Master`s Thesis

Master`s degree in Biomedicine

May 2021

A study of quantitative and qualitative performance of metabolomics compared with targeted analyses used in laboratory diagnostics

Marielle Storbråten

Course code: MABIO5900

60 ECTS

Faculty of Health Sciences
OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Oslo University Hospital

A study of quantitative and qualitative performance of metabolomics compared with targeted analyses used in laboratory diagnostics

By

Marielle Storbråten

A thesis submitted for degree of

Master in Biomedicine, 60 ECTS

Oslo Metropolitan University (OsloMet)

Faculty of Health Sciences

Department of Life Science and Health

Conducted at the Department of Medical Biochemistry at

Rikshospitalet – Oslo University Hospital, Oslo

Supervisor: Katja Benedikte Prestø Elgstøen

Co-supervisors: Per Ola Rønning and Hanne Bendiksen Skogvold

May 2021

OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Oslo University Hospital

Acknowledgement

This master´s thesis was performed at the National Unit for Screening and Diagnosis of Congenital Pediatric Metabolic Disorders at Oslo University Hospital – Rikshospitalet from August 2020 to May 2021. My supervisors were Katja Benedikte Prestø Elgstøen and Hanne Bendiksen Skogvold at Rikshospitalet, and Per Ola Rønning at the Department of Mechanical, Electronic and Chemical Engineering, at Oslo Metropolitan University.

First, I would like to thank Katja for excellent supervising and for sharing useful knowledge every week. The weekly meetings and discussions have been highly valuable to organize all my thoughts. It has been a delight being a part of her research group. Thank you for welcoming me so well, I have always felt so comfortable during this time.

Thank you, Hanne for great supervising and for always being positive, taking time to answering my questions, and all the constructive feedback. Elise and Anja also deserves a special thank you, for all the insights and discussions. Also, thank you Elise and my fellow student Tharsana for teaching me all the powerful tools of Excel and FreeStyle which made my studies easier ☺. Thank you Anja for the time helping me finding all samples and concentrations that I did not have access to. Thank you to everyone at IKB for all conversations and social lunches and a lot of laughter.

Thank you, Per Ola for introducing me for this fun and exciting project and for all the Zoom-calls critically reading this thesis with me, correcting my typos, and the constructive feedback. A huge thank you to Helge Rootwelt for the medical and clinical perspectives and his excitement and feedback on this project. I really have appreciated it.

Finally, I would like to thank my family and friends, especially my fiancé Eirik, for always being supportive, positive, and for always taking good care of me. Marte Eikenes for all the late nights of writing. Mamma and Kiara also deserve a special thanks, for bringing homemade, delicious buns home to me during a stressful week, and endless love.

Marielle Stobråten

Oslo, May 16th, 2021

Abstract

Today, several specific and targeted analyses are used for detection of different chemical compounds and diagnostics of inborn errors of metabolism (IEMs). Metabolomics, an untargeted analytical approach, has been used as a supplement to established targeted analytical approaches in diagnostic laboratories for more effective diagnostics of IEMs. However, little is known regarding the quantitative performance of metabolomics, and a higher level of confidence is needed for metabolite identification. A metabolomics method that is well suited for easily accessible biological materials e.g. urine and plasma is an advantage. The aim of this study was to evaluate the quantitative and qualitative performance of the established metabolomics method HPLC – ESI – Q Exactive Orbitrap MS studying amino acids and amino acid related compounds, diagnostic markers, in urine and plasma samples. Almost all the selected compounds were identified with level 1 identification, the highest level of confidence, using the in-house library that was created in this thesis. The in-house library will be used for identification of amino acids and related compounds in future metabolomics studies. The metabolomics method is a useful diagnostic tool, as the diagnostic metabolite PEA, which is not quantified by established targeted methods in the diagnostic laboratory of IEMs at Oslo University Hospital – Rikshospitalet, was successfully detected. The precision of the metabolomics method was good for most of the selected compounds in urine and in plasma (retention time <2%, and peak area <30% RSD). Urinary metabolites were normalized to creatinine concentration, and the correlation between amino acid peak area from metabolomics and measured concentration ratios from targeted analyses was found. The correlation was better in samples diluted to 0.1 mM creatinine than to higher creatinine concentrations, and approximately linear. This correlation was also found comparing diagnostic markers for cystinuria in normal urine samples to a pathologic sample, and in samples spiked with different concentrations of PEA representing normal samples and HPP. However the qualitative performance of the metabolomics method was better when urine samples were diluted to 2 mM creatinine. Furthermore, several of the selected compounds were affected by chromatographic matrix effects, ion suppression or ion enhancement in plasma. Less variation was found in plasma samples than in urine samples. This study showed that metabolomics is a useful tool for diagnostics of IEMs, e.g. HPP and cystinuria, but it is important to remember that at higher matrix concentrations, matrix effects can cause false positive and negative results.

Sammendrag

I dag brukes flere spesifikke målrettede analysemetoder for deteksjon av ulike kjemiske forbindelser og diagnostikk av medfødte metabolske sykdommer. Metabolomikk – en ikke-målrettet analytisk metode har vært et nyttig supplement for en mer effektiv diagnostisering av medfødte metabolske sykdommer. Lite er kjent for hvor kvantitativ metabolomikk kan være, og en høyere grad av sikkerhet er nødvendig for identifisering av metabolitter. En metabolomikkmetode som er godt egnet for studier av urin og plasma er en fordel da disse biologiske materialene er lett tilgjengelige. Hensikten med denne studien var å evaluere den kvantitative og kvalitative evnen til metabolomikkplattformen for studier av aminosyrer og aminosyre-liknende forbindelser, som er diagnostiske markører, i urin og plasma. Omtrent alle de utvalgte forbindelsene ble identifisert med nivå 1 identifisering, det høyeste nivået av sikkerhet, ved bruk av internbiblioteket som ble laget. Dette biblioteket vil bli brukt for identifisering av aminosyrer og aminosyre-liknende forbindelser i fremtidige metabolomikkstudier. Metabolomikkmetoden er et nyttig diagnostisk verktøy da den diagnostiske markøren fosfoetanolamin (PEA), som ikke kvantiteres med de målrettede metodene som brukes på Oslo Universitetssykehus – Rikshospitalet i dag, ble detektert. Metabolomikkmetoden viste god presisjon for de fleste av de studerte forbindelsene i urin og plasma (retensjonstid <2 %, og toppareal <30 % RSD). Metabolitter i urin ble normalisert til kreatininkonsentrasjonen og sammenhengen mellom aminosyrenes toppareal fra metabolomikk og konsentrasjonsratioene fra målrettede analysemetoder ble funnet. Sammenhengen var bedre når urinprøvene ble fortynnet til 0.1 mM enn 2 mM kreatinin. Denne sammenhengen ble også funnet for de diagnostiske markørene til cystinuri i normale urinprøver og en patologisk prøve, og for prøver som ble spiket med ulike konsentrasjoner av PEA for å representere normale og HPP-prøver. Den kvalitative evnen til metabolomikkplattformen var bedre når urinprøver ble normalisert til 2 mM kreatinin. Videre viste det seg at flere av forbindelsene ble påvirket av kromatografiske matrikseffekter, ionesuppresjon, eller ioneforhøyelse i plasma, men at det var mindre variasjon i plasmaprøver enn i urinprøver. Denne studien viste at metabolomikk er et nyttig verktøy i diagnostikk av medfødte metabolske sykdommer, som HPP og cystinuri, men at for lite fortynnede matrikser kan gi falske positive og negative resultater.

Table of content

Abbreviations

Abbreviation	Term
A	Peak area
APCI	Atmospheric Pressure Chemical Ionization
APPI	Atmospheric Photoionization
C18	Octadecyl Carbon Chain 18
C8	Octadecyl Carbon Chain 8
CAWG	Chemical Analysis Working Group
CT	Helical Computed Tomography
D	Dextrorotatory
Da	Dalton
DBS	Dry Blood Spot(s)
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
DNA	Deoxyribonucleic Acid
EI	Electron Ionization
EIC	Extracted Ion Chromatogram
EMA	European Medicines Agency
ESI	Electrospray Ionization
FT-ICR	Fourier Transform Ion Cyclotron Resonance
GC-MS	Gas Chromatography-Mass Spectrometry
HCD	Higher energy Collision-induced Dissociation
hmdb	human metabolome database
HPLC	High Performance Liquid Chromatography
HPP	Hypophosphatasia
IEMs	Inborn Errors of Metabolism
L	Levorotary
LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
LC-MS/MS	Liquid Chromatography tandem Mass Spectrometry
LLOQ	Lower Level of Quantification

Log P	Partition coefficient P
M	Molecular ion
m/z	mass-to-charge ratio
MALDI	Matrix-Assisted Laser Desorption Ionization
METLIN	The Metabolite and Chemical Entity Database
MF	Matrix Factor
MRM	Multiple Reaction Monitoring
mRNA	messenger Ribonucleic Acid
MS	Mass Spectrometry
MSI	Metabolomics Standards Initiative
MSTUS	MS total useful signal
NBS	Newborn Screening
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
PEA	Phosphoethanolamine
PLP	Pyridoxal-5-Phosphate
PP_i	Inorganic Pyrophosphate
ppm	parts per million
Q-TOF	Quadrupole Time-of-Flight
R	Resolving power
RSD	Relative Standard Deviation
SD	Standard Deviation
TIC	Total Ion Chromatogram
TNSALP	Tissue Non-Specific Alkaline Phosphatase
UHPLC	Ultra-High-Performance Liquid Chromatography
V	Volt
$W_{0.5}$	The width of the peak at half height
WHO	Worlds Health Organization
ΔRT	Retention Time shift

1. Introduction

1.1 Metabolites and metabolism

Metabolites are small chemical compounds (usually <1500 Da) e.g. carbohydrates, nucleotides, lipids, and amino acids that are the starting point, intermediates in, or end points of a diverse range of biochemical reactions and pathways. Many are also used for energy production in the cells (1). These are metabolic processes catalyzed by various enzymes regulated by specific genes. Metabolic processes are highly regulated, but defects in genes or enzymes involved in these processes may cause diseases and, in worst case scenario mortality in humans at a young age if not treated (2, 3). The metabolome can be defined as all the metabolites present in a cell, or in a biological sample. Metabolomics is biochemical profiling, and the study of all the metabolites in a biological sample (4).

1.1.1 Amino acids and related compounds

Amino acids are building blocks in proteins and peptides, and some are precursors to biomolecules like lipids, nucleic acids, and hormones. Amino acids are chiral compounds that typically consist of an amino group, a carboxylic acid group, and a functional group (R group). As such, amino acids can exist as enantiomers, called levorotatory (L) isomers or dextrorotatory (D) isomers, whereas proteins only consist of L isomers (5, 6). Furthermore, amino acids can be classified by the localization of the amino-functional group dependent on which carbon atom the amino group is attached to, e.g. α-, β-, γ-, or δ-amino acids. Other classifications of amino acids can be related to polarity, the types of chemical groups in their side chains, pH level, or they can be different isoforms (7). The ionization state of amino acids varies with pH according to the pK_a value and the partition Coefficient P (Log P) value, which describes the polarity and thereby affects the solubility in water (5, 8).

Essential amino acids cannot be generated in the body and must be supplied by the diet. Nonessential amino acids can be synthesized *de novo* in adequate amounts in the body with optimal requirements and are encoded by the genetic code. There are 22 amino acids involved in protein synthesis, but there are other amino acids with regulatory importance e.g. regulating gene expression, transcription, translation, post-transcriptional modifications, protein turnover,

reproduction, cell signaling, or they may have immune functions (9). Some amino acids are also involved in synthesis of hormones. One example is phenylalanine, which breaks down to tyrosine and further to dopamine, melanin and thyroxine (10). Chemical compounds with almost the same properties as amino acids, amino acid related compounds, are often studied in the same way as amino acids. Creatinine is an endogenous amino acid derivative, and an example of such an amino acid related compound.

Creatinine

Creatinine, a spontaneous biproduct in creatine metabolism in muscles and kidney is excreted in urine and plasma. Creatinine has several applications in clinical chemistry. There are some factors that can affect the creatinine concentration e.g. muscle mass, gender, diet, age, physical activity, and menstrual cycle. Since urinary creatinine excretion rates are relatively constant for an individual, creatinine has been used as a marker to reflect the overall concentration of urinary metabolites for normalization of data obtained from urine samples (11-13). Creatinine is also used as a diagnostic serum marker for acute kidney injury in serum, as well as for chronic renal failure (14).

1.2 Inborn errors of metabolism

A protein is encoded by genes that are segments of hereditary deoxyribonucleic acid (DNA) and where genetic information is stored and decoded to transcribe messenger ribonucleic acid (mRNA). mRNA is translated into polypeptides that subsequently form proteins (15). Inborn errors of metabolism (IEMs) are genetic diseases where the production or the function of an enzyme in a biochemical pathway, transport protein, or other functional protein is defect (16). This is the result of DNA mutations (17). In such metabolic diseases intermediate metabolites may pile up to toxic levels, or downstream metabolites are not produced in sufficient quantities. **Figure 1. 1** is a hypothetical metabolic pathway, and shows that a defective enzyme from a mutated gene leads to accumulation of the substrate (D) and lower levels of an important metabolite product (E). Furthermore, the accumulation of D might lead to new metabolites (D´) that might also be toxic. The arrows illustrate enzymes.

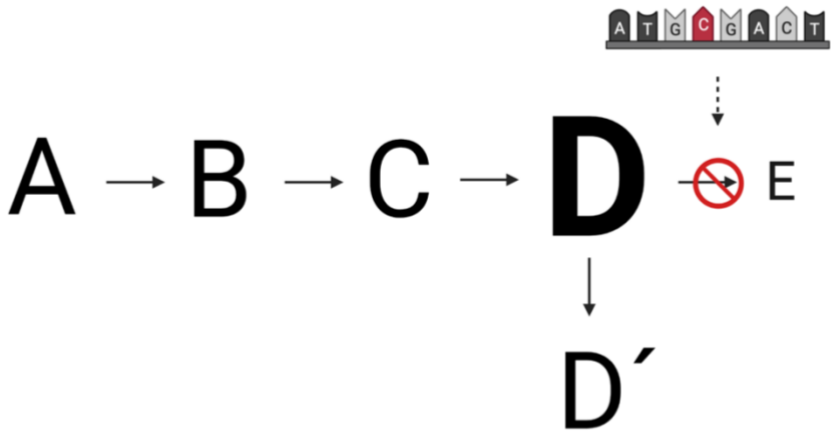


Figure 1. 1: Hypothetical metabolic pathway with an enzyme defect. *The figure is an illustration of a hypothetical metabolic pathway occurring when an enzyme (arrows) converts one metabolite to another in a series of biochemical reactions. A genetic disease can lead to a defect enzyme and can lead to accumulation of upstream metabolite(s) (D) in the biochemical pathway, and low levels of downstream metabolite(s) (E).*

Over 1000 IEMs have been identified (18) and the cumulative incidence has been estimated to be 1 in 2500 (19), which makes them collectively common and cause a significant level of pediatric morbidity and mortality, even though each IEM is individually rare. Hypophosphatasia (HPP) and cystinuria are two IEMs that have been characterized with specific biomarkers identified (2). There are two analytical approaches used for identification of diagnostic biomarkers or metabolites. These are targeted analyses and metabolomics ("untargeted" or "global" analytical approach). Targeted analyses are used in clinical laboratories to measure and quantify one or a few specific metabolites, while metabolomics aim to detect and identify as many metabolites as possible in a biological sample, both known and unknown compounds (20).

1.2.1 Hypophosphatasia

HPP is a heterogeneous disease that can occur at birth in young infants as a life threatening disease, or in young adults. The clinical symptoms vary with bone mineralization, seizures and hypercalcemia in infants, and premature exfoliation of the teeth in younger adults. HPP is caused by mutations in the ALPL gene, which encodes the enzyme tissue non-specific alkaline phosphatase (TNSALP), resulting in decreased enzymatic activity of TNSALP and thereby increased levels of its substrates inorganic pyrophosphate (PP_i), pyridoxal-5-phosphate (PLP)

and phosphoethanolamine (PEA), a polyprotic acid and derivative of the amino acid ethanolamine (21).

1.2.2 Cystinuria

Cystinuria, an autosomal recessive disorder, can become clinically evident at any age. It is caused by a transepithelial transport defect of the di-basic amino acids cystine, lysine, ornithine, methionine and arginine, resulting in an increased excretion of these amino acids and crystal formation of cystine in urine. Therefore, the symptoms of cystinuria are related to stone formation, and include pain, hematuria, lower urinary tract symptoms, and urinary tract infection. Symptoms often appear in childhood or early adulthood (22-24).

1.3 Diagnostic practice for inborn errors of metabolism

The section for IEMs located at the Department of Medical Biochemistry at the Division of Laboratory Medicine at Oslo University Hospital – Rikshospitalet is a part of the Norwegian National Unit for Screening and Diagnostics of Congenital Pediatric Metabolic Disorders. Urine, blood, and cerebrospinal fluid samples are studied to detect these diseases (25).

The outlook for the patient with a diagnosed IEM depends on early diagnosis and quick institution of effective treatment (26, 27). A quick diagnosis can therefore prevent the development of serious illness and initiate early treatment with right medication(s) and specific diet. Some IEMs may not show any signs until later in the childhood or adulthood. Therefore, several IEMs are included in Newborn screening (NBS) programs (27, 28). The Unit for Newborn Screening at Oslo University hospital – Rikshospitalet is a national service for screening of newborns in Norway, where all newborns in Norway are screened for 25 IEMs since 2012 (29). Newborn Screening aims to identify apparently healthy infants with severe congenital disorders that can be partly prevented and followed up regularly or treated by identifying metabolites in abnormally high or low concentrations in a dried blood spot (DBS). The IEMs that are included in this program are determined by The World Health Organization (WHO) Wilson-Junger criteria (28, 30). If other IEMs are suspected, patients are diagnosed at the section for IEMs at Oslo University hospital – Rikshospitalet (31).

Diagnostics of IEMs can be challenging because of the non-specific list of clinical symptoms. The diagnostics depend on clinical observations and phenotypes, laboratory analyses,

radiological imaging, and family history, which will narrow down the long list of possible IEMs. A series of several investigations are performed, which may range from basic clinical tests to more complex diagnostic methods at gene, protein, or metabolite level (32).

Clinical examination and initial analyses, in addition to family history and similar disorders in close family or neonatal deaths, can give an indication of an IEM. The clinical examination includes hepatosplenomegaly, skin lesions, and neurologic deficits to name a few on a very long list. General laboratory investigations are performed, e.g. complete blood cell count, comprehensive metabolic panel like liver and kidney function, electrolytes and uric acid, or urine pH, color, odor, specific gravity, ketones and urine-reducing substances. If the clinical symptoms are non-specific a range of screening analyses and a selection of more advanced diagnostic methods are used (27, 32).

Analysis of metabolites and the metabolome in biological samples e.g. urine, plasma, and cerebrospinal fluid is an important diagnostic tool. Amino acid analysis in several biofluids, organic acids in urine, carnitines and acylcarnitines in DBS, plasma or urine are some of the groups of chemical compounds that are analyzed. In addition to clinical testing and metabolite detection, enzyme analysis on the tissue in which the enzyme is expressed is often used to validate metabolic disorders. Different and specific targeted analyses are used to analyze all the different classes of chemical compounds representing the metabolome, and to quantitatively detect the concentration and the presence of metabolites used as biomarkers. The different analytes and analyses require different sample preparation methods. The quantitatively measured concentrations are compared to established reference ranges (33). Lately, metabolomics is incorporated as a diagnostic tool for screening of multiple metabolites in a sample and thus facilitates more effective diagnostics (32, 34). **Figure 1. 2** illustrates the difference between metabolomics and targeted analytical approaches for studies of the metabolome. While targeted analyses include selection of appropriate biomarkers and sample preparation for analysis, metabolomics is a screening method for metabolites where the total ion chromatogram (TIC) represents the whole metabolome, and extracted ion chromatograms (EICs) for specific metabolites are filtered out based on the exact mass-to-charge ratio (*m/z*).

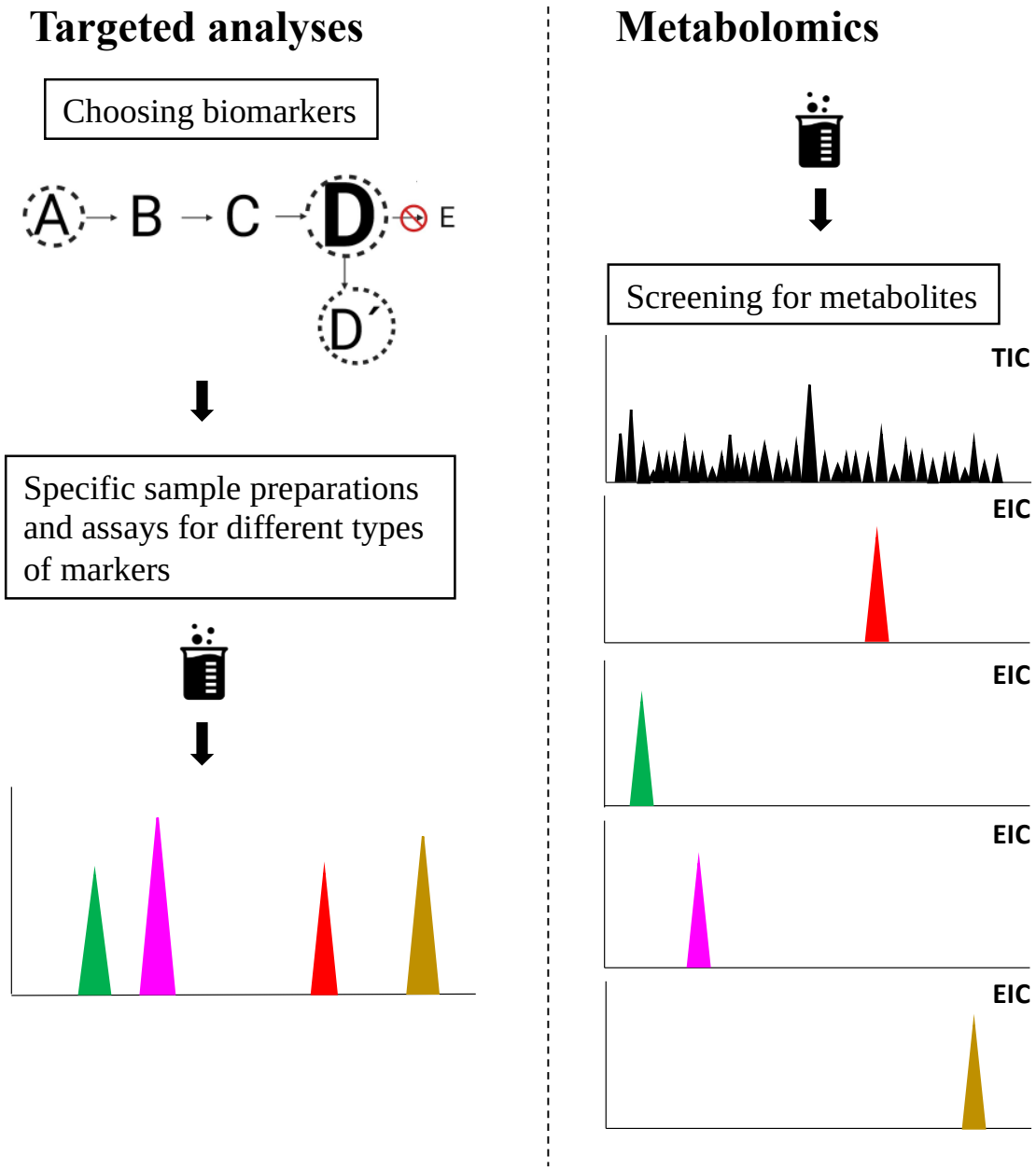


Figure 1. 2: Metabolomics versus targeted analyses for studies of the metabolome. *An illustration of the difference between metabolomics and targeted analyses for studies of the metabolome. Targeted analyses include selecting the appropriate biomarkers and sample preparation for diagnostics and analysis. Metabolomics is a screening method for metabolites. The TIC represents signals from the whole metabolome and EIC for specific metabolites are filtered out based on the exact m/z.*

For many diseases, genotyping of the patient is important to get information about the alleles for a specific gene, and to find mutations. This helps to diagnose the patient and is supported by the clinical symptoms and biochemical findings. Genetic alterations that can cause IEMs are missense or nonsense mutations, and small or large deletions or duplications. Different

polymerase chain reaction (PCR)-based techniques, DNA sequencing by capillary electrophoresis and next generation sequencing techniques are used to detect different types of mutations (32).

1.3.1 Diagnostics of hypophosphatasia

HPP can be diagnosed by the occurrence of a genetic defect in the ALPL gene by sequencing. The diagnostic markers for HPP can also be studied to confirm the diagnosis. These are reduced levels of TNALP in blood and elevated levels of PEA (21). However, the method and equipment used today at Oslo University Hospital – Rikshospitalet cannot detect PEA and is therefore not used as a diagnostic marker. The urinary excretion of PEA depends on age, and there is a higher excretion in newborns than in children from 3 years and older (35). Individuals with HPP have been reported with PEA values >100 μmol/mmol creatinine (36, 37).

1.3.2 Diagnostics of cystinuria

Family history of cystinuria and the presence of cystine crystals in urine raise suspicion for cystinuria. The diagnosis is established by analysis of cystine crystals in urine by helical computed tomography (CT) or ultrasound, and a qualitative cyanide-nitroprusside test. The diagnostic biomarkers for cystinuria are elevated levels of cystine, lysine, ornithine, methionine, and arginine in blood and urine, and are used to confirm the diagnosis (22, 24).

1.3.3 Targeted methods in clinical laboratories

There are a variety of analytical techniques used in targeted analysis, including liquid chromatography tandem mass spectrometry (LC-MS/MS), gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis, ion pair chromatography, and nuclear magnetic resonance (NMR) dependent on the analyte of interest. To assess the most suitable technique relies on correct selection of specific analytes for the individual patient, which depends on relevant clinical information provided by the clinician to the metabolic laboratory. There is a risk of false negatives due to the incomplete description of the symptoms of the patient. In targeted analyses for IEMs the goal is to provide quantitative information of the selected metabolites. Therefore, measured concentrations are compared to established reference ranges that are used in conjunction with formal calibration, validation, internal standards and quality controls, to provide good sensitivity and specificity (38-40). However, these analytical

techniques give little information of other parts of the metabolome, the biochemical reaction paths and how these are affected (33).

1.4 Metabolomics – the untargeted analytical approach

Metabolomics is the study and discovery of all metabolites in a biological sample that are <1500 Da, and is a relatively new "omics" strategy in biomedical research for discovery of diagnostic biomarkers for diseases like IEMs (40, 41). Metabolomics can be referred to as biochemical profiling, global or untargeted analytical approach (4). This analytical method gives a metabolic profile and makes it possible to study how other factors affect the metabolome, generating a fingerprint characteristic for a biological sample. Gertsman et al. (33) describe metabolomics as a "discovery mode process" for research because the bottleneck of metabolomics is to identify known, but also unknown features.

As Kim, S.J., et. al. reviewed (40), different techniques is used in metabolomics e.g. LC-MS, GC-MS, and capillary electrophoresis/MS. For small and volatile metabolites GC-MS is a preferred method. However, GC-MS requires thermally stable analytes and derivatization introduces variability and derivatization artifacts. LC is the dominating separation technique for reduction of complexity prior to detection. Therefore, a broader range of polarity and a broader range of metabolites can be identified (42-44). LC-MS gives high sensitivity, increased versatility and a higher metabolite coverage to be identified, compared to NMR used in targeted analyses (45). For maximum coverage, optimal sample preparation and high resolution mass spectrometer e.g. Orbitrap MS and statistical and bioinformatic analysis is crucial. This allows the accurate mass of metabolites to be determined. It is not yet established to which extent quantitative analyses can be performed using metabolomics (40, 45).

1.4.1 Metabolomics workflow

The metabolomics workflow includes sample preparation of biological samples, data acquisition, data processing, metabolite identification, and data interpretation. **Figure 1. 3** is an illustration of how the metabolomics workflow can look like.

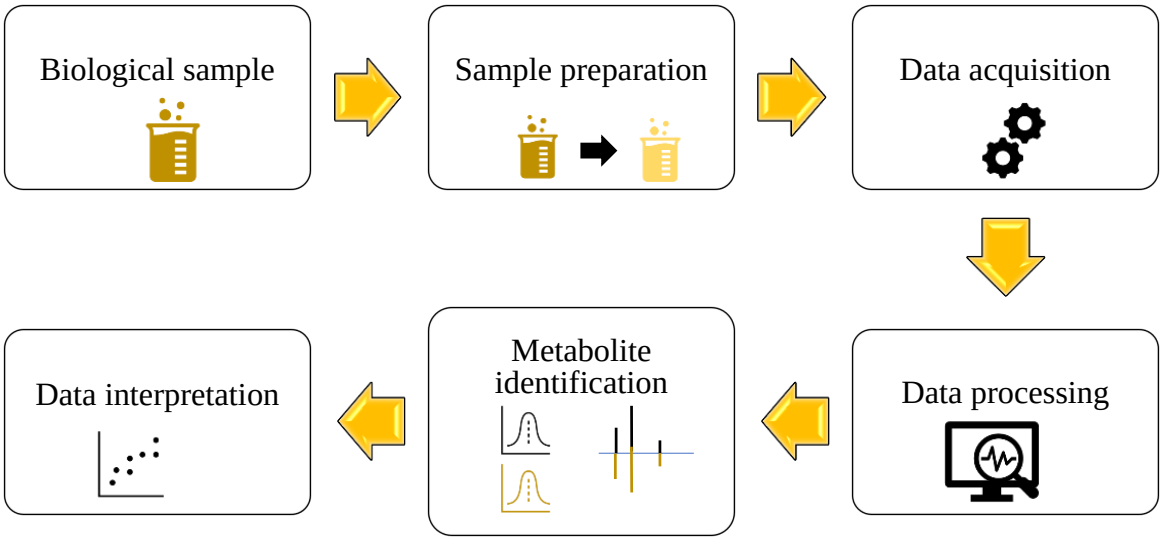


Figure 1. 3: Metabolomics workflow. *The figure illustrates what the workflow in a metabolomics study can look like. A sample preparation of the biological sample is done before data acquisition on the instrument, data is processed, metabolites are identified or detected, and lastly data is interpretated.*

Sample preparation is performed to minimize experimental error and for removal of compounds that can interfere with signals from metabolites in the sample. The extraction method is chosen in accordance with the analytical platform, and the sample preparation varies with the biological sample studied. In many cases, data acquisition starts with chromatographic separation of metabolites, before detection. The massive amount of data collected in metabolomics is then processed in software programs, and is also corrected for mass and retention time shift. Depending on the aim of study, there are a variety of statistical methods for data processing that can be used in the discrimination process and classification of metabolites. For better level of confidence in metabolite identification, the signal peaks are, if possible, aligned to chemical standards and fragmentation spectra from an in-house library. Lastly, there is data interpretation that often includes finding a correlation between data, or significant differences in the metabolome of two or more groups of individuals (46).

