Chemical Isotope Labeling LC-MS for Universal Urine Metabolomics

By

Yingwen Wang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Alberta

Abstract

Metabolomics refers to the characterization and quantification of small molecule metabolic products in a biological specimen. It is an emerging and evolving science studying the practice of precious medicine. Metabolomics study has been used to understand individual variations influenced by both genetic and environmental factors. Thus, measurement of metabolites has played an important role in clinical practice since the concept was introduced. In order to provide qualitative and quantitative information with high metabolite coverage, chemical isotope labeling (CIL) method has been developed. CIL is able to target different submetabolomes by adding isotope tags to improve separation, sensitivity and capability of relative quantification. This “divide and conquer” technology simplifies the platform of metabolomics study and promotes analytical performance of metabolites.

My research focuses on utilizing CIL LC-MS approaches to profile amine/phenol submetabolomes to evaluate matrix effects in universal urine metabolome standard. In the first part of this thesis, CIL LC-MS was used to investigate the effect of urine sample matrix on metabolome analysis by comparing the absolute concentrations of selected metabolites with concentrations from external standards. The extent of matrix effects on labeling was evaluated by comparison of samples at different concentrations that were labeled using the same protocol. In the second part, CIL LC-MS was used to profile the amine and phenol submetabolome of universal urine

metabolome standard. Matrix effects for a large number of identified metabolites were evaluated.

Acknowledgements

I would like to sincerely thank my supervisor, Dr. Liang Li, for his enthusiasm, motivation, support and patience on my research. I am very lucky to have him as my mentor, who led me to the world of analytical chemistry and metabolomics. I am grateful to work and learn in Dr. Li's lab for the past 3 years. Thanks to him for the joyful activities he organized, the great opportunities he provided and the professional guidance he gave. These memories made my studies a wonderful experience.

I would like to thank my supervisory committee, Dr. Michael Serpe and Dr. Lingzi Sang for their guidance and kind suggestions on my research projects.

I would like to thank my group members for their kind help and friendship over the years. I would like to thank the all the current members with special thanks to Shuang Zhao, Xian Luo and Wei Han, who gave me the lab training and were always willing to help me to overcome obstacles. In no particular order, thanks to Xiaohang Wang, Hao Li, Xinyun Gu, Wan Chan, Minglei Zhu, Adriana Zardini Buzatto, Barinder Bajwa, Chufan Wang, Zhan Cheng and Carlos Canez Quijada.

Finally, I would like to thank my family and friends, for their company and support during the course of my degree. I am gratefully indebted to my parents, Zhuocheng Wang and Xiaoyan Wei. I would like to thank them for giving birth to me, for bringing me up with their unconditional love and for supporting me through my whole life. Lastly, I would also be appreciative to all of my friends for their encouragement and awesome friendship.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAS	Amino Acid Standard
ACN	Acetonitrile
ANOVA	Analysis of variance
CE	Capillary electrophoresis
CIL	Chemical isotope labeling
CSF	Cerebrospinal fluid
DmPA	p-Dimethylaminophenacyl
DnsCl	Dansyl chloride
ESI	Electrospray ionization
FTICR	Fourier transform ion cyclotron resonance
GC	Gas chromatography
HILIC	Hydrophilic interaction liquid chromatography
HMDB	Human metabolome database
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
LLE	Liquid-liquid extraction
MeOH	Methanol
MPA	Mobile phase A
MPB	Mobile phase B
MS	Mass spectrometry
NMR	Nuclear magnetic resonance spectroscopy

PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
ppm	Parts per million
QTOF	Quadrupole time of flight
RP	Reversed phase
RPLC	Reversed-phase liquid chromatography
RT	Retention time
SPE	Solid-phase extraction
UHPLC	Ultra high-performance liquid chromatography
UMS	Universal metabolome standard
UUMS	Universal urine metabolome standard

List of Symbols

Da	Dalton
m/z	Mass-to-charge
v/v	Volume to volume ratio

Chapter 1 Introduction

1.1 Multi "omics"

Biological molecules produced by cells and living organisms are complex with multiplex characteristics and are critical to typical biological processes. High-throughput technologies called "omics" have been developed in order to achieve a better understanding about biological molecules.[1] Areas of research used to assess comprehensive biological functions could be classified as "omics", which include genomics, transcriptomics, proteomics and metabolomics.[2] Genomics, as the most mature of the omics fields, focuses on whole genomes and genetic variants associated with disease or medical treatment through DNA sequencing.[3] However, most available approaches have limited scope to determine DNA alterations, such as duplications, deletions and inversions.[4] Transcriptomics is used to examine RNA levels genome-wide, which includes transcription and expression levels, functions, locations, trafficking and degradation.[5,6] Since some knowledge of transcriptome is still from gene predictions, this technology has been criticized that mRNA levels cannot be consistently used to predict the abundance of protein.[7-9]

Proteomics is a new research field focusing on studying proteins on a grand scale, such as quantifying peptide abundance, modification, and interaction.[10] And being revolutionized by MS-based approaches enables this technology to achieve a high-throughput analyses of thousands of proteins in cells and body fluids.[11,12] As the end point of the omics cascade, metabolomics quantitatively studies multiple small molecule types, including amino acids, fatty acids, carbohydrates, and other products of metabolic functions.[2,13] The multi-omics approach provides a pathway to analyze the comprehensive biological processes.

1.2 Metabolomics

1.2.1 Significance of Metabolomics

Metabolomics refers to the characterization and quantification of the small molecule metabolic products in a biological specimen, is an emerging and evolving science studying the practice of precious medicine.[14] With small changes induced by diet or environment, the concentration of metabolite could be varied dramatically.[15] The detection and measurement of a huge profile of large number of metabolites enable the investigation of changes induced by external stimuli (e.g., drug treatment).[16,17] This advantage makes untargeted metabolomics a useful tool in biomarker discovery.[16] Biomarkers are defined as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention".[18] It is reported that, diagnostic biomarkers could demonstrate detectable changes before disease symptoms become noticeable, which is important in early disease diagnosis.[19,20] Therefore, measurements of metabolites has been playing an irreplaceable role in clinical practice since the concept was introduced.

1.2.2 Analytical Platforms for Metabolomics

In general, the analytical platforms for metabolomics studies include two processes: detection and separation. Nuclear magnetic resonance (NMR) and mass spectrometry (MS) as two major platforms are commonly used for detection. As for separation techniques, gas chromatography (GC), liquid chromatography (LC) and capillary electrophoresis (CE) are three high-throughput techniques usually incorporated in MS-based metabolomic analysis. To obtain comprehensive information of metabolites, analytical techniques are chosen based on their unique characteristics.

NMR is known as one of the very first spectroscopic techniques used in metabolomics studies.[21] The principle of NMR is based on atom nuclei absorb and re-emit energy when an external magnetic field is applied.[22] Due to the relatively high metabolome coverage and abundant database resource, it can provide detailed structural information and identify molecules in a short time.[23,24] It does not require going through major sample preparation processes, which simplifies and speeds up the overall data acquisition time.[25,26] Samples can be recovered for further observation and analysis

since NMR is a non-destructive technique. These strengths make NMR a commonly used method for exploring biomarkers in disease studies.[27,28] However, the major drawbacks for NMR analysis are the relatively low sensitivity as detection can only be achieved above micromolar range; requiring a large quantity of samples. Another issue of NMR is the convoluted spectral signals for complex mixtures.[29] In comparison, MS has a much higher sensitivity, and with high-resolution MS technique, the accurate mass measurement can be achieved.

MS has become a widely used platform in metabolomics. There are two major components in a mass spectrometer: an ionization source and a mass analyzer. To have a high metabolome coverage, both positive and negative ion mode data are collected simultaneously. This can be achieved by using electrospray ionization (ESI) as an ionization source, which applies soft ionization energy leading to minimal fragmentation. The overall mechanism of ESI can be summarized as follows: After eluting from LC, the liquid containing analytes of interest is dispersed and charged at ESI capillary tip; the initial droplet is shrunk as heated dry gas is applied; with droplet disintegration processes repeated, small “offspring” droplets are formed. Finally, the field strength at droplet surface becomes large enough to assist ion escape from the liquid phase into the gas phase.

Various mass analyzers can be chosen according to their detectable mass range, sensitivity as well as resolving power. Time of flight mass spectrometer (TOF-MS) has become the most commonly used mass analyzer due to its high sensitivity as well as high resolution. Compared with Fourier-transform ion cyclotron mass spectrometers (FTICR-MS), which can provide extremely high resolutions (100,000-1,000,000), the relatively low cost of TOF-MS is more popular for metabolomics. Each of the existing analyzers has its unique characteristics and shortcomings. The combination of different analyzers into one mass spectrometer has been widely used. For instance, quadrupole-TOF-MS stands out for rapid screening analysis.[30]

For complex mixture analyses, chromatographic separation of metabolites before MS has been extensively used in metabolomics.[31] GC, CE and LC are three predominant chromatographic techniques commonly incorporated with MS. GC-MS

can provide fast scan speeds with mass resolution and a high level of accuracy. However, GC-MS can only apply to volatile, thermally and energetically stable compounds.[32] CE-MS is suitable for the analysis of highly polar and ionic metabolites with a relatively low cost.[33] The limitations have been summarized as low sensitivity, poor reproducibility and electrochemical reactions of metabolites.[34] Among these platforms, LC-MS has been widely used due to its high sensitivity, good compatibility and accurate quantification ability. The development of ultra high-performance liquid chromatography (UHPLC) has increased chromatographic resolution and peak capacity. To achieve a better coverage, both hydrophilic interactions liquid chromatography (HILIC) and reversed-phase liquid chromatography (RPLC) are employed for different separation requirements. For example, HILIC is applied for polar metabolites separation, while RPLC is mainly served for non-polar compounds. Since one type of column can only be used for certain polarity of metabolites, the detection process remains challengeable.

1.2.3 Chemical Isotope Labeling Metabolomics

Metabolomics, a widely used technology, aims to provide qualitative and quantitative information for as many metabolites as possible. In order to achieve this goal, experimental design and sample preparation with high metabolite coverage, sensitive detection, accurate quantification and accurate metabolite identification are crucial in MS-based metabolomics study. However, in practice, metabolomics faces several challenges. Firstly, unlike genomic and proteomic methods, metabolites in biological samples present disparate physical and chemical properties.[35] Since there is no single technique that could resolve all these metabolites simultaneously, comprehensive metabolomic technology platforms are required in complex metabolites separation (such as combining RPLC and HILIC).[35,36] Moreover, ion suppression is found to be an issue when using MS to detect complex mixtures. When ESI is applied as an ionization source, it is easier for ions on the surface of the shrunken droplet to get ionized than those in the center of the droplet. Ion suppression could be caused by heterogeneous ionization efficiency. A proper separation method is likely to reduce ion

suppression, while this could double instrument time and make experiments becoming more complicated.

Our lab has developed a "divide and conquer" technology and applied to divide the whole metabolome into four different submetabolomes based on chemical functional groups. Chemical isotope labeling (CIL) is used to target a particular functional group (submetabolome) of interest with high coverage and the combined results are used for the complete metabolome analysis. For submetabolome containing amine/phenol[37], dansyl chloride (5-(dimethylamino)naphthalene-1-sulfonyl chloride, DnsCl) has been chosen as the labeling reagent. In this approach, experiment samples are labeled by ^{12}C-DnsCl, while an internal reference sample is labeled with ^{13}C-DnsCl. After mixing and LC-MS analysis, metabolites are detected as peak pairs, containing a lighter peak and a heavier peak with a mass-to-charge ratio (m/z) difference of 2.0067 Da. The ratio between two peaks is used for relative or absolute quantification of metabolites. There are several advantages by using dansyl labeling approach. Firstly, the aromatic naphthyl part increases the hydrophobicity of the labeled metabolites which increases retention in RPLC. Secondly, the tertiary amine enhances ESI chargeability and further enhances ionization efficiency. Additionally, the two methyl groups on the tertiary amine can introduce ^{13}C isotope, which can be used for relative quantification. Together, DnsCl labeling approach improves quantification and detection and achieves a better separation for hydrophilic compounds on RPLC column. In the last decade, our lab has also developed ^{12}C-/^{13}C DnsCl with base-activation for the hydroxyl submetabolome[38], ^{12}C-/^{13}C-dimethylaminophenacyl (DmPA) bromide labeling for the carboxyl submetabolome[39] and ^{12}C-/^{13}C dansylhydrazine labeling for the carbonyl submetabolome.[40] These comprehensive techniques for metabolites profiling and biomarker discovery in various samples such as cells[41], serum[42], sweat[43] and cerebrospinal fluid (CSF)[44] have been reported.

1.2.4 Workflow for Metabolomics

1.2.4.1 Sample Preparation

A typical LC-MS based metabolomics study relies heavily on experimental design and sample preparation. In general, experiment steps include sample preparation, chemical isotope labeling, sample normalization, metabolite identification and statistical analysis.[45] Sample preparation is primarily depending on sample types. Some commonly used samples include cell extract, blood, urine, CSF and saliva. For cellular metabolomics, cell harvest and metabolism quenching are always required. When handling protein enriched biofluid such as blood, metabolites are extracted from complex matrix by protein precipitation with organic solvent. When necessary, extraction methods, such as liquid-liquid extraction (LLE) and solid-phase extraction (SPE) are used to pre-concentrate certain metabolites.

1.2.4.2 Chemical Isotope Labeling

As described above, experiment samples are individually labeled with ^{12}C-DnsCl, while an internal reference sample is labeled with ^{13}C-DnsCl. Then the mixed sample is analyzed by direct injection MS, CE-MS, GC-MS or LC-MS, with LC-MS as the most commonly used one. In this process, each metabolite is detected as peak pair rather than a single mass peak. The light peak of the pair is for ^{12}C-labeled individual sample, and the heavy peak belongs to ^{13}C-labeled internal reference. Relative concentration can be obtained by calculating the intensity ratio of chromatographic peak pair. This information is important for metabolomics analysis to find metabolites with significant changes. Moreover, absolute quantification of selected biomarker candidates is able to be conducted later.

1.2.4.3 Sample Normalization

In biomarker discovery study, accurate quantitative analysis relies heavily on the sample normalization step. Some biological samples have large concentration variations, such as urine, depending on water intake, diet and sweating. In this case, concentration differences between urine samples should be excluded but the variations be kept due to biological differences. There are two types of sample normalization method which have been commonly used: pre-acquisition normalization and post-

acquisition normalization. In pre-acquisition normalization, the total concentration of metabolites within each sample is determined, followed by changing injection volume of various samples to ensure equal amount of samples being analyzed.[46-48] Creatinine has been commonly used as a pre-acquisition normalization reference for urine sample.[49] As for post-acquisition normalization, samples are directly loaded on analytical platforms and statistical adjustment will be applied afterwards.[50] In general, the total signal intensity is proportional to the total concentration of metabolite. Therefore, the concentration of individual metabolite is normalized by total intensity of the signal.

1.2.4.4 Metabolite Identification and Statistical Analysis

In disease studying and novel treatment strategies development fields, accurate and complete identification of metabolites play important roles by directly observing metabolic activities.[51] For metabolite identification, putative identification information is firstly obtained by searching accurate mass of metabolites in metabolite database, such as METLIN[52] and Human Metabolome Database (HMDB).[53] Then, further experimental data are needed, including retention time and tandem mass spectrometry (MS/MS) spectra to obtain fragment patterns, followed by the data comparison with standards to get accurate identification of the metabolites of interest. Although metabolome coverage is still limited compared to the number of metabolites existing in biological samples, current metabolomics analyses are being developed to generate a massive amount of data. Finally, statistical tools are applied for sample classification and biomarker discovery studies, such as principal component analysis (PCA), partial least square discriminant analysis (PLS-DA), analysis of variance (ANOVA) and volcano plot.

1.3 Scope of the Thesis

The objective of this research work is to develop chemical isotope labeling (DnsCl labeling) LC-MS platform to analyze matrix effects in universal urine metabolome standard (UUMS).

In Chapter 2, CIL LC-MS is used as a high-coverage and quantitative metabolome profiling technique to investigate the effect of urine sample matrix on metabolome analysis. The absolute concentrations of selected metabolites from standard addition were compared with concentrations from external standard method to evaluate matrix effects. Besides, comparison of samples at different concentrations that were labeled using the same protocol enables the determination of the extent of matrix effect on labeling.

In Chapter 3, CIL LC-MS is used to profile the amine and phenol submetabolome of UUMS. Matrix effects for a large number of identified metabolites were evaluated.

Chapter 2 Chemical Isotope Labeling LC-MS for Matrix Effects Test in Universal Urine Metabolome Standard

2.1 Introduction

In recent years, biomarkers are being used as an important tool for disease studying and novel treatment strategies development, and this is becoming an emerging approach in many medical areas.[54,55] Many strategies have been applied to biomarker discovery including: metabolomics, proteomics and epigenetics.[56,57] Among these technologies, metabolomics has been proved to be a useful technique in a number of research areas ranging from disease biomarker discovery to systems biology studies.[58,59] For biomarker discovery, metabolomics is an emerging field, which performs quantification and identification of many small molecules in biological systems. Through metabolome profiling, metabolites that differentiate diseases can be used as potential biomarker candidates.[60,61] It is important to perform accurate and precise measurement of metabolites concentration changes in biomarker discovery.[62,63]

Chemical isotope labeling liquid chromatography mass spectrometry (CIL LC-MS), is an analytical platform which provides qualitative and quantitative metabolomic information for metabolomics research.[37] CIL is used to target a particular functional group (submetabolome) of interest with proper labeling reagent. After labeling, the chemical and physical properties of metabolites can be altered. In addition, chemical-group-based submetabolome can be efficiently separated using reverse phase liquid chromatography (RPLC), then effectively ionized and detected by MS. This approach improves detection sensitivity and quantification accuracy, enables the generation of a comprehensive profile of the submetabolome.[37]

By using differential isotope labeling of individual samples (e.g., labeled with ^{12}C-reagent) and their control (e.g., labeled with ^{13}C-reagent), CIL LC-MS overcomes matrix effects and ion suppression associated with MS detection to a large extent. [37,39] This method has been used as a powerful tool for relative quantification of metabolites in comparative samples. However, chemical labeling itself may encounter matrix

effects.[64] In order to obtain absolute concentrations of metabolites in individual samples and also eliminate matrix effects at the same time, standard addition method has been used to achieve this goal.[65,66] The process involves adding a series of small increments of the analyte to the sample and measuring the signal. One significant disadvantage of this method is the extra time required for making the additions and measurements.[67,68]

In this work, we investigated the effects of urine sample matrix on metabolome analysis using a high-coverage and quantitative metabolome profiling technique based on differential CIL LC-MS. We compared absolute concentrations of selected metabolites from standard addition strategy with concentrations from external standard method to evaluate matrix effects. Besides, comparison of samples at different concentrations that were labeled using the same protocol enables us to determine the extent of matrix effect on labeling.

