Development of Methods for Global and Targeted Metabolomics Analysis of Biological Fluids

Dmitri Sitnikov

A Thesis

In the Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Chemistry) at

Concordia University

Montreal, Quebec, Canada

November 2020

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: *Dmitri Sitnikov*

Entitled: *Development of Methods for Global and Targeted Metabolomics Analysis of Biological Fluids*

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (*Chemistry*)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair

Dr. Jack A. Kornblatt

_____ External Examiner

Dr. Liang Li

_____ External to Program

Dr. Vladimir Titorenko

_____ Examiner

Dr. Cameron Skinner

_____ Examiner

Dr. Ann English

_____ Thesis Supervisor

Dr. Dajana Vuckovic

Approved by _____

Graduate Program Director, *Dr. Y. Gélinas*

January 13, 2021 _____

Dean, *Dr. P. Sicotte*

Faculty of Arts and Science

Abstract

Development of Methods for Global and Targeted Metabolomics Analysis of Biological Fluids

Dmitri Sitnikov, Ph.D.

Concordia University, 2021

Protein precipitation with organic solvents is the most common extraction method used for global metabolomics of human plasma. This standard method has broad selectivity but does not provide good coverage of low abundance metabolites. Increasing the metabolite coverage by parallel analysis of plasma samples prepared using different organic solvents is inefficient due to the low orthogonality of their selectivity. Moreover, there is an increasing demand for more selective sample preparation methods suitable for targeted and untargeted metabolomics analysis of biofluids that can also provide better coverage of metabolites missed by current standard workflows. To achieve these goals, a novel sequence of orthogonal but complementary extraction methods was designed and evaluated after performing a quantitative systematic side-by-side comparison of seven extraction protocols with different selectivity to establish the most orthogonal combinations with good analytical performance. A novel sequential solid-phase extraction protocol was introduced to fractionate the metabolome into anion, cation, neutral, and zwitterion fractions, followed by a rigorous analytical assessment. The improvement in metabolome coverage in metabolomics study could be achieved also by analyses of alternative samples such as oral fluid. Therefore, the final goals were to evaluate oral fluid as a sample source for targeted metabolomics analysis via the development and validation of a liquid chromatography with tandem mass spectrometry assay and accumulate sufficient background knowledge to handle this sample type in future developments.

The execution of the first objective, i.e., the side-by-side comparison of solvent precipitation methods confirmed their low orthogonality and was accompanied by severe matrix effects in comparison to more selective methods. The comparison of solid-phase extraction (SPE) and liquid-liquid extraction methods revealed their orthogonal metabolite coverage to other methods. However, further experiments showed that methyl tert-butyl ether incompletely removed lipids, which resulted in the significant and undesirable splitting of lipids between aqueous and organic layers. Thus, the final SPE sequential fractionation protocol combined sample deproteinization by methanol followed by the sequential SPE executed on mixed-mode strong anion-exchange and mixed-mode strong cation-exchange polymeric sorbents. This produced four fractions enriched in

metabolites with distinct ionic properties providing flexibility to analyze the fraction individually or in strategic combinations tailored to various mass spectrometry methods. The final protocol resulted in a 1.6-fold increase in total metabolite coverage compared to methanol precipitation with a 2-fold increase in the analysis time. The method demonstrated excellent signal repeatability for both targeted (relative standard deviation < 13%) and global (relative standard deviation < 30% for 75% metabolites) metabolomics. Moreover, excellent separation of anion and cation metabolites (< 4% overlap) and small (< 27%) overlap between other pairs of metabolite classes allow supplementary assignment of ionic properties to unknown metabolites that can aid metabolite identification. The utility of the above sequential solid-phase extraction method for fractionation of the polar metabolome can, in future, extend beyond plasma to other biofluid types, such as oral fluid. Oral fluid allows portable sample acquisition and reflects the composition blood at least for several metabolites. Moreover, the similar first step (methanol precipitation) between two methods promised easier and faster transfer of the solid-phase extraction protocol from plasma to oral fluid. Two model analytes cortisone and cortisol were selected for a preliminary study in oral fluid to serve as standard metabolites in the future implementation of solid-phase extraction protocol. Their selection was driven by their biological and clinical role and a commercial availability of stable isotope labelled standards. Methanol precipitation and a short reversed-phase chromatography run were combined with tandem mass spectrometry analysis. The successful development and validation of the method revealed an accurate, precise (< 15% variability), and sensitive (low limit of quantitation = 0.31 ng/mL) high-throughput assay, whose selectivity outcompeted an immunoaffinity method, while requiring only 30 µL of oral fluid sample.

In summary, this thesis establishes the first steps towards the adaptation of the sequential solid-phase extraction method for simultaneous global and targeted metabolomics analyses of plasma and oral fluid. Such studies will benefit from the systemic, more in-depth characterization of metabolite status, enriched pathway coverage, and faster identification and transition of putative biomarkers from the discovery to the targeted stage. Therefore, my research provided the necessary groundwork for developing and validating universal preparation and analysis workflows, which will permit the execution of pluripotent metabolomics studies.

.

Acknowledgements

A special note of thanks to my supervisor, Dr. Dajana Vuckovic, and committee members Dr. Cameron Skinner and Dr. Ann English for useful comments, guidance, and criticism.

The author would like to thank the Natural Sciences and Engineering Research Council of Canada for research funding and the Department of Chemistry and Biochemistry at Concordia University for material support. The author wants to provide special thanks to the PERFORM Centre at Concordia University for Ph.D. scholarship funding.

I would like to thank all my current and previous laboratory partners: Dr. Shama Naz, Irina Slobodchikova, Alexander Napylov, Cian Monnin, Marianna Russo, Ankita Gupta, Rosalynde Sonnenberg, Oluwatosin Kuteyi, Lise Cougnaud, and Melika Mirabi.

Nobody has been more important to me during these studies than my family. I would like to thank my wife, Tatiana, and other members of my family, who were patiently waiting for the termination of this long journey.

CONTRIBUTION OF AUTHORS

Chapter 2 was published in an article entitled “Systematic assessment of seven solvent and solid-phase extraction methods for metabolomics analysis of human plasma by LC-MS”, authored by Dmitri G. Sitnikov, Cian Monnin, and Dr. Dajana Vuckovic and published in *Sci Rep* 6**,** 38885 (2016).
D.G.S. and C.M. executed all the experimental work and analysis, D.G.S. and D.V. designed the study, interpreted results, and co-wrote the manuscript. All authors reviewed and revised the manuscript.

Chapter 3 describes the study designed by Dmitri Sitnikov and Dajana Vuckovic. Experimental work and analysis were executed by D.S. The lipid analysis was executed by D. S. and Shama Naz, who also analyzed the lipidome data. D.S and D.V. interpreted results and co-wrote the manuscript. This is a first draft of manuscript which will be submitted for publication in March 2021 after revisions by all authors.

Chapter 4 describes the study designed by Dmitri Sitnikov, Dajana Vuckovic, and Peter J. Darlington. Experimental work and data analysis were executed by D.S, Azadeh Ghassemi, Catalina Carvajal Marysol Gonczi, and Cian Monnin. Sample collection and ELISA analysis were executed by A.G., and C. C. M. G. The LC-MS/MS was developed by D.S., who also executed the study validation and sample analysis with C. M. Results interpretation and the manuscript writing were executed by D.S, D.V. and P. J. D. All authors reviewed and revised the manuscript. Chapter 4 has been submitted for publication in December of 2020.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES APPENDIX A

LIST OF FIGURES APPENDIX B

LIST OF FIGURES APPENDIX C

LIST OF TABLES

LIST OF TABLES APPENDIX A

LIST OF TABLES APPENDIX B

LIST OF TABLES APPENDIX C

ABBREVIATIONS

(-ve) ESI	ESI in negative mode
(+ve) ESI	ESI in positive mode
µM	micromole per liter
2-D	two-dimensional
3-D	three-dimensional
3-MHPG	3-methoxy-4-hydroxy-phenylglycol
A	anion
ACN	acetonitrile
ACTH	adrenocorticotropic hormone
AN	mix of A and N fractions
ANOVA	analysis of variance
APCI	atmospheric pressure chemical ionization
aqMTBE NP	aqueous phase of MTBE
AZ	mix of A and Z fractions
BPD	bipolar disorder
C	cation
C18	octadecyl alkyl chain
C8	octyl alkyl chain
CE	capillary electrophoresis
Cer	ceramide
ChEBI	Chemical Entities of Biological interest
CID	collision-induced dissociation
CN	mix of C and N fractions
CNS	central nervous system
CRH	corticotropin-releasing hormone
CSF	cerebrospinal fluid
CV	coefficient of variation
CZ	mix of C and Z fractions
DDA	data-dependent acquisition
DFI	direct flow injection
DG	diglycerides
DIA	data-independent acquisition
DVBP	divinylbenzene-co-N-vinylpyrrolidone
ECD	electron capture dissociation
ECDt	electrochemical detection
EDTA	ethylene diamine tetra-acetic acid
ECLIA	electroluminescent immunoaffinity assay
EI	electron ionization

EIC	extracted ion chromatogram
EL	elution
ELISA	enzyme-linked immunosorbent assay
ESI	electrospray ionization
ETD	electron transfer dissociation
FA	formic acid
FDA	U.S. Food and Drug Administration
FDR	false discovery rate
FIA	flow injection analysis
FT	flow-through
FT-ICR	Fourier transformed ion cyclotron resonance
FWHM	full width at half maximum
GABA	4-aminobutanoic acid (gamma-aminobutyric acid)
GC-MS	gas chromatography with mass chromatography
GPF	gas-phase fractionation
HCD	higher energy-collision induced dissociation
HILIC	hydrophilic interaction chromatography
HLB	hydrophilic-lipophilic balanced (polymer)
HMDB	Human Metabolome Database
HPA	hypothalamic pituitary adrenal (axis)
HPLC	high pressure liquid chromatography
HPT	hypothalamic pituitary thyroid (axis)
HRMS	high resolution mass spectrometry
ICP	inductively coupled plasma ionization
IEX	ion exchange
IMS	ion mobility separation
IPA	isopropanol
IR	infrared
IROA	isotopic ratio outlier analysis
IS	internal standard
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	liquid chromatography
LC-MS	liquid chromatography-mass spectrometry
LLE	Liquid-liquid extraction
LLOQ	low limit of quantitation
LOD	low limit of detection
LPA	lysophosphatidic acid
LPC	lysophosphatidylcholine
LPS	lysophosphatidylserine
Lyso-SM	lysosphingomyelin

MAX	strong anion-exchange/divinylbenzene-co-N-vinylpyrrolidone
MCX	strong cation-exchange/divinylbenzene-co-N-vinylpyrrolidone
MDD	major depressive disorder
MeCN	acetonitrile
MeOH	methanol
MeSH	Medical Subject Heading ontology
MM64	60/40 (v/v) mix of acetonitrile and methanol
MPP	Mass Profiler Professional
MRM	multiple reaction monitoring
MS	mass spectrometer
MS/MS	tandem fragmentation mass spectrometry
MS^1	mass spectrum (no fragmentation)
MS^n	multistage fragmentation mass spectrometry
MTBE	methyl tert-butyl ether
MTBE NP	non-polar phase of MTBE
N	neutral
NMR	nuclear magnetic resonance
OPLS-DA	orthogonal partial least square differential analysis
PC	phosphatidylcholine
PCA	principal component analysis
PE	phosphatidylethanolamine
PEG	polyethylene glycol
PEP2	SPE cartridges with divinylbenzene-pyrrolidone phase
PG	phosphatidylglycerol
PI	phosphatidylinositol
PLS-DA	partial least square differential analysis
POMC	proopiomelanocortin
PS	phosphatidylserine
PSIG	pound-force per square inch
QA	quality assurance
QC	quality control
QCH	quality control high
QCL	quality control low
QCM	quality control medium
QQQ	triple quadrupole
QTOF	quadrupole time of flight
ROC	receiver operating characteristic
RP	reversed-phase

RSD%	relative standard deviation
RT	retention time
S/N	signal-to-noise
SAX	strong anion-exchange
SCX	strong cation-exchange
SD	standard deviation
sGPF	staggered gas-phase fractionation
SIL	stable isotope-labeled
SM	sphingomyelin
SNS	sympathetic nervous system
SOP	standard operation procedure
SPE	solid-phase extraction
SRM	selected reaction monitoring
sSPE	sequential SPE
SWATH	sequential window acquisition of all theoretical mass spectra
T3	triiodothyronine
T4	thyroxine
TG	triglycerides
Th	Thomson
TIC	total ion current
TMB	3,3′,5,5′-Tetramethylbenzidine
TRH	thyrotropin-releasing hormone
TSH	thyroid-stimulating hormone
UHPLC	ultrahigh pressure liquid chromatography
ULOQ	upper limit of quantitation
UV	ultraviolet
Vis	visible
V	volt
Z	zwitterion
ZIC-HILIC	zwitterionic HILIC

1 Introduction

1.1 Introduction to metabolomics

The metabolome is defined as the collection of all low-molecular-weight species (frequently defined by molecular weight < 1,500 Da) present in a given biological system of interest.[1] The comprehensive quantitative and qualitative analysis of metabolome in a given biological system is defined as metabolomics. The assessment of metabolite levels in a given system in response to a particular stimulus or treatment has historically been defined with the term metabonomics.[2] The increasing popularity of metabolomics approaches is illustrated by the number of metabolomics and metabonomics studies found in PubMed using the terms ("metabolomics" OR "metabonomics") AND ("disease" OR "biomarker" OR "medicine") which have increased from 63 in 2006 to 2050 in 2017.[3] In recent years, both terms "metabolomics" and "metabonomics" are used interchangeably[4] while the usage of "metabolomics" is more frequent. For example, a search on Science Direct https://www.sciencedirect.com/search returns 14427 and 1019 manuscripts published between 2014 and 2018 using terms "metabolomics" and "metabonomics", respectively. Therefore, in this thesis, the term metabolomics will be used to refer to both types of analyses.

The metabolome consolidates the influence of genomic, transcriptomic, proteomic, biological, and environmental factors and, therefore, is a reliable indicator of the status and perturbations of a phenotype. Metabolomics can be executed via global (untargeted) and targeted approaches. A global metabolomics analysis focuses on the holistic acquisition of data for as many metabolites as possible in a given sample. This type of analysis may include incomplete or partial qualitative identification of metabolites and utilize relative quantification in order to establish key metabolites affected by a given treatment and use the resulting information to generate new hypotheses.[5] However, the global approach has a high risk of false discoveries because it lacks standard references for many metabolites, frequently exhibits poor quantitative performance for various metabolites, and is plagued by numerous confounding factors that may impact data interpretation.[4–6] In contrast to global metabolomics, a targeted metabolomics analysis provides absolute quantification of a limited number (tens to hundreds) of known metabolites and uses an extensive set of authentic reference standards. This decreases the risk of generation of analytical artifacts but will not be able to detect the markers which are not included in the initial set of targets.[7] A review of metabolomics approaches is provided in **Section 1.2.**

The combination of targeted and global strategies in a single study helps to minimize the intrinsic limitations of each. The current metabolomics discovery paradigm includes a global study to generate a hypothesis using one cohort of samples, followed by targeted analysis of selected differentially regulated metabolites in a second independent cohort to validate the findings. For example, in the metabolomics study of blood plasma from patients treated with rosuvastatin, metabolites whose concentrations were significantly altered by rosuvastatin treatment were detected, and identified and then, the list of *a priori* selected relevant metabolites was quantified via a targeted metabolomics approach.[8]This allows verification of *de novo* discovered putative biomarkers and/or expansion of the list of targets to ensure better coverage of selected pathways using targeted approach. Another strategy used in this study was a parallel analysis of lipid and non-lipid metabolites using distinct sample preparation and liquid chromatography (LC) – mass spectrometry (MS) methods optimal for each. This increased metabolome and pathway coverage and allowed to comprehensively integrate perturbations between lipid and non-lipid pathways. Furthermore, it facilitated a biological interpretation of the data to propose the mechanism of myopathy, a side effect that occurs in hyperlipidemic patients upon rosuvastatin treatment.

Another approach that has gained popularity in recent years is to combine targeted and global analyses during the discovery stage in order to ensure the coverage of metabolites of specific interest to a particular disease or condition to provide more detailed information about the patient health status. This approach was used in the LC-MS profiling of polar metabolites in serum after liquid-liquid extraction (LLE) in order to identify unique biomarker profiles for cervical cancer in patients with human papillomavirus infection.[9] They used a targeted approach for the verification of altered pathways and the establishment of an objective clinical staging of the infection in patients, and a global approach for the detection of putative biomarkers in samples from the same patients. The combination of the results improved the quality of diagnostic biomarker panels and the stratification of patient risk groups and decreased study time. Another example of the combined approach is in the study of the protective effects of traditional Chinese medicine against systemic inflammation in rats.[10] This study combined global profiling of metabolites in serum by GC-MS, targeted profiling of fatty acids by LC-MS, and ELISA-based assay of corticosteroids in urine. The results of the targeted approach helped to filter and improve the quality of the global set of markers with a concomitant discovery of a new putative biomarker (gluconic lactone). Furthermore, the enriched (global and targeted) biomarker set enabled better quality for the

stratification of rats exposed to different anti-inflammation therapy. Therefore, the combined analysis may be extended to different types of biological samples and provide early and more detailed scientific data compared to global metabolomic analysis of a single biofluid.

1.1.1 Proposed metabolomics analysis of bipolar disorder

Based on the above advantages, we intended to apply combined global and targeted metabolomics approaches for the analysis of polar and mid-polar metabolites extracted from the blood of patients with different subtypes of bipolar disorder (BPD). The disease was selected due to its socio-economic impact (affects up to 1% of adult Canadians) and the lack of objective tests[11,12] for the differential diagnosis of BPD from other disorders related to the central nervous system (CNS).[13,14,15] The disorder manifests symptoms of depression, which is occasionally interrupted by switches to manic or sub-manic mood states. The neuroendocrine background demonstrates the dysregulation of the hypothalamic-pituitary-adrenal (HPA) and hypothalamic-pituitary-thyroid axis (HPT)[16] (**Figure 1.1**). For example, the hyperactivity of HPA associated with elevated levels of the corticotropin-releasing hormone (CRH), adrenocorticotropic hormone (ACTH), and cortisol is considered one of the major neuroendocrine abnormalities in depression and in BPD.[16,17] Another abnormality is the HPT dysfunction which is manifested by the enhanced activity of thyrotropin-releasing hormone (TRH) and blunted thyroid-stimulating hormone (TSH) response to injections of TRH.[16] Finally, as it is shown in **Figure 1.1,** the increased activity of the HPA axis has a negative effect on the HPT axis.[17] However, the detailed molecular mechanisms of HPA and HPT dysfunctions, their interactions, and an influence on the entire organism remain unknown in BPD. Therefore, it was necessary to create a new method, which combines the advantages of global and targeted approaches and enables us to study the contribution of the HPA-HPT axis in BPD and discover new associated biomarkers. We proposed to execute these combined metabolomics analyses based on the robust methanol-based precipitation with wide metabolome coverage[18] and analyze samples using reversed-phase (RP) and hydrophilic interaction (HILIC) chromatography coupled to high-resolution quadrupole time-of-flight mass spectrometer (QTOF-MS).

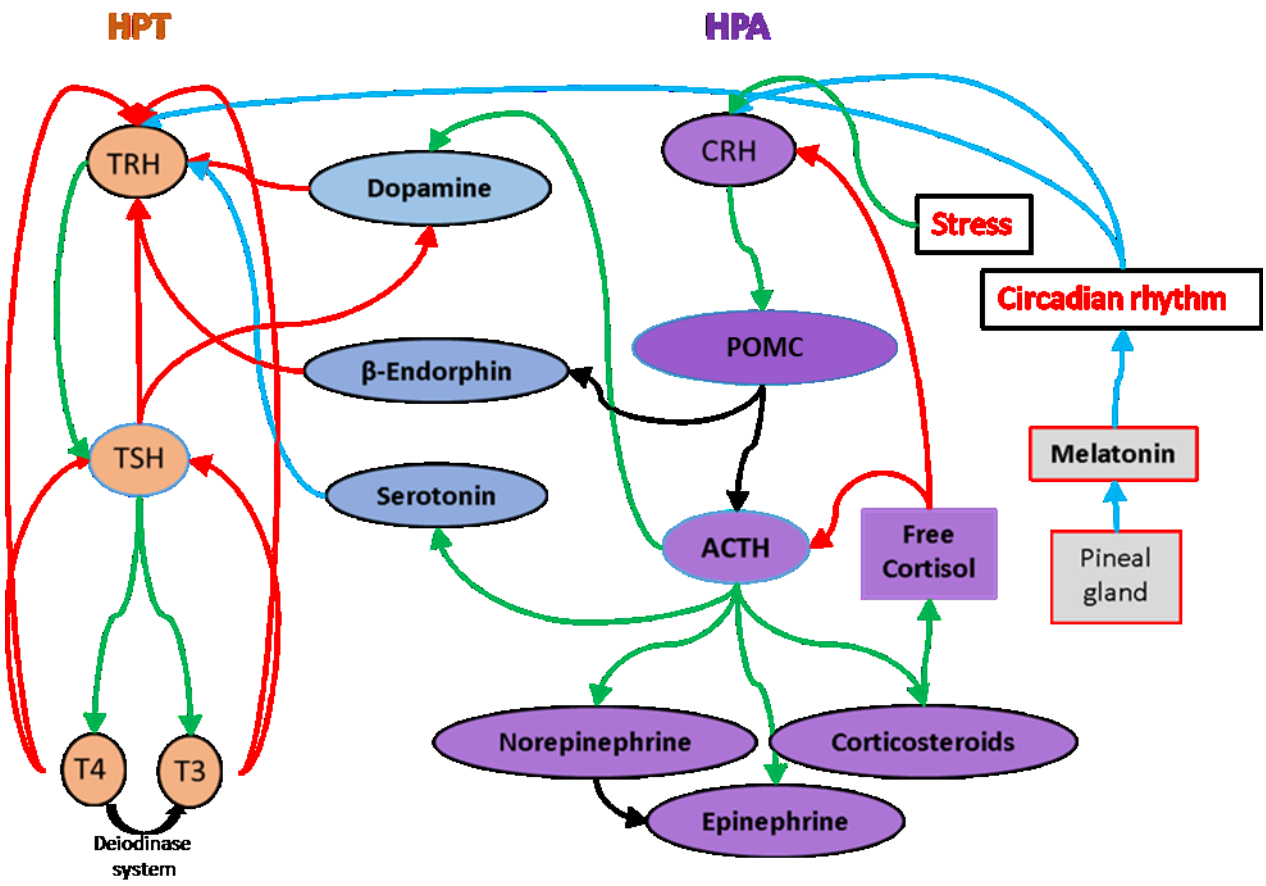


Figure 1.1. Hypothalamic-pituitary thyroid (HPT) and hypothalamic-pituitary-adrenal (HPA) axis, external factors (stress and circadian rhythm), key regulatory metabolites, and peptides involved in the pathogenesis of BPD. Regulation of HPA and HPT activities is executed via activation (green), modulation (blue), inhibition (red), and biotransformation (black) scenarios. Corticotropic releasing (CRH) and thyrotropin-releasing hormones (TRH) are produced in the hypothalamus. TRH stimulates the production of thyroid-stimulating (TSH) hormone in the pituitary gland, which is released to the vascular system and increases the production of thyroid hormones thyroxine (T4) and relatively smaller amounts of triiodothyronine (T3). These hormones are distributed via the bloodstream, regulate multiple peripheral processes, and also provide feedback inhibition of TSH and TRH in the CNS. Inactivation of T4 to T3 is executed by the deiodinase system in various tissues. In the HPA axis, CRH stimulates the truncation of proopiomelanocortin (POMC) into adrenocorticotropic hormone (ACTH). The latter is distributed via the bloodstream to hormonal glands and stimulates the synthesis of cortisol, epinephrine, and norepinephrine. Cortisol is released into the bloodstream and controls multiple functions, including feedback inhibition of HPA. Activated HPA inhibits the HPT axis via increased synthesis of dopamine and beta-endorphin.

To ensure that all metabolites of specific interest participating in HPT and HPA networks (**Figure 1.1**) could be detected, targeted triple quadrupole MS analysis was planned to supplement high-resolution MS as needed. The study aimed to achieve the following goals: (i) detection and

identification of differentially expressed metabolites and corresponding pathways in the global analysis; (ii) absolute quantitation of neurotransmitters and hormones involved in HPA and HPT regulatory networks and circulating in the blood such as TRH, TSH, T4, T3, dopamine, β-endorphin, serotonin, CRH, ACTH, cortisol, melatonin and catecholamines (**Figure 1.1)**; (iii) absolute quantitation of several members of synthesis pathways for metabolites listed in (ii); (iv) verification of putative biomarkers and pathways discovered during the global analysis and finally (v) construction of the targeted-global biomarker sets capable for the stratification of patient cohorts.

This design is different than that of other metabolomics studies of BPD and CNS disorders which used nuclear magnetic resonance (NMR)[19] or LC-MS[20–22] in either targeted[22] or global[19–21] metabolomics approaches. These studies used the same classical solvent precipitation, which was also proposed in our study. These studies assigned a significant number of plasma metabolites representing systemic pathways (choline, myoinositol, amino acids, biotin, pyruvate, linoleic acid, 4-aminobutanoic acid (GABA), 3-hydroxybutyrate, betaine, creatinine, 3-methoxy-4-hydroxy-phenylglycol (3-MHPG), phenylacetic acid) as putative biomarkers. All of these metabolites are highly abundant in plasma, with GABA present at the lowest concentration (approximately 0.1 μM in healthy individuals as reported in the Human Metabolome Database (HMDB) v 4.0.[23,24]

Highly abundant metabolites participate in multiple systemic pathways which may be redundant between different diseases. For example, pathways affected in BPD (acetyl-cholinergic, GABAergic, noradrenergic, dopaminergic, serotonergic) are very similar to the list of those in schizophrenia (dopaminergic, serotonergic, glutamatergic) and in major depression disorder (GABAergic, serotonergic, noradrenergic).[25,26] The most outstanding example of metabolite and pathway redundancy was published recently by Lindahl *et al.*[27] The topic of this research was not focused on CNS disorders and BPD but demonstrated the systematic limitations in sensitivity and coverage of current metabolomics studies. Thus, the LC-MS analysis of methanol-precipitated plasma of several distinct diseases demonstrated an overlap of 61% (out of 178 of all putative biomarkers) between non-related diseases (non-Hodgkin lymphoma, congestive heart failure, and infectious pneumonia). Moreover, even the identified metabolites which appeared as specific to a given disease tested in the study (phenylalanine, lysophosphatidylcholines (O-18:0) and (20:5), and androsterone sulfate) have been reported to be specific to other diseases in previous un-related studies.

In sum, sample preparation and LC-MS methodologies used in current global metabolomics studies of plasma demonstrate a general tendency to discover systemic, highly abundant putative biomarkers with low specificity to a particular disease. Therefore, in my study, the development of new approaches for global metabolomics analysis with better sensitivity and coverage was prioritized over proceeding with the immediate execution of another biomarker study prone to the generation of another list containing nonspecific biomarkers. Ultimately, the limitations of current global metabolomics approaches could not support the long-term goal of discovering and introducing objective BPD diagnostic biomarkers into clinical practice.

A global or targeted metabolomics analysis of biological specimens generally follows the typical workflow depicted in **Figure 1.2** and may be applied to studies of biofluids [28–32], tissues[33,34], cell cultures[32], and breath.[35] The subsequent sections will describe in detail common metabolomics approaches employed for biological fluids.

1.1.2 Study design and selection of biological fluids for metabolomics analysis

Global metabolomics studies can be used to search for metabolite differences between sets of samples from subjects with different phenotypes. Therefore, the study design defines biological or clinical parameters that will be used to segregate study cohorts, cohort sizes, and the enrollment criteria for the subjects of the study.[36,37] The selection of a sample source(s) to use for a given study is determined by the expected availability and reliability of analytes in a sample, by the availability of samples in repositories, and/or the invasiveness of sample collection procedures. These parameters differ between the four most frequently used bio-fluids: urine, cerebrospinal fluid (CSF), saliva, and blood.

Urine is generated by kidneys and is an easily accessible biofluid if collected by urination. The sampling of urine requires privacy, but it is non-invasive and does not require specialized medical attention or space. It contains water-soluble metabolites filtered in kidneys from the blood, such as sugars, electrolytes, amino acids, organic acids, lipids, and trace levels of proteins. The pH of urine in healthy individuals ranges from 4.5 to 7.5, and its osmolarity ranges from 570 to 800 mOsm/kg.[38]

1.2 Overview of common LC-MS metabolomics approaches for the analysis of biological fluids

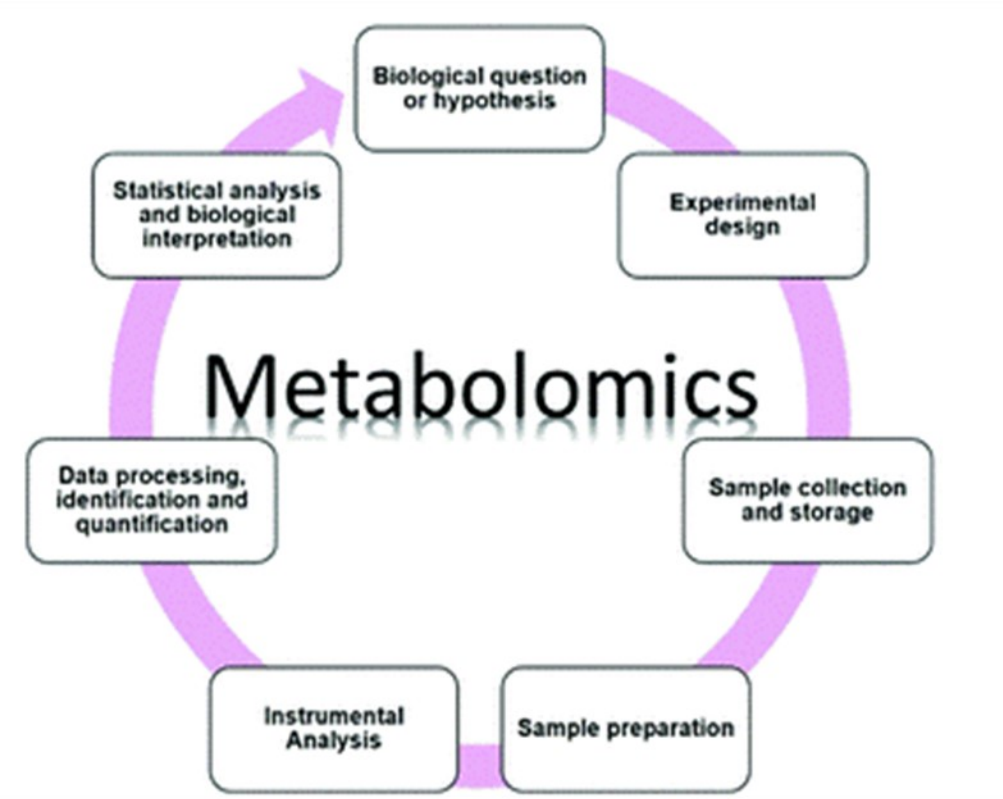


Figure 1.2. General workflow of metabolomics analysis. Reproduced/Adapted from[28] *with permission from The Royal Society of Chemistry.*

The concentration of metabolites in urine varies fold-wise depending on a number of factors, including hydration and kidney status, which mainly determine urine volume and analyte concentrations.[39,40] The inter-patient differences in volumes can be corrected by several approaches, including concentrations of creatinine, osmolality/osmolarity, specific gravity, or total mass spectral signal[39–42]. Urine is widely used in targeted and global metabolomics studies, clinical, toxicological, and doping control analysis.[43,44]

Cerebrospinal fluid (CSF) occupies subarachnoid spaces around the brain, ventricles, cisterns, and sulci inside the brain and the central canal of the spinal cord. The metabolite composition of CSF is closely linked to brain function and provides exclusive biomarker and pathway information for studies on neurodegenerative disorders, drug pharmacokinetics, and CNS regulation.[45–47]

However, high invasiveness and risk of CSF sampling preclude the routine usage of this sample source in clinical and "-omics" analysis.[48]

Saliva is excreted by salivary glands into the mouth cavity where it is mixed with transudate of buccal mucosa, secretion of gingival cavities, and the contents of the mouth, resulting in the generation of oral fluid.[49] Sampling of oral fluid is non-invasive and does not require medical supervision. This improves participants' enrollment and study completion rates. The concentrations of several metabolites in oral fluid are known to correlate well with their concentration in blood. For example, the concentration of cortisol in oral fluid correlates with its free concentration in blood and could be used in studies of mental diseases, a diagnostic of Cushing's syndrome and Addison's disease.[50–53] The ease and feasibility of oral fluid collection enable longitudinal and continuous sample collection during various activities in a non-restricted environment, which would be unfeasible for the sampling of other biofluids.[54] Oral fluid consists of approximately 99% water, 40-140 mM electrolytes, metabolites, proteins, bacteria, epithelial and blood cells, cell debris, and food remnants.[54,55] Total protein content of oral fluid is about 6.7 mg/mL. Its pH is about 7.2, and the average production rate of saliva is between 0.75 and 1.5 L of saliva per day.[56,57] The volume of oral fluid depends on physical activity and disease status and could be corrected between individuals based on calculated blood volume and, potentially, osmolarity.[58,59]

Blood plasma and serum are cell-free derivatives of whole blood and obtained by blocking or permitting a coagulation cascade, respectively. The collection of blood by venipuncture is an invasive procedure and requires specialized space and qualified medical attention. In contrast to other biofluids, blood is in direct contact with a larger number of organs and tissues and acts as a primary metabolite carrier. Therefore, blood metabolome composition may be indicative of a larger number of pathophysiological processes in an organism than other biofluids.[60] In addition, the long-time usage of blood as a diagnostic specimen in clinical analysis and "-omics" studies may provide additional information regarding the interpretation, and known reference ranges for metabolites of historical clinical interest. Subsequently, blood is used widely as a sample source in metabolomics analysis. Blood consists of white and red blood cells, 92-95% of water in addition to gases, electrolytes, salts, lipid and non-lipid metabolites, peptides, proteins (up to 7% w/v), extracellular vesicles, and cell debris.[61,62] The total ion concentration in healthy plasma is 270-295 mM and comprises of major ions such as Na^+, Cl^- and HCO_3^-. The pH of collected plasma is

maintained between 7.35 and 7.45 and determined mostly by the bicarbonate/carbonic acid ratio.x.[63,64] Homeostasis of blood is maintained by multiple mechanisms. Unlike urine and oral fluid, the blood volume is under homeostatic control by several mechanisms, including Na-dependent filtration in kidney glomerulus.[65]

Each of these four biofluids has different metabolite composition, but many metabolites may be detected in more than one of the biofluids. Therefore, the analysis of several biofluids within the course of a single metabolomic study may increase metabolite coverage and provide complementary information. However, the evaluation of all biofluids in a single study is unfeasible. This thesis will focus on the evaluation of oral fluid and blood, while CSF and urine will be omitted from further discussions.

1.2.1 Sample collection, pre-processing, and storage of samples for metabolomics analysis

The collection of a representative sample, its modification without loss of integrity, and the preservation of metabolite profiles until analysis are critically important to ensure the success of a metabolomics study. Typically, sample integrity is preserved via metabolism quenching or, if not executed, by the use of cold temperatures to minimize enzyme activity. Blood samples are usually collected into tubes containing anticoagulants to inhibit clotting. For the production of plasma, tubes are centrifuged at +4°C and the upper, cell-free layer is distributed into clean tubes and stored at below - 20 ºC for further analysis. The most common anticoagulants for plasma preparation are ethylene diamine tetra-acetic acid (EDTA), cationic salts of heparin, citrate, and occasionally - fluorides. All of them effectively stop the coagulation cascade, but the suitability and the influence of a particular anticoagulant on the quality of metabolomics results is the subject of discussion in the literature.[66,67] In contrast to the blood plasma preparation, a serum is prepared from blood without additives. For serum isolation, the coagulation (clotting) process is allowed to proceed for 30-60 min at room temperature. Then, coagulated blood is centrifuged, and the upper, cell-free phase is collected for further analysis.[68] Typically, plasma is more frequently used for metabolomics studies as it avoids metabolite changes associated with room temperature incubation step, which is required for activation of the clotting cascade for serum. Preparation of both plasma and serum requires relatively expensive equipment and a minimal volume of several hundred microliters. In the case of equipment shortage or volume limitations (newborns) the plasma can be

substituted by whole blood collected via finger-prick tests and dried blood spots.[69] However, the assays executed on these samples may suffer from the presence of metabolites from blood cells, which will complicate analysis and interpretation of results due to increased complexity, artifacts and decreased reproducibility. Thus, blood collection remains an invasive and challenging technique, which raises a significant interest in oral fluid as an alternative sample source. Oral fluid can be collected directly into sterile, 50 mL Falcon tubes by spitting or by oral swabs of various shapes after rinsing of the mouth to remove food debris.[54] If the study requires the collection from a specific salivary gland or continuous collection, the acquisition could be executed with special devices or capillary tubes.[70] Collected samples are then immediately placed at +4 ºC or centrifuged if the removal of glycoprotein (mucin) clot is required and then stored at below -20 ^{0}C.[56,57]

1.2.2 Sample preparation and LC-MS analysis

1.2.2.1 Global metabolomics analysis of biological fluids

1.2.2.1.1 *Sample preparation*

The objective of global metabolomics is to analyze the maximum number of metabolites with diverse physical-chemical properties. Various metabolites and metabolite classes may require special considerations in sample preparation and analysis. To avoid such issues, minimal sample manipulation and unselective sample preparation methods predominate in global metabolomics workflows to ensure good reproducibility while minimizing metabolite losses inherent in the use of more selective sample preparation methods.[36] Thus, the analysis of samples in the native state, or after minimal sample manipulation, i.e., sample dilution and/or solvent-based (methanol, acetonitrile, ethanol) protein precipitation is the most frequently used sample preparation approach in global metabolomics of biofluids.[71–75] Protein precipitation using an organic solvent is the leading method for a global metabolomics analysis of blood-based derivatives and oral fluid due to the low cost, high sample throughput, excellent reproducibility, and wide metabolome coverage.[76]

1.2.2.1.2 *LC-MS in metabolomics*

Currently, global metabolomics is frequently executed by NMR or mass spectrometry (MS) with or without coupling to separation techniques. High-resolution NMR is a non-destructive technique that generates quantitative, reproducible signals which are informative of the chemical structure of compounds and provides good inter-laboratory similarity of the data. The main disadvantages of NMR are large sample volumes and low sensitivity/poor limits of detection, which result in lower metabolite coverage (typically < 100 metabolites) than MS-based methods.[19,77] Among separation techniques, gas chromatography (GC) is limited to volatile and thermostable compounds and capillary electrophoresis (CE) to hydrophilic, charged molecules. Besides, CE-MS has a limited sensitivity due to the use of makeup fluid.

Compared with GC-MS and CE-MS, LC-MS is applicable to a wider range of metabolites, which makes it one of the most popular strategies for global metabolomics. The LC-MS analysis is typically executed on high-resolution mass spectrometers (HRMS) such as QTOF MS (routine resolution 10,000 - 30,000 FWHM, with some top models reaching a resolution of up to 60,000 FWHM) or Orbitrap MS (resolution of up to 1,000,000 FWHM depending on the model and acquisition method). In addition to high resolving power, these modern MS instruments also provide high mass accuracy (< 1-2 ppm) when coupled with internal mass calibration. Due to the versatility and potential high metabolite coverage, LC-MS analysis will be used in the current study. Combining multiple analytical techniques beyond just LC-MS can be used to increase metabolite coverage but increases the cost of analysis and is typically employed for deep mining of metabolome composition[57,61]

The versatility of chromatography columns, which, when combined strategically, may allow the retention of molecules with orthogonal properties and increase metabolome coverage via parallel analysis of the same sample. LC reduces sample complexity in the ionization source and subsequent matrix effects,[78] which increases metabolite coverage.[79] A frequent combination in metabolomics includes the analysis of the same sample on a reversed-phase (RP) column consisting of octadecyl alkyl (C18) and hydrophilic interaction liquid chromatography (HILIC) which provide complementary retention and separation of less polar (RP) and more polar/charged (HILIC) metabolites.[2] The RP C18 retains nonpolar compounds via hydrophobic interactions, which are eluted by a mobile phase with gradually decreasing polarity. The process is reversed in HILIC, where the partitioning occurs between the nonpolar mobile phase and a polar water layer

adjacent to a polar stationary phase is considered to be the main retention mechanism. Depending on the stationary phase employed in HILIC, this water layer portioning may be accompanied by interactions driven by electrostatic and/or hydrogen bonding forces. The stationary phase employed in zwitterionic ZIC-HILIC incorporates both positive and negative moieties which provide electrostatic interactions with positively and negatively-charged functional groups of the analytes. [80] ZIC-HILIC separations in performed near neutral pH improve the retention of a large set of anionic, cationic and zwitterionic metabolites,[81] estimated to represent up to 78% of non-lipid metabolome in HMDB in 2013.[82] ZIC-HILIC retained higher number of metabolites and produced signals with better peak shapes compared to several other HILIC columns[81,83] This makes its usage advantageous to increase the metabolome coverage. When coupling RPLC or ZIC HILIC to MS, volatile mobile phases must be used. Thus, formic and acetic acids, ammonium formate and acetate, and other volatile acids or salts are used to improve chromatography and/or ionization efficiency.[84]

1.2.2.1.2.1 Ionization of metabolites in ESI

The ESI source is an interface that connects LC to MS and is the most frequently used ionization method in global metabolomics. Electrospray ionization occurs in the electrostatic field at atmospheric pressure in the interface between a nebulizer and the MS inlet. The field is formed by applying a voltage (typically in the range of 2000-4500 V) between the nebulizer and MS inlet. If a positive voltage is applied on the nebulizer, it generates a "plume" with positively charged ions (+ve) ESI, and negative in (-ve) ESI.[85] An electrospray ionization combines droplet surface ionization, gas-phase ionization and in-solution ionization. If the number of analyte molecules in the ESI source exceeds the limited amount of charge available at surfaces and in gas phase for the formation of ions, the ionization process becomes highly competitive and leads to the generation of matrix effects.[86,87] Thus, ionization matrix effects are manifested by a change in analyte signal intensity produced by the same number of analyte molecules when present in samples of different complexity and composition.

1.2.2.1.3 *LC-MS/MS analysis*

In global metabolomics workflows, there is a need to fragment as many metabolites as possible and a tandem mass spectrometry (MS/MS) is one of the tools used in order to aid the identification. This can be accomplished in two conceptually different ways: data-dependent acquisition (DDA)

and data-independent (DIA) acquisition. For DDA, the selection of precursor ions for fragmentation is executed by data-acquisition software based on user-defined prioritization criteria: for example, the highest relative abundance and isotopic pattern of the precursor ion or the presence of diagnostic ion. The DDA attempts to collect MS/MS spectra from a single metabolite at a time, which enables detection, quantification, and identification of metabolites. However, not all metabolites will be selected for fragmentation, depending on instrument acquisition speed and sample complexity. In this case, processing algorithms as "ms-Purity", "iterative exclusion -omics", and acquisition protocols such as gas-phase fractionation (GPF) can be used.[88,89] They improve DDA limitations such as precursor impurity, limited spectral coverage, and increase the number of features with collected MS/MS spectra.[90] In contrast to DDA, the DIA method fragments all ions within the selected m/z range simultaneously but the assignment of fragments to precursor ion is complicated compared to DDA. Examples of DIA approaches are: (i) an alternating wide m/z range scan collected at low or high collision energies (MS^E) and (ii) SWATH or Sequential Window Acquisition of all Theoretical Mass Spectra, in which fragmentation is applied on groups of ions isolated in sequential windows of 20-50 Da. Compared to MS^E, SWATH approach simplifies the reconstruction of a lineage between precursors and fragment ions and improves spectra quality. Finally, the MS/MS spectra, which are acquired in the course of a global metabolomics study using DDA, DIA, or targeted MS/MS strategies, enable the elucidation of metabolite structures and assist in confident metabolite identification. These operations require specialized computer algorithms, databases, and libraries, which facilitate the processing and analysis of large datasets.[91,90]

1.2.2.1.4 *Processing of MS and MS/MS data*

The analysis of raw LC-MS data is very complex and is executed using computer-based algorithms included in commercial software packages such as MassHunter™ (Agilent Technologies), Xcalibur/Compound Discoverer™ (ThermoFisher Scientific), or free software packages such as XCMS[92] or MetaboAnalyst.[93] The initial analysis generates a list of ions with a unique combination of m/z and retention time, i.e., MS features. A single metabolite may yield many features including different adducts, isotope peaks, and singly/multiply charged ions. After alignment of feature RT across analytical runs, features belonging to background ions or random noise should be filtered out from the dataset, but this process may not be 100% effective. Then,

remaining signals are used for further processing, where isotopes and adducts likely belonging to a single metabolite are grouped together. To these putative metabolites, molecular formula is then assigned. If necessary, the complete elucidation of the structure of metabolite is carried out in a multi-step process and requires the use of additional equipment (**Section 1.3.4.3**). One key difference between global metabolomics and targeted analysis is the large number of analyzed metabolites, and subsequently, the drastically larger size and complexity of the data. In addition, the lack of authentic standards puts a great emphasis on the identification of unknown metabolites of interest, unlike targeted methods where authentic standards are typically included in the analysis and used for immediate confident identification of the targeted analytes.

1.2.2.2 Targeted metabolomics analysis of biological fluids

The goal of a targeted metabolomics study is the measurement of differences in concentrations of *a priori* selected metabolites. The selection of targets can be executed according to their biological roles or in a class- and/or pathway-specific manner depending on the objectives of the study.[94,95] The knowledge of chemical/physical properties of analytes and the overall limited number of target metabolites enables the use of analyte-tailored sample preparation and instrumental techniques. Together, these analyte-specific strategies can yield higher sensitivity of a targeted analysis as compared to a global one. For example, sample preparation may be executed with the use of more selective extraction methods such as solid-phase extraction (SPE). Despite its higher cost, SPE usage can provide a drastic reduction of interferences and higher sample pre-concentration than solvent extractions.[96] Moreover, the knowledge of the targets' properties and their relatively small number facilitate the adjustment of LC methods to resolve targets of interest from coeluting impurities and reduce matrix effects. Samples are typically analyzed using RP or HILIC method in (+ve) and (-ve) ESI, depending on the properties of the target analytes. Among RP methods, the use of columns with C18 stationary phases is the most popular, while for HILIC methods, underivatized silica, amide, and zwitterionic phases are predominant.[76,81]Beyond enrichment during sample preparation and resolving impurities by LC, targeted metabolomics approaches achieve better sensitivity than untargeted metabolomics by using triple quadrupole MS (QQQ). These instruments can be set-up in the selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) modes to monitor selected transition(s) of precursor ion to product ion(s). This

can considerably reduce noise and enhance signal-to-noise (S/N) ratios, thus resulting in lower limits of detection.

The identification of metabolites in targeted metabolomics is facilitated by the availability of standards for metabolites of interest and executed by the comparison between metabolite of interest and its authentic standard. The availability of authentic standards (ideally, also isotopically labeled standards) facilitates absolute quantitation of metabolites using multipoint calibration curves for every target metabolite. A less labor-intensive, semi-quantitative approach is also frequently used in targeted metabolomics studies and utilizes a single, class-specific internal standard (IS) for quantification of multiple metabolites. This approach relies on the similarity of ionization behavior between metabolites of the same class and is frequently used for the quantitation of lipids.[97] However, different retention times between target analytes and the class-specific standard may result in different matrix effects and reduce quantitation accuracy.[98,99]

1.2.3 Quality assurance and quality control

Quality assurance (QA) addresses and recommends the activities the laboratory should undertake to provide confidence that quality requirements will be fulfilled, whereas quality control (QC) describes the measures that are used to minimize factors that negatively influence the data quality (**Figure 1.3**).[100] QA includes activities such as laboratory personnel training, regular instrument maintenance, instrument calibration, and introduction of standard operating procedures and other quality management systems into the laboratory, including QC samples.

Among those are QC samples, which are used to control the performance of a method and equipment and the correction of analytical variability in data.[100] Aside from the analysis of QC samples, QC activities include monitoring of contaminants via analysis of blanks (collection, extraction, mobile phase), system suitability testing, and control of known sources of irreproducibility via the use of replicates (biological and technical). Therefore, QC analysis provides a significant amount of information, which requires appropriate data visualization. Control charts and/or principal component analysis (PCA) speed up the analysis of system stability and help to detect systematic and random variability by visualization of samples or metabolites within the space scaled by these graphs. [100,101]

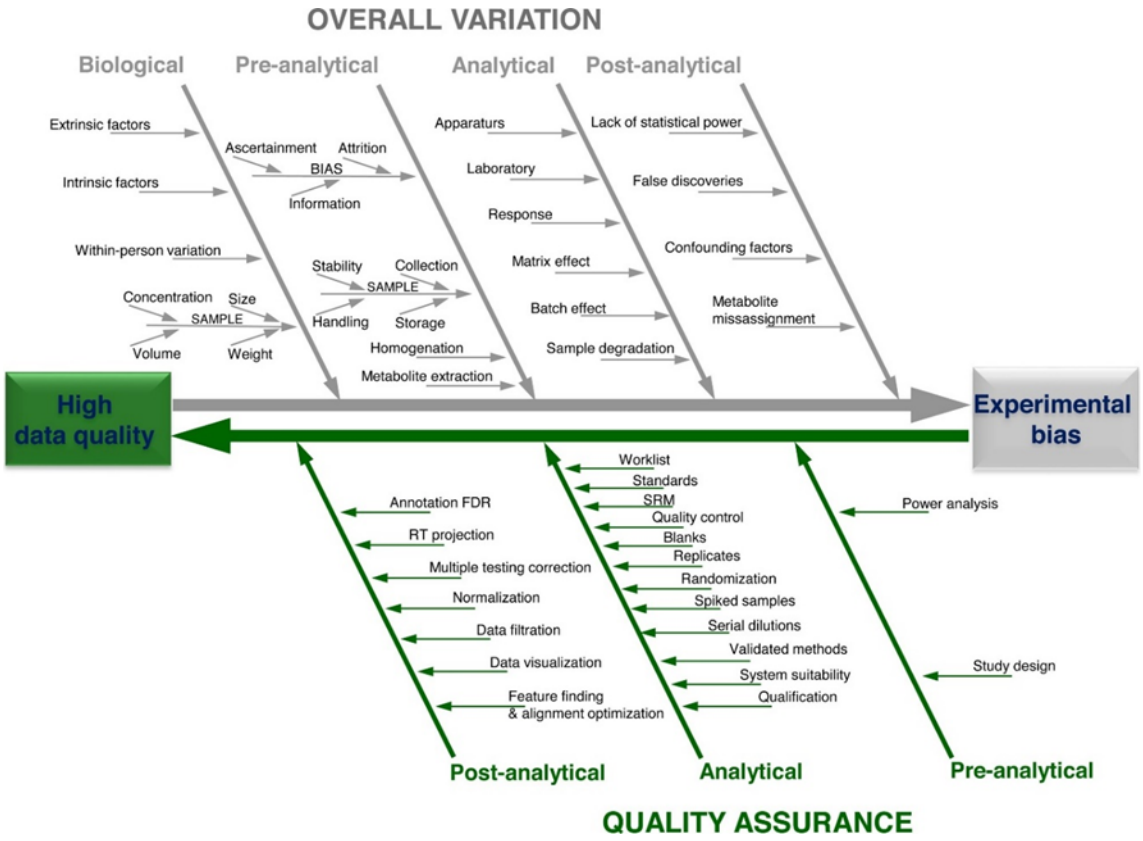


Figure 1.3. Quality assurance, quality control, and sources of variability n the context of the data quality in metabolomics analysis. Re-printed from Dudzik et al., 2018 with permission of Elsevier.[100]

Quality assurance and quality control are similar in many aspects between targeted and global metabolomics studies. However, unlike in global metabolomics, targeted metabolomics methods can be fully validated according to regulatory guidelines to establish their accuracy, precision, and limit-of-quantitation. Although there are currently no specific guidelines for metabolomics, the guidelines of the U.S. Food and Drug Administration (FDA) for bioanalytical method validation are often employed.[102] Another difference used in the targeted approach is the necessity to prepare and inject samples representing a calibration curve together with analyzed samples. In addition, during each targeted metabolomics analysis, additional QC samples are used beyond what is used for global methods to ensure method accuracy.

In particular, FDA guidelines recommend the preparation of QC samples in the matrix identical to the calibration curve at three concentrations (low, mid-range, and high) from stock standards different than the one used for the calibration curve. The frequency of injection of QC samples

may also be adjusted as targeted methods tend to be shorter than global methods. For both global and targeted metabolomics approaches, QA and QC procedures provide tools and practices, which are used to control, correct and accept the dataset prior to submitting it to statistical analysis or biological interpretation.

1.2.4 Statistical analysis

In the simplest approach, applicable when comparing only a few metabolites, the results of relative metabolite quantitation can be evaluated via univariate statistical methods such as the t-test. The testing of multiple hypotheses must be corrected by adjusting p values using correction algorithms such as Bonferroni, false discovery rate (FDR), or others.[103] Volcano plots may be used to visualize metabolite fold-changes and significance, whereas analysis of variance (ANOVA) can be used for comparison of multiple groups. However, these univariate approaches are better suited for hypothesis-testing in targeted metabolomics.

For global metabolomics or targeted methods that cover a large number of metabolites, multivariate statistical approaches are preferred. The selection of individual differential metabolites may be executed using multivariate methods such as PCA and partial least square differential analysis (PLS-DA). PCA is an unsupervised multivariate method and applied to the full dataset. PCA reduces data dimensionality by constructing two or three (or more) orthogonal variables named principal components and allows the visualization of the data and its variability on multidimensional graphs. PCA aids to reveal inter-cohort and inter-sample variability and detect outliers amongst samples and to make decisions whether an analytical batch is suitable for further data analysis and interpretation.[100] In contrast, PLS-DA is a supervised multivariate method, which aims to maximize covariance between independent (signal readings, for example) and dependent (groups) variables to find subsets of exploratory variables, which are subsequently used to build prediction or classification models of dependent variables.[104] In other words, PLS-DA may define groups of metabolites, which can classify samples to groups relevant to the studied biological or clinical phenomena with the advantage of higher specificity and selectivity but requires a heteroscedastic distribution of variables in groups, which is rarely observed in biological experiments.[105,106] These biomarkers should be further verified for the ability to differentiate cohorts using receiver operating characteristic (ROC) analysis. The ROC analysis is a graphical tool that calculates and shows how the true positive rate (i.e., sensitivity)

changes with the false positive rate (i.e., 1−specificity) in a graphical manner. Observation of a sharp increase in the true positive rate versus a minimal increase in a false positive rate is a desirable shape of a ROC curve, and the area under the curve should be as close to 1 as possible.[107] Therefore, a typical outcome of a properly designed, executed, and analyzed study at this stage is just a handful of metabolites. These differentiating metabolites are then prioritized for their identification.

1.2.5 Identification

The identification of metabolites in MS-based metabolomics requires the spectra of ions generated by ESI as described in **Section 1.2.2.1.2.1**. The ions produced from the original molecule form an isotopic pattern (also called "envelope") consisting of (M, M+1, M+2, etc.) ions with m/z and signal strengths dictated by the charge, adducts, elemental composition of a molecule, and the relative natural abundance of isotopes of each atom that comprises the molecular formula. Therefore, the molecular formula can be elucidated from that the envelope based on HRMS's provision of accurate m/z, relative intensity, and spacing of isotope peaks from each other and provide a putative metabolite identification via database search. This information can then be used to obtain tentative matches of an unknown compound to possible candidates in public and commercial databases or "level 3" identification.[108] More confident identification designated as "putative" or "level 2" identification can be achieved in the absence of authentic standard by comparing MS^1 and fragmentation (MS/MS) data of unknown compounds to reference spectra in external or in-house spectral libraries.[5]

Further, the possible existence of many isomers with the same or similar MS/MS spectra may require additional experiments such as ion mobility separation (IMS), spectroscopy, or NMR. Therefore, a minimum of two independent and orthogonal data sets confirming identical structural properties between an unknown compound and authentic standard analyzed under identical experimental conditions are required for a confident identification, which is designated as "validated" or "level 1".[109] Identification at levels 1 and 2 is required for confident interpretation of the experimental data in the context of biological processes and pathways.[5]

The effort spent on the identification is determined by the goal of metabolomics analysis. In biomarker discovery metabolomic studies, statistical analysis is performed at first, and only the resulting differentiating metabolites are then identified or attempted to be identified using existing

databases and/or authentic standard confirmation. On the other hand, global pathway analyses require the identification of as many metabolites as possible in order to provide better pathway coverage and reliable biological interpretation.[28]

1.2.6 Biological interpretation of the data

A comprehensive, computerized biological pathway interpretation of untargeted metabolomics datasets is still under development and currently requires the compilation from several databases. The most widely used ones are public: Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, Cytoscape in MetScape. These can be queried by an assignment of a putative metabolite(s) to a pathway in a given model organism.[93,110,111] In addition, a web-based tool such as MassTRIX can be used for a simultaneous identification of metabolites and their annotation to to pathways using metabolite MS/MS fingerprints for the search in HMDB, KEGG and LipidMAPS.[112] Moreover, chemical class ontologies can be computed for metabolites using Medical Subject Heading ontology (MeSH) or chemical entities of biological interest (ChEBI) by the European Bioinformatics Institute. The chemical ontology may provide a hint regarding the enzymes, that interact with a given metabolite and thus, aid in hypothesizing on the metabolite's biological role.[113,114]Also, metabolites can be assigned to biological and gene ontology terms by querying large repository databases such as UniProt. However, this process is manual and tedious because not all metabolites have yet been linked to gene ontology terms.[115] Finally, metabolites can be assigned to pathways or biological processes using literature searches or by querying a metabolite repository database such as HMDB.[24] The resulting biological function/pathway assignments are used to evaluate and group putative biomarkers according to their biological function and to increase the confidence for hypothesis formulation and planning of further biological or clinical studies.[106]

1.3 **Key challenges in global metabolomics studies**

Several challenges inherent in global metabolomics approaches can reduce data quality and hamper the successful integration of metabolomics data into the systems biology network. These major challenges include: (i) accurate and precise quantitation; (ii) high-confidence identification and biological interpretation and (iii) incomplete metabolome coverage. In a routine study, the data quality depends on the selected analytical method(s), study design, and data variability of

biological and experimental origin. The experimental variability can be assigned to the various stages of a metabolomics study at which it occurs: (i) pre-analytical; (ii) analytical and (iii) post-analytical.

1.3.1 Accurate and precise quantitation

1.3.1.1 Biological variability

Biological intra- and inter-individual variability occur prior to the pre-analytical stage of a study as a result of the influence of intrinsic biological factors, in particular, genetic, epigenetic, gender, age and disease status.[116] External factors such as circadian rhythms, medication or supplement intake, personal habits, and/or compliance to the fasting regime, to name just a few can also alter metabolite levels. For example, a rapid increase in blood concentrations of various molecules can be observed in blood samples if they are collected too soon after a meal or the administration of drugs.[116] Strict control of the time of sample collection is also critical for many metabolites due to large (3-4 fold) circadian changes in blood or oral fluid.[117–119] The biological variability can also arise due to other intrinsic biological factors, including genetic, epigenetic, gender, age, and disease stages.[116] All these factors may produce a sporadic and unpredictable change in concentrations of metabolites or occurrence of matrix effects in individual samples. Preventative measures can be employed at the pre-analytical and analytical stages of the study (study design, sample acquisition, sample, and data analysis) to minimize the influence of biological factors.[119]

1.3.1.2 Pre-analytical variability

To counteract the effects of high biological variability, it is important to determine the appropriate size and patient selection criteria for study cohorts, which can be estimated with preliminary knowledge (literature, pilot studies) of data variability[120] or by using permutation analysis.[121] Study design should include detailed and objective parameters for the cohort stratification (clinical diagnosis or clinical tests), patient enrollment (age, sex, therapy, etc.), and patients' behavior (fasting, synchronization of visit time to circadian rhythms). These measures reduce uncontrolled differences in behavior, physiological, and health status of study participants and increase the detection rate of potential biomarkers.

The variability during sample collection and sample pre-processing (steps that typically occur before the sample arrives at the analytical laboratory) may arise from the inconsistent execution of

withdrawal and pre-processing or due to the absence/poor compliance with the standardized operating procedures for sample collection. For example, differences in serum incubation times (0.5 and 6 hours) resulted in significant changes in signals of 225 metabolites.[122] Another source of variability is the inconsistent usage of blood sampling tubes, i.e., underfilling of tubes, usage of tubes from different manufacturers or different lots, or even with different anticoagulation compounds. This may cause the appearance of inconsistent artifacts (different anticoagulants and/or contaminants) or induce variations in signals of endogenous metabolites. For example, heparin, which is a large molecular weight anti-coagulant, can be easily distinguished from metabolites on the one hand but increases MS background signals across a wide range of masses.[67,123] Heparin may be replaced by potent anticoagulants, including citrates or EDTA. These compounds chelate calcium ions and, therefore, quench the activity of many Ca-dependent enzymes (inhibiting the clotting cascade), and citrate also acts as an anti-oxidant.[124] A disadvantage of both EDTA and citrates is the similarity of their masses to metabolites and relatively high concentrations. These factors may cause ion suppression in their chromatographic elution window and an increase in data variability.[6] Furthermore, citrate is an endogenous metabolite, so the use of this anti-coagulant impedes the measurement of this metabolite Hemolysis is another frequent source of variability which occurs due to traumatic venipuncture, vigorous shaking of a blood tube, and/or high centrifugation forces. Hemolysis has been reported to affect signals of up to 21% of plasma metabolites.[125]

Finally, QC at the preanalytical phase should aim to: (i) monitor the compliance of actual operations to standard operating procedures (SOP); and (ii) assess and document sample quality status/appearance (hemolysis, lipemia, and icterus) and remove samples of poor quality as described in the Canadian "The visual assessment guide".[126] However, instrumental methods are needed for the assessment of a low degree of hemolysis, such as the estimation of free hemoglobin.[127] In addition, lipemia is known to interfere with regular clinical assays, and few instrumental methods for its assessment are available.[116] Furthermore, the elevation of triglycerides and other lipids in a lipemic sample can diminish the quality of LC separation and reduce column lifetime, but the frequency of this occurrence was not reported in the literature to the best of my knowledge. Regarding bilirubinemia (icterus), an isotope dilution LC-MS method may be used as the reference method for the assessment of interference of bilirubin on clinical creatinine tests.[128]

The sources of variability for oral fluid are similar to blood, with a few exceptions. First, anticoagulation agents are not used for pre-processing, which decreases the risk of artifact detection associated with anticoagulant or antioxidant use. However, collection devices that are frequently used for the collection of oral fluid could introduce polymers and contaminants. Second, oral fluid is in direct contact with the environment, which leads to the introduction of exogenous metabolites from food or air (smoking). Third, the composition and concentrations of metabolites in saliva and oral fluid vary significantly between patients due to individual differences in salivary production and health status.[54,129] Finally, saliva has a high concentration of glycoproteins, which form a clot in a collection device and interfere with sample manipulation. This clot can be removed by centrifugation prior to the extraction, or the sample can be extracted as-is. These factors need to be controlled by including into the study the analysis of collection blanks, rigorous control of patient selection and behavior, as well as the normalization of salivary volume production. To the best of my knowledge, the latter has not yet been established.

The efforts to preserve sample integrity continue throughout the whole process and into sample handling during storage, where low temperatures are the preferred approach to preserve sample integrity over the long-term. In addition, the integrity of samples can also be affected by repetitive changes in their physical status, which occur during the freeze-thaw process.[66] Therefore, continuous handling of samples on ice and their storage at or below -70°C in small aliquots to avoid the need for repeated freeze/thaw cycles are strongly recommended practices for all metabolomics studies.[67,125]

1.3.1.3 Analytical variability

Analytical variability includes all steps of sample preparation and LC-MS analysis. The most complex sample manipulations are executed at this stage. Moreover, all QC samples, which were generated in the preceding stages, undergo an actual measurement at this stage. The combination of these two activities requires a meticulous execution of protocols and the use of robust and repeatable sample extraction and analysis methods.

1.3.1.3.1 *Variability during sample extraction*

The extraction procedure itself is the most vulnerable step to operator errors due to complexity and lack of automation. A solvent-based extraction is one of the simplest and most robust methods and is frequently used for global metabolomics. It relies on the relative insolubility of plasma proteins

in organic solvents. Proteins precipitate, forming a solid phase that can be removed by centrifugation, while metabolites remain in the liquid phase. In this workflow, inconsistent metabolite recovery may occur due to (i) inconsistent co-precipitation with proteins; (ii) non-specific adsorption to the surface of the extraction vessel; (iii) poor solubility in the extraction solvent and/or reconstitution solvent; (iv) saturation of solvent by excessive metabolite(s) concentration and (v) inconsistent execution of quenching measures.

Therefore, the extraction procedure should: (i) have comprehensively characterized analytical performance for the sample type; (ii) incorporate blanks, standards, quality control, and reference materials in parallel with processed samples (iii) be reproducible and have a minimum number of steps; and (iv) incorporate metabolism quenching measures to ensure metabolite stability. These measures preserve sample integrity, control the analytical performance of methods, and help to reduce/avoid the introduction of artifacts. They also provide a means to trace back/troubleshoot the extraction process.

1.3.1.3.2 *LC-MS variability in untargeted metabolomics*

The accuracy of relative quantitation in untargeted metabolomics is not generally evaluated. This stems from a lack of authentic standards for all metabolites, poor compatibility of procedures with all classes of metabolites, a large dynamic range of metabolite concentrations, and poor availability of reference materials, to name just a few. Another merit of quantitation, method precision, can be measured if multiple replicates are used. The accumulation of variability from both pre-analytical and analytical (sample handling, preparation, and LC-MS) phases should be considered and segregated, if possible, for troubleshooting purposes.

The variability in data can be caused by constant, repeatable factors and appear as a systematic inaccuracy, which may form a pattern (bias). Examples of such causes include repeatable instrument and sample instability, and consecutive batch effects. The analysis of a sample batch, or a sub-batch (created by dividing sample set if the throughput of analytical methods is limited), can be regarded as the continuous process of analysis of a group of samples under similar analytical conditions executed within a particular time frame. In this case, differences and drifts in retention times, signal intensities, and MS spectra may occur within and between batches across the analysis time and cause inter-and intra-batch signal drift. In untargeted metabolomics, systematic biases can be detected and corrected by the repeated analysis of pooled QC or sample replicates and by

the usage of IS spiked into samples prior to analysis.[130,131] In addition, random error(s) will also occur as a result of sporadic events. There is a particular assumption in the segregation of systematic and random errors, because the "random" is attributed to the repeatable occurrence within observer's perception limits, for example in time or sensitivity. By other words, random error may occur repeatedly but at the observation level appear only randomly. Both, hidden and real random error do not form a particular pattern and may be caused, for example, by a carry-over or uncontrollable LC or MS drifts due to sampling or run-specific factors.[132]

Therefore, in order to eliminate/minimize systematic and decipher random errors at the analytical phase, several concurrent strategies are employed and rely in general on the QA measures, which were discussed in **Section 1.2.4** and metabolism quenching recommendation (**Section 1.3.2.1**). The measures which are specifically relevant to LC-MS analysis include control of contamination (analysis of blank mobile phases) and ensuring signal precision: (i) external and internal MS calibration; (ii) assessment of instrument suitability; (iii) randomized injection order to minimize systematic error and reduction in the size of analysis batches; (iv) regular re-analysis [76] of QC samples throughout analytical batch; (v) use of IS spiked before LC-MS and (vi) control and elimination of a carry-over.[76,133]

The acceptability thresholds for the signal precision are not yet agreed upon throughout the metabolomics community for non-targeted metabolomics studies. The acceptable values for QC samples vary between publications from 20%. to 30% relative standard deviation (RSD).[76,97] For lipidomics, the data quality threshold proposed in the community-initiated white paper by *Burla et al.* is a maximum CV of 20% for pooled QC sample analysis.[134]

Finally, the analytical stage provides an ultimate point to acquire metrics necessary for the identification and troubleshooting of sources of the variability, evaluation of signal repeatability (intra- and inter-day, for example), correction, and elimination of low-quality metabolites or the rejection of the whole dataset.[2,135] In global metabolomics, metabolite quantitation, in the absence of authentic standards, is executed in a relative manner for all of the thousands of metabolites. The detection of a differential biomarker depends critically on ratios of peak areas for a given metabolite across all samples, and a low method variability is crucial for the success of biomarker discovery. In particular, accuracy of biomarker levels across cohorts may be impacted by poor signal linearity, detector saturation, and/or differences in inter-patient matrix effects and may lead to false biomarker detection.

1.3.1.3.3 *LC-MS variability in targeted metabolomics*

In contrast to relative quantitation, absolute quantitation in targeted metabolomics requires the control of both accuracy and precision. One of the most difficult tasks in the preparation of calibration standards is the compensation for composition differences and matrix effects between calibration points and study samples and the fact that metabolites are endogenous analytes. Therefore, it is often not possible to obtain blank biological samples that do not contain the metabolites of interest. In order to use matrix-matched calibration and provide the range of concentrations below the smallest concentration in all study samples, targeted endogenous metabolites need to be removed. This can be executed by various methods, including solvent precipitation, (LLE) and SPE (charcoal stripping, for example). However, these processes are not specific to targeted metabolites and significantly modify the composition of calibration standards compared to samples.[136] To avoid this problem, a standard addition calibration method is often used, but is extremely labor- and time-intensive. In this method, calibration point(s) are created by adding known increasing amounts of analyte(s) to the aliquots of the sample of interest in order to build the calibration curve, thus keeping the matrix similar between calibration points and study sample.[137] Finally, IS calibration is common in metabolomics, especially in targeted workflows. The use of stable isotope-labeled (SIL) IS, where one or several hydrogens, carbon-12, or nitrogen-14 in a molecule are replaced by deuterium, carbon-13, or nitrogen-15, respectively) can help to compensate for matrix effects and improve method accuracy and precision. Ideally, the molecular weights of isotopologues should differ in mass ≥ 3 Da to prevent an overlap of their isotopic envelopes, inaccurate measurement of ion signals and erroneous quantitation. Calibration with SIL IS relies on the introduction of a constant amount for each metabolite of interest into samples and calibration curve points before extraction. The intensities of SIL IS are used for the correction of signal differences caused by recovery or matrix differences and/or LC-MS signal drift over time.[138,139] This approach is also known as stable isotope-dilution. The use of SIL IS compensates well for matrix effects between samples because RT and ionization behavior are similar between SIL IS and its analyte due to similar physical-chemical properties. However, in RP chromatography, RT shifts have been observed for SIL IS containing a high number of deuteriums.[139] Even a slight difference in RT may result in the onset of different matrix effects between the analyte and corresponding SIL IS with the subsequent errors in quantitation .[140]

However, the replacement of deuterated SIL IS by the ones with carbon-13 or nitrogen-15 (which do not demonstrate RT shifts) is feasible but, expensive and not always available solution.

In sum, the selection of the quantitation approach requires detailed attention to analytical performance and must be implemented with sufficient controls. The FDA Guidance for Bioanalytical Method Validation[102] provides a set of measures which ensures accurate and reproducible quantitation during targeted LC-MS analysis: (i) calibration curves and QCs should be included in all analytical runs; (ii) total QCs should number at least five percent of the total samples analyzed, or be at least six in number (low-, mid-, and high-QCs, in duplicate), whichever is greater; (iii) the minimum number of non-zero calibrator (a calibration point with spiked IS) levels ≥ 6, and (iv) experimental concentrations, in at least 67% of all, and 50% of each level QC samples should be ± 15% nominal (theoretical) concentrations except of the low limit of quantitation (LLOQ), where the calibrator should be ± 20% of nominal concentrations in each run. Repeatability with a coefficient of variation (CV) of 15% or less is acceptable for non-zero points, except for concentrations at the lower limit of quantification (LLOQ), where a CV of 20% or less is acceptable.[102,141] Moreover, the execution of these stringent merits could be applied for absolute quantitation of multiple metabolites simultaneously, as it was demonstrated in several studies reviewed by Cajka and Fiehn.[76] However, the labor for the creation of a multitarget calibration curve, material cost (mostly SIL IS), and difficulties to maintain optimal separation requirements on LC limit progressively the maximum number of targeted metabolites that can be used for simultaneous absolute quantitation. On the bright side, a multi-targeted metabolomics LC-MS assay does not require metabolite identification and generate less complex data files for processing in the post-analytical phase.

1.3.1.4 Post-analytical variability

Unlike a targeted metabolomics analysis, global metabolomics experiments lead to highly complex data sets and require proper data handling, which determines the quality of the metabolomics results. As the origin of the observed signals cannot be validated by IS, it is a challenging task to differentiate MS (for example, solvent contaminants, background signal noise, in-source fragments) and chromatography or peak-picking (for example, peaks with RT shifts, peak shoulders, baseline noise) artifacts from real featured signals which represent a particular metabolic feature including adducts, isotope, and multicharged ions. The data analysis in targeted

metabolomics is facilitated by the relatively small number of analyte targets and the use of IS and authentic standards, which validate the identity of signals in study samples. This also facilitates automation of the data quantitation. However, it is a challenging task to establish a robust automated procedure. The reasons for this are similar to those observed in a global approach, i.e., the distinguishing of real signals from MS and LC artifacts, especially at the lowest signal intensities.[142]

The consistent assignment of metabolite ions is required by quantitation procedures in both metabolomics approaches. Signals from isotopes, adducts, and multicharged ions, are grouped together and assigned to a metabolite. Therefore, the inconsistent assembly of metabolite ions across samples, decreases the reproducibility of relative quantification. Subsequently, special attention should be paid to the execution of the robust data analyses and data quality (filtering noise, peak picking, deconvolution, spectra alignment, missing values imputation), before biological interpretation and identification.

1.3.2 High confidence identification and biological interpretation

Unlike nucleic acids and proteins, metabolites are not constructed by the linear assembly of a limited number of building blocks (4 nucleic acids and 20 amino acids, respectively). This makes their MS-based identification difficult. Metabolites (with some exceptions for lipids) do not possess such uniformity of building blocks and can be extremely chemically diverse, which complicates MS-based identification. In addition, the possibility to link a gene and corresponding encoded protein(s) allows assignment and computational investigations of biological functions and pathways. On the other hand, the assignment of biological roles for metabolites is complicated by the lack of comprehensive systematic relationship of the metabolites to their genes or proteins, pathway redundancy, metabolite localization, which may affect its role, and the ability of several proteins/enzymes to interact with a single metabolite. The challenges in an identification of metabolites and interpretation of their biological roles are discussed below.

1.3.2.1 Use of MS spectra to assign molecular formulas

The success of metabolite identification depends on the quality of the acquired MS spectra, which is determined by resolving power, accuracy, and sensitivity of instruments. For example, the assignment of the mass of 382.0832 Da (belonging to celecoxib, which contains one atom of sulfur) using MS^1 at the resolution of 50,000 FWHM can be narrowed down from 37 formulas (at

five ppm accuracy) to five formulas (at one ppm mass accuracy). Moreover, a resolution above 50,000 FWHM can separate M+2 isotope peak of celecoxib into two peaks (one with ^{13}C and another with ^{34}S). This confirms the presence of sulfur in the celecoxib and reduces the number of formulas from five to two but this level of analysis requires manual curation. If the analysis is executed at the resolution 500,000 FWHM (achievable on advanced Orbitrap and Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometers), the software automatically resolves individual isotopes of celecoxib to baseline and provides fine isotopic structure including low abundant ^{15}N and ^{18}O isotopes, that eliminate the redundancy of formula assignments.[143] To further improve molecular assignments, accurate measurements of isotope abundance are also beneficial.[144] The accuracy of isotope ratio calculations deteriorates at low signal intensities due to the increased errors in measurements of mass and relative isotope signal ratios or complete disappearance of signals from M+n isotopes. This problem is prominent for metabolites with low abundance, low recovery and/or low ionization efficiency, especially in complex samples, and can be partially resolved via sample decomplexation approaches.

Another difficulty encountered during assignment of molecular formulas is the propensity of ESI to form various adducts and in-source fragments in addition to protonated and deprotonated ions of precursor metabolite. The ESI process can result in re-arrangements and adduct formation, all of which dramatically expand the number of potential formula candidates and the ambiguity of formula assignments in database searches. A differential metabolic labeling with stable isotope-labeled nutrients can be used for cell or plant metabolomics to distinguish background ions from true metabolite signals. For instance, the isotopic labeling outlier algorithm (IROA)[145] searches for specific isotope signature pairs in the data. Any metabolites of biological origin will be present as a pair, and the difference in mass between the pairs can also help narrow down molecular formulas. Any background ions which do not have the correct isotopic pair pattern are removed from further processing.

Overall, despite great strides in improving mass accuracy and resolution of MS measurements, it is not feasible to routinely run large metabolomics studies on ultra-high-resolution instruments as these acquisition methods are still too slow to comply with peak width in ultra-high pressure liquid chromatography (UHPLC). This means that MS^1 formula assignment with current metabolomics workflows is not always accurate and should not be used alone to infer metabolite identity. For biological matrices of clinical interest such as a plasma or oral fluid, SIL labeling strategies such

as IROA are not available, thus increasing the likelihood of identifying background ions as possible metabolites. The confidence of MS-based metabolite annotations can be improved by including RT as an orthogonal parameter. However, a poor LC method synchronization between different laboratories, and minor differences in chromatographic conditions limited its usage for the creation of in-house LC-MS libraries, and strict QC practices are used to ensure retention times have not drifted significantly. However, the recent progress in the development of RT prediction approaches for RP and HILIC provides a promising possibility for the comprehensive incorporation of this parameter into metabolite identification in untargeted metabolomics. While an optimization of prediction algorithms and protocols of integration of RT into metabolite identification are still under debate [146] the immediate and clear advantages of such integration with existing tools such as Mass Frontier and MetFrag have been already demonstrated.[147] Moreover, Witting *et al*.[146] concluded that the most promising approach would integrate ion RT predictions for RP and HILIC based on training sets of metabolites and machine learning algorithms with the possibility to apply the trained models to new separation methods and systems. The advantage of the approach is in the constantly increasing availability of metabolomics data supplemented with RT, LC method information and structural properties in metabolomics repositories. This approach would address some of the inherent issues with RT variability between laboratories and systems. However, a wide-spread integration of RT into elucidation of metabolite identity still remains elusive.

1.3.2.2 Use of MS/MS and MS^n spectra to assign putative metabolite identifications

The correct assignment of molecular formula does not provide the molecular structure (i.) and must be supplemented by MS/MS, MS^n and/or other analytical techniques. MS/MS data can also help narrow down molecular formula assignments when more than one molecular formula fit the acquired MS^1 data. The MS/MS spectrum of a precursor ion provides fragmentation information, which helps deduce atom connectivity in the molecule. For best results, the spectrum should originate from a presumed precursor metabolite isolated at the highest purity. In a complex sample, this is achieved by filtering out any ions except the ones which fit into the m/z window. This window is limited to approximately 1 Da on modern MS instruments. This minimizes (but does not eliminate) the risk of simultaneous co-fragmentation of different precursors and improves the authenticity of the MS/MS spectrum and the ability to match it correctly to the precursor library

spectrum. The use of the same collision energies and same instrument model as used for library spectrum acquisition also facilitates high-confidence matching. The match, in particular, is defined by the quality of MS/MS spectra, i.e. purity, fragment m/z, number of expected fragments presented and signal intensity of fragment ions. The MS/MS methods are constantly improving via the development of new acquisition and interpretation approaches mentioned in **Section 1.2.2.1.3**. Depending on the instrument configuration, targeted MS/MS for global metabolomics may require two runs: the first run to acquire m/z and RT of molecules of interest, and the second to execute MS/MS analysis. In the absence of target information, an untargeted simultaneous MS and MS/MS analysis in a single run is possible in current MS instruments with fast scan time or two independent mass analyzers via DDA. However, DDA analysis over a wide m/z range provides MS/MS data for only 48-57% of detected metabolites in serum due to the complexity and bias of the DDA algorithm towards the most intense ions.[88] A recent study comparing DDA and DIA reported an increase of the coverage by DIA, while DDA outperforms DIA in the quality of fragmentation spectra.[148]The use of a gas-phase fractionation (GPF), and especially staggered GPF (sGPF) improves the coverage of the DDA method for untargeted metabolomics. Unlike a conventional GPF, which uses a single m/z range, sGPF uses several narrower m/z segments (for example, low, middle, and high ranges) for the simultaneous DDA. The fragmentation coverage increases by > than 80% with the use of sGPF over a conventional DDA.[88,89]

Poor quality of a MS/MS spectra may preclude automated search in spectral libraries. This issue can be partially resolved by obtaining spectra at different standardized collision energies that match the energies of library MS/MS spectra, such as 10, 20, and 40 V used to build the METLIN database. Multistage (MS^n) fragmentation of the precursor of interest and its fragments for the easier elucidation of metabolite composition and structure can also aid correct metabolite identification.[149] Besides, the quality of fragmentation spectra and coverage of metabolites by fragmentation data deteriorates at high sample complexity, which emphasize again the importance of the use of decomplexation methods. Even when good quality MS/MS spectrum of an unknown metabolite is obtained, this metabolite may not be matched to any entries in a particular library. In addition, poor standardization of quality merits for MS/MS spectral matching and insufficient assessment of the false detection rate (FDR) between multiple libraries and publications continue to slow the success of the multi-target search.[3] However, this aspect of metabolomics analysis has greatly improved over recent years due to extensive efforts towards the development of spectral

libraries. Moreover, the use of computational *in silico* fragmentation algorithms such as calculated fragmentation trees combined with machine learning and automated database searches may help to assign putative metabolite identity, but this is still an active area of research.[150]

Finally, the analysis of MS/MS or MS^n spectra may provide bonding patterns of atoms and distinguish structural isomers. However, the differentiation of stereoisomers, despite an occasional detection of subtle differences in fragmentation patterns,[151] is beyond routine MS capabilities. Therefore, comprehensive structure characterization requires the application of multiple techniques, including infrared (IR), NMR, UV-Vis spectroscopies, and various separations (chiral LC, differential ion mobility, CE).[143] In conclusion, low concentrations of some metabolites and lack of preconcentration in standard global metabolomics workflows may preclude the generation of quality MS/MS spectra. It has been shown that approximately 60-70% of all features detectable in MS^1 do not generate good quality fragmentation spectra.[152] As such, I would speculate that the development of enrichment extraction protocols for particular metabolite classes together with specific depletion of highly abundant compounds could provide a promising approach to increase the number of metabolites with usable MS/MS spectra and/or to aid orthogonal characterization such as NMR which requires large quantities of unknown metabolite.

1.3.2.3 Biological interpretation

The simultaneous presence of compounds originating from exogenous, endogenous, and microbiome sources, their instability, inter-individual biological variability together with the analytical variability and identification issues complicate the biological interpretation of metabolomics data and the assignment of biological roles and pathways.[153] Metabolomics is still a relatively young discipline, so the development of tools for the computational investigation of biological and pathway roles of metabolites is still under development. In addition, from the systems biology point of view, metabolites are the subjects of a pool of enzymatic reactions, from which more than 2000 remain to be characterized in humans.[154] Therefore, extensive development of computational tools integrating the knowledge of metabolic networks with the identification and structural characterization of metabolites appears to be the most promising strategy to achieve the full integration of metabolomics into systems biology.[106]

1.3.3 Incomplete metabolome coverage

The metabolomics analysis of human blood plasma and serum is complicated by the wide concentration range and wide chemical diversity of metabolites. The concentrations of metabolites in human serum/plasma span over 9 orders of magnitude (i.e., g/L to ng/L), whereas a non-targeted metabolomic analysis is limited to intermediate to high concentration (> 10 μg/L) metabolites except for those with high ionization efficiency. A wide dynamic range severely limits the potential of untargeted metabolomics analysis of blood for biomarker discovery because many active and regulatory metabolites such as norepinephrine, epinephrine, and dopamine, are present at very low (sub-μg/L) concentrations.[28,155] Another consequence of metabolome complexity and dynamic range is that high amounts of metabolite ions can saturate the detector and affect isotopic abundances, which may result in erroneous metabolite identification and quantitation. This issue can be minimized via sample dilution or the filtering of the ion flow on its way between the ionization source and the detector.[156] However, such approaches will also reduce the detection of low abundance metabolites.

The total number of metabolites present in plasma is high, although the exact number of metabolites is unknown. However, the total number of all compounds in the HMDB (known, expected, and predicted) exceeds 114,000.[24] Chemical properties of blood plasma metabolites cover all ionic classes. For example, the non-lipid metabolites set available in HMDB consists of 34% acids, 14% bases, 30% zwitterions, and 22% neutral molecules.[82] In addition, metabolites also differ by their polarity (**Figure 1.4**), and their predicted octanol/water partition coefficients span across 40 orders of magnitude.

For lipids, 37% are acids, only 2% bases, 31%- zwitterions, and 30% are neutral compounds.[82] The total lipid content of blood plasma (phospholipids, neutral lipids, total cholesterol and cholesterol ester of fatty acids) was reported between 0.4 to 0.7% (w/v) with 33% of this total corresponding to phospholipids[157] More recent estimate of total lipids provides similar values between 0.19 and 0.41% with triglycerides comprising between 0.1 – 0.2% (w/v).[158,159]

Notably, lipids span over 80% of the LogP range and overlap with non-lipid metabolites, which occupy approximately the polar quarter of the range (**Figure 1.4**). This distribution predicts the co-extraction of non-lipid and lipid metabolites by solvent-based extraction protocols and their co-elution in chromatographic separations.

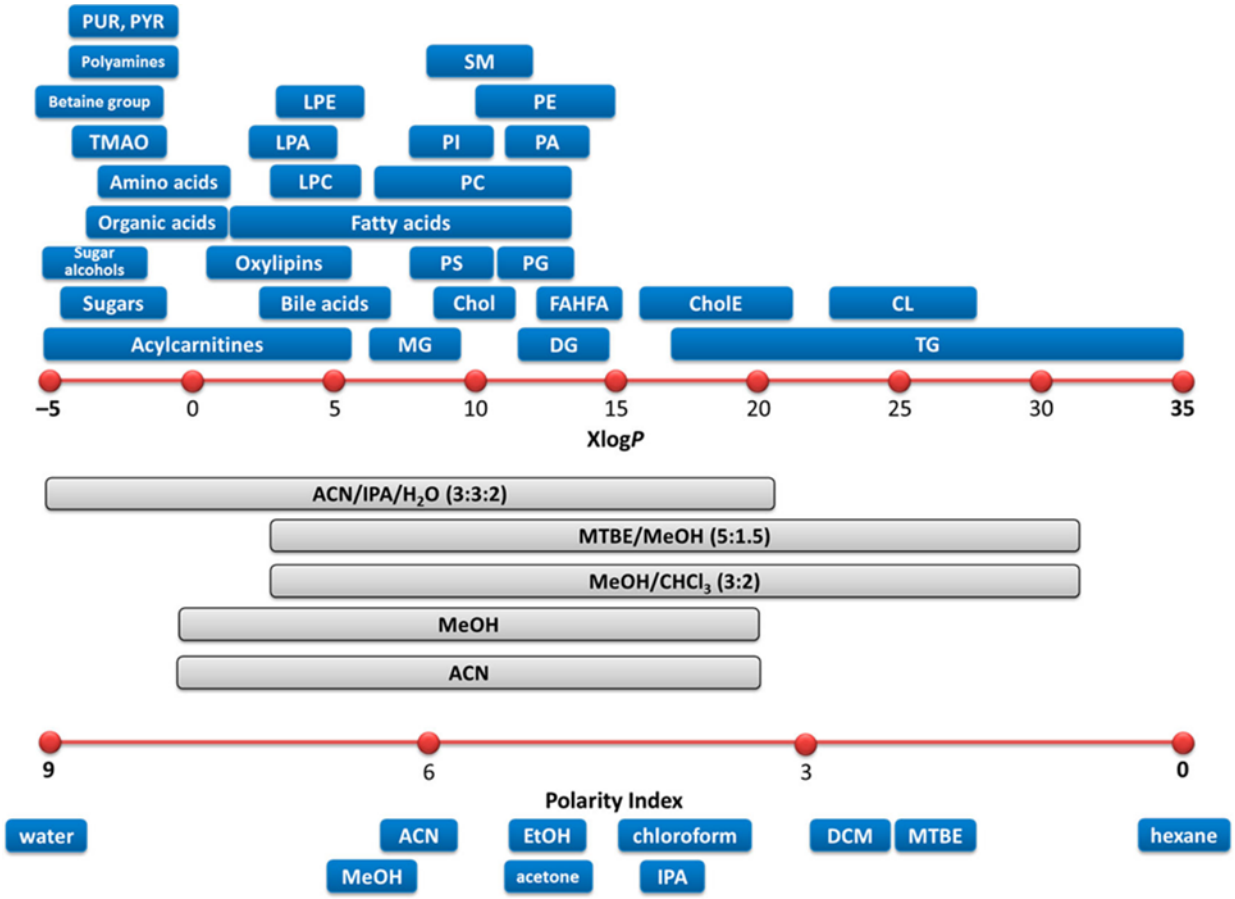


Figure 1.4. Predicted octanol/water partition coefficients (X log P) of metabolites. Reprinted with permission from Cajka et Fiehn.[76] *Copyright © 2016, American Chemical Society.*

One major consequence of such co-elution is its impact on the ESI process and the increased likelihood of matrix effects for various metabolites.

