ms. *Mass spectrometry reviews*, 29(5):671–684, 2010.

[23] V. Thongboonkerd. Recent progress in urinary proteomics. *PROTEOMICS–Clinical Applications*, 1(8):780–791, 2007.

[24] M.T. Nguyen and P. Devarajan. Biomarkers for the early detection of acute kidney injury. *Pediatric nephrology*, 23(12):2151–2157, 2008.

[25] E.J. Saude, D. Adamko, B.H. Rowe, T. Marrie, and B.D. Sykes. Variation of metabolites in normal human urine. *Metabolomics*, 3(4):439–451, 2007.

[26] B.P.B. Do Yup Lee and T.R. Northen. Mass spectrometry–based metabolomics, analysis of metabolite-protein interactions, and imaging. *Biotechniques*, 49(2):557, 2010.

[27] T. Greibrokk, E. Lundanes, and K.E. Rasmussen. *Kromatografi: separasjon og deteksjon*. Universitetsforlaget, 1994.

[28] J.F. Rubinson and K.A. Rubinson. Contemporary chemical analysis. *Chemical Analysis*, 1998.

[29] E. de Hoffmann. Tandem mass spectrometry: a primer. *Journal of mass spectrometry*, 31(2):129–137, 1996.

[30] H.F.N. Kvitvang, T. Andreassen, T. Adam, S.G. Villas-Bo as, and P. Bruheim. Highly sensitive gc/ms/ms method for quantitation of amino and nonamino organic acids. *Analytical chemistry*, 83:2705–2711, 2011.

[31] S.G. Villas-Bôas, K.F. Smart, S. Sivakumaran, and G.A. Lane. Alkylation or silylation for analysis of amino and non-amino organic acids by gc-ms? *Metabolites*, 1(1):3–20, 2011.

[32] S.G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, and J. Nielsen. Mass spectrometry in metabolome analysis. *Mass spectrometry reviews*, 24(5):613–646, 2005.

[33] M.M. Koek, B. Muilwijk, M.J. Van Der Werf, and T. Hankemeier. Microbial metabolomics with gas chromatography/mass spectrometry. *Analytical chemistry*, 78(4):1272–1281, 2006.

[34] S. Cubbon, T. Bradbury, J. Wilson, and J. Thomas-Oates. Hydrophilic interaction chromatography for mass spectrometric metabonomic studies of urine. *Analytical chemistry*, 79(23):8911–8918, 2007.

[35] P. Hemström and K. Irgum. Hydrophilic interaction chromatography. *Journal of separation science*, 29(12):1784–1821, 2006.

[36] Agilent Technologies Deconvolution Reporting Sodtware(DRS). Agilent technolgies. *http://www.agilent.de/labs/features/2010.html*, [cite:04.06.2012].

[37] S.G. Villas-Bôas. *Metabolome analysis: An introduction*. John Wiley & Sons, 2007.

[38] Gunnar G. Løvås. *Statistikk for universiteter og høgskoler*. Universitetsforlaget, 1999.

[39] J. Boccard, J.L. Veuthey, and S. Rudaz. Knowledge discovery in metabolomics: an overview of ms data handling. *Journal of separation science*, 33(3):290–304, 2010.

[40] C.A. Smith, G.O. Maille, E.J. Want, C. Qin, S.A. Trauger, T.R. Brandon, D.E. Custodio, R. Abagyan, and G. Siuzdak. Metlin: a metabolite mass spectral database. *Therapeutic drug monitoring*, 27(6):747, 2005.

[41] The Scripts Research Institute and Metlin. Metlin: a metabolite mass spectral database. *http://metlin.scripps.edu/statistics/*, [cite:04.06.2012].

[42] S.G. Villas-Bôas, D.G. Delicado, M. Akesson, J. Nielsen, et al. Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate derivatives using gas chromatography-mass spectrometry. *Analytical biochemistry*, 322(1):134–138, 2003.

[43] O. Fiehn, G. Wohlgemuth, M. Scholz, T. Kind, D.Y. Lee, Y. Lu, S. Moon, and B. Nikolau. Quality control for plant metabolomics: reporting msi-compliant studies. *The Plant Journal*, 53(4):691–704, 2008.

[44] S.U. Bajad, W. Lu, E.H. Kimball, J. Yuan, C. Peterson, and J.D. Rabinowitz. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *Journal of Chromatography A*, 1125(1):76–88, 2006.

[45] Y.Y. Zhao, J. Liu, X.L. Cheng, X. Bai, and R.C. Lin. Urinary metabonomics study on biochemical changes in an experimental model of chronic renal failure by adenine based on uplc q-tof/ms. *Clinica Chimica Acta*, 2011.

[46] Horst Friebolin. *Basic One- and Two-Dimensional NMR Spectroscopy*. Wiley-VCH Verlag GmbH and Co. KGaA, 2008.

[47] A.L. Lehninger, D.L. Nelson, and M.M. Cox. *Principles of biochemistry*, volume 1. WH Freeman & Co, 2005.

[48] H. Shahbazian, A.Z. Moghadam, A. Ehsanpour, M. Khazaali, et al. Changes in plasma concentrations of hypoxanthine and uric acid before and after hemodialysis. *Iranian Journal of Kidney Diseases*, 3(3):151–155, 2009.

[49] M.J. Sarnak, A.S. Levey, A.C. Schoolwerth, J. Coresh, B. Culleton, L.L. Hamm, P.A. McCullough, B.L. Kasiske, E. Kelepouris, M.J. Klag, et al. Kidney disease as a risk factor for development of cardiovascular disease. *Circulation*, 108(17):2154–2169, 2003.

[50] M. Tarantola, R. Motterlini, M. Beretta, E. Rovida, and M. Samaja. Impairment of the post-anoxic recovery of isolated rat hearts by intravascular hypoxanthine and xanthine. *Artificial Cells, Blood Substitutes and Biotechnology*, 18(2):309–320, 1990.

[51] H. Jiang, J. Jiang, P. Hu, and Y. Hu. Measurement of endogenous uracil and dihydrouracil in plasma and urine of normal subjects by liquid chromatography–tandem mass spectrometry. *Journal of Chromatography B*, 769(1):169–176, 2002.

[52] R.W. Sparidans, TM Bosch, M. Jörger, J.H.M. Schellens, and J.H. Beijnen. Liquid chromatography–tandem mass spectrometric assay for the analysis of uracil, 5, 6-dihydrouracil and β-ureidopropionic acid in urine for the measurement of the activities of the pyrimidine catabolic enzymes. *Journal of Chromatography B*, 839(1):45–53, 2006.

[53] G. Gstraunthaler, T. Holcomb, E. Feifel, W. Liu, N. Spitaler, and N.P. Curthoys. Differential expression and acid-base regulation of glutaminase mrnas in gluconeogenic llc-pk1-fbpase+ cells. *American Journal of Physiology-Renal Physiology*, 278(2):F227–F237, 2000.

[54] A. Tizianello, G. De Ferrari, G. Garibotto, G. Gurreri, and C. Robaudo. Renal metabolism of amino acids and ammonia in subjects with normal renal function and in patients with chronic renal insufficiency. *Journal of Clinical Investigation*, 65(5):1162, 1980.

[55] E.J. Saude and B.D. Sykes. Urine stability for metabolomic studies: effects of preparation and storage. *Metabolomics*, 3(1):19–27, 2007.

[56] Stina K. Lien, Hans Fredrik Nyvold Kvitvang, and Per Bruheim. Utilization of a deuterated derivatization agent to synthesize internal standards for gas chromatography –tandem mass spectrometry quantification of silylated metabolites. *Journal of Chromatography A*, -(0021-9673):doi: 10.1016/j.chroma.2012.05.053, 2012.

A Appendix

A.1 Chemical list

Table A.1: *Chemical list over reagents used during sample preparation*

Chemicals	Batch number	Producer
Urease from *Calavalie ensiformis* (Jack bean)	110M7352	Sigma-Aldrich
Methoxyamine	BCBC344IV	Sigma-Aldrich
Methanol (MeOH)	SZBB232SX	Sigma-Aldrich
Sodium hydroxide (4.0 g per 100 mL yields 1M) (NaOH)	BCBB5919	Sigma-Aldrich
Pyridine	BCBD4443V	Sigma-Aldrich
Chloroform (CHl)	9204A	Sigma-Aldrich
Sodium bicarbonate ($NaHCO_3$) (0.42005 g per 100 mL yields 50 mM)	1346077	Sigma-Aldrich
Sodium sulfate ($NaSO_4$)	120MOOO79V	Sigma-Aldrich
Methyl chloroformate (MCF)	STBB3850V	Sigma-Aldrich
N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) + 1% trimethylchlorosilane (TMCS)	JL127075	Thermo Scientific
d_3-MCF	Batch E PSO AC-425	Lardon Fine Chemicals AB
Myristic-d_{27} acid (98%)	EK1220	Sigma-Aldrich

A.2 MCF and TMSstandard mix

Table A.2: *MCF standard mix*

Compound	TgI	R.T.
Fumarate-1	113	8.320
Fumarate-2	113	8.364
Succinate	55	8.498
d3-alanine	105	10.042
Alanine	102	10.075
Valine	130	12.086
a-ketoglutarate	115	12.097
Threonine	115	13.564
Proline	128	13.941
Aspartate	160	14.963
Citrate	143	15.019
d5-glutamate	179	16.519
Glutamate	114	16.585
BMAA	102	17.508
Phenylalanine	162	18.262
Cadaverine	88	19.285
Ornithine	128	20.485
Lysine	142	21.585
Histidine	210	22.174
Tyrosine	121	23.250
Tryptophan	130	25.062

Table A.3: *TMS standard mix of 2 mM sugar mix*

Sugars
10mM fructose
10mM lactose
10mM maltose
10mM raffinose

75 μL of the 2 mM sukkermixblanding (same volume) are added in a GC/MS vial and concentrated (yielding 150 nmol) . Dissolved in 100 μL 4% methoxyamine HCl in pyridine where 50 μL is used in the derivatization steps.