2.2 Experimental

2.2.1 Chemical and Reagents

All the chemicals and reagents, unless otherwise stated, were from Sigma-Aldrich Canada (Markham, ON, Canada). ^{13}C-dansyl chloride was synthesized in our lab with the procedures published previously[37] and is available from MCID.chem.ualberta.ca. LC-MS grade water, acetonitrile (ACN), and methanol (MeOH) were from Thermo Fisher Scientific (Edmonton, AB, Canada).

2.2.2 Overall Workflow

Figure 2.1 shows the overall workflow of this study. In this work, four amino acids were selected as target metabolites in universal urine metabolome standard (UUMS). For each amino acid, after being injected into LC-MS for analysis, the absolute concentration of ^{12}C-isotope labeled amino acid in UUMS was determined by using a ^{13}C-isotope labeled amino acid added as internal standard. A calibration curve using ^{12}C-isotope labeled amino acid solution was built up, which was used to quantify the amino acids after conducting a serial dilution of ^{12}C-isotope labeled UUMS. By

comparison of the concentrations of selected amino acids from two different approaches, the effect of dilution and matrix effects on amino acids were evaluated. The extent of matrix effect on labeling was determined from the results obtained when samples were prepared at different concentrations. The detailed experimental conditions in the workflow are described below.

2.2.3 Dansylation Labeling

The frozen urine sample was re-dissolved to 1000μL with water to reach its original concentration. Then 12.5 μL of 250 mM sodium carbonate/sodium bicarbonate buffer was added to 25 μL of urine sample. The solution was then mixed with 37.5 μL of freshly prepared ^{12}C-DnsCl solution (18 mg/mL for light labeling, individual samples) or ^{13}C-DnsCl solution (18 mg/mL for heavy labeling, internal standard samples). After incubation for 45 min at 40 °C, 7.5 μL of 250 mM sodium hydroxide solution was added to the reaction mixture, followed by incubation at 40 °C for another 10 min. Finally, 30 μL of formic acid (425 mM) in 1:1 ACN/H_2O was added.

2.2.4 LC-MS

Each ^{12}C-labled individual sample was mixed with a ^{13}C-labled internal standard by equal volume. LC-MS was done using a Themo Scientific Dionex Ultimate 3000 UHPLC System (Sunnyvale, CA) linked to a Bruker Maxis II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker, Billerica, MA). The LC column was an Agilent reversed phase Eclipse Plus C18 Column (2.1 mm x 10 cm, 1.8 μm particle size). LC mobile phase A was 0.1% (v/v) formic acid in water, and mobile phase B was 0.1% (v/v) formic acid in ACN. The gradient elution profile was as follow: t = 0 min, 25% B; t = 10 min, 99% B; t = 13 min, 99% B; t = 13.1 min, 25% B, t = 16 min, 25% B. Flow rate was 400 μL/min. All mass spectra were collected in the positive ion mode. The MS conditions for MS were as follows: dry gas flow, 8 L/min; dry temperature, 230 °C; capillary voltage, 4500 V; nebulizer, 1.0 bar; end plate offset, 500V; spectra rate, 1.0 Hz.

2.2.5 Data Processing, Metabolite Identification and Statistical Analysis

After LC-Q-TOF-MS analysis, all the spectra were first converted to .csv files by Bruker Daltonis Data Analysis 4.4 software. The peak pairs were extracted from .csv files by IsoMS.[69] Data from multiple runs were aligned by retention time and accurate mass. After the alignment, Zerofill program[70] was applied to recover the high-confidence peak pair ratios lost during the previous data processing steps. Based on the accurate mass and retention time, positive metabolite identification was performed using dansyl standard library search, which contains 275 unique dansylated amines/phenols[71] (www.mycompoundid.org). Putative identification or match was performed by searching accurate mass against MyCompoundID library, which contains 8,021 known human metabolites and 375,809 predicted metabolites[72] (www.mycompoundid.org).

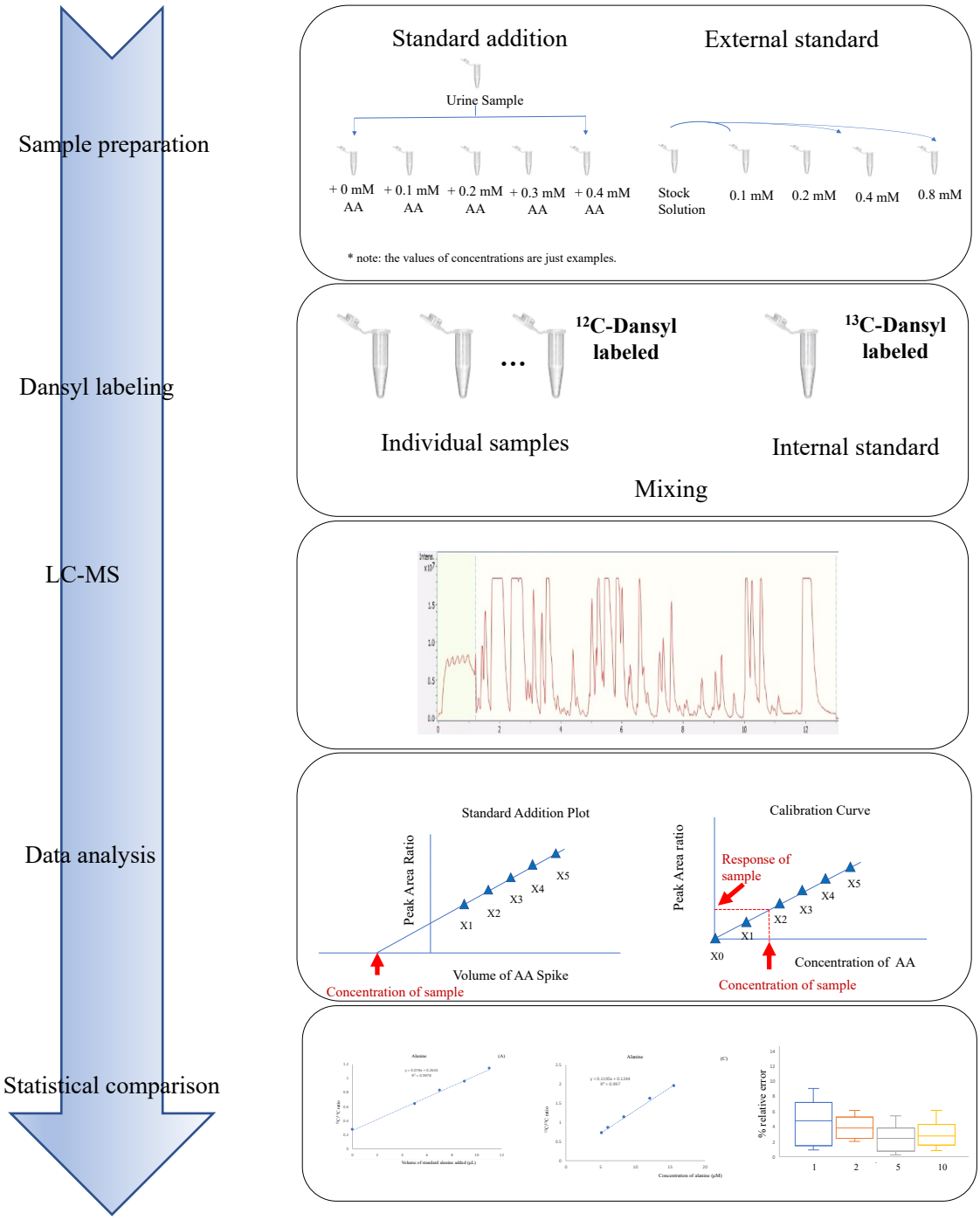


Figure 2.1. Workflow for sample preparation and dansylation isotope labeling LC-MS.

2.3 Result and Discussion

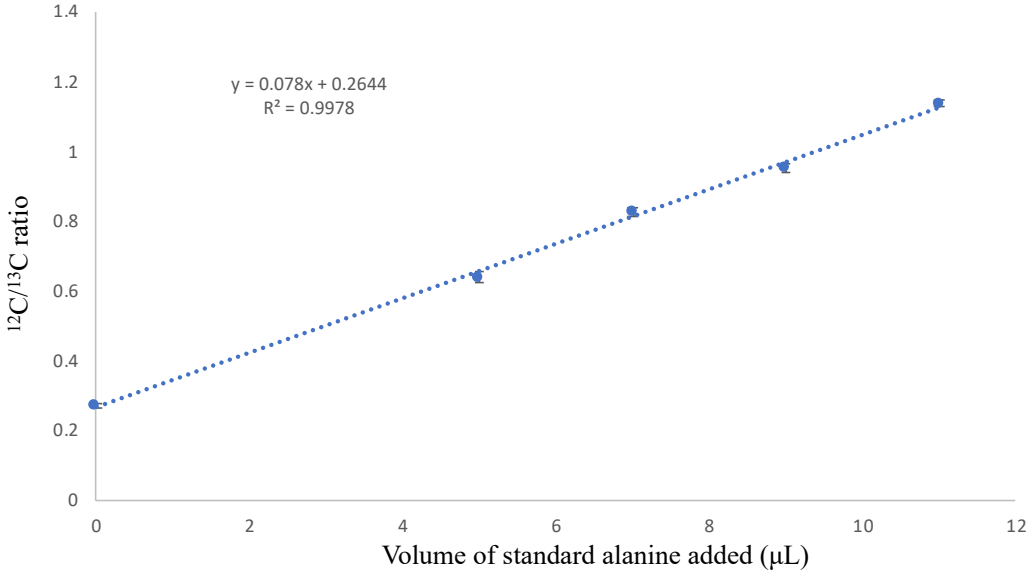
2.3.1 Absolute Concentrations of Metabolites from Standard Addition Method

Accurate measurement of metabolic changes in disease-associated processes and pathways is critical to discover new biomarkers. Although LC-MS is widely used for absolute quantification in targeted metabolomics, it often suffers matrix effects for

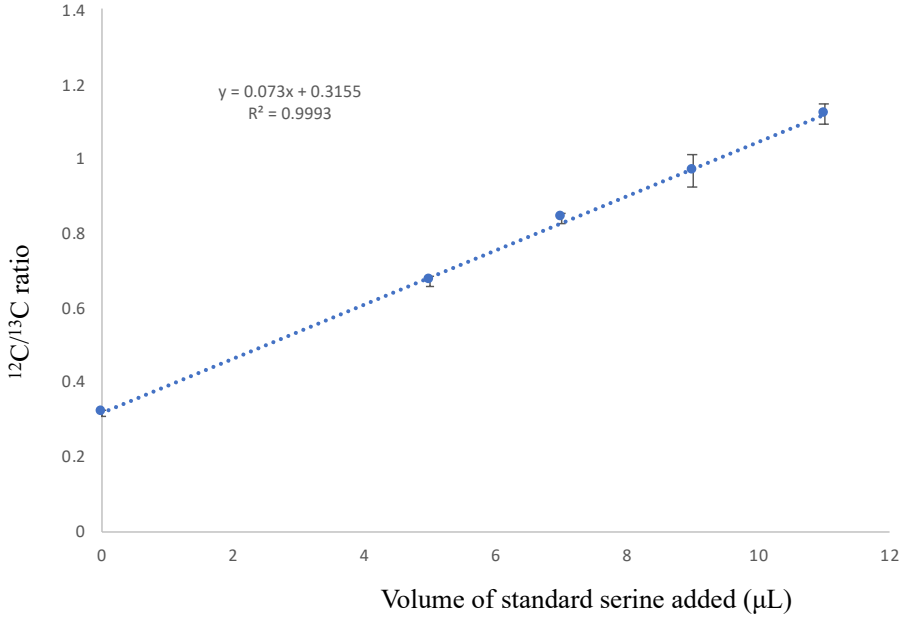
large scale untargeted analysis. To overcome matrix effects, differential CIL LC-MS has been used as a powerful tool for relative quantification of metabolites in comparative samples. A universal metabolome standard (UMS) can be used as a reference sample. By determining the concentrations of metabolites in UMS, absolute concentrations of these metabolites can be obtained. The general strategy of measuring the absolute concentrations of individual metabolites is from standard addition method. In this work, we applied this approach to evaluate the absolute concentrations of four amino acids in UUMS. Briefly, we started with differential CIL for selected metabolites, followed by preforming LC-MS for absolute concentrations determination.

In order to obtain the absolute concentrations of selected amino acids at the same time, amino acid standard (AAS) solution was chosen. It contains 17 different amino acids with known concentrations. The standard addition approach was performed by adding increasing small volume of standard solution directly to the aliquots of analyzed sample. In this work, five different volumes of AAS solution (0, 5, 7, 9, and 11 μL) were added into five identical UUMS solutions. The individual samples were labeled by ^{12}C-dansyl chloride, a ^{13}C-dansylation labeled sample with a 7 μL AAS addition served as internal standard. The ^{13}C-dansylation labeled internal standard was mixed with ^{12}C-dansyl chloride labeled individual samples by same volume. Then the same volume of mixtures prepared from all individual samples was injected into LC-MS. The peak pair of ^{12}C- and ^{13}C-labeled metabolites were integrated and plotted to generate graphs which showed peak area ratio against the volume of targeted amino acids added. From regression equations, the absolute concentrations of selected amino acids in UUMS can be obtained. Figure 2.2 shows the plots generated from standard addition approach for four amino acids: alanine (Figure 2.2A), serine (Figure 2.2B), phenylalanine (Figure 2.2C) and threonine (Figure 2.2D). As calculated from regression equations, the absolute concentrations for alanine, serine, phenylalanine and threonine were: 42.4 μM, 54.1 μM, 8.42 μM and 22.0 μM, respectively (Table 2.1).

Alanine (A)



Serine (B)



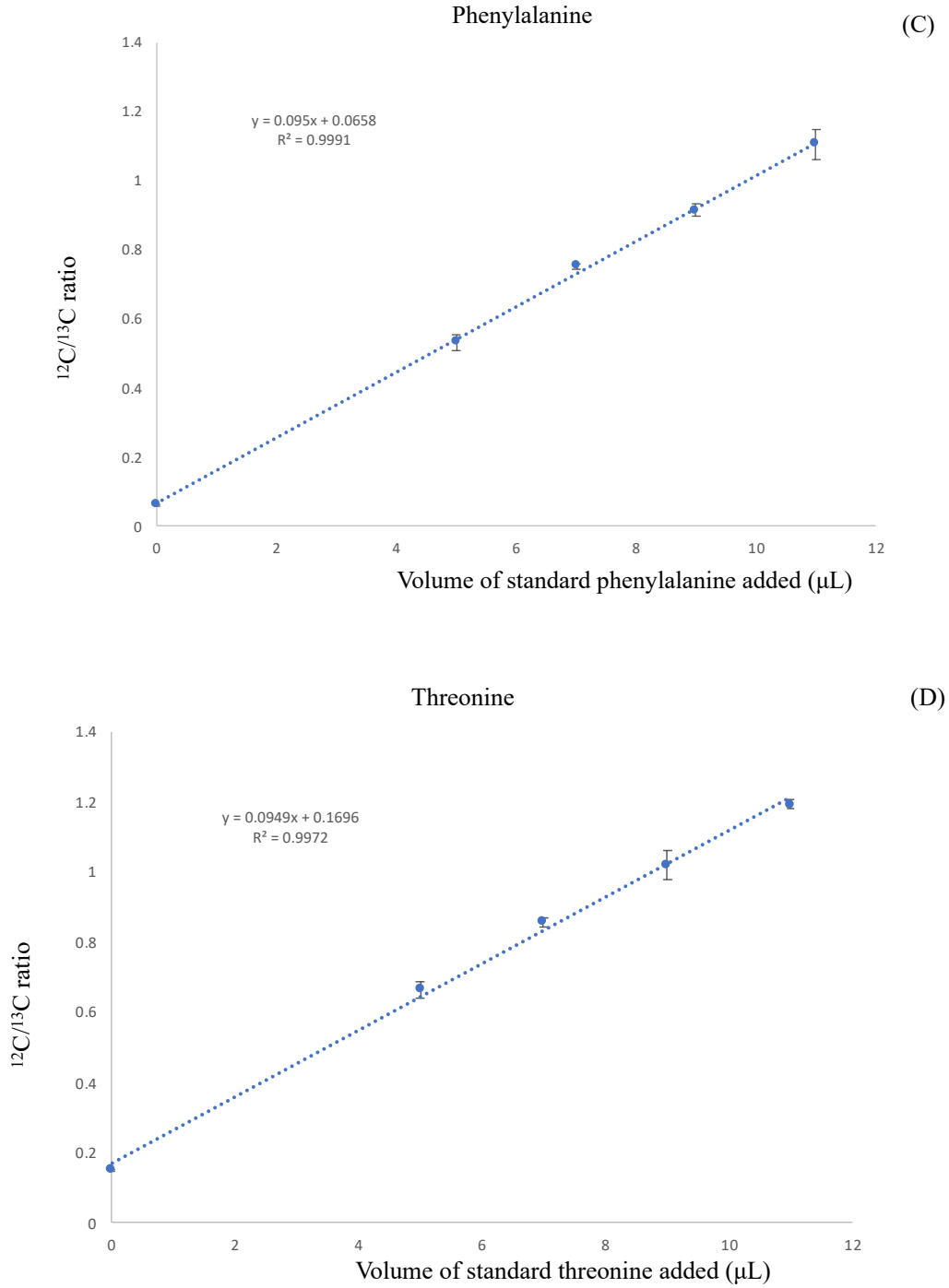


Figure 2.2. Standard addition plots for four dansyl labeled amino acids: (A) alanine, (B) serine, (C) phenylalanine and (D) threonine.

Amino Acid	Absolute Concentration (µM)
Alanine	42.4
Serine	54.1

Phenylalanine	8.42
Threonine	22.0

Table 2.1. Absolute concentrations for alanine, serine, phenylalanine and threonine contained within UUMS obtained from standard addition approach.