Metabolomic analysis of biofluids is not always limited to blood-derived samples, and alternative sample sources such as oral fluid can supplement or replace the analysis of blood. Untargeted salivary metabolomics studies for biomarker discovery have been executed by NMR and MS.[49,160] In general, the oral fluid appears to be less complex than plasma, which is also supported by the smaller number of organs in contact for oral fluid compared to blood. For instance, combined analysis using NMR, DFI-MS, LC-MS/MS, LC-ultraviolet(UV)/fluorescence spectroscopy, GC-MS, and inductively coupled plasma ionization (ICP)-MS resulted in the quantification/identification of 308 salivary metabolites.[57] A second study combining only GC-MS and LC-MS untargeted platforms identified 370 out of 475 total metabolites found.[29] An HMDB[24] query at the end of January 2021 reported 1245 compounds detected in saliva, 885 of

which have been quantified. However, Zheng *et al.*[161] recently demonstrated that a more sensitive approach using chemical derivatization followed by LC-MS analysis resulted in drastic increase of metabolite coverage of oral fluid and revealed a plethora of low abundant metabolites, that were previously undetectable. These findings imply the possibility of using oral fluid to substitute or supplement the analysis of blood with several possible practical advantages. For example, free concentrations of two steroid hormones, cortisol, and cortisone in plasma are more difficult to quantify than in oral fluid because they are close to the limit of detection of the targeted metabolomics analysis. In addition, their concentrations in oral fluid correlate very well with their free concentration in plasma and are informative of the systemic status of an individual such as circadian cycle, which can be used for the QC of sampling, time and for the supplementary assessment of the stress status of an individual.[162,163] Very frequently, enzyme-linked immunosorbent assays (ELISA) are used for such supplementary measurements due to high sensitivity. However, this assay has a high cost per sample and can overestimate corticosteroid concentrations.[164] Finally, the differences in the composition of plasma and oral fluid may result in different matrix effects, whereby one type of biofluid may be preferred for the analysis of a subset of metabolites.

From the metabolite coverage perspective, due to the use of different instruments and data analysis algorithms and subsequent different in-source fragmentation, de-isotoping, de-adducting, peak picking, integration, the direct comparison of metabolite coverage numbers across studies is precluded. The incorporation of stable isotopes into metabolome via metabolic labelling of yeast or *E.coli* cultures (credentialed metabolites)[165] helps to distinguish biologically derived metabolites from contaminant or exogenous ones. but requires a direct side-by-side comparison.

Moreover, the ionization of some analytes, especially low abundance metabolites and metabolites with poor ionization efficiency, may become so suppressed that they are no longer detectable in some samples or some biofluid types. Even if the suppression is not complete, such low-intensity signals may demonstrate too low signal-to-noise (S/N) ratios and be impossible to distinguish from background ions and thus un-processable for automatic retrieval. The decrease of matrix effects via sample decomplexation, the manipulation of matrix effects via changing the composition of a sample, or manipulations of the selectivity of ionization source can be used in metabolomics analysis to improve metabolome coverage.

1.3.3.1 LC strategies to increase metabolome coverage

The coupling of CE, LC, or nano-LC to MS decreases the complexity of the sample entering the ESI interface and can increase metabolome coverage to 10^3 versus 10^2 compounds typically observed by direct flow injection (DFI)-MS. An additional improvement of metabolome coverage is achieved via replacement of HPLC by an UHPLC, which can drastically increase resolution and decrease co-elution phenomena using particle sizes < 2 µm. Proper selection of stationary phase-type and the switch to smaller particle size can together drastically increase metabolome coverage. Two UHPLC methods, CORTECS T3 with C18 phase and CORTECS C8, each with solid core particles of 1.6 µm, provided the detection of 679 and 879 credentialed features, respectively, in extracts from *E. coli* compared to 173 features detected on xBridge™ RP C18 column containing 3.5 µm porous particles.[166] Another frequently-used approach to increase the metabolome coverage is the re-analysis of the same extracts on the highly complementary RP and ZIC-HILIC LC-MS. Contrepois *et al.* compared metabolite coverage during LC-MS profiling of plasma samples on five HILIC and five RP columns.[81] The study reported the combined non-redundant detection of 5188 and 4739 of metabolic features in (+ve) and (-ve) ESI, respectively. Individual LC-MS analyses generated 2956 and 3088 in (+ve) ESI for ZIC-HILIC and RP, respectively. In (-ve) ESI, 3611 and 1913 features were detected on ZIC-HILIC and RP, respectively. The best RP results between tested columns were observed on Zorbax SB column using 0.06% acetic in water (mobile phase A) and in MeOH (mobile phase B). The best ZIC-HILIC results were achieved using 10 mM ammonium acetate in 95/5 water/acetonitrile (mobile phase A) and 5/95 water/acetonitrile (mobile phase B). Therefore, the parallel analysis of complex samples on RP and (IEX)-RP or RP and ZIC-HILIC columns, in both (-ve) and (+ve) ESI, is currently considered the best approach to achieving good metabolome coverage in an untargeted analysis.

1.3.3.2 MS approaches to increase metabolite coverage

In addition to the separation strategies discussed above, improving the efficiency of ionization can also be used to increase metabolite coverage. Thus, MS analysis in (+ve) and (-ve) ESI is an easy and productive approach to increase metabolome coverage, as some analytes ionize preferentially in one mode.[167] In addition, metabolite coverage can be increased using nano(-ve) ESI, the drastically scaled-down (nebulized diameter 1-50 µm, flow rate 1-100s nL/min) version of conventional ESI (100 µm diameter and 10-100s µL/min flow rates).[168] Due to such tiny

dimensions, the droplet diameters are 100–1000 fold smaller than in a conventional ESI and produce a drastic (up to 500 times) increase in the number of ions being formed and entering the mass analyzer. Other strategies can focus on optimizing the composition of the mobile phase and executing analysis with other ionization methods that provide an orthogonal (at least partial) selectivity, such as atmospheric pressure chemical ionization (APCI)[169,170] or atmospheric pressure photoionization.

Other promising strategies in untargeted metabolomics include the addition of an ion mobility separation (IMS) between LC and MS dimensions.[171] While the increase of metabolite coverage via advanced development of MS methodologies is on-going, the improvement of the coverage via advanced sample preparation is another viable option.

1.3.3.3 Sample depletion approaches to increase metabolite coverage

Sample preparation can reduce complexity and the amount of sample entering the ESI source. This can be achieved via dilution (widely used for urine analysis, not plasma), the removal of abundant compounds, and/or fractionation of metabolites by their physical and chemical properties with the subsequent comprehensive LC-MS analysis of every fraction. However, until now, methanol-based extractions of biofluids are the most popular due to high reproducibility (RSD < 20%) and recovery with the detection of 1808 - 2073 molecular features in LC-MS analysis.[18,172–174] However, the excessive redundancy of metabolite compositions in different extracts results in a disproportionally small increase in the metabolome coverage compared to the labor and MS time expenditure. Thus, the metabolome coverage of acetone plasma extract, the most orthogonal out of other nine (including methanol, ethanol, and acetonitrile) tested in the study, adds only 12.8% of new features.[175] On the other hand, successful removal of phospholipids using Ostro (Waters, Milford, MA) and Phree (Phenomenex, Torrance, CA) SPE plates improved sensitivity and repeatability, decreased matrix effects but, unfortunately, was also associated with a loss in recovery for many polar metabolites.[167,172] Similar observations were made when phospholipid removal was achieved using SPE with polymeric or mixed-mode, polymeric-ion-exchange (IEX) phases,[96] and zirconium-based Hybrid SPE® (Sigma-Aldrich, St. Lois, MO) SPE phases.[176]

1.3.3.4 Sample fractionation to increase metabolite coverage

A significant increase in metabolite coverage has been demonstrated by combining sample fractionation with LC-MS analysis. Thus, one study reported the detection of 4,264 molecular

features in deproteinized plasma sub-fractionated into polar, phospholipid, and lipid fractions by SPE, compared to 1,792 features detected in a methanol extract alone.[177] In another study, the analysis of methyl tert-butyl ether (MTBE) polar and SPE fractionated (phospholipids, fatty acids, neutral and hydrophobic lipids) non-polar fractions resulted in the detection of 3,806 compared to 1,851 molecular features on methanol precipitation alone.[174] For non-lipid metabolites, the analysis of SPE fractions enriched in cation and anion metabolites obtained from deproteinized and delipidated (phospholipid removal) fish plasma demonstrated the possibility of at least four-fold increase in sample loading compared to methanol extracts, an increase in signal intensity up to 40-fold for many spiked exogenous metabolites, a fold increase in signal-to-noise ratios and a signal variability below RSD of 30% for 62% of the metabolites.[96] The numbers of molecular features detected in cationic and anionic SPE fractions were similar to each other. However, PCA analysis revealed significant differences between these two extracts, thus indicating potentially orthogonal composition and the possibility to increase metabolite coverage if both methods were run in parallel. Unfortunately, only the bound SPE fractions were analyzed, thus missing any metabolites present in the flow-through fractions. Ultimately, the generation of multiple SPE fractions takes time, which can be reduced via the implementation of on-line SPE-LC-MS.

These examples demonstrate how sample decomplexation can be achieved using fractionation methods with narrow but complementary (preferably orthogonal) selectivity. In general, such methods demonstrate distinct enrichment of compounds with specific properties. Therefore, the number of metabolites in each fraction is reduced compared to a whole sample, and analysis could reveal additional metabolites, which were not found in more complex extracts. Moreover, if a composition of these fractions is complementary to each other, their parallel analysis may recapitulate or even exceed the coverage of a method with wide selectivity. In addition, it is a practical solution, which allows customizing analytical methods to suit the composition of the fraction to the desired goals or make them more compatible with analytical methods. Although promising, these recent developments of sample preparation methods for metabolomics lack the systematic assessment of multiple methods that have good potential for the increase of metabolome coverage.

Another successful way to expand metabolome coverage is chemical derivatization. The approach employs reagents, that derivatize reproducibly specific functional groups of metabolites, including amines, carboxyls, phenols, hydroxyls, carbonyls, and thiols with a subsequent LC-MS and

MS/MS analysis.[178] Several derivatization agents used in the approach (dansyl chloride, 10-methyl-acridone-2-sulfonyl chloride) improve analytical sensitivity several orders of magnitude (0.04 nM for neurotransmitters, for example)[179] thus greatly enhancing metabolite detection in a typical metabolomics study (**Section 1.1.1**). Moreover, this approach can provide other important advantages: improvement of chromatographic separations, applicability for different biofluids and tissues, enhancement of molecular ion fragmentation, and provision of relatively inexpensive SIL IS for semi-quantitative and absolute concentration measurements.[178,180,181] Thus, derivatization is a promising approach for combined global-targeted metabolomics analysis. In particular, the enhanced sensitivity expands the metabolome coverage of low abundance or poorly ionizing metabolites, if a reproducible, uniform and comprehensive derivatization can be established.

In conclusion, the major challenge which drove the current study was the limited coverage and sensitivity provided by canonical sample preparation and analyses currently used in global metabolomics. The effect of sample preparation on data quality and metabolite coverage in global metabolomics has not been fully evaluated, especially when considering more selective sample preparation methods such as SPE and LLE.

1.4 **Thesis objectives**

Metabolome complexity of human blood plasma, or other biofluids, cannot be comprehensively covered using a single untargeted metabolomics analytical method. In LC-MS based untargeted metabolomics research, the current most promising strategy to improve metabolome coverage is a repetitive analysis of the same sample using three LC approaches tailored to polar (IEX-RP or IEX-HILIC), mid-polar (RP), and lipid (RP with isopropanol-based mobile phases) using both ESI modes. Recent data shown in **Sections 1.3.3.3 and 1.3.3.4** suggests that a sequential usage of traditional and untraditional methods may successfully increase metabolome coverage and quality of data. However, the systematic assessments of promising sample preparation protocols have been limited mostly to solvent-based methods.[167,172] There is currently not enough known about the quantitative performance of more selective methods such as SPE and LLE for global metabolomics. This systematic knowledge is indispensable for the development of better sample preparation methods with increased metabolite coverage for untargeted analysis and rational selection of the best sample preparation method for a given metabolomics study.

The first objective of this thesis was to perform a side-by-side comparison of three conventional solvent precipitation methods to test the effect of solvent polarity (methanol, methanol-ethanol, and methanol-MTBE), one LLE (MTBE) method to separate lipids from other metabolites, and three post-deproteinization SPE methods (C18, mixed anion/cation IEX, and divinylbenzene-pyrrolidone (PEP2)). The seven methods will be compared according to absolute recovery, matrix effects, repeatability, selectivity, and metabolome coverage using suitable targeted and untargeted metabolomics approaches after analysis on RP and mixed-mode IEX/RP (Scherzo) LC-MS in (+ve) and (-ve) ESI. The resulting data will provide a basis for the rational selection of the best and most complementary extraction and LC-MS methods. This addresses an important knowledge gap about the quantitative performance of sample preparation methods for global metabolomics and is described in **Chapter 2**.

The second objective of this research was to build on the findings of **Chapter 2** to design a new sequential SPE-based sample preparation method for global metabolomics. First, it is necessary to develop the protocol of sequential de-lipidation, deproteinization, and SPE IEX fractionation to produce separate fractions enriched with anion, cation, zwitterion, and neutral metabolites. The performance of this sequential preparation was to be assessed for metabolome coverage, fraction splitting, recovery, matrix effects, and repeatability using targeted and untargeted metabolomics strategies. The analysis of fractions is executed on RP and ZIC-HILIC, each on (+ve) and (-ve) ESI, as the latter method outperforms the mixed-mode method used in **Chapter 2** for polar untargeted metabolomics. Secondly, analyzing all fractions on all four methods is unnecessary as many metabolites may be detected in both ESI modes or using both chromatographic methods, creating highly redundant datasets. Thus, it is necessary to optimize and integrate sample preparation and LC-MS analysis schemes to obtain the highest metabolome coverage in the minimal number of analytical runs. This new methodology increases the coverage of the polar and mid-polar metabolites while decreasing sample complexity in a predictable fashion, as described in **Chapter 3**. These advantages make it compatible not only with global metabolomics workflows but also for targeted metabolomics assays, whereby this new protocol saves time and labor for the development of targeted sample preparation methods during transition of biomarker(s) assessment from untargeted to targeted metabolomics analysis.

The final objective of the current work aims at the possibility of using saliva as an alternative sample source to substitute or supplement metabolomics analysis of blood. For this preliminary

study, a targeted metabolomics method was set up to monitor cortisol and cortisone concentrations in the oral fluid while permitting easy sampling throughout an intervention. Cortisone and cortisol are biologically important metabolites involved in human stress response, and oral fluid has previously been established as a suitable biospecimen for their measurement. Therefore, in **Chapter 4,** I describe the development and validation of sample preparation and a targeted MRM LC-MS assay for these two hormones in oral fluid. The performance of the final validated method for cortisol was compared side-by-side with the commercial ELISA method, which is traditionally used for this analysis. This example targeted analysis provides critical knowledge on details and limits of metabolomics analysis of oral fluid, with a particular focus on quantitative method performance and the most suitable calibration strategies to use.

2 Systematic assessment of seven solvent and solid-phase extraction methods for metabolomics analysis of human plasma by LC-MS

2.1 Introduction

The objective of global metabolomics is to analyze all small-molecular-weight species (≤1,500 Da) in a biological sample.[1] LC-MS is currently the method of choice for global metabolomics studies because it provides the highest metabolite coverage using a single analytical technique. Typically, several hundred to thousand(s) metabolites can be detected in a single analysis.[182] The size of the human metabolome is currently unknown but is projected to exceed the conservative estimate of 4229 endogenous metabolites in concentrations spanning 11 orders of magnitude.[61] The most recent estimates predict 8500 endogenous metabolites,[183] and up to 40,000 additional exogenous metabolites, such as drugs, additives, and toxins that may be present in human samples.[184] Considering the typical coverage of untargeted metabolomics analysis, it is clear that metabolome complexity is overwhelming the capacity of modern metabolomics methods. Therefore, new strategies to increase metabolome coverage are required.

The most widely used protocol for global metabolomics of plasma is solvent precipitation using cold methanol or methanol/ethanol (1/1, v/v) with a plasma-to-solvent ratio of 1 to 4.[174,185,175,186] Cold solvent is added to minimize the extent of enzymatic conversion of metabolites and to precipitate the proteins. The removal of proteins from plasma also prevents protein build-up on the LC column, which improves LC column lifetime and significantly increases the number of detected metabolites through disruption of protein binding and minimizing the number of signals originating from proteins.[18] Methanol and methanol/ethanol are the solvents of choice due to high metabolite coverage, as shown by several studies.[172,175,185] However, the wide selectivity of such solvent-based precipitations results in highly complex samples, which precludes the detection of lower abundance metabolites. A liquid-liquid extraction using methyl-tertbutyl ether (MTBE) has become a popular alternative in recent years for its ability to provide good coverage of both polar and lipid metabolites and compatibility with robotic systems.[187,188] In contrast, solid-phase extraction (SPE) methods are often avoided in global metabolomics of plasma due to their increased selectivity compared to methanol-based extraction methods. SPE methods, thus, tend to decrease overall metabolite coverage[18] but may improve data quality through improved repeatability [18,189] and reduced matrix effects.[167,172] For instance, optimized HybridSPE™

successfully removed phospholipids in order to lower matrix effects while maintaining acceptable recoveries and repeatability.[18] In order to increase metabolome coverage beyond what can be achieved with methanol-based precipitations, multiple extraction methods with narrow but complementary selectivity can be combined in a sequential manner. The approach demonstrated 2-fold increase in lipidome coverage.[177,190] However, no similar sequential extraction approaches exist to date for non-lipid metabolites. To systematically design such sequential extraction protocol(s), it is necessary to directly compare coverage of various solvent precipitation, LLE, and SPE methods. However, only a limited number of studies compared solvent precipitation methods to SPE and LLE to date.[18,172,190] Based on these published evaluations of extraction methods in real samples, a few limitations should be highlighted. None of these studies examine the orthogonality of SPE and MTBE methods to methanol-based methods in a side-by-side fashion, and comparison across the studies is not possible due to the different instrumentation and data processing strategies used. Most of these studies focus on metabolome coverage and extraction repeatability only, and no simultaneous evaluation of matrix effects and recovery in the biological matrix has been performed to date. Recovery studies are crucial in order to design sequential extraction methods that are fully orthogonal and do not split the signal between multiple fractions. In addition, semi-quantitative comparisons of metabolite signal intensities between extraction methods can be misleading because variations in analyte signals due to matrix effects are not properly taken into account using the addition of SIL analytes[139,191] fully isotopically-labeled complex matrices, or standard addition calibration. The latter approach was successfully employed to monitor and compare the absolute recovery of sequential extraction by hybrid and mixed-mode SPE in untargeted metabolomics.[96] The underappreciated advantages of the standard addition method become obvious when comparing different extraction methods. It is well-established that the slopes of calibration curves for biofluids originating from multiple populations can show significant differences.[192] Similarly, different matrix effects are expected in samples originating from extraction methods with different selectivity. In such cases, signal intensity changes may be driven by matrices alone, leading to erroneous conclusions regarding the extraction performance. Although matrix effects are extensively studied in targeted bioanalysis, this issue has not been addressed in global metabolomics except in one study where a post-column infusion experiment was performed to identify the region of significant ion suppression.[172] However, anecdotal evidence across multiple comparison studies shows potentially significant matrix effects with huge

differences in signal intensity observed when using different extraction methods.[96,189,193] Therefore, the quantification of absolute recovery and matrix effects using a systematic set of standard analytes when evaluating multiple extraction methods is missing from comparisons to date, leaving a critical gap in our knowledge.

Following an extraction, the most frequently used LC separation in global metabolomics is the parallel use of C18 (RP) chromatography and HILIC to achieve good coverage of non-polar and polar metabolome, respectively.[186,194] More recently, mixed-mode chromatographic materials combining RP and ion-exchange mechanisms in low-bleed MS-compatible stationary phases provide improved retention of a broad spectrum of metabolites.[195,196] The major objective of this study was to perform the first side-by-side comparison of three conventional solvent precipitation methods to test the effect of small changes in solvent polarity (methanol, methanol-ethanol, and methanol-MTBE), one LLE (MTBE) method, and three post-deproteinization SPE methods based on cartridges (C18, mixed cation-anion exchange (IEX) and divinylbenzene-pyrrolidone (PEP2)) for LC-MS metabolomics of human plasma (**Figure 2.1**). Absolute recovery and matrix effects for standard analytes were evaluated for all seven extraction methods using the standard addition method. The repeatability and selectivity/orthogonality of extraction methods were compared using both targeted metabolites and on a global basis in combination with RP and mixed-mode IEX/RP (Scherzo) LC-MS. These data pave the way for the rational selection of the best and most complementary extraction methods of the human plasma metabolome and clearly show the effect of using multiple sample preparation methods in a given study design on metabolome coverage.

2.2 **Materials and methods**

2.2.1 Solvents and reagents

LC-MS grade solvents/mobile phase additives and analyte standards were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless stated otherwise. ACTH (1-39) was obtained from Anaspec (Fremont, CA, USA). Norepinephrine (d_6), cholic acid (d_4), epinephrine (d_3), dopamine (d_4), melatonin (d_4), 4-aminobutanoic acid (d_6), and phenylalanine (d_5) were obtained from CDN Isotopes (Point-Claire, QC, Canada), while $^{13}C_6$- thyroxine was purchased from Toronto Research Chemicals (Toronto, ON, Canada). MTBE was bought from Fisher Scientific (Toronto, ON, Canada), while all phospho-lipids were obtained from Avanti Polar Lipids (Alabaster, Alabama, USA).

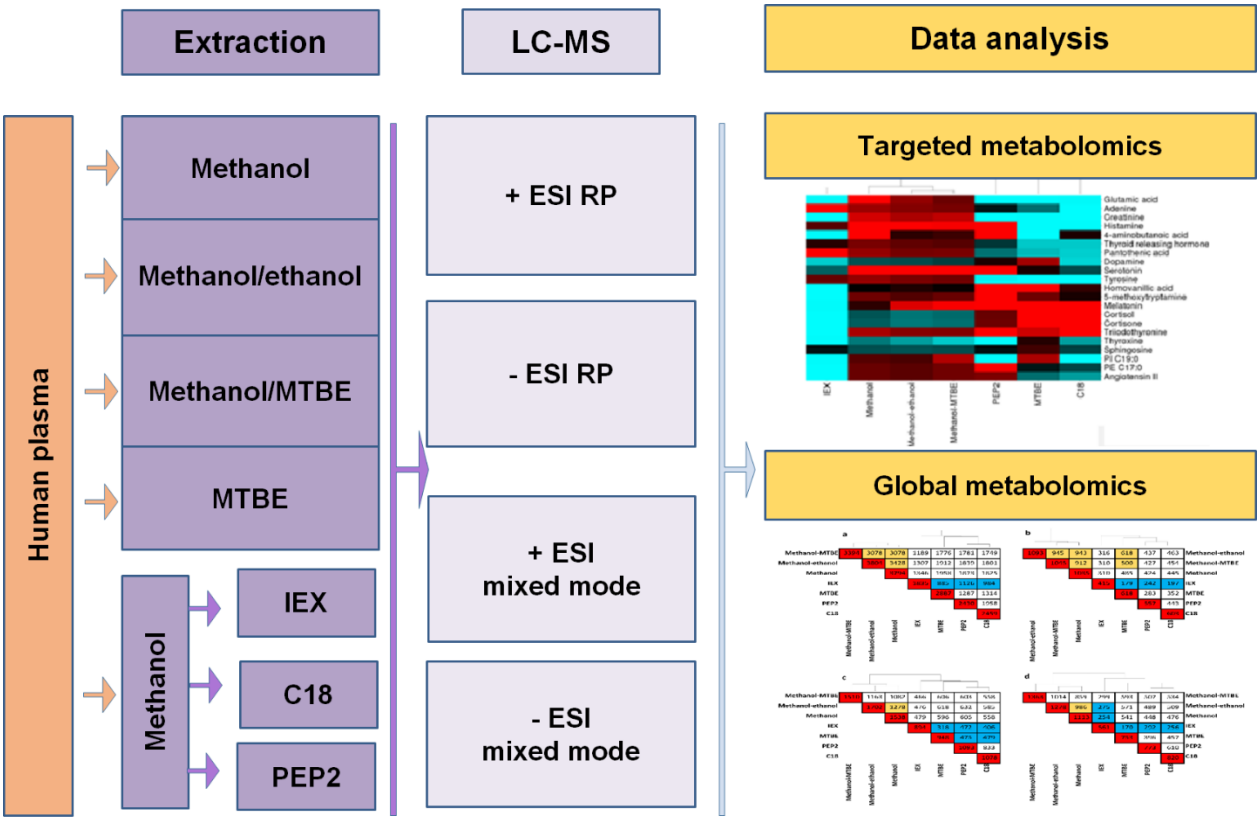


Figure 2.1. Overview of experimental design to compare seven extraction methods for untargeted metabolomics analysis of human plasma.

Kynurenine and D-erythro-sphingosine (further mentioned as sphingosine for brevity) were purchased from Cayman Chemicals (Ann Arbor, MI, USA). Solid stationary phases (PEP2, ODS-C18, and divinylbenzene conjugated with sulfonic acid and quaternary amine moieties (IEX)) were obtained from Agela Technologies (Wilmington, DE, USA. Citrated pooled human plasma was obtained from Bioreclamation (Baltimore, MD, USA), which was collected in accordance with the company's code of ethics. All reagents were of an analytical or higher grade.

2.2.2 Standard analyte mix

The chemical diversity of the metabolome is enormous both in terms of polarity and charge.[76,81] Using predicted octanol/water partition values, metabolites in human plasma cover a polarity range from - 5 (polyamines, amino acids) to 10 (fatty acids) to 35 (triacylglycerides).[76] For the charge state, metabolites can be separated into acidic, basic, neutral, and zwitterion classes. For instance, the study analyzing charge properties of 2553 non-lipid human metabolites from the Human Metabolome Database had found that approximately 22% of metabolites are neutral, while 46.5

and 18.2% contain acidic carboxylic and phosphate groups, respectively. Basic aliphatic amines and aromatic heterocyclic nitrogen groups were found in 16 and 24.5% of non-lipid metabolites, while 13.8% of compounds were zwitterions.[82] The focus of current work is non-lipid metabolites, so standard metabolite selection was confined to metabolites with high to intermediate polarity typically found in blood plasma. A few lipids were also included in the mix to help in the assessment of matrix effects and method selectivity towards lipids, but systematic evaluation of extraction performance for lipids was beyond the scope of this study. Therefore, we evaluated extraction methods using standards with a limited but systematic set of chemical properties that (i) resembled class composition of target samples and included acids, bases, neutrals (steroids), zwitterions, lipids, and small peptides, (ii) was systematic and scalable in terms of chemical properties (Log P range of -3.9 to 11.5, MW range of 105 to 900 Da), and (iii) amenable to the RP and mixed-mode LC-MS analytical methods employed in the study. All individual stock solutions were prepared in appropriate solvents as summarized in **Appendix A, Supplementary Table A1**, divided in aliquots and stored at below -70 ºC, while working standards were prepared at appropriate concentrations prior to analysis. Standard mix was prepared at 5 µg/mL from appropriate stock solutions using 20% methanol unless otherwise specified.

2.2.3 Extraction of standard analytes from a buffer

The standard mix was prepared at 5 µg/mL of each compound in 25 mM ammonium acetate, pH 6.5 buffer. This high concentration was required to avoid non-specific adsorptive losses. Buffer composition was selected to obtain suitable pH and ionic strength for IEX stationary phases in order to achieve maximum recovery of analytes while ensuring MS compatibility. The standard mix was extracted in six replicates by solvent-precipitation (methanol/ethanol (1/1, v/v), methanol, methanol/MTBE (1/1, v/v), LLE (MTBE) and solid-phase extraction (PEP2, C18, IEX). In solvent precipitations and LLE, 100 µL of the standard mix was extracted with 400 µL of ice-cold solvent, vortexed for 30 min, and centrifuged for 15 min at 15000x g. All steps were executed at +4 ºC. After centrifugation, 350 µL of the upper layer was dried and stored at below -70°C until analysis. For SPE, 100 µL of the standard mix was loaded on a 3 mL SPE cartridge containing 100 mg (C18, IEX) or 60 mg (PEP2) sorbent. The cartridges were washed with 1 mL of sample buffer and eluted into glass tubes with 1.5 mL of elution solvent specific for every sorbent: C18 with 0.1% formic acid in 100% acetonitrile, PEP2 with 150 mM ammonium acetate, pH 6.8 in 94% methanol

and IEX with 400 mM ammonium acetate pH 6.8 in 42% methanol. Eluted samples were evaporated to dryness under vacuum and stored at below -70°C. Before analysis, all samples were reconstituted in 10 µL of 20% methanol containing 2.5 mM ammonium acetate (pH 6.5), sonicated at ambient temperature for 5 min, vortexed for 10 min, diluted with 90 µL of 2.5 mM ammonium acetate, pH 6.5, sonicated and vortexed for 5 and 10 min and centrifuged at 15000 x g for 30 s.

2.2.4 Extraction of plasma samples spiked with internal standard analytes

Solvent precipitations and LLE were carried out as described above using (i) the sample buffer (composed of 2% acetonitrile in 2.5 mM ammonium acetate (pH 6.5) to obtain a blank extract for each method, (ii) plasma samples spiked with standard analytes to yield approximately 800 ng/mL before extraction, and 100 ng/mL at LC-MS step in six replicates and (iii) un-spiked plasma samples in 12 replicates to be pooled on the per-method basis and used to build calibration curves for each method. Prior to SPE extraction, replicates of plasma and sample buffer were precipitated using methanol as described above, evaporated to dryness, reconstituted in the sample buffer, pooled as appropriate, and divided into replicates equivalent to 100 µL of plasma. Six of these replicates were spiked with the standard analytes at 800 ng/mL per each of three SPE methods. All samples were extracted by three SPE sorbents in parallel following the protocols described above to generate the sample sets similar to the one prepared for precipitation and LLE, i.e., blank extracts (i), spiked plasma extracts (ii), and un-spiked plasma extracts (iii). All samples were dried under vacuum and stored at below -70 ºC.

2.2.5 Preparation of plasma extracts for LC-MS analysis

All plasma extracts were reconstituted in 30 µL of 20% methanol as described for standard analytes and further dissolved in 270 µL of 2.5 mM ammonium acetate. Standard addition calibration curves were prepared for each extraction method by adding 30 µL of the sample buffer or the mix of standard analytes to yield matrix calibration curve with 0 or 62.5, 125, 250, and 500 ng/mL. For the assessment of matrix effect, an external standard calibration curve in 2.5 mM ammonium acetate pH 6.5, 2% acetonitrile was also prepared in the same concentration range.

2.2.6 LC-MS analysis

All extracts (10 µL) were analyzed on 1290 UHPLC chromatograph (Agilent Technologies, Santa Clara, CA) using the 3.0 µm, mixed-mode Scherzo SM-C18, 2 x 150 mm column (Imtakt, Portland, OR) and 1.8 µm Zorbax Eclipse octadecyl 2 x 200 mm column coupled to Agilent 6550 iFunnel Q-TOF mass spectrometer in positive and (-ve) ESI in the mass range 100–1000 m/z. Additional details, including LC-MS settings, are provided in Appendix A.

2.2.7 Data analysis

TOF Quant software (version B.07.00 SP1, Agilent) was used for the determination of absolute recovery of standard metabolites from buffer and plasma. For that, raw data was extracted at 15 ppm mass accuracy, aligned within ±0.15 min retention time, integrated and corresponding adducts (sodiated adducts were found only for melatonin in positive (+ve) ESI were summarized. Otherwise, protonated and deprotonated ions were used for all other analytes in positive and negative (-ve) ESI, respectively. Quantitation was executed using external calibration curves in the buffer and standard addition calibration in plasma. The recoveries of each analyte for each extraction method were hierarchically clustered using the Euclidian distance method with CIMminer online analysis at http://discover.nci.nih.gov/cimminer/oneMatrix.do.[197] The recoveries of analytes below 5 and above 80% were assigned to 0 and 100%, respectively for a correct visualization. Matrix effect was calculated by dividing the peak area of an analyte in matrix calibration standard spiked post-extraction by the area in the calibration standard prepared in the sample buffer at the same analyte concentration and converting it to percentage. The subtraction of endogenous signals was performed using a signal obtained in un-spiked plasma extracts. The final result reported for matrix effect represented the mean value obtained across four different concentrations tested for each analyte. Pooled QC samples for all target analytes in all analytical batches showed RSD ≤ 25%. QC data showed no evidence of analyte degradation in extracted plasma samples except possibly for histamine and sphingosine, both of which showed a systematic 20-30% decrease of signal intensity throughout the long analytical batches. For the global evaluation of the extraction methods, peak picking, deconvolution, alignment, and integration were executed on Profinder (Agilent) with the following parameters: ion mass threshold of ±15 ppm, the relative height of MS+1/MS isotope abundance 15%, RT threshold ±0.15 min, minimum peak height 200 and 2000 counts for M+1 and M peaks, respectively. The selectivity and repeatability

analysis were carried out on Mass Profiler Professional (MPP, Agilent) with integration and binning parameters similar to Profinder, after removal of low-quality metabolite signals that (i) were not at least 5x higher than the signal in blank and (ii) that were not found in at least 5 out of 6 replicates of a given extraction method. The data were manually verified and found to include 2-3% duplicate entries (a feature split between multiple entries by the peak picking algorithm). Therefore, the accuracy of putative metabolite coverage is ±3%. The orthogonality of extraction methods in the global metabolomics approach was evaluated in a pairwise manner using the above high-quality data. A number of matched features for all possible paired combinations were used for hierarchical clustering using CIMminer online tool.

2.3 Results and discussion

Seven different extraction methods are compared based on the absolute recovery of standard analytes from the buffer and human plasma, repeatability, selectivity, and matrix effects in parallel with global LC-MS based metabolomics analysis. The overall experimental design is shown in **Figure 2.1,** and the list of standards in **Supplementary Table A1**. To the best of our knowledge, this is the first side-by-side comparison of the quality of sample preparation from blood plasma by conventional solvent precipitations (methanol-ethanol, methanol, methanol-MTBE), LLE (MTBE), and SPE (C18, IEX, PEP2) methods in a single study.

2.3.1 Targeted analysis

2.3.1.1 Recovery, repeatability, and selectivity of metabolite extraction from buffer

Analyte recovery is summarized in **Appendix A, Supplementary Table A2**. The analytes are listed by increasing logP values retrieved from the ChemSpider database predicted using the ACD Laboratories algorithm. The extraction methods are arranged according to the results of the hierarchical analysis. As expected, methanol, methanol/ethanol, and methanol/MTBE extractions clustered closely together and provided the broadest coverage and the highest recovery across the wide range of metabolite classes tested. IEX provided high recovery only for polar charged metabolites, while MTBE provided high recovery for hydrophobic neutral metabolites. Among SPE methods, PEP2 provided broader metabolite coverage than C18 (**Figure 2.2)** due to its ability to retain some of the polar metabolites.

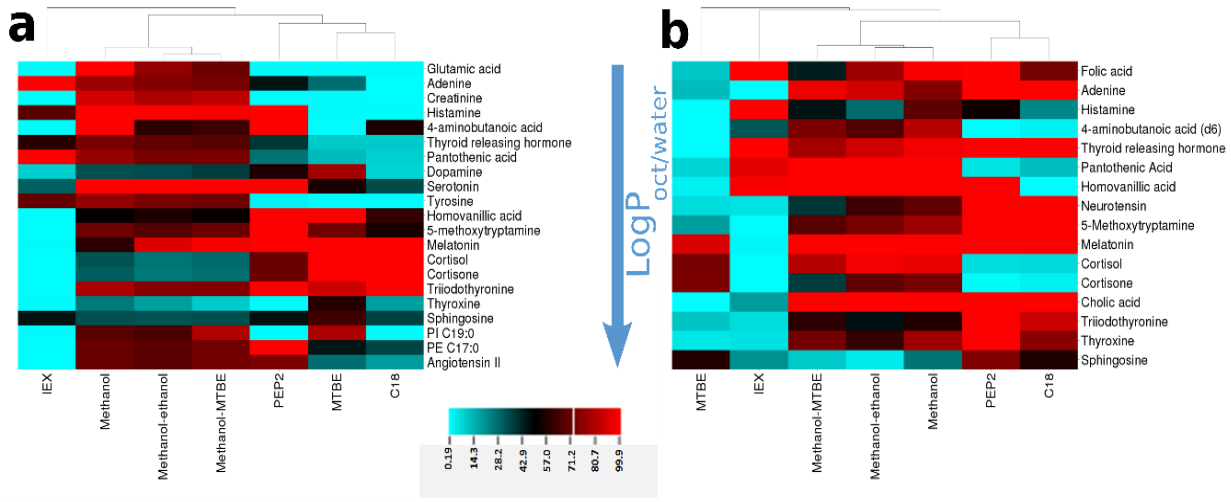


Figure 2.2. Hierarchical analysis and heat maps show the recovery of standard analytes from the buffer (A) and human plasma (B). All extraction methods were hierarchically clustered using the Euclidian distance method. The intensity of each cell represents the range of recovery of an analyte relative to the initial amount spiked prior an extraction. Recoveries below 5 and above 80% were assigned to 0 and 100% for visualization purposes. The order of analytes corresponds to the increase in octanol-water partition coefficients, except for angiotensin II, which did not have a predicted value.

The highest selectivity in buffer was demonstrated by IEX, followed by C18 and MTBE. Moreover, IEX and C18/MTBE methods demonstrated little overlap, which can be exploited in sequential sample preparations for global and targeted metabolomics. None of the tested analytes exhibited ≥ 50% recovery in all extraction methods (**Appendix A**, **Supplementary Table A2**). The **Supplementary Table A6 (Appendix A)** summarizes the main performance characteristics of all extraction methods tested. The recovery of ≥80% is considered exhaustive, and quantitative bioanalytical methods permit method precision of up to ≤20% RSD at a lower limit of quantitation. However, very few metabolites can meet these most stringent criteria for any of the tested methods, as shown in S**upplementary Table A4 (Appendix A**). Global metabolomics methods are considered semi-quantitative, so applying more relaxed criteria of ≥50% recovery and ≤ 30% RSD is a reasonable compromise between method coverage and method performance. Using these criteria, methanol-based precipitations can provide acceptable performance for 17 out of 22 metabolites. Overall, metabolite recovery correlated to the predicted LogP values and the expected selectivity of the extraction methods. Neutral metabolites such as melatonin demonstrated the best quantitative (≥80%) recoveries amongst all standard metabolites. The best repeatability was demonstrated by methanol-based precipitations, while SPE and LLE methods demonstrated lower

repeatability than methanol blends with the poorest performance by MTBE and C18 SPE (**Appendix A, Supplementary Table A3**). This poor repeatability of MTBE is attributed to irreproducible partitioning of some metabolites between organic and aqueous phases and was most pronounced for pantothenic acid, thyroxine, and phenylalanine.

2.3.1.2 Recovery, repeatability, and selectivity of metabolite extraction from plasma

The high recovery (≥80%) was demonstrated by thyrotropin-releasing hormone and melatonin in 6 out of 7 methods tested (**Appendix A, Supplementary Table A4).** Ionic compounds such as histamine, tyrosine, and kynurenine with low molecular weight and LogP values demonstrated quantitative recovery in only 1 out of 7 methods. In addition, better recovery of triiodothyronine, thyroxine, and the large peptide neurotensin (in contrast to tripeptide thyrotropin-releasing hormone) was observed on RP SPE compared to solvent-based extractions (**Appendix A, Supplementary Table A4).** The recovery of some analytes from plasma changed drastically in plasma *versus* buffer. The recovery of neutral compounds (cortisol, cortisone) by MTBE, PEP2, and C18 was decreased in plasma, but the recovery in methanol-based solvents increased (**Appendix A, Supplementary Table A4**). This clearly shows the importance of performing recovery studies in biofluid matrix. The extraction repeatability (**Appendix A, Supplementary Table A5**) showed similar trends in plasma to what was seen in buffer with methanol-based methods outperforming both SPE and LLE methods. However, SPE and LLE methods showed significant deterioration of method precision. For instance, the MTBE method provided acceptable precision (≤30% RSD) for only melatonin, cortisol, and triiodothyronine. Hierarchical clustering results shown in **Figure 2.3** confirm wide metabolite coverage of gold standard methanol-based solvent precipitation methods with high recovery across metabolite classes. The results also demonstrated the selectivity of MTBE towards uncharged species with LogP ≥ 1.4 and confirmed orthogonal selectivity of IEX and MTBE methods previously observed for buffer. This can be used for the removal of hydrophobic compounds in sequential sample preparations. Finally, methanol-based methods provide the most comprehensive and reproducible extraction of standard analytes from plasma as indicated by much lower mean RSD values than observed for other extractions (**Appendix A**, **Supplementary Table A5**). The more selective methods of SPE and LLE show good performance only for a narrow range of metabolites that are best suited to each extraction method depending on their polarity and charge characteristics. In **Supplementary Table A4**

(**Appendix A),** some metabolites show recoveries above 120% in some of the SPE methods. This result was surprising, considering the similar matrix composition of standard addition calibration and unknown samples, so it was investigated further. The first possibility considered was different adduct formation in buffer versus plasma. No such differences were found, so this was eliminated as a contributing factor. Melatonin and melatonin (d_4) both had similar high recoveries in PEP2 and C18 SPE, which pointed to the fact this result may be due to co-suppression of spiked metabolites. The only compositional difference between calibration standards and samples used to evaluate recovery is the number of spiked metabolites present. Calibration standards were spiked after extraction and will therefore contain all metabolites of the standard mix, whereas the recovery samples were spiked before extraction and will remove or incompletely extract some of the metabolites from the standard mix depending on the extraction method selectivity. This was further verified by re-analyzing the same extracts on a longer chromatographic method (60 min analysis time), and proper quantitative recovery (80-120%) was obtained in all instances. These results clearly show that global metabolomics methods are extremely susceptible to matrix effects and that the semi-quantitative performance of these methods can be affected by minor differences in matrix composition.

2.3.1.3 Extraction preferences of standard analytes

Two groups of analytes emerged based on our recovery studies in buffer and plasma (**Figure 2.2**). Analytes with LogP below 0.4 (above tyrosine) show poor recovery in MTBE and good recovery in methods suitable for extraction of polar species such as PEP2 and methanol-based solvent precipitations (Group I). The second group consists of less polar analytes (LogP $\geq$ 0.4), which demonstrate a good recovery in MTBE, PEP2, and C18 (Group II). Interestingly, Group II analytes had recoveries $\geq$ 50% in most solvent and SPE methods in contrast to their recoveries from the buffer. This difference in recovery is attributed to adsorptive losses in the buffer. In the experiments with buffer, we tried to minimize these losses by using high metabolite concentrations, but clearly, adsorptive losses were still considerable, especially for metabolites such as thyroxine, cortisone, and cortisol. In conclusion, a side-by-side systematic comparison of the absolute recovery of extraction methods was possible using standard addition calibration and showed clearly the critical importance of recovery determination in biofluids.

2.3.1.4 Matrix effects

Matrix effects were evaluated using the post-extraction spike method at up to four different concentration levels. This evaluation was performed using both RP and mixed-mode LC-MS methods in positive and (-ve) ESI. Ion suppression was observed for metabolites with lower LogP values (melatonin, 4-aminobutanoic acid, adenine, and homovanillic acid) in methanol-based extractions in RP. Neutral analytes of intermediate polarity with LogP of 1.2 (cortisol and cortisone) were not affected in any extraction method, while signals from more hydrophobic metabolites (LogP ≥1.4, triiodothyronine and thyroxine) were enhanced in all solvent-based extracts, suppressed in IEX and PEP2 and remained unaffected in C18 SPE. The suppression in solvent-based extractions in (–ve) ESI RP (**Appendix A, Supplementary Figure A2, Supplementary Table A7**) affected a wider range of analytes than in (+ve) ESI RP and extended toward neutral mid-hydrophobic analytes (cortisol and cortisone), while organic acids (folic, pantothenic, homovanillic and cholic) remained unaffected. The suppression of polar analytes in RP analysis in both positive and (-ve) ESI is not surprising and could be explained by the co-elution of a large number of un-retained metabolites. However, mixed-mode LC-MS analysis, which is capable of chromatographically separating the majority of these charged species, also shows very significant problems with ionization suppression and/or enhancement depending on the analyte and extraction method tested. Previous studies have also shown that HILIC methods are also highly susceptible to matrix effects even when using microextraction format.[193] Matrix effects could be partially decreased via improved resolution at LC-MS step by the decrease of stationary phase particle size (< 2 μm) or drastic increase of column length and chromatography time to impractical limits.[198] Therefore, there is no simple solution to implement to address this major problem. Finally, methanol-based extracts demonstrated a higher number of analytes affected by matrix effects. In contrast, the more selective MTBE and C18 SPE methods showed matrix effects for fewer metabolites and less pronounced extent of suppression/enhancement if matrix effects were present (**Table 2.1, Supplementary Figure A2**). The higher suppression matrix effect observed for methanol blends is most likely caused by higher matrix complexity as compared to MTBE and C18 SPE methods, which are more selective, as shown by our recovery experiments.

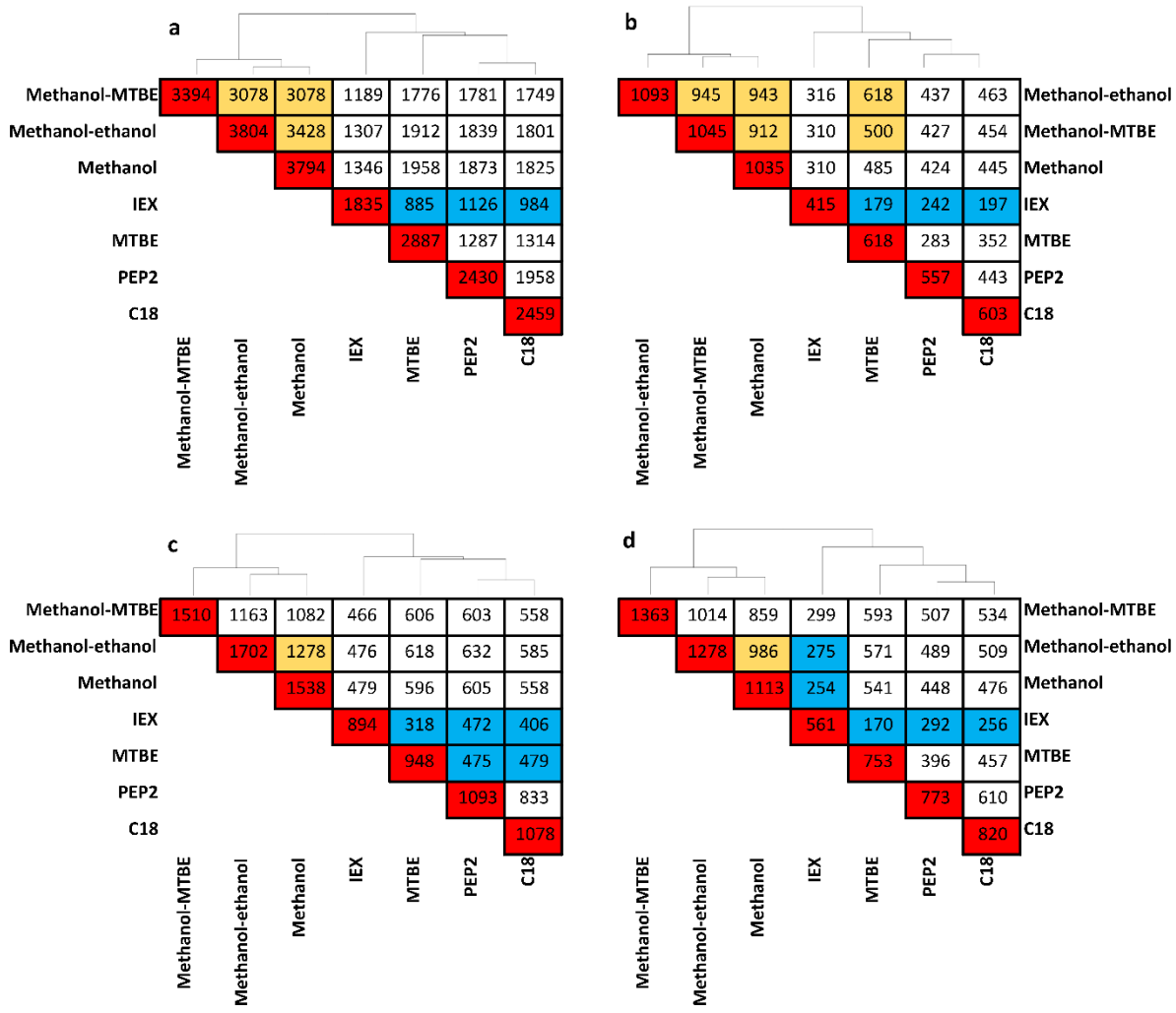


Figure 2.3. Hierarchical clustering of the number of putative metabolites detected in plasma extracts and pairwise overlap coverage of seven extraction methods. Samples were analyzed using (+ve) ESI RP (a) – ESI RP (b), (+ve) ESI Scherzo (c), and – ESI Scherzo (d). Red color boxes located across the diagonal show the total number of putative metabolites detected with that extraction method from plasma. Yellow, white, and cyan blue colors designate high (99.9-80% overlap), medium (79.9-50.0%), and low (50.0 – 0.0% overlap) of putative metabolite populations observed by the two extraction methods specified. Therefore, the methods indicated with cyan blue boxes are the most orthogonal pairs of methods across all of the seven extraction methods tested.

Overall, the results of this comparison show that matrix effects pose a significant challenge in all extraction protocols and can have a significant impact on biomarker discovery efforts. Additional ways to reduce and evaluate matrix effects during such studies should be explored and implemented. **Table 2.1** also shows that the analyses performed using mixed-mode chromatography are more susceptible to signal enhancement. The observed enhancement of signal

response (for example, triiodothyronine and thyroxine in Scherzo analysis (**Appendix A, Supplementary Figure A2**) may be explained by: (i) true matrix effect; (ii) the presence of co-eluting isobaric contaminants whose signals were mistakenly included due to insufficient resolution of QTOF; (iii) formation of a significant amount of adducts in the buffer but not in matrix calibration points and/or (iv) limited solubility of standard analytes in buffer calibration samples versus plasma extracts. The matrix effects experiment was repeated for (+ve) ESI Scherzo LC analysis using Orbitrap Velos^{TM} mass spectrometer with a mass resolution of ≥ 100,000 to evaluate the possibility (ii). The same enhancement results were observed, so we can conclude that the cause of observed ion enhancement is not co-elution of species with similar m/z to the analytes of interest. Next, Na^+ and $NH4^+$ adduct formation was compared for buffer samples *versus* plasma extract samples. No significant differences in adduct formation were found for any analytes except for melatonin, where sodiated adduct with an intensity similar to protonated ions was formed in buffer calibration points. The total area of sodiated and protonated ions was used in calculations of matrix effect for melatonin to correct for sodiated adduct formation. Finally, the observation of large differences between matrix effects in Scherzo and RP despite identical sample preparation protocols between these two methods exclude the involvement of partial solubility.

Thus, it is plausible to conclude that the enhancement effect observed in mixed-mode LC analysis (and occasionally in RP analysis) is based on the true difference in ionization process between plasma-based and buffer-based samples. This is further supported by the fact that for both analytes that exhibition enhancement, their isotopically labeled standards also confirm the same extent of enhancement, showing high-quality of the collected data.

2.3.1.5 Selection of internal standards for global metabolomics.

Our recovery and matrix effects results can be used to guide the selection of the best internal standards for quality control for human plasma metabolomics. Standards that show no susceptibility to matrix effects make ideal internal standards (as SIL) spike before extraction in order to monitor extraction recovery. These include 5-methoxytryptamine, folic acid, thyrotropin-releasing hormone, and cortisol for (+ve) ESI RP; and pantothenic and cholic acids for (-ve) ESI RP, all of which showed no matrix effects across all seven extraction methods tested as shown in **Supplementary Table A7 (Appendix A).**

Table 2.1. Summary of the total number of standard analytes which experienced matrix effect (suppression or enhancement) across different extraction methods in combination with either RP or mixed-mode LC-MS analysis. Analytes (n=24, including SIL for some of the metabolites) were counted if they were detected in the buffer and at least one of the post-extraction spiked calibration standards. For metabolites detected at all concentration levels, the mean matrix effect obtained across all concentration levels is reported. A metabolite was considered to be enhanced if its matrix effect ratio exceeded 120% or suppressed if it was less than 80%. The species that were not detected in either matrix were not counted because matrix effects could not be properly determined for such cases. ***Supplementary Table A6 (Appendix A)*** *shows full results for matrix effect evaluation.*

Evaluation of matrix effects	**RP LC-MS analysis**							**Scherzo mixed-mode LC-MS analysis**						
	Methanol-ethanol	**Methanol-MTBE**	**Methanol**	**MTBE**	**IEX**	**PEP2**	**C18**	**Methanol-ethanol**	**Methanol-MTBE**	**Methanol**	**MTBE**	**IEX**	**PEP2**	**C18**
Suppressed (+ ESI)	6	6	6	4	7	9	5	4	4	4	4	4	7	5
Enhanced (+ ESI)	5	5	5	4	2	4	2	11	11	11	7	7	3	2
Total affected (+ ESI)	11	11	11	8	9	13	7	15	15	15	11	11	10	7
Total unaffected(+ ESI)	6	6	6	12	7	4	10	2	2	1	8	4	7	10
Suppressed (- ESI)	3	3	5	0	1	1	1	7	4	9	1	10	5	3
Enhanced (- ESI)	2	2	2	2	4	7	8	4	3	2	5	1	4	2
Total affected (- ESI)	5	5	7	2	5	8	9	11	7	11	6	11	9	5
Total unaffected (- ESI)	5	5	3	10	9	6	5	2	6	2	8	1	4	9

These can be supplemented with additional standards spiked post-extraction to monitor for matrix effects such as SIL analogues of adenine and thyroxine for (±) ESI RP; neurotensin, melatonin, thyroxine, and cortisol for (+ve) ESI Scherzo and homovanillic acid, melatonin and thyroxine (-ve) ESI Scherzo. These analytes show large differences in matrix effects between different extraction methods as shown in **Supplementary Figure A2 (Appendix A),** which suggests their ionization is susceptible to the presence of possible co-eluting interferences and do not overlap with proposed recovery standards. The use of matrix effect standards is suggested to evaluate relative matrix effects between individual samples, but it should be noted that they would only reflect the extent of ion suppression at that specific moment of chromatographic run.

Finally, the above internal standard suggestions are valid for the exact extraction methods, plasma loading, and LC methods tested in this study. Further testing is required to extrapolate the use of these specific standards to other experimental conditions. In general, for any combination of extraction method and LC-MS analysis, standards with high recovery in that extraction method and no matrix effects across all extractions for the chosen LC-MS method would make ideal

recovery standards, while standards highly susceptible to ion suppression/enhancement would make useful internal standards for monitoring of matrix effects across individual samples.

2.3.2 Global metabolite analysis

Seven extraction protocols were compared using four LC-MS analyses in order to assess metabolite coverage, extraction repeatability, and method overlap (orthogonality) as shown in **Figure 2.3.** In addition, **Table 2.2** summarizes metabolite coverage and extraction repeatability of all extraction methods tested. As expected, the highest number of putative metabolites was extracted by methanol-based solvent precipitation methods, with the highest number of putative metabolites (3804) detected for methanol/ethanol. The analysis of organic MTBE fraction resulted in 2887 putative metabolites as revealed by (+ve) ESI RP analysis. Approximately 30% fewer metabolite features were detected in C18, and PEP2 SPE extracts, while only 1835 putative metabolites were observed for IEX SPE. The table also shows the median RSD of signals across all extraction methods for each LC-MS analysis. Methanol-ethanol and methanol extractions demonstrated the best repeatability versus all other extraction methods independently of the LC-MS method employed. PEP2 and IEX had acceptable repeatability for global metabolomics. On the other hand, MTBE and C18 extraction methods demonstrated the worst repeatability independent of LC-MS analysis (**Table 2.2**). The high proportion of irreproducible features in these two methods requires the application of rigorous quality controls and in-depth investigation for the sources of such irreproducibility.

Previous C18 SPE studies for plasma metabolomics indicate conflicting evidence regarding C18 repeatability for this application. In the first study on this topic, Michopoulos *et al.* showed 48% and 55% of features detected in C18 SPE and methanol precipitation had RSD ≤30%, respectively, which implied both methods have similar repeatability.[189] Rico *et al.* also observed similar repeatability between methanol and C18 SPE with approximately 80% of features, which met 30% RSD criteria for both methods.[186] Our results show that only 42% of features extracted by C18 SPE met the repeatability criteria, which is consistent with Michopoulos's study.[189]

Table 2.2. Metabolite coverage and repeatability of extraction methods assessed by global metabolomics analysis. This table shows the total number of putative metabolites (M) detected in a minimum 5 out of 6 extraction replicates analyzed using RP or Scherzo mixed-mode LC-MS after removal of features present in blank extracts, median RSD of signal intensity across all putative metabolite features detected in the extraction method and the number of metabolites for which extraction method was highly repeatable with RSD ≤ 30% for n=6 independent extractions (M30). RSD for each putative metabolite feature was calculated using raw signal intensities in extraction replicates (n=6). The number of features with RSD ≤ 30% (between replicates) represent high-quality features that could be used for biomarker discovery and pathway analysis in global metabolomics projects.

Extraction method	(+ve) ESI RP			(-ve) ESI RP			(+ve) ESI Scherzo			(-ve) ESI Scherzo		
	M	**Median % RSD**	**M_{30}**	**M**	**Median % RSD**	**M_{30}**	**M**	**Median % RSD**	**M_{30}**	**M**	**Median % RSD**	**M_{30}**
Methanol-ethanol	3804	12.8	3306	1093	13.4	1051	1702	22.3	1087	1278	18.7	930
Methanol-MTBE	3394	17.7	2389	1055	12.7	849	1510	25.2	877	1363	21.0	897
Methanol	3795	11.5	3483	1035	11.4	940	1538	19.0	1089	1113	17.9	802
IEX	1835	17.5	1406	415	12.5	364	894	23.2	571	561	23.1	373
MTBE	2887	37.9	1037	618	26.9	362	948	39.1	326	753	31.7	345
PEP2	2430	21.0	1635	557	14.5	444	1093	24.9	651	773	21.5	498
C18	2459	34.6	1032	603	22.7	394	1078	41.4	318	820	33.2	357
Total metabolome coverage	5853			1466			3072			2229		

However, our results also show vast superiority of methanol repeatability, where 92% of features were highly repeatable with RSD ≤30% for n=6 extraction replicates. Considering this discrepancy across the studies for C18 SPE, further investigation of the factors affecting repeatability is required. Such contributing factors may include the lack of automation used in our study and the exact selection of sorbent characteristics and wash/elution conditions. For instance, our study employed acetonitrile, whereas both Michopoulos *et al.* and Rico *et al.* used methanol as elution solvent which may have contributed to the poor precision.[186,188] Our MTBE results are in contrast to good precision obtained when using MTBE with in-vial dual extraction method where ≥80% of features had RSD ≤30% for n=3 extraction replicates.[173] The same authors observed poor precision of MTBE LLE with evaporation/reconstitution step whereby only 56% of detected features exhibited RSD ≤30%. The latter result is consistent with the results of the current study, where the evaporation/reconstitution step was employed.

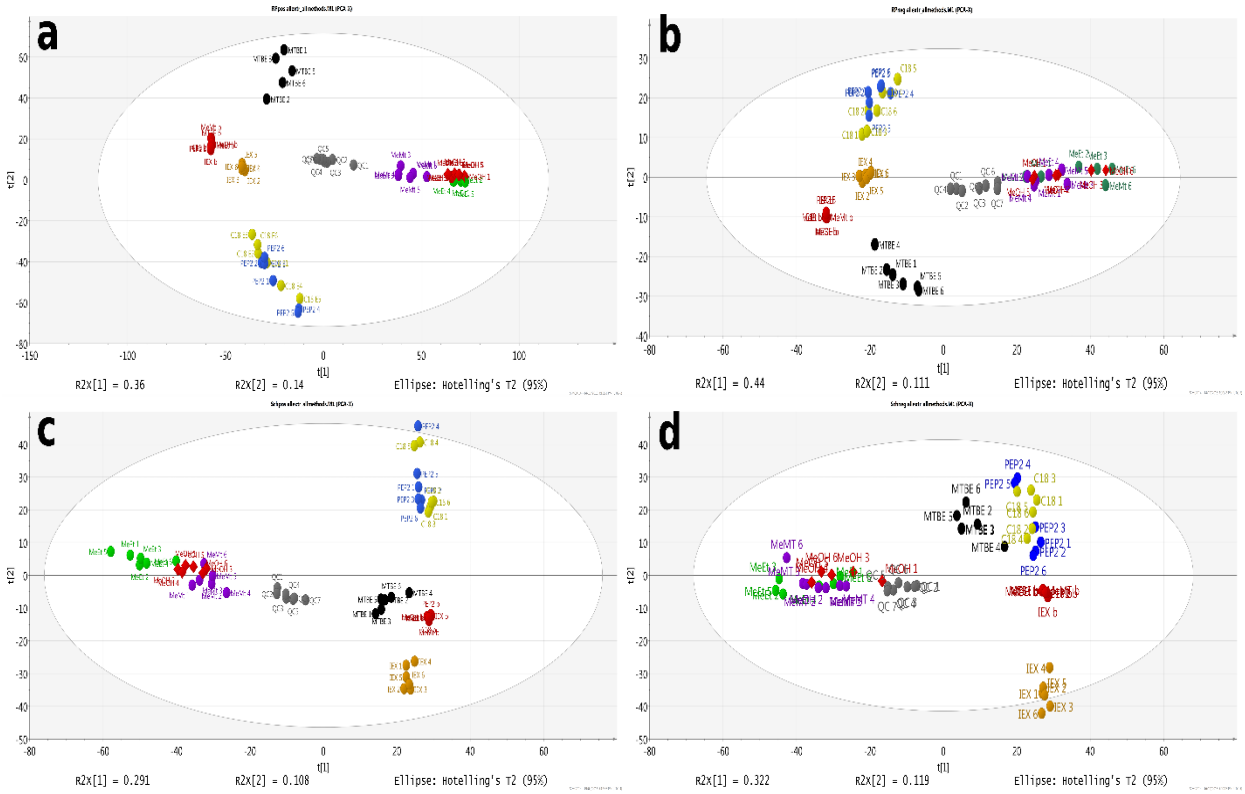


*Figure 2.4. PCA analysis of seven extraction methods and quality control samples analyzed on four LC-MS methods (*a,b *– RP analysis;* c, d *– Scherzo analysis in positive and (-ve) ESI, respectively) executed on all metabolites that satisfy criteria described in the main text (not present in the blank, and present in a minimum of 5 out of 6 replicates of at least one extraction method). The graph displays colored spheres for blank (red), QC (dark grey), IEX (dark gold), C18 (yellow), PEP2 (blue), MTBE (black), methanol-MTBE (violet), methanol-ethanol (green) extraction replicates, and red diamonds for methanol extraction replicates. The plots show the top two principal components. The numbering of replicates corresponds to their sequential injection order in a given LC-MS analysis. The analysis was executed using multivariate analysis software SIMCA (v 14.1.64, Umetrics, San Jose, CA, USA) after Pareto scaling.*

During their evaluation of the optimum method for lipidomics, Sarafian *et al.* also showed poor repeatability of MTBE in comparison to methanol with a similar 2-fold deterioration of mean RSD and the numbers of highly reproducible lipids[199] consistent with what was observed in the current study. During further investigation of MTBE extraction repeatability, the repeatability of solvent pipetting was investigated and found not to be a significant contributing factor to overall method performance. Next, the (+ve) ESI RP LC-MS analysis of newly prepared aqueous and organic layers obtained after MTBE extraction demonstrated repeatability in aqueous layer to methanol extracts (median RSD of 15.7% for MTBE aqueous) versus 16.0% RSD for methanol obtained

during this follow-up experiment. In contrast, organic extracts had a median RSD of 36.4%, which is consistent with the 37.9% median RSD obtained in our initial experiment presented in **Table 2.2.** Detailed investigation of this data showed clear dependence of RSD on retention time: a large proportion of peaks eluting with the retention time of > 20 min had RSDs greater than 50% in both methanol and MTBE extracts. Methanol had a large number of peaks with retention time < 20 min, which exhibited good repeatability, which resulted in good median RSDs observed for the global metabolomics data. MTBE, on the other hand, had only a small number of metabolites eluting with retention time <20 min, and a very high proportion of metabolites with retention time > 20 min, which resulted in overall higher median RSD's observed in **Table 2.2.** Based on this evidence, it is believed that poor MTBE repeatability observed in our study may not arise from the extraction method itself, but from a poor match between the composition of MTBE extract and RP and mixed-mode LC separation methods employed in this study, which would not adequately separate lipids extracted by MTBE.

Further analysis of **Table 2.2** across LC-MS methods demonstrated inferior reproducibility of Scherzo analysis compared to RP. This is attributed to the lower resolution of the Scherzo column (larger particle size than RP column) and larger matrix effect than in RP as shown in matrix effects (**Section 2.3.1.4).** Finally, a hierarchical analysis was performed to determine the pairwise orthogonality of each of the methods tested (**Figure 2.3**). The hierarchical clustering confirms the high orthogonality of IEX and MTBE to other methods observed in targeted analysis, the similarity between methanol-based extractions, and the similarity between C18 and PEP2 SPE methods. Our orthogonality results for methanol and C18 SPE (1825/2459 putative metabolites = 74% overlap using data shown in **Figure 2.3**) are consistent with what was reported by Rico *et al.* who observed 58-68% overlap between the two methods and ability to detect 600 additional features when comparing SPE to methanol precipitation.[186] Using (+ve) ESI RP analysis, a total of 5853 non-redundant putative metabolite features were detected across all seven extraction methods tested. This represents only 54% improvement over the single best extraction method of methanol/ethanol (3804 putative metabolites) or methanol (3795 putative metabolites). Therefore, 7x increase in MS analysis time and the use of LLE and SPE with widely different selectivity mechanisms did not provide a huge boost in our ability to detect low abundance metabolome. Similar results were observed for other LC-MS methods where the increases were 34% ((-ve) ESI RP LC-MS), 80% ((+ve) ESI mixed mode LC-MS), and 74% ((-ve) ESI mixed mode LC-MS). These results clearly

show that simply using multiple extraction methods in parallel is not the best way to increase metabolite coverage and that sequential extractions should be explored to further boost metabolome coverage. In support, **Figure 2.4** shows principal component analysis results for all extraction and LC-MS methods, further illustrating that IEX and MTBE are the most complementary methods to methanol-based solvent precipitation.

Table 2.3. Summary of method performance for extraction of plasma. The number of pluses represents the scoring of method performance where + is the worst and ++++ is the best.

Extraction method	Recovery	Matrix effects	Repeat ability	Metabolome coverage
Methanol/ethanol	+++	++	+++	++++
Methanol	++++	++	++++	++++
Methanol/MTBE	+++	++	+++	+++
MTBE	++	++++	+	++
C18	+++	++++	+	++
PEP2	+++	++	++	++
IEX	+	+++	++	+

2.4 Conclusions

For the first time, absolute analyte recoveries and matrix effects in plasma were systematically assessed for seven solvent precipitations, LLE, and SPE methods using standard addition calibration. In addition, method repeatability, orthogonality, and metabolome coverage were compared in combination with four LC-MS methods. Our results confirm the wide selectivity of methanol-based precipitation methods versus LLE and SPE, with the best results observed using methanol or methanol/ethanol, as shown in **Table 2.3**. However, methanol-based methods suffer from severe matrix effects, which negatively impact data quality and may result in an inaccurate selection of tentative biomarkers. We also show that IEX and PEP2 SPE provide acceptable performance for global metabolomics studies of plasma and can be employed depending on the desired coverage of the metabolome for a given application. Our analysis platform revealed high orthogonality of MTBE and IEX to each other and other methods, providing the possibility of increased metabolome coverage via sequential application of these methods.

3 Development of a sequential SPE-based sample preparation method for global metabolomics analysis of human plasma by LC-MS

3.1 Introduction

The main objective of global metabolomics is a comprehensive analysis of all small-molecular-weight species (≤1,500 Da) in a biological sample.[61] The execution of this task in a complex biofluid such as blood plasma is complicated by the large number of metabolites present, their chemical diversity, and the broad span of their concentrations, exceeding 9-10 orders of magnitude. The chemical diversity of blood plasma metabolome spans the entire polarity range and exceeds any single solvent's extraction capability, thus immediately narrowing down the number of metabolites that can be detected with any single extraction method.[76] Metabolome can be further sub-divided into metabolites with anionic, cationic, zwitterionic, or neutral properties at physiological pH.[93] The inherent size and complexity of metabolome in any biospecimen of interest leads to the co-elution of multiple compounds, which may lead to matrix effects, low signal-to-noise ratios, and overlaps of MS and MS/MS spectra from several compounds.[192,200–203] This results in poor metabolite coverage and can adversely impact data quality and biological interpretation.

In one of the most comprehensive analyses to date, the combined MS analysis of plasma solvent extracts on gas-chromatography (GC-MS), LC-MS, NMR, and revealed a total of 3564 confirmed non-redundant metabolites, 79.9% of which were detected in lipidome profiling in TLC/GC-FID-MS.[61]In addition to combining different analytical techniques, other ways to increase the metabolome coverage include combining orthogonal chromatographic separations in combination with LC-MS, performing LC-MS analysis using both positive and (-ve) ESI modes to ensure adequate coverage of both acidic and basic metabolites, two dimensional (2-D)-LC separations, and/or addition of IMS to LC-MS separations. For example, combining different LC separation methods prior to MS detection provides different elution order of analytes and re-arrangement of the composition of co-eluting compounds and thus enable successful detection of additional non-redundant metabolites. This is commonly achieved by combining highly complementary RP and HILIC methods[166] or, more recently, RP and multimodal ion-exchange (IEX) – RP separations.[173,196] Parallel LC-MS analysis using orthogonal chromatographic separations can significantly increase metabolome coverage. Thus, parallel analysis of the solvent extract of

human blood plasma using RP and zwitterionic HILIC (ZIC-HILIC) resulted in detecting 5188 and 4739 distinct metabolic features in positive and (-ve) ESI, respectively, representing > 50% increase versus using the single LC method.[81] Also, the increase of metabolome coverage can be achieved using an online 2-D LC-MS with BEH C8 and BEH C18 columns, which have complementary selectivity towards aqueous and lipid metabolites, respectively.[204] This approach identified 447 metabolites vs. 213, and 374 metabolites identified off-line during aqueous and lipid analyses, respectively. Finally, the introduction of IMS provides an orthogonal separation to LC and enhances the separation of metabolites with different collisional cross-sections, thus increasing the number of metabolic features in MS analysis of blood serum.[171] However, despite all these advances in analytical instrumentation and methods, the high complexity of the human plasma metabolome remains one of the major analytical challenges, and new strategies to further increase metabolome coverage and data quality in untargeted metabolomics analysis are critically needed.[28] One way to decrease sample complexity with a simultaneous increase in blood plasma metabolome coverage is sample preparation. The key goal of sample preparation in standard untargeted metabolomics is to remove compounds that interfere with MS analysis and equipment lifetime, such as proteins.[174,175,185,189,199] Methanol extraction is currently one of the most popular methods for the preparation of plasma or serum samples because it removes nearly 99% of proteins[199], has good repeatability, and generates the widest metabolome coverage using a single solvent.[18,172–174] Although the use of different solvents results in slightly different metabolome coverage, a parallel LC-MS analysis of extracts prepared using different solvents results in a disproportionally small increase in the metabolome coverage when compared to labor and MS time expenditure. Thus, the metabolome coverage of acetone plasma extract, the most dissimilar out of nine solvents tested in parallel, added only 12.8% of new features to the coverage provided by its most orthogonal rival - methanol/ acetone cocktail.[205] The separation of proteins and metabolites could also be achieved by size exclusion techniques, such as dialysis, ultra-filtration or size-exclusion chromatography and is used in the studies of protein-metabolite interactions.[206] However, the resolution of a size exclusion chromatography is not sufficient for a clear separation of peptides from metabolites. In addition, the recovery losses of multiple metabolites were reported in microdialysis and ultrafiltration.[207] Thus, none of these size-based methods have yet found wide use in global metabolomics analysis. Studies based on SPE removal of phospholipids such as HybridSPE™ from Sigma,[18] Ostro™ from Waters, and Phree™ from Phenomenex[167,172] demonstrated mixed

results, where a decrease in a metabolome coverage of metabolites is concomitant with improvements in sensitivity, repeatability, and matrix effects for metabolites remaining after phospholipid removal. On the other hand, sensitivity, repeatability, and matrix effects were improved significantly for several metabolites remaining after phospholipid removal. At the same time, sequential sample preparation, commencing with methanol precipitation of plasma followed by SPE sub-fractionation of methanol supernatant into phospholipid, lipid, and polar fractions, drastically increased the metabolite coverage to 4,264 molecular features compared to 1,792 detected in methanol extract alone.[177] In addition, the parallel analysis of non-polar and polar fractions generated by LLE of plasma by methyl tert-butyl (MTBE) increased metabolome to 3125 metabolites compared to 1851 metabolites observed in methanol precipitations.[174] The subsequent SPE fractionation of the non-polar MTBE fractions into phospholipid, fatty acid, neutral, and hydrophobic lipid fractions demonstrated a further increase of the metabolome coverage to 3,806. The above SPE sub-fractionation methods focused on extensive fractionation of lipid sub-metabolome. However, the effect of the sub-fractionation of polar compounds on the metabolome coverage had not been demonstrated clearly to date. In one study, fish plasma was deproteinized by methanol and delipidated on Phree^{TM} SPE with a subsequent parallel sub-fractionation using three SPE methods: (i) polymeric-RP SPE consisting of Strata-X sorbent; (ii) a mixed-mode SPE consisting of Strata-X sorbent with strong cation IEX (Strata-X-C) and (iii) mixed-mode SPE of Strata X with weak anion IEX (Strata-X-AW).[96] This approach allowed at least a four-fold increase in sample loading for SPE fractions versus methanol extracts without overloading the nano-LC column used for the analysis. It demonstrated a 40-fold increase in signal intensity for many metabolites, the detection of more than 5900 molecular features, and precision below 30% for at least 62% of metabolites. However, only metabolites bound to SPE were analyzed in this experiment, which made a comprehensive analysis of the effect of ion-exchange sub-fractionation on the coverage of a polar metabolome impossible. The second study compared the performance of seven extraction methods and found that mixed-mode ion-exchange SPE fractionation demonstrated a lower total number of metabolites detected in bound fractions compared to methanol extracts and RP C18 SPE extracts.[173] Besides, the composition of this mixed-mode IEX was highly orthogonal to methanol, thus illustrating its ability to increase polar metabolome coverage. However, to the best of our knowledge, comprehensive studies of the effect of a sequential fractionation on polar metabolome have not been reported.

Therefore, there is a need for a comprehensive assessment of polar and mid-polar plasma metabolite coverage after ion-exchange SPE. The objective of this study was to design, develop and characterize a novel sample preparation protocol that combines methanol deproteinization and sequential SPE fractionation of polar and mid-polar plasma metabolites into anion, cation, zwitterion, and neutral fractions. We systematically assessed this approach's advantages and disadvantages on metabolome coverage, matrix effects, method reproducibility, and metabolite recovery. Finally, we also investigated the compatibility of these anion, cation, zwitterion, and neutral fractions with RP and ZIC-HILIC chromatographic LC-MS methods to design the most optimal combinations of sSPE fractions and LC-MS analysis. This methodology will increase the coverage of the polar and mid-polar metabolome according to the chemical properties of metabolites, the knowledge of which can be partially useful for metabolite identification. In addition, this new approach may facilitate the development of targeted sample preparation methods and improve the integration of global and targeted metabolomics analyses.

3.2 **Materials and methods**

3.2.1 Chemicals and consumables

LC-MS grade solvents/mobile phase additives and analyte standards were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless stated otherwise. Norepinephrine-(d_6), cholic acid (d_4), cortisol (d_4), epinephrine (d_3), dopamine (d_4), melatonin (d_4), 4-aminobutanoic acid (d_6), and phenylalanine (d_5) were obtained from CDN Isotopes (Point-Claire, QC, Canada); $^{13}C_6$- thyroxine and diosmetin (d_3) were purchased from Toronto Research Chemicals (Toronto, ON, Canada). Eicosanoids, kynurenine, and D-erythro-sphingosine (further referred to as sphingosine for brevity) were purchased from Cayman Chemicals (Ann Arbor, MI, USA) or Cedarlane (Burlington, ON, Canada). MTBE, acetic acid, and formic acid were bought from Fisher Scientific (Toronto, ON, Canada). Strong mixed-mode anion-exchange/divinylbenzene-co-N-vinylpyrrolidone (MAX) and strong mixed-mode cation-exchange/divinylbenzene-co-N-vinylpyrrolidone (MCX) Oasis plates were purchased from Waters (Mississauga, ON, Canada). Citrated pooled human plasma was obtained from BioIVT (Baltimore, MD, USA).

Lipids were purchased from Avanti Polar Lipids (Alabaster, AL, USA). All reagents were of analytical or higher grade. A kit of 17 stable isotope labeled amino acids (part# MSK A2 S) was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA).

3.2.2 Overview of method development for sequential sample preparation

The initial sequential sample preparation consisted of three consecutive steps: (i) MTBE extraction of plasma to give an organic phase (designated further as "MTBE NP" fraction) and the aqueous phase (defined further as "aqMTBE" fraction); (ii) MeOH-based precipitation of aqMTBE defined further as "aqMTBE-MeOH" and (iii) sequential SPE mixed-mode strong anion-exchange (SAX) followed by mixed-mode strong cation-exchange (SCX) of aqMTBE-MeOH (**Figure 3.1**). The analysis of non-lipid metabolites was executed on C18 RP column using LC conditions for mid-polar metabolites, i.e., those that are retained (RT > 2 min) on such columns and, elute between 2 and 75% MeCN and require the special C18 RP LC method conditions described in detail in **Section 3.2.3.3.** Experiments were used to assess metabolome coverage, orthogonality to lipid metabolites (lipidome), and to define the final scheme of sequential sample preparation, and results are described in **Section 3.3.1**. An assessment of lipidome coverage was executed using lipid LC-MS method (**Section 3.2.3.4.5**). The lipid composition of extractions was compared to a standard lipid extraction of plasma by isopropanol (IPA).[199] These experiments are described in detail in **Section 3.2.3.4,** and results were used to assess lipid recovery, lipidome coverage, and suitability of sSPE for lipidomics analysis, as described in detail in **Section 3.3.1**. The step of MTBE extraction was removed, so the final protocol, included only two consecutive steps: (i) MeOH extraction of blood plasma and (ii) SPE fractionation (strong anion-exchange Oasis followed by strong cation-exchange Oasis) of MeOH extract. Samples generated using the final extraction methods were analyzed on C18 RP and ZIC-HILIC columns, using targeted and global analysis (**Section 3.2.4)**. Fractionation quality, orthogonality, metabolome coverage, maximum loading amount, and optimal combination of sSPE fractions for faster analysis were assessed in these experiments (**Section 3.3.2**)**.** An additional experiment (**Section 3.2.5)** using the final protocol of sSPE (**Section 3.2.4**) was executed separately with the LC-MS analysis carried out on the HP 1290 Infinity II coupled to a 6545 QTOF MS (Agilent Technologies, Santa Clara, CA). The methods of sample preparation and LC-MS analysis were identical to those described in **Section 3.2.4**, except that (i) LC-MS analysis was executed on sSPE fractions combined in pairs to reduce analysis time 2-fold and (ii) sample loading amounts were maximized to yield the highest metabolome coverage. The parameters of analytical performance (recovery, repeatability, metabolome coverage, matrix effects) were assessed, and results are discussed in **Section 3.3.3.**

3.2.3 Initial protocols of sequential sample preparation consisting of MTBE LLE, MeOH precipitation and sSPE and LC-MS analysis

3.2.3.1 Standard analyte mixture for initial protocol

The systematic evaluation of the extraction methods was performed using a standard mix that (i) resembled the class composition of biological samples and included acids, bases, neutrals, and zwitterions; (ii) covered a wide range of chemical properties ($LogP_{octanol/water}$ range from -10 to 7, MW range from 75 to 783 Da), and (iii) was amenable to the C18 RP LC-MS method employed in the study. All individual stock solutions were prepared in appropriate solvents as summarized in **Appendix B, Supplementary Table B1**, divided into aliquots and stored at below -70ºC.
The standard mix was prepared at 10 µg/mL in 20% MeOH from individual stocks. Working solutions were prepared by dilution of the standard mix to the required concentrations on the day of extraction, incubated on ice for 30 min, and added into samples prior to SPE.

3.2.3.2 Sequential extraction of metabolites by MTBE LLE, MeOH precipitation, and sSPE

3.2.3.2.1 *MTBE LLE and MeOH extractions of the aqueous phase of aqMTBE*

Aliquots (150 µL) of citrated blood plasma or water (blank) were mixed with 600 µL of ice-cold MTBE and shaken at +4 ºC for 30 min with subsequent centrifugation at 17,000 x g for 15 min at +4 ºC. Then, 100 µL of the lower, aqueous phase (aqMTBE) was mixed in a separate Eppendorf tube with 400 µL of ice-cold MeOH and immediately shaken for a few seconds. Tubes were incubated at ambient temperature for 15 min with shaking every 5 min, followed by 15 min incubation at below -70 ºC without shaking and centrifuged for 15 min at 17,000 x g at +4 ºC. Further, 175 µL of the resulting supernatant of aqMTBE-MeOH was transferred into clean tubes and stored on ice until SPE. In addition, aliquots (126 µL) of aqMTBE-MeOH and aliquots (101 µL) of nonpolar (organic, upper phase) fractions of MTBE (MTBE NP) were dried under a vacuum in a speed-vac for 8 h at 30 ºC and stored at below -70 ºC until LC-MS analysis. Volumes of aqMTBE-MeOH and MTBE NP were adjusted to the amount of material, equal to the content of sSPE fractions (25.2 µL of the original plasma)).

3.2.3.2.2 *Strong anion-exchange Oasis SPE*

The principal steps of sSPE are presented in **Figure 3.1.** A SAX SPE procedure was executed on 96-well, 2 mL MAX Oasis plates. The MAX plate was washed sequentially with 1 mL of MeOH, 1 mL of 5 mM ammonium hydroxide in MM64 (60% MeCN and 40% MeOH), and 2 mL of 5 mM ammonium hydroxide in water.

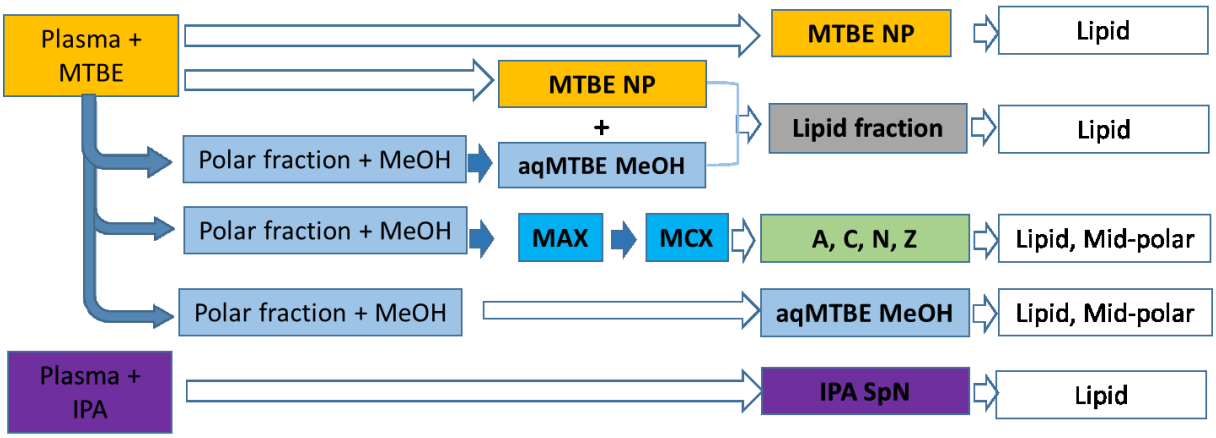


Figure 3.1. Flow chart of initial experiment to evaluate sSPE sample preparation versus MTBE LLE and plasma precipitation by IPA. The sequential sample preparation method included: MTBE extraction of plasma, followed by MeOH protein precipitation of polar fraction, followed by a sequential MAX-MCX SPE to generate four fractions enriched with anions (A), cations (C), neutrals (N), and zwitterions (Z). For the comparison of lipid content, non-polar fractions of MTBE ("MTBE NP"), methanol precipitation of polar MTBE fraction (aqMTBE-MeOH), IPA extraction of plasma ("IPA SpN") and sSPE fractions were compared in lipid analysis. Fractions of sSPE and their predecessor, the aqMTBE-MeOH were compared in mid-polar LC-MS analysis. Both types of analysis were carried on C18 RP columns, though LC-MS methods were modified for an optimal separation and detection of either mid-polar metabolites or lipids.

Aliquots of MeOH extracts (175 µL) of plasma and blanks were mixed with 825 µL of the freshly prepared standard mix (85 ng/mL) in 2.5 mM ammonium hydroxide to generate "pre-spiked" samples and prepared in single replicates. Aliquots of MeOH extracts, which were used for "post-spiked" sSPE fractions and "background," were prepared in single replicates and received the same solution without standards. Then, 750 µL of each plasma sample (equivalent to 26.25 µL of blood plasma) and blanks were transferred to the MAX Oasis SPE plate and loaded by gravity at ambient temperature. The filtrate was reloaded back onto the MAX Oasis SPE plate to ensure complete binding of anion and zwitterion metabolites, and filtrates were collected into a clean plate under

positive pressure (nitrogen pressure 8 bars, hold time 2 min) with the aid of the MPE2 unit of the Nimbus 96 robot (Hamilton, ON, Canada), which was used for all further filtration steps. Wells were washed once with 200 µL of 5 mM ammonium hydroxide in water and two times by 250 µL of 7.5 mM ammonium hydroxide in MM64 and collected into the same plate forming fraction enriched with cations and neutral metabolites (CN). The SPE plate was eluted into another collection plate by two sequential passages (total 600 µL) of 3% of formic acid in MM64, forming the fraction enriched with anion and zwitterion metabolites (AZ). The collected fractions were dried under vacuum in a Speed-vac dryer for 26 hours at ambient temperature and subjected to MCX fractionation on the next day.

