Methodical aspects of urinary metabolomics by liquid chromatography - mass spectrometry

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.) an der Naturwissenschaftlichen Fakultät IV
-Chemie und Pharmazie-
der Universität Regensburg

vorgelegt von

Franziska C. Vogl

aus Eggenfelden

im Jahr 2019

Diese Doktorarbeit entstand in der Zeit von Juli 2012 bis August 2018 am Institut für Funktionelle Genomik an der Universität Regensburg.

Die Arbeit wurde angeleitet von Prof. Dr. Peter J. Oefner.

Promotionsgesuch eingereicht am: 12.04.2019

Für meine Eltern

Danksagung

So, nun ist es soweit und ich darf endlich all denjenigen Menschen meinen Dank aussprechen, die maßgeblich dazu beigetragen haben, dass ich diese Arbeit fertig stellen konnte.

An erster Stelle gilt mein Dank **Prof. Dr. Peter J. Oefner**, der mir es ermöglicht hat, an seinem Lehrstuhl diese interessante Arbeit zu verfassen. Vielen Dank auch für die Übernahme des Erstgutachtens. Die Möglichkeit auf internationalen Konferenzen Erfahrung sammeln zu dürfen, sowie die fachliche Begleitung und Unterstützung habe ich sehr geschätzt.

Vielen Dank auch an **Prof. Dr. Frank-Michael Matysik**, der sich bereitwillig als Zweitgutachter zur Verfügung gestellt hat.

PD Dr. Katja Dettmer-Wilde möchte ich danken für die Bereitschaft als Drittprüfer und die fachliche Betreuung. Danke für deine Geduld, deine konstruktive Kritik und dein offenes Ohr in jeder Lebenslage. Mit deinem unendlichen GC/LC-MS Wissen hattest du immer einen passenden Tipp parat, der mir weitergeholfen hat. Ohne dich, hätte ich nicht gewusst, dass „nach zu, ab kommt“ und was der Unterschied zwischen Schraubenschlüssel und Schraubenzieher ist. Es hat wahnsinnig Spaß gemacht, mit dir gemeinsam am Maxi zu schrauben und dabei manchmal sogar alle Teile wieder richtig verbaut zu bekommen. Nur die Namensgebung deiner Spielgefährten (Georg, Metti, etc.) war am Anfang etwas verwirrend für mich. Unvergesslich bleibt für mich auch „Pinky and the Brain“ und deine CSI Einlage.

Des Weiteren möchte ich **Prof. Dr. Antje Baeumner** danken für die Übernahme des Prüfungsvorsitzes.

Ein ausdrücklicher Dank geht auch an die Nachwuchsförderung der Universität Regensburg für die Promotionsabschlussförderung, die ich für die Fertigstellung meiner Dissertation erhalten habe.

Bei **Prof. Dr. Wolfram Gronwald**, bedanke ich mich für die fachlich sehr kompetenten Ratschläge und den Einblick in die Welt der NMR. An das gemeinsame Skifahren erinnere ich mich gerne zurück.

Für die gute Zusammenarbeit, die gemeinsame Publikation sowie den immer guten Informationsaustausch möchte ich **Dr. Peggy Sekula** recht herzlich danken.

Es war mir eine Ehre **Dr. Magda Waldhier** kennen gelernt zu haben. Die gemeinsamen Yogaabende waren immer sehr gesprächig und lustig. Und auch durch die Niederbayern Fraktion, bestehend aus **Dr. Christian Wachsmuth** und **Raffaela Berger** hab ich mich wie zu Hause gefühlt. Ihr wart meinem niederbayrischen Dialekt gewachsen und habt auch das besondere bayrische Lob „basst scho“ richtig deuten können. Die entstandene und bis heute bestehende Freundschaft mit euch Dreien bedeutet mir sehr viel.

Ein besonderes Dankeschön auch an **Dr. Helena Zacharias, Jochen Hochrein**, **Franziska Taruttis** und **PD. Dr. Julia Engelmann** die mir bei jeglichen statistischen und programmier-technischen Problemen zur Seite standen.

Danke an „Pinky“ alias **Nadine Nürnberger**. Du warst unser farbiger Stern im sonst eher tristen Laboralltag. Leider hat es dich wieder in die Heimat gezogen. Deine Geburtstagsparty mit Tortenschlacht ist jedenfalls legendär. Das Küken, **Claudia Samol,** hat jede Mittagspause zum spannenden Erlebnis gemacht. Als unermüdliche „Pippettöse“ und „Vorthexe“ war es toll mit dir im Labor zu arbeiten, auch wenn die Aliquotierung der Urinproben nie enden wollte. Bei der guten Seele, **Lisa Ellmann**, bedanke ich mich auch recht herzlich für die Hilfestellungen im Labor, vor allem für die Zeit, als ich nicht mehr ins Labor durfte und du meine etwas komplizierten Anweisungen interpretieren musstest.

Dem „PC-Betreuer vor Ort“ **Christian Kohler** möchte ich besonders Danken. Du hast dich immer sofort um alle meine kleineren und auch größeren Computer Probleme und Problemchen gekümmert. Und auch bei der Namensgebung deiner Tochter hast du alles richtig gemacht.

Meinen Kaffee Buddies **Dr. Jörg Reinders** und **Dr. Yvonne Reinders**, möchte ich auch noch danken. Ihr hattet immer Zeit für ein Gespräch, sowohl fachlicher als auch privater Art. Eure kulinarische Verpflegung war ein Highlight, in deren Gunst wir des Öfteren bei so manchen Grillpartys kamen.

Für das sonnige Gemüt**,** die heitere, aufbauende Art und auch dafür, dass du immer einen Übernachtungsplatz hattest, bedanke ich mich bei **Elke Perthen.**

Den beiden NMR-lern, **Dr. Trixi von Schlippenbach** und **Dr. Philipp Schwarzfischer**, sowie den alten Hasen **Dr. Nadine Assmann, Dr. Matthias Klein** und **Dr. Martin Almstetter,** die mir vor allem am Anfang sehr geholfen haben, möchte ich danken.

Bei Dr. Thomas Stempfl und **Dr. Christoph Möhle** bedanke ich mich für das Asyl am Ende meiner aktiven Zeit am IFG. Mit euren Labordamen, **Jutta Schipka** und **Susanne Schwab,** waren so manche Mittagspausen im Nu vorbei. Und auch die beiden Organisationstalenten **Eva Engl** und **Sharon Petersen** möchte ich noch dankend erwähnen.

Allen ehemaligen und aktiven Kollegen am IFG danke ich für die schöne und lehrreiche Zeit an die ich mich gerne erinnere.

Mein besonderer Dank geht an **Marianne Felixberger**. Nicht nur als Oma von Marie bist du immer für uns da. Du hast mir den Rücken frei gehalten und mich motiviert, damit ich diese Arbeit fertig stellen konnte. Danke für alles.

Meinen Eltern **Gisela** und **Wolfgang Vogl**, sowie meiner Lieblingsschwester **Alexandra Lehner** gebührt der größte Dank. Ohne euren Rückhalt hätte ich diese Arbeit nicht zu Ende gebracht. Kein Jammern und Klagen meinerseits war euch zu viel. Kein Weg zu weit, um mir zu helfen. Ihr habt mich bedingungslos unterstütz. Ihr könnt euch nicht vorstellen, wie wichtig das für mich war. Einer für alle und alle für einen!

Johannes und **Marie,** ihr seid mein Ein und Alles, mein Fels in der Brandung. Ihr wart immer verständnisvoll, wenn ich mal wieder wegen meiner Doktorarbeit keine Zeit für euch hatte. Das hat mir manchmal das Herz gebrochen. Eure Liebe war meine wichtigste Stütze, die mir geholfen hat, diese Arbeit abzuschließen. Ich freue mich riesig auf unser Familienleben zu viert und auf die Abendteuer die noch auf uns warten.

An dieser Stelle möchte ich auch an einen sehr guten Freund erinnern, der viel zu früh von uns gegangen ist: **Dr. Alexander Riechers**. Ich kann es immer noch nicht fassen, dass du nicht mehr unter uns bist. Mir fehlen dein Lachen, deine immer fröhliche Art und die Gummibärchen, die du immer für alle bereit hattest. Wie gerne hätte ich diesen Moment mit dir geteilt und mit dir gefeiert.

Mit deiner unbändige Hilfsbereitschaft, deinem immensen Wissen und deiner liebe zur Wissenschaft, warst du immer ein Vorbild für mich. Alex, du fehlst mir!

1. Table of contents

2. Abbreviations and Acronyms

AA	Anthranilic acid
AER	Albumin excretion rate
AKI	Acute chronic kidney injury
CKD	Chronic kidney disease
CMT	C-Mannosyltryptophan
CRE	Creatinine
CRT	Creatine
Da	Dalton
DMX	Dimethylxanthine
DN	Diabetic nephropathy
eGFR	Estimated glomerular filtration rate
eGFRcr	Estimated GFR based on creatinine
ESI	Electrospray ionization
ESRD	End stage renal disease
FDA	Food and Drug Administration
GC	Gas chromatography
GFR	Glomerular filtration rate
GUA	Guanidinoacetate
HAA	Hydroxyanthranilic acid
HIAA	Hydroxyindolacetic acid
HIP	Hippuric acid
HK	Hydroxykynurenine

HMDB	Human Metabolome Database
HPLC	High performance liquide chromatography
HRTOF	High-resolution time-of-flight
HX	Hypoxanthine
IAA	Indole-3-acetic acid
IDA	Isotope dilution analysis
ILA	Indole-3-lactic acid
INC	Indole-3-carboxaldehyde
IPA	Indole-3-propionic acid
IS	Internal standard
IS	3-Indoxylsulfate
KA	Kynurenic acid
KYN	Kynurenine
LC	Liquid chromatography
LLOQ	Lower limit of quantification
LOD	Limit of detection
LOQ	Limit of quantification
m/z	Mass-to-charge ratio
MEL	Melatonin
MG	Methylguanine
MI	1-Methylinosine
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MX	Methylxanthine
NA	Nicotinic acid

NAM	Nicotinamide
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
PCA	Principle component analysis
PLS-DA	Partial least squares discriminant analysis
PSU	Pseudouridine
QA	Quinolinic acid
qMS	Quadrupole mass spectrometry
QqQ-MS	Triple quadrupole MS
qTOF	Quadrupole-Time-of-flight hybrid mass spectrometry
RSD	Relative standard deviation
S/N	Signal to noise ratio
SER	Serotonin
SIL-IS	Stable isotope labeled internal standard
SVM	Support vector machines
TG	Tiglylcarnitine
TIC	Total ion current
TOF	Time of flight
TPO	Tryptophol
TRP	Tryptophan
TRY	Tryptamine
UA	Uric acid
ULOQ	Upper Limit of quantification
XA	Xanthurenic acid
XIC	Extracted ion chromatogram
XT	Xanthine

3. Motivation

Metabolomics aims at the holistic analysis of the qualitative and quantitative composition of low molecular weight compounds in biological systems.[1] The major goal is the determination of metabolites, quantitatively as well as qualitatively, that differ between biological groups. This information is contributing to the understanding of the underlying biological mechanisms under investigation.

A common matrix for large scale metabolomics studies is urine since it is obtained non-invasively.[2,3] However, urinary output can vary widely due to various processes involved in the body's water balance. The hydration is influenced by water uptake, loss of water due to respiration, perspiration and defecation[3]. As a result, variable urine amounts are excreted, which effects the metabolite concentrations measured in the urine. Hence, untargeted LC-MS based metabolomics analyses of urine samples are challenging due to these biological variances that are not necessarily related to the phenotype under investigation. Consequently, to reveal the true biological variances, normalization strategies are necessary.[4] Commonly, creatinine is utilized to normalize for urinary output. Creatinine is an endogenous metabolite, produced at a constant rate and mainly excreted by glomerular filtration.[5] However, creatinine concentration can vary widely due to various factors, such as age, sex, body mass, health, diet or water intake.[4] Alternatively, the osmolality, which is the molar sum of all solutes in urine, can be applied for correcting the urinary output because it reflects the urinary output more closely.[4 6] The common normalization procedure in metabolomics encompasses a uniform dilution of urine samples followed by a post-acquisition normalization to either creatinine concentration or to the sum of all integrals of each sample. Nevertheless, these normalization methods are unable to correct for column overloading, peak overlapping or ion suppression. Furthermore, a uniform dilution may result in failure to detect analytes in urine specimens of low concentration. Therefore, a normalization strategy is needed to overcome these shortcomings.

Chronic kidney disease (CKD) is a major public health concern affecting more than 10% of the population worldwide.[7] Serum creatinine is the most commonly used clinical biomarker of renal dysfunction. However, there are important limi-

tations to its use. For example, it rises only after 50% of kidney function is already lost. Additionally, tubular secretion of creatinine results in overestimation of renal function at lower glomerular filtration rates (GFR) and, as already mentioned above, it reflects differences in muscle mass.[8] Hence, serum creatinine concentrations can vary widely. Therefore, the estimation of GFR can be inaccurate. Cystatin C has been proposed as an alternative and more sensitive marker of eGFR.[9,10] However, it is metabolized by proximal tubular cells.[11] The ideal filtration markers for GFR estimation should exhibit no or only little net reabsorption or secretion in the nephron. This leads to the clinical necessity of additional markers for an early detection of CKD. Additionally, the underlying pathophysiological mechanisms are still not fully understood.[12] Sekula et al. found that C-mannosyltryptophan and pseudouridine were strongly and reproducibly associated with eGRF and CKD in population-based studies.[12] Therefore, a quantitative method for further investigations was required for these metabolites.

For quantification of small molecules and metabolites, the triple quadrupole mass spectrometer in multiple reaction monitoring (MRM) mode has been the golden standard. However, with recent developments regarding the mass accuracy, resolution and scan times, the relative quantification with a quadrupole Time-of-flight-mass spectrometer (qTOF-MS) in full scan mode is possible.[13] Nevertheless, very little is known about the performance of quantification in full scan mode, also called untargeted screening, in biological matrices. Lu et al. compared the performance of a triple quadrupole instrument in MRM mode to a TOF-MS in full scan mode for the analysis of 20 standard compounds (10 in positive ion mode, 10 in negative ion mode) at 5 different concentration levels and a cellular extract.[13] However, they only showed the results for one standard compound. Therefore, a comparison of the absolute quantification via MRM mode and relative quantification via qTOF-MS in full scan mode in urine is needed.

Aim#1: Evaluation of strategies to reduce sources of variance in untargeted LC-MS based metabolomic analysis

As the sample preparation is one of the first steps in the analytical process, the first goal of this thesis was the optimization of the sample preparation of urine to correct for urinary output. For this purpose, the performance of different pre- and post-acquisition normalization methods to correct for urinary output in LC-MS based metabolic fingerprinting was systematically compared. As pre-acquisition methods, sample dilution to either a uniform creatinine concentration, osmolality value, or using a constant factor were used. For the different post-acquisition normalization methods such as normalization to creatinine, osmolality and sum of all integrals, data were acquired with a constant 1:4 dilution. Urine specimens from patients with chronic kidney disease from two different epidemiological studies, namely the German chronic kidney disease study (GCKD)[14] and the Trial to reduce cardiovascular events with Aranesp therapy (TREAT)[15] study and healthy controls were chosen for evaluation of the different normalization and dilution strategies.

Aim#2: Development of a quantitative LC-QqQMS profiling method

The aim was the development of a sensitive and reliable quantitative profiling method for metabolites recently associated with renal disease, e.g. pseudouridine and C-mannosyltryptophan in order to further investigate these metabolites as potential markers. Moreover, seven metabolites identified as discriminating metabolites between patients suffering from CKD and healthy controls within the evaluation of normalization strategies together with the metabolites of the tryptophan pathway were also implemented in this method for further investigations. The validated method was applied for the quantification of 383 urine and plasma samples from the GCKD study to analyze the fractional excretion.

Aim#3: Comparison of untargeted screening LC-qTOFMS and quantitative LC-QqQMS analysis

The goal was the comparison of fingerprinting LC-qTOFMS data and quantitative LC-QqQMS data of urine samples from the GCKD study. Since, the kidney function has a broad impact on circulation metabolite levels, untargeted metabolomics is a promising approach in nephrology research. However, the major drawback of untargeted (full scan) MS-based metabolomics is the semi-quantitative nature of the data.[16] They are in arbitrary units and therefore cannot be compared across studies and further investigations e.g. in different matrices are difficult.[17] In contrast, metabolic profiling methods provide absolute concentrations for the investigated metabolites. Therefore, the generated data from the two different approaches were compared via Bland-Altman plots for a defined set of metabolites. Moreover, the lower limit of quantification (LLOQ), the upper limit of quantification (ULOQ), the linear range as well as the relative standard deviation (RSD) of the standard compounds for different concentrations were calculated and compared for both approaches.

4. Background

This chapter is focusing primarily on untargeted LC-MS based metabolomics.

4.1 Metabolomics

Metabolomics aims to investigate, qualitatively as well as quantitatively, all metabolites in a biological system simultaneously. Metabolites are low molecular weight compounds (<1000 Da) of organic or inorganic origin, which comprise mainly educts, intermediates or products of enzymatic biochemical reactions.[18] Metabolomics encompasses the study of discriminative changes in metabolite profiles due to genetic modifications, physiological stimuli, environmental, nutritional or other factors [18]. Hereby, the ultimate goal is the recognition of discriminative pattern rather than identifying all detected compounds.[19]

4.1.1 Fingerprinting analysis

Untargeted metabolomics, also called fingerprinting analysis, is considered to be the true “omics” approach.[20] Fingerprinting analysis enables a global snapshot of the metabolic phenotype and therefore is also a valuable diagnostic tool. [19,21,22] The metabolome is the end-stage of the “omics”-cascade and hence reflects genetic, disease, life style, as well as environmental responses. Metabolic profiles, obtained by analyzing metabolites, for instance, in body fluids such as serum or urine, provide additional information that cannot be obtained by investigating the genotype, transcriptome, or even proteome. This additional knowledge can contribute to the achievement of personalized drug therapy.[22] The analysis of reflective metabolite pattern enables therapy on a more personalized level. Besides genotype variations, environmental influences like diet, which individually influence drug metabolism or toxicity, are also taken into account.[23] Moreover, the characterization of metabolic networks by multiparametric analysis may enable earlier diagnoses of some diseases.[24]

As a large scale procedure, fingerprinting analysis covers a wide range of metabolites and enables a high number of experiments in a relatively short

time.[20,21] However, the varying molecular weight, hydrophobicity/hydrophilicity, acidity/basicity and boiling points of the metabolites represent major technical challenges. Technical platforms need to fulfill a variety of requirements like high accuracy, sensitivity and reproducibility.[20] Primarily, two different platforms are employed, which are either based on mass spectrometry (MS), coupled to a separation technique, or nuclear magnetic resonance (NMR). In principle, gas-chromatography-mass-spectrometry (GC-MS) and liquid-chromatography-mass-spectrometry (LC-MS) are the mainly used MS based platforms. Mass spectrometry based approaches have the potential not only to quantify metabolites with high sensitivity and selectivity but also to identify them. However, sample preparation is a crucial step for MS based analyzing techniques. Sample preparation should be as simple as possible to minimize the loss of analytes. Therefore, sample matrices that are compatible with the analytical technique, are commonly only diluted and the dilution is adapted with respective to the sample concentration and the dynamic range of the MS. The coupling to a chromatographic separation technique reduces complexity of the matrix by adding the time scale as a further dimension. Moreover, it enables an isobaric separation and achieves additional physicochemical information of the compounds.

Since GC-MS with electron ionization (EI) is a well-established method in metabolomics, robust protocols are available for fingerprinting analysis including maintenance, sample preparation, chromatographic evaluation and identification.[21] The highly reproducible fragmentation with EI enables the implementation of mass spectral libraries which facilitates metabolite identification. Nevertheless, GC-MS is restricted to volatile analytes although chemical derivatization increases the range of detectable metabolites. Derivatization is a major bottleneck of GC-MS analysis because it can lead to different forms of derivatives for the same analyte, the derivatization reaction can be incomplete or it can lead to byproducts or degradation.[25] Furthermore, the additional sample preparation step may increase the analytical variance and can cause the loss of metabolites. However, there are correction methods, e.g. the derivatization of standard compounds and normalization methods, to overcome these shortcomings.[25] Nevertheless, GC-MS is considered to be the golden standard in metabolomics.

Yet, LC-MS has recently enjoyed a growing popularity as platform in metabolomics fingerprinting studies mostly because it can be applied to a larger diversity of molecules[21,26,27]. Moreover, there is no need for a derivatization step, which simplifies the sample preparation. The lower analysis temperatures allow a gentler metabolite handling.[28] However, metabolite identification is more difficult due to the lack of comprehensive databases. Additionally, LC-MS measurements of complex matrices are hampered by ion suppression. The simultaneous presence of multiple metabolites affects the ionization, which impedes quantification of these metabolites. Despite these difficulties recent enhancements in separation, ionization and mass accuracy, have contributed to a new level of performance. These new achievements made LC-MS a complimentary gold standard in metabolomics.

For the analysis of fingerprinting metabolomics data, the whole e.g. LC-MS chromatogram is exported. These signals yield characteristic patterns for each sample which can be tested for significantly different signals of different groups or can be used to classify the samples (see chapter 4.3). The significantly discriminating signals, also called features or markers, are then further investigated and identified (see chapter 4.4.). Therefore, data complexity is reduced and only discriminatory metabolites are further analyzed and biologically interpreted.

Alternatively, NMR is applied as a comprehensive platform for fingerprinting analysis. Despite limitations in sensitivity, NMR provides a rapid, non-destructive, high throughput analysis with minimal sample preparation efforts.[28] Nevertheless, spectra are complex because most metabolites result in multiple signals that often overlap with signals from other metabolites. Therefore, metabolite identification is more complicated than with MS based approaches. Moreover, a relatively large sample amount of a few hundred microliters is required for NMR analysis. Due to the often restricted sample amount this is a major disadvantage of the NMR approach.

4.1.2 Profiling analysis

The second approach in metabolomics is called profiling or targeted analysis. Thereby, defined sets of metabolites of a specific compound class or pathway are investigated quantitatively. For this purpose, analytical methods are special-

ly tailored to the respective metabolite set to be quantified. Moreover, sample preparation is normally more extensive than for fingerprinting analysis in order to minimize sample complexity prior to the measurements or to achieve a sample pre-concentration for low abundant metabolites. Reference compounds with a known amount and an appropriate assortment of internal standards (IS) are applied to establish calibration curves for the quantification of the target analytes. Instead of calibration curves, isotope dilution analysis (IDA) is also utilized for quantification (see Figure 4.1). Therefore, a known amount of the stable isotope labeled internal standard (SIL-IS) is added to the analyte to be quantified before the sample preparation procedure.[29] Hence, the losses of the analyte and SIL-IS are proportional and both are affected equally by ion suppression, matrix effects, injection variability and signal drifts. For the quantification the ratio of the peak area of the analyte to the SIL-IS is multiplied by the concentration of the SIL-IS.[29] However, IDA is not only suited for quantification within a profiling analysis but it is also applicable for fingerprinting analysis (see chapter 6).

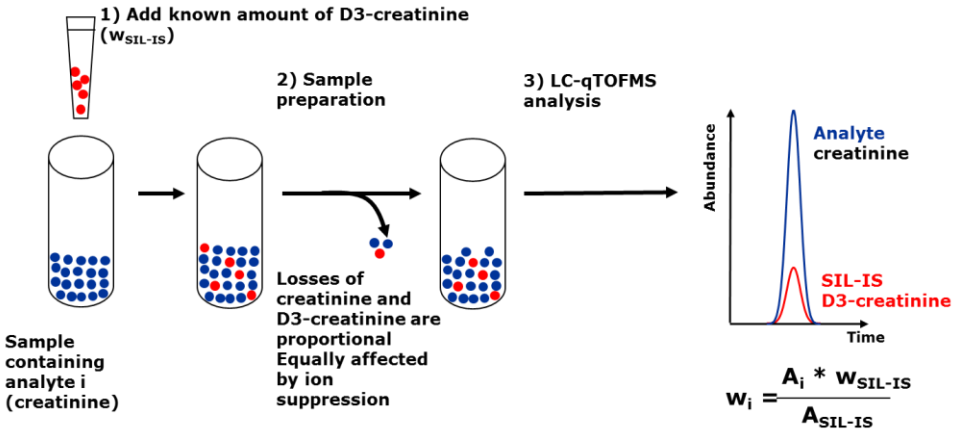


Figure 4.1: Isotope dilution analysis (IDA) exemplarily shown for the quantification of creatinine applying stable isotope labeled D3-creatinine. Adapted by permission from Springer, Practical Gas chromatography by Katja Dettmer-Wilde and Werner Engewald, Springer-Verlag Berlin Heidelberg 2014.[29]

4.2 LC-MS based metabolomics in the study of chronic kidney disease

The kidneys play an important role in the acid-base balance, the regulation of plasma volume and hormone secretion, which all maintain vertebrate homeo-

stasis.[30] However, many kidney diseases can hamper these processes. Kidney diseases which persist over 3 months or longer are considered to be chronic (chronic kidney disease, CKD).[31] Approximately 8-10% of the western population is affected by CKD.[32] The clinical picture of CKD shows a progressive loss of the kidneys ability to filter potentially toxic compounds out of the blood into the urine.[33] Consequently, toxic compounds accumulate in the body, negatively affecting biological functions. Such compounds are also called uremic toxins.[33,34] Ninety uremic retention solutes were classified accordingly following the suggestion of the European uremic toxin work group.[33,35] This list was extended in 2012 by Duration et al. who assigned 56 additional uremic toxins.[35]

Diabetes mellitus and high blood pressure are the most prominent causes of CKD.[36] Patients suffering from diabetic mellitus are at increased risk of developing diabetic nephropathy (DN). 30-40% of type 1 and 15-20% of type 2 diabetes mellitus patients have DN after 20 years. [37,38]

CKD is classified into five stages according to its severity. The classification is based on the estimated glomerular filtration rate (eGFR). The GFR is defined as the volume of plasma filtered by the glomeruli per unit of time.[39] It is equal to the Clearance Rate if the filtered substance is freely filtered and neither reabsorbed nor secreted by the kidney. Therefore, it can be measured using the clearance of filtration markers such as urinary creatinine. Urinary clearance can be calculated applying the formula $CL = U \times V / P$,[40,41] where U represents the urinary concentration, V the urinary flow rate, and P the plasma concentration.[40] GFR can also be estimated from serum levels of endogenous filtration markers (eGFR), such as serum creatinine or cystatin C, without the need of calculating the clearance.[40] A variety of different estimating equations are available using creatinine, all of them come with variable biases across populations and are imprecise.[40] Several estimation equations using cystatin C have also been published.[40] Calculation of eGFR based on cystatin C is not affected by muscle mass and diet. Hence, it is more reliable than calculations based on creatinine and not as strongly associated by sex, age and race.[40] On the downside, there is some evidence that shows cystatin C levels are influenced by tubular excretion.[40] Additionally, a combination of creatinine and cystatin C can also be used for the calculation of the eGFR.

Stage one of the CKD classification, as the mildest CKD class, shows only few symptoms (e.g. proteinuria) with a GFR < 90 mL/min/1.73 m^2, while stage five as the end stage displays severe symptoms (e.g. kidney failure) with a poor life expectancy if untreated (GFR < 15 mL/min/1.73 m^2).[33] End stage renal disease (ESRD) either requires dialysis or renal replacement. In many kidney diseases, kidney damage can also be determined by albuminuria which is diagnosed based on urinary albumin excretion rate (AER).[31,42]

Kidney disease and its associated metabolites were investigated with markedly 'low-tech' methodologies by physicians since the Middle Ages.[43] At that time, color, smell and taste of urine were applied to diagnose and specify renal disease. Nowadays, clinical markers such as serum creatinine have been established for the detection of CKD. [43] While classical assays, like colorimetric (Jaffe reaction) or enzymatic assays,[40] are still the gold standard in routine diagnostics, rapid improvements in analytical techniques made especially LC-MS a popular tool in the search for clinical markers. However, the sensitivity of these markers (e.g. creatinine), only allows the detection of CKD at later stages (see also chapters 6 and 7), leaving the urgent need for the identification of early detection markers. Aside from early detection, identification of those patients at increased risk of progressing rapidly to ESRD is a pressing issue.

In the following, studies using untargeted LC-MS metabolomics for the identification of CKD-biomarkers in human urine, serum and plasma specimens are summarized. In short, untargeted LC-MS has proved to be a very useful tool for the identification of potential novel clinical markers for CKD due to its holistic analysis character. However, the analytical methods and the statistical analysis of the acquired data are critical aspects that need to be addressed in order to achieve valid results (see chapter 6).

In 2015, Sekula et al. used untargeted LC- and GC-MS to investigate human serum specimens of three different cohorts within the general population of the KORA F4 study, the TwinsUK registry and the AASK study.[34] They quantified 493 small metabolites in human serum samples. Moreover, they analyzed the correlation between these molecules and the GFR estimated on the basis of creatinine (eGFRcr) and cystatin C levels of participants in the KORA F4 study and the TwinsUK registry. The statistical analysis yielded 54 metabolites that

were significantly associated with eGFRcr. They also found that C-mannosyltryptophan, pseudouridine, N-acetylalanine, erythronate, myo-inositol, and N-acetylcarnosine show a pairwise correlation ($r \geq 0.50$) with routine kidney function measures.[34] Moreover, they demonstrated that higher C-mannosyltryptophan, pseudouridine, and O-sulfo-L-tyrosine concentrations are related to incident CKD (eGFRcr <60 ml/min per 1.73 m^2) in the KORA F4 study. Additionally, they demonstrated that C-mannosyltryptophan and pseudouridine correlated (0.78) with measured GFR of patients of the AASK study. Adjusting both metabolites to measured GFR resulted in the disappearance of the highly significant relation to ESRD. In summary, Sekula et al. were able to demonstrate that both metabolites could be alternatively used to determine kidney function. However, this study was based on semi-quantitative data. Thus, they needed to verify these findings and specify reference ranges by quantitative data. In order to overcome this shortcoming, in the course of this thesis serum and urine samples from the GCKD study were quantitatively investigated by LC-QqQ-MS which contributed to the study published by Sekula et al in 2017.[44] In this study, the novel kidney function markers C-mannosyltryptophan and pseudouridine were characterized in blood and urine specimens of subjects with and without CKD. Additionally, fractional excretions and the relation to GFR were determined and quantitative and semi-quantitative data were compared.

Boelart et al. 2014 investigated blood serum specimens from 20 patients each with CKD stage 3 and hemodialysis requiring stage 5, and a healthy control group (N=20), by means of high-resolution LC-qTOF-MS and GC-qMS in positive and negative ionization mode.[33] This study put emphasis on the validity of the applied method. They used quality control samples and demonstrated method validity by satisfactory retention time shifts, mass accuracy and peak area fluctuations. In order to identify significantly discriminating metabolites between the investigated groups, the Mann-Whitney unpaired test (Benjamini–Hochberg FDR corrected) was applied. They were able to identify 85 metabolites associated with advanced CKD.[33] Aside from 43 metabolites that had been already reported earlier, 31 unique metabolites were identified, whose serum levels increased significantly with CKD progression, while the serum levels of 11 additional metabolites decreased with CKD progression. Eighteen novel metabolites were identified in positive ionization mode, including acetylhomoserine,

methylglutarylcarnitine, 3-methyluridine/ribothymidine, methyluric acid, nicotinuric acid/isonicotinylglycine, oxoprolylproline and the dipeptide PhePhe. [33]

Plasma samples from 30 non-diabetic men with different CDK stages were investigated applying LC-LTQ-MS and GC-MS by Shah et al. in 2013.[45] Metabolite profiles of CKD stages 2, 3 and 4 (each N= 10) were compared to identify novel biomarkers for the respective CKD stage.[45] CKD stages were determined based on the eGFR. Different sample groups were examined on the basis of Welch's t-test corrected for multiple testing by the FDR and random forest classification. Statistical analysis yielded 62 significant different metabolites between stage 3 and 2, 111 metabolites between stages 4 and 2 and 11 metabolites between stage 4 and 3. Within this study major metabolic differences were revealed, reflecting inter alia alterations in arginine metabolism.[45]

Untargeted LC-ESI-TOF-MS was applied in the study of Sato et al. in 2011 to investigate plasma specimens from patients with end-stage renal disease (ESRD), who were treated with hemodialysis (N=10).[46] As a control group, samples from healthy subjects (N=16) were used. The investigation of the plasma samples before and after hemodialysis yielded 54 metabolites whose concentrations were affected by hemodialysis. Significant differences were determined utilizing analysis of variance (ANOVA). The Tukey–Kramer's multiple comparison test for pairwise comparisons was applied for further analysis. According to the authors, the discovery of methylinosine and two unknown molecules with an m/z ratio of 257.1033 and 413.1359 as potential biomarkers could be helpful to identify the appropriate hemodialysis dose. However, these potential biomarkers need to be confirmed.[46]

Untargeted LC-qTOF-MS was also used in the study of Jia et al. in 2008.[47] The authors examined serum samples from 32 patients with chronic renal failure without renal replacement therapy and 30 healthy volunteers. They intended to discover novel biomarkers and shed light on their pathophysiological changes. Statistical analysis revealed 7 potential biomarkers: creatinine, tryptophan, phenylalanine, kynurenine and three lysophosphatidylcholines. This study emphasized the importance of untargeted LC-MS metabolomics and its future in clinical diagnostics.[47]

LC-MS was also exploited to reveal early biomarkers for DKD.[42] In order to differentiate progressive albuminuria from non-progressive albuminuria, the authors used urine specimens from patients suffering from type 1 diabetes with a normal urinary albumin excretion rate. After 5.5 years of follow-up, half of the patients had progressed from normoalbuminuria to microalbuminuria whereby the other half had remained normoalbuminuric.[42] They employed both LC-LTQ-FT-MS and GC-MS. Multivariate logistic regression analysis yielded a profile of metabolites that discriminated patients with deteriorating albumin excretion rate (microalbuminuric) from normoalbuminuric patients with an accuracy of 75% and a precision of 73%. The discriminating profile included acyl-carnitines, acyl-glycines and metabolites linked to the tryptophan metabolism. Moreover, the discriminating profile included metabolites already linked to DKD.[42]

In 2009, Zhang et al. focused on the metabolic research of diabetic nephropathy and type 2 diabetes mellitus.[37] An UPLC-TOF-MS system was utilized to differentiate global serum profiles of 8 patients suffering from diabetic nephropathy (DN), 33 type 2 diabetes mellitus (T2DM) patients and 25 healthy volunteers.[37] Moreover, the authors intended to identify potential biomarkers for DN. Principle component analysis was implemented for group separation. Distinctive clusters between patients and healthy volunteers were observed. Further, DN and T2DM patients were separated in the scores plot. An independent t-test yielded 8 metabolites significantly differentiating patients and controls. However, they only tentatively identified 3 of these metabolites: leucine, dihydrosphingosine and phytosphingosine.[37]

4.3 Data analysis of untargeted large scale metabolomic studies

Data analysis of untargeted LC-MS data encompasses several steps including data pre-processing, statistical analysis, identification of metabolites, validation of the results followed by biological classification and interpretation. Untargeted LC-qTOF-MS measurements of large scale metabolomic studies generate highly complex data. Commonly data are generated over several months or years and sample sets comprise hundreds to thousands of samples. Therefore, data variability originates not only from biological variance but also from technical

variance and batch effects. Resulting issues that need to be addressed are e.g., retention time (RT) shifts, peak alignment and missing values, just to name a few. Hence, a special focus must be placed on data processing and analysis to obtain solid and robust data for statistical and biological interpretation. Here, the basic concepts of data processing and analysis are covered as they were applied throughout this thesis. Moreover, recommendations for the instrumental analysis to avoid batch- and technical effects are described.