2.3.2 Concentrations of Metabolites from External Standard Method

The general strategy of finding a way to rapidly measure the concentrations of individual metabolites in UMS is to compare the standard addition method with the calibration method to evaluate the extent of matrix effect on metabolites of interest. Because of the lack of isotope standards for many metabolites, we use chemical isotope labeling to create isotope standard. However, chemical labeling itself may encounter matrix effects. Comparison of samples at different concentrations that are labeled using the same protocol can determine the extent of matrix effect on labeling. In this work, calibration curves were built up for four selected amino acids by using ^{12}C-isotope labeled AAS while ^{13}C-isotope labeled sample served as internal standard, followed by conducting serial dilutions of ^{12}C-isotope labeled UUMS and quantifying amino acids by using calibration curves. In order to achieve an accurate measurement for selected amino acids contained in different concentrations of UUMS, calibration curves were generated individually for serial diluted UUMS to ensure every amino acid lay in the middle of calibration curve. Then the effect of dilution and matrix effects on amino acids were evaluated by comparing the results from different diluted UUMS. Alanine, serine, phenylalanine and threonine were four amino acids to be evaluated, and the UUMS sample was diluted by 1-, 2-, 5- and 10-fold, respectively to test the extent of matrix effect.

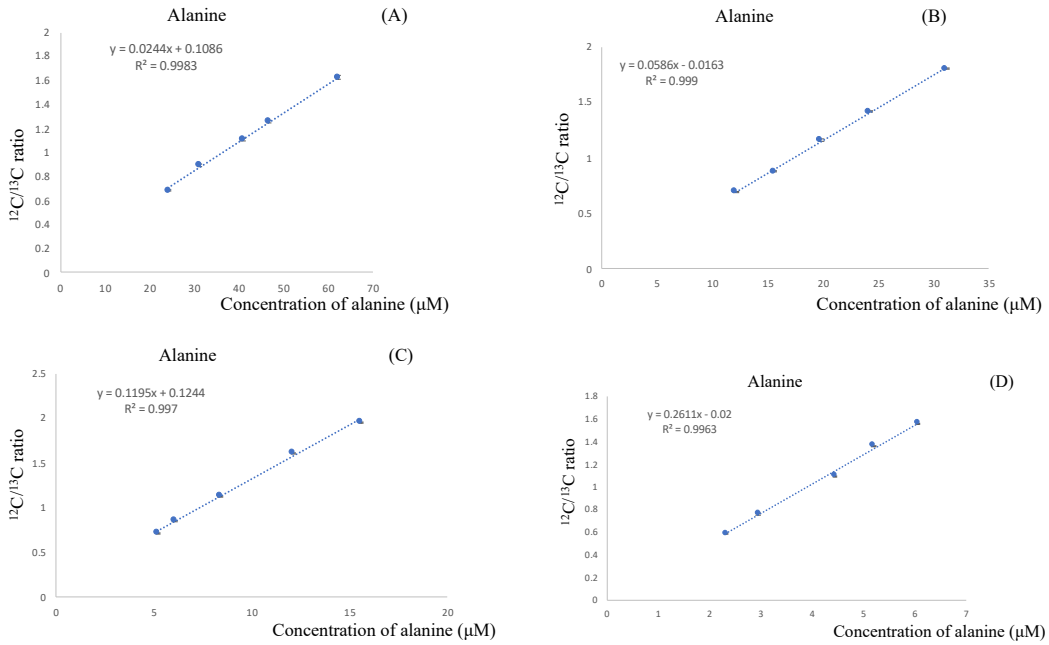


Figure 2.3. Calibration curves of alanine for (A) 1-, (B) 2-, (C) 5-, and (D) 10-fold diluted UUMS.

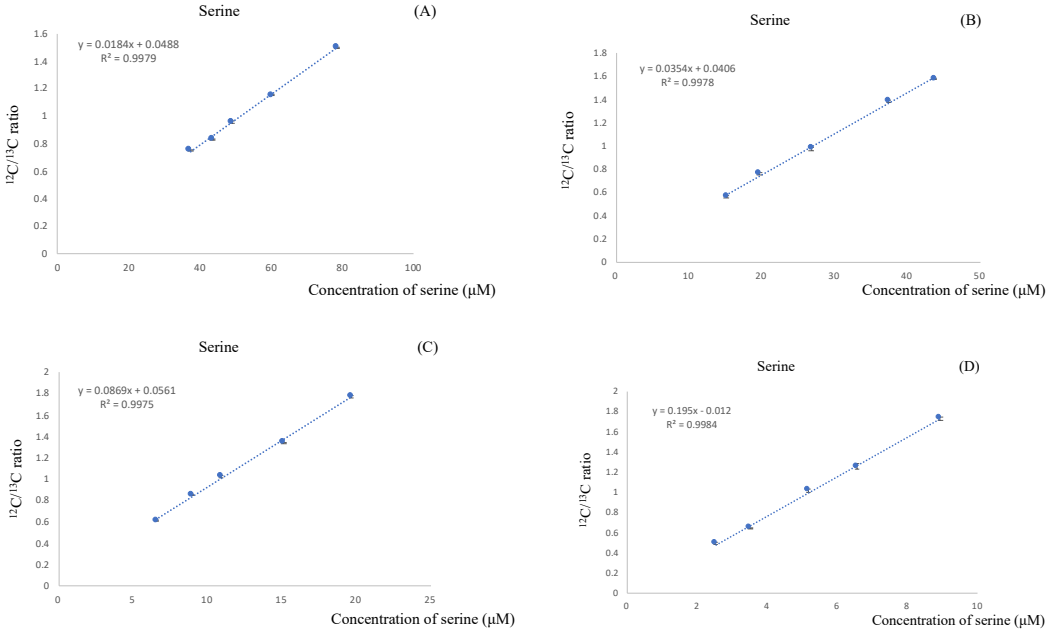


Figure 2.4. Calibration curves of serine for (A) 1-, (B) 2-, (C) 5-, and (D) 10-fold diluted UUMS.

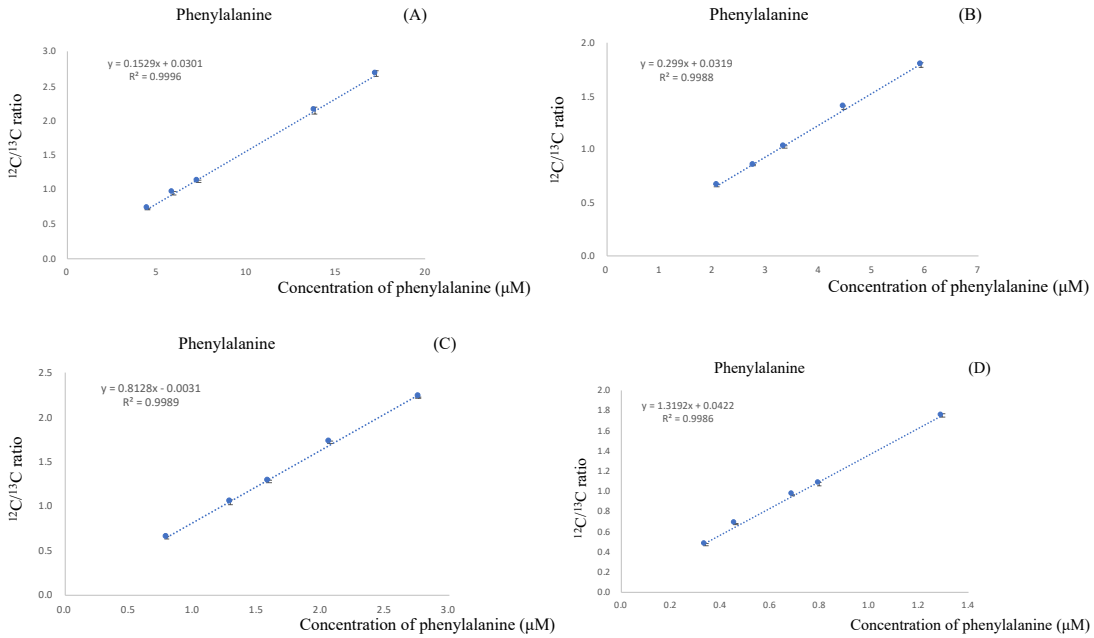


Figure 2.5**.** Calibration curves of phenylalanine for (A) 1-, (B) 2-, (C) 5-, and (D) 10-fold diluted UUMS.

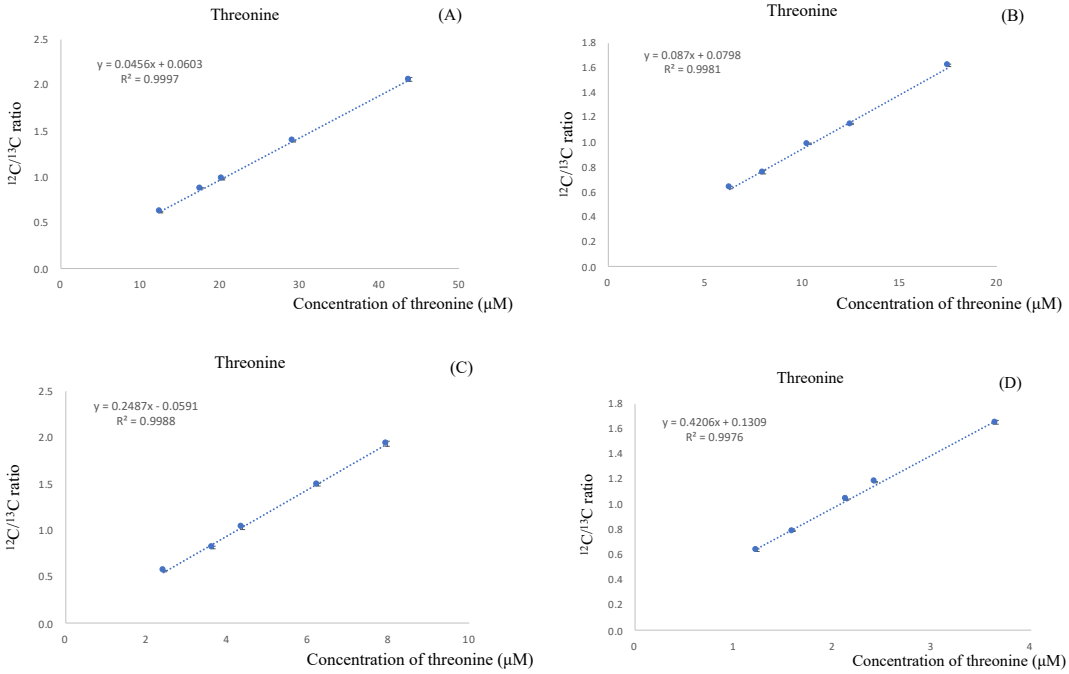


Figure 2.6. Calibration curves of threonine for (A) 1-, (B) 2-, (C) 5-, and (D) 10-fold diluted UUMS.

Alanine, serine, phenylalanine and threonine concentrated standard solutions were separately labeled with ^{12}C-dansyl chloride and then diluted into various concentrations. Furthermore, UUMS sample was diluted by 1-, 2-, 5- and 10-fold and ^{12}C-isotope labeling was performed under the same protocol. Calibration curves

corresponding to serial diluted UUMS were built up for four selected amino acids with ^{13}C-isotope labeled sample used as an internal standard. ^{13}C-isotope labeled internal standard was mixed with ^{12}C-dansyl labeled individual samples by equal volume. The same volume of mixtures was injected in LC-MS. From Figure 2.3 to Figure 2.6, calibrations curves for each amino acid contained in serial diluted UUMS were plotted as ^{12}C and ^{13}C peak area ratio against the concentration of individual amino acids in µM. The regression equations were shown for each calibration curve. For every figure, A, B, C and D represent dilution factors of 1-, 2-, 5- and 10-fold, respectively.

After quantitative analysis, concentrations for each of the selected amino acids contained within four diluted UUMS were calculated from calibration curves. For alanine, the concentrations within 1-, 2-, 5- and 10-fold diluted UUMS were: 42.1 ± 0.11, 42.1 ± 0.11, 8.77 ± 0.22 and 4.54 ± 0.10 µM, respectively (Table 2.2A). The standard deviation for each concentration measurement was the result of combined variations in biological triplicate and experimental triplicate (n=9). The results for serine within four serials diluted UUMS were: 59.4 ± 0.13, 30.2 ± 0.03, 11.5 ± 0.13 and 5.88 ± 0.16 µM, respectively (Table 2.2B). For phenylalanine, the concentrations for 1-, 2-, 5- and 10-fold diluted UUMS were 8.07 ± 0.19, 4.35 ± 0.14, 1.69 ± 0.07 and 0.88 ± 0.04 µM, respectively (Table 2.2C). For threonine, these were 23.3 ± 0.74, 12.0 ± 0.20, 4.72 ± 0.13 and 2.39 ± 0.06 µM in 1-, 2-, 5- and 10-fold diluted UUMS, respectively (Table 2.2D).

Alanine (A)

Dilution Factor	Concentration (µM)
1	42.1±0.11
2	21.2±0.03
5	8.77±0.22
10	4.54±0.10

Serine (B)

Dilution Factor	Concentration (µM)
1	59.4±0.13
2	30.2±0.03
5	11.5±0.13
10	5.88±0.16

Phenylalanine (C)

Dilution Factor	Concentration (µM)
1	8.07±0.19
2	4.35±0.14
5	1.69±0.07
10	0.88±0.04

Threonine (D)

Dilution Factor	Concentration (µM)
1	23.3±0.74
2	12.0±0.20
5	4.72±0.13
10	2.39±0.06

Table 2.2. Average concentrations (n=9) for (A) alanine, (B) serine, (C) phenylalanine and (D) threonine contained within 1-, 2-, 5-, and 10-fold diluted UUMS calculated from calibration curves.

2.3.3 Matrix Effects Analysis

In order to evaluate matrix effects, we compared absolute concentrations of alanine, serine, phenylalanine and threonine from standard addition strategy with concentrations from external standard method. The result in Table 2.3 presents the relative differences of concentrations obtained using external standard method from absolute concentrations in percentage for four selected amino acids.

Amino acid	Absolute Concentration (µM)	Concentration from external standard method(µM)	Relative error (%)
Alanine	42.4	42.1	0.71
Serine	54.1	59.4	9.8
Phenylalanine	8.42	8.07	4.2
Threonine	22.0	23.3	5.9

Table 2.3. Comparison of absolute concentrations from standard addition strategy with concentrations from external standard method for alanine, serine, phenylalanine and threonine, and the relative differences of concentrations obtained using external standard method from absolute concentrations in percentage.

It is clear from Table 2.3, for alanine and phenylalanine, the concentrations obtained from calibration curve were very close to absolute values with relative errors being 0.71% and 4.2%, respectively. As for threonine, the concentration from external

standard method was about ~6% different from the value generated from standard addition approach. Serine presented the largest relative error of ~10% for comparison between concentrations from calibration curve with standard addition method.

Moreover, the extent of matrix effect on labeling can be determined by comparing UUMS samples at different concentrations that were labeled using the same protocol. As shown in Table 2.4, for alanine, serine, phenylalanine and threonine, the concentrations without dilution were calculated based on the values obtained from calibration curves (Table 2.2) times the dilution factors. And for different dilution factors, the relative differences between concentrations obtained using external standard method and standard addition were determined first (Table 2.4), then presented in both scatter plot (Figure 2.7) and box plot (Figure 2.8).

Alanine (A)

Dilution factor	Concentration without dilution (µM)	Relative error (%)
1	42.1	0.71
2	42.4	0
5	43.9	3.4
10	45.4	7.1

Serine (B)

Dilution factor	Concentration without dilution (µM)	Relative error (%)
1	59.4	9.8
2	60.4	11.6
5	57.5	6.5
10	58.8	8.9

Phenylalanine (C)

Dilution factor	Concentration without dilution (µM)	Relative error (%)
1	8.07	4.2
2	8.70	3.3
5	8.45	3.6
10	8.80	4.5

Threonine (D)

Dilution factor	Concentration without dilution (µM)	Relative error (%)
1	23.3	5.9
2	24	9.1
5	23.6	7.3
10	23.9	8.6

Table 2.4. Comparison of absolute concentrations from standard addition strategy with concentrations from external standard method for (A) alanine, (B) serine, (C) phenylalanine and (D) threonine in UUMS under 1-, 2-, 5-, and 10-fold dilution factors, and also the relative differences of concentrations obtained using external standard method from absolute concentrations in percentage.

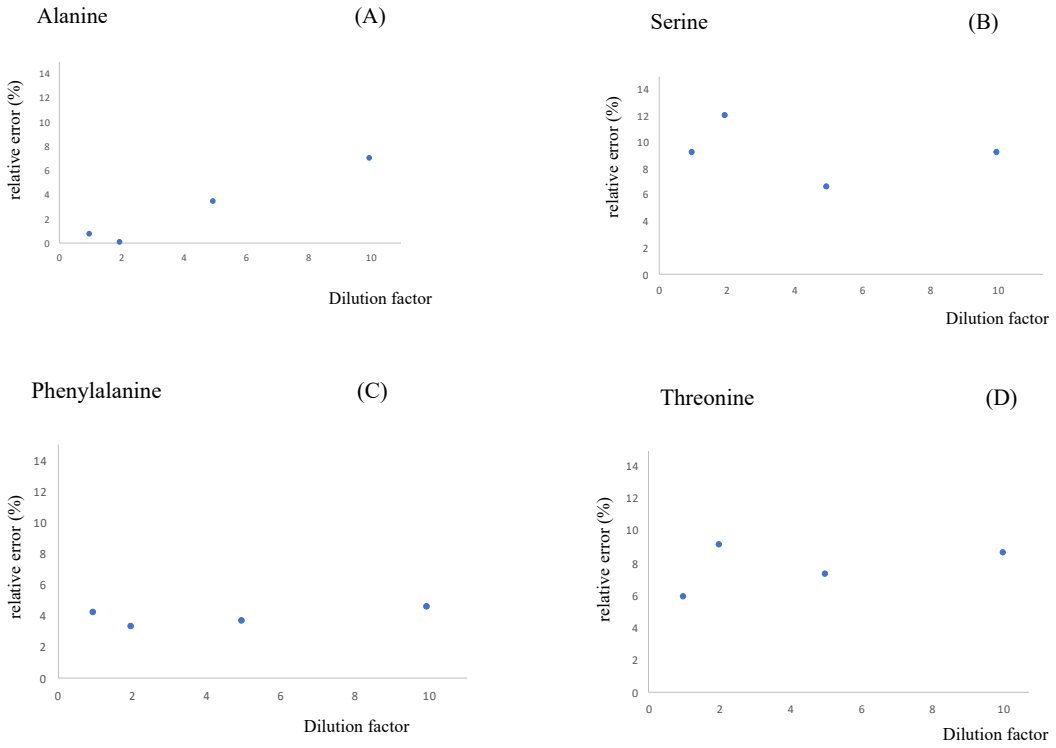


Figure 2.7. Scatter plots of the relative errors in percentage between absolute concentrations from standard addition strategy and concentrations from external standard method for (A) alanine, (B) serine, (C) phenylalanine and (D) threonine in UUMS under 1-, 2-, 5-, and 10-fold dilution factors.