A.3 Fatty acid methyl esters (FAME) standard mix

	Fatty acid methyl ester
1	Methyl caprylate (C8) Methyl octanoate $CH_3(CH_2)_6COOCH_3$
2	Methyl perlargonate (C9) Methyl nonanoate
3	Methyl caprate (C10) Methyl decanoate
4	Methyl laurate (C12) Methyl dodecanoate
5	Methyl myristate (C14) Methyl tetradecanoate
6	Methyl palmitate (C16) Methyl hexadecanoate
7	Methyl stearate (C18) Methyl octadecanoate
8	Methyl eicosanoate (C20) Methyl arachidate
9	Methyl docosanoate (C22) Methyl behenate
10	Methyl linocerate (C24) Methyl tetracosanoate
11	Methyl hexacosanoate (C26)
12	Methyl octacosanoate (C28)
13	Methyl triacontanoate (C30)

A.4 Standard mix for absolute quantifiaction

Table A.4: *Amino acids*

Amino acid and non-amino acid	RT	MRM TgI	ISTD MRM TgI
4-Methylvalerate	6,438	43,0	43,0
Malonic acid	6,769	68,8	69,0
Pyruvate	6,967	56,8	60,0
(+-)-3-Methyl-2-oxovalerate	7,220	67,0	67,0
Fumarate	7,740	84,8	88,0
d4-succinateHVL	7,782	57,8	57,9
Lactate/Methylglyoxal	7,803	30,8	61,9
Succinate	7,813	54,8	62,0
Citraconate/Itaconate	8,369	41,0	42 / 62
Benzoate	8,491	77,0	77,0
Citramalate	8,584	43,0	43,0
Glyoxylate	8,687	75,0	53,0
β-3-hydroxybutyrate	8,696	41,0	41,0
d3-AlanineHVL	8,711	60,8	108,0
Glycine	8,720	43,8	47,1
Alanine	8,727	101,9	105,1
O-Acetyl-L-Serine	8,735	42,0	42,0
Nicotinate	8,875	77,9	77,9
Phenylacetate	9,185	91,0	91,0
Salicylate	9,378	121,0	121,0
2-Aminobutyrate	9,422	57,1	60,0
m-Toluate	9,460	91,0	91,0
Beta-Alanine	9,495	44,0	47,0
OH-Glutarate	9,659	98,7	101,9
Adipate	9,679	73,0	76,0
d8-Valine	9,777	61,9	62,0
Valine	9,825	70,9	74,0
aKG	9,848	55,0	55,0
2-Isopropylmalate	9,943	43,0	43,0
beta-hydroxypyruvate	10,049	74,9	81,1
alpha-ketoadipate	10,363	68,9	61,9
Leucine	10,466	43,1	46,8

Continued on next page

Table A.4 – *Continued from previous page*

Amino acid and non-amino acid	RT	MRM TgI	ISTD MRM TgI
Isoleucine1	10,497	68,9	36,0
gamma-aminobutyrate	10,544	69,0	62,0
Isoleucine2	10,548	68,9	36,0
Threonine	10,621	55,8	55,8
Malate	10,703	80	81,0
L-Homoserine	10,718	30,0	61,1
Proline	10,836	83,8	87,0
OAA (Oxaloacetate)	10,847	68,8	68,9
Aspatate	11,345	74,8	81,0
Citrate	11,381	100,8	61,9
5-aminovalerate	11,412	55,0	55,0
Threo-beta-hydroxyaspartate	11,583	81,9	85,0
Serine	11,819	99,7	102,8
Anthranilate	11,827	89,9	89,9
Allantoin	11,922	41,9	45,1
N-Acetyl-L-Glutamate	12,169	84,0	84,0
d5-GlutamateHVL	12,177	102,9	106,0
Glutamate	12,204	97,8	100,9
Hydroxyproline	12,323	81,8	85,0
Methionine	12,342	60,8	61,0
Beta-Methylamino-L-alanine	12,679	98,0	115,9
2-oxobutyrate	12,809	67,0	95,0
Cysteine	12,917	88,7	62,0
Isocitrate	13,003	75,0	80,9
Putrescine	13,006	55,0	55,0
Hippurate	13,069	104,9	105,0
Phenylalanine	13,126	127,9	128,0
Phenylpyruvate	13,139	115,0	62,0
2,4-Diaminobutyrate	13,465	70,0	73,0
4-imidazoleacrylate	13,503	107,0	110,1
Cadaverine	13,642	69,0	69,0
4-Aminobenzoate	13,742	90,8	94,0
Histamine	13,963	109,1	67,9
p-Coumarate	14,092	132,9	136,0

Continued on next page

Table A.4 – *Continued from previous page*

Amino acid and non-amino acid	RT	MRM TgI	ISTD MRM TgI
Ornithine	14,256	128,0	131,1
N-GlycylL-Proline	14,708	70,0	70,0
Lysine	14,833	142,0	145,0
Ferulate	15,104	176,0	175,9
Histidine	15,154	150,0	153,1
Tyrosine	15,736	158,0	161,0
2,6-diaminopimelate	15,844	139,9	143,1
Tryptophane	16,713	185,0	185,2
Cystein2	17,543	73,9	
Serotonin	18,184	157,1	157,0

A.5 Data from alkylation and metabolic application of GC-SQ-MS system

Table A.5: *Results from MCF derivatization of polar components from sample 1 to 5 and 101 og 105. The results are given in peak area per 100 µL urine*

Urine Sample	1	2	3	4	5	101	102	103	104	105
Malanoic acid	0	57765	0	0	0	253998	286925	193932	140922	224925
Fumarate-2	288576	112461	127483	219113	54923	327785	88000	141644	148150	216582
Lactate	601916	444771	620334	1025708	254749	1518651	603125	856797	583515	776096
Methylglyoxal	601916	444771	620334	1025708	254749	1518651	603125	856797	583515	776096
Succinate	626774	420813	855691	1074644	337667	4015095	1368262	3012492	2490105	1358604
Hexachloroethane	0	0	843174	2188507	946278	0	0	1905835	1214875	1661038
b-hydroxypyruvat	1186518	1076329	831378	2144178	941005	1757268	954780	1891646	1192882	1682046
Citraconate	0	23559	180868	398973	46405	1436476	260622	391112	882679	894663
Benzoat	3490292	1444252	1528334	1542216	729813	567032	316063	554017	836016	492006
Glycine	2883835	5774668	4959497	16212987	1109594	59250755	4260912	25310129	30744913	25965035
Alanine	2383481	4427673	1933215	6226214	813302	23006768	1390076	8238873	9212284	12151025
Phenylacetat	5607305	4697756	2109676	4691125	0	2147393	483173	0	3851238	0
Valine	720311	582846	572091	1271905	192452	3194155	398438	0	0	3366107
Leucine	218923	704738	116906	450389	42472	983362	86933	569704	259057	847021
Isoleucine-1	221314	123052	122281	390657	42783	967220	87384	429914	280257	824060
Serine-1	130849	83912	84164	516609	33073	0	135489	437857	0	499454
Isoleucine-2	0	0	121698	0	0	0	86806	408652	0	797886
Malate-1	75569	101815	116830	175963	49903	188596	85595	91457	95848	116755
Threonine	0	0	378003	1495064	0	0	0	2831152	2575535	2931256
Proline	533770	511069	733785	2718576	194454	1058460	138649	537255	321540	497753
Asparagine	539038	185869	382111	957990	196866	3749985	392744	1993577	1814123	2259167
O-acetyl-L-serine	17697461	6462748	8084915	11724607	5006498	5138945	157983	8003266	12538360	7245013
Aspartate	290802	114207	174253	184067	67044	514338	0	273537	295355	539737
Citrate	432618	395193	5214777	11458124	1037072	67663878	10269438	11554282	42617012	50737947
Serine-2	320765	287304	224765	785253	106516	2510320	171498	1231643	1300247	1234007
Glutamine	0	170411	202529	275949	60811	1491292	146406	986004	571745	675580
Cysteine	921811	795015	417459	1375555	590661	1597497	133862	1360124	932323	1773957
Isocitrate	0	257460	276663	513391	275307	1503445	434259	807439	819542	1243904
Phenylalanine	0	899605	698918	1332670	147788	4200984	682922	1821103	0	11202931
Putresine	246872	299901	59453	313596	0	187693	0	97172	136416	157600
Ornithine	1035957	524927	344432	0	311516	1271271	182155	1406214	947344	1775005
Lysine	1721206	1734536	1189875	3361693	522473	12790967	844589	21292637	4462140	7916935
Histidine	722304	1016204	621552	2363536	149982	29115457	1149661	21162214	8399349	19945860
Ferulate	113839	110208	118242	200259	0	0	216936	146218	140337	358004
Tyrosine	887102	850865	665052	1522981	237834	0	925925	5961987	3160414	12445150
Tryptophan	1735375	2135651	1245612	2761436	340869	18727080	1678883	8992756	5397958	19372534

A.6 Statistical analysis with student-t

Statistical analysis was performed on 36 metabolites and calculations of five of the metabolites are presented in A.6. Metabolites found to be statistical different was o-acetyl-L-serine, glycine, tryptophane, ferulate, isocitrate, serine-2, aspartate, aspergine, citraconate, succinat, methyl glyoxal and lactate. Metabolites found not to be statistical different were hexacloroethane, leucine, isoleucine, tyrosine, proline, cystein, lysine and putrecine.

Table A.6: *Student-t test between the CKD group(1-5) and control group (101-105). The results are given in peak area per 100 μL urine*