1.4.2 Liquid Chromatography-Mass Spectrometry

LC-MS has become a common technique for data acquisition in metabolomics studies so that a broad range of metabolites can be detected (47, 48). The high-efficiency separations with LC coupled to the sensitive and selective MS as a detector with accurate mass measurements provides high metabolite coverage (33, 49).

Liquid Chromatography

In LC, the equilibrium of analyte between the mobile phase and stationary phase leads to separation based on the chemical properties e.g. solubility, polarity and structure of the analyte and determines the retention time. The retention time is the time an analyte needs to travel through the chromatographic column and is one of the parameters that can be used for identification of metabolites. Other factors e.g. mobile phase, flow rate, column dimensions, back pressure, particle technology, particle size, and temperature can affect the chromatographic separation. The column material is packed into a chromatographic column. Pentaflourodiphenyl, octadecyl carbon chain 18 (C18), octadecyl carbon chain 8 (C8), and C18-bonded silica particles are examples of reverse phase colum materials, while normal phase column materials often contain silica, amino or cyano groups (33, 50-52). Separated metabolites from the chromatographic column are presented in a chromatogram, a graph where the detector response is plotted against the retention time in minutes. The detector response, or signal peak data are assumed to be Gaussian normally distributed and can therefore be characterized by the mean and standard deviation (SD) (33, 47, 53, 54).

High Performance Liquid Chromatography

High performance liquid chromatography (HPLC) uses high pressure to force the mobile phase through the column that contains fine particles and gives high resolution separation. The HPLC system consists of a sample injection valve, a pump system, a chromatographic column, a degasser and a detector. **Figure 1. 4** is an illustration of the HPLC system.

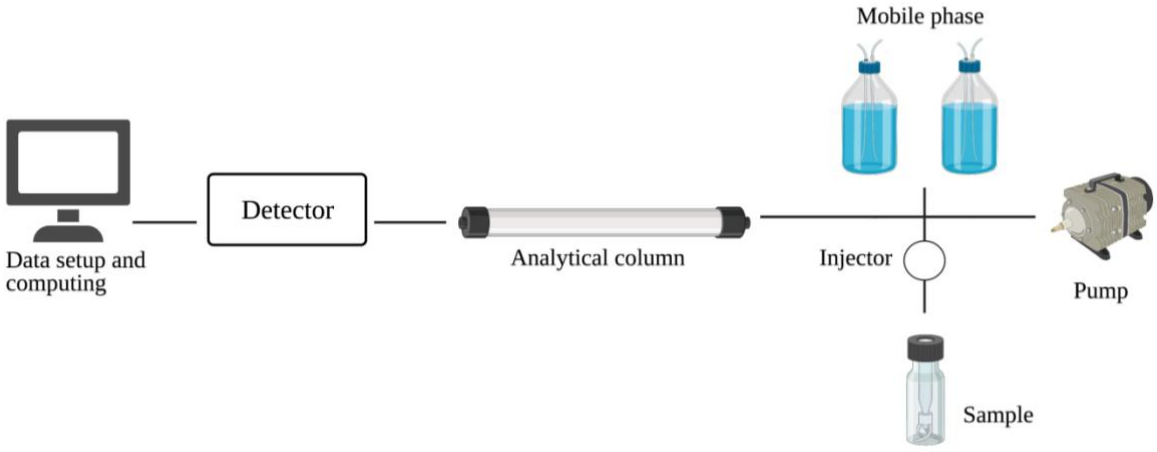


Figure 1. 4: The high performance liquid chromatography system. *The HPLC system consists of a sample injection valve, a pump system, an analytical column, a degasser, detector and a computer for data setup and processing.*

The sample is injected through a sample injection valve, blends into the mobile phase and is pumped by a pumping system throughout the system with constant flow. The mobile phase leads the sample into a chromatographic column for separation. For increased separation efficiency and separation of complex samples, a gradient elution can be used by changing the mobile phase composition throughout the analysis. After separation, the chemical compounds are sent to the detector and the signal is computed. A degasser in the HPLC system removes dissolved gas like oxygen in the mobile phase to prevent reaction with the mobile or stationary phase (53, 55, 56).

Manufacturers usually offer chromatographic columns with an inner diameter of 1-5 mm, column length between 3-30 cm, and typical particle size 1.7-5 μm in diameter. The efficiency of the column increases as the size of the stationary phase particle decreases. The entrance to the main part of the column is protected with a guard column containing the same stationary phase as the main column, and increases the lifetime of the column. Stationary material depends on the components it is desired to separate, but most used HPLC technique is reversed phase LC with non-polar stationary phase material and polar mobile phase. Ultra-high-performance liquid chromatography (UHPLC) provides better chromatographic separation because the particle diameter can be less than 2 μm (53). For highly polar and ionic compounds hydrophilic interaction liquid chromatography (HILIC) is better suited. However, the disadvantage of HILIC is the high consumption of organic solvent like acetonitrile, and unsuitable performance. It has been reported that UHPLC therefore is a good alternative (45, 56-58).

Mass Spectrometry

MS is a technique that is used to study the mass of atoms, molecules or molecular fragments, and is one of the most powerful detectors. It is often used together with LC because it gives both qualitative and might give quantitative information with high sensitivity. Based on the exact mass, or the *m/z* it is possible to separate different substances with the same mass, even though they have the same chromatographic retention time. There are several different types of MSs, but mainly it consists of a sample inlet, ion source, a mass analyzer, computer, and a detector, as illustrated in **Figure 1. 5**.

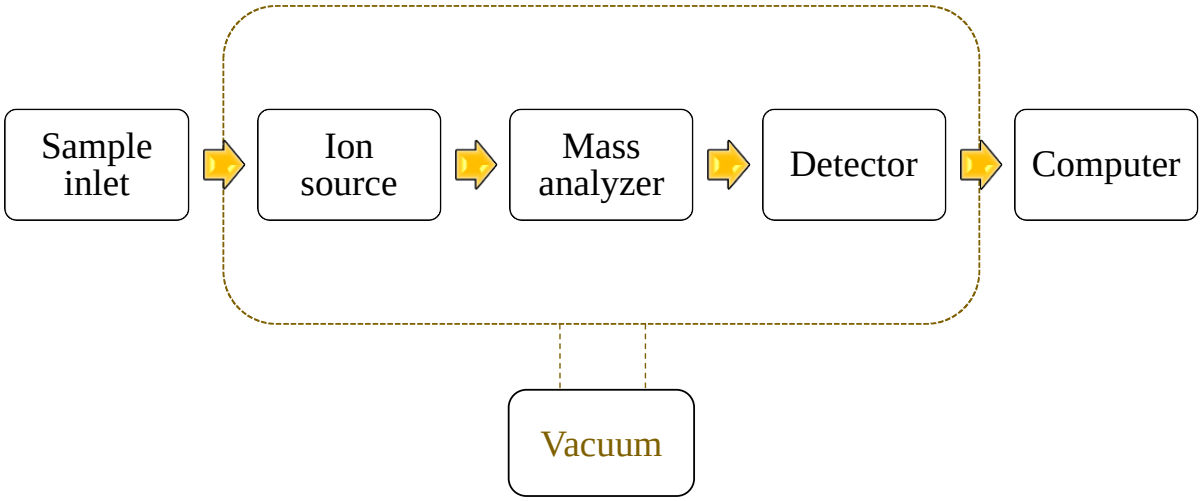


Figure 1. 5: The main components in a mass spectrometer. *Molecules or substances enter the mass spectrometer in the sample inlet into the ion source. In the ion source the molecular ion is formed, before they are filtered in the mass analyzer by vacuum based on their m/z. The filtered ions that have passed the mass filter are detected.*

Molecules enter the MS in a sample inlet, into the ion source where a charge is applied to them by exposure to high energic electrons accelerated through a high electrical potential. The electrical potential causes instability of electrons in the molecule, and results in formation of a cation with one unpaired electron, called the molecular ion ($M^{\bullet+}$) (53). In the entrance to the mass analyzer, vacuum is used to separate the ions by a mass filter. The ions are filtered according to their *m/z* ratio. Furthermore, the filtered ions that have passed the mass filter reach the detector. The signal from the detector is processed and a mass spectra is constructed by a computer, providing the number of ions detected with the same *m/z* value (59). The molecular formula can be derived from full-scan accurate mass data. Some mass analyzers uses tandem MS technology, where several MS are coupled for mass filtering, fragmentation and detection to obtain the specific fragmentation spectra for molecules. Tandem MS where two MSs are ensembled in tandem are called MS/MS or MS^2, and if three MSs are coupled they are called MS/MS/MS or MS^3 (60). An example of tandem MS is multiple reaction monitoring (MRM) on a triple quadrupole MS, and is one of the MS technologies used for targeted analysis in clinical laboratories today (61). In metabolomics, fragmentation spectra can be obtained by a data dependent acquisition (DDA) or data independent acquisition (DIA). In DDA metabolites with the highest signal intensities measured in full-scan MS are selected for fragmentation. The fragments are then detected. In DIA the instrument acquires fragmentation data from one specific mass range at a time (62, 63).

The resolution and the mass accuracy vary dependent on the LC-MS platform. The resolution, or resolving power describe how close two signal peaks can be resolved, see **Equation 1**. When a mass spectrometer has higher resolution it is better suited to separate two peaks for metabolites or other chemical compounds with the same mass (53, 64).

$$R = \frac{m}{W_{0.5}} \qquad \textbf{\textit{Eq. 1}}$$

where R is the resolving power, m is the m/z value for the analyte, and $W_{0.5}$ is the width of the peak at half height.

The mass accuracy, or the mass error is measured in parts per million (ppm) and is the difference between the accurate, experimental mass and the exact mass that is calculated or expected. MS detectors can have different ppm (47, 53, 65), see **Equation 2.**

$$Mass\ accuracy\ (ppm) = \left(\frac{Exact\ mass - Accurate\ mass}{Accurate\ mass}\right) \cdot 10^6 \qquad \textbf{\textit{Eq. 2}}$$

Molecules that have very similar mass can be separated using MS with high resolution and is separated in decimal level. The mass accuracy and resolution in high resolution instruments for metabolomics is usually <5 ppm and above 10 000, respectively, and the low mass accuracy can be achieved by mass calibration (47, 66).

1.4.3 The Orbitrap High Resolution Mass Spectrometer

Orbitrap MS is a high resolution MS and in this thesis a Q Exactive Orbitrap MS was used for detection with electrospray ionization (ESI) coupled to UHPLC.

Electrospray Ionization - ESI

Before entering the mass analyzer, the molecules separated in the LC are converted to gas phase and then to ions in the ion source. In metabolomics, different soft ionization techniques are often used for ionization. ESI, atmospheric pressure chemical ionization (APCI), electron ionization (EI), and atmospheric photoionization (APPI) are often used (67). APCI and APPI are better suited for neutral or less polar compounds and induce little or less in-source fragmentation, but covers smaller metabolites under 1000 Da (68). Matrix-assisted laser desorption ionization (MALDI) has also been used in LC-MS metabolomics. MALDI does not show the same problem with ion suppression as other techniques and provides low selectivity

and chromatographic separation. However, the lack good signal precision makes it not suitable for quantitative analysis (43, 68, 69). In this thesis the ESI technique was used.

ESI is a type of atmospheric pressure ionization, and is suitable for ionization of polar compounds that has acidic or basic properties, or that makes adducts when binding to stable ions e.g. Na+ or NH4+. It has the advantage that there is no limitation in mass, which makes it suitable even for large protein complexes. The basis for ESI is to produce ions by an electrospray with high voltage applied to a liquid and to form an aerosol. High voltage is applied from the capillary to the MS inlet through an atmospheric pressure region. The liquid forms a Taylor cone, and a fine liquid jet is ejected from the Taylor cone in a region known as the jet region. A coaxial stream of nebulizing nitrogen gas and the high voltage potential will transform the liquid into a spray of small droplets in the ion evaporation region creating an aerosol of charged particles. From this the molecular ion is formed by evaporation and eventually moves towards the ion acceleration chamber into the inlet of the MS by an electric potential. The molecular ion could be either positively or negatively charged dependent on whether the voltage applied is positive or negative (53, 70). For ESI it is common to observe adducts in both positive and negative ionization mode. The most common adduct are protonated molecules, but also sodium, potassium, ammonium, acetate, and chloride have been observed originated from the mobile phase, plastic tubes, or the sample itself (53, 71). Furthermore, fast polarity switching between positive and negative ionization mode is often used in MSs and gives better comprehensive coverage for identification and for saving time. However, this can reduce the sensitivity. The disadvantage with ESI is ion suppression, ion enhancement and in-source fragmentation (43, 45, 72, 73).

Orbitrap Mass Spectrometer

The mass analyzer used in this thesis was a Q Exactive Orbitrap MS. The main components of the Q Exactive Orbitrap MS is a bent flatapole, a quadrupole mass filter, a C-Trap, a higher energy collision-induced dissociation (HCD) cell and an orbitrap mass detector. **Figure 1. 6** shows the main components of the Q Exactive Orbitrap MS.

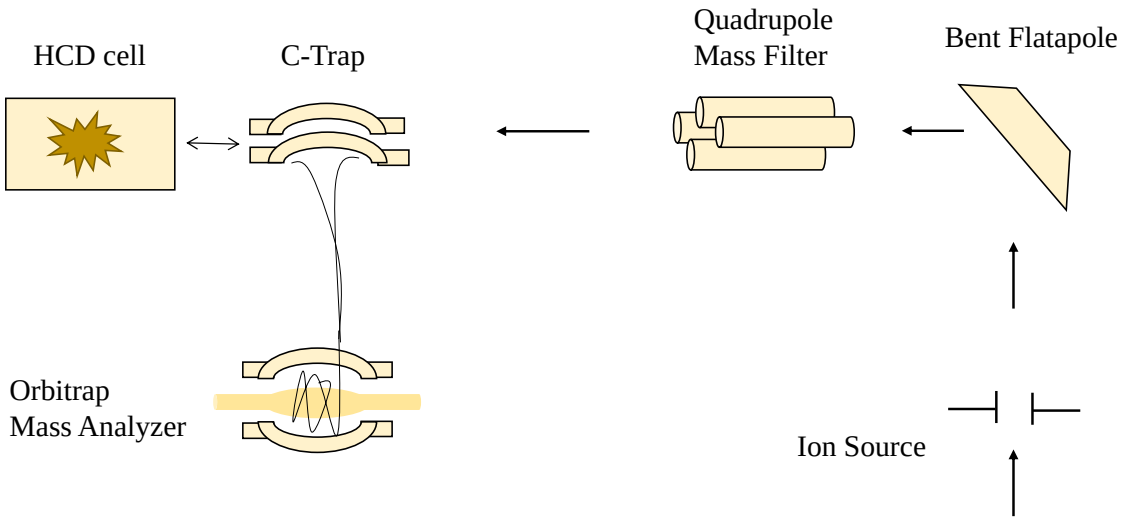


Figure 1. 6: Q Exactive Orbitrap Mass Spectrometer. *The figure shows the main components in the Q Exactive Orbitrap Mass spectrometer. It consists of an ion source, a bent flatapole, a quadrupole mass filter, C-trap, HCD cell, and the orbitrap mass analyzer.*

Intact molecule ions made in the electrospray enter the bent flatapole where the neutral molecules are filtered out and the ions are sent to the quadrupole mass filter. A quadrupole use an electrical field to filter the desired molecule ions in the desired *m/z* range by oscillation. The molecule ions with stable oscillation are sent to and collected in the C-Trap before they are detected by the Orbitrap mass analyzer. Here, the ions oscillate in an axial homogeneous electrostatic field so that the *m/z* obtains a specific oscillation used for determine the exact *m/z* of the metabolite. DDA is one of the possible techniques for this instrument to obtain the fragmentation pattern of the compounds with the highest signal intensities at each specific chromatographic retention time. These are sent to the HCD cell for fragmentation, and sent to the orbitrap for detection (74, 75). Orbitrap is a newer analyzer and has better resolution than other analyzers, e.g. quadrupole time-of-flight (Q-TOF) that is also used for metabolomics. While orbitrap has resolution 150 000 and mass accuracy 1-5 ppm, Q-TOF has resolution 40 000 and mass accuracy <1 ppm. Fourier transform ion cyclotron resonance (FT-ICR) MS has an ultra-high resolving power over 1 000 000 and has better mass accuracy (<1 ppm) than orbitrap MS. However, FT-ICR needs a superconducting magnet and is more expensive (45, 76).

1.4.4 Data processing

Data acquired from metabolomics platforms are generated in large numbers, and the data processing technique is dependent on data acquisition. System suitability tests can be used to follow the systems drift over time, and pooled quality controls are used to correct for retention time shifts, baseline drift, background noise, artifacts, and correct mass to obtain quantitative data. In LC-MS based metabolomics the basic steps for data processing are data alignment, filtering, and normalization. The matrix of a biological sample can vary between individuals, and is everything other than the analyte that is unknown. The matrix can affect the extent of the signal from the analytes and cause a change in the analytical signal (53, 77-79).

Data quality in metabolomics studies

A method validation should be done for an estimation of the method variability. According to the European Medicines Agency (EMA) Guideline on bioanalytical method validation, the precision is defined as "*the closeness of agreement (degree of scatter) between a series of measurements obtained under the prescribed conditions. Precision is defined as the ratio of standard deviation/mean (%)*" (80). Statistical parameters like the mean, standard deviation and the relative standard deviation (%RSD) can be used to express the precision. The repeatability is the within- run precision, and represents the smallest variation in the results. Repeatability is a measurement of the variation between data from the same sample by the same analyst using the same method over a short period of time. The intermediate precision is the between-run precision, and represents the largest variation between data from the same sample from analysis by different analysts, temperature, mobile phase, and analytical column, using the same method over a longer period of time. According to EMA the variation should be <15% RSD, except for the lower level of quantification (LLOQ) the variation should be <20% RSD. For the repeatability the sample size should be minimum five samples per concentration level, and for the intermediate precision at least three runs should be analyzed on at least two different days (80). However, these values are adapted for validation of results from targeted analyses, and cannot be directly transferred to bioanalytical method validation in metabolomics studies. This is due to the different detector systems used in metabolomics which have different amount of variation. Other studies have showed that acceptable limits for retention time and peak area variation is <2% and <30% RSD, respectively (81-85).

Contribution of matrix effects on data variation in metabolomics studies

Matrix effects can be chromatographic effects, ion suppression, or ion enhancement from the electrospray. The matrix effect, for determination of ion suppression and ion enhancement, is calculated by the matrix factor (MF). MF <1 it indicates ion suppression, and MF >1 it indicates ion enhancement. Ion suppression and ion enhancement occur during the electrospray ionization and influence the mass detector recovery. Ion suppression appears when the metabolite reaches the ESI simultaneously with many other molecules in the matrix causing a competition for ionization. Metabolites might not be ionized for the benefit for other molecules, or in a less extent. The signal intensity will be smaller compared to the chemical standard. This might appear due to salts, ion pairing reagents, endogenous compounds, drugs, or metabolites forming droplets and increasing the amount of loaded ions in the gas phase. The mass and charge are also factors that affect whether or not a metabolite becomes a candidate for ion suppression. Previous findings show molecules with a higher mass will suppress the signal of smaller molecules, and polar compounds are more likely to suppress the signal from other molecules. Chromatographic matrix effects are everything in the LC that leads to changes in the retention time, so chromatographic peaks are detected at a later time than in a chemical standard. This phenomenon appears when there is a competition between the analyte and other molecules in the matrix of binding to the stationary phase in the chromatographic column, or if other molecules in the matrix interact with the metabolite, making it move slower through the column. This can be captured finding a wider peak than usual, and can complicate the chromatographic separation of analytes with similar or identical retention time (86, 87). Chromatographic matrix effects can also be captured by retention time shift compared to a chemical standard. Chaleckis, R. et al. proposed that the retention time shift (ΔRT) should not exceed 0.1 min compared to a chemical standard (88).

1.4.5 Metabolite identification and data annotation

LC-MS data of metabolites has to be annotated to convert them into biological information. The metabolomics community and the Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI) has proposed defined metrics for annotation which reflects how good the identification is (89). This is often referred to as the level of confidence, and is the bottleneck of metabolomics. The level of confidence gives some criteria for metabolite identification e.g. accurate mass, retention time, fragmentation pattern and information on the study design. Dependent on how many criteria that contribute for

identification, the stronger the level of confidence become. Levels of confidence can be defined differently. In this thesis, the lowest level is level 5 and the highest is level 1. Level 5 is the feature annotation without annotation, level 4 is annotation by one parameter, level 3 is confirmation by class-specific standard, and level 2 requires matching fragmentation or MS/MS spectra to online databases or literature. The annotation can also be supported by retention time predictions and projection to similar LC methods. The challenge with retention time is that it is not the same in every analysis, it can vary with chromatographic column or from sample to sample because of different matrixes, and varies between different laboratories that uses different LC systems. Finally, there is level 1 that is the strongest level of confidence and includes stereochemistry discrimination. In some cases there is also level 0, and is a level where other analytical methods or techniques with other analytical properties are used to verify the identification or characterization (47, 88, 90, 91). **Figure 1. 7** is an overview of the five levels of confidence with increasing level of confidence from level 5 to level 1.

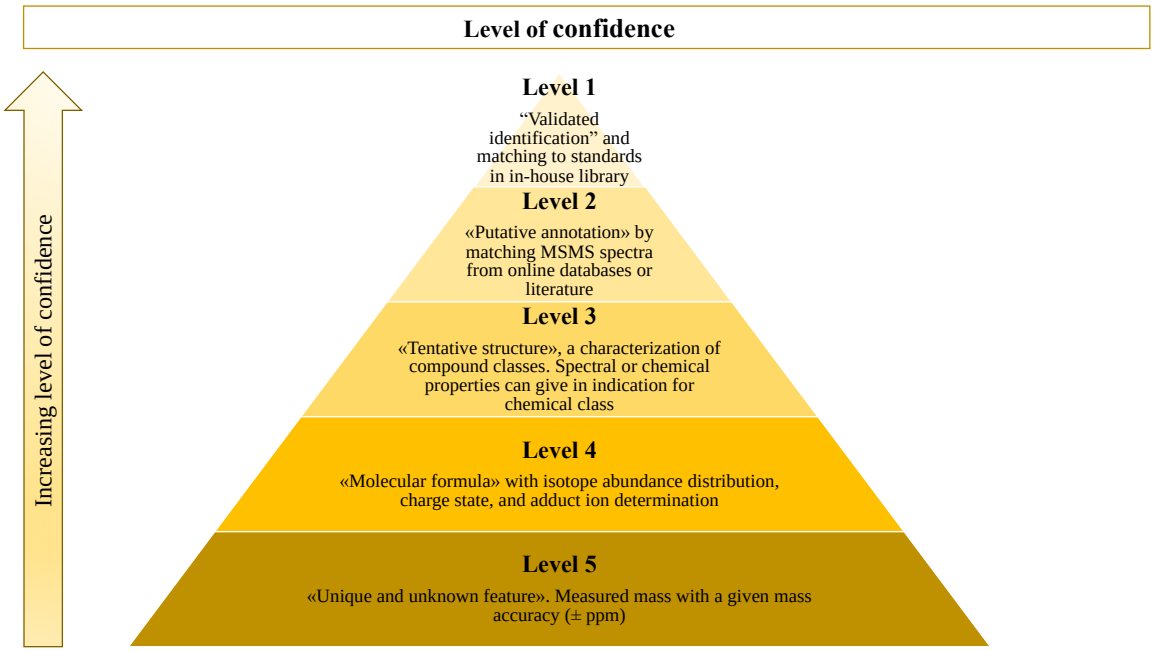


Figure 1. 7: Level of confidence. *The figure is an illustration of the five levels of confidence with level 5 at the bottom and increased level to level 1 as the highest level of confidence on top. Level of confidence is used for metabolite annotation and identification. The figure is inspired by (90).*

Metabolomics detection usually starts with the full scan accurate mass data and a conversion between the chemical formula and *m/z* ratio within 5 ppm to limit the possibilities. An *in-silico* prediction of metabolite properties is matched to retention time and fragmentation spectra for chemical standards. Together with experimental fragmentation spectra matched to databases

and libraries metabolite annotation and identification can be done. The prediction of the *m/z* value is not always without complication because in-source fragmentation can occur in addition to the formation of adducts, increasing the number of possibilities for other candidates (47, 88). For identification of and distinguishing between closely eluting compounds from the LC the fragmentation pattern can be used comparing experimental fragmentation spectra to spectra in online databases e.g. Human metabolome Database (hmdb) (92) and The Metabolite and Chemical Entity Database - METLIN (93-95), but also with spectra obtained from chemical standards from an in-house library. To evaluate the annotation of the metabolite and to increase the level of confidence for identification, it is important that as many fragments as possible match with the fragments in the library spectra. Different metabolites have different optimal collision energy for fragmentation. For the most informative fragmentation spectra different collision energies has to be used. The in-house library is therefore important equipment and is unique to the specific chromatographic and detection method used. However, many metabolites in a sample has low abundance and do not qualify for fragmentation because they are not always the components with highest abundance. A possible solution for this is by predefining with chemical formula in the inclusion list or exclusion list on the instrument (47, 88).

1.5 Metabolomics analysis of urine and plasma

A metabolomics method that is well suited for studies of urine and plasma samples is an advantage. Urine is one of the most commonly used biological fluids in metabolomics because of its ease of collection that does not require technical skills, and because it contains a wide range of metabolites from different biochemical classes, representing numerous of biochemical pathways (11, 96). However, unlike other biological fluids, e.g. plasma, the concentration of urinary metabolites varies widely between samples and can vary up to 15-fold (11, 13, 97). Therefore, analysis result from metabolomics of urine samples needs to be normalized to minimize the effect of variable urine concentration and to improve the quality of the statistical analyses of metabolomics data (97). There are several normalization techniques for urine samples, e.g. normalization to creatinine concentration, osmolality, urine volume, and MS total useful signal (MSTUS). Normalization of the urinary metabolome to the creatinine concentration is a common normalization method because creatinine reflects the overall concentration of metabolites in the urine. Creatinine is also excreted in relatively constant amounts dependent on gender, age, diet, hydration, muscle mass, and activity level. Urinary metabolites are expressed as µmol/mmol creatinine. Plasma on the other hand, is more

physiologically controlled and therefore a more stable matrix that provides good predictability and a snapshot of metabolites in a sample. Plasma samples are therefore not normalized (13, 96, 98). The challenge with plasma is the sample preparation and the complex matrix that consist of both lipophilic and hydrophilic compounds (13, 99, 100).

1.6 Aim of study

Today, several targeted analyses are used diagnostically in the study of the metabolome, specific for different types of biochemical compounds when an IEM is suspected. Metabolomics is the study of all metabolites, or as many biochemical reactions as possible using one single high performance mass spectrometer that is incorporated into clinical laboratories for a more effective diagnostics. However, it is not yet established to which extent quantitative analyses can be performed using metabolomics, and a higher level of confidence is needed for metabolite identification (38, 40, 82, 83). It is an advantage and need to have a metabolomics method that is well suited for easy accessible biological materials e.g. urine and plasma to provide better level of confidence for selected compounds and test the matrix effects.

The aim of this study is to evaluate the quantitative and qualitative performance of metabolomics studying amino acids and amino acid related compounds in urine and plasma samples. This includes comparing peak areas determined by the established metabolomics method HPLC – ESI – Q Exactive Orbitrap MS with quantitatively measured concentrations obtained with established targeted analyses used in laboratory diagnostics, and evaluate the quantitative and qualitative results from metabolomics on the selected compounds studying chemical standards, urine, plasma and matrix effects.

Specific objectives

1. Identify the selected compounds in standard solutions to construct an in-house library for better level of confidence in the study of amino acids and related compounds using metabolomics.
2. The annotated selected compounds in the in-house library will be the base for qualitative and quantitative studies of the selected compounds in plasma and urine samples, with the main focus on urine.
3. Comparing peak areas from metabolomics with quantitative concentration measurements from established targeted methods for the selected compounds in urine.
4. Study the matrix effects of the selected compounds in plasma.

2. Experimental

2.1 Chemicals

2.1.1 Solvents

Methanol (CH_3OH) with LCMS purity grade was from Rathburn Chemical Lrd (Walkerburn, Scotland). Formic acid (CH_2O_2) with LCMS purity, 98%-100% that was acquired from Sigma Aldrich (Darmstadt, Germany), and Thermo Scientific (Rockford, USA). The water used in this thesis was Type 1 water obtained from a Millipore MilliQ purification system with 18.2 MΩ•cm resistivity at 25 °C, with a Q-guard cartridge, quantum cartridge, and 0.2 μm pore filter membrane obtained from Merck (Darmstadt, Germany).

2.1.2 Reagents

DL-α-amino-n-butyric acid, gamma-amino-n-butyric acid (GABA; 4-aminobutanoic acid; Peptide acid) and 3-Methyl-L-Histidine ([S]-1-Methylimidazole-5-alanine; π-Methyl-L-histidine) were obtained from Sigma Aldrich. L-allo-isoleucine, L-Carnosine (β- Alanyl-L-Histidine), L-alanine, L-Isoleucine β – alanine and L-Leucine (α-Aminoisocaproic acid) was obtained from Sigma Aldrich.

2.1.3 Solutions

Mobile phases

The mobile phases made in this project were called Mobile phase A and Mobile phase B. Mobile phase A consisted of 0.1% formic acid in MilliQ water and mobile phase B consisted of 0.1% formic acid in methanol. All mobile phases were stored at room temperature.

Calibration solutions

Pierce LTQ Velos ESI Positive Ion Calibration solution and Pierce LTQ Velos ESI Negative Ion Calibration solution from Thermo Fisher Scientific and were stored at -20 °C.

Amino acid standard solutions

The amino acid standard A9906 (0.5 µmol/mL) was from Sigma Aldrich. PEA stock solution was 5 mM and diluted to 20 µM for analysis. The amino acid standard solutions called Special mix and 3AA were made at Oslo University Hospital – Rikshospitalet. All the amino acids studied in this thesis are listed in **Appendix A**, **Table A. 1**. The standard solution A9906 was diluted to 20 µM. **Table 2. 1** list the content of the amino acid standard solution 3AA made at Oslo University Hospital – Rikshospitalet. The monoisotopic molecular weights were found using hmdb.

Table 2. 1: Amino acid standard 3AA content. *The table list the amino acids in the chemical standard 3AA, the molecular weight, and concentration (µM).*

Amino acid	Monoisotopic molecular weight (Da)	Concentration (µM)
L-Glycine	75.032028409	500
L-Glutamine	146.069142196	500
L-Asparagine	132.053492132	100

Table 2. 2 list the content of the amino acid standard solution Special mix made at Oslo University Hospital – Rikshospitalet. The monoisotopic molecular weights were found using hmdb.

