Metabolomic Assessment of Dietary Interventions in Obesity by Capillary Electrophoresis Mass Spectrometry

Metabolomic Assessment of Dietary Interventions in Obesity by Capillary Electrophoresis Mass Spectrometry

By

Karen Phoebe Lam, H.B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

Doctor of Philosophy (2018) McMaster University
(Department of Chemisty & Chemical Biology) Hamilton, Ontario

TITLE: Metabolomic Assessment of Dietary Interventions in Obesity by Capillary Electrophoresis Mass Spectrometry

AUTHOR: Karen Phoebe Lam, H.B.Sc. (University of Toronto), M.Sc. (McMaster University)

SUPERVISOR: Professor Philip Britz-McKibbin

NUMBER OF PAGES: xxix, 347

Abstract

Capillary electrophoresis mass spectrometry (CE-MS) is a versatile instrumental method for metabolomics, which allows for comprehensive metabolite profiling of volume-limited biological specimens in order to better understand the molecular mechanisms associated with chronic diseases, including an alarming epidemic of obesity worldwide. Multiplexed CE separations enable high-throughput metabolite screening with quality assurance to prevent false discoveries when combined with rigorous method validation, robust experimental designs, complementary statistical methods, and high-resolution tandem mass spectrometry (MS/MS) for unknown metabolite identification. In this thesis, multiplexed CE-MS technology is applied for both targeted and untargeted metabolite profiling of various biological fluids, including covalently bound thiol-protein conjugates, as well as free circulating metabolites in serum and plasma, and excreted/bio-transformed compounds in urine due to complex host-gut microflora co-metabolism. This work was applied to characterize aberrant metabolic responses of obese subjects in response to dietary challenges, and measure the benefits of dietary interventions that reduce adiposity without deleterious muscle loss. Chapter 2 presents, a simple, sensitive yet robust analytical protocol to expand metabolome coverage in CE-MS for the discovery of labile protein thiols in human plasma using a rapid chemical derivatization method based on *N-tert*-butylmaleimide (NTBM). Chapter 3 describes targeted metabolite profiling of serum and plasma to investigate the differential metabolic responses between healthy and unhealthy obese individuals before and after consumption of a standardized high-caloric meal, respectively. Chapter 4 of this thesis describes an untargeted metabolite profiling strategy for urine using multisegment-injection (MSI)-CE-MS for elucidating the effects

of protein supplementation following a short-term dietary weight-loss intervention study. This work revealed six urinary metabolites that were classified as top-ranking treatment response biomarkers useful for discriminating between subjects consuming carbohydrate (control), soy, and whey supplemented diets. In summary, this thesis demonstrated the successful implementation of multiplexed CE-MS technology for biomarker discovery in nutritional-based metabolomic studies as required for more effective treatment and prevention of obesity for innovations in public health.

Acknowledgements

This thesis would not have been possible without the collective contributions of the many people whom I have encountered throughout the course of my graduate research. First, I would like to sincerely thank my supervisor, Prof. Philip Britz-McKibbin, who welcomed me into his research group and introduced me to the wonderful world of CE. He has provided me with many opportunities throughout my graduate studies, in which I have gained a lot of invaluable hands-on training and experience. He has guided my research journey with unparalleled optimism and an unforgettable adventure-seeking spirit, in the same way that he has guided me and fellow group members through rugged treks to find zen (aka, "the waterfall") during our memorable annual Britz-group lab hikes. I would also like to express my appreciation to supervisory committee members, Prof. Murray Potter, Prof. Jose Moran-Mirabal and the late Prof. Brian McCarry, as well as Kirk Green from the McMaster Regional Centre for Mass Spectroscopy, for their insightful questions, advice, support, encouragement, and technical assistance over the years. Additionally, I would like to thank my collaborators, Prof. David Mutch (University of Guelph) and Prof. Stuart Phillips (Department of Kinesiology), for their contributions and involvement in the various studies described in this thesis, aside from the stimulating discussions during our project-update meetings.

The cordial lab atmosphere, collaborative group dynamics, and sense of camaraderie amongst the Britz group members have played an important role in enhancing my

graduate student experience and cannot be overstated. I would like to thank past and present lab members in the Britz group for, not only sharing their advice, ideas and providing assistance when needed, but also, for their individual contributions to creating a warm, supportive, and friendly working atmosphere: Lisa D'Agostino, Naomi Kuehnbaum, Meera Shanmuganathan, Alicia DiBattista, Nadine Wellington, Adriana Macedo, Michelle Saoi, Mai Yamamoto, Jennifer Wild, Stellena Mathiaparanam, Ritchie Ly, Sandi Azab, and Biban Gill. Outside of the lab environment, I am very grateful for all the friends that I have made at McMaster, in addition to my long-time friends from Toronto, as well as those living out-of-province. Thank you for all the fun times we had together, whether it be going for a stroll or a bike ride; enjoying a movie, bowling or board game night; going out for a nice dinner or being invited for potluck or a delicious home-cooked meal; indulging in desserts; or even, as simple as a telephone call or text message. And of course, thank you for always rooting for me. All these gestures mean so much to me and I am very blessed to have your friendship. Special thanks go to my friend, Fan Fei, for her kind support and graciously offering her time and expertise to help me with particular revisions of this thesis.

My parents, Connie and Joseph, my brother, Eugene, and all other family members have played a significant role in my life and I would not be here without them. I thank them all for their unconditional love, sacrifices, patience, and continuous support throughout this journey. In addition, I thank my Mom for giving me the gift of music, along with its natural healing powers; I thank my Dad for teaching me patience, forgiveness, and the art of prayer; and I thank my brother for making me feel valued and important in this world as an older sister. As well, I also thank Patrick Wilson and his parents, Anna and Robert, for their care, kindness, love, guidance, and encouragement. In particular, I express my deepest heartfelt gratitude to Patrick, who has been by my side for most of this adventure and especially, at the most

difficult times. He has guided and supported me tremendously in the completion of this journey in more ways than words can ever describe. He has selflessly devoted countless hours, days, weeks, and months of his personal time and energy to my cause, including assisting with multivariate data analysis and completing multiple rounds of meticulous typesetting, formatting, and editing of this entire thesis. Most importantly, he has encouraged and emboldened me to reach the finish line of this pursuit with his abundant love and immense support – and by firmly believing in me, much more than I believe in myself. Finally, I thank God for giving me endurance and personal strength, and for everyone in my life that has helped me to be where I am today.

This thesis is dedicated to three very special people in my life:

My 90-year-old energetic grandmother, Chiu Lai Har, a recent breast cancer survivor whose love, support, courage, and fearless determination have inspired me to complete this work;

My long-time mentor and friend of 20 years, Toronto-based pianist and performer Peter Longworth, whom I will always remember as being a major influence in my life and on my music as a pianist, up until his sudden passing at an early age due to kidney cancer on June 26, 2018; and

My late Godfather, Vincent Lam, who passed away on February 1, 2018 after a heroic, yet painful, battle with lymphoma. He celebrated life up until his last day and was the one to witness my first steps as a toddler and would later inspire me to pursue and develop a passion for the sciences.

Table of Contents

List of Figures

List of Tables

List of Acronyms & Abbreviations

–SH	sulfhydryl	49
–SiOH	silanol	9
2D	two-dimensional	195
3-Cl-Tyr	3-chlorotyrosine	52
3-MeHis	3-methylhistidine	45
3D	three-dimensional	54
4-EP	4-ethylphenol	225
4-EPS	4-ethylphenyl sulfate	45
4-MPS	*p*-cresol sulfate (4-methylphenyl sulfate)	224
AA	amino acid	140
Ace-K	acesulfame potassium	45
ACN	acetonitrile, CH_3CN	53
ADMA	asymmetric dimethylarginine	219
AGAT	arginine:glycine amidinotransferase	213
α	probability rate of a false positive or type I error	39
AM1	Austin Model 1	59
ANOVA	analysis of variance	29
APCI	atmospheric pressure chemical ionization	13
APPI	atmospheric pressure photoionization	13

ASA	argininosuccinic acid	212
Asp-His	aspartyl-histidine	204
AUC	area under the curve	126
BCAA	branched-chain amino acid	121
BGE	background electrolyte	9
BGS	bile acid-glycine-sulphate conjugate	196
BMI	body mass index	22
BP	blood pressure	160
BQB	ω-bromoacetonylquinolinium bromide	51
C0	free carnitine	204
C_4H_8	2-methylpropene	103
C_4H_9	tertiary butyl moiety	103
CAT	cysteine aminotransferase	216
CE	capillary electrophoresis	7
CEC	capillary electrochromatography	8
CEM	chain ejection model	14
CGE	capillary gel electrophoresis	8
CHO	carbohydrate	172
CID	collision-induced dissociation	100
CIEF	capillary isoelectric focusing	8
CITP	capillary isotachophoresis	8
CKD	chronic kidney disease	231
CO	carbon monoxide	100
CoA	coenzyme A	136
CPT1	carnitine palmitoyltransferase I (carnitine acyltransferase I)	207
CRF	chronic renal failure	225

CRM	charge residue model	14
Crn	creatinine	27
CV	coefficient of variation	178
CV-ANOVA	cross-validated residuals analysis of variance	120
CVD	cardiovascular disease	141
Cys	cysteine	48
Cys34	the free cysteine residue of human serum albumin	67
CysGly	cysteinylglycine	48
CZE	capillary zone electrophoresis	8
DA	discriminant analysis	33
DF	discriminant function	201
DFA	discriminant function analysis	201
DHD	dihydrodaidzein	222
DI	deionized	52
DNA	deoxyribonucleic acid	49
DRA	Diabetes Risk Assessment	111
DTNB	5,5-dithio-bis2-nitrobenzoic acid, also known as Ellman's reagent	50
DTT	dithiothreitol	68
E	electric field applied	11
EDTA	ethylenediaminetetraacetic acid	23
EE^+	even-electron	102
EI	electron impact	6
EIE	extracted ion electropherogram	57
EOF	electroosmotic flow	9
ε_0	permittivity of vacuum	11

ε_r	dielectric constant	11
ESI	electrospray ionization	13
η	viscosity	11
FA	fatty acid	140
FDA	Food and Drug Administration	229
FDR	false discovery rate	41
FEM	*N*-(2-ferrocene-ethyl)maleimide	51
FL	fluorescence	50
FT-ICR	Fourier-transform inductively coupled resonance	7
FWER	family-wise error rate	41
γ-GluCys	γ-glutamyl-cysteine	48
GC	gas chromatography	3
GFR	glomerular filtration rate	46
glog	generalized logarithm	29
Glu	glutamic acid	49
Gly	glycine	49
GSA	guanidinosuccinic acid	198
GSH	glutathione	48
H_2O	water	52
HbA1c	glycated hemoglobin	108
HCA	hierarchical cluster analysis	32
Hci	homocitrulline	174
HCl	hydrochloric acid	145
Hcy	homocysteine	48
HDL-c	high-density lipoprotein cholesterol	112
HMDB	Human Metabolome Database	95

HOMA-IR	homeostatic model assessment of insulin resistance	108
HOMA%B	homeostatic model assessment for β-cell function	160
HP-0321	hexamethoxyphosphazine	175
HP-0921	hexakis(2,2,3,3-tetrafluoro-propoxy)phosphazine	175
HPC	hydroxypropyl cellulose	8
HPLC	high-performance liquid chromatography	11
HSA	human serum albumin	67
ID	identification	111
IDC	identification confidence	239
IEM	ion evaporation model	14
IS	internal standard	52
iso-C4	isobutyrylcarnitine	196
IT	ion-trap	54
KCl	potassium chloride	146
KEGG	Kyoto Encyclopedia of Genes and Genomes	119
KM	K-means	32
LC	liquid chromatography	7
LDL-c	low-density lipoprotein cholesterol	112
LH	lean healthy	108
LIF	laser-induced fluorescence	12
LMW	low-molecular-weight	13
LOD	limit of detection	78
LOESS	locally weighted smoothed scatterplot	180
LV1	first latent variable, predictive component	114
LV2	second latent variable, orthogonal component	114
m/z	mass-to-charge ratio	7

MANOVA	multivariate analysis of variance	41
MEKC	micellar electrokinetic chromatography	8
MeOH	methanol	52
MESNA	2-mercaptoethane-sulfonate	52
Met	methionine	48
MFE	Molecular Feature Extractor	88
MFG	Molecular Formula Generation	89
MHO	metabolically healthy obese	108
MLCD	3-mercaptolactate-cysteine-disulfide	198
MPS	myofibrillar protein synthesis	170
MS	mass spectrometry	5
MS/MS	tandem mass spectrometry	45
MSC	Molecular Structure Correlator	89
MSI	multisegment-injection	16
MST	3-mercaptopyruvate sulfurtransferase	216
MSTFA	*N*-methyl-trimethylsilyl-trifluoroacetamide	112
MTBE	methyl-tert butyl ether	26
μ_{ep}	electrophoretic mobility	9
MUFA	monounsaturated fatty acid	150
MUO	metabolically unhealthy obese	108
MV	molecular volume	9
MW	molecular weight	1
MWCNT	multi-walled carbon nanotube	50
N_2	nitrogen	54
NaOH	sodium hydroxide	52
NH_3	ammonia	100