3.2.3.2.3 *Strong cation-exchange Oasis SPE*

SCX was executed in 2 mL MCX Oasis 96-well plates. Wells of MCX plates were washed sequentially with 1 mL of MeOH, 1 mL of 4 mM acetic acid in MM64, and 2 mL of 4 mM acetic acid in water prior to applying a sample. Dried CN and AZ fractions were reconstituted in 100 µL of 8 mM of acetic acid in 30% of MM64, sonicated for 10 min, shaken on the rocker shaker at the frequency of 450 rpm/min, and diluted with 300 µL water. Sonication and shaking were repeated, and plates were centrifuged at 5,000 x g, +4 ºC for 1 min. Supernatants (350 µL) were transferred into the MCX Oasis SPE plate and filtered by gravity. Then, filtrates were reloaded back into the same MCX Oasis SPE wells and collected with the aid of the Hamilton Nimbus as described above for MAX SPE. These filtrates were combined with several washes. The first wash was executed by 100 µL of 2.5 mM acetic acid in 7.5% MM64, followed by 600 µL of 2 mM acetic acid in 100% MM64, and produced N and A fractions (for CN and AZ fractions, respectively). The wells were then eluted two times using 300 µL of 2.5 M NH_4OH diluted in 100% MM64, and the eluent was collected into a clean 96-well plate resulting in C and Z fractions (for CN and AZ fractions, respectively). The collected fractions were dried under vacuum in a speed-vac dryer for 18 h at ambient temperature and stored at or below -70 ºC until LC-MS analysis the next day.

3.2.3.3 Analysis of mid-polar metabolome

3.2.3.3.1 *LC-MS analysis of mid-polar metabolites*

Dried samples received 40 µL of 20% MeCN and were subjected to the reconstitution procedure, which included two operations executed at ambient temperature: (i) sonication in a water bath for 10 min and (ii) shaking (450 rpm) for 20 min in an orbital shaker. Samples destined to serve as "post-spiked" were reconstituted with the same solvent containing standard metabolites at concentrations 5x higher than expected 100% recovery of standards spiked into "pre-spiked" samples. Then, 160 µL of water was added, the reconstitution procedure repeated, and samples were centrifuged at 10,000 x g and +4 ºC for 60 s. Samples were analyzed using 10 µL injection volumes on 1290 UHPLC chromatograph (Agilent Technologies, Santa Clara, CA) using 2.1 x 100 mm UHPLC Zorbax Eclipse Plus^{TM} (octadecylsilane particles, 1.8 µm diameter, 95 Å pore size, equipped with 2.1 x 5 mm guard column. Mobile phases for (-ve) ESI were: solvent A: 0.02% (v/v) acetic acid in water, solvent B: 0.02% (v/v) acetic acid in MeCN. For (+ve) ESI, acetic acid concentrations in both A and B were increased to 0.05% (v/v). The gradient started from isocratic 2% B between 0 and 2 min, followed by a linear increase of B from 2% to 100% between 2 and 22 min, followed by 100% B isocratic for 2 min and finalized at 2% B for the last 6 min for column equilibration. The chromatography was carried out at 0.4 mL/min at +35°C. The LC column was coupled to Agilent 6550 iFunnel Q-TOF mass spectrometer equipped with the source containing two nebulizers: (i) main-for samples and (ii) auxiliary-for calibration standards).

MS analysis was executed with the internal mass calibration using reference standards supplied via the auxiliary nebulizer in positive (+ve) and negative (-ve) ESI in the mass range 100–1000 m/z. Parameters in (+ve) ESI Q-TOF were set to: capillary voltage of 3800 V for the entire run and nozzle voltage of 200 V for the first 4 min and 1500 V from the 4th to the 28th minute of the analysis. In (- ve) ESI, capillary and nozzle voltages were set to 3500 and 500 V, respectively, during the first 5.5 min of a run and to 4200 and 800 V between 5.5 and 29 min. Drying and sheath gas temperatures were set to +250 and +275°C and flow rates to 15 and 12 L/min, respectively, disregarding ESI mode. Nebulizer pressure was set to 30 PSIG, and the fragmentor voltage to 175 V. Data was acquired in both centroid and profile mode at the rate of 3 spectra per s in the extended dynamic range mode (2 GHz). Resolution of 12,000 FWHM (full width at half maximum) at m/z 121 and 24,000 FWHM at m/z 922 was achieved. To assure acceptable mass accuracy of recorded

ions, continuous internal calibration was executed using signals at m/z 121.0509 (protonated purine) and m/z 922.0098 (protonated hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine (HP-921)) in(+ve) ESI mode. Ions with m/z 119.0363 (deprotonated purine) and m/z 966.0007 (formate adduct of HP-921) were used for calibration in (-ve) ESI.

Due to the large number of samples, each type of LC-MS analysis was executed in a separate analytical batch. Time intervals between batches did not exceed 36 h. Sample injections were executed in a temperature-controlled (+6 ºC) injector in random order. Pooled QC samples were injected at the beginning (n=3), at every 11^{th} injection, and at the end (n=3) of each batch. Pooled QC samples were created by mixing equal volumes of all plasma extracts destined for the batch unless stated otherwise.

3.2.3.3.2 *Data analysis*

For the targeted data analysis, peaks with known mass and RT were picked and integrated using the targeted method in Profinder (version 08.01 SP) application from Agilent Technologies (Santa Clara, CA, USA) with an ion mass accuracy threshold of ± 20 ppm, retention time ± 0.5 min and intensity cutoff of 7500 counts. Peak selection and integration were verified for each standard.

The analysis of recovery was preceded by subtracting signals in "background" samples from "pre-spiked" and "post-spiked" samples. Then, the recovery was calculated in Excel by the formula: (signal in "pre-spiked" * 100%) /signal in "post-spiked."

For the global data analysis, peak picking, deconvolution, alignment, and integration were executed on the Profinder (version 08.01 SP1) application of Mass Hunter™ suite (Agilent, Santa Clara, CA) with the following parameters: retention time 2-15.2 min, ion mass accuracy threshold of ± 20 ppm, the relative height of MS+1/MS isotope abundance to 15%, RT threshold ±0.15 min, minimum peak height 200 and 2000 counts for M+1 and M peaks, respectively. The analysis of global metabolome coverage in MeOH and sSPE extracts were carried out on Mass Profiler Professional (version 14.1) application of Mass Hunter™ suite with integration and binning parameters similar to the settings on the Profinder method. Metabolites with low quality were removed if they met the following criteria: (i) were not at least 5x higher than the signal in a blank, or (ii) were not found in at least 4 out of 6 replicates of a given extraction method. Manual curation of a subset of data found 2-3% duplicate entries (a compound split between multiple entries by the peak picking algorithm). The metabolome coverage of aqMTBE-MeOH and sSPE was compared

in Venn analysis. In addition, TIC chromatograms of MTBE NP, aqMTBE-MeOH, and sSPE fractions were visualized side-to-side to evaluate the orthogonality of metabolome coverage in C18 RP analysis in (+ve) and (-ve) ESI to MTBE LLE and MeOH extraction. The identification of metabolites was not carried out in this thesis. Therefore, the global results are based on the analysis of putative unidentified metabolites. The term "metabolites" mentioned throughout the analysis and discussion of global analysis was used for brevity.

3.2.3.4 Lipidomics analysis

3.2.3.4.1 *IPA extraction for lipidomics analysis*

Ice-cold IPA (300 µl) was mixed with 100 µl of ice-cold plasma in a 1.5 mL centrifuge Eppendorf tube and vortexed immediately for 1 minute at ambient temperature. Then, the mix was incubated at below -70°C for 1 h without shaking and centrifuged at 25,000 x g at +4°C for 20 min. Aliquots of the supernatant (100 µl) were transferred into 1.5 mL Eppendorf polypropylene tubes and dried for 2 hours at +30°C in a vacuum dryer.

3.2.3.4.2 *Generation of the lipid fraction*

An aliquot (100 µL) of MTBE NP of plasma was mixed with 125 µL of the aqMTBE-MeOH, dried for 8 h at +30°C in a vacuum dryer, and stored at below -70°C until lipidome analysis.

3.2.3.4.3 *Preparation of extracts for lipid LC-MS analysis in (-ve) ESI*

Dried samples were reconstituted in 40 µl (36.3 µl for SPE extracts) of a buffer containing 0.02% (v/v) acetic acid, 10% (v/v) MeOH, 90% (v/v) IPA and 3 µM of lipid internal standard mix (**see Appendix B, Table B2**). After 10 min of sonication in a water bath at ambient temperature and 15 min of the shaking in Genie Fisher shaker at 450 rpm, samples were centrifuged at 17,000 x g and +4 °C for 1 min. Then, samples were diluted with 160 µL (145.2 µL for SPE extracts) of 0.02% v/v acetic acid in 60/40 water/MeOH v/v, sonicated, and shaken again as described above.

3.2.3.4.4 *Preparation of extracts for lipid LC-MS analysis in (+ve) ESI*

Aliquots (10 µL) of samples prepared in **Section 3.2.3.4.3** were further diluted with 190 µl of the same solvent, sonicated and shaken again as described above in glass HPLC inserts. All samples were vortexed for 1 min and centrifuged for 1 min at 10000 x g at ambient temperature.

3.2.3.4.5 *LC-MS methods for lipid analysis*

LC-MS analysis was executed on LTQ Orbitrap Velos™ (Thermo Fisher Scientific, Mississauga ON, Canada) using Waters CSH C18 (130 Å, 2.5 µm, 2.1x75 mm) column equipped with a 2.1 x 5 mm guard-column (XSelect™ CSH C18 XP VanGuard™ Cartridge, 130 Å, 2.5 µm). The sample was eluted at a flow rate of 0.25 mL/min and temperature 55ºC. For the analysis in (+ve) ESI, mobile phase A contained 10 mM of ammonium acetate in 60/40 water/MeOH (v/v) while B was 90/10 IPA/MeOH (v/v). For the analysis in (-ve) ESI, ammonium acetate in both phases was replaced by 0.02% (v/v) aqueous acetic acid. The same gradient was used for both ESI modes (**Appendix B, Supplementary Table B3**). The MS parameters were set to: capillary temperature +275°C, a source at +300°C, and the voltage of 3.5 kV in (+ve) ESI or 3 kV in (-ve) ESI. Sheath gas flow was set to 10 for (+ve) ESI or 15 for (-ve) ESI; aux gas flow was set to 10 (arbitrary units) and resolution to 60,000 FWHM (MS^1). Internal mass calibration was performed using the mass lock set to the mass of the ubiquitous background ion: 391.28428 in (+ve) ESI or 311.16859 in (-ve) ESI. The data-dependent CID fragmentation was executed, and spectra collected spectra for eight highest signals above 5000 counts in mass ranges 385-1200 m/z for positive and 280-1200 m/z for negative ionization modes with an isolation width of 2, activation time 50 milliseconds, and normalized collision energy of 35 V.

3.2.3.4.6 *Data analysis*

Lipid identification was carried out using LipidSearch 4.1 (Thermo Fisher Scientific) by matching spectra with a mass tolerance of precursor +/- 10 ppm and product at +/- 0.5 Da against the in-silico online LipidSearch library. The identified lipids were aligned between samples using +/- 5 ppm mass accuracy and RT threshold of +/- 0.5 min. The data was curated by removing duplicates, and signals with a signal ratio plasma/ blank < 5. Further data analysis was executed in Excel, and Venn analysis was carried in MPP (MassHunter™, Agilent Technologies).

3.2.4 Optimization and evaluation of sSPE for analysis of polar and mid-polar metabolites in individual sSPE fractions

The evaluation of sequential sample preparation method was simplified (**Figure 3.2**) compared to the initial protocol due to the removal of MTBE LLE extraction.

3.2.4.1 Standard analyte mix for the analysis of single fractions

Standards were prepared as described in **Section 3.2.3.1** except that the number of targeted metabolites was increased to improve the systematic representation of metabolite classes and amenability of targets to RP and ZIC-HILIC methods. (**Appendix B, Supplementary Table B1**).

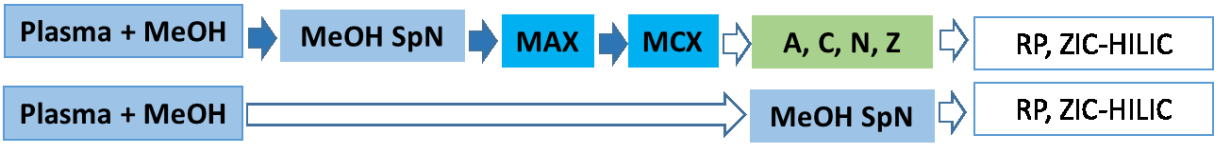


Figure 3.2. Flow chart of the comparison of sSPE fractions (A, C, N and Z) to MeOH extract ("MeOH SpN"), for mid-polar (RP) and polar (ZIC-HILIC) LC-MS analysis. A sequential sample preparation includes MeOH extraction of plasma with subsequent sequential SPE (MAX -> MCX) approach. Sequential SPE generated fractions presumably enriched with a single class of metabolites: anion (A), cation (C), neutral (N), zwitterion (Z), These fractions were analyzed in RP and ZIC-HILIC LC-MS with (+ve) and (-ve) ESI in parallel to MeOH extract of plasma.

3.2.4.2 MeOH extraction of plasma

The flow chart of the final protocol for sample preparation with emphasis on sSPE is shown in **Figure 3.3**. The standard mix was added into plasma or water (blank) to yield 70 ng/mL concentration to form "pre-spiked" samples. Samples, which were processed without standard analytes (and which will serve as "background" and "post-spiked"), received equal volumes of 20% MeOH. For MeOH extraction, 200 µL of the sample was mixed with 800 µL of ice-cold MeOH and processed identically to the initial protocol (**Section 3.2.3.2**). 572 µL of the supernatant were transferred into clean tubes and dried in a vacuum at +36°C for 4 h in a speed-vac. Dried samples were kept at below -70°C until SPE extractions.

3.2.4.3 Strong anion-exchange Oasis SPE

The preparation of a MAX SPE plate was identical to the initial protocol and described in **Section 3.2.3.2.2**. Dried MeOH extracts were reconstituted in 175 µL of 80% MM64 using the reconstitution procedure described in **Section 3.2.3.3**, centrifuged (15,000 x g, 15 s, +4 °C) and diluted with 825 µL of 5 mM ammonium hydroxide in water of ambient temperature, followed by the reconstitution procedure and centrifugation at 10,000 x g, 15 s, +4 °C.

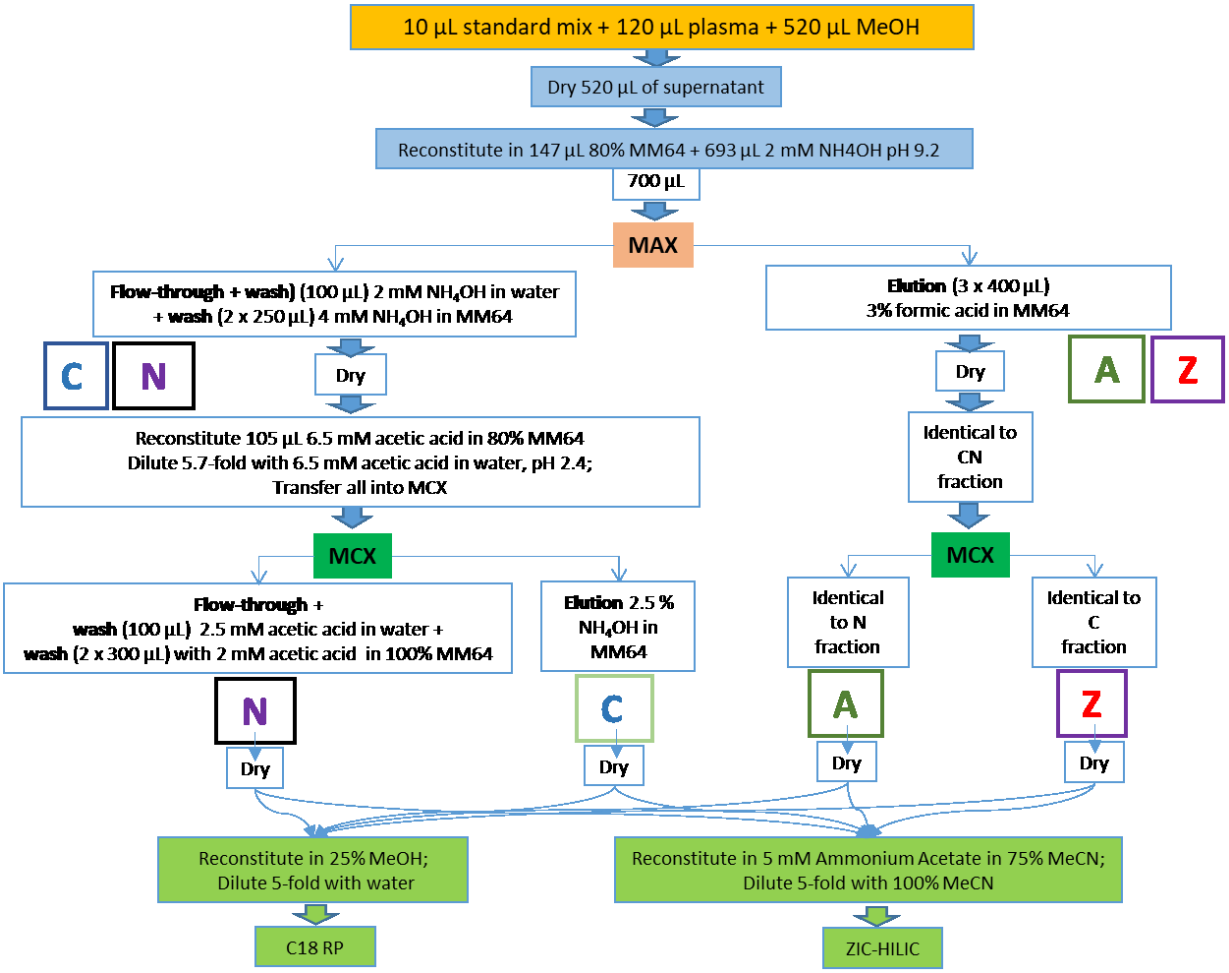


*Figure 3.3. Flow diagram of the final protocol of sSPE for the analysis of individual fractions. The flow and parameters were identical between the analysis of individual and combined sSPE fractions (**Section 3.2.5**) except that in the latter, fractions were reconstituted and mixed as described in **Section 3.2.5.3**. The ideal fractionation would provide segregation of metabolites (A, C, N, and Z) into correspondent sSPE fractions, according to their charge depending on pKa of standards and pH of SPE (MAX, pH 9.2; MCX, pH 2.4). For example, cations (C) neutral at pH 9.2 of at the MAX loading-> CN fraction, from which due to the presence of a charge, will bind to MCX and eluted into cation (C) fraction. Sample loading and the first wash (100 µL) were aqueous, while the second wash and elution solutions were 100% organic and prepared by blending 60% methanol and 40% acetonitrile ("MM64"). These measures were aimed to reduce hydrophobic interactions due to DVBP stationary base used in MAX and MCX Oasis SPE plates.*

Aliquots (700 µL) equivalent to 80 µL of plasma were transferred into wells of MAX SPE plate and processed identically to the initial protocol (**Section 3.2.3.2.2**) except for the use of wash buffers with lower ionic strength. Wells were washed sequentially with: (i) 100 µL of 2 mM ammonium hydroxide in water and (ii) two volumes (250 µL each) of 4 mM ammonium hydroxide

in 100% MM64. The plate with the collected CN fractions was sealed and placed on ice. The elution was carried with three volumes (400 µL, each) of 3% formic acid in MM64 and collected into a clean 2 mL, 96-well plate, producing AZ fractions. The collected fractions were transferred into 1.5 mL Eppendorf tubes and dried under vacuum at +30°C for 12 hours in a Labconco centrifugal vacuum drier (model 7982010). Dried samples were kept at below -70°C until strong cation-exchange SPE, which took place no more than 48 h after drying.

3.2.4.4 Strong cation-exchange Oasis SPE of CN and AZ fractions

Dried MAX SPE fractions were reconstituted in 105 µL of 6.5 mM acetic acid in 80% MM64, processed through the reconstitution procedure, diluted with 495 µL of 6.5 mM acetic acid in the water, and the reconstitution procedure (**Section 3.2.3.3**) was repeated. After centrifugation (15,000 x g, 15 sec, +4°C), reconstituted MAX fractions were loaded into wells of the MCX plate, and SPE was executed identically to the process described in **Section 3.2.3.2.3**. Collected SPE fractions (A, C, N, and Z) were transferred into Eppendorf tubes, dried as described in **Section 3.2.4.3,** and stored at below -70°C until LC-MS analysis, which was executed within 48 h from the time of drying.

3.2.4.5 RP LC-MS analysis

Dried MeOH extracts were reconstituted using 28.6 µL of 25% (v/v) MeOH in water and were subjected to the reconstitution procedure followed by centrifugation (17,000 x g, 15 s, +4°C). Then, samples were diluted with 114.3 µL of water, and the reconstitution procedure was repeated. SPE extracts were reconstituted identically using 20 µL and diluted in 80 µL of 25% MeOH, respectively. Samples were analyzed in random order with QC (mix of equal volumes of al samples) injected at each 11^{th} injection. The instrument and parameters of LC-MS were identical to the ones described in the LC-MS analysis of mid-polar metabolites in **Section 3.2.3.3**.

3.2.4.6 ZIC-HILIC LC-MS analysis

Dried MeOH extracts were reconstituted using 28.6 µL of 5 mM ammonium acetate in 75% MeCN and diluted with 114.3 µL of 100% MeCN using the reconstitution procedures described in **Section 3.2.3.3**. SPE extracts were reconstituted identically but using 20 µL of 5 mM ammonium acetate in 75% of MeCN/water followed by 80 µL of 100% MeCN. The chromatographic separation was executed on a 2.1 x 100 mm SeQuant^{TM} ZIC-HILIC column (Millipore Sigma, ON, Canada)

packed with 3.5 µm particles with a pore diameter of 200Å equipped with a 2.1 x 20 mm guard at 0.4 mL/min and +35°C. Mobile phase A contained 5 mM ammonium acetate in 95/5 (v/v) water/MeCN and B contained 5 mM ammonium acetate in 5/95 (v/v) water/MeCN). The gradient started at isocratic 100% B and held between 0 and 2 min, followed by an 18 min linear gradient from 0 to 20% A and a 7 min gradient from 20 to 50% of A. The gradient was continued to an isocratic mode at 50% A between 27 and 29 min. The column was re-equilibrated for 5 min in 100% B prior to the next injection. Samples were analyzed in random order with QC (mix of equal volumes of al samples injected at every 11^{th} injection. MS analysis was carried out at VCap of 3500 V, nozzle voltage of 400 V, drying gas temperature of +250°C supplied at 15 L/min, and nebulizer pressure of 35 PSIG. Sheath gas temperature was set to +275°C and the flow to 12 L/min, for both ionization modes. Internal mass calibration was executed as described for the RP analysis of mid-polar metabolites in **Section 3.2.3.3.1**.

3.2.4.7 Optimization of sample loading

All samples prepared for LC-MS contained 0.8 µL of the original plasma sample / 1 µL. Samples were injected in variable volumes to evaluate the effect of different sample loads on metabolome coverage. MeOH extracts, as the most complex sample, were loaded in volumes 1, 2, and 4 ul, which were equivalent to the amount of material equivalent to 0.8, 1.6, and 3.2 µL of original plasma volume, respectively. Individual sSPE fractions were loaded in volumes 4, 8, and 12 ul, which were equivalent to the amount of material equivalent to 3.2, 6.4, and 12.8 µL of original plasma volume, respectively.

3.2.4.8 Targeted data analysis

Targeted data analysis was performed using selected identified metabolites:(**Appendix B, Supplementary Table B1**). Raw LC-MS data were processed identically to the initial protocol (**Section 3.2.3.3.2)**. Standards of two types were identified and integrated: (i) reference standards (IR), which were spiked into plasma prior to MeOH extraction, (ii) background reference standards (ER), which were detected in plasma samples using masses and RT of known standards in solvents analyzed in the same analytical batch. The quality of the fractionation was evaluated by examining in how many sSPE fractions each metabolite was detected in. Further interpretation of fractionation behavior was executed by comparing the observed metabolite fractionation behavior with their predicted physical-chemical properties, i.e., pKa's, hydrogen bonding, and LogD (at pH

of MAX and MCX), which can be found in **Appendix B, Supplementary Tables B1** and **B4**. These values were calculated *in silico* by Chemicalize calculator in November of 2019. (www.chemicalize.com) built by ChemAxon (Budapest, Hungary) www.chemaxon.com. In total, 93 metabolites were tested, and 62 were detected in this experiment (**Appendix B, Supplementary Table B4**).

3.2.4.9 Global data analysis

The global analysis of the extraction methods, peak picking, deconvolution, alignment, and integration was executed by Profinder (version 10.0) software from Agilent within the retention time interval of 1 - 16 min (RP) and 0.5 - 20 min (ZIC-HILIC), ion mass accuracy threshold of ± 20 ppm, the ratio of MS+1/MS isotope abundance 12.5%, RT threshold ±0.1 min, minimum ion height of 450 and compound height of 4500. The RT and integration were verified (and corrected, if necessary) manually for every compound detected across all samples. Duplicated compounds and compounds with poor peak shapes were removed. In addition, low-quality metabolite signals that: (i) were not at least 5x higher than the signal in blank and (ii) that were not found in at least 66% of replicates of a given extraction method were also removed. Metabolome coverage and orthogonality of fractions were analyzed using a Venn analysis in Mass Profiler Professional™ (v. 13.0, Agilent). Datasets from each LC-MS analysis for MeOH extracts and sSPE fractions were processed through log transformation and Pareto scaling. A 2-D hierarchical clustering of metabolites and samples was executed using the Euclidian distance and average linkage approach in the GENE-E™ application developed by the Broad Institute (https://software.broadinstitute.org/GENE-E/). The effect of the sample loading amount was analyzed by the assessment of the amount at which the increase in metabolome coverage and signal growth (MeOH/fraction signal ratios) between increasing loads was saturating.

3.2.5 Assessment of the analytical performance of the final sequential protocol in combined sSPE fractions using targeted and global metabolomics approaches

3.2.5.1 Preparation of the standard analyte mix for the analysis of matrix effects and recovery in combined fractions

The analysis of combined fractions was executed using a separate experiment of the sequential sample preparation executed identically to the final protocol developed and described in **Section**

3.2.4, except the composition of the mix of standard metabolites. Due to the necessity to assess recovery and matrix effects, the mix for the experiment was prepared at higher concentrations for several metabolites to yield significant signal differences from their endogenous levels. In addition, the kit with 17 amino acids isotopically labeled (^{15}N and ^{13}C) was added to the standard analyte mix to improve the representation of polar metabolome. These metabolites will be designated by additional sign "SIL" from here and further. The concentrations of amino acids in the mix yielded between 1.3 µg/mL (glycine) to 3.13 µg/mL (tyrosine). The mix contained a total of 62 metabolites in 20% (v/v) MeOH in water (**Appendix B, Table B7**).

3.2.5.2 MeOH extraction and sSPE fractionation

Frozen citrated human plasma was thawed on ice for 30 min, distributed into aliquots (120 µL), which were spiked with 10 µL of standard mix stock ("pre-spiked") or 10 µL of 20% MeOH ("post-spiked" and "background") and vortexed twice for approximately 2 seconds on Genie Fisher 2 (ThermoFisher Scientific, MA, USA). Water in volumes of 130 µL was used for the generation of blank samples. Sufficient numbers of MeOH extracts were prepared to ensure the availability of 6 "post-spiked", 3 "pre-spiked", 3 "background", and one blank for MeOH extraction and each sSPE fraction at LC-MS analysis. After adding ice-cold MeOH (520 µL), aliquots were immediately vortexed (2 seconds) and processed using the identical extraction procedure described in **Section 3.2.4.2.** Supernatants of MeOH extraction (520 µL) equivalent to the amount of material in 96 µL of plasma were dried in a centrifugal vacuum dryer at +30°C for 4 hours and stored at -70°C. On the day of SPE fractionation, MeOH extracts destined for SPE were reconstituted in 147 µL of 80% MM64 and diluted with 693 µL of 5 mM NH_4OH in water. Diluted samples (700 µL) were processed through MAX and then, MCX sSPE identically to the methods which are described in the final protocol (**Sections 3.2.4.3 and 3.2.4.4)** and resulted in the generation of four fractions (A, C, N, and Z) containing material equivalent to 80 µL of the original plasma.

3.2.5.3 LC-MS analysis

3.2.5.3.1 *Reconstitution of extracts and fraction combination for RP LC- MS analysis*

On the day of RP analysis, dried sSPE fractions destined to serve as "post-spiked" were reconstituted in 20 µL of 30% MeCN, which contained the standard mix at a concentration 3.75 x

higher than the expected 100% recovery of metabolites in SPE samples spiked before MeOH extraction (3-OH kynurenine, for example, 151.7 ng/mL). Therefore, upon 3.75-fold dilution by 55 µL of water, these samples represented an idealized 100% reference for recovery calculations. Dry fractions of sSPE, which were destined to serve as "background" samples, were reconstituted in 20 µL of 30% MeCN without standards, followed by the dilution with 55 µL of water. Then, samples were vortexed for 20 min at ambient temperature. After dilution in 55 µL of water, samples were vortexed for 20 min and centrifuged at 17,000 x g at ambient temperature for 10 min. Finally, solvent (20 µL 30% MeCN with the standard mix), which was used for the generation of "post-spiked" samples, was mixed with 55 µL of water to generate neat standards in a solvent, which served as 100% reference for matrix effects assessments. The pooling of fractions was carried out by mixing 25 µL of each fraction in pairs C with N and A with Z for RP analysis in (+ve) ESI. For RP analysis in (-ve) ESI, fractions were mixed to form CZ and AN combinations. All samples were analyzed using an injection volume of 12 µL, which was equivalent to 6.4 µL of original plasma.

On the RP analysis day, dry MeOH extracts were reconstituted in 96 µL of 30% MeCN in water and vortexed for 20 min (450 rpm) at ambient temperature. After dilution with 264 µL of water, samples were vortexed for 20 min and centrifuged at 17,000 x g at ambient temperature. MeOH extracts desired to be used as "post-spiked" samples were reconstituted in 96 µL of 30% MeCN in water, which contained the standards mix at the concentration 3.75 x higher than 100% recovery of metabolites in samples spiked before MeOH extraction (for example, the concentration of 3-OH kynurenine = 75.75 ng/mL). Neat standards, which will serve as 100% references for matrix effects assessments, were prepared on the day as described in detail in **Section 3.2.4.5**. Therefore, after 3.75-fold dilution by 264 µL of water, these samples represented 100% reference for recovery calculations. All MeOH samples were analyzed using 6 µL injection volume, which corresponded to 1.6 µL of plasma. Due to different amounts of sample materials in sSPE and MeOH extracts, separate QC samples were generated for sSPE and MeOH samples. Samples for QC sSPE were created by mixing equal volumes of replicates of combined sSPE fractions. Samples for QC MeOH were created by equal volumes of reconstituted MeOH extracts. Otherwise, experiments were executed as described in **Section 3.2.4.5.**

3.2.5.3.2 *Reconstitution of extracts and fraction combination for ZIC-HILIC analysis*

On the day of ZIC-HILIC analysis, dry sSPE C and A fractions destined to serve as "pre-spiked", "background", and "blank" were reconstituted in 20 µL of 5.5 mM ammonium acetate in 60 % aqueous MeCN and vortexed on an orbital shaker (Thermo Fisher, Mississauga, ON, Canada) at the maximum speed for 25 min at ambient temperature. After centrifugation (10 seconds at 17000 x g at ambient temperature), the content of C and A was transferred into N and Z, respectively, shaken and centrifuged as described in the previous step. Wells N and Z received 80 µL of 100% MeCN, followed by another round of shaking and centrifugation as in previous steps. Then, samples were transferred from N and Z into C and A, respectively, followed by shaking and centrifugation. Therefore, AZ and CN combinations were analyzed in ZIC-HILIC, in both ESI mode. Fractions of sSPE destined to serve as "post-spiked" were reconstituted identically to "pre-spiked" ones, except that the reconstitution buffer contained the standard mix (3-OH kynurenine, for example, = 404.5 ng/mL). After dilution, standard metabolites yielded concentrations corresponded to an idealized 100% recovery for metabolites in pre-spiked samples. All samples were analyzed in 8 µL injection volumes, which corresponded to the amount of material equivalent to 6.4 µL of plasma. Neat standards in a solvent (100% reference for matrix effects assessment) were prepared by mixing reconstitution spiked with the standard mix as described for "post-spiked" samples with 80 µL of 100% MeCN.

On the day of ZIC-HILIC analysis, dried MeOH extracts ("pre-spiked", "background" and "blank") were reconstituted in the reconstitution solvent (96 µL of 5.5 mM ammonium acetate in 60% aqueous MeCN) and vortexed identically to sSPE fractions. After dilution with 384 µL of 100% MeCN, samples were vortexed for 20 min and centrifuged for 10 sec at 17,000 x g at ambient temperature. MeOH extracts, which were destined to serve as "post-spiked" samples, were processed identically to other samples, except that the 5.5 mM ammonium acetate in 60% aqueous MeCN was spiked with the standard mix which, after dilution of samples with 384 µL of 100% MeCN yields the concentration equal to 100% recovery of metabolites in "pre-spiked" samples. All samples were analyzed using 8 µL injection volumes, equivalent to 1.6 µL of plasma material. Therefore, concentrations of standards were strictly proportional to the amount of loaded plasma material for all LC-MS samples regardless of LC mode **(Appendix B, Supplementary Table B7**). Due to the different amounts of sample materials in sSPE and MeOH extracts, separate QC was

generated for sSPE and MeOH samples, as described in **Section 3.2.5.3.2**. Otherwise, experiments were executed, as described in **Section 3.2.4.6.**

3.2.5.4 Data analysis

For the targeted data analysis, raw LC-MS data was processed as described in the final protocol (**Section 3.2.4.8**). Calculations of recovery, repeatability, and matrix effects were executed in Excel. Repeatability was calculated as RSD% across signals in six replicates in "pre-spiked" after subtraction of an average background signal (n=3) from each "pre-spiked" replicate.

For the calculation of recovery, the average background signal (n=3) was subtracted from signals in each "pre-spiked" replicate (n=6) and each "post-spiked" replicate (n=3). The recovery was calculated as the mean of recoveries in each "pre-spiked" replicate(n=6), which in its turn were calculated by the formula: signal in subtracted pre-spiked * 100% / mean of subtracted "post-spiked".

The matrix effects were calculated for post-extraction spiked replicates (n=3) after subtraction of average "background" signals using the formula: (signal in "post-spiked" * 100% / mean of signals in replicates (n=3) of neat standards in reconstitution solvent. Finally, the mean of matrix effects across three "post-spiked" replicates for each standard analyte was calculated.

The global analysis was executed as described in **Section 3.2.4.9** on metabolites that satisfy the following criteria: (i) detected in 4 or more metabolites out of 6 replicates, and (ii) ratio of the average signal in plasma 5-fold higher than in blank, except that 2-D hierarchical cluster analysis and PCA were executed on Statistical Analysis of MetaboAnalyst 4.0.[208] after log transformation and Pareto scaling.

3.3 Results and discussion

3.3.1 Assessment of initial protocols of sequential sample preparation

The initial experiments aimed to define an optimal sequence of sample preparation methods for the global analysis of metabolites in blood plasma, including the fractionation of metabolites based on charge. The flow chart of the entire experiment is depicted in **Figure 3.1**, while the principal steps of sSPE can be found in **Figure 3.3**. The initial protocol combined the most orthogonal methods from our previous work.[173] The sequential sample preparation was designed to fractionate the metabolites into distinct fractions: (i) neutral, hydrophobic metabolites in the organic phase of

MTBE extraction ("MTBE- NP") and (ii) charged and/or neutral polar and mid-polar metabolites, fractionated by sSPE into anion, cation, neutral and zwitterionic groups from the aqueous phase of MTBE LLE deproteinized with MeOH ("aqMTBE-MeOH"). The extraction of plasma by MTBE was intended to remove lipids in order to improve the analysis of polar metabolome and protect chromatographic columns from the damage caused by strong binding of hydrophobic lipids and enable a separate analysis of lipids in MTBE NP fraction. Methanol precipitation was executed on an aqueous phase of MTBE LLE (aqMTBE) to remove proteins prior to sSPE. Mixed-mode SPE phases provide desired ionic interactions with either positively (MAX plates) or negatively (MCX plates) charged moieties complicated by undesirable hydrophobic interactions from the divinylbenzene-co-N-vinylpyrrolidone (DVBP) matrix. Therefore, sSPE procedures were adjusted to minimize hydrophobic interactions by the high content of organic solvents in wash and elution steps (**Figure 3.3)**. The main objectives of the initial protocol of sequential sample preparation (**Section 3.2.3**) were to demonstrate, that sSPE fractionation improves metabolome coverage compared to MeOH extraction and generates orthogonal fractionation of metabolites based on their charge properties. The compatibility of sequential sample preparation with efficient lipid removal was also evaluated and used to guide the development of the final protocol.

3.3.1.1 Lipid content of fractions generated using the sequential sample preparation method

The untargeted lipidomic analysis was executed on all of the extracts generated by sequential sample preparation, including (i) aqMTBE after the deproteinization with MeOH, i.e., aqMTBE-MeOH; (ii) MTBE NP without additional treatment and (iii) the mix of (i) and (ii) the "lipid fraction", which represented combined lipidome coverage of (i) and (ii).

3.3.1.2 Analysis of mid-polar metabolite coverage in sequential sample preparation combining MTBE LLE, MeOH and sSPE

The targeted analysis demonstrated an expected distribution of the majority of standards into sSPE fractions (**Appendix B, Supplementary Figure B2**). The exceptions were found amongst anions (pantothenic acid, split between A and N fractions, and resveratrol detected in N only). In the group of neutral metabolites, adenine was detected in A and creatinine – in C fraction. Zwitterions were fractionated as expected, except for saccharopine, which was detected in the C fraction. The ionization status of these metabolites assumed from predicted pKa values for pH's of sSPE steps (**Appendix B, Supplementary Table B1)** does not explain their unexpected fractionation.

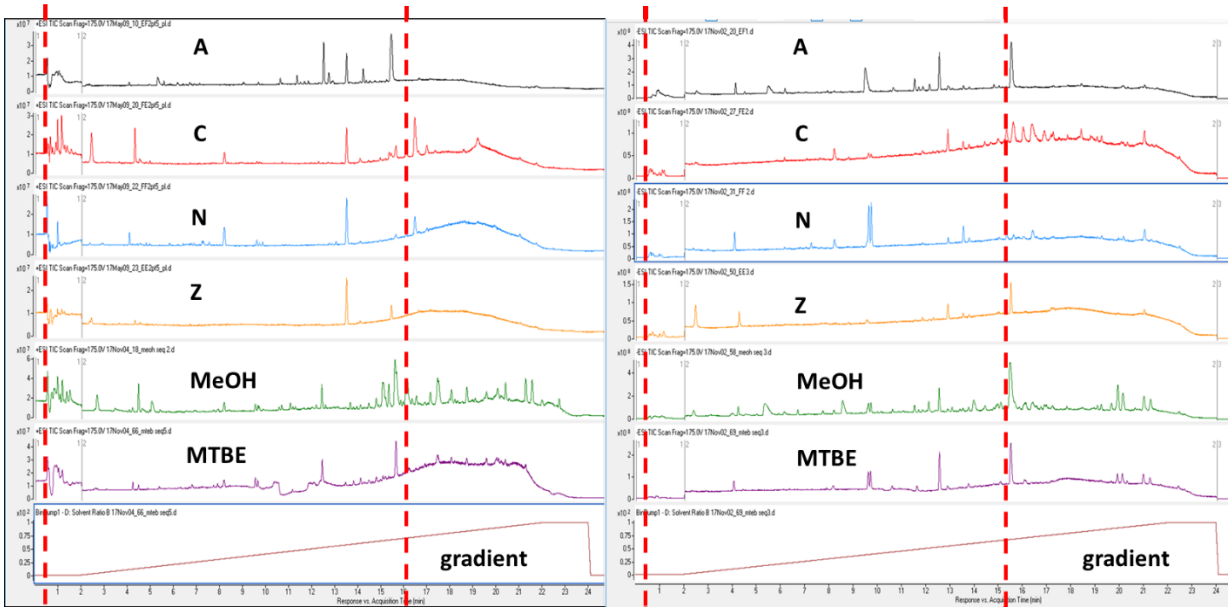


Figure 3.4. TIC profiles of RP LC-MS mid-polar analysis in (+ve) ESI (left panel) and (-ve) ESI (right panel) for plasma metabolites detected in MTBE NP(designated by "MTBE"), aqMTBE-MeOH (designated as "MeOH"), and sSPE fractions: anion (A), cation (C), neutral (N) and zwitterion (Z). The "gradient" panels show increase in mobile phase B, which is equal to 72% at 15.2 min of RT designated by the red dashed line. Vertical red dashed lines show the RT region used for the mid-polar data analysis.

The targeted analysis has also revealed insufficient representation of the cation group of standards (only TRH), which precluded the analysis of the fractionation of this group. The analysis of the fractionation will be investigated in further experiments with an extended standard mix and in more detail (**Section 3.3.2.2**).

The global analysis demonstrated the increase of metabolome coverage in sSPE fractions for 1.5 – fold in (+ve) and 1.4 – fold in (-ve) ESI on RP C18 (RP from here and further) compared to aqMTBE-MeOH (**Appendix B, Supplementary Figure B3**). This confirms our expectations and provided merits for further development of sequential sample preparation. In addition, the comparison of RP TIC profiles of individual sSPE fractions, aqMTBE-MeOH, and MTBE NP extracts clearly demonstrates a limited selectivity of the current design. At first, RP analysis demonstrated a limited selectivity toward polar metabolites. A notable surface area is covered by TIC between 0.5 and 2 min of RT, disregarding extraction type, and ESI mode (**Figure 3.4**). In the second, sSPE fractions demonstrated presumably low recovery of hydrophobic compounds. Profiles of TIC in A, Z, and N fractions are very low at RT > 15.2 min (> 72% MeCN) compared to the sSPE predecessor, i.e., aqMTBE-MeOH.

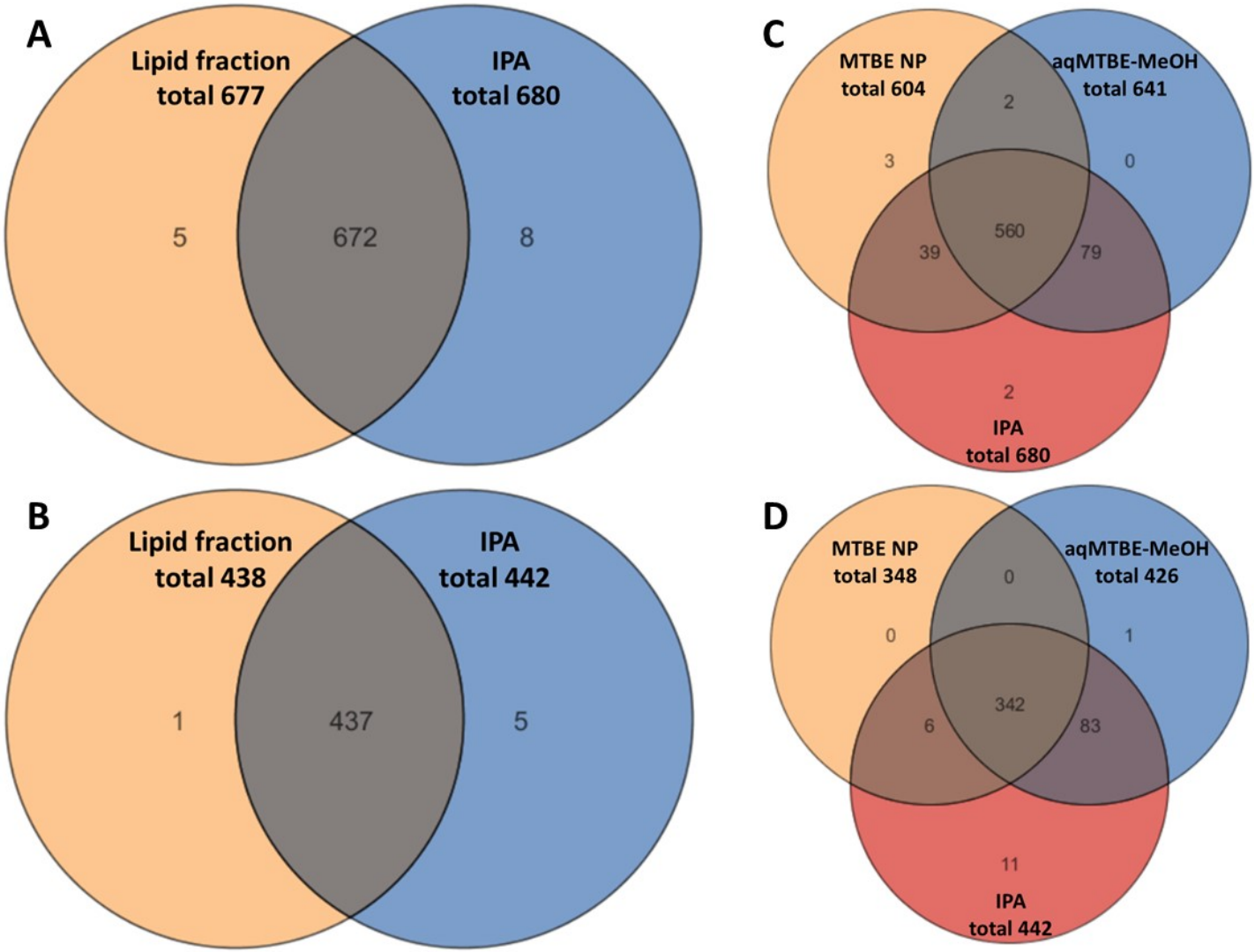


Figure 3.5. Analysis of lipid coverage in "MTBE NP" and "aqMTBE-MeOH", lipid fraction and IPA extract in (+ve) ESI (panels A, C) and (-ve) ESI (panels B, D) using untargeted lipidomic analysis on LTQ-Orbitrap Velos.

These problems will be addressed in more detail in **Section 3.3.2.2.** Overall, the evaluation of the initial sequential sample preparation protocol towards non-lipid metabolites demonstrated increased mid-polar metabolome coverage by sSPE compared to the predecessor: aqMTBE-MeOH extract.

Fractions A, C, N, and Z demonstrated orthogonal composition in a global metabolomic analysis. The results were supported by the expected fractionation of standard metabolites in a targeted analysis. The experimental data supported the continuation of the further development and evaluation of sequential sample preparation.

3.3.1.3 Lipid coverage in MTBE non-polar fraction

The first objective of lipidome analysis was to establish if MTBE NP (organic phase of MTBE LLE) may provide sufficient lipidome coverage comparable to an established lipidomics method for plasma, such as IPA.[199] It was also unclear if sSPE may provide an improved lipidome coverage compared to its predecessor, aqMTBE-MeOH.

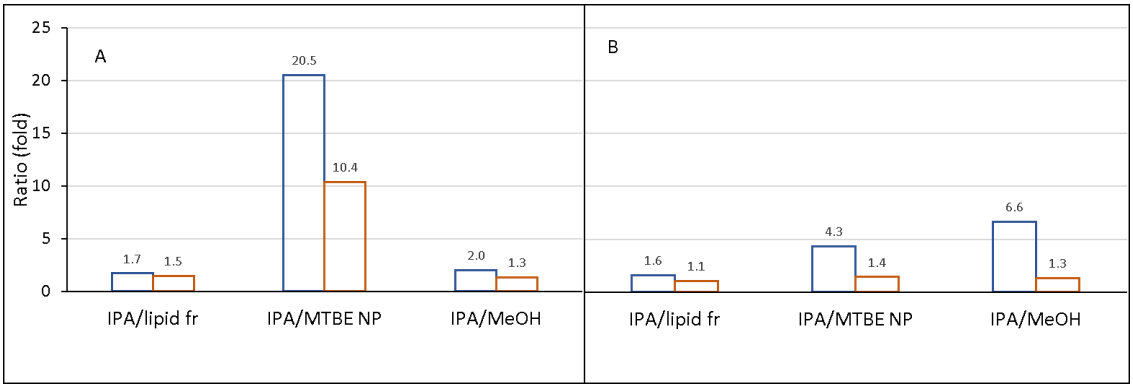


Figure 3.6. Assessment of intensities of lipids in (+ve) (A) and (-ve) (B) ESI of MTBE NP ("MTBE NP"), aqMTBE- MeOH ("MeOH"), and the lipid fraction ("lipid fr."). Blue bars represent average and orange bars - median (n=3) of signal ratios for lipids in three sequential extracts relative to IPA.

Therefore, a lipid analysis was executed to assess metabolome coverage and relative analysis of signals between sequential extracts (including sSPE fractions) and IPA precipitation, which is considered a reference method for comprehensive and reproducible lipidome coverage in plasma.[199] The comparison of lipidome coverages in MTBE NP, aqMTBE-MeOH, their mix ("lipid fraction"), and IPA precipitation is shown in **Figures 3.5 C and 3.5 D.** The coverage in MTBE NP and aqMTBE-MeOH was found to be similar to IPA and is confirmed by the analysis of the lipid

fraction, which was demonstrated similar to IPA lipid composition (**Figure 3.5 A and 3.5 B**). At the same time, the recovery of lipids was lower in MTBE NP and aqMTBE-MeOH compared to IPA (**Figure 3.6 A and B**), but it was restored in their mix ("lipid fraction") disregarding ESI mode. Nevertheless, the recovery in the lipid fraction was lower than in IPA, which is demonstrated by high average IPA/lipid fraction signal ratios of 1.7 and 1.6 in (+ve) and (-ve) ESI, respectively (**Figure 3.6**).

*Table 3.1. Lipidome coverage and signal reproducibility in extracts of sequential sample preparation in nonpolar MTBE fractions (MTBE_NP), MeOH extract of polar MTBE fraction (aqMTBE-MeOH), their mix (Lipid fraction) and in IPA in (+ve) and (-ve) ESI. Signals of lipids detected in all replicates were curated manually prior to the analysis to remove false-positive integrations and duplicates. After the removal of irreproducible (CV > 30%) signals, the median CV was calculated (third row). Furthermore, signals with CV < 30% were removed if their intensities were not five times higher than in correspondent blank extracts (the lowest row). Low coverage for the lipid fraction in this column in (+ve) ESI has been caused by the failed injection of one of three replicates and explained in detail in the text (**Section 3.3.1.1**). The results of lipidomic analysis MTBE NP fraction demonstrated drastic (10.4x) loss of recovery in (+ve) ESI (**Figure 3.5A**) compared to IPA. This indicates that the de-lipidation step is not as successful as predicted and suggests this step may be omitted from the finalized protocol.*

Data cleanup stages	(+ve) ESI				(-ve) ESI			
	aqMTBE-MeOH	MTBE NP	Lipid fraction	IPA	aqMTBE-MeOH	MTBE NP	Lipid fraction	IPA
Number of identified lipids after manual curation	641	604	677	680	426	348	438	442
Number of lipids with CV≤30%	538	540	475	581	338	250	403	385
Median CV of remaining lipids (%)	17.4	15	10.1	17.3	19.9	20.5	15	14
Number of lipids with signal to blank ratio ≥ 5	185	186	130	264	295	195	362	348

These differences in signal intensities can be explained by (i) lower lipid extraction efficiency of MTBE LLE and aqMTBE-MeOH compared to IPA and (ii) more complex manipulation and higher non-specific sample losses during the production of the lipid fraction. Surprisingly, the repeatability was not significantly affected in the lipid fraction. For example, the number of lipids with CV ≤30% in (+ve) ESI was 475 (median RSD 10.1%) in the lipid fraction and 581 (median RSD 17.3%) in IPA. In (-ve) ESI, the number of lipids with CV ≤30% was 403 (median RSD 15.0%) in the lipid fraction and 385 (median RSD 14.0%) in IPA extract, respectively (**Table 3.1**).

The very low number of signals of non-blank lipids in the lipid fraction in (+ve) ESI (**Table 3.1**) has been caused by the drop of all intensities in one of three extraction replicates and was not observed in (-ve) ESI in the same replicate indicating a failed injection issue and not an extraction repeatability issue.

Therefore, the analytical performance of the combined lipid extraction from plasma by MTBE NP and aqMTBE-MeOH is approaching the quality of the "gold standard" method, i.e., IPA.

3.3.1.4 Applicability of sSPE for the fractionation of lipids

Sequential SPE did not fractionate lipids into each of 4 fractions. Most of the lipids were split between cation and neutral fractions (**Figure 3.7**). In addition, the lipid recovery in sSPE fractions was lower than in the lipid fraction (**Figure 3.8**). The low recovery is not surprising, considering the hydrophobic nature of phases from which MAX and MCX SPE were manufactured. In addition, the strong binding of lipids to various SPE phases is well documented and is sometimes used for the removal of lipids from non-lipid metabolites in a sample cleanup procedure.[209,210]

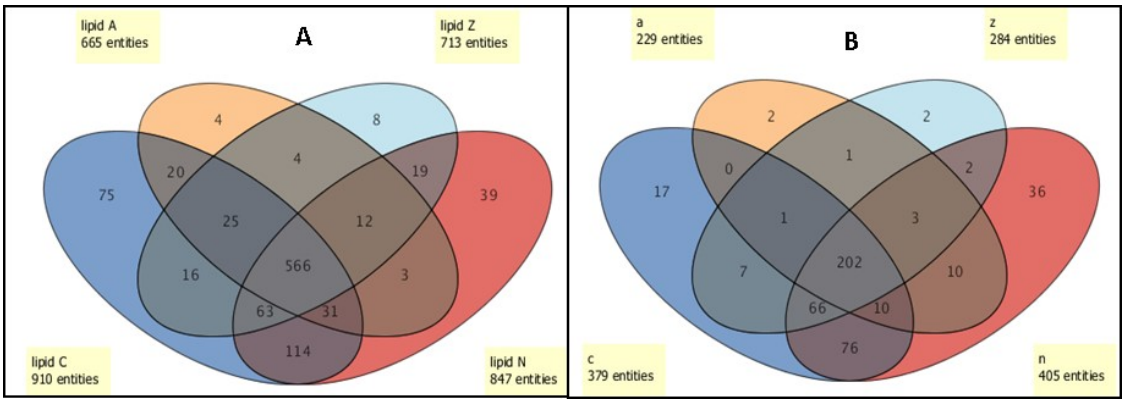


Figure 3.7. Distribution of lipids between fractions of sSPE assessed using positive (A) and negative (B) ESI. Fractions were prepared and analyzed in lipidomic LC-MS at (+ve) and (-ve) ESI, as described in ***Sections3.2.3.4.3 -3.2.3.4.5****.*

Therefore, sSPE demonstrated poor lipid recovery, which confirms the observation of low TIC profiles (**Figure 3.4)** for sSPE fractions on RP LC-MS analysis after 15.2 min (72% MeCN).

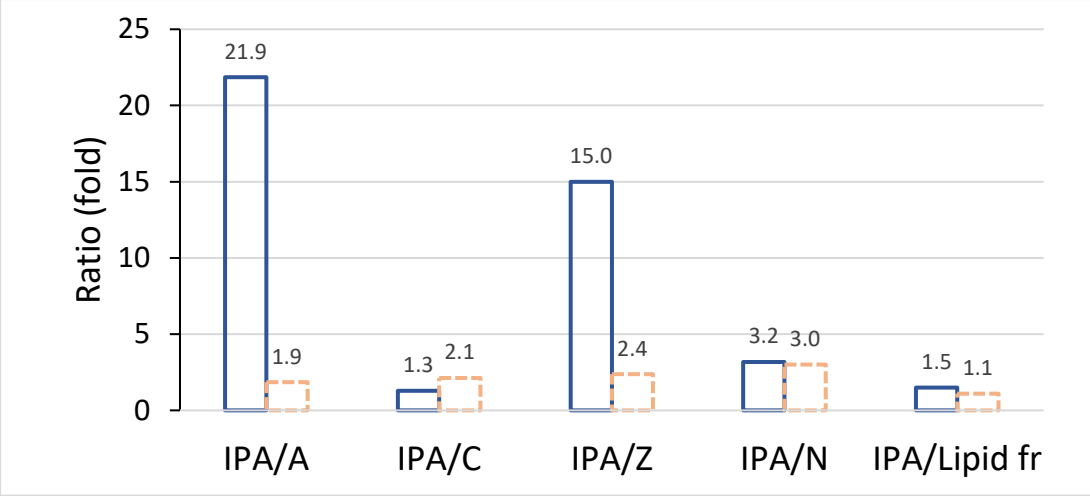


Figure 3.8. Median (n=3) of lipid signal ratios in sSPE fractions and the lipid fraction to IPA in positive (orange, dashed) and negative (blue, solid) ESI. Ratio values are designated by numbers on the top of each bar. Fractions were prepared and analyzed in lipidomic LC-MS, as described in ***Sections 3.2.3.4.3 – 3.2.3.4.5****.*

The most prominent decrease in signal intensities was observed in (-ve) ESI (**Figure 3.8**) for A and Z fractions. Such loss of signals was not observed in aqMTBE-MeOH and pinpoints the possibility for lipid losses in the sSPE step. The losses in A and Z fractions using (-ve) ESI indicate the involvement of fatty acids, phospholipids, and sulfur-containing lipids due to strong/irreversible binding to the MAX phase is a plausible explanation. Another surprising observation was detecting a large number (more than half of all lipids) of di- and triglycerides in sSPE fractions (**Table 3.2**).

Table 3.2. Analysis of signals (median of IPA/fraction ratio) and composition (% of di-(DG) and triglycerides (TG) to all lipids in sSPE fractions. Fractions were prepared and analyzed in lipidomic LC-MS at (+ve) ESI, as described in ***Sections 3.2.3.4.4*** *and* ***3.2.3.4.5****.*

Glycerides in SPE fractions	IPA/A	IPA/C	IPA/Z	IPA/N
DG, TG median signal ratios	3.5	3.5	3.5	3.7
DG TG % to all lipids in fraction	65.2	51.7	62.1	53.4

This coincided with low lipid recovery, which was demonstrated by MTBE NP (**Figure 3.6**) and indicated an inefficiency of MTBE in the depletion of lipids and impaired protection of chromatographic RP columns from compounds with may undergo the irreversible binding with C18 stationary phases. This brings another reason to remove MTBE LLE from the sequential sample process. Finally, the low recovery of lipids in sSPE is not surprising. The strong binding of

lipids to various SPE phases is well documented and is used for the removal of lipids from non-lipid metabolites in a sample cleanup procedure.[209,210]

3.3.2 Optimization and evaluation of sequential sample preparation for analysis of polar and mid-polar metabolites in individual sSPE fractions using targeted and global metabolomics approach

Next, we assessed the fractionation of polar and mid-polar metabolites into sSPE fractions and examined the limitations of the method towards metabolites belonging to various chemical classes. The coverage of the newly designed sSPE method for polar and mid-polar metabolome was then compared with the standard method of MeOH precipitation to see how much the new approach can increase method performance. The effect of the sample loading amount on metabolite coverage was also examined in detail. The third aim was the integration of sSPE fractions with the most suitable LC-MS methods to maximize metabolome coverage in the shortest analytical time.

3.3.2.1 Performance of sequential fractionation method for selected metabolites

The performance of the final sequential fractionation method was first evaluated using 64 standards listed in **Supplementary Table B4 (Appendix B)**. The main evaluation parameters included metabolite coverage and fraction overlap. The ideal result of sSPE fractionation would be a clear fractionation of metabolites (A, C, N, and Z) into correspondent sSPE fractions (**Figure 3.3**). Any splitting <10% was considered negligible and is not further discussed below.

Overall, the results of the analysis of spiked plasma demonstrated orthogonal separation of anionic and cationic metabolites from each other (**Figure 3.9**). However, several anionic metabolites (homovanillic acid, pantothenic acid, arachidonic acid, and biotin) were split between A and N fractions, which may result from incomplete binding of these metabolites to MAX sorbent. This observed split contradicts the pKa values of these anionic compounds (**Appendix B, Supplementary Table B1**), which predicts their complete retention on MAX. Arachidonic and pantothenic acids should be completely protonated at pH 2.4 of MCX (**Appendix B, Supplementary Table B1**), which makes ionic interactions improbable.

Several metabolites (arachidonic acid, eicosapentaenoic acid, pyridoxal, and cholesteryl acetate) were split between A and Z fractions (Fig. 9). This can be caused by the unexpected retention of

these anion metabolites on the MCX phase, except zwitterion pyridoxal, which was just partially retained on MCX against the expected complete retention.

For cations that are comprehensively charged at the pH of MAX (creatinine, neopterin to histamine), unsplit fractionation into C fraction was observed as expected (**Figure 3.9**). Neopterin possesses a weak basic group, behaves as a cation, and was fractionated into the C fraction due to the presence of a positive charge. This behavior was used previously for an SCX chromatography of this compound.[211,212] The results also revealed an unexpected but welcome fractionation of carnitines into the C fraction without any splitting. Carnitines, due to the permanent positive charge of quaternary ammonium, behave as strict cations (i.e., do not bind to MAX) despite the presence of an ionized carboxyl group (**Figure 3.9**). The standards between kynurenine and aspartic acid had predicted pKa values that characterize them clearly as zwitterions. However, these standards were found predominantly in C, instead of the Z fraction (kynurenine to valine). Trans-4-OH-proline and methionine were highly split between C and N fractions, while standards between S-adenosylhomocysteine to aspartic acid shown in **Figure 3.9** were split between C and Z fractions. These results indicate a poor binding of these zwitterions via positive moieties to MCX or poor binding of negative ion moieties to MAX for metabolites with C/N and C/Z split, respectively.

Neutral metabolites were predictably fractionated into N fraction, and the only minor C/N split (< 10%) was observed for guanosine, beta-estradiol, cortisol, and cortisone. All neutral standards demonstrate pKa values distinct from other metabolites, which predicted their neutral charge in both SPE protocols (**Figure 3.9**). The results supported that the primary mechanism of fractionation was ionic and that the designed protocol is working well. The observation of minor splits suggests the minor involvement of other types of interactions (hydrophobic or polar), which provide the retention of neutral metabolites on MCX with subsequent elution into Z fraction.

Only five zwitterion metabolites were exclusively detected in the Z fraction. Such retention is possible only in the case of their complete retention on both MAX and MCX sorbents. The profiles of pKa for these standards do not differ from zwitterions found in C or split into C fraction, which suggests the involvement of other types of interactions in the tested SPE plates. Moreover, hydrophobicity or hydrophilicity alone do not determine such behavior because the metabolites detected in Z exclusively contain distinctively polar (glutamine, $\log D_{pH\ 9.2} = -4.4$) and nonpolar (thyroxine, $\log D_{pH\ 9.2} = 1.9$) standards (**Appendix B, Supplementary Table B4**).

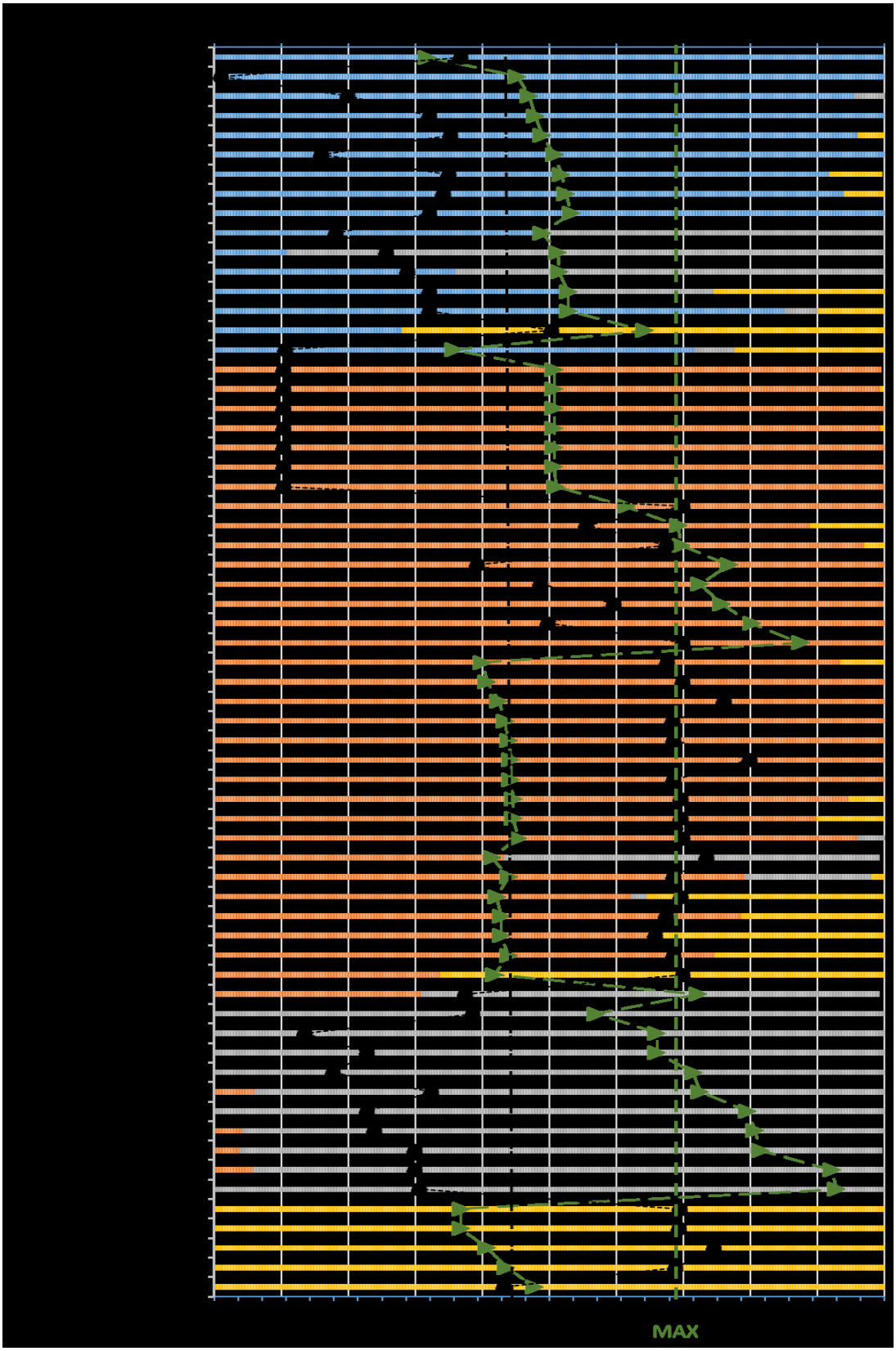

MAX

Figure 3.9. Fractionation of selected metabolites from blood plasma into fractions (blue-anion, orange-cation, grey-neutral and yellow-zwitterion) using sSPE. The length of each bar is proportional to the sum of peak areas (equal to 100%) in four sSPE fractions and is designated by the primary (top) X axis. Peak area of each metabolite in a fraction, where it was detected is represented by the corresponding color-coded portion of a stacked bar. It is necessary to note, that relative proportion of peak areas does not guarantee similar proportion of actual amounts in different fractions due to unknown matrix effects. The secondary X axis (bottom) shows pH for MAX (9.2, green dashed line), MCX (2.4, black dashed line) and pKa's of the strongest acidic (green triangles) and the strongest basic (black round dots) moieties of each metabolite. The location of pKa symbols on pH lines with similar color indicates equal ratio of conjugate base and acids, i.e. equal ratio of charged/neutral moieties. The value of basic pKa's above (to the right) MCX pH, as well as the value of acidic pKa's below (to the left) pH of MAX corresponds to an increased proportion of charged moieties over neutral conjugate and vice-versa. Values of pKas were obtained on Chemicalize™ calculator (https://chemicalize.com/#/calculation) available on the website of ChemAxon in December of 2019.

A broader number of conditions needed to be assessed to understand and predict the fractionation behavior of zwitterions.

The experiment demonstrated five groups of metabolites with distinct pKa profiles (**Figure 3.9**), which encounter one group in A, three groups in C, and 1 group in N. One of three pKa profiles in the C is similar to pKa profiles of metabolites in the Z fraction. The fractionation performance of all metabolites except the group of zwitterions (from alanine to aspartic acid) can be predicted by pKa, based on the set of standard metabolites evaluated. The fractionation behavior of zwitterions was the most complex and is attributed to the presence of multiply charged groups, as it is impossible to find pH values to have all zwitterions in the same overall charge state and minor contribution of mixed-mode interactions with the sorbent. Moreover, not surprisingly, a fractionation behavior of zwitterions could not be fully predicted using simple descriptors such as log P, log D, hydrogen bonding potential, and pKa values (**Appendix B, Supplementary Table B4**). This suggests the necessity of a wider and more comprehensive analysis of the interactions between metabolites and complex, mixed-mode sSPE, which is outside the scope of the current study. More importantly for our objectives, these results confirm good and predictable separation of most cations, anions, and neutrals in their respective fractions, thus achieving our goal of sample decomplexation.

3.3.2.2 Fractionation of unknown metabolites in plasma using sSPE

3.3.2.2.1 *RP analysis in (+ve) ESI*

Next, the fractionation performance of all metabolites present in plasma was examined using an untargeted dataset containing all curated metabolite features, which comply with requirements (**Sections 3.2.4.9 and 3.2.5.4**).

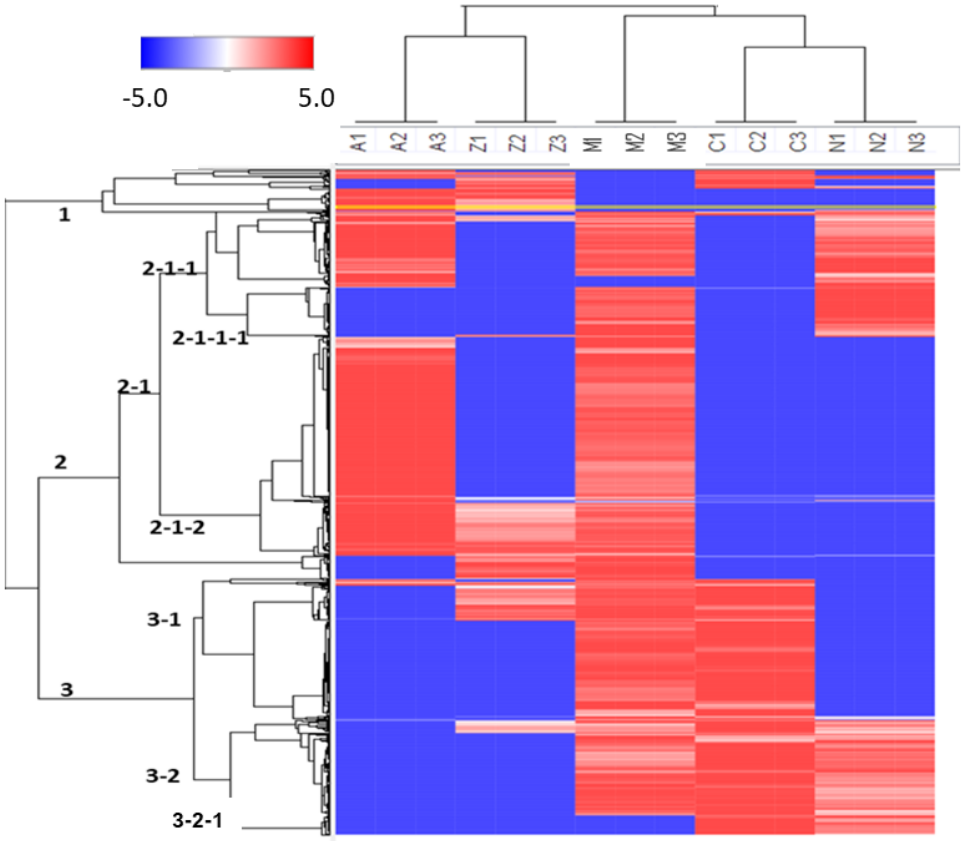


Figure 3.10. Heat map and hierarchical analysis of all metabolites f (in rows) detected in sSPE fractions (columns) and MeOH analyzed by RP (+ve) ESI. Fractions of sSPE are designated by A - anion, C – cation, N- neutral, and Z – zwitterion. MeOH extracts are designated by "M". Extractions and metabolites are clustered using the Euclidian distance and average linkage method. The main metabolites cluster are designated by numbers to facilitate discussion in the text. Normalized metabolite signals are depicted by heat map colors, from dark blue to dark red, representing the scale between -5.0 to 5.0 from not detected to the highest signal, respectively.

This analysis's main objective was to investigate if the trends that were observed for standard metabolites prepared as described in **Section 3.2.4.1** also hold across the full dataset and examine the extent of splitting in the entire dataset. **Figure 3.10** shows the hierarchical clustering analysis of A, C, N, and Z fractions and methanol precipitation, which was also included in the comparison. The results show a clear decrease in the sample complexity of fractions versus methanol extract. The metabolite coverage observed for methanol extract is mostly split between A, C, and N fractions.

The full dataset used for clustering analysis included identified metabolites discussed in **Section 3.3.2.1**. Unknown metabolites that cluster with known identified metabolites are likely to have similar physicochemical properties. Also, **Figure 3.10** visually illustrates several clusters where fractions improve the metabolite coverage in comparison to methanol: cluster: (1) 87 metabolites; (2-1-1) 44 metabolites and (3-2-1) with 44 metabolites. Just a few N metabolites in cluster 1 were unique to N, which others were detected in ionic fractions, which suggests a negligible role of neutral fraction in the expansion of metabolome coverage (**Figure 3.10**). In addition, (+ve) RP LC-MS analysis did not reveal any metabolites which were uniquely detected in MeOH, showing that the proposed fractionation was able to recapitulate the composition of typical methanol plasma extract fully.

3.3.2.2.2 *RP analysis in (-ve) ESI*

Two major clusters are formed by metabolites in RP neg ESI (**Figure 3.11**). Similar to what was observed for selected known metabolites in **Section 3.3.2.1** and demonstrated in **Figure 3.9**, this analysis also confirmed that a very small number of metabolites were split between A and C or N and Z fractions and that these fraction pairs were highly orthogonal**.**

Compared to MeOH, sSPE enhances metabolite coverage in three sub-clusters. The sub-cluster 1-1-1-1 contains 64 compounds detected in A and N fractions, including alfa-ketoglutaric and pantothenic acids. Sub-cluster 1-2 includes 94 compounds detected in A, Z, and C fractions, including phenylalanine. The third cluster (#2-2-2-1) contains 60 metabolites (no standards were detected in this cluster) in C, Z, and N fractions, most of which were split between C and N.

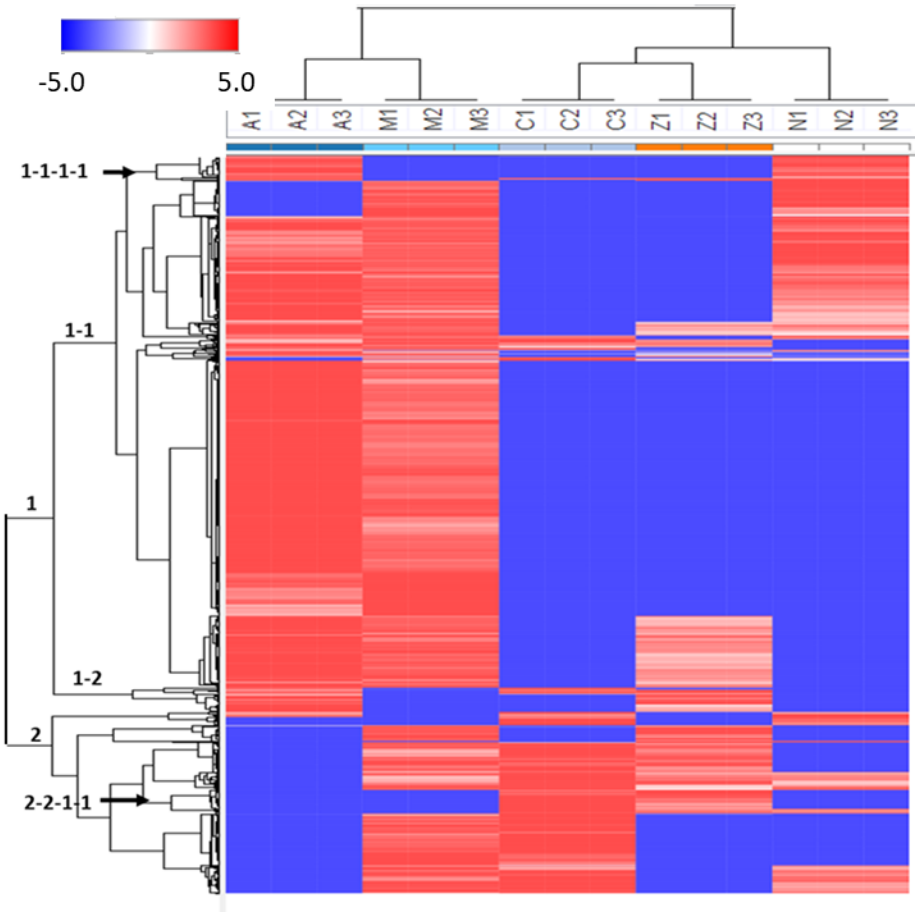


Figure 3.11. Heat map of metabolites (in rows) in fractions (columns) of sSPE and MeOH analyzed by RP (-ve) ESI. Fractions of sSPE are designated by A - anion, C – cation, N- neutral, and Z – zwitterion. MeOH extracts are designated by "M". Extractions and metabolites are clustered using the Euclidian distance and average linkage method. Metabolites clusters which are discussed in the text designated by numbers. Normalized metabolites signals depicted by colors, representing the scale between -5.0 to 5.0 from the dark blue for not detected to the dark red for the highest signal.

3.3.2.2.3 *ZIC-HILIC analysis in (+ve) ESI*

Two major clusters are formed by metabolites, which were detected in ZIC-HILIC at (+ve) ESI (**Figure 3.12**).

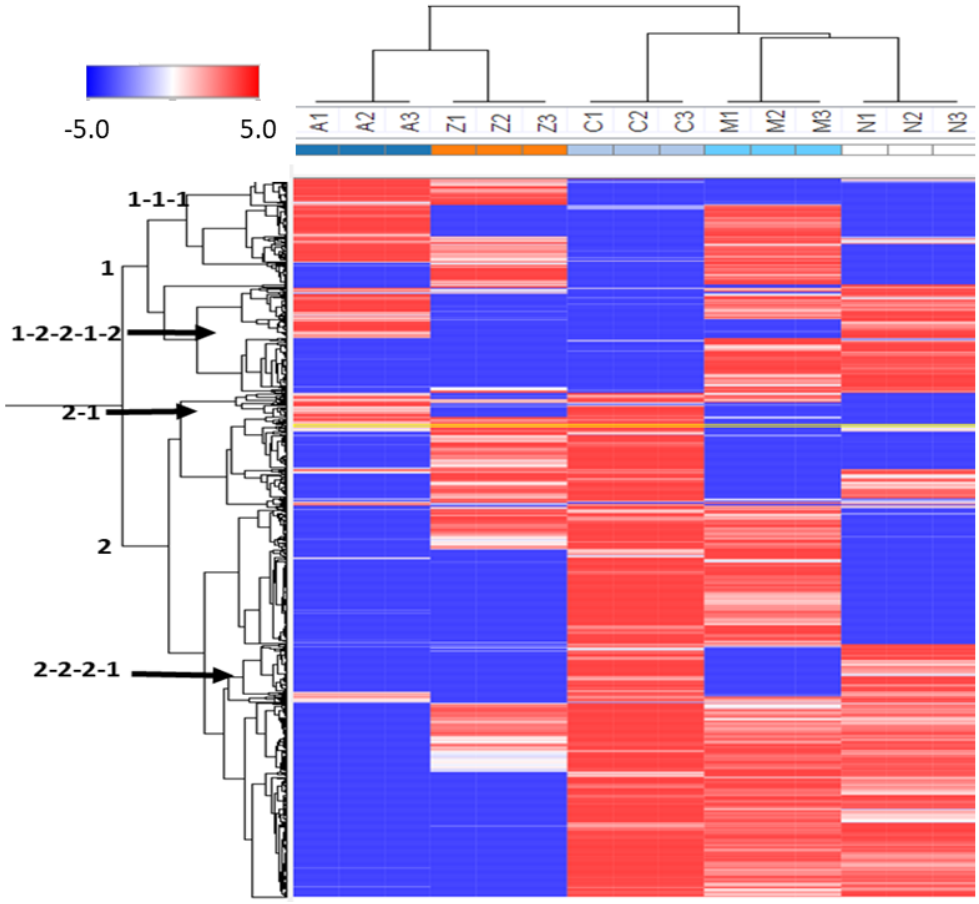


Figure 3.12. Heat map of metabolites (in rows) in fractions (columns) of sSPE and MeOH analyzed by ZIC-HILIC (+ve) ESI. Fractions of sSPE are designated by A - anion, C – cation, N- neutral, and Z – zwitterion. MeOH extracts are designated by "M". Extractions and metabolites are clustered using the Euclidian distance and average linkage method. Metabolites clusters which are discussed in the text designated by numbers. Normalized metabolites signals depicted by colors, from the dark blue to dark red, representing the scale between -5.0 to 5. from the dark blue for not detected to the dark red for the highest signal.

These sub-clusters were: (1-1-1) with 46 compounds and A/Z split; (1-2-2-1-2) with 34 metabolites and A/N split; (#2-1) with 178 metabolites, including cation S-adenosylhomocysteine and zwitterions alanine, serine and C/Z split, and (2-2-2-1) with 90 metabolites including guanosine and trans-4-hydroxy-proline and split into C/N. These results further confirm that N fraction is only a minor contributor to expand metabolome coverage beyond the metabolites that can be observed using methanol precipitation.

3.3.2.2.4 *ZIC-HILIC analysis in (-ve) ESI*

Two major clusters are observed (**Figure 3.13**) and formed by metabolites detected in A and C fractions. In a ZIC-HILIC (-ve) ESI mode the sSPE expands the coverage over methanol precipitation as shown by four sub-clusters: 1-1-1-2, 1-2-2, 2-1-1-2-1, and 2-2, respectively. Standards (asparagine, aspartic acid, citrulline) were found only in 2-1-1-2-1 and were split between C and Z fractions as all other metabolites in the cluster.

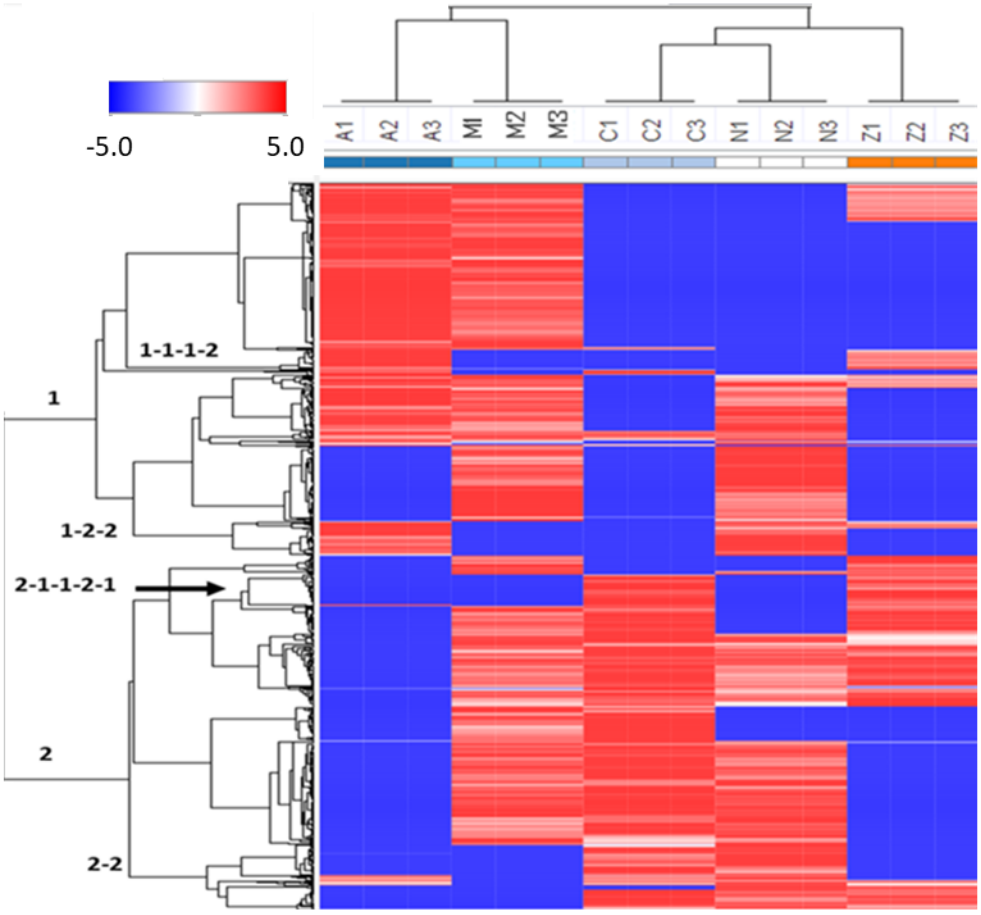


Figure 3.13. Heat map of metabolites (in rows) in fractions (columns) of sSPE and MeOH analyzed by ZIC-HILIC (-ve) ESI. Fractions of sSPE are designated by A - anion, C – cation, N- neutral, and Z – zwitterion. MeOH extracts are designated by "M". Extractions and metabolites are clustered using the Euclidian distance and average linkage method. Metabolites clusters which are discussed in the text designated by numbers. Normalized metabolites signals depicted by colors, from the dark blue to dark red representing the scale between -5.0 to 5.0 from the dark blue for not detected to the dark red for the highest signal.

3.3.2.2.5 *Quality of metabolite fractionation in sSPE across the entire metabolite set*

The global analysis reported in **Sections 3.1.2.1 to 3.1.2.4** has confirmed a small overlap between anion and cation fractions, confirming the results of our targeted analysis that anion and cation fractionation works very well. In fact, as summarized in **Table 3.3**, more than a half of metabolites were detected in only one fraction (**Table 3**), with 61.8-62.4% and 51.9-58.5% metabolites detected in a single fraction in RP and ZIC-HILIC analysis, respectively. The most common overlaps were observed between 2 fractions and ranged from 31.0% in (-ve) RP to 35.8% in (+ve) ZIC-HILIC (**Table 3**). The number of metabolites, which overlapped between 3 and more fractions, were higher in ZIC-HILIC (12.3-12.6%) than in RP analysis (4.1-5.7%). The overlap values do not take signal intensity into account, so even a minor detectable signal is considered as overlap, but could be due to a negligible split for the majority metabolites. Further assessment of split signals was performed to reveal how profound the splitting phenomena is, and the necessity to prioritize the correction of split over other issues with the method. To further investigate how many of metabolites demonstrated minor (negligible) splits, the ratio of the area of a minor component peak area to the total sum peak area was calculated. These results are shown in **Table 3.3.** The proportion of metabolites with a negligible split in 2 fractions was small and ranged between 29.4% (+ve) ZIC-HILIC to 16.5% in (-ve) ZIC-HILIC, which suggests that the majority of split metabolites experienced a profound split with the smaller signal > 10% of the total one. Metabolites detectable in (-ve) ESI demonstrated even a larger proportion of metabolites with profound split, compared to (+ve) ESI. This may indicate a more profound split of in-solution anion compared to in-solution cation metabolites.

In addition, the hierarchical cluster analysis demonstrated a considerable C/N overlap of metabolites in three LC-MS analyses, except for RP (-ve) ESI, where A/N overlap exceeds C/N (**Appendix B, Supplementary Table B5**). However, the most massive split was demonstrated by the Z fraction.

The drastic split of metabolites between fractions may affect relative and absolute quantitation if the splitting is not reproducible. There is also limited gain in analyzing fractions that do not contain many unique metabolites, such as Z fraction, opening up the necessity to combine fractions strategically.

Table 3.3. Analysis of the overlap of metabolites in sSPE fractionation. The number of metabolites detected in one (Not split), and 2, 3, and 4 sSPE fractions were obtained from Venn analysis. The proportion of split metabolites (%) to the total number of metabolites (100%) in the correspondent fractions were calculated for each LC-MS analysis. The split was considered to be negligible if the ratio of the signal in one of two overlapped fractions to the total signal in both fractions was ≤ 0.1. The proportion (%) of metabolites was calculated only for metabolites split into 2 fractions, assuming the total, non-redundant number of metabolites in both fractions = 100%.