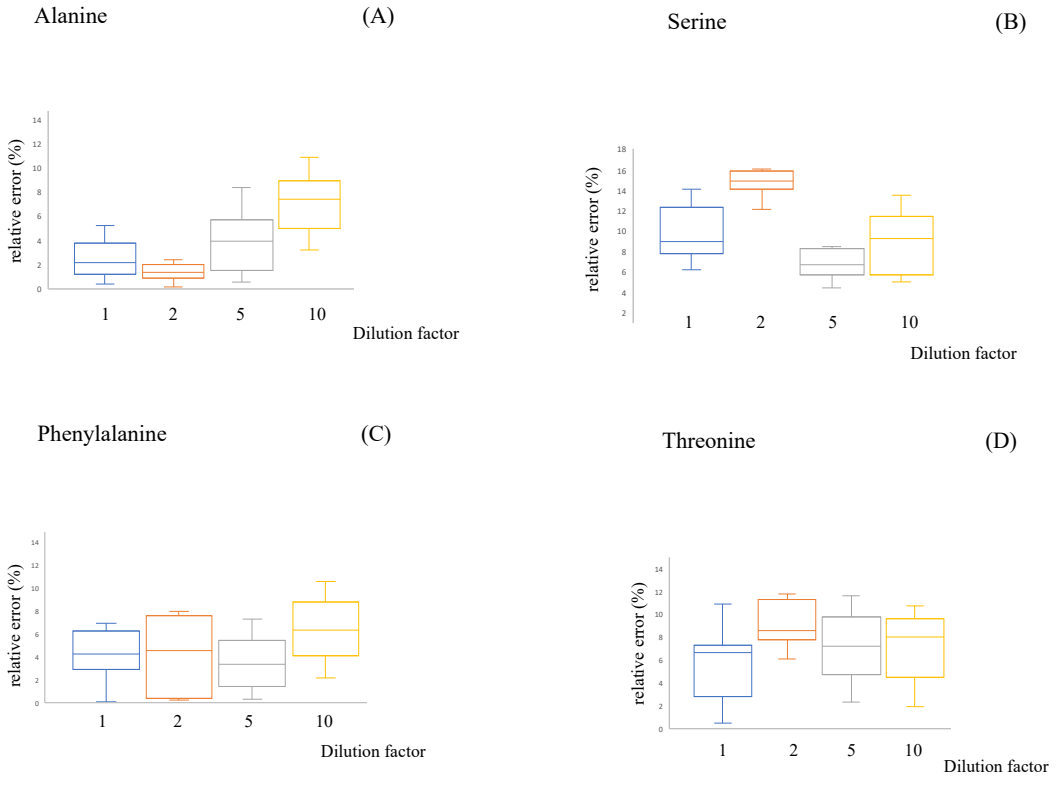


Figure 2.8. Box plots of the relative errors in percentage between absolute concentrations from standard addition strategy and concentrations from external standard method for (A) alanine, (B) serine, (C) phenylalanine and (D) threonine in UUMS under 1-, 2-, 5-, and 10-fold dilution factors.

As shown in Table 2.4, the relative errors between concentrations obtained from external standard method and absolute concentrations were determined in percentage and presented in the third column for each of the selected amino acids: alanine (Table 2.4A), serine (Table 2.4B), phenylalanine (Table 2.4C) and threonine (Table 2.4D). From Table 2.4, for four selected amino acids under 1-, 2-, 5- and 10-fold dilution factors, only serine in 2-fold diluted UUMS showed a 11.6% difference from its absolute concentration, other values were all within 10%. Furthermore, the relationships between relative errors and dilution factors for individual amino acids were presented in both scatter plot and box plot. As shown in Figure 2.7 and Figure 2.8, for four selected amino acids, there were no obvious tendency showing the relative errors changed proportionally with dilution effects. Besides, without performing any dilution, the relative errors between concentrations obtained from external standard method and absolute concentrations were all lower than 10%.

2.4 Conclusion

We have used a high-coverage and quantitative metabolome profiling technique based on differential chemical isotope labeling LC-MS to determine the absolute concentrations from standard addition method for four amino acids (alanine, serine, phenylalanine and threonine) in UUMS. The matrix effects were evaluated by comparing absolute concentrations of selected amino acids from standard addition strategy with concentrations from external standard method. Comparison of UUMS at different concentrations that were labeled using the same protocol was used to determine the extent of matrix effect on labeling. The results show that without any dilution, the relative errors between absolute concentrations and concentrations from external standard method for four amino acids in UUMS were all lower than 10%. Moreover, from the results of comparing UUMS at different concentrations, there was

no tendency indicating that the extent of matrix effect on labeling would be lowered by dilution for these four selected amino acids. For future work, it is also possible to evaluate other metabolites of interest to increase the overall matrix effects study coverage.

Chapter 3 Chemical Isotope Labeling LC-MS for Matrix Effects of Identified Metabolites in Universal Urine Metabolome Standard

3.1 Introduction

As an easily obtained biological fluid, urine is becoming an important source for biomarker study in disease.[73,74] Compared with other sample types such as serum and plasma, urine sample collections are less aggressive and easier to process.[75,76] What is more, it is reported that urine is sensitive to show the internal changes within the body and is to be found as a place for early biomarkers study.[77-79] It is expected that urine metabolomics would help us understand better dynamic changes in the body and identify potential disease biomarkers.[80,81] However, the matrix contained in urine sample could interfere the accuracy and precision when measuring the change of metabolites concentrations in biomarkers discovery.[82,83]

Different techniques have been introduced to lower the interferences caused by matrix, such as removing interferences during sample preparation[84-86] and improving chromatographic selectivity to reduce the coelution of metabolites of interest with matrix components.[87,88] When matrix phenomenon cannot be completely eliminated, calibration approaches are used, such as standard addition[89] and quantification with matrix-matched calibration curves.[90] One major disadvantage of calibration approaches is the extra time required for making the additions and measurements.[91] In order to avoid matrix effects, dilution is found as an easy and effective method, with less interfering compounds existing during sample treatment and loading into chromatographic system.[92-94] With reduced number of analytes being injected into system, a method for quantitative metabolomics profiling with high coverage is needed.[95-97]

The development of new analytical techniques with high coverage and instrument sensitivity makes sample dilution a common approach for analysis.[98-101] Our group has reported a differential chemical isotope labeling (CIL) LC-MS method for quantitative metabolomics profiling with high coverage. Dansylation labeling

technique can be used for profiling amine/phenol submetabolome.[37] What is more, this technique improves MS sensitivity and quantification accuracy, enables us to generate a comprehensive profile of the submetabolome. In this work, we applied dansylation LC-MS for evaluating the effectiveness of different dilutions in metabolites analysis in urine matrix.

3.2 Experimental

3.2.1 Chemical and Reagents

All the chemicals and reagents, unless otherwise stated, were from Sigma-Aldrich Canada (Markham, ON, Canada). ^{13}C-dansyl chloride was synthesized in our lab with the procedures published previously[37] and is available from MCID.chem.ualberta.ca. LC-MS grade water, acetonitrile (ACN), and methanol (MeOH) were from Thermo Fisher Scientific (Edmonton, AB, Canada).

3.2.2 Overall Workflow

Figure 3.1 shows the overall workflow of this study. In this work, universal urine metabolome standard (UUMS) sample was firstly diluted into three different folds: 5-, 10- and 20-fold. Each diluted UUMS sample was labeled by ^{12}C-dansyl chloride and ^{13}C-isotope labeled 20-fold diluted UUMS served as internal standard. Then ^{12}C-labled sample and ^{13}C-labled internal standard were mixed by equal volume. After being mixed and injected into LC-MS for analysis, the peak pairs detected in MS were extracted by IsoMS. The individual peak pairs from different LC-MS runs were aligned together based on accurate mass and retention time to produce a metabolite peak pair table. The relative concentration for each metabolite was obtained by IsoMS. And identifications of metabolites were performed by searching MycomoundID library (www.mycompoundid.org). Finally, comparison of the experimental concentrations for every identified metabolite under different dilution factors with theoretical values were performed to determine the effect of dilution and matrix effects on identified metabolites. The detailed experimental conditions in the workflow are described below.

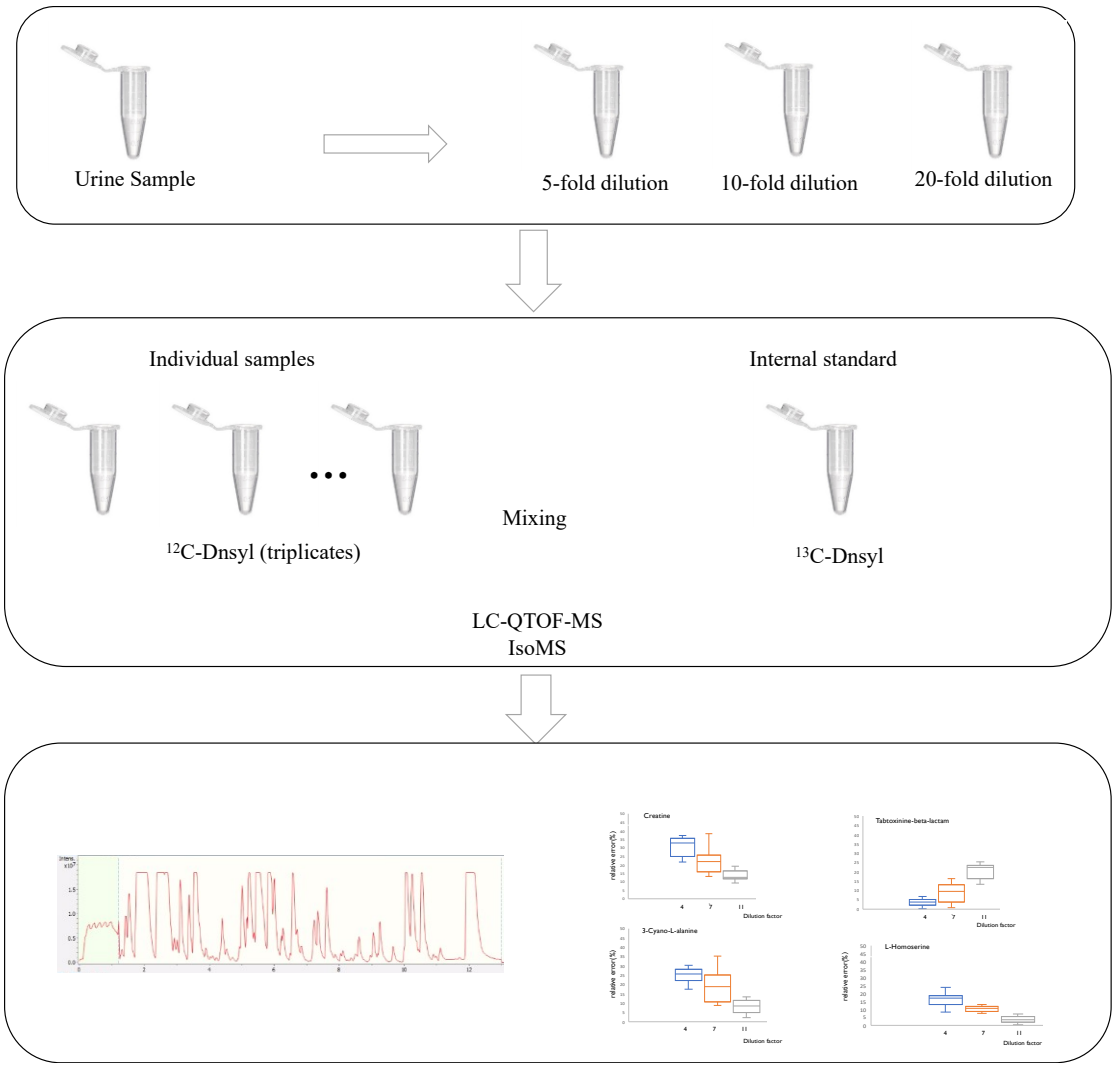


Figure 3.1. Workflow for sample preparation and dansylation isotope labeling LC-MS.

3.2.3 Matrix Dilution Folds

In order to study the matrix effect behavior when the matrix amount is reduced, three different dilutions of the UUMS were prepared, with dilution factors of 5, 10 and 20. Half of the matrix was introduced from ^{12}C-labled individual samples and the other half came from ^{13}C-labled internal standard. ^{12}C-labled individual samples were diluted by 5-, 10-, and 20-fold, respectively, and ^{13}C-labled internal standard was under a 20-fold dilution. As a result, the matrix was diluted by 4-, 7- and 11-fold, respectively.

3.2.4 Dansylation Labeling

The frozen urine sample was re-dissolved to 1000μL with water to reach its original concentration. Then 12.5 μL of 250 mM sodium carbonate/sodium bicarbonate buffer was added to 25 μL of urine sample. The solution was mixed with 37.5 μL of freshly prepared ^{12}C-DnsCl solution (18 mg/mL for light labeling, individual samples) or ^{13}C-DnsCl solution (18 mg/mL for heavy labeling, internal standard samples). After incubation for 45 min at 40 °C, 7.5 μL of 250 mM sodium hydroxide solution was added to the reaction mixture, followed by incubation at 40 °C for another 10 min. Finally, 30 μL of formic acid (425 mM) in 1:1 ACN/H_2O was added.

3.2.5 LC-MS

Each ^{12}C-labled individual sample was mixed with ^{13}C-labled internal standard by equal volumes. LC-MS was done using a Themo Scientific Dionex Ultimate 3000 UHPLC System (Sunnyvale, CA) linked to a Bruker Maxis II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker, Billerica, MA). The LC column was an Agilent reversed phase Eclipse Plus C18 Column (2.1 mm x 10 cm, 1.8 μm particle size). LC mobile phase A was 0.1% (v/v) formic acid in water, and mobile phase B was 0.1% (v/v) formic acid in ACN. The gradient elution profile was as follow: t = 0 min, 25% B; t = 10 min, 99% B; t = 13 min, 99% B; t = 13.1 min, 25% B, t = 16 min, 25% B. Flow rate was 400 μL/min. All mass spectra were collected in the positive ion mode. The MS conditions for MS were as follows: dry gas flow, 8 L/min; dry temperature, 230 °C; capillary voltage, 4500 V; nebulizer, 1.0 bar; end plate offset, 500V; spectra rate, 1.0 Hz.

3.2.6 Data Processing, Metabolite Identification and Statistical Analysis

After LC-Q-TOF-MS analysis, all the spectra were first converted to .csv files by Bruker Daltonis Data Analysis 4.4 software. The peak pairs were extracted from .csv files by IsoMS.[69] Data from multiple runs were aligned by retention time and accurate mass. After the alignment, Zerofill program[70] was applied to recover the high-confidence peak pair ratios lost during the previous data processing steps. Based on the accurate mass and retention time, positive metabolite identification was performed

using dansyl standard library search, which contains 275 unique dansylated amines/phenols[71] (www.mycompoundid.org). Putative identification or match was performed by searching accurate mass against MyCompoundID library, which contains 8,021 known human metabolites and 375,809 predicted metabolites[72] (www.mycompoundid.org).

3.3 Result and Discussion

3.3.1 Metabolites Identification

In this work, UUMS sample was diluted under three different dilution factors: 5-, 10-, and 20-fold, and then labeled with ^{12}C-dansyl chloride. Then 20-fold diluted UUMS was also labeled by ^{13}C-dansyl chloride which served as the internal standard. For each concentration of ^{12}C-labled individual samples, three aliquots were prepared. In the LC-MS analysis, the ^{13}C-dansylation labeled internal standard was mixed with ^{12}C-dansylation labeled individual samples by same volume. Then the same volume of mixtures prepared from all individual samples was injected into LC-MS. After LC-MS acquisition and data processing, a total of 202 peak pairs were detected. As can be seen from Table 3.1 there were totally 77 metabolites being identified. Using CIL library search, 24 metabolites were positively identified according to accurate mass and retention time matched against a labeled standard compound library. For the remaining metabolites, using Li library searched with accurate mass and predicted retention time matched against a linked identity library, 53 metabolites were putatively identified with high confidence. As a result, we achieved a good submetabolome coverage and about 40% of metabolites can be matched with different confidence levels for various diluted UUMS samples.

RT	Compound ID	Metabolite
72	HMDB00168	Asparagine
78.1	AFA00126000	LL-2,6-Diaminoheptanedioate/meso-2,6-Diaminoheptanedioate
81.4	AXA00234000	Symmetric dimethylarginine
83.7	AAA00616000	Creatine

87	AXA00042000	Biopterin
90	HMDB00641	Glutamine
109.1	HMDB00187	Serine
122.8	AXA00795000	Beta-Guanidinopropionic acid
122.8	HMDB00725	4-Hydroxyproline
127.3	AFA00118000	Tabtoxinine-beta-lactam
133.3	AAA00624000	Threonine
144.5	AAA00727000	Aminoadipic acid
144.8	AAA00630000	L-Homoserine
152.4	AAA01327000	Ethanolamine
153.1	AFA00145000	3-Cyano-L-alanine
159.3	HMDB00123	Glycine
161.5	AAA00770000	4-Guanidinobutanal
172.0	HMDB00161	Alanine
172.7	AFA00145000	3-Cyano-L-alanine
175.3	AFA00143000	Gamma-Amino-gamma-cyanobutanoate
187.0	HMDB00271	Sarcosine
192.0	HMDB00112	Gamma-Aminobutyric acid
213.3	AXA03615000	N-Ethylglycine
219.7	AXA00148000	7-Methylguanine
226.5	HMDB01906	2-Aminoisobutyric acid
230.8	AAA00437000	N-Acetylornithine
233.0	AAA00459000	Hypoxanthine - Isomer
233.7	AFA00216000	Methylamine
242.5	HMDB00300	Uracil
261.3	HMDB00883	Valine
263.0	AAA00589000	(Z)-3-Peroxyaminoacrylate
266.6	AFA00098000	4-Hydroxyphenylglyoxylate
272.0	APA00823000	Homomethionine
273.6	AAA00640000	Tryptophan
277.0	HMDB00159	Phenylalanine
293.6	HMDB00118	Homovanillic acid
300.9	AAA00887000	2-Carboxy-2,3-dihydro-5,6-dihydroxyindole
300.9	HMDB00840	Salicyluric acid
301.0	HMDB00112_2	Gamma-Aminobutyric acid - H_2O
304.7	AAA00840000	Urocanic acid
306.6	AAA00691000	Leucine
309.7	AAA00287000	3-(4-Hydroxyphenyl) lactate
311.4	AAA01692000	4-Aminobenzoate
313.9	APA00767000	5-Aminopentanal
326.4	AAA00674000	Cystine
327.8	AAA01010000	5-Hydroxyindoleacetic acid
329.9	HMDB03911_2	3-Aminoisobutanoic acid - H_2O

354.0	AAA00902000	3-Methoxy-4-hydroxyphenylglycolaldehyde
360.6	AAA00425000	Ornithine
376.5	AAA00739000	Lysine
377.3	AXA00705000	(6R)-6-(L-Erythro-1,2-Dihydroxypropyl)-5,6,7,8-tetrahydro-4a-hydroxypterin - 2 tags
378.0	HMDB00500	4-Hydroxybenzoic acid
389.3	AAA00896000	3-Methoxy-4-hydroxyphenylacetaldehyde
390.0	HMDB00669	4-Hydroxy-3-methylbenzoic acid
390.8	AXA00322000	N8-Acetylspermidine - 2 tags
393.1	AAA00864000	4-Hydroxyphenylacetic acid
394.9	AAA00832000	Histidine
402.3	APA00873000	S-(Hydroxyphenylacetothiohydroximoyl)-L-cysteine - 2 tags
402.3	AAA00635000	D-Lombricine - 2 tags
404.8	AXA03600000	Diethylamine
412.2	AAA01719000	7,8-Dihydroxanthopterin - 2 tags
431.0	AAA00956000	4-Hydroxy-3-methoxy-benzaldehyde
433.6	HMDB0000750_2	3-Hydroxymandelic acid - COOH
456.0	ABA00515000	1,2-Dihydroxynaphthalene-6-sulfonate - 2 tags
458.7	HMDB0006050	Tyrosine
458.8	HMDB00819	Normetanephrine
468.7	AAA00873000	Phenol
502.6	HMDB01858	p-Cresol
516.0	HMDB00292	Xanthine
518.7	AAA00916000	3,4-Dihydroxyphenylpropanoate - 2 tags
518.7	AAA00929000	3-(2,3-Dihydroxyphenyl) propanoate - 2 tags
526.0	AAA00868000	2,5-Dihydroxybenzoate - 2 tags
536.9	AAA00871000	Tyramine
554.5	ABA00156000	Resorcinol - 2 tags
554.9	HMDB04811	2,4-Dichlorophenol
578.8	AXA02817000	4-tert-Butylphenol
603.4	APA00413000	3,4-Dihydroxystyrene - 2 tags

Table 3.1. Positive and high confidence putative identification results in UUMS. Order from lowest to highest retention time.