Table 2. 2: Amino acid standard special mix content. *The table list the amino acids in the chemical standard Special mix, the molecular weight, and the concentration (µM).*

Amino Acid	Monoisotopic molecular weight (Da)	Concentration (µM)
Glycylproline	172.08479226	10
4-Hydroxy-L-glutamic acid	163.048072403	50
3-Hydroxy-D,L-kynurenine	224.079706882	20
S-adenosyl-L-homocysteine	384.12158847	50
Formiminoglutamic acid	174.064056818	50
Argininosuccinic acid	290.122634328	50
Alloisoleucine	131.094628665	20
S-Sulfocysteine	200.976563719	20
Saccharophine	276.132136382	10
Homocitrulline	189.111341361	20
Homocarnosine	240.122240398	10
Aspartylglucosamine	335.132864663	50

2.2 Materials and equipment

2.2.1 Small equipment

The analytical weight used was AG 245 from Mettler-Toledo (Columbus, OH, USA). Pipets used in this thesis were Pipetman Classic line P1000 from Gilson (Middleton, USA) and Fisherbrand Elite adjustable-volume single-channel pipettes from Fisher Scientific (Waltham, MA, USA). Centrifugation at 4 °C was done using Heraeus Fresco 21 Centrifuge from Thermo Scientific and centrifugation at room temperature was done using megafuge 1.0R from Heraeus Instruments (Hanau, Germany).

2.2.2 Liquid Chromatography – Mass Spectrometry instrumentation and settings

The LC instrument used was UHPLC Dionex UltiMate 3000 from Thermo Fischer Scientific, consisting of a pump, column department, and an autosampler. The LC was coupled to a Q Exactive Orbitrap MS from Thermo Fisher Scientific. Instrumentation settings that was used in this thesis has been optimized by Arnesen (101), Skogvold (102) and Sandås (103) to achieve separation and detection of as many metabolites as possible, and for maximizing signal intensity in DBS.

The analytical column used was Pursuit XRs Diphenyl column (250 x 2.0 mm, particle size 3 µm) from Agilent Technologies (Santa Clara, CA, USA). These are shown in **Table 2. 3**. The mobile phase gradient is shown in **Figure 2. 1** and **Table 2. 4**.

Table 2. 3: Liquid chromatography instrumentation settings.

Parameter	Setting
Mobile phase A	0.1% formic acid in water
Mobile Phase B	0.1% formic acid in methanol
Flow rate (µL/min)	300
Injection volume (µL)	2
Column temperature (°C)	30
Analysis time (min)	27.5
Re-equilibration time (min)	10

Table 2. 4: Gradient elution during the analysis.

Time (min)	Mobile phase B (%)
0	2
6	10
8.5	75
25	100
27.5	100

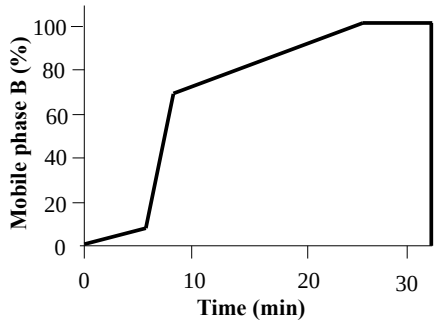


Figure 2. 1: Mobile phase B gradient throughout the analysis.

The MS and ESI instrumentation settings used in this thesis are shown in **Table 2. 5** and **Table 2. 6**, respectively. The analysis was done in positive and negative ionization mode for a comprehensive coverage of identification.

Table 2. 5: Mass spectrometry instrumentation settings.

Parameter	Setting
Scan types	Full MS and data dependent acquisition (DDA), Top 5
Scan ranges (*m/z*)	50-750
Resolution	Full MS: 70 000 DDA, Top 5: 17 500
Stepped normalized collision energy	20, 50 and 80
Polarity	Positive and negative
Micro scans	1
Lock masses	Off
Automatic gain control target value	Full MS: $1 \cdot 10^6$ DDA, Top 5: $5 \cdot 10^6$
Maximum injection time (ms)	Full MS : 250 DDA, Top 5: 100
Analysis time (min)	32.5
Re-equilibrium time (min)	10

Table 2. 6: Electrospray instrumentation settings.

Parameter	Setting
Electrospray needle position	C
Capillary temperature (°C)	250
Electrospray voltage (kV)	+/- 3.5
Sheat gas (N_2) flow rate (a.u.)	40
Auxiliary gas (N_2) flow rate (a.u.)	10
Auxiliary gas heater temperature (°C)	300
Sweep gas (N_2) flow rate (a.u.)	2
S-lens RF level	50

2.2.3 Computer Software

Xcalibur (version 4.2.47) from Thermo Fischer Scientific was used for data acquisition, and FreeStyle 1.5 and 1.6 from Thermo Fischer Scientific was used for data processing. The exact *m/z* and 5 ppm was filtered out in the mass range *m/z* 50-750 by typing the chemical formula to observe the chromatographic peaks and retention times. Template files were made with the exact *m/z* and automatic integrated chromatographic peaks to obtain peak area. If needed manual integration was used.

mzVault (version 2.2 and 2.3) from Thermo Fischer Scientific was used to make the in-house library for amino acid standard solutions. The raw data file that contained the mass spectral data was selected, and the query spectrum was found. In the Trace Type list the mass, and the tolerance of 5 ppm was chosen to filter the chromatogram. The specific time point was chosen in the Chromatogram pane. The Compound Class, ChemSpider ID, HMDB ID, KEGG ID, and PubChem ID were added in the Compound List pane.

Tune (version 2.11) from Thermo Fischer Scientific was used for calibration and observation of the mass spectrometric parameters.

2.3 Sample preparation

The biological material used in this thesis was obtained from the diagnostic biobank (Diagnostikk- og behandlingsbiobank no. 329 in the Biobank Registry, the Norwegian Institute of Public Health) from which biological material can be used for development and quality assurance of diagnostic laboratory analyses. System suitability samples were analyzed in all metabolomics analyses to assess the suitability of the system.

2.3.1 Sample preparation of urine

Urine samples stored at -20 °C were thawed before vortexed for blending and centrifuged at 3600 rpm for ten minutes at room temperature. Then, the samples were diluted to the appropriate creatinine concentration before filtration at 4800 rpm and 4 °C for ten minutes in 0.22 µM Cellulose Acetate Spin Filters from Agilent Technologies. The filtrate was used further. **Figure 2. 2** is an illustration of the general workflow for the sample preparation of urine for this thesis.

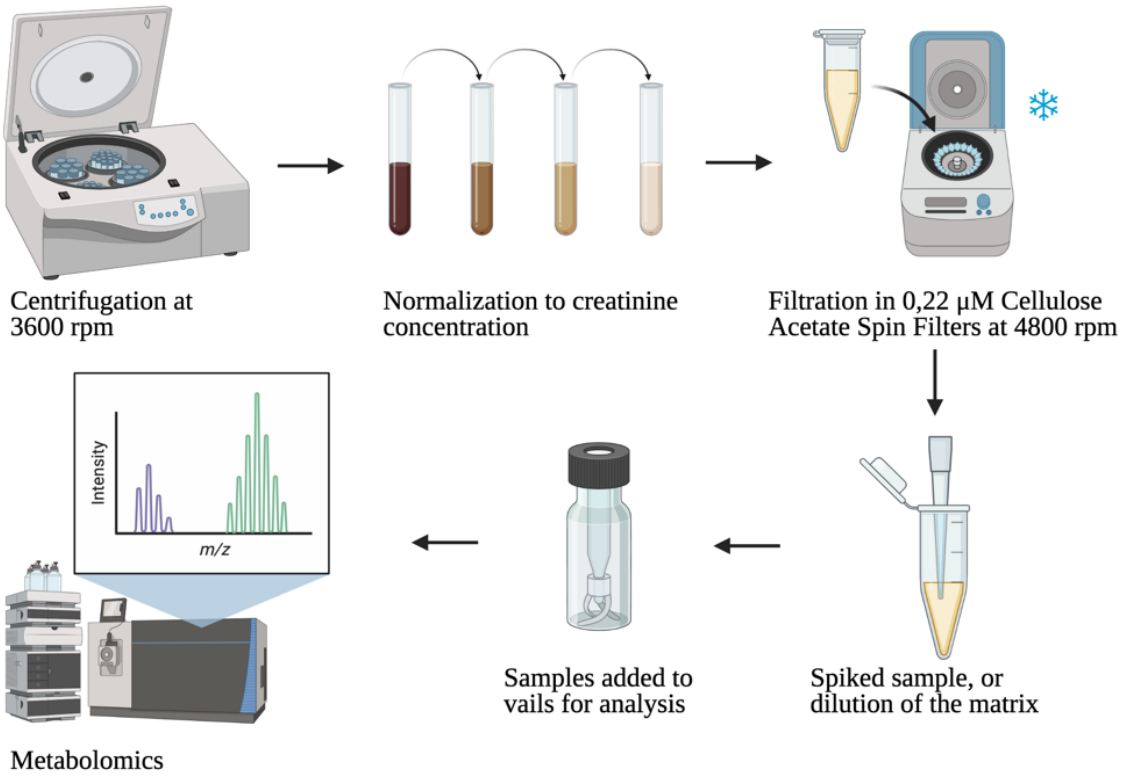


Figure 2. 2: Sample preparation of urine samples. *The general workflow for urine sample preparation used in this thesis. Samples were centrifuged at 3600 rpm, normalized to the creatinine concentration, filtered in 0.22 µM Cellulose Acetate Spin Filters at 4 ℃, before transferred to vials and analyzed using metabolomics.*

2.3.2 Sample preparation of plasma

Blood was drawn and centrifuged for ten minutes in 3600 rpm and room temperature. Proteins from the plasma were pelleted by adding three times more methanol than plasma (vol/vol) to the sample and centrifuged at 4 °C and 4800 rpm for ten minutes. **Figure 2. 3** is an illustration of the general workflow for sample preparation of plasma samples for this thesis.

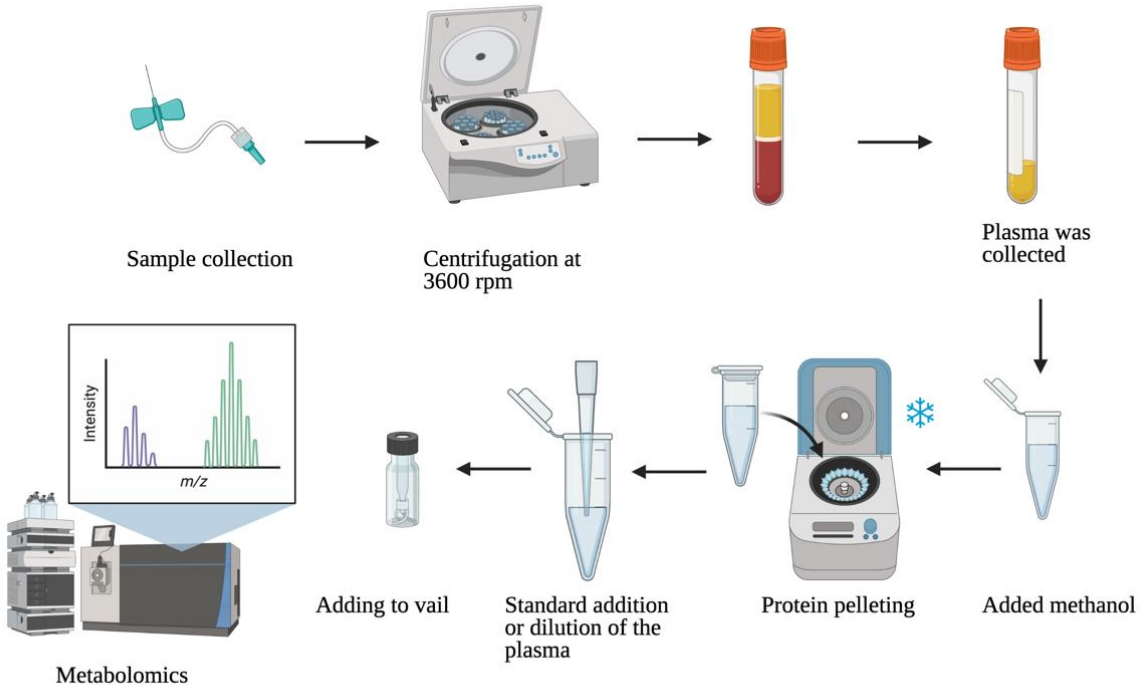


Figure 2. 3: Sample preparation of plasma. *The general workflow for plasma sample preparation used in this thesis. Blood sample was collected and centrifuged at 3600 rpm. The plasma was collected and protein was pelleted by centrifugation with methanol at 4 ℃ and 4800 rpm. The sample was spiked with chemical standard or the matrix was diluted, before adding to vials and analyzed by metabolomics.*

2.4 Analysis of selected compounds in urine

2.4.1 Creatinine measuring using Cobas c702 from Roche

The creatinine concentration was measured prior to metabolomics using Cobas c702 by Roche Diagnostics AS (Basel, Switzerland). All measurements of the creatinine concentration for urine samples were performed by this accredited method as a part of the routine analysis of creatinine concentration at the Department of Medical Biochemistry at Oslo University Hospital – Rikshospitalet.

2.4.2 Metabolomics detection of selected compounds in urine

Creatinine was detected using the retention time and the exact mass with an added hydrogen atom $[M + H]^+ \pm 5$ppm in positive ionization mode, and deprotonated $[M - H]^- \pm 5$ppm in negative ionization mode, extracted from the TIC in FreeStyle 1.6 from Thermo Fischer Scientific. The retention time and the MS/MS fragmentation spectra for creatinine were compared to the chemical standard in the in-house library. The peak area for the creatinine signal was found using FreeStyle 1.6 from Thermo Fisher Scientific. Three urine samples were spiked with creatinine to study the creatinine signal peak.

Amino acids were detected in urine. Retention time and fragmentation spectra from the in-house library in mzVault 2.2 were used. The protonated molecular ion for each amino acid $[M + H]^+ \pm 5$ppm in positive ionization mode, and the deprotonated molecular ion $[M -H]^- \pm 5$ppm in negative ionization mode, were extracted from the TIC in FreeStyle 1.6. The peak areas for the amino acid signals were found in FreeStyle 1.6.

2.4.3 Precision

The precision of the metabolomics method for detection of the selected compounds in urine samples was found by studying peak area and retention time repeatability and intermediate precision in samples containing 2 mM creatinine.

Repeatability

Creatinine and amino acid peak area and retention time repeatability were found by analyzing ten injections of one urine sample distributed evenly throughout an analysis sequence of 45 injections. The mean, standard deviation, and the %RSD were calculated for each of the selected compounds.

The intermediate precision

Creatinine and amino acid peak area and the retention time intermediate precision were found by analyzing nine aliquots of one urine sample for nine different analyses performed in different days. The samples were stored at -80 °C in 200 µL vials from NanoTemper (München, Germany) and sample preparation was done before each analysis. The mean, standard deviation, and %RSD were calculated for each of the selected compounds.

2.4.4 Quantitative comparison of metabolomics and targeted analysis

Peak area from the metabolomics method was compared to concentrations measured with targeted analyses for a quantitative study of selected compounds in urine.

Evaluation of the qualitative and quantitative results for creatinine in urine samples from metabolomics

Creatinine was detected in ten different samples from different individuals with 2 mM creatinine using the metabolomics method. The mean, standard deviation and %RSD were calculated for the creatinine peak area and retention time. The samples were also diluted to 1, 0.5, and 0.1 mM creatinine, respectively.

Creatinine peak area in four urine samples from different individuals was found using the metabolomics method. The creatinine concentrations in the four samples from different individuals were 2, 10, 15, and 20 mM, respectively. The concentrations were measured with the targeted analysis prior to metabolomics as a part of the routine analysis, as mentioned in **Chapter 2.4.1**. Creatinine peaks with the same retention time were compared, and the regression coefficient was found.

Quantitative study of selected compounds in urine

Metabolomics data from urine samples were normalized to 2 mM creatinine, as illustrated in **Figure 2. 4**. Samples that contained higher creatinine concentrations than 2 mM were diluted to 2 mM creatinine before data acquisition. Samples that contained lower concentrations than 2 mM creatinine was injected with a higher volume of sample into the instrument.

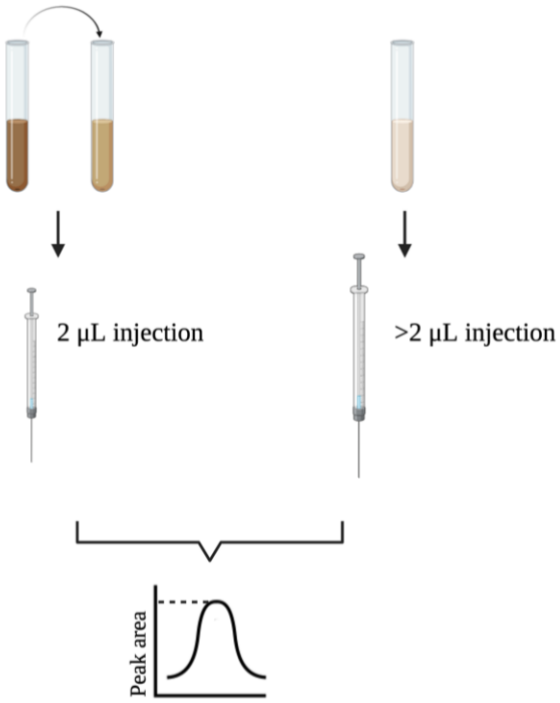


Figure 2. 4: Normalization to 2 mM creatinine for urine samples in metabolomics studies. *The figure illustrates the normalization strategy to the creatinine concentration for urine samples in metabolomics studies. Sample with higher creatinine concentration than 2 mM were diluted to 2 mM before analysis, and samples with lower creatinine concentrations were injected with higher injection volume into the instrument.*

Amino acid peak area from metabolomics was studied for five samples normalized to 2 mM creatinine and 0.1 mM creatinine were compared to measured concentrations from targeted analyses. Scatter plots were created to find the correlation between peak area and concentration. One pathogenic cystinura sample was analyzed and the biomarkers L-cystine, L-arginine, L-lysine and L-ornithine were quantitatively studied.

For the metabolomic study of HPP, two normal newborn urine samples were spiked with PEA by adding 10 µL of 25 and 50 µM PEA standard solutions, respectively. To represent HPP, two newborn samples were spiked with PEA by adding 10 µL of 100 and 500 µM PEA, respectively.

2.5 Analysis of selected compounds in plasma

2.5.1 Metabolomics detection of selected compounds in plasma

Amino acids were detected in plasma. Retention time and fragmentation spectra from the in-house library in mzVault 2.2 were used. The exact mass for each amino acid with an added hydrogen atom $[M + H]^+ \pm 5$ppm in positive ionization mode, and the deprotonated molecular ion $[M -H]^- \pm 5$ppm in negative ionization mode, were extracted from the TIC in FreeStyle 1.6. The peak areas for the amino acid signals were found using FreeStyle 1.6.

2.5.2 Precision

The precision of the metabolomics method on the selected compounds in plasma was found studying peak area and retention time repeatability and intermediate precision in plasma samples spiked with 20 µM amino acid standard solution A9906.

Repeatability

Amino acid retention time and peak area repeatability was found by analyzing ten injections of one plasma sample distributed evenly throughout an analysis sequence of 20 injections. The mean, standard deviation, and the %RSD were calculated for each of the selected compounds.

Intermediate precision

Amino acid retention time and peak area to the signal intermediate precision was found by analyzing ten aliquots of one plasma sample in ten different analysis sequences. Ten aliquots of 100 µL of the plasma sample were stored at -80 °C in Micro tube 0.5 mL, blue from Sarstedt AG & CO. KG (Nümbrecht, Germany), and the sample preparation was done before each analysis. The mean, standard deviation, and the %RSD were calculated for each of the selected compounds.

2.5.3 Matrix effects for selected compounds in plasma using metabolomics

Chromatographic matrix effects, ion suppression, and ion enhancement were studied for amino acids in plasma. Plasma was spiked with amino acid standards with final concentration 20 µM. A plasma sample with added water was made with the same plasma concentration as for the spiked plasma, and a standard solution with 20 µM amino acids was made. Matrix effects were

studied by comparing amino acid retention time and peak area in the spiked plasma with amino acid retention time and peak area in the standard solution and the normal plasma sample with water added. The difference in retention time (ΔRT) was calculated, and the MF (%) was calculated for each amino acid using **Equation 3**.

$$MF = \frac{A_{Spiked\ plasma}}{A_{normal\ plasma} + A_{standard\ solution}} \qquad \textbf{\textit{Eq.3}}$$

where A is the peak area for the amino acid signal, and MF is the matrix factor in %.

3. Results

Selected metabolites were analyzed with the established metabolomics method HPLC – ESI - Q Exactive Orbitrap MS method and detected in chemical standards, urine and plasma. A qualitative characterization of selected metabolites was performed, and an in-house library was created. The precision of the metabolomics method on the selected compounds in urine and plasma was tested. Lastly, the quantitative performance of the metabolomics method was evaluated, studying data from metabolomics compared to concentration obtained from targeted analyses.

3.1 Detection and characterization of selected metabolites

A study of the qualitative performance of the metabolomics method was performed for detection of amino acids and amino acid related compounds. An in-house library was created for better level of confidence for identification of the selected compounds in future metabolomics studies. The selected compounds were analyzed using the metabolomics method in both positive and negative ionization mode. The chemical data and overview of the selected compounds detected can be found in **Appendix A, Table A. 1**. The table gives the chemical formula, molecular weight, structure, and logP value. **Figure 3. 1** shows the EIC of the selected compounds detected in positive ionization mode. For the identification and annotation of the selected compounds, the fragmentation spectra for each amino acid in the standard solutions were found in positive and negative ionization mode. The experimental fragmentation spectra were compared to spectra in databases e.g. Metlin, HMDB, and mzCloud, and used to make the in-house library in mzVault.

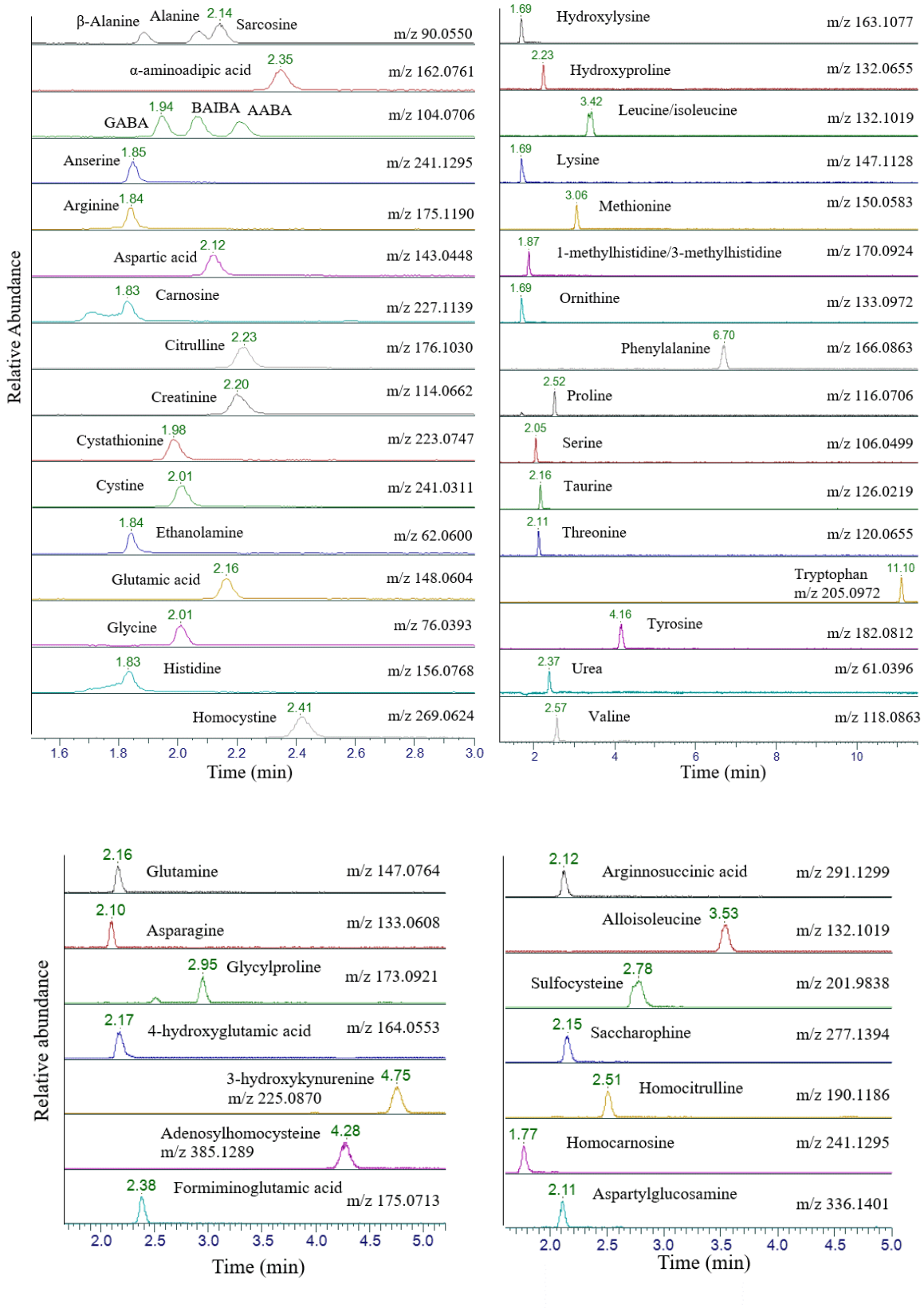


Figure 3. 1: Detection and characterization of selected metabolites. *The figure shows the chromatograms for the selected metabolites in chemical standards studied in this thesis.*

All the selected amino acids were detected in standard solutions as described in **Chapter 2.1.3** using the metabolomics method. Ethanolamine was detected in positive ionization mode, but not in negative ionization mode. Carnosine, L-histidine, and S-sulfocysteine had characteristic chromatograms. With the metabolomics method PEA was also detected. **Figure 3. 2** shows the TIC, EIC and fragmentation spectra in negative ionization mode for PEA.

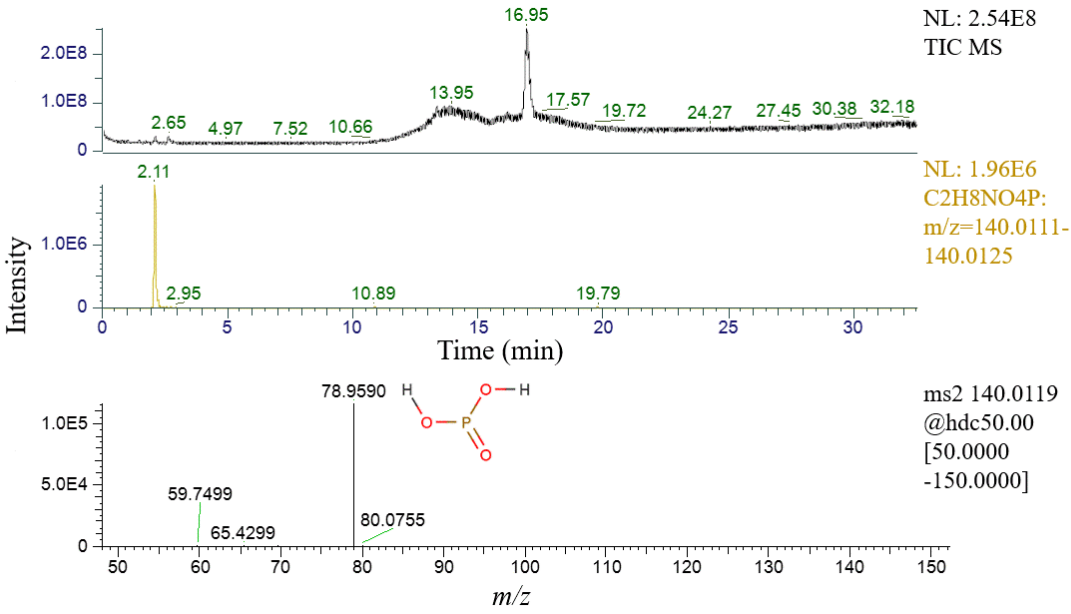


Figure 3. 2: Detection of phosphoethanolamine using metabolomics. *The figure shows the PEA chromatogram and fragmentation spectra obtained from the analysis with the metabolomics platform detected in negative ionization mode.*

As **Figure 3. 2** shows, PEA had retention time 2.11 minutes, and had signature fragment 78.9590. PEA was detectable in positive ionization mode and negative ionization mode, but the signal was observed to be better in negative ionization mode.

3.1.1 Identification of chromatographically separated amino acids with the same exact mass

Amino acids with the same exact *m/z* ± 5 ppm that were chromatographically separated were identified by spiking with each of the amino acids to the standard solution. β-alanine, L-alanine, and L-sarcosine have the chemical formula $C_3H_7NO_2$ and the same mass. **Figure 3. 3 A** shows elevated signal intensity of β-alanine, **Figure 3. 3 B** shows elevated signal intensity of L-alanine, and **Figure 3. 3 C** shows elevated signal intensity of L-sarcosine.

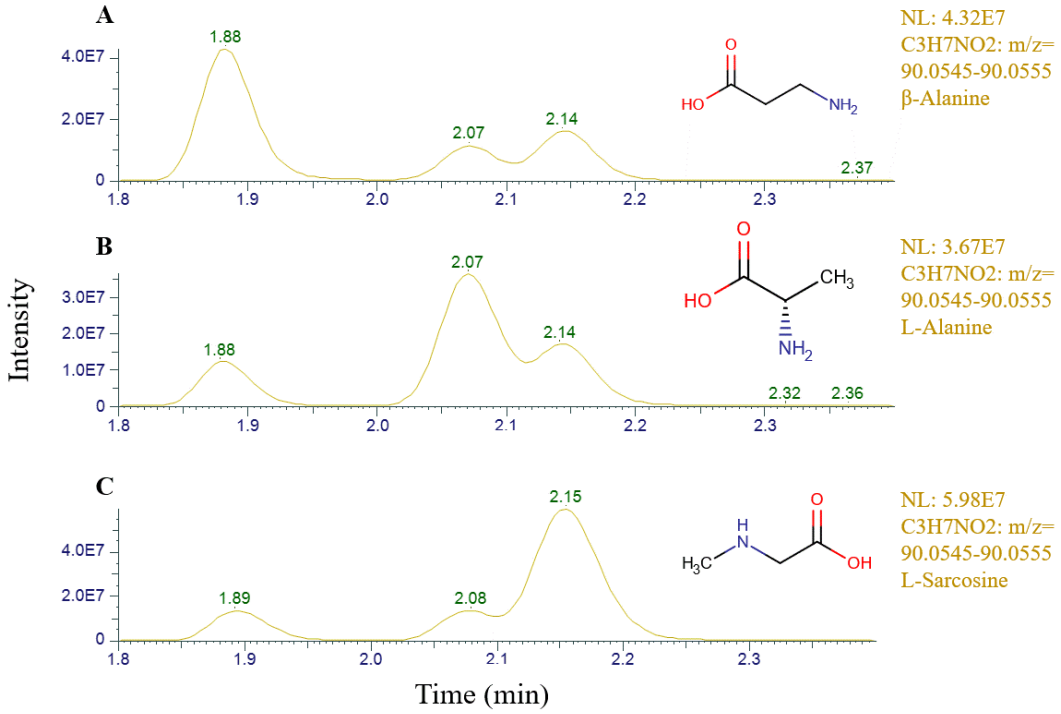


Figure 3. 3: Detection of metabolites in the same chromatogram with the same exact mass. *Three samples were spiked with β-alanine, L-alanine, and L-sarcosine, respectively.* ***Figure 3.3 A*** *show β-Alanine,* ***Figure 3.3 B*** *show L-Alanine and* ***Figure 3.3 C*** *show L-sarcosine. The fragmentation spectra were also used for identification.*

The same characterization was done using the standard addition method for L-α-amino-n-butyric acid, D,L-β-aminoisobutyric acid and γ-amino-n-butyric acid. These compounds have the chemical formula $C_4H_9NO_2$ and the same mass.