4.3.1 Data processing

Data processing starts with the recalibration of the mass scale (see chapters 4.4.1 and 5.3.1). Then, a peak picking algorithm searches automatically for peaks in the chromatographic trace. Bruker Daltonics Find Molecular Feature algorithm (Bremen, Germany) searches for ions that belong to one compound (peak). The search is based on a high correlation in time and m/z distances (e.g. isotopic distances) of these ions. The resulting compound is therefore the average of these clustered ions. Afterwards, peaks across the chromatograms of different samples are aligned in a single matrix. In the data matrix, which is also called bucket table, every peak is represented by its *m/z* value, retention time and area integral.

However, as already mentioned, in large scale metabolomics studies pitfalls like RT shifts occur which affect peak alignment and may increase the number of missing values. RT shifts can result in the alignment of wrong peaks throughout the samples and batches. Moreover, peaks found in some samples may not be found in other samples of the same or different batches resulting in missing values in the bucket table. In that case, it is difficult to distinguish between “true” missing values, because the signal was below the limit of detection, and peaks missed due to retention time shifts or incorrect alignment. Identifying the origin of a missing value would require the manual inspection of all signals, which is not feasible in the case of thousands of features detected. However, missing values constitute a serious problem in statistical data analysis (see chapter 6).

Accordingly, measurements must be performed with great care (see chapter 4.3.3). For severe (between-) batch effects we developed a correction strategy for retention time shifts to reduce missing values. Each batch is aligned sepa-

rately by the commercially available software which was in our case the Bruker Daltonics ProfileAnalysis software. Additionally, we developed a tool for correcting RT shifts according to one or more prominent compounds in the sample matrix. Differences in retention time for these house-keeping compounds are used to correct those of the remaining compounds (see chapter 5.4.3). Afterwards, with subsequent pairwise bucketing via an in-house written bucket assigner all the measured batches are combined (see chapter 5.4.4). The described procedure was applied in chapter 8. The Bucket assigner alone was also implemented in chapter 6. This procedure reduces missing values. Nevertheless, some missing values still remain in the data matrix because they are either true missing values or the detected area is below the signal-to-noise threshold. Consequently, this requires imputation of the missing values prior to statistical analysis. Therefore, 1/10th of the minimal detected area integral can be inserted. Alternatively, methods like missing completely at random or k-nearest neighbor imputation method (knn) can be utilized.[48,49] Moreover, prior to multivariate data analysis, it is necessary to reduce any variance in the data which is not biologically induced. Strategies to avoid batch effects and technical variations in advance are described in chapter 4.3.3. Different normalization methods were tested in this thesis to reduce any contribution from unwanted biases and experimental variance before and after the measurement of the urine samples (for details see chapter 6). However, normalization to the sum of all buckets can be utilized independent of the sample origin.

4.3.2 Data visualization and statistical analysis

Statistical analysis approaches applied for high throughput data like LC-MS data are mainly adapted from earlier developed omic technologies.[50]. Commonly, univariate or multivariate approaches are used in order to search for group separation. Classical univariate analysis are the two-sided t-test or the analysis of variance (ANOVA) followed by a post-hoc test. Especially for large scale metabolomics where thousands of features are detected, it is important to correct for multiple testing. Adjusting for multiple testing by controlling the false discovery rate according to Benjamini and Hochberg[51] is an example for such a correction. Multivariate analysis methods use not only single discriminating me-

tabolites but also dependency structures to differentiate sample groups.[50] Most prominently utilized methods are principle component analysis (PCA), cluster analysis and classification methods like the random forest (RF) classifier. As a starting point for data analysis, the PCA is a very valuable tool to decrease complexity of large scale LC-MS data and to visualize differences between samples. Thereby, loading plots can help to identify discriminating metabolites. Another popular method in metabolomics data analysis is the partial least squares discriminant analysis (PLS-DA).Instead of PCA, it is a supervised discriminant analysis method, which is dependent on class labels.[52] In order to analyze high-throughput data in more detail, cluster analysis like hierarchical clustering is a valuable unsupervised method. For more details see chapter 5.5.1. The RF classifier is a supervised classification method based on decision trees. Therefore, the data set is split in two sets, training and test set.[53] Both are selected randomly applying resampling with replacement which is also known as bootstrapping. As the names intend, the training set is used to develop the decision tree and the test set is applied to calculate the classification accuracy. [53] According to the literature, the RF is a very accurate and robust classification method.[53] Support vector machine (SVMs) or regularized linear regression are alternative machine learning algorithms widely used in metabolomics data analysis. These algorithms also use training sets to learn rules and form patterns, which then can be utilized to analyze new data.[54]

4.3.3 Handling batch effects and technical variations

Batch effects, also called between-block effects, occur in large scale studies where samples are measured in several batches over an extended time period. In order to prevent any kind of batch effects and technical variation, it is advisable to consider the following experimental setup recommendations. Firstly, the samples should be measured in random order. Secondly, the instrument should be cleaned and calibrated before every measurement and the column should be equilibrated by injecting blanks and/or a standard before analyzing the sample sequence. Attention should also be paid to the eluent preparation and usage. It should be prepared exactly the same way every time. Blank samples must be measured repeatedly throughout the sequence in order to check for sample carry over and contaminations originating from the column. Moreover, quality con-

trol samples should be regularly interspersed throughout the measurement to investigate inter- and intra- batch effects. These QC samples can be used to check for batch effects including retention time shifts after the measurement.

4.4 Identification of metabolites by LC-MS analysis

As mentioned above, untargeted metabolomics is a retrospective analysis. In order to interpret the acquired data biologically, it is necessary to identify at least those metabolites that differentiate the biological groups the most. However, this is a very challenging and labor-intensive step especially for untargeted metabolomics of complex mixtures like urine, which contain exogenous as well as endogenous metabolites originating from individual diet, medication, life-style and environmental influences.[55] Moreover, the available databases (e.g. Human Metabolome Database (HMDB)[56], Metlin[57], Madison Metabolomics Consortium Database (MMCD)[58], MassBank[59], LIPID MAP[60]) for metabolite identification from LC-MS data are limited in their content and a comprehensive database is missing.[61] Therefore, it is almost impossible to identify all metabolites measured by an untargeted analysis. Sumner et al. introduced four levels of metabolite identification confidence.[62] Metabolites with at least two orthogonal parameters e.g. accurate mass and RT identical to an authentic chemical standard measured under the same analytical conditions are considered to be confidently identified (level 1).[55,62] However, some metabolites and especially stereoisomers are very hard to distinguish even by comparison with an authentic chemical standard.[55] This is due to almost identical RT and m/z values particularly in untargeted metabolomics where the analytical methods are needed to be very robust and fast. Therefore, these methods are not perfectly optimized.[55] In these cases, alternative analytical methods like NMR or GC-EI-MS need to be applied for a positive identification. In comparison to confidentially identified compounds, putatively identified compounds or classes (level 2 and level 3, respectively) are not confirmed by an authentic standard. Instead only one or two parameters are utilized to putatively identify metabolites by comparing them to data from libraries or other analytical properties from different laboratories.[55,62] In case of LC-MS measurements, such properties are accurately measured m/z values, RT or fragmentation patterns. Comparing the isotopic pattern to in-silico generated

pattern provides additional confidence. Level 4 metabolites are unknown compounds, which can nevertheless be differentiated by the chromatogram or mass spectra. Moreover, the relative quantification of level 4 metabolites is feasible. However, a confident or putative identification is not possible. [55,62]

4.4.1 Identification workflow

The first step in the identification workflow is the calculation of the elemental formula of a certain accurate mass. This leads to a limited number of alternatives for the identification. However, for a reliable identification this number needs to be further reduced. Therefore, different filters and approaches are applied.

Firstly, the number of possible alternatives for the elemental composition can be limited by reducing the mass error. The more accurate the mass the lower is the number of possible elemental formulas.[63,64] Therefore, the detected masses should be internally recalibrated before calculating the elemental formulas. For this purpose, a sodium formate cluster injected via a six-port valve before every sample run can be utilized. The sodium formate cluster contains representative masses across the required mass range (see Fig.4.2 A). After measurement the mass spectrum of the sodium formate cluster is used to recalibrate the mass scale, improving the mass accuracy of the measured compounds (see Fig.4.2. B).

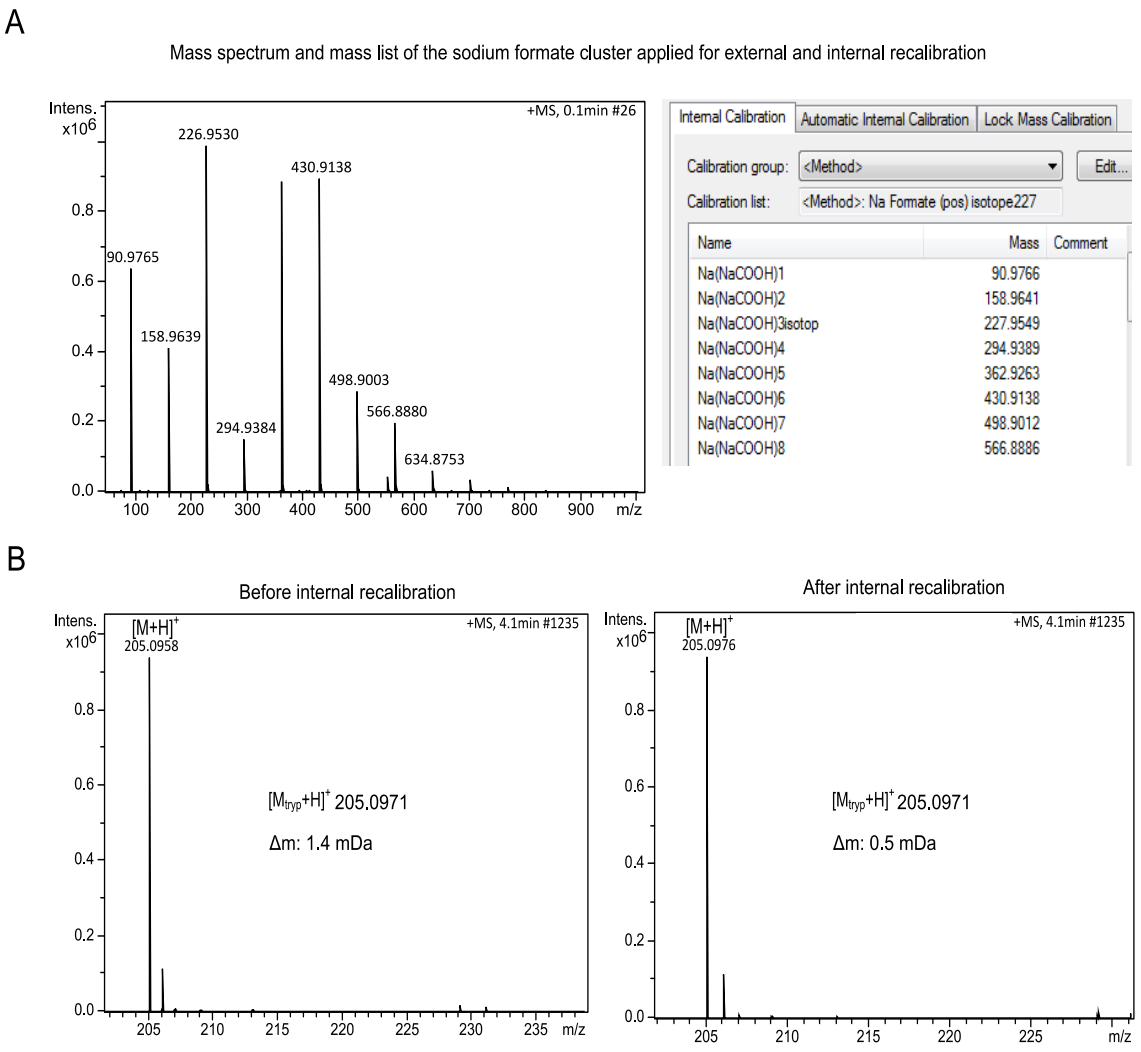


Figure 4.2: A sodium formate cluster is applied for external and internal mass calibration of HPLC-ESI-qTOF-MS measurements. (A) The mass spectrum of the sodium formate cluster, which is infused at the beginning of each sample analysis, is depicted on the left. The corresponding mass list of the recalibration method is shown on the right. (B) The improvement of the mass accuracy after recalibration is illustrated for tryptophan exemplarily.

Secondly, potential elements and adducts should be initially selected before calculating the elemental formula of an accurate mass in order to avoid false alternatives from including all elements from the periodic table.[64] Of course, certain knowledge about the source of the unknown is necessary.

Furthermore, tools like SmartFormula from Bruker Daltonics (Bremen, Germany) calculate the elemental composition from accurate masses by considering the rings-plus-double-bonds equivalent (RDBE), the nitrogen rule, the isotopic pattern as well as heuristic and chemical rules as introduced by Kind and Fiehn in 2007.[65] Consequently, these tools further constrain the composition of ele-

mental formulas fitting an accurate mass. The RDBE is calculated by the following formula:

$$RDBE = N_C + N_{Si} - \frac{1}{2}(N_H + N_F + N_{Cl} + N_{Br} + N_I) + \frac{1}{2}(N_N + N_P) + 1$$

Here, N stands for the amount of atoms of the corresponding element.[65] The nitrogen rule states that an odd nominal molecular mass also implies an odd number of nitrogen atoms.[65]

In 2006, Kind and Fiehn showed the advantage of using isotopic abundance pattern as an additional filter for calculating elemental formulas.[63] They compared a MS with 3 ppm mass accuracy applying the isotopic pattern filter (2% error) to a MS with less than 1 ppm mass accuracy omitting the isotopic pattern as a filter. They demonstrated that high mass accuracy alone is insufficient in order to reduce the number of possible elemental formulas especially for candidates with complex elemental compositions.[63] Even a hypothetical mass spectrometer with 0.1 ppm mass accuracy was outperformed by the additional isotopic pattern filter.[63] Therefore, the implementation of filters using the isotopic abundance pattern, removes most of the falsely assigned elemental formulas.[63] The isotopic pattern fit is evaluated by the calculation of the mSigma value in the SmartFormula tool from Bruker Daltonics. In the case of a perfect match the mSigma value is 0, whereas a value of 1000 means no match.

After limiting the number of possible molecular formulas, whereby preferably only one elemental formula is left, available data bases like HMDB[56], Metlin[57] or ChEBI (Chemical Entities of Biological Interest)[66] are searched for the chemical structure. However, the knowledge of the mass spectrometrist about the samples under investigation is an essential and mandatory requirement for the correct assignment of a chemical structure. As mentioned above, another aspect to positively identify an unknown signal is the corresponding fragmentation pattern. Holcapek et al. summarized the fragmentation behaviors of individual functional groups in 2010.[67] Hence, MS/MS experiments can increase the knowledge on the identification of the signal under investigation.[64]

Finally, the comparison to an authentic chemical standard is the last step in the identification workflow. The measured RT, isotopic pattern and fragmentation pattern from the standard should be compared to the unknown. The unambigu-

ous identification of an unknown requires the additional measurement with an independent analytical method such as NMR.

4.4.2 Concrete examples for LC-MS based identification

Here, the LC-MS based identification workflow and its significant role in metabolomics is exemplarily shown.

In the following, the identification procedures are partly adapted from Dettmer et al. 2013 and Zacharias et al. 2015, respectively.[30,68]

In Dettmer et al. 2013, metabolic footprinting of the cell culture supernatants of 20 human cancer cell lines and 4 primary cell cultures was conducted by means of high-resolution LC-QTOF-MS.[68] Statistical analysis resulted in 391 differential features after correction for multiple testing. The 49 most significant features and five additional features were then identified according to Fig. 4.3. Here, exemplarily the identification workflow of arginine, which was the most discriminating feature, is shown. Furthermore, MS/MS experiments were performed and the MS/MS spectra were compared to the METLIN library. The identification of the most significant features showed interesting insights into cancer metabolism. It revealed extracellular arginine and nicotinamide as major discriminants between normal and neoplastic hepatocytes. Further, the observed significant differences in the assimilation of di- and tripeptides appeared to underscore the increased bioenergetic and biosynthetic demands of many cancers.[68] These findings emphasize the importance of metabolite identification via high resolution LC-MS.

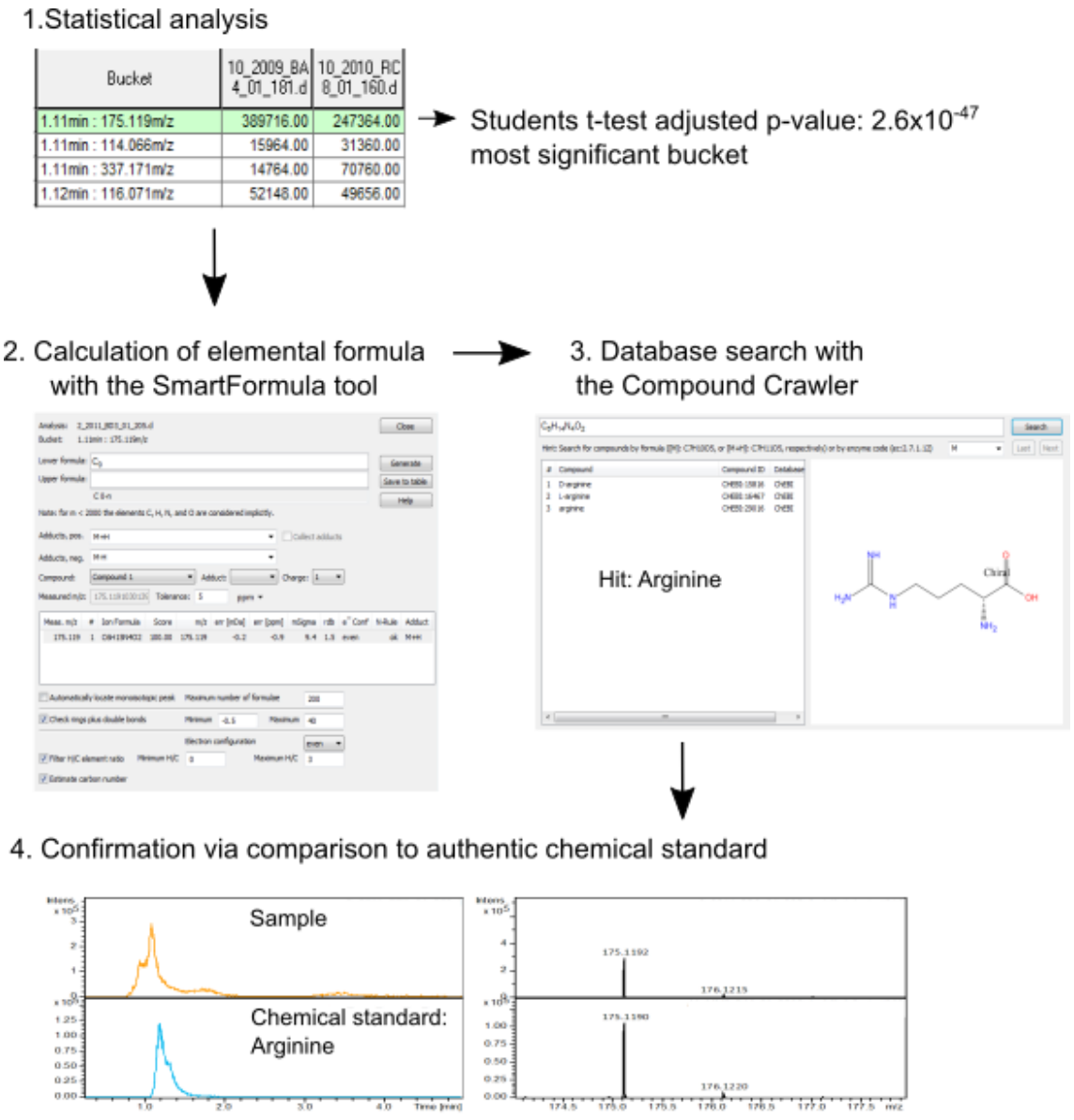


Figure 4.3: Exemplary workflow for the identification of unknown LC-ESI-qTOF-MS features that discriminate between different cell types. The Bruker tools SmartFormula and Compound Crawler were employed for the identification. The tentative identification of arginine was confirmed with an authentic chemical standard.

In the study by Zaccharias et al. (2015), LC-QTOF-MS contributed to the identification of the most significant metabolite for the prognostication of acute kidney injury (AKI) patients after cardiac surgery with cardiopulmonary bypass.[30] A well resolved NMR signal was found in plasma samples that distinguished patients, who developed AKI, from those that did not experience AKI. However, it was not possible to identify this signal by database searches or by 2D 1H TOCSY,1H−13C HSQC, and 1H−13C HMBC spectra, respectively. Therefore, 5 AKI and 5 non-AKI plasma samples were investigated by LC-QTOF-MS. After

correction for multiple testing, 11 significant features remained. Identification was performed as shown exemplarily in Fig. 4.3 and standards were purchased for the most promising hits. MS/MS experiments were performed on both the standards and the plasma specimens for additional verification. Among the most discriminating features, the propofol metabolites propofol glucuronide and 4-hydroxy-propofol-1-OH-D-glucuronide were positively identified. 1D 1H NMR reference spectra were acquired on these compounds and unambiguously verified the assignment of the NMR signal.[30]

5. Experimental section

5.1 Materials

Deionized water (PureLab Plus system, ELGA LabWater, Celle, Germany) creatinine-D3 (C/D/N Isotopes, Pointe-Claire, Canada), HPLC-grade acetonitrile (ACN) were purchased from VWR International (Vienna, Austria). Formic acid, uric acid, 1-methyluric acid, phenylacetic acid, phenyllactic acid, 3-(4-hydroxyphenyl)propionic acid, indoxyl-sulfate, creatine, creatinine, hippuric acid, guanidinoacetic acid, hypoxanthine, methylxanthine, kynurenic acid, DL-kynurenine, DL-tryptophan, xanthurenic acid, 3-hydroxyanthranillic acid, 3-indolacetic acid, 3,4-dihydroxy-L-phenylalanine, indole, DL-3-indolelactic acid, 3-indolepropionic acid, indole-3-carboxaldehyde, 3-indolepyruvic acid, tryptamine, tryptophol, nicotinic acid, melatonin, 3-hydroxy-DL-kynurenine, 5-hydroxyindole-3-acetic acid, 1,7-dimethylxanthine, serotonine hydrochloride, xanthine, nicotinamide, quinolinic acid, anthranilic acid, ammonium hydroxide and sodium hydroxide were purchased from Sigma-Aldrich/Fluka (Taufkirchen, Germany). Isopropanol (VWR), propofol (Toronto Research Chemicals, Toronto, Ontario, Canada).

5.2 Sample preparation

If not indicated otherwise, all urine specimens were diluted 1:4 with deionized water either directly in 1.5-mL glass vials with 0.2-mL micro-inserts (Machery-Nagel, Düren, Germany) or in micro-reaction tubes and subsequently transferred into the glass vials.

5.3 Instrumentation

5.3.1 HPLC-ESI-TOFMS

A Thermo Scientific Dionex Ultimate 3000 UHPLC system (Idstein, Germany) consisting of the HPG3400 RS pumping system, the TCC-3000 RS column ov-

en and the WPS3000TFC autosampler was coupled to a Maxis Impact QTOF-MS (Bruker Daltonics, Bremen, Germany) through an ESI source. A KinetexTM C18 column (100 mm x 2.1 mm id x 2.6 µm C18 1000, Phenomenex, Aschaffenburg, Germany) with a SecurityGuard ULTRA C18 cartridge (Phenomenex) as a pre-column was applied to separation in chapter 6 and to the fingerprinting measurements in chapter 8. The column oven temperature was set at 35°C. The chromatographic separation was accomplished by applying 0.1% (v/v) formic acid in water as mobile phase A and 0.1% (v/v) formic acid in ACN as mobile phase B at a flow rate of 0.3 mL/min. Elution was accomplished by the following ACN gradient unless stated otherwise in the text: 0-40% in 10 min, 40-100% in 2 min, 100% for 5 min, back to 0% in 0.1 min, 5 min equilibration. An injection volume of 5 µL was used unless otherwise noted in the text. The ESI source was operated in positive mode applying the following settings for the source and the mass spectrometer: drying gas: nitrogen with a temperature of 220 °C and a flow rate of 10 L/min; pressure of the nebulizer gas (nitrogen): 2.6 bar; end plate offset: 500 V; capillary voltage: 4500 V; mass range: 50-1000 m/z; acquisition rate: 5 spectra/s. Before the measurements, an external calibration of the mass spectrometer was implemented using sodium formate clusters (10 mM). Therefore, 12.5 mL of water, 12.5 mL of isopropanol, 50 µL of formic acid (conc.) and 250 µL of 1M NaOH were mixed. Additionally, each run was started with an injection of the sodium formate solution by means of a six-port valve for internal recalibration (see section 4.4.1).

5.3.2 HPLC-ESI-QqQMS

A 1200 SL HPLC (Agilent, Böblingen,Germany) was coupled to an Applied Biosystems 4000 QTRAP mass spectrometer (Sciex, Darmstadt, Germany) via a TurboV ESI source operating either in positive or negative ionization mode. For separation a Waters (Eschborn, Germany) Atlantis T3 column (2.1 x 150 mm i.d., 3 µm) at 25°C was used, applying 0.1% (v/v) formic acid in water as mobile phase A and 0.1% (v/v) formic acid in ACN as mobile phase B at a flow rate of 0.4 mL/min. The elution of the analytes detected in positive and negative mode was achieved by the following ACN gradient: 0% to 50% in 1 min, 50% for 5 min, back to 0% in 0.1 min, 4 min equilibration. This gradient is equivalent to the

gradient used by Zhu et al. (2011).[69] An injection volume of 10 µL was utilized. The MS parameters for each metabolite were optimized by direct infusion via a syringe pump (Harvard Apparatus, Holliston, MA, USA).

5.3.3 Miscellaneous

In the course of this doctoral research work the following other lab equipment was utilized: a heater with two heating blocks (Haep Labor Consult, Bovenden, Germany), a vortexer (lab dancer, IKA-Werke GmbH, Staufen, Germany) and a model himac CT15RE centrifuge from Hitachi (Düsseldorf, Germany).

5.4 Data analysis

5.4.1 Software

DataAnalysis version 4.1 (Bruker Daltonics) was utilized for manual examination and processing of the HPLC-TOF-MS chromatograms and mass spectra, compound extraction, internal recalibration of the mass spectra, and calculation of accurate masses, elemental formulas and mSigma values as well as for feature extraction from the chromatograms by the Find Molecular Feature (FMF) algorithm. ProfileAnalysis version 2.2 (64-bit) (Bruker Daltonics) was employed for feature alignment. Bucket tables were then exported and further analyzed with the R package version 3.1.1[70] and Excel (Microsoft Corporation, Redmond, WA). Bland Altman plots as well as basic statistics e.g. the calculation of relative standard deviations (RSDs), paired student's t test, spearman rank correlation coefficients (SCC) as well as Pearson correlation coefficients was performed with Excel. R was used to generate box plots and Venn diagrams as well as to perform principle component analysis (PCA), hierarchical cluster analysis by using Pvclust[71], Shapiro-Wilk test and to calculate false discovery rates (FDRs) according to Benjamini and Hochberg[51] using MULTTEST[72].

MassLynx (Waters, Eschborn, Germany) was used for manual reintegration and quantification of specific analytes in chapters 6 and 8. The Analyst (Applied Biosystems/MDS Sciex, Darmstadt, Germany) software was used for calculating

calibration curves, quantitative analysis and manual reintegration of analytes in chapter 7.

5.4.2 Calibration curves

For each standard the corresponding SIL-IS or an appropriate IS, which was structurally similar or was closely eluting, was used to normalize the peak area. Calibration curves were generated for each standard by plotting the ratio of the standard area and the IS area against the ratio of the absolute standard concentration and the IS concentration. This was followed by a linear regression analysis with a 1/x weighting to determine the coefficient of correlation R. LLOQ and ULOQ limited the linear range for each standard, whereby a S/N of 8/1 and a deviation smaller than 20% were utilized according to the FDA Guide for Bioanalytical Method Validation.[73]

5.4.3 Retention time corrector (RTcorrector)

The RTcorrector is an in-house written tool implemented by Alexandra Holler. The tool corrects for retention time shifts between different measurement batches. It was implemented for untargeted LC-MS data. In order to correct the RT of a certain bucket table as shown in Fig. 4.3, a bucket table is loaded as reference table into the tool. Then a bucket table will be corrected to fit the bucket table loaded as reference. Next, metabolites that serve as reference to compare the RTs are selected (Main Metabolites function). These should be metabolites which are abundant in all samples and not in detector saturation. They can be found in the given bucket tables by using a m/z and RT window. The Main Metabolite shift is calculated in order to predict the respective shift of all buckets. Further it can be selected if the shift should be handled as an absolute value (e.g. all peaks are shifted 2 min towards higher RTs) or as ratios (e.g. the retention times of the peaks are doubled). If there are more than one Main Metabolite, one can select if either a linear model to predict the shift depending on retention time or a constant shift using the median or mean of the Main Metabolite shifts are applied to correct the bucket tables. This tool was successfully used for the untargeted LC-MS bucket table generated in chapter 8.3.2. The Python interface is shown in Figure 5.1.

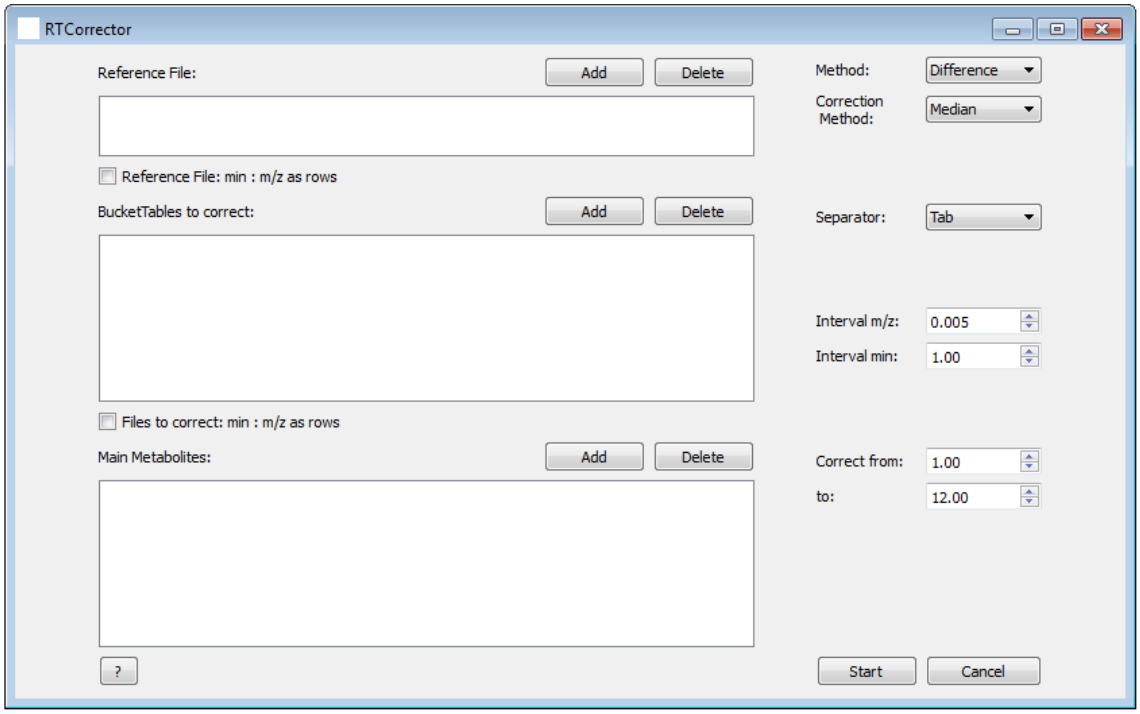
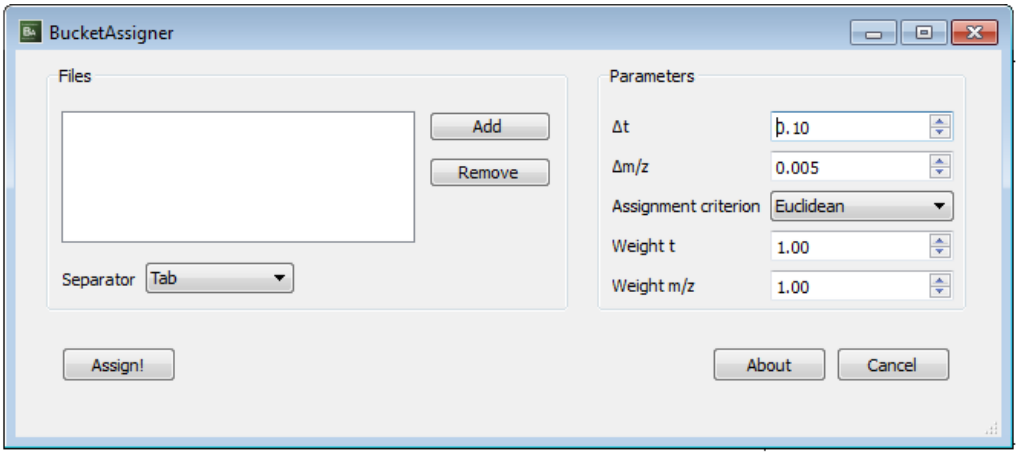


Figure 5.1: RT corrector Python interface. On the left side the reference file, the bucket table which should be corrected and the main metabolites are entered. On the right side, settings for the calculation are selected.