NH_4Ac	ammonium acetate	52
NH_4OH	ammonium hydroxide	145
NMR	nuclear magnetic resonance spectroscopy	5
NMS	2-naphthalene(mono)sulfonate	174
NOS	nitric oxide synthases	213
NS	nonsignificant	123
NTAM	*N*-[2-(trimethylammonium)-ethyl]maleimide	48
NTBM	*N-tert*-butylmaleimide	44
O-DMA	*O*-desmethylangolensin	222
OGTT	oral glucose tolerance test	142
OPLS	orthogonal projection to latent structures	33
OSC	orthogonal signal correction	33
OTC	ornithine transcarbamylase	213
PAA	polyacrylamide gel	8
PC	principal component	31
PCA	principle component analysis	31
PEG	polyethylene glycol	8
%CV	percent coefficient of variation	62
%PP	percent postprandial change	145
%RSD	percent relative standard deviation	71
pH	potential of hydrogen	9
pK_a	logarithmic acid dissociation constant	9
PLS	partial least squares	33
Post	the morning following the intervention period	173
PQ	probabilistic quotient	179
PQN	probabilistic quotient normalization	27

Pre	the morning prior to start of intervention	173
PRMT	protein-arginine methyltransferase	219
PUFA	polyunsaturated fatty acid	150
PVA	polyvinyl alchohol	8
Q	charge state	93
Q-Q	quantile-quantile	40
Q-TOF	quadrupole-time-of-flight	88
QC	quality control	27
RF	random forest	37
RMT	relative migration time	92
RNA	ribonucleic acid	3
RNS	reactive nitrogen species	67
ROC	receiver operating characteristic	117
ROS	reactive oxygen species	67
RPA	relative peak area	57
RPH	relative peak height	71
SAA	sulfur amino acid	217
SAT	subcutaneous adipose tissue	108
SDMA	symmetric dimethylarginine	204
SEM	standard error of the mean	160
SFA	saturated fatty acid	150
SNR	signal-to-noise ratio	88
SOP	standard operating procedure	23
SOY	soy protein	172
SPE	solid-phase extraction	26
SSI	sonic spray ionization	13

SUS	shared and unique structures	114
T120	postprandial time point, corresponding to 120 min following dietary intervention	144
T2D	type 2 diabetes	109
TCA	tricarboxylic acid	45
TCEP	tris(2-carboxyethyl)phosphine hydrochloride	68
TG	triglycerides	112
TIC	total ion chromatogram	114
TMCS	trimethylsilyl chloride	112
TML	trimethyllysine	204
TOF	time-of-flight	7
Total-c	total cholesterol	112
UMDB	Urine Metabolome Database	206
UPLC	ultra-performance liquid chromatography	86
UV	ultraviolet	12
v_{EOF}	velocity of the electroosmotic flow	11
V_{cap}	capillary voltage	88
VIP	variable importance to the projection	35
VIS	visible	64
WHY	whey protein	172
ζ	zeta potential	11

Chapter 1

Defining Metabolomics and its Role in Health and Nutrition

1.1 Overview of Metabolomics

In simplest terms, metabolomics may be defined as the comprehensive assessment of endogenous or exogenous metabolic products ('metabolites') with molecular weight (MW) of up to 1.5 kDa in a biological system, including a single cell, tissue extract, biofluid, or fecal matter (Matysik et al., 2016; Tebani et al., 2016; Wishart, 2005). Metabolites represent the downstream end-products of a complex array of metabolic processes that are closely associated with phenotype. Additionally, metabolites are tightly regulated by proteins, encoded by the genome, that are susceptible to modifications by the presence of pathophysiological conditions, as well as from the exposure to various external stressors (e.g., environment, drug, and diet) (Deidda et al., 2015; Sébédio and Brennan, 2014; Spratlin et al., 2009; Wishart, 2016). As a result, metabolomics provides a snapshot in time of the current physiological state of an organism (Hoerr and Vogel, 2013; Peng et al., 2015). When combining metabolite

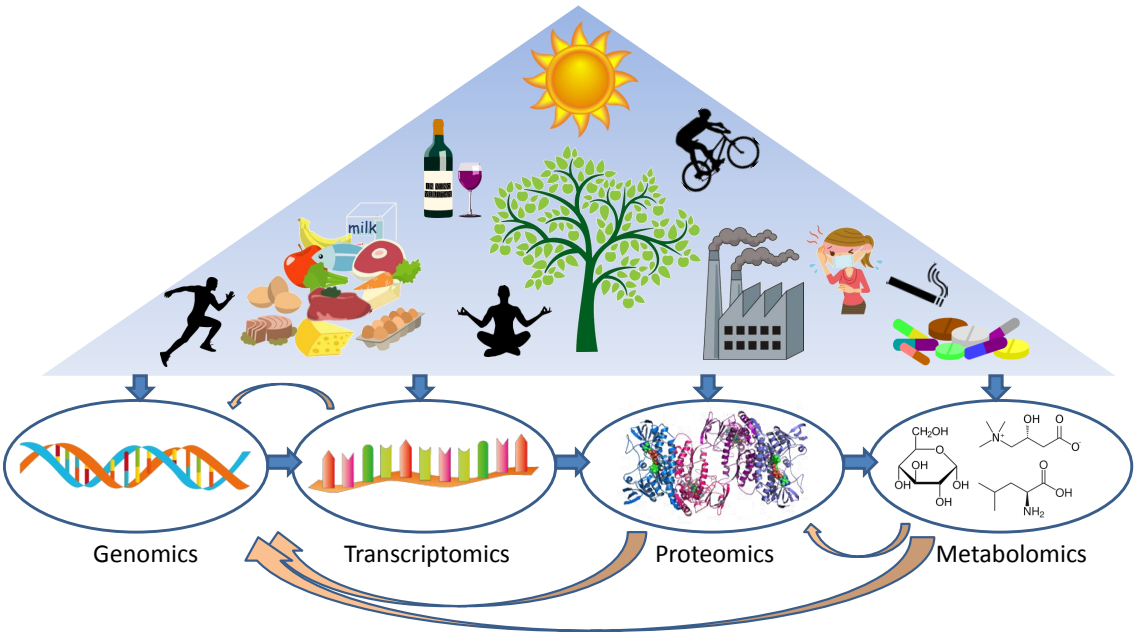


Figure 1.1: An overview of the four major "omics" fields, indicating the flow of biological information from genomics to metabolomics via biochemical signaling. Each "omics" level of systems biology is influenced by lifestyle and environmental factors, leading to downstream products (i.e., metabolites) of cellular regulatory processes. Metabolites can in turn, modulate the activity of proteins and regulate gene expression upstream in a feedback loop through protein modification. Metabolomics, being the final step in the omics cascade, is most closely related to the biological phenotype.

profiling with multivariate statistical methods and robust experimental designs, deeper insight into a biological system's underlying health status and condition can be obtained, and the discovery of novel predictive, prognostic, or diagnostic markers of disease may be revealed. As a result, metabolomics has become an invaluable tool in clinical medicine for the early screening and detection of a wide range of human disorders, including monitoring of treatment responses to therapeutic interventions and understanding of the mechanisms of disease pathogenesis (Gowda et al., 2008). The latter are examples of metabonomics, a subfield of metabolomics that is concerned with the multiparametric quantitative modeling of biological responses to genetic manipulation or biological stimuli, such as diet, drugs, disease, or environmental stressors (Nicholson and Lindon, 2008; Ramsden, 2009).

The concept of metabolomics was first introduced in the work by Pauling et al. in 1971, in which over 500 organic compounds in human breath and urine vapour were quantified by gas chromatography (GC). However, the term 'metabolome' was first coined in 1998 by Oliver et al., and was used to refer to the set of metabolites produced by an organism mainly as a way to elucidate the function of unknown genes in support of expanding genomic sequencing projects. Shortly afterwards, 'metabolomics' rapidly established itself to be an important functional genomics tool and the newest member of the family of holistic "-omics" technologies in systems biology, which include genomics (analysis of genes), transcriptomics (analysis of transcripts: messenger and functional ribonucleic acids (RNAs), and proteomics (analysis of proteins) (Demetrowitsch and Schwarz, 2014) (Figure 1.1). Even after its debut 20 years ago, metabolomics is still quickly evolving in terms of technological development for biological/biomedical applications and continues to be regarded as a rapidly emerging field while addressing major technical hurdles, including unknown metabolite identification (Clish, 2015; Wishart, 2016). Ever since its introduction to the scientific community in 1998, the field of metabolomics has generated nothing short of an astonishing upsurge in publications, which has only been growing at an exponential rate, year after year (Figure 1.2). Metabolomics is increasingly used in diverse applications of translational research, including the monitoring of organ transplants (Wishart, 2005), development of drugs (Wishart, 2016); understanding metabolism of drugs and xenobiotics (Das et al., 2016), outcome prediction of drug interventions ('pharmacometabolomics') (Burt and Nandal, 2016), assessment of exposure to the environment, toxins, and contaminants ('exposomics') (Johnson et al., 2017; Lankadurai et al., 2013; Wagner et al., 2017), and monitoring of exercise interventions (Bally et al., 2017; Duft et al., 2017; Kuehnbaum et al., 2015; Muhsen Ali et al., 2016), as well as dietary interventions (Badoud et al., 2015a; Brennan, 2013; Gibbons and Brennan, 2017; Hanhineva, 2015;

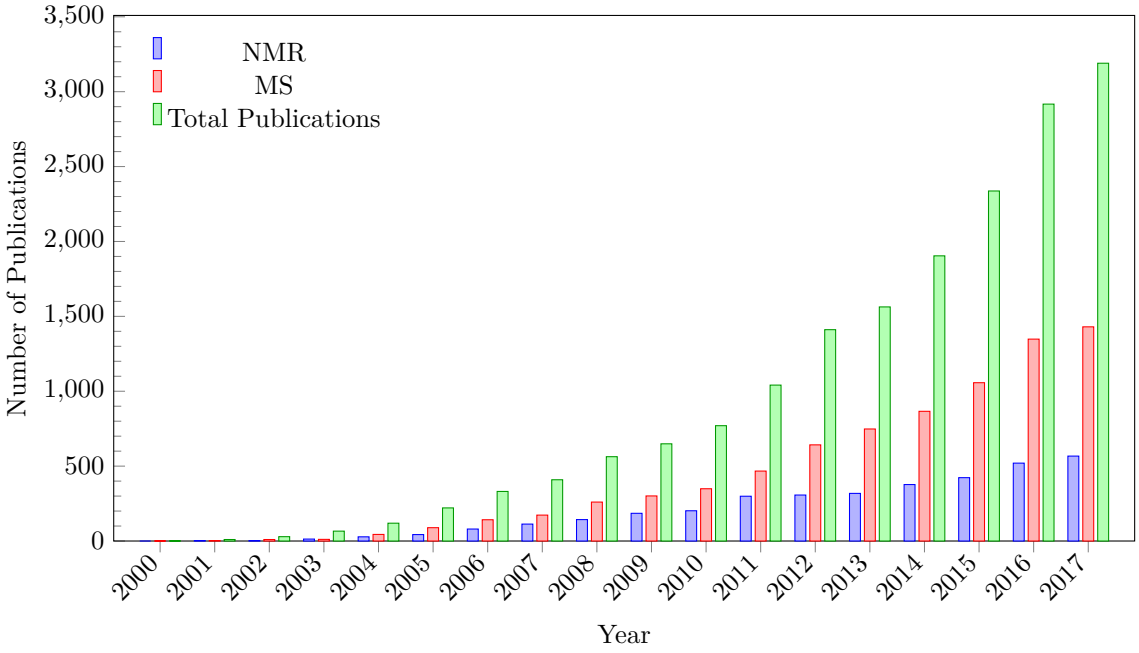


Figure 1.2: Representation of the number of published metabolomics journal articles and comparison of the number of MS- and NMR-based metabolomics publications from 2000 to 2017 when using keywords "metabolomics," "mass spectrometry metabolomics," and "NMR metabolomics," respectively, in Web of Science search.

Hector et al., 2015; Vázquez-Fresno et al., 2015).