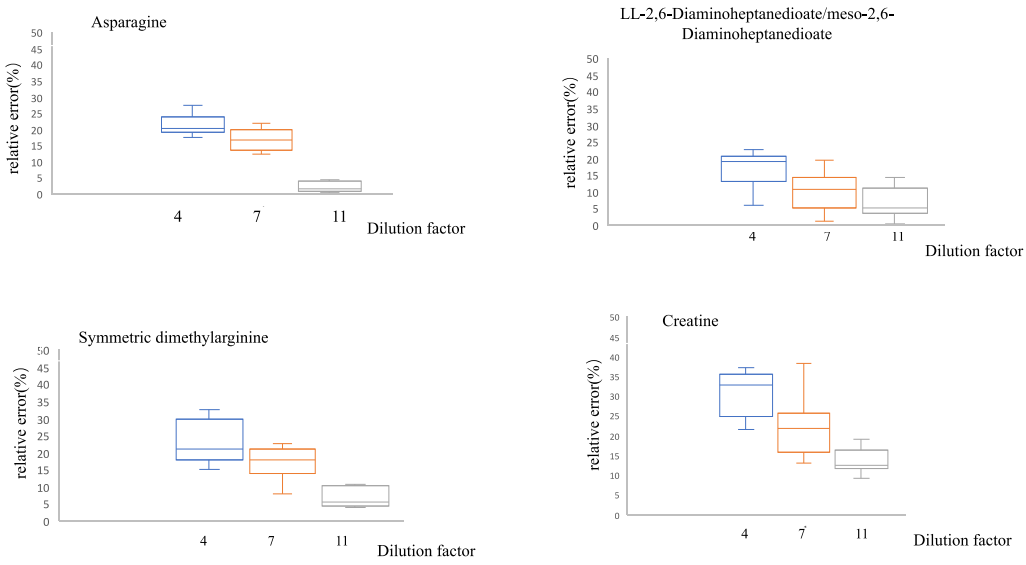
3.3.2 Comparison Between Experimental and Theoretical Values

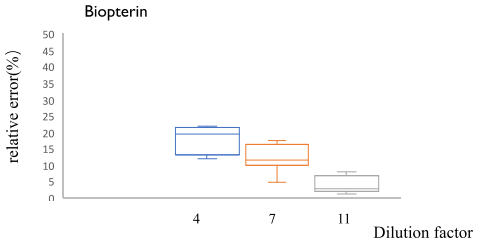
After metabolites being identified, for each individual metabolite, the peak pair ratio of ^{12}C and ^{13}C between experimental and theoretical values were compared in order to test matrix effect behavior with reduction of the matrix amount. Individual samples were diluted by 5-, 10-, and 20-fold, respectively and labeled with ^{12}C-dansyl

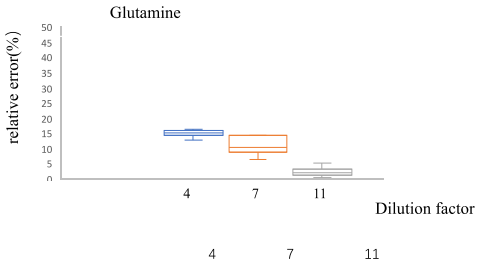
chloride, while ^{13}C-labled internal standard was under 20 dilution factors. In this LC-MS analysis, internal standard was mixed with individual samples by same volume, as a result, the theoretical peak pair ratio between ^{12}C and ^{13}C for 5-, 10-, and 20-fold diluted UUMS were 4, 2 and 1, respectively. As for the change in concentrations of matrix, since half of the matrix was introduced from ^{12}C-labled individual samples and the other half came from ^{13}C-labled internal standard, the matrix was eventually diluted by 4-, 7- and 11-fold, respectively. An absolute percentage of the difference between experimental and theoretical values within 30% was considered as no matrix effect, since variation would be close to the repeatability values.[95]

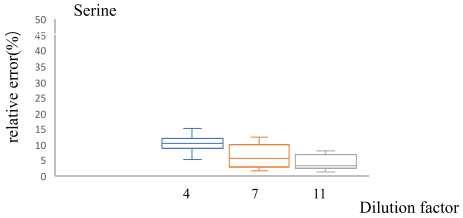
In UUMS sample, there were totally 77 metabolites being identified, the relationships between relative errors and dilution factors for identified metabolites were presented in box plot (Figure 3.2). The standard deviation for each concentration measurement was the result of combined variations in biological triplicate and experimental triplicate (n=9).

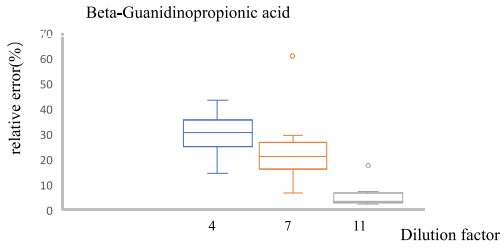
Table 3.2 summarizes the identifications of 64 metabolites in UUMS with absolute percentage differences under 30% between experimental and theoretical values when matrix was diluted by 4-fold. Table 3.3 presents the remaining 13 metabolites which had more than 30% difference between experimental and theoretical values after being diluted by 4-fold.

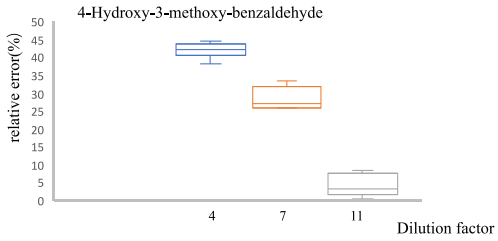


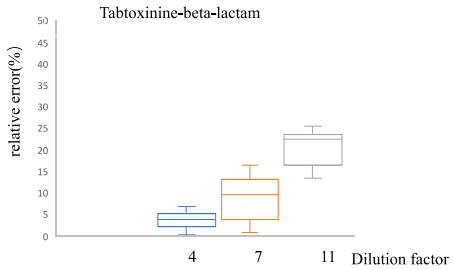

Biopterin
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

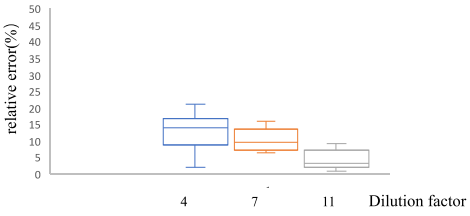

Glutamine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

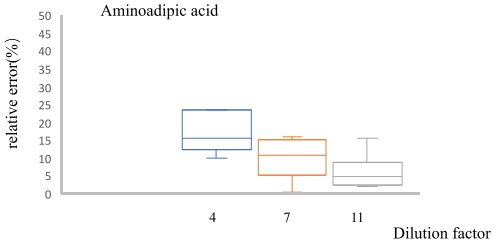

Serine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

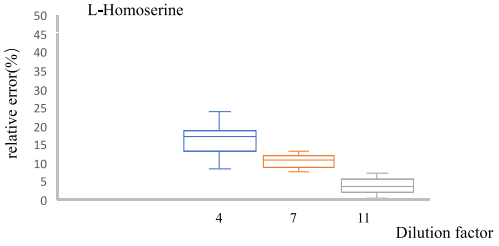

Beta-Guanidinopropionic acid
relative error(%)
70
60
50
40
30
20
10
0
4
7
11
Dilution factor

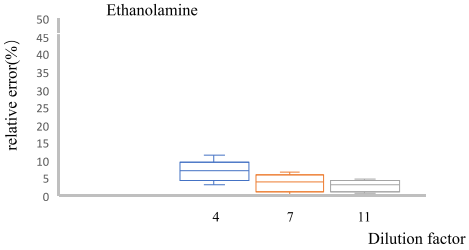

4-Hydroxy-3-methoxy-benzaldehyde
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

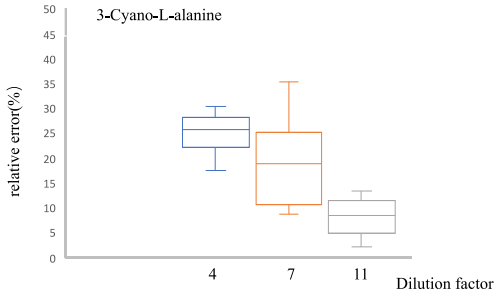

Tabtoxinine-beta-lactam
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

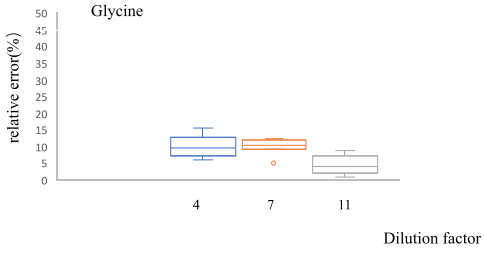

relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

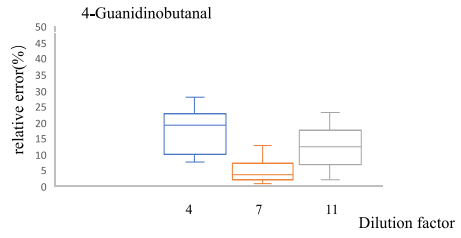

Aminoadipic acid
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

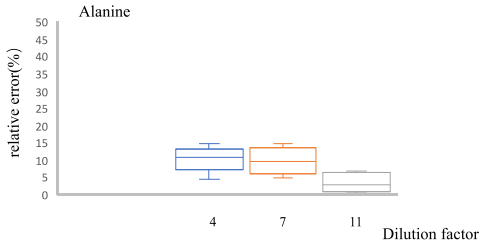

L-Homoserine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

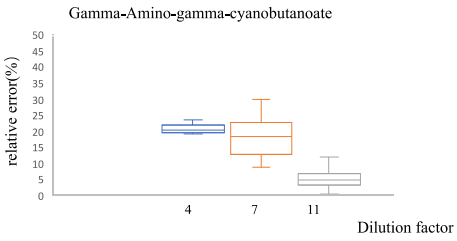

Ethanolamine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

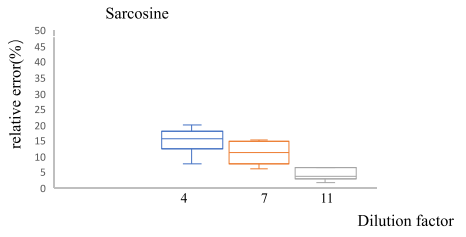

3-Cyano-L-alanine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

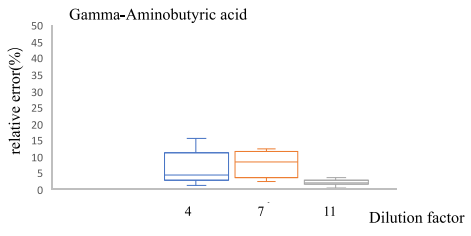

Glycine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

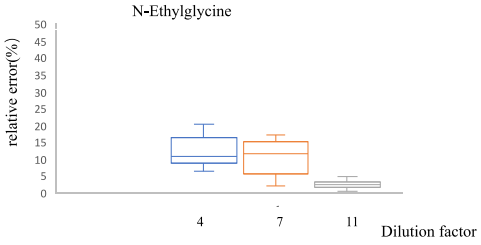

4-Guanidinobutanal
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

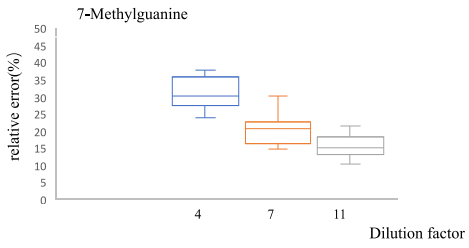

Alanine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

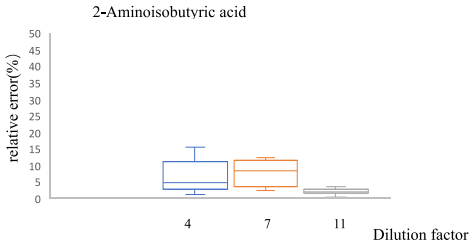

Gamma-Amino-gamma-cyanobutanoate
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