5.4.4 Bucket assigner

For comparison of samples from more than one batch, an in-house written bucket assigner (written by Sebastian Mehrl and Leonhard Heiziger) was used to align the bucket tables of the respective sample sets (see Fig. 5.2). This was applied to all experiments were the samples were measured in more than one batch. Features were individually aligned based on closeness within a manually specified m/z (e.g. 5 mDa) and retention time window (e.g. 0.1 min). Features in both bucket tables were marked if they were not assigned or multiple times assigned.[74]



	A	B	C	D	E	F	G	H	I
1	t_BT1	mz_BT1	t_BT2	mz_BT2	dist	n.selected	nc1	nc2	nc3
2	0.63	132.998	0.62	132.998	0.01	1	11956	14064	24600
3	0.63	86.992	0.63	86.993	0.001	1	8844	9464	15368
4	0.63	114.987	0.62	114.987	0.01	1	14444	18312	31336
5	0.63	128.019	0.63	128.019	0	1	39140	44884	87508
6	0.68	258.899	0.67	258.899	0.01	1	87708	92724	59708
7	0.68	190.912	0.67	190.912	0.01	1	16712	13684	11924
8	0.71	122.924	0.71	122.924	0	1	96324	81372	104796

Figure 5.2: BucketAssigner Python interface with resulting bucket table is shown.

5.5 Validation methods

5.5.1 Hierarchical cluster analysis

Unsupervised clustering is a widely used statistical tool in metabolomics. An unsupervised method is independent of class label information or example based grouping properties of the data. Therefore, it is an ideal method to identify novel patterns in data.[75] For clustering, a statistical distance function is required in order to partition a dataset into several subclasses or clusters.[76] Hierarchical and non-hierarchical are the two main clustering algorithms, whereby both divide the dataset into subgroups so that data with the same metabolic profile are clustered together.[76] In hierarchical clustering the grouped dataset is represented by a binary tree-like dendrogram.[75] The similarity of two observations is determined by a specific metric ,e.g. Euclidian, defined as geometric distance between two components, or Manhattan, the sum of the absolute dis-

tances of two vector values.[75] Iteratively the most similar observations are either grouped in one cluster or as neighbors. Ward's method clustering is a very efficient algorithm for hierarchical clustering. It tries to minimize the sum of squares of any two clusters that can be formed at each step.[75]

5.5.2 Bland-Altman plot

A Bland-Altman plot is a graphical approach to study the extent of agreement or disagreement between two different analytical techniques.[77,78] This method can also be applied to check the repeatability of one analytical technique or to compare the measurement by two operators.[77] Bland and Altman introduced this method in 1983 due to the lack of an appropriate analysis method to investigate the agreement of two measurements that cannot be fully captured by correlation coefficient and linear regression.[79]

In a Bland-Altman plot the y-axis represents the difference between the output variable x of the two paired measurements (x_1-x_2) and the corresponding average is plotted on the x-axis (x_1+x_2)/2).[2] For an easy visualization the mean difference ($\bar{d}$), and the limits of agreement, which are defined as the mean difference ±1.96 times the standard deviation ($\bar{d}$±1.96*σ), are depicted as horizontal lines in the plot.[77] These limits were determined under the assumption that the differences are normally distributed and therefore 95% of the differences are expected to lie within these limits.[77] As stated by Kaspar et al., six different types of Bland-Altman plots can be categorized.[80] In order to distinguish the plots graphically, the absolute mean differences, the scattering of the differences (randomly or proportionally) and the relative mean differences (ratio of the mean difference and the averaged mean of all couples, on a percentage basis) are used. According to Kaspar et al., a type A plot represents the mean difference of almost zero, a random scattering of all differences and a relative difference of ≤15%.[80]

6. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics

6.1 Introduction

Urine is a common matrix employed in large-scale human metabolomics studies, because it can be obtained non-invasively.[3,81] As products of cellular processes and responses, urinary metabolites are tightly related to phenotypes.[82] However, urinary output can differ greatly due to various processes involved in regulating the body's water and solute content. Factors affecting urinary metabolite abundance include among others water intake by drinking and eating, water loss due to respiration, perspiration and defecation, age, solute intake and the urine-concentrating ability of the kidneys.[3] As a result, urinary metabolite concentrations can vary widely. Therefore, untargeted analysis of urine is challenging, because inter-individual differences in metabolite abundance are not necessarily related to the phenotype under investigation and may actually mask "true" metabolic differences. Consequently, normalization strategies are necessary to reveal true biological variances.[4] Commonly, creatinine is utilized for this task, as it is produced at a constant rate and excreted - at least in healthy individuals - by glomerular filtration only without being reabsorbed or secreted in the renal tubule.[5] However, urinary creatinine concentration is influenced by various factors, including age, sex, muscle mass, diet, kidney function and water excretion.[5] Alternatively, the osmolality, i.e. the molar sum of all solutes in urine, can be applied to correct for urinary output.[4,6]

The procedure most commonly employed in LC-MS based metabolomics of urine uses a uniform pre-acquisition dilution of the urine specimens followed by post-acquisition normalization to either creatinine concentration or to the sum of all integrals of each sample. However, the selection of a proper pre-acquisition dilution factor is often difficult. On the one hand, analysis of highly concentrated specimens may result in column overloading, peak overlapping or ion suppression, which cannot be corrected by post-acquisition normalization. On the other

hand, uniform dilution may result in overly dilution of less concentrated urine specimens and, hence, failure to detect low-abundant analytes.

Lutz et al. suggested pre-analytical adjustment of urine specimens to an equal creatinine concentration by individual dilution, observing improved repeatability of retention times and area integrals in the profiling of dextromethorphan glucuronides.[83] Chen et al. proposed the adjustment of injection volumes to obtain uniform creatinine and, therefore, roughly equal overall metabolite amounts on the column for every sample. The acquired data were then normalized to the "MS total useful signal" (MSTUS).[84] However, urinary creatinine concentration can easily differ by a factor of 20 up to 100. Hence, injection volume range and column capacity can be limiting factors. Mattarucchi et al. diluted urine specimens according to the MSTUS measured in a preceding run.[85] More recently, Edmands et al. proposed pre-acquisition normalization to specific gravity, i.e. the weight of solids in urine, and evaluated different post-acquisition normalization methods to improve the identification of biomarkers, with post-acquisition normalization to specific gravity yielding the highest number of discriminant features.[86] Furthermore, Chetwynd et al. recently recommended pre-analytical adjustment to uniform osmolality.[87]

To date, the various strategies proposed to correct for variation in urinary output prior to MS-based metabolic fingerprinting have not been compared systemically with regard to their capacity of keeping the number of missing values in the data to a minimum and detecting a maximum of significantly discriminating features that allow accurate classification of samples. Here, we compared pre-acquisition dilution of spot urinary specimens to either a uniform creatinine concentration or osmolality without any further normalization of the acquired data to the common practice of uniform dilution followed by post-acquisition normalization to creatinine, osmolality, or sum of all integrals. Urine osmolality was chosen over specific gravity, as the former is less affected by the presence of large and heavy molecules such as glucose and proteins, thus reflecting more the number than the weight of solutes in urine. To test the effect of the various strategies on sample classification, we investigated spot urine specimens from both an apparently healthy cohort as well as two different cohorts suffering from chronic kidney disease (CKD).

Preparation of samples from batch 4 was carried out by Nadine Nürnberger from batch 5 by Lisa Ellmann. For the Random Forest classification as well as the heat map, the R script from Helena U. Zaccharias was applied.

This chapter was published in the Journal Analytical and Bioanalytical Chemsitry in 2016.[74] Text passages were taken verbatim from this publication.

6.2 Materials and Methods

6.2.1 Chemicals.

Deionized water (PureLab Plus system, ELGA LabWater, Celle, Germany) was used to prepare mobile phase A, a 1.5 mM stock solution of creatinine-D3 (C/D/N Isotopes, Pointe-Claire, Canada), and dilutions of the urine specimens. HPLC-grade acetonitrile (ACN) was purchased from VWR International (Vienna, Austria), formic acid from Sigma-Aldrich/Fluka (Taufkirchen, Germany).

6.2.2 Urine specimens

Human urine specimens were collected with informed consent from two different patient cohorts. The initial patient group comprised baseline urine specimens from 25 randomly selected patients from the clinical Trial to Reduce Cardiovascular Events with Aranesp Therapy (TREAT).[15] The TREAT had enrolled at 623 sites in 24 countries from the Americas, Australia and Europe a total of 4038 patients with type 2 diabetes, CKD with an estimated glomerular filtration rate (eGFR) of 20-60 mL/min/1.73 m^2 of body-surface area, and anemia (hemoglobin level, ≤11.0 g/dL). As a control group, we used 25 randomly chosen urine specimens from a group of 228 apparently healthy volunteers that had been selected at random through the Regensburg population registry within the framework of the German National Cohort (NC) (www.nationale-kohorte.de). To validate the obtained classification model and features identified to distinguish healthy from affected individuals, a second set of 25 urine specimens each from the TREAT and the NC cohort was analyzed. Finally, 50 urine specimens from patients enrolled into the German Chronic Kidney Disease (GCKD) study were

examined.[14] The group included patients with an eGFR below 60 mL/min/1.73 m^2, with and without type 2 diabetes as well as with and without increased albuminuria (> 300 mg/mL and < 30mg/ml, resp.).

6.2.3 Sample preparation

Urine samples were prepared in 1.5-mL vials with a 200-µL conical glass insert (Machery-Nagel, Düren, Germany). A first set of 25 urine specimens each from the TREAT and the NC cohort were diluted uniformly 1:4 with deionized water (batch 1) or adjusted before measurement to either a uniform osmolality (batch 2) or a uniform creatinine concentration (batch 3). For uniform 1:4 dilution in batch 1, 10 µL of an aqueous creatinine-D3 solution (1.5 mM) was added to 10 µL of urine followed by 20 µL of water resulting in a creatinine-D3 concentration of 375 µM. For batches 2 and 3, volumes of urine and water added to 17 µL of an aqueous creatinine-D3 solution (1.5 mM) were varied for each specimen to reach a final volume of 50 µL and either a uniform osmolality of 100 mOsmol (batch 2) or a uniform endogenous creatinine concentration of 1.5 mM (batch 3). The target values of 1.5 mM creatinine or 100 mOsmol were chosen, because these were, with a few exceptions (see below), the lowest measured values in the sample set. The resulting concentration of creatinine-D3 was 510 µM. Due to limited sample availability seven urine specimens (N=3 for batch 2, N=4 for batch 3) were individually diluted to a final volume of 25 µL each while maintaining the target creatinine or osmolality values. In these cases the creatinine-D3 standard (8.5 µL) was added beforehand to the insert and the solvent was evaporated.

In batch 3, two specimens (NC 6, NC 25) were not diluted due to low creatinine concentrations of 0.9 mM and 1.5 mM, respectively. For the dilution to a uniform osmolality, one specimen (NC 25) already possessed an osmolality of 100 mOsmol and, therefore, was measured unmodified. These three samples were prepared with a final volume of 25 µL with pre-evaporated internal standard (IS) (8.5 µL) as described above.

To assess dilution repeatability, 3 NC samples with low, medium and high creatinine concentrations (0.9 mM, 12.0 mM, 33.4 mM) were diluted 5 times each

and analyzed. Additionally, an independent urine quality control sample (QC) as well as a blank (H_2O) were injected repeatedly over the analysis of each batch (n=6 and n=5, respectively). The QC sample was prepared once (1:4 dilution).

In order to validate markers whose concentrations differed significantly between patients and controls, a second set of 25 urine specimens each from the TREAT and the NC cohort, respectively, was analyzed (batch 4). Samples were adjusted to a uniform target creatinine concentration of 2 mM because this was the lowest measured value in this sample set. Due to sample limitation, 4 specimens were diluted to a final volume of 25 µL and 3 specimens were not diluted at all due to low creatinine levels, but prepared in 25 µL as described above. In addition, 50 urine specimens from the GCKD study and the 25 urine specimens from the NC cohort that had already been part of batch 4, were examined (batch 5). The urine specimens were diluted to the lowest creatinine concentration, which for this batch was 1mM. Due to low creatinine levels, 2 GCKD specimens were not diluted at all. 2 NC samples were pre diluted 1/10 due to high creatinine concentrations prior to final dilution to 1 mM. An overview of all employed samples is given in Table 6.1.

Table 6.1. Overview of all the employed sample sets, their respective batch numbering and dilution methods. Reprinted from Vogl et al. 2016.[74]

Sample set	Batch	Dilution method
Normalization set 25 TREAT - 25 NC	1	1:4
	2	const. osmolality value
	3	const. c(creatinine)
Validation set 1 25 TREAT - 25 NC	4	const. c(creatinine)
Validation set 2 25 GCKD and 25* NC	5	const. c(creatinine)

* Identical with batch 4

6.2.4 Creatinine quantification

For the TREAT and NC urine specimens, creatinine was determined by LC-MS based isotope dilution analysis (IDA) upon addition of creatinine-D3 to each sample. The endogenous creatinine concentration was calculated by the ratio of the peak area of endogenous creatinine to the peak area of creatinine-D3 and multiplied by the concentration of creatinine-D3. Integration of creatinine and

creatinine-D3 peaks was performed using QuantLynx V4.1 (Waters Inc., Milford, MA). For batch 5, the creatinine concentrations determined by Synlab (SYNLAB Holding Deutschland GmbH, Augsburg, Germany) were used for dilution to a uniform creatinine concentration.

6.2.5 Osmolality

Measurements of urine osmolality were performed at the Department of Clinical Chemistry and Laboratory Medicine of the University Clinic Regensburg by determination of the freezing point depression using the Advanced Osmometer model 2020 (Advanced Instruments, INC, Norwood, MA).

6.2.6 LC-MS Analysis

A Thermo Scientific Dionex Ultimate 3000 UHPLC system (Idstein, Germany) was coupled to a Maxis Impact QTOF-MS (Bruker Daltonics, Bremen, Germany) through an ESI source. Analytes were separated on a Kinetex^{TM} C18 column (100 mm x 2.1 mm id x 2.6 µm C18 1000, Phenomenex, Aschaffenburg, Germany) at 35 °C applying 0.1% (v/v) formic acid in water as mobile phase A and 0.1% (v/v) formic acid in ACN as mobile phase B at a flow rate of 0.3 mL/min. Elution was accomplished by the following ACN gradient: 0-40% in 10 min, 40-100% in 2 min, 100% for 5 min, back to 0% in 0.1 min, 5 min equilibration. A shorter ACN gradient was implemented for creatinine quantification: 0-100% in 2 min, 100% for 5 min, back to 0% in 0.1 min, 5 min equilibration. An injection volume of 5 µL was used for metabolite fingerprints, whereas for creatinine quantification only 0.1 µL were injected.

Electrospray ionization was performed in positive mode using the following settings to operate the source and the mass spectrometer: drying gas: nitrogen with a temperature of 220 °C and a flow rate of 10 L/min; pressure of the nebulizer gas (nitrogen): 2.6 bar; end plate offset: 500 V; capillary voltage: 4500 V; mass range: 50-1000 m/z; acquisition rate: 5 spectra/s. Before the measurements, an external calibration of the mass spectrometer was implemented using sodium formate clusters (10 mM sodium formate in 50:50 v/v water/ isopropanol). Additionally, each run was started with an injection of the sodium formate

solution by means of a six-port valve for internal recalibration. Mass spectral resolution of R=27000 was obtained. Samples were measured in random order.

6.2.7 Data Analysis

Mass spectra were internally recalibrated based on the sodium formate clusters analyzed at the beginning of each run using Bruker Data Analysis V4.1 (Bruker Daltonics, Bremen, Germany). Features were then extracted from the chromatograms with the “find molecular feature” algorithm using the following parameters: signal-to-noise threshold: 2; minimum compound length: 20 (minimum number of spectra per compound); correlation coefficient: 0.7 (minimal time correlation needed for peak clusters to be combined to one charge state); smoothing width: 0 (number of spectra used by the chromatographic peak finder for smoothing). Profile Analysis V2.1 (64-bit) (Bruker Daltonics) was employed for feature alignment. All features, called buckets by the software, between 0.01-12 min and 50-1000 m/z were aligned. Each sample batch was aligned separately employing the advanced bucketing option with a time and mass window of 0.2 min and 5 mDa, respectively, except for batch 1, for which a mass window of 20 mDa was used. Buckets were only incorporated in the final bucket table if they were detected in at least 10% of the samples (bucket filter “value count of bucket” in Profile Analysis). Accuracy of bucketing was verified by recovery of 4 common urine metabolites (creatinine, uric acid, hippuric acid, tryptophan). Bucketing was considered to be accurate if all 4 metabolites were found in every sample. Bucket tables were exported from Profile Analysis as txt-files and further reduced by removing background signals and features found only in a limited number of samples. Peaks detected in the blank samples (N=5) with an intensity exceeding 20% of respective signals in the urine samples were excluded from the data. For this, the median of the respective peak in the blanks was compared to the median of the peak in the urine samples. Additionally, only features found in at least 80% of the samples (40 out of 50 samples for batch 1-4, respectively 60 out of 75 for batch 5), were used for statistical analysis. All missing values (zero values) were substituted with NA for “not available”. Statistical data analysis was performed using the R package version 3.1.1[70] and Excel (Microsoft Corporation, Redmond, WA). A two-sided Welch t-statistic was

calculated to obtain raw p-values with the R package MULTTEST[72]. These p-values were adjusted for multiple testing by controlling the false discovery rate at the 5% level according to Benjamini and Hochberg.[88] The significant features were displayed in a heat-map generated with an in house written method. Venn diagrams were created using the R package GPLOTS.[89] Hierarchical clustering of data was performed with the R package PVCLUST[71] employing Manhattan distances and the Ward clustering method. Before clustering, missing values were imputed by the minimum intensity found in the respective batch divided by ten and the data were scaled (mean centered). To limit the computation time, the number of bootstrap samples was set to 1000.

For comparison of samples from more than one batch, an in house written bucket assigner was used to align the bucket tables of the respective sample sets. This was applied to the two sample sets diluted to a uniform creatinine concentration (batch 3 and 4). Similarly, batch 5, consisting of samples from an independent patient cohort, was aligned to the previously generated bucket table of batches 3 and 4. Features were individually aligned based on closeness within a manually specified m/z (5 mDa) and retention time window (0.1 min). Features in both bucket tables were marked if they were not assigned or multiple times assigned. For the combined batches 3 and 4, 210 features were not assigned and excluded from further analysis, while the corresponding number of features excluded for the combined batches 3, 4 and 5 was 522, resulting in bucket tables consisting of 524 and 437 features, respectively. None of the features was assigned multiple times. For further analysis, the mean retention times of the assigned features and their mean *m/z* value were used. The aligned batches 3, 4, and 5 were combined, tested for the presence of normal distribution applying the Shapiro-Wilk normality test in R, and then Quantile normalized using the R package AFFY[90]. Classification employing a Random Forests (RF) classifier[91] provided in the R-package RANDOMFOREST[92] was only applied to assigned features. The internal parameters of the RF classifier were set to default. (number of trees used: 500, number of variables tried at each split: 20) Missing values were imputed as described above.

For batches 3 and 4, permutation tests, as outlined in Zacharias et al. [93], enabled an estimation of the significance of the obtained classification results by

comparing the classification accuracies obtained for the original non-permuted and the permutated data. The class labels of the training set were randomly permutated for each run and the trained RF classifier was used for subsequent classification of the test set. The permutation tests were performed 10 times in total. The mean values and standard deviations of the average overall classification accuracy and the corresponding error rates for both groups were calculated for the training and test set, respectively.

For the analysis of retention time shifts, retention times of creatinine, uric acid, tryptophan, uric acid, and hippuric acid were determined using the Quant Analysis software from Bruker.

6.3 Results and Discussion

6.3.1 Creatinine Quantification

Creatinine was quantified by isotope dilution analysis. Creatinine concentrations for batch 1 (constant 1:4 dilution) ranged from 0.9 to 33 mM. The analysis of the QC sample in this batch yielded a relative standard deviation (RSD) of 3.7% (N=6). The respective RSD values for the QC samples in batches 2 (dilution to constant osmolality) and 3 (dilution to constant creatinine) were 4.6% and 3.9%. The inter-batch precision of the QC sample analysis for all three batches was 4.3% (N=18). Moreover, the analysis of the dilution replicates within batches 1 – 3 yielded RSD values below 7.5%.

6.3.2 Basic characteristics of the first sample set

The osmolality in the first sample set (see Table 6.1, normalization set) of 25 NC and TREAT specimens each ranged from 107 to 1119 mOsmol. A significant correlation ($r = 0.657$, $p = 2.24*10^{-7}$) between creatinine and osmolality across all samples was observed (Figure 6.1 A). Examining the NC and the TREAT separately increased the correlation between creatinine and osmolality in the former ($r=0.793$, $p = 2.24*10^{-6}$), while a weaker, albeit still significant correlation was observed for the latter ($r=0.5$, $p = 0.011$) (Figure 6.1 B).

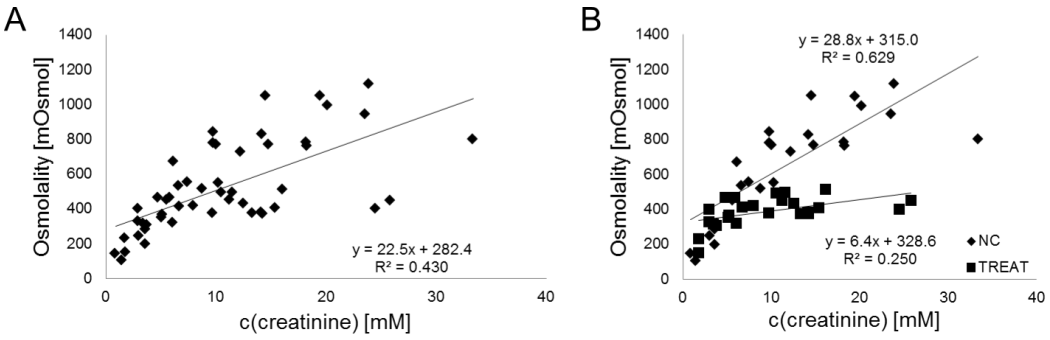


Figure 6.1: Linear correlation between osmolality and creatinine concentrations of the measured samples. (A) For TREAT and NC specimens combined. (B) Separately for TREAT and NC specimens. Reprinted from Vogl et al. 2016.[74]

Furthermore, osmolality was significantly lower (t-test, p= $4.89*10^{-5}$) in the TREAT (389 mOsmol ± 83 mOsmol) than in the NC specimens (660 mOsmol ± 293 mOsmol) (Figure 6.2 A). Urinary osmolality values below 400 mOsmol, as observed for the TREAT patients, are a diagnostic marker for tubular damage.[94]

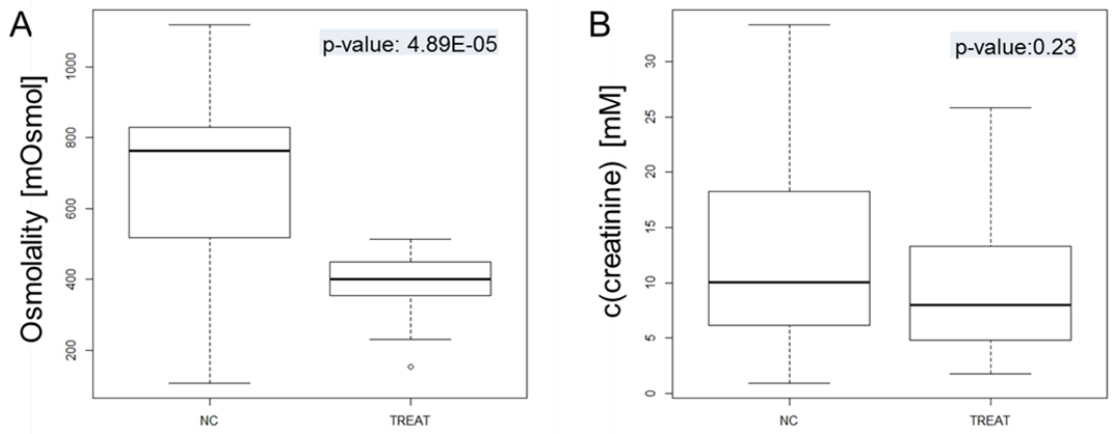


Figure 6.2: Boxplots representing the distribution of osmolality values (A) and creatinine concentrations (B) for TREAT and NC urine specimens. Reprinted from Vogl et al. 2016.[74]

Urinary creatinine concentrations between the two groups were comparable (p>0.05) (Figure 6.2 B). The urinary creatinine concentrations of the TREAT patients were not decreased compared to the NC cohort. This is likely due to the fact that in patients with CKD creatinine is not only excreted by glomerular filtration but also by renal tubular secretion.[95,96] Furthermore, patients with reduced glomerular filtration often also show a reduction in urinary output. After

uniform 1:4 dilution of the specimens, creatinine concentrations ranged from 0.2-8.3 mM, whereas dilution to a uniform osmolality of 100 mOsmol yielded a more narrow range of 0.6-5.4 mM creatinine (Figure 6.3).

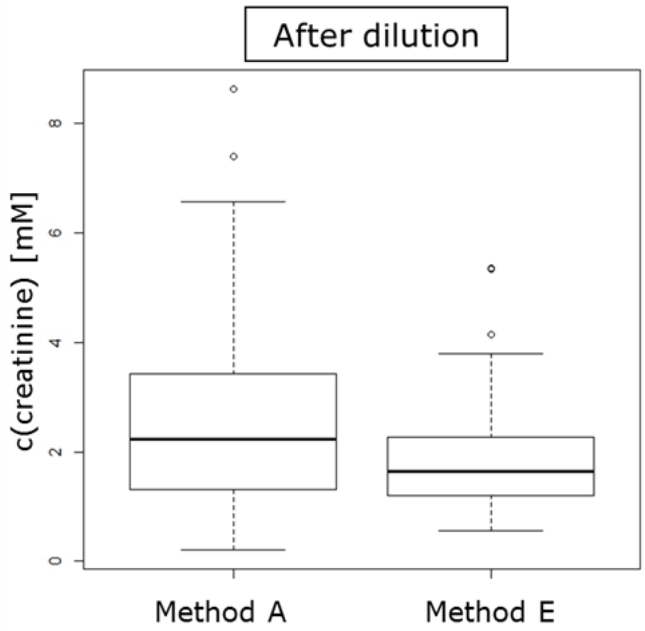


Figure 6.3: Boxplot illustrating the distribution of urinary creatinine concentrations for the TREAT and NC specimens after uniform 1:4 dilution (method A) or dilution to a uniform osmolality value (method E). Reprinted from Vogl et al. 2016.[74]

Lutz et al. observed in contrast to the use of crude urine smaller retention time shifts over time for urine specimens diluted to a uniform creatinine concentration.[83] However, one must keep in mind that the authors employed a sample enrichment step by means of an online trap column. In our study, retention did not depend on the dilution strategy but was stable for all tested dilution methods as illustrated in Figure 6.4, which shows the retention times of 4 common urine metabolites (creatinine (Panel A), uric acid (Panel B), hippuric acid (Panel C), and tryptophan (Panel D)) vs. the respective original creatinine concentrations of the urine specimens of batches 1, 2 and 3. Repeatability (RSD) of retention times across all 3 batches for the 4 metabolites ranged between 0.6% and 8.3%. The retention time shift of approximately 0.2 min observed for creatinine and uric acid resulted from elution of these metabolites close to the hold-up time.

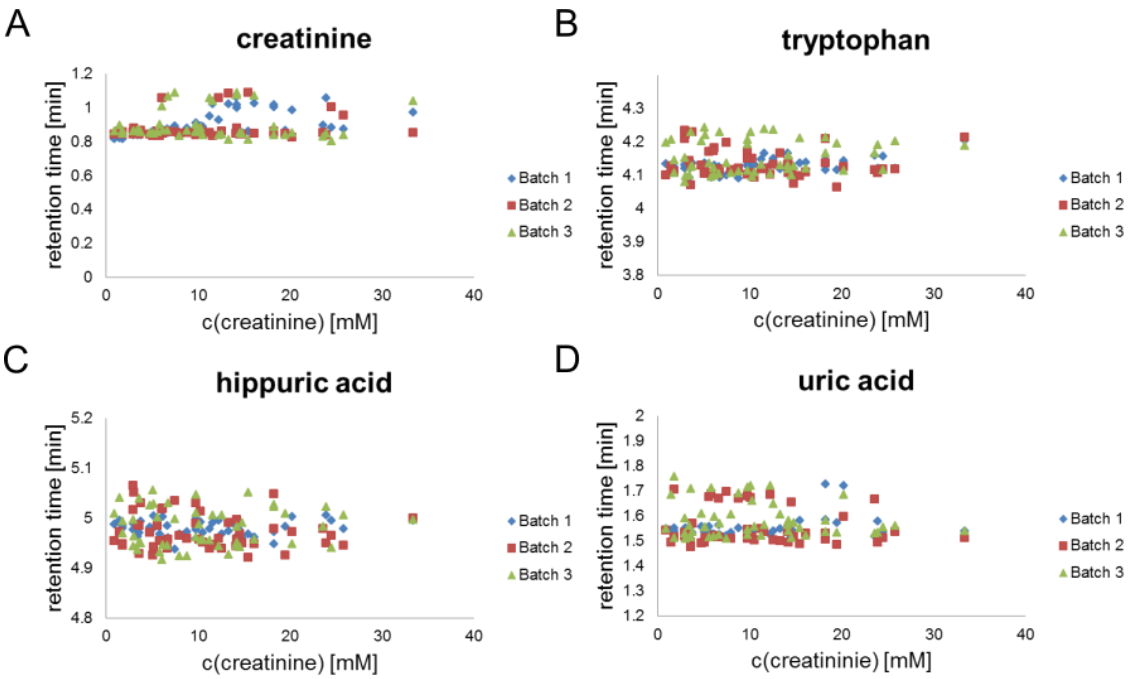


Figure 6.4: Retention time of 4 common urine metabolites creatinine (Panel A), uric acid (Panel B), hippuric acid (Panel C), tryptophan (Panel D) vs. the respective original creatinine concentration of the urine sample from batch 1, 2 and 3. Reprinted from Vogl et al. 2016.[74]

6.3.3 Effects of different dilution strategies on missing values

After feature extraction and alignment of the sample batches prepared by uniform dilution (batch 1), dilution to uniform osmolality (batch 2) or uniform creatinine concentration (batch 3), the resulting raw bucket tables were corrected for background signals and then examined for the number of missing values (MVs), which may originate from errors by the peak picking and alignment algorithm or may indicate that the feature was not detected in the sample above the limit of detection.

The evaluation of MV counts was performed with the background corrected bucket tables that incorporated all features (buckets) that were detected in at least 10% of the samples. This resulted in a high number of missing values with an average of 4833 MVs per sample in batch 1, 4818 MVs for batch 2, and 3374 MVs for batch 3. However, the number of extracted features was also the lowest for batch 3 due to the overall higher dilution of the samples. The inter-sample variability of the MV count was also decreased with a relative standard deviation (RSD) of 6 % in batch 3 compared to 7.6 % for batch 2 and 16.7 % for batch 1. The MV count per sample correlated significantly with the creatinine

concentration and the osmolality in the undiluted specimens. For the uniformly diluted samples (batch 1), the number of MVs decreased with higher creatinine concentrations ($r = -0.843$, $p = 1.69*10^{-14}$) or higher osmolality values ($r = -0.775$, $p = 3.78*10^{-11}$), respectively (see Figure 6.5 panel A and B). This indicates that samples with a low osmolality or creatinine concentration were overly diluted, resulting in failure to detect lower abundant features. In contrast, these correlations were less distinctive for the samples diluted to a constant osmolality (see Figure 6.5 panel C and D). Noticeably, the correlation of the MV count with the creatinine concentration in undiluted samples ($r= -0.634$, $p = 7.38*10^{-7}$) and the creatinine concentration in samples diluted to uniform osmolality ($r= -0.755$, $p = 2.38*10^{-10}$; Figure 6.5C) were still significant, while a correlation with the osmolality in the original specimens was no longer observed ($r = -0.179$, $p = 0.212$; Figure 6.5D). The still observed correlation with the creatinine concentration indicates that this dilution strategy does not completely normalize for urinary output, but some features seem to be lost due to the dilution. Only the samples diluted to a uniform creatinine concentration (batch 3) revealed neither a correlation of the MV count with the original creatinine concentration ($r = 0.216$, $p = 0.133$) nor with osmolality ($r = -0.074$, $p = 0.608$) (see Figure 6.5 panel D and E).