3.1.2 Not all amino acids with the same exact mass were chromatographically separated

Some of the amino acids with exact *m/z* were not chromatographic separated using the metabolomics method and were represented in the same peak. L-leucine, L-isoleucine, and alloisoleucine were identified with fragmentation spectra and two not chromatographic separated peaks were observed. The same characterization was done for 1-methyl-L-histidine and 3-methyl-L-histidine, except the peak was slimmer and no split peak was observed. **Figure 3. 4 A** shows the L-isoleucine, L-leucine, and alloisoleucine chromatogram, and **Figure 3. 4 B** shows the 1-mehtyl-L-histidine and 3-methyl-L-histidine chromatogram.

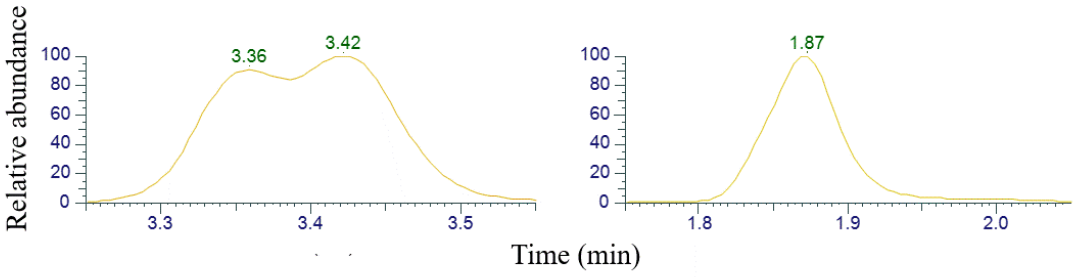


Figure 3. 4: Amino acids that were not chromatographically separated. Figure 3.4 A *gives the L-isoleucine, L-leucine and alloisoleucine chromatogram.* ***Figure 3.4 B*** *gives the 1-methyl-L-histidine and 3-methyl-L-histidine chromatogram.*

L-isoleucine and alloisoleucine fragmentation spectra were found in the first part of the peak, and L-leucine fragmentation spectra was found in the last part of the peak. **Figure 3. 5** shows the fragmentation spectra for L-isoleucine, L-leucine, and alloisoleucine, respectively.

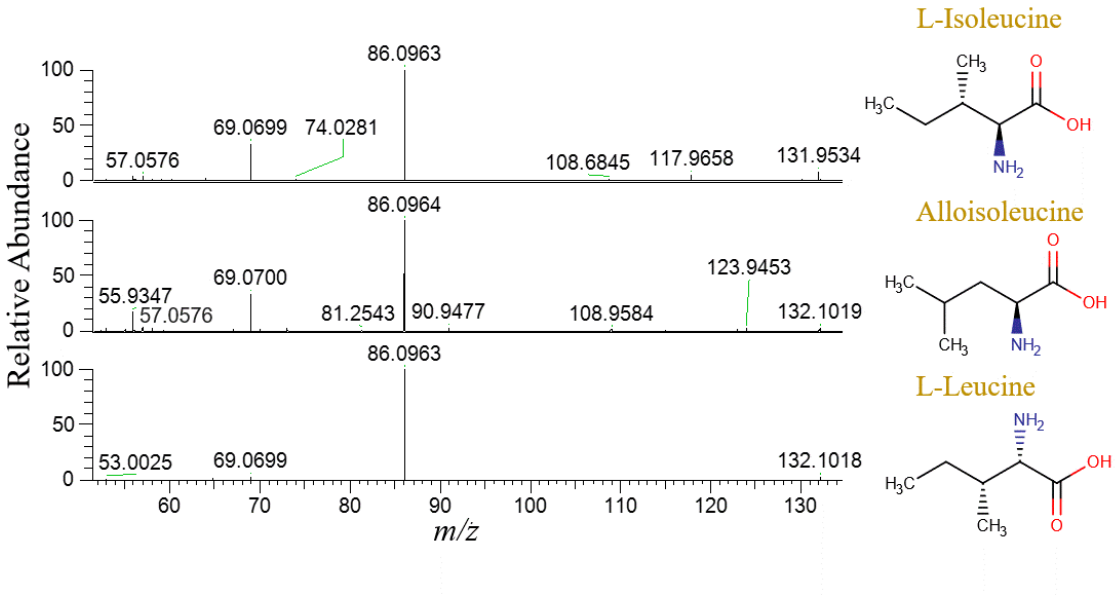


Figure 3. 5: L-isoleucine, L-leucine, and alloisoleucine fragmentation spectra. *The figure shows the L-isoleucine, L-leucine, and alloisoleucine fragmentation spectra, respectively. The fragment at m/z 69.0699 was common for all three amino acids. L-isoleucine and alloisoleucine had in addition a fragment m/z 57.0576.*

Figure 3. 5 shows that L-isoleucine, L-leucine and alloisoleucine had the fragment *m/z* 69.0699, but L-leucine had lower occurrence of this fragment than L-isoleucine and alloisoleucine. L-isoleucine and alloisoleucine had also fragment *m/z* 57.0576. L-isoleucine and alloisoleucine spectra were found at retention time 3.27 minutes, while L-leucine fragmentation spectra was found at retention time 3.45 minutes.

3.1.3 No obvious correlation was found between the amino acid structure and retention time

A study of a possible correlation between amino acid structure and retention time using a diphenyl UHPLC column was performed by evaluating different parameters of the amino acid structure. The different properties that were examined were the total number of rings in the structure, the number of benzene rings, the total number of double bonds, the number of conjugated double bonds, the number of sulfur and nitrogen atoms, solubility, logP-value, and the molecular weight. **Figure 3. 6 A** shows the correlation between the numbers of conjugated double bonds and the retention time. **Figure 3. 6 B** shows the correlation between the number of benzene rings and retention time.

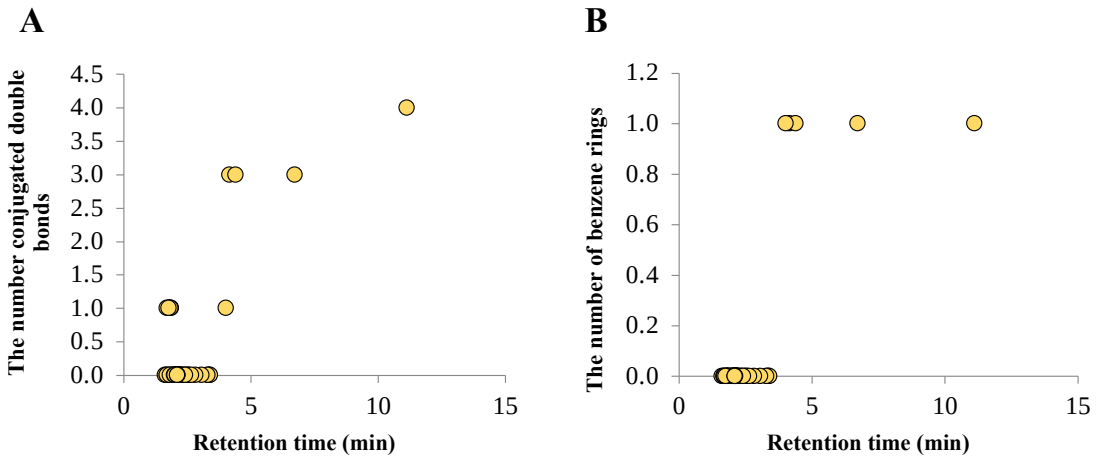


Figure 3. 6: Correlation between amino acid structure and the retention time (min). Figure 3.6 A *shows the correlation between the number of conjugated double bonds and the retention time measured in minutes for amino acids and related compounds studied in this thesis.* ***Figure 3.6 B*** *shows the number of conjugated double bonds and the retention time measured in minutes*.

Amino acids with benzene rings in their structure had the longest retention time, and with increasing number of conjugated double bonds the retention time was longer. L-isoleucine and L-leucine have no rings in their structure. However, the retention time was relatively long (3.40 minutes). L-histidine and L-anserine have a ring structure, but the retention time was short (1.80 minutes and 1.84 minutes, respectively). Based on the parameters evaluated in this thesis, there was no obvious correlation between amino acid structure and retention time.

3.3 Analysis of selected compounds in urine

Selected amino acids and amino acid related compounds were detected in urine using the metabolomics platform. The precision of the metabolomics method was evaluated, and a comparison of the metabolomics method with targeted methods for measurement concentration of the selected compounds was performed.

3.3.1 Detection of selected compounds in urine

Selected compounds were detected in urine samples diluted to 2 mM creatinine using the metabolomics method, and the signal peaks were studied and identified. Some samples displayed peak splitting for the creatinine signal. **Figure 3. 7** shows the EIC for creatinine $[M + H]^+ \pm 5$ ppm in positive ionization mode before and after spiking of creatinine. **Figure 3. 7 A** shows the result before spiking, and **Figure 3. 7 B** shows the result after spiking.

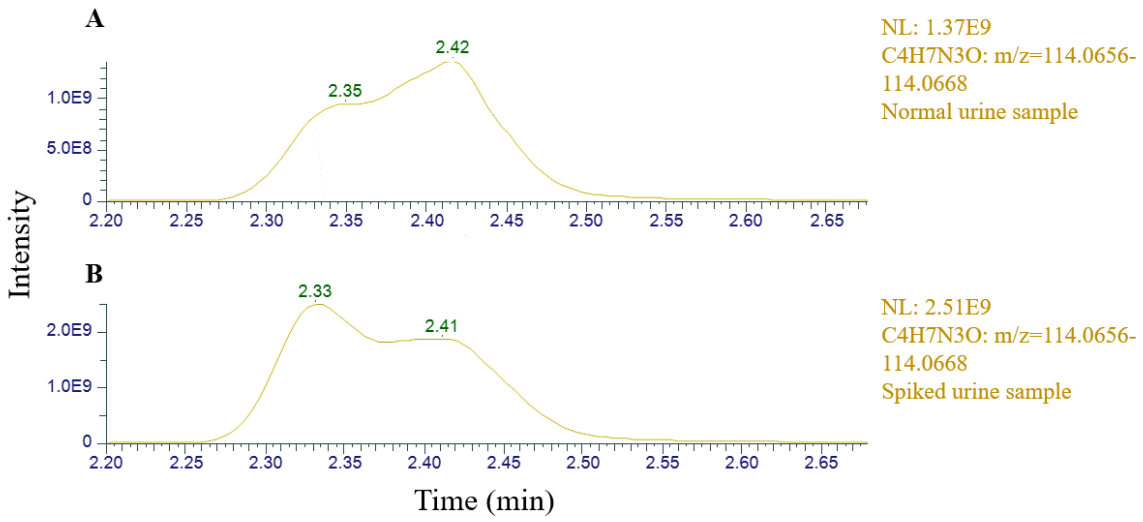


Figure 3. 7: Identification of the creatinine signal. Figure 3.7 A *is the result from the identification of creatinine in one of the urine samples included in this study that shows that there were two not fully chromatographically separated signals for creatinine in this sample. After standard addition of 2 mM creatinine, there was an increase in the foremost signal (2.35 minutes), as seen in* ***Figure 3.7 B****.*

Figure 3. 7 B shows an increase of the foremost signal by a ratio of 2.6. The retention time and peak area varied for several of the selected compounds studied.

3.3.2 Precision

The precision of the metabolomics method for the detection of selected compounds in urine samples was characterized studying peak area and retention time repeatability and intermediate precision in samples containing 2 mM creatinine.

Repeatability

The metabolomics repeatability and the stability of the samples for ten injections of a urine sample were determined by studying the amino acid peak area and retention time repeatability. The mean, standard deviation and %RSD were calculated for each of the selected compounds. The results for creatinine repeatability in positive ionization mode are shown in **Table 3. 1**, and the amino acid repeatability is shown in **Appendix B**, **Table B. 1**.

Table 3. 1: Creatinine repeatability. *Creatinine peak area and retention time repeatability were found by ten injections of one urine sample that contained 2 mM creatinine. The peak area and retention time mean, standard deviation and %RSD were calculated.*

	Retention time (min)	**Peak area (a.u.)**
Mean	2.306	$1.6 \cdot 10^{10}$
SD	0.004	$7.9 \cdot 10^{8}$
%RSD	0.2	5.0

The %RSD for creatinine retention time and peak area repeatability were 0.2 and 5% RSD, respectively. **Table B. 1** show that the amino acid retention time and peak area repeatability in urine had low variation (<2% and <30% RSD, respectively). For example the %RSD for L-anserine retention time was 0.1%, and the %RSD for L-anserine peak area was 6.99% RSD.

Intermediate precision

Amino acid peak area and retention time intermediate precision were studied nine aliquots of a urine sample, prepped and analyzed nine different days. The mean, standard deviation and %RSD were calculated for peak areas and retention times for each of the selected compounds. The results for creatinine repeatability in positive ionization mode are shown in **Table 3. 2**, and the amino acid intermediate precision is shown in **Appendix B**, **Table B. 2**.

Table 3. 2: Creatinine intermediate precision. *Creatinine peak area and retention time intermediate precision were found studying nine aliquots of a urine sample, containing 2 mM creatinine, for nine different days. The peak area and retention time mean, standard deviation and %RSD were calculated.*

	Retention time (min)	**Peak area (a.u.)**
Mean	2.47	$6.27 \cdot 10^{12}$
SD	0.08	$9.39 \cdot 10^{12}$
%RSD	3.07	149.8

The %RSD for creatinine retention time and peak area intermediate precision was 3.07% and 149.8% RSD, respectively. As **Table B. 2** showed low variation of <2% and <30% RSD for amino acid retention time and peak area intermediate precision in urine. For example L-anserine retention time and peak area was 2.89% and 14.1% RSD, respectively. L-α-aminoadipic acid, hydroxykynurenine, and S-adenosyl-L-homocysteine peak area repeatability had high variation and was 47.1%, 31.9%, and 32.6% RSD, respectively.

3.3.3 Quantitative comparison of the metabolomics method with the targeted method for concentration measurement

The quantitative performance of the metabolomics method was evaluated studying peak area from metabolomics compared to measured concentrations from targeted analyses. Four to ten urine samples were studied to evaluate the quantitative and the qualitative results from the metabolomics method.

Evaluation of the qualitative and quantitative results from metabolomics

Creatinine was detected using the metabolomics method in ten different samples from different individuals with creatinine concentrations diluted to 2, 1, 0.5, and 0.1 mM, respectively. The mean, standard deviation and %RSD were calculated for creatinine peak area and retention time. **Figure 3. 8** shows the correlation between creatinine peak area and retention time for samples with different creatinine concentrations.

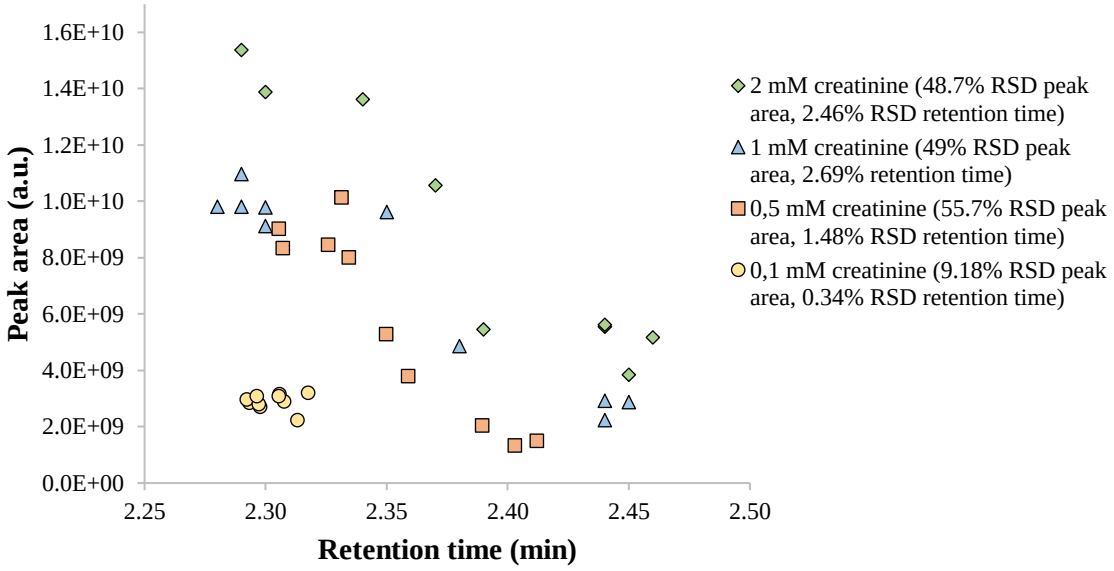


Figure 3. 8: Correlation between creatinine peak area and retention time. *The figure shows that lower creatinine concentration gave lower retention time and peak area, and that the peak area variation was low (<30% RSD).*

The %RSD for creatinine peak area in samples with 2, 1, 0.5, and 0.1 mM creatinine were 48.7, 49.0, 55.7, and 9.18% RSD, respectively, and for creatinine retention time it was 2.6, 2.8, 1.6, and 0.35% RSD, respectively. In samples with 2, 1, and 0.5 mM creatinine the retention time varied from 2.30 to 2.45 minutes. When the samples were diluted to 0.1 mM creatinine the peak area and retention time variation was low (<30% and <2%, respectively). No splitting of the peaks for the creatinine signal was observed in samples diluted to 0.1 mM creatinine, and the %RSD was <30%. **Figure 3. 8** show that creatinine peak area was low in samples with high creatinine retention time and high in samples with low retention time for samples with 2, 1, and 0.5 mM creatinine. These results indicated that just creatinine peaks with the same retention time should be compared.

Quantitative study of selected compounds in urine

Peak area for selected compounds from metabolomics was compared to measured concentrations from targeted analyses. First, the results for creatinine were studied. Furthermore, the selected amino acids and amino acid related compounds were studied quantitatively by normalization to the creatinine concentration.

Creatinine peak area in four urine samples from different individuals was found using the metabolomics method. The creatinine concentrations were 2, 10, 15, and 20 mM and were measured with the targeted analysis of creatinine concentration. Creatinine peak areas with the

same retention time were compared, and the regression coefficient was found. The result is shown in **Figure 3. 9**.

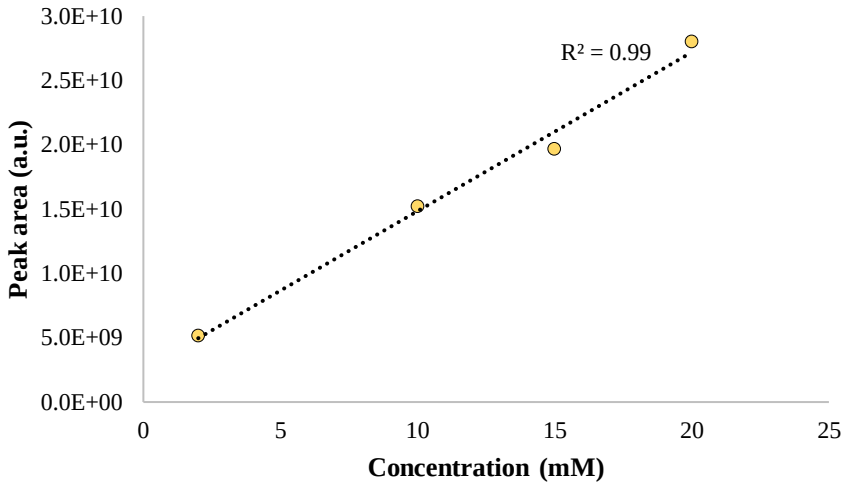


Figure 3. 9: Quantitative comparison of creatinine peak area and concentration for samples with the same creatinine retention time. *The figure shows that there was a linear correlation between the creatinine peak area from metabolomics and the measured creatinine concentration, with regression coefficient 0.99.*

The regression coefficient was 0.99 and a linear correlation was found between creatinine peak area from metabolomics and the measured concentrations form targeted analysis. This correlation was found when peaks with the same creatinine retention time from the metabolomics method were compared.

The quantitative performance of the metabolomics method was studied for the selected amino acids and amino acid related compounds in five samples with 2 and 0.1 mM creatinine, respectively. The peak areas from metabolomics and concentrations from targeted analyses were compared. **Appendix C**, **Figure C. 1** shows the results that are not shown here, and shows that most of the selected amino acids had an approximate linear correlation between peak area from metabolomics and quantitatively measured concentrations from targeted analysis. This correlation was more often observed and more prominent when samples were diluted to 0.1 mM rather than 2 mM creatinine, where increased concentration of amino acid gave increased peak areas. **Figure 3. 10 A** and **B** show the results in positive ionization mode for L-α-aminoadipic acid in samples that contained 2 mM creatinine, and in samples diluted to 0.1 mM creatinine, respectively. **Figure 3. 10 C** and **D** show the results in positive ionization mode for glycine in samples that contained 2 mM creatinine and samples diluted to 0.1 mM creatinine, respectively.

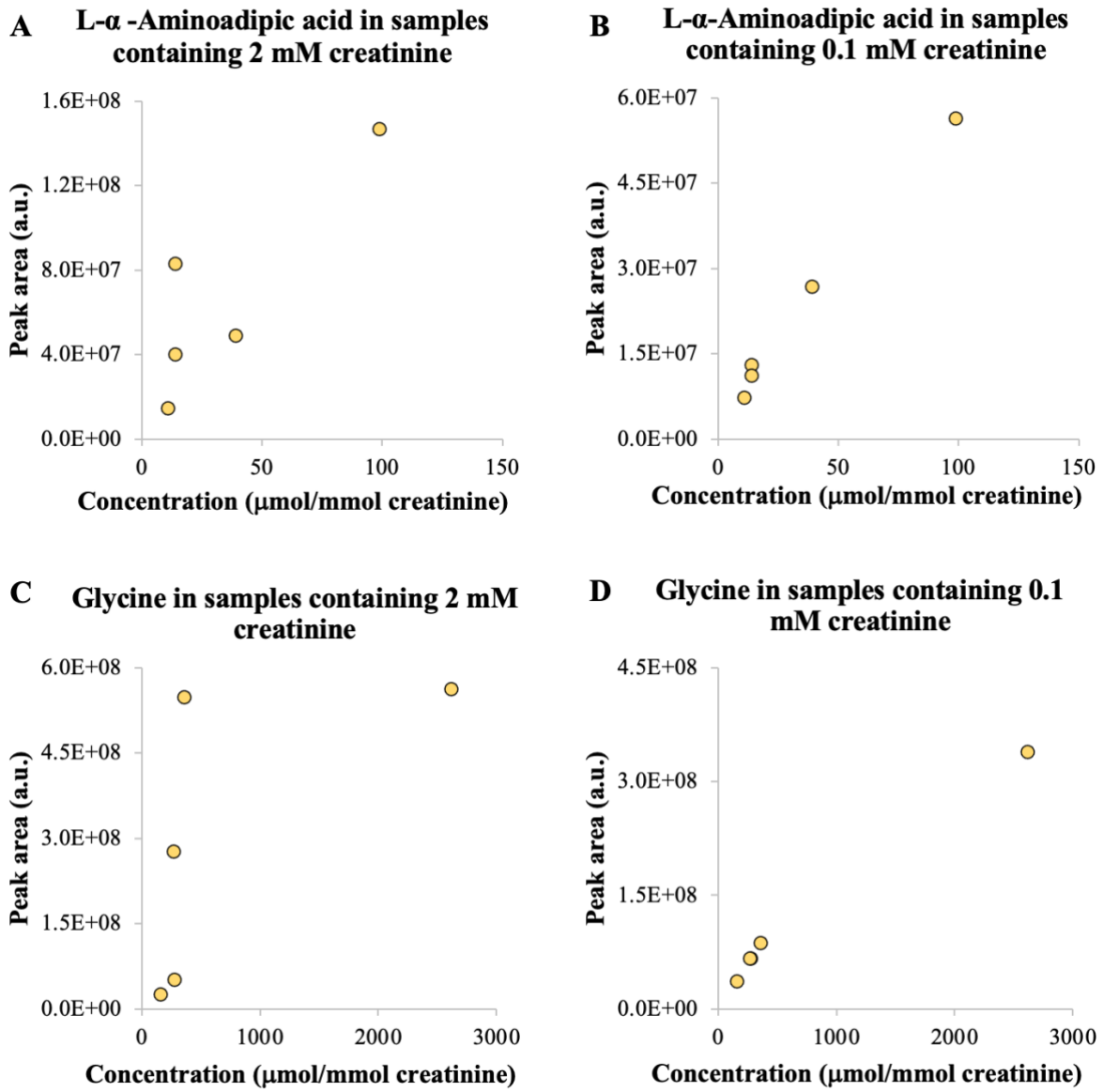


Figure 3. 10: Correlation between peak area and concentration ratios of amino acids relative to creatinine. *The figure shows that the best correlation between amino acid peak area and concentration ratios were found in samples diluted to 0.1 mM creatinine.*

The correlation between L-α-aminoadipic acid peak area and concentration ratios were better in samples diluted to 0.1 mM creatinine than 2 mM creatinine. The same evaluation was done for glycine peak area and concentration ratios. The ratio between the peak areas matched the ratio between the concentrations, and gave the impression of an approximate linear correlation.

Figure 3. 11 A and **B** show the results in positive ionization for L-tyrosine in samples that contained 2 mM creatinine, and sampled diluted to 0.1 mM creatinine, respectively. **Figure 3.**

11 C and **D** shows the results in positive ionization for formiminoglutamic acid in samples that contained 2 mM creatinine, and samples diluted to 0.1 mM creatinine, respectively.

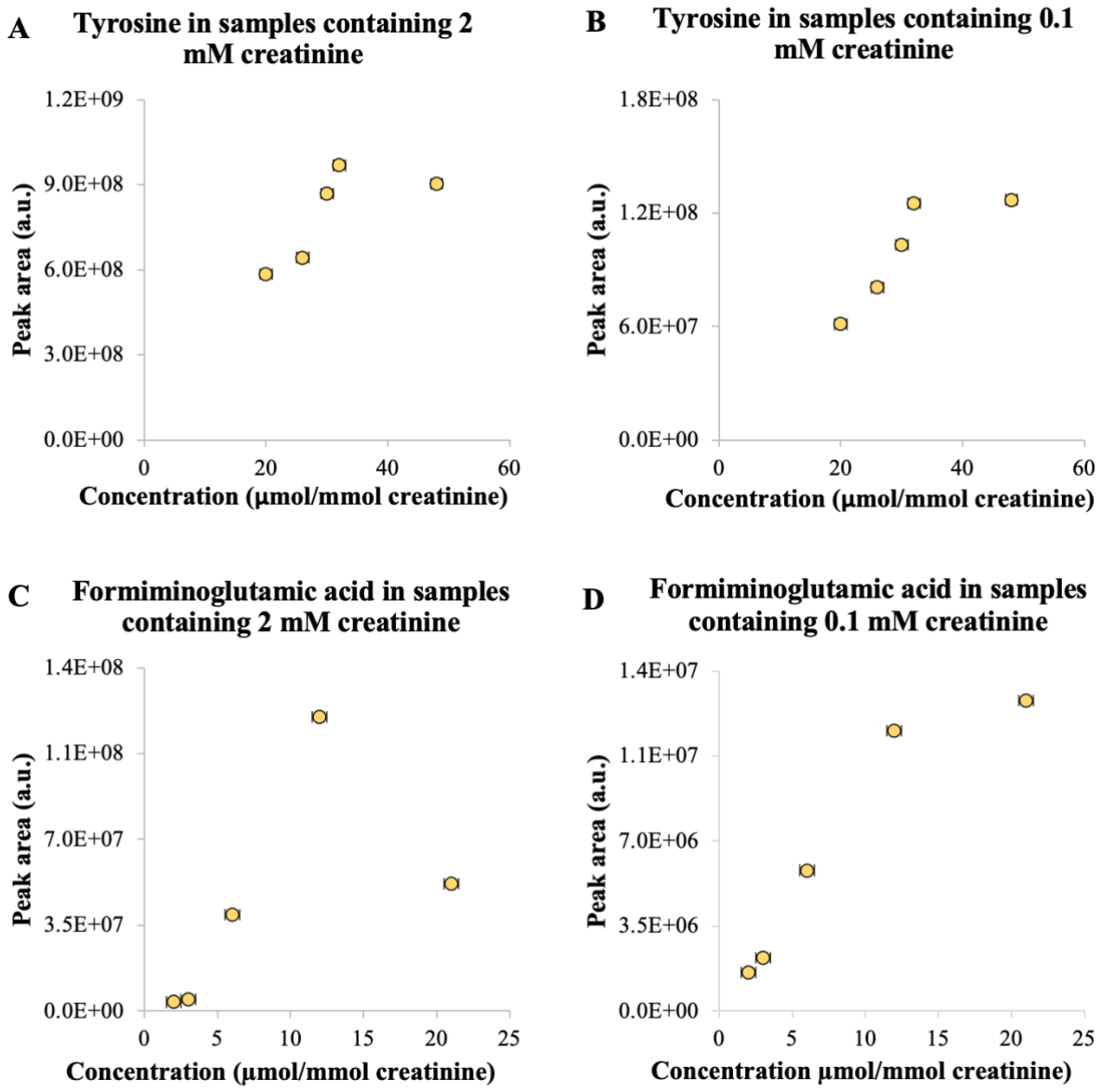


Figure 3. 11: L-tyrosine and formiminoglutamic acid peak area and concentration relative to creatinine. *There was a correlation between peak area from metabolomics and concentration from targeted analyses, but the ratio between peak areas did not match the ratios between the concentrations.*

L-tyrosine and formiminoglutamic acid had better correlation between peak area and concentration when the samples were diluted to 0.1 mM creatinine than 2 mM creatinine. Nonetheless, the ratios between peak areas did not match the ratios between the concentrations for samples with the highest concentrations after normalization to 0.1 mM creatinine. The correlation between peak area and concentration were therefore not linear for L-tyrosine and fomiminoglutamic acid.

Figure 3. 12 A and **B** show the results in positive ionization mode for L-histidine in samples that contained 2 mM creatinine, and in samples diluted to 0.1 mM creatinine, respectively. **Figure 3. 12 C** and **D** show the results in positive ionization mode for L-valine in samples that contained 2 mM creatinine, and in samples diluted to 0.1 mM creatinine, respectively.