Metabolomics studies may be conducted in a targeted or untargeted manner (Begou et al., 2017; Wang et al., 2010). Targeted approaches are hypothesis-driven, meaning that they aim to address particular biological questions or validate critical pathways by selectively quantifying a specific set of known and well-annotated compounds related to the biochemical pathway(s) in question (Dudley et al., 2010; Roberts et al., 2012; Vinaixa et al., 2012; Wang et al., 2010). In contrast, untargeted methods are hypothesis-free and aim to generate new hypothesis by providing a more universal and global characterization of a large number of known and unknown metabolites in an unbiased manner (Alonso et al., 2015; Begou et al., 2017; Horgan and Kenny, 2011; Schrimpe-Rutledge et al., 2016). Due to the lack of pre-specified hypothesis, untargeted approaches are often referred to as 'top-down' methods (Alonso et al.,

2015), in which the analyst aims to maximize the number of detected metabolites for revealing new insights into biological mechanisms (Patti et al., 2012) and for discovering unexpected metabolite changes in terms of relative concentrations, in response to the manipulation of a biological system's condition (Newgard, 2017). However, recently, metabolomics studies are also beginning to explore the merging of targeted (hypothesis-testing) and untargeted (hypothesis-generating) approaches in order to independently validate known metabolites or pathways while also revealing unknown and poorly understood processes associated with complex human disorders (Cajka and Fiehn, 2016).

1.2 Instrumental Methods in Metabolomics

Technological advancements in instrument sensitivity, selectivity, and throughput have resulted in the resolution and efficient detection of hundreds, or thousands, of metabolites simultaneously – which is still only a small fraction of the approximately 114 000 detectable metabolites that the human body is believed to contain (Wishart et al., 2018). Owing to the wide ranges of chemical complexity and metabolite concentrations, there is no single analytical platform that can characterize the human metabolome in its entirety (Zhang et al., 2012a). Currently, mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) are the leading technologies in metabolomics research (Markley et al., 2017). Complementary to each other, each technology has its own strengths and weaknesses, and may be optimally selected based on the goal of the study and the chemical composition of the metabolites of interest (Patti, 2011). Of course, using a combination of platforms enables greater coverage of various metabolite classes and sample types (Bouatra et al., 2013; Psychogios et al., 2011). Although NMR is recognized as the gold standard in metabolite structural

Table 1.1: Comparison of major analytical strategies used in metabolomics.

	NMR	GC-MS	LC-MS	CE-MS
Applicability: Metabolite Class & Sample Type	• Non-polar and polar metabolites • Ideal for biofluids, but also suitable for tissue/cell extracts and tissue samples	• Non-polar and thermally stable metabolites • Ideal for volatile organics/head-space analysis	• Non-polar and polar metabolites • Ideal for biofluids and tissue/cell extracts	• Polar and ionic metabolites • Ideal for biofluids and tissue/cell extracts
Strengths	• High sample throughput • Minimal sample preparation • Very robust • Non-destructive • Excellent reproducibility • Qualitative identification • Absolute quantification	• High separation efficiency/peak capacity by GC×GC • Excellent reproducibility • Moderate sample volumes: <0.2 mL • Good sensitivity • Extensive electron impact (EI)-MS library for identification	• High separation efficiency with new column technology • Complementary separation mechanisms • Low sample volumes: 10–100 µL • Excellent sensitivity	• Flexibility in sample throughput (low/high) as needed • Minimal sample preparation • High separation efficiency • Very low sample injection volumes: 1–20 nL[*] • Nearly non-destructive • Minimal solvent consumption/low operating costs • Signal pattern recognition with quality assurance[†]
Weaknesses	• Large sample volumes: <0.5 mL • low sensitivity (LOD = 1 µM) • Expensive infrastructure & operation	• Low sample throughput due to long total analysis times with temperature programs • Complicated sample workup and chemical derivatization for polar metabolites	• Low sample throughput due to long total analysis times with gradient elution • Heavy consumption of solvents • Long-term system drift and column batch variability	• Poor robustness with few long-term validation studies • Poor migration-time precision • Poor concentration sensitivity

[*] Current instrumentation technology requires 20 µL of sample for nL sample withdrawal for analysis.
[†] Method developed by DiBattista et al. (2017).
References: Boizard et al. (2016); Demetrowitsch and Schwarz (2014); Emwas (2015); Gomez-Casati et al. (2013); Gowda and Djukovic (2014); Kuehnbaum et al. (2013); Scalbert et al. (2009); Wishart (2016); Zhang et al. (2012a).

elucidation (Lei et al., 2011), it continues to have lower sensitivity when compared to modern MS, despite significant and continued improvements in the sensitivity of NMR based on higher magnetic field strengths and cyroprobes for noise reduction (Emwas, 2015). On the other hand, MS is known for its superior sensitivity and high mass accuracy when using modern time-of-flight (TOF), Fourier-transform inductively coupled resonance (FT-ICR), or Orbitrap mass analyzers, the latter of which was recently shown to achieve sub-ppm detection (Weidt et al., 2016). When comparing the prevalence of MS and NMR in the metabolomics literature, it is evident that MS progressively dominates the field of metabolomics (Figure 1.2). To reduce interferences and enhance selectivity, MS is often coupled to various high efficiency separation techniques, including GC, liquid chromatography (LC), and capillary electrophoresis (CE). Furthermore, the use of separation methods provides additional information on elution/migration time aside from accurate mass (i.e., mass-to-charge ratio (m/z)), which is beneficial to metabolite identification, as recommended by the Metabolomics Standards Initiative (Dunn et al., 2013). The major advantages and disadvantages of the main instrumental platforms in metabolomics are summarized in Table 1.1. While GC and LC are the most widely used separation technologies in metabolomics (Wishart, 2016), recent advances in CE technology development are becoming increasingly recognized as important tools in metabolomic analysis (Begou et al., 2017; Kuehnbaum et al., 2013; Ramautar et al., 2013).

1.3 Principles of CE

First introduced in 1981 by Jorgenson and Lukacs, CE is a high efficiency microscale separation technique based on the differential electrophoretic mobilities of ionic solutes within a narrow (10–100 µm internal diameter; 375 µm outer diameter; 10–100 cm

Table 1.2: Summary of various operating modes in CE.

CE Operating Modes	Analysis	Separation media	Reference
Capillary Zone Electrophoresis (CZE)	Ionic and charged compounds; most commonly used	Background electrolyte: • formate, acetate, ammonium ions (MS detection); • phosphate or borate buffers (UV detection)	García et al. (2017); Bonvin et al. (2012)
Micellar Electrokinetic Chomratography (MEKC)	Ionic and neutral analytes	Electrolyte solution containing surfactant (sodium dodecyl sulfate) above critical micellar concentration	Hancu et al. (2013)
Capillary Gel Electrophoresis (CGE)	Proteins and nucleic acids	Sieving/coating matrix: • polyacrylamide gel (PAA); • polyethylene glycol (PEG); • hydroxypropyl cellulose (HPC); • polyvinyl alchohol (PVA)	Zhu et al. (2012); Nakazumi and Hara (2017)
Capillary Isoelectric Focusing (CIEF)	Protein/peptide separation based on isoelectric point	Immobilized pH gradient gel (PAA, PVA, cellulose derivative gel) with ampholyte solution	Silvertand et al. (2008)
Capillary Electrochromatography (CEC)	Ionic and neutral analytes	Immobilized stationary phase and background electrolyte solution	Dittmann and Rozing (1998)
Capillary Isotachophoresis (CITP)	Analysis of cations or anions separately	Heterogenous buffer system: leading and terminating electrolytes	Udseth et al. (1989)

length) buffer-filled fused silica capillary in the presence of a constant electric field (Vaidyanathan et al., 2005). While several separation modes of CE exist (Table 1.2), for a wide range of solutes (e.g., metals, metabolites, nucleic acids, proteins, intact cells, etc.) including CGE, MEKC, CIEF, and CITP, CZE is one of the most commonly used modes for small molecule analysis in metabolomic studies. There are two distinct electrokinetic processes that occur during CE separations, namely the electrophoretic mobility (μ_{ep}) of discrete solute ions and the electroosmotic flow (EOF) of the bulk solution. Each species possesses a characteristic μ_{ep}, or migration velocity proportional to the electric field, for a given buffer solution and temperature. The mobility of each species is determined by dielectric and hydrodynamic frictional forces, which are dependent on the effective charge and hydrodynamic radius of the species, as described by the Hubbard-Onsager model (Hubbard and Onsager, 1977). As a result, the μ_{ep} of an ion represents a characteristic physicochemical property of an ion that may be predicted for defined experimental conditions based on the ion's valence charge (as determined by its logarithmic acid dissociation constant (pK_a) and given buffer potential of hydrogen (pH)), and molecular volume (MV) (Lee et al., 2007).

Most CE applications use an uncoated/bare fused silica capillary, in which acidic silanol (–SiOH) groups are exposed along the capillary wall and readily dissociate and ionize upon contact with an aqueous background electrolyte (BGE) (Kok, 2000). Depending on the purity of the silica material and heterogeneity of the silica surface, the pK_a of the –SiOH groups have been reported in the literature to vary between 5.3 and 6.3 (Cazes, 2001; Landers, 1997). Thus, for buffer solutions of pH < 3, the –SiOH groups will be only partially dissociated, while at pH > 7, the capillary wall will be completely ionized and negatively charged (Cazes, 2001). The presence of ionized $–SiO^-$ along the inner wall forms an electrostatic layer of positive counter-ions from the solution, thereby, generating an electric double layer and a potential difference (known

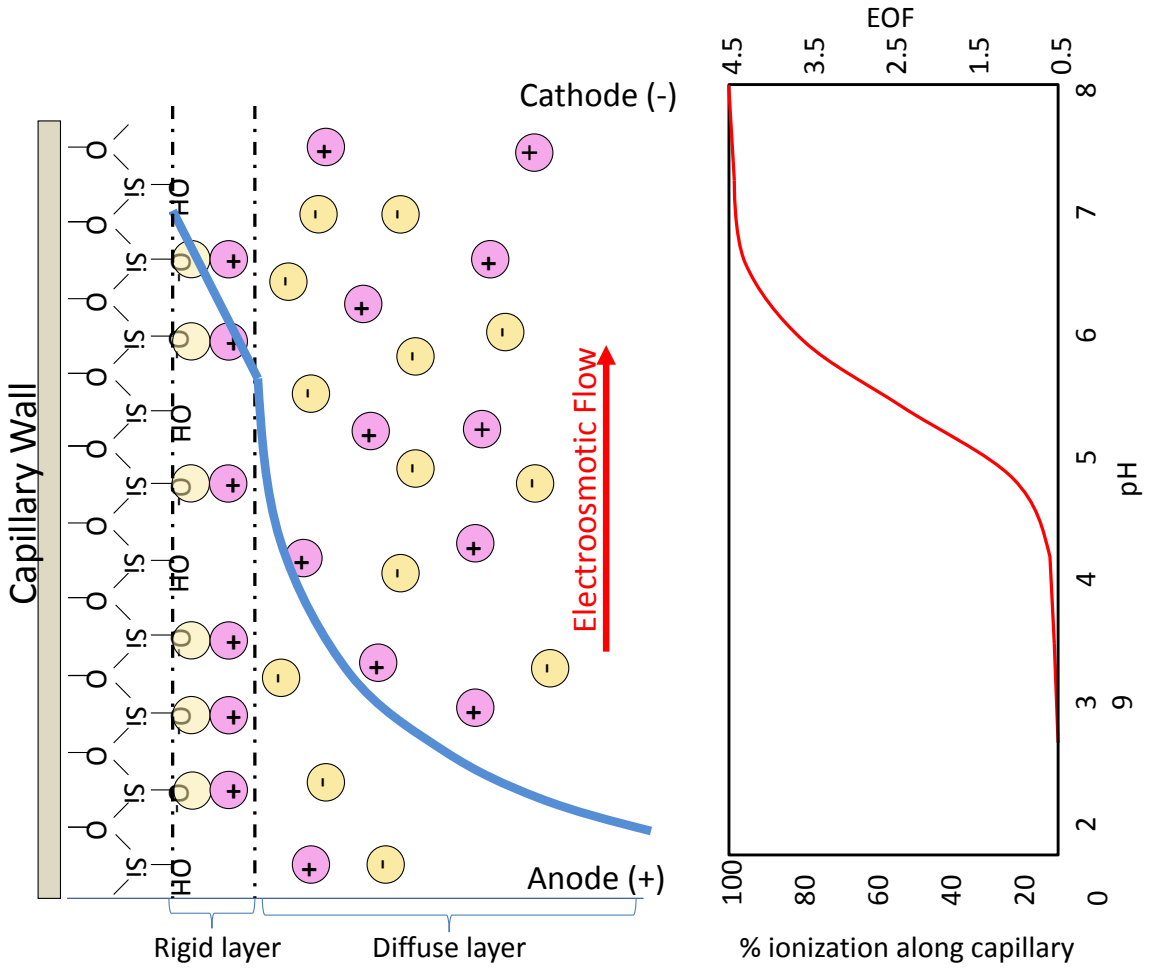


Figure 1.3: (Left) Deprotonated silanol groups along the capillary wall lead to the formation of an electric double layer and a zeta potential, which creates electroosmotic flow of the bulk solution towards the cathode upon application of an electric field. The blue line indicates the decrease in the zeta potential as distance from the capillary wall is increased (Dziubakiewicz and Buszewski, 2013). (Right) Effect of pH on EOF and percent ionization of silanol groups along capillary wall (adapted from Whatley (2001)).