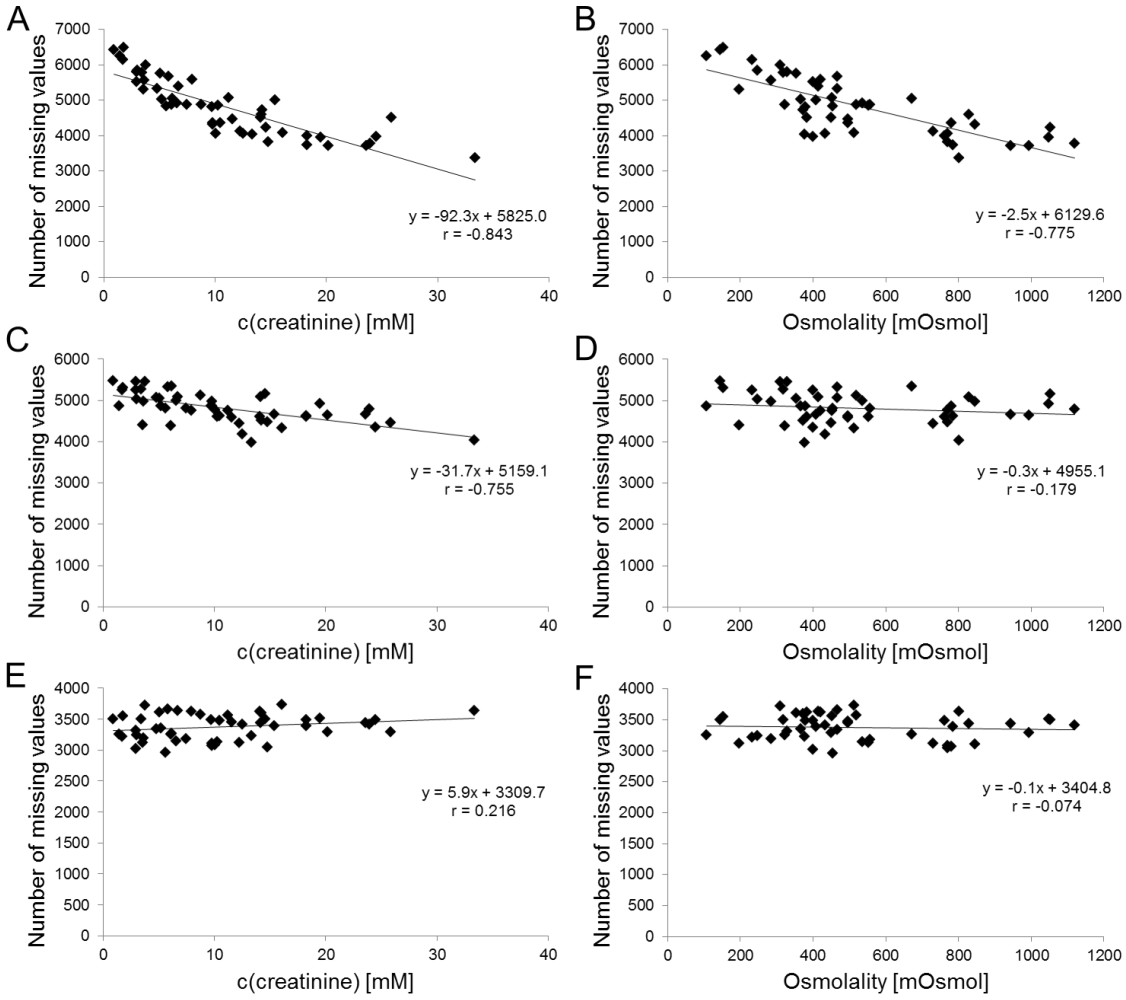


Figure 6.5: Impact of dilution strategies (A, B: uniform 1:4 dilution; C, D: dilution to a uniform osmolality value of 100 mOsmol; E, F: dilution to a uniform creatinine concentration of 1.5 mM) on correlation of missing values with creatinine concentration (A,C,E) and osmolality (B,D,F), respectively. Reprinted from Vogl et al. 2016.[74]

6.3.4 Evaluation of different dilution strategies and normalization methods

To reduce the MV count in the data matrices for further data evaluation, a stringent feature filter was applied including only features in the bucket table that were found in at least 40 out of 50 samples. The uniformly diluted samples (batch 1) were then normalized after LC-HRTOFMS analysis testing different post-acquisition normalization methods. In detail, no normalization (A) was compared to normalization to creatinine concentration (B), osmolality (C) and sum of all integrals (D). These post-acquisition methods were compared to the two pre-acquisition methods (batch 2 – E and batch 3 – F). Table 6.2 lists all tested methods (A-F) with their corresponding number of features, number of

readings (number of features × number of samples), number of MV relating to the number of readings, and number of features without MV. Furthermore, t-tests with correction for multiple testing (FDR < 0.05) according to Benjamini and Hochberg[88] were performed to test for features that were significantly different between NC and TREAT samples for each method. For methods A-D (batch 1), 726 features were extracted from all 50 samples, whereas for methods E (batch 2) and F (batch 3) 762 and 617 features, respectively, were extracted. The lower number of extracted features for method F (batch 3) is the result of the greater dilution of these samples.

Table 6.2. Figures of merit for the different batches and normalization methods tested for the normalization set (batch 1-3) measured with LC-HRTOFMS. No normalization (batch 1, method A) was compared to normalization to creatinine concentration (batch 1, method B), osmolality (batch 1, method C) and the sum of all integrals (batch 1, method D). These post-acquisition methods were compared to the two pre-acquisition methods (batch 2 – E and batch 3 – F). For all the tested methods (A-F) the corresponding number of features, number of readings (number of features × number of samples), number of MV relating to the number of readings, and number of features without MV are shown. Furthermore, t-tests with correction for multiple testing (FDR < 0.05) according to Benjamini and Hochberg[88] were performed to test for features that were significantly different between NC and TREAT samples for each method. Reprinted from Vogl et al. 2016.[74]

Batch	Method	Dilution method	Normalization after measurement	Number of features[a]	Number of readings[b]	Number of MV[f] (percentage)	Number of features without MV[f] (percentage)	Number (percentage) of significant features[c]
Batch 1	A	1/4	no normalization	726	36300	3290 (9.1)	117 (16.1)	194 (26.7)
	B		creatinine concentration	726	36300	3290 (9.1)	117 (16.1)	192 (26.4)
	C		osmolality	726	36300	3290 (9.1)	117 (16.1)	237 (32.6)
	D		sum of all integrals	726	36300	3290 (9.1)	117 (16.1)	249 (34.3)
Batch 2	E	uniform osmolality[d]	no normalization	762	38100	3227 (8.5)	126 (16.5)	211 (27.7)
Batch 3	F	uniform c(creatinine)[e]	no normalization	617	30850	2614 (8.5)	110 (17.8)	208 (33.7)

a Only features found in 40 out of 50 samples with additional blank subtraction.
b Represents the matrix dimension.
c FDR < 0.05 according to Benjamini and Hochberg.
d Samples were diluted to a uniform osmolality of 100 mOsmol.
e Samples were diluted to a uniform creatinine concentration of 1.5 mM.
f missing values

The numbers of significant features for the post-acquisition methods (batch 1, method A-D) indicate that no normalization (A) and normalization to creatinine (B) yielded a highly similar number of significant features, namely 194 and 192,

respectively, whereas normalization to osmolality (batch 1, method C) yielded 237 significant features. The best performance was observed with normalization to the sum of all integrals with 249 significant features. The latter approach does not only correct for different urinary output, but it also accounts for injection variability. Our results are in accordance with the results of Warrack et al., who observed no improvement in the principal component analysis (PCA) group differentiation with normalization to creatinine in comparison to no normalization, whereas the normalization to osmolality resulted in an improvement.[4] Chen et al. described that the normalization to the sum of all integrals yielded the best intergroup clustering and the lowest intragroup bias in the PCA score plot amongst the tested post-acquisition methods with identical injection volume. Interestingly, normalization to creatinine concentration was the least effective method.[84]

Although the total number of extracted features was the smallest for method F (batch 3), the number of significantly different features (FDR < 0.05) that distinguished NC from TREAT samples was, in terms of percentage, only slightly lower (33.7%) than for method D (batch 1). Interestingly, for method E (batch 2), only 211 (27.7%) of the detected features were significantly different between TREAT and NC, resembling the results obtained without normalization (batch 1, method A) (194/26.2%). This might have been caused by an insufficient correction for urinary output, as indicated by the remaining correlation with the MVs (see Figure 6.5 panel C), thus resulting in insufficient consideration of varying degrees of ion suppression or detector saturation in MS analysis and of failure to detect features due to overly dilution.

We further tested how many significantly different features were shared by application of the different normalization approaches. Comparing the results of the post-acquisition methods (B-D) yielded 137 significant features that were shared by all three post-acquisition methods (Figure 6.6A). Comparing the pre-acquisition normalization methods with post-acquisition method D, 91 significant features were shared by all three methods (Figure 6.6B). Consequently, the Venn diagrams imply that the compared methods obtain similar sets of significant features. However, some of the detected significant features are character-

istic for the respective methods. Nonetheless, the overlap of significant features represents features that distinguish robustly the investigated phenotypes.

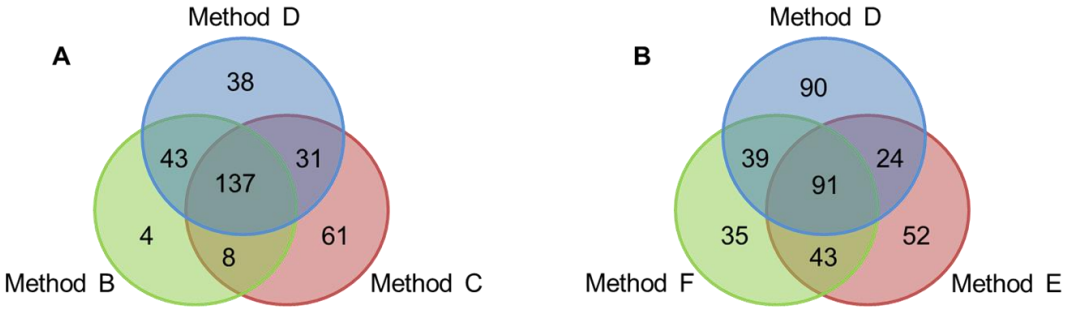


Figure 6.6. Venn diagrams of the numbers of significant features that discriminate between CKD patients and healthy controls for the different normalization methods employed. (A) Overlap of post-acquisition methods applied to LC-HRTOFMS fingerprints recorded for uniformly 1:4 diluted urine specimens: method B, normalization to creatinine concentration; method C, normalization to osmolality; method D, normalization to sum of all integrals. (B) Overlap of pre-acquisition methods E (dilution to a uniform osmolality value of 100 mOsmol) and F (dilution to a uniform creatinine concentration of 1.5 mM) with the post-acquisition method to the sum of all integrals (method D). Reprinted from Vogl et al. 2016.[74]

Additionally, the sum of all integrals was examined for batches 1, 2 and 3 using box plots (Figure 6.7). The sum of all integrals can be considered as a measure for the overall sample concentration as a composite of all metabolite concentrations. The spread of the sum of all integrals over all samples is the largest with uniform dilution, decreases with dilution to a uniformly osmolality value, and is the smallest for dilution to a uniform creatinine concentration. Furthermore, the median for batch 3 is lower than for the other batches, because the samples of batch 3 had been diluted more. Overall, this plot illustrates that the dilution to a uniform creatinine value (batch 3) balanced the highly varying sample concentrations more than the dilution to a uniform osmolality value or a uniform 1:4 dilution. This result is in agreement with the lower inter-sample variability of the MV count and the absent correlation of the MV count with creatinine concentration or osmolality in batch 3 as described above.

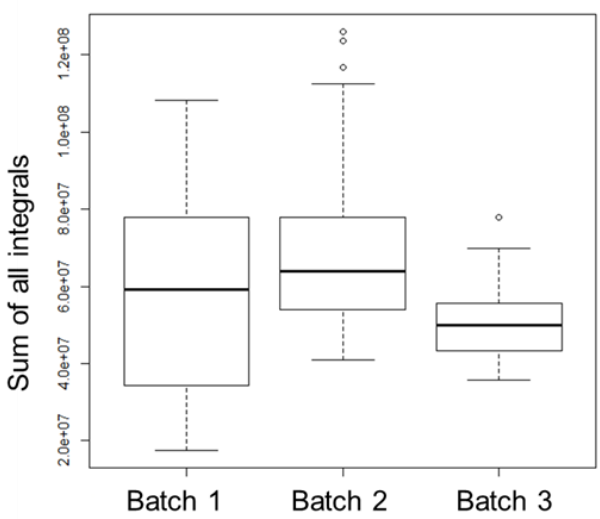
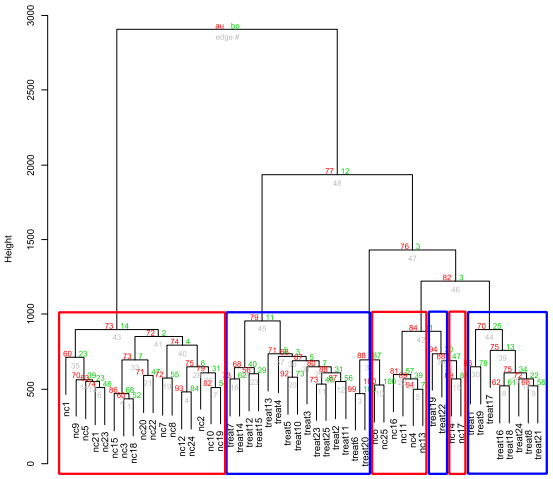
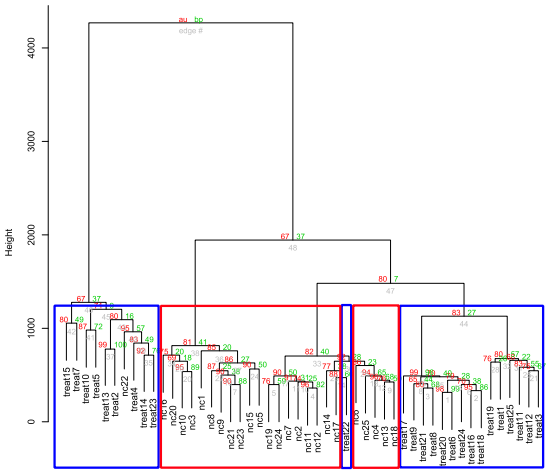


Figure 6.7: Boxplots representing the distribution of the sum of all integrals of the urine specimens analyzed as a function of the pre-acquisition method used. Batch 1: uniform 1:4 dilution, no normalization; batch 2: dilution to a uniform osmolality of 100 mOsmol; batch 3: dilution to a uniform creatinine concentration of 1.5 mM. Reprinted from Vogl et al. 2016.[74]

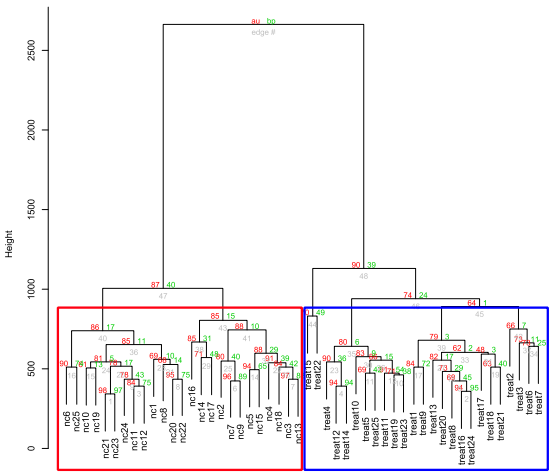
To test the degree to which the different normalization methods impact group separation, hierarchical cluster analysis was employed. To assess the uncertainty in hierarchical cluster analysis either multiscale bootstrapping resampling or ordinary bootstrapping resampling was used, allowing the calculation of approximately unbiased (AU) p-values and bootstrap probability (BP) values, respectively.[71] Both measures adopt values between 0 and 1 indicating how strong a given cluster is supported by the data. Clusters with an AU p-value >0.95 are existing at a significance level of alpha.[97] To keep computation time reasonably short, the bootstrap sample size was set to 1000. For hierarchical cluster analysis, the data matrices obtained with method D (batch 1), E (batch 2) and F (batch 3) were used (Figure 6.8). Here, only Method D was used as the only post-acquisition normalization method because it performed best amongst the post-acquisition methods. Two distinct clusters were anticipated for the NC and the TREAT. However, only dilution to a uniform creatinine concentration (method F) yielded two distinct clusters representing only NC and TREAT samples, respectively.



Method D



Method E



Method F

Figure 6.8. Hierarchical cluster analysis of fingerprinting data employing the Manhattan distance as a distance measure. Method D: uniform 1:4 dilution followed by a normalization to the sum of all integrals; method E: dilution to a uniform osmolality of 100 mOsmol without any additional normalization; method F: dilution to a uniform creatinine concentration of 1.5 mM without any additional normalization. The TREAT and NC samples are marked with a blue and red box, respectively. Reprinted from Vogl et al. 2016.[74]

Contrary to dilution to a fixed creatinine concentration (method F), pre-acquisition dilution to uniform osmolality (method E) and post-acquisition normalization to sum of all integrals (method D) failed to separate CKD patients clearly from healthy controls. This seems to indicate that only method F can account effectively for the analytical consequences resulting from highly varying metabolite concentrations, such as ion suppression, detector saturation and column overloading. Nevertheless, method D should be considered as an alternative to method F because this method allows correction for injection volume variability. We also tested a combination of methods D and F but failed to detect any significant improvement over dilution to a fixed creatinine concentration only (data not shown).

Recently, Chen et al. tested normalization methods with differential injection volumes calibrated by creatinine followed by either MSTUS normalization or normalization to sum of all integrals. Both methods were the most effective methods regarding inter- and intragroup clustering in PCA amongst all investigated methods.[84] Nevertheless, the adjustable injection volume range of the utilized autosampler as well as the column capacity could be limiting factors of this method due to the widely varying creatinine concentrations of urine specimens. For example, the creatinine concentrations in our small sample set of 50 specimens varied by a factor of 37.1 from 0.9 – 33.4 mM, In another study, urine specimens were normalized to specific gravity, measured by a refractometer, prior to LC/MS analysis.[86] However, depending on the predominant solutes, e.g. salts, glucose or urea, present in urine, the refractometer readings can be influenced.[6,98,99] Consequently, different diets of patients could influence the measurement of the specific gravity as well as certain clinical conditions, such as diabetes mellitus, nephrotic syndrome or saline diuresis.[100,101] In 2011, Mattarucchi et al. proposed a pre-acquisition normalization method applying MSTUS as a corrective factor for the dilution of the samples. Nonetheless, the MSTUS are affected by the precision and reproducibility of the peak picking process and missing values lower the efficacy of this normalization strategy.[85]

6.3.5 Classification of a second sample set

To test robustness of sample classification based on feature integrals obtained for urinary specimens diluted to a uniform creatinine concentration, a RF-based classifier was trained on the data obtained in batch 3 and then tested with an independent test set (validation set 1, batch 4) that contained 25 additional NC and TREAT specimens each that were different from those of batch 3 (see Table 6.1). The RF classification was performed one time. For the training set, the RF classifier obtained an overall classification accuracy of 96% with a corresponding error rate of 4% for both groups. For the test set, the RF classifier obtained an overall classification accuracy of 100% with a corresponding error rate of 0% for both groups. In order to exclude the possibility that the diagnostic accuracies for these biological groups had been obtained by chance, permutation tests were performed 10 times. To achieve this, the class labels were randomly perturbed. For the training set, an averaged overall classification accuracy of 45.0 (+/- 8.2)% with averaged error rates of 56.0 (+/- 9.8)% and 54.0 (+/- 7.6)%, respectively, for the NC and the TREAT group were obtained. The respective averaged overall classification accuracy for the test set was 44.8 (+/- 23.2)%, with respective averaged error rates of 56.0 (+/- 32.2)% and 54.4 (+/- 23.8)% for the NC and the TREAT group. The far lower classification accuracy of the permutated data indicates that the results for the non-permutated data were with high probability not obtained by chance.

Next, a comparison of NC versus TREAT samples was performed employing two-sided Welch t-statistics separately for training and test set, yielding 166 and 226 significantly discriminant features with Benjamini-Hochberg-adjusted p-values below 0.05, respectively. The number of discriminant features common to both training and test set was 130. Figure 6.9 displays a heat-map representation of these 130 features for both training and test data. The groups, each comprising 25 samples, were arranged from left to right: NC training set, NC test set, TREAT training set, and TREAT test set, while rows are ordered in correlation with disease status. Figure 6.9 shows the homogeneity of the up and down regulation of the significantly different features within each group as well as the divergence between the groups. This emphasizes the diagnostic rele-

vance of the detected metabolic features in the identification of patients with chronic kidney disease.

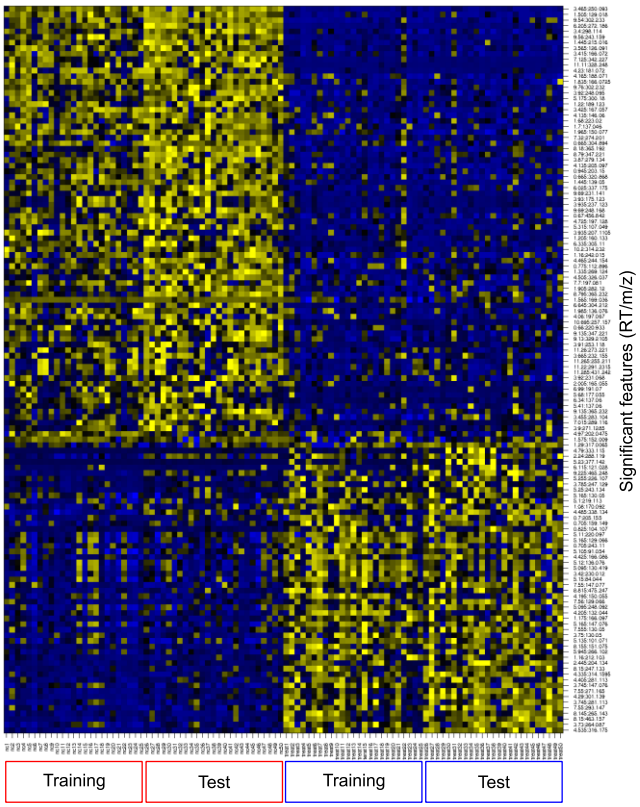


Figure 6.9: Heat-map representation of the 130 features that distinguish healthy controls from patients with chronic kidney disease. Rows are ordered in correlation with disease status. Urine specimens of the first and second set were diluted to uniform creatinine concentrations of 1.5 and 2.0 mM, respectively. The up- and down-regulation of features is color coded in yellow and blue, respectively. NC samples and TREAT samples are marked with a red and blue box, respectively. Reprinted from Vogl et al. 2016.[74]

6.3.6 Feature identification

We next attempted to assign metabolites to the N=130 discriminant features. It should be noted that each feature does not necessarily equal an individual metabolite. Rather, metabolites, for example due to in-source fragmentation, may be represented by more than one feature. Therefore, we sorted features according to retention time and grouped them. This resulted in 72 potential metabolites, which were subjected to further identification as recently described.[68] Sum formulas were calculated using the Smart Formula tool (Bruker Daltonics). The calculated mass error in mDa and the mSigma value as a measure for the isotopic fit are listed in Table 11.1 (Appendix). The average mass error was be-

low 1.5 mDa. Each feature was searched in the HMDB[102], METLIN[103] and ChEBI (Chemical Entities of Biological Interest) databases[104], respectively, and wherever possible a tentative identification was made (Table 11.1, Appendix). Reference compounds were analyzed to verify the identification of 3-methylhistidine, citric acid, uric acid, hypoxanthine, methyladenosine, methylxanthine, tryptophan, 1,7-dimethylxanthine, phenylalanine and hippuric acid. Among the identified metabolites, N-acetylcarnosine, citric acid, hypoxanthine, tryptophan, and hippuric acid, just to name a few, are metabolites that have been already associated with renal disease in the literature.[12,105]

6.3.7 Analysis of an independent patient cohort

To validate our results on an independent data set (validation set 2) and to exclude the possibility that discrimination of CKD and healthy specimens was driven by factors other than renal disease and diabetes, such as differences in diet and life style between the multinational TREAT cohort and the German NC, we analyzed a set of 50 baseline urine specimens from the German CKD study that had been selected to represent patients with reduced GFR with and without diabetes and/or increased albuminuria, respectively. These and 25 urine specimens from the NC cohort, which had already been analyzed before as part of batch 4, were diluted to a uniform creatinine concentration of 1 mM prior to LC-HRTOFMS analysis (batch 5). Hierarchical cluster analysis yielded again distinct clusters for GCKD and NC (Figure 6.10). However, this time two apparently healthy NC samples (NC 39, 33) clustered within the GCKD cluster and a single GCKD (GCKD 8) sample clustered within the NC cluster. Nevertheless, the results clearly demonstrated once more the ability of LC-HRTOFMS based metabolite fingerprinting to distinguish between urine specimens of healthy individuals and those with CKD.

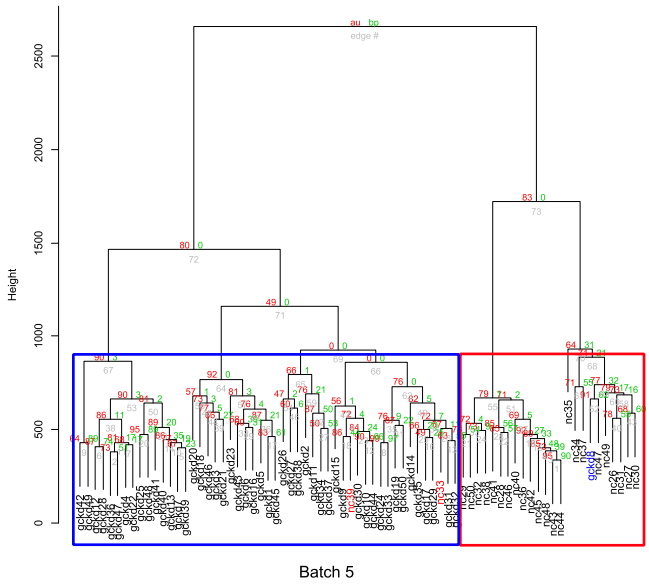


Figure 6.10: Hierarchical cluster analysis of fingerprinting data of an independent, more diverse sample cohort (GCKD, batch 5) employing the Manhattan distance as a distance measure. Samples were diluted to a fixed creatinine concentration of 1.0 mM before the measurement. GCKD and NC samples are marked with a blue and red box, respectively. Reprinted from Vogl et al. 2016.[74]

Next, the diagnostic accuracy for samples from patients with CKD of two independent patient cohorts (TREAT and GCKD) was investigated. To that end, the 25 TREAT and 25 NC samples from batch 3 (normalization set), the 25 TREAT and 25 NC samples from the independent test set (batch 4, validation set 1) and the samples from the independent patient cohort (50 GCKD and 25 NC samples, batch 5, validation set 2) were aligned to test different training and test sets with the RF classifier (see Table 11.2 A-H, Appendix). Prior to classification, normal distribution of batches 3, 4 and 5 was verified by the Shapiro-Wilk test (batch 3: p=0.123, batch 4: p =0.248, batch 5: p=0.146). In order to correct for the different fixed creatinine concentrations to which batches 3, 4 and 5 were diluted to (1.5 mM, 2.0 mM, and 1.0 mM, respectively), the aligned bucket table was Quantile normalized and log2 transformed prior to further data analysis. In Figure 6.11, a hierarchical cluster analysis of the aligned batches (batch 3, 4 and 5) is depicted.

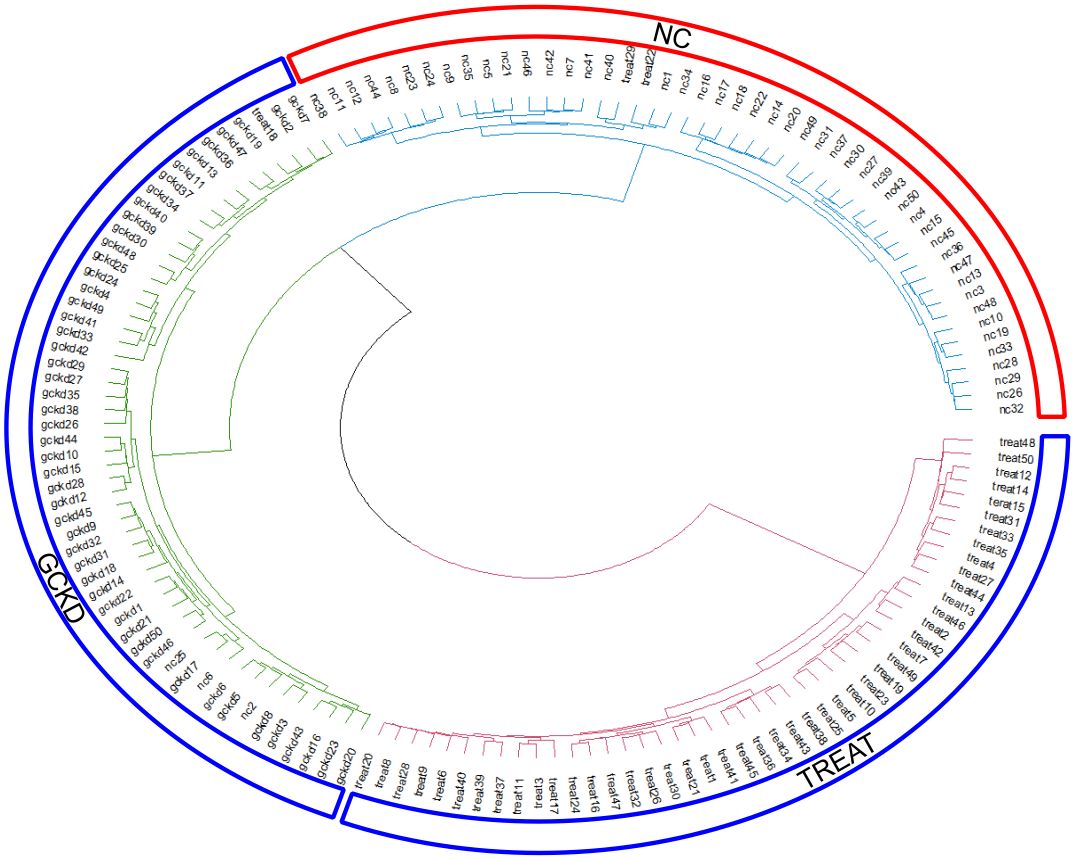


Figure 6.11. Hierarchical cluster analysis of the aligned data set of first and second sample set (TREAT and NC, batch 3, 4) and the independent patient cohort (GCKD and NC, batch 5) including the NC samples measured within batch 5. Urine samples were diluted to uniform creatinine concentrations of 1.5 mM, 2.0 mM ,and 1.0 mM, respectively. TREAT/GCKD (diseased patients) and NC (apparently healthy) specimens are marked with a blue and red box, respectively. Cluster branches of GCKD, TREAT and NC are colored green, red and blue, respectively. The R package dendextend and circlize was used for the visualization.[106] [107] Reprinted from Vogl et al. 2016.[74]

Three distinctive clusters were obtained corresponding to the three sample groups TREAT, GCKD and NC. While the TREAT cluster comprised TREAT samples only, the GCKD cluster contained in addition three NC samples (NC 2, 6, 25) from apparently healthy individuals and a TREAT sample (TREAT 18). Two additional TREAT samples (TREAT 22, 29) clustered within the NC cluster. The distinct clusters observed for the TREAT and the GCKD indicates that factors other than a reduced glomerular filtration rate may distinguish the two CKD cohorts.

Next, we used a RF classifier to classify samples into diseased and healthy. For this, the above bucket table consisting of batches 3, 4 and 5 was divided multi-

ple times into different training and test sets as shown in Table 11.2 (Appendix). For the RF classification, duplicate NC samples measured independently within batch 4 and batch 5 were excluded either from the training or test set. The RF classification was performed three times per test (A-H). Mean values are shown for the prediction accuracies. Results are summarized in Table 11.2. Training the RF classifier on the data obtained in batch 5 and employing batch 4 and 5 as test set (Test A, Table 11.2, Appendix) resulted in an overall classification accuracy of 92%, whereby the sensitivity (TREAT) and specificity (NC) amounted to 100% and 75%, respectively. This result revealed that training the RF classifier on batch 5, which consisted of GCKD and NC patients, classified TREAT samples with perfect accuracy (100%). This also holds true for tests B and D where the predication accuracies for TREAT samples were 100% for both tests, whereas the prediction accuracies for NC samples were 75% and 73%, respectively (Table 11.2). For test F, batch 5 was split (Batch 5/2). One half was used together with batches 3 and 4 for training and the other half for testing. Results revealed perfect predictions for both NC individuals and GCKD patients. In addition, the training set of test G used one half of batch 5 and batch 3 as training set whereas test H used batch 4 and one half of batch 5. Here again, perfect predictions of GCKD and TREAT samples were achieved. NC prediction accuracy was 82% for test G and 76% for test H, which led to an overall prediction accuracy of 92% (test G) and 90% (test H), respectively. However, for tests C and E, where training was based solely on NC and TREAT samples, inferior results were obtained for the prediction of GCKD samples. This clearly indicates that training data must include samples of this cohort to allow reliable classification of GCKD samples, as it was the case for tests F, G and H.

Finally, a comparison of the discriminant features of GCKD versus NC (batch 5) and TREAT versus NC (batch 3 and batch 4) with Benjamini-Hochberg-adjusted p-values below 0.05 yielded 49 features common to both patient cohorts (TREAT and GCKD). The corresponding potential metabolites are marked in Table 11.1. This result further indicates that the metabolic profiles of both, the TREAT and the GCKD share metabolites whose abundance differs significantly from that in the urine of apparently healthy individuals.

6.4 Conclusions

The present study demonstrates the advantages of pre-acquisition dilution of urine to a fixed creatinine concentration in the acquisition of metabolic fingerprints by LC-MS. This was demonstrated by the absence of significant correlations of MVs with either creatinine concentration or osmolality values. Moreover, this method yielded among all investigated approaches the highest percentage of significantly discriminant features between CKD and apparently healthy urine specimens. The second best method in terms of percentage of discriminant features was uniform dilution in combination with post-acquisition normalization to the sum of all integrals. The superiority of pre-acquisition dilution to a uniform creatinine concentration was also evident from hierarchical cluster analysis, which reliably assigned the urine specimens of controls and CKD patients to their respective clusters. In contrast, the other dilution and normalization methods tested yielded indistinct clusters. Hence, uniform dilution (e.g. 1:4) in combination with normalization to creatinine or osmolality, respectively, is not recommended. The analytical consequences of high variability in metabolite abundance between samples, such as ion suppression, detector saturation and column overloading, are not sufficiently accounted for by these normalization methods.

Furthermore, the proposed normalization method allows reliable sample classification as demonstrated for several different patient cohorts. Also, as demonstrated for batches 3 and 4, metabolites corresponding to differential features were for the most part identified. The results obtained underscore the superior reproducibility and applicability of pre-acquisition dilution to a uniform creatinine concentration and, thus, justify the associated time and labor.

7. Development of a quantitative LC-QqQMS method for renal disease-associated metabolites

7.1 Introduction

The problematic nature of current kidney disease markers was already discussed in chapters 4.2 and 6. Shortly, serum creatinine, the most applied clinical marker of renal dysfunction, rises only after 50% of kidney function is already lost. Moreover, the amount of tubular secretion of creatinine results in overestimation of renal function at lower glomerular filtration rates (GFR) and it reflects differences in muscle mass.[8] Ideally, filtration markers for GFR estimation should eliminate renal net reabsorption or secretion in the nephron.[34]

In the search of better alternatives to the so far established kidney disease markers, Takahira et al. in 2001 showed that C-mannosyltryptophan (CMT) is a potential novel marker of renal function.[108] Moreover, Yonemura et al. demonstrated that CMT is a more reliable diagnostic parameter than serum creatinine.[109] Furthermore, pseudouridine (PSU) was associated with renal decline and discussed as renal marker in several studies.[110-112] Additionally, Sekula et al. showed in 2016 that CMT and PSU are strongly and reproducibly associated with eGFR and CKD in population-based studies.[34] Therefore, a quantitative method was urgently needed for those markers in order to investigate the fractional excretion (FE) of CMT and PSU and to obtain an accurate picture of their renal clearance (see also chapter 4.2). In more recent studies, such as the Cooperative Health Research in the Region of Augsburg S4/F4 Study[113], the Framingham Heart Study[114], and the Atherosclerosis Risk in Communities Study[115] the relation between baseline metabolite profiles and the subsequent progression of CKD was investigated. In all of these studies different tryptophan metabolites were associated with CKD.[116] Therefore, the comprehensive study of tryptophan metabolites in urine appeared as highly promising in the search of kidney disease markers, which I addressed by the development of a quantitative method for an extended panel of tryptophan metabolites.