(+ve) RP	Number of metabolites	% of metabolites	(-ve) RP	Number of metabolites	% of metabolites
Split in 2	660	32.3	**Split in 2**	745	31
Negligible split in 2	191	28.9	**Negligible split in 2**	141	18.9
Split in 3	74	3.6	**Split in 3**	124	5.2
Split in 4	11	0.5	**Split in 4**	10	0.4
Not split	1262	61.8	**Not split**	1526	63.5
Total	**2042**	**100**	**Total**	**2405**	**100**
(+ve) ZIC-HILIC	**Number of metabolites**	**% of metabolites**	**(-ve) ZIC-HILIC**	**Number of metabolites**	**% of metabolites**
Split in 2	646	35.8	**Split in 2**	510	31.2
Negligible split in 2	190	29.4	**Negligible split in 2**	84	16.5
Split in 3	202	11.2	**Split in 3**	168	10.3
Split in 4	20	1.1	**Split in 4**	42	2.6
Not split	935	51.9	**Not split**	956	58.5
Total	**1803**	**100**	**Total**	**1634**	**100**

3.3.2.2.6 *Combination of sSPE fractions for maximum metabolome coverage in the shortest analysis time*

The desire to repair the split of metabolites between individual fractions **(Section 3.3.2.2.4)** coincides with the need to combine fractions to shorten an analysis time. The split between fractions varies between LC-MS methods (**Appendix B, Supplementary Table B5**). Subsequently, LC-MS analysis should be executed by combining the most affected fractions: (i) C&Z fractions in RP (-ve) with the subsequent combination of A&N fractions and (ii) C&N fractions in RP (+ve) and both ZIC-HILIC LC-MS methods with the subsequent combination of A&Z fractions. The theoretical analysis of the composition of combined sSPE fractions indicated a plausible increase of metabolome coverage compared to MeOH (**Figure 3.4**)

3.3.2.2.7 *Effect of sample load on metabolome coverage and signal intensities using sSPE*

The reduction of sample complexity in sSPE fractions opened the door to increase the amount of analyzed material.

Table 3.4. Comparison of metabolome coverage in MeOH and predicted coverage in sSPE fraction combinations using different LC-MS methods. The table summarizes the number of metabolites detected in extracts of MeOH (at the load equivalent to 1.6 µL plasma) and the theoretical coverage in the proposed combination of sSPE fractions (at the load equivalent to 3.2 µL plasma) obtained from Venn analysis of individual fractions in the correspondent LC-MS methods. The number of "Uniquely detected" in all sSPE combinations has been obtained from Venn analysis of the combination against MeOH. "Uniquely detected" for MeOH were obtained from Venn analysis of MeOH against CZ+AN for (-ve) RP or CN+AZ combination for other LC-MS methods. The total number of unique metabolites across all LC-MS methods is the sum of unique metabolites in each LC-MS method and may be redundant due to the absence of confident tools for matching global metabolites between different LC methods and different ESI regimes.

LC-MS analysis	**Metabolites**	**Fractions**			
RP (+ve) ESI		**MeOH**	**CN**	**AZ**	**CN+AZ**
	Total	1648	1070	1246	2019
	Uniquely detected	264	264	411	638
RP (-ve) ESI		**MeOH**	**CZ**	**AN**	**CZ+AN**
	Total	2022	989	1893	2383
	Uniquely detected	409	527	354	806
ZIC-HILIC (+ve) ESI		**MeOH**	**CN**	**AZ**	**CN+AZ**
	Total	1052	1398	844	1772
	Uniquely detected	91	566	383	811
ZIC-HILIC (-ve) ESI		**MeOH**	**CN**	**AZ**	**CN+AZ**
	Total	1041	1116	946	1627
	Uniquely detected	171	497	394	757
Sum of uniquely detected metabolites across LC-MS methods		**935**			**3012**

The increase of the metabolome coverage with the analyzed sample amount was investigated in sSPE and compared to MeOH. Doubling the amount of biological material injected from 0.8 to 1.6 µL plasma equivalents for MeOH increased the number of metabolites detected by 299 and 388 metabolites in RP (+ve) ESI and (-ve) ESI (**Table 3.5**).

A further increase to 3.2 µL of plasma only increased metabolite coverage by less than 150 additional metabolites (**Table 3.5**). For sSPE, predicted metabolite coverage combined in Venn analysis of fractions A and C also showed only minor increases in the number of detected

metabolites between 3.2 and 12.8 µL loads. However, theoretical, combined metabolite coverage in fractions N and Z demonstrated a much higher increase in the number of metabolites.

For example, as shown in **Table 3.5**, at 6.4 µL load, metabolite coverage of fraction N increased by 268 and 661 metabolites for (+ve) and (-ve) ESI, respectively. For fraction Z, the coverage increased moderately by 134 and 237 metabolites for (+ve) and (-ve) ESI, respectively.

Table 3.5. Metabolome coverage in MeOH and sSPE fractions at different loading amounts of material of blood plasma analyzed using (+ve) (A) and (-ve) RP analysis. The number of metabolites in combined fractions are non-redundant. The sample load indicates the load of material equivalent to the amount in the original plasma volume (µL). The experiment was executed as described in ***Section 3.2.5*** *and generated four sSPE fractions designated by A for anion, C – cation, N – neutral, and Z – zwitterion. Note, that due to a high overlap between A and Z or C and N, metabolite coverage in theoretically (Venn analysis) reconstructed (CN, AZ, CZ, AN) fractions is lower than in LC-MS analysis of individual fractions.*

A. Number of metabolites in (+ve) RP ESI							
Sample load (µL)	**MeOH**	**A**	**C**	**N**	**Z**	**CN**	**AZ**
0.8	1349						
1.6	1648						
3.2	1775	1008	788	616	436	1070	1246
6.4		1076	778	750	570	1153	1329
12.8		1122	825	886	706	1266	1410
B. Number of metabolites in (-ve) RP ESI							
Sample load (µL)	**MeOH**	**A**	**C**	**N**	**Z**	**CZ**	**AN**
0.8	1634						
1.6	2022						
3.2	2162	1556	601	684	587	989	1893
6.4		1605	680	1117	815	1485	1979
12.8		1676	733	1248	1023	1592	2076

The further increase in the load to 12.8 µL demonstrated the saturation in metabolite coverage, which implies that the maximum sample loading for RP analysis should be equivalent to 6.4 µL of plasma. The analysis of MeOH extracts on ZIC-HILIC (**Table 3.6**) demonstrated a negligible increase in metabolome coverage in both ESI modes between 0.8 and 1.6 µL loads and saturation between 1.6 and 3.2 µL.

This limits the loading amount of MeOH extract onto the ZIC-HILIC column to 1.6 µL plasma. In sSPE, increasing the loading from 3.2 and 6.4 µL showed no improvement in metabolite coverage, and a further increase to 12.8 µL showed only a very negligible increase in metabolite coverage.

As expected, when the sample loading amount was increased, the signal intensities of metabolites also generally increased for all LC-MS methods tested (**Figure 3.14**). The rate of the signal increase

was higher between 3.2 and 6.4 µL and slowed down or stopped between 6.4 and 12.8 µL, indicating some metabolites reached saturation.

Table 3.6. Metabolome coverage in MeOH and sSPE fractions at different loading amounts of material of blood plasma analyzed using (+ve) (A) and (-ve) ZIC-HILIC analysis. The number of metabolites in combined fractions are non-redundant. The sample load indicates the load of material equivalent to the amount in the original plasma volume (µL). The experiment was executed as described in Section 3.2.5 and generated four sSPE fractions designated by A for anion, C – cation, N – neutral, and Z – zwitterion. Note, that due to a high overlap between A and Z or C and N, metabolite coverage in theoretically (Venn analysis) reconstructed (CN, AZ, CZ, AN) fractions is lower than in LC-MS analysis of individual fractions.

A. Number of metabolites in (+ve) ZIC-HILIC							
Sample load	**MeOH**	**A**	**C**	**N**	**Z**	**CN**	**AZ**
0.8	950						
1.6	1052						
3.2	1082	481	1074	856	502	1398	844
6.4		457	1067	890	537	1416	853
12.8		505	1064	904	578	1427	910
B. Number of metabolites detected in (-ve) ZIC-HILIC							
Sample load	**MeOH**	**A**	**C**	**N**	**Z**	**CN**	**AZ**
0.8	875						
1.6	1041						
3.2	1100	673	657	842	434	1116	946
6.4		711	694	749	496	1087	1008
12.8		748	718	934	513	1198	1041

Metabolites in the Z demonstrated the lowest signal compared to other metabolite classes at the same loads. In addition, RP analysis demonstrated lower signals for metabolites in the Z, than ZIC-HILIC at the same loads (**Figure 3.14**). Signals of anions and zwitterions were drastically improved in 6.4 µL loads, compared to 3.2 µL for both RP and ZIC-HILIC analysis. Thus, taking into account both metabolite coverage and signal saturation, the maximum sample load equivalent to material in 6.4 µL of original plasma was selected for the final sSPE protocol.

Moreover, ZIC-HILIC appears to be the method of choice for LC-MS analysis of zwitterion and neutral fractions because it provides signal intensities closer to what is observed in MeOH extracts than RP in either ESI mode. Sample loading for sSPE fractions should be established at 6.4 µL loads to achieve high signal intensities and good metabolome coverage. Sample loading for MeOH extracts should be limited to 1.6 µL because the increase of metabolome coverage is saturated

between 1.6 and 3.2 µL. Besides, higher loads of MeOH extracts increase the chance to damage chromatographic columns because they contain drastically higher amounts of non-polar metabolites, including triglycerides and phospholipids, than SPE fractions (**Figure 3.8**). These compounds may bind irreversibly and/or block chromatographic columns.

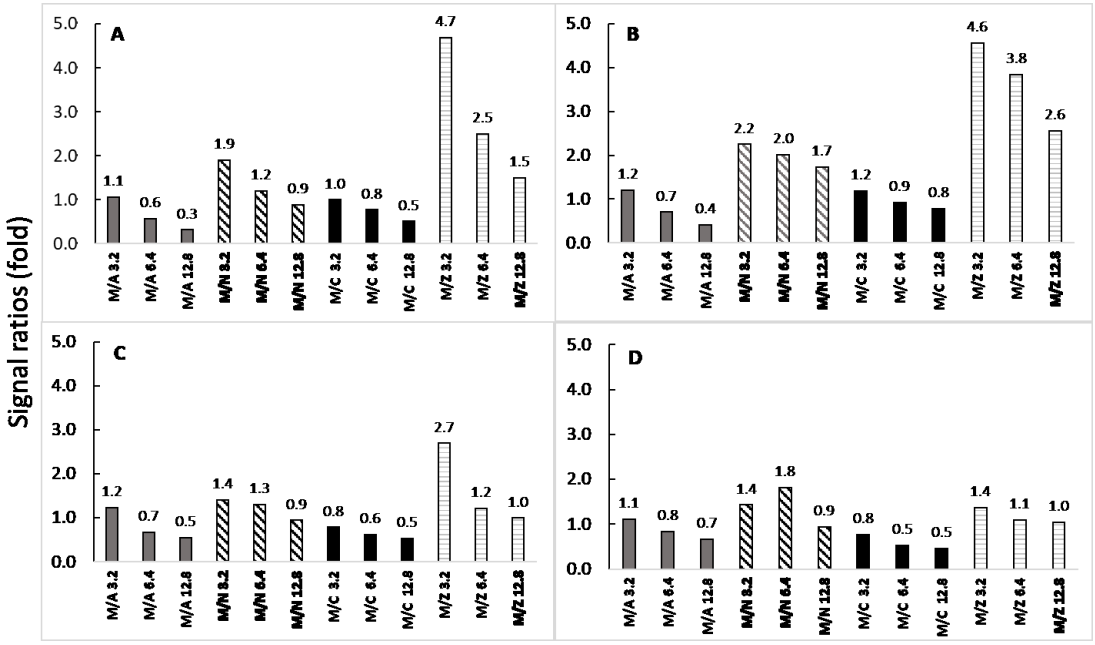


Figure 3.14. Analysis of signal intensities in sSPE fractions on RP at (+ve) ESI (A), (-ve) ESI (B), (+ve) ZIC-HILIC (C) and (-ve) ZIC-HILIC (D). Ratios (MeOH/sSPE fraction) of signal intensities designated by Y-axis for identical metabolites in MeOH (M) at the load of 3.2 µL and an individual sSPE fraction (A, C, N or Z) were calculated for each of three (3.2, 6.4 and 12.8 µL) loads of sSPE. Median values for each sSPE fraction/load are displayed on top of the bars.

3.3.3 Assessment of the analytical performance of the final fractionation protocol in combined sSPE fractions using targeted and global metabolomics approaches

The goal of these experiments was to establish how well the final sSPE protocol works against MeOH precipitation, one of the most commonly employed methods in untargeted metabolomics. Fractions of sSPE (n=6) were prepared and analyzed as described in **Section 3.2.5.3** in combinations and in amounts recommended in **Sections 3.3.2.2.5 and 3.3.2.2.6**. Three main parameters were evaluated: (i) method precision as evaluated by signal repeatability of spiked standards; (ii) extraction recovery and the splitting of spiked standards between sSPE fractions and (iii) matrix effects in combined SPE fractions and MeOH extracts.

3.3.3.1 Analysis of the final protocol using a targeted approach

3.3.3.1.1 *Extraction repeatability of spiked metabolites*

Fractionation of metabolites was assessed by the fact of detection in one or more (splitting) fractions. In the case of metabolite split, a preferable fraction was determined by a higher recovery. Seven standards were found split between combined sSPE fractions. The inter-fraction split with similar recovery between CN and AZ was demonstrated by homovanillic acid (d_3), phenylalanine "SIL", and biotin with recoveries of 32.2 and 39.3, 16.6 and 11.4 and 13.3 and 5.0%, respectively. These results match very well with the results shown in **Figure 3.9,** where homovanillic acid demonstrated an even split between A and N fractions. In contrast, biotin showed higher intensity in N fraction versus A fraction. Similarly, phenylalanine showed a split between C and Z fractions with higher intensity in C fraction. The inter-fraction split with high recovery in one fraction and easy designation of a preferential fraction was demonstrated by deoxycholic acid (recovery of 1.2 and 51.5% in CZ and AN, respectively) and pyridoxin, neopterin, and tyrosine "SIL" with CN and AZ recoveries of 40.4 and 2.0, 28.7 and 5.5 and 25.6 and 7.8, respectively (**Figure 3.9**). Besides, the study confirmed fractionation preferences obtained in the previous experiment for the most metabolites (**Figure 3.9** and **Appendix B, Supplementary Table B7**).

The recovery of the sSPE method was calculated in plasma samples (n=6) spiked with the standard mix before MeOH extraction against samples (n=3), which were spiked after extraction after subtraction of the average if signals in background samples (n=3) as described in detail in **Section 3.2.5.2** and **3.2.5.3**. For the composition of the mix and concentration of standards, please see **Appendix B, Supplementary Table B7.** Out of 61 spiked metabolites, eight (kynurenine, glycine "SIL", 3-hydroxy-DL-kynurenine, threonine "SIL", aspartic acid "SIL", saccharopine, lysine "SIL", and arginine "SIL") were not detected in plasma samples but were detected as neat standards in the reconstitution solvent, which was used as a 100% reference for matrix effects analysis (**Section 3.2.5.3)**. Out of 53 metabolites detected in plasma samples, a similar detection rate (45 were detected in sSPE and 43 in MeOH) were observed. The majority of metabolites in sSPE on RP analysis demonstrated recovery higher than 50% in both ESI modes (**Figure 3.16**). The exception was five metabolites (arachidonic and taurocholic acids, melatonin, resveratrol, and biotin).

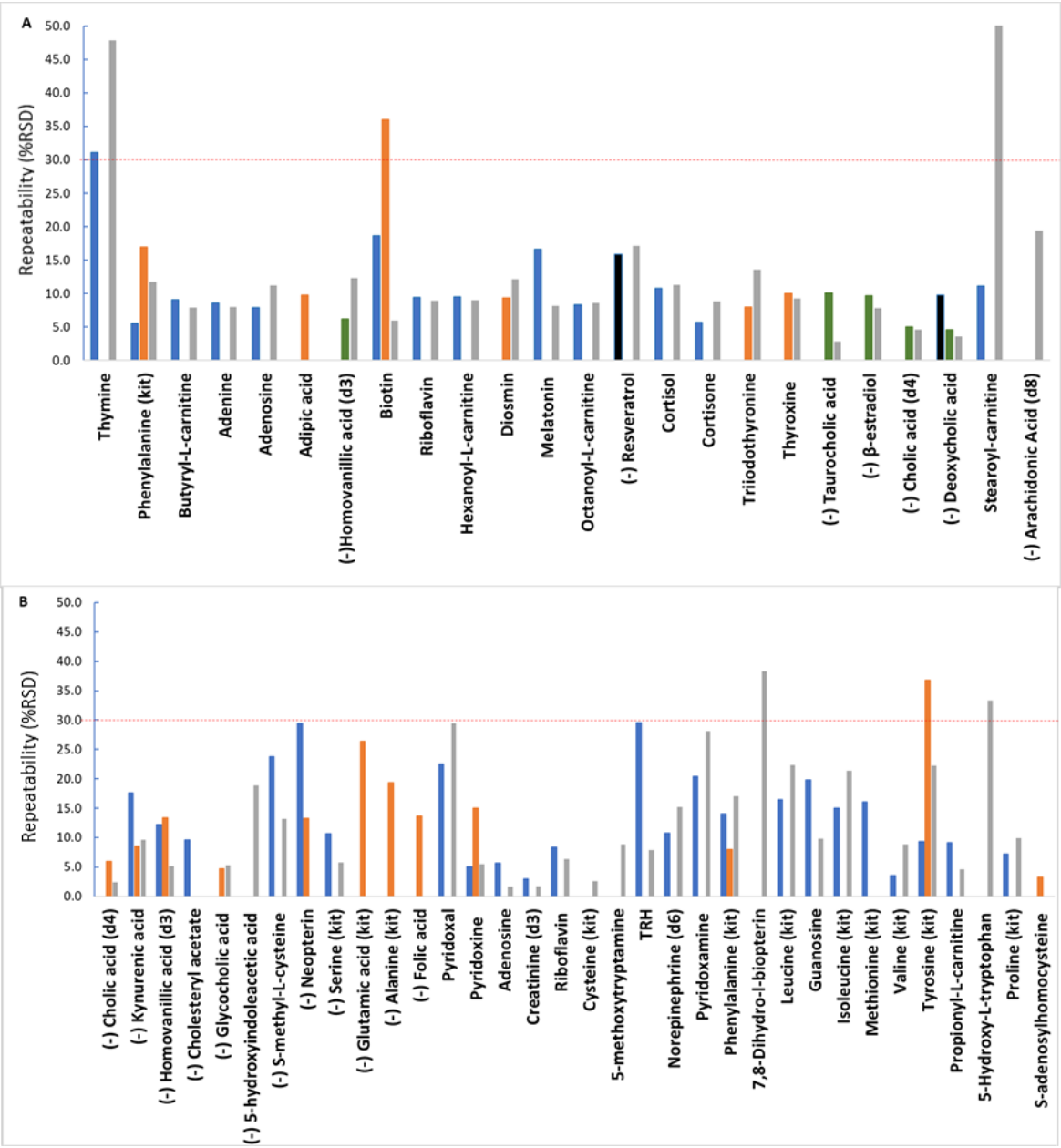


Figure 3.15. Signal repeatability (n=6) evaluated using standard metabolites detected in sSPE (blue (CN), orange (AZ)) and MeOH(grey) extracts of blood plasma. In (-ve) RP CZ (black) and AN (green) fractions were analyzed. Repeatability was calculated as RSD% of peak areas in RP (A) and ZIC-HILIC (B) LC-MS of 6 extraction replicates after subtracting background signal as described in ***Section 3.2.5.4****. Red dashed lines indicate acceptance threshold of 30% RSD for global metabolomics. Standards detected in (-ve) ESI designated by (-) at the metabolite' name. "TRH" stands for a thyrotropin releasing hormone.*

Notably, the loss of taurocholic acid in sSPE resembled the loss of presumably negatively charged lipids in lipidome analysis (**Section 3.3.1.4, Table 3.8**). Taurocholic acid may undergo a strong/irreversible binding to the MAX phase via very strong (pKa = -1.1) sulfur acid moiety.

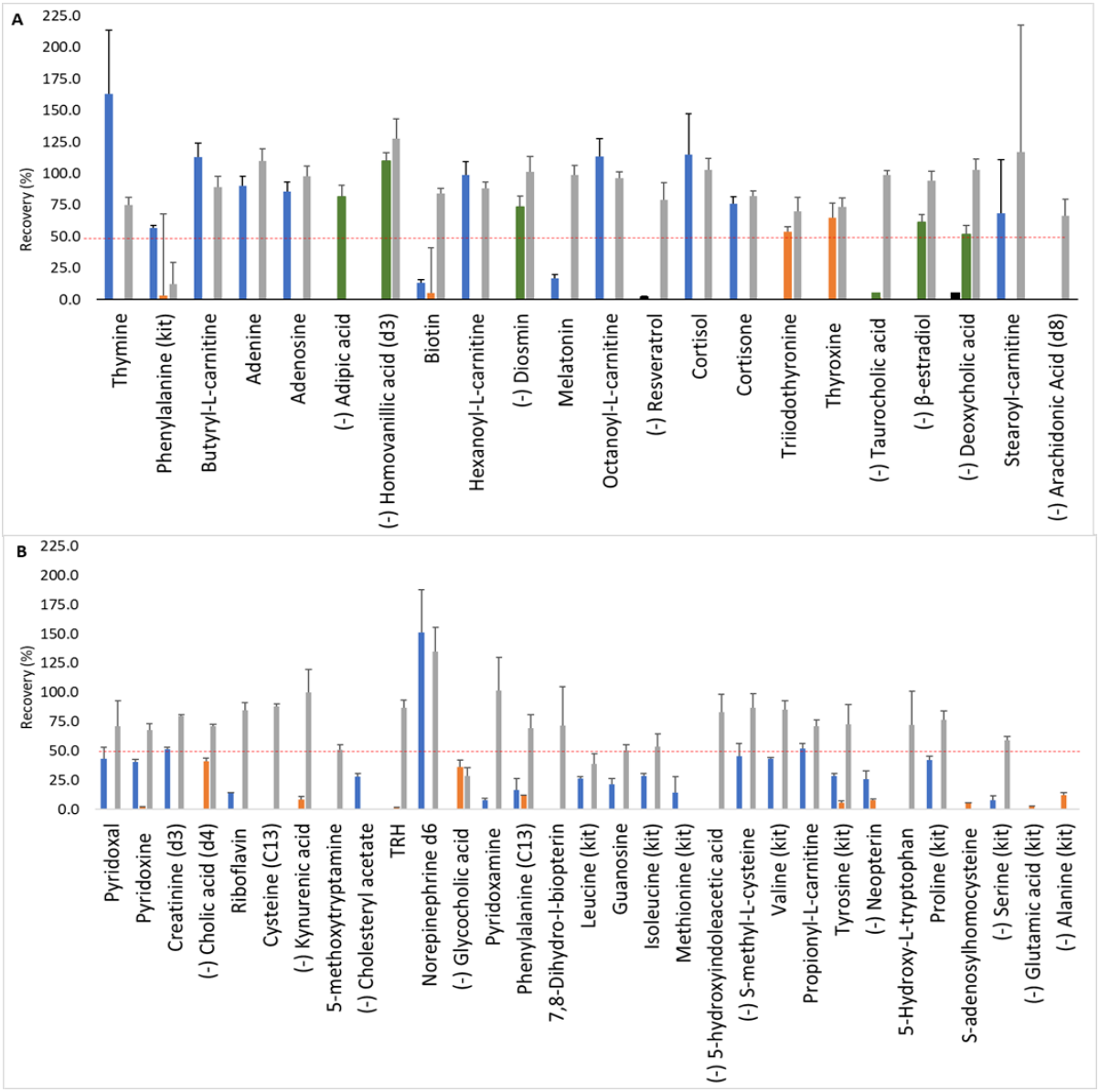


Figure 3.16. Recovery of spiked metabolites using sSPE (CN (blue) and/or AZ (orange) or MeOH (grey) precipitation after analysis using RP (A) and ZIC-HILIC (B). Metabolites detected in (-ve) RP analysis, where analyzed in sSPE fractions CZ (black) and AN (green) instead. Metabolites detected in a (-ve) ESI mode are designated by the minus sign, which is placed in parenthesis at the metabolite name. Red dashed lines indicate 50% recovery. Standards are shown from left to right according to their elution order (RT) on RP and ZIC-HILIC. "TRH" stands for a thyrotropin releasing hormone.

This differentiates taurocholic acid form structurally similar cholic and deoxycholic acids, which incorporated weak carboxyl (pKa = 4.5) groups and were detected in sSPE. Therefore, the recovery of strong anions requires a detailed investigation and improvements in the future.

The current experiment demonstrated a poor recovery of metabolites in sSPE compared to MeOH. The losses may be caused by a poor solubility of polar compounds at high concentrations of organic solvents in sSPE, the complex nature of interactions provided by a DVBP-IEX phase, and a large number of transfer and drying steps in the sSPE procedure. A detailed investigation of the recovery revealed the majority of metabolites (30 out of 34) with recovery < 50% were detected in ZIC-HILIC, while in RP, only 5 out of 22 metabolites demonstrated < 50% recovery (**Figure 3.16**).

Non-specific adsorptive losses of polar compounds indicated that the most likely culprit for the low sSPE recovery of polar compounds in ZIC-HILIC. For example, from a total of 22 metabolites, which were not detected in MeOH or/and sSPE, 19 metabolites encountered problems on ZIC-HILIC (**Table 3.7**). Moreover, the vast majority (14) of 19 metabolites were very polar (RT > 14 min) zwitterions (**Appendix B, Supplementary Table B1**). The cause of the limited recovery of highly polar metabolites in sSPE and/or ZIC-HILIC was probably their insolubility or precipitation at high concentrations (≥ 95% MeCN) of organic solvents.

The solubility of highly polar compounds (amino acids) tested individually in solvents drops to below low µg/mL at high (> 75%) concentrations of organic solvents.[213] Concentrations of spiked amino acids (81-267 ng/mL) in sSPE extracts prepared for a ZIC-HILIC analysis (**Appendix B, Supplementary Table B7**) were lower than the published solubility data. However, a very high sample complexity could be the factor that decreases analyte solubility even at concentrations used in the current experiments. Additional investigations on the relationship between solubility, sample complexity, and sample composition are required. Therefore, simplification of sSPE and optimization of the content of organic solvents, without affecting the recovery of mid-polar metabolites should be considered in future optimization experiments. Alternatively, two-step wash and elution (aqueous and non-aqueous) in sSPE may be investigated to improve the recovery of metabolites.

The analysis of fractionation of standards demonstrated that several fractions and/or LC-MS methods could be avoided without affecting the coverage of standard metabolites used in the experiment. For example, the analysis of CN in the (-ve) ZIC-HILIC did not add to the detection of standards (**Appendix B, Supplementary Table B8)**. Therefore, a further reduction in the analysis time for the current experiment could be achieved by limiting the analysis to only 5LC-MS runs instead of 8 : (i) (+ve) RP - AZ and CN; (ii) (-ve) RP - AN; (ii) (+ve) ZIC-HILIC – CZ and (iv) (-ve) ZIC-HILIC - AZ.

*Table 3.7. Standards detected in combined sSPE fractions (A-anion, C- cation, N-neutral, Z-zwitterion) or MeOH extracts-(Y) using preparation protocols described in **Section 3.2.5.2**. Metabolites that were not detected are designated by "ND". Concentrations in LC-MS samples ("at LC-MS step") were calculated based on the assumption of 100% recovery of standard and did not include endogenous levels. Amino acids with incorporated stable isotopes (commercial kit of "SIL" amino acids are designated by the word "SIL" at the name of a standard. For metabolites not detected in sSPE and MeOH extracts, RT at LC-MS were obtained from the analysis of neat standards dissolved in the reconstitution solvents used as 100% reference for analysis of matrix effects (**Section 3.2.5.3**).*

Standards	Preferable LC-MS	RT (min) RP in standard solution	RT (min) ZIC-HILIC in standard solution	Class (charge at physiological pH)	preferred combined SPE fraction	MeOH detection	Conc in MeOH at LC-MS (ng/mL)
Kynurenine	(+ve) RP+	2.5	ND	Z	ND		20.2
Adipic acid	(-ve) RP	4.3	6.8	A	AN	ND	20.2
Arachidonic Acid (d_8)	(-ve) RP	20	1	A	ND	Y	34.7
Cysteine "SIL"	(+ve)ZIC-HILIC	0.6	6.8	Z	ND	Y	46.5
5-methoxytryptamine	(+ve) ZIC-	ND	9.8	C	ND	Y	20.2
Cholesteryl acetate	(-ve) ZIC-HILIC	17.5	10.2	N	CN	ND	20.2
7,8-Dihydro-l-biopterin	(+ve) ZIC-	0.6	12.2	N	ND	Y	20.2
Methionine "SIL"	(+ve) ZIC-	0.9	13.8	Z	CN	ND	57.3
5-Hydroxy-indoleacetic acid	(-ve) ZIC-HILIC	5.5	14.2	A	ND	Y	20.2
3-Hydroxy-DL-kynurenine	(+ve) ZIC-	ND	14.8	Z	ND		20.2
Neopterin	(-ve) ZIC-HILIC	1	16	N	CN/AZ	ND	20.2
5-Hydroxy-L-tryptophan	(-ve) ZIC-HILIC	2	16.5	Z	ND	Y	20.2
Threonine "SIL"	(-ve) ZIC-HILIC	0.8	18.9	Z	ND		45.8
Glycine "SIL"	(-ve) ZIC-HILIC	0.6	19	Z	ND		28.9
S-adenosylhomocysteine	(+ve) ZIC-	3	19.1	Z	AZ	ND	150.2
Glutamic acid "SIL"	(-ve) ZIC-HILIC	0.7	22.6	Z	AZ	ND	56.5
Alanine "SIL"	(-ve) ZIC-HILIC	0.7	23	Z	AZ	ND	34.2
Aspartic acid "SIL"	(-ve) ZIC-HILIC	0.7	23.1	Z	ND		51.1
Folic acid	(-ve) ZIC-HILIC	5.3	23.2	A	AZ	ND	20.2
Saccharopine	(-ve) ZIC-HILIC	0.6	24.1	Z	ND		20.2
Lysine "SIL"	(+ve) ZIC-	0.4	24.2	Z	ND		59.6
Arginine "SIL"	(+ve) ZIC-	0.5	25.1	Z	ND		66.9

This approach reduces analysis time when a metabolomics study aims at the limited classes of metabolites. The optimized protocol also minimizes splitting across the fractions, with only 8 out of 61 metabolites tested showing significant splitting (>10%) in two and only one metabolite (arachidonic acid) splitting across three fractions (**Appendix B, Supplementary Tables B12,** and **B13**).

3.3.3.1.2 *Analysis of matrix effects*

Matrix effects were evaluated using the post-extraction spike method at a single concentration level equal to the expected concentration of standards at 100% recovery at the LC-MS step (**Appendix B, Supplementary Table B7**). The analysis of sSPE and MeOH extracts on RP revealed nine metabolites for which matrix effects were similar between sSPE and MeOH, five metabolites for which matrix effects were improved in sSPE, and seven metabolites which demonstrated better matrix effects in MeOH than in sSPE (**Figure 3.17**). The majority of metabolites in sSPE, which were affected on RP by matrix effects, eluted after 7.9 min (i.e., after diosmin) (**Figure 3.17**).

The analysis of matrix effects using ZIC-HILIC demonstrated seven metabolites with smaller matrix effects in sSPE than in MeOH and five metabolites with matrix effects smaller in MeOH *versus* sSPE. Out of 61 standards, matrix effects (below 80% or above 120% on **Figure 3.17**) affected 42 metabolites in sSPE and 32 metabolites in MeOH (**Appendix B, Supplementary Table B9**). The higher number of affected metabolites in sSPE was likely due to the higher sample load on LC-MS analysis for sSPE (equivalent to 6.4 µL of original plasma) compared to MeOH (equivalent to 1.6 µL of original plasma).

The analyses of signal strength and matrix effects between ZIC-HILIC and RP were also executed using five standards that were detected in both LC-MS methods (**Appendix B, Supplementary Tables B11,** and **12**). The data demonstrated overall lower peak areas in ZIC-HILIC and higher matrix effects compared to RP (**Appendix B, Supplementary Table B12**).

Notably, signals of these five standard metabolites were detected in different LC-MS methods and fractions at different intensities and matrix effects. This provides a flexibility in the selection of fractions and LC-MS methods for the analysis of such metabolites. For example, when a signal strength (for quality of fragmentation spectrum, for example) becomes a priority, the analyzed fraction with the highest signal can be used. Besides, the list of standards in this analysis could be

supplemented by larger number of highly polar/charged standards such as mono- and di-saccharides and polyols to execute a more systematic assessment in a future study.

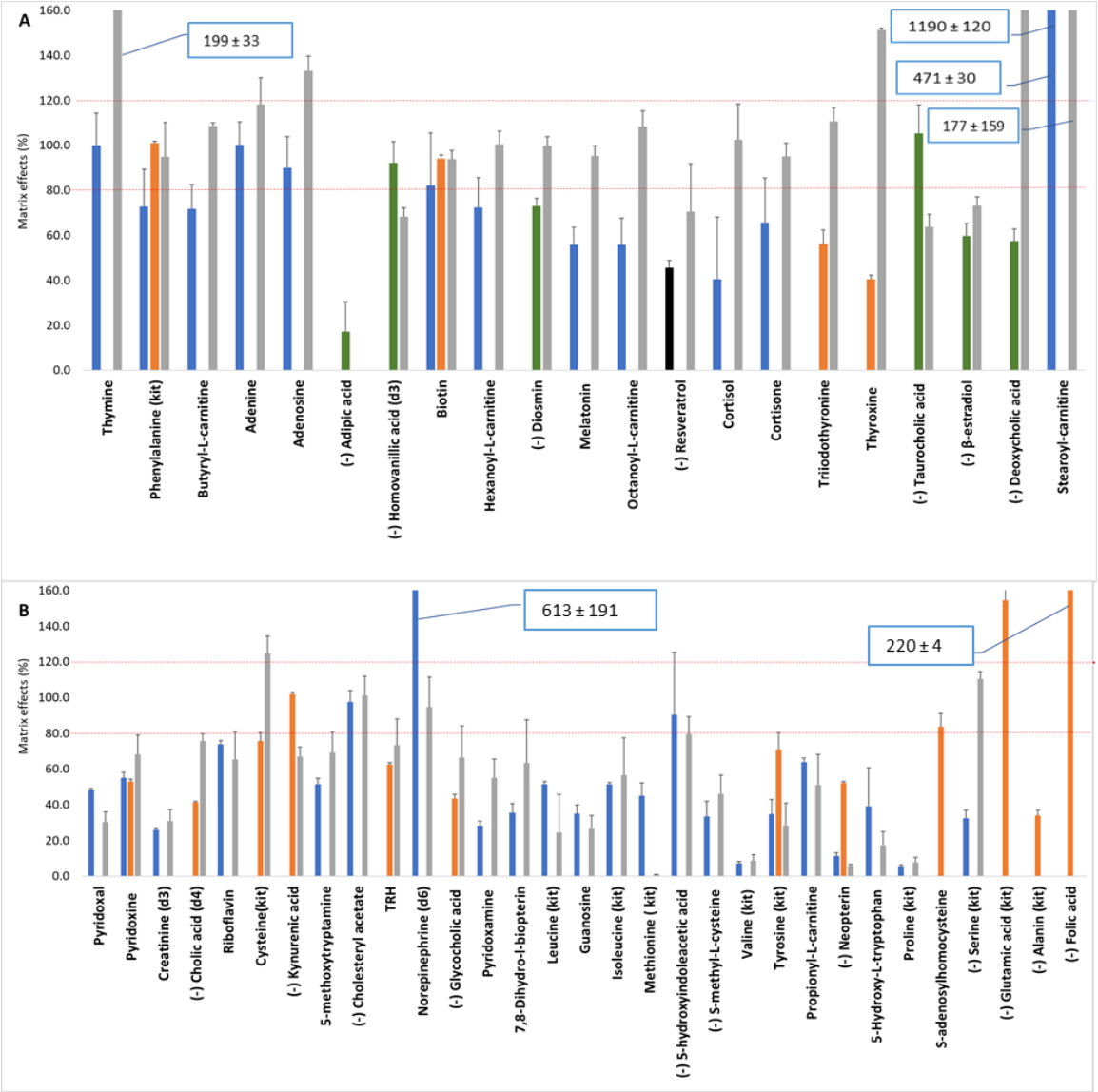


*Figure 3.17. Evaluation of matrix effects for metabolites spiked into sSPE and MeOH extracts (**Section 3.2.5.3**) at RP (A) and ZIC-HILIC (B) LC-MS analyses. Metabolites assessed in CN (blue) and/or AZ (orange), and/or in MeOH extracts in all LC-MS except (-ve) RP analysis, where analyzed sSPE fractions CZ (black) and AN (green). Metabolites detected in (-ve) ESI mode are designated by the minus sign, which is placed in parenthesis before metabolite name. Red dashed lines indicate an 80-120% acceptance interval, which indicates negligible matrix effects. Standards are shown from left to right according to their elution order on RP and ZIC-HILIC. Metabolites with matrix effects above 160% range are annotated.*

The proportion of metabolites affected by matrix effects (15 out of 21) in the RP analysis of sSPE in the current experiment (**Appendix B, Supplementary Table B9**) increased compared to the IEX SPE (14 out of 30) in the systematic analysis of seven extraction methods **(Section 2.3.1.4**). The trend coincided with the increased sample load (6.4 µL plasma equivalent) in the current experiment compared to IEX SPE (2.9 µL plasma equivalent) (**Section 2.3.1**).

The decrease to 1.6 µL of original plasma, from 2.9 µL, produced an improvement of matrix effects for RP analysis of MeOH in the current experiment (8 out of 21 affected) as described in **Section 2.3.1.4,** compared to 18 out of 27 metabolites. The targeted results suggested that the decrease of the analyzed sample amount/volume improved solubility and matrix effects. Also, the increase in the length of the LC separation may provide positive impact in future developments.

3.3.3.2 Evaluation of the final protocol using a global approach

3.3.3.2.1 *Analysis of repeatability for all metabolites*

After confirmation of the final protocol's appropriate performance using standards, the performance of the sSPE was analyzed using a global approach. The main results of this analysis (**Table 3.8)** demonstrated acceptable repeatability and good control over sSPE performance. For example, the median repeatability of all signals was in the range 13.7 - 28.4% RSD (average 18.7 – 32.2%, n=6), which is comparable to the traditional MeOH extraction (median 14.7-25.2%, average 19.9-30.4%, n=6).

The worst performance, with median RSD% higher than 25%, was encountered in the analysis of CN in all methods and MeOH in RP (+ve) analysis. This also resulted in a decrease of metabolites with RSD ≤ 30% in these extracts, as shown in **Table 3.8**. In addition, the decrease in signal repeatability between LC-MS methods coincided with the decrease of signal strength (**Appendix B, Supplementary Figure B5**).

There are several key sample processing steps, which can be a source of irreproducibility: MeOH extraction, sSPE fractionation; reconstitution and mixing fractions prior to an LC-MS analysis; and LC-MS analysis itself. The poor repeatability of signals in either of the first two should result in systematic deterioration of all samples across LC-MS methods, while the irreproducibility of LC-MS analysis should be revealed by QC samples and becomes visible on PCA. However, PCA analysis (**Figure 3.18**) demonstrated no outliers, and signals in QC are highly repeatable. In addition, the poor repeatability was isolated to CN fractions only in three LC-MS methods except

for (-ve) RP (**Table 3.8**). Therefore, the issue was not related to sSPE or LC-MS irreproducibility. Therefore, a sample reconstitution and fractions combining remained as the most probable source of the poor signal repeatability and a primary suspected reason is human error caused by the excessively laborious procedure.

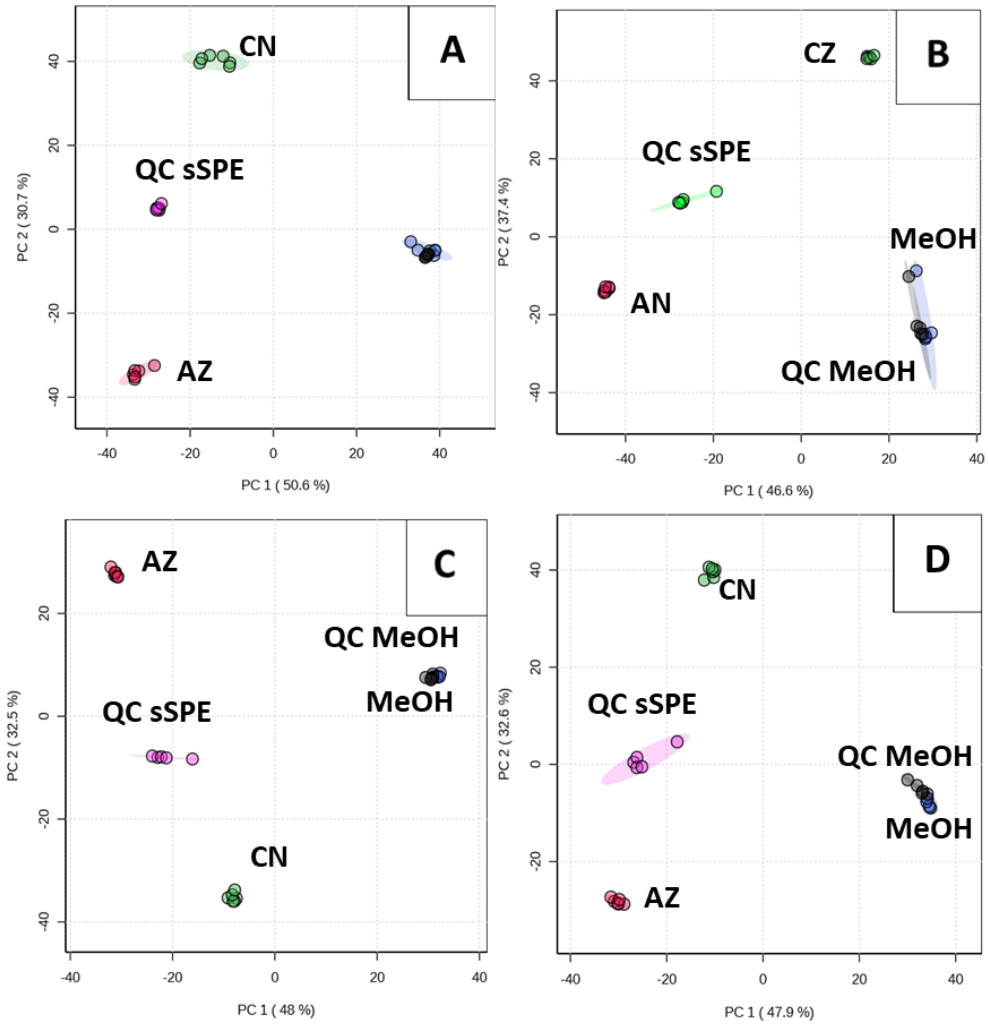


*Figure 3.18. PCA analysis of fractions and MeOH extracts. Two QC sample sets were used (one of SPE and another for MeOH extracts) due to extremely different loading amounts of these sample sets. PCA executed for each LC-MS method and displays QC for sSPE fractions ("QC sSPE"), QC for MeOH ("QC MeOH") and all combined fractions AN (anion + neutral,) CZ (cation + zwitterion, AZ (anion + zwitterion) and CN (cation + neutral) for all metabolites that satisfy criteria described **Section 3.2.5.4**. Plots are designated to LC-MS methods by letters positioned in the upper right corners: A-(+ve)RP, B – (-ve) RP, C, and D for ZIC-HILIC (+ve) and (-ve) ESI, respectively, and shows the top two principal components. The analysis was executed using the statistical package of MetaboAnalyst 4.0, as described in **Section 3.2.5.4**.*

The poor repeatability of metabolite signals in MeOH in (+ve) RP occurred only once on the single sample preparation day specific to this LC-MS analysis (see **Section 3.2.3.3.1** for setting up analytical batches). Aside from these four cases, the repeatability of signals in sSPE fractions and MeOH extracts (**Table 3.8**) was just slightly lower but comparable to results observed in the parallel assessment of seven extraction methods,[173] where MeOH extracts and IEX SPE demonstrated between 11 - 19% and 17 - 23% of median RSD (n=6), respectively.

The LC-MS sample preparation was laborious, which may explain the occurrence of errors due to fatigue. Subsequently, a simplification of sample preparation should be considered to improve the performance of the sSPE. Overall, the detailed global metabolomics analysis confirmed the acceptable analytical performance of the sequential sample fractionation, reported in the targeted analysis (**Section 3.3.3.1**).

*Table 3.8. Repeatability of metabolite signals in the experiment on combined sSPE fractions in four LC-MS methods: (+ve) RP- A; (+ve) ZIC-HILIC – B; (-ve) RP – C; (-ve) ZIC-HILIC – D. To assess repeatability, RSD% was calculated for each metabolite in an sSPE fraction and MeOH extract (n-=6) and for QC samples (n=5). The batch analysis was executed injected as described in **Section 3.2.3.3.1** in detail. Mean and median values were calculated across all metabolites, which satisfy conditions described in **Section 3.2.5.4** in the LC-MS method. The proportion (%) of metabolites with RSD ≤ 30% to the total number of metabolites in an sSPE fraction or MeOH extract were also calculated.*

Samples	Mean (%)	Median (%)	(%) of metabolites with RSD ≤ 30%	Samples	Mean (%)	Median (%)	(%) of metabolites with RSD ≤ 30%
A. (+ve) RP				**B. (+ve) ZIC-HILIC**			
AZ	25.7	19.6	75.7	**AZ**	29	24	63.8
CN	29.9	26.1	55.8	**CN**	32.3	25.8	58.3
MeOH	30.4	25.2	58.6	**MeOH**	22	17	74.3
QC sSPE	14.3	10.1	90.6	**QC sSPE**	23.1	17.3	73.3
QC MeOH	13.2	8.9	91.4	**QC MeOH**	16.7	11	83.8
C. (-ve) RP				**D. (-ve) ZIC-HILIC**			
AN	18.7	13.7	83.9	**AZ**	23.6	19.4	75
CZ	22.7	17.6	86.7	**CN**	32.2	28.4	54.3
MeOH	21.8	16.2	78.5	**MeOH**	19.9	14.7	80
QC sSPE	21.5	16.7	78.5	**QC sSPE**	18.9	13.8	82.9
QC MeOH	20.5	14.7	80.2	**QC MeOH**	16.8	11.8	87.1

3.3.3.2.2 *Analysis of fractionation and metabolome coverage*

The fractionation and metabolome coverage were analyzed on combined fractions of sSPE and on MeOH extracts, at amounts of material equivalent to 6.4 µL and 1.6 µL of original plasma, respectively. The experiment demonstrated the 1.6-fold increase of total and 3.4-fold of unique metabolite coverage in combined sSPE fractions compared to MeOH extract, which was comparable to an average improvement of the metabolome coverage (1.4-fold) in the analysis of individual sSPE fractions (**Table 3.9**). As expected, combined fractions demonstrated a reduced split (25.3%) of metabolites on average compared to 40.7% of metabolites affected by the split in individual sSPE fractions (**Table 3.9**).

The orthogonality of combined sSPE fractions to each other and to MeOH was demonstrated by PCA analysis (**Appendix B, Supplementary Figure B4**) and confirmed by hierarchical cluster analysis (**Appendix B, Supplementary Figure B5**). The prominent orthogonality of MeOH to combined fractions AZ and CZ (**Appendix B, Supplementary Figure B5)** resembles the one demonstrated by MeOH and the Z fraction in the individual fraction analysis (**Section 3.2.2**). This suggests the prominent involvement of zwitterionic metabolites for enriching the metabolome coverage in sSPE compared to MeOH. Indeed, the increase of metabolome coverage for very polar metabolites (RT> 14 min in ZIC HILIC and between 4 and 6 mins on RP) in sSPE compared to MeOH was demonstrated by metabolite maps (**Appendix B, Supplementary Figures B6 and B7**). Therefore, polar zwitterions (including amino acids, which elute after 14th min) are associated with increased metabolome coverage in sSPE (and the loss of their detection in MeOH).

Subsequently, a poor ZIC-HILIC detection of highly polar metabolites in MeOH compared to sSPE confirmed the results discussed in **Section 3.3.1.2**. The total number of metabolites detected in this study varies from 17852 (the sum is corrected for the redundancy between MeOH and sSPE) in the latest experiment with combined fractions to 8775 non-redundant metabolites in the previous experiment with the analysis of individual fractions (**Table 3.9**). The most plausible explanation is that the experiment with combined fractions was executed on the QTOF 6545, while all previous experiments were run on QTOF 6550 before its replacement. The manual curation of complete data sets in both experiments and the consistency in the increase of metabolite counts across all extractions ensure the consistency of observations towards metabolite coverage and the split.

In conclusion, the global analysis demonstrated acceptable analytical performance, increased metabolome coverage, and repaired the metabolite split. However, the method requires further

improvements to minimize matrix effects, simplification of the protocol and the decrease of the metabolite split via improved control of metabolites' fractionation behavior.

Table 3.9. Comparison of metabolome coverage of the sequential sample preparation method analyzed using a global approach for an individual (A) and combined (B) sSPE fractions on QTOF 6550 and 6545, respectively. The comparison was executed using data from analyses of sSPE and MeOH at sample loads equivalent to 6.4 and 1.6 µL of original plasma, respectively. Reasons for different loads of sSPE and MeOH were discussed in ***Section 3.3.2.2.****7. The number of total metabolites was calculated using approaches and methods described in* ***Sections 3.2.4.9 and 3.2.5.4*** *for A and B table sections. The number of unique metabolites was obtained from Venn comparisons of MeOH against the non-redundant metabolite composition of sSPE.*

A. Individual fractions					
Parameter	**(+ve) RP**	**(-ve) RP**	**(+ve) ZIC HILIC**	**(-ve) ZIC HILIC**	**Study average**
Total metabolites in sSPE	2019	2383	1772	1627	
Total MeOH	1648	2022	1052	1041	
Metabolites unique to sSPE	638	806	811	757	
Metabolites unique to MeOH	267	445	91	171	
SPE/MeOH total coverage ratio	1.2	1.2	1.7	1.6	1.4
% of all split metabolites in sSPE fractions	36.4	36.6	48.1	41.5	40.7
B. Combined fractions					
Parameters	**(+ve) RP**	**(-ve) RP**	**(+ve) ZIC HILIC**	**(-ve) ZIC HILIC**	**Study average**
Total metabolites in sSPE	4370	4472	3037	3545	
Total MeOH	2787	2925	1713	2165	
Metabolites unique to sSPE	2184	2340	1721	2017	
Metabolites unique to MeOH	601	793	397	637	
SPE/MeOH total coverage ratio	1.6	1.5	1.8	1.6	1.6
% of all split metabolites in sSPE fractions	29.9	16.9	30.4	23.9	25.3

3.3.3.3 Limitations and advantages of the sSPE method

Limited metabolome coverage and low sensitivity of standard metabolomics workflow had driven the current method development, which results had achieved the study objectives. However, the advantages and limitations of the sSPE needed to be discussed from the perspective of an entire metabolomics study. Low recovery of a part of the metabolome pool is a complex scientific challenge for every extraction method used in global metabolomics, including sSPE or MeOH. However, this pitfall of the sSPE is amenable to improvements, which could be achieved using relatively simple and practical approaches. At the same time, it is impractical to expect that the sSPE may provide equally great recovery for each and all metabolites. This is due to the complexity

of the chemical properties of metabolites, which cannot be satisfied by the limited number of conditions employed in sSPE.

For the same reasons, it will not be possible to completely resolve another sSPE pitfall, i.e., compromised fractionation orthogonality. While the decrease/adjustment of the split phenomena is amenable to the reparation via practical approaches like fraction combinations or change of SPE (**Section 3.3.3**), splitting can also provide advantages, where one fraction may provide better quantitative results for a given metabolite or analyses in different fractions can confirm the accuracy of relative quantitation of the analyte. However, the irreproducible splitting of metabolites across different samples may jeopardize accuracy of quantitation. Therefore, the split reproducibility (constant ratio of metabolite signals between overlapping fractions) should be assessed in a routine manner.

The third pitfall of the sSPE is in the complexity of the sample preparation process. While the majority of labor could be outsourced to automaton devices, the preparation still requires close attention to details and will limit sample throughput compared to standard workflows.

The final disadvantage of the sSPE is an inefficient fractionation and lower recovery of blood lipidome, which was discussed in **Section 3.3.1**. Due to these reasons, sSPE may not be recommended for lipidome analysis, despite similar lipidome coverage compared to IPA. Concluding the discussion of pitfalls, sSPE demonstrated several limitations, which can be feasibly eliminated, or minimized, with further method development. Most of these limitations are common between extraction methods and are outscored by the advantages of sSPE.

The achievement of key objectives in the development provided several advantages to sSPE compared to other sample preparation methods, including MeOH precipitation. The prominent advantage of the sSPE is a 1.6-fold increase of the metabolome coverage compared to the most popular solvent precipitation method (MeOH) but at the cost of a 2-fold increase in analysis time. We demonstrated the improved detection of very hydrophilic metabolites in sSPE fractions analyzed on a ZIC-HILIC. Several physiologically important metabolites belong to this group of metabolites (e.g. neurotransmitters and amino acids). Therefore, sSPE can facilitate the global approach to discover additional low abundance metabolites and their physiological functions, which are largely inaccessible in studies based on MeOH extracts.

The sSPE provides more comprehensive metabolome coverage and higher enrichment factor for metabolites than MeOH precipitation. Moreover, the sSPE provides the possibility to sub-

fractionate (for example, to execute elution using aqueous or solvent-based fluids). Therefore, the sSPE offers the advantage of running either global or targeted metabolomics analysis within the same study. This saves labor time and facilitates the transfer of putative biomarkers to a targeted stage. Even further reduction of the study's time can be achieved using the sSPE fraction combination's flexibility. For example, suppose a study aims at a limited set of metabolite classes with known fractionation behavior. In that case, there may be no need to analyze every fraction (or every fraction combination), as demonstrated in **Section 3.3.3.1.2**. Therefore, the complete set of advantages of sSPE will be available in studies, employing both global and targeted approaches. There is a growing demand for the analysis of several biofluids in the same study, and in the future, the applicability of this workflow for oral fluid, urine, and CSF can be tested. Considering the method is designed for polar and mid-polar metabolite coverage, its use for urine analysis could be particularly beneficial

Another advantage is that the method successfully reduces the lipid content due to the strong binding of lipids to SPE phases. Its use for the removal of lipids from non-lipid metabolites in a sample cleanup procedure is well documented.[75,172,177]

In addition, if individual analysis of sSPE fractions is employed, the method helps to assign ionic properties to metabolite based on their fractionation behavior. In the current final protocol, the nearly orthogonal separations of cations and anions are already achieved. This supports improving the confidence in metabolite identifications by differentiating the presence of a particular atomic element in the formula to a functional group. For example, if a metabolite demonstrates the fractionation behavior of a cation, the presence of a nitrogen atom in the molecular formula suggests the presence of a cationic (primary or quaternary) amine within the structure.

From the instrumental point of view, the advantage of sSPE is its adaptation to a 96-well plate and the use of entirely volatile buffers. These enable a high throughput processing of sSPE, an advanced degree of automation via robots, and flexible 2-D online SPE or LC setups. In conclusion, sSPE provides new opportunities in metabolomics analysis.

3.4 Conclusions and further work

In conclusion, the evaluation of the analytical performance of the proposed sSPE protocol demonstrated a reproducible extraction of spiked metabolites with an average RSD equal to 13.1%, which was comparable to the repeatability in MeOH extractions and to the results (CV=12.1%) of

the targeted metabolomics SPE-based analysis executed by Tulipani *et al.*[172] These results were confirmed by the global metabolomics analysis of combined fractions despite an elaborate sample preparation in the sSPE procedure. However, the risk of an occasional sample mishandling is higher in complex procedures, which indeed was likely observed in the final experiment. This could be mitigated by using a set of internal standards to monitor extraction quality and exclude any failed samples. Otherwise, the repeatability of signals in the global analysis (63-87% of metabolites with RSD ≤ 30%) was comparable or better than the one reported by Liu *et al.,* which reported 55-75% of reproducible metabolites in SPE.[75] Therefore, from the repeatability point of view, the prevention of occasional human-related errors should be addressed via the simplification and/or automation of the sample preparation. Considering the repeatability of current sSPE protocol, metabolite fold-changes of 2x could be detected for the majority of metabolites, in line with other metabolomics studies where 1.5-2.0x are routinely used. However, future development and success of metabolomics technology requires further improvement of repeatability to enable detection of even smaller changes in metabolite levels, especially in biospecimens such as plasma which are under homeostatic control.

The relative quantification of metabolites such as ketogenic amino acids and their degradation products after physical exercise.[214] The detection of such small signal differences require a further improvement of method repeatability, which can be achieved in multiple approaches. Thus, the simplification of the process (reduction of transfer and drying steps) and automation (reduction of human errors) and consequent expected improvement in repeatability should be given a priority in further development in parallel to splitting phenomena. In addition, the involvement of other factors that affect the recovery of metabolites was observed in the study. In particular, the low recovery of spiked standards was especially prominent for ZIC-HILIC compared to RP LC-MS. The reason for that could be the limited solubility of polar metabolites at high concentrations of organic solvents in sSPE (100%) fractionation and ZIC-HILIC analysis (95%). Moreover, the global metabolomics analysis revealed a drastically lower detection rate of highly hydrophilic metabolites (elute after 14th minute in ZIC-HILIC) in MeOH compared to sSPE, which points to the possible involvement of a more diverse sample composition of MeOH relative to sSPE. An additional outcome of a very high load was the observation of profound matrix effects in all samples, which should also be addressed in future improvements, for example, by coupling fractions with nano or capillary LC where lower flowrates promote more efficient ionization.

The poor recovery as the consequence of strong/irreversible binding of metabolites to MAX and MCX phases was expected. Indirect evidence of such a strong binding via ionic or hydrophobic interactions came from the loss of taurocholic acid in sSPE fractions in the targeted assay (**Section 3.3.3.1.2**). The poor recovery of hydrophobic species was also confirmed in the assessment of sSPE for a lipidome analysis, which demonstrated (**Section 3.3.1.4**) a drastic loss of lipids in Z and A fractions (anions, as taurocholic acid) at (-ve) ESI compared to the classical IPA precipitation. The strong binding of lipids to various SPE phases was also described in experiments on the removal of lipids from non-lipid metabolites in SPE cleanup procedures.[209,210] Moreover, this process was accompanied by the irreversible loss of recovery for select hydrophilic metabolites as well.[172,75] However, our selection of polymer-based SPE plates was rationalized by the impression of this DVBP phase's better stability at very alkaline (pH 9.2) conditions and fewer chances of bleeding compared to silica-based media. Just recently, lipid recovery investigations have demonstrated the advantages of silica-IEX and HILIC-IEX SPE matrices relative to SPE with hydrophobic phases.[215,216] Finally, improved recovery of neurotransmitters and neuropeptides from plasma was presented in silica HILIC SPE compared to porous graphitic carbon phases at low (70% MeCN) concentrations of solvent in the sample loading solvent.[217]

There are multiple causes of poor metabolite recovery, and the metabolite split is also involved. The reduction of the split to below 31% in our analysis of combined fractions was achieved mechanistically and did not address the gap in the understanding of ionic interactions and our control of it. Without that, the prediction of ionic properties of metabolites and the advantage of the analysis of individual sSPE fractions will be diminished. In sum, the sSPE method needs a further reduction of sample preparation complexity, the increase in recovery, the decrease of matrix effects, and metabolite split. Therefore, several application improvements are recommended for future work.

The simplification, automation, and miniaturization of the method should also improve nonspecific losses of metabolites. However, there is also the need to improve conditions of sSPE, and ZIC-HILIC LC-MS to increase the recovery of polar metabolites. The reduction of organic solvents in the wash and elution fluids of sSPE, sample preparation, and at the ZIC-HILIC analysis to 70-75% should increase the solubility and recovery of highly polar metabolites. The increase of the analyzed sample amount above 1.6 µL for MeOH and 6.4 µL for sSPE would not be recommended because it will increase matrix effects. Alternatively, the use of microflow/nanoflow ionization

sources, the increase in the length of a chromatographic gradient and/or an increase in the resolution of columns used in the analysis would be recommended. Such approaches will help to avoid reducing the sample amount and metabolome coverage.

The combination of fractions decreases the metabolite split and analysis time but precludes the analysis of individual sSPE fractions and the assignment of metabolites' ionic properties from their fractionation behavior. To comprehend fractionation behavior and to decrease the split, several actions must be carried out in future work: (i) inclusion of additional (to pKa) physical-chemical parameters for the interpretation of sSPE fractionation; (ii) increase in the number of metabolites used for targeted evaluation/method optimization to increase the representation of molecular properties; (iii) simplification of fractionation conditions by using less complex phases (silica-based IEX instead of Oasis SPE). In addition, the ability of sSPE to sustain repeatable fractionation and metabolite splitting across various samples reflective of typical inter-patient variability has not been investigated in the current study. Moreover, it is important to consider a relatively limited binding capacity of IEX cartridges, which could be overwhelmed by ionic compounds with high abundance such as small peptides, phospholipids and xenobiotics. The concentrations of these compounds can widely vary between individuals due to hemolysis, inappropriate enzymatic quenching or post-prandial effects and medical procedures. The methodological and QC/QA measures should be implemented and tested prior the usage of the sSPE in real studies.

Finally, the MAX-MCX sSPE approach appears to be promising sample preparation technique, which increases metabolite coverage of plasma while maintaining acceptable repeatability compared to MeOH and provides orthogonal metabolite separation. With improvements, the advantages of sSPE could make it the method of choice in the practice of metabolomic assays.

4 Validated LC-MS/MS method for the quantitation of cortisol and cortisone in human oral fluid

4.1 Introduction

The analysis of cortisol and cortisone in oral fluid was driven by the physiological and clinical importance of these steroid hormones. In addition, oral fluid was selected as a sample source due to its portable, non-invasive and simple sample collection. Finally, this biofluid may provide an alternative/expanded metabolome coverage compared to blood, if analyzed using sSPE method in future.

Cortisol (hydrocortisone) is synthesized and released in response to the activation of the hypothalamic-pituitary axis (HPA) by various arousal factors or stressors. The hypothalamus releases corticotrophin-releasing hormone, which travels to the anterior pituitary gland and signals cells to secrete adrenocorticotropic hormone (ACTH) into the blood. ACTH travels to the adrenal gland and signals the adrenal cortex to secrete cortisol into circulation.[218] Cortisol levels follow a daily rhythm, with levels peaking in the morning upon awakening, decreasing throughout the day, and reaching a nadir within a few hours of sleep onset.[219] Due to the lack of enzymes, which can metabolize cortisol in oral fluid outside of salivary glands, its levels remain stable at room temperature for up to one week.[129,220] Salivary cortisol correlates closely with the free cortisol fraction in serum, with a correlation coefficient ranging from r= 0.71 to r= 0.96, making saliva and oral fluid useful and less-invasive surrogate for studies aimed at measuring plasma levels of cortisol.[221] Unlike plasma cortisol, which is highly bound to carrier proteins, only free and unbound cortisol appears in saliva, which reduces some of the issues related to sample processing of plasma.[222] The patterns of change in the blood and oral fluid cortisol levels are similar. The strong correlation is due to the passive diffusion of free blood cortisol into oral cavity, which is independent of the production rate of oral fluid.[218,223] In addition, a substantial amount of diffused cortisol is converted into cortisone in the parotid gland at a constant rate. Therefore, the concentration of cortisone in oral fluid is also a potential indirect measure of blood-free cortisol.[224,225] However, the inter-individual variability in parotid cortisol-cortisone conversion rates also makes the simultaneous measurement of cortisol and cortisone of interest.

Cortisol is commonly measured using commercially available immunoassays, including automated electrochemiluminescence immunoassays and enzyme-linked immunosorbent assays (ELISA).

Because cortisol is a steroid-based hormone and much smaller than a typical protein, the ELISA kits use a competitive format rather than a more selective sandwich assay format. In this assay, cortisol that is conjugated to an enzyme competes with cortisol standards or cortisol from unknown samples for binding to a cortisol-specific polyclonal enzyme-labeled antibody. An enzyme, usually horseradish peroxidase, converts 3,3',5,5"-tetramethylbenzidine (TMB) to a product that absorbs at 650 nm or 450 nm. The absorbance intensity is inversely proportional to the concentration of cortisol in the standard or sample. The specificity of an immunoassay for cortisol is highly dependent on the specificity of the antibody. If the antibody can bind with other structurally similar hormones and/or drugs, then the quantitation of cortisol will not be accurate. Competitive ELISA assays have been reported to overestimate the concentration of cortisol from 1.4 to 2.4-fold (depending on the kit) in bio-specimens such as human saliva, urine,[226] and camel hair[227] as compared to LC-MS/MS. In humans, cortisol is much more potent than cortisone, so that cortisone is considered an inactive form. Cortisone is structurally similar to cortisol and differs only by the presence of a ketone group instead of hydroxyl at the C11 of ring C. due to the removal of hydrogen by 11-beta-steroid dehydrogenase enzymes which reside in salivary glands. Indeed, cortisol-cortisone cross-reactivity has been suspected as the major reason for the overestimation of cortisol concentration by ELISA assays. The cortisone cross-reactivity was estimated at 7% for quantitation of cortisol concentrations below 5 mmol/L (which are still of clinical relevance), but the contribution of this cross-reactivity to inaccuracy diminished at higher concentrations of cortisol in saliva.[164] This is attributed to higher cortisone concentration in oral fluids than that of cortisol, which in turn results in higher inaccuracy at lower cortisol levels. Moreover, the same study reported that excessively high values of cortisol calibration setpoints were the major contributing factor for the ELISA error, which could be decreased by correcting setpoints using MS measurements. Finally, the cross-reactivity of anti-cortisol ELISA to other similar structures such as dihydro- and tetrahydro metabolites of cortisol has been reported for urine analysis, where the presence of these interferences increased ELISA response several-fold.[226] Therefore, further LC-MS assessment of interferences and their contribution to ELISA cortisol quantitation in different biofluids and for different kits remains an important research area. Finally, immunoassays also show significant inter-assay differences, making it difficult to compare cortisol results across different labs and tests.[228,229] In fact, the differences between clinical immunoassays for serum cortisol measurements are so large that custom diagnostic cut-off values had to be implemented for

each assay type. This limitation has resulted in a great push for better standardization of cortisol methods and the increased need for validated LC-MS/MS methods as the preferred approach to replace immunoassay measurements of cortisol in various biofluids.[230]

There are several validated LC-MS/MS assays for cortisol or both cortisol and cortisone in oral fluid reported to date. Most of these methods use Salivette[164,224,231–235] oral fluid collection devices, although these may introduce contaminants and adversely impact cortisol recovery and long-term stability.[129,228] Mezzullo *et al*. compared direct spitting and Salivette and found both methods were comparable with a small negative bias (12-16%) for Salivette devices due to incomplete recovery but better precision presumably due to removal of mucin by the cotton pad, which facilitated clear layer separation during LLE.[236] In terms of sample preparation, LLE with MTBE, dichloromethane[233,236] or ethyl acetate[232,237] as well as online polymeric or C18 SPE8–10,23, are currently the most popular methods for cortisol determination. Automated 96-well plate solid-phase extraction was also recently proposed to further increase automation and sample throughput by reducing manual manipulations.[235] Cortisol is typically separated from cortisone and other glucocorticoids using C18,[164,225,231,233–235,237,238] C8,[236] pentafluorophenyl,[229] phenyl,[239] or phenyl-hexyl[232] RP LC with analysis times ranging from 4 to 14 min. Fustinoni *et al.* proposed online turbulent flow LC for automated high-throughput determination of cortisone, cortisol, and melatonin in oral fluid.[240] McBride *et al*. combined protein precipitation with dual polarity LC-MS/MS to expand the panel of measured metabolites beyond cortisol to include steroids, alkaloids, and neurotransmitters.[241] Magda *et al*. pushed method sensitivity for cortisol by combining ethyl acetate LLE, charge derivatization with 2-hydrazino-1-methylpyridine, and online SPE to achieve lower limits of quantitation (LLOQ) of 5 and 10 pg/mL for cortisol and cortisone, respectively.[232] Cao *et al.* proposed a simpler ethyl acetate/butanol LLE-based method that achieved LLOQ of 60 pg/mL for cortisol but did not measure cortisone.[242] Such ultra-sensitive methods are required for late-night salivary cortisol determinations but not necessary for daytime cortisol measurements. In this study, we report a simple and sensitive, low-cost, fully validated LC-MS/MS assay to measure cortisol and cortisone in oral fluid of healthy individuals and compare the results to those obtained using a commercially available ELISA kit for salivary cortisol.

4.2 **Materials and methods**

4.2.1 Materials and reagents

Cortisol (purity ≥ 98%), cortisone (purity ≥ 99%), dextran-coated charcoal, LC-MS grade methanol, cortisone (d_8), water, and acetic acid were purchased from Sigma-Aldrich (Oakville, ON, Canada). Cortisol ELISA kit, cat# KA1885, was purchased from Abnova Corporation (Taipei, Taiwan). Cortisol-d_4 was purchased from Cambridge Isotope Laboratories (Tewksbury, MA, US). Protein micro-BCA assay kit which also included Bovine Serum Albumin (BSA) standard for calibration was obtained from Fisher Scientific (Ottawa, ON, Canada)

4.2.2 Participants

Certification of ethical acceptability for research involving human subjects (Certificate 30001940) was obtained from the Concordia University Research Ethics Committee. Participants (age range: 20-30 years old) were healthy, medication-free, and non-smokers, which was determined by a health questionnaire. Exclusion criteria included chronic or acute diseases (e.g., heart disease and autoimmune disease) and/or prior or current use of immunomodulatory medications. Smoking, alcohol, and caffeine affect HPA axis activity and cortisol levels, which is why smoking was an exclusion factor.[57,243] All participants were instructed to abstain from alcohol consumption and sexual activity for at least 24 hours prior to their visits and to refrain from caffeine and food intake for 2 hours before their scheduled session.

4.2.3 Testing protocol

The experiments commenced at 10:00 am (or as close to that time as possible, with an average start time of 11:00 am) to minimize the confounding factor of diurnal rhythms. The room temperature was kept at 20°C to minimize the effects of thermal stress on the HPA axis and the sympathetic nervous system (SNS).[244] The oral fluid was collected at three time points over an hour: at time 0, 35 min and 60 min. The participants filled a microcentrifuge tube with oral fluid (~1.5 mL) using a sterile plastic transfer pipette. The samples were vortexed for 10 s and centrifuged at 1,100 x g for 10 min to sediment the debris. The supernatants were removed carefully without disturbing the pellet at the bottom of the tube, and the supernatants were aliquoted into several appropriately labeled cryotubes for storage at below -70°C until analysis.

4.2.4 Cortisol analysis by ELISA

Levels of cortisol were determined using competitive a colorimetric enzyme-linked immunosorbent assay (ELISA) according to the manufacturer's instructions. First, the samples were diluted 3-fold with standard mix A (blank calibrant with 0.0 ng/mL cortisol). Next, the standards or oral fluid samples were added to the coated plate. The hormone-conjugate was added at room temperature for 60 minutes to allow the competition for binding sites to occur. The plate was washed to remove unbound material, TMB substrate added, and the plate incubated in the dark at ambient temperature for 30 min. Samples were quantified against standards on an ELISA plate reader BioTek Gen5 (BioTek, VT, USA) at 450 nm. All samples of the same subject were analyzed in duplicates.

4.2.5 LC-MS/MS assay for measurement of cortisol and cortisone

4.2.5.1 Preparation of cortisone and cortisol stock standard solutions

Cortisol and cortisone stock standards were prepared in methanol to yield 3.7 and 1.8 mg/mL, respectively. Commercially pre-weighed (5 mg) cortisol-(d_4) (purity > 97%) was reconstituted in methanol to yield 5 mg/mL and commercially pre-weighed (1 mg) cortisone (d_8) (purity > 98.2%) to yield 1 mg/mL. Cortisol and cortisone standards were then diluted in 44% methanol to 200 µg/mL, while cortisol (d_4) and cortisone (d_8) internal standards were diluted to 50 µg/mL. These standards were aliquoted and stored at -70 ºC. On the day of analysis, an aliquot of each standard was placed on ice for 10-15 min and then used to prepare two individual standards of 2.0 and 1.8 µg/mL concentration in 44% methanol. Cortisol (d_4) and cortisone (d_8) internal standard were prepared at 500 ng/mL with the same diluent.

4.2.5.2 Preparation of blank oral fluid matrix for validation and calibration of LC-MS assay

The oral fluid was collected via spitting into 15 mL polypropylene centrifuge tubes from two male and two healthy female volunteers and kept below -70°C until the day of matrix preparation. On the preparation day, 2.5 mL of oral fluid from each individual were thawed on ice, pooled in a 50 mL polypropylene centrifuge tube, and mixed with 40 mL of ice-cold methanol. Then, the tube was vigorously shaken for 15 sec and incubated at ambient temperature for 15 min with periodic shaking (every 5th min) followed by incubation at -70 ºC for 15 min without shaking. After 30-min centrifugation at 4000 x g at 4ºC, 45 mL of the resulting supernatant was distributed into nine 5

mL aliquots in glass tubes and dried in a Speed-Vac (Centrivap, Labconco, cat#7812013) overnight. Next morning, the content of each tube was reconstituted to 1.0 mL with 5% methanol, combined into a single polypropylene 15 mL centrifuge tube and stripped using 0.5 g of charcoal (activated as described in **Supplementary materials and methods (Appendix C**)) followed by overnight incubation on a rocker platform (10 oscillations per min) at ambient temperature. The stripped oral fluid was centrifuged at 4000 x g at ambient temperature for 15 min. The supernatant was transferred into a new tube and then centrifuged one more time using the same conditions. The aliquots (400 µL) of the resulting supernatant (called the blank oral fluid matrix) were stored at -70 ºC. The success of the stripping procedure was confirmed by the absence of cortisol and cortisone during LC-MS/MS analysis of this supernatant.

4.2.5.3 Preparation of calibration curve

Aliquots of the stripped frozen oral fluid were thawed in a water bath at ambient temperature with periodic manual shaking on the day of the extraction. Once thawed, they were sonicated for 5 minutes and placed on ice. Then, using 2 µg/mL stock standards, a 40 ng/mL combined cortisol/cortisone standard in the oral fluid matrix was prepared in 6 replicates, each of which was further diluted by 2-fold serial dilutions in the stripped oral fluid matrix resulting in calibration points from 20.0 to 0.156 ng/mL. Samples were stored on ice.

4.2.5.4 Extraction and reconstitution of samples prior to LC-MS/MS analysis

Individual oral fluid samples (aliquots of 100 µL) from study participants (n=20) were thawed in a water bath, sonicated for 5 min, and vortexed for 5 min. After centrifugation at 10,000 x g for 30 s at 4 °C, samples were placed on ice. 30 µL of water (for blank extracts) or study samples or calibration curve standards prepared in stripped oral fluid matrix were extracted with 120 µL of ice-cold methanol containing 1.25 ng/mL of cortisol (d_4). After 15 min incubation at ambient temperature with periodic (every 5th min) shaking and 15 min incubation at -70ºC, samples were centrifuged at 15,000 x g at 4 ºC for 15 min. An aliquot (120 µL) of supernatant was transferred into a clean 1.5 mL polypropylene Eppendorf tube and dried overnight in a Speed-Vac (Centrivap, Labconco, cat#7812013). On the next day, samples were re-solubilized in 24 µL of 44% methanol, sonicated for 5 min, and vortexed for 5 min. Then, 24 µL of water were added, and samples were sonicated for 5 min, vortexed for 10 min, and finally centrifuged at 15,000 x g for 30 s to bring the liquid to the bottom. Samples were stored on ice prior to LC-MS analysis. Two QC samples were

created after reconstitution (i) one pooled sample by mixing equal volume from each study sample and (ii) one QC at the concentration of 5 ng/mL by mixing equal volumes of all calibration points prepared as described in **Section 4.2.5.3**.

4.2.5.5 LC-MS/MS analysis

LC-MS/MS analysis was executed on triple quadrupole mass spectrometer QQQ 6460 connected to UHPLC 1290 from Agilent Technologies (Santa Clara, CA, US). Cortisol and cortisone were chromatographically separated using gradient elution and Zorbax Eclipse Plus C18 column (2.1 x 50 mm, 1.8 µm, Agilent) at 30ºC and 0.4 mL/min flow rate. Mobile phase A was 0.1% (v/v) acetic acid in the water, and mobile phase B was 0.1% (v/v) acetic acid in methanol. The following gradient was used: 22% to 55% B between 0-0.5 min, a linear increase from 55% to 62.5% B between 0.5 and 2.0 min, increase to 90% B at 2.1 min, hold at 90% B for 1.5 min and final re-equilibration to the initial conditions between 3.7 to 6.0 min. The injection volume was 15 µl representing 7.5 µl of the original sample volume.

MS analysis was executed using (+ve) ESI with capillary and nozzle voltages set to 3500 and 800 V, nebulizing and sheath gas temperatures to 325 and 385°C, and gas flow and nebulizer pressure to 10 l/min and 35 psi, respectively. The following SRM parameters were used for quantitation: 367.2 → 121.2, 363.2 → 121.2 and 361.2 → 163.1; the fragmentor voltage: 126, 126, 132 V and collision energy: 16, 16, 20 V for cortisol-d_4, cortisol and cortisone, respectively. The isolation window was set to 0.7 FWHM (full width at half maximum) at Q1 and Q3; cell accelerator voltage to 3 V and dwell time to 300 ms for all quantifier transitions and 80 ms for qualifier transitions. For additional confirmation of analyte identity second SRM transition can be used: 363.2 → 327.6 and 361.2 → 121.2 for cortisol and cortisone, respectively. The final optimized method used separate time segments where cortisone quantifier, qualifier, and IS (369.3 → 169.2, collision energy 27) were monitored between 0 and 1.9 min, and cortisol quantifier, qualifier, and IS were monitored between 1.91 and 6 min. Please, refer to **Supplementary Figure C1 (Appendix C)** for a comprehensive view of all chromatograms. The addition of cortisone (d_8) as IS for cortisone was necessary to provide appropriate correction of recovery because charcoal-stripping of oral fluid affected recovery as shown during validation, and cortisol standard was not able to compensate adequately for this effect.

All LC-MS analyses were performed as follows: 10 equilibration injections of the stripped oral fluid matrix were performed at the beginning of all runs. This was followed by the injection of QC

samples, which were re-injected after every 10 injections and at the end of the injection sequence to monitor LC-MS stability. Calibration points were injected twice, at the beginning of the batch, after QC samples, and at the end of the batch, before the last QC samples. All other study samples were analyzed in randomized order throughout the sequence.

The same LC-MS/MS method with modifications was used for the assessment of adduct formation and contribution of cortisol (d_4) to the SRM channel of cortisol. Adduct formation was analyzed with Q3 set to scan in the range 75 - 410 m/z and Q1 set to ion masses of cortisol, cortisone, and cortisol (d_4) forming adduct ions ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ and $[M+NH_4]^+$) as shown in supplementary materials. The contribution of cortisol (d_4) to SRM channel of cortisol (363.2 → 121.2) was assessed with the following settings: Q1 set to 367.2, 366.2, 365.2, 364.2, and 363.2 (precursor ions of cortisol (d_4) containing 4, 3, 2, 1 and 0 deuterium atoms, respectively) and Q3 set to 121.1 for each precursor.

4.2.5.6 Data analysis

Data acquisition, processing, and quantification were performed using Acquisition and Quant applications from Mass Hunter Suite (Version B 07.00, SP1). Peaks were integrated using the Agile 2 algorithm, and area ratios of analytes to the internal standard were calculated using spiked cortisol (d_4) until cortisone (d_8) was added to the method. Calibration curves were built using 1/x weighted regression and used to quantify concentrations in unknown and validation samples.

4.2.6 Method validation

4.2.6.1 Preparation of the pooled samples for the assessment of recovery and matrix effects

The recovery of cortisol, cortisone, and cortisol (d_4) was evaluated using non-stripped oral fluid generated by pooling eight samples from different individuals. Samples were distributed into three aliquots (A, B, and C), each consisted of three replicates of 30 µL. Replicate A was extracted with 120 µL of methanol containing 2.5 ng/mL of cortisol, cortisone, and cortisol (d_4). Replicate B was extracted with methanol containing cortisol (d_4) at 2.5 ng/mL. Replicates C were extracted with methanol without standards to determine endogenous levels of cortisone and cortisol in this pooled sample and for matrix experiments. Samples were processed according to **Section 4.2.5.4** and reconstituted on the day of LC-MS/MS analysis as in the procedure except that replicates B were reconstituted in 44% methanol containing 10 ng/mL of cortisol and cortisone to serve as 100%

recovery references, while reconstitution solvent for replicates A contained no standards. Replicates C were either reconstituted in 44% methanol without a standard to control endogenous signals or in 44% methanol spiked with all standards at concentration 10 ng/mL to be used in the assessment of matrix effects as post-extraction spiked samples. The stocks of 1.8 µg/mL of cortisol and cortisone were used for the preparation of spiked standards.

4.2.6.2 Preparation of QC samples for validation experiments and for the quantitation of individual study samples

For the intra- and inter-day analyses, high and low QC samples were prepared fresh from 1.8 µg/mL stock at concentrations 1.5 and 11.5 ng/mL concentration in the stripped oral fluid matrix. The concentrations were selected (**Section 4.2.6.2**) for temporary use in the course of the validation experiments until more appropriate concentrations of QC could be selected.

For the analysis of study samples, the stock QC solution was prepared at the concentration of 60 ng/mL for cortisol and 200 ng/mL for cortisone in stripped oral fluid matrix directly from 1.8 µg/mL stocks, prepared as described in **Section 4.2.5.1**. This sample was diluted in oral fluid matrix 10-fold, then 3-fold, and then 2-fold to obtain high (QCH), medium (QCM), and low (QCL) samples with concentrations of (6 and 20 ng/mL), (2 and 6.67 ng/mL), and (1 and 3.33 ng/mL) for cortisol and cortisone, respectively. Five replicates of each QC sample were extracted and processed identically to other samples and calibration curve.

4.2.6.3 Preparation of calibration curve and assessment of LLOQ, linearity, and intra-day accuracy and precision

Preparation started by thawing aliquots of the stripped frozen oral fluid in a water bath at ambient temperature with periodic manual shaking. Once thawed, the calibration curve was prepared as described in **Section 4.2.5.3** using stocks of cortisol and cortisone with a concentration of 2 µg/mL in six replicates. The calibration samples were used to assess the LLOQ, the upper limit of quantitation (ULOQ), calibration linearity, accuracy and precision of calibration curve according to FDA Bioanalytical Method Validation guidelines.[102]

4.2.6.4 Preparation of validation samples for the assessment of intra-day accuracy and precision

Five sets of validation samples were prepared from 1.8 mg/mL individual stocks of cortisol and cortisone in a stripped oral fluid matrix to cover the range 0.54, 1.09, 4.38. 8.75 and 17.5 ng/mL in

n=8 replicates per concentration level and six of those extracted as described **Section 4.2.5.4**. Several replicates of a stripped oral fluid matrix without the addition of standards "blank oral fluid matrix" were also extracted to provide the signals in the matrix without spiking. For the assessment of LLOQ, intra-day accuracy, and precision, six replicates of each calibration point or validation samples were reconstituted as described in **Section 4.2.5.4** and analyzed on LC-MS/MS as described in **Section 4.2.5.5**.

4.2.6.5 Assessment of short and long-term storage stability and freeze-thawing stability

For the assessment of long-term storage stability, two replicates of validation samples (**Section 4.2.6.4**) with concentrations of 1.09 and 4.38 ng/mL were stored for 6 months at below -70 ºC as a dry extract or after reconstitution for LC-MS analysis (wet extract). After 6 months, stability samples were either reconstituted (dry extract) or thawed (wet extract) and analyzed in the LC-MS/MS. Samples for short-term stability were reconstituted as for LC-MS/MS analysis (**Section 4.2.5.4**) and stored for 2 h on the bench at room temperature and for 24 h at +7 °C in the autosampler, respectively. All samples for the assessment of storage stability were quantitated using calibration curves freshly prepared as described in **Section 4.2.5.3**. Freeze-thaw stability was assessed by exposing three 30 µL replicate aliquots of the 4.38 ng/mL validation sample (**Section 4.2.6.4**) to 30 min freezing at -70 ºC followed by thawing in a water bath at ambient temperature. A single aliquot out of three was exposed to either 1, 2, or 3 freeze-thawing cycles, while the validation sample with a concentration of 4.38 ng/mL was used as the no freeze-thawing control. (**Section 4.2.5.3**).

4.2.6.6 Assessment of recovery and matrix effects

For the assessment of recovery, duplicates of samples prepared as described in **Section 4.2.6.1** were used. Recovery was calculated by the formula: (area in test samples-area in background sample) * 100% / (area in reference sample-area in background sample).

For the analysis of matrix effects, samples prepared by post-extraction spiking of cortisol and cortisone into extracts of oral fluid aliquots C (**Section 4.2.6.1**) were used. For comparison, neat standards in 22% methanol were prepared at 5 ng/mL concentrations to match post-extraction spiked samples. Matrix effects were calculated using the following formula: (signal intensity in a post-extraction spiked sample of stripped oral fluid*100%)/signal intensity in 22% methanol neat standard.

4.2.6.7 Assessment of inter-day accuracy and precision

Validation samples originated from 1.8 µg/mL individual stocks (**Section 4.2.6.4**) and calibration curve standards prepared as described in **Section 4.2.5.3** originated from 2.0 µg/mL individual stocks, were prepared, dried overnight, reconstituted, and analyzed on LC-MS/MS as described in **Sections 4.2.6.1, 4.2.6.3, 4.2.6.4**, and **4.2.5.5,** respectively. The entire procedure was repeated on different non-consecutive days resulting in a total of 7 inter-day batches. During inter-day validation, an injector leak and a subsequent irreproducible performance of the sample injector were discovered. After replacement of the needle seat capillary and injection needle, highly reproducible autosampler performance was restored.

4.2.6.8 Stability of deuterated cortisol

Individual aliquots of each analyte standard at 5 ng/mL in 22% methanol and oral fluid matrix were prepared and analyzed as described in procedures for the assessment of the formation of ions with $[M+H]^+$, $[M+K]^+$, and $[M+NH_4]^+$ adducts and contribution of deuterated cortisol (to the signal in MRM channel of cortisol (363.2→121.1) which may occur due to hydrogen replacement of all deuterium atoms. The LC-MS/MS batch included 10 equilibration injections of a stripped oral fluid matrix, followed by sequential injections of the low, medium, and high QC samples with additional injections of QC samples at each 11th run and at the end of injection sequence. Other samples were loaded onto LC-MS/MS in random order.

4.2.6.9 Addition of cortisone (d_8) as IS for cortisone

To verify that cortisone (d_8) is an appropriate IS for cortisone in non-stripped oral fluid in order to improve method accuracy for this analyte, oral fluid from 10 individuals was evaluated. Samples were distributed into 2 x 95 µL aliquots, A and B. Replicates A were spiked with 5 µL of cortisol and cortisone to yield a 2.5 ng/mL increase in concentration for cortisol and 5 ng/mL for cortisone and then spiked with both IS to yield concentrations of 5 ng/mL. Replicates B were spiked with 5 µl of solvent and then were spiked with a mix of IS to yield 5 ng/mL. Calibration standards were prepared in a stripped oral fluid matrix, and 95 µL of it were spiked with 5 µL of the IS mix to yield 5 ng/mL. QC samples were prepared as described in **Section 4.2.6.2.** After 15 min of incubation on ice, all samples were extracted, dried, and analyzed after reconstitution as described in **Section 4.2.5.4**.