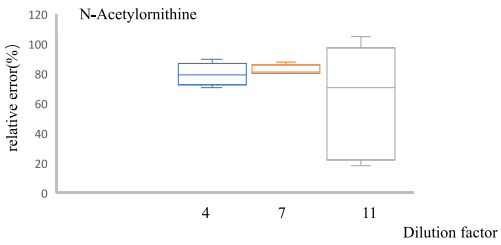

Sarcosine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

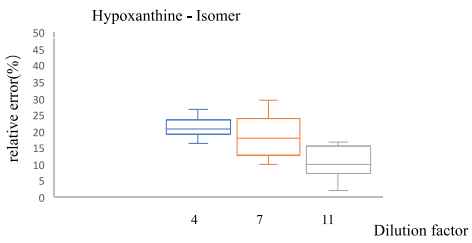

Gamma-Aminobutyric acid
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

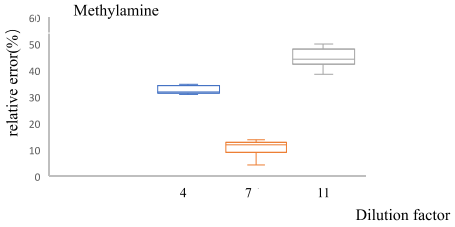

N-Ethylglycine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

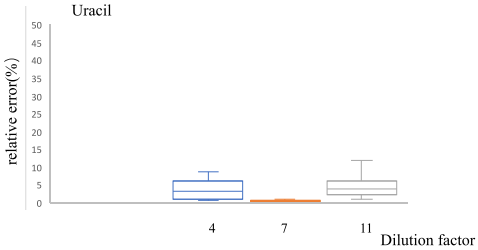

7-Methylguanine
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

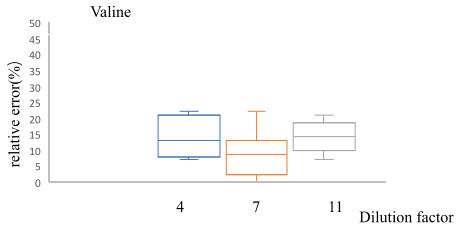

2-Aminoisobutyric acid
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

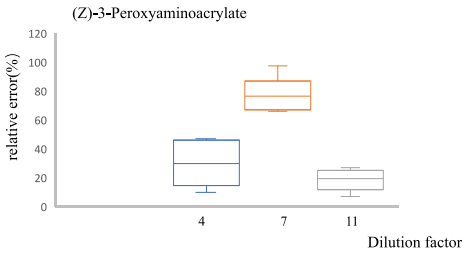

N-Acetylornithine
relative error(%)
120
100
80
60
40
20
0
4
7
11
Dilution factor

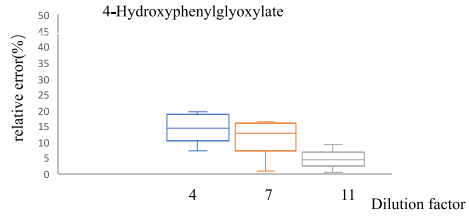

Hypoxanthine - Isomer
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

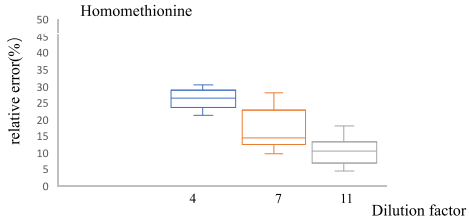

Methylamine
relative error(%)
60
50
40
30
20
10
0
4
7
11
Dilution factor

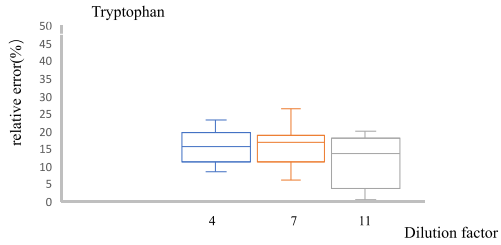

Uracil
relative error(%)
50
45
40
35
30
25
20
15
10
5
0
4
7
11
Dilution factor

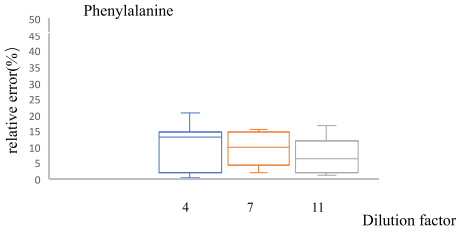

Valine
relative error(%)
Dilution factor

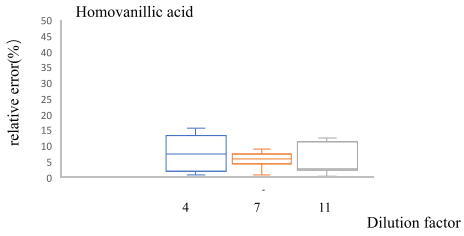

(Z)-3-Peroxyaminoacrylate
relative error(%)
Dilution factor

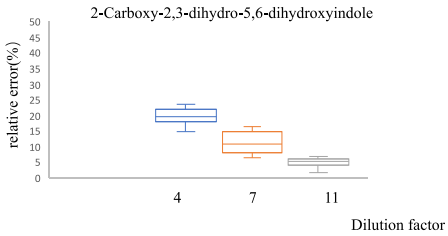

4-Hydroxyphenylglyoxylate
relative error(%)
Dilution factor

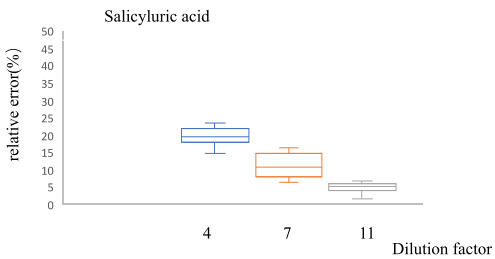

Homomethionine
relative error(%)
Dilution factor

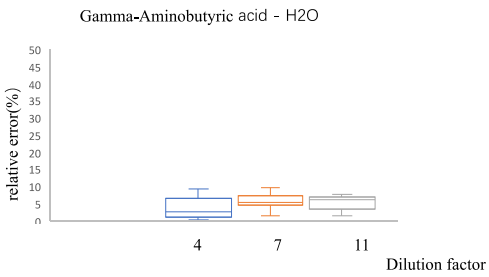

Tryptophan
relative error(%)
Dilution factor

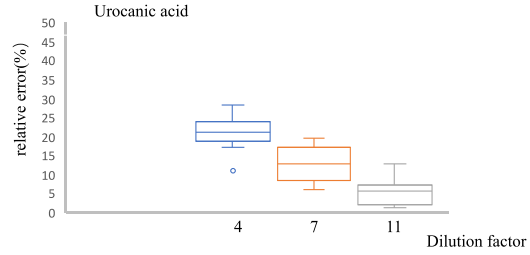

Phenylalanine
relative error(%)
Dilution factor

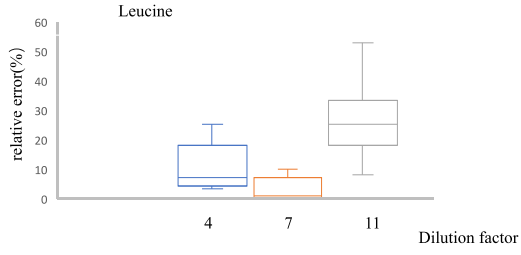

Homovanillic acid
relative error(%)
Dilution factor

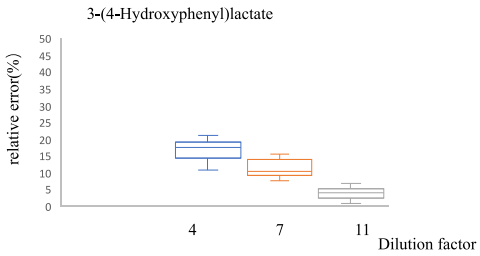

2-Carboxy-2,3-dihydro-5,6-dihydroxyindole
relative error(%)
Dilution factor

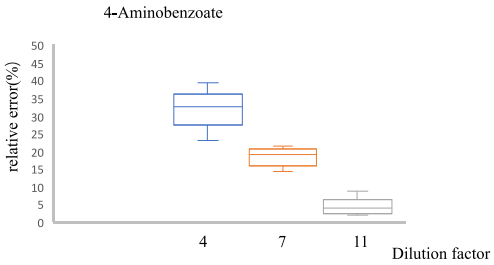

Salicyluric acid
relative error(%)
Dilution factor

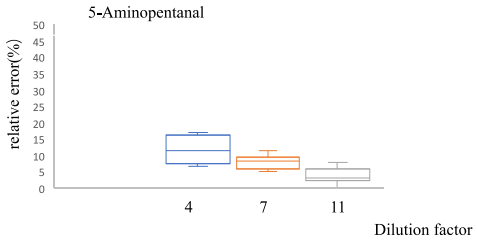

Gamma-Aminobutyric acid - H2O
relative error(%)
Dilution factor

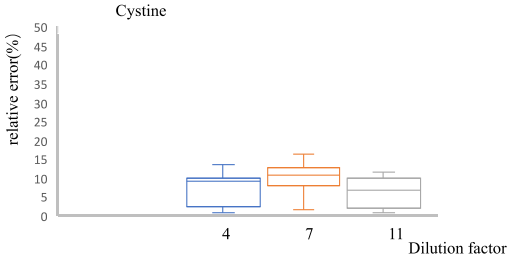

Urocanic acid
relative error(%)
Dilution factor

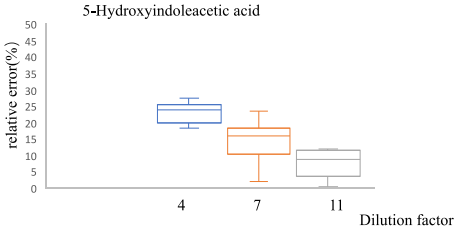

Leucine
relative error(%)
Dilution factor

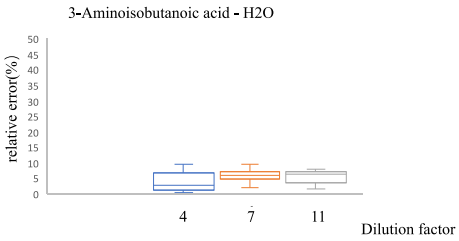

3-(4-Hydroxyphenyl)lactate
relative error(%)
Dilution factor

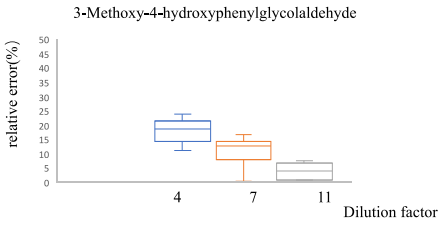

4-Aminobenzoate
relative error(%)
Dilution factor

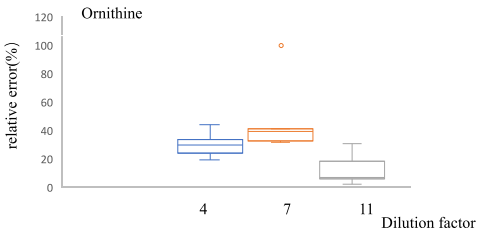

5-Aminopentanal
relative error(%)
Dilution factor

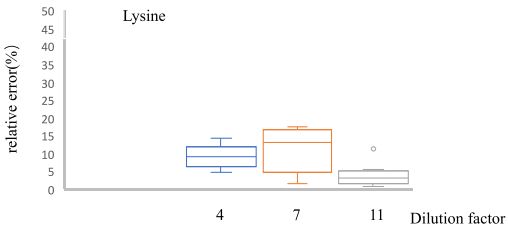

Cystine
relative error(%)
Dilution factor

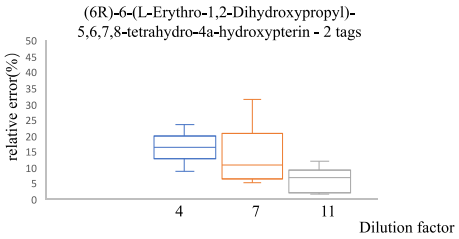

5-Hydroxyindoleacetic acid
relative error(%)
Dilution factor

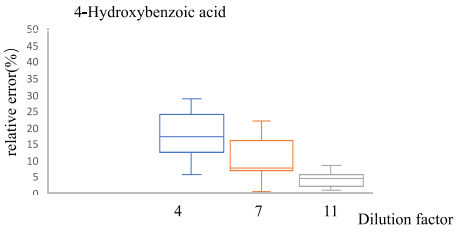

3-Aminoisobutanoic acid - H2O
relative error(%)
Dilution factor

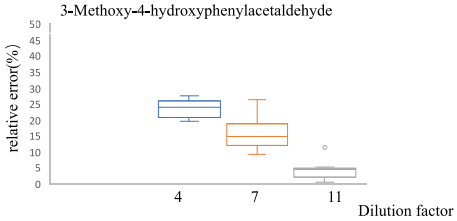

3-Methoxy-4-hydroxyphenylglycolaldehyde
relative error(%)
Dilution factor

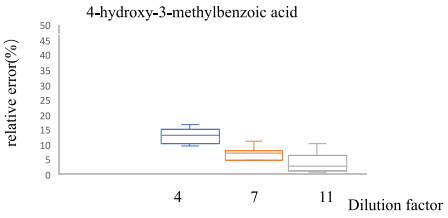

Ornithine
relative error(%)
Dilution factor

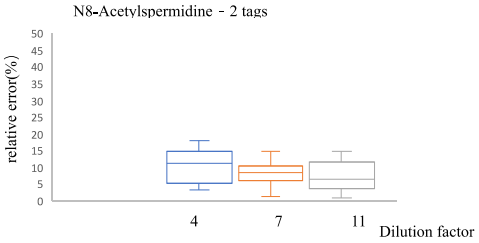

Lysine
relative error(%)
Dilution factor

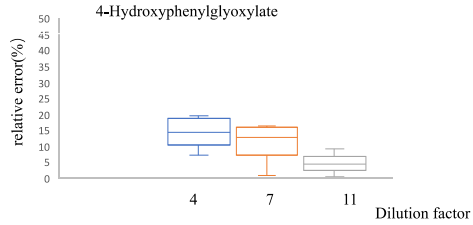

(6R)-6-(L-Erythro-1,2-Dihydroxypropyl)-
5,6,7,8-tetrahydro-4a-hydroxypterin - 2 tags
relative error(%)
Dilution factor

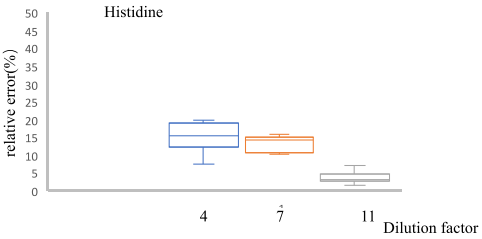

4-Hydroxybenzoic acid
relative error(%)
Dilution factor

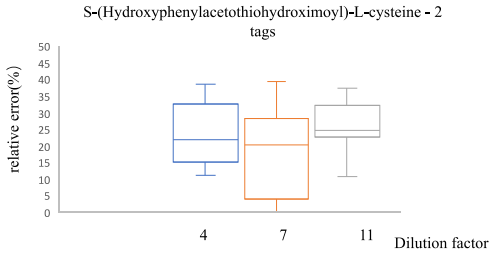

3-Methoxy-4-hydroxyphenylacetaldehyde
relative error(%)
Dilution factor

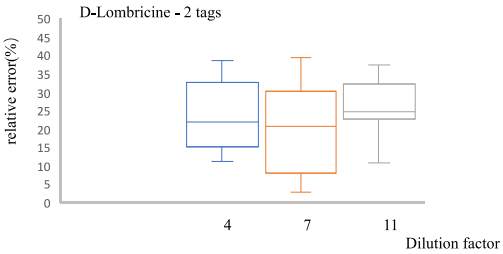

4-hydroxy-3-methylbenzoic acid
relative error(%)
Dilution factor

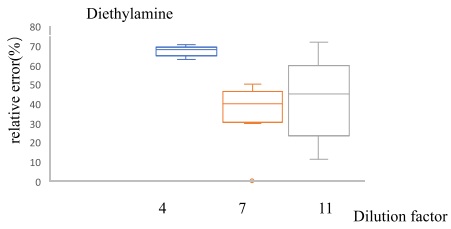

N8-Acetylspermidine - 2 tags
relative error(%)
Dilution factor

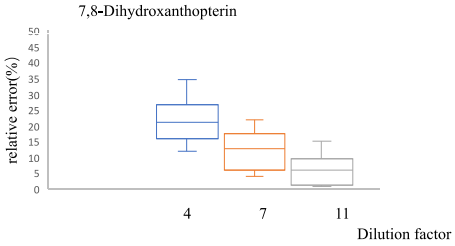

4-Hydroxyphenylglyoxylate
relative error(%)
Dilution factor

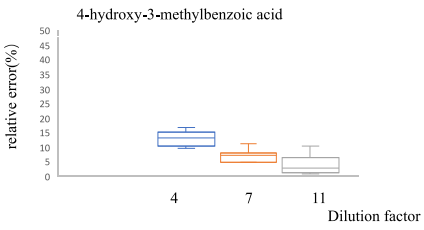

Histidine
relative error(%)
Dilution factor

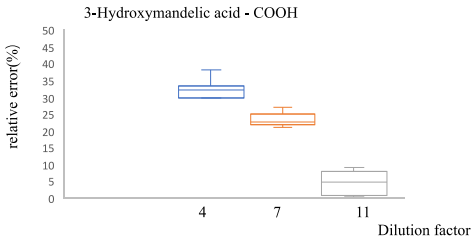

S-(Hydroxyphenylacetothiohydroximoyl)-L-cysteine - 2 tags
relative error(%)
Dilution factor

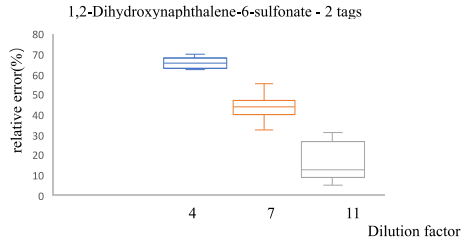

D-Lombricine - 2 tags
relative error(%)
Dilution factor

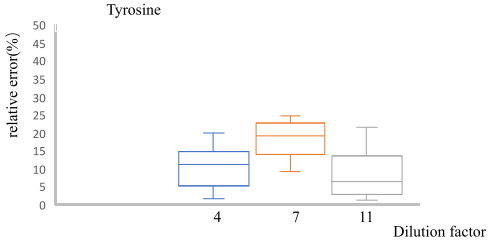

Diethylamine
relative error(%)
Dilution factor

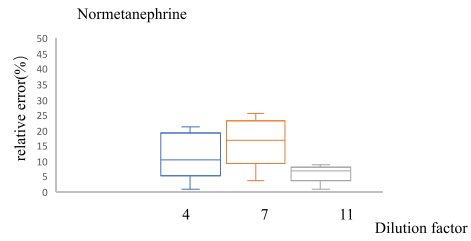

7,8-Dihydroxanthopterin
relative error(%)
Dilution factor

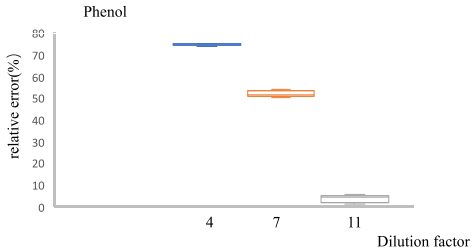

4-hydroxy-3-methylbenzoic acid
relative error(%)
Dilution factor

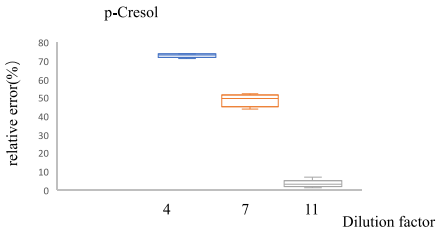

3-Hydroxymandelic acid - COOH
relative error(%)
Dilution factor

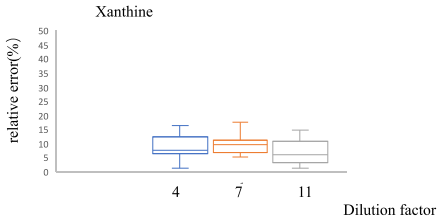

1,2-Dihydroxynaphthalene-6-sulfonate - 2 tags
relative error(%)
Dilution factor

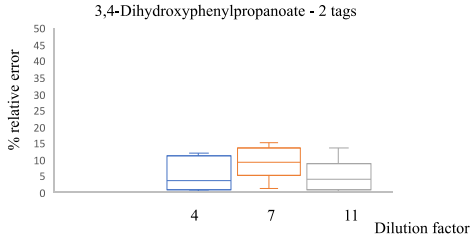

Tyrosine
relative error(%)
Dilution factor

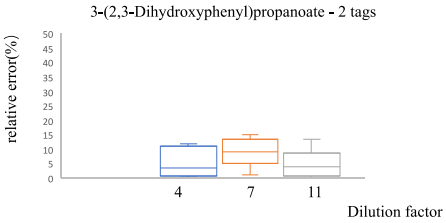

Normetanephrine
relative error(%)
Dilution factor

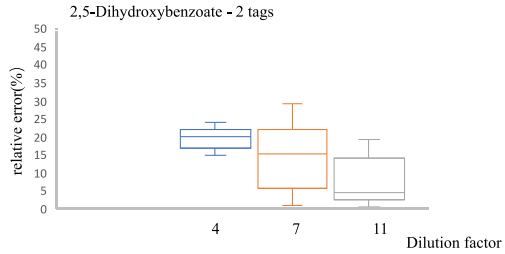

Phenol
relative error(%)
Dilution factor

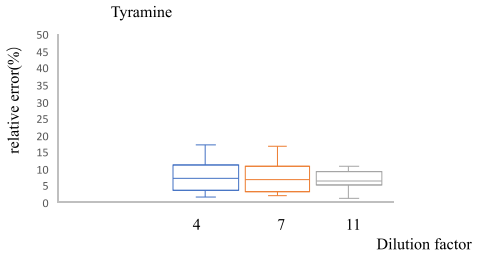

p-Cresol
relative error(%)
Dilution factor

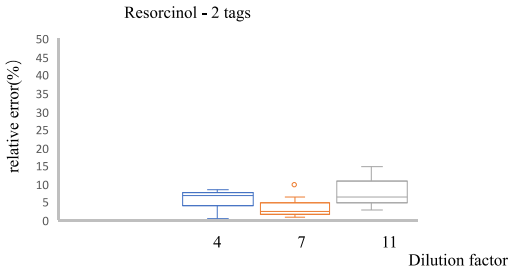

Xanthine
relative error(%)
Dilution factor

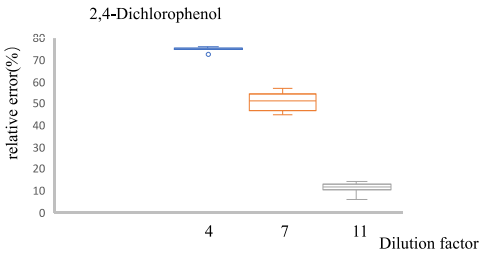

3,4-Dihydroxyphenylpropanoate - 2 tags
% relative error
Dilution factor

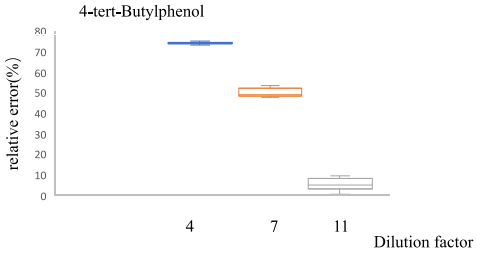

3-(2,3-Dihydroxyphenyl)propanoate - 2 tags
relative error(%)
Dilution factor


2,5-Dihydroxybenzoate - 2 tags
relative error(%)
Dilution factor


Tyramine
relative error(%)
Dilution factor


Resorcinol - 2 tags
relative error(%)
Dilution factor


2,4-Dichlorophenol
relative error(%)
Dilution factor


4-tert-Butylphenol
relative error(%)
Dilution factor

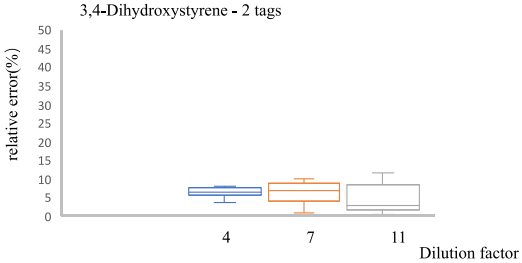

3,4-Dihydroxystyrene - 2 tags
relative error(%)
Dilution factor

Figure 3.2. Box plots of absolute percentage differences between experimental and theoretical values for 76 metabolites in UUMS when matrix was under 4-, 7- and 11-fold dilution factors. Order from lowest to highest retention time.