In 2011, Zhu et al. published a comprehensive quantitative HPLC-MS based method for the determination of tryptophan metabolites.[69] In the course of this thesis, the quantitative method from Zhu et al was combined with a newly implemented quantitative HPLC-MS method for CMT, PSU as well as metabolites which were identified as discriminating metabolites between patients suffering from CKD and healthy controls within the evaluation of normalization strategies (see chapter 6). Amongst these are creatine (CRT), creatinine (CRE), hippuric acid (HIP), xanthine (XT), hypoxanthine (HX), uric acid (UA) and guanidinoacetate (GUA), methylxanthine (MX), and dimethylxanthine (DMX), all of which have been described previously in the literature as being associated with renal disease.[32,117-120]

Preparation of standard solutions and IS solutions were carried out by Lisa Ellmann. Additionally, preparation of calibration curves and spike-in experiments as well as manual recalibration of both was performed by Lisa Ellmann and Katja Dettmer-Wilde.

7.2 Experimentals

7.2.1 Chemicals

Uric acid, 1-methyluric acid, 3-indoxyl-sulfate, 2,5-furandicarboxylic acid, creatine, creatinine, hippuric acid, guanidinoacetic acid, hypoxanthine, methylxanthine, kynurenic acid, DL-kynurenine, DL-tryptophan, xanthurenic acid, 3-hydroxyanthranilic acid, indole-3-acetic acid, 3,4-dihydroxy-L-phenylalanine, DL-indole-3-lactic acid, indole-3-propionic acid, indole-3-carboxaldehyde, indole-3-pyruvic acid, tryptamine, tryptophol, nicotinic acid, melatonin, 3-hydroxy-DL-kynurenine, 5-hydroxyindole-3-acetic acid, 1,7-dimethylxanthine, indole-3-acetic-2-2-d2* acid, uric acid 15N*, methylxanthine 13C,D3* were purchased from Sigma Aldrich (Taufkirchen, Germany). Xanthine, ß-pseudouridine, nicotinamide, quinolinic acid, methylguanine, pseudouric acid 13C-15N*, indoxyl-sulfate d4*, hippuric acid d5*, hypoxanthine $13C_2$,15N*, serotonine d4*, hydroxykynurenine, 1-methylinosine, tiglylcarnitine were purchased from Toronto Research Chemicals (Toronto, Canada). Anthranilic acid, serotonin hydrochlo-

ride were ordered from Fluka (Munich, Germany). Creatinine d3*, 5-hyroxyindole-3-acetic acid d5*, kynurenic-3,5,6,7,8-d5* acid, nicotinamide d4*, nicotinic acid d4*, tryptamine d4*, tryptophan d5* were purchased from CDN Isotopes (Quebec, Canada). C-mannosyltryptophan d4* was ordered at Santa Cruz (Heidelberg, Germany). C-mannosyltryptophan was kindly supplied by Shino Manabe (Synthetic Cellular Chemistry Laboratory Hirosawa, Japan). 3-hydroxyanthranilic acid d3*, anthranilic acid d4*, kynurenine d4*, quinolinic acid d3* were purchased from Buchem (Apeldoorn, the Netherlands).

7.2.2. Solutions

Stock solutions of each internal standard were prepared in water with 0.1% (v/v) formic acid and stored at -80°C. The stable isotope labeled standards were combined in two different internal standard mixtures (IS-mix 1 and IS-mix). IS mix 1 contained all labeled internal standards from the tryptophan pathway and IS mix 2 contained all other labeled internal standards. These two IS-mixes were combined 1:1 to a IS master mix. In the IS master mix concentrations were 5 µM for the IS-mix 1 standards except quinolinic acid-d3 and tryptophan-d5, which had a concentration of 50 µM and 100 µM, respectively. All IS from IS-mix 2 had a concentration of 50 µM.

Stock solutions of each unlabeled standard were prepared in either water, methanol, water/methanol, water/formic acid 0.1% or water/ 1 M NaOH mixtures (for details see Tab 11.3, Appendix). Different solvents were used according to the solubility of the analytes. Formic acid 0.1% and 1 M NaOH were added only in necessary amounts. Three different analyte mixes (AM) were produced. AM 1 contained all metabolites from the tryptophan pathway, AM 2 contained uremic toxins and AM 3 contained methylinosine, 1,7-dimethylxanthine, tiglylcarnitine and methylguanine (see Tab. 7.1, 7.2 and 7.3). Many of these metabolites were determined in chapter 6 as discriminating metabolites between patients with CKD and healthy volunteers. A master mix was generated from AM 1, AM 2 and AM 3. Therefore, 500 µL of AM 1 was mixed with 400 µL of AM 2 and 100 µL of AM 3. This master mix was serially diluted with 0.1 % formic acid covering 19 points over a concentration range of 0.04-11,317 nM in AM 1 except for TRP

where the range was 0.42-110,799 nM, 0.69-180,198 nM in AM 2 except for CRE, where the range was 2.06-539,839 nM and 0.26-67,505 nM in AM 3.

Table 7.1. Concentrations of metabolites included in analyte mix 1 (AM 1).

AM 1	c [µM]
Kynurenic acid	0.05
DL-Kynurenine	0.05
Nicotinamide	0.05
DL-Tryptophan	0.50
Xanthurenic acid	0.05
Quinolinic acid	0.05
Anthranilic acid	0.05
3-Hydroxyantranilic acid	0.05
3-Indoleacetic acid	0.05
Serotonin	0.05
DL-Indole-3-lactic acid	0.05
Indole-3-propionic acid	0.05
Indole-3-carboxaldehyde	0.05
Indole-3-pyruvic acid	0.05
Tryptamine	0.05
Tryptophol	0.05
Nicotinic acid	0.05
Melatonin	0.05
3-Hydroxy-DL-Kynurenine	0.05
5-Hydroxyindole-3-acetic acid	0.05

Table 7.2: Concentrations of metabolites included in analyte mix 2 (AM 2).

AM 2	c [µM]
Uric acid	1000
3-Indoxylsulfate	1000
Creatine	1000
Creatinine	3000
ß-Pseudouridine	1000
Hippuric acid	1000
Guanidinoacetic acid	1000
C-Mannosyltryptophan	1000
Xanthine	1000
Hypoxanthine	1000
Methylxanthine	1000

Table 7.3: Concentrations of metabolites included in analyte mix 3 (AM 3).

AM 3	c [µM]
1-Methylinosine	1500
1,7-Dimethylxanthine	1500
Tiglylcarnitine	1500
Methylguanine	1500

7.2.3 Sample preparation

For calibration, 10 µL of the internal standard solution were transferred to a 96 well plate with 0.5 mL wells (Agilent Technologies, Waldbronn, Germany) and then diluted to 100 µL with the respective aqueous calibration standard. The well plate was covered with a silicone closing mat for 96 well plates (Agilent Technologies).

For the spike-in experiment, endogenous concentrations were determined beforehand for five different urine samples. Three urine samples from the GCKD study as well as two samples from apparently healthy volunteers were used. Urine samples were pre-diluted 1:5 with water (PURELAB Plus water). The urine samples from the volunteers were measured in replicates (N=18), whereas for the GCKD samples single measurements were used. The mean of this concentration was used to calculate the concentrations of the metabolites for the spike-in mix which was then generated for urine. It contained all analytes in 10-fold higher concentration than the endogenous urine concentration levels.

For the actual spike in experiment, two urine samples from apparently healthy volunteers and two samples from patients with chronic kidney disease from the GCKD study were used (for detailed information see chapter 6.2.2). In order to generate the spiked matrix samples with 1-fold of the endogenous concentration for the spike in experiments, 10 µL of IS-master mix were added to 10 µL of pre-diluted urine and 10 µL of the urine spike-in mix. Next, samples were diluted to a final volume of 100 µL with water. The spike-in experiment was performed in a 96 well plate with 0.5 ml wells (Agilent) covered with a silicone closing mat for 96 well plates (Agilent).

As calibration check sample the master mix solution (AM 1, AM 2 and AM 3) was diluted 1:50 with 0.1% formic acid. Then 10 µL internal standard solution was added to 90µl of this diluted master mix dilution.

7.2.4 Instrumentation

An Agilent 1200 SL HPLC system (Böblingen, Germany) coupled to a 4000 QTrap mass spectrometer with a TurboV electrospray ion source (Sciex, Darmstadt, Germany) was utilized. MS parameters were optimized using direct infusion. The ion source was operated employing the following specifications: turbo ion spray voltage 5,500 or -4,500 V, curtain gas 10 psig, ion source temperature 500 °C, ion source gas 1 and 2 at 50 psig, collision gas medium. Measurements were performed in positive and negative ionization mode and multiple reaction monitoring applying the parameters summarized in Tab. 7.4. Spike in experiments were measured in both positive and negative ionization mode.

Chromatographic separation was performed using an Atlantis T3 (2.1×150 mm i.d, 3 µm, Waters, Eschborn, Germany) reversed-phase column equipped with a C18 security guard column (Phenomenex, Aschaffenburg, Germany). Elution was carried out by gradient elution with mobile phases A (0.1% formic acid in water, v/v) and mobile phase B (0.1% formic acid in acetonitrile, v/v) with the following gradient: from 0% to 40% B in 10 min, 40% to 95% B in 2 min (10-12 min) and from 12.0 to 12.1 min 0% B, which was hold for 4.9 min. The column temperature was 30°C and an injection volume of 10 µL was used. Samples were measured randomized. After 20 samples a urine reference sample from a healthy volunteer and the calibration check sample were measured.

Table 7.4: List of applied MRM parameters, DP (declustering potential), CE (collusion energy) and CXP (collusion cell exit potential) for each analyte measured. a) Metabolites newly added to the method of Zhu et al.[69] b) Metabolites included within the method of Zhu et al.[69]

Analyte	**Ab-brevi-ation**	**Q1 mass**	**Q3 mass**	**DP**	**CE**	**CXP**
Period1						
Creatinine[a]	CRE	114.0	86.0	26	17	14
Creatinine D3		117.0	89.0	26	17	14
Guanidinoacetate[a]	GUA	117.0	43.0	36	35	8
Nicotineamide[b,]	NAM	123.0	80.0	46	31	4
NAM D4		127.0	84.0	46	31	4
Nicotinic acid[b]	NA	124.0	80.0	56	31	12
NA D4		128.0	84.0	56	31	12
Creatine[a]	CRT	132.2	90.0	21	19	14
Hypoxanthine[a]	HX	137.1	110.3	66	33	6
Hypoxanthine-13C,15N		139.1	112.3	66	33	6
Methylguanine[a]	MG	166.0	149.0	66	27	8
Quinolinic acid[b]	QA	168.0	149.9	41	15	12
Quinolinic acid D3		171.0	153.0	41	15	12
Hydroxykynurenine[b]	HK	225.2	208.0	41	13	12
Hydroxikynurenine D3		228.2	211.0	41	13	12
Pseudouridine[a]	PSU	245.1	209.1	20	15	10
Pseudouridine 13C15N1		248.1	212.1	20	15	10
Period 2						
Xanthine[a]	XT	153.0	136.0	86	21	6
Methylxanthine[a]	MX	167.2	110.2	60	52	9
Methylxanthine-13C,D3		171.2	110.2	60	52	9
Serotonin[b]	SER	177.2	160.1	41	15	12
Serotonin D4		181.1	164.0	41	15	12
Kynurenine[b]	KYN	209.1	192.0	40	13	14
Kynurenine D4		213.1	196.0	40	13	14
1-Methylinosine[a]	MI	283.1	151.0	40	19	12
C-Mannosyltryptophan[a]	CMT	367.2	156.2	61	47	18
C-Mannosyltryptophan D4		371.2	160.2	61	47	18
Period 3						
Hydroxianthranilic acid[b]	HAA	154.2	80.0	46	36	14
Hydroxianthranilic acid D2		156.2	82.0	46	36	14
Tryptamine[b]	TRY	161.0	144.0	36	15	12
Tryptamine D4		165.0	148.0	36	15	12
Hippuric acid[a]	HIP	180.1	105.0	46	23	2
Hippuric acid D5		185.1	110.1	46	23	2
Dimethylxanthine[a]	DMX	181.1	124.1	41	29	20
Dimethylxanthine-D6		187.1	127.1	41	29	20
Kynurenic acid[b]	KA	190.2	144.0	51	30	10
Kynurenic acid D5		195.2	149.0	51	30	10

Hydroxyindolacetic acid[b]	HIAA	192.2	146.0	31	21	10
Hydroxyindolacetic acid D5		197.2	151.0	31	21	10
Tryptophan[b]	TRP	205.1	118.1	39	26	11
Tryptophan D5		210.2	122.0	39	26	11
Xanthurenic acid[b]	XA	206.2	160.0	56	27	10
Tiglylcarnitine[a]	TG	244.1	84.9	56	27	6
Period 4						
Anthranilic acid[b]	AA	138.2	120.0	36	15	8
Anthranilic acid D4		142.1	124.0	36	15	8
Indol3-carboxaldehyde[b]	INC	146.2	118.2	51	21	8
Tryptophol[b]	TPO	162.2	115.0	51	41	6
Indole-3-acetic acid[b]	IAA	176.2	130.0	56	25	8
Indole-3-acetic acid D2		178.2	132.0	56	25	8
Indole-3-propionic acid[b]	IPA	190.2	130.1	51	22	8
Indole-3-lactic acid[b]	ILA	206.2	130.0	36	39	8
Melatonin[b]	MEL	233.1	174.0	56	21	14
Melatonin D4		237.1	178.0	56	21	14
Negative ion mode						
3-Indoxylsulfate	IS	212.0	80.0	-60	-30	-11
3-Indoxylsulfate-D4		216.0	80.0	-60	-30	-11
Uric acid[a]	UA	167.1	124.0	-65	-20	-19
Uric acid-15N2		169.1	125.1	-65	-20	-19

7.3 Method validation

7.3.1 Quantification

For GUA, CRT, MG, XT, MI and TG, no stable isotope labeled standards were available. Therefore, stable isotope labeled standards eluting nearby were applied for calculating the respective area ratios. For GUA and CRT, CRE D3 was used, for MG hypoxanthine-13C,15N, for XT and MI methylxanthine-13C,d3, for TG 1,7-dimethylxanthine-d6, for XA KA-d5, for ILA, IPA, melatonin and tryptophol IAA-d2. This assignment was based on chemical and retention time similarity.

Quantification was performed using the respective stable isotope labeled internal standards. Data analysis was carried out using Analyst version 1.6.2 (Sciex).

7.3.2 Linear range, LOD and LOQ

The method was validated testing linearity, limit of detection and quantification, recovery and precision. The linear range described by the lower and upper limit of quantification (LLOQ and ULOQ, respectively) was assessed for each analyte. The LLOQ and the ULOQ are the lowest and the highest concentration for which quantification can be performed with a certain precision and trueness. According to the FDA Guide for Bioanalytical Method Validation[73], the LLOQ and the ULOQ are defined as the lowest and highest points of the calibration curve which can be quantified with accuracy between 80% and 120% and an imprecision smaller than 15%, except at the LLOQ where it should be smaller than 20%. Therefore, a series of standard solution accompanied by the respective internal standards was used. Each standard solution was measured at least one time. Nine standard solutions of the linear dilution series (calibration points 3, 5, 7, 9, 11, 13, 14, 17, 19) were measured three times in order to assess the precision over the calibration range. The limit of detection (LOD) is the lowest detected concentration which is detectable but not necessarily quantified. The LODs were determined from the linearity data as a signal to noise ratio (S/N) of ≥3.

7.3.3 Recovery and precision

Analyte recovery and precision for the analysis of urine were determined using a standard working solution of all candidates that was spiked into aliquots of urine. In order to determine the spike concentration for each metabolite, four urine samples were first measured unspiked with 10 µL injection volumes. The mean endogenous concentration for each metabolite was calculated. Then, a standard solution was added that contained the analytes at the initially determined endogenous urine concentration levels resulting in a 1-fold increase of the analyte concentrations. This was performed in triplicate for four different urine samples with an injection volume of 10 µL (spike-in experiment).

In order to determine the analyte concentration for the spiked urine the respective calibration curves were applied. Then, the endogenous concentration of the analyte, which was determined in the unspiked sample, was added to the theo-

retical spike concentration (desired value), which was determined beforehand. Finally the ratio between the measured spiked concentration (actual value) and this sum was calculated and expressed as recovery in percent (see following equation).

$$\left(\frac{c_{measured}(spiked\ analyte)}{(c_{measured}(unspiked\ analyte) + c_{calculated}(spiked\ analyte))}\right) x\ 100\%$$

7.4 Results and Discussion

7.4.1 Chromatographic separation and MS detection of analytes

The mobile phase and the gradient were adopted from Zhu et al..[69] Due to the large number of metabolites baseline separation was not feasible for all metabolites. It was a compromise between short analysis time and baseline separation. This method was intended to be used for high throughput analysis therefore a short analysis time was essential. Consequently, baseline separation of all included metabolites is almost impossible within the applied gradient of 15 min. In part, chemical structures and retention behaviors of the included metabolites are too similar. However, stable isotope labeled IS were used in order to minimize ion-suppression for co-eluting metabolites. The following metabolites co-elute (see Fig.7.1 and Fig. 7.2.): CRE, GA and CRT; NA, PSU, NAM and QA; MG, HX and HK; KYN and CMT; MI and MX; TG, TRP, DMX, TRY and XA; HIAA and HIP; MEL and INC; IAA and TPO. However, all these metabolites have different mass transitions. Therefore, the quantification is not affected by interferences.

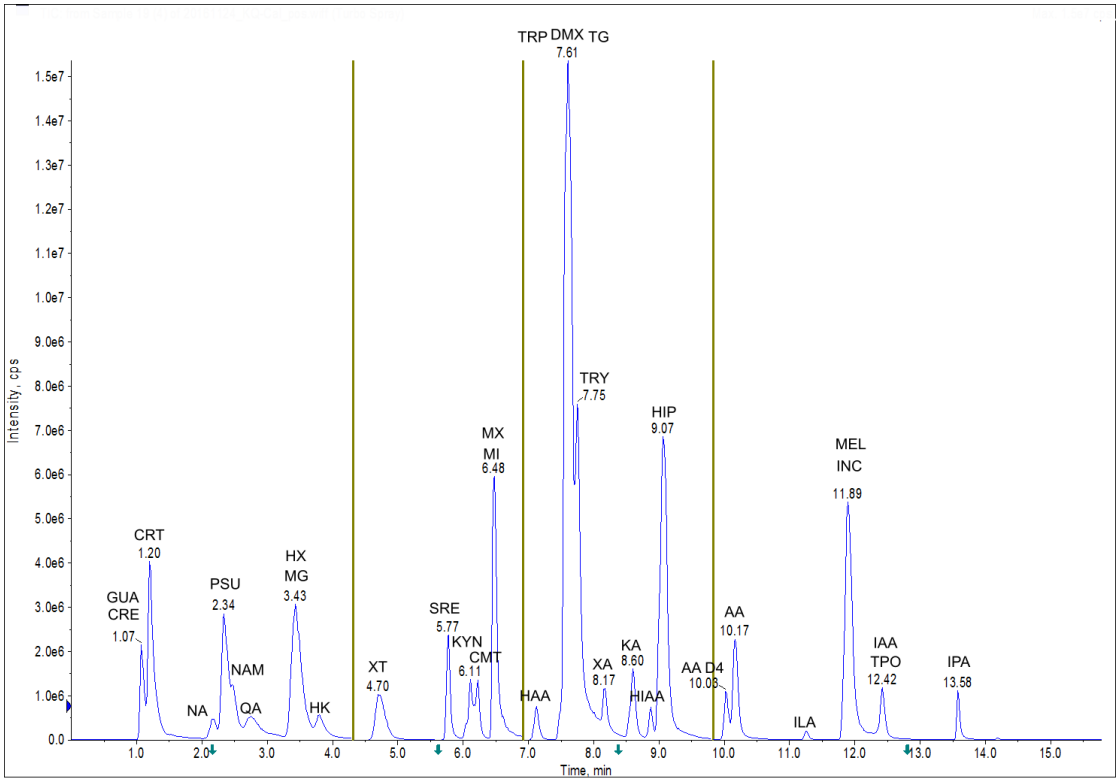


Figure 7.1.: TIC of a LC-QqQ-MS measurement of a representative standard mix in positive ionization mode. Bold lines represent scan periods.

The chromatographic run was divided into 4 scan periods in positive ionization mode (see Fig. 7.1 and Fig. 7.2). Thus, minimizing the number of transitions monitored in parallel. This resulted in 10-19 transitions per period.

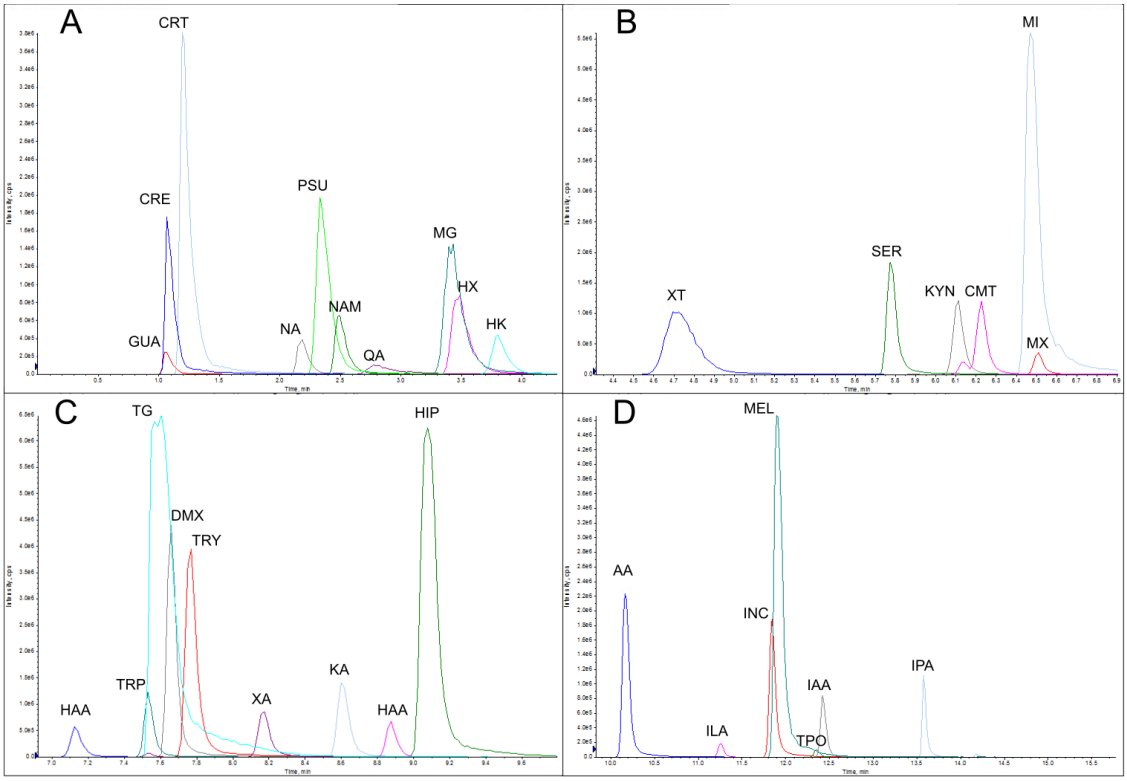


Figure 7.2: XIC of a LC-QqQ-MS measurement of a representative standard mix in positive ionization mode. A: scan period 1, B: scan period 2, C: scan period 3 and C: scan period 4.

Constancy of retention times for all analytes was checked for 100 urine samples from apparently healthy volunteers from the national cohort study (detailed description of this cohort see chapter 6.2.2). However, some metabolites were not detected in all samples. In Tab 7.5 the retention times are summarized for all metabolites measured in positive and negative mode. The RSD of retention times were below 1.3% for all analytes in both ionization modes. Small deviations in RT can be due to contamination buildup or general column deterioration. Therefore, RT may drift for samples measured over a long period.

For the calibration curves as well as for quantification, stable isotope labeled standards were used. As the stable isotope labeled standard will experience ion suppression in the same way as the respective endogenous analyte it will correct for ion suppression and accurate quantification should be achieved. Since several corresponding stable-isotope labeled standards were not available, data can only be used as semi-quantitative measures rather than quantitative measures because of the lack of an appropriate internal standard to correct for ion suppression.

UA and IS were measured in negative ionization mode because these metabolites only were ionizable in negative mode. Consequently, one scan period for both analytes was sufficient (see Fig 7.3).

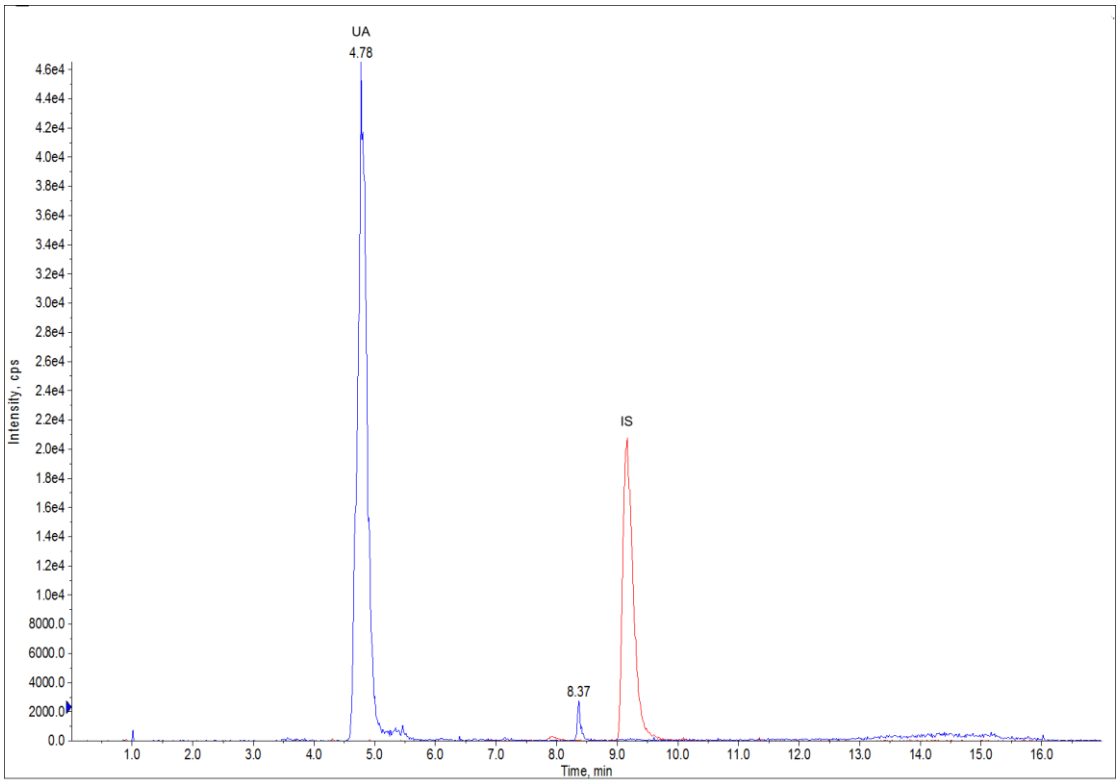


Figure 7.3: XIC of a LC-QqQ-MS measurement of a representative standard mix in negative ionization mode.

7.4.2 Calibration

The quality of the calibration curves depends on the instrument status. For instance, contamination of the MS instrument may result in higher background and/or lower signal intensities and a concomitant increase in LOD values. Furthermore, column aging may negatively affect peak shapes and hence the quality of the calibration curves. Therefore, calibration curves need to be measured regularly or better before every measurement. Moreover, calibration curves need to be adapted to the concentration range of the metabolites in the matrix under investigation. Calibration curves were determined for each metabolite, whereby a 1/x weighting was applied. In Tab. 7.5 calibration results are summarized.

Table 7.5: Calibration parameters.

Analyte	RT (min)	LOD (nM)	Linear range (nM)	Slope	Intercept	R^2	OOM
GUA	1.05	0.686	5.49-45000	0.44	0.0111	0.9998	4
CRE	1.06	2.06	65.9-270000	0.96	0.0074	0.9995	4
CRT	1.19	0.684	700-44800	7.29	0.6650	0.9987	2
NA	2.12	0.687	0.687-2810	0.91	0.0011	0.9997	4
PSU	2.33	0.684	1.37-44800	0.57	0.0003	0.9991	4
NAM	2.44	0.687	2.75-2810	3.01	0.0057	0.9988	3
QA	2.80	11.1	11.1-11300	0.93	0.0010	0.9998	3
MG	3.29	0.257	2.06-527	142	0.0222	0.9969	2
HX	3.45	2.75	87.8-22500	18	0.1460	0.9977	3
XT	4.69	0.687	1.38-90100	2.38	0.0129	0.9995	4
SER	5.70	0.043	0.343-1410	1.12	0.0002	0.9990	4
KYN	6.05	0.043	1.37-2810	1.21	0.0023	0.9986	3
CMT	6.17	0.687	2.75-11300	0.52	0.0001	0.9982	4
MI	6.45	0.257	1.03-132	63.2	0.0028	0.9950	2
MX	6.48	1.38	11.0-45000	0.26	0.0001	0.9997	3
HAA	7.09	0.043	2.75-5630	0.63	0.0148	0.9990	3
TRP	7.49	1.69	6.76-13800	1.62	0.0009	0.9977	4
TG	7.50	0.258	527-4220	3.37	0.8170	0.9877	1
DMX	7.64	0.515	1.03-2110	0.98	0.00003	0.9983	3
TRY	7.67	0.086	0.686-702	0.64	0.0003	0.9975	3
XA	8.20	0.344	22-2820	0.71	0.0194	0.9987	2
KA	8.65	0.086	2.75-704	1.19	0.0004	0.9989	2
HIAA	8.87	0.344	2.75-2820	4.43	0.0040	0.9990	3
HIP	9.07	2.75	11.0-1410	1.30	0.0010	0.9985	2
AA	10.16	0.345	0.345-353	0.93	0.0009	0.9995	3
ILA	11.27	1.37	1.37-5620	0.23	0.0002	0.9996	3
INC	11.81	0.043	0.172-176	3.57	0.0013	0.9977	3
MEL	11.90	0.043	0.0430-44.0	23.10	0.0002	0.9987	3
IAA	12.40	0.172	0.344-1410	1.02	0.0005	0.9987	4
TPO	12.41	1.37	2.75-5620	0.13	0.0005	0.9999	3
IPA	13.60	0.172	0.172-2810	0.94	0.0004	0.9997	4
Nega-tive mode							
IS	9.27	0.687	11.0-11300	0.65	0.0007	0.9991	3
UA	4.72	0.687	176-45000	1.57	0.3970	0.9992	2

The LLOQ ranged from 0.04 nM to 11.0 nM for most metabolites except for CRE, CRT, HX and XA with an LLOQ of 65.9 nM, 700 nM, 87.8 nM and 22.0 nM, respectively. Most metabolites had an ULOQ of 44.0 nM to 90.0 µM only for CRE the ULOQ ranged up to 270 µM. For the two metabolites measured in negative mode, IS and UA, the linear range reached from 11.0 nM to 11.3 µM

and from 176 nM to 45.0 µM, respectively. For all analytes the R^2 value of the regression analysis over the linear range was bigger than 0.99. Order of magnitudes (OOM) for all metabolites ranged between 1 and 4.

RSDs of triplicate measurements of the nine selected standard solutions which were within the respective linear range were below 20% for all metabolites.

The LOD of the metabolites measured in positive ionization mode ranged between 0.043 and 2.75 nM only QA had a LOD of 11.1 nM. The LOD of UA and IS was 0.687 nM for both.

In comparison to the values Zhu et al. 2011 reported for the linear range of the tryptophan metabolites, the values here are similar, with some metabolites having smaller LLOQ but also smaller ULOQ.[69] For example, Zhu et al determined for TRP a linear range of 50-200,000 nM, whereas a linear range of 6.76-13800 nM was found here. For other metabolites, e.g. KYN and XA, the linear range was narrower (1.0-7500 nM and 1.0-5000 nM, respectively). As mentioned above small variances can be explained by a different instrument status. However, one should keep in mind that the implemented method in positive mode includes more metabolites than the original method. Because more transitions per period were measured, MS dwell time had to be reduced for each transition to a greater extent than for the original method. This explains in part the more narrow linear ranges. Moreover, the final concentration of the IS differed for the two methods (see Tab. 7.6). This could also explain in part the differences in linear range.

Table 7.6. Comparison of internal standard (IS) concentration, linear range and order of magnitude (OOM) of the extended quantitative LC-QqQ-MS method and the original method implemented by Zhu et al. in 2011.For all other ISs not mentioned in the table, Zhu et al. used a yeast extract for which no exact IS concentrations were available.

	Extended method			**Original method (Zhu et al.)**		
Analyte	IS conc. (nM)	Linear range (nM)	OOM	IS conc. (nM)	Linear range (nM)	OOM
NA	500	0.687-2810	4	1000	5.0-10000	4
NAM	500	2.75-2810	3	1000	1.0-10000	4
SER	500	0.343-1410	4	200	1.0-4000	3
TRY	500	0.686-702	3	200	0.5-1000	4
HIAA	500	2.75-2820	3	1000	5.0-10000	4
AA	500	0.345-353	3	200	1.0-1000	3

7.4.3 Spiking experiment

Mean recoveries of the spike in level for a subset of the 36 included metabolites were determined in urine. Four urine samples (2GCKD and 2 apparently healthy volunteers) were measured in triplicates. CMT, PSU, XT, HX, GA, MX, XA and TRP were analyzed in positive mode, UA and IS in negative ionization mode. These are metabolites newly included within the original method of Zhu et al.. TRP was selected because of its key role in the tryptophan pathway. Moreover, XA is an interesting metabolite also from the tryptophan pathway. UA and IS were the only measured metabolites in negative ionization mode. The other metabolites newly added in positive ionization mode were either below LLOQ or above ULOQ.