4.2.6.10 Creation of the calibration curve for routine implementation of the method

Based on validation results, for the routine implementation of the method, the preparation of the calibration curve (dilution factors and concentration limits) were slightly modified. The calibration curve was prepared by 1.5x serial dilution while keeping LLOQ and ULOQ of cortisol to maintain the validated linear range to increase the number of calibration points. For the measurement of study samples, the concentration ranges of calibration curves were 0.16-16.0 for cortisol (16.0, 10.7, 7.1, 4.7, 3.2, 2.1, 1.4, 0.94, 0.62, 0.31 and 0.16 ng/mL) and 0.21-21.3 for cortisone (21.3, 14.2, 9.5, 6.3, 4.2, 2.8, 1.9, 1.2, 0.83, 0.42, 0.21 ng/mL). Aliquots (95 µL) of calibration samples were spiked with 5 µL of the IS mix to yield concentration of 5 ng/mL, extracted, dried, and reconstituted as described in the correspondent **Section 4.2.6.3** except that reconstitution solvents did not contain IS.

4.2.7 Comparison of LC-MS/MS and ELISA assays

In addition to 19 test samples, several standards (0.63, 2.5, and 10 ng/mL) from separate calibration curves (from 10 to 0.31 ng/mL) of cortisol and cortisone in water and a QC at 1 ng/mL cortisol in a stripped oral fluid matrix using in-house standards were analyzed on ELISA and extracted (including all standards from the calibration curves) on the same day for LC-MS/MS. In addition, individual calibration curves of cortisol and cortisone were prepared in the commercial ELISA blank ("0.0" calibration point supplied with ELISA kit) and extracted in parallel to 4 ELISA commercial calibration standards (0.0, 0.1, 0.4, 1.7, and 7 ng/mL cortisol) for LC-MS/MS analyses on the same day. The ELISA analysis and extractions of samples for LC-MS/MS analysis were executed as described in **Sections 4.2.4** and **4.2.5.4**, respectively. All extracted samples were dried, reconstituted, and analyzed on the next day after extraction as described in **Sections 4.2.5.4** and **4.2.5.5**.

4.3 **Results and discussion**

The objective of this research was to develop and validate a sensitive low-cost, high-throughput LC-MS/MS assay for cortisol and cortisone in oral fluid and to minimize the sample volume required for this analysis. Although most validated methods reported to date required the use of LLE or SPE, protein precipitation with methanol was selected for this work to increase sample throughput and decrease sample volume requirements. The only other successful protein

precipitation methods for this analysis utilized acetonitrile but required 240-250 µL sample volume.[49,234] A more recent protein precipitation method which used acetonitrile:methanol:acetone cocktail and 100 µL of the oral fluid showed poor LOQ (8.6 ng/mL) for cortisol and did not have cortisone included in its panel.[241] Majority of other SPE and LLE-based validated methods also required 100 to 500 µL of the sample.[231–233,242,245,246] Our final validated method requires only 30 µL of oral fluid, making it ideally suited for time-course studies, and the sensitivity of the method is suitable for all time points except late-night cortisol determinations. In terms of MS analysis, we opted for (+ve) ESI, although successful methods with APCI[247,248] and (-ve) ESI[236] have also been reported.[164]

4.3.1 Validation of LC-MS/MS assay

The finalized method was validated for selectivity, accuracy, precision, stability, recovery, matrix effects, linearity, and LOD/LLOQ according to the procedures and acceptance criteria for Bioanalytical Method Validation established by FDA[102]. Calibration curves of cortisol and cortisone prepared in charcoal-stripped oral fluid demonstrated excellent linearity ($R^2 = 0.998$) in the range 0.15-20 ng/mL, acceptable intra-day accuracy (maximum inaccuracy < 7.8%) and low imprecision (< 15%) using cortisol (d_4) as IS for both analytes (**Table 4.1**). Inter-day comparison of calibration curves also showed acceptable performance, with at least 6 out of 8 calibration points meeting FDA requirements on all days (**Table 4.2**). Slopes of calibration curves remained similar between 7 days of the inter-day validation for each cortisol and cortisone (0.18±0.03 and 0.27±0.05, n=7, respectively), and intercepts did not deviate from zero (**Appendix C, Supplementary Table C1**).

The analysis of inter-day accuracy and precision for cortisol and cortisone (**Table 4.2**) demonstrated acceptable accuracy and precision in the interval of 0.31-20 ng/mL. The lowest concentration tested (0.15 ng/mL) provided the most inaccurate daily results and did not routinely meet LLOQ requirements set by FDA of RSD ≤20% and inaccuracy within ± 20% for both cortisol and/or cortisone on several of the 7 days tested.

*Table 4.1. Intra-day accuracy and precision of calibration curves (**Section 4.2.6.3**) and validation samples (**Section 4.2.6.4**) in charcoal-stripped oral fluid. Six replicates of each calibration standard were run to obtain calibration curves on the first day of validation and establish LLOQ.*

Calibration standard (ng/mL)	**Average accuracy (%)**		**Precision (n=6, RSD (%))**	
	Cortisol	**Cortisone**	**Cortisol**	**Cortisone**
0.15	98.9	99.6	13.3	7.9
0.31	102.7	106.8	8.0	4.0
0.63	96.9	92.2	13.0	16.9
1.25	97.4	97.7	10.2	11.2
2.5	103.9	104.4	5.7	2.1
5	102.7	102.4	4.6	5.0
10	96.7	98.1	6.1	4.8
20	100.7	100.5	4.2	3.7
Validation samples (ng/mL)	**Average accuracy (%)**		**Precision (n=6, RSD (%)**	
	Cortisol	**Cortisone**	**Cortisol**	**Cortisone**
0.54	98.9	101.5	6.2	6.5
1.09	96.0	101.5	6.0	7.7
4.38	100.1	105.9	10.3	12.3
8.75	87.6	90.4	6.3	7.6
17.5	88.9	93.0	7.2	8.8

As a result, 0.15 ng/mL was selected as the limit of detection (LOD) of the method, while LLOQ was determined to be 0.31 ng/mL for both cortisol and cortisone as this concentration consistently met FDA requirements on all days tested. The highest concentration tested was 20 ng/mL as higher concentrations were not expected for the late morning- mid-day samples such as the ones collected in the current study. These LLOQ's compare well with other methods reported in the literature, which generally ranged from 0.3-0.8 nM for both cortisol and cortisone.[164,224,235,237,240,245,249] Despite the sporadic inaccuracy of some mid calibration points during the first 5 days, the accuracy of at least 6 calibration points remained suitable for quantitation every day for both standards, except for day one of the cortisone test (**Table 4.2**).

Table 4.2. Summary of inter-day accuracy and precision for cortisol and cortisone in charcoal-stripped oral fluid over non-consecutive days (n=7) of analysis. Samples were prepared as described in ***Section 4.2.5.3.*** *Based on several gross autosampler mis-injections denoted by asterisks in the table, injector needle and needle seat capillary were changed between 5th and 6th days, which decreased injection variability and improved the accuracy.*

Concentration (ng/mL)	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average accuracy (n=7 days, %)	Inter-day precision (n=7 days, RSD %)
Cortisol accuracy (%)									
0.16	97.6	95.8	113	119.6	105.8	118.7	100.2	**107.2**	**9.9**
0.31	109.5	101.5	99.4	107.3	113.8	102.7	116.9	**107.3**	**6.5**
0.63	80.7	111.6	94.4	94.0	69.4*	94.9	102.1	**96.3**	**10.2**
1.25	102.6	91.3	101.1	90.8	106.1	92.4	89.4	**96.2**	**6.8**
2.5	113.0	95.0	89.9	89.9	104.5	92.5	88.4	**96.2**	**9.2**
5.0	97.1	64.7*	99.5	93.1	96.4	97.2	100.4	**97.3**	**2.6**
10.0	78.5	107.9	102.7	103.6	108.3	100.7	101.9	**100.5**	**10.1**
20.0	99.4	96.9	100.1	101.7	96.5	101.8	100.7	**99.6**	**2.1**
Cortisone accuracy (%)									
0.16	121.3	124.7	118.4	113.9	105.2	119.4	113.7	**116.7**	**6.4**
0.31	97.1	92.8	95.6	116.5	111.1	102.5	105.7	**103**	**8.7**
0.63	79.4	96.3	88.2	88.7	69.9*	92.5	107.0	**92.0**	**9.3**
1.25	95.2	84.6	104	93.8	108.2	98.3	86.7	**95.8**	**8.6**
2.5	111.7	94.0	93.9	84.9	102.1	90.5	85.6	**94.7**	**9.5**
5.0	94.7	53.1*	96.7	95.9	95.0	99.1	98.7	**96.7**	**1.9**
10.0	66.1*	111.7	102.8	106.2	116.5	93.7	100.1	**105.2**	**8.2**
20.0	100.7	95.9	100.2	100.2	93.0	104.7	102.5	**99.6**	**3.9**

Several calibration points demonstrated very low repeatability (**Table 4.2**), which seems to affect both cortisol and cortisone quantitation despite the correction by IS. This could be explained by irregular injections, which was confirmed by the disappearance of the problem after replacement of the injection needle and needle seat capillary between the 5th and 6th days. The recovery of cortisol and cortisone spiked into the charcoal-stripped oral fluid at various concentrations ranged from 75.2 to 86.4% for cortisol and from 69.1 to 81.4% for cortisone (**Appendix C, Supplementary Table C2**). The calculation of recovery using either solvent calibration curve or comparison of pre-and post-extraction spiked samples yielded similar results. Extraction repeatability was excellent, with an RSD ranging from 2.6 to 10.2% for cortisol and from 1.9 to 9.5% for cortisone. Cortisol stability of up to 1 year[129] and 3 years[164] when stored at -80°C was

previously established, so it was not examined in the current study. **Table 4.3** summarizes the results for short-term and long-term stability tests. Both cortisol and cortisone in wet extracts of oral fluid demonstrated satisfactory stability for bench manipulations (2 hours at ambient temperature).

The stability test of cortisone in autosampler conditions for 24 hours (**Table 4.3**) demonstrated surprising failure for the cortisone, and the reasons for that were unclear. The inconsistent performance of the autosampler may be involved because the analysis was done a day before the injector failure detection and the replacement of the injector assembly. However, the injector failure was not corrected by the IS (cortisol (d_4)), which implies the necessity to investigate the compliance of cortisol (d_4) as IS for cortisone.

*Table 4.3.Summary of short-term and long-term stability results. Stability was assessed by comparing the measured concentration assessed by the freshly prepared calibration curve (**Section 2.5.3**) of stability samples to the expected theoretical concentrations. Samples with concentration 1.09 ng/mL destined for the assessment of the stability at 2 hours, ambient temperature were lost during manipulation. A single replicate per condition was tested.*

Expected concentration in samples (ng/mL)	2 h, ambient temperature, wet extracts		24-h autosampler, +7°C, wet extracts		6 months, < - 70 °C, dried extracts		6 months, < - 70 °C, wet extracts	
	Cortisol	Cortisone	Cortisol	Cortisone	Cortisol	Cortisone	Cortisol	Cortisone
1.09	NA	NA	95.5	120.3	34.4	12.4	141.7	115.7
4.38	99.6	115.1	114.4	144.0	18.0	8.3	124.4	107.6

Moreover, in the analysis of samples from study participants, the cortisone demonstrated acceptable stability in the batch consisted of 120 injections with a total length of 14 h (**Figure 4.3**). In the validation, however, the longest incubation of standard samples was carried in the intraday analysis batch, which was consisted of 72 injections with a total run time of 8.5 h. In this batch, the concentration of cortisone was stable across the entire time (**Appendix C, Supplementary Table C3**).

In addition, cortisone showed acceptable stability for long-term storage when reconstituted extracts were stored directly (wet extracts) (6 months storage at below -70°C). Cortisol wet extracts showed higher variability than expected, thus not meeting acceptance criteria. In contrast, dried extracts stored after evaporation indicated 3 to 12 times lower concentrations after 6 months of storage at

below -70 °C followed by reconstitution prior to LC-MS/MS analysis. This suggests either degradation and/or precipitation (**Table 4.3**). Finally, the stability of cortisol and cortisone in oral fluid during freeze-thawing was evaluated. All results met the 80-120% stability acceptance criteria (**Table 4.4**). In sum, long-term storage of prepared extracts is not recommended, so samples should be analyzed immediately after preparation. Therefore, the length of the analytical batch is not recommended to be extended beyond 14 hours. Unfortunately, the entire set of samples with a concentration of 1.09 ng/mL was lost during sample manipulation, and the experiment was executed with only one (4.38 ng/mL) sample.

*Table 4.4. Cortisol and cortisone freeze-thawing stability (using validation samples (**Section 2.2.6.4**) with a concentration of cortisol and cortisone at 4.38 ng/mL) was assessed as the accuracy of quantitation in single replicates. (Expected theoretical concentration = 100%).*

Number of freeze-thaw cycles	Concentration accuracy (%)	
	Cortisol	Cortisone
1	101.6	106
2	90.6	92.0
3	94.8	96.3

The evaluation of absolute matrix effects for cortisol and cortisone in charcoal-stripped oral fluid showed minor matrix effects with two and one samples, respectively exceeding 80-120% acceptance criteria, where absolute matrix effects are considered negligible (**Figure 4.1**). When cortisol (d_4) internal standard was used to correct matrix effects, all results meet acceptance criteria except at the LOD level of 0.15 ng/mL. Moreover, as expected, the internal standard correction reduced data variability for cortisol (SD = 9.4 for n=12 as compared to SD = 23 before correction) and cortisone (SD = 4.0 for n=12 as compared to SD = 14 before correction). This indicates that even though cortisol and cortisone do not co-elute, cortisol (d_4) may serve as an internal standard for both cortisol and cortisone in the stripped oral fluid matrix (**Figure 4.1**). However, in subsequent experiments, we further examined the dependence of recovery and matrix effects on oral fluid composition by evaluating these parameters in eight different lots of oral fluid. The results of the recovery experiment are shown in **Table 4.5**.

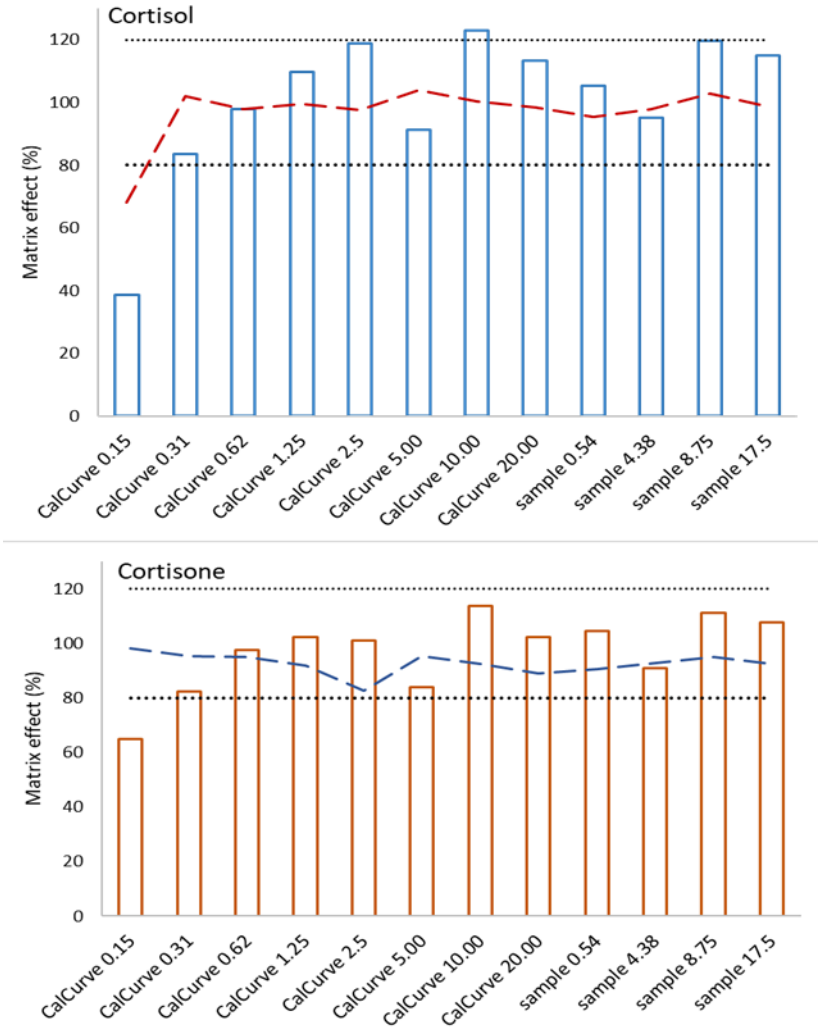


Figure 4.1. Correction of matrix effects in charcoal-stripped oral fluid for the calibration curve (Cal Curve) and validation samples (sample) at different concentrations. The columns show absolute matrix effects without internal standard correction, while the dashed line shows the matrix effects after the correction of analyte peak areas by cortisol (d4) internal standard. Values of 80-120% shown by dotted lines depict acceptance criteria where matrix effects are considered negligible. Values above 120% show ionization enhancement, and values below 80% show ionization suppression. Digits in the sample names on the x-axis correspond to their concentrations in ng/mL.

The data demonstrates that the recovery of both cortisol and cortisone in real oral fluid samples is similar to each other and to the recovery of IS (cortisol (d_4) across different lots of oral fluids. The t-test comparing cortisone recovery versus cortisol and cortisol (d_4) recoveries showed no significant difference with p-values of 0.56 and 0.97, respectively. More importantly, the recovery results also show that there is a high variability of recovery from lot to lot of oral fluid when

comparing these results to **Supplementary Table C2 (Appendix C),** but that the internal standard may be able to correct for this effect.

*Table 4.5. Recoveries of cortisone, cortisol, and cortisol (d_4) and accuracy of their quantification in eight lots of non-striped oral fluid after correction by cortisol (d_4). Samples were spiked with cortisol and cortisone to yield 2 ng/mL in the extraction tube. Recovery was calculated for each lot of oral fluid using peak areas in three extracts: (i) pre-extraction spiked; (ii) not spiked (endogenous concentration) and (iii) post-extraction spiked replicates by formula (pre-extraction spike area- endogenous area)*100/(post-extraction spike area-endogenous area). Analyte concentrations in each sample were calculated using a calibration curve built-in charcoal-stripped oral fluid matrix. The difference between pre-extraction spiked and non-spiked sample concentrations was compared to the theoretical amount of cortisol and cortisone spiked to assess the accuracy of the final method.*

Sample ID	Recovery (%)			Accuracy of assessment (%)	
	Cortisone	Cortisol	Cortisol (d_4)	Cortisone spike	Cortisol spike
Lot1	87.7	86.5	87.4	108.7	105.7
Lot2	56.7	68.0	62.7	76.9	102.8
Lot3	35.5	41.7	40.1	81.8	100.4
Lot4	108.1	102.4	92.1	90.7	103.2
Lot5	69.1	80.8	70.6	86.8	106.4
Lot6	74.3	88.4	79.7	62.6	104.8
Lot7	95.0	100.3	81.0	83.7	105.8
Lot8	70.7	79.9	86.6	63.6	100.5

However, during the evaluation of accuracy using cortisol (d_4) as an internal standard for both analytes, it was discovered the accuracy of quantitation of cortisone is much poorer than for cortisol (mean accuracy of 81.9±14.9 for cortisone versus 103.7±2.4 for cortisol, significant t-test with a p-value of 0.001). Although all validation results met FDA criteria in charcoal-stripped pooled oral fluid when faced with compositional variability of real oral fluid samples and corresponded matrix effects (**Figure 4.2**), the accuracy of cortisone quantitation was inferior compared to cortisol and was not sufficient as judged by the standard addition method. The combined effect of recovery differences and different matrix effects resulted in poor accuracy of cortisone when using cortisol (d_4) internal standard. Thus, to achieve acceptable accuracy for cortisone, cortisone (d_8) internal standard was added to the finalized method before any analysis of real samples.

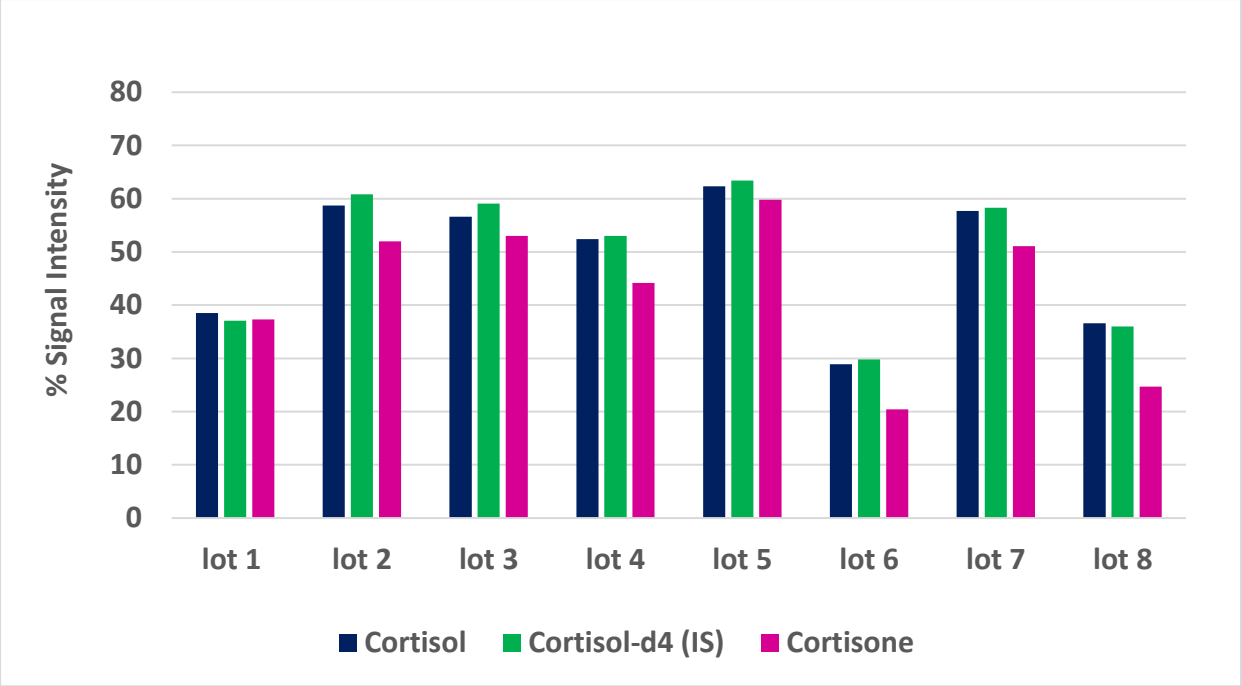


Figure 4.2. Summary of matrix effect results obtained for eight lots of non-stripped oral fluid. The signal intensity of 80-120% would indicate no significant matrix effects, and values below < 80% show significant ion suppression.

As expected, the incorporation of cortisone (d_8) as IS for cortisone reduced the quantitation inaccuracy in real oral fluid samples to the acceptable 85-115% range, which is comparable to cortisol (**Figure 4.2**). In contrast, other studies successfully used a single internal standard for both cortisol and cortisone but combined it with more selective sample preparation such as LLE[231] and LLE-derivatization-online SPE.[232] Our LC-MS/MS SRM method has been improved without compromising its robustness and analytical performance validated as described in **Section 4.2.6**. The overall improvement is reflected by acceptable levels of accuracy (**Figure 4.3)** demonstrated by the method in the analysis of oral fluid samples from study participants. This example of poor cortisone accuracy illustrates a key difficulty during the validation of methods for endogenous biomarkers such as cortisol and cortisone, where it is usually not possible to obtain a matrix that does not contain the analytes of interest.

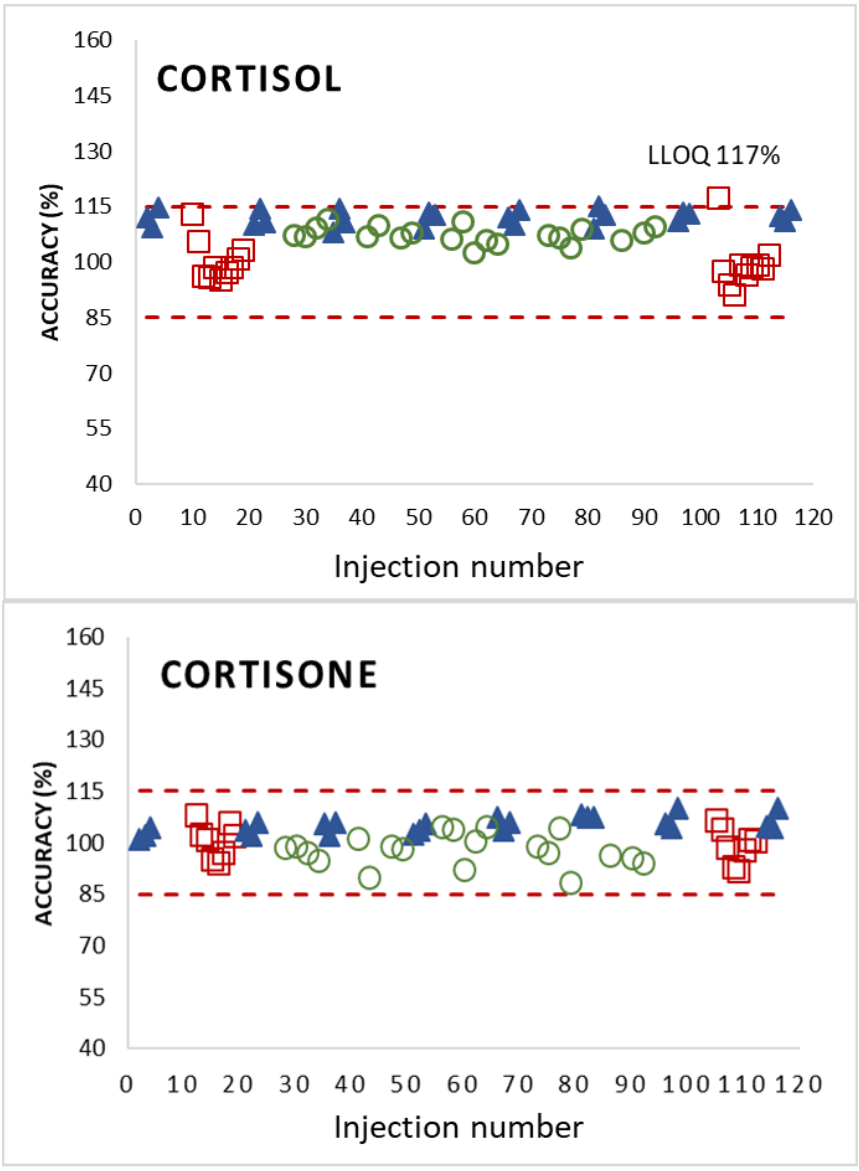


*Figure 4.3. Accuracy of cortisol and cortisone quantitation across an analytical batch of 51 participant's samples. All samples were quantified using the averaged external calibration curve (squares) injected at the beginning and at the end of the batch. Three QC samples: low, medium, and high (triangles) were injected one after another at the beginning, after every 10 injections, and at the end of the analytical batch. Circles represent individual saliva samples spiked with known amounts of cortisol and cortisone, extracted and quantified in parallel to aliquots of the same samples that were not spiked (not shown on the graph). The accuracy was evaluated using the formula: (concentration in spiked samples - concentration in not spiked samples)*100% /theoretical spiked concentration, that still satisfies limits of FDA (less than 120%). Except for one sample at LLOQ level of cortisol (0.31 ng/mL, at LLOQ 80-120% accuracy is allowed according to FDA), all samples were quantified within 85-115% accuracy range (red dashed lines).*

The systematic positive bias of the calibration curve was observed for cortisol but not for cortisone in standards and QC samples **(Figure 4.3).** This indicates a tolerable error in the preparation of spiking solutions of authentic or/and SIL cortisol from the working stock.

Charcoal-stripped saliva was previously used for method validation in several studies,[232,234] but matrix effect evaluation was performed only in charcoal-stripped oral fluid, without comparison to real oral fluid samples. Meszaros *et al*. did assessed overall method accuracy in four lots of non-striped saliva, but the results obtained for both cortisol and cortisone were acceptable with mean values of 97.9% and 99.7%.[234] It is possible that their method, which utilized protein precipitation with acetonitrile and (-ve) ESI LC-MS/MS analysis, had slightly different selectivity or that the four lots tested were not sufficient to detect higher variability and poorer accuracy of cortisone. Other clinical LC-MS/MS assays have used artificial solutions such as 0.2% w/v bovine serum albumin,[233] phosphate buffered saline,[238] artificial saliva,[242] or water[240,250] for the calibration of cortisol assay. However, Fustinoni *et al*. did supplement their determination by evaluation of relative and absolute matrix effects in six lots of oral fluid and did not find any significant matrix effects.[240] Jonsson *et al*. compared water and oral fluid for one lot and did not find significant matrix effects for cortisol, but cortisone was not measured in that method.[250]

4.3.2 Cortisol quantitation by LC-MS/MS and ELISA

Using the validated LC-MS/MS assay, cortisol was detected at concentrations between 0.34 and 1.78 ng/mL (0.9 and 5 nmol/L) (**Appendix C, Supplementary Table C4**). The concentrations of cortisol below LLOQ (0.31 ng/mL) of LC-MS/MS assay were detected in 4 out of 51 samples and fell between LOD and LLOQ of the assay (0.15-0.31 ng/mL). The concentrations of cortisone in oral fluids of healthy individuals varied between 3.2 and 17.3 ng/mL (8.9-48.0 nmol/L). These values match, in general, the results obtained by Perogamvros *et al.* in oral fluids of healthy individuals.[238] They also reported decreasing cortisol-cortisone ratios from a mean of 0.2 in the morning to a mean of 0.1 in the evening. Cortisol-to-cortisone ratios obtained in this study ranged from 0.05 to 0.18, with a mean of 0.12. These results match well the ratios obtained at 8:30 am (mean 0.16 to 0.23 for determinations on three different days) and noon (mean 0.12-0.14 on three different days), which are the time points that are the closest to our collection times.[247] These results are also comparable to the reported mean ratios of 0.13 at noon in healthy individuals and 0.16 reported for 9:30 am sampling time in late adolescent and young women.[251] Thus, it can be

concluded that there is a 1.9x standardization error between the two assays, and to resolve this standardization error requires comparison to traceable reference material. Others have also reported standardization errors between ELISA and LC-MS/MS resulting in the overestimation of cortisol concentrations[164], including blood cortisol measurements.[252] However, the standardization of assays for oral fluid has lagged behind and requires further attention.

The measurement of cortisol concentrations in oral fluid samples collected during the study using our validated LC-MS/MS method and ELISA revealed a drastic overestimation of cortisol concentrations by competitive ELISA as compared to LC-MS/MS (**Figure 4.4**).

Table 4.6. Quantitation of ELISA cortisol calibration standards (supplied with the ELISA kit) by LC-MS/MS method using an in-house cortisol calibration curve. The curve was prepared by serial dilutions of in-house cortisol stock standard using ELISA calibration blank supplied with the kit.

ELISA calibration	ELISA setpoint (ng/mL)	Concentration by LC-MS (ng/mL)	Ratio ELISA / LC-MS/MS
ELISA calibration blank	0.0	ND	NA
ELISA calibration point 1	0.1	ND	NA
ELISA calibration point 2	0.4	< LOQ	NA
ELISA calibration point 3	1.7	0.78	2.2
ELISA calibration point 4	7.0	3.42	2.0

Two main possible reasons for ELISA to yield systematically higher results than LC-MS/MS are antibody cross-reactivity and improper standardization. To investigate both of these possibilities, we first analyzed ELISA blanks and calibration standards using LC-MS/MS. The results of this comparison are shown in **Table 4.6** and show that the concentrations of ELISA standards are 1.9x lower than stated when quantified against an in-house cortisol calibration curve built in the same buffer.

Next, cortisol and cortisone calibration standards prepared in water were quantitated using ELISA and its calibration curve. As expected, these standards showed 1.9x higher results when quantitated using the ELISA calibration curve (**Table 4.7**).

Thus, it can be concluded that there is a 1.9x standardization error between the two assays, and to resolve this standardization error requires comparison to traceable reference material. Others have also reported standardization errors between ELISA and LC-MS/MS resulting in ELISA

overestimation of cortisol concentrations[164], including blood cortisol measurements.[252] However, the standardization of assays for oral fluid has lagged behind and requires further attention.

Table 4.7. ELISA quantitation of cortisol and cortisone using in-house standards prepared in water from the same stock solution (2 µg/mL) whose aliquots were used for the preparation of calibration curve in LC-MS/MS method.

Standard composition	Cortisol ELISA results (ng/mL)		
	Replicate 1	Replicate 2	Mean
Cortisol in water 0.63 ng/mL	1.0	1.2	1.1
Cortisol in water 2.5 ng/mL	5.4	5.4	5.4
Cortisol in water 10 ng/mL	18.7	20.1	19.4
Cortisone in water 0.63 ng/mL	0.2	0.2	0.2
Cortisone in water 2.5 ng/mL	0.4	0.3	0.35
Cortisone in water 10 ng/mL	0.5	0.3	0.4

The measurement of individual cortisone standards prepared in water by ELISA also allowed us to investigate antibody cross-reactivity to cortisone. As shown in **Table 4.7**, antibody cross-reactivity to cortisone may be contributing approximately 5-20% inaccuracy to ELISA assay, with lower concentrations of cortisol more significantly impacted. This level of cross-reactivity is higher than stated by the manufacturer at 0.8%, but it is not known which concentration(s) the manufacturer used to evaluate the cross-reactivity.

In sum, the cross-reactivity to cortisone yielded around 17.1 % of cortisol response at low (0.63 ng/mL (1.8 nmol/L) concentration and becomes more negligible at higher concentrations. These results are similar to the cross-reactivity results reported by Bae *et al*. for IBL immunoassay, where cross-reactivity to cortisone contributed from 2.6 to 43% of inaccuracy for similar cortisol concentrations and cortisol-cortisone ratios.[164] Indeed, the overestimation of cortisol concentrations by ELISA compared to LC-MS/MS was observed in 18 (out of 19) individual samples (**Figure 4.4**).

Moreover, the non-linear ($r = 0.094$, $n = 46$) increase in the overestimation of cortisol concentrations by ELISA compared to increasing cortisol concentrations in LC-MS/MS was demonstrated by the Bland-Altman plot (**Appendix C, Supplementary Figure C2**) for 46 study samples. This suggests the involvement of other factors beyond standardization and cortisone cross-reactivity, one of which could be the cross-reactivity to other structures similar to cortisone.[226]

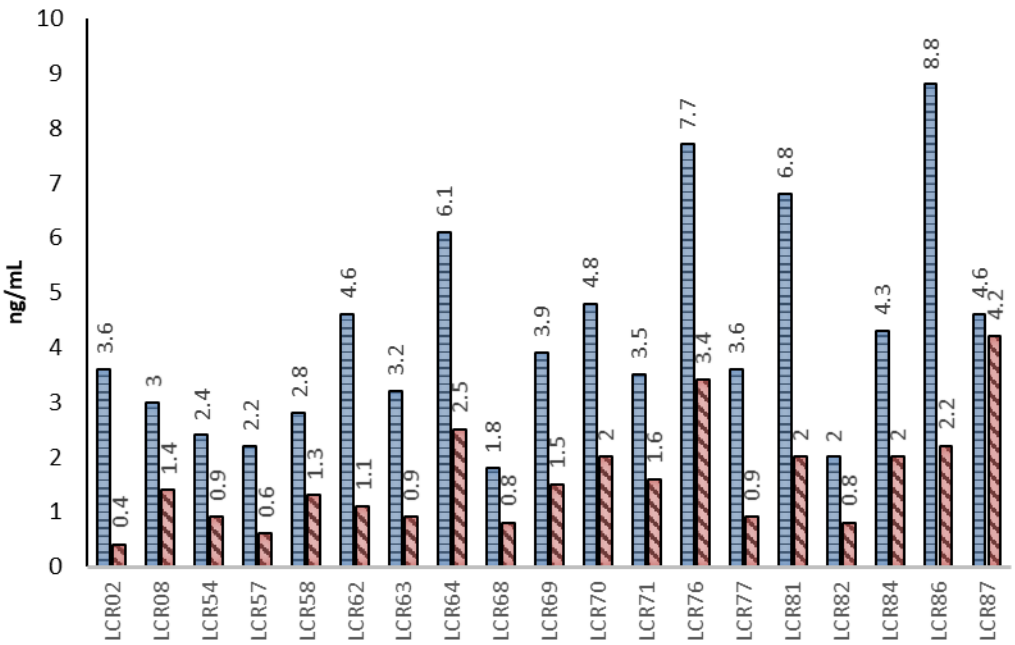


Figure 4.4. Quantitation of cortisol in oral fluid samples from participants by competitive ELISA (columns with horizontal stripes) and in-house LC-MS/MS method (columns with diagonal stripes). Concentration is represented by Y-axis. Measurements executed in single replicates. The results of two-tail, pair-wised t-test (p = 5.73E-07, (n=19)) support the confidence in observation of higher quantitation results by ELISA.

Based on the manufacturer's data, other structurally related hormones also show cross-reactivity, among which 11-deoxycortisol and 17-hydroxyprogesterone demonstrate higher cross-reactivity than cortisone.

A more comprehensive investigation of ELISA interferences in real samples was beyond the scope of this study, but our results indicate that cortisone is not the main/only source of inaccuracy, at least for the AbNova kit. More than 2-fold differences were observed between results of ELISA and LC-MS/MS in test samples from participants **(Figure 4.4**). In addition, correlation analysis shown in **Figure 4.5** demonstrates poor linearity (r^2 = 0.27), indicating poor agreement between the two methods. Finally, the correlation analysis shows huge positive Y intercept which indicates the presence of interfering (non-analyte) cross-reacting factor. Altogether, this data indicated, that cross-reactivity to cortisone and 2x standardization error do not explain all the reasons for inaccurate cortisol quantitation by ELISA kit.

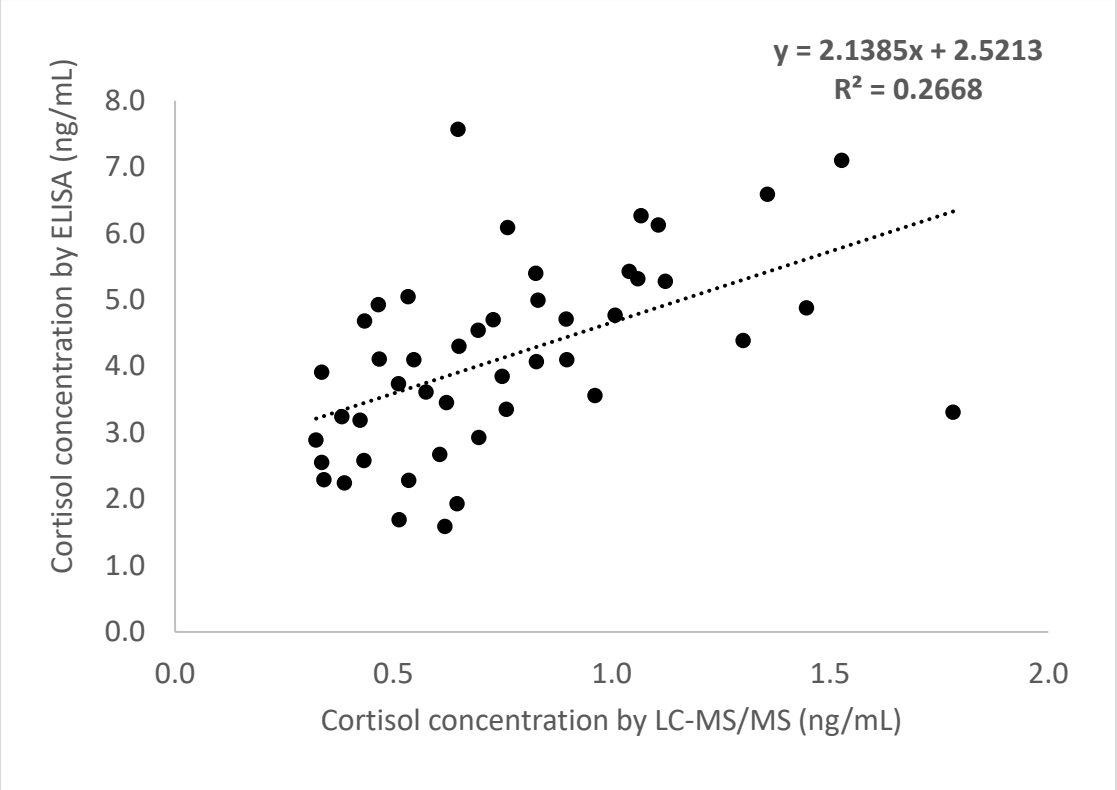


*Figure 4.5. Correlation analysis between cortisol concentrations obtained using LC-MS/MS and ELISA for n=46 oral fluid samples collected at three time points (**Section 4.2.3).** Only samples with signals above LLOQ in both assays were used for this plot. Linear least-squares regression results are shown in the top right corner of the figure.*

In an early study, Jonsson *et al*. showed radioimmunoassay overestimated cortisol concentrations by 2.7x.[49] Similarly, Bae *et al*. showed IBL luminescence immunoassay overestimated cortisol levels by 2.0-2.6x but showed much higher correlation values of ~0.9.[164] Our Bland Altman analysis shown in **Figures 4.6 A and 4.6 B** indicates clearly the presence of proportional bias: as the concentration of cortisol increases, the agreement between the two methods becomes closer.

Considering that LC-MS/MS assays are accepted as a gold standard method for cortisol analysis, Bland-Altman plots in **Figure 4.6 A** were plotted both against the mean of the two assays and against LC-MS/MS assays values as a gold standard method. The latter results show very weak proportional bias with a much lower r^2 value of 0.09. Electroluminescent ECLIA assay showed good agreement with LC-MS/MS (r^2 of 0.892)[234] across the entire cortisol concentration range tested, but only r^2 of 0.02 for cortisol values below 2 nmol/L. Similarly, the Bland-Altman plot in a study by Bae *et al*. revealed a wide dispersion of results below 5 nmol/L.[164] In sum, our results

for the comparison of LC-MS/MS against the Abnova immunoassay are consistent with other literature comparisons at low cortisol levels, indicating poor agreement between the two methods.

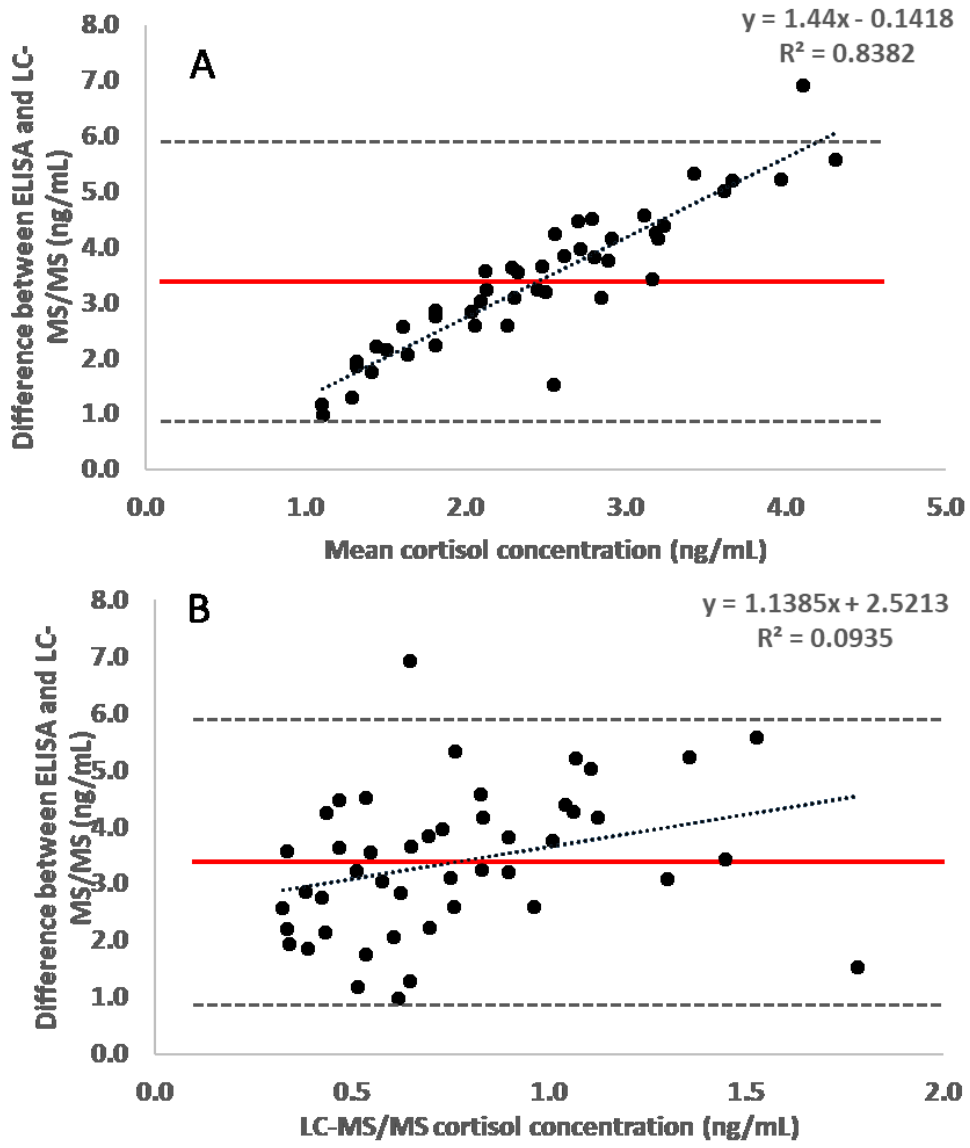


Figure 4.6. Bland-Altman analysis for cortisol measurements using LC-MS/MS and ELISA prepared by plotting (A) mean concentration of both assays on the x-axis or (B) LC-MS/MS assay concentration on x-axis. In both cases, the y-axis shows the absolute concentration differences (in ng/mL) obtained by two methods. The red line shows the mean difference between the assays, while the dashed lines show limits of agreement as determined by ± 1.96xSD. Linear least-squares regression was also performed, and its results are shown in the top right corner.

Clinical automated immunoassays such as Roche Cortisol II assay that uses monoclonal antibodies showed better agreement with LC-MS/MS for cortisol measurement in serum (n=405) with the

slope of 1.02 and intercept of 4.473 nmol/L, but an agreement for saliva (n=253) was poorer with the slope of 1.12 and the intercept 0f 0.825 nmol/L.[253] Nevertheless, none of these studies included ELISA tested in our study. Our results demonstrate greater specificity of LC-MS/MS method than ELISA assay, confirm excessive values of ELISA calibration setpoints, and call for further investigation of ELISA interferences in real samples. In conclusion, there is an obvious necessity to correct the values of ELISA calibration setpoints according to LC-MS/MS quantitation, and validated LC-MS/MS assays such as the one presented provide a better alternative for more selective cortisol and cortisone measurements.

Our results of Bland-Altman analysis in **Figure 4B** show that immunoassay results cannot be converted into correct cortisol concentrations by a simple function. Miller *et al*. compared five immunoassays with LC-MS/MS using 195 saliva specimens.[248] Despite high correlation coefficients ranging from 0.92-0.96 against LC-MS/MS, they found non-linear relationships between different immunoassays and LC-MS/MS assay and proposed a set of conversion functions to facilitate inter-assay and inter-study comparisons of cortisol concentrations. To examine if both ELISA and LC-MS/MS analysis would provide the same conclusions in a given study, we have plotted relative change at t1 and t2 versus cortisol concentrations at the onset next to each other. The results of this analysis are shown as a mirror plot in **Figure 4.7** and indicate that the use of ELISA or LC-MS/MS would have resulted in incorrect conclusions about the change in cortisol levels in at least 14 instances out of 34 tests even when using the conservative value of 50% to indicate up- or down-regulation of cortisol in an individual during the test.

Furthermore, since 3-4 activity tests were performed per individual, we examined the dataset to see if only particular individuals showed these discrepancies; however, no trends, according to the individual, could be established. Based on these results, ELISA testing cannot be used to provide correct results for acute laboratory psychosocial stress tests. To the best of our knowledge, no one else examined relative assay results obtained by the two methods. Bae *et al.* examined mean trends across time and found both immunoassay and LC-MS/MS could provide correct conclusions about cortisol levels.[164] Thus, it is believed that despite standardization errors and interferences often observed for ELISA, the results for an individual test could still be usable by looking at relative changes. Our results dispute this supposition, at least for studies where low levels of cortisol are being measured in healthy individuals (< 2 ng/mL).

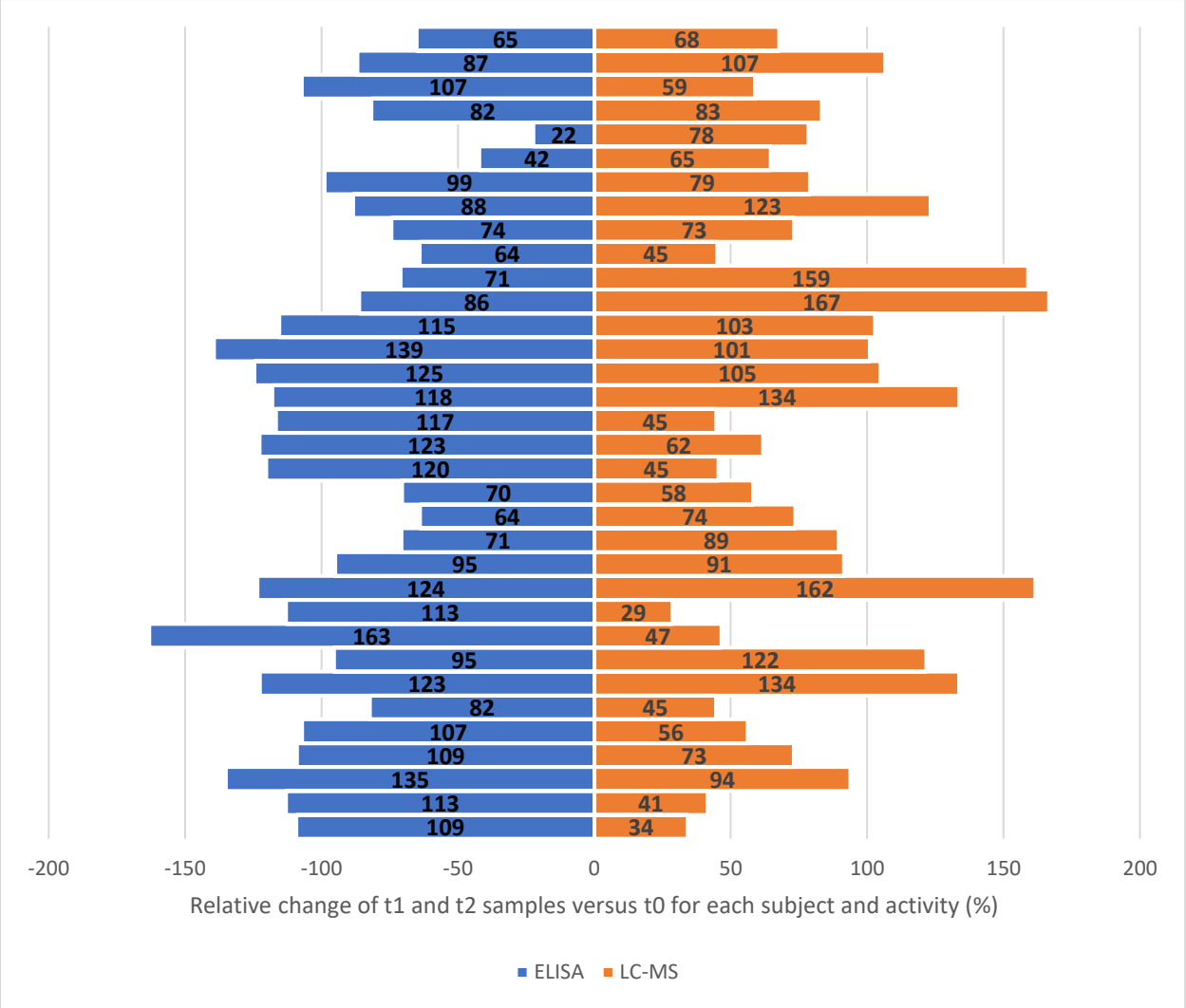


Figure 4.7. Mirror plot showing relative changes in cortisol concentration at t1 and t2 (versus t0) obtained by ELISA and LC-MS/MS for each individual collected in our study.

4.4 Conclusions and future recommendations

Cortisol oral fluid measurement is preferred to serum measurement, especially for acute laboratory psychosocial stress tests due to non-invasive sample collection. We have developed and implemented an inexpensive, simple, accurate, robust, and high-throughput LC-MS/MS assay for simultaneous absolute quantitation of cortisol and cortisone in oral fluids. The method replaces LLE and SPE with simpler and less expensive methanol extraction while ensuring excellent quantitative performance despite limited sample clean-up. The developed LC-MS/MS method out-competes the ELISA approach in throughput and accuracy. The throughput of LC-MS/MS is high, and considering all necessary calibration, blank, and QC injections, the method allows quantitation

of both target steroids in 72 samples per 14 h batch. In comparison to other LC-MS/MS methods, this method reduces the volume of oral fluid required for analysis to 30 μL increasing its applicability. However, the sensitivity of the current LC-MS/MS method was not sufficient to quantify cortisol in 4 out of 51 late morning samples. Therefore, further improvements in LC-MS/MS sensitivity towards cortisol are required if late-night salivary cortisol measurements are of interest. The comparison of Abnova ELISA and LC-MS/MS assay showed that ELISA significantly overestimates actual cortisol concentrations in oral fluid samples primarily due to excessively high values of calibration setpoints. Therefore, the accuracy of the cortisol quantitation by ELISA with corrected calibration setpoints should be investigated on a large set of samples originating from different individuals and representing the wide range of cortisol concentrations. Moreover, a partial (17% of cortisol signal) cross-reactivity of ELISA kit between cortisol and cortisone at lower concentrations cannot explain the increase of ELISA error with increasing concentrations of cortisol, indicating other interferences must be present. Therefore, further investigation of the presence of compounds with structures similar to cortisol and their involvement in the generation of overestimated results is necessary. This can be accomplished by performing LC-MS identification of compounds which are bound to the antibodies used in the ELISA kit used in the study. Overall, our results demonstrated that at cortisol concentrations tested in this study (< 2 ng/mL), ELISA does not yield either correct relative or absolute results limiting the utility of this technique for daytime (late morning to evening) time-point measurements. The poor correlation between ELISA and actual salivary cortisol concentrations requires urgent attention, as it may result in serious consequences for patient health and reputation of manufacturing companies and test laboratories.

5 Conclusions and future directions

5.1 Conclusions

No single analytical method can successfully measure all metabolites present in a given biological sample. Although numerous strategies can be used to increase metabolite coverage, my thesis focused specifically on the role of sample preparation in metabolite coverage of untargeted metabolomics. Unselective organic solvent precipitations dominate in metabolomics sample preparation due to their simplicity and broad metabolome coverage. Other sample preparation methods such as SPE were considered not suitable for a global metabolomics analysis due to narrow selectivity, which should reduce metabolome coverage compared to solvent precipitations. However, selective methods and SPE, in particular, can enhance the coverage for select metabolites by enrichment and sample decomplexation compared to solvent precipitations. Therefore, it was postulated that the combination of several methods with narrow but complementary selectivity might increase metabolome coverage compared to solvent precipitations. In my thesis, I examined the performance of sample preparation methods with narrow selectivity and the effects of their combinations on the increase in the coverage of polar and mid-polar metabolomes. In my thesis, I examined the performance of sample preparation methods with narrow selectivity and the effect of their combinations on the increase in the coverage of polar and mid-polar metabolomes.

Therefore, the first study was focused on a detailed comparison of solvent precipitations, LLE, and several SPE methods to identify the most complementary sample techniques for global metabolomics. Orthogonality, repeatability, absolute recovery, and matrix effects were evaluated using simple standard solutions and human plasma. The results confirmed the broad selectivity of solvent precipitations, which were superior to MTBE LLE and SPE, and provided the highest metabolite coverage in a single method but suffered from severe matrix effects. Furthermore, changing the composition of the extraction solvent from MeOH to methanol-ethanol cocktail provided a negligible increase in the coverage disregarding LC-MS conditions. Thus, RP (+ve) ESI demonstrated the increase of metabolome coverage in a methanol-ethanol cocktail to 3804 against 3795 in MeOH and an overlap of 3428 metabolites between the two methods. Comparisons of all extraction methods to each other revealed the most orthogonal metabolite composition in MTBE LLE and mixed-mode IEX SPE. In this experiment, IEX SPE included the mix of SAX and SCX stationary phases based on DVBP support. The IEX SPE presented an acceptable analytical

performance with good recovery and repeatability. The results provided sufficient data to design the sequential sample preparation, which combined MTBE LLE, MeOH protein precipitation, and IEX SPE, which was evaluated in the second study.

The need for the maximum decomplexation prompted the design of the sequential SPE in which plasma was fractionated in two steps. In the first step, the enrichment of cation and neutral metabolites in unbound fraction (CN) and anions and zwitterions into a bound fraction (AZ) in SAX (MAX plates) with binding pH of 9.2 and elution pH of 2.4 is executed. Int the second step, the fractionation of CN and AZ in SCX (MCX plates) with binding pH of 2.4 and elution pH of 9.2 generated four individual fractions N, A (MCX unbound) and C, Z (MCX bound), enriched with neutral, anion, cation and zwitterion metabolites, respectively. Wash and elution media contained 100% organic solvents to counteract hydrophobic interaction between metabolites and the DVBP stationary phases of the SPE plates. The initial protocol demonstrated good orthogonality of individual fractions to each other in standard and global metabolomics analysis and up to a 1.5-fold increase in metabolome coverage in sSPE compared to MeOH. Next, the initial inclusion of MTBE LLE into sequential sample preparation was reconsidered due to the incomplete removal of many lipids, making it not very effective for LC column protection. These results shaped the final design of sequential sample preparation, consisting of MeOH protein precipitation and sSPE. The protocol explicitly focused on polar and mid-polar metabolites because other sequential SPE methods were already successfully reported the effective fractionation of lipidome. In addition to metabolome coverage, the sequential (MeOH-sSPE) sample preparation was evaluated for the critical parameter, namely, the split of metabolites between fractions. The protocol demonstrated an orthogonal composition of cation and anion fractions (< 4% overlap) and < 36% metabolite overlap between pairs of other fractions. Moreover, the analysis of individual fractions of sSPE demonstrated a 1.6-fold increase in total and 3.4-fold increase in unique metabolome coverage compared to MeOH. Finally, optimal combinations of fractions that provide increased metabolome coverage and reduce metabolite overlap were selected for each LC-MS analysis to reduce analysis time 2-fold.

During the analysis of combined fractions the repeatability, split of metabolites, and metabolome coverage in sSPE were thoroughly assessed. The developed method is highly repeatable and demonstrated an average (n=6) RSD less than 30% for nearly 75% of metabolites in global metabolomics analysis. Total metabolome coverage in sSPE was 1.6-fold higher compared to

MeOH. Notably, a combined sSPE provides excellent separation of anion and cation metabolites, with less than 27% of these metabolites showing detectable splitting. However, our method demonstrated several issues, which could be improved in future developments. One of these issues was low metabolome recovery in sSPE. The group of very polar metabolites was the most affected. Another slight issue was observed with the split of zwitterion and neutral metabolites into A and C fractions. The most feasible resolution for those issues could be found via testing different SPE plates and modifications of ionic and organic compositions of SPE and LC-MS solvents. Overall, the performance of the method was robust, comparable to the MeOH extraction by all other evaluation criteria and satisfies the requirements of global metabolomics studies with the above noted improvement in metabolome coverage.

In parallel to the execution of the first and second study's objectives, the development of targeted analysis of cortisol and cortisone in oral fluid was carried out. Although the thesis's main focus was on plasma as the biological matrix of choice, the rationale for the analysis of oral fluid was an easier detection of some metabolites, which reflect the metabolome status of blood but are difficult to detect in plasma or serum. Besides, the analysis of two fluids in a single study provides the potential to increase the metabolome coverage and the knowledge of the metabolic status of the whole organism. Moreover, the oral fluid allows less invasive sample collection for metabolomics outside the laboratory.

The MRM LC-MS method for the absolute quantitation of cortisol and cortisone in the oral fluid was developed and validated. The performance of the LC-MS method was compared with a commercial cortisol ELISA kit. The developed and validated method provides a low-cost, accurate, robust, and high-throughput MRM LC-MS/MS quantitation of cortisol and cortisone in oral fluids and demonstrated LLOQ of 0.31 ng/mL. The comprehensive validation of the method according to regulatory guidelines confirmed intra- and inter-day accuracy in the range of 85-115% and precision (≤15% RSD) at all concentrations above LLOQ, which fit into performance limits recommended by FDA. The developed LC-MS/MS method out-competes ELISA in specificity, and probably in cost per analysis at high number of analyses executed at once. The comparison of this validated method to ELISA showed that the latter overestimated the concentrations of cortisol in the samples due to antibody cross-reactivity and improper standardization, especially at low concentrations. However, the LC-MS/MS method sensitivity was insufficient to quantify cortisol in four out of 51 samples, indicating the need for further improvement especially if late-night

cortisol salivary values are of interest. The development of sSPE provides a feasible solution for that. Thus, the recovery of cortisol in CN fraction from plasma in the (+ve) RP C18 analysis of combined sSPE was 20% higher than in MeOH extract of oral fluid with a yield of 100%. Therefore, a single step MCX SPE should be tested for the targeted analysis of cortisol in oral fluid in order to potentially improve sensitivity.

5.2 Importance to the field and future directions

One of the big challenges in global metabolomics is its untargeted nature and the desire to discover/characterize as many biomarker metabolites as possible via a semi-quantitation approach, which typically analyzes hundreds to thousands of unknown metabolites simultaneously. This means that global metabolomics methods cannot be validated in the same way as methods targeting one or several analytes. This inability to perform full validation studies also brought hesitancy in the metabolomics community to execute *any* quantitative studies of method performance such as recovery and matrix effect studies. Although these studies cannot be executed for all metabolites present, obtaining such data on even a subset of metabolites provides invaluable information on method performance, aid in rational method selection, and help to move the field forward.

Thus, one of the key accomplishments of my thesis research was the demonstration of the advantages and limitations of both well-established methods such as solvent precipitations as well as infrequently employed ones, such as SPE for global metabolomics. Hence, my thesis challenges the wide-spread hesitance to use SPE in global metabolomics because of its selectivity and narrow metabolome coverage. Therefore, my second accomplishment was the employment of a sequential mixed-mode SAX-SCX SPE, which produces fractions with narrow, but complementary composition and a > 1.6-fold broader metabolome coverage compared to MeOH, including very hydrophilic/charged metabolites. Another major accomplishment of the study was the establishment of the orthogonal sSPE fractionation of metabolites into anion, cation, neutral, and zwitterion fractions with a degree of orthogonality sufficient to assign anionic and cationic properties to metabolites based on their detection in a particular sSPE fraction. Although sSPE demonstrated increased metabolome coverage, exploration of the full potential of the method required its optimal integration with LC-MS analysis.

Therefore, another major accomplishment of the study was the optimal integration of combined sSPE fractions into RP and ZIC-HILIC LC-MS methods with a 2-fold decrease in analysis time, and a decrease in the split rate to below 30%.

The practical impact of this new sequential SPE method in metabolomics is that it can lead to the increased pathway coverage via the increased detection of polar and mid-polar metabolites compared to classical solvent precipitations. In addition, the knowledge of the fractionation behavior of newly discovered biomarkers enables their faster transition from a global to a targeted assay. Finally, sSPE fractionation utilizes only volatile solvents and buffers, thus allowing direct MS analysis of samples or facile adaptation to 2-D on-line sample analysis in LC-MS and SPE-MS approaches.

Throughout this research, the limitations of the chosen sets of metabolite standards became apparent and were revised to be more representative of metabolite diversity in plasma. As noted at the beginning of this chapter, a targeted approach always follows fewer compounds than a global one. However, metabolite standard cost, solubility issues, and a lack of a priori knowledge about method selectivity of some methods impeded the design of optimal metabolite sets to use for these types of evaluations. This necessitated the adjustments of the composition of standard sets according to the results of previous experiments and the goals of future experiments. The lesson learned from this challenge calls for more active usage of additional metabolites to expand targeted metabolite mixes to better reflect chemical diversity of biospecimen type of interest. For example, the enrollment of chemically pure standard compounds can be guided by the identity (even partial) of those unknown metabolites, which demonstrate a particular behavior during the development (for example, metabolites with a split, or with particular RT in ZIC-HILIC, etc.). Such an approach will be more beneficial on instruments with higher mass resolution (≥500 K), which speed-up/facilitate elucidations (confirmations) of metabolites' structures and identities.

Despite the limited metabolite coverage, targeted analysis remains the primary approach for exploratory investigations and the development of sample preparation and LC-MS analysis methods, including samples from novel sources. This targeted approach was applied to successfully develop a cortisol/cortisone quantitation method in oral fluid. The major challenge in this development was the lack of knowledge of small but numerous essential details (for example, interference from glycoproteins, stability, matrix effects, etc.) due to the lack of published information at the time. The developed MRM method provides a specific, quick assessment of the

steroid concentrations in oral fluid within a day from the time of sample acquisition. The method quantifies cortisol and cortisone simultaneously, which establishes the tool to control cortisol-cortisone conversion. This makes the method more reliable for the assessment of free cortisol levels in the blood than a single cortisol immunoassay. The technique uses a cheap and straightforward methanol extraction with broad selectivity, which in the future can permit facile incorporation of additional metabolites into the current assay. The ease of oral fluid sample collection outside of the laboratory provides possibilities to quantitate steroids in this biofluid alternatively when plasma collection is impossible (for example, physical exercise, real-time monitoring of stress, etc.). Finally, this study on cortisol demonstrated poor selectivity of cortisol immunoassays due to the inability to create a more selective ELISA sandwich approach for analytes smaller than ~1000 Da. Although newer immunoassays have reduced antibody cross-reactivity to cortisone, our results show that important interferences still remain unaccounted for and merit further investigation. Alternately, switch from small immunoassays to targeted LC-MS/MS metabolomics assays such as ours can provide better quantitation due to the improvements in selectivity. Combining sSPE with these targeted LC-MS/MS provides further opportunities for additional sample clean-up and enrichment to further push LLOQs in such methods. Finally, one of the key ideas that drove the current research was as the goal to create an integrated targeted/global workflow with increased metabolome coverage. Such an approach allows selected target and unknown metabolites to be measured simultaneously in an absolute and relative manner, respectively, in the same sample extract. A prerequisite for this is the detailed knowledge of properties and restrictions of sample preparation and analysis methods, which allow safe and controlled adjustments of extract compositions unavailable for solvent precipitations. Moreover, the satisfactory analytical performance of the two developed methods promises a quick and feasible integration of both in a combined, targeted/global sSPE method for the analysis of plasma and oral fluid. The use of the same method for the analysis of both biofluids saves time and labor on method development. In addition, simultaneous analysis of more than one biofluid may benefit from the direct comparison of the data between biological fluids and the expansion of metabolite coverage of such combined study.

Considering the knowledge and methods, which were acquired during the execution of the thesis, the metabolomics study of BPD descried in **Section 1.1.1** would be executed via combined targeted/global approach and with using oral fluid as a supplementary sample source. In such future

study, I would execute a targeted approach towards low abundant members of HPT and HPA axes. Immunoaffinity-based methods would be used for absolute quantitation of peptide hormones (ACTH and TSH). The absolute quantification of other members of HPT and HPA axes requires a development of an immunoaffinity enrichment or chemical derivatization sample preparation method with a subsequent analysis via LC-MS/MS. The untargeted analysis in this study would be executed as an on-line 2-D-sSPE-LC-MS or with off-line sSPE sample preparation fully automated with the aid of a robotic liquid handler. Considering the observation of high number of low abundant metabolites in saliva, I would prioritize the comparison of metabolome coverage of HPT and HPA metabolites in oral fluid to blood, using sSPE and/or chemical derivatization analysis.

By outlining the strategic vision on the further development/execution of metabolomics analysis of BPD, I am intentionally demonstrating that the knowledge and findings accumulated during the execution of my thesis, provide an important groundwork for comprehensive and strategic planning of prospective workflows for metabolomics analysis of samples such as plasma and oral fluid.

REFERENCES

(1) Ceglarek, U.; Leichtle, A.; Brügel, M.; Kortz, L.; Brauer, R.; Bresler, K.; Thiery, J.; Fiedler, G. M. Challenges and Developments in Tandem Mass Spectrometry Based Clinical Metabolomics. *Mol. Cell. Endocrinol.* **2009**, *301* (1–2), 266–271. https://doi.org/10.1016/j.mce.2008.10.013.

(2) Theodoridis, G.; Gika, H. G.; Wilson, I. D. LC-MS-Based Methodology for Global Metabolite Profiling in Metabonomics/Metabolomics. *TrAC - Trends Anal. Chem.* **2008**, *27* (3), 251–260. https://doi.org/10.1016/j.trac.2008.01.008.

(3) Cui, L.; Lu, H.; Lee, Y. H. Challenges and Emergent Solutions for LC-MS/MS Based Untargeted Metabolomics in Diseases. *Mass Spectrom. Rev.* **2018**, No. May 2017, 1–21. https://doi.org/10.1002/mas.21562.

(4) Vuckovic, D. Current Trends and Challenges in Sample Preparation for Global Metabolomics Using Liquid Chromatography – Mass Spectrometry. *Anal. Bioanal. Chem.* **2012**, *403*, 1523–1548. https://doi.org/10.1007/s00216-012-6039-y.

(5) Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; Mclean, J. A. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *J. Am. Soc. Mass Spectrom* **2016**, 1897–1905. https://doi.org/10.1007/s13361-016-1469-y.

(6) Barri, T.; Dragsted, L. O. UPLC-ESI-QTOF/MS and Multivariate Data Analysis for Blood Plasma and Serum Metabolomics: Effect of Experimental Artefacts and Anticoagulant. *Anal. Chim. Acta* **2013**, *768* (1), 118–128. https://doi.org/10.1016/j.aca.2013.01.015.

(7) Roberts, L.; Souza, A.; Gerszten, R.; Clish, C. Targeted Metabolomics. *Curr. Protoc Mol. Biol.* **2012**, 1–34. https://doi.org/10.1002/0471142727.mb3002s98.Targeted.

(8) Lee, H.; Choi, J. M.; Cho, J.-Y.; Kim, T.-E.; Lee, H. J.; Jung, B. H. *Regulation of Endogenic Metabolites by Rosuvastatin in Hyperlipidemia Patients: An Integration of Metabolomics and Lipidomics*; Elsevier Ireland Ltd, 2018. https://doi.org/10.1016/j.chemphyslip.2018.05.005.

(9) Khan, I.; Nam, M.; Kwon, M.; Seo, S. S.; Jung, S.; Han, J. S.; Hwang, G. S.; Kim, M. K. Lc/Ms-Based Polar Metabolite Profiling Identified Unique Biomarker Signatures for

Cervical Cancer and Cervical Intraepithelial Neoplasia Using Global and Targeted Metabolomics. *Cancers (Basel).* **2019**, *11* (4). https://doi.org/10.3390/cancers11040511.

(10) Zhao, L.; Zhao, A.; Chen, T.; Chen, W.; Liu, J.; Wei, R.; Su, J.; Tang, X.; Liu, K.; Zhang, R.; Xie, G.; Panee, J.; Qiu, M.; Jia, W. Global and Targeted Metabolomics Evidence of the Protective Effect of Chinese Patent Medicine Jinkui Shenqi Pill on Adrenal Insufficiency after Acute Glucocorticoid Withdrawal in Rats. *J. Proteome Res.* **2016**, *15* (7), 2327–2336. https://doi.org/10.1021/acs.jproteome.6b00409.

(11) Zimmerman, M.; Morgan, T. A. The Relationship between Borderline Personality Disorder and Bipolar Disorder. *Dialogues Clin. Neurosci.* **2013**, *15* (2), 155–169.

(12) Nesse, R. M.; Stein, D. J. Towards a Genuinely Medical Model for Psychiatric Nosology. *BMC Med.* **2012**, *10* (1), 5. https://doi.org/10.1186/1741-7015-10-5.

(13) Duval, F.; Mokrani, M.; Schulz, P.; Champeval, C.; Macher, J. Neuroendocrine Predictors of the Evolution of Depression. *Dialogues Clin. Neurosci.* **2005**, *7* (3), 273–282.

(14) Chakrabarti, S. Thyroid Functions and Bipolar Affective Disorder. *J. Thyroid Res.* **2011**, 1–13. https://doi.org/10.4061/2011/306367.

(15) Quinones, M.; Kaddurah-Daouk, R. Neurobiology of Disease Metabolomics Tools for Identifying Biomarkers for Neuropsychiatric Diseases. *Neurobiol. Dis.* **2009**, *35* (2), 165–176. https://doi.org/10.1016/j.nbd.2009.02.019.

(16) Min, W.; Liu, C.; Yang, Y.; Sun, X.; Zhang, B.; Xu, L.; Sun, X. Progress in Neuro-Psychopharmacology & Biological Psychiatry Alterations in Hypothalamic – Pituitary – Adrenal / Thyroid (HPA / HPT) Axes Correlated with the Clinical Manifestations of Depression. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2012**, *39*, 206–211. https://doi.org/10.1016/j.pnpbp.2012.06.017.

(17) Feng, G. H.; Kang, C. Y.; Yuan, J.; Zhang, Y.; Wei, Y. J.; Xu, L.; Zhou, F.; Fan, X.; Yang, J. Z. Neuroendocrine Abnormalities Associated with Untreated First Episode Patients with Major Depressive Disorder and Bipolar Disorder. *Psychoneuroendocrinology* **2019**, *107* (May), 119–123. https://doi.org/10.1016/j.psyneuen.2019.05.013.

(18) Rico, E.; González, O.; Blanco, M. E.; Alonso, R. M. Evaluation of Human Plasma Sample Preparation Protocols for Untargeted Metabolic Profiles Analyzed by UHPLC-ESI-TOF-

MS. *Anal. Bioanal. Chem.* **2014**, *406* (29), 7641–7652. https://doi.org/10.1007/s00216-014-8212-y.

(19) Sussulini, A.; Prando, A.; Da, M.; Rj, P.; Tasic, L.; Ce, B.; Ma, A. Metabolic Profiling of Human Blood Serum from Treated Patients with Bipolar Disorder Employing 1H NMR Spectroscopy and Chemometrics . *Anal Chem* **2009**, *81*, 9755–9763. https://doi.org/10.1021/ac901502j.

(20) Burghardt, K.; Evans, S.; Wiese, K.; Ellingrod, V. An Untargeted Metabolomics Analysis of Antipsychotic Use in Bipolar Disorder. *Clin. Transl. Sci.* **2015**, *8* (5), 432–440. https://doi.org/10.1111/cts.12324.

(21) Villaseñor, A.; Ramamoorthy, A.; Silva Dos Santos, M.; Lorenzo, M. P.; Laje, G.; Zarate, C.; Barbas, C.; Wainer, I. W. A Pilot Study of Plasma Metabolomic Patterns from Patients Treated with Ketamine for Bipolar Depression: Evidence for a Response-Related Difference in Mitochondrial Networks. *Br. J. Pharmacol.* **2014**, *171* (8), 2230–2242. https://doi.org/10.1111/bph.12494.

(22) Setoyama, D.; Kato, T. A.; Hashimoto, R.; Kunugi, H.; Hattori, K.; Hayakawa, K.; Sato-Kasai, M.; Shimokawa, N.; Kaneko, S.; Yoshida, S.; Goto, Y. I.; Yasuda, Y.; Yamamori, H.; Ohgidani, M.; Sagata, N.; Miura, D.; Kang, D.; Kanba, S. Plasma Metabolites Predict Severity of Depression and Suicidal Ideation in Psychiatric Patients-a Multicenter Pilot Analysis. *PLoS One* **2016**, *11* (12), 1–16. https://doi.org/10.1371/journal.pone.0165267.

(23) Campollo, O MacGillivray, B McIntyre, N. Association of Plasma Ammonia and GABA Levels and the Degree of Hepatic Encephalopathy. *Rev. Invest. Clin.* **1992**, *44* (4), 483–490.

(24) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D608–D617. https://doi.org/10.1093/nar/gkx1089.

(25) Sethi, S.; Brietzke, E. Omics-Based Biomarkers: Application of Metabolomics in Neuropsychiatric Disorders. *Int. J. Neuropsychopharmacol.* **2015**, *19* (3), pyv096.

https://doi.org/10.1093/ijnp/pyv096.

(26) Kaddurah-Daouk, R.; Krishnan, K. R. R. Metabolomics : A Global Biochemical Approach to the Study of Central Nervous System Diseases. *Neuropsychopharmacol. Rev.* **2009**, *34*, 173–186. https://doi.org/10.1038/npp.2008.174.

(27) Lindahl, A.; Forshed, J.; Nordström, A. Overlap in Serum Metabolic Profiles between Non-Related Diseases: Implications for LC-MS Metabolomics Biomarker Discovery. *Biochem. Biophys. Res. Commun.* **2016**, *478* (3), 1472–1477. https://doi.org/10.1016/j.bbrc.2016.08.155.

(28) Vuckovic, D. Improving Metabolome Coverage and Data Quality: Advancing Metabolomics and Lipidomics for Biomarker Discovery. *Chem. Commun.* **2018**, *54* (50), 6728–6749. https://doi.org/10.1039/C8CC02592D.

(29) Barnes, V. M.; Kennedy, A. D.; Panagakos, F.; Devizio, W.; Trivedi, H. M.; J??nsson, T.; Guo, L.; Cervi, S.; Scannapieco, F. A. Global Metabolomic Analysis of Human Saliva and Plasma from Healthy and Diabetic Subjects, with and without Periodontal Disease. *PLoS One* **2014**, *9* (8), 1–8. https://doi.org/10.1371/journal.pone.0105181.

(30) Wu, Y.; Streijger, F.; Wang, Y.; Lin, G.; Christie, S.; Mac-Thiong, J. M.; Parent, S.; Bailey, C. S.; Paquette, S.; Boyd, M. C.; Ailon, T.; Street, J.; Fisher, C. G.; Dvorak, M. F.; Kwon, B. K.; Li, L. Parallel Metabolomic Profiling of Cerebrospinal Fluid and Serum for Identifying Biomarkers of Injury Severity after Acute Human Spinal Cord Injury. *Sci. Rep.* **2016**, *6* (July), 1–14. https://doi.org/10.1038/srep38718.

(31) Mena-Bravo, A.; Luque de Castro, M. D. Sweat: A Sample with Limited Present Applications and Promising Future in Metabolomics. *J. Pharm. Biomed. Anal.* **2014**, *90*, 139–147. https://doi.org/10.1016/J.JPBA.2013.10.048.

(32) Vorrink, S. U.; Ullah, S.; Schmidt, S.; Nandania, J.; Velagapudi, V.; Beck, O.; Ingelman-Sundberg, M.; Lauschke, V. M. Endogenous and Xenobiotic Metabolic Stability of Primary Human Hepatocytes in Long-Term 3D Spheroid Cultures Revealed by a Combination of Targeted and Untargeted Metabolomics. *FASEB J.* **2017**, *31* (6), 2696–2708. https://doi.org/10.1096/fj.201601375R.

(33) Vorkas, P. A.; Isaac, G.; Anwar, M. A.; Davies, A. H.; Want, E. J.; Nicholson, J. K.; Holmes,

E. Untargeted UPLC-MS Profiling Pipeline to Expand Tissue Metabolome Coverage: Application to Cardiovascular Disease. *Anal. Chem.* **2015**, *87* (8), 4184–4193. https://doi.org/10.1021/ac503775m.

(34) Fernández-Peralbo, M. A.; Priego-Capote, F.; Luque de Castro, M. D.; Casado-Adam, A.; Arjona-Sánchez, A.; Muñoz-Casares, F. C. LC–MS/MS Quantitative Analysis of Paclitaxel and Its Major Metabolites in Serum, Plasma and Tissue from Women with Ovarian Cancer after Intraperitoneal Chemotherapy. *J. Pharm. Biomed. Anal.* **2014**, *91*, 131–137. https://doi.org/10.1016/J.JPBA.2013.12.028.

(35) Peralbo-Molina, A.; Calderón-Santiago, M.; Priego-Capote, F.; Jurado-Gámez, B.; Luque de Castro, M. D. Identification of Metabolomics Panels for Potential Lung Cancer Screening by Analysis of Exhaled Breath Condensate. *J. Breath Res.* **2016**, *10* (2), 026002. https://doi.org/10.1088/1752-7155/10/2/026002.

(36) Luque de Castro, M. D.; Priego-Capote, F. The Analytical Process to Search for Metabolomics Biomarkers. *J. Pharm. Biomed. Anal.* **2018**, *147*, 341–349. https://doi.org/10.1016/j.jpba.2017.06.073.

(37) Barnes, S.; Benton, H. P.; Casazza, K.; Cooper, S. J.; Cui, X.; Du, X.; Engler, J.; Kabarowski, J. H.; Li, S.; Pathmasiri, W.; Prasain, J. K.; Renfrow, M. B.; Tiwari, H. K. Training in Metabolomics Research. I. Designing the Experiment, Collecting and Extracting Samples and Generating Metabolomics Data. *J. Mass Spectrom.* **2016**, No. November 2015, 461–475. https://doi.org/10.1002/jms.3782.

(38) Waters, W. E.; Sussman, M.; Asscher, A. W. Community Study of Urinary PH and Osmolality. *Br. J. Prev. Soc. Med.* **1967**, *21* (3), 129–132. https://doi.org/10.1136/jech.21.3.129.

(39) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. Use of a Pre-Analysis Osmolality Normalisation Method to Correct for Variable Urine Concentrations and for Improved Metabolomic Analyses. *J. Chromatogr. A* **2015**, *1431*, 103–110. https://doi.org/10.1016/j.chroma.2015.12.056.

(40) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.;

Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. The Human Urine Metabolome. *PLoS One* **2013**, *8* (9), e73076. https://doi.org/10.1371/journal.pone.0073076.

(41) Zhang, T.; Watson, D. G. A Short Review of Applications of Liquid Chromatography Mass Spectrometry Based Metabolomics Techniques to the Analysis of Human Urine. *Analyst* **2015**, *140* (9), 2907–2915. https://doi.org/10.1039/c4an02294g.

(42) Mattarucchi, E.; Guillou, C. Critical Aspects of Urine Profiling for the Selection of Potential Biomarkers Using UPLC-TOF-MS. *Biomed. Chromatogr.* **2012**, *26* (4), 512–517. https://doi.org/10.1002/bmc.1697.

(43) Ahrens, B. D.; Kucherova, Y.; Butch, A. W. Detection of Stimulants and Narcotics by Liquid Chromatography-Tandem Mass Spectrometry and Gas Chromatography-Mass Spectrometry for Sports Doping Control. In *Methods in molecular biology*; 2016; Vol. 1383, pp 247–263. https://doi.org/10.1007/978-1-4939-3252-8_26.

(44) Meyer, G. M.; Maurer, H. H.; Meyer, M. R. Multiple Stage MS in Analysis of Plasma, Serum, Urine and *in Vitro* Samples Relevant to Clinical and Forensic Toxicology. *Bioanalysis* **2016**, *8* (5), 457–481. https://doi.org/10.4155/bio.16.15.

(45) Trushina, E.; Dutta, T.; Persson, X. T.; Mielke, M. M.; Petersen, R. C. Identification of Altered Metabolic Pathways in Plasma and CSF in Mild Cognitive Impairment and Alzheimer ' s Disease Using Metabolomics. *PLoS One* **2013**, *8* (5). https://doi.org/10.1371/journal.pone.0063644.

(46) Autmizguine, J Moran, C Gonzalez, D Capparelli, E Smith, B Grant, G Benjamin, D Cohen-Wolkowiez, M Watt, K. Vancomycin Cerebrospinal Fluid Pharmacokinetics in Children with Cerebral Ventricular Shunt Infections Julie. *Pediatr. Infect. Dis. J.* **2014**, *33* (10), 280–272. https://doi.org/10.1021/nl061786n.Core-Shell.

(47) Kaddurah-daouk, R.; Yuan, P.; Boyle, S. H.; Matson, W.; Wang, Z.; Zeng, Z. B.; Zhu, H.; Dougherty, G. G.; Yao, J. K.; Chen, G.; Guitart, X.; Carlson, P. J.; Neumeister, A.; Zarate, C.; Krishnan, R. R.; Manji, H. K.; Drevets, W. Cerebrospinal Fluid Metabolome in Mood Disorders-Remission State Has a Unique Metabolic Profile. *Sci. Rep.* **2012**, 1–10. https://doi.org/10.1038/srep00667.

(48) Nunes de Paiva, M. J.; Menezes, H. C.; de Lourdes Cardeal, Z. Sampling and Analysis of

Metabolomes in Biological Fluids. *Analyst* **2014**, *139* (15), 3683–3694. https://doi.org/10.1039/c4an00583j.

(49) Figueira, J.; Jonsson, P.; Nordin Adolfsson, A.; Adolfsson, R.; Nyberg, L.; Öhman, A. NMR Analysis of the Human Saliva Metabolome Distinguishes Dementia Patients from Matched Controls. *Mol. BioSyst.* **2016**, *12* (8), 2562–2571. https://doi.org/10.1039/C6MB00233A.

(50) Lee, S.; Kwon, S.; Shin, H.; Lim, H.; Singh, R. J.; Lee, K.; Kim, Y. Simultaneous Quantitative Analysis of Salivary Cortisol and Cortisone in Korean Adults Using LC-MS / MS. *BMB Rep.* **2010**, *43* (7), 506–511.

(51) Hayes, L. D.; Sculthorpe, N.; Cunniffe, B.; Grace, F. Salivary Testosterone and Cortisol Measurement in Sports Medicine: A Narrative Review and User???S Guide for Researchers and Practitioners. *Int. J. Sports Med.* **2016**, *37*, 1007–1018. https://doi.org/10.1055/s-0042-105649.

(52) Raff, H.; Raff, J. L.; Findling, J. W. Late-Night Salivary Cortisol as a Screening Test for Cushing's Syndrome [1]. *J. Clin. Endocrinol. Metab.* **1998**, *83* (8), 2681–2686. https://doi.org/10.1210/jcem.83.8.4936.

(53) Ross, I.; Levitt, N.; Van der Walt, J.; Schatz, D.; Johannsson, G.; Haarburger, D.; Pillay, T. Salivary Cortisol Day Curves in Addison's Disease in Patients on Hydrocortisone Replacement. *Horm. Metab. Res.* **2012**, *45* (01), 62–68. https://doi.org/10.1055/s-0032-1321855.

(54) Elmongy, H.; Abdel-Rehim, M. Saliva as an Alternative Specimen to Plasma for Drug Bioanalysis. A Review. *TrAC - Trends Anal. Chem.* **2016**, *83*, 70–79. https://doi.org/10.1016/j.trac.2016.07.010.

(55) Hildes, J. A.; Ferguson, M. H. The Concentration of Electrolytes in Normal Human Saliva. *Biochem. Cell Biol.* **1955**, *33* (2), 217–225. https://doi.org/10.1139/o55-031.

(56) Agha-Hosseini, F.; Dizgah, I. M.; Amirkhani, S. The Composition of Unstimulated Whole Saliva of Healthy Dental Students. *J. Contemp. Dent. Pract.* **2006**, *7* (2), 104–111.

(57) Dame, Z. T.; Aziat, F.; Mandal, R.; Krishnamurthy, R.; Bouatra, S.; Borzouie, S.; Guo, A. C.; Sajed, T.; Deng, L.; Lin, H.; Liu, P.; Dong, E.; Wishart, D. S. The Human Saliva Metabolome. *Metabolomics* **2015**, *11* (6), 1864–1883. https://doi.org/10.1007/s11306-015-

0840-5.

(58) Lu, Y.; Huang, J.; Lee, I.; Weng, R.; Lin, M.; Yang, J.; Lin, C. OPEN A Portable System to Monitor Saliva Conductivity for Dehydration Diagnosis and Kidney Healthcare. *Sci. Rep.* **2019**, 1–9. https://doi.org/10.1038/s41598-019-51463-8.

(59) Cummings, O. T.; Morris, A. A.; Enders, J. R.; Mcintire, G. L. Normalizing Oral Fluid Hydrocodone Data Using Calculated Blood Volume. *J. Anal. Toxicol.* · **2016**, *40* (July), 486–491. https://doi.org/10.1093/jat/bkw057.