as zeta potential) close to the capillary inner wall (Frazier et al., 2000). According to Stern's model (Stern, 1924), the double layer is comprised of a rigid, compact layer and a diffuse layer, which, upon application of voltage, creates the bulk flow, or EOF, of the electrolyte buffer towards the cathode by movement of the cations from the diffuse layer carrying water molecules (Landers, 1997). Therefore, as buffer pH increases, the ionization along the capillary walls increases, which leads to a greater zeta potential, thereby, enhancing the magnitude of the EOF (Figure 1.3). On the contrary, when buffer solutions of high ionic strength are used, the thickness of the double layer becomes decreased, thereby decreasing the zeta potential and in turn, suppressing the EOF (Dziubakiewicz and Buszewski, 2013). Based on Smoluchowski's equation (Smoluchowski, 1905), the velocity of the electroosmotic flow (v_{EOF}) depends on the μ_{ep} and the electric field applied (E) and is expressed in Equation (1.1) below:

$$v_{EOF} = \mu_{ep} E = -\left(\frac{\varepsilon_0 \varepsilon_r \zeta}{\eta}\right) E \tag{1.1}$$

where ε_0 is the permittivity of vacuum, ε_r is the dielectric constant of the solution, ζ is the zeta potential, and η is the viscosity of the solution. From Equation (1.1), it can be seen that the EOF is also dependent on the viscosity of the buffer solution, which is temperature-dependent, such that faster EOFs are favoured with solutions of lower viscosity (at elevated temperatures), lower ionic strength, and higher pH – which all lead to higher zeta potential. As a result, the EOF is highly dependent on both surface properties of the capillary and buffer properties of the solution, which can contribute large variations in apparent migration times in CE.

Since EOF is charge-driven, the EOF flow across a narrow capillary is uniform, thereby, resulting in a flat flow profile in CE. This is in contrast to the parabolic laminar flow seen in high-performance liquid chromatography (HPLC), which is driven by an

applied pressure from an external pump. The key advantage of the flat flow profile is the reduction of band broadening, which allows for greater separation efficiencies to be possible in CE (Frazier et al., 2000). Under most operating conditions when the flow of the EOF is greater than that of the individual ions originating from the sample injected into the buffer, all species (cations, neutrals, and anions) migrate towards the cathode, with differences in their electrophoretic mobilities allowing for their separation through time (Vaidyanathan et al., 2005).

1.4 Interfacing CE to MS

While CE was initially considered an alternative to gel electrophoresis or LC in the early 80s, it has become an increasingly used microseparation technique over time in forensic, environmental, and bioanalytical applications due to its high separation efficiency, tuneable selectivity and low operation costs for analysis of clinically relevant biomolecules, including chiral drugs, metals, metabolites, protein, and nucleic acids (García et al., 2017; Ramos-Payán et al., 2018). The major evolutionary turning point of CE was its successful coupling to MS-based detection methods in the late 1980s to early 1990s which was complementary to traditional optical detectors based on ultraviolet (UV) absorbance and laser-induced fluorescence (LIF) detection, as it offered greater selectivity for resolution and identification of analytes in complex biological samples for the first time (Johansson et al., 1991; Mück and Henion, 1989; Smith et al., 1989; Thibault et al., 1991). The coupling of CE to MS allowed for analytes within solution to first be separated in time based on their effective charge-to-size ratio in CE prior to orthogonal separation in the gas-phase based on their m/z by MS. Because CE-MS is particularly adapted for the analysis of hydrophilic/charged compounds in aqueous buffer solutions (although non-polar and neutral organic

Table 1.3: Ionization techniques for interfacing CE to MS for analysis of LMW species.

CE-MS Ionization Techniques	Mechanism	Comments	Reference
Electrospray ionization (ESI)	High voltage, corona discharge, heated nitrogen	Most commonly used; available in sheathless, sheath-liquid and liquid-junction interfaces	Hommerson et al. (2011)
Sonic spray ionization (SSI)	Nebulizer nitrogen gas flow close to sonic speed	Band-broadening and loss of resolution due to siphoning effect at capillary outlet	Hirabayashi et al. (1994)
Atmospheric pressure chemical ionization (APCI)	Gas-phase ion-molecule reactions with reagent ions (formed by collision of corona-discharged $N_2^{+\bullet}$ and $N_4^{+\bullet}$ with vapour molecules, ammonia and methanol)	Better compatibility with non-volatile background electrolytes than spray techniques	Takada et al. (1995a); Takada et al. (1995b)
Atmospheric pressure photoionization (APPI)	Photoionization by vacuum UV photons (10.0–10.6 eV), emitted by gas-discharge lamp	Analyte signal intensities may be enhanced with addition of a dopant (e.g., acetone, toluene) to the sheath liquid	Nilsson et al. (2003)

compounds can also be analyzed by CE-MS in certain aqueous/non-aqueous buffer systems (de Oliveira et al., 2014; Ding and Fritz, 1998; Vaher et al., 2001)), it provides complementary selectivity to reversed-phase LC-MS (Ramautar et al., 2013). As electrospray ionization (ESI) led to the first successful coupling of CE to MS (Olivares et al., 1987), it has since seen significant progress in interface design, sensitivity, robustness, and user-friendliness. Despite ESI continuing to be the most widely used ion source, other ionization techniques have also been developed over the years for the analysis of low-molecular-weight (LMW) species by CE-MS, which are briefly summarized in Table 1.3.

ESI is a soft ionization technique, developed by Yamashita and Fenn (1984), in which a fine aerosol of charged liquid droplets is generated from the needle tip, with the assistance of coaxial nebulizer gas flow, in the presence of a strong electrical potential of several kV, as a result of electrostatic repulsion (Konermann et al., 2013). As droplets become smaller with solvent evaporation, the charge density in the droplet

increases until electrostatic repulsion forces overcome the liquid surface tension. This results in a Coulombic burst or explosion of the droplet, in which smaller charged droplets are formed. The subsequent formation of gas-phase ions may be explained by three different mechanisms, depending on the analyte type. According to the ion evaporation model (IEM) (Iribarne and Thomson, 1976), the ejection of LMW ions (i.e., metabolites) from the droplet surface into the gas phase occurs as a result of strong Coulombic repulsions when the droplet radius becomes sufficiently small. Alternatively, the charge residue model (CRM) (Dole et al., 1968) explains that gaseous ions are formed through repeated Coulombic explosions, until a droplet contains only a single molecule of charged analyte. Large globular proteins are thought to be released into the gas phase via this route (Iavarone and Williams, 2003; Kebarle and Verkerk, 2009). Lastly, the chain ejection model (CEM) has recently been proposed by Konermann et al. (2013) to be a viable mechanism for the ejection of unfolded proteins or disordered polymers following molecular dynamics simulations.

Due to the low flow rates (nL/min) required by CE, many commercial ESI sources developed for LC systems, which have flow rates of several µL/min, were initially incompatible for coupling CE to MS (Bonvin et al., 2012). Today, ESI sources compatible with CE are available, either via modifications of existing sources or through the fabrication of new CE-dedicated sources, which permit the stable formation of droplet ions in the nL/min range using either coaxial sheath-liquid or sheathless interfaces (García et al., 2017) (Table 1.4). Although coaxial sheath-liquid interfaces generally result in loss of sensitivity when compared to sheathless interfaces due to post-capillary dilution effects by the sheath liquid, the coaxial sheath-liquid interface continues to be the most popular CE-MS interface due to its versatility and high robustness. ESI-CE-MS has also seen recent progress in miniaturization (e.g., microfluidic CE), on-line microextraction techniques, various capillary coatings for different

Table 1.4: Types of ESI interfaces developed for interfacing CE to MS.

ESI interfaces	Configuration	Reference
Coaxial sheath-flow interface	Nebulizing gas and sheath liquid (1–10 µL/min) delivered in concentric tubes orthogonal to MS inlet; most common interface *Advantages:* Robust and high spray-stability *Disadvantages:* Sheath liquid dilutes CE effluent and lowers sensitivity	Bonvin et al. (2012); Smith et al. (1988a); Smith et al. (1988b);
Liquid-junction interface	Make-up liquid present at T-junction of separation capillary and metal transfer capillary *Advantages:* No dilution effect by sheath liquid *Disadvantages:* Difficult to align CE capillary and spray tip; Improper adjustment of make-up liquid flow rate may can result in dead volume effects at junction	Cai and Henion (1995); Pleasance et al. (1992); Lee et al. (1988); Lee et al. (1989)
Sheathless interface	(a) Conductive material deposited at tapered tip of separation capillary or use of junction with conductive tip to achieve electrical contact; (b) electrode wire inserted into CE capillary outlet or through a drilled hole; (c) split flow of CE effluent through drilled hole or ion diffusion through a porous, etched capillary wall, into coaxial metal sleeve; (d) direct contact between CE effluent and metal sleeve via junction *Advantages:* No dilution or dead volume effects enhances sensitivity; maintains electrophoretic flow rate of few nL/min *Disadvantages:* Time-consuming, complex fabrication of delicate junction and miniaturized parts can be hard to reproduce; difficulty maintaining stable spray with low flow rates	Bonvin et al. (2012); Maxwell and Chen (2008)
Low sheath-flow (low-dilution) interface	(a) Separation capillary and make-up liquid capillary inserted into tapered emitter tip allowing for mixing of effluents at needle tip micro-vial region; (b) Separation capillary inserted into beveled stainless steel needle with T-union allowing for delivery of a spray-stabilizing modifier liquid *Advantages:* Low flow rate (<1 µL/min); dilution effects minimized; stable spray achieved with beveled tip; chemical modifier may also be added to stabilize flow *Disadvantages:* Requires careful optimization of composition and flow rate of chemical modifier; laminar flow occurs in micro-vial region between end of capillary and inner wall of beveled tip	Liu et al. (2005); Maxwell et al. (2010); Zhong et al. (2011)

classes of analytes, single-cell metabolomic studies, sample throughput enhancement with multisegment-injection (MSI), as well as, implementation in two-dimensional separation techniques with high peak capacity (Bonvin et al., 2012; Kohl et al., 2015; Kuehnbaum et al., 2013; Onjiko et al., 2017; Ramos-Payán et al., 2018). Recently, Soga et al. (2009) revealed significant sensitivity improvement of up to 63-fold in the analysis of anionic metabolites with the use of a platinum ESI spray needle over the conventional stainless steel needle with improved long-term robustness. As a result of these technological developments, CE-MS has been gaining recognition as a versatile and robust separation tool in clinical metabolomic studies (Ramautar, 2016; Zhang et al., 2017), as well as in food analysis and nutritional research (García et al., 2017; Ramos-Payán et al., 2018). In fact, CE-MS was recently used for the first time in a large-scale and long-term epidemiological metabolomics study, which involved the analysis of more than 10 000 plasma samples by Harada et al. (2018).