Results are depicted in Fig. 7.4 and Table 7.7. In urine, the selected metabolites (positive mode) showed mean recoveries over all 4 urine samples between 86-121%, which indicates a very good recovery. RSD values ranged between 2% and 7%. This represents good repeatability of the spike-in experiments. For XA, only three (1 GCKD and 2 apparently healthy volunteers) of the four urine samples were in the linear range. A recovery of 89% with a RSD of 5.3% was obtained for the spike-in experiment of these samples.

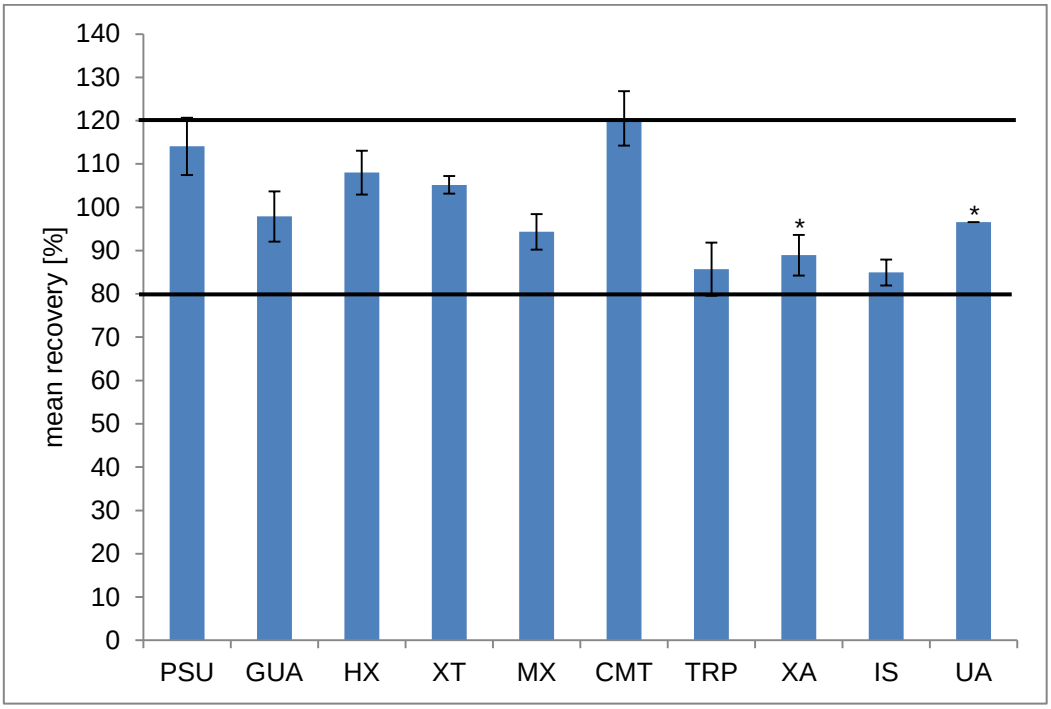


Figure 7.4: Mean recoveries ± SD over four samples of the spike-in experiment of the selected metabolites in urine. The black solid lines represent mean recoveries of 80% and 120%, respectively (triplicate measurements of each sample).* For XA only three (1 GCKD and 2 apparently healthy volunteers) and for UA only two (one GCKD sample) of the four urine samples were in the linear range and were analyzed.

Table 7.7: Mean recoveries and respective RSD values over four urine samples of the selected subset of metabolites from the spike-in experiments (N=3).

	PSU	**GUA**	**HX**	**XT**	**MX**	**CMT**	**TRP**	**IS**	**UA**
Mean recovery [%]	114.1	97.9	108.0	105.2	94.3	120.5	85.7	84.9	96.6
RSD [%]	5.8	5.9	4.7	2.0	4.4	5.2	7.2	3.5	0.0

For UA the concentrations of all the reference samples and one GCKD sample were above the ULOQ. However, for 2 GCKD samples the spike-in yielded a recovery of 96.6%. For IS the recovery over all urine samples (2 GCKD and 2 reference samples) was 85% (RSD 3.5%). Therefore, for both metabolites measured in negative ionization mode recoveries were very good.

In order to check the quality of the measurements, calibration check samples as well as reference samples were measured every 20 measurements. Accuracies of the calibration check measurements ranged between 95% and 111% except for CMT, which had a slightly increased accuracy (123%). RSDs of the reference samples ranged between 1.1 and 5.6% for the selected metabolites. These results imply the good quality of the method.

In order to validate the metabolites, which exceeded the ULOQ or were below the LLOQ in the spiked samples, additional lower and higher spike-in levels would be necessary. Moreover, according to the FDA more urine samples as well as more replicates would be necessary. However, obtained recoveries for the selected subset were very good. Moreover, for creatinine and creatine a comparison with independent quantification methods (NMR and enzymatic assay) yielded a very good agreement for the CRE and CRT quantification (see 7.4.4). Therefore, for these metabolites the implemented method is highly feasible for the analysis of urine samples.

7.4.4 Application

The developed method was applied to the targeted analysis of CRE, PSU and CMT in plasma and urine specimens collected by the Qatar Metabolomics Study on Diabetes (QMDiab) (n=111) and the GCKD study (n=328), respectively.[44] These measurements were part of a study published in Scientific Reports by Sekula et al. in 2017.[44] In this study, CMT and PSU were investigated in detail as potential kidney function markers and CRE was used for normalization of urine samples as well as for the calculation of the fractional excretion.

The quality of measurements was stringently controlled via several reference samples, calibration check samples and blank samples, which were measured every 20 samples. As reference samples, a urine and serum sample were used for the QMDiab study, whereas a urine and plasma sample were used for the GCKD study.[44]

Due to the study sizes, samples were measured in several batches using six sample plates for the GCKD and three plates for the QMDiab. Intra- and inter-

batch variability were checked for the reference samples. In urine samples from the GCKD and QMDiab, the overall RSD of the reference measurements were below 8.9% for CRE, PSU and CMT. In the plasma samples from the GCKD study, the overall RSD values for CRE and PSU were below 18.5%. CMT was not detected in the plasma reference sample. However, the pooled human male plasma sample used as reference (purchased form Sigma-Aldrich/Fluka Taufkirchen, Germany) showed very low concentrations for all metabolites in comparison to CKD samples. In the serum reference samples measured within the samples of the QMDiab study, the overall RSD was below 15.3%.[44]

Ranges of metabolite quantification and measures of quality (accuracy, coefficients of variations) per study population as well as batch-related intra-/inter-variability in reference measurements are provided in Tables 7.8 and 7.9.[44] Per study population and metabolite, accuracy of measurements was in the range of 80% to 120%.[44] All measurements were within the range of quantification. Coefficients of variation were < 20% for all measurements except for PSU in plasma of GCKD participants (20.8%).[44] These results demonstrate the excellent quality of the developed method in different matrices.

Table 7.8. Information of targeted measurements adapted from the paper of Sekula et al. 2017.[44] Targeted measurements were performed using the implemented quantitative method described in this chapter. Reprinted from Sekula et al. 2017.[44]

	Plasma				Urine			
	Mode	**Accu-racy**	**CV**	**ROQ***	**Mode**	**Accu-racy**	**CV**	**ROQ***
QMDiab								
Creatinine	+	98.5%	1.4	0.055-900	+	98.3%	2.2	1.375-22,500
C-mannosyltryptophan	+	94.4%	2.4	0.0274-450	+	93.8%	1.9	0.685-11,250
Pseudouridine	+	99.2%	6.9	0.055-900	+	114.1%	4.5	1.375-22,500
GCKD								
Creatinine	+	95.5%	5.4	0.0658 - 1,080	+	88.7%	3.6	3.295-27,000
C-mannosyltryptophan	+	95.8%	11.8	0.0440 - 360	+	97.3%	7.7	0.2745-9,000
Pseudouridine	+	86.9%	20.8	0.0109-179.2	+	89.6%	10.0	0.1365-1,120

* unit: µmol/L (transformed limits: for plasma: *2/1000, for urine: *50/1000)
CV: coefficient of variation; ROQ: range of quantification

Table 7.9: Intra- and inter-batch variability of targeted measurements of reference samples (adapted from the paper of Sekula et al 2017).[69] Targeted measurements were performed using the quantitative method described in this chapter. Reprinted from Sekula et al. 2017.[44]

Relative Standard Deviation (RSD)	Batch	Creatinine	Pseudo-uridine	C-mannosyl-tryptophan
QMDiab				
Urine reference	1	1.36	8.36	5.78
	2	1.86	7.06	2.63
	3	-*	-*	-*
	Overall†	2.41	7.03	6.02
Serum reference	1	1.66	3.11	6.79
	2	-*	-*	-*
	3	2.28	7.15	-
	Overall†	1.7	7.2	15.27
GCKD				
Urine reference, male	1	1.1	11,2	4,7
	2	-*	-*	-*
	3	3.9	9.0	2.3
	4	3.7	8.6	3.8
	5	2	5.8	2.8
	6	-*	-*	-*
	Overall†	3.4	8.3	7.8
Urine reference, female	1	2.4	2.3	1.7
	2	-*	-*	-*
	3	3.5	6.1	6.9
	4	6.4	8.5	8.3
	5	1.8	8.2	3.6
	6	-*	-*	-*
	Overall†	3.9	8.7	8.9
Plasma reference	1	7.4	14.0	-‡
	2	4.5	10.9	-‡
	3	3.4	12.1	-‡
	4	1.7	4.5	-‡
	5	-*	-*	-‡
	Overall†	7.0	18.5	-‡

Overall, measurements were done on 6 plates for GCKD and on 3 plates for QMDiab.

* Batch variability is not presented for plates with <3 reference measurements.

† RSD calculation based on all reference measurements.

‡ No reference values for C-mannosyltryptophan in plasma available as the measurements were at the LLOQ.

The developed quantitative method was also used in Wallmeier et al. (2017).[121] In this study, quantification of metabolites by NMR in the presence of protein were investigated. In order to validate the NMR measurements of CRE and CRT in plasma samples from patients with CKD, these samples were also measured by the LC-MS/MS method. Moreover, the CRE values of both the NMR and LC-MS/MS measurements were compared to an enzymatic CRE assay performed by a commercial clinical chemistry laboratory (Synlab GmbH, Augsburg, Germany). For the comparison of CRE values, 50 samples from the GCKD study were taken. The Bland−Altman plot revealed for the LC−MS/MS data slightly higher values with a mean difference of 10.5 µmol/L and 95% confidence intervals of 38.5 and−17.4 µmol/L, respectively. The mean difference between LC-MS/MS and NMR was 12.0 µmol/L. Moreover, the mean difference in CRE values obtained by LC-MS/MS and the enzymatic assay was 1.1 µmol/L with 95% limits of agreement ranging from 14.9 to.17.1 µmol/L. Additionally, the comparison of CRT values showed a good agreement with the NMR measurements demonstrated by a Bland-Altman plot with a mean difference of −1.37 µmol/L with 95% limits of agreement ranging from +11.1 to −13.8µmol/L. These results demonstrate an excellent agreement between the LC-MS/MS method and the enzymatic assay and a good agreement with the NMR measurements.[121] These findings demonstrated the quality of the developed quantitative LC-MS/MS method for CRE and CRT.

7.5 Conclusions

An LC-QqQMS method for the quantification of metabolites currently associated with renal disease was established. Therefore, the method established by Zhu et al. in 2011 was merged with a newly established method for metabolites currently associated with renal disease. The dynamic range of the merged method showed adequate sensitivity of all metabolites included for the analysis of urine samples.

Further, a spike-in experiment showed very good recoveries for a subset of metabolites. Moreover, several published applications proved the efficiency of the

implemented quantitative method in urine and plasma matrices for selected metabolites.

However, further experiments are necessary in order to validate all the metabolites included in this method. Nevertheless, the presented results demonstrate the high potential of this method regarding the quantification of various metabolites in urine from CKD patients. The combined HPLC-QqQMS-based targeted method enables the parallel quantitative investigation of tryptophan metabolites and other interesting metabolites associated with renal disease.[117] Additionally, the method could shed new insights into the underlying pathophysiological mechanisms of CKD.

8. Comparison of fingerprinting LC-HRTOF-MS and LC-QqQMS profiling data

8.1 Introduction

The triple quadrupole mass spectrometer in multiple reaction monitoring (MRM) mode is the well-established gold standard for quantification of small molecules and metabolites. For quantification via qTOF-MS in MS/MS mode using stable isotope labeled internal standards, comparability to classical targeted analysis has already been demonstrated by several publications, e.g. Park et al. 2016.[122] However, for this technique a preselection of metabolites is necessary, limited by the availability of stable isotope labeled standards. Yet, recent instrumental improvements raise the question whether quantification in untargeted full scan mode using a quadrupole time-of-flight-mass spectrometer (qTOF-MS) is reliable enough to replace targeted quantification in some cases.[13,123]

Untargeted screening measurements result in semi-quantitative data while targeted analysis yields absolute quantitative data. However, the reliability of quantitative data is dependent on the applied calibration method. For example, multi-point calibration using external calibration standards which are normalized by stable isotope labeled analogues is surely more reliable than a single-point calibration. Although, single-point calibration using a stable isotope labeled analogue certainly is more reliable than using non-similar IS.[123]

One tremendous advance of untargeted analysis is the parallel detection of thousands of metabolites with high resolution and mass accuracy. Instead, targeted analysis is limited to detect only several hundreds of metabolites in parallel.[124,125] For certain applications the limited number of metabolites that can be simultaneously quantified in MRM mode is insufficient. Hence, quantification in full scan untargeted mode is a valuable alternative.[126] The advantages of untargeted analysis in full scan mode make it highly interesting to investigate this technique for its quantitative performance.[13,127]

However, current knowledge about the performance of quantification in full scan mode, also called untargeted screening, in biological matrices is rather limited in comparison to the targeted quantification by MRM mode.

In 2008 Lu et al. conducted such a comparison between a triple quadrupole instrument in MRM mode and a TOF-MS in full scan mode for 20 standard compounds (10 in positive ion mode, 10 in negative ion mode) at 5 different concentration levels, including a cellular extract.[13] The published results were however limited to one standard compound.

Similarly, in 2011 Ramanathan conducted a comparison of a hybrid QTOF mass spectrometer with a triple quadrupole MS/MS system.[127] The comparison was performed with calibration curves for 25 metabolites including the drugs buspirone, prednisolone, prednisone, nefazodone and reserpine. The molecular weights of the compounds ranged from 230 to 747 Da.[127] After protein precipitation a drug free human plasma sample was spiked with these 25 metabolites in order to generate calibration curves ranging from 0.1 to 2000 ng/ml plasma. As internal standards, stable-isotope labeled (SIL) compounds and appropriate structural analogues were used. In addition to the full scan mode, the calibration curves also analyzed by LC-HRMS in information-dependent acquisition (IDA) MS/MS scan mode. Similar LOQs and linear ranges for the measurements of the samples with the combined TOF-MS and TOF-MS/MS (2 MS/MS scans). For the most part, the LLOQ was similar or better than the SRM mode.[127] Moreover, they yielded a similar range to the SRM mode while the % CV and %bias were acceptable for the HRMS analysis.[127] However, these results surely dependent on the selected instruments.

Recently, only few other papers further investigated the quantitative performance of targeted versus untargeted LCMS analysis.[128-133] Suhre et al compared three different metabolomics platforms, Biocrates Life Sciences AG (Austria), Chenomx Inc. (Canada), and Metabolon Inc. (USA) for diabetes research. They used 100 blood samples from the KORA study (Cooperative Health Research in the Region of Augsburg) consisting of 40 diabetes and 60 control samples for the investigation by all three platforms. Biocrates and Metabolon provided MS-based analysis, whereas Chenomx used a NMR-based analysis method. Moreover, Biocrates used a targeted MRM approach, while Metabolon

used an untargeted approach combining GC and LC chromatography. The final data set consisted of 482 distinct values of absolute (Biocrates and Chenomx) or relative (Metabolon) metabolite concentrations. Hereof, 68 metabolites could be quantified by two different platforms. These duplicate measurements yielded a median correlation coefficient (R) of 0.61, with the best results showing strong correlations up to R = 0.95. Suhre et al concluded that these results show the possibility for cross-platform replication.[128]

Yet et al. compared serum samples from twins using both targeted (Biocrates,) and untargeted (Metabolon) mass spectrometry platforms. They found 43 metabolites which were detected with both approaches. These 43 metabolites were compared and yielded a mean correlation coefficient of 0.44. The maximum correlation of r = 0.92 was found for octanoylcarnitine, the minimum correlation of r = 0 for 1-docosahexaenoylglycero-phosphocholine. Seven metabolites, including lipids and an amino acid, showed weak correlations between r = 0 and r = 0.2.[132]

It would constitute a tremendous advantage to use a qTOFMS platform not only for semi-quantitative fingerprinting but also for quantitative analysis. One would get the opportunity to have a platform which is able to accumulate quantitative information of thousands of metabolites in parallel within a single run. Given the current scientific knowledge, a more thorough comparison of the two techniques, absolute quantification via MRM mode and relative quantification via qTOF-MS in full scan mode, for urine samples is required in order to address this question. In this chapter, first the multi-point calibration of the quantitative method described in chapter 7 was compared to qTOF-MS measurements in untargeted mode, using the same calibration standards and chromatographic conditions. With this analysis the quantification performance using stable isotope labeled analogues for both methods (qTOF-MS and QqQ-MS) were checked. Analytical values like LLOQ, ULOQ, LOD, linear range and RSD were compared between the two techniques. Next, urine samples from the GCKD study and NC samples measured with the quantitative method described in chapter 7 were measured in semi-quantitative untargeted mode (qTOF-MS). Linear correlation, Spearman's rank correlation coefficients (SCCs) and Bland-Altmann plots were generated for selected metabolites in order to investigate

the quantitative performance of the untargeted platform without using ISs representing the routine approach of untargeted analysis.

As mentioned in chapter 7, targeted measurements and manual reintegration of the GCKD samples as well as the calibration measurements were done by Lisa Ellmann and Katja Dettmer-Wilde. Manual reintegration of the QmDiab samples published in Sekula et al 2017 was executed by Lisa Ellmann.[44] Data analysis of the corresponding calibration and samples was performed together with Katja Dettmer-Wilde. Statistical comparison of targeted and non-targeted measurements (published in Sekula et al 2017) was accomplished by Peggy Sekula.

8.2 Materials and Methods

8.2.1 Calibration curves and urine sample preparation

Calibration curves were used in order to compare LOD, LLOQ, ULOQ and linear range for both targeted and untargeted MS platform. Preparation of the dilution series, stock solutions of the internal standards and the chromatographic conditions were the same for both analysis methods. The preparation of the dilution series used for both calibrations (fingerprinting and profiling method) is described in chapter 7.2.2. Twice the volume of the dilution series was prepared and then equally split before the analysis.

200 urine specimens from the GCKD study as well as 100 NC samples were analyzed with both approaches in order to compare the performance. A detailed description of the sample cohorts can be found in chapter 6.2.2. For the untargeted analysis of the GCKD and NC specimens, samples were measured in 5 batches consisting of 60 samples each (40 GCKD and 20 NC samples). The GCKD and the NC samples were measured in a randomized fashion. For the targeted quantitative analysis, urine specimens (GCKD and NC) were pre-diluted 1:5 with water (PURELAB Plus water). Afterwards, 10 µL of internal standard mix were added to 10 µL of pre-diluted urine (for the detailed description see chapter 7.2.2). Finally, the sample was diluted to a final volume of 100 µL with water. A 1:4 dilution with water was used for the untargeted fingerprint-

ing analysis of the urine samples. Diluted samples were transferred in 1 mL HPLC vials with a 0.5 mL insert (Machery-Nagel, Düren, Germany.

8.2.2 Instrumentation

For both, targeted quantitative and untargeted semi-quantitative methods the calibration samples were measured with the Atlantis T3 (2.1×150 mm i.d., 3 µm, Waters, Eschborn, Germany) reversed-phase column equipped with a C18 security guard column (Phenomenex, Aschaffenburg, Germany). However, GCKD and NC urine samples were measured semi-quantitatively with the KinetexTM C18 column. Both calibration and urine samples were only measured in positive ionization mode. A detailed description of the applied instruments, parameters and gradients can be found in chapter 5.3.1, 5.3.2 and 7.2.4.

8.2.3 Data analysis

Data processing for the calibration of the untargeted measurements was performed using DataAnalysis Version 4.1 (Bruker) and QuanLynx V4.1 (Waters). First the data were recalibrated and processed in DataAnalysis, and then the data were exported in QuanLynx where the selected metabolites were reintegrated manually. Data processing of the untargeted measurements of the GCKD and NC samples was carried out as described in chapter 6.2.7. Data of the five batches were assigned using the in-house RTcorrector and bucket assigner (see chapter 5.4.3). Then only the data of selected metabolites were extracted from the bucket table for the comparison with the targeted data. In order to exclude possible batch effects within the untargeted measurements of the GCKD samples, hierarchical clustering was performed of all samples as described in chapter 6.2.7. Data analysis of the targeted measurements was accomplished using Analyst version 1.6.2 (Sciex). Statistical analysis of targeted and untargeted measurements was performed using the software described in chapter 5.4.1.

8.3 Results and Discussion

8.3.1 Comparison of calibration curves

In order to compare targeted and untargeted quantification performance, the calibration curves measured under the same chromatographic conditions (same column, same gradient, same column temperature) were compared. The calibration curves described in chapter 7 for the targeted quantification method were compared to untargeted measurements of the same calibration samples. Peaks were manually reintegrated for both methods. The manual peak reintegration applied to both methods is usually not used for untargeted measurements due to the huge amount of analytes measured in parallel. However, here only six metabolites (PSU, CRE, HIP, CMT, XA and TRP) were selected and manually reintegrated for the comparison. Firstly, the comparison was focused on these metabolites because PSU, CMT and HIP are interesting metabolites associated with renal disease. Secondly, XA and TRP are metabolites from the TRP pathway which are also related to renal disease (see chapter 7). Moreover, XA was interesting to compare because no appropriate IS was available for the quantification. Since serum CRE is the current clinical measure for renal disease, CRE was also included. For the targeted method the measurements and results generated in chapter 7.4.2 were used for the comparison with the fingerprinting method. The calibration results for both methods are summarized in Table 8.1 and 8.2.

Not surprisingly, the targeted quantification resulted in better calibration performance than the untargeted method. Nevertheless, the performance of the untargeted approach was fairly good. Due to the instrument characteristics (MRM vs qTOF), targeted analysis was more sensitive. This means that for the targeted method a small increase in analyte concentration causes a larger signal increase. Therefore, the LLOQ for all metabolites were higher when measured with the untargeted platform (see Tab. 8.1). The LLOQs ranged from 5.49 nM to 1,730 nM, whereas for the targeted analysis LLOQs ranged from 1.37 nM to 65.9 nM. The ULOQs were also higher. In numbers, the ULOQ ranged from 2,810 nM to 179,000 nM and from 1,140 nM to 270,000 nM for the untargeted and targeted analysis, respectively. Due to the higher sensitivity of the targeted

instrument, detector saturation is reached earlier in comparison to the untargeted platform. However, CRE is an early eluting metabolite which is only slightly retarded. Despite these results, general assumptions are rather critical due to the dependency on the used instruments. Moreover, one should keep in mind that the QTRAP 4000 used here for the comparison is not state of the art technology regarding linearity and sensitivity.

The slopes of the calibration curves from the untargeted analysis were distinctly lower in comparison to the targeted method. Only for XA the slope was steeper. This is surely the result of lacking an appropriate IS. In particular, slopes ranged from 0.3337 to 1.0050 whereas the targeted platform yielded slopes for the calibration curves ranging from 0.522 to 1.62. Additionally, the order of magnitudes (OOM) of the linear ranges were lower in comparison to the targeted method of all selected metabolites. This also reflects the sensitivity of the two methods. In order to check reproducibility of both approaches, 9 calibration levels over the complete calibration range were measured in triplicates. For all six metabolites measured with the untargeted instrument, RSDs were below 10% in the linear range. Therefore, RSDs of the untargeted MS were comparable with the RSDs measured in this concentration range with the targeted MS. Except for XA for which the LLOQ of the targeted measurements was higher (22 nM) than for the untargeted (5.49 nM) (see Tab. 8.1 and 8.2). There the RSD of the targeted measurements was 14% in comparison to 5% of the untargeted measurements. For XA no appropriate IS was available, therefore KA-d5 was selected instead. Consequently, it is more prone to ion suppression which could explain the increased RSDs of the targeted measurements.

However, it should be kept in mind that for the targeted analysis the Agilent LC was used whereas for the untargeted analysis a Dionex instrument was applied. Therefore, instrumental differences in void volume, injection mode or needle care also influence the chromatographic separation which impacts the calibration performance. Moreover, differences in electronics and construction designs can also affect quantitative performance. Additionally, in untargeted (screening) mode, background noise is also detected, whereas in targeted MRM mode background noise is suppressed.

Table 8.1 Calibration parameters untargeted method.

Analyte	LOD (nM)	Linear range (nM)	Slope	Intercept	R^2	OOM
CRE	527	1050-33,700	0.7382	170.665	0.9998	1
PSU	700	1400-179,000	0.5329	174.894	0.9996	2
CMT	22	44-11,300	0.3337	9.44343	0.9995	3
TRP	108	1730-111,000	0.3905	-204.639	0.9935	2
XA	0.172	5.49-2,810	11.716	68.2329	0.9987	3
HIP	176	176-45,000	1.005	-88.8484	0.9989	2

Table 8.2 Calibration parameters of the targeted method. Adapted from chapter 7.4.2

Analyte	LOD (nM)	Linear range (nM)	Slope	Intercept	R^2	OOM
CRE	2.06	65.9-270,000	0.96	0.00736	0.9995	4
PSU	0.684	1.37-44,800	0.573	0.000294	0.9991	4
CMT	0.687	2.75-11,300	0.522	0.000103	0.9982	4
TRP	1.69	6.76-13,800	1.62	0.000896	0.9977	4
XA	0.344	22-2,820	0.71	0.0194	0.9987	2
HIP	2.75	11.0-1,410	1.3	0.000951	0.9985	2

While analyzing these results, one should keep in mind that these results stem from data reintegrated by hand. Further, IS were available for the selected analytes (except XA) and were used for calibration for both methods (targeted and untargeted). Commonly, fingerprinting measures thousands of metabolites in parallel. It is a retrospective analysis method for which it is not possible to get IS for all measured metabolites. However, if performance of quantification via untargeted analysis without IS compares well to targeted quantification with IS this would be a tremendous advantage.

Lu et al., who also investigated targeted an untargeted MS analysis obtained similar results.[13] The reproducibility of all their tested standards were comparable to their targeted approach with RSD <10% (N=4). TOF (Agilent 6220 TOF) sensitivity was also comparable to the triple quadrupole (Thermo Scientific TSQ Quantum Ultra). However, it was operated with an unheated ESI source. The linear ranges of all compounds ranged over 2 to 3 orders of magnitude on the TOF-MS. Lu et al. concluded that the TOF-MS likely has reasonable quantitative performance for many targeted metabolomics applications. However, the

untargeted approach provides not automatically fully comparable quantitative results to targeted analysis. Although, 10 compounds were measured in positive mode the results of only one compound were shown. Moreover, the RSD of only one concentration level was analyzed and no further details were shown.

8.3.2 Comparison of population based measurements on targeted and untargeted platforms

Targeted and untargeted measurements of 200 CKD and 100 non-CKD urine samples were compared in order to examine the quantitative performance of the untargeted platform. Here, the untargeted measurements were performed without ISs for the metabolites, which is the routine approach in untargeted analysis.

First, the quality of the untargeted measurements was investigated via hierarchical clustering of all 200 GCKD and 100 NC samples. The GCKD and NC samples were measured in five batches consisting of 40 GCKD and 20 NC samples each. All the samples were measured randomized (for details see section 8.2.1). The clustering showed no obvious batch effects since all the samples, GCKD and NC, from all five batches clustered randomly and not batch wise (see Figure 8.1). The main resulting clusters are marked in blue, green and red. Hence, the quality of the data was suitable for a comparison with the targeted data. Quality of the quantitative data was shown by the calibration parameter and recovery described in chapter 7.

For the comparison between the two platforms, TRP, XA, PSU and CMT were used. These are prominent metabolites which were found in most of the 200 GCKD and 100 NC samples with concentrations below detector saturation. CRE and HIP were not analyzed because these metabolites were above ULOQ in both methods in the GCKD and NC samples.

The linear range of the untargeted method determined in chapter 8.3.1 was compared to the selected metabolites concentrations of the GCKD and NC samples. The concentration for PSU and CMT lay within the linear range for all GCKD and NC samples. For TRP, 77 GCKD samples (39% of the samples) and 20 NC samples (20% of the samples) were below the LLOQ. Moreover, for one

GCKD sample TRP was not detectable. For XA, 72 GCKD (39% of the samples) and 30 NC (32% of the samples) samples were below the LLOQ. Additionally, for 17 GCKD and 6 NC samples XA was not detectable. Therefore, most of the GCKD and NC samples were in the linear range of the untargeted platform. However, more GCKD samples were outside the linear range of the untargeted platform than NC samples. For PSU, CMT, TRP as well as XA one GCKD sample each was below the LLOQ of the targeted method. All NC samples were within the linear range of the targeted platform for all selected metabolites. Yet, all samples were used for the comparison because LLOQ and ULOQ are usually not available for routine fingerprinting data.

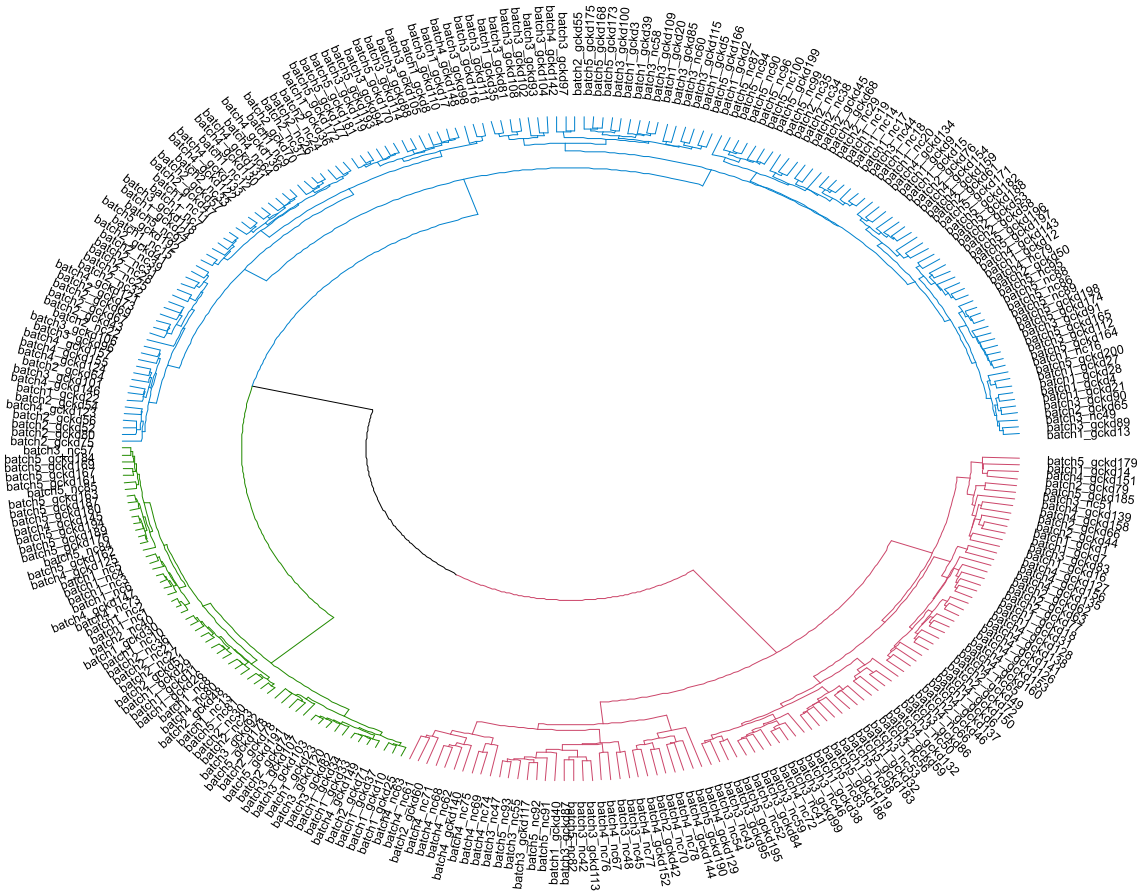


Figure 8.1:Hierachical cluster analysis of fingerprinting data of 200 GCKD and 100 NC samples measured with untargeted LC-qTOFMS. The main resulting clusters are marked in blue, green and red.

In order to determine the quantification performance of the untargeted platform, Spearman's rank correlation coefficients (SCC) were calculated for the selected

metabolites. For the GCKD samples, targeted and untargeted measurements of PSU yielded a SCC of 0.67, while the corresponding coefficients for XA, CMT and TRP were 0.83, 0.87, and 0.85, respectively. The lower SCC for PSU is probably due to its low retention time (1.3 min), at which many metabolites elute simultaneously. Thus, ion suppression might affect the quantification. Bland-Altmann plots of the ranks of targeted and untargeted measurements revealed a mean difference of -1.0, -0.7, 0.0 and 0.0 for PSU, CMT, TRP and XA, respectively. 95% confidence intervals ranged from 90.8 to -93.8, 56.1 to -57.5, 60.5 to -60.5, and 57.7 to -57.7 for PSU, CMT, XA and TRP, respectively (see Fig. 8.3 E-H). Confidence intervals were rather high, but no systematic bias was observed.