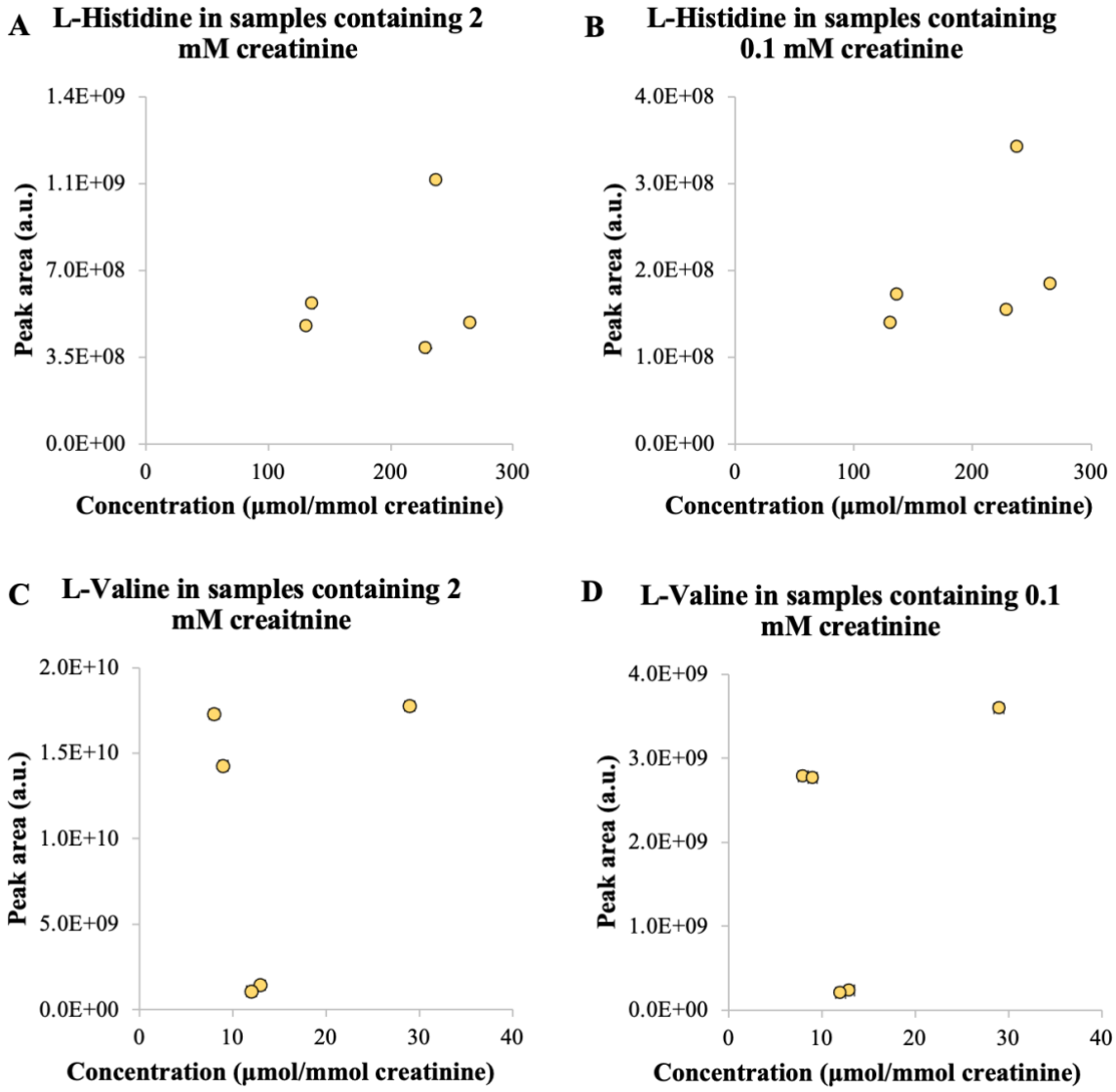


Figure 3. 12: No correlation between peak area and concentration. *L-histidine, L-valine, β-alanine, γ-amino-n-butyric acid, L-α-amino-n-butyric acid, and L-aspartic acid did not show a correlation between peak area and concentration.*

L-histidine, L-valine, L-aspartic acid, L-α-amino-n-butyric acid acid, γ-amino-n-butyric acid, and β-alanine did not show a linear correlation between peak area from metabolomics and concentration measured from targeted analysis in neither samples with 2 mM creatinine or 0.1 mM creatinine. **Figure 3. 12** show that samples containing the lowest L-histidine and L-valine

concentration (131 and 8 µmol/mmol, respectively) did not have lowest peak areas, and that samples that contained the highest L-histidine and L-valine concentration (265 and 29 µmol/mmol creatinine) did not have the highest peak areas.

Quantitative study of biomarkers for cystinuria and hypophosphatasia in urine samples

The biomarkers L-cystine, L-lysine, L-ornithine, and L-arginine for cystinuria were studied in five normal samples and in one pathologic cystinuria sample. Peak areas from metabolomics were compared quantitatively to measured concentrations from targeted analyses. The results are shown in **Figure 3. 13 A-D**.

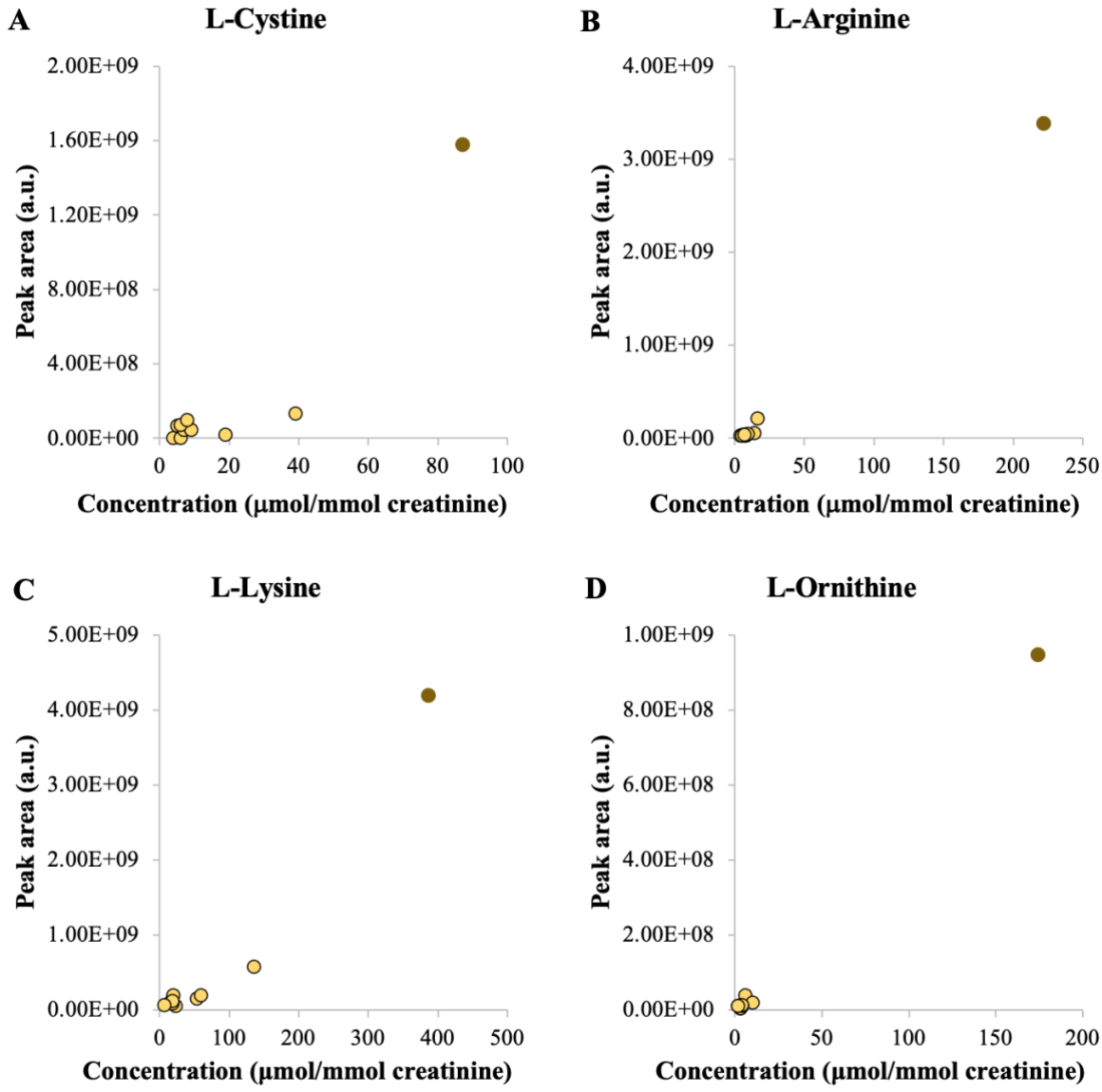


***Figure 3. 13: Quantitative study of biomarkers for cystinuria in urine samples. Figure 3.13 A, B, C** and **D** shows peak area from metabolomics compared to concentration ratios of L-cystine, L-arginine, L-lysine, and L-*

ornithine, respectively in normal samples and a pathologic sample. The yellow dots represent normal samples and the brown dots represent the cystinuria sample.

Figure 3. 13 shows that compared to the normal samples (yellow dots) the pathologic samples (brown dots) had higher peak area for all biomarkers for cystinuria studied in this thesis. The ratio between the peak area in the pathologic sample and normal samples matched the ratios between the concentration in the pathologic sample and normal samples for all cystinuria biomarkers. This means that biomarkers in pathologic levels, outside the reference ranges for normal levels, can be detected for cystinuria using the metabolomics method.

PEA was spiked into urine samples. Two of the samples were made to represent HPP and three normal samples were made. Peak areas from metabolomics were compared to known concentrations. **Figure 3. 14** shows the result.

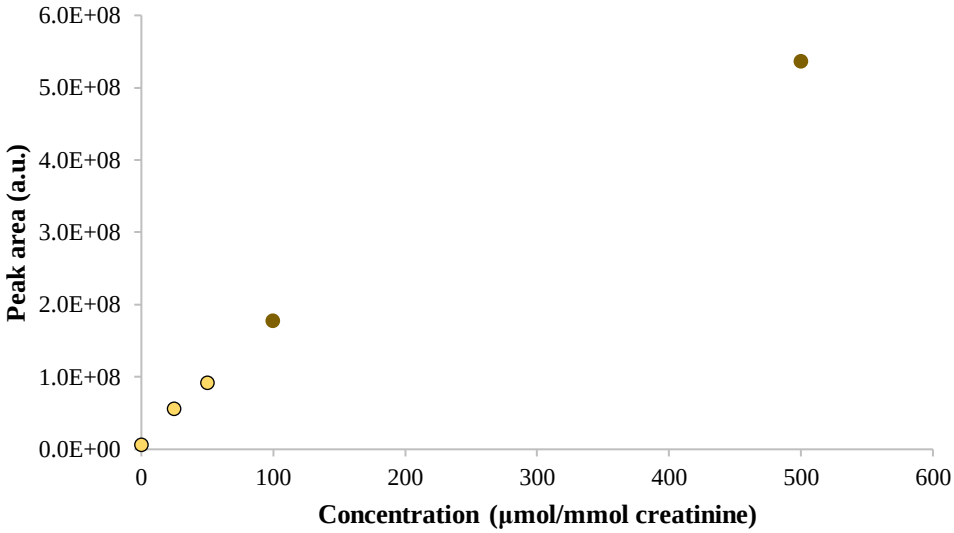


Figure 3. 14: Quantitative study of phosphoethanolamine spiked in urine samples. *Peak areas from metabolomics were compared to concentrations of PEA spiked to urine samples. The yellow dots represent normal samples and the brown dots represent pathologic levels of PEA and HPP.*

Figure 3. 14 shows that compared to normal samples (yellow dots) pathologic samples (brown dots) had higher PEA peak area, and that there was a linear correlation between peak area and concentration. This means that PEA in pathologic levels, outside the reference ranges for normal levels in newborns, can be detected for HPP using the metabolomics method.

3.5 Analysis of selected compounds in plasma

Selected amino acids and amino acid related compounds were detected in plasma samples using metabolomics. The precision of the metabolomics method for the selected compounds was determined, and matrix effects on the selected compounds were evaluated.

3.5.1 Detection of amino acids using metabolomics

All the selected compounds were detected in plasma using metabolomics, as well as PEA. The in-house library was used to distinguish between the chromatographic peaks for detection and characterization of amino acids in plasma samples.

3.5.2 Precision

The precision of the metabolomics method for detection of the selected amino acids in plasma samples was characterized studying amino acid peak area and retention time to the signal repeatability and intermediate precision for one plasma sample.

Repeatability

Amino acid peak area and retention time repeatability were studied in a plasma sample. The mean, standard deviation and %RSD were calculated for each amino acid. The result for amino acid repeatability is shown in **Appendix D**, **Table D. 1**. The metabolomics method showed low variation for retention time repeatability for all selected compounds and low peak area repeatability for most of the selected compounds (<30% RSD). For example L-anserine retention time repeatability had 0.27% RSD and peak area repeatability 2.09% RSD. L-α-amino-n-butyric acid and glycine peak area repeatability had high variation and was 297.0% and 134.3%, respectively.

Intermediate precision

Amino acid peak area and retention time intermediate precision were studied in a plasma sample. The mean, standard deviation and %RSD were calculated and the results for each amino acid are shown in **Appendix D**, **Table D. 2**. The metabolomics method showed high variation for retention time intermediate precision (>2% RSD), while low variation for peak area intermediate precision for most of the selected compounds (<30% RSD). For example L-anserine retention time and peak area intermediate precision was 3.2% and 26.0%, respectively.

Creatinine, L-homocysteine, L-anserine, L-cystine, ethanolamine, and L-sarcosine had high variation for peak area intermediate precision (>30%).

3.5.3 A study of matrix effects on selected compounds in plasma using metabolomics

The matrix effects on selected amino acids in plasma using metabolomics were evaluated. The amino acids included in this study are listed in **Appendix E, Table E. 1**, as well as the results. The table gives the retention time difference, ΔRT (min), and the matrix factor, MF (%). Chromatographic matrix effects (ΔRT >0.1 minutes) were found for L-arginine, L-anserine, ethanolamine, L-histidine, and 1-methyl-L-histidine and 3-methyl-L-histidine. Ion enhancement was found for β-alanine, with an elevated peak area in plasma spiked compared to the standard solution (MF >1, >30% RSD) comparing plasma sample and standard solution with the same amount amino acids added. Ion suppression was found for L-α-aminoadipic acid, L-anserine, L-arginine, L-carnosine, L-cystathionine, L-cystine, ethanolamine, glycine, L-histidine, L-homocysteine, δ-hydroxylysne, L-lysine, L-methyl-L-histidine and 3-methyl-L-histidine, L-ornithine, and L-proline, with a reduced peak area in plasma compared to the standard solution (MF <1, >30% RSD). **Figure 3. 15** show the chromatographic matrix effects and ion suppression of ethanolamine in plasma compared to the chemical standard.

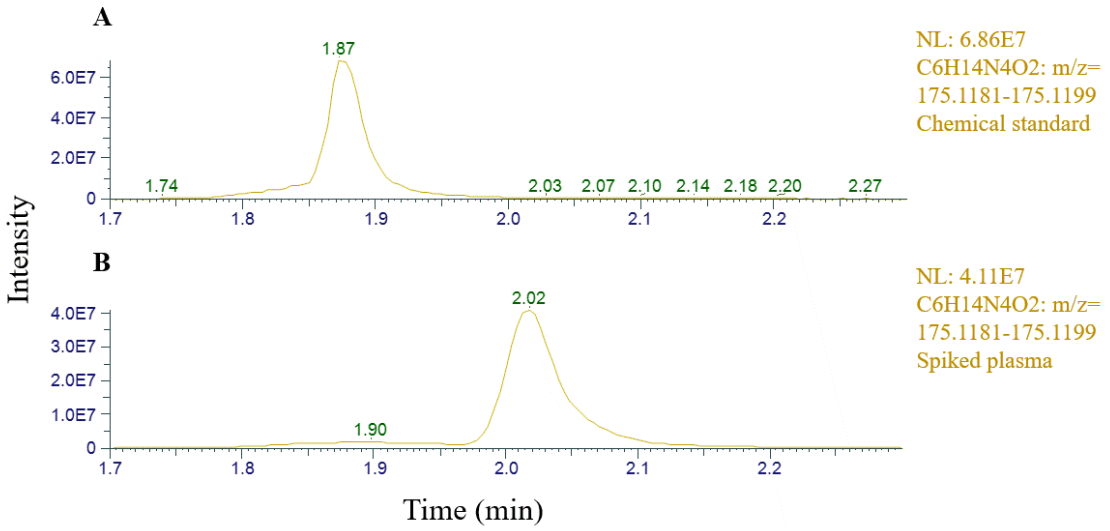


Figure 3. 15: Matrix effects for ethanolamine in plasma. *The figure shows the chromatographic matrix effects, and the ion suppression on ethanolamine in plasma (**B**) compared to standard solution (**A**), as an example of one amino acid affected by the matrix.*

Figure 3. 15 A shows ethanolamine retention time and peak area in the chemical standard, while **Figure 3. 15 B** shows ethanolamine retention time and peak area in plasma. In plasma ethanolamine had higher retention time (2.02 minutes) than in the chemical standard (1.87 minutes). Ethanolamine peak area was lower in the plasma sample than in the chemical standard due to ion suppression.

3.6 Summary of the results

Table 3. 3: Summary of the results for the study of amino acids in chemical standards, urine, and plasma samples.

Amino acid	Detectable in positive ionization mode	Detectable in negative ionization mode	Endogenous in plasma	Endogenous in urine	Retention time shift in plasma	Ion suppression /enhancement plasma	Linear correlation in 0.1 mM creatinine urine
Ethanolamine	Yes	No	Yes	Yes	Yes	Yes	N.D.*
Glycine	Yes	Yes	Yes	Yes	No	Yes	Yes
β-Alanine	Yes	N.S.	No	Yes	No	Yes	No
L-Alanine	Yes	N.S.	Yes	No	No	No	N.D.*
L-Sarcosine	Yes	N.S.	No	No	No	No	N.D.*
L-α-amino-n-butyric acid	Yes	N.S.	Yes	Yes	No	No	No
γ-amino-n-butyric acid	Yes	N.S.	No	Yes	No	No	No
D,L-β-aminoisoburytic acid	Yes	N.S.	Yes	Yes	No	No	N.D.*
L-Serine	Yes	Yes	Yes	Yes	No	No	Yes
Creatinine	Yes	Yes	Yes	Yes	No	No	Yes
L-Proline	Yes	Yes	Yes	Yes	No	Yes	Yes
L-Valine	Yes	Yes	Yes	Yes	No	No	No
L-Threonine	Yes	Yes	Yes	Yes	No	No	Yes
Taurine	Yes	Yes	Yes	Yes	No	No	N.D.*
Hydroxy-L-proline	Yes	Yes	Yes	Yes	No	No	Yes
L-Isoleucine, L-Leucine, Alloisoleucine	N.S.	N.S.	Yes	Yes	No	No	Yes
Asparagine	Yes	Yes	N.D.***	Yes	N.D.***	No	Yes
L-Ornithine	Yes	Yes	Yes	Yes	No	No	No
L-Aspartic acid	Yes	Yes	Yes	Yes	No	No	No
Phospho-ethanolamine	Yes	Yes	Yes	Yes	No	No	Yes
L-Glutamine	Yes	Yes	N.D.***	Yes	No	No	Yes
L-Lysine	Yes	Yes	Yes	Yes	No	Yes	Yes
L-Glutamic Acid	Yes	Yes	Yes	Yes	No	No	Yes
L-Methionine	Yes	Yes	Yes	Yes	No	Yes	No
L-Histidine	Yes	Yes	Yes	Yes	Yes	No	No
L-α-aminoadipic acid	Yes	Yes	Yes	Yes	No	Yes	Yes

Table 3. 3 *continues on the next page.*

Amino acid	Detectable in positive ionization mode	Detectable in negative ionization mode	Endogenous in plasma	Endogenous in urine	Retention time shift in plasma	Ion suppression /enhancement plasma	Linear correlation in 0.1 mM creatinine urine
δ-Hydroxylysine	Yes	Yes	No	Yes	No	Yes	N.D.*
4-Hydroxy-L-glutamic acid	Yes	Yes	N.D.***	Yes	N.D.***	No	Yes
L-Phenylalanine	Yes	Yes	Yes	Yes	No	No	Yes
1-Methyl-L-histidine, 3-Methylhistidine	N.S.	N.S.	Yes	Yes	Yes	Yes	No
Glycylproline	Yes	Yes	N.D.***	Yes	N.D.***	No	N.D.*
Formimino-glutamic acid	Yes	Yes	N.D.***	Yes	N.D.***	No	No
L-Arginine	Yes	Yes	Yes	Yes	Yes	Yes	Yes
L-Citrulline	Yes	Yes	Yes	Yes	No	No	Yes
L-Tyrosine	Yes	Yes	Yes	Yes	No	No	No
Homocitrulline	Yes	Yes	N.D.***	Yes	N.D.***	No	N.D.**
S-Sulfocysteine	Yes	Yes	N.D.***	Yes	N.D.***	No	No
L-Tryptophan	Yes	Yes	Yes	Yes	No	No	Yes
L-Cystathionine	Yes	Yes	No	Yes	No	No	Yes
3-Hydroxy-D,L-kynurenine	Yes	Yes	N.D.***	Yes	N.D.***	No	N.D.**
L-Carnosine	Yes	Yes	No	Yes	No	No	N.D.*
L-Anserine	Yes	Yes	No	Yes	Yes	No	N.D.*
L-Cystine	Yes	Yes	Yes	Yes	No	Yes	No
Homocarnosine	Yes	Yes	N.D.***	Yes	N.D.***	No	N.D.**
L-Homocysteine	Yes	Yes	No	No	No	No	N.D.*
Saccharopine	Yes	Yes	N.D.***	Yes	N.D.***	No	Yes
Argininosuccinic acid	Yes	Yes	N.D.***	Yes	N.D.***	No	No
Aspartyl-glucosamine	Yes	Yes	N.D.***	Yes	N.D.***	No	Yes
S-Adenosyl-L-homocysteine	Yes	Yes	N.D.***	Yes	N.D.***	No	N.D.**

* *= Concentration was not measured quantitatively using targeted analyses.*

** *= Was not detected in urine samples diluted to 0.1 mM creatinine using the metabolomics method.*

*** *= Was not included in this study.*

N.D. = Not detected.

N.S. = Not chromatographic separated.

4. Discussion

Metabolomics is the study of the metabolome in a biological sample, giving high coverage of metabolites compared to targeted analyses used in diagnostics of IEMs today. A metabolomics method with high precision for detection of metabolites in easily accessible biological materials e.g. urine and plasma is useful. The level of confidence for identification in metabolomics gives some criteria for metabolite identification and can be provided by creating an in-house library (38, 40, 82, 83). The aim of this study was to evaluate the quantitative and qualitative performance of metabolomics studying amino acids and amino acid related compounds in urine and plasma samples. This included comparing peak areas determined by the established metabolomics method HPLC – ESI – Q Exactive Orbitrap MS with quantitatively measured concentrations obtained with established targeted analyses used in laboratory diagnostics. Furthermore, this included evaluating the quantitative and qualitative results from metabolomics on the selected compounds studying chemical standards, urine, plasma and matrix effects.

4.1 Detection and characterization of selected metabolites

Chemical standards were analyzed using the metabolomics method for qualitative analysis, and all the 52 selected amino acids and amino acid related compounds were detected, identified and annotated in an in-house library. The findings of this study show that all selected compounds can be detected using this metabolomics method, also PEA that cannot be detected by targeted analyses in the diagnostic laboratory of IEMs at Oslo University Hospital – Rikshospitalet today. This will extend the number of compounds that might be detected, providing a greater coverage of metabolomics studies. Studying the chemical standards, all selected compounds were detected in positive and negative ionization mode, except ethanolamine that was not detected in negative ionization mode. Therefore, metabolomics studies of ethanolamine will lack the comprehensive coverage from negative ionization mode. Not surprisingly, this could be because the hydroxyl group is not easily protonated.

With the chromatographic parameters used in this thesis L-isoleucine, L-leucine, and alloisoleucine were not fully chromatographically separated, nor were 1-methyl-L-histidine and 3-methyl-L-histidine separated. However, in **Figure 3. 5** L-isoleucine, L-leucine, and alloisoleucine shows hint of chromatographic separation, and that the fragmentation spectra

could be used to tell if there is more of one or the other amino acid if one side of the peak is higher than the other. Nevertheless, since neither of them had any signature fragments none of them can be excluded to be present in the sample. To distinguish between L-leucine, L-isoleucine and alloisoleucine another separation method should be used. As described in a review by Ferré et al., research show that ion pair chromatography can be used to separate L-isoleucine, L-leucine and alloisoleucine because of the differences in hydrophobicity (57). Even though ion pairing chromatography is a robust separation method for polar compounds and therefore used in targeted analyses of these isomeric amino acids in the diagnostic laboratory at Rikshospitalet, the disadvantage is system contamination. This is due to the use of organic ion pairing reagents in the mobile phase and system contamination, and gradient elution with high ratio of organic solvent for separation of hydrophobic components (43, 57, 104). HILIC is another alternative, and other studies show that L-isoleucine and L-leucine were separated chromatographically with HILIC-columns (105, 106). However, both ion pair chromatography and HILIC can require a high amount of organic solvent, e.g. acetonitrile which is not Eco-friendly (58). Furthermore, other studies showed that 1-methyl-L-histidine and 3-methyl-L-histidine has been chromatographic separated with capillary electrophoresis (107, 108). Capillary electrophoresis can provide high resolution efficiency for separation of amino acids (109). However, the disadvantage is that the robustness and sensitivity are lower than for LC which limits its applications for metabolomics analyses (110, 111). The former results suggest that ion pair chromatography should be used for chromatographic separation of L-isoleucine and L-leucine, and that capillary electrophoresis should be used for separation of 1-methyl-L-histidine and 3-methyl-L-histidine if chromatographic separation and further information is required, other than the information obtained from metabolomics.

To summarize, almost all the selected compounds from this study were identified at level 1 identification, described in **Chapter 1.4.5**, as these metabolites will be annotated in the in-house library. The exception is for L-leucine, L-isoleucine, alloisoleucine, 1-methyl-L-histidine, and 3-mehtyl-L-histidine which were not chromatographically separated using the metabolomics method.

For prediction of retention time and identification of new metabolites in future metabolomics studies, the correlation between structure for the selected compounds was evaluated. The different parameters of the molecular structure studied were the total number of rings in the structure, the number of benzene rings, the total number of double bonds, the number of

conjugated double bonds, the number of sulfur and nitrogen atoms, solubility, logP-value, and the molecular weight. In this study, no correlation was found between the structural parameters evaluated and the retention time for the selected compounds. However, the selected compounds with benzene in their structure and less polar compounds had the longest retention times, as shown in **Figure 3. 6**. These amino acids have similar structural geometry to the diphenyl material in the stationary phase in the analytical column due to the benzene rings in their structure. Therefore, these amino acids had stronger affinity to the diphenyl stationary phase material and thereby longer retention time. Compounds with shorter retention time, e.g. ethanolamine did not have ring structure, were polar and thereby more soluble in the mobile phase than other less polar amino acids. Nevertheless, L-isoleucine, L-leucine, and alloisoleucie did not have any ring structure, but had longer retention times than many of the less polar compounds, at 3.42 minutes, as shown in **Figure 3. 1**. L-histidine and L-anserine had aromatic ring structures but had shorter retention times, at 1.80 and 1.84 minutes, respectively. L-histidine and L- anserine also do not have benzene rings, but a five-ring that does not interact as strongly with diphenyl. The weakness of this study was that the selected compounds had very similar chemical properties. Other studies show that the use of retention time as a supplement to the accurate mass identification has been of significant advantage (112-114). However, the retention time could vary between samples or between analyses making it challenging to be used for identification in metabolomics. This could probably be avoided using retention time indexes to normalize the retention times. Zheng et al. (115) demonstrated a chemical labeling-based HPLC retention index strategy as a promising technology for metabolomics. Newer tools have also been developed for prediction of retention time based on structure using R-based data program (116) and other computer programming approaches (117) that has been shown to be useful tools for a more effective metabolite annotation and identification. In future studies the correlation between structure and retention time should be evaluated for several other chemical groups with other chemical properties and the different groups should be compared.

4.2 Precision of the metabolomics method on selected compounds in urine and plasma

Urine and plasma are the most widely used biological materials for laboratory diagnostics. To have a metabolomics method with good precision for analysis of these sample types is therefore a great advantage. Most of the selected compounds were identified at level 1 identification

using the in-house library in urine and plasma, except for the amino acids that were not chromatographically separated. The precision of the metabolomics method was investigated for amino acids and amino acid related compounds in urine and plasma. The metabolomics method showed low variation for retention time repeatability (<2% RSD) for both urine and plasma, as shown in **Appendix B Table B. 1** and **Appendix D Table D. 1**, respectively. It also showed that the peak area repeatability was good for the selected compounds in urine (<30%). The peak area repeatability was also good for most of the selected compounds in plasma, except for L-α-amino-n-butyric acid and glycine, an unexpected finding with no obvious reason. The stability of the system over time was studied with system suitability samples in all analyses, and to conclude the systematic variation from the instrument was found to be low.

Furthermore, the metabolomics method showed good peak area intermediate precision (<30% RSD) for most of the selected compounds in urine as shown in **Appendix B Table B. 2**. Surprisingly, L-α-aminoadipic acid, hydroxykynurenine, and S-adenosyl-L-homocysteine had high variation for peak area intermediate precision (>30% RSD) in urine. Creatinine also showed high variation for peak area intermediate precision, as shown in **Table 3. 2**. In plasma, creatinine, L-homocysteine, L-anserine, L-cystine, ethanolamine, and L-sarcosine had high variation for peak area intermediate precision. One explanation for these findings could be the high peak area, which means these metabolites were present at high concentrations, leading to detector saturation. The retention time intermediate precision had high variation for most of the selected compounds in urine and plasma (>2% RSD), except for L-α-aminoadipic acid in urine and L-methionine and L-tryptophan in plasma. There was no obvious reason for these findings. Other studies with acceptable retention time variation <5% RSD and compared to these studies, the retention time variations found in this study were low (84, 118, 119). The variation for peak area intermediate precision also agreed with the findings in this study, that variation >30% RSD was found for some of the selected compounds studied. As mentioned, retention time is an important tool for metabolite identification in metabolomics comparing results from the same lab. However, it will not be used by itself, but in addition to fragmentation spectra for distinguishing between isobaric compounds and the localization of chromatographic peaks relative to each other using the in-house library. Internal standards and calibration curves for each metabolite cannot be used in metabolomics, unlike targeted analyses, because of the aim of discover as many metabolites as possible. However, as reviewed by Broadhurst et al. (79), increased data quality and better intermediate precision in metabolomics are obtained

conditioning the analytical system and stabilizing the detector response using pooled quality controls. The results from this characterization of the precision for the metabolomics method and compared to other studies, suggest that the retention time intermediate precision <5% RSD should be considered as acceptable limits for retention time variation. This assertion is based on the non-systematic errors that can occur in metabolomics studies and that variation <2% RSD is too strict comparing results from different metabolomics analyses.

4.3 Quantitative analyses of urine and plasma

The quantitative performance of the metabolomics method was evaluated for selected compounds in urine samples, comparing data from metabolomics and targeted analyses. Lastly, matrix effects were studied for the selected compounds in plasma.