(60) Jové, M.; Collado, R.; Luís Quiles, J.; Ramírez- Tortosa, M.-C.; Sol, J.; Ruiz-Sanjuan, M.; Fernandez, M.; de la Torre Cabrera, C.; Ramírez-Tortosa, C.; Granados-Principal, S.; Sánchez-Rovira, P.; Pamplona, R.; Jové, M.; Collado, R.; Luís Quiles, J.; Ramírez-Tortosa, M.-C.; Sol, J.; Ruiz-Sanjuan, M.; Fernandez, M.; de la Torre Cabrera, C.; Ramírez-Tortosa, C.; Granados-Principal, S.; Sánchez-Rovira, P.; Pamplona, R. A Plasma Metabolomic Signature Discloses Human Breast Cancer. *Oncotarget* **2017**, *8* (12), 19522–19533. https://doi.org/10.18632/oncotarget.14521.

(61) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T. L.; Steven, R.; Bamforth, F.; Greiner, R.; Mcmanus, B.; Newman, J. W.; Goodfriend, T.; Wishart, D. S. The Human Serum Metabolome. *PLoS One* **2011**, *6* (2), 1–23. https://doi.org/10.1371/journal.pone.0016957.

(62) Yuana, Y.; Sturk, A.; Nieuwland, R. Blood Reviews Extracellular Vesicles in Physiological and Pathological Conditions. *Blood Rev.* **2013**, *27* (1), 31–39. https://doi.org/10.1016/j.blre.2012.12.002.

(63) Porter, R. S., Kaplan, J. L., M. S. & D. *The Merck Manual of Diagnosis and Therapy*, 19th ed.; Whitehouse Station, N.J: Merck Sharp & Dohme Corp, 2011.

(64) Page, M. J.; Di Cera, E. Role of Na+ and K+ in Enzyme Function. *Physiol. Rev.* **2006**, *86* (4), 1049–1092. https://doi.org/10.1152/physrev.00008.2006.

(65) Brod, J. Volume Homeostasis, Renal Function and Hypertension. *Ulster Med. J.* **1985**, *54* (SUPPL. AUG.), 20–33.

(66) Jobard, E.; Trédan, O.; Postoly, D.; André, F.; Martin, A.-L.; Elena-Herrmann, B.; Boyault,

S. A Systematic Evaluation of Blood Serum and Plasma Pre-Analytics for Metabolomics Cohort Studies. *Int. J. Mol. Sci.* **2016**, *17* (12), 2035. https://doi.org/10.3390/ijms17122035.

(67) Yin, P.; Lehmann, R.; Xu, G. Effects of Pre-Analytical Processes on Blood Samples Used in Metabolomics Studies. *Anal. Bioanal. Chem.* **2015**, No. 407, 4879–4892. https://doi.org/10.1007/s00216-015-8565-x.

(68) Chetwynd, A. J.; Dunn, W. B.; Rodriguez-blanco, G. Chapter 2 Collection and Preparation of Clinical Samples for Metabolomics. In *Metabolomics: From Fundamentals to Clinical Applications*; 2017; Vol. 965, pp 19–44. https://doi.org/10.1007/978-3-319-47656-8.

(69) Malsagova, K.; Kopylov, A.; Stepanov, A.; Butkova, T.; Izotov, A.; Kaysheva, A. Dried Blood Spot in Laboratory: Directions and Prospects. *Diagnostics* **2020**, *10* (4). https://doi.org/10.3390/diagnostics10040248.

(70) Khurshid, Z.; Zohaib, S.; Najeeb, S.; Zafar, M. S.; Slowey, P. D.; Almas, K. Human Saliva Collection Devices for Proteomics: An Update. *Int. J. Mol. Sci.* **2016**, *17* (6). https://doi.org/10.3390/ijms17060846.

(71) Cai, H.-L.; Li, H.-D.; Yan, X.-Z.; Sun, B.; Zhang, Q.; Yan, M.; Zhang, W.-Y.; Jiang, P.; Zhu, R.-H.; Liu, Y.-P.; Fang, P.-F.; Xu, P.; Yuan, H.-Y.; Zhang, X.-H.; Hu, L.; Yang, W.; Ye, H.-S. Metabolomic Analysis of Biochemical Changes in the Plasma and Urine of First-Episode Neuroleptic-Naïve Schizophrenia Patients after Treatment with Risperidone. *J. Proteome Res.* **2012**, *11* (8), 4338–4350. https://doi.org/10.1021/pr300459d.

(72) Yan, S.; Wei, B.; Lin, Z.; Yang, Y.; Zhou, Z.; Zhang, W. A Metabonomic Approach to the Diagnosis of Oral Squamous Cell Carcinoma , Oral Lichen Planus and Oral Leukoplakia. *Oral Oncol.* **2008**, 477–483. https://doi.org/10.1016/j.oraloncology.2007.06.007.

(73) Weiss, N.; Barbier, P.; Hilaire, S.; Colsch, B.; Isnard, F. F.; Attala, S.; Schaefer, A.; Amador, M.; Rudler, M.; Lamari, F.; Sedel, F.; Thabut, D.; Junot, C. Cerebrospinal Fluid Metabolomics Highlights Dysregulation of Energy Metabolism in Overt Hepatic Encephalopathy. *J. Hepatol.* **2016**, *65* (6), 1120–1130. https://doi.org/10.1016/j.jhep.2016.07.046.

(74) Satomi, Y.; Hirayama, M.; Kobayashi, H. One-Step Lipid Extraction for Plasma Lipidomics Analysis by Liquid Chromatography Mass Spectrometry. *J. Chromatogr. B* **2017**, *1063*

(December 2016), 93–100. https://doi.org/10.1016/j.jchromb.2017.08.020.

(75) Liu, R.; Chou, J.; Hou, S.; Liu, X.; Yu, J.; Zhao, X.; Li, Y.; Liu, L.; Sun, C. Evaluation of Two-Step Liquid-Liquid Extraction Protocol for Untargeted Metabolic Profiling of Serum Samples to Achieve Broader Metabolome Coverage by UPLC-Q-TOF-MS. *Anal. Chim. Acta* **2018**, *1035*, 96–107. https://doi.org/10.1016/j.aca.2018.07.034.

(76) Cajka, T.; Fiehn, O. Towards Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Anal. Chem.* **2016**, *88* (1), 524–545. https://doi.org/10.1021/acs.analchem.5b04491.

(77) Aretz, I.; Meierhofer, D. Advantages and Pitfalls of Mass Spectrometry Based Metabolome Profiling in Systems Biology. *Int. J. Mol. Sci.* **2016**, *17* (5). https://doi.org/10.3390/ijms17050632.

(78) Van De Steene, J. C.; Lambert, W. E. Comparison of Matrix Effects in HPLC-MS/MS and UPLC-MS/MS Analysis of Nine Basic Pharmaceuticals in Surface Waters. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 713–718. https://doi.org/10.1016/j.jasms.2008.01.013.

(79) Wilson, I. D.; Nicholson, J. K.; Castro-Perez, J.; Granger, J. H.; Johnson, K. A.; Smith, B. W.; Plumb, R. S. High Resolution "Ultra Performance" Liquid Chromatography Coupled to Oa-TOF Mass Spectrometry as a Tool for Differential Metabolic Pathway Profiling in Functional Genomic Studies. *J. Proteome Res.* **2005**, *4* (2), 591–598. https://doi.org/10.1021/pr049769r.

(80) Greco, G.; Letzel, T. Main Interactions and Influences of the Chromatographic Parameters in HILIC Separations. *J. Chromatogr. Sci.* **2013**, *51* (7), 684–693. https://doi.org/10.1093/chromsci/bmt015.

(81) Contrepois, K.; Jiang, L.; Snyder, M. Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)–Mass Spectrometry. *Mol. Cell. Proteomics* **2015**, *14* (6), 1684–1695. https://doi.org/10.1074/mcp.M114.046508.

(82) Manallack, D. T.; Dennis, M. L.; Kelly, M. R.; Prankerd, R. J.; Yuriev, E.; Chalmers, D. K. The Acid/Base Profile of the Human Metabolome and Natural Products. *Mol. Inform.* **2013**, *32* (5–6), 505–515. https://doi.org/10.1002/minf.201200167.

(83) Sonnenberg, R. A.; Naz, S.; Cougnaud, L.; Vuckovic, D. Comparison of Underivatized Silica and Zwitterionic Sulfobetaine Hydrophilic Interaction Liquid Chromatography Stationary Phases for Global Metabolomics of Human Plasma. *J. Chromatogr. A* **2019**, *1608*, 460419. https://doi.org/10.1016/j.chroma.2019.460419.

(84) Wang, J.; Aubry, A.; Bolgar, M. S.; Gu, H.; Olah, T. V.; Arnold, M.; Jemal, M. Effect of Mobile Phase PH, Aqueous-Organic Ratio, and Buffer Concentration on Electrospray Ionization Tandem Mass Spectrometric Fragmentation Patterns: Implications in Liquid Chromatography/Tandem Mass Spectrometric Bioanalysis. *Rapid Commun. Mass Spectrom.* **2010**, *24* (22), 3221–3229. https://doi.org/10.1002/rcm.4748.

(85) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, *2012*, 1–40. https://doi.org/10.1155/2012/282574.

(86) Loris Toninandel, R. S. Matrix Effect, Signal Suppression and Enhancement in LC–ESI–MS. *Adv. LC-MS Instrum.* **2007**, *72*, 193.

(87) Liigand, J.; Laaniste, A.; Kruve, A. PH Effects on Electrospray Ionization Efficiency. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (3), 461–469. https://doi.org/10.1007/s13361-016-1563-1.

(88) Calderón-Santiago, M.; Priego-Capote, F.; Luque De Castro, M. D. Enhanced Detection and Identification in Metabolomics by Use of LC-MS/MS Untargeted Analysis in Combination with Gas-Phase Fractionation. *Anal. Chem.* **2014**, *86* (15), 7558–7565. https://doi.org/10.1021/ac501353n.

(89) Yan, Z.; Yan, R. Improved Data-Dependent Acquisition for Untargeted Metabolomics Using Gas-Phase Fractionation with Staggered Mass Range. *Anal. Chem.* **2015**, *87* (5), 2861–2868. https://doi.org/10.1021/ac504325x.

(90) Fenaille, F.; Barbier Saint-Hilaire, P.; Rousseau, K.; Junot, C. Data Acquisition Workflows in Liquid Chromatography Coupled to High Resolution Mass Spectrometry-Based Metabolomics: Where Do We Stand? *J. Chromatogr. A* **2017**, *1526* (March), 1–12. https://doi.org/10.1016/j.chroma.2017.10.043.

(91) Hilaire, P. B. Saint; Rousseau, K.; Seyer, A.; Dechaumet, S.; Damont, A.; Junot, C.; Fenaille, F. Comparative Evaluation of Data Dependent and Data Independent Acquisition

Workflows Implemented on an Orbitrap Fusion for Untargeted Metabolomics. *Metabolites* **2020**, *10* (4). https://doi.org/10.3390/metabo10040158.

(92) Spicer, R.; Salek, R. M.; Moreno, P.; Cañueto, D.; Steinbeck, C. Navigating Freely-Available Software Tools for Metabolomics Analysis. *Metabolomics* **2017**, *13* (9), 1–16. https://doi.org/10.1007/s11306-017-1242-7.

(93) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. MetaboAnalyst 4.0: Towards More Transparent and Integrative Metabolomics Analysis. *Nucleic Acids Res.* **2018**, *46* (W1), W486–W494. https://doi.org/10.1093/nar/gky310.

(94) Plewa, S.; Horała, A.; Dereziński, P.; Nowak-Markwitz, E.; Matysiak, J.; Kokot, Z. J. Wide Spectrum Targeted Metabolomics Identifies Potential Ovarian Cancer Biomarkers. *Life Sci.* **2019**, *222* (December 2018), 235–244. https://doi.org/10.1016/j.lfs.2019.03.004.

(95) Liu, C.; Gu, C.; Huang, W.; Sheng, X.; Du, J.; Li, Y. Targeted UPLC-MS/MS High-Throughput Metabolomics Approach to Assess the Purine and Pyrimidine Metabolism. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2019**, *1113* (November 2018), 98–106. https://doi.org/10.1016/j.jchromb.2019.03.008.

(96) David, A.; Abdul-Sada, A.; Lange, A.; Tyler, C. R.; Hill, E. M. A New Approach for Plasma (Xeno)Metabolomics Based on Solid-Phase Extraction and Nanoflow Liquid Chromatography-Nanoelectrospray Ionisation Mass Spectrometry. *J. Chromatogr. A* **2014**, *1365*, 72–85. https://doi.org/10.1016/j.chroma.2014.09.001.

(97) Ribbenstedt, A.; Ziarrusta, H.; Benskin, J. P. Development, Characterization and Comparisons of Targeted and Non-Targeted Metabolomics Methods. *PLoS One* **2018**, *13* (11), 1–18. https://doi.org/10.1371/journal.pone.0207082.

(98) Gruber, F.; Bicker, W.; Oskolkova, O. V.; Tschachler, E.; Bochkov, V. N. A Simplified Procedure for Semi-Targeted Lipidomic Analysis of Oxidized Phosphatidylcholines Induced by UVA Irradiation. *J. Lipid Res.* **2012**, *53* (6), 1232–1242. https://doi.org/10.1194/jlr.D025270.

(99) Zhou, J.; Yin, Y. Strategies for Large-Scale Targeted Metabolomics Quantification by Liquid Chromatography-Mass Spectrometry. *Analyst* **2016**, *141* (23), 6362–6373. https://doi.org/10.1039/C6AN01753C.

(100) Dudzik, D.; Barbas-Bernardos, C.; García, A.; Barbas, C. Quality Assurance Procedures for Mass Spectrometry Untargeted Metabolomics. a Review. *J. Pharm. Biomed. Anal.* **2018**, *147*, 149–173. https://doi.org/10.1016/j.jpba.2017.07.044.

(101) Gika, H. G.; Zisi, C.; Theodoridis, G.; Wilson, I. D. Protocol for Quality Control in Metabolic Profiling of Biological Fluids by U(H)PLC-MS. *J. Chromatogr. B* **2016**, *1008*, 15–25. https://doi.org/http://dx.doi.org/10.1016/j.jchromb.2015.10.045.

(102) Food and Drug Administration (FDA). Guidance for Industry: Bioanalytical Method Validation. *Bioanal. Method Validation. Guid. Ind.* **2018**, No. May, 1–22.

(103) Kirwan, J. A.; Broadhurst, D. I.; Davidson, R. L.; Viant, M. R. Characterising and Correcting Batch Variation in an Automated Direct Infusion Mass Spectrometry (DIMS) Metabolomics Workflow. *Anal. Bioanal. Chem.* **2013**, *405* (15), 5147–5157. https://doi.org/10.1007/s00216-013-6856-7.

(104) Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R. A Tutorial Review: Metabolomics and Partial Least Squares-Discriminant Analysis - a Marriage of Convenience or a Shotgun Wedding. *Anal. Chim. Acta* **2015**, *879*, 10–23. https://doi.org/10.1016/j.aca.2015.02.012.

(105) Engskog, M. K. R.; Haglöf, J.; Arvidsson, T.; Pettersson, C. LC–MS Based Global Metabolite Profiling: The Necessity of High Data Quality. *Metabolomics* **2016**, *12* (7). https://doi.org/10.1007/s11306-016-1058-x.

(106) Barupal, D. K.; Fan, S.; Fiehn, O. Integrating Bioinformatics Approaches for a Comprehensive Interpretation of Metabolomics Datasets. *Curr. Opin. Biotechnol.* **2018**, *54*, 1–9. https://doi.org/10.1016/j.copbio.2018.01.010.

(107) Xi, B.; Haiwei, Gu Baniasadi, Hamid Raftery, D. Statistical Analysis and Modeling of Mass Spectrometry-Based Metabolomics Data. *Methods Mol Biol* **2014**, *1198*, 333–353. https://doi.org/10.1002/9781119977438.ch12.

(108) 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.

(109) Sumner, L. W.; Samuel, T.; Noble, R.; Gmbh, S. D.; Barrett, D.; Beale, M. H.; Hardy, N.

Proposed Minimum Reporting Standards for Chemical Analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI) Lloyd. *Metabolomics* **2007**, *3* (3), 211–221. https://doi.org/10.1007/s11306-007-0082-2.

(110) Jassal, B.; Viteri, G.; Sidiropoulos, K.; Haw, R.; Shamovsky, V.; D'Eustachio, P.; Milacic, M.; Hermjakob, H.; Korninger, F.; Weiser, J.; Roca, C. D.; Stein, L.; Sevilla, C.; Jupe, S.; Garapati, P.; Matthews, L.; Fabregat, A.; Varusai, T.; Rothfels, K.; Wu, G.; May, B.; Shorser, S.; Gillespie, M. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **2017**, *46* (D1), D649–D655. https://doi.org/10.1093/nar/gkx1132.

(111) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44* (D1), D457–D462. https://doi.org/10.1093/nar/gkv1070.

(112) Marques, A. P.; Serralheiro, M. L.; Ferreira, A. E. N.; Freire, A. P.; Cordeiro, C.; Silva, M. S. Metabolomics for Undergraduates: Identification and Pathway Assignment of Mitochondrial Metabolites. *Biochem. Mol. Biol. Educ.* **2016**, *44* (1), 38–54. https://doi.org/10.1002/bmb.20919.

(113) Tsuyuzaki, K.; Morota, G.; Ishii, M.; Nakazato, T.; Miyazaki, S.; Nikaido, I. MeSH ORA Framework: R/Bioconductor Packages to Support MeSH over-Representation Analysis. *BMC Bioinformatics* **2015**, *16* (1). https://doi.org/10.1186/s12859-015-0453-z.

(114) Grosse, I.; Taruttis, F.; Dornfeldt, S.; Muthukrishnan, V.; Steinbeck, C.; Neumann, S.; Harsha, B.; Tudose, I.; Dekker, A.; Moreno, P.; Hastings, J.; Beisken, S. BiNChE: A Web Tool and Library for Chemical Enrichment Analysis Based on the ChEBI Ontology. *BMC Bioinformatics* **2015**, *16* (1), 1–7. https://doi.org/10.1186/s12859-015-0486-3.

(115) Consortium, T. U. UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515. https://doi.org/10.1093/nar/gky1049.

(116) Nikolac, N. Lipemia : Causes , Interference Mechanisms , Detection and Management. *Biochem. Medica* **2014**, *24* (1), 57–67.

(117) Kumar, S.; Wu, H. Recent Advances in Mass Spectrometry for the Identification of Neuro-Chemicals and Their Metabolites in Biofluids. *Curr. Neuropharmacol.* **2013**, *11*, 436–464.

(118) Dallmann, R.; Viola, A. U.; Tarokh, L.; Cajochen, C.; Brown, S. A. The Human Circadian

Metabolome. *PNAS* **2012**, *109* (7), 1–5. https://doi.org/10.1073/pnas.1114410109.

(119) Eckel-Mahan, K.; Sassone-Corsi, P. Metabolism and the Circadian Clock Converge. *Physiol. Rev.* **2013**, *93*, 107–135. https://doi.org/10.1152/physrev.00016.2012.

(120) Blaise, B. J.; Tin, A.; Young, J. H.; Vergnaud, A.; Lewis, M.; Pearce, J. T. M.; Elliott, P.; Nicholson, J. K.; Holmes, E.; Ebbels, T. M. D. Power Analysis and Sample Size Determination in Metabolic Phenotyping. *Anal. Chem.* **2016**, *88*, 5179–5188. https://doi.org/10.1021/acs.analchem.6b00188.

(121) Nyamundanda, G.; Gormley, I. C.; Fan, Y.; Gallagher, W. M.; Brennan, L. MetSizeR: Selecting the Optimal Sample Size for Metabolomic Studies Using an Analysis Based Approach. *BMC Bioinformatics* **2013**, *14* (1), 1–8. https://doi.org/10.1186/1471-2105-14-338.

(122) Kamlage, B.; Neuber, S.; Bethan, B.; Gonz, S.; Wagner-golbs, A.; Peter, E.; Schmitz, O.; Schatz, P. Impact of Prolonged Blood Incubation and Extended Serum Storage at Room Temperature on the Human Serum Metabolome. *Metabolites* **2018**, *8* (1), 6. https://doi.org/10.3390/metabo8010006.

(123) Mei, H.; Hsieh, Y.; Nardo, C.; Xu, X.; Wang, S.; Ng, K.; Korfmacher, W. a. Investigation of Matrix Effects in Bioanalytical High-Performance Liquid Chromatography/Tandem Mass Spectrometric Assays: Application to Drug Discovery. *Rapid Commun. Mass Spectrom.* **2003**, *17* (1), 97–103. https://doi.org/10.1002/rcm.876.

(124) Choe, E.; Min, D. B. Mechanisms of Antioxidants in the Oxidation of Foods. *Compr. Rev. Food Sci. Food Saf.* **2009**, *8* (4), 345–358. https://doi.org/10.1111/j.1541-4337.2009.00085.x.

(125) Yin, P.; Peter, A.; Franken, H.; Zhao, X.; Neukamm, S. S.; Rosenbaum, L. Preanalytical Aspects and Sample Quality Assessment in Metabolomics Studies of Human Blood METHODS : *Clin. Chem.* **2013**, *59* (5), 833–845. https://doi.org/10.1373/clinchem.2012.199257.

(126) Canadian Blood Services. Visual Assessment Guide. Canadian Blood Services 2009.

(127) Shah, J. S.; Soon, P. S.; Marsh, D. J. Comparison of Methodologies to Detect Low Levels of Hemolysis in Serum for Accurate Assessment of Serum MicroRNAs. *PLoS One* **2016**,

1–12. https://doi.org/10.1371/journal.pone.0153200.

(128) Nah, H.; Lee, S.; Lee, K.; Won, J.; Ok, H.; Kim, J. Evaluation of Bilirubin Interference and Accuracy of Six Creatinine Assays Compared with Isotope Dilution – Liquid Chromatography Mass Spectrometry. *Clin. Biochem.* **2016**, *49* (3), 274–281. https://doi.org/10.1016/j.clinbiochem.2015.10.015.

(129) Garde, A. H.; Hansen, Å. M. Long-term Stability of Salivary Cortisol. *Scand. J. Clin. Lab. Invest.* **2005**, *65* (5), 433–436. https://doi.org/10.1080/00365510510025773.

(130) Watrous, J. D.; Henglin, M.; Claggett, B.; Lehmann, K. A.; Larson, M. G.; Cheng, S.; Watrous, Jeramie D. Henglin, Mir Claggett, Brian Lehmann, Kim A. Larson, Martin G. Cheng, S. Visualization, Quantification, and Alignment of Spectral Drift in Population Scale Untargeted Metabolomics Data. *Anal. Chem.* **2017**, *89* (3), 1399–1404. https://doi.org/10.1021/acs.analchem.6b04337.Experimental.

(131) Sanchez-Illana, A.; Pineiro-Ramos, J.; Sanjuan-Herraez, D.; Daniel, J.; Vento, M.; Quintas, Guillermo Kuligowski, J. Evaluation of Batch Effect Elimination Using Quality Control Replicates in LC-MS Metabolite Profiling. *Anal. Chim. Acta* **2018**, *1019*, 38–48. https://doi.org/10.1016/j.aca.2018.02.053.

(132) Dunn, W. B. I. D.; W, W. A.; Nicholls David; Broadhurst. The Importance of Experimental Design and QC Samples in Large-Scale and MS-Driven Untargeted Metabolomic Studies of Humans. *Bioanalysis* **2012**, *4* (18), 2249–2264.

(133) Matsuda, F. Technical Challenges in Mass Spectrometry-Based Metabolomics. *Mass Spectrom.* **2016**, *5* (2), S0052–S0052. https://doi.org/10.5702/massspectrometry.S0052.

(134) Burla, B.; Arita, M.; Arita, M.; Bendt, A. K.; Cazenave-Gassiot, A.; Dennis, E. A.; Ekroos, K.; Han, X.; Ikeda, K.; Liebisch, G.; Lin, M. K.; Loh, T. P.; Meikle, P. J.; Orešič, M.; Quehenberger, O.; Shevchenko, A.; Torta, F.; Wakelam, M. J. O.; Wheelock, C. E.; Wenk, M. R. MS-Based Lipidomics of Human Blood Plasma: A Community-Initiated Position Paper to Develop Accepted Guidelines. *J. Lipid Res.* **2018**, *59* (10), 2001–2017. https://doi.org/10.1194/jlr.S087163.

(135) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-mcintyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.;

Kell, D. B.; Goodacre, R.; Human, T.; Metabolome, S.; Consortium, H. Procedures for Large-Scale Metabolic Profiling of Serum and Plasma Using Gas Chromatography and Liquid Chromatography Coupled to Mass Spectrometry. **2011**. https://doi.org/10.1038/nprot.2011.335.

(136) Thakare, R.; Chhonker, Y. S.; Gautam, N.; Alamoudi, J. A.; Alnouti, Y. Quantitative Analysis of Endogenous Compounds. *J. Pharm. Biomed. Anal.* **2016**, *128*, 426–437. https://doi.org/10.1016/j.jpba.2016.06.017.

(137) Butter, J. J.; Koopmans, R. P.; Michel, M. C. A Rapid and Validated HPLC Method to Quantify Sphingosine 1-Phosphate in Human Plasma Using Solid-Phase Extraction Followed by Derivatization with Fluorescence Detection. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2005**, *824* (1–2), 65–70. https://doi.org/10.1016/j.jchromb.2005.06.040.

(138) Abdel Rahman, A. M.; Pawling, J.; Ryczko, M.; Caudy, A. A.; Dennis, J. W. Targeted Metabolomics in Cultured Cells and Tissues by Mass Spectrometry: Method Development and Validation. *Anal. Chim. Acta* **2014**, *845*, 53–61. https://doi.org/10.1016/j.aca.2014.06.012.

(139) Berg, T.; Karlsen, M.; Oiestad, A. M. L.; Johansen, J. E.; Liu, H.; Strand, D. H. Evaluation of ^{13}C- and ^{2}H-Labeled Internal Standards for the Determination of Amphetamines in Biological Samples, by Reversed-Phase Ultra-High Performance Liquid Chromatography-Tandem Mass Spectrometry. *J. Chromatogr. A* **2014**, *1344*, 83–90. https://doi.org/10.1016/j.chroma.2014.04.020.

(140) Wang, S.; Cyronak, M.; Yang, E. Does a Stable Isotopically Labeled Internal Standard Always Correct Analyte Response?. A Matrix Effect Study on a LC/MS/MS Method for the Determination of Carvedilol Enantiomers in Human Plasma. *J. Pharm. Biomed. Anal.* **2007**, *43*, 701–707. https://doi.org/10.1016/j.jpba.2006.08.010.

(141) FDA. *Guidance for Industry. Bioanalytical Method Validation*, Rev. 1.; U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), 2013.

(142) Jacob, M.; Malkawi, A.; Albast, N.; Bougha, S. Al; Lopata, A.; Dasouki, M.; Rahman, A. M. A. A Targeted Metabolomics Approach for Clinical Diagnosis of Inborn Errors of

Metabolism. *Anal. Chim. Acta* **2018**, *1025*, 141–153. https://doi.org/10.1016/j.aca.2018.03.058.

(143) De Vijlder, T.; Valkenborg, D.; Lemière, F.; Romijn, E. P.; Laukens, K.; Cuyckens, F. A Tutorial in Small Molecule Identification via Electrospray Ionization-Mass Spectrometry: The Practical Art of Structural Elucidation. *Mass Spectrom. Rev.* **2018**, *37* (5), 607–629. https://doi.org/10.1002/mas.21551.

(144) Kind, T.; Fiehn, O. Metabolomic Database Annotations via Query of Elemental Compositions: Mass Accuracy Is Insufficient Even at Less than 1 Ppm. *BMC Bioinformatics* **2006**, *7*. https://doi.org/10.1186/1471-2105-7-234.

(145) Carey, J.; Nguyen, T.; Korchak, J.; Beecher, C.; De Jong, F.; Lane, A. L. An Isotopic Ratio Outlier Analysis Approach for Global Metabolomics of Biosynthetically Talented Actinomycetes. *Metabolites* **2019**, *9* (9). https://doi.org/10.3390/metabo9090181.

(146) Witting, M.; Böcker, S. Current Status of Retention Time Prediction in Metabolite Identification. *J. Sep. Sci.* **2020**, *43* (9–10), 1746–1754. https://doi.org/10.1002/jssc.202000060.

(147) Samaraweera, M. A.; Hall, L. M.; Hill, D. W.; Grant, D. F. Evaluation of an Artificial Neural Network Retention Index Model for Chemical Structure Identification in Nontargeted Metabolomics. *Anal. Chem.* **2018**, *90* (21), 12752–12760. https://doi.org/10.1021/acs.analchem.8b03118.

(148) Guo, J.; Huan, T. Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography-Mass Spectrometry Based Untargeted Metabolomics. *Anal. Chem.* **2020**, *92* (12), 8072–8080. https://doi.org/10.1021/acs.analchem.9b05135.

(149) Qi, Y.; Volmer, D. A. Structural Analysis of Small to Medium-Sized Molecules by Mass Spectrometry after Electron-Ion Fragmentation (ExD) Reactions. *Analyst* **2016**, *141*, 794–806. https://doi.org/10.1039/c5an02171e.

(150) Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* **2018**, *8* (2). https://doi.org/10.3390/metabo8020031.

(151) Jaiswal, R.; Matei, M. F.; Glembockyte, V.; Patras, M. A.; Kuhnert, N. Hierarchical Key for the LC − MSn Identification of All Ten Regio- and Stereoisomers of Caffeoylglucose. *Agric. Food Chem.* **2014**, *62*, 9252–9265. https://doi.org/10.1021/jf501210s.

(152) Chaleckis, R.; Meister, I.; Zhang, P.; Wheelock, C. E. ScienceDirect Challenges , Progress and Promises of Metabolite Annotation for LC – MS-Based Metabolomics. *Curr. Opin. Biotechnol.* **2019**, *55*, 44–50. https://doi.org/10.1016/j.copbio.2018.07.010.

(153) Dias, D. A.; Jones, O. A. H.; Beale, D. J.; Boughton, B. A.; Benheim, D.; Kouremenos, K. A.; Wolfender, J.; Wishart, D. S. Current and Future Perspectives on the Structural Identification of Small Molecules in Biological Systems. *Metabolites* **2016**, *6* (46), 1–29. https://doi.org/10.3390/metabo6040046.

(154) Ellens, K. W.; Christian, N.; Singh, C.; Satagopam, V. P.; May, P.; Linster, C. L. Confronting the Catalytic Dark Matter Encoded by Sequenced Genomes. *Nucleinc acids Res.* **2017**, *45* (20), 11495–11514. https://doi.org/10.1093/nar/gkx937.

(155) Lambert, G.; Naredi, S.; Edén, E.; Rydenhag, B.; Friberg, P. Sympathetic Nervous Activation Following Subarachnoid Hemorrhage: Influence of Intravenous Clonidine. *Brain Res. Bull.* **2002**, *58* (1), 77–82. https://doi.org/10.1034/j.1399-6576.2002.

(156) Knolhoff, A. M.; Callahan, J. H.; Croley, T. R. Mass Accuracy and Isotopic Abundance Measurements for HR-MS Instrumentation: Capabilities for Non-Targeted Analyses. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1285–1294. https://doi.org/10.1007/s13361-014-0880-5.

(157) Boyd, E. Lipid Analysis of Blood Plasma. *J. Biol. Chem.* **1933**, *101*, 323–336.

(158) Morais, D. R. De; Eliete, J.; Visentainer, L.; Matsushita, M. Evaluation of Lipid Extraction and Fatty Acid Composition of Human Plasma. *Brazilian J. Hematol.* **2010**, *32* (6), 439–443.

(159) Miller, M.; Stone, N. J.; Ballantyne, C.; Bittner, V.; Criqui, M. H.; Ginsberg, H. N.; Goldberg, A. C.; Howard, W. J.; Jacobson, M. S.; Kris-Etherton, P. M.; Lennie, T. A.; Levi, M.; Mazzone, T.; Pennathur, S. Triglycerides and Cardiovascular Disease: A Scientific Statement from the American Heart Association. *Circulation* **2011**, *123* (20), 2292–2333. https://doi.org/10.1161/CIR.0b013e3182160726.

(160) Alvarez-sanchez, B.; Priego-capote, F.; Castro, M. D. L. De. Study of Sample Preparation

for Metabolomic Profiling of Human Saliva by Liquid Chromatography – Time of Flight / Mass Spectrometry. *J. Chromatogr. A* **2012**, *1248*, 178–181. https://doi.org/10.1016/j.chroma.2012.05.029.

(161) Zheng, J.; Dixon, R. A.; Li, L. Development of Isotope Labeling LC − MS for Human Salivary Metabolomics and Application to Pro Fi Ling Metabolome Changes Associated with Mild Cognitive Impairment. **2012**.

(162) Meulenberg ', P. M. M.; Hofman2, J. A. Differences between Concentrationsof Salivary Cortisol and Cortisone and of Free Cortisol and Cortisone in Plasma during Pregnancy and Postpartum. *Clin. Chem.* **1990**, *36136* (1), 70–75.

(163) Gadad, B. S.; Jha, M. K.; Czysz, A.; Furman, J. L.; Mayes, T. L.; Emslie, M. P.; Trivedi, M. H. Peripheral Biomarkers of Major Depression and Antidepressant Treatment Response: Current Knowledge and Future Outlooks. *J. Affect. Disord.* **2017**, No. March, 1–12. https://doi.org/10.1016/j.jad.2017.07.001.

(164) Bae, Y. J.; Gaudl, A.; Jaeger, S.; Stadelmann, S.; Hiemisch, A.; Kiess, W.; Willenberg, A.; Schaab, M.; Von Klitzing, K.; Thiery, J.; Ceglarek, U.; Döhnert, M.; Kratzsch, J.; VonKlitzing, K.; Thiery, J.; Ceglarek, U.; Döhnert, M.; Kratzsch, J. Immunoassay or LC-MS/MS for the Measurement of Salivary Cortisol in Children? *Clin. Chem. Lab. Med.* **2016**, *54* (5), 811–822. https://doi.org/10.1515/cclm-2015-0412.

(165) Mahieu, N. G.; Huang, X.; Chen, Y. J.; Patti, G. J. Credentialing Features: A Platform to Benchmark and Optimize Untargeted Metabolomic Methods. *Anal. Chem.* **2014**, *86* (19), 9583–9589. https://doi.org/10.1021/ac503092d.

(166) Naser, F Mahieu, Nathaniel Lingjue, Wnag Spalding, Jonathan Johnson, Stephen Patti, G. Two Complementary Reversed-Phase Separations for Comprehensive Coverage of the Semipolar and Nonpolar Metabolome. *Anal. Bioanal. Chem.* **2018**, *410*, 1287–1297. https://doi.org/10.1007/s00216-017-0768-x.

(167) Tulipani, S.; Mora-Cubillos, X.; Jáuregui, O.; Llorach, R.; García-Fuentes, E.; Tinahones, F. J.; Andres-Lacueva, C. New and Vintage Solutions To Enhance the Plasma Metabolome Coverage by LC-ESI-MS Untargeted Metabolomics: The Not-So-Simple Process of Method Performance Evaluation. *Anal. Chem.* **2015**, *87* (5), 2639–2647.

https://doi.org/10.1021/ac503031d.

(168) Chetwynd, A. J.; David, A. A Review of Nanoscale LC-ESI for Metabolomics and Its Potential to Enhance the Metabolome Coverage. *Talanta* **2018**, *182* (February), 380–390. https://doi.org/10.1016/j.talanta.2018.01.084.

(169) Commisso, M.; Anesi, A.; Dal Santo, S.; Guzzo, F. Performance Comparison of Electrospray Ionization and Atmospheric Pressure Chemical Ionization in Untargeted and Targeted Liquid Chromatography/Mass Spectrometry Based Metabolomics Analysis of Grapeberry Metabolites. *Rapid Commun. Mass Spectrom.* **2017**, *31* (3), 292–300. https://doi.org/10.1002/rcm.7789.

(170) Riaz, N.; Wolden, S. L.; Gelblum, D. Y.; Eric, J. Biological Matrix Effects in Quantitative Tandem Mass Spectrometry-Based Analytical Methods: Advancing Biomonitoring. *Crit. reveiws Anal. Chem.* **2016**, *46*, 93–105. https://doi.org/10.1002/cncr.27633.Percutaneous.

(171) Tebani, A.; Schmitz-afonso, I.; Rutledge, D. N.; Gonzalez, B. J.; Bekri, S.; Afonso, C. Analytica Chimica Acta Optimization of a Liquid Chromatography Ion Mobility-Mass Spectrometry Method for Untargeted Metabolomics Using Experimental Design and Multivariate Data Analysis *. *Anal. Chim. Acta* **2016**, *913*, 55–62. https://doi.org/10.1016/j.aca.2016.02.011.

(172) Tulipani, S.; Llorach, R.; Urpi-Sarda, M.; Andres-Lacueva, C. Comparative Analysis of Sample Preparation Methods To Handle the Complexity of the Blood Fluid Metabolome: When Less Is More. *Anal. Chem.* **2013**, No. 85, 341–348.

(173) Sitnikov, D. G.; Monnin, C. S.; Vuckovic, D. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Sci. Rep.* **2016**, *6* (November), 38885. https://doi.org/10.1038/srep38885.

(174) Yang, Y.; Cruickshank, C.; Armstrong, M.; Mahaffey, S.; Residorph, R.; Reisdorph, N. New Sample Preparation Approach for Mass Spectrometry-Based Profiling of Plasma Results in Improved Coverage of Metabolome. *J. Chromatogr. A* **2013**, *1300*, 217–226. https://doi.org/10.1016/j.chroma.2013.04.030.New.

(175) Want, E. J.; Maille, G. O.; Smith, C. A.; Brandon, T. R.; Uritboonthai, W.; Qin, C.; Trauger, S. A.; Siuzdak, G. Solvent-Dependent Metabolite Distribution, Clustering, and Protein

Extraction for Serum Profiling with Mass Spectrometry. *Anal. Chem.* **2006**, *78* (3), 743–752.

(176) Tsakelidou, E.; Virgiliou, C.; Valianou, L.; Gika, H. G.; Raikos, N.; Theodoridis, G. Sample Preparation Strategies for the Effective Quantitation of Hydrophilic Metabolites in Serum by Multi-Targeted HILIC-MS/MS. *Metabolites* **2017**, *7* (2). https://doi.org/10.3390/metabo7020013.

(177) Skov, K.; Hadrup, N.; Smedsgaard, J.; Frandsen, H. LC–MS Analysis of the Plasma Metabolome—A Novel Sample Preparation Strategy. *J. Chromatogr. B* **2015**, *978–979*, 83–88. https://doi.org/10.1016/j.jchromb.2014.11.033.

(178) Zhao, S.; Li, L. Chemical Derivatization in LC-MS-Based Metabolomics Study. *TrAC - Trends Anal. Chem.* **2020**, *131*. https://doi.org/10.1016/j.trac.2020.115988.

(179) Zheng, L.; Zhao, X. E.; Zhu, S.; Tao, Y.; Ji, W.; Geng, Y.; Wang, X.; Chen, G.; You, J. A New Combined Method of Stable Isotope-Labeling Derivatization-Ultrasound-Assisted Dispersive Liquid–Liquid Microextraction for the Determination of Neurotransmitters in Rat Brain Microdialysates by Ultra High Performance Liquid Chromatography Tandem Mas. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2017**, *1054*, 64–72. https://doi.org/10.1016/j.jchromb.2017.03.039.

(180) Malec, P. A. Derivatization Methods for Improved Metabolome Analysis by LC-MS/MS, 2018.

(181) Li, Y.; Li, L. Retention Time Shift Analysis and Correction in Chemical Isotope Labeling Liquid Chromatography/Mass Spectrometry for Metabolome Analysis. *Rapid Commun. Mass Spectrom.* **2020**, *34* (S1), 1–8. https://doi.org/10.1002/rcm.8643.

(182) Kraly, J. R.; Holcomb, R. E.; Guan, Q.; Henry, C. S. Review: Microfluidic Applications in Metabolomics and Metabolic Profiling. *Anal. Chim. Acta* **2009**, *653* (1), 23–35. https://doi.org/10.1016/j.aca.2009.08.037.Review.

(183) Wishart, D. S. Advances in Metabolite Identification. *Bioanalysis* **2011**, *3* (15), 1769–1782. https://doi.org/10.4155/bio.11.155.

(184) Wishart, D. S.; Jewison, T.; Guo, a. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.;

Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41* (D1), D801–D807. https://doi.org/10.1093/nar/gks1065.

(185) Bruce, S. J.; Tavazzi, I.; Parisod, V.; Rezzi, S.; Kochhar, S.; Guy, P. a. Investigation of Human Blood Plasma Sample Preparation for Performing Metabolomics Using Ultrahigh Performance Liquid Chromatography/Mass Spectrometry. *Anal. Chem.* **2009**, *81* (9), 3285–3296. https://doi.org/10.1021/ac8024569.

(186) Theodoridis, G.; Gika, H. G.; Wilson, I. D.; Park, A.; Spec, M.; Introduction, I.; Metabolomics, A. Mass Spectrometry-Based Holistic Analytical Approaches for Metabolite Profiling in Systems Biology Studies. *Mass Spectrom Rev.* **2011**, No. 30, 884–906.

(187) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. Lipid Extraction by Methyl-Tert-Butyl Ether for High-Throughput Lipidomics. *J. Lipid Res.* **2008**, *49* (5), 1137–1146. https://doi.org/10.1194/jlr.D700041-JLR200.

(188) Whiley, L.; Godzien, J.; Ruperez, F. J.; Legido-Quigley, C.; Barbas, C. In-Vial Dual Extraction for Direct LC-MS Analysis of Plasma for Comprehensive and Highly Reproducible Metabolic Fingerprinting. *Anal. Chem.* **2012**, *84* (14), 5992–5999. https://doi.org/10.1021/ac300716u.

(189) Michopoulos, F.; Lai, L.; Gika, H.; Theodoridis, G.; Wilson, I. UPLC-MS-Based Analysis of Human Plasma for Metabonomics Using Solvent Precipitation or Solid Phase Extraction. *J. Proteome Res.* **2009**, *8* (4), 2114–2121. https://doi.org/10.1021/pr801045q.

(190) Yang, J.; Chen, T.; Sun, L.; Zhao, Z.; Qi, X.; Zhou, K.; Cao, Y.; Wang, X.; Qiu, Y.; Su, M.; Zhao, A.; Wang, P.; Yang, P.; Wu, J.; Feng, G.; He, L.; Jia, W.; Wan, C. Potential Metabolite Markers of Schizophrenia. *Mol. Psychiatry* **2013**, 67–78. https://doi.org/10.1038/mp.2011.131.

(191) Bueschl, C.; Krska, R.; Kluger, B.; Schuhmacher, R. Isotopic Labeling-Assisted Metabolomics Using LC – MS. *Anal. Bioanal. Chem.* **2013**, *405*, 27–33. https://doi.org/10.1007/s00216-012-6375-y.

(192) Matuszewski, B. K. Standard Line Slopes as a Measure of a Relative Matrix Effect in Quantitative HPLC-MS Bioanalysis. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.*

2006, *830* (2), 293–300. https://doi.org/10.1016/j.jchromb.2005.11.009.

(193) Vuckovic, D.; Pawliszyn, J. Systematic Evaluation of Solid-Phase Microextraction Coatings for Untargeted Metabolomic Profiling of Biological Fluids by Liquid Chromatography-Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 1944–1954. https://doi.org/10.1021/ac102614v.

(194) Noga, S.; Buszewski, B. Hydrophilic Interaction Liquid Chromatography (HILIC) — a Powerful Separation Technique. *Anal. Bioanal. Chem.* **2012**, *402* (1), 231–247. https://doi.org/10.1007/s00216-011-5308-5.

(195) Van De Steene, J.; Lambert, W. Validation of a Solid-Phase Extraction and Liquid Chromatography – Electrospray Tandem Mass Spectrometric Method for the Determination of Nine Basic Pharmaceuticals in Wastewater and Surface Water Samples. *J. Chromatogr. A* **2008**, *1182*, 153–160. https://doi.org/10.1016/j.chroma.2008.01.012.

(196) Yanes, Oscar Tautenhahn, Ralf Patti J., Suizdak, G. Expanding Coverage of the Metabolome for Global Metabolite Profiling. *Anal. Chem.* **2011**, *83* (6), 2152–2161. https://doi.org/10.1021/ac102981k.Expanding.

(197) Weinstein, J. N.; Myers, T. G.; O'Connor, P. M.; Friend, S. H.; Fornace, a J.; Kohn, K. W.; Fojo, T.; Bates, S. E.; Rubinstein, L. V; Anderson, N. L.; Buolamwini, J. K.; van Osdol, W. W.; Monks, a P.; Scudiero, D. a; Sausville, E. a; Zaharevitz, D. W.; Bunow, B.; Viswanadhan, V. N.; Johnson, G. S.; Wittes, R. E.; Paull, K. D. An Information-Intensive Approach to the Molecular Pharmacology of Cancer. *Science* **1997**, *275* (5298), 343–349. https://doi.org/10.1126/science.275.5298.343.

(198) Shi, T.; Fillmore, T. L.; Gao, Y.; Zhao, R.; He, J.; Athena, A.; Nicora, C. D.; Wu, C.; Chambers, J. L.; Moore, R. J.; Kagan, J.; Srivastava, S.; Liu, A. Y.; Rodland, K. D.; Liu, T.; David, G.; Ii, C.; Smith, R. D.; Qian, W. Long-Gradient Separations Coupled with Selected Reaction Monitoring for Highly Sensitive, Large Scale Targeted Protein Quantification in a Single Analysis. *Anal. Chem.* **2013**, *85* (19), 1–16. https://doi.org/10.1021/ac402105s.Long-gradient.

(199) Sarafian; Magali H; Gaudin, M.; Lewis, M. R.; Martin, F.; Holmes, E.; Nicholson, J. K.; Dumas, M.; Sarafian, M. H.; Gaudin, M.; Lewis, M. R.; Martin, F.; Holmes, E.; Nicholson, J. K.; Dumas, M. Objective Set of Criteria for Optimization of Sample Preparation

Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Pro Fi Ling by Ultra Performance Liquid Chromatography − Mass Spectrometry. *Anal. Chem.* **2014**, *86*, 5766–5774. https://doi.org/10.1021/ac500317c.

(200) Michopoulos, F.; Gika, H.; Palachanis, D.; Theodoridis, G.; Wilson, I. D. Solid Phase Extraction Methodology for UPLC-MS Based Metabolic Profiling of Urine Samples. *Electrophoresis* **2015**, *36* (18), 2170–2178. https://doi.org/10.1002/elps.201500101.

(201) Hewavitharana, A.; Shaw, N.; Swee, T. Strategies for the Detection and Elimination of Matrix Effects in Quantitative LC–MS Analysis. *LCGC North Am.* **2014**, *32* (1), 54–64.

(202) Patterson, R. E.; Ducrocq, A. J.; McDougall, D. J.; Garrett, T. J.; Yost, R. A.; He, Q.; Johnston, J.; Zeitlinger, J. Comparison of Blood Plasma Sample Preparation Methods for Combined LC-MS Lipidomics and Metabolomics. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2015**, *33* (4), 395–401. https://doi.org/10.1038/nbt.3121.ChIP-nexus.

(203) Kaufmann, A.; Walker, S. Evaluation of the Interrelationship between Mass Resolving Power and Mass Error Tolerances for Targeted Bioanalysis Using Liquid Chromatography Coupled to High-Resolution Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2013**, *27* (2), 347–356. https://doi.org/10.1002/rcm.6454.

(204) Wang, S.; Zhou, L.; Wang, Z.; Shi, X.; Xu, G. Simultaneous Metabolomics and Lipidomics Analysis Based on Novel Heart-Cutting Two-Dimensional Liquid Chromatography-Mass Spectrometry. *Anal. Chim. Acta* **2017**, *966*, 34–40. https://doi.org/10.1016/j.aca.2017.03.004.

(205) Want, E. J.; Maille, G. O.; Smith, C. A.; Brandon, T. R.; Uritboonthai, W.; Qin, C.; Trauger, S. A.; Siuzdak, G. Clustering , and Protein Extraction for Serum Profiling with Mass Spectrometry. *Anal Chem* **2006**, *78* (3), 743–752.

(206) Veyel, D.; Sokolowska, E. M.; Moreno, J. C.; Kierszniowska, S.; Cichon, J.; Wojciechowska, I.; Luzarowski, M.; Kosmacz, M.; Szlachetko, J.; Gorka, M.; Méret, M.; Graf, A.; Meyer, E. H.; Willmitzer, L.; Skirycz, A. PROMIS, Global Analysis of PROtein-Metabolite Interactions Using Size Separation in Arabidopsis Thaliana. *J. Biol. Chem.* **2018**, *293* (32), 12440–12453. https://doi.org/10.1074/jbc.RA118.003351.

(207) Mchugh, C. E.; Flott, T. L.; Schooff, C. R.; Smiley, Z.; Puskarich, M. A.; Myers, D. D.;

Younger, J. G.; Jones, A. E.; Stringer, K. A. Rapid , Reproducible , Quantifiable NMR Metabolomics : Methanol and Methanol : Chloroform Precipitation for Removal of Macromolecules in Serum and Whole Blood. 1–15. https://doi.org/10.3390/metabo8040093.

(208) Pang, Z.; Chong, J.; Li, S.; Xia, J. Metaboanalystr 3.0: Toward an Optimized Workflow for Global Metabolomics. *Metabolites* **2020**, *10* (5). https://doi.org/10.3390/metabo10050186.

(209) Ahmad, S.; Kalra, H.; Gupta, A.; Raut, B.; Hussain, A.; Rahman, M. A Novel Technique to Reduce Phospholipid-Based Matrix Effect in LC-ESI-MS Bioanalysis. *J. Pharm. Bioall. Sci.* **2012**, *4*, 267–275.

(210) Carmical, J.; Brown, S. The Impact of Phospholipids and Phospholipid Removal on Bioanalytical Method Performance. *Biomed. Chromatogr.* **2016**, *30* (5), 710–720. https://doi.org/10.1002/bmc.3686.

(211) Lindsay, A.; Janmale, T.; Draper, N.; Gieseg, S. P. Measurement of Changes in Urinary Neopterin and Total Neopterin in Body Builders Using SCX HPLC. *Pteridins* **2014**, *25* (2), 53–62. https://doi.org/10.1515/pteridines-2014-0003.

(212) Wernisch, S.; Pennathur, S. Evaluation of Coverage, Retention Patterns, and Selectivity of Seven Liquid Chromatographic Methods for Metabolomics. *Anal. Bioanal. Chem.* **2016**, *408* (22), 6079–6091. https://doi.org/10.1007/s00216-016-9716-4.

(213) Needham, T. E. The Solubility of Amino Acids in Various Solvent Systems, 1970.

(214) Schranner, D.; Kastenmüller, G.; Schönfelder, M.; Römisch-Margl, W.; Wackerhage, H. Metabolite Concentration Changes in Humans After a Bout of Exercise: A Systematic Review of Exercise Metabolomics Studies. *Sport. Med. - Open* **2020**, *6* (1). https://doi.org/10.1186/s40798-020-0238-4.

(215) Ferreiro-vera, C.; Priego-capote, F.; Castro, M. D. L. De. Comparison of Sample Preparation Approaches for Phospholipids Profiling in Human Serum by Liquid Chromatography – Tandem Mass Spectrometry. *J. Chromatogr. A* **2012**, *1240*, 21–28. https://doi.org/10.1016/j.chroma.2012.03.074.

(216) Fauland, A.; Trötzmüller, M.; Eberl, A.; Afiuni-Zadeh, S.; Köfeler, H.; Guo, X.; Lankmayr, E. An Improved SPE Method for Fractionation and Identification of Phospholipids. *J. Sep.*

Sci. **2013**, *36* (4), 744–751. https://doi.org/10.1002/jssc.201200708.

(217) Johnsen, E.; Leknes, S.; Wilson, S. R.; Lundanes, E. Liquid Chromatography-Mass Spectrometry Platform for Both Small Neurotransmitters and Neuropeptides in Blood, with Automatic and Robust Solid Phase Extraction. *Sci. Rep.* **2015**, *5*, 1–8. https://doi.org/10.1038/srep09308.

(218) Pruessner, J.; Hellhammer, D.; Kirschbaum, C. Burnout , Perceived Stress , and Cortisol Responses to Awakening . *Psychosom. Med.* **1999**, *61* (2), 197–204. https://doi.org/10.1097/00006842-199903000-00012.

(219) Papacosta, E.; Nassis, G. Saliva as a Tool for Monitoring Steroid , Peptide and Immune Markers in Sport and Exercise Science . *J. Sci. Med. Sport* **2011**, *14* (5), 3–4. https://doi.org/10.1016/j.jsams.2011.03.004.

(220) Eller, N.; Netterstrøm, B.; Hansen, A. Psychosocial Factors at Home and at Work and Levels of Salivary Cortisol . *Biol. Psychol.* **2006**, *73* (3), 280–287. https://doi.org/10.1016/j.biopsycho.2006.05.003.

(221) Hackney, A.; Mcmurray, R.; Ondrak, K.; VanBruggen, M. The Relationship between Serum and Salivary Cortisol Levels in Response to Different Intensities of Exercise . *Int. J. Sport. Physiol. Perform.* **2011**, *6* (3), 396–407. https://doi.org/10.1123/ijspp.6.3.396.

(222) Kirschbaum, C.; Hellhammer, D. H. Salivary Cortisol in Psychoneuroendocrine Research: Recent Developments and Applications. *Psychoneuroendocrinology* **1994**, *19* (4), 313–333. https://doi.org/10.1016/0306-4530(94)90013-2.

(223) Aardal, E.; Holm, A. Cortisol in Saliva--Reference Ranges and Relation to Cortisol in Serum . *Eur. J. Clin. Chem. Clin. Biochem.* **1995**, *33* (12), 927–932. https://doi.org/10.1515/cclm.1995.33.12.927.

(224) Perogamvros, I.; Keevil, B. G.; Ray, D. W.; Trainer, P. J. Salivary Cortisone Is a Potential Biomarker for Serum Free Cortisol. *J. Clin. Endocrinol. Metab.* **2010**, *95* (11), 4951–4958. https://doi.org/10.1210/jc.2010-1215.

(225) Harrison, R. F.; Debono, M.; Whitaker, M. J.; Keevil, B. G.; Newell-Price, J.; Ross, R. J. Salivary Cortisone to Estimate Cortisol Exposure and Sampling Frequency Required Based on Serum Cortisol Measurements. *J. Clin. Endocrinol. Metab.* **2018**, *104* (3), 765–772.

https://doi.org/10.1210/jc.2018-01172.

(226) Wood, L.; Ducroq, D.; Fraser, H.; Gillingwater, S.; Evans, C.; Pickett, A.; Rees, D.; John, R. Measurement of Urinary Free Cortisol by Tandem Mass Spectrometry and Comparison with Results Obtained by Gas Chromatography-Mass Spectrometry and Two Commercial Immunoassays. *Ann. Clin. Biochem.* **2008**, *45* (Pt 4). https://doi.org/10.1258/acb.2007.007119.

(227) Ashraf, S. S.; El-gahany, W.; Alraeesi, A.; Al-hajj, L.; Al-maidalli, A. Analysis of Illicit Glucocorticoid Levels in Camel Hair Using Competitive ELISA – Comparison with LC-MS / MS. *Drug Test. Anal.* **2020**. https://doi.org/10.1002/dta.2750.

(228) El-Farhan, N.; Rees, D. A.; Evans, C. Measuring Cortisol in Serum, Urine and Saliva – Are Our Assays Good Enough? *Ann. Clin. Biochem.* **2017**, *54* (3), 308–322. https://doi.org/10.1177/0004563216687335.

(229) Hawley, J. M.; Owen, L. J.; Mackenzie, F.; Mussell, C.; Cowen, S.; Keevil, B. G. Candidate Reference Measurement Procedure for the Quantification of Total Serum Cortisol with LC-MS/MS. *Clin. Chem.* **2016**, *62* (1), 262–269. https://doi.org/10.1373/clinchem.2015.243576.

(230) Handelsman, D. J.; Wartofsky, L. Requirement for Mass Spectrometry Sex Steroid Assays in the Journal of Clinical Endocrinology and Metabolism. *J. Clin. Endocrinol. Metab.* **2013**, *98* (10), 3971–3973. https://doi.org/10.1210/jc.2013-3375.

(231) Krumbholz, A.; Schönfelder, M.; Hofmann, H.; Thieme, D. The Plasma Protein Binding of the Endogenous Glucocorticosteroids Is of Vital Importance for the Concentrations in Hair and Saliva. *Forensic Sci. Int.* **2018**, *286*, 23–30. https://doi.org/10.1016/j.forsciint.2018.01.030.

(232) Magda, B.; Dobi, Z.; Mészáros, K.; Szabó, É.; Márta, Z.; Imre, T.; Szabó, P. T. Charged Derivatization and On-Line Solid Phase Extraction to Measure Extremely Low Cortisol and Cortisone Levels in Human Saliva with Liquid Chromatography–Tandem Mass Spectrometry. *J. Pharm. Biomed. Anal.* **2017**, *140*, 223–231. https://doi.org/10.1016/j.jpba.2017.03.028.

(233) Sturmer, L. R.; Dodd, D.; Chao, C. S.; Shi, R. Z. Clinical Utility of an Ultrasensitive Late

Night Salivary Cortisol Assay by Tandem Mass Spectrometry. *Steroids* **2018**, *129* (November 2017), 35–40. https://doi.org/10.1016/j.steroids.2017.11.014.

(234) Mészáros, K.; Karvaly, G.; Márta, Z.; Magda, B.; Tőke, J.; Szücs, N.; Tóth, M.; Rácz, K.; Patócs, A. Diagnostic Performance of a Newly Developed Salivary Cortisol and Cortisone Measurement Using an LC–MS/MS Method with Simple and Rapid Sample Preparation. *J. Endocrinol. Invest.* **2018**, *41* (3), 315–323. https://doi.org/10.1007/s40618-017-0743-6.

(235) Antonelli, G.; Padoan, A.; Artusi, C.; Marinova, M.; Zaninotto, M.; Plebani, M. Automated Saliva Processing for LC-MS/MS: Improving Laboratory Efficiency in Cortisol and Cortisone Testing. *Clin. Biochem.* **2016**, *49* (6), 518–520. https://doi.org/10.1016/j.clinbiochem.2015.12.006.

(236) Mezzullo, M.; Fanelli, F.; Fazzini, A.; Gambineri, A.; Vicennati, V.; Di Dalmazi, G.; Pelusi, C.; Mazza, R.; Pagotto, U.; Pasquali, R. Validation of an LC–MS/MS Salivary Assay for Glucocorticoid Status Assessment: Evaluation of the Diurnal Fluctuation of Cortisol and Cortisone and of Their Association within and between Serum and Saliva. *J. Steroid Biochem. Mol. Biol.* **2016**, *163* (2015), 103–112. https://doi.org/10.1016/j.jsbmb.2016.04.012.

(237) García-Blanco, A.; Vento, M.; Diago, V.; Cháfer-Pericás, C. Reference Ranges for Cortisol and α-Amylase in Mother and Newborn Saliva Samples at Different Perinatal and Postnatal Periods. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2016**, *1022*, 249–255. https://doi.org/10.1016/j.jchromb.2016.04.035.

(238) Perogamvros, I.; Owen, L. J.; Newell-Price, J.; Ray, D. W.; Trainer, P. J.; Keevil, B. G. Simultaneous Measurement of Cortisol and Cortisone in Human Saliva Using Liquid Chromatography-Tandem Mass Spectrometry: Application in Basal and Stimulated Conditions. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2009**, *877* (29), 3771–3775. https://doi.org/10.1016/j.jchromb.2009.09.014.

(239) Ionita, I. A.; Fast, D. M.; Akhlaghi, F. Development of a Sensitive and Selective Method for the Quantitative Analysis of Cortisol, Cortisone, Prednisolone and Prednisone in Human Plasma. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2009**, *877* (8–9), 765–772. https://doi.org/10.1016/j.jchromb.2009.02.019.

(240) Fustinoni, S.; Polledri, E.; Mercadante, R. High-Throughput Determination of Cortisol, Cortisone, and Melatonin in Oral Fluid by on-Line Turbulent Flow Liquid Chromatography Interfaced with Liquid Chromatography/Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2013**, *27* (13), 1450–1460. https://doi.org/10.1002/rcm.6601.

(241) McBride, E. M.; Lawrence, R. J.; McGee, K.; Mach, P. M.; Demond, P. S.; Busch, M. W.; Ramsay, J. W.; Hussey, E. K.; Glaros, T.; Dhummakupt, E. S. Rapid Liquid Chromatography Tandem Mass Spectrometry Method for Targeted Quantitation of Human Performance Metabolites in Saliva. *J. Chromatogr. A* **2019**, *1601*, 205–213. https://doi.org/10.1016/j.chroma.2019.04.071.

(242) Cao, Z. (Tim); Wemm, S. E.; Han, L.; Spink, D. C.; Wulfert, E. Noninvasive Determination of Human Cortisol and Dehydroepiandrosterone Sulfate Using Liquid Chromatography-Tandem Mass Spectrometry. *Anal. Bioanal. Chem.* **2019**, *411* (6), 1203–1210. https://doi.org/10.1007/s00216-018-1549-x.

(243) Steptoe, A.; Ussher, M. Smoking, Cortisol and Nicotine. *Int. J. Psychophysiol.* **2006**, *59* (3), 228–235. https://doi.org/10.1016/j.ijpsycho.2005.10.011.

(244) Wright, H. E.; Selkirk, G. A.; McLellan, T. M. HPA and SAS Responses to Increasing Core Temperature during Uncompensable Exertional Heat Stress in Trained and Untrained Males. *Eur. J. Appl. Physiol.* **2010**, *108* (5), 987–997. https://doi.org/10.1007/s00421-009-1294-0.

(245) Gaudl, A.; Kratzsch, J.; Bae, Y. J.; Kiess, W.; Thiery, J.; Ceglarek, U. Liquid Chromatography Quadrupole Linear Ion Trap Mass Spectrometry for Quantitative Steroid Hormone Analysis in Plasma, Urine, Saliva and Hair. *J. Chromatogr. A* **2016**, *1464*, 64–71. https://doi.org/10.1016/j.chroma.2016.07.087.

(246) Antonelli, G.; Ceccato, F.; Artusi, C.; Marinova, M.; Plebani, M. Salivary Cortisol and Cortisone by LC-MS/MS: Validation, Reference Intervals and Diagnostic Accuracy in Cushing's Syndrome. *Clin. Chim. Acta* **2015**, *451*, 247–251. https://doi.org/10.1016/j.cca.2015.10.004.

(247) Zhang, Q.; Chen, Z.; Chen, S.; Xu, Y.; Deng, H. Intraindividual Stability of Cortisol and Cortisone and the Ratio of Cortisol to Cortisone in Saliva, Urine and Hair. *Steroids* **2017**, *118*, 61–67. https://doi.org/10.1016/j.steroids.2016.12.008.

(248) Miller, R.; Plessow, F.; Rauh, M.; Gröschl, M.; Kirschbaum, C. Comparison of Salivary Cortisol as Measured by Different Immunoassays and Tandem Mass Spectrometry. *Psychoneuroendocrinology* **2013**, *38* (1), 50–57. https://doi.org/10.1016/j.psyneuen.2012.04.019.

(249) Debono, M.; Harrison, R. F.; Whitaker, M. J.; Eckland, D.; Arlt, W.; Keevil, B. G.; Ross, R. J. Salivary Cortisone Reflects Cortisol Exposure under Physiological Conditions and after Hydrocortisone. *J. Clin. Endocrinol. Metab.* **2016**, *101* (4), 1469–1477. https://doi.org/10.1210/jc.2015-3694.

(250) Jönsson, B. A. G.; Malmberg, B.; Amilon, Å.; Garde, A. H.; Ørbæk, P. Determination of Cortisol in Human Saliva Using Liquid Chromatography-Electrospray Tandem Mass Spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2003**, *784* (1), 63–68. https://doi.org/10.1016/S1570-0232(02)00753-5.

(251) Mezzullo, M.; Fanelli, F.; Di Dalmazi, G.; Fazzini, A.; Ibarra-Gasparini, D.; Mastroroberto, M.; Guidi, J.; Morselli-Labate, A. M.; Pasquali, R.; Pagotto, U.; Gambineri, A. Salivary Cortisol and Cortisone Responses to Short-Term Psychological Stress Challenge in Late Adolescent and Young Women with Different Hyperandrogenic States. *Psychoneuroendocrinology* **2018**, *91* (February), 31–40. https://doi.org/10.1016/j.psyneuen.2018.02.022.

(252) Lee, S.; Lim, H. S.; Shin, H. J.; Kim, S. A.; Park, J.; Kim, H. C.; Kim, H.; Hyung, J. K.; Tae, KimHyogyeong Hyung Chul, Y.; Lee, K.-R.; Kim, Y. J. Simultaneous Determination of Cortisol and Cortisone from Human Serum by Liquid Chromatography-Tandem Mass Spectrometry. *J. Anal. Methods Chem.* **2014**, *2014*. https://doi.org/10.1155/2014/787483.

(253) Vogeser, M.; Kratzsch, J.; Ju Bae, Y.; Bruegel, M.; Ceglarek, U.; Fiers, T.; Gaudl, A.; Kurka, H.; Milczynski, C.; Prat Knoll, C.; Suhr, A. C.; Teupser, D.; Zahn, I.; Ostlund, R. E. Multicenter Performance Evaluation of a Second Generation Cortisol Assay. *Clin. Chem. Lab. Med.* **2017**, *55* (6), 826–835. https://doi.org/10.1515/cclm-2016-0400.

(254) Pragst, F. Chapter 13. High Performance Liquid Chromatography in Forensic Toxicological Analysis. In *Handbook of Analytical Separations*; 2008; Vol. 3, pp 447–489.

Appendix A

Supplementary information for Chapter 2

Supplementary materials and methods

Plasma extraction

Plasma extractions were analyzed on C18 or Scherzo columns coupled to QTOF via positive and (-ve) ESI. Therefore, each sample was analyzed in four LC-MS modes: (i) RP C18 UHPLC in positive and (-ve) ESI and, (ii) mixed /ion-exchange Scherzo HPLC in positive and (-ve) ESI. Buffer extractions were analyzed on Orbitrap Velos^{TM} mass spectrometer in (+ve) ESI.

LC-MS reversed-phase method for the analysis of extractions in plasma

C18 separation was executed at 0.4 mL/min flow rate at 35°C using binary solvent system consisting of 0.1% formic acid, 2% acetonitrile in water (solvent A) and 0.1% formic acid in 100% acetonitrile (solvent B). For LC-MS analysis in (-ve) ESI, concentration of formic acid in both solvents was reduced to 0.05%. The following gradient was used for separation: 2% B for 3 min, then increase of B from 2 to 100 % for 20 min followed by 2 min hold at 100% B and 4 min of re-equilibration to 100% A.

In addition, (+ve) ESI QTOF settings: capillary voltage of 3800 V for the entire run and nozzle voltage of 200 V for the first 4 min and 1500 V from 4th to 28th min of the analysis. In (-ve) ESI, capillary and nozzle voltages were set to 3500 and 500 V, respectively during the first 5.5 minutes of run and to 4200 and 800 V, respectively between 5.5 and 29th minutes. For both positive and (-ve) ESI, drying and sheath gas temperatures were set to 250 and 275°C and flow rates to 15 and 12 L/minutes, respectively. Nebulizer pressure was set to 30 psig and the fragmentor voltage to 175 V. Data was acquired in both centroid and profile mode at the rate of 3 spectra per second in the extended dynamic range mode (2 GHz). Resolution of 12,000 FWHM (full width at half maximum) at m/z 121 and 24,000 FWHM at m/z 922 was achieved. To assure the desired mass accuracy of recorded ions, continuous internal calibration was executed using signals at m/z 121.0509 (protonated purine) and m/z 922.0098 (protonated hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine (HP-921)) in positive ion mode. For (-ve) ESI analysis, ions with m/z 119.0363 (deprotonated purine) and m/z 966.0007 (formate adduct of HP-921) were used.

LC-MS mixed-mode method for the analysis of extractions in plasma

The stationary phase of Scherzo SM-C18 HPLC column is composed of C18 alkyl and weak cation and anion moieties. Therefore, solvent was supplemented with ammonium acetate buffer to execute mixed mode (RP/ion exchange) chromatography and elute ionic species from the column. The binary solvent system consisted of 2 mM ammonium acetate buffer, pH 3.5 in 2% acetonitrile and 98% of water (solvent A) and 100 mM ammonium acetate buffer, pH 3.5 in 96% acetonitrile and 4% of water (solvent B). The samples were separated at 0.22 mL/min flow rate at 35°C using the following gradient: 0% B for the first 2 min, increase of B from 0 to 20% between 2^{nd} and 10^{th} min, increase of B from 20 to 100% between 10^{th} and 25^{th} min, isocratic hold at 100% B for 6 min and 6 min of re-equilibration to 100% A. In (+ve) ESI, the capillary voltage was kept at 3500 V for the entire run, while nozzle voltage was set to 200 V for the first 5 min of run and to 800 V between 5^{th} and 35^{th} min. In (-ve) ESI, nozzle and capillary voltages were held at 250 V and 3750 V for the entire run. All other MS settings were the same as described for LC-MS RP in previous section.

For targeted analysis, all samples were analyzed in random order with QC samples (mixture of all analyzed samples (iii) loaded at each 11^{th} injection. XICs were extracted with 15 ppm accuracy and aligned within ± 0.15 min retention time interval. Metabolite identity was confirmed by the match of retention time and m/z between signals in analyzed samples and signals in calibration points. The concentrations of standard analytes were quantitated using external calibration curves obtained as described in materials and methods. For both targeted and global analyses, at least 10 runs of QC samples preceded batch runs in order to stabilize chromatography and ESI performance.

For the global analysis, to ensure confident comparisons across extraction methods in global metabolomics analysis, samples from solvent, LLE and SPE extractions were analyzed side by side in 6 randomly organized sub-batches containing a single replicate from each of 7 extractions and one QC sample. In order to analyze seven blank samples (type i) from seven methods in six sub-batches, five of sub-batches comprised one and the sixth sub-batch – two blank samples. PCA was performed using SIMCA 14 (Umetrics, Sweden) on all high-quality data described in Methods section, after Pareto scaling. PCA was used to verify the stability of LC-MS signals throughout the analysis by checking the clustering of QC signals for each analytical batch **(Figure 4.3**). It can also be used to visualize repeatability and similarity between different extraction methods.

LC-MS mixed-mode method for the analysis of extractions in buffer

The column and buffer composition were identical to the analysis in plasma. Samples were separated at 0.22 mL/min flow rate at 35 ºC using the following gradient: 0% B for the first 2 min, increase of B from 0 to 100% between 2nd and 18th min, isocratic hold at 100% B for 6 min and 7 min of re-equilibration to 100% A. Column was coupled to Orbitrap Velos^{TM} mass spectrometer in (+ve) ESI, at the following settings: ESI voltage 3.25 kV, heater and capillary temperature at 350 ºC, sheet gas, auxiliary gas and sweep gas flow rates at 40, 10 and 0 arbitrary units, respectively; detection of signals was executed in the range 50-1000 m/z at 250 ms activation time, normalized collision energy 35 V, resolution 60,000 FWHM and mass accuracy 10 ppm. Deconvolution of raw data and quantitative analysis was executed using Xcalibur ^{TM}, v. 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) and statistical analysis in Microsoft Excel unless otherwise specified.

Supplementary tables and figures

*Supplementary Table A1. Standard analytes used in the study. Monoisotopic masses and predicted ACD/LogP values were obtained from ChemSpider (http://www.chemspider.com). LogP values predicted by ACDLabs algorithm (**Section 2.2.2**). LogP predicted values could not be found for some metabolites in ChemSpider database and these entries are shown as not available (N/A) in the table. Retention times provided are for 500 ng/mL buffer calibration (= highest concentration analyzed) point analyzed in either ESI mode. ND stands for not detected and (-ve) designates analytes detected at (-ve) ESI mode. ACTH, epinephrine, norepinephrine and PC (19:0/19:0) were not detected in any conditions and are removed from subsequent tables. For details on the usage and fate of analytes in experiments see **Supplementary Tables A2, A3 and A7.** Other abbreviations: AmAc - ammonium acetate, FA- formic acid, NH4OH-ammonium hydroxide, IPA-isopropanol, DMSO-dimethylsulfoxide.*

Target analyte	Chemical group	Monoisotopic mass	RT on RP column (min)	RT on Scherzo column (min)	ACD/LogP	Buffer composition in the Individual stock
4-aminobutanoic acid	Zwitterion	103.0633	0.7	1.4	-0.9	50% methanol, 2 mM AmAc
4-aminobutanoic acid (d_6)	Zwitterion	109.1043	0.6	1.4	-0.9	50% methanol, 2 mM AmAc
5-methoxytryptamine	Zwitterion	190.1106	6.6	8.6	1.3	Water
ACTH	Zwitterion	4508.0410	ND	ND	N/A	Water
Adenine	Positive	135.0545	0.9	2.8	-2.1	Water
Cholic acid	Negative	408.2875	13.2(-ve)	19.1 (-ve)	2.3	0.1% NH_4OH
Cholic acid (d_4)	Negative	412.3167	13.2(-ve)	19.1 (-ve)	2.3	0.1% NH_4OH
Cortisol	Neutral	362.2093	10.7	16.8	1.4	50% methanol
Cortisol (d_4)	Neutral	366.2366	10.7	16.8	1.4	50% methanol
Cortisone	Neutral	360.1937	10.8	16.9	1.4	50% methanol
Creatinine	Zwitterion	113.0589	0.7	1.6	-1.6	50% methanol
Sphingosine	Zwitterion	299.2824	14.6	20.4	6.4	50% methanol
Dopamine	Positive	153.0790	0.7	1.7	0.1	20% methanol, 0.1% FA
Dopamine (d_4)	Positive	157.1063	0.7	1.7	0.1	See Dopamine
Epinephrine	Negative	183.0895	ND	ND	-0.6	See Dopamine
Epinephrine (d_3)	Negative	186.1100	ND	ND	-0.6	See Dopamine
Folic acid	Zwitterion	441.1397	6.6	14.7	-3	0.1% NH_4OH, 50% methanol
Glutamic acid	Zwitterion	147.0531	1.1	6.4	-3.9	water
Histamine	Zwitterion	111.0796	1.5	2	-0.7	water
Homovanillic acid	Negative	182.0579	7.3	14.2	1.1	as Folic acid
Homovanillic acid (d_3)	Negative	185.0784	7.3	14.2	1.1	as Folic acid
Kynurenine	Zwitterion	208.0848	2.9	5.8	1.1	50% methanol, 0.1% FA
Melatonin	Neutral	232.1210	9.2	15.5	1.2	50% methanol
Melatonin (d_4)	Neutral	236.1485	9.2	15.5	1.2	50% methanol
Neurotensin	Zwitterion	1671.9097	8.2	12.7	N/A	0.1% FA

Supplementary Table A1 (cont'd).

Target analyte	Chemical group	Monoisotopic mass	RT on RP column (min)	RT on Scherzo column (min)	ACD/LogP	Buffer composition in the Individual stock
Neurotensin	Zwitterion	1671.9097	8.2	12.7	N/A	0.1% FA
Norepinephrine (d_6)	Positive	175.1149	ND	ND	-0.9	see epinephrine
Pantothenic acid	Negative	219.1107	5.0	8.6	-0.4	See Folic
PC (19:0/19:0)	Positive	818.5540	ND	ND	11.5	90% methanol, 10% IPA
PE (17:0/17:0)	Positive	719.5465	21.1	ND	11.5	
Phenylalanine (d_5)	Zwitterion	170.1155	3.0	5.3	1.1	See Dopamine
Angiotensin II	Zwitterion	1045.5345	8.9	ND	N/A	See Dopamine
PI (18:3/22:4)	Neutral	908.5391	21.4	ND	11.5	See PC
Serotonin	Zwitterion	176.0950	2.5	6.8	0.2	See Dopamine
Thyroxine	Negative	776.6867	11.8	18.2	5.9	50-50 DMSO-methanol
Thyroxine ($^{13}C_6$)	Negative	782.6860	11.8	18.2	5.9	
Triiodothyronine	Negative	650.7900	11	17.2	5.1	
Tyrosine	Zwitterion	181.0739	1.4	3	0.4	See Dopamine
Thyrotropin-releasing hormone	Zwitterion	362.1703	1.5	2.6	N/A	See Dopamine

*Supplementary Table A2. Recovery of extractions of standard analytes from buffer. The table displays average recovery calculated from individual replicates (n=6, amount of standard analyte spiked before extraction = 100%) and a relative standard deviation (standard deviation/mean*100%, n=6) of recovered amounts. "ND" stands for "not detected", "N/A" stands for "not applicable. LogP was obtained from ChemSpider (http://www.chemspider.com/Default.aspx) and represents predicted octanol-water partition coefficients (predicted ACD/LogP, ACD/Labs, Toronto, Canada. For details on the usage and fate of analytes* ***see Supplementary Table A10.***

Analytes	Methanol-Ethanol	Methanol	Methanol-MTBE	MTBE	C18	PEP2	IEX
Glutamic acid	79.0	106.3	70.8	ND	ND	25.7	73.4
Tyrosine	73.9	79.1	71.4	ND	ND	2.3	70.0
Creatinine	83.7	91.6	87.6	ND	ND	ND	ND
Thyrotropin releasing hormone	68.8	74.1	66.9	10.4	10.2	39.3	57.7
Pantothenic Acid	73.2	79.5	73.9	13.1	9.1	27.7	120.9
Histamine	112.2	122.4	98.7	ND	ND	133.6	67.0
4-aminobutanoic acid	58.6	123.7	61.7	ND	55.5	141.9	ND
Adenine	75.9	80.4	73.5	44.6	ND	46.1	108.4
Dopamine	35.2	34.4	38.0	82.0	8.1	55.5	9.8
Serotonin	106.2	117.5	110.5	54.7	35.3	107.0	31.2
Homovanillic acid	56.0	50.3	51.6	102.7	60.0	122.2	ND
Phenylalanine (d_5)	64.7	69.2	66.5	14.3	11.1	22.8	128.8
5-methoxytryptamine	65.9	71.7	70.3	71.4	54.6	134.2	ND
Cortisol	26.7	33.6	29.1	112.6	108.9	70.3	ND
Cortisone	25.9	31.6	28.0	101.1	113.7	71.1	ND
Melatonin	93.6	58.8	99.0	104.1	109.8	122.4	ND
Triiodothyronine	76.3	83.3	75.4	90.5	105.4	142.8	ND
Thyroxine	18.8	25.9	10.3	57.2	18.6	ND	ND
Sphingosine	33.7	34.4	34.7	62.6	74.4	46.8	46.5
PI (19:0)	64.5	66.9	85.0	83.8	ND	ND	ND
PE (17:0)	89.6	103.6	81.7	45.5	ND	ND	ND
Angiotensin II	67.6	70.4	71.9	28.4	19.5	74.7	ND

*Supplementary Table A3. Precision of extractions of standard analytes from buffer. The table displays average recovery calculated from individual replicates (n=6, amount of standard analyte spiked before extraction = 100%) and a relative standard deviation (standard deviation/mean*100%, n=6) of recovered amounts. "ND" stands for "not detected", "N/A" stands for "not applicable. LogP was obtained from ChemSpider (http://www.chemspider.com/Default.aspx) and represents predicted octanol-water partition coefficients (predicted ACD/LogP, ACD/Labs, Toronto, Canada. For details on the usage and fate of analytes* ***see Supplementary Table A10.***

Analytes	Methanol-Ethanol	Methanol	Methanol-MTBE	MTBE	C18	PEP2	IEX
Glutamic acid	25.5	26.4	6.2	N/A	N/A	14.6	41.5
Tyrosine	4.6	12.2	5.7	N/A	N/A	9.0	30.8
Creatinine	9.0	8.7	5.8	N/A	N/A	N/A	N/A
Thyrotropin releasing hormone	4.0	12.6	4.0	42.0	3.0	14.0	11.2
Pantothenic Acid	3.9	8.6	1.5	65.0	26.8	13.2	18.9
Histamine	7.1	3.3	3.6	N/A	N/A	10.4	27.5
4-aminobutanoic acid	5.5	8.6	5.4	N/A	31.0	14.7	N/A
Adenine	14.0	20.4	14.1	46.9	N/A	17.1	13.1
Dopamine	8.4	11.0	23.4	24.6	25.9	20.8	40.5
Serotonin	6.4	2.9	4.1	17.3	22.6	11.7	10.9
Homovanillic acid	19.1	6.0	16.1	25.2	52.3	10.2	N/A
Phenylalanine (d_5)	4.9	2.8	3.0	61.0	21.3	16.1	17.6
5-methoxytryptamine	5.5	3.2	4.6	12.3	25.2	3.6	N/A
Cortisol	24.6	23.8	27.3	6.8	18.6	29.4	N/A
Cortisone	16.9	24.9	18.9	8.7	13.9	27.1	N/A
Melatonin	14.4	9.6	12.1	9.0	20.5	16.8	N/A
Triiodothyronine	5.2	3.5	2.8	6.1	11.7	6.6	N/A
Thyroxine	15.0	56.3	15.6	55.6	18.3	N/A	N/A
Sphingosine	1.2	3.0	3.5	3.0	31.9	1.0	0.3
PI (19:0)	16.0	10.4	20.4	25.0	N/A	N/A	N/A
PE (17:0)	46.8	21	63.7	16.8	N/A	N/A	N/A
Angiotensin II	3.9	5.3	3.0	10.5	6.8	2.7	N/A

*Supplementary Table A4. Recovery of extractions of standard analytes from plasma. The table displays average recovery calculated from individual replicates after subtraction of any endogenous level of metabolite present (n=6, amount of standard analyte spiked before extraction = 100%) and a relative standard deviation (standard deviation/mean*100%, n=6) of recovered amounts. LogP was obtained online from ChemSpider website (http://www.chemspider.com/Default.aspx) and represents predicted octanol-water partition coefficients (predicted ACD/LogP, ACD/Labs, Toronto, Canada). "ND" stands for "not detected", "N/A" stands for "not applicable". For details on the usage and fate of analytes* ***see Supplementary Table A10.***

Analytes	Methanol-Ethanol	Methanol	Methanol-MTBE	MTBE	C18	PEP2	IEX
Folic acid	81.5	99.2	44.2	10.8	73.1	139.0	185.7
Adenine	91.1	75.3	96.6	13.0	244.0	245.6	ND
Histamine	28.8	68.0	46.3	ND	23.2	53.2	129.7
4-aminobutanoic acid (d_6)	65.6	86.3	74.5	ND	2.4	ND	33.2
4-aminobutanoic acid	ND	ND	ND	ND	15.5	ND	ND
Pantothenic Acid	110.3	109.3	102.9	7.9	12.6	5.4	95.1
Tyrosine	ND	ND	ND	101.6	53.7	18.6	ND
Homovanillic acid	98.0	111.6	92.7	8.5	ND	115.6	99.1
Homovanillic acid (d_3)	105.1	115.0	101.7	8.7	5.0	108.5	96
Kynurenine	ND	ND	ND	1.1	152.7	ND	ND
Melatonin (d_4)	118.0	121.0	119.0	103.9	163.9	158.8	2.2
Melatonin	106.1	106.8	104.1	92.6	150.7	157.3	1.5
5-Methoxytryptamine	73.1	80.4	66.8	18.8	107.1	109.4	ND
Cortisol (d_4)	93.4	90.7	80.7	69.4	5.2	2.7	ND
Cortisol	98.1	96.4	86.2	73.7	7.3	6.4	ND
Cortisone	69.3	74.0	38.3	74.5	2.9	0.4	ND
Cholic acid	114.4	123.5	121.4	ND	19.6	138.8	142.9
Cholic acid (d_4)	114.3	114.6	110.6	ND	12.7	139.7	147.6
Triiodothyronine	46.7	56.8	58.8	11.7	90.0	125.3	6.9
Thyroxin ($^{13}C_6$)	56.7	77.5	68.5	3.8	84.9	117.0	6.2
Thyroxine	60.3	80.8	72.4	4.6	77.1	105.7	5.8
Sphingosine	3.9	27.4	10.5	56.6	56.0	75.0	21.4
Neurotensin	62.0	68.9	39.3	6.9	143.3	154.7	5.2
Thyrotropin-releasing hormone	90.7	98.0	82.6	ND	102.9	125.8	134.9

*Supplementary Table A5. Precision of extraction of standard analytes from plasma. The table displays average recovery calculated from individual replicates after subtraction of any endogenous level of metabolite present (n=6, amount of standard analyte spiked before extraction = 100%) and a relative standard deviation (standard deviation/mean*100%, n=6) of recovered amounts. LogP was obtained online from ChemSpider website (http://www.chemspider.com/Default.aspx) and represents predicted octanol-water partition coefficients (predicted ACD/LogP, ACD/Labs, Toronto, Canada). "ND" stands for "not detected", "N/A" stands for "not applicable". For details on the usage and fate of analytes* ***see Supplementary Table A10.***

Analytes	Methanol - Ethanol	Methanol	Methanol - MTBE	MTBE	C18	PEP2	IEX
Folic acid	12.8	8.5	15.3	103.0	110.0	46.0	40.0
Adenine	8.7	16.3	15.6	69.2	81.2	46.0	N/A
Histamine	26.9	13.4	13.7	N/A	60.1	46.0	20.6
4-aminobutanoic acid (d_6)	8.4	20.8	7.7	N/A	12.1	N/A	14.2
4-aminobutanoic acid	N/A	N/A	N/A	N/A	59.0	N/A	N/A
Pantothenic Acid	5.4	3.0	3.6	123.8	38.9	42.7	18.0
Tyrosine	N/A	N/A	N/A	115.6	61.4	63.8	N/A
Homovanillic acid	8.7	5.8	4.0	245.0	N/A	52.1	11.4
Homovanillic acid (d_3)	5.9	7.0	8.7	235.0	0.5	48.7	12.8
Kynurenine	N/A	N/A	N/A	161.9	143.5	N/A	N/A
Melatonin (d_4)	9.7	6.4	3.1	18.8	6.3	9.5	244.9
Melatonin	8.6	4.5	3.4	19.1	6.6	9.9	244.9
5-methoxytryptamine	7.0	4.1	5.3	58.2	14.6	22.6	N/A
Cortisol (d_4)	5.2	5.4	8.9	35.0	91.6	131.4	N/A
Cortisol	6.2	7.8	8.8	36.7	58.5	53.8	N/A
Cortisone	9.2	7.4	42.5	29.4	93.0	N/A	N/A
Cholic acid	11.8	10.7	9.9	N/A	113.3	9.1	40.6
Cholic acid (d_4)	12.6	9.0	9.8	N/A	110.6	9.6	38.0
Triiodothyronine	36.0	23.8	10.4	23.3	10.1	32.4	28.6
Thyroxin ($^{13}C_6$)	32.0	27.6	27.1	51.2	13.5	40.9	2.7
Thyroxine	33.0	28.8	27.1	45.3	13.7	41.7	1.9
Sphingosine	18.2	15.4	20.1	55.6	19.2	81.0	54.3
Neurotensin	10.7	7.4	11.0	111.2	43.1	24.7	52.3
Thyrotropin releasing hormone	23.8	27.1	22.8	N/Ap	68.4	37.1	10.5

*Supplementary Table A6. Efficiency of extraction methods in buffer and plasma for metabolite standards using the data obtained in RP analysis in positive or (-ve) ESI mode. The first two columns compare the extraction methods using the number of analytes recovered in extractions from buffer and plasma at different thresholds of precision and recovery (R). The number of analytes recovered with RSD less than 20% and R≥80% and can be considered as metabolites where the extraction method provides excellent performance. The metabolites which have RSD less than 30%, and recovery higher than 50% can be considered as metabolites with acceptable performance in semi-quantitative methods such as global metabolomics. The table also shows mean recovery and repeatability (expressed as mean RSD) across all metabolites observed in a given extraction method. The total number of standard analytes detected from standard mixture by at least one extraction method is 22 metabolites for buffer and 24 metabolites for plasma extractions. Appendix A. Table 7 summarizes which metabolites were detected/analyzed in only one matrix. *Higher mean recovery of PEP2 is caused by the larger number of analytes with enhanced matrix effect (Table 2 in the Chapter 2) in RP analysis in (+ve) ESI compared to other extractions.*

Extraction method	Buffer				Spiked plasma			
	# of analytes RSD≤20 R≥80%	# of analytes RSD≤30 R≥50%	Mean R (%)	Mean RSD (%)	# of analytes RSD≤20 R≥80%	# of analytes RSD≤30 R≥50%	Mean R (%)	Mean RSD (%)
Methanol-ethanol	4	15	65.9	11.9	11	16	80.4	14.3
Methanol	5	16	73.1	12.9	11	19	89.6	12.4
Methanol-MTBE	5	15	66.2	12.0	10	19	77.1	13.3
MTBE	4	10	63.5	25.6	2	3	37.1	85.4
C18	4	5	52.9	22.0	5	7	69.8	53.4
PEP2	6	11	77.0	13.3	4	6	100.1*	42.5
IEX	3	5	71.4	21.2	4	5	69.6	52.2

*Supplementary Table A7. Extraction quality and resistance of standard analytes to matrix effects across LC-MS methods. Analytes ordered top to down accordingly to the LogP oct/water coefficient predicted by ACD laboratories algorithm. "ND" stands for not detected. "N" stands for neutral, "NA" stands not available as standard analyte at the time of extraction, "N/Ap" designates analytes for which matrix effect could not be calculated in any extraction method, "zwit." – for zwitterions, "neg." - for negative, "pos." - for positive considering the presence of either both, acidic or basic groups, respectively. Two columns "Number of methods for which metabolite had ≥80% recovery and ≤20% RSD:" provide the number of extraction methods in which an analyte was detected with RSD less than 20 and recovery higher than 80% in plasma or buffer. The matrix effect section (**2.3.1.4**) represents the number of extraction methods (total =7) in which matrix effect was not observed for the standard analyte. For the standard analytes marked with asterisk, the matrix effect for isotope analog was just slightly outside 80% limit. Thus, 4-aminobutanoic acid demonstrated 71.6% matrix effect in MTBE, while its deuterated analog demonstrated 89.4%, ** recovery of thyroxine was slightly lower (77.1%) then for its deuterated analog (84.9%). Both of these results lie within normal LC-MS experimental error and are not statistically significantly different. Histamine also showed no matrix effects in 2 and 5 extraction methods in RP and Scherzo analysis in (+ve) ESI and in 4 extraction methods in Scherzo (-ve) ESI analysis, respectively. *** Cholic acid was not found in any of liquid extractions (4 in total) in RP analysis at (-ve) ESI in contrary to its deuterated analog which was detected in all extractions.*

Analytes	Charge class	Number of methods for which metabolite had ≥80% recovery and ≤20% RSD:		Number of methods without matrix effect			
		in buffer	in plasma	RP (+ve) ESI	Scherzo (+ve) ESI	RP (-ve) ESI	Scherzo (-ve) ESI
Folic acid	zwit.	NA	2	7	2	4	N/Ap
Adenine	pos.	2	2	2	3	1	5
Histamine	zwit.	4	0	2	5	N/Ap	4
4-aminobutanoic acid (d6)	zwit.	NA	0	1*	0	N/Ap	N/Ap
4-aminobutanoic acid	zwit.	3	0	0	0	N/Ap	0
Pantothenic Acid	neg.	1	4	0	1	7	3
Thyrotropin releasing hormone	zwit.	0	0	7	6	2	4
Tyrosine	zwit.	0	0	3	3	1	N/Ap
Homovanillic acid	neg.	1	4	2	0	6	3
Homovanillic acid (d_3)	neg.	NA	4	0	0	6	3
Neurotensin	zwit.	NA	0	1	0	N/Ap	N/Ap
Kynurenine	zwit.	NA	0	2	2	N/Ap	N/Ap
Melatonin (d_4)	N	NA	6	0	0	4	2
Melatonin	N	4	6	0	0	4	2
5-Methoxytryptamine	zwit.	1	2	7	4	N/Ap	1
Cortisol	N	2	3	7	2	1	2
Cortisol (d_4)	N	2	3	7	N/Ap	N/Ap	N/Ap
Cortisone	N	2	0	7	2	0	3
Triiodothyronine	neg.	4	1	1	3	0	2
Thyroxine ($^{13}C_6$)	neg.	NA	1**	1	1	0	0
Thyroxine	neg.	0	0**	1	1	N/Ap	0
Cholic acid	neg.	NA	4	ND	ND	3***	2
Cholic acid (d_4)	neg.	NA	4	ND	ND	7	2
Sphingosine	zwit.	0	0	0	0	0	0

Supplementary Table A8. Summary of matrix effects observed for all metabolites across all extraction methods and RP LC-MS analyses. Not all analytes were successfully detected in buffer (NDB) or in an extract (NDM) which made the calculation of matrix effect impossible as specified. Analytes which did not respond to increased concentrations of spiked standards or which exhibited negative signal response were considered to be suppressed or saturated (SS) and were not used for calculation of matrix effects.