	Benzoat	Glutamine	Phenylalanine	Citrate	Histidine
1	$3{,}49\times10^{6}$	$1{,}80\times10^{5}$		$4{,}33\times10^{5}$	$7{,}22\times10^{5}$
2	$1{,}44\times10^{6}$	$1{,}70\times10^{5}$	$9{,}00\times10^{5}$	$3{,}95\times10^{5}$	$1{,}02\times10^{6}$
3	$1{,}53\times10^{6}$	$2{,}03\times10^{5}$	$6{,}99\times10^{5}$	$5{,}21\times10^{6}$	$6{,}22\times10^{5}$
4	$1{,}54\times10^{6}$	$2{,}76\times10^{5}$	$1{,}33\times10^{6}$	$1{,}15\times10^{7}$	$2{,}36\times10^{6}$
5	$7{,}30\times10^{5}$	$6{,}08\times10^{4}$	$1{,}48\times10^{5}$	$1{,}04\times10^{6}$	$1{,}50\times10^{5}$
Average, $\bar{x}_{patient}$	$1{,}75\times10^{6}$	$1{,}77\times10^{5}$	$7{,}70\times10^{5}$	$3{,}71\times10^{6}$	$9{,}75\times10^{5}$
St.dev, $\sigma_{patient}$	$1{,}03\times10^{6}$	$8{,}94\times10^{4}$	$4{,}92\times10^{5}$	$4{,}77\times10^{6}$	$8{,}37\times10^{5}$
SEM, $\bar{\sigma}_{patient}$	$4{,}61\times10^{5}$	$4{,}47\times10^{4}$	$2{,}46\times10^{5}$	$2{,}14\times10^{6}$	$3{,}74\times10^{5}$
101	$5{,}67\times10^{5}$	$1{,}49\times10^{6}$	$4{,}20\times10^{6}$	$6{,}77\times10^{7}$	$2{,}91\times10^{7}$
102	$3{,}16\times10^{5}$	$1{,}46\times10^{5}$	$6{,}83\times10^{5}$	$1{,}03\times10^{7}$	$1{,}15\times10^{6}$
103	$5{,}54\times10^{5}$	$9{,}86\times10^{5}$	$1{,}82\times10^{6}$	$1{,}16\times10^{7}$	$2{,}12\times10^{7}$
104	$8{,}36\times10^{5}$	$5{,}72\times10^{5}$		$4{,}26\times10^{7}$	$8{,}40\times10^{6}$
105	$4{,}92\times10^{5}$	$6{,}76\times10^{5}$	$1{,}12\times10^{7}$	$5{,}07\times10^{7}$	$1{,}99\times10^{7}$
Average, $\bar{x}_{control}$	$5{,}53\times10^{5}$	$7{,}74\times10^{5}$	$4{,}48\times10^{6}$	$3{,}66\times10^{7}$	$1{,}60\times10^{7}$
St.dev, $\sigma_{control}$	$1{,}87\times10^{5}$	$5{,}01\times10^{5}$	$4{,}72\times10^{6}$	$2{,}51\times10^{7}$	$1{,}11\times10^{7}$
SEM, $\bar{\sigma}_{control}$	$8{,}37\times10^{4}$	$2{,}24\times10^{5}$	$2{,}11\times10^{6}$	$1{,}12\times10^{7}$	$4{,}96\times10^{6}$
Student t-test:					
$\bar{x}_{patient}$ - $\bar{x}_{control}$	$1{,}19\times10^{6}$	$5{,}97\times10^{5}$	$3{,}71\times10^{6}$	$3{,}29\times10^{7}$	$1{,}50\times10^{7}$
$\sqrt{\bar{\sigma}^2_{patient} - \bar{\sigma}^2_{control}}$	$4{,}54\times10^{5}$	$2{,}20\times10^{5}$	$2{,}10\times10^{6}$	$1{,}10\times10^{7}$	$4{,}95\times10^{6}$
t =	2,63	2,72	1,77	2,98	3,03
Degree of freedom, Df	8	8	6	8	8
Significant different	yes	yes	no	yes	yes

Table A.7: *Student-t test between the CKD group(1-5) and control group (101-105) for one unidentified compound. The results are given in peak area per 100 μL urine*

CKD	Peak height	Control	Peak height
1	800000	101	1400000
2	900000	102	800000
3	800000	103	900000
4	900000	104	900000
5	800000	105	800000
Average,$\bar{x}$	850000		960000
St.dev,σ	547722		224499
SEM, $\bar{\sigma}$	24495		100399

Student t-test:	
$\bar{x}_{patient}$ - $\bar{x}_{control}$	120000
$\sqrt{\bar{\sigma}^2_{patient} - \bar{\sigma}^2_{control}}$	97365
t =	1.23
Degree of freedom, Df	8
Significant different	no

A.7 Scan of unidentified compound

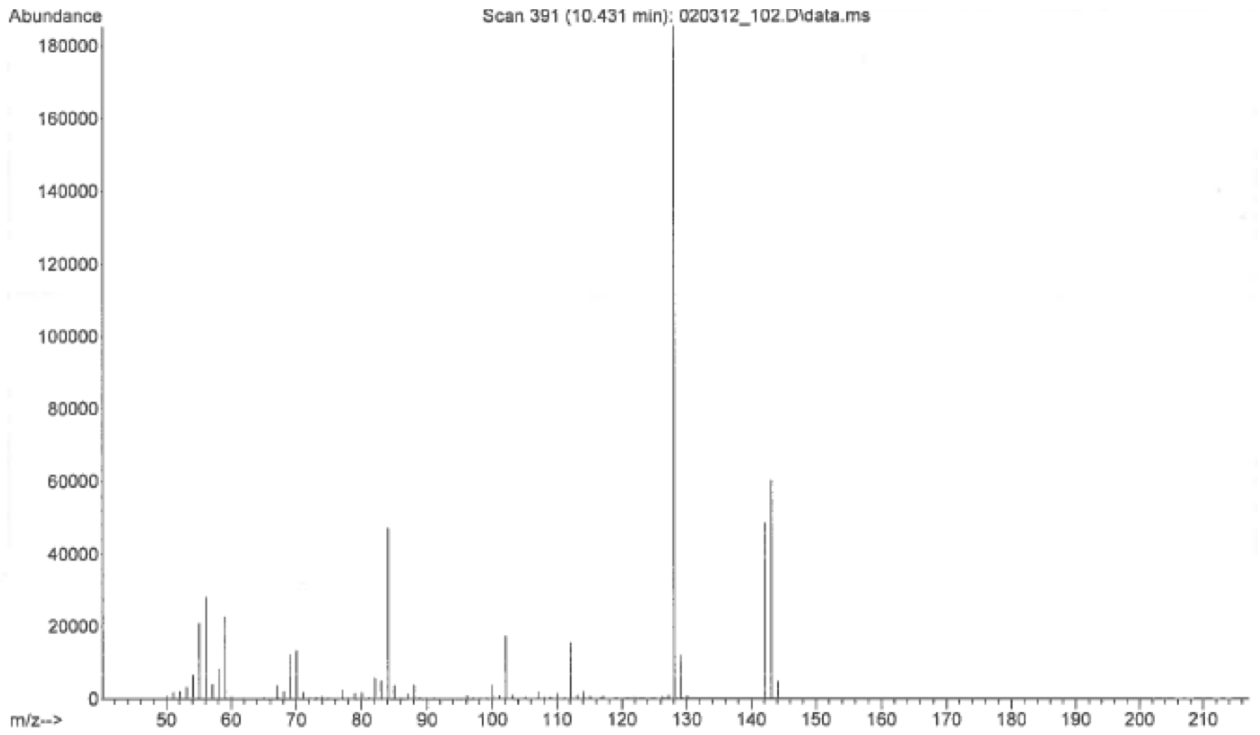


Figure A.1: *Scan of unidentifed compound at retention time 10.4 minutes*

A.8 Results from silylation and metaboloic application of GC-SQ-MS system

Table A.8: *Results from silylation of urine samples obtained from CKD patients. The results are given in peak area per 200 µL urine*

Urine samples	1	2	3	4	5	6	7	8	9	10
Glycolic acid	207152	0	0	238909	0	0	0	323019	276775	0
L-valine 2	0	0	0	0	0	290776	112001	78140	156848	0
Urea	0	67324002	449880	460398	211936	330006	373126	406199	406328	232987
Glycerol	303385093	54252360	366111657	394147628	387617082	374803013	334545569	394058223	353521432	377714120
Phosphoric acid	70777083	71412393	28337474	0	13374686	32497354	60408211	64230954	32433253	15143932
L-proline 2	0	0	0	226065	0	701114	685427	163561	222851	76299
Glycine	360730	911778	377560	2569121	0	2213925	1618891	2435586	4170092	538658
Succinic acid	0	12075	20550	52658	0	155510	38461	27163	30574	52794
Glyceric acid	0	0	0	0	0	0	0	143187	55263	0
Porphine 1	145012	202558	199842	225974	223123	180530	181153	217087	149805	165302
Uracil	36631	45177	0	0	0	0	0	0	0	0
L-serine 2	337597	122697	141460	1390579	110150	82687	651475	1119588	977452	202218
L-threonine 2	59887	0	31016	230883	0	167647	204644	238536	550331	0
DL-3-aminoisobutyric acid 2	252499	49698	0	82033	76257	139143	41422	212746	70593	51327
D-threitol	1341832	407627	368494	459620	241611	186706	333695	402871	178099	213834
L-pyroglutamic acid	1211032	903875	481907	0	640342	0	899748	1432706	0	0
L-glutamic acid 3	1208709	911509	475865	1405101	644333	1059494	899513	1432686	1060146	1935781
trans-4-hydroxy-L-proline 2	0	0	0	187433	0	137155	229185	288285	258952	0
Creatinine	1087489	634767	178051	1013870	471014	338823	573397	729919	568773	552346
4-hydroxybenzoic acid	0	114814	0	104111	0	92058	0	89066	57187	0
Tartaric acid	726497	0	35727	117082	0	117073	115504	61754	223496	0
4-hydroxyphenylacetic acid	0	193420	0	0	0	0	21047	72479	39205	38172
D-lyxose 1	679375	0	0	0	0	0	376974	568442	0	0
D-lyxosylamine 2	0	427209	0	0	0	0	0	0	178791	0
D-lyxose 2	0	0	0	0	168284	0	0	0	162354	252280
D-lyxose 2	424551	0	0	0	0	0	0	226502	157524	187806
Xylitol	538635	807609	0	204314	0	482545	0	235324	415031	189845
6-deoxy-D-glucose 1	0	0	0	0	0	0	0	0	0	0
2-amino-1-phenylethanol	0	0	0	0	0	266493	0	0	0	0
Trans-Aconitic acid	68910	87612	46584	176610	64073	74194	0	137703	75350	58021
Glycerol 1-phosphate	0	0	41383	104109	0	41739	52813	82660	53819	0
L-glutamine 3	0	0	0	223918	0	512104	192691	246346	273691	0
Hypoxanthine 2	0	0	0	0	0	0	0	0	0	0
Citric acid	56895	60211	271409	1242077	74378	0	385800	736759	1022577	531016
Myristic Acid-d27	9054732	14782134	9730447	13212973	12137920	9919011	10293080	11264631	9386168	11058101
Hippuric acid 2	26226739	13358872	7147891	13667879	6581357	8494132	9644825	5359879	4363523	4655346
Tagatose 1	0	0	0	296369	0	0	149213	0	0	0
Quinic acid	1671714	2164551	1270325	2018347	1460106	2019313	1517088	1612032	371686	1841554
DL-4-hydroxy-3-methoxy-mandelic acid	262662	206307	67584	287075	112764	0	0	0	0	0
Fructose 1	203429	151766	233786	221081	0	0	0	575567	196141	0
Tagatose 2	90486	0	200933	163801	82804	0	0	0	0	82537
D-allose 1	519129	475965	92214	6330826	43382	16489969	22692870	2748704	65450058	136027
D-mannose 1	520633	1468919	409351	6416899	91347	15466683	22238373	2620878	79312737	0
N-carbamyl-L-glutamic acid	182072	57268	0	0	0	0	0	0	0	0
D-glucose 1	513834	2099147	399907	6405630	258276	16667813	22921204	2751811	80801251	553103
L-histidine 3	0	0	0	333142	0	349212	180685	574499	547291	0
D-mannitol	4679272	1449514	516488	2000426	461155	18466797	1734321	1042065	1342574	1215675
tyrosine 2	0	0	62482	173842	0	344643	150860	323036	535251	0
Galacturonic acid 4	305831	0	96892	0	0	0	0	214344	0	107135
Xanthine	110828	94100	0	0	0	0	0	0	0	0
D-saccharic acid	282666	242564	0	192898	69965	85748	130503	158035	0	0
N-acetyl-D-mannosamine 1	0	0	0	0	0	0	0	0	0	0
Uric acid 1	1613056	2046608	1687799	9629463	1110410	3462248	2948279	9808018	7097444	3316049
5-hydroxyindole-3-acetic acid	0	0	0	0	0	0	0	0	0	0
L-tryptophan 2	0	190900	0	173456	0	0	0	237327	332559	0
Dioctyl phthalate	0	0	18573	33488	25729	23259	24340	25327	0	0
Sucrose	5088792	685736	248796	1043489	1956184	1029756	151261	508579	55816	142028
Cellobiose 1	176293	0	225509	316716	200929	273053	178120	256344	204598	199525
Lactose 2	0	0	298350	485510	294434	347358	0	392641	0	0