Quantitative analysis comparing data from metabolomics and targeted analyses

Peak areas for the selected compounds determined by metabolomics were compared with concentrations measured obtained from targeted analyses to characterize the quantitative performance of the metabolomics method. Creatinine retention time and peak area in samples from different individuals normalized to 2 mM creatinine gave high variation (>2% and >30% RSD, respectively). This might be because the matrix varies between samples causing different retention times, and that the matrix contain more chemical compounds with similar chemical properties and with the same retention time as creatinine (2.40 minutes). This could lead to ion suppression of creatinine and lower creatinine peak area. Other studies have also found creatinine tautomers, and this could be an explanation for the split creatinine peak for and for the high variation of creatinine retention time (120, 121), and could be an explanation to why the first part of the peak was elevated after the sample was spiked with creatinine, as shown in **Figure 3. 7**. Nonetheless, a lower creatinine retention time and peak area variation was found in samples diluted to 0.1 mM creatinine than samples with 2, 1, and 0.5 mM creatinine, as shown in **Figure 3. 8**. These results indicate that the split creatinine peak was due to matrix effects rather than creatinine tautomers, and explain the high variation for creatinine peak area intermediate precision from samples diluted to 2 mM creatinine. Furthermore, the correlation between peak area determined by metabolomics and measured concentrations obtained from targeted analyses for the selected compounds in urine was also better when samples were normalized to 2 mM creatinine than 1 mM creatinine, and the correlation were observed to be

approximately linear. Few samples were studied. Therefore, more samples should be studied for a linear regression analysis to be performed. These results still pens the idea of diluting urine samples to 0.1 mM creatinine instead of 2 mM creatinine. Nevertheless, there is a possibility that diluting urine samples too much can result in the loss of signal for some metabolites, making them undetectable. This problem occurred for S-adenosyl-L-homocysteine, and homocarnosine that were not detected in urine samples diluted to 0.1 mM creatinine. Amino acids that did not show a linear correlation between peak areas from metabolomics and concentrations from targeted analyses were γ-amino-n-butyric acid, α-amino-n-butyric acid, aspartic acid, L-histidine, L-valine, and β-alanine. The signal intensities for these amino acids were high ($>10^8$), and as for the creatinine results, this gave the impression that urine samples should be diluted even more for better quantitative data of the selected compounds. Other amino acids that did not show a linear correlation between peak areas and concentrations were present in lower concentrations e.g. L-glutamic acid, L-methionine, and S-sulfocysteine. This might be because the measured concentrations from targeted analyses were available with SD of 0.5 which affect the results studying lower concentrations of metabolites in samples. To summarize, the study suggests that quantitative data from the metabolomics method should be obtained for the selected amino acids and amino acid related compounds in urine samples diluted to 0.1 mM creatinine or less.

For evaluation of the quantitative performance of the metabolomics method in a diagnostic perspective, one cystinuria sample and two urine samples spiked with different concentrations of PEA to represent HPP were analyzed using the metabolomics method. The correlation of PEA peak area and concentration showed good linearity, as shown in **Figure 3. 13** and **Figure 3. 14**, respectively. Biomarkers for cystinuria e.g. L-ornithine and L-cystine did not show a good correlation between peak area from metabolomics and concentrations from targeted analyses when the five normal samples were normalized to 0.1 mM creatinine, as shown in **Appendix C**, **Figure C. 1**. However, cystinuria markers in the pathologic sample showed peak area that were a lot higher than in normal samples, and that the ratios between peak areas matched the ratios between the concentrations. Therefore, an approximate linear correlation between peak area from metabolomics and concentrations from targeted analyses was found, even though these samples were normalized to 2 mM creatinine. In future testing of the metabolomics method several biomarkers of IEMs in pathologic samples should be normalized to different creatinine concentrations to decide what concentration is optimal to obtain quantitative data, and for detection of as many metabolites as possible.

The creatinine concentration can be used for normalization of metabolite concentrations in urine (96, 98), and results from this study suggest that urine samples should be diluted to a lower concentration than 2 mM creatinine. Normalization methods other than creatinine concentration used are osmolality, urine volume, and MSTUS. The advantage of urinary metabolite normalization to the creatinine concentration is the stable creatinine excretion in urine dependent on gender, age, muscle mass, hydration, diet and activity level, and reflects the overall concentration of metabolites in the urine (11-13). As seen in this study, the creatinine signal peak is visibly split into two overlapping peaks in some samples, and the peak area and retention time had high variation in urine samples that were not diluted. Even though several studies show that normalization to creatinine concentration gives acceptable results (96, 122, 123), other studies have found that normalization to creatinine concentration gives high variation and therefore suggest that urine samples should be normalized by other techniques e.g. osmolality and MSTUS (11, 97, 124, 125). On the other hand, these studies had larger number of samples and Q-TOF were used, while in this thesis a sample size of five to ten was analyzed with an Orbitrap mass analyzer. However, Orbitrap mass analyzers provides higher resolution than Q-TOF. Which normalization method that should be used might therefore depend on the study design. Other studies suggest that the combination of two normalization techniques should be used because of the variation between urine samples from different individuals (13, 126). Based on the results from this study it cannot be excluded that other normalization methods should be tested to see if both qualitative and quantitative results can be obtained and improved for amino acids and related compounds in urine samples.

Matrix effects on selected compounds in plasma

Chromatographic matrix effects, ion suppression, and ion enhancement on the selected compounds in plasma were studied. L-anserine and ethanolamine had retention time shift, or chromatographic matrix effects (ΔRT >0.1 minutes). Fronting and wider peaks were observed for L-methionine and L-tyrosine in plasma. Many chemical compounds in a plasma sample probably have the same retention time as L-anserine, ethanolamine, L-methionine, and L-tyrosine leading to competition of binding to the stationary phase material, and thus longer retention time compared to the standard solution. It also gives wider peaks because the chemical compounds spend a longer time in the chromatographic column. In other studies retention time shifts have been defined as either 0.2, 0.5, and 1 minutes variation (127, 128). Compared to these studies, retention time shift of <0.1 minutes used in this study might be too strict.

Furthermore, L-α-aminoadipic acid, L-cystine, glycine, δ-hydroxylysine, L-lysine, L-ornithine, L-proline, and 1-methyl-L-histidine and 3-methyl-L-histidine were exposed to ion suppression, electrospray matrix effects, as shown in **Appendix E**, **Table E. 1**. Amino acids studied in this thesis had very similar retention times, around two minutes or less. Probably there are more chemical compounds with the same retention time that will compete for protonation and ionization in the ESI, resulting in ion suppression of the amino acids mentioned. The degree of ion suppression is dependent on the chromatographic separation of metabolites and the ratio between metabolite and matrix concentrations (86). Therefore, it might be higher abundance of other chemical compounds in the matrix with the same retention time that leads to ion suppression of the amino acids mentioned. Formic acid in the mobile phase also contributes to ion suppression. As described in a review by Annesley, T., M. (86) and found by King, R. et al. (129), the main reason for ion suppression is the presence high concentrations of nonvolatile or less volatile solutes in the spray droplet with the analyte, and occurs using ESI, but also other ion sources like APCI, and can lead to less effective evaporation of the solvent and less ions are produced for analysis. MALDI does not show the same problem with ion suppression. However, the lack good signal precision makes MALDI less suitable for quantitative analysis (68).

To summarize, these results indicate that less matrix effects are found in plasma compared to urine. This agrees with other studies, that has shown that the urine matrix is more variable than the plasma matrix (130). However, this observation is based on the results from urine samples diluted to 2 mM creatinine, and the fact that urine samples were not spiked with the selected compounds. Therefore, chromatographic matrix effects, ion suppression, and ion enhancement in urine should be studied in future projects.

5. Conclusion

A qualitative metabolomics study of amino acids and amino acid related compounds in urine and plasma samples was performed using the HPLC – ESI – Q Exactive Orbitrap MS metabolomics method. Almost all the selected compounds were detected at level 1 identification except for L-leucine, L-isoleucine, 1-methyl-L-histidine, and 3-methyl-L-histidine. The metabolomics method is a useful diagnostic tool, as the diagnostic metabolite PEA, which is not quantified by established targeted methods in the diagnostic laboratory of IEMs at Oslo University Hospital – Rikshospitalet, was successfully detected. The precision of the metabolomics method on retention time and peak area for amino acids in urine and plasma was good with the %RSD within acceptable limits for retention time and peak area variation for most of the selected compounds studied (<2% and <30% RSD, respectively). However, in urine samples normalized to 2 mM creatinine the creatinine retention time and peak area had higher variation than in samples normalized to 0.1 mM creatinine. A correlation between peak area from metabolomics and concentration from targeted analyses was found to be approximately linear for most of the selected compounds in urine when samples were diluted to 0.1 mM creatinine. Therefore, this study suggests that urine samples should be diluted to 0.1 mM creatinine or less to obtain quantitative data from the metabolomics method in studies of amino acids and amino acid related compounds. The qualitative performance of the metabolomics method in studies of the selected compounds in urine was better when samples were normalized to 2 mM creatinine than 0.1 mM because some of the selected compounds were not detected in samples diluted to 0.1 mM creatinine. Lastly, chromatographic matrix effects, ion suppression and ion enhancement on the selected amino acids and amino acid related compounds in plasma were evaluated. It was concluded that most of the selected compounds were affected by matrix effects in plasma, especially ion suppression. Nevertheless, less variation was observed in plasma samples compared to urine samples based on the study of precision of the metabolomics method.

6. Future perspectives

A qualitative analysis of amino acids and related compounds was performed in this thesis and an in-house library was made for the selected compounds. If level 1 identification is wanted for other metabolites, more chemical standards of other chemical compounds should be identified and annotated. Furthermore, the correlation between metabolite structure and retention time should be evaluated comparing different chemical compounds with different chemical properties for retention time prediction and identification of new metabolites in future metabolomics studies.

The quantitative performance of the metabolomics method was tested, mainly for the selected compounds in urine. However, a larger sample size should be tested and studied to decide if the correlation between peak area from metabolomics and concentration from targeted analyses is linear, as the results in this thesis indicate for most of the selected compounds in urine. The results from this study also indicated that less matrix effects are found in plasma samples using the metabolomics method than urine samples, but qualitative analysis comparing peak areas from metabolomics to concentrations from targeted analyses needs to be done for the selected compounds in plasma studying a larger sample size.

PEA spiked to urine samples showed promising results for diagnosis of HPP using the metabolomics method. This should be confirmed in by studying samples from a patient confirmed with HPP. Furthermore, metabolomics method should be tested for its usefulness in diagnostics of IEMs.

7. References

1. Milne SB, Mathews TP, Myers DS, Ivanova PT, Brown HA. *Sum of the parts: mass spectrometry-based metabolomics.* Biochemistry. 2013;52(22):3829-40.
2. Vangala S, Tonelli A. *Biomarkers, metabonomics, and drug development: Can inborn errors of metabolism help in understanding drug toxicity?* The AAPS Journal. 2007;9(3):E284-E97.
3. Madhava M. *A clinical guide to inherited metabolic diseases.* Arch Dis Child. 2006;91(10):875-.
4. Beecher CWW. The Human Metabolome. In: Harrigan GG, Goodacre R, editors. *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Boston, MA: Springer US; 2003. p. 311-9.
5. Tymoczko JL, Berg JM, Stryer L. *Biochemistry: A Short Course*: 3rd. ed.: Macmillan Learning; 2015. p. 38-9.
6. Neuberger A. *Stereochemistry of Amino Acids.* In: Anson ML, Edsall JT, editors. Adv Protein Chem. 4: Academic Press; 1948. p. 297-383.
7. McMurry J. *Organic chemistry : with biological applications.* 3rd. ed. Stamford, Conn: Cengage Learning; 2015. p. 678-9, 682, 119.
8. Manners CN, Payling DW, Smith DA. *Distribution coefficient, a convenient term for the relation of predictable physico-chemical properties to metabolic processes.* Xenobiotica. 1988;18(3):331-50.
9. Wu G. *Amino acids: metabolism, functions, and nutrition.* Amino Acids. 2009;37(1):1-17.
10. Williams RA, Mamotte CD, Burnett JR. *Phenylketonuria: an inborn error of phenylalanine metabolism.* Clin Biochem Rev. 2008;29(1):31-41.
11. Mervant L, Tremblay-Franco M, Jamin EL, Kesse-Guyot E, Galan P, Martin JF, et al. *Osmolality-based normalization enhances statistical discrimination of untargeted metabolomic urine analysis: results from a comparative study.* Metabolomics. 2021;17(1):2.
12. Narayanan S, Appleton HD. *Creatinine: a review.* Clin Chem. 1980;26(8):1119-26.
13. Warrack BM, Hnatyshyn S, Ott KH, Reily MD, Sanders M, Zhang H, et al. *Normalization strategies for metabonomic analysis of urine samples.* J Chromatogr B Analyt Technol Biomed Life Sci. 2009;877(5-6):547-52.

14. Nickolas TL, Barasch J, Devarajan P. *Biomarkers in acute and chronic kidney disease.* Curr Opin Nephrol Hypertens. 2008;17(2):127-32.
15. Strachan T, Read AP. *Human Molecular Genetics*: Taylor & Francis Group; 2019. p. 3.
16. Garg U, Smith LD. Chapter 1 - Introduction to laboratory diagnosis and biomarkers in inborn error of metabolism. In: Garg U, Smith LD, editors. *Biomarkers in Inborn Errors of Metabolism.* San Diego: Elsevier; 2017. p. 1-24.
17. General principles. In: Clarke JTR, editor. *A Clinical Guide to Inherited Metabolic Diseases.* 3rd. ed. Cambridge: Cambridge University Press; 2005. p. 1-27.
18. Ferreira CR, van Karnebeek CDM, Vockley J, Blau N. *A proposed nosology of inborn errors of metabolism.* Genet Med. 2019;21(1):102-6.
19. Tebani A, Abily-Donval L, Afonso C, Marret S, Bekri S. *Clinical Metabolomics: The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era*. Int J Mol Sci. 2016;17(7):1167.
20. Cajka T, Fiehn O. *Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics.* Anal Chem. 2016;88(1):524-45.
21. Linglart A, Biosse-Duplan M. *Hypophosphatasia.* Curr Osteoporos Rep. 2016;14(3):95-105.
22. Mattoo A, Goldfarb DS. *Cystinuria.* Semin Nephrol. 2008;28(2):181-91.
23. Milliner DS. *Cystinuria.* Endocrinol Metab Clin North Am. 1990;19(4):889-907.
24. Rogers A, Kalakish S, Desai RA, Assimos DG. *Management of Cystinuria.* Urol Clin North Am. 2007;34(3):347-62.
25. Seksjon for medfødte metabolske sykdommer: Oslo Universitetssykehus - Rikshospitalet; 2020 [Available from: https://oslo-universitetssykehus.no/fag-og-forskning/laboratorietjenester/medisinsk-biokjemi/seksjon-for-medfodte-metabolske-sykdommer.
26. Mak CM, Lee HC, Chan AY, Lam CW. *Inborn errors of metabolism and expanded newborn screening: review and update.* Crit Rev Clin Lab Sci. 2013;50(6):142-62.
27. Kamboj M. *Clinical approach to the diagnoses of inborn errors of metabolism.* Pediatr Clin North Am. 2008;55(5):1113-27.
28. Pourfarzam M, Zadhoush F. *Newborn Screening for inherited metabolic disorders; news and views.* J Res Med Sci. 2013;18(9):801-8.
29. Rikshospitalet OUh-. Nyfødtscreening 2021 [Available from: https://oslo-universitetssykehus.no/avdelinger/barne-og-ungdomsklinikken/nyfodtscreeningen/nyfodtscreening.

30. Andermann A, Blancquaert I, Beauchamp S, Déry V. *Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years.* Bull World Health Organ. 2008;86(4):317-9.
31. Rikshospitalet OUh-. Nasjonal behandlingstjeneste for screening av nyfødte og avansert laboratoriediagnostikk ved medfødte stoffskiftesykdommer 2018 [Available from: https://oslo-universitetssykehus.no/fag-og-forskning/nasjonale-og-regionale-tjenester/nasjonal-behandlingstjeneste-for-screening-av-nyfodte-og-avansert-laboratoriediagnostikk-ved-medfodte-stoffskiftesykdommer.
32. Ezgu F. *Chapter Seven - Inborn Errors of Metabolism.* In: Makowski GS, editor. Adv Clin Chem. 73: Elsevier; 2016. p. 195-250.
33. Gertsman I, Barshop BA. *Promises and pitfalls of untargeted metabolomics.* J Inherit Metab Dis. 2018;41(3):355-66.
34. Tebani A, Afonso C, Marret S, Bekri S. *Omics-Based Strategies in Precision Medicine: Toward a Paradigm Shift in Inborn Errors of Metabolism Investigations.* Int J Mol Sci. 2016;17(9):1555.
35. Licata AA, Radfar N, Bartter FC, Bou E. *The urinary excretion of phosphoethanolamine in diseases other than hypophosphatasia*. Am J Med. 1978;64(1):133-8.
36. Akiyama T, Kubota T, Ozono K, Michigami T, Kobayashi D, Takeyari S, et al. *Pyridoxal 5'-phosphate and related metabolites in hypophosphatasia: Effects of enzyme replacement therapy.* Mol Genet Metab. 2018;125(1-2):174-80.
37. Eastman JR, Bixler D. *Urinary phosphoethanolamine: normal values by age.* Clin Chem. 1980;26(12):1757-8.
38. Lu W, Bennett BD, Rabinowitz JD. *Analytical strategies for LC-MS-based targeted metabolomics.* J Chromatogr B Analyt Technol Biomed Life Sci. 2008;871(2):236-42.
39. Coene KLM, Kluijtmans LAJ, van der Heeft E, Engelke UFH, de Boer S, Hoegen B, et al. *Next-generation metabolic screening: targeted and untargeted metabolomics for the diagnosis of inborn errors of metabolism in individual patients.* J Inherit Metab Dis. 2018;41(3):337-53.
40. Kim SJ, Kim SH, Kim JH, Hwang S, Yoo HJ. *Understanding Metabolomics in Biomedical Research.* Endocrinol Metab (Seoul). 2016;31(1):7-16.
41. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. *HMDB: the Human Metabolome Database.* Nucleic Acids Res. 2007;35(suppl_1):D521-D6.

42. Ivanisevic J, Zhu Z-J, Plate L, Tautenhahn R, Chen S, O'Brien PJ, et al. *Toward 'Omic Scale Metabolite Profiling: A Dual Separation–Mass Spectrometry Approach for Coverage of Lipid and Central Carbon Metabolism.* Anal Chem. 2013;85(14):6876-84.
43. Ren J-L, Zhang A-H, Kong L, Wang X-J. *Advances in mass spectrometry-based metabolomics for investigation of metabolites.* RSC Advances. 2018;8(40):22335-50.
44. Xiao JF, Zhou B, Ressom HW. *Metabolite identification and quantitation in LC-MS/MS-based metabolomics.* Trends Analyt Chem. 2012;32:1-14.
45. Lei Z, Huhman DV, Sumner LW. *Mass Spectrometry Strategies in Metabolomics*.* J Biol Chem. 2011;286(29):25435-42.
46. Krastanov A. *Metabolomics—The State of Art.* Biotechnology & Biotechnological Equipment. 2010;24(1):1537-43.
47. Nash WJ, Dunn WB. *From mass to metabolite in human untargeted metabolomics: Recent advances in annotation of metabolites applying liquid chromatography-mass spectrometry data.* TrAC, Trends Anal Chem. 2019;120:115324.
48. Brown M, Dunn WB, Dobson P, Patel Y, Winder CL, Francis-McIntyre S, et al. *Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics.* Analyst. 2009;134(7):1322-32.
49. Metz TO, Zhang Q, Page JS, Shen Y, Callister SJ, Jacobs JM, et al. *The future of liquid chromatography-mass spectrometry (LC-MS) in metabolic profiling and metabolomic studies for biomarker discovery.* Biomark Med. 2007;1(1):159-85.
50. Jandera P, Janás P. *Recent advances in stationary phases and understanding of retention in hydrophilic interaction chromatography. A review.* Anal Chim Acta. 2017;967:12-32.
51. Contrepois K, Jiang L, Snyder M. *Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)–Mass Spectrometry*.* Mol Cell Proteomics. 2015;14(6):1684-95.
52. Shen Y, Zhang R, Moore RJ, Kim J, Metz TO, Hixson KK, et al. *Automated 20 kpsi RPLC-MS and MS/MS with Chromatographic Peak Capacities of 1000−1500 and Capabilities in Proteomics and Metabolomics.* Anal Chem. 2005;77(10):3090-100.
53. Harris DC. *Quantitative Chemical Analysis* 9ed. W.H. Freeman and Company Kate Parker; 2016. 792 p. 106, 559-63, 579, 668, 670-1, 683.

54. Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O. *A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. Metabolites*. 2012;2(4):775-95.
55. Jandera P. *Gradient elution in normal-phase high-performance liquid chromatographic systems.* J Chromatogr A. 2002;965(1-2):239-61.
56. Kazakevich Y, LoBrutto R. Introduction. *HPLC for Pharmaceutical Scientists;* 2007. p. 1-24.
57. Ferré S, González-Ruiz V, Guillarme D, Rudaz S. *Analytical strategies for the determination of amino acids: Past, present and future trends.* J Chromatogr B. 2019;1132:121819.
58. Olives AI, González-Ruiz V, Martín MA. *Sustainable and Eco-Friendly Alternatives for Liquid Chromatographic Analysis.* ACS Sustainable Chem Eng. 2017;5(7):5618-34.
59. Katajamaa M, Orešič M. *Data processing for mass spectrometry-based metabolomics.* J Chromatogr A. 2007;1158(1):318-28.
60. de Hoffmann E. *Tandem mass spectrometry: A primer.* J Mass Spectrom. 1996;31(2):129-37.
61. Kitteringham NR, Jenkins RE, Lane CS, Elliott VL, Park BK. *Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics.* J Chromatogr B. 2009;877(13):1229-39.
62. Kalli A, Smith GT, Sweredoski MJ, Hess S. *Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers.* J Proteome Res. 2013;12(7):3071-86.
63. Guo J, Huan T. *Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics.* Anal Chem. 2020;92(12):8072-80.
64. Xian F, Hendrickson CL, Marshall AG. *High Resolution Mass Spectrometry*. Anal Chem. 2012;84(2):708-19.
65. Brenton AG, Godfrey AR. *Accurate mass measurement: terminology and treatment of data.* J Am Soc Mass Spectrom. 2010;21(11):1821-35.
66. Junot C, Fenaille F, Colsch B, Bécher F. *High resolution mass spectrometry based techniques at the crossroads of metabolic pathways*. Mass Spectrom Rev. 2014;33(6):471-500.

67. Metz TO, Zhang Q, Page JS, Shen Y, Callister SJ, Jacobs JM, et al. *The future of liquid chromatography-mass spectrometry (LC-MS) in metabolic profiling and metabolomic studies for biomarker discovery.* Biomark Med. 2007;1(1):159-85.
68. Aszyk J, Byliński H, Namieśnik J, Kot-Wasik A. *Main strategies, analytical trends and challenges in LC-MS and ambient mass spectrometry–based metabolomics*. TrAC, Trends Anal Chem. 2018;108:278-95.
69. Putri SP, Yamamoto S, Tsugawa H, Fukusaki E. *Current metabolomics: Technological advances.* J Biosci Bioeng. 2013;116(1):9-16.
70. Wilm M. *Principles of Electrospray Ionization.* Mol Cell Proteomics. 2011;10(7):M111.009407.
71. Kruve A, Kaupmees K. *Adduct Formation in ESI/MS by Mobile Phase Additives*. J Am Soc Mass Spectrom. 2017;28(5):887-94.
72. Zhou Bin XJF, Tuli Leepika, Ressom Habtom W. *LC-MS-based metabolomics.* Mol Biosyst. 2011;2012(8):470-81.
73. El-Aneed A, Cohen A, Banoub J. *Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers*. Applied Spectroscopy Reviews. 2009;44(3):210-30.
74. Michalski A, Damoc E, Hauschild J-P, Lange O, Wieghaus A, Makarov A, et al. *Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer**. Mol Cell Proteomics. 2011;10(9):M111.011015.
75. Eliuk S, Makarov A. *Evolution of Orbitrap Mass Spectrometry Instrumentation.* Annu Rev Anal Chem (Palo Alto Calif). 2015;8(1):61-80.
76. Brown SC, Kruppa G, Dasseux J-L. *Metabolomics applications of FT-ICR mass spectrometry*. Mass Spectrom Rev. 2005;24(2):223-31.
77. Hendriks MMWB, Eeuwijk FAv, Jellema RH, Westerhuis JA, Reijmers TH, Hoefsloot HCJ, et al. *Data-processing strategies for metabolomics studies.* TrAC, Trends Anal Chem. 2011;30(10):1685-98.
78. Lamichhane S, Sen P, Dickens AM, Hyötyläinen T, Orešič M. Chapter Fourteen - An Overview of Metabolomics Data Analysis: Current Tools and Future Perspectives. In: Jaumot J, Bedia C, Tauler R, editors. *Comprehensive Analytical Chemistry*. 82: Elsevier; 2018. p. 387-413.
79. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, et al. *Guidelines and considerations for the use of system suitability and quality control*

samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. Metabolomics. 2018;14(6):72.

80. European Medicines Agency. *Bioanalytical method validation*. United Kingdom2011. p. 8, 22.
81. Dudzik D, Barbas-Bernardos C, García A, Barbas C. *Quality assurance procedures for mass spectrometry untargeted metabolomics. a review.* J Pharm Biomed Anal. 2018;147:149-73.
82. Godzien J, Alonso-Herranz V, Barbas C, Armitage EG. *Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample.* Metabolomics. 2015;11(3):518-28.
83. Naz S, Vallejo M, García A, Barbas C. *Method validation strategies involved in non-targeted metabolomics*. J Chromatogr A. 2014;1353:99-105.
84. Ng B, Quinete N, Gardinali PR. *Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis*. Sci Total Environ. 2020;713:136568.
85. Begou O, Gika HG, Theodoridis GA, Wilson ID. Quality Control and Validation Issues in LC-MS Metabolomics. In: Theodoridis GA, Gika HG, Wilson ID, editors. *Metabolic Profiling: Methods and Protocols.* New York, NY: Springer New York; 2018. p. 15-26.
86. Annesley TM. *Ion Suppression in Mass Spectrometry*. Clin Chem. 2003;49(7):1041-4.
87. Taylor PJ. *Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry*. Clin Biochem. 2005;38(4):328-34.
88. Chaleckis R, Meister I, Zhang P, Wheelock CE. *Challenges, progress and promises of metabolite annotation for LC-MS-based metabolomics.* Curr Opin Biotechnol. 2019;55:44-50.
89. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. *Proposed minimum reporting standards for chemical analysis.* Metabolomics. 2007;3(3):211-21.
90. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. *Untargeted Metabolomics Strategies-Challenges and Emerging Directions*. J Am Soc Mass Spectrom. 2016;27(12):1897-905.
91. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. *Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence.* Environ Sci Technol. 2014;48(4):2097-8.

92. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. *HMDB 4.0: the human metabolome database for 2018.* Nucleic Acids Res. 2018;46(D1):D608-d17.
93. Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR, et al. *METLIN: A Metabolite Mass Spectral Database.* Ther Drug Monit. 2005;27(6).
94. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, et al. *METLIN: a metabolite mass spectral database*. Ther Drug Monit. 2005;27(6):747-51.
95. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, et al. *METLIN: A Technology Platform for Identifying Knowns and Unknowns.* Anal Chem. 2018;90(5):3156-64.
96. Kennedy AD, Miller MJ, Beebe K, Wulff JE, Evans AM, Miller LA, et al. *Metabolomic Profiling of Human Urine as a Screen for Multiple Inborn Errors of Metabolism.* Genet Test Mol Biomarkers. 2016;20(9):485-95.
97. Chetwynd AJ, Abdul-Sada A, Holt SG, Hill EM. *Use of a pre-analysis osmolality normalisation method to correct for variable urine concentrations and for improved metabolomic analyses*. J Chromatogr A. 2016;1431:103-10.
98. Ryan D, Robards K, Prenzler PD, Kendall M. *Recent and potential developments in the analysis of urine: A review.* Anal Chim Acta. 2011;684(1):17-29.
99. Yeum K-J, Russell RM, Krinsky NI, Aldini G. *Biomarkers of antioxidant capacity in the hydrophilic and lipophilic compartments of human plasma.* Arch Biochem Biophys. 2004;430(1):97-103.
100. Wawrzyniak R, Kosnowska A, Macioszek S, Bartoszewski R, Jan Markuszewski M. *New plasma preparation approach to enrich metabolome coverage in untargeted metabolomics: plasma protein bound hydrophobic metabolite release with proteinase K*. Sci Rep. 2018;8(1):9541.
101. Arnesen CE. *Optimization of liquid chromatographic parameters for untargeted metabolomics of dried blood spots* [Master´s Thisis]. Oslo, Norway: Oslo University; 2017.
102. Skogvold HB. *Laboratory diagnostics: Maximizing sensitivity of a Q-Exactive Orbitrap mass spectrometer for untargeted metabolomics of dried blood spots* [Master´s Thesis]. Oslo, Norway Oslo University 2017.
103. Sandås EM. *Liquid chromatography - Orbitrap mass spectrometry is a useful tool in untargeted metabolomics analysis of dried blood spots in clinical chemistry* [Master´s Thesis]. Oslo, Norway: Oslo University; 2018.

104. Kord AS, Khaledi MG. *Chromatographic characteristics of surfactant-mediated separations: micellar liquid chromatography vs ion pair chromatography.* Anal Chem. 1992;64(17):1901-7.
105. Sun L, Jiao H, Gao B, Yuanzi Q, Zhang H, Wang Y, et al. *Hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry method for the simultaneous determination of l-valine, l-leucine, l-isoleucine, l-phenylalanine, and l-tyrosine in human serum.* J Sep Sci. 2015;38(22):3876-83.
106. Socia A, Foley JP. *Direct determination of amino acids by hydrophilic interaction liquid chromatography with charged aerosol detection*. J Chromatogr A. 2016;1446:41-9.
107. Tůma P, Samcová E, Balínová P. *Determination of 3-methylhistidine and 1-methylhistidine in untreated urine samples by capillary electrophoresis.* J Chromatogr B. 2005;821(1):53-9.
108. Tůma P, Samcová E, Opekar F, Jurka V, Štulík K. *Determination of 1-methylhistidine and 3-methylhistidine by capillary and chip electrophoresis with contactless conductivity detection.* Electrophoresis. 2007;28(13):2174-80.
109. Soga T, Heiger DN. *Amino Acid Analysis by Capillary Electrophoresis Electrospray Ionization Mass Spectrometry.* Anal Chem. 2000;72(6):1236-41.
110. Hirayama A, Wakayama M, Soga T. *Metabolome analysis based on capillary electrophoresis-mass spectrometry.* TrAC, Trends Anal Chem. 2014;61:215-22.
111. Maxwell EJ, Zhong X, Zhang H, van Zeijl N, Chen DDY. *Decoupling CE and ESI for a more robust interface with MS.* Electrophoresis. 2010;31(7):1130-7.
112. Boswell PG, Schellenberg JR, Carr PW, Cohen JD, Hegeman AD. *A study on retention "projection" as a supplementary means for compound identification by liquid chromatography–mass spectrometry capable of predicting retention with different gradients, flow rates, and instruments.* J Chromatogr A. 2011;1218(38):6732-41.
113. Wen Y, Amos RIJ, Talebi M, Szucs R, Dolan JW, Pohl CA, et al. *Retention Index Prediction Using Quantitative Structure–Retention Relationships for Improving Structure Identification in Nontargeted Metabolomics.* Anal Chem. 2018;90(15):9434-40.
114. Creek DJ, Jankevics A, Breitling R, Watson DG, Barrett MP, Burgess KEV. *Toward Global Metabolomics Analysis with Hydrophilic Interaction Liquid Chromatography–Mass Spectrometry: Improved Metabolite Identification by Retention Time Prediction.* Anal Chem. 2011;83(22):8703-10.