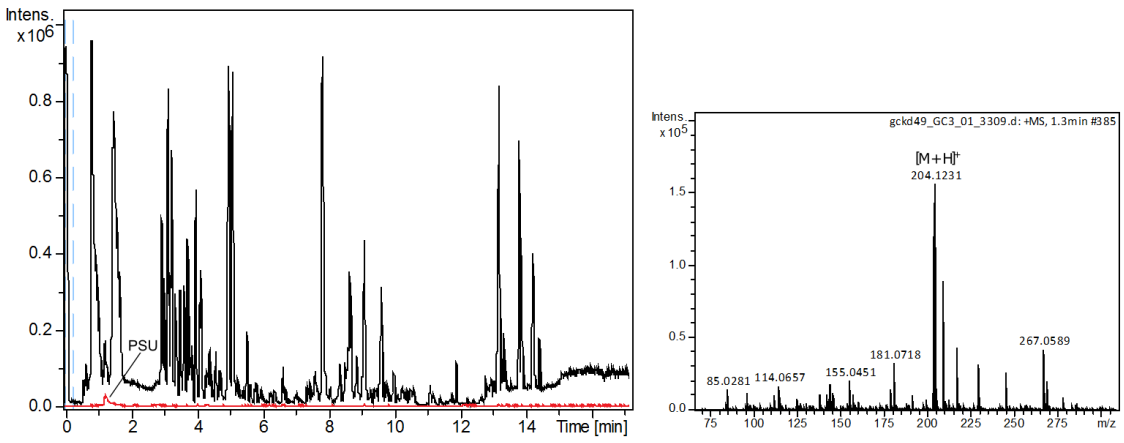


Figure 8.2: Representative chromatogram of a GCKD urine sample measured in untargeted mode by LC-qTOF MS. XIC of PSU (red line) is depicted on the left and the corresponding mass spectrum on the right. At 1.3 min coelution of unknown metabolites with PSU is observable.

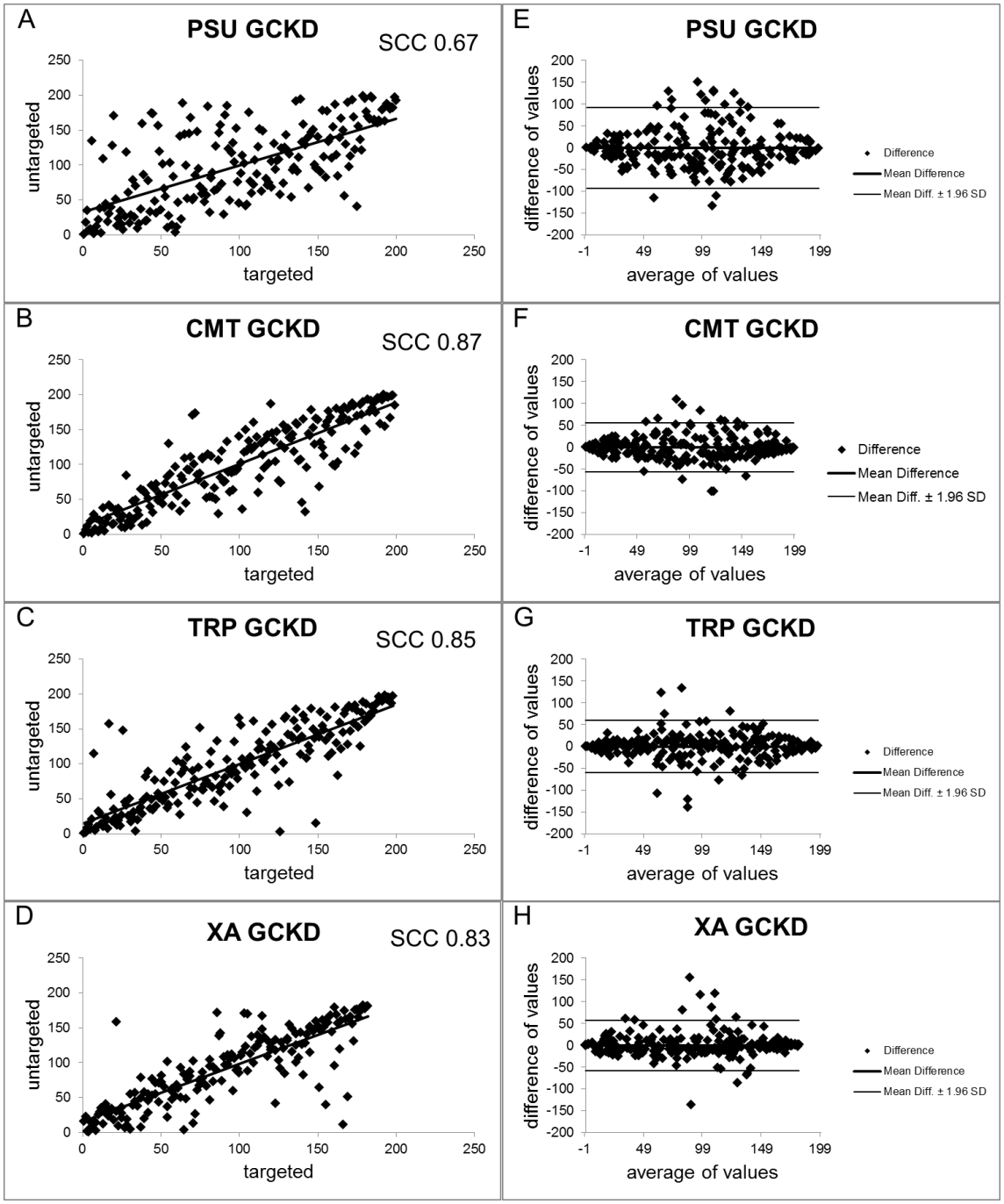


Figure 8.3: Comparison of ranks of targeted and untargeted mass spectrometric measurements of pseudouridine (PSU), C-mannosyltryptophan (CMT), tryptophan (TRP) and xanthurenic acid (XA) in individuals with chronic kidney disease (GCKD) with the corresponding Bland-Altmann plots (E-H). SCC: Spearman correlation coefficient corresponding to slope of displayed regression line of rank transformed measurements.

As mentioned above for fingerprinting analysis no IS were used to correct for ion suppression. The SCC of PSU, CMT, XA and TRP for NC samples were investigated. These samples represent average diet, environmental exposures and medication. The SCC of PSU for the NC samples is 0.73. For CMT the

SCC was 0.95, for XA 0.94 and for TRP 0.90. This means that targeted and non-targeted measurements showed a very good correlation. In comparison to the SCCs of the GCKD samples, targeted and non-targeted measurements of NC samples correlate much better. Bland-Altmann plots of the ranks of targeted and untargeted MS measurements revealed a mean difference of -1.0, 0.0, -0.6 and 0.0 for PSU, CMT, XA and TRP, respectively. 95% confidence intervals ranged from 36.6 and -38.5, 18.4 and -18.4, 18.1 and -19.3 as well as 26.0 and -26.0, PSU, CMT, XA and TRP respectively (see Fig. 8.4 E-H). Confidence intervals are also rather high. However, no systematic bias was observed.

Chromatographic separation as well as the concentration range of the analytes in the samples affect correlation. However, also the availability of a stable isotope labeled analog as IS for the targeted analysis, which was not the case for XA, influence the correlation between targeted and untargeted measurements. Nevertheless, the results indicate that relative quantification by untargeted screening is reliable in large studies with this instrument setup.

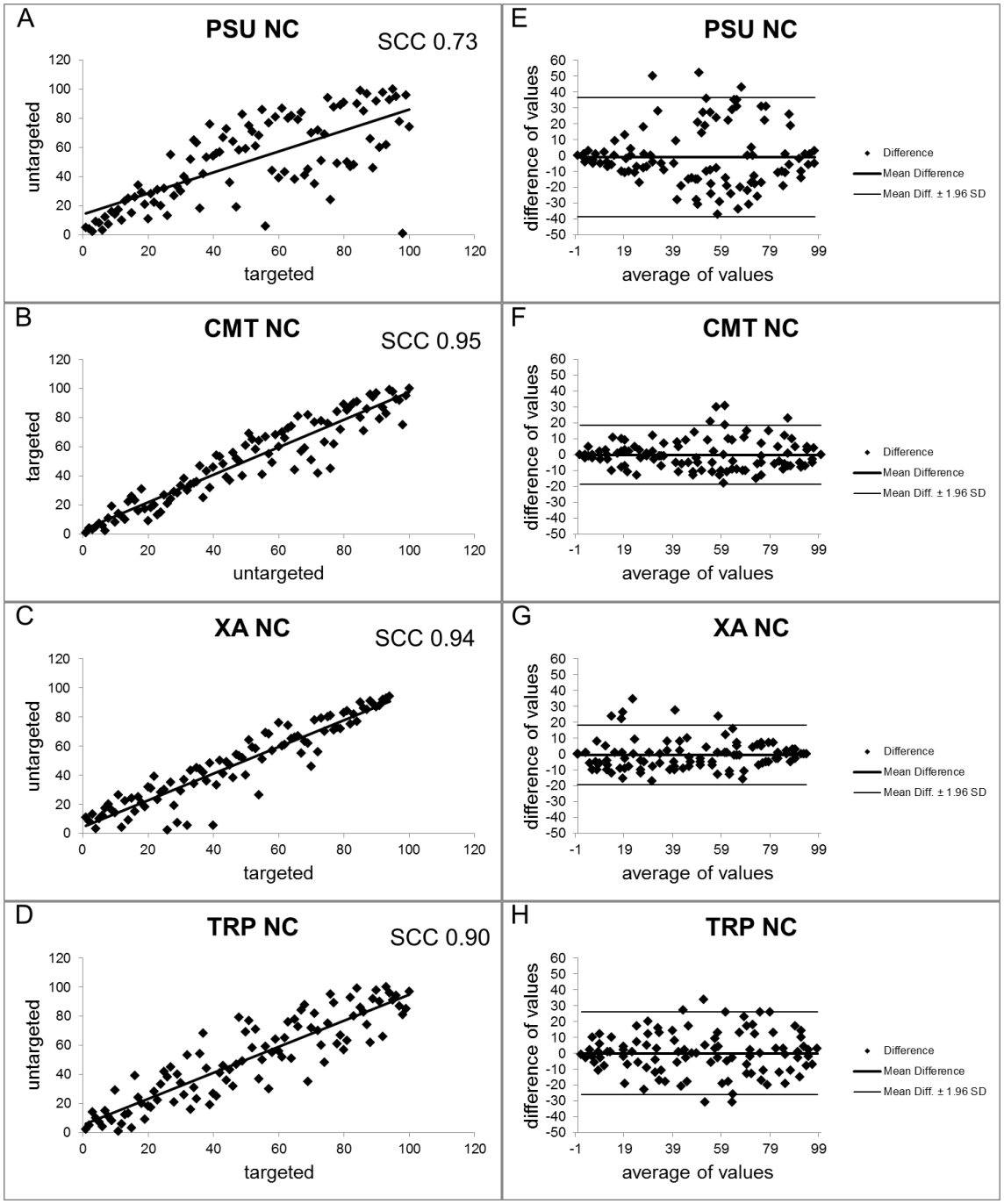


Figure 8.4: Comparison of ranks of targeted and untargeted mass spectrometric measurements of pseudouridine (PSU), C-mannosyltryptophan (CMT), xanthurenic acid (XA) and tryptophan (TRP) in apparently healthy individuals from the national cohort study (NC) with the corresponding Bland-Altmann plots (E-H). SCC: Spearman correlation coefficient corresponding to slope of displayed regression line of rank transformed measurements.

In Sekula et al. (2017), targeted and untargeted measurements of 111 samples of non-CKD patients from the QMDiab study were compared.[44] These targeted measurements were performed using the quantitative method described in chapter 7, whereas untargeted measurements were carried out by Metabolon.

For pseudouridine they found a SCC of 0.67 for the correlation of targeted and untargeted measurements in plasma. The correlation for urine samples was even better with a SCC of 0.90. For CMT the SSC was 0.69 in plasma and 0.89 in urine. For CRE the SCC was 0.68 in plasma and 0.95 in urine. This is especially impressive as targeted and untargeted measurements were performed in different laboratories using different equipment and sample preparation protocols.

Sekula et al also highlighted previous studies which compared targeted and untargeted measurements.[44,129-132] They came to the conclusion that biomarker discovery by untargeted methods should be validated quantitatively by targeted approaches.[44] Besides, they pointed out that general assumption across metabolomics platforms are challenging. Further they emphasize the importance of using absolute quantification results in order to validate biomarker discoveries.[44]

8.4 Conclusions

The linear range determined with selected metabolites for the untargeted platform proved to be sufficient in comparison to the targeted platform. Reproducibility in the linear range is very good in comparison to the targeted approach. Moreover, targeted and untargeted measurements of urine samples from the NC yielded a very good agreement, whereas the agreement of GCKD samples was lower but still adequate, which was attributed to stronger matrix effects for the GCKD samples.

Overall the correlation of targeted and untargeted platforms is surely instrument, metabolite and concentration dependent. Moreover, chromatographic separation and ion-suppression are also affecting the correlation. However, it certainly is impossible to separate all detected metabolites of urine samples measured with the untargeted method, in order to reduce ion suppression. By adding ISs to the matrix under investigation, untargeted measurements could provide supplementary quantitative insights into preselected metabolites. Nevertheless, quantification by fingerprinting without using ISs could be sufficient enough for some metabolites to replace targeted. This surely depends on the aim and

scope of the analytical question under investigation. Nevertheless, general assumptions across different platforms are critical and are surely dependent on the type of instruments used. Since even small variations in measurements can result in drastic deviations of the results as demonstrated by Sekula et al. for the calculation of the fractional excretion, it would be advisable, as it is the case for every quantification, to obtain a secondary analysis with an independent platform to get unbiased and reliable quantitative results. Despite, this approach would combine both quantitative and untargeted information which is a tremendous advantage over targeted analysis.

9. Conclusion and Perspectives

This thesis clearly demonstrates the vast potential of high-resolution LC-MS in untargeted metabolomics of complex biological matrices like urine. Through the introduction of pre-acquisition dilution to a uniform creatinine concentration, significant improvements were achieved in both sample classification and metabolite identification. Further promising accomplishments were the implementation of a quantitative LC-MS method to investigate urinary metabolites associated with renal disease in order to reveal new insights into pathophysiological mechanisms and the quantitative comparison of LC-qTOF-MS and LC-qQq-MS in order to investigate the combined potential of targeted and untargeted LC-MS metabolomics on a single platform.

The comparison of different pre-acquisition dilution strategies of urine clearly demonstrated that dilution to a uniform creatinine concentration outperforms all other tested strategies. This pre-acquisition strategy was the only method which showed no correlation of MVs with creatinine concentration or osmolality values. It yielded the highest percentage of significant differential metabolites and was the sole strategy for which the hierarchical cluster analysis reliably assigned the urine specimens of controls and CKD patients to their respective clusters. Furthermore, the proposed normalization method allows reliable sample classification as demonstrated for several different patient cohorts. Apparently, dilution to a uniform creatinine concentration is the only tested method which balances the high variability in metabolite abundance between samples sufficiently enough to eliminate analytical consequences like ion suppression, detector saturation and column overloading. The results obtained underscore the improved reproducibility and applicability of pre-acquisition dilution to a uniform creatinine concentration and, thus, justify the associated time and labor.

The quantitative LC-qQqMS method implemented in this thesis comprised significantly discriminating metabolites between CKD patients and apparently healthy specimen which had been identified by the aforementioned project. This quantitative method expands the one published by Zhu et al. in 2011. It was successfully applied to the analysis of CRE, CMT and PSU in the study of Sekula et al. 2017 and for the analysis of CRE and CRT in the study of Wall-

meier et al. 2017. Within these studies the excellent quality of the developed method was demonstrated in different matrices.

A further key aspect of this thesis was the comparison of targeted LC-qQq-MS and untargeted LC-qTOF-MS measurements of urine samples elucidating the quantitative performance of the used untargeted platform. Merging targeted and untargeted analysis of samples within one run would be a tremendous advantage over separate quantitative MRM- and untargeted measurements. Results from the comparison showed that for the selected metabolites high correlations were yielded. However, the correlations are dependent on the instruments, the metabolites and on the chromatographic separation. Therefore, the application of untargeted analysis of urine samples without using IS for the relative quantification is limited. Still, quantification via fingerprinting LC-MS results in semi-quantitative data, thus verification by an LC-MS measurement in MRM mode is advisable. Moreover, as already stated by Sekula et al. 2017 general assumptions across platforms are critical. However, quantification by fingerprinting without using ISs could be sufficient enough for some metabolites to replace targeted analysis. This surely would be dependent on the aim and scope of the analytical question under investigation.

For fingerprinting analysis by LC-MS further studies need to focus on the implementation of identification tools to simplify the identification of unknowns. Therefore, more comprehensive MS metabolite libraries are required as well as more comprehensive identification tools like the compound crawler from Bruker. In order to distinctly identify unknown metabolites, the supply of proper reference compounds must be increased. Additionally, data storage becomes more and more an issue due to increasing data file sizes with LC-MS measurements (full scan and MS/MS LC-MS).

10. References

(1) Zelena, E.; Dunn, W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. M.; Begley, P.; O'Hagan, S.; Knowles, J. D.; Halsall, A.; Wilson, I. D.; Kell, D. B. *Analytical Chemistry* **2009**, *81*, 1357-1364.

(2) Barr, D. B.; Wilder, L. C.; Caudill, S. P.; Gonzalez, A. J.; Needham, L. L.; Pirkle, J. L. *Environmental Health Perspectives* **2005**, *113*, 192-200.

(3) Ryan, D.; Robards, K.; Prenzler, P. D.; Kendall, M. *Analytica Chimica Acta* **2011**, *684*, 8-20.

(4) Warrack, B. M.; Hnatyshyn, S.; Ott, K.-H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M. *Journal of Chromatography B* **2009**, *877*, 547-552.

(5) Boeniger, M. F.; Lowry, L. K.; Rosenberg, J. *American Industrial Hygiene Association Journal* **1993**, *54*, 615-627.

(6) Chadha, V.; Garg, U.; Alon, U. S. *Pediatric Nephrology* **2001**, *16*, 374-382.

(7) Eckardt, K. U.; Coresh, J.; Devuyst, O.; Johnson, R. J.; Kottgen, A.; Levey, A. S.; Levin, A. *Lancet* **2013**, *382*, 158-169.

(8) Stevens, L. A.; Coresh, J.; Greene, T.; Levey, A. S. *New England Journal of Medicine* **2006**, *354*, 2473-2483.

(9) Shlipak, M. G.; Matsushita, K.; Arnlov, J.; Inker, L. A.; Katz, R.; Polkinghorne, K. R.; Rothenbacher, D.; Sarnak, M. J.; Astor, B. C.; Coresh, J.; Levey, A. S.; Gansevoort, R. T. *New England Journal of Medicine* **2013**, *369*, 932-943.

(10) Inker, L. A.; Schmid, C. H.; Tighiouart, H.; Eckfeldt, J. H.; Feldman, H. I.; Greene, T.; Kusek, J. W.; Manzi, J.; Van Lente, F.; Zhang, Y. L.; Coresh, J.; Levey, A. S. *New England Journal of Medicine* **2012**, *367*, 20-29.

(11) Poge, U.; Stoschus, B.; Stoffel-Wagner, B.; Gerhardt, T.; Klehr, H. U.; Sauerbruch, T.; Woitas, R. P. *Kidney Blood Press Res* **2003**, *26*, 55-60.

(12) Sekula, P.; Goek, O. N.; Quaye, L.; Barrios, C.; Levey, A. S.; Romisch-Margl, W.; Menni, C.; Yet, I.; Gieger, C.; Inker, L. A.; Adamski, J.; Gronwald, W.; Illig, T.; Dettmer, K.; Krumsiek, J.; Oefner, P. J.; Valdes, A. M.; Meisinger, C.; Coresh, J.; Spector, T. D.; Mohney, R. P.; Suhre, K.; Kastenmuller, G.; Kottgen, A. *Journal of the American Society of Nephrology* **2016**, *27*, 1175-1188.

(13) Lu, W.; Bennett, B. D.; Rabinowitz, J. D. *J Chromatogr B Analyt Technol Biomed Life Sci* **2008**, *871*, 236-242.

(14) Eckardt, K. U.; Barthlein, B.; Baid-Agrawal, S.; Beck, A.; Busch, M.; Eitner, F.; Ekici, A. B.; Floege, J.; Gefeller, O.; Haller, H.; Hilge, R.; Hilgers, K. F.; Kielstein, J. T.; Krane, V.; Kottgen, A.; Kronenberg, F.; Oefner, P.; Prokosch, H. U.; Reis, A.; Schmid, M.; Schaeffner, E.; Schultheiss, U. T.; Seuchter, S. A.; Sitter, T.; Sommerer, C.; Walz, G.; Wanner, C.; Wolf, G.; Zeier, M.; Titze, S. *Nephrology, Dialysis, Transplantation* **2012**, *27*, 1454-1460.

(15) Pfeffer, M. A.; Burdmann, E. A.; Chen, C. Y.; Cooper, M. E.; de Zeeuw, D.; Eckardt, K. U.; Feyzi, J. M.; Ivanovich, P.; Kewalramani, R.; Levey, A. S.; Lewis, E. F.; McGill, J. B.; McMurray, J. J.; Parfrey, P.; Parving, H. H.; Remuzzi, G.; Singh, A. K.; Solomon, S. D.; Toto, R. *New England Journal of Medicine* **2009**, *361*, 2019-2032.

(16) Berg, M.; Vanaerschot, M.; Jankevics, A.; Cuypers, B.; Breitling, R.; Dujardin, J.-C. *Computational and Structural Biotechnology Journal* **2013**, *4*, 1-8.

(17) Scalbert, A.; Brennan, L.; Fiehn, O.; Hankemeier, T.; Kristal, B. S.; van Ommen, B.; Pujos-Guillot, E.; Verheij, E.; Wishart, D.; Wopereis, S. *Metabolomics* **2009**, *5*, 435-458.

(18) Dunn, W. B.; Broadhurst, D. I.; Atherton, H. J.; Goodacre, R.; Griffin, J. L. *Chem Soc Rev* **2011**, *40*, 387-426.

(19) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrometry Reviews* **2007**, *26*, 51-78.

(20) Dettmer, K.; Hammock, B. D. *Environmental Health Perspectives* **2004**, *112*, A396-A397.

(21) Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. *Nat Rev Mol Cell Biol* **2004**, *5*, 763-769.

(22) Holmes, E.; Wilson, I. D.; Nicholson, J. K. *Cell* **2008**, *134*, 714-717.

(23) Nicholson, J. K. *Mol Syst Biol* **2006**, *2*, 52.

(24) Hood, L.; Heath, J. R.; Phelps, M. E.; Lin, B. *Science* **2004**, *306*, 640-643.

(25) Kanani, H.; Chrysanthopoulos, P. K.; Klapa, M. I. *J Chromatogr B Analyt Technol Biomed Life Sci* **2008**, *871*, 191-201.

(26) Theodoridis, G. A.; Gika, H. G.; Want, E. J.; Wilson, I. D. *Anal Chim Acta* **2012**, *711*, 7-16.

(27) Zhou, B.; Xiao, J. F.; Tuli, L.; Ressom, H. W. *Mol Biosyst* **2012**, *8*, 470-481.

(28) Dunn, W. B.; Ellis, D. I. *Trac-Trends in Analytical Chemistry* **2005**, *24*, 285-294.

(29) Dettmer-Wilde, K.; Engewald, W. *Practical Gas Chromatography*; Springer, 2014.

(30) Zacharias, H. U.; Hochrein, J.; Vogl, F. C.; Schley, G.; Mayer, F.; Jeleazcov, C.; Eckardt, K. U.; Willam, C.; Oefner, P. J.; Gronwald, W. *J Proteome Res* **2015**, *14*, 2897-2905.

(31) Levey, A. S.; Eckardt, K. U.; Tsukamoto, Y.; Levin, A.; Coresh, J.; Rossert, J.; De Zeeuw, D.; Hostetter, T. H.; Lameire, N.; Eknoyan, G. *Kidney International* **2005**, *67*, 2089-2100.

(32) Zhao, Y. Y. *Clinica Chimica Acta* **2013**, *422*, 59-69.

(33) Boelaert, J.; t'Kindt, R.; Schepers, E.; Jorge, L.; Glorieux, G.; Neirynck, N.; Lynen, F.; Sandra, P.; Vanholder, R.; Sandra, K. *Metabolomics* **2014**, *10*, 425-442.

(34) Sekula, P.; Goek, O. N.; Quaye, L.; Barrios, C.; Levey, A. S.; Romisch-Margl, W.; Menni, C.; Yet, I.; Gieger, C.; Inker, L. A.; Adamski, J.; Gronwald, W.; Illig, T.; Dettmer, K.; Krumsiek, J.; Oefner, P. J.; Valdes,

A. M.; Meisinger, C.; Coresh, J.; Spector, T. D.; Mohney, R. P.; Suhre, K.; Kastenmuller, G.; Kottgen, A. *J Am Soc Nephrol* **2016**, *27*, 1175-1188.

(35) Duranton, F.; Cohen, G.; De Smet, R.; Rodriguez, M.; Jankowski, J.; Vanholder, R.; Argiles, A. *J Am Soc Nephrol* **2012**, *23*, 1258-1270.

(36) Levin, A.; Tonelli, M.; Bonventre, J.; Coresh, J.; Donner, J. A.; Fogo, A. B.; Fox, C. S.; Gansevoort, R. T.; Heerspink, H. J. L.; Jardine, M.; Kasiske, B.; Kottgen, A.; Kretzler, M.; Levey, A. S.; Luyckx, V. A.; Mehta, R.; Moe, O.; Obrador, G.; Pannu, N.; Parikh, C. R.; Perkovic, V.; Pollock, C.; Stenvinkel, P.; Tuttle, K. R.; Wheeler, D. C.; Eckardt, K. U. *Lancet* **2017**.

(37) Zhang, J.; Yan, L.; Chen, W.; Lin, L.; Song, X.; Yan, X.; Hang, W.; Huang, B. *Clinica Chimica Acta* **2009**, *650*, 16-22.

(38) *Diabetes Care* **2003**, *26*, s94-s98.

(39) Ghosh, S.; Collier, A. *Churchill's Pocketbook of Diabetes E-Book*; Elsevier Health Sciences, 2012.

(40) Levey, A. S.; Inker, L. A. *Clin Pharmacol Ther* **2017**, *102*, 405-419.

(41) Levey, A. S.; Inker, L. A.; Coresh, J. *Am J Kidney Dis* **2014**, *63*, 820-834.

(42) Van der Kloet, F.; Tempels, F.; Ismail, N.; Van der Heijden, R.; Kasper, P.; Rojas-Cherto, M.; Van Doorn, R.; Spijksma, G.; Koek, M.; Van der Greef, J. *Metabolomics* **2012**, *8*, 109-119.

(43) Weiss, R. H.; Kim, K. M. *Nature Reviews Nephrology* **2012**, *8*, 22-33.

(44) Sekula, P.; Dettmer, K.; Vogl, F. C.; Gronwald, W.; Ellmann, L.; Mohney, R. P.; Eckardt, K. U.; Suhre, K.; Kastenmuller, G.; Oefner, P. J.; Kottgen, A. *Sci Rep* **2017**, *7*, 17400.

(45) Shah, V. O.; Townsend, R. R.; Feldman, H. I.; Pappan, K. L.; Kensicki, E.; Vander Jagt, D. L. *Clin J Am Soc Nephrol* **2013**, *8*, 363-370.

(46) Sato, E.; Kohno, M.; Yamamoto, M.; Fujisawa, T.; Fujiwara, K.; Tanaka, N. *Eur J Clin Invest* **2011**, *41*, 241-255.

(47) Jia, L.; Chen, J.; Yin, P.; Lu, X.; Xu, G. *Metabolomics* **2008**, *4*, 183-189.

(48) Baraldi, A. N.; Enders, C. K. *Journal of School Psychology* **2010**, *48*, 5-37.

(49) Hrydziuszko, O.; Viant, M. R. *Metabolomics* **2012**, *8*, 161-174.

(50) Bartel, J.; Krumsiek, J.; Theis, F. J. *Computational and Structural Biotechnology Journal* **2013**, *4*, e201301009.

(51) Benjamini, Y.; Hochberg, Y. *Journal of the Royal Statistical Society: Series B (Methodological)* **1995**, *57*, 289-300.

(52) Worley, B.; Powers, R. *Curr Metabolomics* **2013**, *1*, 92-107.

(53) Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R. *Clinica Chimica Acta* **2015**, *879*, 10-23.

(54) Mahadevan, S.; Shah, S. L.; Marrie, T. J.; Slupsky, C. M. *Analytical Chemistry* **2008**, *80*, 7562-7570.

(55) Dunn, W. B.; Erban, A.; Weber, R. J. M.; Creek, D. J.; Brown, M.; Breitling, R.; Hankemeier, T.; Goodacre, R.; Neumann, S.; Kopka, J.; Viant, M. R. *Metabolomics* **2013**, *9*, S44-S66.

(56) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. *Nucleic Acids Res* **2013**, *41*, D801-807.

(57) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther Drug Monit* **2005**, *27*, 747-751.

(58) Cui, Q.; Lewis, I. A.; Hegeman, A. D.; Anderson, M. E.; Li, J.; Schulte, C. F.; Westler, W. M.; Eghbalnia, H. R.; Sussman, M. R.; Markley, J. L. *Nature Biotechnology* **2008**, *26*, 162-164.

(59) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. *J Mass Spectrom* **2010**, *45*, 703-714.

(60) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H., Jr.; Murphy, R. C.; Raetz, C. R.; Russell, D. W.; Subramaniam, S. *Nucleic Acids Res* **2007**, *35*, D527-532.

(61) Vinaixa, M.; Schymanski, E. L.; Neumann, S.; Navarro, M.; Salek, R. M.; Yanes, O. *TrAC Trends in Analytical Chemistry* **2016**, *78*, 23-35.

(62) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. *Metabolomics* **2007**, *3*, 211-221.

(63) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2006**, *7*, 234.

(64) Moco, S.; Bino, R. J.; De Vos, R. C. H.; Vervoort, J. *Trac-Trends in Analytical Chemistry* **2007**, *26*, 855-866.

(65) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2007**, *8*, 105.

(66) Degtyarenko, K.; de Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcantara, R.; Darsow, M.; Guedj, M.; Ashburner, M. *Nucleic Acids Res* **2008**, *36*, D344-350.

(67) Holcapek, M.; Jirasko, R.; Lisa, M. *Journal of Chromatography A* **2010**, *1217*, 3908-3921.

(68) Dettmer, K.; Vogl, F. C.; Ritter, A. P.; Zhu, W.; Nurnberger, N.; Kreutz, M.; Oefner, P. J.; Gronwald, W.; Gottfried, E. *Electrophoresis* **2013**, *34*, 2836-2847.

(69) Zhu, W.; Stevens, A. P.; Dettmer, K.; Gottfried, E.; Hoves, S.; Kreutz, M.; Holler, E.; Canelas, A. B.; Kema, I.; Oefner, P. J. *Anal Bioanal Chem* **2011**, *401*, 3249-3261.

(70) Team, R. C.; ISBN 3-900051-07-0, 2014.

(71) Suzuki, R.; Shimodaira, H. *Bioinformatics* **2006**, *22*, 1540-1542.

(72) Pollard, K. S.; Dudoit, S.; van der Laan, M. J. In *Bioinformatics and computational biology solutions using R and bioconductor*; Springer, 2005, pp 249-271.

(73) Food, U.; Administration, D.

(74) Vogl, F. C.; Mehrl, S.; Heizinger, L.; Schlecht, I.; Zacharias, H. U.; Ellmann, L.; Nurnberger, N.; Gronwald, W.; Leitzmann, M. F.; Rossert, J.; Eckardt, K. U.; Dettmer, K.; Oefner, P. J. *Anal Bioanal Chem* **2016**, *408*, 8483-8493.

(75) Matthiesen, R. *Bioinformatics Methods in Clinical Research*; Springer, 2010.

(76) Sugimoto, M.; Kawakami, M.; Robert, M.; Soga, T.; Tomita, M. *Curr Bioinform* **2012**, *7*, 96-108.

(77) Bland, J. M.; Altman, D. *Lancet* **1986**, *327*, 307-310.

(78) Giavarina, D. *Biochem Med (Zagreb)* **2015**, *25*, 141-151.

(79) ltman, D. G.; Bland, J. M. *The Statistician* **1983**, 307-317.

(80) Kaspar, H.; Dettmer, K.; Chan, Q.; Daniels, S.; Nimkar, S.; Daviglus, M. L.; Stamler, J.; Elliott, P.; Oefner, P. J. *J Chromatogr B Analyt Technol Biomed Life Sci* **2009**, *877*, 1838-1846.

(81) Barr, D. B.; Wilder, L. C.; Caudill, S. P.; Gonzalez, A. J.; Needham, L. L.; Pirkle, J. L. *Environmental Health Perspectives* **2005**, 192-200.

(82) Zhang, A.; Sun, H.; Wu, X.; Wang, X. *Clinica Chimica Acta* **2012**, *414*, 65-69.

(83) Lutz, U.; Bittner, N.; Lutz, R. W.; Lutz, W. K. *Journal of Chromatography B* **2008**, *871*, 349-356.

(84) Chen, Y.; Shen, G.; Zhang, R.; He, J.; Zhang, Y.; Xu, J.; Yang, W.; Chen, X.; Song, Y.; Abliz, Z. *Analytical Chemistry* **2013**, *85*, 7659-7665.

(85) Mattarucchi, E.; Guillou, C. *Biomedical Chromatography* **2012**, *26*, 512-517.

(86) Edmands, W. M.; Ferrari, P.; Scalbert, A. *Analytical Chemistry* **2014**, *86*, 10925-10931.

(87) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. *Journal of Chromatography A* **2016**, *1431*, 103-110.

(88) Benjamini, Y.; Hochberg, Y. *Journal of the Royal Statistical Society Series B-Methodological* **1995**, *57*, 289-300.

(89) Warnes, G. R.; Bolker, B.; Bonebakker, L.; Gentleman, R.; Huber, W.; Liaw, A.; Lumley, T.; Maechler, M.; Magnusson, A.; Moeller, S. *R package version* **2009**, *2*.

(90) Bolstad, B. M.; Irizarry, R. A.; Åstrand, M.; Speed, T. P. *Bioinformatics* **2003**, *19*, 185-193.

(91) Breiman, L. *Machine Learning* **2001**, *45*, 5-32.

(92) Liaw, A.; Wiener, M. *R news* **2002**, *2*, 18-22.

(93) Zacharias, H. U.; Schley, G.; Hochrein, J.; Klein, M. S.; Köberle, C.; Eckardt, K.-U.; Willam, C.; Oefner, P. J.; Gronwald, W. *Metabolomics* **2013**, *9*, 697-707.

(94) Reddi, A. S. In *Fluid, Electrolyte and Acid-Base Disorders*; Springer New York, 2014, pp 13-19.

(95) Zhao, Y.-Y. *Clinica Chimica Acta* **2013**, *422*, 59-69.

(96) Snyder, S.; Pendergraph, B. *Interventions* **2005**, *100*, 24-25.

(97) Suzuki, R.; Shimodaira, H. In *The Fifteenth International Conference on Genome Informatics*, 2004.