Table A.9: *Results from silylation of urine samples obtained from CKD patients. The results are given in peak area per 200 µL urine*

Urine samples	11	12	13	14	15	16	17	18	19	20
Glycolic acid	106092	0	460148	100200	0	0	202676	117279	118570	71904
L-valine 2	95482	0	95913	69808	337495	292756	118516	238029	0	128103
Urea	13108057	8249988	471753	279807	312742	295890	445932	327754	240329	159790
Glycerol	337648116	0	400195239	342156997	377362299	313389123	388450906	363473212	385238730	242223863
Phosphoric acid	3697057	23414721	97530576	33202772	53685244	678273	67493252	32817209	57873137	50719599
L-proline 2	99389	0	0	377991	1836246	302461	290537	1265276	186092	975104
Glycine	1144804	226417	1712301	4172153	3215269	1296380	3148315	2344754	733027	3902896
Succinic acid	0	0	0	25643	28428	0	62722	52855	18223	0
Glyceric acid	0	0	0	0	0	0	22657	0	0	0
Porphine 1	149467	144582	216437	134263	159040	145968	152499	156487	163446	125334
Uracil	0	0	58562	0	0	0	0	0	0	0
L-serine 2	425640	49902	1192597	418260	1042983	506202	535134	868590	393099	682927
L-threonine 2	127443	0	140545	83303	460717	238648	113172	554874	119080	267446
DL-3-aminoisobutyric acid 2	455390	180433	3126669	1205416	38533	0	1658830	111573	66672	61667
D-threitol	798996	137598	377765	123132	290184	304681	537008	174327	185238	99635
L-pyroglutamic acid	1394542	554537	1832575	0	0	0	2243338	0	0	0
L-glutamic acid 3	1395160	554428	1828275	554873	995488	882927	2244119	910267	1141093	817468
trans-4-hydroxy-L-proline 2	0	0	0	442795	330445	172416	206220	155804	229342	300913
Creatinine	475039	754707	1096020	345308	452927	421253	950507	329692	447876	446964
4-hydroxybenzoic acid	266665	103537	171056	77362	0	0	0	53691	60218	0
Tartaric acid	233605	52667	1161579	0	78714	327478	199203	0	79366	0
4-hydroxyphenylacetic acid	44325	0	55776	0	0	0	74085	0	60678	0
D-lyxose 1	417816	0	223701	204212	0	282581	1296972	0	161397	0
D-lyxosylamine 2	0	0	0	0	0	0	0	111547	0	0
D-lyxose 2	406778	0	0	0	0	0	0	0	0	0
D-lyxose 2	0	0	259048	0	0	0	192625	32800	0	0
Xylitol	273283	0	0	73575	433518	447532	82388	0	436544	66335
6-deoxy-D-glucose 1	0	0	0	0	0	0	0	0	0	0
2-amino-1-phenylethanol	0	0	0	0	0	0	0	0	0	0
Trans-Aconitic acid	81997	31822	209305	20016	0	0	247582	36581	48880	0
Glycerol 1-phosphate	60117	0	219667	36544	52846	38944	0	40052	70339	36998
L-glutamine 3	0	0	458087	66230	130928	271615	0	422234	0	116122
Hypoxanthine 2	0	0	0	0	0	0	0	0	0	0
Citric acid	56990	32855	812775	52627	135608	52786	365039	513971	71588	101269
Myristic Acid-d27	9274297	9813131	12716863	7739835	9383887	9680622	10335158	9340560	10008522	6125215
Hippuric acid 2	0	3262221	27974994	1244246	2187700	3614181	12162098	1047876	0	4674358
Tagatose 1	0	63747	0	127449	292231	90594	847158	78577	0	0
Quinic acid	2447311	160303	5556150	1220168	272051	697392	4391132	0	1396207	856038
DL-4-hydroxy-3-methoxy-mandelic acid	0	70678	414143	0	104280	0	627736	60552	131086	0
Fructose 1	0	0	0	0	278181	0	0	0	0	0
Tagatose 2	105363	71577	0	0	0	0	0	0	0	0
D-allose 1	365797	-24094	688879	1988446	21678019	33994332	2987769	6585776	1455133	2330850
D-mannose 1	3177931	0	0	0	21114418	33479688	0	0	2409962	4839811
N-carbamyl-L-glutamic acid	124441	39568	131141	0	0	0	0	0	112496	0
D-glucose 1	3354648	518644	895426	2000870	21879265	34172898	3001760	9575622	2425549	4883131
L-histidine 3	168680	0	892747	118749	198447	212413	0	292557	100181	186669
D-mannitol	1322954	414549	1266893	245057	1557253	495844	741876	203269	378582	396585
tyrosine 2	184271	0	341120	0	136116	130733	215860	286238	0	0
Galacturonic acid 4	377401	0	361495	113500	0	0	524904	78359	179480	86409
Xanthine	0	0	68931	0	37812	0	0	0	0	0
D-saccharic acid	248005	0	589790	230175	95621	0	255506	0	136352	58411
N-acetyl-D-mannosamine 1	0	0	59495	0	0	0	0	0	0	0
Uric acid 1	5455869	1285789	11981684	3529734	3420067	1887225	6635251	2639411	3434458	3189476
5-hydroxyindole-3-acetic acid	0	0	0	0	0	0	335906	0	0	0
L-tryptophan 2	180595	0	310364	0	154571	0	0	97717	0	128474
Dioctyl phthalate	26525	21013	40324	0	21605	23490	0	20132	23581	0
Sucrose	2623503	75446	639162	749439	265256	182431	10828749	199006	326258	195428
Cellobiose 1	219234	0	266290	198385	226615	151472	0	225563	245635	104658
Lactose 2	300143	0	356774	282306	318149	197719	0	299953	333258	162901

Table A.10: *Results from silylation of urine samples obtained from healthy volunteers. The results are given in peak area per 200 µL urine*