115. Zheng S-J, Liu S-J, Zhu Q-F, Guo N, Wang Y-L, Yuan B-F, et al. *Establishment of Liquid Chromatography Retention Index Based on Chemical Labeling for Metabolomic Analysis*. Anal Chem. 2018;90(14):8412-20.
116. Bonini P, Kind T, Tsugawa H, Barupal DK, Fiehn O. *Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics*. Anal Chem. 2020;92(11):7515-22.
117. Wolfer AM, Lozano S, Umbdenstock T, Croixmarie V, Arrault A, Vayer P. *UPLC–MS retention time prediction: a machine learning approach to metabolite identification in untargeted profiling*. Metabolomics. 2015;12(1):8.
118. Virgiliou C, Theodoridis G, Wilson ID, Gika HG. *Quantification of endogenous aminoacids and aminoacid derivatives in urine by hydrophilic interaction liquid chromatography tandem mass spectrometry*. J Chromatogr A. 2021;1642:462005.
119. Cui L, Lu H, Lee YH. *Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases*. Mass Spectrom Rev. 2018;37(6):772-92.
120. Valadbeigi Y, Ilbeigi V, Tabrizchi M. *Effect of mono- and di-hydration on the stability and tautomerisms of different tautomers of creatinine: A thermodynamic and mechanistic study*. Comput Theor Chem. 2015;1061:27-35.
121. Gao J, Hu Y, Li S, Zhang Y, Chen X. *Tautomeric equilibrium of creatinine and creatininium cation in aqueous solutions explored by Raman spectroscopy and density functional theory calculations*. Chem Phys. 2013;410:81-9.
122. Hou H, Xiong W, Zhang X, Song D, Tang G, Hu Q. *LC-MS-MS Measurements of Urinary Creatinine and the Application of Creatinine Normalization Technique on Cotinine in Smokers' 24 Hour Urine*. J Anal Methods Chem. 2012;2012:245415.
123. Saude EJ, Adamko D, Rowe BH, Marrie T, Sykes BD. *Variation of metabolites in normal human urine*. Metabolomics. 2007;3(4):439-51.
124. Garde AH, Hansen AM, Kristiansen J, Knudsen LE. *Comparison of uncertainties related to standardization of urine samples with volume and creatinine concentration*. Ann Occup Hyg. 2004;48(2):171-9.
125. Rosen Vollmar AK, Rattray NJW, Cai Y, Santos-Neto ÁJ, Deziel NC, Jukic AMZ, et al. *Normalizing Untargeted Periconceptional Urinary Metabolomics Data: A Comparison of Approaches*. Metabolites. 2019;9(10):198.
126. Chen Y, Shen G, Zhang R, He J, Zhang Y, Xu J, et al. *Combination of Injection Volume Calibration by Creatinine and MS Signals' Normalization to Overcome Urine Variability in LC-MS-Based Metabolomics Studies*. Anal Chem. 2013;85(16):7659-65.

127. Meister I, Zhang P, Sinha A, Sköld CM, Wheelock ÅM, Izumi T, et al. *High-Precision Automated Workflow for Urinary Untargeted Metabolomic Epidemiology*. Anal Chem. 2021;93(12):5248-58.
128. Gika HG, Theodoridis GA, Wingate JE, Wilson ID. *Within-Day Reproducibility of an HPLC−MS-Based Method for Metabonomic Analysis: Application to Human Urine*. J Proteome Res. 2007;6(8):3291-303.
129. King R, Bonfiglio R, Fernandez-Metzler C, Miller-Stein C, Olah T. *Mechanistic investigation of ionization suppression in electrospray ionization*. J Am Soc Mass Spectrom. 2000;11(11):942-50.
130. Thévenot EA, Roux A, Xu Y, Ezan E, Junot C. *Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses*. J Proteome Res. 2015;14(8):3322-35.

Appendix

Appendix A - Chemical data of the amino acids and related compounds

Table A. 1: Chemical data of the amino acids and related compounds. *The table list all amino acids and related compounds that were studied by metabolomics for this thesis. The chemical formula, molecular weight, structure and logP value are also listed for each chemical compound. The amino acid structure and the monoisotopic molecular weight were obtained from hmdb, and LogP values were obtained from ChemAxon.*

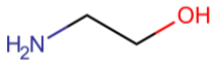
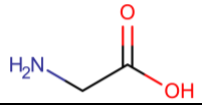
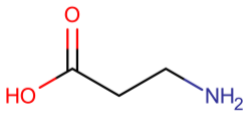
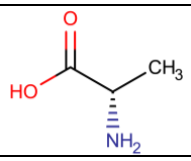
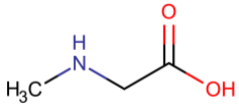
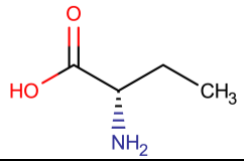
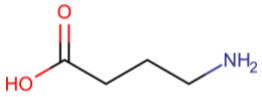
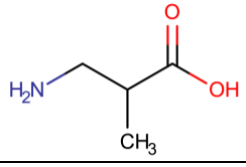
Amino acid	Chemical formula	Monoisotopic Molecular Weight	Structure	LogP value	Standard solution
Ethanolamine	C_2H_7NO	61.052763851		-1.3	A9906
Glycine	$C_2H_5NO_2$	75.032028409		-3.4	A9906
β-Alanine	$C_3H_7NO_2$	89.047678473		-3.2	A9906
L-Alanine	$C_3H_7NO_2$	89.047678473		-2.8	A9906
L-Sarcosine	$C_3H_7NO_2$	89.047678473		-3.2	A9906
L-α-amino-n-butyric acid	$C_4H_9NO_2$	103.063328537		-2.3	A9906
γ-amino-n-butyric acid	$C_4H_9NO_2$	103.063328537		-2.9	A9906
D,L-β-aminoisoburytic acid	$C_4H_9NO_2$	103.063328537		-2.6	A9906

Table A. 1 *continues on the next page.*

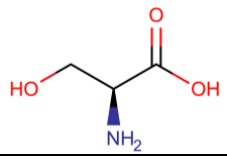
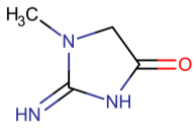
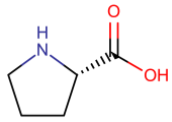
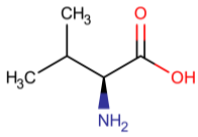
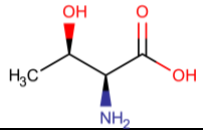
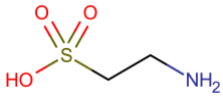
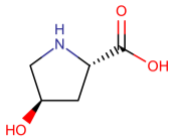
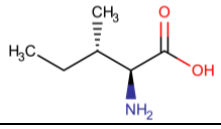
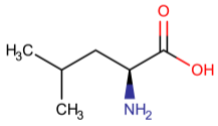
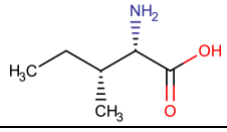
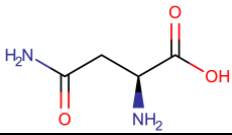
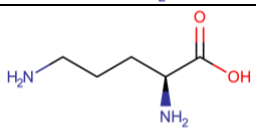
Amino acid	Chemical formula	Monoisotopic Molecular Weight	Structure	LogP value	Standard solution
L-Serine	$C_3H_7NO_3$	105.042593095		-3.9	A9906
Creatinine	$C_4H_7N_3O$	113.058911861		-1.1	A9906
L-Proline	$C_5H_9NO_2$	115.063328537		-2.6	A9906
L-Valine	$C_5H_{11}NO_2$	117.078978601		-2.0	A9906
L-Threonine	$C_4H_9NO_3$	119.058243159		-3.5	A9906
Taurine	$C_2H_7NO_3S$	125.014663785		-2.6	A9906
Hydroxy-L-proline	$C_5H_9NO_3$	131.058243159		-3.7	A9906
L-Isoleucine	$C_6H_{13}NO_2$	131.094628665		-1.5	A9906
L-Leucine	$C_6H_{13}NO_2$	131.094628665		-1.6	A9906
Alloisoleucine	C6H13NO2	131.094628665		-1.5	Special mix
Asparagine	$C_4H_8N_2O_3$	132.053492132		-4.3	Special mix
L-Ornithine	$C_5H_{12}N_2O_2$	132.089877638		-3.7	A9906

Table A. 1 *continues on the next page.*

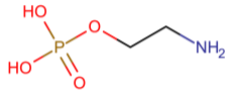
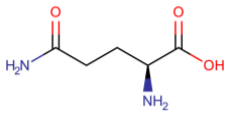
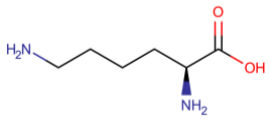
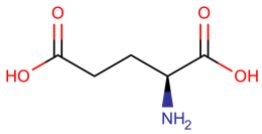
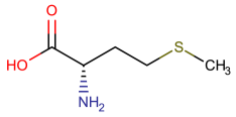
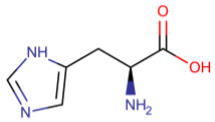
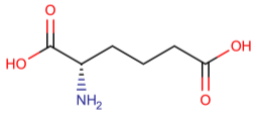
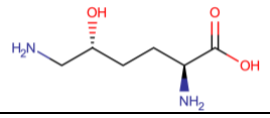
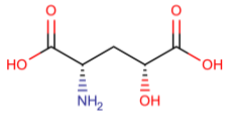
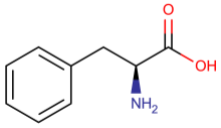
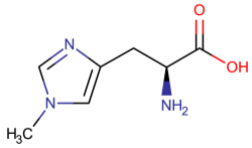
Amino acid	Chemical formula	Monoisotopic Molecular Weight	Structure	LogP value	Standard solution
L-Aspartic acid	$C_4H_7NO_4$	133.037507717		-3.5	A9906
Phosphoethanolamine	$C_2H_8NO_4P$	141.019094261		-2.5	
L-Glutamine	$C_5H_{10}N_2O_3$	146.069142196		-4.0	Special mix
L-Lysine	$C_6H_{14}N_2O_2$	146.105527702		-3.2	A9906
L-Glutamic Acid	$C_5H_9NO_4$	147.053157781		-3.2	A9906
L-Methionine	$C_5H_{11}NO_2S$	149.051049291		-2.2	A9906
L-Histidine	$C_6H_9N_3O_2$	155.069476547		-3.6	A9906
L-α-aminoadipic acid	$C_6H_{11}NO_4$	161.068807845		-2.8	A9906
δ-Hydroxylysine	$C_6H_{14}N_2O_3$	162.100442324		-4.4	A9906
4-Hydroxy-L-glutamic acid	$C_5H_9NO_5$	163.048072403		-4.2	A9906
L-Phenylalanine	$C_9H_{11}NO_2$	165.078978601		-1.2	A9906
1-Methyl-L-histidine	$C_7H_{11}N_3O_2$	169.085126611		-3.1	A9906

Table A. 1 *continues on the next page.*

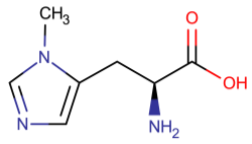
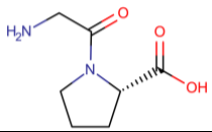
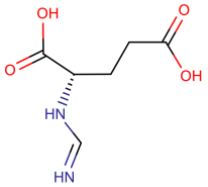
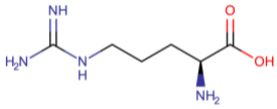
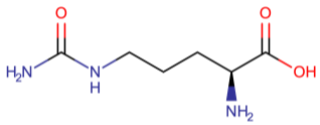
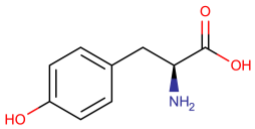
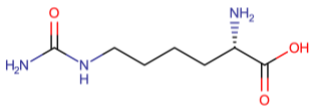
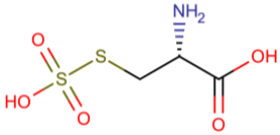
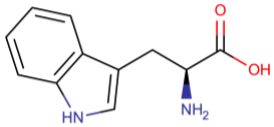
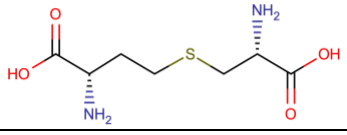
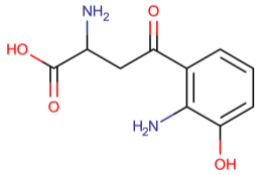
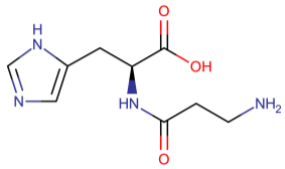
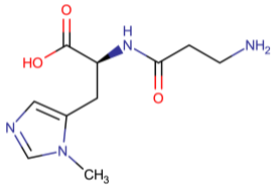
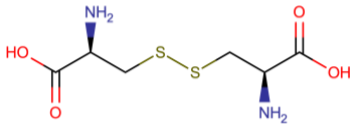
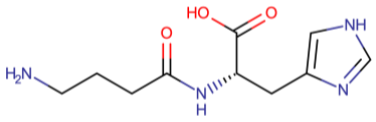
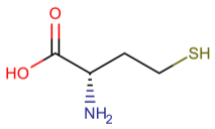
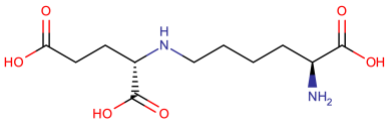
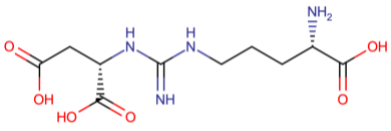
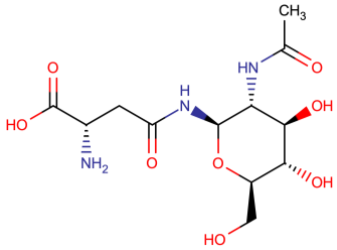
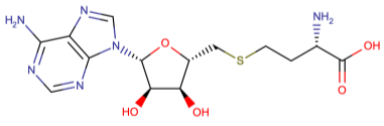
Amino acid	Chemical formula	Monoisotopic Molecular Weight	Structure	LogP value	Standard solution
3-Methyl-L-histidine	$C_7H_{11}N_3O_2$	169.085126611		-3.4	A9906
Glycylproline	$C_7H_{12}N_2O_3$	172.08479226		-3.7	Special mix
Formiminoglutamic acid	$C_6H_{10}N_2O_4$	174.064056818		-2.8	Special mix
L-Arginine	$C_6H_{14}N_4O_2$	174.111675712		-3.2	A9906
L-Citrulline	$C_6H_{13}N_3O_3$	175.095691297		-3.9	A9906
L-Tyrosine	$C_9H_{11}NO_3$	181.073893223		-1.5	A9906
Homocitrulline	$C_7H_{15}N_3O_3$	189.111341361		-3.5	A9906
S-Sulfocysteine	$C_3H_7NO_5S_2$	200.976563719		-2.3	Special mix
L-Tryptophan	$C_{11}H_{12}N_2O_2$	204.089877638		-1.1	A9906
L-Cystathionine	$C_7H_{14}N_2O_4S$	222.067427636		-5.8	A9906

Table A. 1 *continues on the next page.*

Amino acid	Chemical formula	Monoisotopic Molecular Weight	Structure	LogP value	Standard solution
3-Hydroxy-D,L-kynurenine	$C_{10}H_{12}N_2O_4$	224.079706882		-2.2	Special mix
L-Carnosine	$C_9H_{14}N_4O_3$	226.106590334		-4.5	A9906
L-Anserine	$C_{10}H_{16}N_4O_3$	240.122240398		-4.3	A9906
L-Cystine	$C_6H_{12}N_2O_4S_2$	240.023848262		-5.9	A9906
Homocarnosine	$C_{10}H_{16}N_4O_3$	240.122240398		-3.9	Special mix
L-Homocysteine	$C_4H_9NO_2S$	135.035399227		-2.6	Special mix
Saccharopine	$C_{11}H_{20}N_2O_6$	276.132136382		-5.4	Special mix
Argininosuccinic acid	$C_{10}H_{18}N_4O_6$	290.122634328		-5.8	Special mix
Aspartylglucosamine	$C_{12}H_{21}N_3O_8$	335.132864663		-6.8	Special mix
S-Adenosyl-L-homocysteine	$C_{14}H_{20}N_6O_5S$	384.12158847		-4.0	Special mix

Appendix B - Precision of metabolomics for amino acids in urine

Table B. 1: Amino acid peak area and retention time repeatability in urine. *The amino acid retention time and peak area repeatability was found for one urine sample. The mean, standard deviation (SD), and the relative standard deviation (%RSD) were calculated.*

	Retention time (min)			Peak area (a.u.)		
Amino acid	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
L-α-aminoadipic acid	2.39	0.002	0.08	5.37E+07	2.89E+06	5.38
L-Anserine	2.03	0.002	0.10	7.99E+06	5.58E+05	6.99
L-Arginine	2.03	0.002	0.11	3.21E+07	9.60E+05	2.99
L-Aspartic acid	2.17	0.003	0.13	2.03E+07	4.95E+05	2.44
L-Carnosine	2.03	0.002	0.09	1.68E+08	8.46E+06	5.02
L-Citrulline	2.27	0.002	0.07	6.48E+06	2.60E+05	4.01
L-Cystathionine	2.07	0.002	0.12	1.24E+07	8.03E+05	6.47
L-Cystine	2.09	0.003	0.12	1.16E+08	7.44E+06	6.41
Ethanolamine	2.03	0.002	0.10	2.31E+08	9.64E+06	4.17
L-Glutamic acid	2.22	0.003	0.14	2.16E+08	7.81E+06	3.62
Glycine	2.09	0.003	0.13	8.25E+08	2.77E+07	3.35
L-Histidine	2.03	0.004	0.19	2.85E+08	1.45E+07	5.09
L-Hydroxyproline	1.78	0.002	0.11	1.38E+08	5.01E+06	3.62
L-Isoleucine, L-leucine	3.45	0.003	0.10	5.43E+08	2.83E+07	5.22
L-Lysine	1.79	0.002	0.12	8.24E+07	5.94E+06	7.20
L-Methionine	3.10	0.003	0.10	7.60E+07	3.20E+06	4.21
1-Methyl-L-histidine, 3-MethylL-histidine	2.03	0.002	0.09	2.64E+08	1.07E+07	4.06
L-Ornithine	1.79	0.003	0.15	1.20E+07	3.81E+05	3.18
L-Phenylalanine	6.64	0.005	0.08	6.25E+08	2.66E+07	4.26
L-Proline	2.58	0.003	0.11	3.62E+08	2.70E+07	7.46
L-Serine	2.11	0.003	0.12	5.48E+08	2.49E+07	4.54
Taurine	2.22	0.003	0.12	4.44E+08	1.93E+07	4.35
L-Threonine	2.17	0.002	0.08	3.43E+08	1.55E+07	4.51
L-Tryptophan	11.16	0.01	0.11	7.68E+08	5.26E+07	6.85
L-Tyrosine	4.15	0.002	0.04	6.33E+08	4.40E+07	6.95
L-Valine	2.85	0.002	0.08	2.00E+10	1.29E+09	6.46
Glutamine	2.19	0.002	0.09	1.83E+09	7.38E+07	4.03
Asparaginine	2.13	0.003	0.12	1.64E+08	3.04E+06	1.86
4-hydroxy-L-glutamic acid	2.16	0.004	0.17	1.53E+07	7.81E+05	5.10

Table B. 1 *is continued on the next page.*

	Retention time (min)			Peak area (a.u.)		
Amino acid	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
L-Alanine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
β-Alanine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
γ–Amino-n-butyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
β–Amino-isobutyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
α–Amino-n-butyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
Homocysteine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
Hydroxykynurenine	4.51	0.003	0.06	1.51E+07	1.11E+06	7.83
S-Adenosyl-L-homocysteine	4.12	0.006	0.15	1.44E+06	1.91E+05	13.3
Formominoglutamic acid	2.38	0.003	0.11	4.32E+07	3.09E+06	7.15
Argininosuccinic acid	2.15	0.003	0.12	2.65E+07	1.85E+06	6.99
S-sulfocysteine	2.70	0.008	0.28	4.92E+06	4.15E+05	8.44
Saccarophine	2.17	0.004	0.17	3.61E+06	3.59E+05	9.94
Homocarnosine	2.03	0.002	0.10	7.99E+06	5.58E+05	6.99
Aspartylglucosamine	2.13	0.002	0.11	3.06E+07	1.81E+06	5.91

* N.D. = Not detectable. Means here “not found endogenously”.

Table B. 2: Amino acid peak area and retention time intermediate precision in urine. *The amino acid retention time and peak area intermediate precision was found for one urine sample. The mean, standard deviation (SD), and the relative standard deviation (%RSD) were calculated.*

	Retention time (min)			Peak area (a.u.)		
Amino acid	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
L-α-aminoadipic acid	2.43	0.02	0.88	4.14E+07	1,95E+07	47.1
L-Anserine	2.14	0.06	2.89	9.20E+06	1.30E+06	14.1
L-Arginine	2.14	0.06	2.87	4.82E+07	7.74E+06	16.0
L-Aspartic acid	2.29	0.06	2.81	2.53E+07	4.53E+06	17.9
L-Carnosine	2.14	0.06	2.87	1.96E+08	2.56E+07	13.0
L-Citrulline	2.40	0.07	2.89	1.18E+07	2.50E+06	21.2
L-Cystathionine	2.19	0.06	2.92	1.16E+07	1.84E+06	16.0
L-Cystine	2.20	0.06	2.86	1.27E+08	1.75E+07	13.8
Ethanolamine	2.14	0.06	2.86	1.90E+08	2.13E+07	11.3
L-Glutamic acid	2.34	0.07	2.91	2.70E+08	4.50E+07	16.7
Glycine	2.20	0.06	2.85	8.43E+08	1.00E+08	11.9
L-Histidine	2.14	0.06	2.79	4.30E+08	6.37E+07	14.8
δ–Hydroxylysine	1.90	0.06	3.06	1.59E+08	2.91E+07	18.3
Hydoxy-L-proline	2.40	0.07	2.83	4.40E+09	1.03E+09	23.4
L-Isoleucine, L-leucine	3.56	0.10	3.50	7.30E+08	1.53E+08	21.0

Table B. 2 *is continued on the next page.*

	Retention time (min)			Peak area (a.u.)		
Amino acids	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
L-Lysine	1.92	0.07	3.86	9.06E+07	1.12E+07	12.4
L-Methionine	3.24	0.10	3.24	8.88E+07	1.64E+07	18.4
1-Methyl-L-histidine, 3-Methyl-Lhistidine	2.15	0.06	2.88	3.43E+08	5.27E+07	15.4
L-Ornithine	1.91	0.06	3.09	1.11E+07	1.66E+06	14.9
L-Phenylalanine	6.81	0.30	4.29	9.10E+08	2.03E+08	22.3
L-Proline	2.71	0.08	2.79	4.09E+08	6.39E+07	15.6
L-Serine	2.23	0.06	2.88	6.87E+08	1.12E+08	16.3
Taurine	2.34	0.07	2.82	6.24E+08	1.09E+08	17.5
L-Threonine	2.29	0.07	2.85	4.01E+08	6.10E+07	15.2
L-Tryptophan	11.37	0.3	2.52	9.73E+08	1.72E+08	17.7
L-Tyrosine	4.31	0.20	4.10	1.04E+09	2.29E+08	22.0
L-Valine	2.99	0.08	2.59	2.81E+10	5.69E+09	20.3
Glutamine	2.32	0.07	3.08	2.22E+09	3.69E+08	16.6
Asparaginine	2.25	0.06	2.80	1.80E+08	2.58E+07	14.3
4-hydroxy-L-glutamic acid	2.27	0.06	2.69	2.10E+07	3.59E+06	17.1
L-Alanine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
β-Alanine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
γ–Amino-n-butyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
β–Amino-isobutyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
α–Amino-n-butyric acid	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
Homocysteine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
Hydroxykynurenine	4.37	0.20	3.87	2.52E+06	8.21E+05	31.9
S-Adenosyl-L-homocysteine	4.37	0.20	3.87	2.52E+06	8.21E+05	32.6
Formominoglutamic acid	2.51	0.07	2.91	3.30E+07	5.03E+06	15.3
Argininosuccinic acid	2.27	0.07	2.90	3.71E+07	6.82E+06	18.4
S-sulfocysteine	2.75	0.08	2.99	6.82E+06	1.25E+06	18.3
Saccharopine	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*	N.D.*
Homocarnosine	2.14	0.06	2.89	9.20E+06	1.30E+06	14.1
Aspartylglucosamine	2.25	0.06	2.85	4.16E+07	8.73E+06	21.0

* N.D. = Not detectable. Means here “not found endogenously”.

Appendix C - Quantitative analysis of metabolomics compared to targeted analysis of amino acids in urine

The results in **Figure C. 1** are the quantitative analysis of amino acids peak area from metabolomics and concentration ratios from targeted analyses for amino acids detected in urine samples diluted to 0.1 mM creatinine.