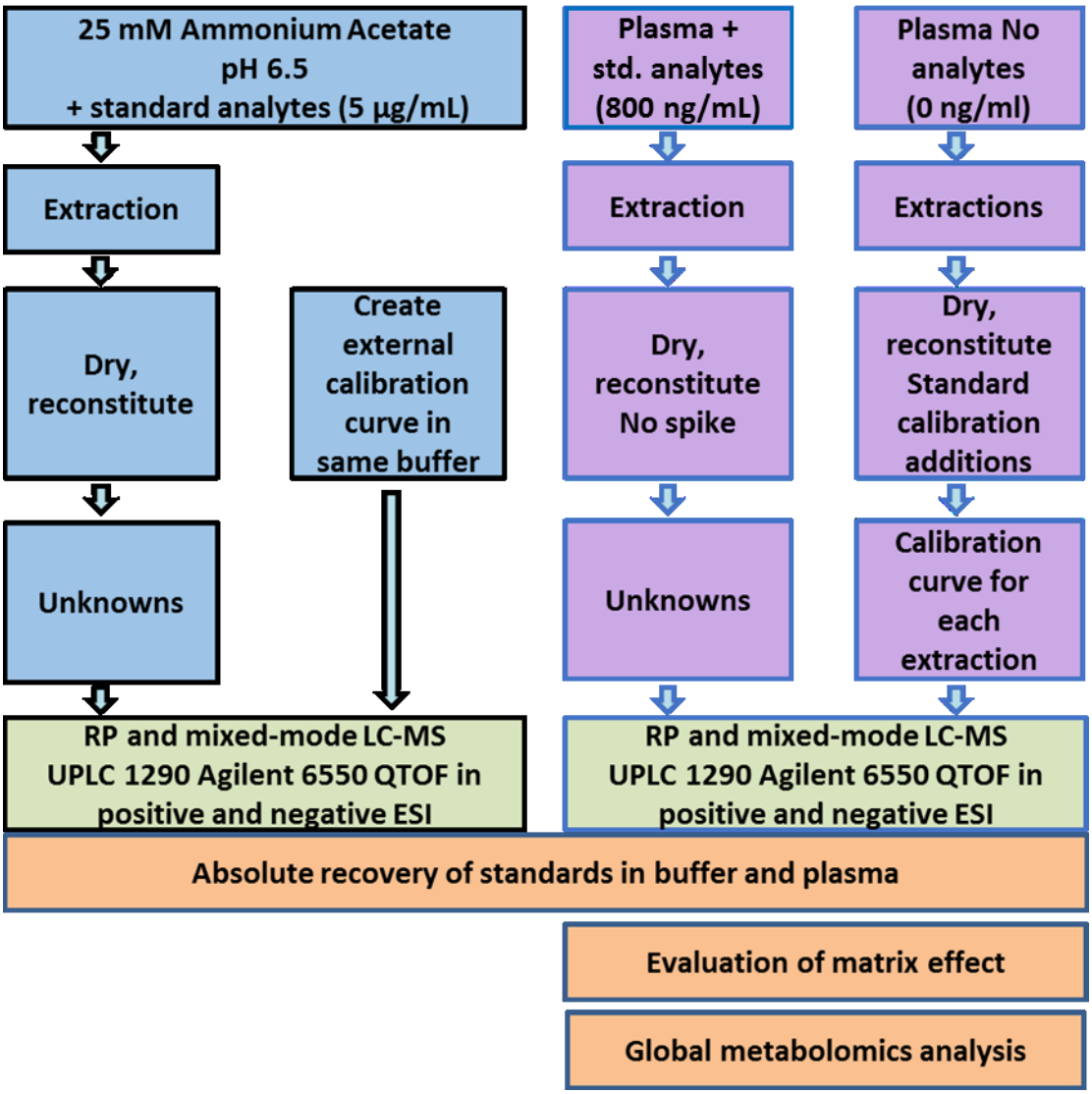
Analytes	RP (+ve) ESI							Scherzo (+ve) ESI						
	Methanol-ethanol	Methanol-MTBE	Methanol	MTBE	IEX	PEP2	C18	Methanol-ethanol	Methanol-MTBE	Methanol	MTBE	IEX	PEP2	C18
Folic acid	95.4	108.2	109.1	106.4	109.9	114.4	111.9	372.8	344.0	400.6	100.5	119.5	***NDM***	***NDM***
Adenine	42.4	43.3	52.8	106.2	87.2	176.9	121.9	77.6	62.8	68.9	101.3	57.9	97.6	103.5
Histamine	198.9	196.1	168.7	88.4	***SS***	129.9	110.0	99.9	98.9	75.5	82.3	36.4	84.3	93.9
4-aminobutanoic acid (d_6)	6.2	12.6	4.3	89.4	0.3	62.1	55.2	0.8	1.7	0.7	41.3	***NDM***	30.3	60.1
4-aminobutanoic acid	0.7	0.7	0.4	71.2	***NDM***	24.8	40.9	0.2	0.2	***NDM***	75.8	***NDM***	29.1	59.6
Pantothenic Acid	***SS***	***SS***	***SS***	77.5	***SS***	26.7	46.6	209.0	213.7	205.9	198.9	***SS***	***SS***	***SS***
Thyrotropin-releasing hormone	97.5	96.1	95.5	101.0	101.2	106.6	101.5	80.6	97.2	86.3	99.3	450.2	100.3	100.0
Tyrosine	***SS***	***SS***	***SS***	105.2	***SS***	92.5	82.9	***NDM***	***NDM***	***NDM***	101.2	***NDM***	99.6	91.0
Homovanillic acid	72.7	43.2	32.5	91.8	17.0	50.0	116.6	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***
Homovanillic acid (d3)	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Neurotensin	***SS***	***SS***	***SS***	***SS***	148.6	152.7	104.8	2578.3	2750.1	3976.0	2542.4	346.6	250.5	138.7
Kynurenine	***SS***	***SS***	***SS***	104.2	***SS***	***SS***	101.7	***NDM***	***NDM***	***NDM***	112.7	115.9	199.8	47.9
Melatonin (d_4)	45.3	46.1	50.0	50.7	43.8	47.5	45.9	161.7	171.0	157.3	46.6	140.2	35.0	35.3
Melatonin	45.2	45.9	47.4	48.3	43.9	45.4	45.7	158.8	168.0	155.7	45.3	172.3	43.8	44.8
5-Methoxytryptamine	88.4	92.7	83.8	100.7	86.9	93.9	91.3	57.4	65.7	52.6	92.7	107.3	92.3	97.6
Cortisol	95.1	105.1	93.8	111.2	92.1	99.9	99.9	442.4	742.5	240.8	481.7	1433.7	107.3	106.0
Cortisol (d4)	93.8	100.6	95.7	108.9	92.5	87.8	94.2	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***
Cortisone	87.4	99.7	89.4	111.6	93.7	99.3	98.6	322.6	575.1	244.8	495.4	2391.1	104.7	104.2
Triiodothyronine	215.8	270.8	223.4	253.5	51.4	53.4	94.4	123.5	124.1	123.9	102.0	114.3	44.8	96.2
Thyroxine ($^{13}C_6$)	219.3	295.1	187.1	261.0	44.3	47.4	98.6	206.3	216.3	519.5	201.7	43.6	40.4	100.9
Thyroxine	229.7	296.7	179.2	257.5	42.5	47.2	98.1	202.5	217.2	471.9	193.6	43.2	41.6	101.8
Cholic acid (d_4)	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Cholic acid	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Sphingosine	745.4	825.8	618.1	695.2	352.1	713.4	1677.1	1075.2	1102.0	962.4	313.0	265.1	989.7	3217.0

Supplementary Table A9. Summary of matrix effects observed for all metabolites across all extraction methods and Scherzo LC-MS analyses. Not all analytes were successfully detected in buffer (NDB) or in an extract (NDM) which made the calculation of matrix effect impossible as specified. Analytes which did not respond to increased concentrations of spiked standards or which exhibited negative signal response were considered to be suppressed or saturated (SS) and were not used for calculation of matrix effects.

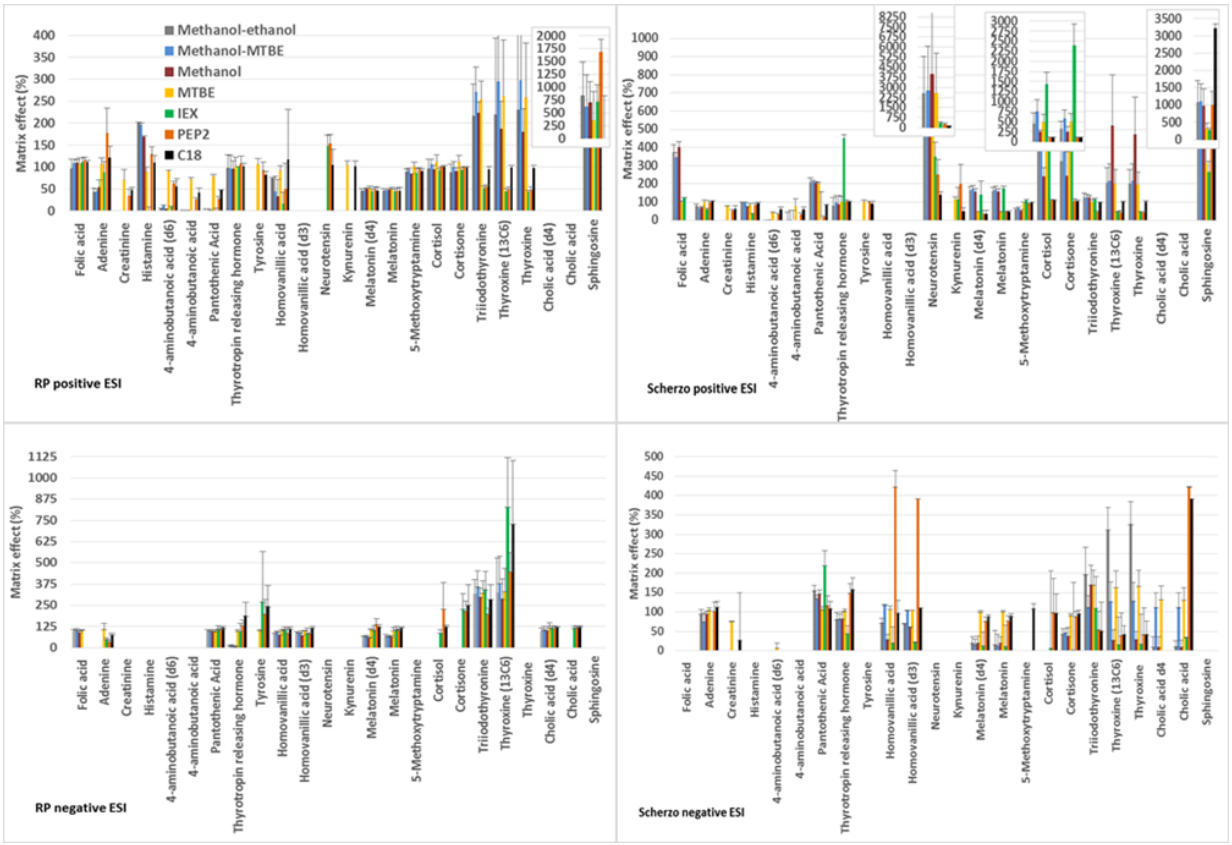
Analytes	RP (-ve) ESI							Scherzo (-ve) ESI						
	Methanol-ethanol	Methanol-MTBE	Methanol	MTBE	IEX	PEP2	C18	Methanol-ethanol	Methanol-MTBE	Methanol	MTBE	IEX	PEP2	C18
Folic acid	96.6	102.5	90.1	102.8	***NDM***	***NDM***	***NDM***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Adenine	***NDM***	***NDM***	***NDM***	108.1	47.2	32.9	75.4	95.0	74.4	94.7	104.2	***NDM***	102.5	112.7
Histamine	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
4-aminobutanoic acid (d_6)	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***	***NDM***	***NDM***	***NDM***	8.7	***NDM***	***NDM***	***NDM***
4-aminobutanoic acid	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***	***SS***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Pantothenic Acid	94.2	95.9	92.0	97.6	108.7	106.4	115.0	155.3	134.5	147.0	105.6	219.7	116.8	108.8
Thyrotropin releasing hormone	8.1	10.2	7.4	97.7	96.5	129.6	187.3	80.7	82.3	82.0	103.3	44.3	149.3	158.2
Tyrosine	***SS***	***SS***	***SS***	100.9	269.4	197.6	242.5	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***
Homovanillic acid	81.7	91.5	75.6	101.2	107.0	85.7	110.3	71.5	118.1	29.1	106.7	20.5	422.1	97.5
Homovanillic acid (d_3)	87.6	87.6	72.5	101.1	81.2	85.7	116.4	70.6	104.4	61.0	105.3	22.7	392.2	111.2
Neurotensin	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Kynurenine	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***
Melatonin (d_4)	63.3	67.7	60.2	105.6	101.5	136.4	121.8	19.3	15.5	19.9	99.5	12.9	75.8	88.3
Melatonin	63.1	64.8	62.1	100.2	106.5	103.4	116.9	15.9	13.2	19.9	100.6	11.8	76.9	89.2
5-Methoxytryptamine	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	109.3
Cortisol	***NDM***	***NDM***	***NDM***	***NDM***	86.5	227.1	123.6	***NDM***	***NDM***	***NDM***	***NDM***	6.1	98.3	97.5
Cortisol (d_4)	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***	***NDM***
Cortisone	***NDM***	***NDM***	***NDM***	***NDM***	225.9	214.8	250.8	42.8	46.8	38.1	90.9	2.8	88.3	95.9
Triiodothyronine	315.4	359.2	297.3	320.6	341.6	198.8	286.2	197.0	112.4	170.4	169.9	109.6	54.8	51.5
Thyroxine ($^{13}C_6$)	322.5	377.5	286.4	330.7	828.3	447.4	728.4	312.0	127.1	27.4	163.1	16.3	39.4	42.3
Thyroxine	***NDM***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	326.2	128.0	29.6	167.2	17.8	42.3	42.5
Cholic acid (d_4)	94.9	105.0	99.0	116.4	114.5	112.1	119.0	9.9	112.3	9.6	131.9	***NDB***	***NDB***	***NDB***
Cholic acid	***NDB***	***NDB***	***NDB***	***NDB***	115.0	116.5	121.3	10.8	112.4	10.2	130.8	34.7	422.1	392.2
Sphingosine	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***	***NDB***

Supplementary Table A10. Usage and fate of standard analytes in targeted quantitation experiments in buffer and plasma. NA-not available at the time of analysis; ND – not detected. Inability to detect PE and PI phospholipids in plasma and PC in buffer and plasma is most likely is due to a high suppression effect at the end of chromatogram caused by LC-MS methodology which was not optimized for lipid metabolites.

Analytes	Detection in extractions from buffer	Detection in extractions from plasma
4-aminobutanoic acid (d_6)	NA	+
Angiotensin II	+	Replaced by Neurotensin
Cholic acid	NA	Added to supplement analysis in (-ve) ESI
Cholic acid (d_4)	NA	Added to supplement analysis in (-ve) ESI
Cortisol (d_4)	NA	+
Creatinine	+	Excluded (too high endogenous concentration)
Dopamine	+	ND
Dopamine (d_4)	NA	ND
Folic acid	ND	+
Glutamic acid	+	Excluded (too high endogenous concentration)
Homovanillic acid (d_3)	NA	Added to supplement analysis in (-ve) ESI
Kynurenine	NA	+
Melatonin (d_4)	NA	+
Neurotensin	NA	Added to represent large peptides
Phenylalanine (d_5)	+	Used as internal reference
PI (18:3/22:4)	+	ND
PE (17:0/17/0)	+	ND
Serotonin	+	ND
Thyroxine ($^{13}C_6$)	NA	+



Supplementary Figure A1. Study design for the targeted and global metabolomics analysis of extraction methods. Blue and purple colors designate experiments in buffer and blood plasma, respectively. Green color designates LC-MS analyses step executed on RP (reversed-phase) and mixed mode (weak anion/weak cation/RP) mode columns. Peach color designates major data analysis blocks.



*Supplementary Figure A2. Matrix effects observed in human plasma in all LC-MS analysis across all analytes and extraction methods. On Y axis, graph displays area ratios between analytes of the same concentrations in buffer and a matrix calibration points ((area in spiked matrix-area in blank matrix) / (area in buffer-area in blank)*100%)) assuming the area of analyte in the buffer as 100%. Analytes are displayed along X-axis from left to right according to the increase of their predicted ACD/LogP, ACD Laboratories Toronto, Canada). Legend designating extraction methods on RP positive panel is same for all panels. Error bars represent relative standard deviation of matrix effects between different concentration levels tested. Due to very high matrix effect of some analytes in RP and Scherzo in (+ve) ESI, full scale inserts were placed. Missing bars indicate that matrix effects could not be calculated due to one of three possible events: (i) analyte is not detected in buffer, (ii) analyte is not detected in a particular matrix or (iii) analyte signal was saturated (signal did not increase together with increased concentration at the concentration levels tested). The differences in matrix effect between 4-aminobutanoic acid and its deuterated analog in RP (+ve) ESI (NDM) and Scherzo in (+ve) ESI (methanol) is most likely due to the strong suppressing matrix effect. The differences in matrix effects between cholic acid, thyroxine and their deuterated analogs in RP in (-ve) ESI are most likely due to the error in preparation of calibration points in the correspondent samples where it was not detected. Cortisol (d_4) was removed from the graph because it was not detected in neither analysis except RP in (+ve) ESI. See Appendix A. Table 5 for details.*

Appendix B

Supplementary information for Chapter 3.

Supplementary Table B1. Standard metabolites used in the analysis of single sSPE fractions obtained in the initial and final protocols of sequential sample preparation ***(Sections 2.3 and 2.4, respectively)*** *included into the standard mix spiked before sSPE (IR) or detected (ER) in LC-MS data files using known masses and RT retrieved from parallel LC-MS analysis runs of standards in solvents. Standards that were not detected are designated by "ND". The ionic class of standards was assumed from the presence of positive (C), negative (A), no charge (N), or both charges (Z) at physiological pH (7.2) based on corresponding pKa values predicted by ACD Labs (ON, Canada) and obtained from ChemSpider Database (https://www.chemspider.com) in February of 2019. Carnitines are designated as Z due to the presence of both charges, negative from the carboxyl group, and pH-independent, the positive charge of the quaternary amine. The presence of quaternary amine determines choline designation as C.*

Metabolite	Type of reference	Class (physiological pH)	Formula	Monoisotopic mass	RT (min) RP	RT (min) ZIC-HILIC	pKa strongest acid	pKa strongest base	LogP (octanol/water)	Stock standard solvent
3-hydroxy-DL-kynurenine	IR	Z	$C_{10}H_{12}N_2O_4$	224.0797	5.4	14.4	1.0	8.9	1.1	water
4-aminobenzoic acid	IR	A	$C_7H_7NO_2$	137.0477	ND	ND	4.8	2.7	0.8	water
4-aminobutanoic acid	ER	A	$C_4H_9NO_2$	103.0633	ND	ND	4.5	1.0	-2.9	water
5-hydroxy-indoleacetic acid	IR	A	$C_{10}H_9NO_3$	191.0582	ND	8.5	4.2	-5.5	0.3	water
5-hydroxy-L-tryptophan	IR	Z	$C_{11}H_{12}N_2O_3$	220.0848	2.2	16.1	2.2	9.2	-1.4	water
5-methoxy-tryptamine	IR	C	$C_{11}H_{14}N_2O$	190.1106	ND	9.8	ND	9.8	1.3	MeOH
7,8-dihydro-L-biopterin	IR	N	$C_9H_{13}N_5O_3$	239.1018	1.2	11.3	13.0	3.7	-2.0	water
Acetyl-DL-carnitine	IR	Z	$C_9H_{17}NO_4$	203.1158	ND	ND	4.1	-7.1	-4.5	water
Adenine	IR	N	$C_5H_5N_5$	135.0545	3.9	ND	10.3	3.7	-0.5	water
Adenosine	IR	N	$C_{10}H_{13}N_5O_4$	267.0968	3.9	4.6	12.5	3.9	-2.1	water
Adenosine monophosphate	ER	A	$C_{10}H_{14}N_5O_7P$	347.0631	ND	14.3	1.2	3.2	-4.7	water
Adipic acid	IR	A	$C_6H_{10}O_4$	146.0579	7.4	ND	3.4	Nav	0.5	20% MeOH
Alanine	ER	Z	$C_3H_7NO_2$	89.0476	ND	15.7	2.5	9.5	-2.8	water
Anthranilic acid	ER	A	$C_7H_7NO_2$	137.0477	ND	ND	4.9	2.0	0.9	MeOH
Arachidonic acid	IR	A	$C_{20}H_{32}O_2$	304.2402	20.1	1.2	4.8	Nav	6.6	MeOH
Arginine	ER	Z	$C_6H_{14}N_4O_2$	174.1117	ND	27.1	2.4	12.4	-3.2	water
Asparagine	ER	Z	$C_4H_8N_2O_3$	132.0526	ND	20.0	2.0	8.4	-4.3	water
Aspartic acid	ER	Z	$C_4H_7NO_4$	133.0379	ND	22.1	1.7	9.6	-3.5	water

Supplementary Table B1 (cont'd).

Metabolite	Type of reference	Class (physiological pH)	Formula	Monoisotopic mass	RT (min) RP	RT (min) ZIC-HILIC	pKa strongest acid	pKa strongest base	LogP (octanol/water)	Stock standard solvent
Bilirubin	ER	A	$C_{33}H_{36}N_4O_6$	584.2629	ND	ND	4.0	-2.8	3.1	water
Biotin	IR	A	$C_{10}H_{16}N_2O_3S$	244.0882	6.4	11.6	4.4	-1.9	0.3	MeOH
Butyryl-L-carnitine	IR	Z	$C_{11}H_{21}NO_4$	231.1471	4.0	13.5	4.3	-7.1	-3.3	MeOH
Carnitine	IR	Z	$C_7H_{15}NO_3$	161.1052	0.9	21.1	4.2	-7.1	-4.9	water
Cholesterol	ER	N	$C_{27}H_{46}O$	386.3549	ND	2.9	18.2	-1.4	7.1	MeOH
Cholesteryl acetate	IR	N	$C_{29}H_{48}O_2$	428.3654	5.7	ND	Nav	-7.0	7.4	20% MeOH
Cholic acid	IR	A	$C_{24}H_{40}O_5$	408.2876	12.4	5.7	4.5	-0.2	2.5	20% MeOH
Choline	ER	C	$C_5H_{14}NO$	104.1075	ND	ND	14.0	14.0	-4.7	water
Citrulline	ER	Z	$C_6H_{13}N_3O_3$	175.0965	0.7	21.6	2.3	9.2	-3.9	water
Coenzyme Q10	ER	N	$C_{59}H_{90}O_4$	862.6839	ND	ND	Nav	-4.7	17.2	MeOH
Cortisol	IR	N	$C_{21}H_{30}O_5$	362.2093	9.6	1.0	12.8	-1.6	1.8	MeOH
Cortisone	IR	N	$C_{21}H_{28}O_5$	360.1937	9.7	1.0	12.6	-3.3	1.7	MeOH
Creatinine	ER	N	$C_4H_7N_3O$	113.0587	0.8	5.6	9.2	5.0	-1.1	water
Decanoyl-L-carnitine	IR	Z	$C_{17}H_{33}NO_4$	315.2410	10.6	10.8	4.2	-7.1	2.9	MeOH
Deoxycholic acid	IR	A	$C_{24}H_{40}O_4$	392.2927	14.8	ND	4.7	-0.4	3.8	MeOH
Diosmin	IR	A	$C_{28}H_{32}O_{15}$	608.1741	8.0	7.2	8.5	-3.6	-0.4	EtOH
Docosahexaenoic acid	IR	A	$C_{22}H_{32}O_2$	328.4883	19.8	ND	4.9	Nav	6.8	MeOH
Eicosapentaenoic acid	IR	A	$C_{20}H_{30}O_2$	302.2246	19.1	ND	4.8	Nav	6.2	water
Folic acid	IR	A	$C_{19}H_{19}N_7O_6$	441.1397	5.5	ND	3.4	2.1	-0.7	20% MeOH
Glutamine	ER	Z	$C_5H_{10}N_2O_3$	146.0674	ND	19.9	2.2	9.3	-4.0	water
Glutathione oxidized	IR	Z	$C_{20}H_{32}N_6O_{12}S_2$	612.1520	ND	18.6	1.4	9.6	-10.0	water
Glycine	ER	Z	$C_2H_5NO_2$	75.0324	0.6	19.0	2.3	9.2	-3.4	MeOH
Glycocholic acid	IR	A	$C_{26}H_{43}NO_6$	465.3090	11.1	10.0	3.7	-0.1	1.4	water
Guanosine	IR	N	$C_{10}H_{13}N_5O_5$	283.0917	0.7	12.0	10.2	0.5	-2.7	water
Guanosine 3'-phosphate	ER	A	$C_{10}H_{14}N_5O_8P$	363.0580	ND	ND	1.3	-0.4	-3.1	water
Guanosine 5'-phosphate	IR	A	$C_{10}H_{14}N_5O_8P$	363.0580	ND	ND	1.1	0.4	-3.1	MeOH
Hexanoyl-L-carnitine	IR	Z	$C_{13}H_{25}NO_4$	259.1783	6.8	11.8	4.2	-7.1	ND	MeOH
Histamine	ER	C	$C_5H_9N_3$	111.0797	ND	24.4	14.5	9.6	-0.7	water
Homovanillic acid	IR	A	$C_9H_{10}O_4$	182.0579	6.2	ND	3.7	-4.9	1.2	water
Kynurenic acid	IR	A	$C_{10}H_7NO_3$	189.0426	4.9	5.7	3.2	-4.4	1.6	water
Kynurenine	IR	Z	$C_{10}H_{12}N_2O_3$	208.0848	2.6	9.2	1.2	9.0	-1.9	water
Lysine	ER	Z	$C_6H_{14}N_2O_2$	146.1055	ND	ND	2.7	10.3	-3.2	water
Maleic acid	IR	A	$C_4H_4O_4$	116.0110	ND	ND	3.1	Nav	0.0	water

Supplementary Table B1 (cont'd).

Metabolite	Type of reference	Class (physiological pH)	Formula	Monoisotopic mass	RT (min) RP	RT (min) ZIC-HILIC	pKa strongest acid	pKa strongest base	LogP (octanol/water)	Stock standard solvent
Methionine	ER	Z	$C_5H_{11}NO_2S$	149.0514	ND	13.1	2.3	9.2	-1.9	water
Melatonin	IR	N	$C_{13}H_{16}N_2O_2$	232.1212	8.2	1.0	15.8	-1.6	1.2	MeOH
Melatonin (d_4)	IS	N	$C_{13}H_{12}D_4N_2O_2$	236.1485	8.2	1.0	15.8	-1.6	1.2	MeOH
Myo-Inositol	ER	N	$C_6H_{12}O_6$	180.0634	ND	14.3	12.3	-3.6	-3.8	water
Myoinositol 1-phosphate	ER	A	$C_6H_{13}O_9P$	260.0297	ND	ND	1.2	-3.6	-3.9	water
N-Acetyl-L-cysteine	IR	A	$C_5H_9NO_3S$	163.1950	ND	ND	3.8	-1.5	-0.7	water
Neopterin	IR	N	$C_9H_{11}N_5O_4$	253.0811	1.0	15.3	10.0	1.0	-2.8	DMSO
Nervonic acid	ER	A	$C_{24}H_{46}O_2$	366.3498	ND	ND	5.0	Nav	9.5	MeOH
NAD	ER	Z	$C_{21}H_{27}N_7O_{14}P_2$	663.4300	ND	22.4	1.9	4.0	-4.9	water
Norepinephrine	IR	C	$C_8H_{11}NO_3$	169.0739	1.1	2.3	9.5	8.9	-0.9	water
Octanoyl-L-carnitine	IR	Z	$C_{15}H_{29}NO_4$	287.2097	9.0	11.1	4.2	-7.1	n/av	MeOH
Pantothenic acid	IR	A	$C_9H_{17}NO_5$	219.1107	4.2	ND	4.4	-2.8	-1.4	20% MeOH
Phenylalanine	IR	Z	$C_9H_{11}NO_2$	165.0790	2.5	ND	2.5	9.5	-1.2	20% MeOH
Pregnenolone	IR	N	$C_{21}H_{32}O_2$	316.2402	ND	ND	18.2	-1.4	3.6	20% MeOH
Proline	ER	Z	$C_5H_9NO_2$	115.0632	ND	15.6	1.9	11.3	-2.6	water
Propionyl-L-carnitine	IR	Z	$C_{10}H_{19}NO_4$	217.1314	1.6	ND	4.2	-7.1	-3.6	MeOH
Pyridoxal	IR	A	$C_8H_9NO_3$	167.0582	1.1	5.7	8.0	4.1	0.2	water
Pyridoxamine	IR	Z	$C_8H_{12}N_2O_2$	168.0899	1.1	ND	7.8	9.8	-1.6	water
Pyridoxine	IR	N	$C_8H_{11}NO_3$	169.0739	2.6	ND	9.4	5.6	-1.0	water
Resveratrol	IR	A	$C_{14}H_{12}O_3$	228.0786	8.9	2.2	8.5	-6.2	3.4	MeOH
Riboflavin	IR	A	$C_{17}H_{20}N_4O_6$	376.1383	6.3	5.9	6.0	-2.6	-1.0	20% MeOH
Saccharopine	IR	Z	$C_{11}H_{20}N_2O_6$	276.2863	ND	23.6	1.4	10.9	-5.4	20% MeOH
S-adenosylhomocysteine	ER	Z	$C_{14}H_{20}N_6O_5S$	384.1214	1.3	18.5	1.8	9.5	-4.0	water
S-adenosylmethionine	ER	Z	$C_{15}H_{23}N_6O_5S$	399.1451	ND	ND	1.7	9.4	-5.3	water
Serine	ER	Z	$C_3H_7NO_3$	105.0426	ND	21.1	2.0	8.9	-3.9	water
Serotonin	ER	C	$C_{10}H_{12}N_2O$	176.0950	ND	ND	9.3	10.0	0.5	water
S-methyl-L-cysteine	ER	Z	$C_4H_9NO_2S$	135.0359	ND	14.1	2.4	9.2	-2.2	water
Spermidine	IR	C	$C_7H_{19}N_3$	145.1579	ND	ND	Nav	10.7	-1.0	water
Spermine	IR	C	$C_{10}H_{26}N_4$	202.2157	ND	ND	Nav	10.8	-1.5	water
Stearoyl-carnitine	ER	Z	$C_{25}H_{49}NO_4$	427.3662	ND	10.1	4.2	-7.1	2.9	MeOH

Supplementary Table B1 (cont'd).

Metabolite	Type of reference	Class (physiological (pH)	Formula	Monoisotopic mass	RT (min) RP	RT (min) ZIC-HILIC	pKa strongest acid	pKa strongest base	LogP (octanol/water)	Stock standard solvent
Taurocholic acid	IR	A	$C_{26}H_{45}NO_7S$	515.2917	8.9	ND	-1.1	0.3	-0.2	MeOH
Taurodeoxycholic acid	IR	A	$C_{26}H_{44}NNaO_6S$	521.2787	ND	ND	-0.8	-0.2	1.1	MeOH
Threonine	ER	Z	$C_4H_9NO_3$	119.0582	ND	18.8	2.6	10.4	-2.9	water
Triiodothyronine	IR	Z	$C_{15}H_{12}I_3NO_4$	650.7900	9.9	8.7	0.3	9.5	2.8	MeOH
Thymine	IR	N	$C_5H_6N_2O_2$	126.0429	1.8	7.8	10.0	-5.0	-0.5	water
TRH	IR	C	$C_{16}H_{22}N_6O_4$	362.1703	1.4	9.6	11.2	6.7	-3.3	water
Thyroxine	IR	Z	$C_{15}H_{11}I_4NO_4$	776.6867	10.7	8.1	0.3	9.4	0.7	MeOH
Trans-4-hydroxy-proline	ER	Z	$C_5H_9NO_3$	131.0582	ND	18.4	1.6	10.6	-3.7	water
Tryptamine	ER	C	$C_{10}H_{12}N_2$	160.1000	ND	ND	17.2	9.7	1.5	water
Valine	ER	Z	$C_5H_{11}NO_2$	117.0790	0.7	14.1	2.7	9.6	-2.0	water
α-Ketoglutaric acid	IR	A	$C_5H_6O_5$	146.0215	1.5	ND	2.7	-9.7	-0.1	water
β-estradiol	IR	N	$C_{18}H_{24}O_2$	272.1776	9.6	6.1	10.3	-0.9	3.8	MeOH

Supplementary Table B2. List of lipid standards used to evaluate the lipid content of fractions including labeled with stable isotopes designated by "dX", where X indicates the number of deuterium atoms) using the LC-MS lipidomics method described in 2.4.6. Standards were prepared in 100% MeOH at 10 µg/mL.

Name	Monoisotopic mass	Formula	RT (+ESI) (min)	RT (-ESI) (min)
Cer (d18:1/18:0)	565.5434	$C_{36}H_{71}NO_3$	24.2	24.2
DG (16:1/0:0/16:1) (d_5)	569.5068	$C_{35}H_{59}D_5O_5$	25.5	ND
DG (18:0/0:0/18:0) (d_5)	629.6007	$C_{39}H_{71}D_5O_5$	27.4	ND
DG (18:1/0:0/18:1) (d_5)	625.5694	$C_{39}H_{67}D_5O_5$	27.4	ND
DG (18:2/0:0/18:2) (d_5)	621.5381	$C_{39}H_{63}D_5O_5$	26.0	ND
LPC (17:0)	509.3481	$C_{25}H_{52}NO_7P$	12.8	12.8
LPC (17:1)	507.3325	$C_{25}H_{50}NO_7P$	ND	11.0
LPS (17:1)	509.2754	$C_{23}H_{44}NO_9P$	9.3	6.7
Lyso-SM (d17:1)	450.3223	$C_{22}H_{47}N_2O_5P$	ND	20.2
PA (18:0/18:0)	704.5356	$C_{39}H_{77}O_8P$	ND	17.4
PC (17:0/14:1)	717.5309	$C_{39}H_{76}NO_8P$	ND	21.2
PC (19:0/19:0)	817.6561	$C_{46}H_{92}NO_8P$	26.2	26.6
PC (17:0/17:0)	761.5935	$C_{42}H_{84}NO_8P$	24.1	24.0
PE (12:0/13:0)	593.4057	$C_{30}H_{60}NO_8P$	17.8	17.9
PE (17:0/17:0)	719.5465	$C_{39}H_{78}NO_8P$	24.2	24.2
PG (17:0/17:0)	750.5411	$C_{40}H_{79}O_{10}P$	22.1	16.5
PG (18:0/18:0)	778.5724	$C_{42}H_{83}O_{10}P$	23.3	18.0
PI (17:0/14:1)	794.4945	$C_{40}H_{75}O_{13}P$	ND	ND
PS (17:0/17:0)	763.5363	$C_{40}H_{78}NO_{10}P$	21.9	18.8
SM (d18:1/12:0)	646.5050	$C_{35}H_{71}N_2O_6P$	26.4	ND
TG (17:0/17:0/17:0) (d_5)	851.7990	$C_{54}H_{97}D_5O_6$	34.6	ND

Supplementary Table B3. LC gradient used in lipidomic LC-MS analysis.

Time (min)	%A	%B
0.0	80	20
2.0	80	20
4.0	70	30
25.0	20	80
35.0	15	85
38.0	5	95
41.0	5	95
41.1	80	20
50.1	80	20

Supplementary Table B4. Parameters of metabolites used in the analysis of single fractions of sSPE. Fractionation was executed identically to the final protocol ***(Section 3.2.4).*** *Values were obtained from Chemicalize™ calculator (https://chemicalize.com/#/calculation) available on the website of ChemAxon (Budapest, Hungary) in December of 2019.*

Standard	Hydrogen bond donor/acceptor	LogP o/w	LogD pH2.4	LogD pH9.2	Standard	Hydrogen bond donor/acceptor	LogP o/w	LogD pH 2.4	LogD pH 9.2
Taurocholic acid	5/7	-0.2	-0.9	0.0	Glutathione oxidized	6/7	-10	-4.9	-8.6
α-Ketoglutaric acid	2/5	-0.11	-0.3	-7.2	Proline	2/3	-2.6	0.3	0.2
Kynurenic acid	2/4	1.6	1.3	-1.7	5-Hydroxy-L-tryptophan	4/4	-1.4	-1.5	-1.7
Adipic Acid	2/4	0.49	0.5	-6.5	Glycine	2/3	-3.4	-3.6	-3.6
Glycocholic acid	5/6	1.4	1.4	-2.1	Arginine	5/6	-3.2	-6.4	-3.4
5-hydroxyindoleacetic acid	3/3	0.26	1.4	-2.2	S-methyl-L-cysteine	2/3	-2.2	-2.7	-2.7
Cholic acid	4/5	2.5	2.5	-1.0	Alanine	2/3	-2.8	-3.1	-3.0
Deoxycholic acid	3/4	3.8	3.8	0.3	Phenylalanine	2/3	-1.2	-1.4	-1.4
Docosahexaenoic Acid	1/2	6.8	6.8	3.3	Valine	2/3	-2	-2.3	-2.0
Homovanillic acid	2/4	1.2	1.1	-2.4	Trans-4-hydroxy-proline	3/4	-3.7	-3.8	-3.7
Pantothenic acid	4/5	-1.4	0.0	-1.0	Methionine	2/3	-1.9	-2.4	-2.5
Biotin	3/3	0.3	0.0	-1.0	S-adenosylhomocysteine	5/10	-4	-5.5	-4.2
Arachidonic Acid	1/2	6.6	6.6	3.1	Serine	3/4	-3.9	-4.0	-4.3
Eicosapentaenoic Acid	1/2	6.2	6.2	2.7	Asparagine	3/4	-4.3	-4.4	-5.0
Pyridoxal	2/4	0.2	-1.1	-1.1	Citrulline	4/4	-3.9	-4.1	-4.2
Cholesteryl acetate	0/1	7.6	7.6	7.6	Aspartic acid	3/5	-3.5	-3.5	-7.0
Carnitine	1/3	-4.9	-4.8	-4.1	Guanosine	5/8	-2.7	-2.7	-2.7
Decanoyl-L-carnitine	1/3	-0.85	-0.6	0.1	Riboflavin	5/9	-1	-0.9	-2.8
Hexanoyl-L-carnitine	1/3	-0.6	-2.4	-1.6	Resveratrol	3/3	3.4	3.4	2.3
Octanoyl-L-carnitine	1/3	-1.5	-1.5	-0.7	Diosmin	8/15	-0.4	-0.4	-2.3
Propionyl-L-carnitine	1/3	-3.75	-3.7	-2.3	Thymine	2/2	-0.5	-0.5	-0.8
Stearoyl-carnitine	1/3	-3.7	2.9	3.7	β-estradiol	2/2	3.75	3.7	3.7
Butyryl-L-carnitine	1/3	-3.7	-3.7	-2.3	Myo-Inositol	6/6	-3.8	-3.8	-3.8
Creatinine	1/4	-1.1	-3.4	-1.6	Cortisone	2/5	1.7	1.7	1.7
Pyridoxine	3/4	-1	-2.4	-1.2	Cortisol	3/5	1.8	1.3	1.3
Norepinephrine	4/4	-0.9	-3.2	-0.7	Melatonin	2/2	1.2	1.2	1.2
Neopterin	6/9	-2.8	-2.6	-2.2	Cholesterol	1/1	7.1	7.1	7.1
Adenine	2/4	-0.53	-2.1	-0.6	Triiodothyronine	3/4	2.8	2.8	1.8
TRH	4/5	-3.3	-3.9	-3.3	Thyroxine	3/4	0.7	3.7	1.9
Adenosine	4/8	-2.09	-3.5	-2.1	Saccharopine	5/8	-5.4	-5.6	-9.0
Histamine	2/2	-0.7	-4.9	-1.1	Glutamine	3/4	-4	-4.2	-4.2
Kynurenine	3/5	-1.9	-2.8	-2.3	Folic acid	6/12	-0.7	-0.7	-7.2

Supplementary Table B5. The split of metabolites between sSPE fractions for RP and ZIC-HILIC analysis in positive (+ve) and negative (-ve) ESI. Colors designate the number of metabolites, which were split between (detected in more than one) sSPE fractions to facilitate visualization. The intense red indicates the highest, while the dark green indicates the lowest number of split metabolites in the correspondent LC-MS method.

(+ve) RP	A	C	N	Z	(+ve) ZIC-HILIC	A	C	N	Z
A	640	47	176	290	**A**	189	121	168	139
C		370	313	134	**C**		371	532	319
N			232	69	**N**			209	215
Z				149	**Z**				89
(-ve) RP	**A**	**C**	**N**	**Z**	**(-ve) ZIC-HILIC**	**A**	**C**	**N**	**Z**
A	944	53	347	297	**A**	365	92	214	161
C		284	177	199	**C**		167	392	250
N			183	104	**N**			303	200
Z				115	**Z**				64

*Supplementary Table B6. Analysis of the metabolite split between individual sSPE fractions in theoretical combinations, which were proposed for LC-MS analysis of combined fractions (combinations designated by fraction letters and "&" symbol) at different loading amounts, which expressed in µL of original plasma. The table demonstrates the expected percentage (% to total) of split metabolites to the nonredundant sum of metabolites in both fractions (100%). The expected number of split metabolites in proposed fraction combinations is also displayed (two columns on the right). The number of redundant metabolites was obtained from Venn analysis (**Section 2.4.9**).*

Sample load (equivalent to µL of original plasma)	**Split metabolites (% to total)**		**Split metabolites (numbers)**	
RP (+ve) ESI	**C & N**	**A & Z**	**C & N**	**A & Z**
3.2	28.7	29.4	313	389
6.4	32.5	23.9	375	317
12.8	35.2	29.6	445	418
RP (-ve) ESI	**C & Z**	**A & N**	**C & Z**	**A & N**
3.2	20.1	18.3	199	347
6.4	21.9	16.6	269	302
12.8	23.9	25.4	339	592
ZIC-HILIC (+ve) ESI	**C & N**	**A & Z**	**C & N**	**A & Z**
3.2	38.1	16.5	532	139
6.4	37.1	16.5	526	141
12.8	38.3	19.0	545	171
ZIC-HILIC (-ve) ESI	**C & N**	**A & Z**	**C & N**	**A & Z**
3.2	35.4	17.0	392	161
6.4	32.8	19.7	356	199
12.8	37.9	21.1	454	220

*Supplementary Table B7. Concentrations of the standards which were used for the analysis of combined sSPE fractions in the final protocol (**Section 2.5.2**). Atoms with stable isotopes are designated by"15NX"for nitrogen -15, "13CX" for carbon -13, and "DX" for deuterium, where the subscript "X" indicates the number of atoms if more than one. TRH stands for thyrotropin-releasing hormone. Concentrations at LC-MS steps are calculated based on the assumption of 100% recovery of standard. Concentrations of spiked metabolites do not include endogenous levels. Amino acids from the kit are designated by the "SIL" positioned near the name of an acid.*

Standards	Formula	Monoisotopic Mass	Conc. of spiking mix (ng/mL)	Conc. in sample prior to MeOH extraction (ng/mL)	Conc. in SPE samples at LC-MS (ng/mL)	Conc. in MeOH at LC-MS (ng/mL)
3-Hydroxy-DL-kynurenine	$C_{10}H_{12}N_2O_4$	224.0797	910	70.0	80.9	20.2
5-hydroxy-indoleacetic acid	$C_{10}H_9NO_3$	191.0582	910	70.0	80.9	20.2
5-hydroxy-L-tryptophan	$C_{11}H_{12}N_2O_3$	220.0848	910	70.0	80.9	20.2
5-methoxy-tryptamine	$C_{11}H_{14}N_2O$	190.1106	910	70.0	80.9	20.2
7,8-dihydro-l-biopterin	$C_9H_{13}N_5O_3$	239.1018	910	70.0	80.9	20.2
Adenine	$C_5H_5N_5$	135.0545	910	70.0	80.9	20.2
Adenosine	$C_{10}H_{13}N_5O_4$	267.0968	1820	140.0	161.7	40.4
Adipic acid	$C_6H_{10}O_4$	146.0579	910	70.0	80.9	20.2
Alanine "SIL"	$^{13}C_3H_7{}^{15}NO_2$	93.0548	1540	118.5	136.9	34.2
Arachidonic Acid (d_8)	$C_{20}H_{24}D_8O_2$	312.2985	1236	120.0	138.6	34.7
Arginine "SIL"	$^{13}C_6H_{14}{}^{15}N_4O_2$	184.1197	3011	231.6	267.5	66.9
Aspartic acid "SIL"	$^{13}C_4H_7{}^{15}NO_4$	138.0480	2301	177.0	204.5	51.1
Biotin	$C_{10}H_{16}N_2O_3S$	244.0882	1820	140.0	161.7	40.4
Butyryl-L-carnitine	$C_{11}H_{21}NO_4$	231.1471	1300	100.0	115.5	28.9
Cholesteryl acetate	$C_{29}H_{48}O_2$	428.3654	910	70.0	80.9	20.2
Cholic acid (d_4)	$C_{24}H_{36}D_4O_5$	412.3167	1300	100.0	115.5	28.9
Cortisol	$C_{21}H_{30}O_5$	362.2093	910	70.0	80.9	20.2
Cortisone	$C_{21}H_{28}O_5$	360.1937	910	70.0	80.9	20.2
Creatinine (d_3)	$C_4H_4D_3N_3O$	116.0807	910	70.0	80.9	20.2
Cysteine "SIL"	$^{13}C_6H_7{}^{15}NO_2S$	125.0268	2093	161.0	186.0	46.5
Decanoyl-L-carnitine	$C_{17}H_{33}NO_4$	315.2410	910	70.0	80.9	20.2
Deoxycholic acid	$C_{24}H_{40}O_4$	392.2927	910	70.0	80.9	20.2
Diosmin	$C_{28}H_{32}O_{15}$	608.1741	910	70.0	80.9	20.2
Folic acid	$C_{19}H_{19}N_7O_6$	441.1397	910	70.0	80.9	20.2
Glutamic acid "SIL"	$^{13}C_5H_9{}^{15}NO_4$	153.0671	2543	195.6	225.9	56.5
Glycine "SIL"	$^{13}C_2H_5{}^{15}NO_2$	78.0357	1298	99.9	115.4	28.9
Glycocholic acid	$C_{26}H_{43}NO_6$	465.3090	910	70.0	80.9	20.2
Guanosine	$C_{10}H_{13}N_5O_5$	283.0952	2860	220.0	254.1	63.6
Hexanoyl-L-carnitine	$C_{13}H_{25}NO_4$	259.1783	910	70.0	80.9	20.2

Supplementary Table B7 (cont'd).

Standards	Formula	Monoisotopic Mass	Conc. of spiking mix (ng/mL)	Conc. in sample prior to MeOH extraction (ng/mL)	Conc in SPE samples at LC-MS (ng/mL)	Conc in MeOH at LC-MS (ng/mL)
Homovanillic acid (d_3)	$C_9H_7D_3O_4$	185.0784	910	70.0	80.9	20.2
Isoleucine "SIL"	$^{13}C_6H_{13}{}^{15}NO_2$	138.1119	2268	174.5	201.6	50.4
Kynurenic acid	$C_{10}H_7NO_3$	189.0426	910	70.0	80.9	20.2
Kynurenine	$C_{10}H_{12}N_2O_3$	208.0848	910	70.0	80.9	20.2
Leucine "SIL"	$^{13}C_6H_{13}{}^{15}NO_2$	138.1119	2268	174.5	201.6	50.4
Lysine "SIL"	$^{13}C_6H_{14}{}^{15}N_2O_2$	154.1197	2527	206.4	238.4	59.6
Melatonin	$C_{13}H_{16}N_2O_2$	232.1212	910	70.0	80.9	20.2
Methionine "SIL"	$^{13}C_5H_{11}{}^{15}NO_2S$	155.0650	2579	198.4	229.2	57.3
Neopterin	$C_9H_{11}N_5O_4$	253.0811	910	70.0	80.9	20.2
Norepinephrine (d_6)	$C_8H_5D_6NO_3$	175.1176	910	70.0	80.9	20.2
Octanoyl-L-carnitine	$C_{15}H_{29}NO_4$	287.2097	910	70.0	80.9	20.2
Phenylalanine "SIL"	$^{13}C_9H_{11}{}^{15}NO_2$	175.1065	2856	219.7	253.8	63.5
Proline "SIL"	$^{13}C_5H_9{}^{15}NO_2$	121.0772	1990	153.1	176.8	44.2
Propionyl-L-carnitine	$C_{10}H_{19}NO_4$	217.1314	1300	100.0	115.5	28.9
Pyridoxal	$C_8H_9NO_3$	167.0582	910	70.0	80.9	20.2
Pyridoxamine	$C_8H_{12}N_2O_2$	168.0899	3640	280.0	323.4	80.9
Pyridoxine	$C_8H_{11}NO_3$	169.0739	910	70.0	80.9	20.2
Resveratrol	$C_{14}H_{12}O_3$	228.0786	1300	100.0	115.5	28.9
Riboflavin	$C_{17}H_{20}N_4O_6$	376.1383	1300	100.0	115.5	28.9
Saccharopine	$C_{11}H_{20}N_2O_6$	276.2863	910	70.0	80.9	20.2
S-adenosylhomocysteine	$C_{14}H_{20}N_6O_5S$	384.1214	910	70.0	80.9	20.2
Serine "SIL"	$_{13}C_3H_{715}NO_3$	109.0497	1817	139.8	161.5	40.4
S-methyl-L-cysteine	$C_4H_9NO_2S$	135.0359	6760	520.0	600.6	150.2
Stearoyl-carnitine	$C_{25}H_{49}NO_4$	427.3662	2730	210.0	242.6	60.7
Taurocholic acid	$C_{26}H_{45}NO_7S$	515.2917	910	70.0	80.9	20.2
Threonine "SIL"	$^{13}C_4H_9{}^{15}NO_3$	124.0687	2059	158.4	183.0	45.8
Thymine	$C_5H_6N_2O_2$	126.0429	910	70.0	80.9	20.2
TRH	$C_{16}H_{22}N_6O_4$	362.1703	910	70.0	80.9	20.2
Thyroxine	$C_{15}H_{11}I_4NO_4$	776.6867	910	70.0	80.9	20.2
Triiodothyronine	$C_{15}H_{12}I_3NO_4$	650.7900	910	70.0	80.9	20.2
Tyrosine "SIL"	$^{13}C_9H_{11}{}^{15}NO_3$	191.1014	3132	241.0	278.4	69.6
Valine "SIL"	$^{13}C_5H_{11}{}^{15}NO_2$	123.0929	2024	155.7	179.8	45.0
β-estradiol	$C_{18}H_{24}O_2$	272.1776	1625	125.0	144.4	36.1

Supplementary Table B8. Fractionation preferences of standards detected in the evaluation of the analytical performance of combined sSPE fractions in plasma or standard solvent solutions. Detectability of standards demonstrated for both individual and n combined sSPE fractions and MeOH. LC-MS analysis on RP and ZIC-HILIC columns were executed in positive (+) and negative (-) ESI modes. The class of a standard is determined by its ability to carry negative (A), positive (C), both (Z) charges or neutral (N) at the physiological pH of human plasma. Fractionation preferences of standards were studied in individual sSPE ***(Section 3.2)*** *and combined sSPE fractions (****Section 3.3****). Fractionation of metabolites into more than one individual or combined fraction is designated by the "/" symbol. Several metabolites were not detected in either SPE or MeOH extracts ("ND"). The retention time of such metabolites and the LC-MS method was obtained from the analysis of neat standards in reconstitution solvent or from optimization and evaluation experiments (****Section 2.4****). Fractionation preferences are designated by fraction symbols: A – anion, C - cation, Z – zwitterion, N – neutral. Combined fractions were used for the assessment of analytical performance: AN (A+N), CZ (C+Z) in RP (-ve) ESI and AZ (A+Z), CN (C+N) in other three LC-MS methods. "TRH" stands for thyrotropin-releasing hormone.*

Standards	LC-MS	Monoisotopic Mass	RT (min) RP in standard solution	RT (min) ZIC-HILIC in standard solution	Class (charge at physiological pH)	preferred individual sSPE fraction(s)	preferred combined SPE fraction	MeOH detection
Thymine	RP+	126.0429	1.9	ND	N	N	CN	Y
Kynurenine	RP+	208.0848	2.5	ND	Z	ND		
Phenylalanine "SIL"	RP+	175.1065	2.6	11.5	Z	C	CN/AZ	Y
Butyryl-L-carnitine	RP+	231.1471	3.3	13.3	Z	C	CN	Y
Adenine	RP+	135.0545	4.1	4.0	N	C	CN	Y
Adenosine	RP+	267.0968	4.1	4.9	N	C	CN	Y
Adipic acid	RP-	146.0579	4.3	6.8	A	A	AN	ND
Homovanillic acid (d_3)	RP+	185.0784	6.2	8.0	A	A/N	CN/AZ	Y
Biotin	RP+	244.0882	6.3	12.3	A	A/N	CN/AZ	Y
Hexanoyl-L-carnitine	RP+	259.1783	6.5	11.3	Z	C	CN	Y
Diosmin	RP-	608.1741	7.9	7.3	N	A/N	AN	Y
Melatonin	RP+	232.1212	8.2	0.9	N	N	CN	Y
Octanoyl-L-carnitine	RP+	287.2097	8.7	10.9	Z	C	CN	Y
Resveratrol	RP-	228.0786	8.9	1.5	N	N	AN	Y
Cortisol	RP+	362.2093	9.5	1.0	N	N	CN	Y
Cortisone	RP+	360.1937	9.7	1.0	N	N	CN	Y
Triiodothyronine	RP+	650.7900	9.9	8.9	Z	Z	AZ	Y
Thyroxine	RP+	776.6867	10.7	8.6	Z	Z	AZ	Y

Supplementary Table B8 (cont'd).

Standards	LC-MS	Monoisotopic Mass	RT (min) RP in standard solution	RT (min) ZIC-HILIC in standard solution	Class (charge at physiological pH)	preferred individual sSPE fraction(s)	preferred combined SPE fraction	MeOH detection
Taurocholic acid	RP-	515.2917	11.2	7.8	A	A	AN	Y
β-estradiol	RP-	272.1776	11.9	6.4	N	N	AN	Y
Deoxycholic acid	RP-	392.2927	14.6	2.5	A	A	CZ/AN	Y
Arachidonic Acid (d_8)	RP-	312.2985	20.0	1.0	A	A	ND	Y
Glycine "SIL"	ZIC-HILIC-	78.0357	0.6	19.0	Z	ND		
Pyridoxal	ZIC-HILIC+	167.0582	1.1	2.1	A	Z	CN	Y
Pyridoxine	ZIC-HILIC+	169.0739	1.2	2.6	N	C/Z	CN/AZ	Y
Creatinine (d_3)	ZIC-HILIC+	116.0807	1.1	6.2	N	C	CN	Y
Riboflavin	ZIC-HILIC+	376.1383	6.3	6.5	A	N	CN	Y
Cholic acid (d_4)	ZIC-HILIC+	412.3167	12.4	6.5	A	A	AN	Y
Cysteine "SIL"	ZIC-HILIC+	125.0268	0.6	6.8	Z	ND	ND	Y
kynurenic acid	ZIC-HILIC-	189.0426	4.9	6.8	A	A	AZ	Y
5-methoxytryptamine	ZIC-HILIC+	190.1106	nd	9.8	C	N	ND	Y
Cholesteryl acetate	ZIC-HILIC-	428.3654	17.5	10.2	N	A	CN	ND
TRH*	ZIC-HILIC+	362.1703	1.2	10.2	C	C	AZ	Y
Norepinephrine (d(C_6)	ZIC-HILIC+	175.1176	0.6	10.6	C	C	CN	Y
Glycocholic acid	ZIC-HILIC-	465.3090	10.1	11.2	A	A	AZ	Y
Pyridoxamine	ZIC-HILIC+	168.0899	0.7	11.2	Z	ND	CN	Y
7,8-Dihydro-l-biopterin	ZIC-HILIC+	239.1018	0.6	12.2	N	N	ND	Y
Leucine "SIL"	ZIC-HILIC+	138.1119	1.2	12.2	Z	C	CN	Y
Guanosine	ZIC-HILIC+	283.0952	ND	12.73	N	C	CN	Y
Isoleucine "SIL"	ZIC-HILIC+	138.1119	1.2	12.9	Z	C	CN	Y
Methionine "SIL"	ZIC-HILIC+	155.0650	0.9	13.8	Z	C	CN	ND
5-hydroxy-indoleacetic acid	ZIC-HILIC-	191.0582	5.5	14.2	A	A	ND	Y
3-Hydroxy-DL-kynurenine	ZIC-HILIC+	224.0797	ND	14.8	Z	ND		
S-methyl-L-cysteine	ZIC-HILIC-	135.0359	0.8	14.9	Z	C	CN	Y
Valine "SIL"	ZIC-HILIC+	123.0929	0.6	15.0	Z	C	CN	Y
Tyrosine "SIL"	ZIC-HILIC+	191.1014	1.2	15.5	Z	ND	CN/AZ	Y
Propionyl-L-carnitine	ZIC-HILIC+	217.1314	1.3	15.5	Z	C	CN	Y
Neopterin	ZIC-HILIC-	253.0811	1.0	16.0	N	C	CN/AZ	ND
5-Hydroxy-L-tryptophan	ZIC-HILIC+	220.0848	2.0	16.5	Z	N	ND	Y

Supplementary Table B8. (cont'd).

Standards	LC-MS	Monoisotopic Mass	RT (min) RP in standard solution	RT (min) ZIC-HILIC in standard solution	Class (charge at physiological pH)	preferred individual sSPE fraction(s)	preferred combined SPE fraction	MeOH detection
Proline "SIL"	ZIC-HILIC+	121.0772	0.5	16.7	Z	C	CN	Y
Threonine "SIL"	ZIC-HILIC-	124.0687	0.8	18.9	Z	ND		
S-adenosylhomocysteine	ZIC-HILIC+	384.1214	3.0	19.1	Z	C/Z	AZ	ND
Serine "SIL"	ZIC-HILIC-	109.0497	0.7	21.1	Z	C	CN	Y
Glutamic acid "SIL"	ZIC-HILIC-	153.0671	0.7	22.6	Z	A	AZ	ND
Alanine "SIL"	ZIC-HILIC-	93.0548	0.7	23.0	Z	C/Z	AZ	ND
Aspartic acid "SIL"	ZIC-HILIC-	138.0480	0.7	23.0	Z	ND		
Folic acid	ZIC-HILIC-	441.1397	5.3	23.2	A	A	AZ	ND
Saccharopine	ZIC-HILIC-	276.2863	0.6	24.1	Z	ND		
Lysine "SIL"	ZIC-HILIC+	154.1197	0.4	24.2	Z	ND		
Arginine "SIL"	ZIC-HILIC+	184.1197	0.5	25.1	Z	ND		

Supplementary Table B9. The total number of standards affected by matrix effects at the maximum amount loaded in SPE fractions (equivalent to 6.4 µL of plasma) and MeOH extracts (equivalent to 1.6 µL of plasma). Matrix effects were compared between SPE fraction(s) and MeOH extracts within the same LC-MS method, which promotes the strongest signal for neat standards in the sample reconstitution solvent (used as 100% reference for calculation of matrix effects. Please see ***Section 3.3.4*** *for details. Therefore, each standard was accounted for once for the best LC-MS method. Matrix effects were calculated in combined fractions CN and AZ (CZ and AN, respectively, in (-ve) RP) and MeOH extracts. In the case of split metabolites, the fraction with higher recovery (****Figure 3.16****) was chosen for this table to avoid redundancy. Matrix effects for split metabolites are demonstrated for both fractions in* ***Figure 3.17****.*

Matrix effects	**(+ve) RP**			**(+ve) ZIC HILIC**		
	sSPE	**MeOH**	**Total**	**sSPE**	**MeOH**	**Total**
Signal suppression (< 80%)	9	0	**9**	18	17	**35**
Signal enhancement (> 120%)	1	3	**4**	1	1	**2**
No matrix effects (80-120%)	4	11	**15**	1	1	**2**
Total affected	10	3	**13**	19	18	**37**
Total not affected	4	11	**15**	1	1	**2**
Matrix effects	**(-ve) RP**			**(-ve) ZIC HILIC**		
	sSPE	**MeOH**	**Total**	**sSPE**	**MeOH**	**Total**
Signal suppression (<80%)	5	4	**9**	6	6	**12**
Signal enhancement (> 120%)	0	1	**1**	2	0	**2**
No matrix effects (80-120%)	2	1	**3**	3	2	**5**
Total affected	5	5	**10**	8	6	**14**
Total not affected	2	1	**3**	3	2	**5**

Supplementary Table B10. Total metabolome coverage in combined sSPE fractions (AN and CZ for (-ve) RP, AZ, and CN for other 3 LC-MS methods) at the loading amount equivalent to equivalent to 6.4 µL plasma. The study's total is a direct sum of metabolites across all LC-MS methods. Due to technical challenges, the search for identical metabolites across methods is unfeasible, and the redundancy assessment was not executed. Metabolites in which signal ratio to the sum of signals in both was below or equal to 0.1 were considered as metabolites with the minor split.

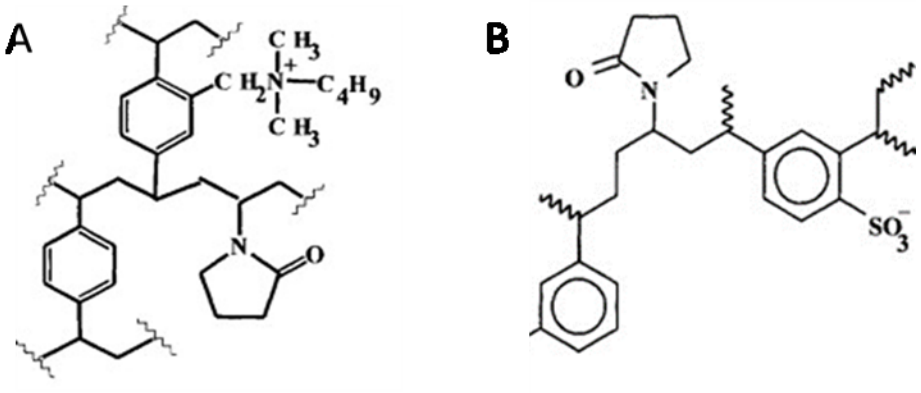
Parameters	(+ve) RP	(-ve) RP	(+ve) ZIC-HILIC	(-ve) ZIC-HILIC	Study' total
Total AX *(X = N for (-ve) RP; X = Z for (+ve) RP)*	3184	3099	1918	2411	10612
Total CX *(X = Z for (-ve) RP; X = N for (+ve) RP)*	2556	2131	2041	1981	8709
Total metabolites sSPE	4370	4472	3037	3545	15424

*Supplementary Table B11. Peak areas for standard metabolites, which were detected in several LC-MS methods and sSPE fractions. Peak areas were obtained as described in **Sections 3.2.4.8**, respectively. The analysis was executed in combined sSPE fractions, as described in **Section 3.2.5.1**.*

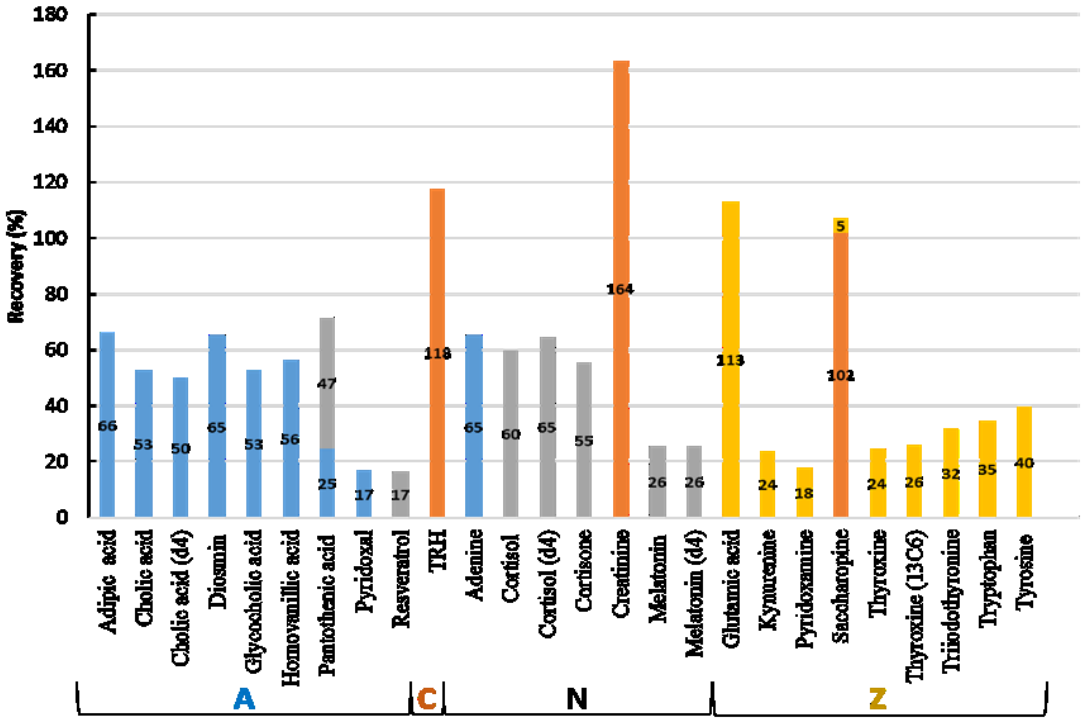
Standard	MeOH (+ve) RP	AZ (+ve) RP	CN (+ve) RP	MeOH (-ve) RP	AN (-ve) RP	MeOH (-ve) ZIC-HILIC	AZ (-ve) ZIC-HILIC	MeOH (+ve) ZIC-HILIC	CN (+ve) ZIC-HILIC
Stearoyl-carnitine			353934						
Adenine	41763		40500						
Kynurenine	176315	138741	354044						
Butyryl-L-carnitine	96616		617029						
Hexanoyl-L-carnitine	84630		404593						
Octanoyl-L-carnitine	436599		2770546						
Decanoyl-L-carnitine	677404		2652091						
Cortisone	33871		109074						
Cortisol	247073		986892						
Riboflavin	15089		91531					426474	759397
Thyroxine	77960	114479				440397	970519		
Adenosine	89710		277100					1919525	4779214
Taurocholic acid	47272	253303		185153		984996	60291		
Glycocholic acid	753358	2126382		1208507	3389056	2351640	5953047		
Kynurenic acid	31326					947299			
Deoxycholic acid				634441	470584				
Folic acid				14873	26373				
Pyridoxine								5554910	6396087
Pyridoxal								119342	
Propionyl-L-carnitine								3100486	10055419

*Supplementary Table B12. Comparison of matrix effects ("M. effects") and peak areas ("Peak area") for standard metabolites, which were detected in several LC-MS methods and sSPE fractions. The analysis was executed in combined sSPE fractions, as described in **Section 3.2.5.1**. Peak areas and matrix effects were calculated as described in **Sections 3.2.4.8 and 3.2.5.4**, respectively. Coloring of peak areas was executed in Excel using tri-color formatting function to facilitate signal differentiation from the lowest (dark blue) to the highest (dark red) in a row.*

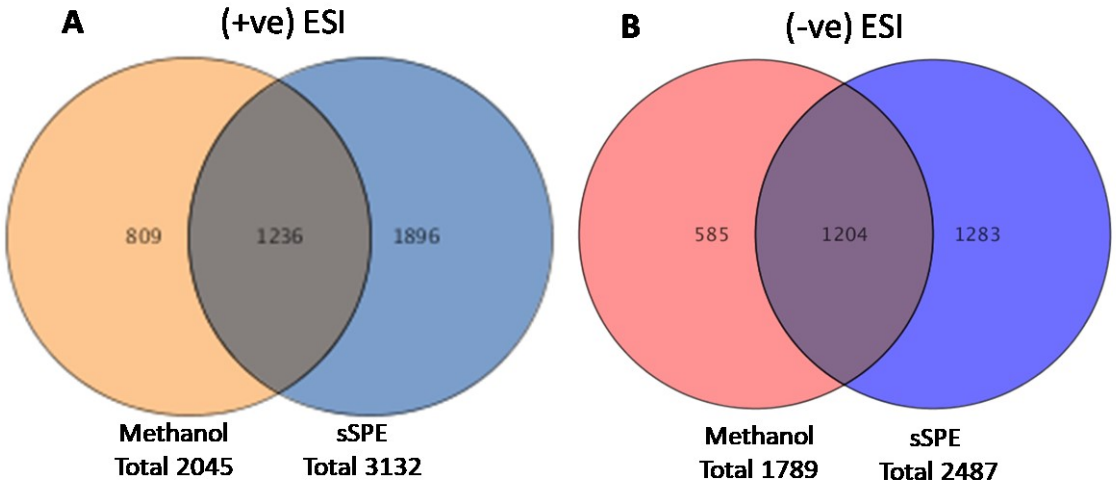
Standard	Parameter	MeOH (+ve) RP	AZ (+ve) RP	CN (+ve) RP	MeOH (-ve) RP	AN (-ve) RP	MeOH (-ve) ZIC-HILIC	AZ (-ve) ZIC-HILIC	MeOH (+ve) ZIC-HILIC	CN (+ve) ZIC-HILIC
Riboflavin	**Peak area**	15,089		91,531					426,474	759,397
Riboflavin	**M. effects**	32		59					65	74
Thyroxine	**Peak area**	77,960	114,479				440,397	970,519		
Thyroxine	**M. effects**	151	41				59	50		
Adenosine	**Peak area**	89,710		277,100					1,919,525	4,779,214
Adenosine	**M. effects**	133		90					68	63
Taurocholic acid	**Peak area**	47,272	253,303		185153		984,996	60,291		
Taurocholic acid	**M. effects**	49	90		64		39	56		
Glycocholic acid	**Peak area**	753,358	2,126,382		1,208,507	3,389,056	2,351,640	5,953,047		
Glycocholic acid	**M. effects**	96	83		106	56	66	43		



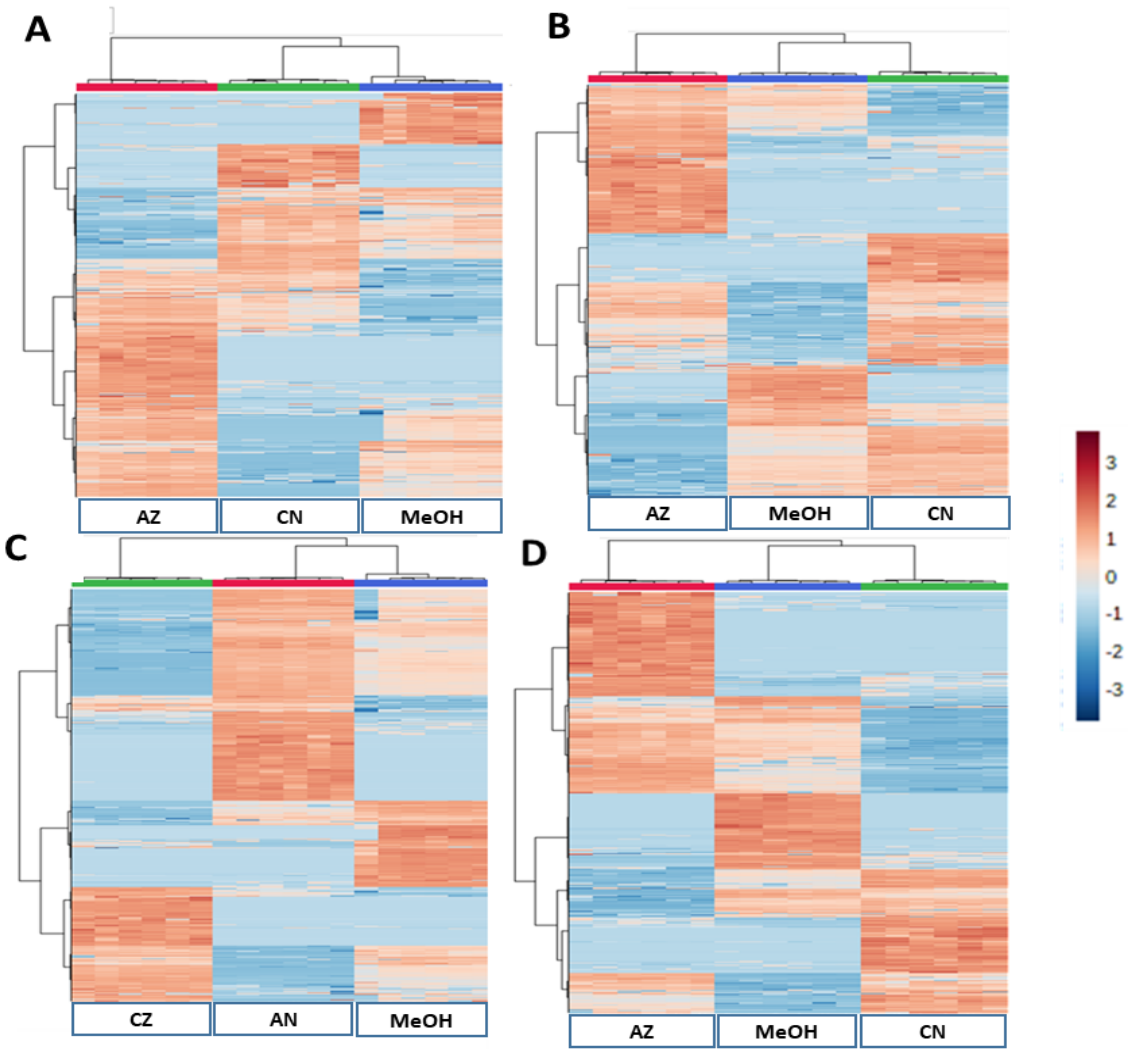
Supplementary Figure B1. Approximate representation of strong anion and strong cation-exchange sorbents in MAX and MCX Oasis SPE plates. Sorbents consist of DVB linked by short alkyl chains, conjugated with pyrrolidone and containing quaternary ammonium (A) or benzenesulfonate (B) for MAX and MCX plates, respectively. Reprinted with permission.[254]



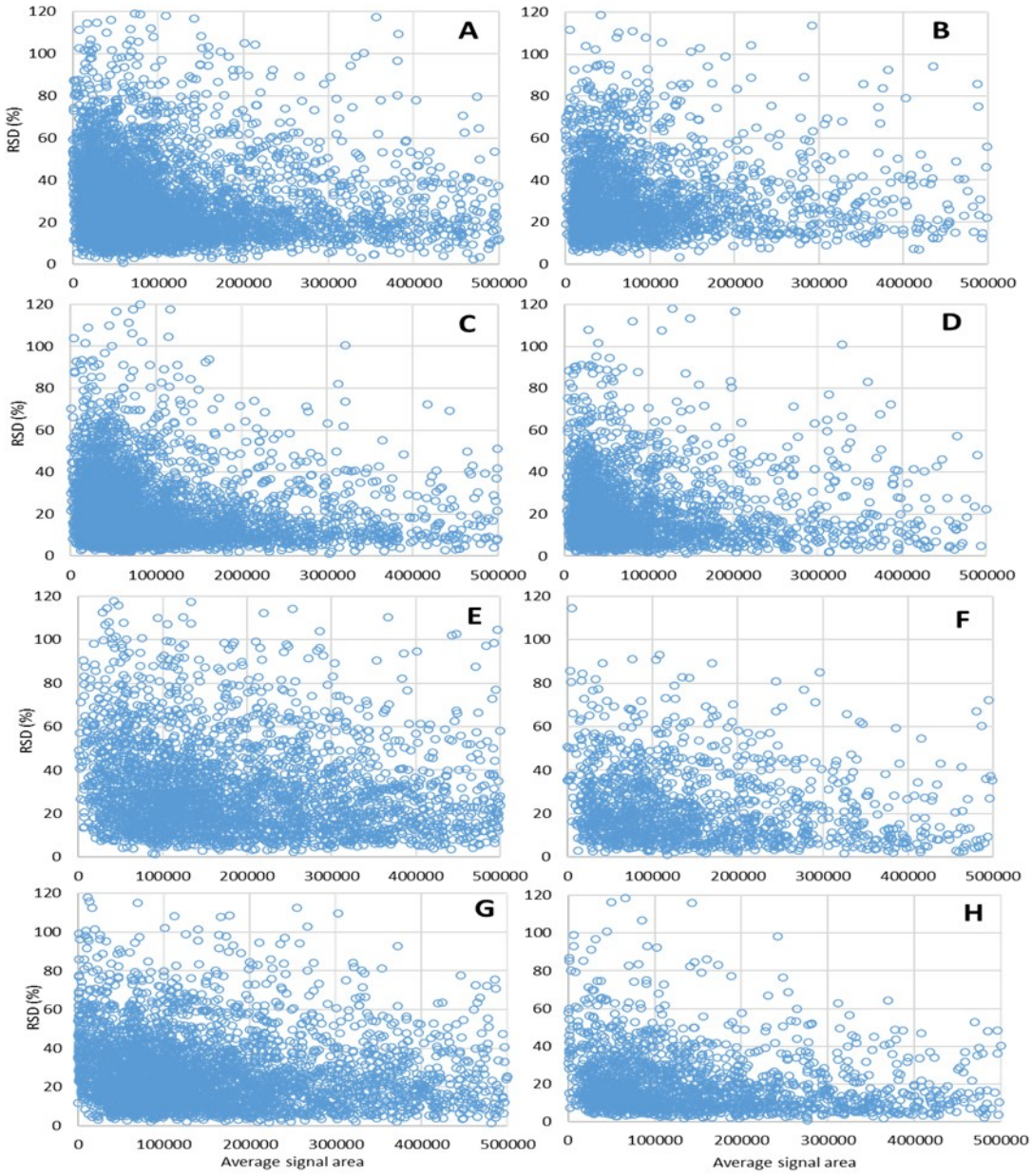
Supplementary Figure B2. Recovery of standard metabolites spiked into aqMTBE-MeOH immediately prior to sSPE and detected in RP analysis in (+ve) and (-ve) ESI. Preferential ionization mode for metabolites can be found in Supplementary ***Table B8.*** *sSPE fractions by the color of bars: blue-anion; grey-neutral; orange-cation and yellow-zwitterion. The recovery (% of signal in "pre-spiked" relatively to "post-spiked") is indicated on Y-axis and by digits within bars. Standards are distributed across X-axis in groups by their charge state at physiological pH (7.2), which are designated by letters on the insert below the graph: A - anion, C - cation, N – neutral, and Z - zwitterion.*



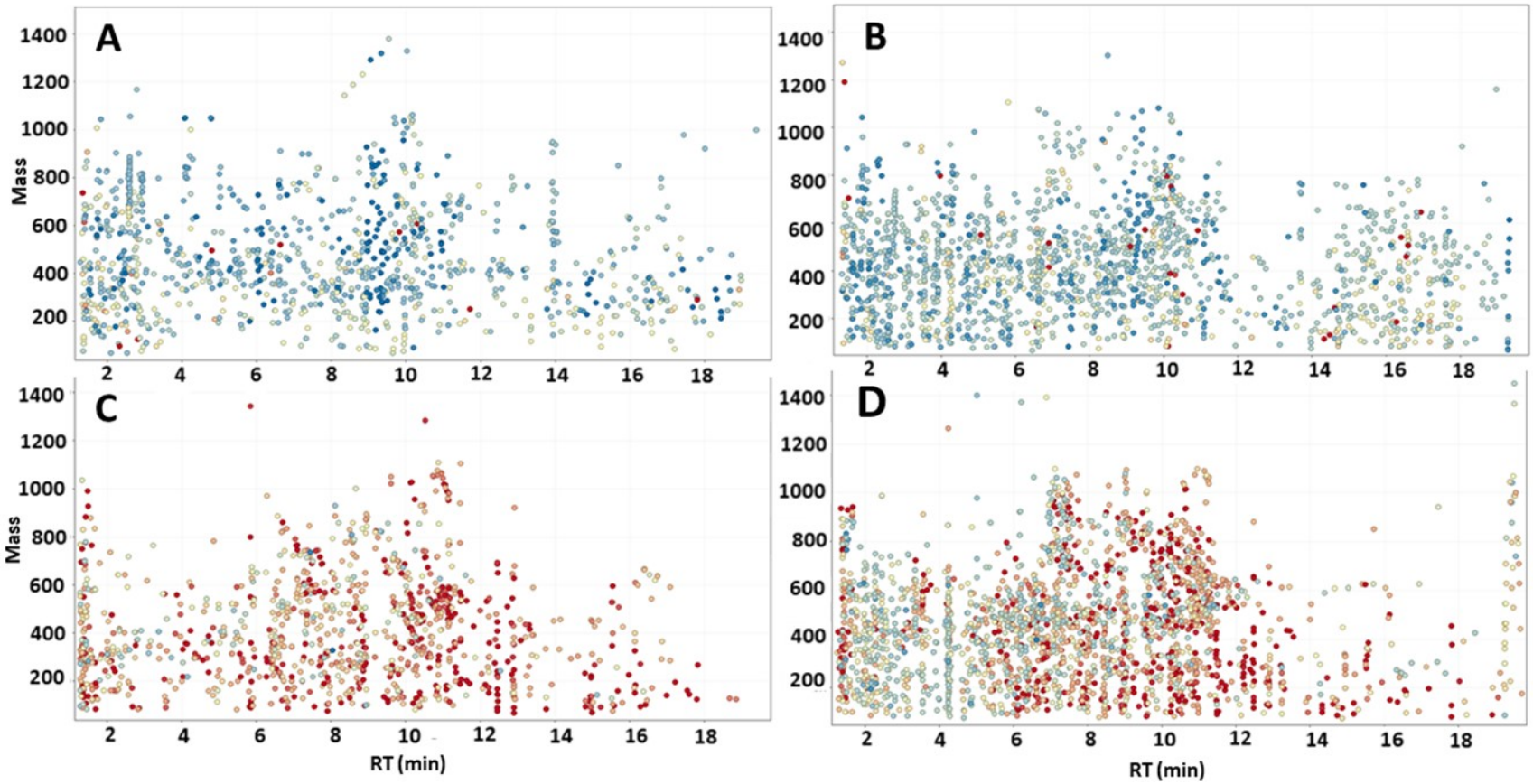
Supplementary Figure B3. Analysis of the coverage of mid-polar metabolome by aqMTBE-MeOH and sSPE in RP LC-MS at (+ve) (panel A) and (-ve) (panel B) ESI. Details for the raw data processing and analysis can be found in ***Section 3.2.3.3.2****. The total number of metabolites calculated for sSPE and aqMTBE-MeOH is non-redundant.*



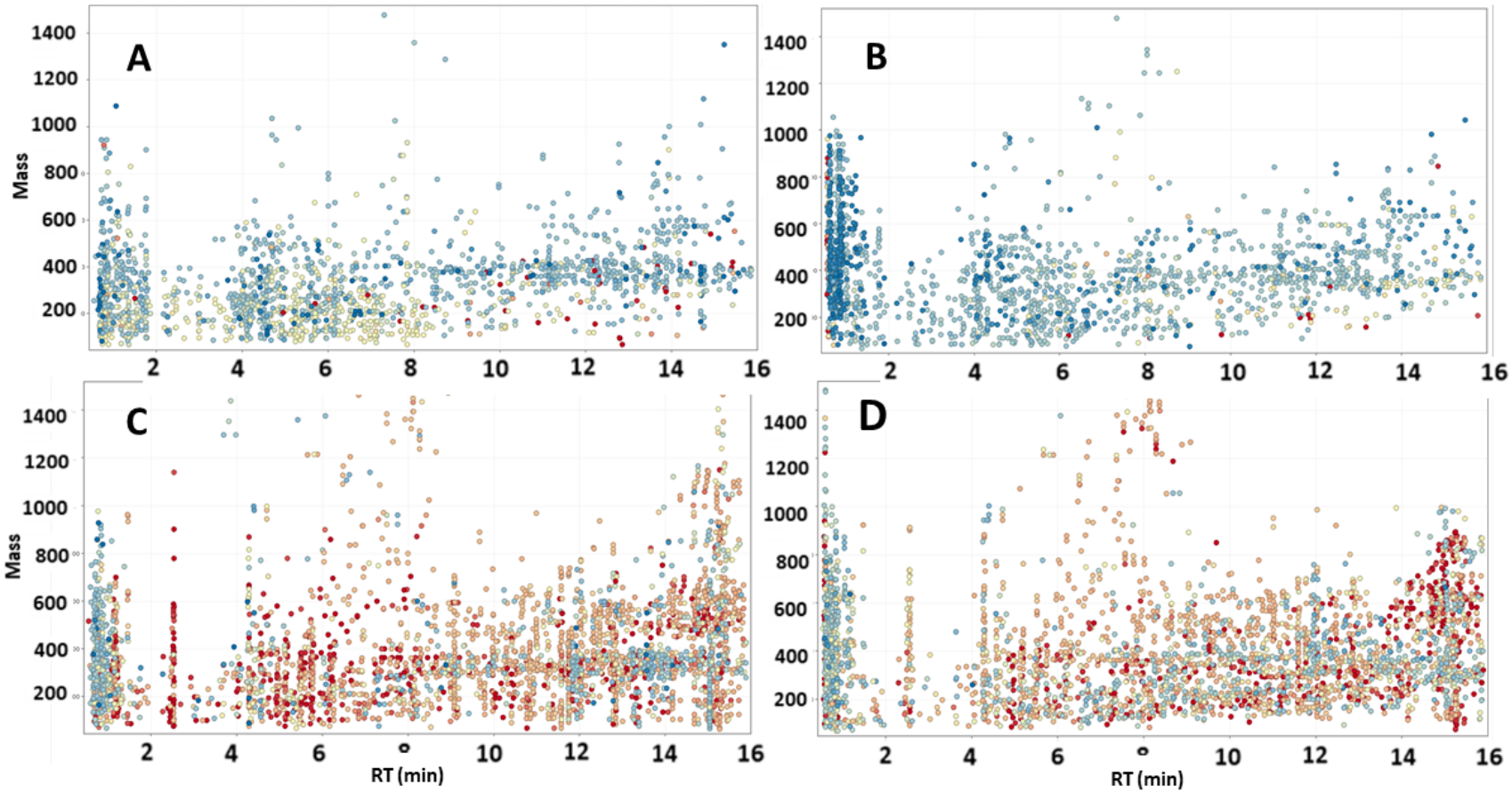
Supplementary Figure B4. Hierarchical cluster analysis of combined sSPE fractions and MeOH extracts for LC-MS analysis A-(+ve) RP, C – (-ve) RP, B, and D for ZIC-HILIC (+ve) and (-ve) ESI, respectively. Extracts designated by color bands at the top of each plot, blue -MeOH, red - AN (-ve)RP and AZ (+ve) RP and green - CZ (-ve)RP and CN (+ve) R. Plots display metabolites that satisfy criteria of the data analysis described in the ***Section 3.2.5.4****. The strength of signals is designated by the range from 4 to – 4, which corresponds to the range from the maximum signal (red) to no detection status (dark blue).*



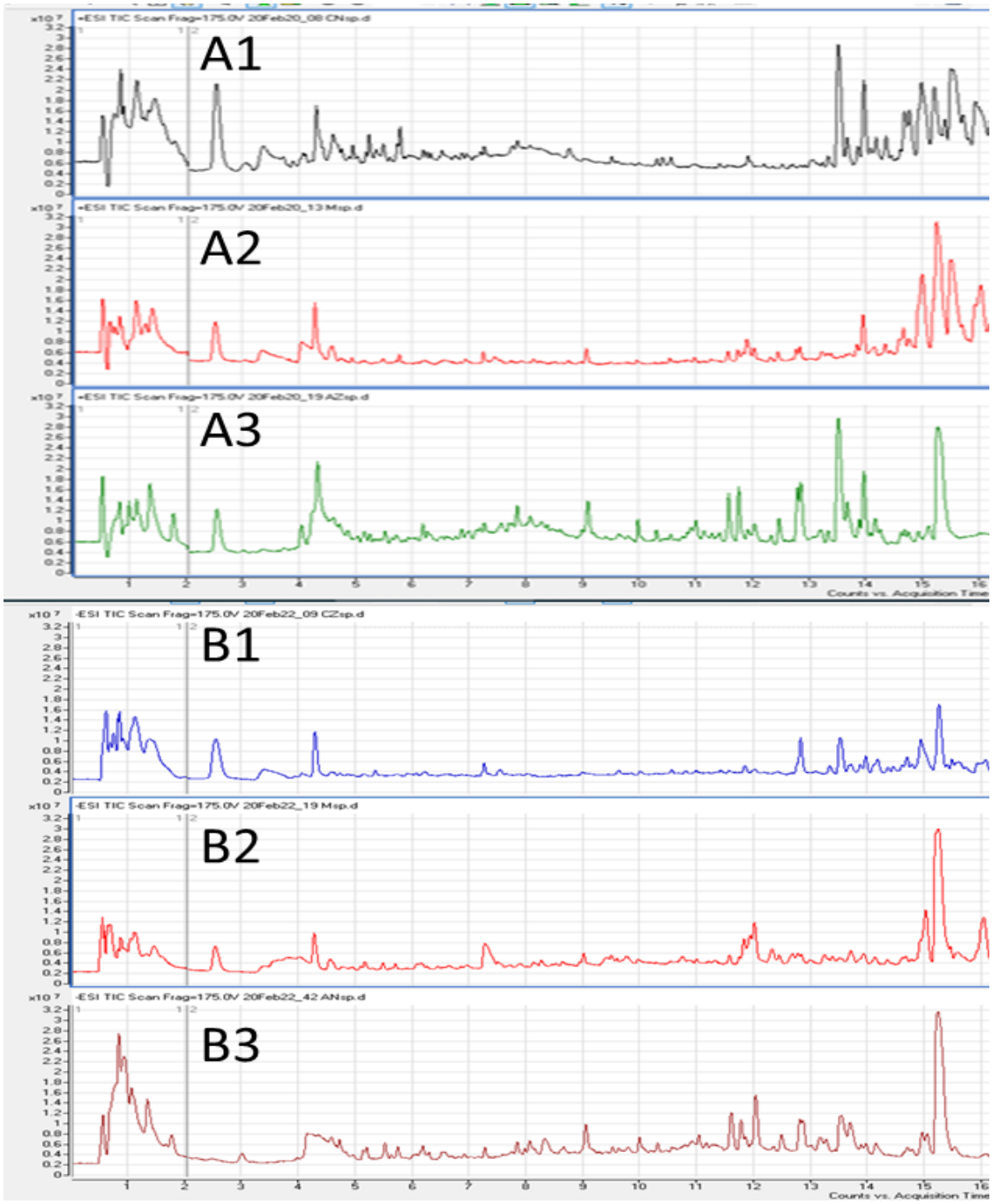
Supplementary Figure B5. Repeatability (Y-axis "RSD%") over the average signal area (X-axis, "Average signal area") in replicates (n=6) of combined sSPE (A, C, E, and G) and MeOH extracts (B, D, F, H) in RP analysis at (+ve) ESI (A, B); (-ve) ESI (C, D) and in ZIC-HILIC analysis at (+ve) ESI (E, F) and (-ve)ESI (G, H). sSPE represented by the nonredundant (Venn analysis) content of both combined fractions: ((AN + CZ) for (-ve) RP; (AZ+CN) for other LC-MS methods.