1.5 Metabolomics in Nutrition and Dietary Intervention Studies

In the past few decades, the composition of the modern diet has evolved to become 'Westernized', in which the consumption of red meat, high-fat foods, and processed sugars has greatly increased (Oriach et al., 2016). Combining this diet with sedentary lifestyles, obesity has become "the disease" or "the greatest epidemic of the 21st century" (Pêgo-Fernandes et al., 2011; Rössner, 2002), in which there are currently greater than 2 billion adults and children worldwide who are overweight or obese (The GBD 2015 Obesity Collaborators, 2017). Because nutrition impacts metabolism, oxidation and inflammation, which in turn, influence health or disease status, metabolomics has

recently been found to be a powerful implement for:

1. the determination of the complex metabolic effects of dietary components, nutrients, and foods in the human metabolome (Badimon et al., 2017; Khakimov et al., 2016; Pellis et al., 2012; Ryan et al., 2013; Scalbert et al., 2009);

2. the profiling of the food metabolome ('food-omics') and biomarkers of dietary exposure and consumption in both controlled feeding studies and large epidemiological studies involving free-living populations (Astarita and Langridge, 2013; Bhupathiraju and Hu, 2017; Claus, 2014; Guertin et al., 2014; Hanhineva, 2015; Lloyd et al., 2013; O'Gorman and Brennan, 2017; O'Gorman et al., 2013; O'Sullivan et al., 2011; Pallister et al., 2016; Ryan et al., 2013);

3. the characterization of health- and composition-altering chemical modifications in food resulting from various methods of food processing and preparation (Astarita and Langridge, 2013; Barnes et al., 2013; Beleggia et al., 2011; Dixon et al., 2006; Heuberger et al., 2010);

4. the identification of unique metabolic profiles ('metabotypes') that result in differential responses to dietary interventions (Brennan, 2017; de Roos and Brennan, 2017; Riedl et al., 2017);

5. the monitoring of treatment and metabolic outcomes with various dietary and nutritional interventions (Astarita and Langridge, 2013; Badoud et al., 2015a; Nicholson et al., 2012; Poesen et al., 2015; Vázquez-Fresno et al., 2015); and

6. the improved assessment of dietary patterns and subject compliance in nutritional interventions, as compared to self-reported food frequency questionnaires (Andersen et al., 2014; Brennan, 2017; Guasch-Ferré et al., 2018).

The application of metabolomics in nutrition research and dietary intervention studies is illustrated in Figure 1.4 and allows for better understanding of the metabolic mechanisms of diet and nutrients in health and disease (Astarita and Langridge, 2013; Badimon et al., 2017; Scalbert et al., 2009). The nutritional effects of diet on health are typically studied through observational (epidemiological) studies and interventional studies (Sébédio and Brennan, 2014). While epidemiological studies generally involve the cross-sectional study of a large population involving hundreds of individuals, dietary intervention studies involve a smaller number of participants in a controlled environment, in which food intake and nutrient content are controlled throughout the study duration (Sébédio and Brennan, 2014). Due to the high operating costs, dietary intervention studies are limited to a small number of participants, in which the accuracy of results is dependent on the compliance of the free-living participants (Sébédio and Brennan, 2014). Unlike drug intervention studies in toxicology and pharmacology, in which obvious changes are induced in blood and urine metabolites with drug administration (Pujos-Guillot et al., 2013), dietary intervention studies usually produce subtle metabolic differences and small cumulative effects, which can be difficult to measure and detect without adequate study power (Scalbert et al., 2009; Solanky et al., 2003). For this reason, pre-analytical procedures for sample processing need to be clearly defined since biological samples are unstable ex vivo, and can lead to changes and artifacts in sample composition due to oxidation or degradation (Jobard et al., 2016). Furthermore, stringent and careful data pre-processing methods are necessary to maintain high standards of quality control and avoid data extraction and integration errors, which can have substantial impact on study outcomes (Scalbert et al., 2009). Last but not least, meticulous study design is essential to take into account large inter-individual variation in metabolic profiles of biofluids that may be dependent on various factors, including age, sex, genetics, body composition, lifestyle,

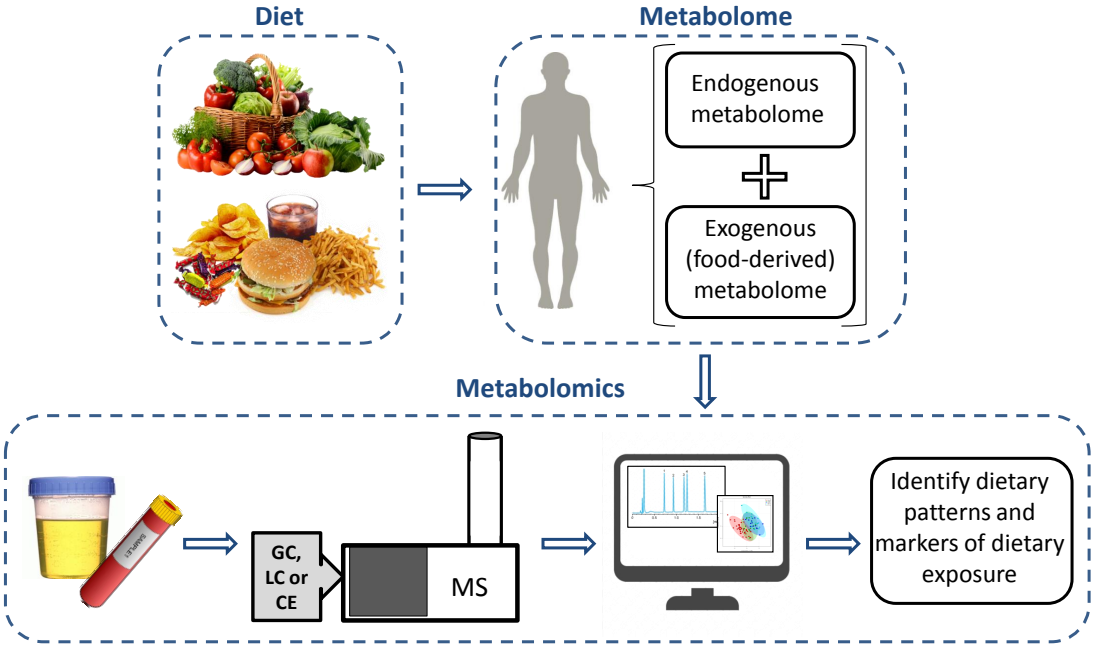


Figure 1.4: Illustration of the role of metabolomics in nutritional and dietary intervention studies. Adapted from O'Gorman and Brennan (2015).

and long-term environmental exposures (Johnson and Gonzalez, 2012). It is critical that confounding variables be minimized as much as possible, in order to obtain valid conclusions from the study that can be replicated independently among different cohorts (Smilowitz et al., 2013).

1.6 Functional and Compositional Differences Between Biofluids

In metabolomics, urine and blood (plasma or serum) represent the most commonly used biospecimens due to their collection being low-cost, relatively non-invasive, efficient, and simple to perform (Holmes et al., 2008; Kosmides et al., 2013; Nicholson and Lindon, 2008). These features are particularly advantageous for sample collection from

critically ill patients, as well as for the measurement of temporal changes, especially in time-series studies, when high-frequency sampling is needed (Kosmides et al., 2013; Nicholson and Lindon, 2008). Capable of reflecting metabolic perturbations, and physiological interactions between organs, urine and plasma exist in dynamic equilibrium with the host (Nicholson et al., 1999), in which plasma provides an instant 'snapshot' of the host's current state, while urine is a time-averaged representation (Kosmides et al., 2013; Maher et al., 2007). In contrast to cells and tissues, with each providing its own unique organ-specific metabolic fingerprint (Lin et al., 2007), blood and urine may be thought of as 'pools' of organ-produced metabolites (Yin et al., 2015), in which blood is a pool of homeostasis-regulated endogenous metabolites, while urine is highly variable and consists of terminal metabolic waste products, derived from excess or metabolized nutrients, exogenous compounds, drugs, and xenobiotics (Bouatra et al., 2013; Gibney et al., 2005; Kaddurah-Daouk et al., 2008; Yin et al., 2015). Given their distinct composition, urine is useful for reporting on exposures (such as dietary interventions, consumption of xenobiotics, and environmental challenges) and imbalances of biochemical pathways within the body (Khamis et al., 2017; Scalbert et al., 2009), while blood plasma is reflective of endogenous processes related to energy metabolism, inflammation and disease state in circulation (Scalbert et al., 2009). Because of the complementary nature of their respective metabolites, blood, and urine together could therefore reflect the state of a system as a whole at a given time point (Yin et al., 2015). Less commonly used biofluids in metabolomic studies include saliva, cerebrospinal fluid, sweat, fecal water, semen, breast milk, nipple aspirate, and tears (Bieniek et al., 2016; Delgado-Povedano et al., 2018; Di Venere et al., 2018; Huynh and Mohan, 2017; Mikkonen et al., 2016; Yen et al., 2018). In all cases, standard operating protocols need to be used for sample collection, transportation, and storage since they are critical elements to reduce biological variation and potential bias when

performing discovery-based metabolomics research.

Urine is composed mostly of water (95%), and contains organic compounds, electrolytes, metabolites, with little protein content relative to the blood-derived serum and plasma samples (Khamis et al., 2017). Besides the matrix differences between urine and blood, serum and plasma also exhibit differences from each other in their composition and metabolite profiles. Serum is the liquid component in whole blood following blood coagulation, and contains various metabolites, lipids, and enzymes that are released by activated platelets and the metabolic activity of red and white blood cells during coagulation (Yin et al., 2015). As a result, several studies have generally found increased levels of peptides, proteins, lipids, and amino acids in serum compared to plasma (Barri and Dragsted, 2013; Liu et al., 2010; Yu et al., 2011). Another contributing factor for the compositional differences between serum and plasma is the exposure of serum samples to ambient room temperatures, which is needed for proper coagulation to occur (Wung and Howell, 1980).

1.7 Pre-Analytical Challenges in the Collection, Handling, and Storage of Biofluids

1.7.1 Sample Collection and Handling

To ensure sample quality and integrity, various aspects need to be considered when dealing with different types of biofluids, since specific challenges are often encountered during the pre-analytical stages of sample collection, handling, storage, and processing (Figure 1.5). For example, time of sampling is an important aspect to consider when designing a metabolomic study since diurnal variation and circadian rhythm have been shown to impact metabolite concentrations measured in both urine and blood from

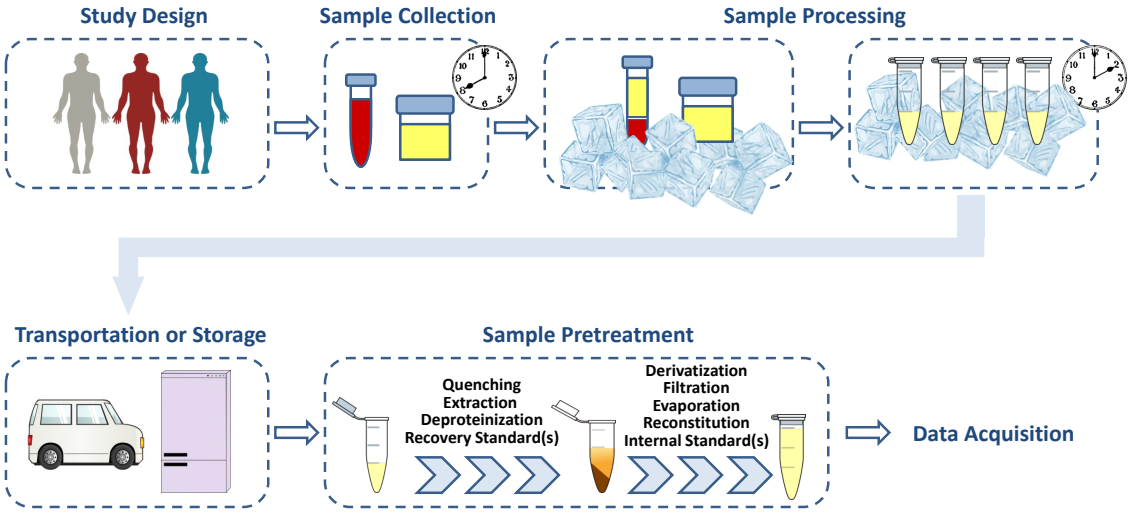


Figure 1.5: Overview of the pre-analytical workflow in metabolomic studies from sample collection to sample analysis. Adapted from Yin et al. (2015).

human subjects (Ang et al., 2012; Dallmann et al., 2012; Maher et al., 2007; Walsh et al., 2006). The work of Maher et al. (2007) reported changes between morning fasting urine samples collected within a two-hour time window, which emphasizes the importance of specifying clearly the collection time of urine samples whenever possible, in order to minimize confounding variation. Meanwhile, the work of Ang et al. (2012) revealed that 19% of the metabolites measured in blood samples exhibited significant differences associated with time of day, including acylcarnitines, cortisol, and amino acids, which cautions against the pooling together of samples collected at different times of day. Lifestyle factors (e.g., physical exercise, stress, and smoking), age, sex, and body mass index (BMI) are also known to influence absolute metabolite concentrations in biofluids (Yin et al., 2015), therefore, study designs should optimally control for such confounding variables by careful screening and matching of study participants, ensuring subject compliance to detailed instructions specified by the study, and collection of participant metadata to minimize false discoveries. Previous-day consumption of food and drugs have been reported to influence next-day morning

urinary profiles (Maher et al., 2007), therefore, diet intervention studies need to be careful to restrict the intake of food, dietary supplements, multivitamins, and drugs extrinsic to the study by the participants. Similarly, dynamic changes have been shown to occur in the blood metabolite profile for several hours after meal consumption (Brauer et al., 2011; Gillio-Meina et al., 2013). Thus, it is critical to establish the time of postprandial blood collection in dietary intervention studies. The effect of food intake may be normalized prior to sample collection by performing a one-day wash-out period with dietary standardization using pre-packaged meals, as shown by (Winnike et al., 2009).