Urine samples	101	102	103	104	105	106	107	108	109	110	112	113	114	115	116
Glycolic acid	2875528	384710	2641760	1239901	1899237	993500	2080428	1069373	594197	478986	1012514	2112386	2045268	1522797	1468595
L-valine 2	350350	0	204982	152598	469175	236416	404544	227210	0	0	181111	453887	323160	316764	183525
Urea	0	1644584	0	324616	0	34355733	0	756835	521361	459636	728451	0	1106481	40547102	766369
Glycerol	312949451	373831942	368351586	404763932	430860905	402531390	472353949	381762473	405806670	428796124	403219609	608789026	416409096	484652070	399844974
Phosphoric acid	53102979	65991004	26335688	56877179	101690485	10300654	83350030	147117644	40996547	38180939	140931321	76492939	131339903	6053163	89185232
L-proline 2	89037	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Glycine	6981926	736230	4470636	4743086	6303296	2441271	6298967	2213591	1457519	1628007	3628835	7499997	4317175	2790309	11039638
Succinic acid	224819	47559	130998	90820	83334	161749	132716	62362	33995	61958	94604	128216	695440	119011	238657
Glyceric acid	52036	108375	50669	0	35819	249101	29418	0	196221	140901	0	57165	140458	90674	304328
Porphine 1	186533	183329	221233	205470	217161	197961	183724	187140	187436	177240	176734	202979	192899	208435	177652
Uracil	253528	57839	160873	162525	291202	316068	340129	227927	42265	107348	84328	228823	366396	285327	259759
L-serine 2	1879657	198579	1455543	1449338	1526692	1334353	1691399	963759	432135	460099	1334624	2379466	1110706	1431422	2626552
L-threonine 2	333856	50873	312242	242761	416616	194994	438213	230187	106024	87884	275714	599501	199927	291129	339151
DL-3-aminoisobutyric acid 2	410735	142954	139021	268298	791790	448422	522867	311524	110134	137533	636973	645227	697040	312486	666194
D-threitol	2689433	1129284	495411	1292985	2663169	308110	2132297	401070	437262	496367	478926	2959990	5316982	2097059	508897
L-pyroglutamic acid	3738133	0	0	0	4309308	2078307	3137708	0	0	0	0	4739982	5195634	3906853	2941005
L-glutamic acid 3	3746492	876998	2727108	1740848	4310121	0	3136228	2348022	817785	1299002	2434417	4728332	0	3903208	2938674
trans-4-hydroxy-L-proline 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Creatinine	1623346	418557	1347295	571175	3146336	958737	1110279	1442679	425060	331119	1004095	2289151	2819539	1492896	1334584
4-hydroxybenzoic acid	0	0	0	0	0	0	0	0	0	0	0	0	125449	0	76593
Tartaric acid	129739	0	0	78424	0	0	811948	0	0	0	0	2407949	683993	0	1145851
4-hydroxyphenylacetic acid	206232	41939	155359	73163	140409	146232	77020	83179	54965	59311	242797	235767	421182	143121	157808
D-lyxose 1	386682	268993	0	386296	0	0	0	293861	0	447645	0	665374	862915	0	705930
D-lyxosylamine 2	0	318116	0	389965	0	0	0	0	178213	236988	522391	304961	0	393856	0
D-lyxose 2	404743	0	512242	384369	283492	0	563225	0	0	229301	525885	303368	855544	0	0
D-lyxose 2	393498	1474995	520182	156204	258757	237222	554831	283152	0	226544	446661	312566	773013	319814	512312
Xylitol	384205	0	283718	175420	310540	386905	1805439	309446	0	0	322978	384797	3271800	781985	1366331
6-deoxy-D-glucose 1	1322039	0	678747	490173	0	560221	1498116	0	0	266012	721319	773183	1354180	588821	460224
2-amino-1-phenylethanol	46472	0	0	0	73639	0	52472	0	0	0	43444	78828	47520	47519	38302
Trans-Aconitic acid	395943	67938	258369	230354	364531	271484	300907	157807	56616	100862	270167	411924	843286	374704	347227
Glycerol 1-phosphate	173442	54335	167546	76307	305787	187975	100071	185084	53150	66558	160455	223454	174905	106864	150521
L-glutamine 3	0	0	0	285811	0	0	0	0	0	0	0	0	0	0	0
Hypoxanthine 2	309641	0	0	307642	893797	154160	379393	0	66083	106694	0	287610	0	302315	0
Citric acid	5880200	715570	1440767	3936208	5939248	2578960	5204002	1270481	913325	1439083	2728556	5598763	10720075	5655069	6837417
Myristic Acid-d27	9367689	10086124	9435324	10077335	11183847	12314936	11152395	11188784	10575424	10850239	10740092	13419482	12343446	14243418	10225303
Hippuric acid 2	7440203	0	10788373	19756806	13477022	19401212	7394311	0	7469990	4326733	23079898	14746941	32938332	15799645	28700855
Tagatose 1	496571	0	0	260667	0	0	468497	0	0	106572	433026	546832	0	0	325974
Quinic acid	452550	1015714	2140669	1653562	2654684	6488344	1918658	1820409	1349308	324226	6523834	1240175	3522363	3849129	3469123
DL-4-hydroxy-3-methoxymandelic acid	469488	94998	0	148102	397297	270387	327075	249370	50966	155154	329108	538138	813099	345611	221331
Fructose 1	0	0	0	213586	0	0	684137	0	0	0	390176	0	959346	0	638077
Tagatose 2	0	54179	0	195938	0	87027	0	0	0	71830	313280	0	964095	87506	537430
D-allose 1	2486262	176357	324007	381918	402128	1183054	2342660	361355	491650	62770	1417095	1916625	762192	1987088	1253052
D-mannose 1	2537057	472602	324612	382456	0	0	2247981	0	0	0	0	2124036	0	1888494	0
N-carbamyl-L-glutamic acid	294719	47155	161895	89938	591942	117496	189607	132241	0	0	0	0	0	265852	0
D-glucose 1	2495091	466478	1443715	1069009	2182513	1520762	2331978	1275090	494270	617811	1649134	2200053	3146562	1951442	1551779
L-histidine 3	1045914	0	1490149	558818	1852386	187352	2670268	156028	161068	0	1006724	3279270	0	626499	1008997
D-mannitol	729645	3941724	4743506	3083324	372996	298181	389044	223896	3143542	725479	1433575	489943	437410	286294	236253
tyrosine 2	1176290	0	0	0	1553805	120892	1085351	0	0	0	0	733354	445546	0	0
Galacturonic acid 4	356375	107247	277351	230945	693508	475582	380141	385968	97905	138592	430793	556210	764093	386731	337017
Xanthine	126529	0	0	161504	160927	160969	141557	0	0	0	0	158461	293367	140204	117461
D-saccharic acid	209600	0	250061	172824	424821	350466	0	183082	0	78445	578290	431310	401958	281134	393878
N-acetyl-D-mannosamine 1	134655	0	69270	47810	144054	0	94870	77283	0	0	78316	124759	177966	71725	0
Uric acid 1	4414577	3684255	9457138	9141783	5261920	8729747	11017946	2648603	5127921	2921821	12221070	13240114	17055591	12997345	8557728
5-hydroxyindole-3-acetic acid	214507	0	123862	83857	157762	122174	0	121599	0	43125	90884	290942	442207	520312	280457
L-tryptophan 2	782124	78017	316805	190253	743593	226448	845356	216543	48916	0	350311	535004	477056	450734	178908
Dioctyl phthalate	24099	0	29555	23847	38335	25529	0	0	19478	17624	21806	31354	0	36197	0
Sucrose	292852	286946	393758	298811	345338	47643	142394	36412	147895	341490	1306220	4297747	5984882	1027570	1329224
Cellobiose 1	174509	192255	253935	266784	240623	292223	216132	260964	222483	257337	316943	400756	403840	339951	264894
Lactose 2	195242	313909	313447	391622	332819	368863	333869	301857	299025	361982	363193	486512	0	523788	325535