Metabolite	Relative error (%)
Asparagine	21
LL-2,6-Diaminoheptanedioate/meso-2,6-Diaminoheptanedioate	17
Symmetric dimethylarginine	23
Biopterin	21
Glutamine	15
Serine	10
Beta-Guanidinopropionic acid	30
4-Hydroxyproline	30
Tabtoxinine-beta-lactam	3.6
Threonine	13
Aminoadipic acid	17
L-Homoserine	16
Ethanolamine	6.8
3-Cyano-L-alanine	25
Glycine	9.9
4-Guanidinobutanal	18
Alanine	10
3-Cyano-L-alanine	23
Gamma-Amino-gamma-cyanobutanoate	21
Sarcosine	15
Gamma-Aminobutyric acid	6
N-Ethylglycine	12
2-Aminoisobutyric acid	6.1
Hypoxanthine – Isomer	21
Uracil	0.4
Valine	14
(Z)-3-Peroxyaminoacrylate	29
4-Hydroxyphenylglyoxylate	14
Homomethionine	26
Tryptophan	15
Phenylalanine	10
Homovanillic acid	7.5
2-Carboxy-2,3-dihydro-5,6 dihydroxyindole	19
Salicyluric acid	19
Gamma-Aminobutyric acid - H_2O	2
Urocanic acid	21
Leucine	8.1

3-(4-Hydroxyphenyl) lactate	17
5-Aminopentanal	12
Cystine	5.8
5-Hydroxyindoleacetic acid	23
3-Aminoisobutanoic acid - H_2O	1.9
3-Methoxy-4-hydroxyphenylglycolaldehyde	18
Ornithine	29
Lysine	5.5
(6R)-6-(L-Erythro-1,2-Dihydroxypropyl)-5,6,7,8-tetrahydro-4a-hydroxypterin - 2 tags	16
4-Hydroxybenzoic acid	17
3-Methoxy-4-hydroxyphenylacetaldehyde	23
4-hydroxy-3-methylbenzoic acid	13
N8-Acetylspermidine - 2 tags	9.1
4-hydroxyphenylacetic acid	6.7
Histidine	15
D-Lombricine - 2 tags	24
S-(Hydroxyphenylacetothiohydroximoyl)-L-cysteine - 2 tags	24
7,8-Dihydroxanthopterin	22
Tyrosine	9.1
Normetanephrine	11
Xanthine	8.6
3,4-Dihydroxyphenylpropanoate - 2 tags	3
3-(2,3-Dihydroxyphenyl) propanoate - 2 tags	3
2,5-Dihydroxybenzoate - 2 tags	19
Tyramine	4.7
Resorcinol - 2 tags	1.6
3,4-Dihydroxystyrene - 2 tags	2

Table 3.2. The identifications of metabolites in UUMS with absolute percentage difference between experimental and theoretical values under 30% when matrix was diluted by 4-fold. Order from lowest to highest retention time.

Metabolite	**Relative error (%)**
Creatine	31
7-Methylguanine	31
N-Acetylornithine	79
Methylamine	33
4-Aminobenzoate	32
Diethylamine	67
4-Hydroxy-3-methoxy-benzaldehyde	42
3-Hydroxymandelic acid - COOH	32

1,2-Dihydroxynaphthalene-6-sulfonate - 2 tags	66
Phenol	75
p-Cresol	73
2,4-Dichlorophenol	75
4-tert-Butylphenol	74

Table 3.3. The identifications of metabolites in UUMS with absolute percentage difference between experimental and theoretical values above 30% when matrix was diluted by 4-fold. Order from lowest to highest retention time.

As summarized in Figure 3.3, in UUMS sample, more than 83% of the metabolites showed less than 30% difference from theoretical values, indicating that there was no large difference between experimental and theoretical values with matrix being diluted by 4 factors. And more than half of the remaining compounds (7%) showed that the matrix effects were improved after matrix being diluted by 7-fold, and half of the remaining compounds (7%) needed to be diluted by 11-fold. After being diluted by 11-fold, only 3% of the metabolites did not present any improvement which may be caused by their low detectable intensities.

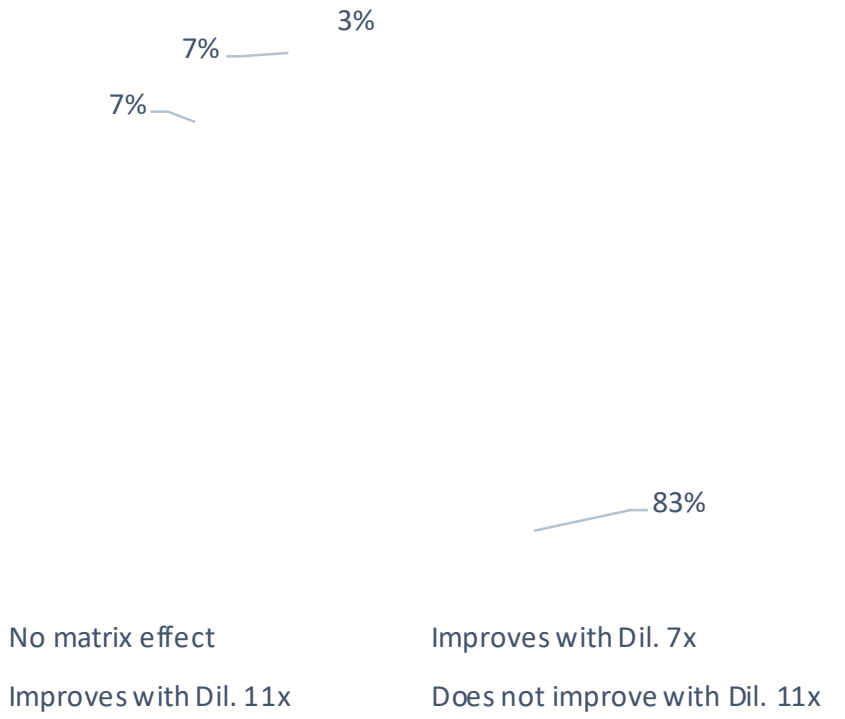


Figure 3.3. Distribution of identified metabolites in UUMS matrix studied according to how matrix effects were affected with reduction of matrix amount.

For the rest of unidentified metabolites in UUMS sample, as presented in Figure 3.4, about half (49%) of the metabolites were free of matrix effects when matrix was diluted by 4 factors, 13% needed dilution up to a dilution factor of 7. A relatively large percentage of compounds (29%) showed matrix effects improvement after matrix being diluted by 11-fold, and 9% of the metabolites did not improve, even with a dilution factor of 11.

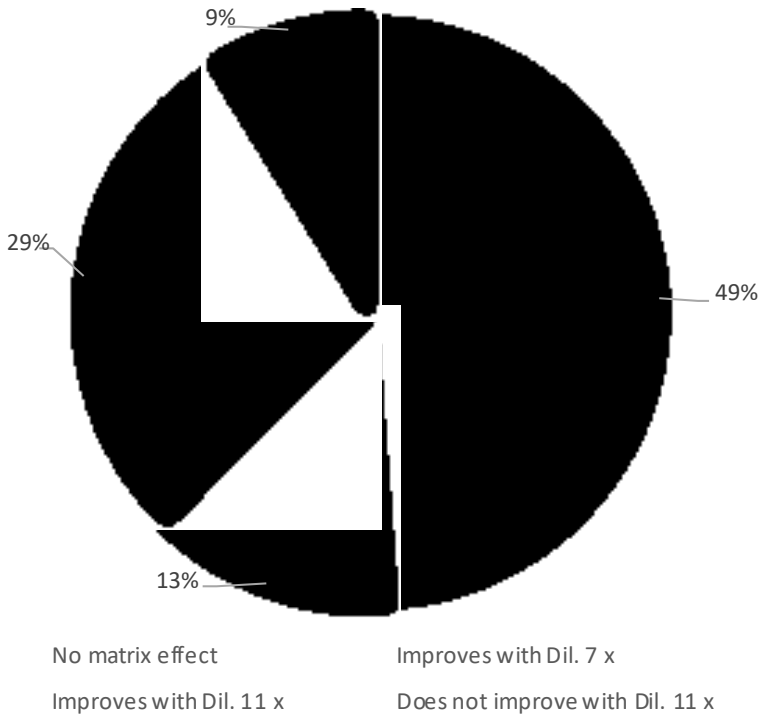


Figure 3.4. Distribution of unidentified metabolites in UUMS matrix studied according to how matrix effects were affected with reduction of matrix amount.

3.4 Conclusion

We have used a high-coverage and quantitative metabolome profiling technique based on differential CIL LC-MS to evaluate the effectiveness of different dilutions in metabolites analysis in UUMS. After LC-MS acquisition and data processing, there

were totally 202 peak pairs and 77 metabolites were identified. In order to test matrix effect behavior with reduction of the matrix amount, UUMS sample was diluted under three different factors, and peak pairs between experimental and theoretical values were compared for individual metabolites. Among identified metabolites, more than 83% of the metabolites showed no large difference between experimental and theoretical values with matrix being diluted by 4 factors. And more than half of the remaining compounds (7%) showed matrix effects improvement after matrix being diluted by 7-fold and half of the remaining compounds (7%) needed to be diluted by 11-fold. After being diluted by 11-fold, only 3% of the metabolites did not present any improvement. As for unidentified metabolites in UUMS sample, about half (49%) of the metabolites were free of matrix effects when matrix was diluted by 4 factors, 13% needed dilution up to a factor of 7. A relatively large percentage of compounds (29%) showed matrix effects improvement after matrix being diluted by 11-fold, and 9% of the metabolites did not improve, even with a dilution factor of 11. In the future, different approaches could be applied to obtain more results from the remaining unidentified metabolites. And also, we will evaluate the extent of matrix effect on different isotopic labeling reagents, such as carboxylic acids[39], hydroxyls[40], ketones and aldehydes[38], as well as different types of biofluids on protocol development for CIL LC-MS based metabolomics.

Chapter 4 Conclusions

By the use of high-performance chemical isotope labeling (CIL) LC-MS method, matrix effects in universal urine metabolome standard (UUMS) have been investigated. In Chapter 2, the absolute concentrations from standard addition method for four amino acids (alanine, serine, phenylalanine and threonine) in UUMS were determined. And matrix effects were evaluated by comparing absolute concentrations of selected amino acids from standard addition strategy with concentrations from external standard method. Comparison of UUMS at different concentrations that were labeled using the same protocol allows us to determine the extent of matrix effect on labeling. Without any dilution, the relative errors between absolute concentrations and concentrations from external standard method for four amino acids in UUMS were all lower than 10%. From the results of comparing UUMS at different concentrations, there was no indication that the extent of matrix effect on labeling would be lowered by dilution for these four selected amino acids.

In Chapter 3, the high-coverage and quantitative metabolome profiling technique based on differential chemical isotope labeling (CIL) LC-MS was used to evaluate the effectiveness of different dilutions in metabolite analysis in UUMS matrix. After LC-MS acquisition and data processing, there were totally 202 peak pairs with 77 metabolites identified. In order to test matrix effect behavior with reduction of the matrix amount, UUMS sample was diluted under three different factors, and peak pairs between experimental and theoretical values were compared for individual metabolites. Among identified metabolites, more than 83% of the metabolites showed no large difference between experimental and theoretical values with matrix being diluted by 4 factors. And more than half of the remaining compounds (7%) showed matrix effect improvement after matrix being diluted by 7-fold and half of the remaining compounds (7%) needed to be diluted by 11-fold. After being diluted by 11-fold, only 3% of the metabolites did not present any improvement. As for unidentified metabolites in UUMS sample, about half (49%) of the metabolites were free of matrix effects when matrix was diluted by 4 factors, 13% needed dilution up to a factor of 7. A relatively large

percentage of compounds (29%) showed matrix effects improvement after matrix being diluted by 11-fold, and 9% of the metabolites did not improve, even with a dilution factor of 11. These results suggest that, for most of identified metabolites, a dilution factor of four would be enough to minimize the influence of matrix effect, opening the possibility to perform quantitative analysis for these identified metabolites in urine.

References

1. Micheel, C. M.; Nass, S. J.; Omenn, G. S. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials. **2003**.
2. Hasin, Y., Seldin, M.; Lusis, A. Multi-Omics Approaches to Disease. *Genome Biol.* **2017**, *83* (18).
3. Lockhart, D. J.; Winzeler, E. A. Genomics, Gene Expression and DNA Arrays. *Nature.* **2000,** *405* (6788), 827.
4. Fiehn, O. Metabolomics—the Link Between Genotypes and Phenotypes. *Functional Genomics*. **2002**, *10* (48), 155-171.
5. Martin, J. A.; Wang, Z. Next-Generation Transcriptome Assembly. *Nature Reviews Genetics.* **2011,** *12* (10), 671.
6. Milward, E. A.; Shahandeh, A.; Heidari, M.; Johnstone, D. M.; Daneshi, N.; Hondermarch, H. Transcriptomics. *Encyclopedia of Cell Biology.* **2016**, *10* (4), 160-165.
7. Schwanhäusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global Quantification of Mammalian Gene Expression Control. *Nature.* **2011,** *473* (7347), 337.
8. Anderson, L.; Seilhamer, J.; A Comparison of Selected mRNA and Protein Abundances in Human Liver. *Electrophoresis*. **1997**, *18*, 533-537.
9. Gygi, S.P.; Rochon, Y.; Franza, B.R.; Aebersold, R. Correlation Between Protein and mRNA Abundance in Yeast. *Mol Cell Biol.* **1999**, *19*, 1720-1730
10. Vitzthum, F.; Behrens, F.; Anderson, N. L.; Shaw, J. H. Proteomics: From Basic Research to Diagnostic Application. A review of Requirements & Needs. *Journal of Proteome Research.* **2005,** *4* (4), 1086-1097.
11. Hein, M.Y.; Sharma, K.; Cox, J.; Mann, M. Proteomic Analysis of Cellular Systems. Handbook of Systems Biology: Concepts and Insights. **2013**, *978* (12), 3–25.
12. Selevsek, N.; Chang, C. Y.; Gillet, L. C.; Navarro, P.; Bernhardt, O. M.; Reiter, L. Reproducible and Consistent Quantification of the Saccharomyces Cerevisiae

Proteome by SWATH-Mass Spectrometry. *Mol Cell Proteomics*. **2015**, *14*, 739–749.

13. Dettmer, K.; Hammock, B. D., Metabolomics--a New Exciting Field Within the" Omics" Sciences. *Environ Health Persp.* **2004,** *112* (7), A396.
14. Clish, C. B. Metabolomics: An Emerging but Powerful Tool for Precision Medicine. *Cold Spring Harb Mol Case Stud*. **2015**, *1* (1), a000588.
15. Orešič, M. Metabolomics, a Novel Tool for Studies of Nutrition, Metabolism and Lipid Dysfunction. *Nutrition, Metabolism and Cardiovascular Diseases.* **2009,** *19* (11), 816-824.
16. Clarke, C. J.; Haselden, J. N. Metabolic Profiling as a Tool for Understanding Mechanisms of Toxicity. *Toxicol Pathol*. **2008**, *36* (1), 140-147.
17. Yanes, O.; Tautenhahn, R.; Patti, G.J.; Siuzdak, G. Expanding Coverage of the Matabolome for Global Metabolite Profiling. *Anal. Chem*. **2011**, *83* (6), 2152-2161.
18. Naylor, S. Biomarkers: Current Perspectives and Future Prospects. *Expert Rev. Mol. Diagn*. **2003**, *3* (5), 525-529.
19. Belinsky, S. A.; Nikula, K. J.; Palmisano, W. A.; Michels, R.; Saccomanno, G.; Gabrielson, E.; Baylin, S. B.; Herman, J. G. Aberrant Metthylation of P16(INK4a) Is an Early Event in Lung Cancer and a Potential Biomarker for Early Diagnosis. *Proc. Natl. Acad. Sci.* U.S.A. **1998**, *95* (20), 11891-11896.
20. Hagmar, L.; Bonassi, S.; Strömberg, U.; Brøgger, A.; Knudsen, L.E.; Norppa, H.; Reuterwall, C.; Chromosomal Aberrations in Lymphocytes Predict Human Cancer: A Report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH). *Cancer Res*. **1998,** *58* (18), 4117-4121.
21. Manach, C.; Hubert, J.; Llorach, R.; Scalbert, A. The Complex Links Between Dietary Phytochemicals and Human Health Deciphered by Metabolomics. *Mol. Nur. Food Res*. **2009**, *53* (10), 1303-1315.
22. Bothwell, J. H.; Griffin, J. L. An introduction to Biological Nuclear Magnetic Resonance Spectroscopy. *Biological Reviews.* **2011,** *86* (2), 493-510.