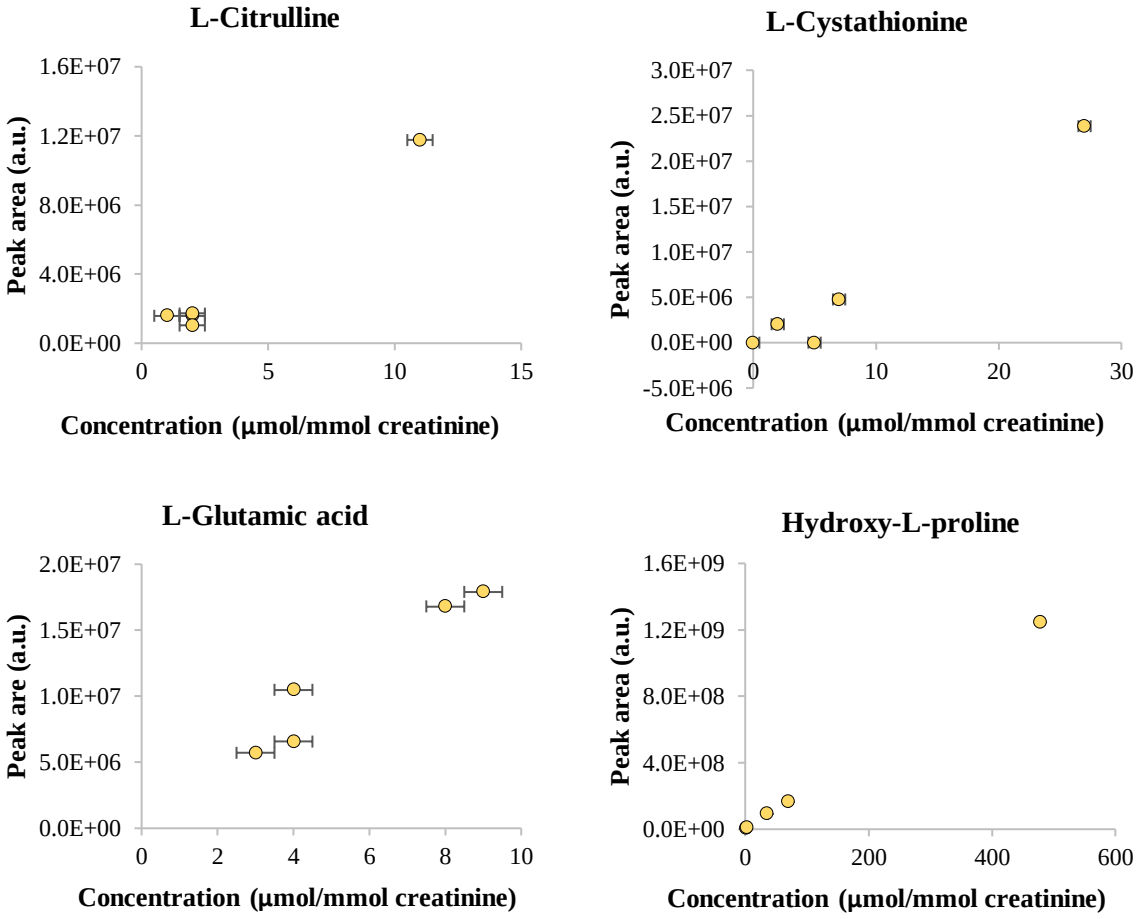


Figure C. 1 *is continued on the next page.*

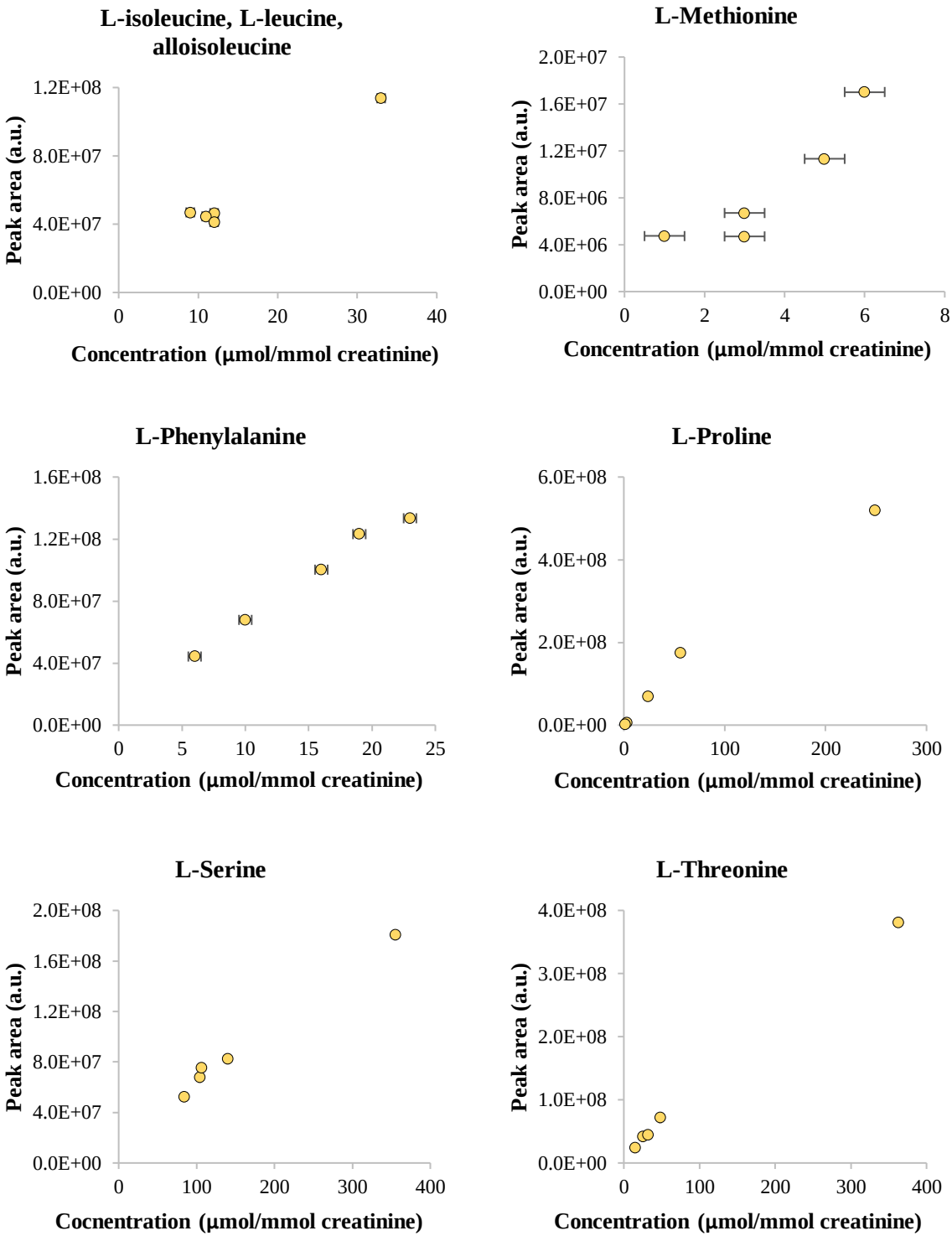


Figure C. 1 *is continued on the next page.*

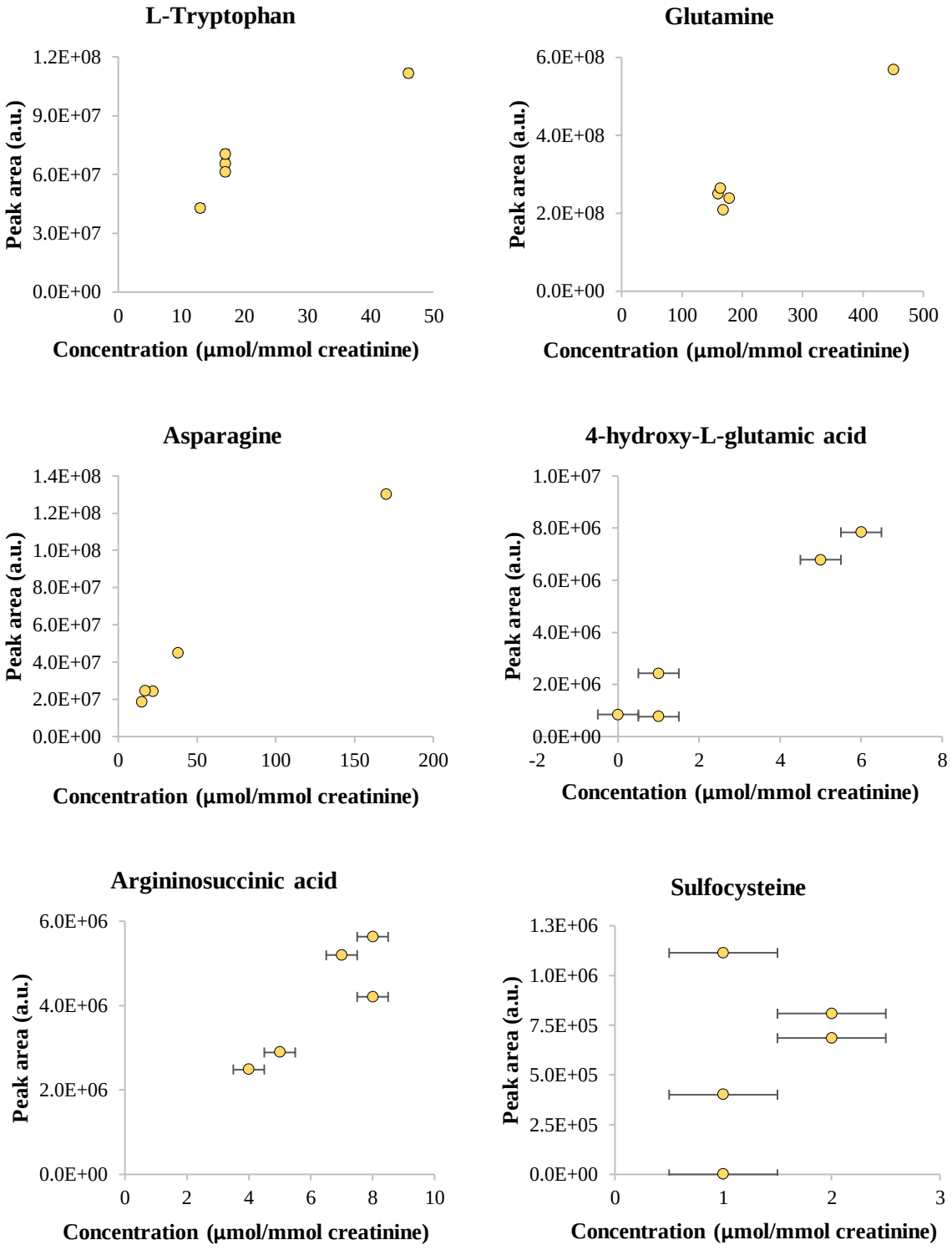


Figure C. 1 *is continued on the next page.*

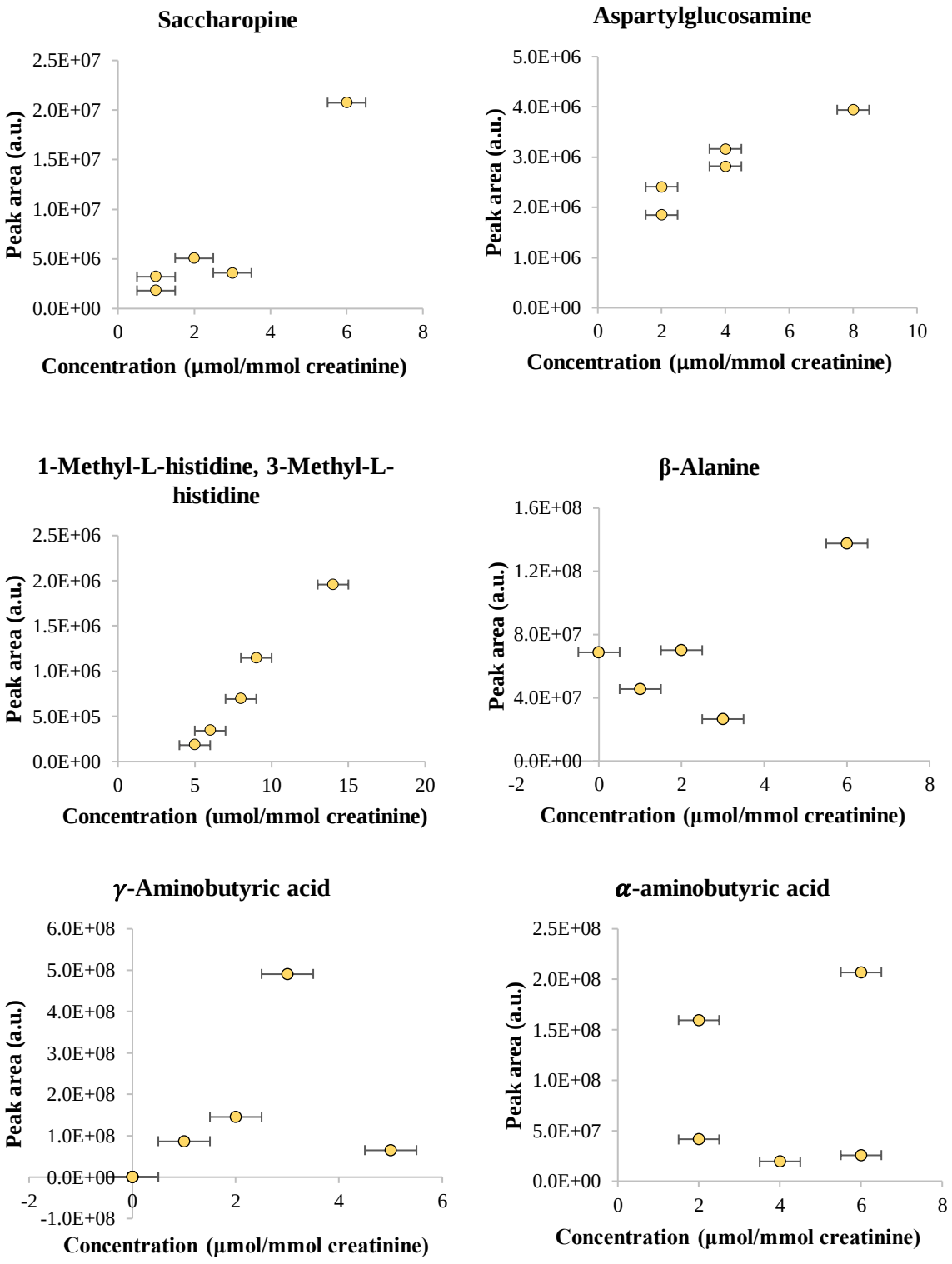


Figure C. 1 *is continued on the next page.*

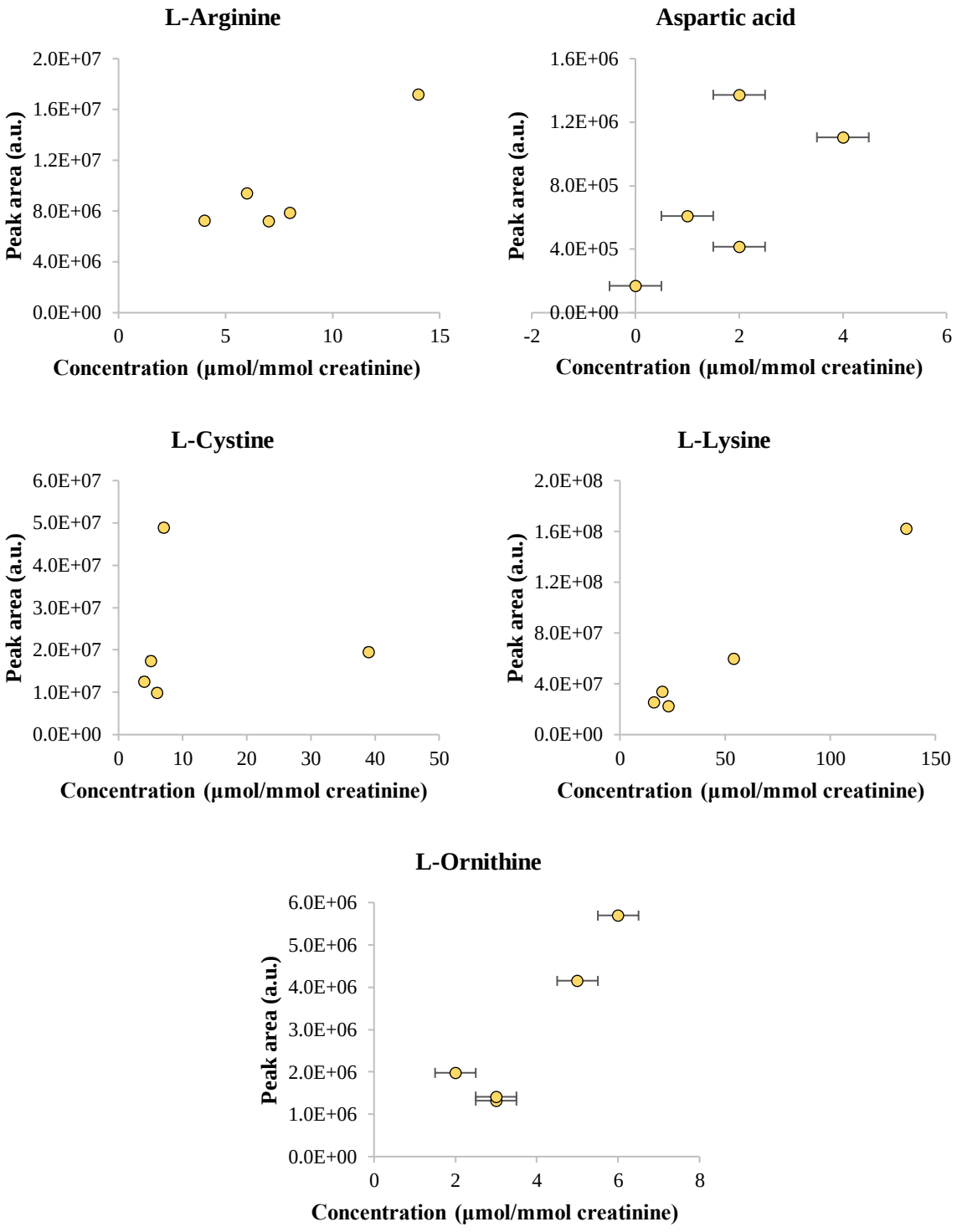


Figure C. 1: Quantitative comparison of peak area from metabolomics and concentration from targeted amino acid analyzes using samples diluted to 0.1 mM creatinine. *The figure shows the quantitative results for amino acids comparing peak areas from metabolomics to concentration ratios by targeted analyses using samples diluted to 0.1 mM creatinine.*

The results in **Figure C. 2** are the quantitative analysis of amino acids peak area from metabolomics and concentration ratios from targeted analyses for amino acids detected in urine samples diluted to 2 mM creatinine, that were not detectable in samples diluted to 0.1 mM creatinine.

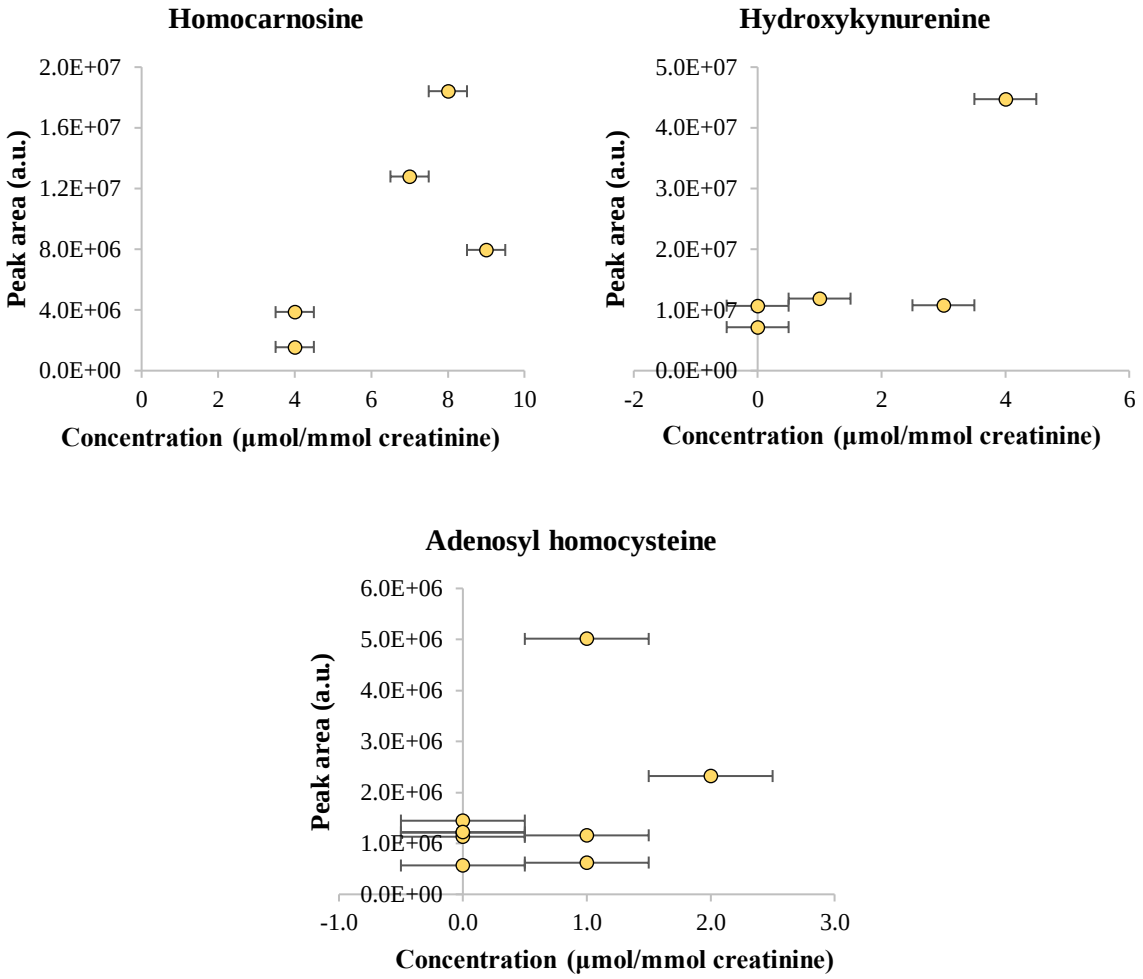


Figure C. 2: Quantitative comparison of peak area from metaolomics and concentration from targeted amino acid analyses using samples diluted to 2 mM creatinine. *The figure shows the quantitative results for amino acids comparing peak areas from metabolomics to concentration ratios by targeted analyses using samples diluted to 2 mM creatinine.*

Table C. 1 shows the concentrations from targeted analyses from the diagnostic laboratory of IEMs at Oslo University Hospital – Rikshospitalet, used to compare with peak areas from metabolomics of the selected compounds in urine.

Table C. 1: Measured concentrations of amino acids from targeted analyses. *The table shows amino acid concentrations from targeted analyses from the diagnostic laboratory of IEMs at Oslo University Hospital – Rikshospitalet.*

	Concentration (µmol/mmol creatinine)								
Amino acids	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**	**Sample 6**	**Sample 7**	**Sample 8**	**Sample 9**
L-Aspartic acid	2	0	1	2	4	1	2	3	0
Hydroxy-L-proline	2	3	36	479	70	133	12	129	0
L-Threonine	48	15	25	363	32	51	39	37	9
L-Serine	104	84	106	355	140	157	193	137	27
L-Asparagine	38	22	15	170	17	17	24	N.D.*	6
L-Glutamic acid	4	3	4	8	9	5	7	7	2
L-Glutamine	160	168	164	451	178	244	192	167	25
L-Proline	3	1	24	249	56	59	8	56	0
Glycine	277	163	271	2620	363	547	563	548	120
L-Alanine	78	49	110	325	118	167	117	165	34
L-Citrulline	2	2	1	11	2	1	3	2	1
L-α-amino-n-butyric acid	6	4	2	6	2	3	N.D.*	N.D.*	N.D.*
L-Valine	13	12	8	29	9	7	15	12	5
L-Cystine	6	5	7	39	9	19	6	8	4
L-Methionine	5	3	1	6	3	4	2	2	1
L-Cystathionine	5	2	N.D.*	27	7	5	4	8	5
L-Isoleucine	4	3	4	7	5	4	4	6	1
L-Leucine	8	8	5	26	7	11	8	12	3
L-Tyrosine	32	26	30	48	20	36	30	24	5
L-Phenylalanine	19	16	10	23	6	13	15	12	2
γ-amino-n-butyric acid	1	0	0	2	0	0	3	5	0
L-Ornithine	3	3	2	5	6	10	4	4	2
L-Lysine	54	23	16	136	20	59	19	18	7
1-Methyl-L-histidine	5	6	14	8	9	21	11	11	34
L-Histidine	229	265	131	237	136	269	251	125	41

Table C. 1 *is continued on the next page.*

	Concentration (µmol/mmol creatinine)								
Amino acids	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**	**Sample 6**	**Sample 7**	**Sample 8**	**Sample 9**
L-Homocysteine	0	0	0	0	0	0	0	1	0
L-Tryptophan	17	17	17	46	13	24	22	19	3
L-Arginine	8	7	4	14	6	9	5	7	16
S-Adenosyl-L-homocysteine	0	0	1	2	0	0	1	0	1
Argininosuccinic acid	8	7	4	8	5	5	9	5	3
Homocitrulline	12	17	1	6	1	11	4	3	4
Hydroxykynurenine	0	3	1	4	0	1	0	1	0
Saccharopine	1	3	2	6	1	4	2	1	0
L-α-aminoadpic acid	39	11	14	99	14	25	20	22	10
S-Sulfocysteine	1	1	1	2	2	2	2	2	0
L-Sarcosine	2	0	2	17	11	9	2	6	0
4-hydroxy-L-glutamic acid	1	0	1	6	5	6	1	3	0
Homocarnosine	4	9	8	4	7	7	18	3	1
Formiminoglutamic acid	3	2	6	21	12	7	6	14	2
Aspartylglucosamine	2	2	4	8	4	2	5	3	1
Glycylproline	11	28	18	28	17	25	11	14	2
β-Alanine	1	3	0	6	2	N.D.*	3	3	0

* N.D.= Not detectable. The analysis failed.

Appendix D - Precision of metabolomics for amino acids in plasma

Table D. 1: Repeatability of metabolomics for amino acids in plasma. *The amino acid retention time and peak area repeatability was found for one plasma sample. The mean, standard deviation (SD), and the relative standard deviation (%RSD) were calculated.*

	Retention time (min)			Peak area (a.u.)		
Amino acid	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
β-Alanine	2.01	0.005	0.23	1.77E+07	0.00E+00	0.00
L-Alanine	2.11	0.005	0.23	2.78E+08	5.75E+06	2.07
L-α-Aminoadipic acid	2.41	0.003	0.12	1.44E+08	1.73E+06	1.20
L-α-Amino-n-butyric acid	2.26	0.004	0.20	1.36E+8	2.69E+6	297.0
D,L-β-Aminoisobutyric acid	2.13	0.004	0.19	1.41E+08	2.63E+06	1.86
γ-Aminobutyric acid	2.03	0.004	0.22	5.50E+07	2.43E+06	4.43
L-Anserine	2.00	0.005	0.27	1.91E+07	3.99E+05	2.09
L-Arginine	2.00	0.005	0.24	1.14E+08	1.99E+06	1.74
L-Aspartic acid	2.16	0.005	0.21	1.01E+08	2.17E+06	2.14
L-Carnosine	1.75	0.008	0.46	1.39E+07	3.67E+05	2.65
L-Citrulline	2.26	0.004	0.18	2.95E+08	3.98E+06	1.35
Creatinine	2.24	0.007	0.31	7.93E+08	2.80E+07	3.54
L-Cystathionine	2.04	0.005	0.24	1.22E+08	2.47E+06	2.03
L-Cystine	2.06	0.004	0.22	2.33E+08	5.07E+06	2.18
Ethanolamine	2.00	0.006	0.30	1.07E+07	4.37E+05	4.10
L-Glutamic acid	2.21	0.004	0.18	2.92E+08	3.31E+06	1.13
Glycine	2.06	0.004	0.22	5.38E+08	1.09E+6	134.3
L-Histidine	2.00	0.005	0.24	7.41E+07	3.20E+06	4.32
L-Homocysteine	2.49	0.004	0.18	1.74E+08	6.84E+06	3.93
δ-Hydroxylysine	1.71	0.003	0.18	1.06E+08	9.66E+06	9.15
Hydroxy-L-proline	2.26	0.005	0.20	3.50E+08	4.24E+06	1.21
L-Isoleucine, L-leucine	3.59	0.01	0.29	2.43E+09	1.04E+08	4.30
L-Lysine	1.72	0.005	0.29	2.04E+08	3.52E+07	17.3
L-Methionine	3.19	0.007	0.22	7.44E+08	1.95E+07	2.62
1-Methyl-L-histidine, 3-Methyl-L-histidine	2.01	0.005	0.23	1.78E+08	4.73E+06	2.66
L-Ornithine	1.72	0.005	0.27	5.92E+07	7.15E+06	12.1
L-Phenylalanine	7.28	0.009	0.12	1.48E+09	2.67E+07	1.80
L-Proline	2.54	0.005	0.21	3.51E+08	1.06E+07	3.01
L-Sarcosine	2.17	0.005	0.21	9.82E+07	4.60E+06	4.69
L-Serine	2.09	0.004	0.19	1.23E+08	2.40E+06	1.95

Table D. 1 *is continued on the next page.*

	Retention time (min)			Peak area (a.u.)		
Amino acids	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
Taurine	2.20	0.004	0.18	7.80E+07	9.49E+05	1.22
L-Threonine	2.16	0.004	0.19	4.39E+08	1.10E+07	2.51
L-Tryptophan	11.36	0.004	0.04	1.06E+09	1.80E+07	1.70
L-Tyrosine	4.51	0.009	0.20	5.55E+08	1.34E+07	2.42
L-Valine	2.41	0.005	0.20	3.92E+08	8.63E+06	2.20

Table D. 2: Intermediate precision of metabolomics for amino acids in plasma. *The amino acid retention time and peak area intermediate precision was found for one plasma sample. The mean, standard deviation (SD), and the relative standard deviation (%RSD) were calculated.*

	Retention time (min)			Peak area (a.u.)		
Amino acid	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
β-Alanine	2.06	0.05	2.36	1.82E+07	2.61E+06	14.4
L-Alanine	2.15	0.05	2.17	3.40E+08	6.73E+07	19.8
L-α-Aminoadipic acid	2.42	0.04	1.61	1.61E+08	4.25E+07	26.4
L-α-Amino-n-butyric acid	2.29	0.05	2.03	1.64E+08	2.99E+07	18.3
D,L-β-Aminoisobutyric acid	2.16	0.06	2.62	1.66E+08	2.80E+07	16.9
γ-Aminobutyric acid	2.07	0.04	1.86	6.14E+07	1.76E+07	28.7
L-Anserine	2.05	0.05	2.21	2.86E+09	6.95E+09	242.7
L-Arginine	2.06	0.04	2.13	1.32E+08	2.81E+07	21.2
L-Aspartic acid	2.19	0.04	2.04	1.34E+08	2.27E+07	16.9
L-Carnosine	1.93	0.16	8.45	1.89E+07	4.92E+06	26.0
L-Citrulline	2.26	0.10	4.54	3.69E+08	3.85E+07	10.4
Creatinine	2.27	0.12	5.35	8.63E+08	2.68E+08	31.1
L-Cystathionine	2.10	0.05	2.56	1.13E+08	2.30E+07	20.4
L-Cystine	2.10	0.05	2.42	3.49E+10	8.50E+10	243.8
Ethanolamine	2.09	0.09	4.47	2.80E+09	6.83E+09	243.8
L-Glutamic acid	2.25	0.05	2.28	3.10E+08	3.62E+07	11.7
Glycine	2.10	0.05	2.19	5.18E+07	5.66E+06	10.9
L-Histidine	2.06	0.05	2.24	8.59E+07	2.01E+07	23.4
L-Homocysteine	2.50	0.04	1.76	6.60E+07	2.35E+07	35.6
δ-Hydroxylysine	1.75	0.05	2.61	1.24E+08	2.88E+07	23.3
Hydroxy-L-proline	2.30	0.05	2.00	4.62E+08	6.16E+07	13.3
L-Isoleucine, L-leucine	3.47	0.07	2.11	2.66E+09	4.19E+08	15.8
L-Lysine	1.75	0.05	2.58	1.87E+08	3.50E+07	18.7
L-Methionine	3.14	0.06	1.83	7.75E+08	1.31E+08	16.9
1-Methyl-L-histidine, 3-Methyl-L-histidine	2.08	0.05	2.29	2.18E+08	6.13E+07	28.1

Table D. 2 *is continued on the next page.*

	Retention time (min)			Peak area (a.u.)		
Amino acids	**Mean**	**SD**	**%RSD**	**Mean**	**SD**	**%RSD**
L-Ornithine	1.75	0.05	2.58	6.26E+07	1.31E+07	21.0
L-Phenylalanine	6.84	0.30	4.34	1.77E+09	3.59E+08	20.3
L-Proline	2.58	0.05	1.83	2.98E+08	4.00E+07	13.4
L-Sarcosine	2.23	0.06	2.77	9.28E+07	5.71E+07	61.5
L-Serine	2.13	0.05	2.22	1.26E+08	7.59E+06	6.04
Taurine	2.24	0.05	2.20	8.46E+07	1.04E+07	12.3
L-Threonine	2.19	0.05	2.18	3.95E+08	3.80E+07	9.6
L-Tryptophan	11.25	0.17	1.54	1.21E+09	2.42E+08	19.9
L-Tyrosine	4.27	0.16	3.72	7.63E+08	2.28E+08	29.9
L-Valine	2.44	0.04	1.73	3.97E+08	5.36E+07	13.5

Appendix E - Matrix effects on amino acid peak area and retention time in plasma

Table E. 1: Matrix effects on amino acid peak area and retention time in plasma. *This table list amino acids included in the study of matrix effects in plasma. Chromatographic matrix effects or ΔRT (min) was studied, as well as ion suppression and ion enhancement (MF, %).*

Amino acid	ΔRT (min)	MF (%)	Endogenous	Comment
β-Alanine	0.01	994.4	No	Elevated peak area
L-Alanine	0.01	85.8	Yes	
L-α-Aminoadipic acid	0.00	68.2	Yes	Peak area reduction
L-α-Amino-n-butyric acid	-0.01	82.6	Yes	
D,L-β-Aminoisobutyric acid	0.02	80.0	Yes	
γ-Aminobutyric acid	0.04	82.6	No	
L-Anserine	0.13	68.1	No	Retention time shift
L-Arginine	0.14	66.2	Yes	Peak area reduction, retention time shift
L-Aspartic acid	0.00	81.3	Yes	
L-Carnosine	0.00	63.8	No	
L-Citrulline	0.00	94.0	Yes	
Creatinine	0.01	85.9	Yes	
L-Cystathionine	0.03	40.0	No	
L-Cystine	0.02	59.3	Yes	Peak area reduction
Ethanolamine	0.14	57.1	Yes	Peak area reduction, retention time shift
L-Glutamic acid	0.00	91.7	Yes	
Glycine	0.02	65.0	Yes	Peak area reduction
L-Histidine	0,14	57.4	Yes	Peak area reduction, retention time shift
L-Homocysteine	-0.03	31.9	No	
δ-Hydroxylysine	0.00	37.3	No	Tailing and peak area reduction
Hydroxy-L-proline	0.00	91.0	Yes	
L-Isoleucine, L-leucine	0.00	99.0	Yes	
L-Lysine	0.00	63.2	Yes	Tailing and peak area reduction
L-Methionine	-0.02	110.7	Yes	Fronting, wide peak
1-Methyl-L-histidine, 3-Methyl-L-histidine	0.12	41.4	Yes	Peak area reduction
L-Ornithine	0.00	39.6	Yes	Tailing and peak area reduction
L-Phenylalanine	-0.01	110.0	Yes	

Table E. 1 *is continued on the next page.*

Amino acid	ΔRT (min)	MF (%)	Endogenous	Comment
L-Proline	-0.02	35.7	Yes	Peak area reduction
L-Sarcosine	0.00	82.6	No	
L-Serine	0.01	73.0	Yes	
Taurine	0.00	72.8	Yes	
L-Threonine	0.00	87.7	Yes	
L-Tryptophan	0.00	107.1	Yes	
L-Tyrosine	-0.01	91.2	Yes	Wide peak
Urea	0.00	111.1	Yes	
L-Valine	-0.02	78.8	Yes	