In the case of blood collection, it is important that samples be collected by following strict guidelines defined in a standard operating procedure (SOP), in order to achieve consistent conditions between different collection sites when multiple institutions are involved in a study (Yin et al., 2015). Generally, hemolysis should be avoided by careful drawing and handling of whole blood, since hemolysis results in the release of intracellular metabolites and enzymes, thereby, altering the sample's metabolite profile (Yin et al., 2015). Treatment procedures for blood vary depending on whether serum or plasma is to be acquired. When obtaining serum, clotting time and temperature need to be standardized and kept consistent for all samples, since both factors can influence the metabolite profile (Teahan et al., 2006). Although clotting on ice has been shown to reduce unwanted metabolite alterations (Teahan et al., 2006), serum usually requires room temperatures to form clots within 30–60 min to avoid cell lyses (Timms et al., 2007; Tuck et al., 2009; Yin et al., 2015). Meanwhile, the preparation of plasma requires the collection of whole blood in tubes containing a specific anticoagulant (i.e., ethylenediaminetetraacetic acid (EDTA), heparin, citrate, or fluoride) prior to centrifugation. Since certain anticoagulants may introduce contaminant or interfering peaks in mass spectrometry, it is highly recommended that various anticoagulant

tubes are pre-tested before deciding on the type of tube to be used consistently for metabolomics studies (Yin et al., 2015). As well, since differences in manufacturer, plastic, composition, and anticoagulant additive purity may produce varying degrees of matrix effects and interferences, an adequate number of blood collection tubes available to complete a metabolomics study should be ensured to minimize batch differences across samples (Yin et al., 2015). Following anticoagulation, plasma preparation requires that whole blood be immediately placed on ice or in ice-water, until refrigerated centrifugation is performed to separate out the blood cells/platelets from plasma, in order to minimize artifactual oxidation and metabolic activity of blood cells and enzymes (Yin et al., 2015). Recent MS work by Yin et al. (2013) and Kamlage et al. (2014) have shown that plasma only exhibited minor changes after whole blood was stored for up to 4 h on ice and up to 6 h in ice water, respectively. Once clotting or centrifugation of whole blood has been completed, aliquots of serum or plasma should be stored in a timely manner since certain metabolites are unstable at room temperature (Yang et al., 2013).

1.7.2 Sample Storage

Long-term storage of biofluids at −80 °C or lower is generally preferred (Vaught, 2006), since storage of serum and plasma at temperatures of −20 and −25 °C have revealed significant changes in metabolites, including glucose, proline, methionine, and B vitamins, after an extended period of time (Hustad et al., 2012; Pinto et al., 2014). On the other hand, long-term stability of urine has been found to be maintained with storage at −25 °C, with higher temperatures resulting in metabolic profile changes caused by bacterial enzymatic activity (Maher et al., 2007). In cases where urine is to be stored at room temperature for extended periods, various preservatives are

available (such as boric acid, hydrochloric acid, and acetic acid). However, for analysis by NMR, the addition of NaN_3 has been recommended as a preservative over NaF to slow down bacterial activity, due to less interference with NMR spectra (Saude and Sykes, 2007). Therefore, attention to the temperature at collection, during processing, and during long-term storage plays an important role in limiting unwanted metabolic alterations and false discoveries in metabolomics by applying standard operating protocols without delays to sample processing and storage. Since repeated freezing and thawing of samples can contribute to undesirable changes in apparent concentrations of labile metabolites prone to oxidation or hydrolysis, it should be avoided by distributing biological samples into small aliquots prior to storage, which are then ideally thawed slowly in ice water, as opposed to at ambient temperatures (Yin et al., 2015).

1.7.3 Sample Pretreatment

Sample pretreatment is the final, yet most important step in the pre-analytical process used in metabolomics. The quenching of bacterial and enzymatic reactions, as well as non-biogenic oxidation reactions, in urine samples may be accomplished by freezing samples immediately upon collection (Khamis et al., 2017). Processing of thawed urine samples is then easily performed by centrifugation, filtering, and in certain cases, desalting, enzyme deconjugation, and/or chemical derivatization following liquid/solid-phase extraction. In the simplest case, urine may just be diluted with the appropriate solvent with addition of internal/recovery standards prior to analysis, which makes urine preparation one of the simplest procedures in the analysis of biological matrices (Khamis et al., 2017). Improved preservation of the urinary metabolic profile during storage has been shown with ultrafiltration, using MW cut-off

filters of 3 kDa, as it assists in the removal of bacteria if chemical preservatives are not used (Saude and Sykes, 2007). On the other hand, blood samples (i.e., serum, plasma, or dried blood spots) require more extensive sample workup procedures, including quenching, deproteinization, and metabolite extraction (Yin et al., 2015). Deproteinization can be achieved with ultrafiltration, solid-phase extraction (SPE), or addition of an organic solvent as chemical denaturant, such as acetonitrile, methanol, chloroform, or methyl-tert butyl ether (MTBE), which also allows for the simultaneous extraction of metabolites (Yin et al., 2015). Selection of the procedure or type of organic solvent depends on how well the technique is able to extract the desired class or polarity of metabolites to be analyzed. Chapter 2 will present and compare the results of metabolite extraction from plasma samples using methanol and acetonitrile as organic solvents for deproteinization followed by ultracentrifugation of extract prior to metabolomic studies.

1.8 Pretreatment of Data

1.8.1 Data Normalization, Batch Correction, and Quality Controls

Considering that human biofluids, such as urine, plasma, and serum, are complex mixtures consisting of hundreds to thousands of endogenous and exogenous compounds (Bouatra et al., 2013; Guo et al., 2015; Psychogios et al., 2011), the resulting datasets generated by modern analytical technology in untargeted metabolomics studies can be quite vast and overwhelming. The analysis of information-rich and complex datasets is, therefore, known to be an exhaustive, time-consuming aspect and thus, major bottleneck in biomarker discovery for metabolomics (Scalbert et al., 2009;

Tugizimana et al., 2016). In order to separate biological variability from unwanted chemical noise, which has important implications in pattern- and variance-based multivariate classification techniques (Eriksson et al., 2001), data normalization is a necessary step that is initially performed, prior to further processing of metabolomic data, to eliminate and correct for technical variation in sample collection, sample preparation, and analytical measurements that contribute to analytical bias and inter-batch variability (Trezzi et al., 2017).

Monitoring and correcting for technical variation may be performed by the repeated measurement of quality control (QC) samples at regular intervals within a data workflow. QC samples are typically obtained by pooling all samples of a given study together in equal amounts and are critical for assessing and evaluating the quality of metabolomics data sets, in addition to serving as a reference sample for batch-to-batch or sample-to-sample variations (Brunius et al., 2016). Batch correction is generally required only when systematic differences between measurements from different groups or batches of experiments are present ('batch effects'), especially in large-scale metabolomics data sets (Nygaard et al., 2016), and may be performed by various algorithms based on single value decomposition (Alter et al., 2000), distance-weighted discrimination (Benito et al., 2004), empirical Bayes method (Johnson et al., 2007), mean-centering, standardization, and ratiometric approaches (Luo et al., 2010). On the other hand, normalization is most often, if not always, required to correct for undesired differences in samples between individuals, such as hydration status, urine volume, and water content in fecal and tissue samples. For example, the analysis of single-spot random urine specimens have conventionally required normalization to urinary creatinine (Crn) levels and osmolality, however, recently, the probabilistic quotient normalization (PQN) method that involves the use of QC samples has been introduced by Dieterle et al. (2006). PQN was applied to the work presented in

Chapter 4 and will be discussed in further detail therein, along with the limitations of Crn and osmolality correction.

1.8.2 Data Transformation

Once metabolomic data has been appropriately normalized or batch-corrected as needed, transformation of data, scaling, and centering may be performed in order to reduce skewness and unequal variability (i.e., heteroscedasticity) in the data (Carmen and Hardiman, 2006; Tugizimana et al., 2016), as well as have measurements put on a comparable scale with equal importance due to the wide dynamic range of metabolite responses in MS (van den Berg et al., 2006). Various transformation and data scaling procedures are available, and, can greatly influence the resulting output of the data analysis, depending on the chosen method (Tugizimana et al., 2016; van den Berg et al., 2006). Thus, it is crucial to have a good understanding of the various data processing and statistical methods, in order to fully exploit the value of metabolomic data and achieve statistically sound and biologically relevant conclusions. Since the choice of the data pretreatment method depends on several factors, including the data set properties and the biological question to be answered (van den Berg et al., 2006), and can affect subsequent options on data analysis strategies, it is also important to acknowledge that there is no single route, or 'one-size-fits-all formula,' to follow for the analysis of untargeted metabolomic data (Tugizimana et al., 2016).

Transformation procedures may be performed to prevent misleading outcomes as a result of technical artifacts in high-throughput metabolomics data. Data transformation results in non-linear conversions, with the goal of correcting for heteroscedasticity, in which larger signals commonly exhibit larger variations (Russell et al., 2008; van den Berg et al., 2006). Through transformation, the data is made less skewed and more

symmetrically distributed to conform to normality, in order to satisfy the criteria and assumptions of many traditional parametric statistical procedures (including regression and analysis of variance (ANOVA)), since violation of normality would otherwise increase the probability of false positives or false negatives (Carmen and Hardiman, 2006). A popular method is the log transformation, in which the logarithm of each measurement replaces the original value. The log transform is useful for reducing the range of data spanning several orders of magnitude, and converts multiplicative (ratio) noise to additive (difference) noise (Russell et al., 2008; Sussulini, 2017). While the log transformation is ideal for eliminating heteroscedasticity when the standard deviation is proportional to the mean of the signal (Kvalheim et al., 1994), unfortunately, most data in reality do not follow this trend (van den Berg et al., 2006). Other limitations of the log transform are that it is undefined for negative and zero values, and can artificially increase the variance for near-background low-intensity signals (Russell et al., 2008).

A transformation that can accept positive, negative and zero values, as well as optimally stabilizing the variance for both low- and high-intensity signals is the generalized logarithm (glog) transformation (Rocke and Durbin, 2003). Although the glog transform was initially introduced for the analysis of gene-expression microarray datasets (Durbin et al., 2002; Huber et al., 2002; Munson, 2001), glog transform may also be applied to data acquired from other high-throughput methods, including metabolomic data (Di Guida et al., 2016; Parsons et al., 2007; Purohit et al., 2004), in which the dimension of the data greatly exceeds sample size. Parsons et al. (2007) found that glog transformation was not only able to stabilize the technical variance in metabolomic data, but also significantly improved between-group discrimination with greater classification accuracy relative to unscaled, autoscaled, or Pareto-scaled data. Additionally, the work of Di Guida et al. (2016) has shown that glog transformation

on metabolomic data was optimal without scaling when searching for biologically relevant metabolites independent of their measured abundance. The glog function for a measurement, x, is defined in Equation (1.2) (Durbin et al., 2002; Huber et al., 2002; Munson, 2001):

$$f(x) = \ln\left(x + \sqrt{x^2 + \frac{a^2}{b^2}}\right) \tag{1.2}$$

in which a is the standard deviation of untransformed data at low intensities and b is the standard deviation of log-transformed data at high intensities. Alternatively, $\frac{a^2}{b^2} = \lambda$ may be estimated directly as a transformation parameter from a series of technical replicates, such that the remaining variation in the dataset is derived mainly from biological sources (Parsons et al., 2007). For large values of x, the transformation converges to $\ln x$ and is approximately linear at $x = 0$ (Durbin et al., 2002).

1.8.3 Data Scaling

Scaling may be performed following transformation, and autoscaling is one of the most commonly used techniques in metabolomics, which is defined by Equation (1.3):

$$\tilde{x} = \frac{x - \bar{x}}{s} \tag{1.3}$$

in which, $\tilde{x}$, x, $\bar{x}$, and s represent: the measurement of a metabolite following scaling; the measurement of a single metabolite; the mean of measurements for a metabolite; and the standard deviation of the measurements for a metabolite, respectively. With autoscaling, all metabolites in a dataset become equally important, however, division of the mean-centred value by the standard deviation of metabolite measurements, makes autoscaling sensitive to large measurement errors (van den Berg et al., 2006), especially if QC filters are not used to reject highly variable metabolites (RSD > 40%)

from the original data set. To reduce error sensitivity, Pareto scaling may be used alternatively, which replaces the denominator in the equation with the square root of the standard deviation. As a result, original data structure is better conserved with Pareto scaling (van den Berg et al., 2006). Although details on other scaling techniques, including range scaling, vast scaling, and level scaling, may be found in the work by van den Berg et al. (2006), autoscaling and Pareto scaling are amongst the most often used scaling methods in metabolomics prior to multivariate statistical analysis and feature selection/ranking (Parsons et al., 2007).