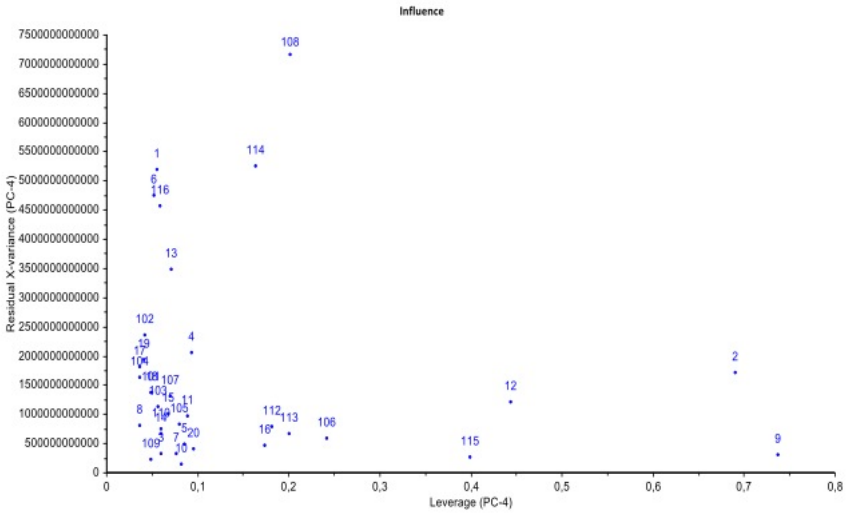


Figure A.2: *Influence plot of the urine samples before weighting*

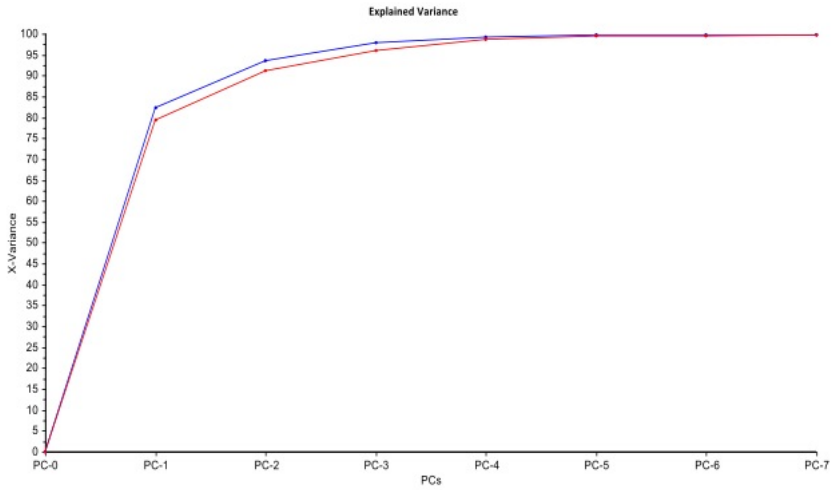


Figure A.3: *Explained variance before weighting*

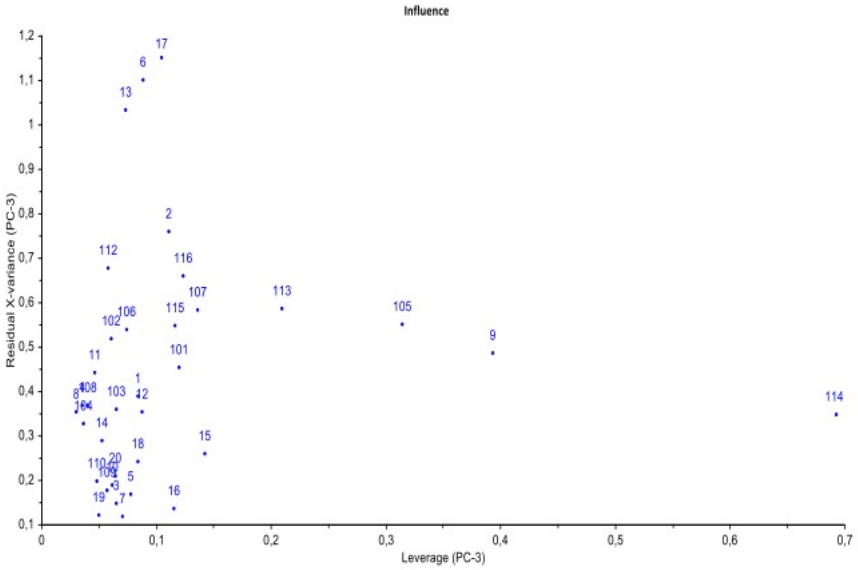


Figure A.4: *Influence plot of the urine samples after weighting*

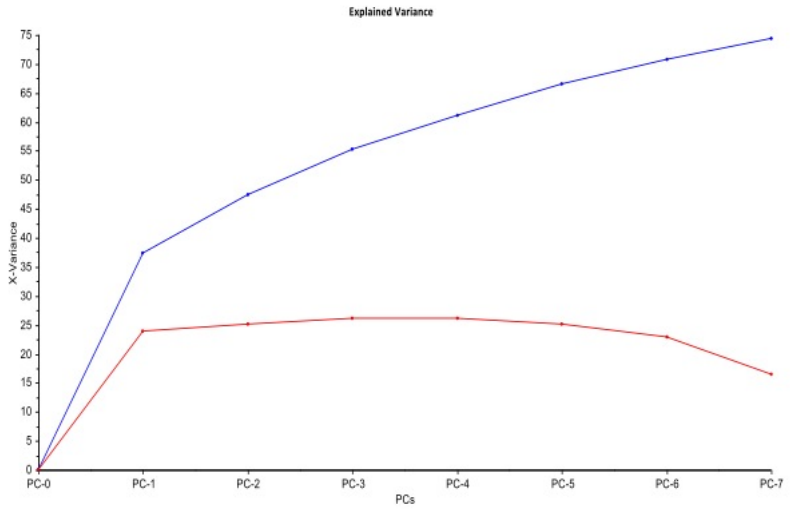


Figure A.5: *Explained variance after weighting*

A.9 Data from absolute quantification on GC-QqQ-MS

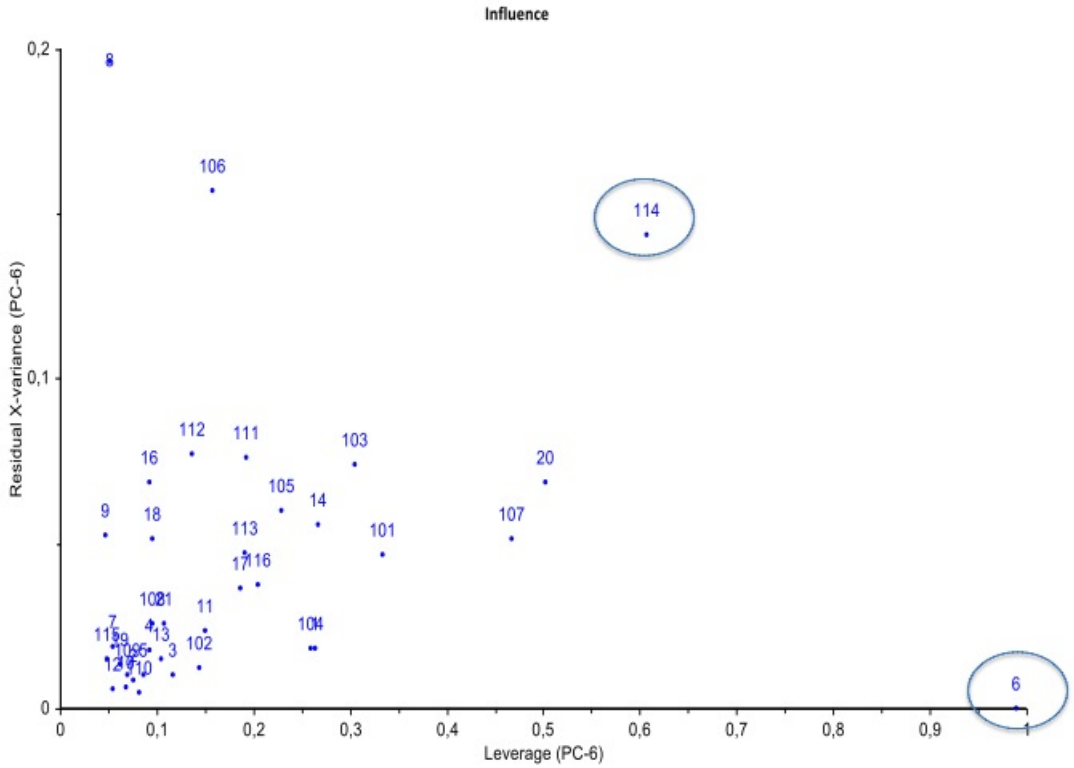


Figure A.6: *An influence and leverage plot. Shows that sample that might contribute to the variation and effect the model*

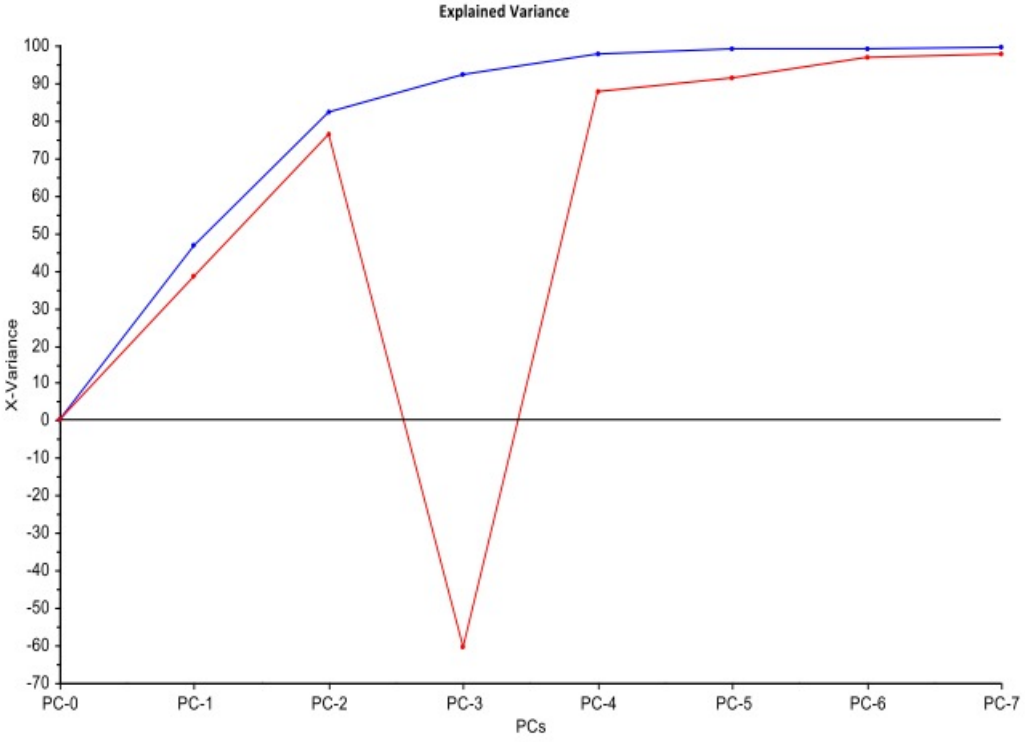


Figure A.7: *Explained variance plot*

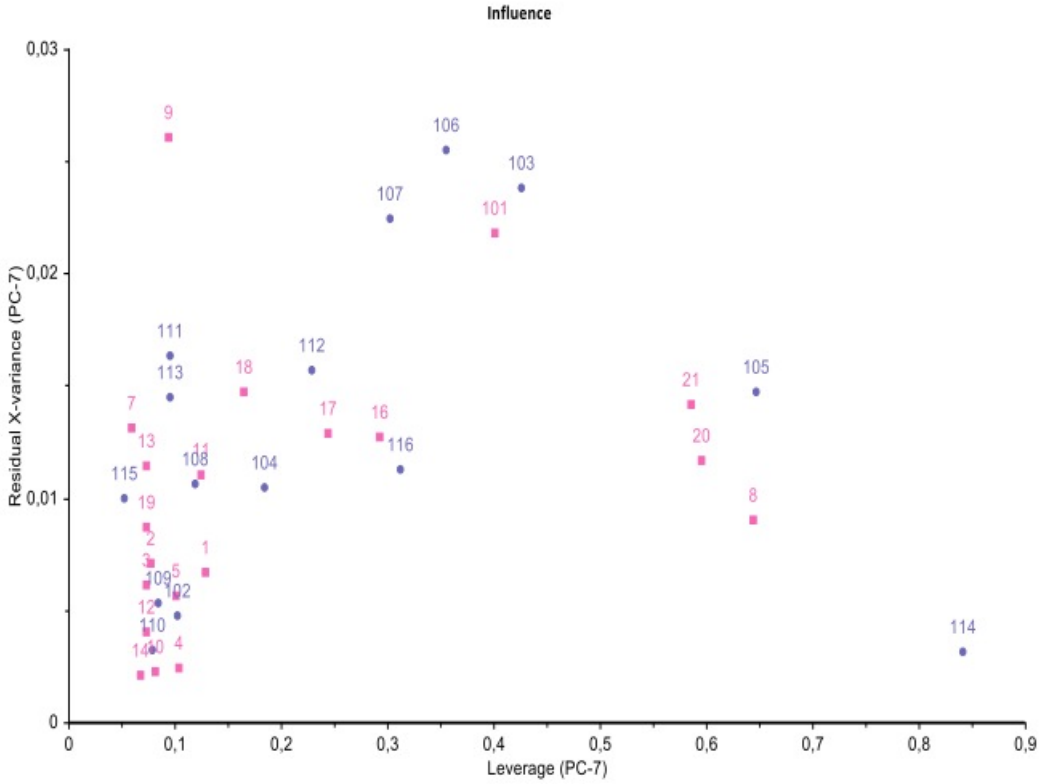


Figure A.8: *An influence and leverage plot. Shows samples that might contribute the variation and effect the model*

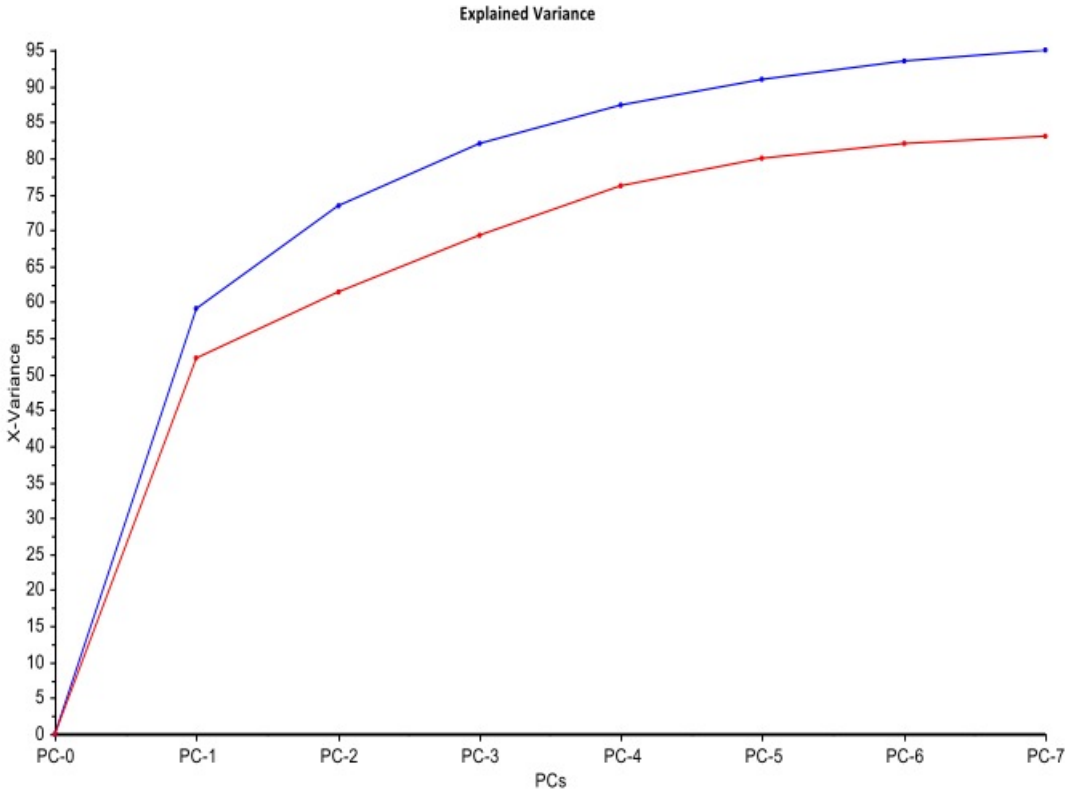


Figure A.9: *Explained variance plot*

Table A.11: *Results from the MCF derivatization of urine samples from CKD patients. The results are given in picomole per μL urine*