(98) Free, A.; Johnston, K. In *Clinical Chemistry*; AMER ASSOC CLINICAL CHEMISTRY 2101 L STREET NW, SUITE 202, WASHINGTON, DC 20037-1526, 1981, pp 1116-1116.

(99) George, J. W. *Veterinary Clinical Pathology* **2001**, *30*, 201-210.

(100) Voinescu, G. C.; Shoemaker, M.; Moore, H.; Khanna, R.; Nolph, K. D. *American Journal of the Medical Sciences* **2002**, *323*, 39-42.

(101) Miller, R. C.; Brindle, E.; Holman, D. J.; Shofer, J.; Klein, N. A.; Soules, M. R.; O'Connor, K. A. *Clinical Chemistry* **2004**, *50*, 924-932.

(102) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E. *Nucleic Acids Research* **2012**, gks1065.

(103) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Therapeutic Drug Monitoring* **2005**, *27*, 747-751.

(104) Degtyarenko, K.; De Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcántara, R.; Darsow, M.; Guedj, M.; Ashburner, M. *Nucleic Acids Research* **2008**, *36*, D344-D350.

(105) Zhao, Y. Y. *Clinica Chimica Acta* **2013**, *422*, 59-69.

(106) Galili, T. *Bioinformatics* **2015**, btv428.

(107) Gu, Z.; Gu, L.; Eils, R.; Schlesner, M.; Brors, B. *Bioinformatics* **2014**, btu393.

(108) Takahira, R.; Yonemura, K.; Yonekawa, O.; Iwahara, K.; Kanno, T.; Fujise, Y.; Hishida, A. *Am J Med* **2001**, *110*, 192-197.

(109) Yonemura, K.; Takahira, R.; Yonekawa, O.; Wada, N.; Hishida, A. *Kidney Int* **2004**, *65*, 1395-1399.

(110) Niewczas, M. A.; Sirich, T. L.; Mathew, A. V.; Skupien, J.; Mohney, R. P.; Warram, J. H.; Smiles, A.; Huang, X.; Walker, W.; Byun, J.; Karoly, E. D.; Kensicki, E. M.; Berry, G. T.; Bonventre, J. V.; Pennathur, S.; Meyer, T. W.; Krolewski, A. S. *Kidney International* **2014**, *85*, 1214-1224.

(111) Dzurik, R.; Lajdova, I.; Spustova, V.; Opatrny, K., Jr. *Nephron* **1992**, *61*, 64-67.

(112) Niwa, T.; Takeda, N.; Yoshizumi, H. *Kidney Int* **1998**, *53*, 1801-1806.

(113) Goek, O. N.; Prehn, C.; Sekula, P.; Romisch-Margl, W.; Doring, A.; Gieger, C.; Heier, M.; Koenig, W.; Wang-Sattler, R.; Illig, T.; Suhre, K.; Adamski, J.; Kottgen, A.; Meisinger, C. *Nephrol Dial Transplant* **2013**, *28*, 2131-2138.

(114) Rhee, E. P.; Clish, C. B.; Ghorbani, A.; Larson, M. G.; Elmariah, S.; McCabe, E.; Yang, Q.; Cheng, S.; Pierce, K.; Deik, A.; Souza, A. L.; Farrell, L.; Domos, C.; Yeh, R. W.; Palacios, I.; Rosenfield, K.; Vasan, R.

S.; Florez, J. C.; Wang, T. J.; Fox, C. S.; Gerszten, R. E. *J Am Soc Nephrol* **2013**, *24*, 1330-1338.

(115) Yu, B.; Zheng, Y.; Nettleton, J. A.; Alexander, D.; Coresh, J.; Boerwinkle, E. *Clin J Am Soc Nephrol* **2014**, *9*, 1410-1417.

(116) Kalim, S.; Rhee, E. P. *Kidney Int* **2017**, *91*, 61-69.

(117) Hocher, B.; Adamski, J. *Nature Reviews Nephrology* **2017**, *13*, 269-284.

(118) Liu, J.; Wang, D.; Chen, Y.; Sun, H.; He, S.; Wang, C.; Yang, G.; Shi, M.; Zhang, J.; Ren, Y.; Wang, L.; Lu, Y.; Cheng, J. *Mol Biosyst* **2013**, *9*, 2645-2652.

(119) Choi, J. Y.; Yoon, Y. J.; Choi, H. J.; Park, S. H.; Kim, C. D.; Kim, I. S.; Kwon, T. H.; Do, J. Y.; Kim, S. H.; Ryu, D. H.; Hwang, G. S.; Kim, Y. L. *Nephrol Dial Transplant* **2011**, *26*, 1304-1313.

(120) Schoots, A. C.; De Vries, P. M.; Thiemann, R.; Hazejager, W. A.; Visser, S. L.; Oe, P. L. *Clin Chim Acta* **1989**, *185*, 91-107.

(121) Wallmeier, J.; Samol, C.; Ellmann, L.; Zacharias, H. U.; Vogl, F. C.; Garcia, M.; Dettmer, K.; Oefner, P. J.; Gronwald, W. *J Proteome Res* **2017**, *16*, 1784-1796.

(122) Park, M. H.; Lee, M. W.; Shin, Y. G. *Biomedical Chromatography* **2016**, *30*, 625-631.

(123) Cajka, T.; Fiehn, O. *Anal Chem* **2016**, *88*, 524-545.

(124) Tsugawa, H.; Ohta, E.; Izumi, Y.; Ogiwara, A.; Yukihira, D.; Bamba, T.; Fukusaki, E.; Arita, M. *Frontiers in genetics* **2015**, *5*, 471.

(125) Tsugawa, H.; Arita, M.; Kanazawa, M.; Ogiwara, A.; Bamba, T.; Fukusaki, E. *Analytical chemistry* **2013**, *85*, 5191-5199.

(126) Dillen, L.; Cools, W.; Vereyken, L.; Lorreyne, W.; Huybrechts, T.; de Vries, R.; Ghobarah, H.; Cuyckens, F. *Bioanalysis* **2012**, *4*, 565-579.

(127) Ramanathan, R.; Jemal, M.; Ramagiri, S.; Xia, Y. Q.; Humpreys, W. G.; Olah, T.; Korfmacher, W. A. *J Mass Spectrom* **2011**, *46*, 595-601.

(128) Suhre, K.; Meisinger, C.; Doring, A.; Altmaier, E.; Belcredi, P.; Gieger, C.; Chang, D.; Milburn, M. V.; Gall, W. E.; Weinberger, K. M.; Mewes, H. W.; Hrabe de Angelis, M.; Wichmann, H. E.; Kronenberg, F.; Adamski, J.; Illig, T. *PLoS One* **2010**, *5*, e13953.

(129) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T. L.; Smith, S. R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J. W.; Goodfriend, T.; Wishart, D. S. *PLoS One* **2011**, *6*, e16957.

(130) Mandal, R.; Guo, A. C.; Chaudhary, K. K.; Liu, P.; Yallou, F. S.; Dong, E.; Aziat, F.; Wishart, D. S. *Genome Med* **2012**, *4*, 38.

(131) Klepacki, J.; Klawitter, J.; Karimpour-Fard, A.; Thurman, J.; Ingle, G.; Patel, D.; Christians, U. *Clin Biochem* **2016**, *49*, 955-961.

(132) Yet, I.; Menni, C.; Shin, S. Y.; Mangino, M.; Soranzo, N.; Adamski, J.; Suhre, K.; Spector, T. D.; Kastenmuller, G.; Bell, J. T. *PLoS One* **2016**, *11*, e0153672.

(133) Fung, E. N.; Xia, Y. Q.; Aubry, A. F.; Zeng, J.; Olah, T.; Jemal, M. *J Chromatogr B Analyt Technol Biomed Life Sci* **2011**, *879*, 2919-2927.

11. Appendix

Table 11.1. List of 72 tentative metabolite identifications out of 130 features that differed significantly in abundance between CKD patients and healthy controls. Each metabolite is characterized by retention time and the *m/z* value. For each tentative metabolite, a sum formula was calculated and the mass error in mDa and the mSigma value as a measure for the isotopic fit are listed. Wherever possible, a tentative identification was made. Therefore, each potential metabolite was searched in ChEBI. Some metabolites were additionally found in HMDB and METLIN. For some features, two or more assignments were possible. Reprinted from Vogl et al. 2016.[74]

RT [min]:m/z [Da]	Sum formula [M+H+]	Error [mDa]	mSigma value	Metabolite	Identified by
0.7:205.155	C9H21N2O3	0.2	8.0	3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	Metlin/ChEBI
0.83:104.107	C5H14NO	0	2.5		
0.95:203.15	C8H19N4O2	-0.1	10.5	Asymmetric/symmetric dimethylarginine[a,b]*	METLIN/ChEBI/ST
1.08:170.092	C7H12N3O2	0.0	5.8	3-Methylhistidine[a]	ChEBI/ST
1.16:212.103	C9H14N3O3	0.0	8.1	*	
1.21:160.133	C8H18NO2	-0.1	2.6	DL-2-Aminooctanoic acid/N-(hexanoyl)ethanolamine	ChEBI
1.22:189.123	C8H17N2O3	0	7.8	N6-Acetyl-L-lysine/2-Keto-6-acetamidocaproate*	ChEBI/HMDB
1.29:317.0065	C14HN6O4	-0.9	16.0		
1.34:269.124	C11H17N4O4	-0.1	9.8	N-acetylcarnosine	ChEBI
1.45:215.016	C6H8NaO7	0.0	0.3	Citric acid[a]*	ChEBI/HMDB/ METLIN/ST
1.57:169.036	C5H5N4O3	-0.3	11.4	Uric acid[a]*	ChEBI/ST
1.68:223.02	C8H8NaO6	0.9	12.5	Maleylacetoacetic acid/ 4-Fumarylacetoacetic acid	ChEBI
1.7:137.046	C5H5N4O	0.0	5.7	Hypoxanthine[a]	ChEBI/ST
1.84:166.0725	C6H8N5O	0.0	6.3	7-Methylguanine[a]*	ChEBI/ST
1.91:282.12	C11H16N5O4	-0.1	7.8	Methyladenosine[a]*	ChEBI/ST
1.97:150.077	C6H8N5	-0.1	8.8	1-Methyladenine*	ChEBI/METLIN
1.99:136.076	C8H10NO	0.0	1.0	Phenylacetamide/N-phenylacetamide*	ChEBI
2.24:288.119	C11H18N3O6	0.5	13.7		
2.45:204.134	C8H18N3O3	-0.1	4.1	Lys-Gly, Gly-Ala*	ChEBI
3.4:298.114	C11H16N5O5	0.1	5.2	1-Methylguanosine*	ChEBI/METLIN
3.43:167.057	C6H7N4O2	-0.1	6.9	1-Methylxanthine[a]	ChEBI/ST
3.46:283.104	C11H15N4O5	-0.2	9.5	1-Methylinosine[a]	ChEBI/METLIN/ST
3.57:126.091	C7H12NO	0.0	4.3		ChEBI
3.67:232.155	C11H22NO4	-0.7	8.8	O-isobutyrylcarnitine/O-butanoylcarnitine/C4-acylcarnitine/N-heptanoyl-L-homoserine	ChEBI
3.75:281.113	C13H17N2O5	0.0	5.6	Phe-Asp, Asp-Phe*	ChEBI
3.79:247.129	C10H19N2O5	0.0	4.2	Leu-Asp/Asp-Leu/Ile-Asp/Asp-Ile/Val-Glu/Glu-Val	ChEBI
3.87:279.134	C14H19N2O4	0.4	7.0	Pro-Tyr, Tyr-Pro*	ChEBI
3.91:253.118	C12H19N2O4	0.0	2.3		
3.92:248.095	C10H18NO4S	0.1	5.2		
3.94:237.123	C12H17N2O3	-0.3	1.4	Phe -Ala, Ala-Phe	ChEBI
3.94:207.1105	C7H11N8	-0.3	4.0		
3.9:271.1285	C14H16N3O3	0.3	14.2		
4.06:197.067	C7H9N4O3	-0.5	11.2	1,3-dimethyluric acid, 1,7-dimetyluric acid, 3,7-dimethyluric acid	ChEBI
4.14:205.097	C11H13N2O2	-0.1	5.8	Tryptophan[a]*	ChEBI/ST/METLIN
4.2:150.055	C8H8NO2	0.0	5.2	2,3-dihydroxyindole/ 5,6-dihydroxyindole	ChEBI
4.23:181.072	C7H9N4O2	-0.2	1.0	1,7-dimethylxanthine[a]	ChEBI/ST
4.29:301.139	C13H21N2O6	0.1	6.7		
4.34:314.1595	C15H24NO6	0.1	7.1		
4.41:281.113	C13H17N2O5	0.2	12.3	Asp-Phe/Phe-Asp	ChEBI
4.43:166.086	C9H12NO2	-0.1	8.4	Phenylalanin[a]	ChEBI/ST
4.47:244.154	C12H22NO4	-0.2	6.0	Tiglylcarnitine[a]*	ChEBI/ST
4.49:338.134	C15H20N3O6	0.2	17.9	N-acetyl-L-tyrosylglycylglycine	ChEBI
4.54:316.175	C15H26NO6	0.7	6.3	(2E)-7-carboxyhept-2-enoylcarnitine	ChEBI
4.73:197.128	C10H17N2O2	-0.1	4.2		
4.79:333.115	C12H22NaO9	0.1	3.1	α-paratopyranosyl-(1→3)-α-D-mannopyranose	ChEBI
4.97:202.0475	C9H9NNaO3	-0.2	4.9	Hippuric acid[a]	ChEBI/ST
5.1:219.113	C12H15N2O2	0	7.9	N-acetylserotonin/5-methyltryptophan	HMDB/ChEBI
5.1:248.092	C13H14NO4	-0.2	8.4	N-(4-coumaroyl)-L-homoserine lactone	ChEBI
5.17:147.076	C5H11N2O3	0.0	6.2		
5.18:300.18	C14H23N5NaO	-0.8	11.1	erythro-9-(2-hydroxy-3-nonyl)adenine (EHNA)*	ChEBI
5.23:377.142	C14H26NaO10	0.3	2.0	N-acetyl-β-D-glucosaminyl-(1→4)-β-D-glucosamine	ChEBI
5.25:243.134	C11H19N2O4	0.1	3.3		
5.95:266.102	C13H16NO5	0.4	12.5		
6.03:337.175	C17H25N2O5	0.1	2.4		
6.21:272.186	C14H26NO4	0.4	4.3	O-heptenoylcarnitine*	ChEBI
6.34:305.11	C13H18N2NaO5	0.5	13.6	Tyr Thr Na-Adduct*	ChEBI/Metlin
6.65:304.212	C15H30NO5	0.2	5.7	3-hydroxyoctanoyl carnitine*	ChEBI/HMDB
7.02:289.116	C11H13N8O2	0.2	8.8		
7.13:342.227	C18H32NO5	0.6	17.4	*	
7.32:274.201	C14H28NO4	0.5	3.5	O-heptanoylcarnitine[a]	ChEBI/ST
7.55:271.165	C13H23N2O4	0.1	7.2		
8.15:463.157	C21H28NaO10	0.6	5.2		
8.18:365.192	C16H33N2O3S2	0.8	10.3	*	
8.82:475.247	C28H36NaO5	1.3	21.5		
9.23:465.248	C26H33N4O4	1.5	14.1		
9.54:302.233	C16H32NO4	0.1	13.8		
9.69:248.168	C12H26NO2S	0.1	6.2	*	
9.76:302.232	C16H32NO4	-0.1	6.1	*	
10.2:314.232	C17H34NO4	0.3	6.0	*	
11.11:328.248	C18H34NO4	0.1	8.1	(9E)-10-(nitro)octadecenoic acid*	ChEBI
11.22:291.2315	C19H31O12	0.4	14.7	Androsterone*	ChEBI
11.29:431.242	C25H35O6	0.7	1.6	6β-hydroxyprogesterone hemisuccinate*	ChEBI

Table 11.2. Figures of merit of the different tested trainings and test sets employed for the random forest classification. Reprinted from Vogl et al. 2016.[74]

Test	Training set (batch number)	Test set (batch number)	Overall prediction accuracy[a] [%]	NC prediction accuracy[a] [%] (specificity)	TREAT prediction accuracy[a] [%] (sensitivity)	GCKD prediction accuracy[a] [%] (sensitivity)
A	5	3+4	92*	75	100	-
B	5+3	4	87*	73	100	-
C	3+4	5	60*	100	-	40
D	5+4	3	87	75	100	-
E	3	4+5	68	100	100	21
F	3+4+5/2	5/2	100*	100	-	100
G	3+5/2	4+5/2	92*	82	100	100
H	4+5/2	3+5/2	90*	76	100	100

* NC samples of batch 4 were the same NC samples as measured in batch 5 but were measured independently. For the random forest classification duplicate NC samples measured independently within batch 4 and batch 5 were excluded either from the training or test set.
a Random Forest classification was performed three times. Mean values are shown for the prediction accuracies.

Table 11.3. Stock solutions of the metabolites included in the extended quantitative method with the applied solvent and the respective concentration.

AM 1	**solvent**	**c [mM]**
Kynurenic acid	100% H_2O	12.20
DL-Kynurenine	100% H_2O	11.24
Nicotinamide	100% MeOH	46.70
DL-Tryptophan	100% MeOH	2.64
Xanthurenic acid	10% 1M NaOH	56.90
Quinolinic acid	50% MeOH	60.60
Anthranilic acid	50% MeOH	54.00
3-Hydroxyantranilic acid	85% MeOH	9.66
Indole-3-acetic acid	50% MeOH	30.77
Serotonin	50% MeOH	9.12
DL-Indole-3-lactic acid	50% MeOH, 50% 0.1% HCOOH	9.40
Indole-3-propionic acid	50% MeOH, 50% 0.1% HCOOH	45.80
Indole-3-carboxaldehyde	50%MeOH, 48% H_2O, 2% HCOOH	54.70
Indole-3-pyruvic acid	100% MeOH	12.75
Tryptamine	100% MeOH	30.46
Tryptophol	100% MeOH	31.90
Nicotinic acid	100% MeOH	81.55
Melatonin	100% MeOH	20.60
3-Hydroxy-DL-Kynurenine	100% MeOH	4.90
5-Hydroxyindole-3-acetic acid	100% MeOH	15.80
AM 2		
Uric acid	20% 1M NaOH	11.18
3-Indoxylsulfate	0.1% HCOOH	18.15
Creatine	100% H_2O	14.01
Creatinine	50% MeOH	98.10
ß-Pseudouridine	100% H_2O	102.14
Hippuric acid	50% MeOH	64.96
Guanidinoacetic acid	100% H_2O	19.98
c-Mannosyltryptophan	100% H_2O	10.00
Xanthine	30% NH_4OH (28% NH_3 in H2O)	23.50
Hypoxanthine	100% H_2O	5.80
Methylxanthine	20% NH_4OH (28% NH_3 in H2O)	23.40
AM 3		
1-Methylinosine	100% H_2O	4.10
1,7-Dimethylxanthine	100% H_2O	5.72
Tiglylcarnitine	100% H_2O	9.76
Methylguanine	20% MeOH, 5% 1M NaOH	7.71

12. Publications and Presentations

12.1 Publications

1 Sekula P, Dettmer K, **Vogl FC**, Gronwald W, Ellmann L, Mohney RP, Eckardt KU, Suhre K, Kastenmüller G, Oefner PJ, Köttgen A (2017).
From Discovery to Translation: Characterization of C-Mannosyltryptophan and Pseudouridine as Markers of Kidney Function.
Scientific Reports 7(1):17400.

2 Wallmeier J, Samol C, Ellmann L, Zacharias HU, **Vogl FC**, Garcia M, Dettmer K, Oefner PJ, Gronwald W; GCKD Study Investigators (2017).
Quantification of Metabolites by NMR Spectroscopy in the Presence of Protein.
Journal of Proteome Research 16(4):1784-1796.

3 **Vogl FC**, Mehrl S, Heizinger L, Schlecht I, Zacharias HU, Ellmann L, Nürnberger N, Gronwald W, Leitzmann MF, Rossert J, Eckardt KU, Dettmer K, Oefner PJ; GCKD Study Investigators (2016).
Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics.
Analytical and Bioanalytical Chemistry 408(29):8483-8493.

4 Liebherr RB, Hutterer A, Mickert MJ, **Vogl FC**, Beutner A, Lechner A, Hummel H, Gorris HH (2015).
Three-in-one enzyme assay based on single molecule detection in femtoliter arrays.
Analytical and Bioanalytical Chemistry 407(24):7443-52, Epub ahead of print on August 8 2015.

5 Zacharias HU, Hochrein J, **Vogl FC**, Schley G, Mayer F, Jeleazcov C, Eckardt KU, Willam C, Oefner PJ, Gronwald W (2015).
Identification of Plasma Metabolites Prognostic of Acute Kidney Injury after Cardiac Surgery with Cardiopulmonary Bypass.
Journal of Proteome Research 14(7):2897-905. Epub ahead of print on June 12 2015.

6 Dettmer K, **Vogl FC**, Ritter AP, Zhu W, Nürnberger N, Kreutz M, Oefner PJ, Gronwald W, Gottfried E (2013).
Distinct metabolic differences between various human cancer and primary cells.
Electrophoresis 34(19):2836-47. Epub ahead of printing on September 12 2013.

12.2 Book Chapter

Wachsmuth CJ, **Vogl FC**, Oefner PJ, Dettmer K (2013).
Gas chromatographic methods in metabolomics.
In T. Hyötylläinen & S. Wiedmer (Eds.), *Chromatographic methods in metabolomics* (pp- 87-113), Cambridge: RSC Publishing.

12.3 Oral Presentation

Normalization and Batch Correction in LC-MS.
25th Postgraduate Seminar of the GDCh's AK Separation Science, Hohenroda, Germany, 11 January 2015.

12.4 Poster Presentation

Batch Effects in Liquid Chromatography: Setting the Data Straight.
Related Techniques, Amsterdam, The Netherlands, 16-20 June 2013.

13. Summary

With recent instrumental improvements, untargeted high-resolution liquid chromatography mass spectrometry (LC-MS) has been increasingly applied to the analysis of complex biological matrices such as urine. The latter is a preferred matrix in large scale metabolomics studies, because it can be obtained non-invasively in sufficient quantities. However, the widely varying concentrations of urine metabolites, mostly due to different fluid intake, pose analytical challenges, such as ion suppression, detector saturation, column overload or failure to detect low abundant metabolites. In this doctoral thesis, both targeted and untargeted analysis of urine by LC-MS was investigated to address some of these challenges. Further, the question was addressed if untargeted fingerprinting of urine specimens is adequate for the quantitative analysis of these specimens.

First, five different dilution and normalization strategies of spot urinary specimens were compared. Specimens were adjusted either to a uniform creatinine concentration or osmolality without any further normalization of the acquired data, or they were uniformly diluted and post-acquisition normalized to creatinine, osmolality, or sum of all integrals. Spot urine specimens from both an apparently healthy cohort and two different cohorts suffering from chronic kidney disease (CKD) were investigated to test the effect of the various strategies on sample classification. The advantages of pre-acquisition dilution of urine to a fixed creatinine concentration in the acquisition of metabolite fingerprints by LC coupled to a high resolution time-of-flight mass spectrometer (LC-HRTOFMS) were demonstrated. These included the absence of significant correlations of missing values with either the original urine creatinine concentration ($r = 0.216$, $p = 0.133$) or osmolality ($r = - 0.074$, $p = 0.608$). Moreover, pre-acquisition dilution to a uniform creatinine concentration was the only method to correctly assign urine specimens of controls (national cohort, NC) and CKD patients to their respective clusters in hierarchical cluster analysis, while the other dilution and normalization methods yielded indistinct clusters. Therefore, time and labor associated with pre-acquisition dilution to a uniform creatinine concentration are more than justified.

Next, a targeted liquid chromatography triple quadrupole mass spectrometry (LC-QqQMS) method for the quantification of urine metabolites associated with CKD was developed. Some of these metabolites had been identified during the testing of different pre-acquisition dilution and normalization strategies. The dynamic range of the method showed adequate sensitivity for all urine metabolites investigated. The lower limits of quantification (LLOQs) ranged between 0.04 nM to 11 nM except for creatinine (CRE), creatine (CRT), hypoxanthine (HX) and xanthurenic acid (XA), which featured LLOQs of 65.9 nM, 700 nM, 87.8 nM, and 22.0 nM, respectively. Spike-in experiments yielded recoveries of 86% to 121%. The method was applied successfully to the determination CRE, C-mannosyltryptophan (CMT) and pseudouridine (PSU) in urine (Sekula *et al*. 2017, Scientific Reports 7(1): 17400) and to CRE and CRT in plasma (Wallmeier *et al*. 2017, Journal of Proteome Research 16(4): 1784-1796).

Finally, targeted and untargeted measurements of urine specimens were compared with respect to the quantitative performance of untargeted measurements. Combining both targeted and untargeted analysis of urine in a single analysis would constitute a tremendous saving of time and resources. First, calibration curves of selected metabolites were compared, measured in both full scan and MRM mode on a Bruker Maxis Impact quadrupole time-of-flight and a Sciex 4000 QTRAP mass spectrometer, respectively. The LLOQs on the former instrument ranged from 5.49 nM to 1730 nM, whereas the corresponding range on the latter instrument was 1.37 nM to 65.9 nM. The linear ranges were also narrower for the untargeted than the targeted method. Reproducibility over the linear range compared well to the targeted approach. All relative standard deviation (RSD) values for the selected metabolites were below 10%. Next, targeted and untargeted measurements of 200 CKD and 100 non-CKD urine specimens were compared to examine the quantitative performance of the untargeted platform. For the CKD samples, targeted and untargeted measurements of PSU, XA, CMT and tryptophan (TRP) yielded Spearman's rank correlation coefficients ranging from 0.67 to 0.87, while the corresponding coefficients for urine specimens from apparently healthy individuals ranged from 0.73 to 0.95. PSU yielded consistently the weakest correlation coefficient, most likely due to matrix effects as a consequence of its low retention and, thus, co-elution with many other analytes. It is concluded, that the correlation of quantitative data gained by

targeted and non-targeted analysis depends on the instruments and methods used as well as the metabolites and their abundance.

14. Zusammenfassung

Auf Grund von instrumentellen Verbesserungen hat die hochauflösende, nicht zielgerichtete Flüssigchromatographie - Massenspektrometrie (LC-MS Fingerprinting-Analyse) rasant an Bedeutung in der Analyse von komplexen biologischen Matrices wie Urin gewonnen. Urin ist eine sehr interessante Probenmatrix für große Kohorten-Studien, da sie nicht invasiv und in ausreichenden Mengen gesammelt werden kann. Allerdings birgt die Analyse von Urin auch viele Herausforderungen. Die stark variierenden Konzentrationen an Metaboliten, die unter anderem aus unterschiedlicher Flüssigkeitsaufnahme resultieren, führen zum Beispiel zu Ionenunterdrückung, Detektorsättigung, Überladung der Säule oder dem Unvermögen, niedrig abundante Metaboliten zu detektieren. In dieser Doktorarbeit wurden sowohl die zielgerichtete als auch die nicht zielgerichtete Analyse von Urinproben mittels LC-MS untersucht, um einige dieser Herausforderungen zu adressieren. Außerdem wurde der Frage nachgegangen, ob eine nicht zielgerichtete Analyse von Urinproben die gezielte quantitative Analyse von Metaboliten im Urin ersetzen kann.

Zuerst wurden fünf Verdünnungs- und Normalisierungsstrategien für die Analyse von Spontanurin mittels der LC gekoppelt an ein hochauflösendes Flugzeitmassenspektrometer (LC-HRTOFMS) verglichen. Die Proben wurden entweder auf eine konstante Kreatinin-Konzentration beziehungsweise Osmolalität eingestellt, ohne anschließende Normalisierung der aufgenommenen Daten, oder wie es die übliche Praxis ist, einheitlich verdünnt und nach der Messung wurden die Daten auf Kreatinin, Osmolalität oder die Summe über alle Integrale normalisiert. Dabei wurde der Effekt dieser verschiedenen Strategien auf die Probenklassifizierung getestet. Als Proben wurde Spontanurin von einer vermeintlich gesunden Kohorte (Nationale Kohorte, NC) und von zwei verschiedenen Patientenkohorten, die unter chronischer Nierenerkrankung (CKD) litten, verwendet. Hierbei konnten die Vorteile der Verdünnung von Urin vor der Datenaufnahme auf eine konstante Kreatinin-Konzentration für einen metabolischen Fingerprint mittels LC-MS demonstriert werden. Nur die Proben, die auf eine konstante Kreatinin-Konzentration verdünnt wurden, zeigten weder eine Korrelation der Anzahl an fehlenden Messwerten mit der ursprünglichen Kreatinin-Konzentration ($r = 0{,}216$, $p = 0{,}133$) noch mit der Osmolalität ($r = -0{,}074$, $p =$

0,608). Zudem erzielte diese Methode unter allen getesteten Vorgehensweisen prozentual die meisten signifikant diskriminierenden Features zwischen Patienten mit CKD und Kontrollen. Die zweitbeste Methode hinsichtlich der Prozentzahl an diskriminierenden Features war die konstante Verdünnung in Kombination mit der nachfolgenden Normalisierung auf die Summe aller Integrale. Die Überlegenheit der Verdünnung vor der Messung auf eine einheitliche Kreatinin-Konzentration wurde zusätzlich durch eine hierarchische Cluster-Analyse gezeigt. Diese konnte die Kontrollen und die CKD Patienten zuverlässig clustern. Im Gegensatz dazu erzielten die anderen Verdünnungs- und Normalisierungsmethoden keine einheitlichen Cluster. Deshalb kann die einheitliche Verdünnung (z.B. ¼) in Kombination mit der Normalisierung auf Kreatinin oder Osmolalität nicht empfohlen werden. Die Auswirkungen der großen Varianz der Metabolit-Konzentration zwischen den Proben auf Ionenunterdrückung, Detektorsättigung und Überladung der Säule werden durch diese Normalisierungsmethoden nicht ausreichend erfasst. Darüber hinaus ermöglicht die vorgeschlagene Normalisierungsmethode die zuverlässige Probenklassifizierung, was anhand verschiedener Patienten-Kohorten gezeigt wurde**.** Außerdem konnten die meisten differenziellen Features, die CKD Patienten von vermeintlich gesunden Kontrollen unterschieden, bekannten Metaboliten zugeordnet werden. All diese Ergebnisse unterstreichen die überragende Reproduzierbarkeit und Anwendbarkeit der Verdünnung auf eine einheitliche Kreatinin-Konzentration vor der eigentlichen Messung und rechtfertigen den damit verbundenen Zeit- und Arbeitsaufwand.

Ein weiterer Fokus dieser Arbeit lag in der Weiterentwicklung einer LC-QqQMS basierten Methode, die ursprünglich von Zhu et al. (*Analytical and Bioanalytical Chemistry* 401: 3249-3261) für die Bestimmung von Tryptophanmetaboliten beschrieben worden war, für die gezielte Quantifizierung von insgesamt 33 Metaboliten im Urin, welche mit CKD in Verbindung gebracht werden. Der dynamische Bereich der weiterentwickelten Methode zeigte eine angemessene Nachweisempfindlichkeit für alle hinzugefügten Metaboliten in der Analyse von Urinproben. Die Bestimmungsgrenzen (LLOQs) lagen bei 0,04 nM bis 11 nM, außer für Kreatinin (CRE), Kreatin (CRT), Hypoxanthin (HX) und Xanthurensäure (XA), welche eine untere Bestimmungsgrenze von 65,9 nM, 700,0 nM, 87,8 nM bzw. 22,0 nM aufwiesen. Die in einem Spike-in Versuch ermittelten Wiederfin-

dungsraten für einen Teil der in der Methode enthaltenen Metaboliten lagen zwischen 86% und 121%. Die Methode wurde erfolgreich für die Quantifizierung von Kreatinin, C-Mannosyltrptophan (CMT) und Pseudouridin (PSU) in Urin (Sekula et al. 2017, *Scientific Reports* 7(1): 17400) sowie von Kreatinin und Kreatin in Plasma (Wallmeier et al. 2017, *Journal of Proteome Research* 16(4): 1784-1796) verwendet.

Zuletzt wurden zielgerichtete und nicht zielgerichtete Messungen von Urinproben in Hinblick auf die quantitative Leistungsfähigkeit der nicht zielgerichteten Messungen verglichen. Es wäre ein enormer Vorteil, wenn nicht zielgerichtete Analysen einer Probe auch eine zuverlässige Quantifizierung der identifizierten Metaboliten erlauben würden. Dafür wurden zuerst die Kalibrationskurven von ausgewählten Metaboliten für die beiden Plattformen verglichen. Es wurde gezeigt, dass der lineare Bereich der beiden Plattformen gut übereinstimmt. Die unteren Bestimmungsgrenzen (LLOQ) lagen zwischen 5,49 nM und 1.730 nM, wohingegen jene für die gezielten Analysen mit 1,37 nM bis 65,9 nM erwartungsgemäß niedriger waren. Für alle ausgewählten Metaboliten waren die Größenordnungen der linearen Bereiche der nicht-zielgerichteten Methode niedriger im Vergleich zur zielgerichteten Methode, wodurch die höhere Sensitivität der zielgerichteten Methode aufgezeigt wurde. Die Reproduzierbarkeit im linearen Bereich war mit relativen Standardabweichungen von <10% im Vergleich zur gezielten Herangehensweise sehr gut. Als nächstes wurden zielgerichtete und nicht zielgerichtete Messungen von 200 CKD und 100 nicht-CKD Urinproben verglichen, um die quantitative Leistungsfähigkeit der nicht zielgerichteten Plattform zu untersuchen. Die zielgerichteten und nicht zielgerichteten Messungen von PSU, CMT, XA und Tryptophan (TRP) in den CKD Urinproben zeigten eine gute Übereinstimmung mit Spearman‘schen Rangkorrelationskoeffizienten (SCC) von 0,67 bis 0,87. Für Proben gesunder Probanden wurde eine noch bessere Korrelation zwischen der zielgerichteten und nicht zielgerichteten Plattform festgestellt; die Rangkorrelationskoeffizienten für PSU, CMT, XA und TRP betrugen 0,73, 0,95, 0,94 und 0,90. Insgesamt kann festgehalten werden, dass die Korrelation der zielgerichteten und nicht zielgerichteten Plattformen sicherlich abhängig von den angewendeten Instrumenten und Methoden sowie den Metaboliten und deren Konzentration ist. Daher ist eine Verallgemeinerung der Resultate nur sehr eingeschränkt möglich.