1.9 Dealing with Massive Data Using Multivariate Statistical Techniques

1.9.1 Statistical Data Mining

The analysis and interpretation of large and complex metabolomic datasets is strongly dependent on pattern recognition techniques, multivariate statistical methods, discriminant analysis techniques, and robust computational techniques (Lindon and Nicholson, 2008), which must be carefully applied to extract meaningful biological interpretations from the data (Kosmides et al., 2013). The most common starting point in metabolomics for exploratory data analysis and outlier detection is principle component analysis (PCA), an unsupervised method in which there is no assumed prior knowledge of group membership (Alonso et al., 2015; Gowda and Djukovic, 2014; Kosmides et al., 2013). In PCA, the dimensionality of the data is reduced while still allowing for variation in the data to be explained, as much as possible, along the projected axes of maximum variance (Kosmides et al., 2013; Parsons et al., 2007). Briefly, principal components (PCs) are orthogonal, or uncorrelated vectors, in

which the first few PCs explain the largest amount of data variation, with subsequent components explaining decreasing amounts of variance (Alonso et al., 2015; Kosmides et al., 2013). Each PC is represented by a linear combination of the variables, or metabolites, in this case. Visually, a PCA map (i.e., scores plot) reveals the grouping and clustering of samples based on similarities in their metabolite profiles, thus, also enabling the identification of any outliers (Kosmides et al., 2013; Sébédio and Brennan, 2014). By using a loadings plot, in which the eigenvector component is plotted against each metabolite variable, the metabolites that contribute the greatest to each PC or differ the most between groups can be determined (Kosmides et al., 2013). PCA is also useful for the visual inspection and graphical representation of the reproducibility of QC replicates, which ideally should show tight clustering with minimal variance as compared to biological variance, since they are replicate measurements performed on the same QC sample over time by the same instrumental platform. Therefore, PCA may be used to assess data quality and technical variation (Alonso et al., 2015), including rejection of samples prone to bias, as well as assessment of batch correction algorithms needed to correct for long-term system drift in large-scale MS-based metabolomic studies (Dudzik et al., 2018).

An alternative unsupervised technique to visualize the classification of samples and metabolites is clustering, in which hierarchical cluster analysis (HCA) and K-means (KM) clustering are two methods of cluster analysis that have been used in metabolomics (Bartel et al., 2013). Briefly, clustering methods assemble or group samples based on similarities in their metabolomic data, which may lead to data-based sample reclassification, or confirmation of known groupings and associations (Čuperlović-Culf, 2013). Recently, HCA was demonstrated to be a new tool for assessing the effectiveness of data filtering prior to statistical analysis, which revealed that technical replicates clustered in all data sets only after sufficient data filtering

had been performed (Caesar et al., 2018). However, as cluster analysis is difficult to interpret, in addition to having poor reproducibility due to its sensitivity to noise and outliers, it is rarely used in the metabolomics literature when compared to its prevalence in genomics and transcriptomics applications (Čuperlović-Culf, 2013). More detail on clustering methods may be found in the review by Andreopoulos et al. (2009).

To further maximize separation between groups and reveal group-discriminating metabolites, supervised multivariate statistical methods, such as partial least squares discriminant analysis (PLS-DA) and orthogonal projection to latent structures discriminant analysis (OPLS-DA), are frequently used in metabolomics. Supervised statistical techniques allow for class membership to be predicted when a priori information on sample classification is provided for the construction of predictive models (Kosmides et al., 2013; Sébédio and Brennan, 2014). Model performance is then validated and tested for robustness and predictive accuracy, which is quantified by Q^2, using such internal cross-validation procedures as leave-one-out, k-fold, or Monte Carlo permutation tests, when an independent data set is unavailable for external cross-validation (Girard, 1989; Heather et al., 2013). Since the presence of noise in data sets can hamper the interpretation and accuracy of PLS-DA models, orthogonal signal correction (OSC) may be used to improve model accuracy (Gavaghan et al., 2002), by removal of the variation in the classification (descriptor) matrix that is uncorrelated to the variation in the metabolite (results) matrix (Bylesjö et al., 2006; Kosmides et al., 2013). The incorporation of OSC into PLS-DA results in OPLS-DA, which allows for easier interpretation of multivariate statistical models.

1.9.2 Variable Selection and Filtering

Using supervised multivariate statistical methods, metabolites that contribute strongly to group separation/classification can then be determined through the ranking of metabolites based on their measured responses, from which a subset of metabolites may then be selected for development of a predictive model. Simplification of a model to only a few variables is helpful to understanding and explaining the underlying metabolic phenomena, by eliminating the majority of 'noisy' metabolites whose variation is not related to the response of interest, and which are, therefore, irrelevant to the investigation (Farrés et al., 2015). This is especially advantageous in untargeted metabolomic studies for biomarker discovery when the number of features is much larger than the number of samples – such that only a small number of metabolites are associated with the underlying pathophysiology of a disease or serve as treatment markers of therapeutic interventions (Degenhardt et al., 2017). Generally, a parsimonious model is preferred since model interpretation becomes substantially easier with simplicity of description, while allowing for improved model performance and predictive accuracy (Farrés et al., 2015; Huberty and Olejnik, 2006; Mehmood et al., 2012). Even in cases when the performance of a model may be compromised by its interpretability, ease of interpretation may still be favoured at the expense of predictive accuracy of the model (Mehmood et al., 2011).

Variable Importance in Projection (VIP) in PLS Regression

Within the context of PLS regression, several types of methods are available for variable ranking and selection, such as filter methods, wrapper methods, and embedded methods, as described by Mehmood et al. (2012) and Andersen and Bro (2010). Filter methods identify important variables (i.e., metabolites) using the output from the

PLS regression algorithm from which variables are then selected depending on a user-defined threshold level (Mehmood et al., 2012). On the other hand, wrapper and embedded methods perform iterative procedures between model fitting and variable selection, or select variables as part of an integrated process in the modified PLS algorithm, respectively (Mehmood et al., 2012). Although filter methods do not provide any measure or indication on the predictive abilities of the selected variables, they are quick and computationally much more simple compared to the wrapper and embedded methods (Mehmood et al., 2012). Currently, the most popular approach is the variable importance to the projection (VIP) method (Farrés et al., 2015) – a filter method that was introduced by Wold et al. (1993). In the VIP approach, the importance of each metabolite (x variable) is measured with respect to its correlation or 'global contribution' to the response (y variance) in a PLS model as the VIP score, which is given by Equation (1.4):

$$\mathrm{VIP}_j = \sqrt{\frac{\sum_{f=1}^{F} w_{j,f}^2 \cdot SSY_f \cdot J}{SSY_{total} \cdot F}} \tag{1.4}$$

where j is the variable being considered, F is the total number of extracted components, $w_{j,f}$ is the weight value for the j^{th} variable and f^{th} component, SSY_f is the sum of squares of explained variance by the f^{th} component, J is the total number of variables, and SSY_{total} is the total sum of squares of explained variance from all components (Čuperlović-Culf, 2013; Farrés et al., 2015). A VIP score greater than one is generally accepted as a threshold for significance (although not statistically justified) to indicate importance of a feature to the model, and its candidacy in the variable selection process (Chong and Jun, 2005; Farrés et al., 2015). However, previous studies has shown that the VIP method can sometimes be inclusive of irrelevant variables and lead to false positive candidates (Farrés et al., 2015; Tran et al., 2014). Another

filter approach that is also commonly used for the ranking of variables is the use of regression coefficients, which is the measured association between each variable and the response (Mehmood et al., 2012). Variables possessing small absolute values of regression coefficients may be removed (Frenich et al., 1995), while the selection of more important variables may be assisted with the application of statistical re-sampling procedures (i.e., bootstrapping or jackknifing) when considering the probabilistic distribution of regression coefficients, at the expense of increased computational time (Mehmood et al., 2012).

OPLS-DA Loadings Plot: S-plot

For mean-centred or pareto-scaled data, the S-plot, which is a loadings plot that combines covariance and correlation vectors from OPLS-DA, has been shown to be a useful tool for the extraction of potentially significant metabolites by Wiklund et al. (2008), and has since been widely used due to its ease of interpretation and usage. The S-plot is so-called due to its S-shaped scatter plot, which allows for visualization of variable influence in a model, in which the x-axis visualizes the contribution (covariance) of the metabolite to the OPLS-DA score vector, while the y-axis is dependent on the variability of metabolite measurements, and is a reliability indicator of the correlation between the metabolite and score vector. Metabolites with high covariance have high influence on the model, while metabolites with high correlation have high reliability due to lower risk of spurious correlation (Tugizimana et al., 2016). Since the selection of metabolites based solely on either high covariance or high correlation indices will result in the biased selection of metabolites with either high or low concentrations, respectively, the selection of metabolites from the S-plot should ideally favour those metabolites with both high covariance and correlation (i.e., located near the upper right or lower left corners in the S-plot, as shown in Figure 1.6, which have been

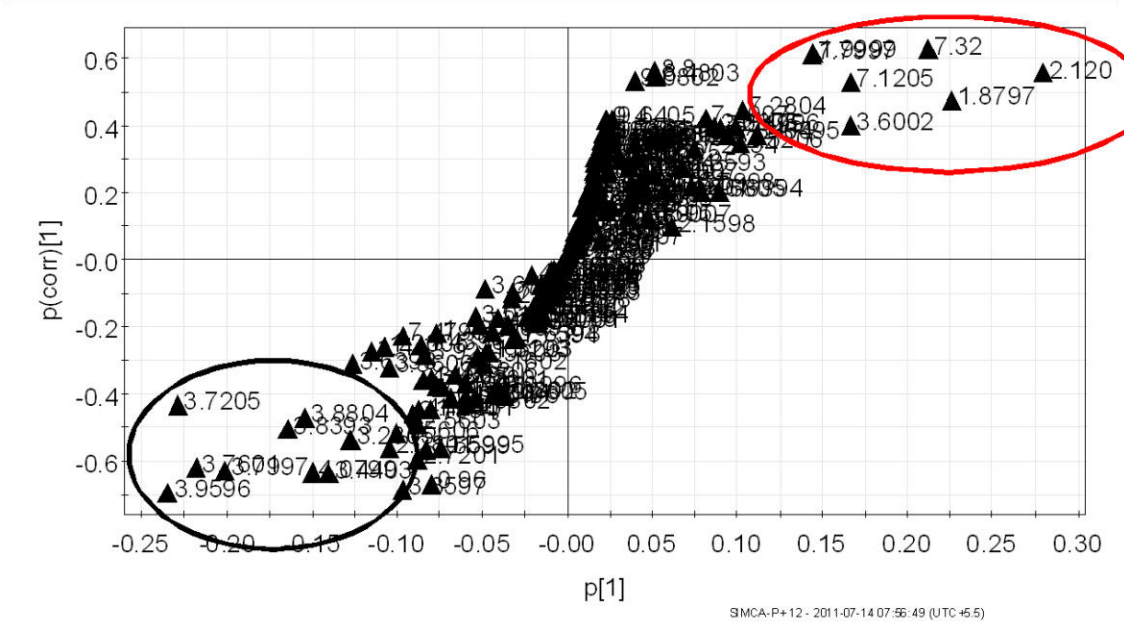


Figure 1.6: Example of an OPLS-DA loadings S-plot (reprinted from Sengupta et al. (2011); with permission from https://creativecommons.org/licenses/by/2.0/legalcode), in which circled regions indicate features selected for further analysis.

shown to be statistically significant as potential discriminatory features (Wiklund et al., 2008).