Urine Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	17	18	19	20	21
4-methylvalerate	0,0084	0,0071	0,0076	0,0071	0,0069	0,0066	0,0069	0,0073	0,0072	0,0071	0,0068	0,0077	0,0071	0,0067	0,0061	0,0075	0,0064	0,0070	0,0081	0,0057
Malonate	0,0470	0,0235	0,0392	0,0380	0,0252	0,0273	0,0295	0,0408	0,0321	0,0238	0,0322	0,0295	0,0390	0,0203	0,0202	0,0439	0,0184	0,0265	0,0316	0,0275
Pyruvate	0,1397	0,1065	0,1055	0,1049	0,0576	0,8047	0,1133	0,1345	0,1058	0,0902	0,0836	0,0702	0,1320	0,0603	0,0667	0,2006	0,1023	0,0997	0,1400	0,1520
3-methyl-oxovalerate	0,0205	0,0167	0,0208	0,0166	0,0197	0,0304	0,0165	0,0160	0,0270	0,0133	0,0133	0,0123	0,0152	0,0116	0,0180	0,0214	0,0237	0,0120	0,0280	0,0195
Fumarate	0,0432	0,0426	0,0487	0,0761	0,0210	0,6459	0,1243	0,0682	0,1133	0,0662	0,0498	0,0423	0,0527	0,0508	0,0508	0,1446	0,1323	0,0457	0,0774	0,1083
Lactate/Methylglyoxal	0,4764	0,2902	0,3618	0,5009	0,1542	32,9034	1,1334	1,0673	1,1799	0,3156	0,5354	0,2630	0,4652	0,7161	1,7640	1,4967	0,8486	0,9634	0,9706	2,5815
Succinate	0,2628	0,2259	0,4795	0,4703	0,1577	2,2126	0,4773	0,3419	0,6426	0,5203	0,3105	0,2934	0,4093	0,3196	0,1550	0,6396	0,4793	0,2011	0,3359	0,6586
Citraconate/Itaconate	0,0822	0,0681	0,3581	0,6441	0,1218	0,0460	0,3014	0,3741	0,6580	0,4377	0,0836	0,1017	0,3605	0,0858	0,0781	0,3119	0,3197	0,0702	0,2524	0,2128
Benzoate	0,1501	0,0956	0,1825	0,1529	0,0977	0,1869	0,0965	0,0692	0,6388	0,0586	0,0636	0,0845	0,1141	0,0418	0,0372	0,0861	0,0414	0,0815	0,1023	0,9963
Citramalate	0,2000	0,0957	0,0840	0,0694	0,0546	0,0644	0,1320	0,0910	0,0361	0,0506	0,1140	0,0399	0,0508	0,0410	0,0676	0,2520	0,0501	0,0786	0,0667	0,0277
Glycine	0,6485	1,3783	0,8689	2,8495	0,0000	4,1183	2,5989	3,9311	5,0363	1,0759	2,4196	0,9966	2,9659	10,2671	0,0000	9,2326	2,7814	0,0000	17,2682	1,8058
Alanine	0,6526	1,0132	0,3669	1,0348	0,1547	2,9735	1,4296	0,7148	1,6829	0,4109	1,2554	0,5061	0,6892	0,9306	1,6834	1,5862	1,7200	1,0079	3,7712	0,7248
O-acetyl-L-serine	1,8712	1,8464	0,7342	1,0835	0,1923	3,0725	1,4921	0,8896	2,1653	0,4159	3,1688	0,8731	1,4776	1,3569	1,2239	4,2184	1,1068	0,9281	4,6046	0,6156
Phenylacetate	0,3522	0,4694	0,2635	0,3243	0,3368	0,1190	0,1918	0,1310	0,2263	0,1671	0,4535	0,6018	0,4357	0,1563	0,0871	0,2692	0,1434	0,2585	0,2903	0,3946
Salicylate	0,2336	0,3847	0,3087	0,4548	0,1991	0,1097	0,2982	0,4655	0,6572	0,1819	0,5396	0,3562	0,3236	0,2590	0,1889	0,9867	0,0773	0,2633	0,3068	0,1160
2-aminobutyrate	0,0242	0,0217	0,0196	0,0220	0,0076	0,0285	0,0202	0,0294	0,0336	0,0149	0,0171	0,0188	0,0335	0,0184	0,0359	0,0226	0,0432	0,0207	0,0534	0,0508
Adipate	0,0405	0,0342	0,0356	0,0614	0,0000	0,0473	0,0240	0,0319	0,0340	0,0196	0,0293	0,0255	0,0805	0,1555	0,0197	0,1018	0,0186	0,0209	0,0405	0,0330
Valine	0,1657	0,1345	0,1038	0,1576	0,0514	0,5999	0,2039	0,1149	0,1898	0,0730	0,1800	0,1334	0,1366	0,1370	0,4414	0,1887	0,3436	0,1067	0,5009	0,0642
β-Hydroxypyruvate	0,1890	0,1430	0,0906	0,1739	0,0888	0,2621	0,4122	0,8710	1,0327	0,2666	0,1475	0,1622	0,1790	0,2369	0,4079	0,7723	0,3763	0,2091	0,4382	0,1420
alpha-Ketoadipate	1,9071	1,0035	0,5925	1,7412	0,4816	1,3192	0,8312	3,6395	0,7016	0,6230	1,0027	0,8449	0,7622	1,1380	0,3198	1,8342	0,3552	0,3380	0,5200	0,2921
Leucine	0,0451	0,0526	0,0200	0,0275	0,0101	0,3048	0,0478	0,0265	0,0674	0,0107	0,0808	0,0267	0,0224	0,0229	0,1674	0,0809	0,0604	0,0194	0,2025	0,0033
Threonine	0,2908	0,1863	0,1851	0,5021	0,0839	0,0000	0,5572	0,5266	1,2905	0,1080	0,5174	0,2547	0,2808	0,3630	0,7471	0,5159	1,2519	0,0000	1,7761	0,1539
Malate	0,0082	0,0083	0,0081	0,0176	0,0016	0,7234	0,0364	0,0276	0,0463	0,0172	0,0124	0,0100	0,0079	0,0129	0,0178	0,0520	0,0525	0,0110	0,0194	0,0206
Proline	0,0381	0,0268	0,0421	0,1135	0,0106	0,6650	0,5069	0,0850	0,1696	0,0417	0,0618	0,0612	0,0147	0,3746	0,2176	0,2044	1,1438	0,1353	2,2184	0,0249
Aspartate	0,0790	0,0600	0,0750	0,0885	0,0395	0,1509	0,0776	0,1213	0,1297	0,0436	0,0563	0,0759	0,0579	0,0428	0,0381	0,1099	0,0592	0,0729	0,1464	0,0735
Citrate	0,3665	0,4082	3,4564	5,4410	0,6493	0,3836	3,8160	3,0067	7,0271	2,9744	0,4794	0,9044	3,0759	0,6097	0,4595	2,6815	2,8409	0,4983	2,2927	2,4157
Serine	0,5659	0,4186	0,2830	1,0724	0,1477	0,0848	0,4635	0,9749	1,2829	0,1932	0,6699	0,3751	0,7657	0,5633	0,4504	1,2042	0,6193	0,4765	2,1256	0,0000
Anthralinate	0,1067	0,1050	0,0157	0,0404	0,0332	0,0657	0,0599	0,0485	0,0621	0,0513	0,0948	0,0714	0,0770	0,0248	0,0382	0,1690	0,0162	0,1043	0,0842	0,0169
Allantoin	0,9398	0,3460	0,1868	0,4596	0,1969	0,3884	0,3705	0,3229	0,1968	0,3607	0,6794	0,4771	0,4201	0,2504	1,7167	0,6497	0,1058	0,5206	0,7607	0,0969
N-Acetyl-L-glutamate	0,4426	0,4587	0,5774	0,4234	0,4761	0,4540	0,5263	0,4141	0,5264	0,3464	0,3592	0,4404	0,3649	0,3412	0,3560	0,3642	0,2982	0,3429	0,4752	0,4574
Glutamate	0,1738	0,1349	0,1470	0,1511	0,1033	0,4366	0,1424	0,2235	0,2524	0,0916	0,1718	0,1451	0,1431	0,1306	0,1353	0,2111	0,1867	0,1240	0,2515	0,1907
Methionine	0,0129	0,0107	0,0038	0,0073	0,0044	0,0349	0,0082	0,0035	0,0054	0,0033	0,0109	0,0069	0,0059	0,0063	0,0195	0,0180	0,0123	0,0065	0,0289	0,0041
Isocitrate	0,7539	0,4286	0,4630	0,6809	0,3012	0,0767	0,4060	0,6736	0,5421	0,2746	0,3411	0,3666	1,0516	0,1932	0,1049	0,7180	0,1835	0,3950	0,3762	0,2492
Hippurate	44,9766	27,5396	30,1407	31,7291	19,5520	30,2356	26,6603	13,1187	14,1970	11,9307	9,4151	20,0164	33,3945	6,0842	11,0825	22,8952	4,9713	17,4565	28,9692	5,1334
Phenylalanine	0,1220	0,0938	0,0556	0,0973	0,0295	0,3170	0,1370	0,0977	0,1068	0,0374	0,1399	0,0524	0,1164	0,0710	0,1900	0,1706	0,1504	0,0501	0,3672	0,0405
2.4-diaminobutyrate	0,0180	0,0163	0,0167	0,0162	0,0000	0,0165	0,0194	0,0200	0,0154	0,0137	0,0150	0,0173	0,0223	0,0141	0,0000	0,0156	0,0000	0,0169	0,0169	0,0000
4-imidazoleacrylate	0,0427	0,0253	0,0209	0,0341	0,0112	0,0561	0,0184	0,0160	0,0241	0,0166	0,0161	0,0142	0,0284	0,0114	0,0000	0,0330	0,0000	0,0136	0,0140	0,0000
p-coumarate	0,0083	0,0054	0,0043	0,0041	0,0000	0,0050	0,0041	0,0017	0,0032	0,0032	0,0068	0,0000	0,0067	0,0037	0,0000	0,0049	0,0000	0,0021	0,0041	0,0000
Ornithine	0,1662	0,0789	0,0436	0,1519	0,0528	0,1594	0,2088	0,1997	0,0541	0,1048	0,1073	0,0743	0,1465	0,1303	0,1778	0,2520	0,0820	0,0846	0,1865	0,0920
Lysine	0,1707	0,1594	0,0830	0,1918	0,0544	0,2800	0,2863	0,3639	0,2410	0,0930	0,2781	0,1414	0,1433	0,3195	0,4245	0,4107	0,2362	0,1667	0,5961	0,0907
Ferulate	0,0305	0,0646	0,0695	0,0387	0,0000	0,0286	0,0555	0,0166	0,0280	0,0341	0,1005	0,0000	0,0986	0,0499	0,0000	0,0687	0,0000	0,0000	0,0000	0,0000
Histidine	0,1399	0,1504	0,0640	0,2795	0,0000	0,2275	0,1484	0,4313	0,2460	0,0000	0,2861	0,1375	0,3655	0,1577	0,1311	0,3011	0,1564	0,2633	0,9753	0,1200
Tyrosine	0,1410	0,1276	0,1199	0,1440	0,0386	0,3198	0,1747	0,2127	0,2489	0,0577	0,1336	0,0838	0,1961	0,1096	0,1367	0,1688	0,2372	0,0868	0,4630	0,0926
Thryptophane	0,1814	0,2060	0,1405	0,2256	0,0662	0,2354	0,2307	0,2422	0,2847	0,0948	0,2171	0,1344	0,3302	0,1722	0,1593	0,2002	0,2796	0,1598	0,5825	0,1846
Cystein2	1,2062	1,1720	0,7468	1,3086	0,6854	1,0670	1,7960	1,4735	0,9523	0,9150	1,6933	1,2628	1,8327	1,2051	1,0555	1,8960	0,8148	0,9787	1,1717	0,9351