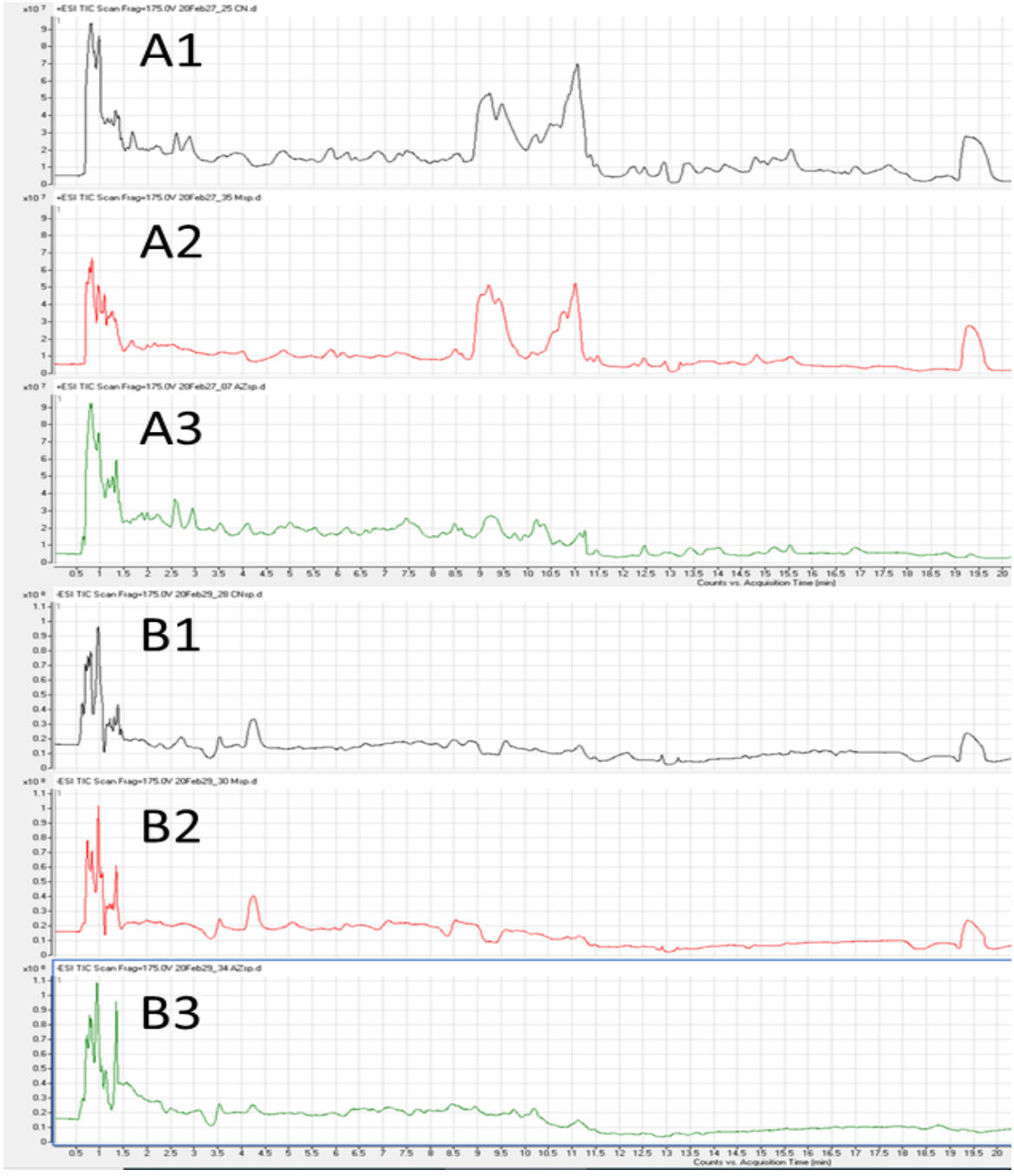
Supplementary Figure B6. An enrichment of polar metabolites on mass maps of metabolites detected in the ZIC-HILIC analysis of combined sSPE fractions: A, B -unique (compared to MeOH) sSPE metabolites in (+ve) (A) and (-ve) (B) ZIC HILIC analysis. C, D – all metabolites detected in MeOH in(+ve) ZIC HILIC) and (-ve) ZIC HILIC, respectively. Colors correspond to the detection frequency in the data set between 29 (dark red) and 4 (dark blue). Note the enrichment of highly polar (RT > 14 min) metabolites in sSPE (A, B) vs. MeOH (C, D).



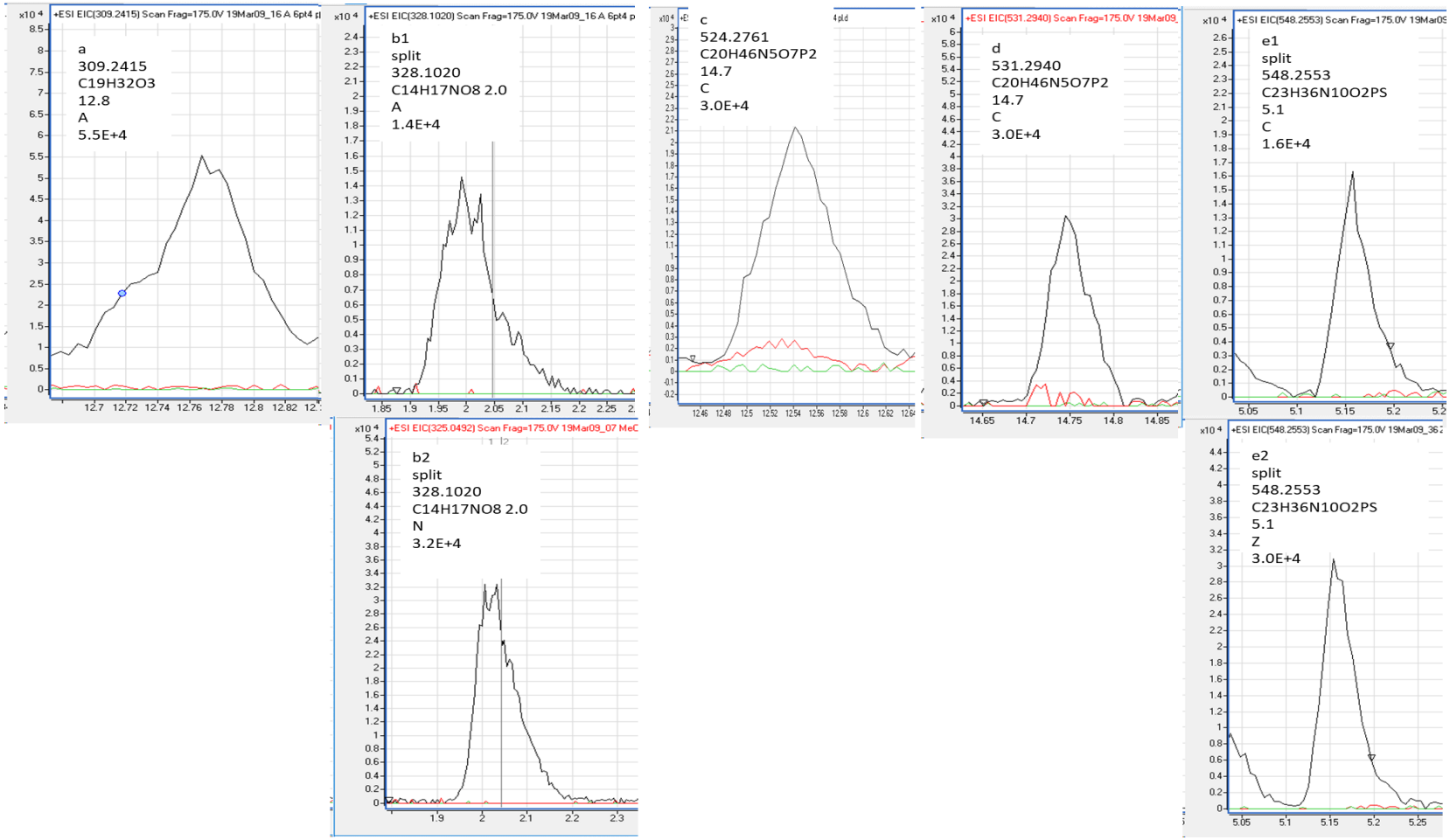
Supplementary Figure B7. An enrichment of polar metabolites on mass maps of metabolites detected in the RP analysis of combined sSPE fractions: A, B -unique (compared to MeOH) sSPE metabolites in (+ve) (A) and (-ve) (B) RP analysis. C, D – all metabolites detected in MeOH in(+ve) RP and (-ve) RP, respectively. Colors correspond to the frequency of the detection in the data set between 29 (dark red) and 4 (dark blue). Note the enrichment of highly polar (RT between 4 and 6 min) metabolites in sSPE (A, B) vs. MeOH (C, D).



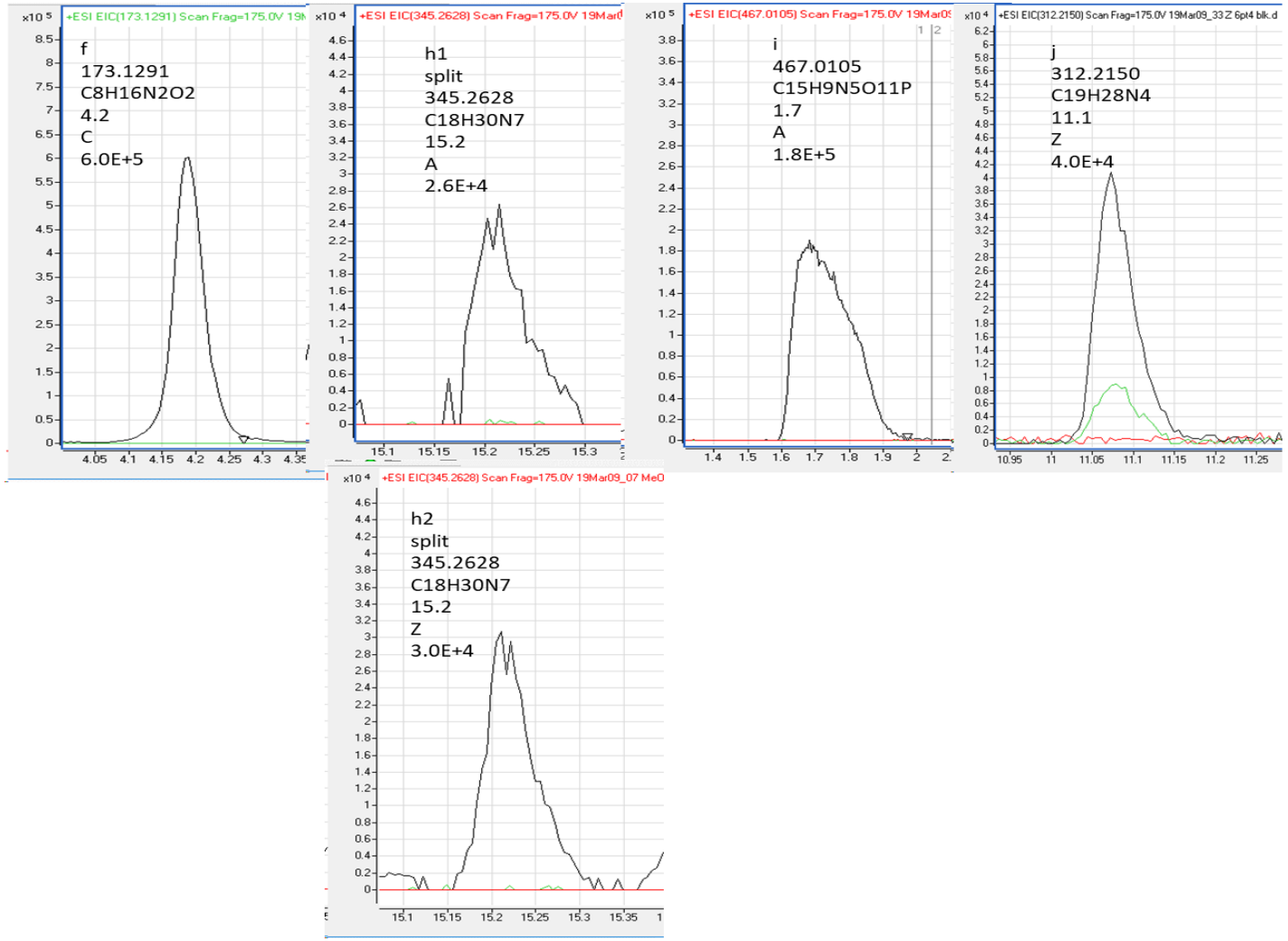
Supplementary Figure B8. TIC profiles of sSPE fractions: CN (A1), AZ (A3), and MeOH extract (A2) in (+ve) RP and CZ (B1), AN (B3) and MeOH (B2) in (-ve) RP analysis.



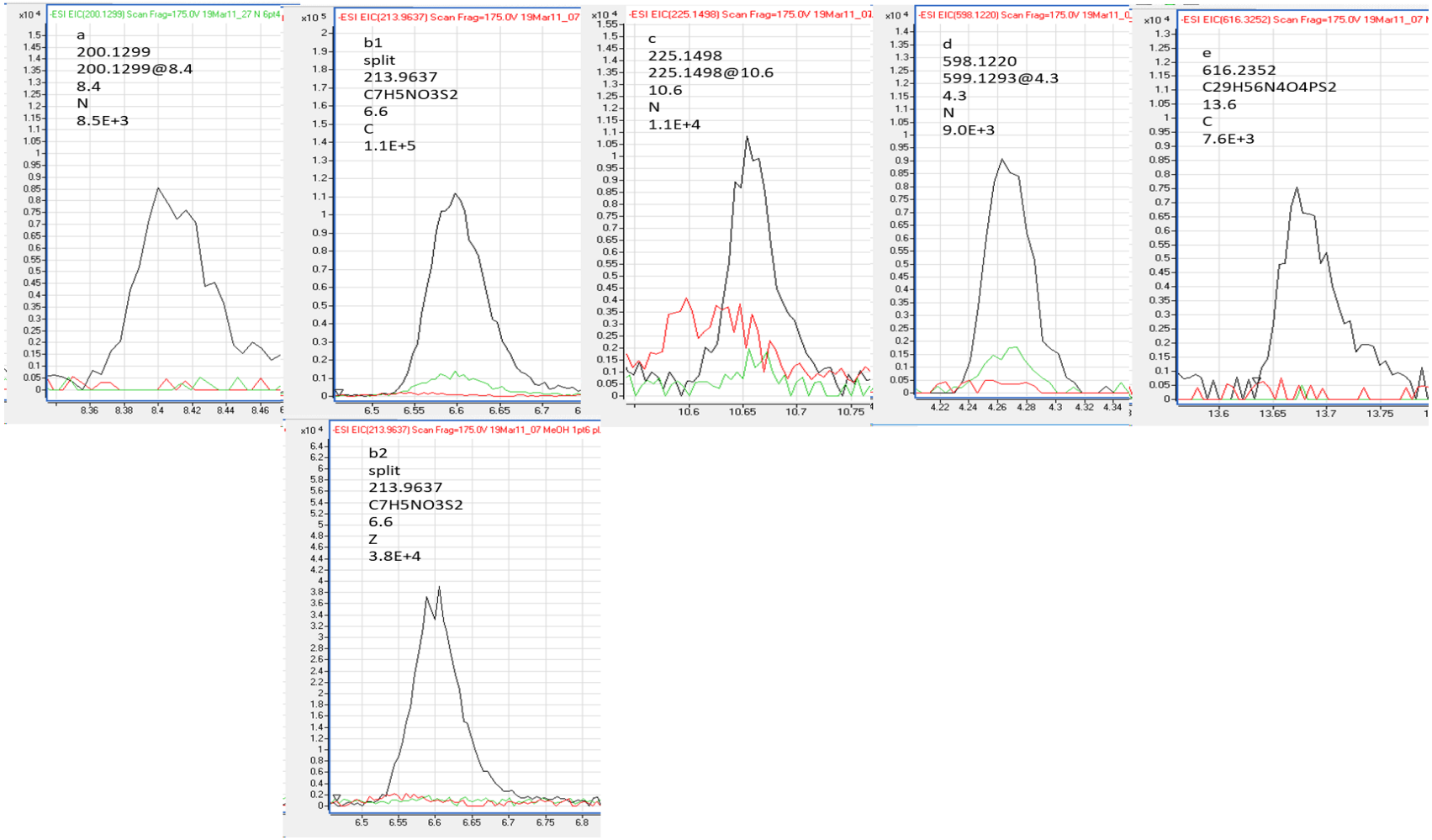
Supplementary Figure B9. TIC profiles of sSPE fractions: CN (A1), AZ (A3) and MeOH extract (A2) in (+ve) ZIC-HILIC and CN (B1), AZ (B3) and MeOH (B2) in (-ve) ZIC-HILIC analysis.



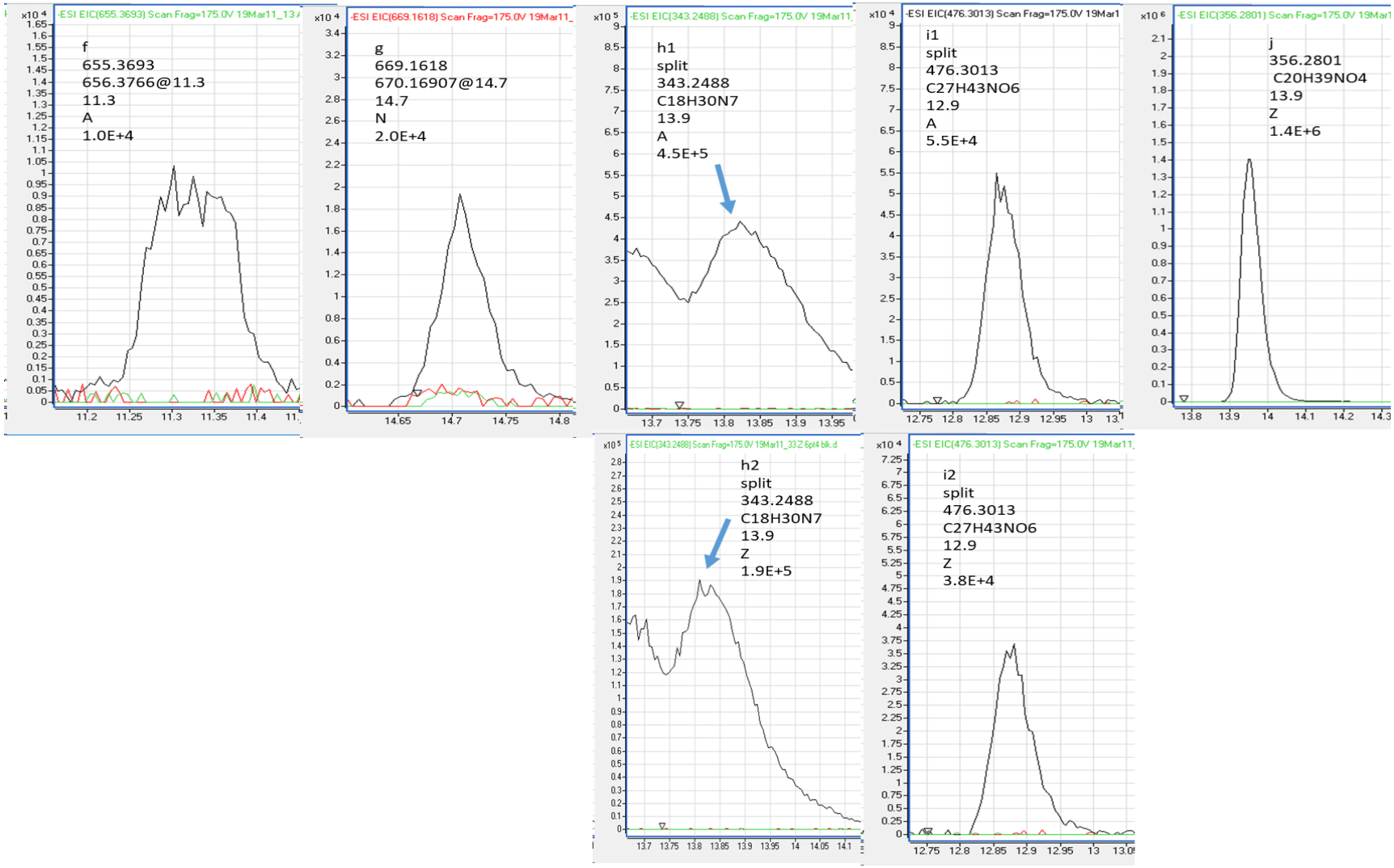
Supplementary Figure B10. Extracted ion chromatograms (EIC)s of example unknown metabolites detected only in sSPE fractions in (+ve) RP global metabolomics analysis. An m/z value, molecular formula, RT (min), fraction name, and height of the peak are indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm. Chromatograms b1, b2, and e1, e2 demonstrate metabolites split between A, N, and C, Z fractions, respectively.



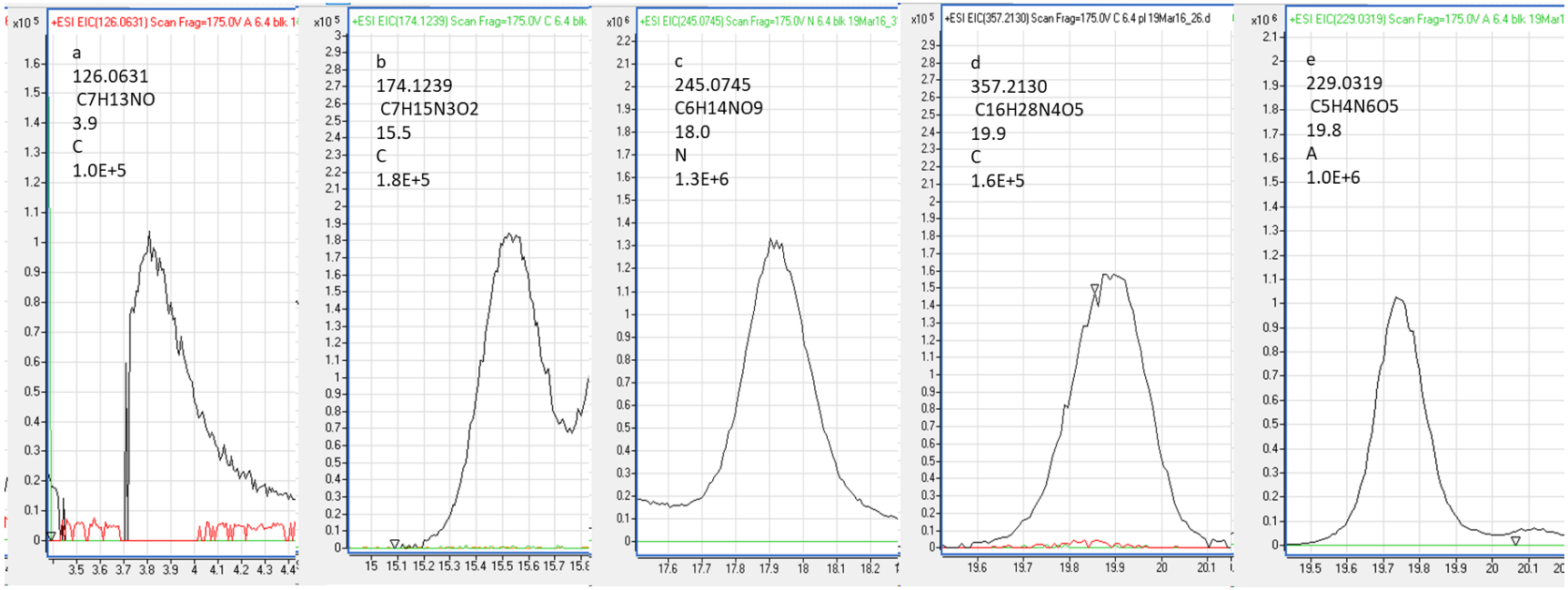
Supplementary Figure B10 (cont'd). Chromatograms h1 and h2 demonstrate the metabolite, which split between A and Z fractions.



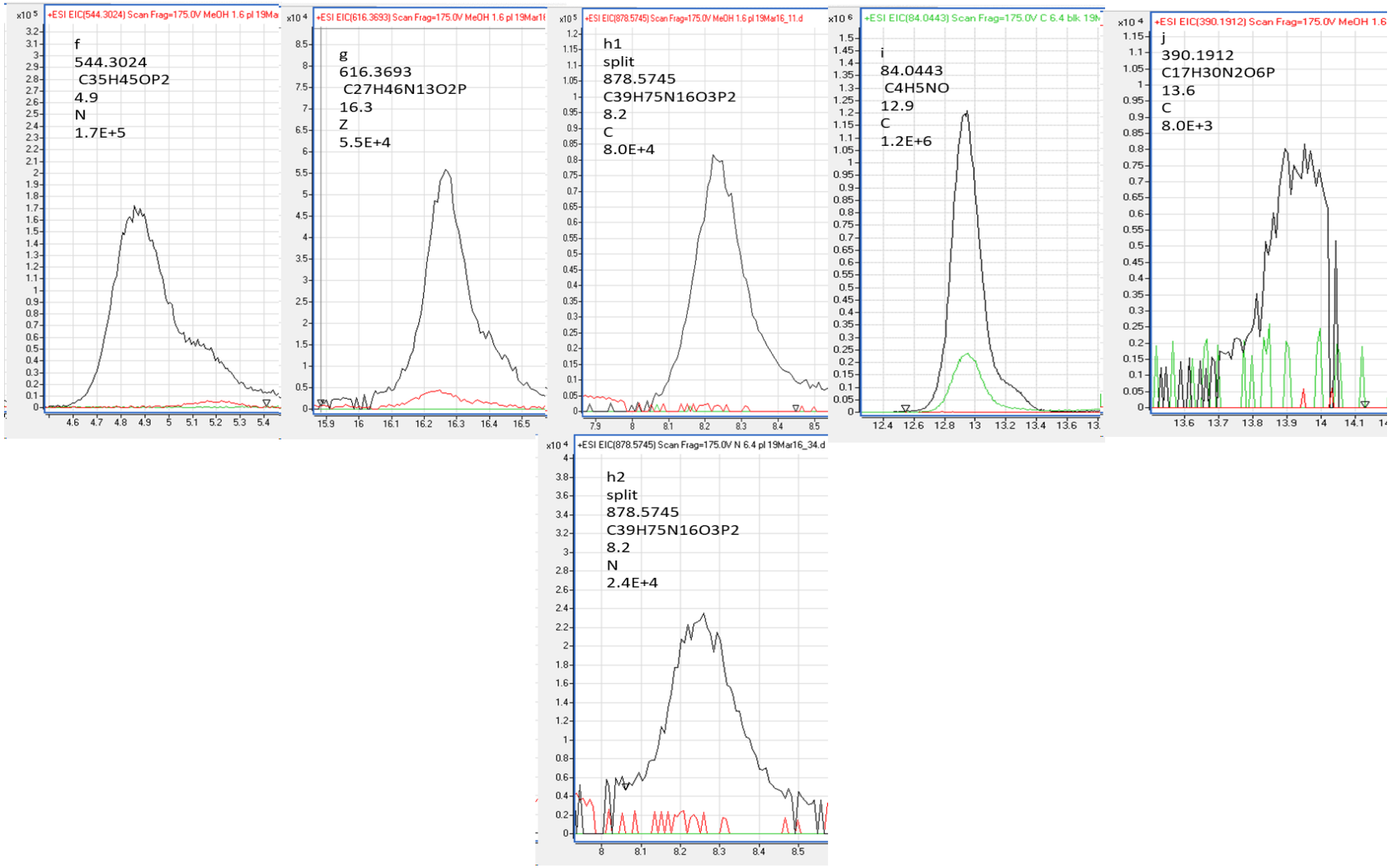
Supplementary Figure B11. EICs of example unknown metabolites detected only in sSPE fractions in (-ve) RP global metabolomics analysis. An m/z value, molecular formula, RT (min), fraction name, and height of the peak are indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm. Chromatograms b1, b2 demonstrate metabolites split between C, Z fractions.



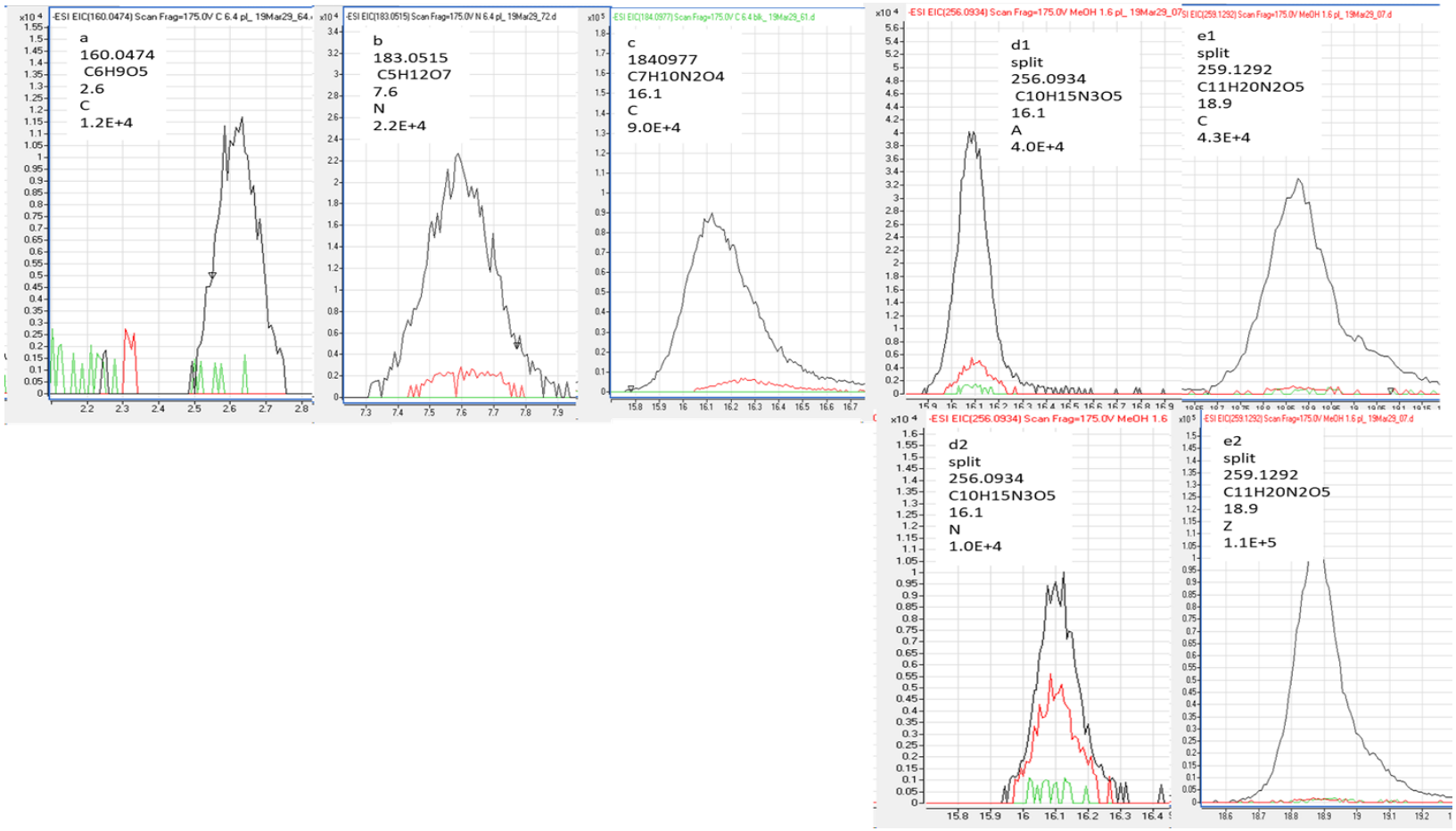
Supplementary Figure B11 (cont'd). Chromatograms h1, h2, and i1, i2 demonstrate metabolites split between A and Z fraction, respectively. Arrows in h1 and h2 point to the peak of interest.



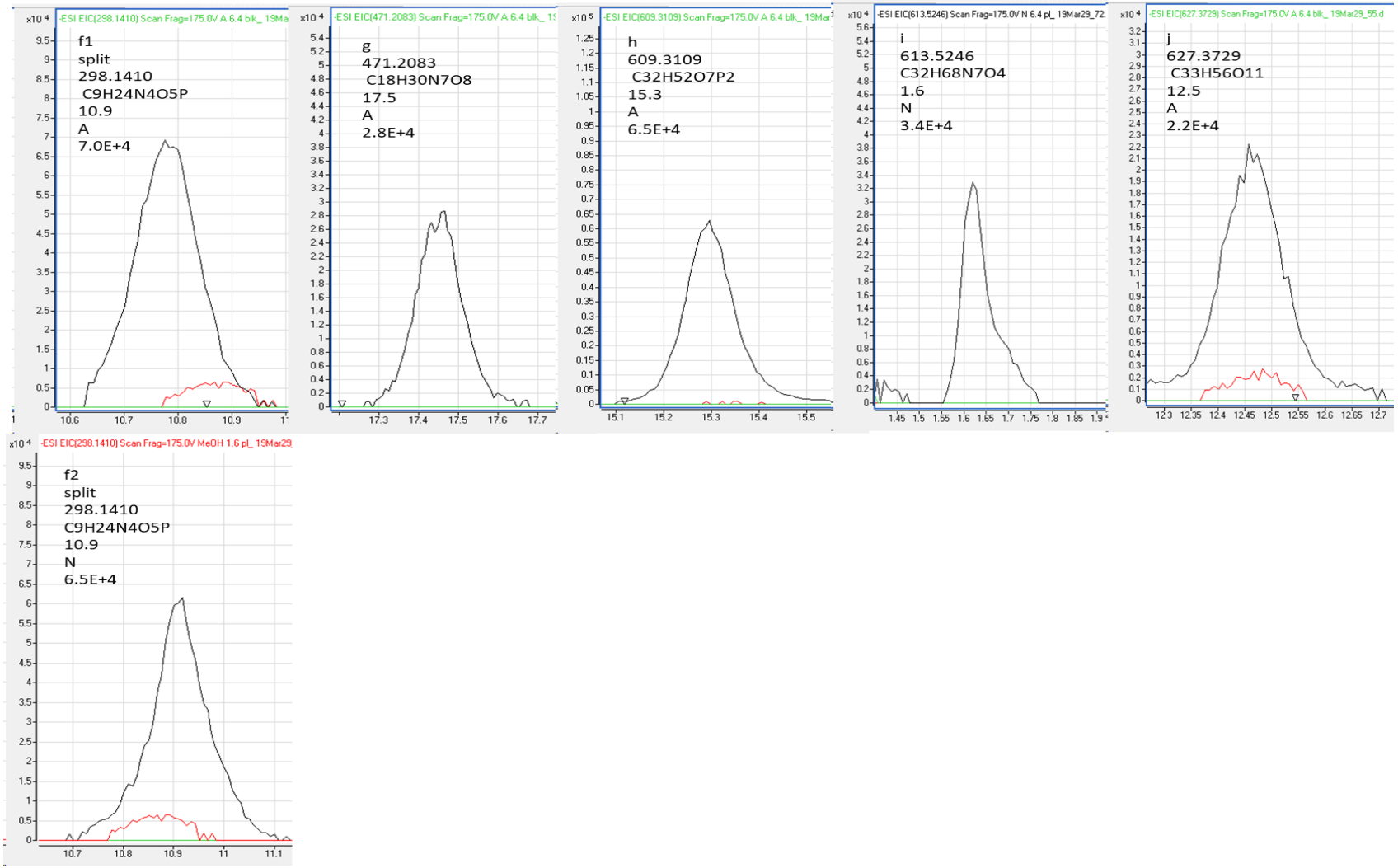
Supplementary Figure B12. EIC of unknown metabolites detected only in sSPE fractions in (+ve) ZIC-HILIC in the global metabolomics analysis of individual sSPE fractions. An m/z value, molecular formula, RT (min), fraction name, and height of the peak are indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm.



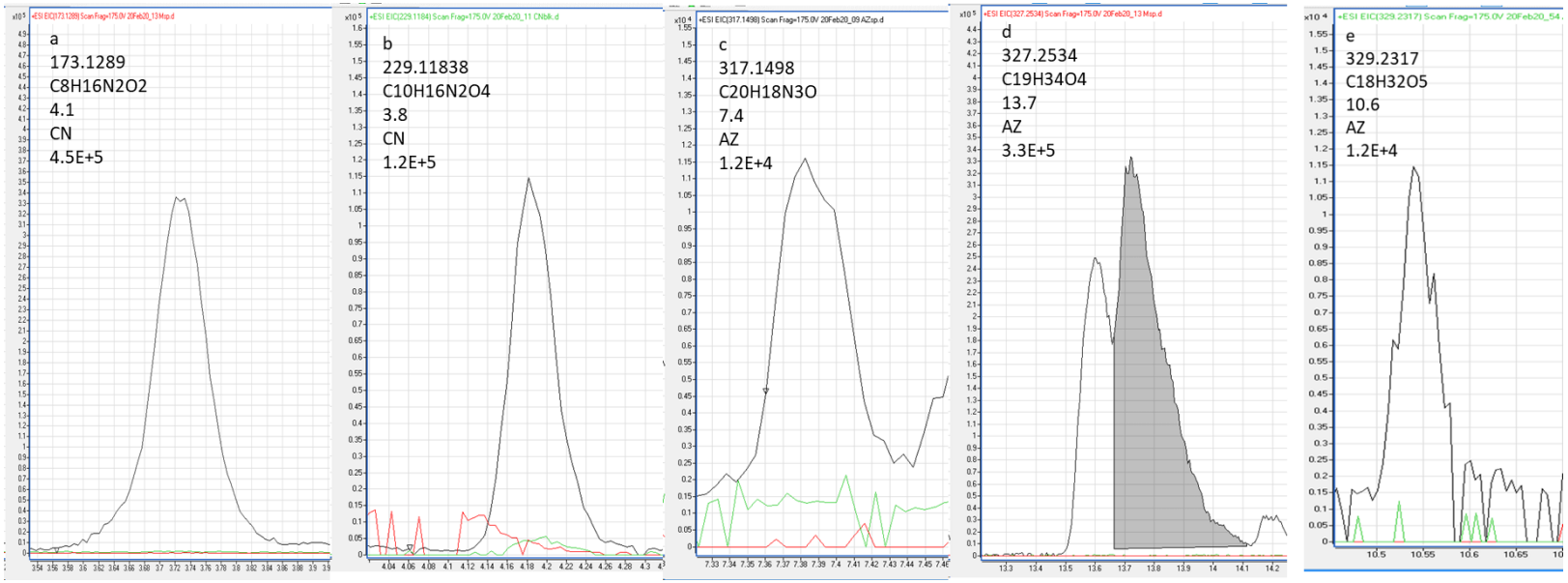
Supplementary Figure B12 (cont'd). Chromatograms h1 and h2 demonstrate the metabolite split between fractions C and N, respectively. The p*eaks in sSPE fraction C (chromatogram I, green trace) was below the intensity threshold (7500) and was rejected from further analysis.*



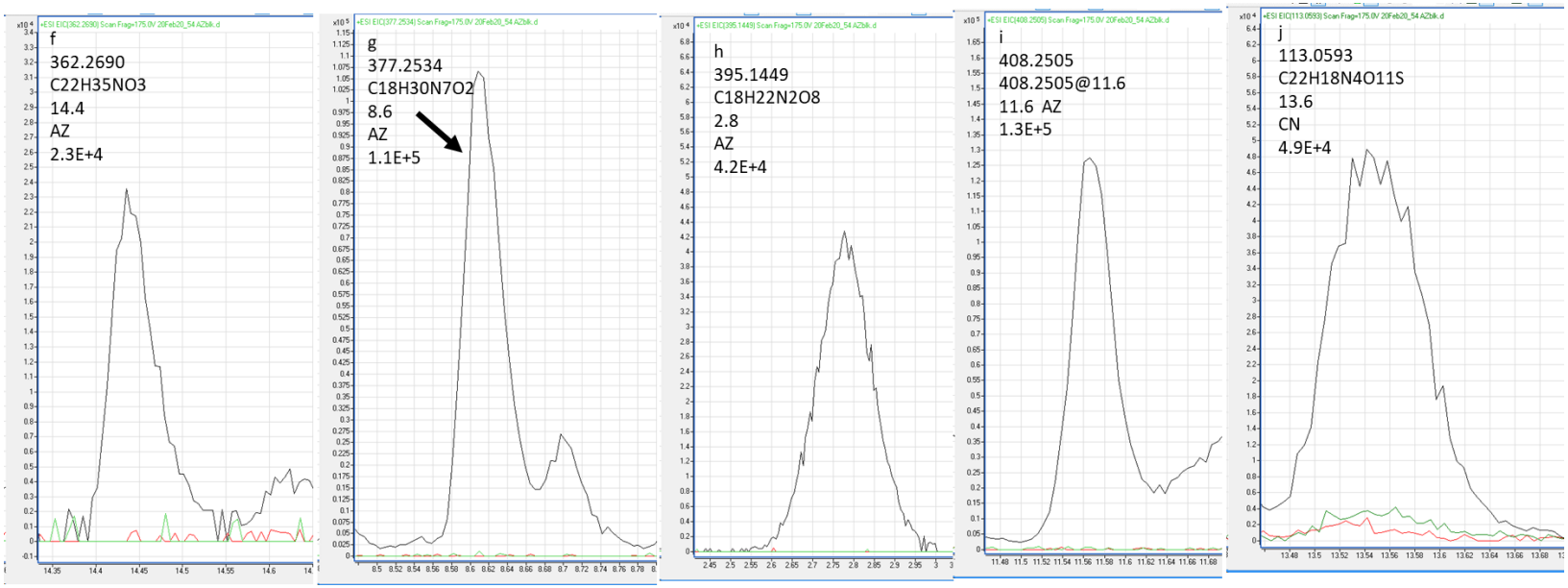
Supplementary Figure B13. EIC of unknown metabolites detected only in sSPE fractions in (-ve) ZIC-HILIC in the global metabolomics analysis of individual fractions. An m/z value, molecular formula, RT (min), fraction' name, and height of the peak indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm. Chromatograms d1, d2, and e1, e2 demonstrate metabolites split between A, N, and C, Z, respectively. Peaks in MeOH (chromatograms c, d1, and d2, red EIC) were below the intensity threshold (7500) and were rejected from further analysis.



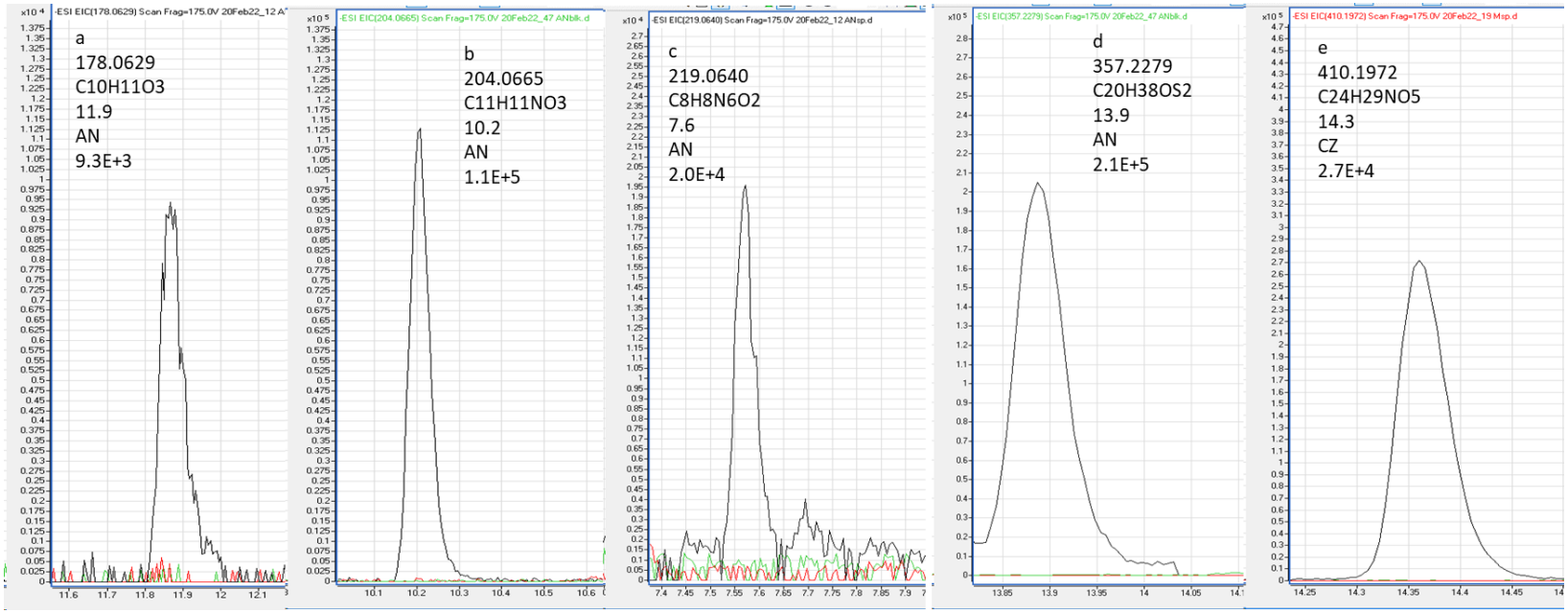
Supplementary Figure B13 (cont'd). Chromatograms f1 and f2 demonstrate the metabolite split between A and N fractions.



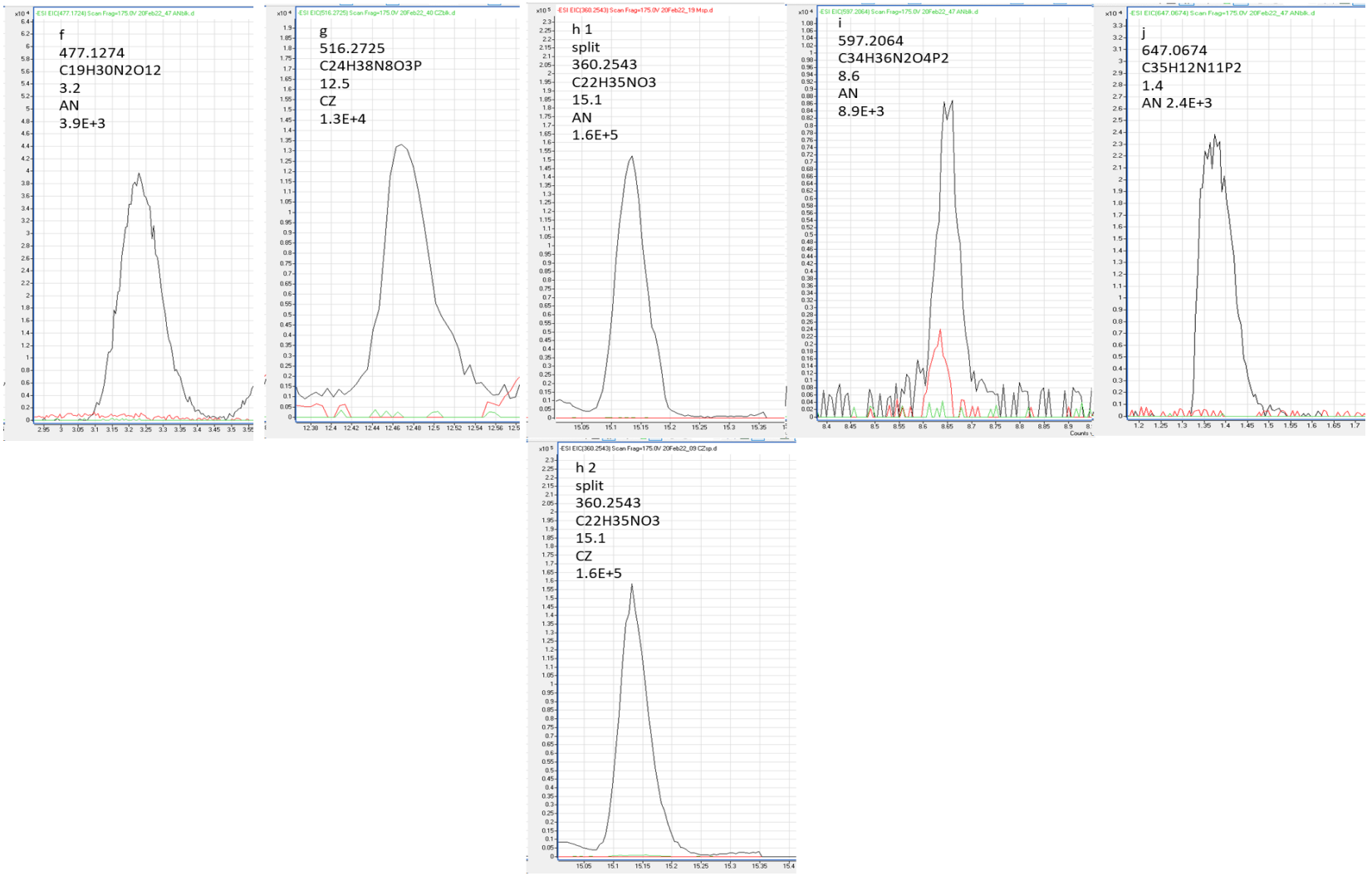
Supplementary Figure B14. EIC of unknown metabolites detected only in sSPE fractions in (+ve) RP in the global metabolomics analysis of combined sSPE fractions. An m/z value, molecular formula, RT (min), fraction' name, and height of the peak indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm. The peak in the "d" was colored to designate the peak of interest.



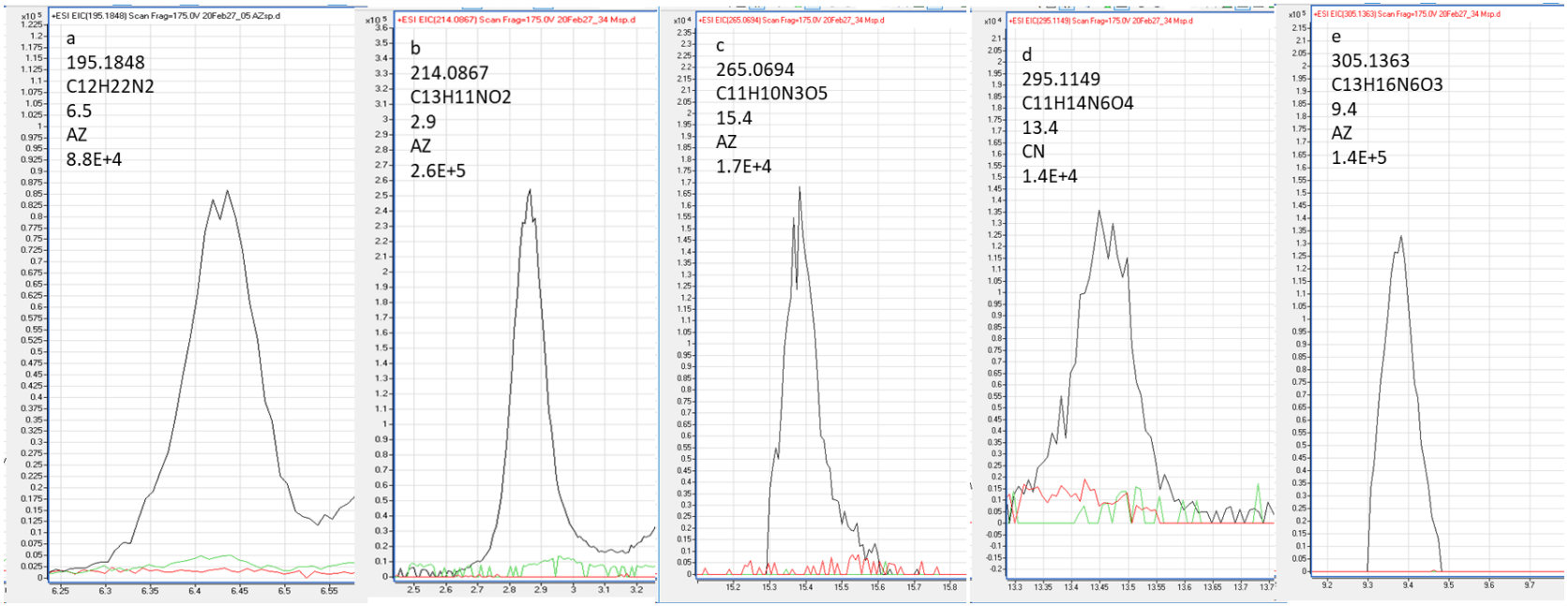
Supplementary Figure B14 (cont'd).



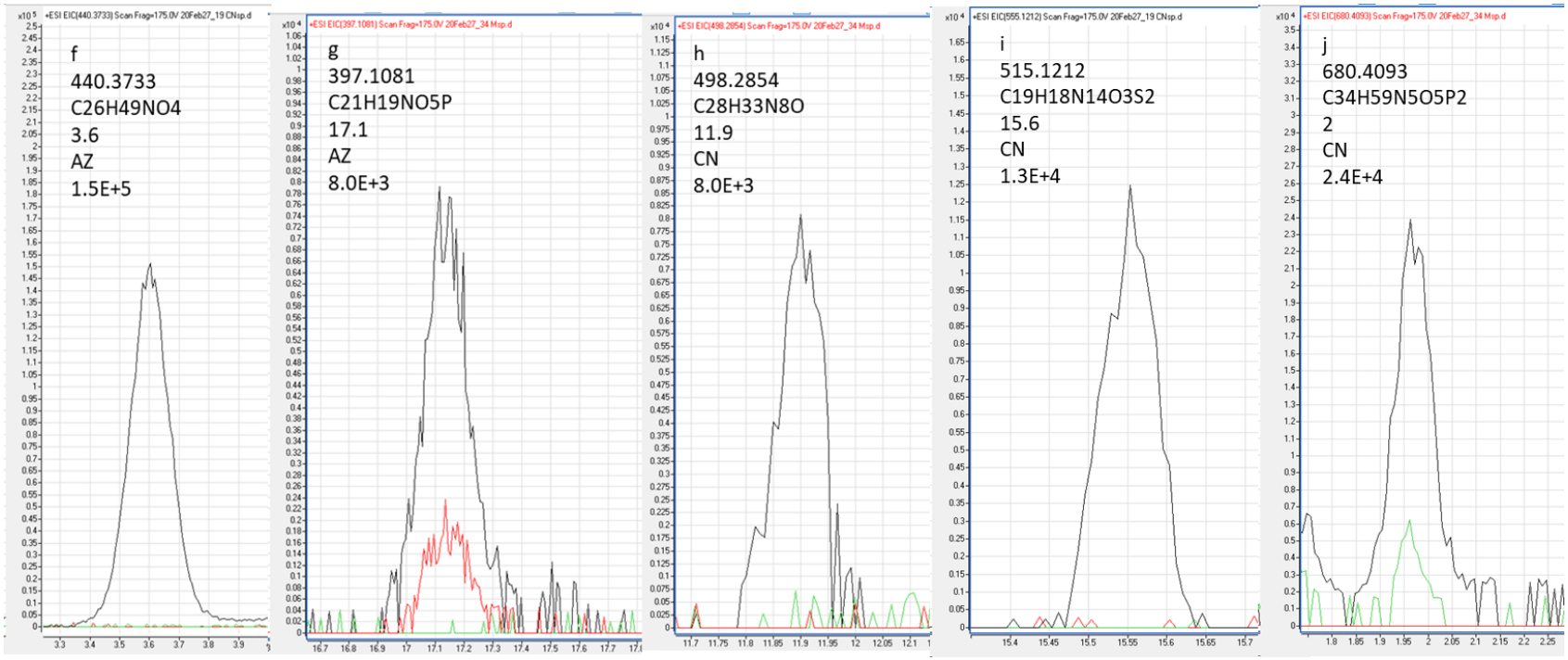
*Supplementary Figure B15. EIC of unknown metabolites detected only in sSPE fractions in (-ve) RP in the global metabolomics analysis of combined sSPE fractions. An m/z value, molecular formula, RT (min), fraction' name, and height of the peak indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted using m/z provided by the Mass Profiler Professional (**see Section 2.3.3.2**) with accuracy ± 20 ppm.*



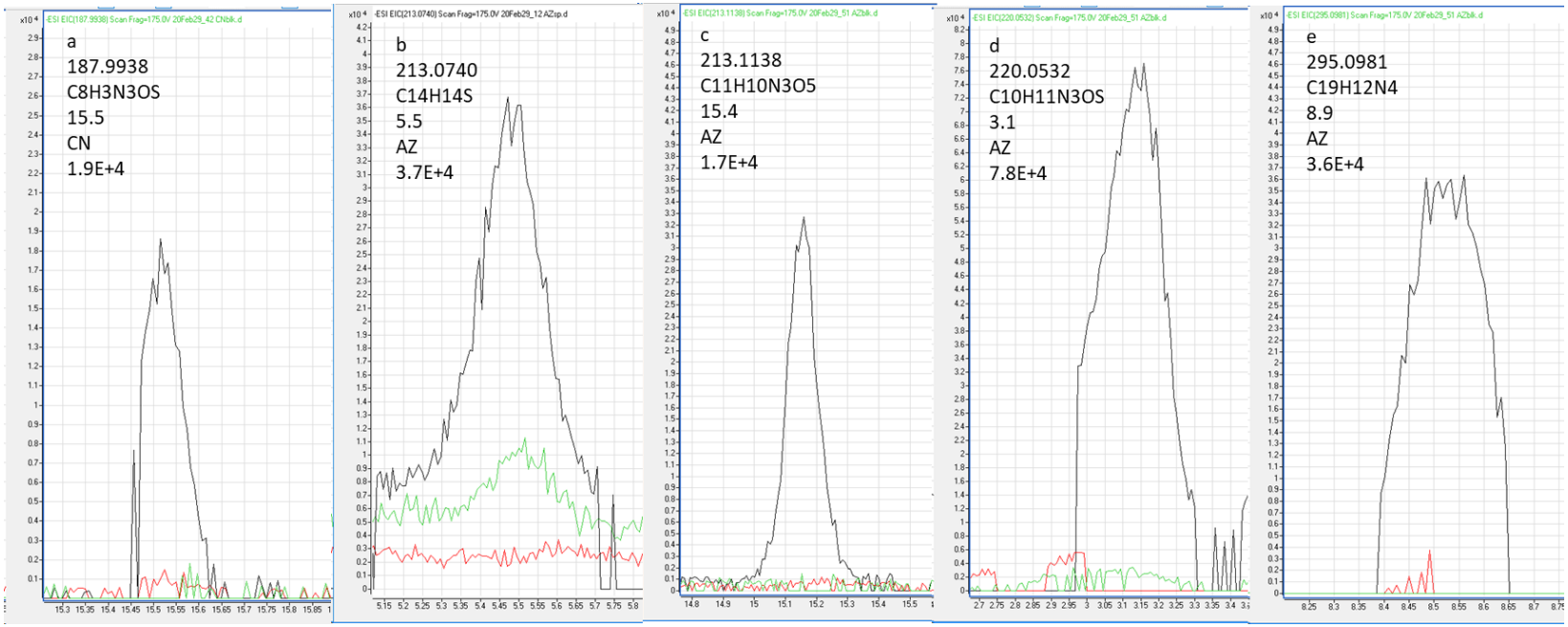
Supplementary Figure B15 (cont'd). The metabolite in h1 and h2 was split (detected in two sSPE fractions) between AN and CZ.



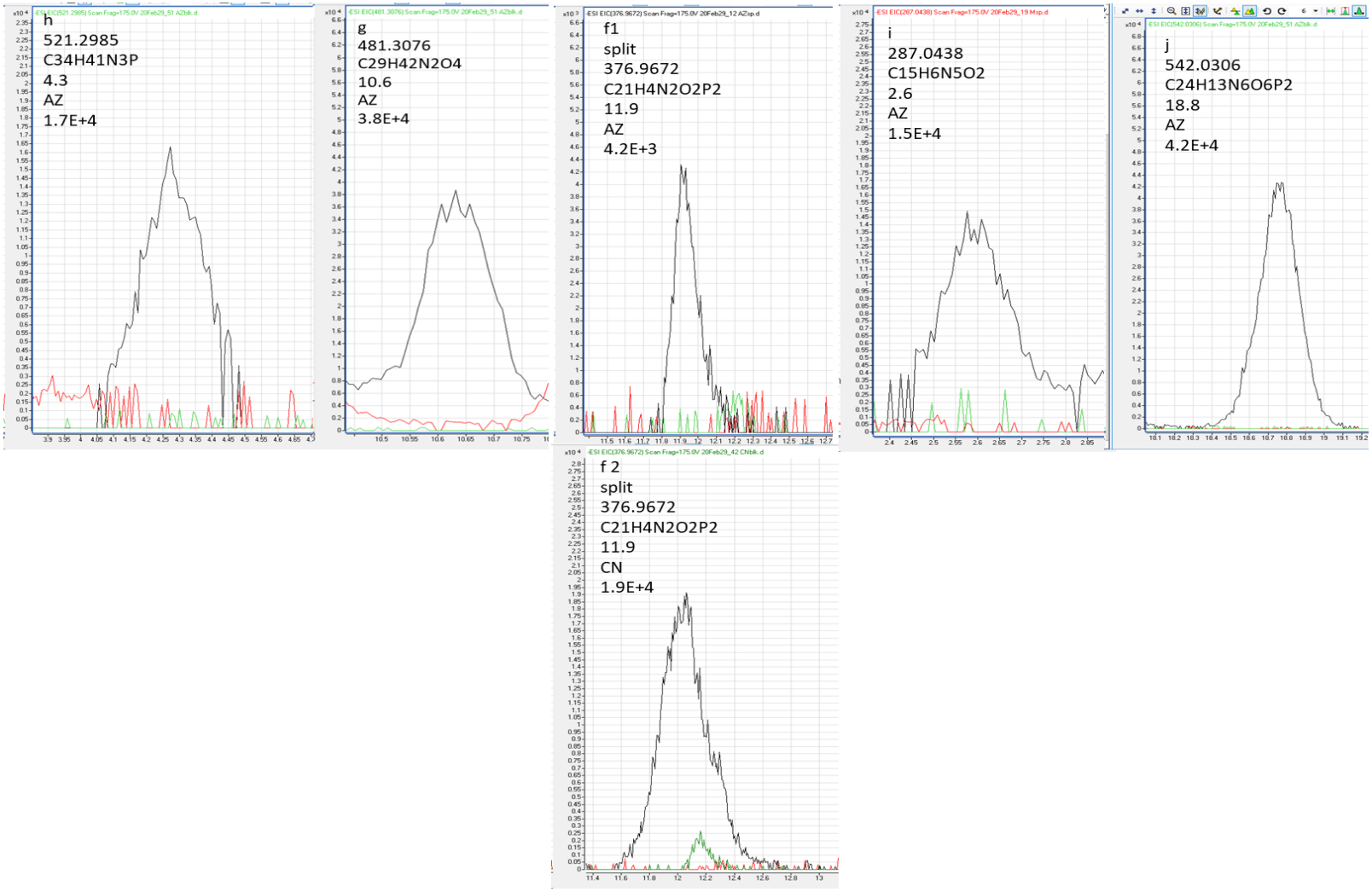
Supplementary Figure B16. EIC of unknown metabolites detected only in sSPE fractions in (+ve) ZIC-HILIC in the global metabolomics analysis of combined sSPE fractions. An m/z value, molecular formula, RT (min), fraction' name, and height of the peak indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm.



Supplementary Figure B16 (cont'd). The height of the peak in the MeOH fraction (chromatogram g, red EIC) is below the intensity threshold (7500) and was rejected from the analysis.



Supplementary Figure B17. EIC of unknown metabolites detected only in sSPE fractions in (-ve) ZIC-HILIC in the global metabolomics analysis of combined sSPE fractions. An m/z value, molecular formula, RT (min), fraction' name, and height of the peak indicated on the top of each chromatogram. The EIC of metabolites in sSPE fraction (black), sSPE blank of the same fraction (green), and MeOH extract (red) were extracted with accuracy ± 20 ppm. The height of the peak in the sSPE blank (chromatogram b, green EIC) is below the intensity threshold (7500) and was rejected from the analysis.



Supplementary Figure B17 (cont'd).

Appendix C

Supplementary information for Chapter 4

Supplementary materials and methods

Activation of charcoal

Charcoal was activated by incubation of 1 gram in 10 mL methanol for 30 mins followed by centrifugation of charcoal-methanol slurry at 4000 x g for 15 mins at ambient temperature, removal of 9 mL of supernatant followed by addition of 10 mL water on the top of the charcoal. The procedure was repeated one more time with water and produced 1 mL of charcoal slurry. One mL of this slurry (equivalent to 0.5 grams of dry charcoal) was used for stripping 9 mL of reconstituted extracted oral fluid.

Supplementary tables and figures

Supplementary Table C1. Repeatability of calibration curve slopes and intercepts during inter-day validation. Calibration curves were prepared fresh daily in a stripped oral fluid matrix.

Experimental day	Cortisol		Cortisone	
	Slope	**Intercept**	**Slope**	**intercept**
Day 1	0.16	-0.002	0.24	-0.009
Day 2	0.16	0.010	0.26	-0.020
Day 3	0.17	-0.002	0.25	-0.001
Day 4	0.15	-0.007	0.23	-0.020
Day 5	0.17	0.008	0.27	0.010
Day 6	0.24	0.010	0.39	0.010
Day 7	0.19	0.001	0.27	0.005
Mean of slope (n=7)	0.18	0.003	0.27	-0.004
SD of slope ((n=7)	0.03	0.007	0.05	0.013

.

*Supplementary Table C2. Analysis of recovery and extraction repeatability in the charcoal-stripped oral fluid matrix. Section A. Recovery was calculated using 6 replicates for each validation sample, and the calibration curve was made in a solvent using the formula: actual concentration in the validation sample*100%/expected concentration. Section B. Recovery was calculated by comparing peak areas in samples spiked pre-extraction (n=6) vs. samples spiked post-extraction (n=1 for each concentration) after correction by the peak area of internal standard, cortisol (d_4). The recovery was calculated by the formula: (peak area ratio in pre-extraction spiked validation sample)*100% / (peak area ratio in a post-extraction spiked sample).*

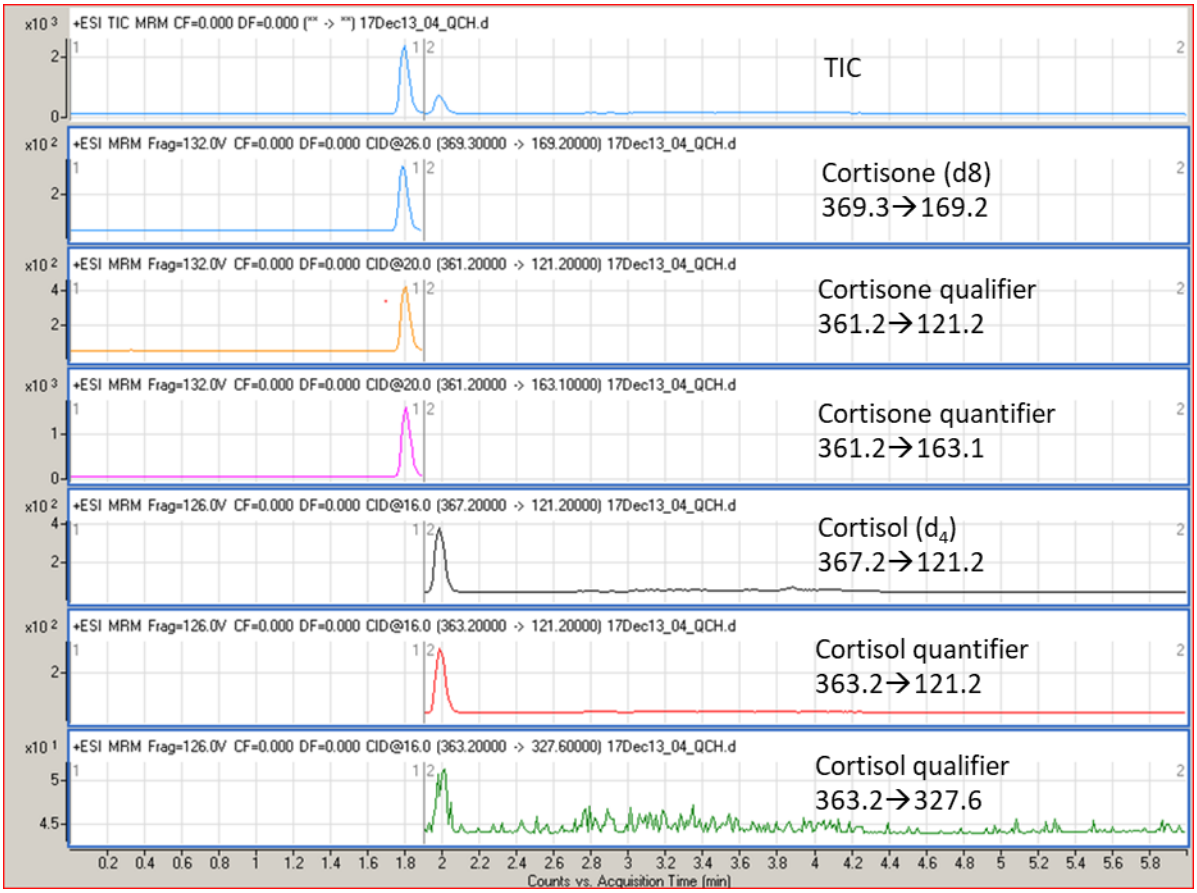
Validation samples (ng/mL)	Cortisol		Cortisone	
	Mean recovery (%)	RSD (%)	Mean recovery (%)	RSD (%)
A. Recovery by quantitation using calibration curve in solvent				
0.54	84.7	5.3	76.9	5.1
1.09	82.3	4.9	77.3	6.0
4.38	86.4	8.9	81.4	10
8.75	75.2	4.7	69.1	5.3
17.5	76.4	5.5	71.1	6.2
B. Recovery by areas in pre- and post-extraction spiked samples				
0.54	84.0	6.0	71.4	6.4
1.09	78.9	5.6	79.1	6.2
4.38	87.5	10.2	75.8	12.3
8.75	76.4	6.3	64.4	7.6
17.5	79.0	7.2	67.5	8.7

Supplementary Table C3. The stability of cortisol and cortisone in the intraday analytical batch consisted of 72 injections with a total length of 8.5 hours. QC samples and calibration curve replicates with different concentrations were prepared as described in ***Sections 3.2.6.2 and 3.2.6.3*** *and injected multiple times during analysis. The first injection and second sequential injections of the batch were QC low (1.5 ng/mL) and QC high (11.5 ng/mL) at "0.0" and 0.11 h, respectively. The replicates of QC low and high were also the two last injections of the batch (71 and 72) injected at 8.3 and 8.4 h of the batch analysis time, respectively. The sequential position of injections is designated in "Seq. injection#". The accuracy of quantitation was calculated against the calibration curve, assuming expected concentrations equal to 100%.*

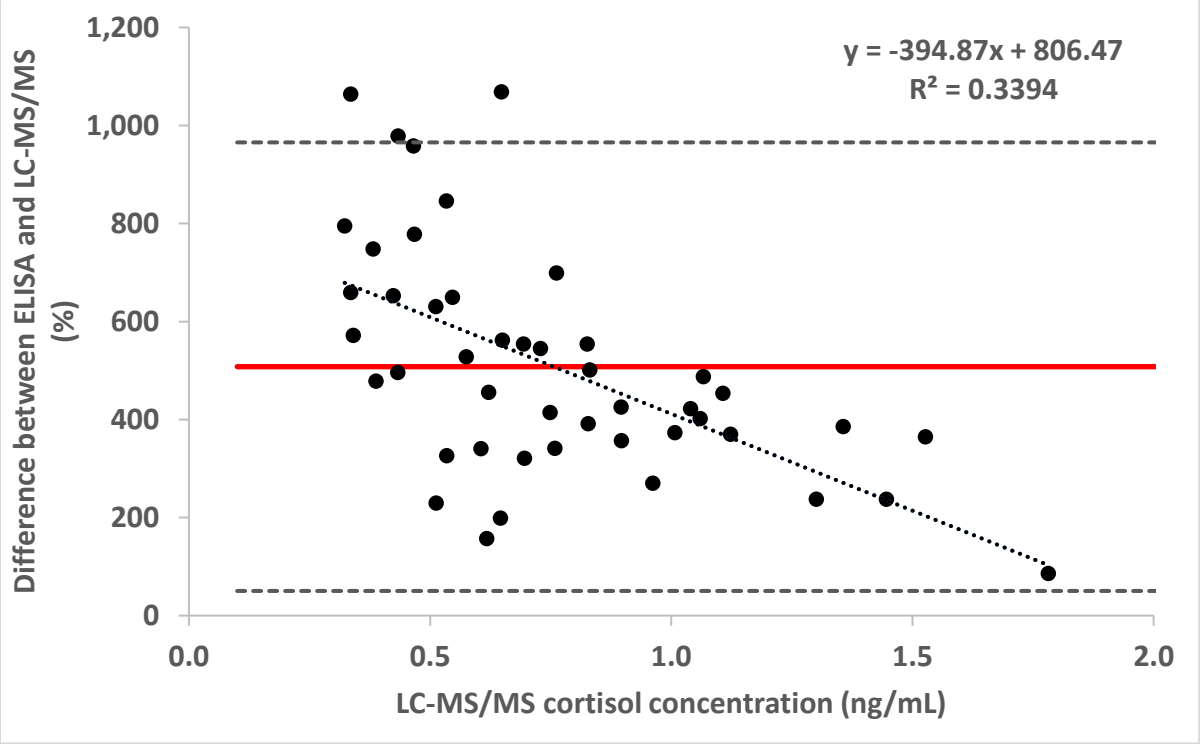
Samples	Expected concentration of cortisol or cortisone (ng/mL)	Sequential injection #	Time of injection (h)	Cortisol concentration ng/mL	Accuracy %	cortisone concentration ng/mL	Accuracy %	Area cortisol (d_4)
Calibration curve 1	1.25	31	3.6	1.2	100.0	1.4	108.4	756
Calibration curve 1	1.25	33	3.9	1.35	110.0	1.2	95.3	693
Calibration curve 1	1.25	40	4.7	1.3	106.3	1.3	101.9	719
Calibration curve 1	1.25	50	5.8	1.2	95.6	1.2	99.1	762
Calibration curve 1	1.25	64	7.5	1.1	91.5	1.2	95.9	728
Calibration curve 2	10	8	0.9	10.6	105.9	9.6	95.8	727
Calibration curve 2	10	19	2.2	9.1	91.0	9.5	95.4	697
Calibration curve 2	10	32	3.7	10.1	101.4	9.7	97.4	715
Calibration curve 2	10	39	4.6	9.2	92.2	8.7	86.6	669
Calibration curve 2	10	49	5.7	9.2	91.5	9.2	92.3	726
Calibration curve 2	10	62	7.2	9.8	97.6	10.0	99.6	729
Calibration curve 3	20	9	1.1	21.3	106.7	20.4	101.8	714
Calibration curve 3	20	20	2.3	19.6	98.2	19.6	97.8	617
Calibration curve 3	20	37	4.3	18.8	94.0	19.2	95.8	684
Calibration curve 3	20	57	6.7	20.4	102.0	18.7	93.3	730
Calibration curve 3	20	66	7.7	20.2	100.9	19.7	98.5	735
Calibration curve 3	20	68	7.9	20.4	101.8	18.5	92.3	735
QC high	11.5	2	0.1	11.2	97.6	11.8	102.9	744
QC high	11.5	23	2.7	11.3	98.1	11.9	103.6	727
QC high	11.5	45	5.3	12.1	105.5	12.7	110.1	690
QC high	11.5	67	7.8	11.9	103.5	12.2	105.9	707
QC high	11.5	72	8.4	11.7	101.8	12.1	105.4	730
QC low	1.5	1	0.0	1.5	101.7	1.6	105.5	711
QC low	1.5	12	1.4	1.5	98.9	1.6	105.0	694
QC low	1.5	34	4.0	1.5	101.9	1.6	105.7	700
QC low	1.5	56	6.5	1.6	103.7	1.6	109.3	680
QC low	1.5	71	8.3	1.5	101.3	1.6	108.8	702

Supplementary Table C4. Comparison of cortisol concentrations measured in oral fluid samples by LC-MS/MS and ELISA and side-by-side comparison to cortisone concentrations measured by LC-MS/MS.

Volunteer ID	Activity and time point	LC-MS cortisol (ng/mL)	LC-MS cortisone (ng/mL)	ELISA cortisol (ng/mL)	Volunteer ID	Activity and time point	LC-MS cortisol (ng/mL)	LC-MS cortisone (ng/mL)	ELISA cortisol (ng/mL)
LCR 20	**act1 A**	0.70	5.21	2.93	**LCR 16**	**act3 C**	0.34	5.44	3.91
LCR 20	**act1 B**	< LLOQ	5.07	3.20	**LCR 16**	**act4 A**	0.62	6.00	3.45
LCR 20	**act1 C**	< LLOQ	3.34	3.31	**LCR 16**	**act4 B**	0.83	5.59	4.07
LCR 20	**act2 A**	1.45	9.68	4.88	**LCR 16**	**act4 C**	0.65	5.92	4.30
LCR 20	**act2 B**	1.36	10.46	6.59	**LCR 18**	**act1 A**	1.04	6.79	5.43
LCR 20	**act2 C**	1.06	7.42	5.32	**LCR 18**	**act1 B**	1.05	17.29	7.57
LCR 20	**act3 A**	1.30	9.67	4.39	**LCR 18**	**act1 C**	1.07	8.72	6.27
LCR 20	**act3 B**	0.73	6.14	4.70	**LCR 18**	**act2 A**	0.39	3.54	2.24
LCR 20	**act3 C**	0.58	7.77	3.61	**LCR 18**	**act2 B**	0.65	6.14	1.93
LCR 19	**act2 A**	0.83	5.47	5.00	**LCR 18**	**act2 C**	0.62	5.09	1.59
LCR 19	**act2 B**	1.11	7.90	6.13	**LCR 18**	**act4 A**	1.53	8.33	7.10
LCR 19	**act2 C**	1.01	8.29	4.77	**LCR 18**	**act4 B**	0.69	5.39	4.54
LCR 19	**act3 A**	1.78	10.17	3.31	**LCR 18**	**act4 C**	1.12	7.68	5.28
LCR 19	**act3 B**	0.83	7.80	5.40	**LCR 14**	**act1 A**	0.43	3.49	2.58
LCR 19	**act3 C**	0.51	5.39	3.74	**LCR 14**	**act1 B**	0.53	4.83	2.28
LCR 19	**act4 A**	0.47	4.17	4.93	**LCR 14**	**act1 C**	0.34	3.17	2.55
LCR 19	**act4 B**	0.76	6.43	6.09	**LCR 14**	**act2 A**	0.65	4.34	7.57
LCR 19	**act4 C**	0.43	4.53	4.68	**LCR 14**	**act2 B**	0.42	3.56	3.19
LCR 16	**act1 A**	0.38	4.83	3.24	**LCR 14**	**act2 C**	0.51	3.74	1.69
LCR 16	**act1 B**	0.34	3.72	2.29	**LCR 14**	**act3 A**	0.90	7.56	4.71
LCR 16	**act1 C**	< LLOQ	3.20	2.07	**LCR 14**	**act3 B**	0.75	7.61	3.85
LCR 16	**act2 A**	0.55	4.20	4.10	**LCR 14**	**act3 C**	0.53	4.75	5.05
LCR 16	**act2 B**	0.32	4.15	2.89	**LCR 14**	**act4 A**	0.90	6.54	4.10
LCR 16	**act2 C**	< LLOQ	3.15	4.93	**LCR 14**	**act4 B**	0.96	5.73	3.56
LCR 16	**act3 A**	0.76	5.61	3.35	**LCR 14**	**act4 C**	0.61	4.47	2.67
LCR 16	**act3 B**	0.47	4.33	4.11					



Supplementary Figure C1. Example chromatograms of QCH (quality control high) with cortisone and cortisol concentrations of 5 ng/mL prepared in a stripped oral fluid matrix.



Supplementary Figure C2. Bland-Altman analysis for cortisol measurements using LC-MS/MS and ELISA by plotting LC-MS/MS assay concentration on the X-axis and a relative concentration differences (in %) obtained by two methods on the Y-axis. The red line shows the mean relative difference between the assays, while the dashed lines show limits of agreement as determined by ± 1.96xSD. Linear least-squares regression was also performed, and its results are shown in the top right corner.