23. Wang, J. H.; Byun, J.; Pennathur, S. Analytical Approaches to Metabolomics and Applications to System Biology. *Semin. Nephrol*. **2010**, *30* (5), 500-511.
24. Smolke, C. D. The metabolic Pathway Engineering Handbook; CRC, **2010**.
25. Weiss, R. H.; Kim, K. Metabolomics in the Study of Kidney Diseases. *Nat. Rev. Nephrol.* **2012**, *8* (1), 22-33.
26. Nicholson, J. K.; Lindon, J. C. Metabolomics. *Nature*. **2008**, *455* (7216), 1054-1056.
27. Barba, I.; Fernandez-Montesinos, R.; Garcia-Dorado, D.; Pozo, D. Alzheimer's Disease Beyond the Genomic Era: Nuclear Magnetic Resonance (NMR) Spectroscopy- Based Metabolomics. *Journal of Cellular and Molecular Medicine* **2008,** *12* (5a), 1477-1485.
28. Jung, J. Y.; Lee, H. S.; Kang, D. G.; Kim, N. S.; Cha, M. H.; Bang, O. S.; Hwang, G. S. 1H-NMR-Based Metabolomics Study of Cerebral Infarction. *Stroke.* **2011**, *42* (5), 1282-1288.
29. Spratlin, J. L.; Serkova, N, J.; Eckhardt, S. G. Clinical Applications of Metabolomics in Oncology; *A Review. Clin. Cancer Res*. **2009**, *15* (2), 431-440.
30. Scholz, M.; Gatzek, S.; Sterling, A.; Fiehn, O.; Selbig, J., Metabolite Fingerprinting: Detecting Biological Features by Independent Component Analysis. *Bioinformatics.* **2004,** *20* (15), 2447-2454.
31. Lei, Z.; Huhman, D.; Summer, L. Mass Spectrometry Strategies in Metabolomics. *J Biol Chem*. **2011**, *286* (29), 25435-25442.
32. Dunn, W. B.; Ellis, D. I. Metabolomics: Current Analytical Platforms and Methodologies. *Trends in Analytical Chemistry.* **2005,** *24* (4), 285-294.
33. Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry. *Journal of Proteome Research.* **2003,** *2* (5), 488-494.
34. Ohashi, H.; Hasegawa, M.; Wakimoto, K.; Miyamoto-Sato, E. Next-Generation Technologies for Multiomics Approaches Including Interactome Sequencing. *Biomed Res Int.* **2015.**

35. Kuehnbaum, N. L.; Britz-McKibbin, P. New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era. *Chem Rev.* **2013**, *113* (4), 2437-2468.
36. Horgan, R. R.; Kenny, L. C. 'Omic' Technilogies: Genomics, Transcriptomics, Proteomics and Metabolomics. *Obstet. Gynaecol*. **2011**, *13* (3), 189-195.
37. Guo, K.; Li, L. Differential 12C-/13C-Isotope Dansylation Labeling and Fast Liquid Chromatography/Mass Spectrometry for Absolute and Relative Quantification of the Metabolome. *Anal Chem.* **2009,** *81* (10), 3919-3932.
38. Zhao, S.; Luo, X.; Li, L. Chemical Isotope Labeling LC-MS for High Coverage and Quantitative Profiling of the Hydroxyl Submetabolome in Metabolomics. *Anal. Chem.* **2016,** *88* (21), 10617-10623.
39. Guo, K.; Li, L. High-Performance Isotope Labeling for Profiling Carboxylic Acid-Containing Metabolites in Biofluids by Mass Spectrometry. *Anal. Chem.* **2010,** *82* (21), 8789-8793.
40. Zhao, S.; Dawe, M.; Guo, K.; Li, L. Development of High-Performance Chemical Isotope Labeling LC–MS for Profiling the Carbonyl Submetabolome. *Anal. Chem.* **2017,** *89* (12), 6758-6765.
41. Luo, X.; Gu, X.; Li, L. Development of a Simple and Efficient Method of Harvesting and Lysing Adherent Mammalian Cells for Chemical Isotope Labeling LC-MS- Based Cellular Metabolomics. *Anal. Chim. Acta.* **2017**.
42. Han, W.; Sapkota, S.; Camicioli, R.; Dixon, R. A.; Li, L. Profiling Novel Metabolic Biomarkers for Parkinson's Disease Using in-Depth Metabolomic Analysis. *Movement Disord.* **2017,** *32* (12), 1720-1728.
43. Wu, Y.; Streijger, F.; Wang, Y.; Lin, G.; Christie, S.; Mac-Thiong, J.-M.; Parent, S.; Bailey, C. S.; Paquette, S.; Boyd, M. C. Parallel Metabolomic Profiling of Cerebrospinal Fluid and Serum for Identifying Biomarkers of Injury Severity After Acute Human Spinal Cord Injury. *Scientific Reports.* **2016,** *6*, 38718.
44. Hooton, K.; Li, L. Nonocclusive Sweat Collection Combined with Chemical Isotope Labeling LC–MS for Human Sweat Metabolomics and Mapping the

Sweat Metabolomes at Different Skin Locations. *Anal. Chem.* **2017,** *89* (15), 7847-7851.

45. Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass Spectrometry-Based Metabolomics. *Mass Spectrometry Reviews.* **2007,** *26* (1), 51-78.
46. Wu, Y.; Li, L. Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Anal. Chem.* **2012**, *84* (24), 10723-10731.
47. Wu, Y.; Li, L. Dansylation Metabolite Assay: A Simple and Rapid Method for Sample Amount Normalization in Metabolomics. *Anal. Chem.* **2014**, *86* (19), 9428-9433.
48. Wu, Y.; Li, L. Sample Normalization Methods in Quantitative Metabolomics. *J. Chromatogr. A*. **2016**, *1430*, 80-95.
49. Zhou, H.; Yuen, P. S. T.; Pisitkun, T.; Gonzales, P. A.; Yasuda, H.; Dear, J. W.; Gross, P.; Knepper, M. A.; Star, R. A. Collection, Storage, Preservation, and Normalization of Human Urinary Exosomes for Biomarker Discovery. *Kidney Int.* **2006**, *69* (8), 1471-1476.
50. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L. Proposed Minimum Reporting Standards for Chemical Analysis. *Metabolomics*. **2007**, *3* (3), 211-221.
51. Kwak, M.; Kang, K.; Wang, Y. Methods of Metabolite Identification Using MS/MS Data. *Journal of Computer Information Systems*. **2019**, *10* (8).
52. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. METLIN: A Metabolite Mass Spectral Database. *Ther. Drug. Monit.* **2005,** *27* (6), 747-751.
53. Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S. HMDB: The Human Metabolome Database. *Nucleic Acids Research.* **2007,** *35*, 521-526.
54. Guest, P.C. The Importance of Biomarkers: The Required Tools of the Trade. *Biomarkers and Mental Illness*. **2017**, *46* (8), 31-41.

55. Mamas, M., Dunn, W.B., Neyses, L. *et al.* The Role of Metabolites and Metabolomics in Clinically Applicable Biomarkers of Disease. *Arch Toxicol.* **2011***, 85***,** 5–17.
56. Rifai, N.; Gillette, M. A.; Carr, S. A. Protein Biomarker Discovery and Validation: The Long and Uncertain Path to Clinical Utility. *Nat Biotechnol.* **2006***, 24*, 971–983.
57. Antoranz, A.; Sakellaropoulos, T.; Saez-Rodriguez, J.; Alexopoulos, L.G. Mechanism-Based Biomarker Discovery. *Drug Discovery Today*. **2017**, *22* (8), 1209-1215.
58. Newgard, C. B. Metabolomics and Metabolic Diseases: Where Do We Stand? *Cell Metab*. **2017**, *25* (1), 43-56.
59. Xia, J., Broadhurst, D.I., Wilson, M. *et al.* Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial. *Metabolomics.* **2013**, *9***,** 280–299.
60. Armitage, E. G.; Barbas, C. Metabolomics in Cancer Biomarker Discovery: Current Trends and Future Perspectives. *J. Pharm. Biomed. Anal*. **2014**, *87* (41), 1-11.
61. Han, W.; Sapkota, S.; Camicioli, R.; Dixon, R. A.; Li, L. Profiling Novel Metabolic Biomarkers for Parkinson's Disease Using In-depth Metabolomic Analysis. *Mov. Disord*. **2017**, *32* (12), 1720-1728.
62. Parker, C. E.; Borchers, C. H. Mass Spectrometry Based Biomarker Discovery, Verification, and Validation – Quality Assurance and Control of Protein Biomarker Assays. *Molecular Oncology*. **2014**, *8* (4), 840-858.
63. Horvatovich P. L.; Bischoff, R. Current Technological Challenges in Biomarker Discovery and Validation. *European Journal of Mass Spectrometry*. **2010**, *16* (1), 101-121.
64. Han, W., Li, L. Matrix Effect on Chemical Isotope Labeling and Its Implication in Metabolomic Sample Preparation for Quantitative Metabolomics. *Metabolomics.* **2015**, *11***,** 1733–1742.

65. Ito, S.; Tsukada, K. Matrix Effect and Correction by Standard Addition in Quantitative Liquid Chromatographic–Mass Spectrometric Analysis of Diarrhetic Shellfish Poisoning Toxins. *Journal of Chromatography A*. **2002**, *943* (1), 39-46.
66. Gergov, M.; Nenonen, T.; Ojanperä, I.; Ketola, R. A. Compensation of Matrix Effects in a Standard Addition Method for Metformin in Postmortem Blood Using Liquid Chromatography–Electrospray–Tandem Mass Spectrometry. *Journal of Analytical Toxicology*. **2015**, *39* (5), 359–364.
67. Sulistyaningrum, I.; Utami, M. P. G.; Istiningrum, R. B. Comparison Between the Calibration and the Standard Addition Methods in Determining Dissolved Lead in Borobudur's Control Tanks Water by Flame Atomic Absorption Spectrophotometry (F-AAS). *Procedia Chemistry*. **2015**, *17*, 70-74.
68. Pang, S.; Cowen, S. A generic Standard Additions Based Method to Determine Endogenous Analyte Concentrations by Immunoassays to Overcome Complex Biological Matrix Interference. *Sci Rep.* **2017***, 7*.
69. Zhou, R.; Tseng, C. L.; Huan, T.; Li, L. IsoMS: Automated Processing of LC-MS Data Generated by a Chemical Isotope Labeling Metabolomics Platform. *Anal. Chem.* **2014,** *86* (10), 4675-4679.
70. Huan, T.; Li, L. Counting Missing Values in a Metabolite-Intensity Data Set for Measuring the Analytical Performance of a Metabolomics Platform. *Anal. Chem*. **2015**, *87* (2), 1306-1313.
71. Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. DnsID in MyCompoundID for Rapid Identification of Dansylated Amine- and Phenol-Containing Metabolites in LC-MS-Based Metabolomics. *Anal.Chem*. **2015**, *87* (19), 9838-9845.
72. Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stainislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; et al. MyCompoundID: Using an Evidence-Based Metabolome Library for Metabolite Identification. *Anal. Chem*. **2013**, *85* (6), 3401-3408.
73. Jing, J.; Gao, Y. Urine Biomarkers in the Early Stages of Diseases: Current Status and Perspective. *Discov Med*. **2018,** *25* (136), 57-65.

74. Lowry, L. K. Role of Biomarkers of Exposure in the Assessment of Health Risks. *Toxicology Letters*. **1995**, *77* (1-3), 31-38.
75. Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Recent and Potential Developments of Biofluid Analyses in Metabolomics. *J Proteomics*. **2012**, *75* (4), 1079-88.
76. Lindh, C. H.; Littorin, M.; Johannesson, G.; Jönsson, B. A. G. Analysis of Chlormequat in Human Urine as a Biomarker of Exposure Using Liquid Vhromatography Triple Quadrupole Mass Spectrometry. *Journal of Chromatography B*. **2011**, *879* (19), 1551-1556.
77. Li, M.; Zhao, M.; Gao, Y. Changes of Proteins Induced by Anticoagulants can be More Sensitively Detected in Urine Rather than Plasma. *Sci China Life Sci.* **2014**, *57* (7), 649-656.
78. Wang, L.; Li, F.; Sun, W.; Wu, S.; Wang, X.; Zhang, L.; Zheng, D.; Wang, J.; Gao, Y. Concanavalin a Captured Glycoproteins in Healthy Human Urine. *Mol Cell Proteomics.* **2016**, *5* (3), 560-562.
79. Khamis, M. M.; Adamko, D. J.; El-Aneed, A. Mass Spectrometric Based Approaches in Urine Metabolomics and Biomarker Discovery. *Mass Spectrometry Reviews*. **2017**, *36* (2), 115-134.
80. Gao Y. Are Human Biomarker Studies Always More Valuable than Animal Ones? *MOJ Proteomics Bioinform.* **2015**, *2* (1), 35.
81. Zhao, M.; Liu, K.; Gao, Y. Phosphoproteins with Stability Against allUrinary Phosphatases as Potential Biomarkers in Urine. *Protein Pept Lett.* **2015**, *22* (9), 795-800.
82. Marchi, I.; Rudaz, S.; Veuthey, J. L. Sample Preparation Development and Matrix Effects Evaluation for Multianalyte Determination in Urine. *Journal of Pharmaceutical and Biomedical Analysis*. **2009**, *49* (2), 459-467.
83. Taylor, T. P.; Janech, M. G.; Slate, E. H.; Lewis, E. C.; Arthur, J. M.; Oates, J. C. Overcoming the Effects of Matrix Interference in the Measurement of Urine Protein Analytes. *Biomarker insights*. **2012**, *7*, 1-8.

84. Chambers, E.; Wagrowski-Diehl, D. M.; Lu, Z.; Mazzeo, J. R. Systematic and Comprehensive Strategy for Reducing Matrix Effects in LC/MS/MS Analyses. J *Chromatogr B Analyt Technol Biomed Life Sci*. **2007**, *852* (1-2), 22-34
85. Matuszewski, B. K.; Constanzer, M. L.; Chavez-Eng, C. M. Strategies for the Assessment of Matrix Effect in Quantitative Bioanalytical Methods Based on HPLC−MS/MS. *Anal. Chem.* **2003**, *75* (13), 3019–3030.
86. Kmellár, B.; Pareja, L.; Ferrer, C.; Fodor, P.; Fernández-Alba, A. R. Study of the Effects of Operational Parameters on Multiresidue Pesticide Analysis by LC-MS/MS. *Talanta*. **2011,** *84* (2), 262-273.
87. Pascoe, R.; Foley, J. P.; Gusev, A. I. Reduction in Matrix-Related Signal Suppression Effects in Electrospray Ionization Mass Spectrometry Using online two-Dimensional Liquid Chromatography. *Anal Chem*. **2001**, *73* (24), 6014-6023.
88. Bogialli, S.; Curini, R.; Corcia, A.; Laganà, A.; Nazzari, M.; Tonci, M. Simple and Rapid Assay for Analyzing Residues of Carbamate Insecticides in Bovine Milk: Hot Water Extraction Followed by Liquid Chromatography-Mass Spectrometry. *J Chromatogr A*. **2004**, *1054* (1-2), 351-357.
89. Kmellár, B.; Fodor, P.; Pareja, L, et al. Validation and Uncertainty Study of a Comprehensive List of 160 Pesticide Residues in Multi-Class Vegetables by Liquid Chromatography-Tandem Mass Spectrometry. *Journal of Chromatography. A.* **2008,** *1215* (1-2), 37-50.
90. Stachniuk, A., Fornal, E. Liquid Chromatography-Mass Spectrometry in the Analysis of Pesticide Residues in Food. *Food Anal. Methods.* **2016**, *9***,** 1654–1665.
91. Shariati-Rad, M.; Irandoust, M.; Amin, N.; Shamsipur, M. Solving Matrix Effect, Spectral Interferences and Nonlinearity by Generalized Standard Addition Method Coupled with Radial Basis Functions-Partial Least Squares: Application to Simultaneous Determination of Drugs in Urine. *Chemometrics and Intelligent Laboratory Systems*. **2013**, *120*, 77-83.

92. Kruve, A.; Leito, L.; Herodes, K. Combating Matrix Effects in LC/ESI/MS: The Extrapolative Dilution Approach. *Anal. Chim. Acta*. **2009**, *651* (1), 75-80.
93. García-Valcárcel, A. I.; Tadeo, J. L. A Combination of Ultrasonic Assisted Extraction with LC-MS/MS for the Determination of Organophosphorus Pesticides in Sludge. *Anal Chim Acta*. **2009**, *641* (1-2), 117-123.
94. Villagrasa, M.; Guillamón, M.; Eljarrat, E.; Barceló, D. Matrix Effect in Liquid Chromatography-Electrospray Ionization Mass Spectrometry Analysis of Benzoxazinoid Derivatives in Plant Material. *J Chromatogr A*. **2007,** *1157* (1-2), 108-114.
95. Ferrer, C.; Lozano, A.; Agüera, A.; Girón, A. J.; Fernández-Alba, A. R. Overcoming Matrix Effects Using the Dilution Approach in Multiresidue Methods for Fruits and Vegetables. *J Chromatogr A*. **2011**, *1218* (42), 7634-7639.
96. Koek, M.M., Jellema, R.H., van der Greef, J. et al. Quantitative Metabolomics Based on Gas Chromatography Mass Spectrometry: Status and Perspectives. *Metabolomics.* **2011***, 7*, 307–328.
97. Bian, Y., Zheng, R., Bayer, F.P. et al. Robust, Reproducible and Quantitative Analysis of Thousands of Proteomes by Micro-Flow LC–MS/MS. *Nat Commun.* **2020**, *11* **(**157).
98. Miggiels, P.; Wouters, B.; Westen, G. J. P.; Dubbelman, A. C.; Hankemeier, T. Novel Technologies for Metabolomics: More for Less. *TrAC Trends in Analytical Chemistry*. **2019**, *120*.
99. Xu, R. N.; Vaca, P.; Rieser, M. J.; El-Shourbagy, T. A. Highly Sensitive LC–MS–MS Analysis of a Pharmaceutical Compound in Human Plasma Using Monolithic Phase-Based On-line Extraction. *Journal of Chromatographic Science*. **2009**, *47* (6), 473-477.
100. De Nys, S., Putzeys, E., Vervliet, P. et al. A novel High Sensitivity UPLC-MS/MS Method for the Evaluation of Bisphenol A Leaching from Dental Materials. *Sci Rep*. **2018**, *8* **(**6981).

101. Chrchwell, M. I.; Twaddle, N.; Meeker, L. R.; Doerge, D. R. Improving LC-MS Sensitivity Through Increases in Chromatographic Performance: Comparisons of UPLC-ES/MS/MS to HPLC-ES/MS/MS. *Journal of Chromatography B*. **2005**, *25* (2), 134-143.