Table A.12: *Results from the MCF derivatization of urine samples from healthy . The results are given in picomole µL urine*

Urine Sample	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116
4-methylvalerate	0,009	0,007	0,007	0,008	0,009	0,008	0,009	0,007	0,006	0,007	0,007	0,008	0,009	0,009	0,008	0,007
Malonate	0,094	0,048	0,067	0,055	0,090	0,069	0,092	0,048	0,034	0,052	0,067	0,071	0,066	0,117	0,046	0,060
Pyruvate	0,275	0,071	0,155	0,079	0,343	0,149	0,255	0,185	0,069	0,066	0,236	0,177	0,151	0,211	0,092	0,087
3-methyl-oxovalerate	0,057	0,014	0,018	0,017	0,065	0,023	0,059	0,026	0,014	0,011	0,024	0,019	0,025	0,030	0,016	0,015
Fumarate	0,111	0,048	0,058	0,058	0,083	0,100	0,088	0,051	0,026	0,046	0,096	0,089	0,045	0,157	0,043	0,091
Lactate/Methylglyoxal	1,584	0,343	0,587	0,413	0,779	0,989	1,333	0,509	0,357	0,363	0,697	0,647	0,574	1,261	0,469	0,508
Succinate	1,887	0,564	1,316	1,009	0,738	1,348	1,383	0,635	0,303	0,527	0,950	0,977	0,590	2,839	0,669	1,954
Citraconate/Itaconate	1,834	0,506	0,520	1,907	2,171	0,876	2,268	0,502	0,469	0,535	1,827	0,970	0,971	1,462	0,670	1,956
Benzoate	0,107	0,045	0,072	0,321	0,084	0,136	0,079	0,046	0,088	0,043	0,082	0,205	0,106	0,163	0,116	0,096
Citramalate	0,273	0,049	0,116	0,058	0,151	0,071	0,169	0,049	0,010	0,049	0,100	0,091	0,113	0,188	0,065	0,078
Glycine	22,080	1,038	7,625	10,600	11,321	3,328	20,202	3,396	1,380	1,522	7,396	5,708	6,415	6,985	4,241	10,812
Alanine	5,281	0,254	1,659	2,075	3,202	1,244	4,529	1,347	0,449	0,588	2,756	1,459	2,533	1,437	1,775	1,637
O-acetyl-L-serine	5,436	0,376	6,114	1,374	3,658	4,550	5,206	1,308	0,449	0,404	2,169	2,583	1,818	1,855	1,453	1,658
Phenylacetate	0,384	0,074	0,272	0,736	0,365	0,466	0,419	0,127	0,122	0,134	0,307	0,408	0,663	0,748	0,204	0,212
Salicylate	0,182	0,174	0,236	0,363	0,231	0,160	0,239	0,100	0,227	0,381	0,297	0,993	0,270	0,087	0,121	0,259
2-aminobutyrate	0,160	0,019	0,087	0,095	0,204	0,070	0,248	0,125	0,026	0,025	0,116	0,065	0,098	0,065	0,077	0,065
Adipate	0,090	0,022	0,054	0,038	0,102	0,102	0,085	0,076	0,020	0,028	0,069	0,044	0,046	0,166	0,061	0,056
Valine	0,547	0,073	0,297	0,260	0,712	0,311	0,758	0,339	0,090	0,098	0,360	0,251	0,381	0,248	0,226	0,194
beta-Hydroxypyruvate	0,270	0,211	0,317	0,165	0,672	1,656	0,824	0,312	0,628	0,471	0,254	0,351	0,204	0,418	0,198	0,414
alpha-Ketoadipate	1,985	0,346	1,430	0,831	1,287	0,728	0,899	1,111	0,333	0,464	1,042	1,011	0,999	2,718	0,842	0,998
Leucine	0,169	0,008	0,077	0,041	0,204	0,038	0,200	0,077	0,012	0,011	0,062	0,043	0,071	0,038	0,039	0,026
Threonine	1,477	0,154	0,935	0,865	1,453	0,547	1,731	0,701	0,251	0,243	1,145	0,663	1,105	0,598	0,537	0,639
Malate	0,020	0,004	0,006	0,008	0,016	0,016	0,016	0,005	0,004	0,004	0,009	0,010	0,005	0,040	0,005	0,012
Proline	0,062	0,007	0,022	0,015	0,034	0,023	0,046	0,026	0,009	0,008	0,019	0,023	0,022	0,028	0,014	0,018
Aspartate	0,223	0,052	0,098	0,143	0,220	0,161	0,180	0,104	0,039	0,036	0,160	0,113	0,119	0,151	0,091	0,089
Citrate	38,093	4,319	5,191	17,145	28,709	11,274	25,095	5,956	4,931	5,118	27,255	10,524	8,274	31,332	11,588	23,197
Serine	4,467	0,167	2,125	1,996	2,438	1,810	4,885	1,233	0,341	0,377	1,605	1,645	1,851	1,689	1,251	1,566
Anthralinate	0,098	0,009	0,064	0,070	0,264	0,170	0,153	0,146	0,032	0,071	0,044	0,111	0,068	0,125	0,028	0,048
Allantoin	0,443	0,148	0,406	0,219	0,712	0,413	0,699	0,351	0,140	0,168	0,217	2,866	0,298	5,820	0,202	0,255
N-Acetyl-L-glutamate	0,467	0,329	0,415	0,315	0,387	0,319	0,393	0,293	0,252	0,246	0,282	0,303	0,237	0,231	0,188	0,274
Glutamate	0,516	0,142	0,320	0,250	0,554	0,269	0,475	0,286	0,140	0,132	0,306	0,290	0,279	0,254	0,209	0,235
Methionine	0,044	0,003	0,018	0,016	0,047	0,008	0,042	0,029	0,003	0,003	0,034	0,012	0,025	0,019	0,011	0,004
Isocitrate	2,208	0,597	1,070	0,858	2,395	1,556	1,810	1,151	0,342	0,347	1,047	1,076	0,927	1,788	0,734	1,259
Hippurate	19,498	0,528	21,616	34,631	23,939	24,578	18,334	5,352	16,872	9,040	13,316	35,147	18,033	19,950	15,354	23,665
Phenylalanine	0,619	0,057	0,323	0,216	0,999	0,254	1,131	0,347	0,064	0,057	0,272	0,276	0,356	0,214	0,209	0,179
2.4-diaminobutyrate	0,018	0,010	0,019	0,018	0,021	0,022	0,019	0,018	0,021	0,017	0,021	0,023	0,017	0,063	0,020	0,020
4-imidazoleacrylate	0,191	0,048	0,154	0,128	0,151	0,093	0,167	0,052	0,058	0,034	0,168	0,090	0,207	0,121	0,079	0,071
p-coumarate	0,007	0,006	0,010	0,005	0,008	0,011	0,009	0,007	0,004	0,000	0,008	0,024	0,012	0,009	0,058	0,008
Ornithine	0,156	0,030	0,158	0,123	0,307	0,197	0,293	0,153	0,057	0,059	0,171	0,159	0,149	0,116	0,079	0,096
Lysine	0,902	0,068	1,619	0,320	0,706	0,379	0,636	0,443	0,181	0,108	0,535	0,428	0,453	0,316	0,270	0,290
Ferulate	0,000	0,075	0,054	0,022	0,096	0,283	0,083	0,169	0,059	0,024	0,000	0,151	0,000	0,076	0,030	0,029
Histidine	9,857	0,094	4,666	1,572	8,793	1,279	13,633	2,000	0,312	0,176	3,145	3,233	6,730	1,268	1,974	2,140
Tyrosine	0,845	0,100	0,675	0,359	2,101	0,334	1,151	0,827	0,155	0,104	0,472	0,594	0,562	0,232	0,350	0,193
Thryptophane	0,983	0,187	0,572	0,384	1,064	0,465	1,107	0,681	0,167	0,179	0,546	0,598	0,497	0,399	0,394	0,235
Cystein2	1,881	0,806	1,913	1,420	1,920	1,822	2,060	1,640	0,986	1,075	1,817	1,897	2,029	2,323	1,823	1,751