Variable Importance Measures in Random Forest (RF) Analysis

Outside of (O)PLS regression, measures of variable importance for feature ranking may also be obtained from random forest (RF) analysis, which is an alternative supervised technique for generating classification models based on decision trees (Breiman, 2001). In RF analysis, the variable importance indices allows identification of important features based on their predictive ability in classification models (Degenhardt et al., 2017), which provides a means to distinguish relevant from irrelevant variables (Hapfelmeier and Ulm, 2013). As a classification method, RF analysis is highly accurate and robust to over-fitting, since a set of variables (typically, the square root of the total number of variables) is randomly selected at each node in a decision tree,

in which a training set of data is split in a binary manner according to the small subset of input variables, until sample groups are recognized (Breiman, 2001; Gromski et al., 2015). This process is repeated with multiple training sets (one tree per training set) that are generated from bootstrapping (i.e., re-sampling with replacement) on two-thirds of the original data set, in which the number of trees is entirely up to the user and can vary from as little as a single digit number of trees to several thousands of trees (Zhang and Wang, 2009). While the current literature does not suggest an optimal number of trees within an RF (i.e., the threshold beyond which an increased number of trees does not significantly improve classification performance at the cost of greater computational time), the work of Oshiro et al. (2012) suggested that a range between 64 and 128 trees is sufficient for obtaining a good balance between classification accuracy, processing time, and computer memory usage. However, given the modern advances in processing power, algorithm optimization, and the availability of large computing clusters, this range appears to be low by today's standards and is easily exceeded, as demonstrated by the thousands of trees generated within a few seconds by the internet-based tool, MetaboAnalyst (Xia and Wishart, 2016).

Another advantage of RF analysis is that cross-validation is not needed on an independent set of data, since the classification error of the test is internally estimated in an unbiased fashion during the analysis on test sets of data that are generated by bootstrapping on the remaining one-third ('out-of-bag' data) of the original dataset that was not used to generate the initial training sets (Breiman, 2001). It is through the out-of-bag data that variable importance is estimated by assessing the difference in classification performance averaged over all trees following random permutations in the values of one variable at a time (Degenhardt et al., 2017). Therefore, those variables whose randomly permutated values result in the greatest number of misclassifications are important for prediction and will have large variable importance measures. RF

Table 1.5: Summary of non-parametric and parametric tests.

Number of Groups to be Compared	Experimental Design	Non-Parametric Test (Non-Normal Distribution)	Parametric Test (Normal Distribution)
Two	Unpaired groups Paired groups	Mann-Whitney U test Wilcoxon rank sum test	Unpaired t-test Paired t-test
Greater than two	Unmatched groups Matched groups	Kruskal-Wallis test Friedman's test	One-way ANOVA Repeated-measures ANOVA

analysis in variable selection has been shown to be successful (Fan et al., 2011; Patterson et al., 2011) and outperform that by PLS-DA, while PLS-DA has been shown to outperform RF analysis in terms of classification (Menze et al., 2009). As a result, the usage and exploration of complementary methods for variable selection and classification in metabolomic studies is recommended.

1.9.3 Comparison Between Multiple Groups

Following the simplification of a data set to a smaller number of important variables, hypothesis testing is used to determine whether observed differences between means of two or more groups are due to random chance or conditions imposed by the study. Hypothesis testing is performed at a specified probability rate of a false positive or type I error (α), typically set at 5%, to investigate two opposing hypotheses: the null hypothesis (i.e., there is no difference between group means) and the alternative hypothesis (i.e., there is a difference between group means) (Vinaixa et al., 2012). Depending on the number of experimental groups and the experimental design, various statistical tests for quantitative data are available (Table 1.5), however, selection of the appropriate multivariate statistical technique also depends heavily on the underlying distribution of the data, which will dictate use of an appropriate parametric or an equivalent non-parametric statistical test.

Non-parametric tests do not rely on assumptions about the data distribution, while

Table 1.6: Comparison of non-parametric and parametric tests.

Parameters	Non-Parametric Test	Parametric Test
Assumptions:		
1) Data distribution	Any	Normal (Gaussian)
2) Variance	Any	Homogenous
3) Independence of observations	Not required	Required
Data type	Ratio or interval	Ordinal or nominal
Statistical comparator	Median	Mean
Correlation testing	Spearman correlation	Pearson correlation
Advantages	Greater flexibility; easier to use; less sensitive to outliers	Greater statistical power (i.e., reduces type II error)

parametric tests require at least three basic assumptions to be valid, namely that: 1) data is normally distributed; 2) data exhibit homogeneity of variances; and 3) observations are independent. Table 1.6 summarizes some of the key differences between non-parametric and parametric tests. In addition to visual inspection of histograms, normality may also be evaluated by probability plots, quantile-quantile (Q-Q) plots, calculation of skewness and kurtosis, Shapiro-Wilk tests, or Kolmogorov-Smirnov tests (Carmen and Hardiman, 2006; Vinaixa et al., 2012). It should be noted, however, that for large metabolite datasets, visual inspection for each metabolite variable becomes impractical. Similarly, the evaluation of homoscedasticity, or homogeneity of within-group variances, by visual inspection of boxplots is not realistic for large metabolomics datasets. Therefore, alternative statistical methods using Levene's and Bartlett's tests are preferred, which test for group variances being equal as the null hypothesis (Vinaixa et al., 2012). In metabolomic studies, the assumption that observations are independent is fulfilled when the measurement of a sample is independent of, or not influenced by, measurements from other samples.

Given that metabolomic datasets rarely satisfy the rigid assumptions of parametric tests due to the large number of measured variables, and the much greater ease

and flexibility with which "assumption-free" non-parametric tests may be applied on both normally- and non-normally distributed data, non-parametric statistical methods are widely used in metabolomics studies, as reflected by the weak presence of parametric applications for multivariate analysis, such as multivariate analysis of variance (MANOVA), in the metabolomics literature. It is possible that the time required to carefully check assumptions may also serve as a deterrent to the usage of parametric tests. Despite these challenges, parametric tests are more robust and have greater statistical power over non-parametric methods, which provide a greater chance of correctly detecting significant differences between groups that may otherwise be missed by non-parametric tests and lead to a false negative result, or type II error (Vinaixa et al., 2012). On the other hand, non-parametric tests have better performance on non-normally distributed data with unequal variances (Vinaixa et al., 2012).

In metabolomics, a large number of metabolites as variables in a statistical model are often tested simultaneously for significance. As a result of multiple hypothesis tests, inflation of type I error occurs, in which the probability of incorrectly rejecting the null hypothesis and finding a false positive increases, due to random chance. The probability of finding at least one false positive when testing all hypotheses is called the family-wise error rate (FWER), which is given by Equation (1.5):

$$FWER = 1 - (1 - \alpha)^k \tag{1.5}$$

in which α is the pre-specified level of significance for each test, and k is the total number of tests performed (Eichstaedt et al., 2013). To correct for multiple hypothesis testing in multivariate statistical models, the probability threshold for significance may be re-calculated or adjusted using Bonferroni correction or false discovery rate (FDR)

procedures. Bonferroni correction maintains the integrity of the FWER, and instead, makes the α for each test more stringent by dividing the FWER by the number of tests (Eichstaedt et al., 2013) as shown in Equation (1.6), where a metabolite is considered significant when its p-value is $\leq \alpha$:

$$\alpha = \frac{FWER}{k} \tag{1.6}$$

In contrast, the FDR method, also known as the Benjamini-Hochberg procedure, recalculates the p-value of each metabolite to represent the probability of the metabolite being a false positive in the event that the metabolite is deemed significant (Benjamini and Hochberg, 1995). The corrected p-value (or q-value) takes into account the original p-value and its ordered rank in the distribution of all p-values that are being considered. For a FDR at 0.05, the q-value must be ≤ 0.05 for the metabolite to be significant. To avoid confusion between the false positive rate and FDR, it is important to understand that the false positive rate is the probability of significant features being found when the null hypothesis is true, while the false discovery rate represents the proportion of significant features being false positives (Storey and Tibshirani, 2003). While both methods are often used in metabolomics, the selection of the method depends on the implications of the study results. Bonferroni correction imposes more stringent control over type I errors at the expense of inflating type II errors (Eichstaedt et al., 2013), while the FDR procedure is less conservative, and provides increased power while still controlling for false discoveries (Narum, 2006). Therefore, in certain clinical studies where false positives are a concern and can potentially have costly implications, the conservative Bonferroni correction would be a more suitable method towards these applications. However, for exploratory studies in which further research is to be conducted on potential findings of interest, the FDR approach is more appropriate

since it reduces the chance of false negatives, and avoids interesting features from being overlooked.

1.10 Thesis Motivation, Objectives, and Contributions

Metabolomics in dietary intervention studies plays a critical role in understanding the metabolism and importance of various nutrient and dietary components, as well as in the elucidation of differential metabolic responses that underlie different phenotypes between individuals. Especially with the increasing global prevalence of obesity in the 21st century, which often leads to the progression of several physical, mental, and social disorders (including type 2 diabetes, cardiovascular diseases, cancer, depression, and social stigmatization), there is an urgent need to treat current widespread obesity, as well as reduce the development of obesity and its risk in future populations. Regardless of whether metabolomics is performed in a targeted or non-targeted manner, the discovery of biologically significant metabolites and mechanisms relevant to the obese condition can help expand our knowledge of the disease to guide and improve the development of interventions for the treatment of obesity.

In this thesis, CE-MS is demonstrated to be a highly efficient analytical platform suitable for the high-throughput metabolomic analysis of various complex biological matrices, namely plasma, serum, protein, and urine, in the context of dietary intervention studies involving obesity. The objectives of this thesis are to: a) develop and validate a new CE-MS strategy for the quantification and discovery of novel thiols in protein via derivatization by a maleimide (Chapter 2); b) perform targeted metabolic profiling in serum and plasma to assess metabolic differences in individuals

of different obese metabolic phenotypes following a high-caloric dietary challenge (Chapter 3); and c) identify discriminating urinary metabolites to reveal differences in energy metabolism attributed to different types of protein supplementation when used in a weight-loss dietary intervention through untargeted metabolomic analysis (Chapter 4).

The introductory chapter of this thesis has presented an overview of metabolomics and its importance in health and nutrition. The main analytical platforms used in the field of metabolomics were briefly discussed and compared, with an emphasis on the advantages of and recent developments in CE-ESI-MS. Analytical challenges that are typically encountered during the collection, analysis, and storage of different biofluid types were highlighted. Finally, the various multivariate statistical methods available to deal with the wealth of data generated by metabolomics were discussed.

The work described in Chapter 2 presents a novel thiol-labeling strategy for CE-ESI-MS that was developed to overcome the analytical obstacles associated with the quantification of labile thiols, which are generally low in abundance, exhibit low signal responses, and are highly susceptible to oxidation artifacts ex vivo. The analytical strategy involves chemical derivatization of thiols using a maleimide, *N*-*tert*-butylmaleimide (NTBM), which yields 20-fold signal enhancement relative to unlabeled reduced thiols. When applied to extracellular protein-bound thiols, the CE-MS method was found to be comparable to previously established HPLC methods in terms of accuracy, but clearly outperformed the other methods in terms of precision, sensitivity, and simplicity of the method protocol. By implementing the MSI-CE-MS strategy (Kuehnbaum and Britz-McKibbin, 2013) and an untargeted feature extraction approach, 8 unknown novel thiol compounds were detected in human plasma protein, which may be derived from diet or environmental exposures.

In Chapter 3, the targeted metabolomic analysis of serum and plasma from healthy

lean and metabolically healthy and unhealthy obese subjects following consumption of a high-fat and high-caloric meal is described for the evaluation of differences in metabolism between different metabolic phenotypes, or metabotypes. Results showed that amino acid levels were capable of discriminating between all three groups, such that the metabolically healthy obese group exhibited an amino acid profile and homeostasis that was intermediate between the healthy lean and unhealthy obese subjects, which is believed to be a result of differences in the functionality of the tricarboxylic acid (TCA) cycle.

Chapter 4 presents the untargeted analysis of urine samples from obese subjects following a two-week hypocaloric weight-loss dietary intervention with supplementation by soy protein, whey protein or carbohydrate. The goal of this work was to elucidate underlying differences in energy metabolism as a result of the supplementation component, since earlier work by Hector et al. (2015) had shown that whey protein was more effective in conserving myofibrillar protein synthesis and attenuating the unwanted loss of muscle mass. The untargeted metabolomic analysis resulted in the detection of 167 unique features in both positive- and negative-ion mode MS, including 4 metabolites which were found to be dietary markers specific to the soy supplementation. Of the four soy-specific metabolites, 4-ethylphenyl sulfate (4-EPS) has been tentatively identified for the first time in human urine after comparison of tandem mass spectrometry (MS/MS) spectra to literature spectra acquired from rats and an authentic standard. Following discriminant function analysis, six metabolites were identified as the top markers in collectively discriminating between the three supplementation groups, in which urinary 3-methylhistidine (3-MeHis) was a clear indicator of increased muscle breakdown in the soy and carbohydrate groups relative to the whey group. Serendipitously, this work has also revealed the potential of the artificial sweetener, acesulfame potassium (Ace-K), to be a potential indicator of

kidney function or glomerular filtration rate (GFR). Finally, the thesis concludes with a discussion on the future outlook of the studies described herein, as well, as recommendations for additional future studies and follow-up work.

Chapter 2

Development of a Novel CE-MS Method for the Discovery of Protein-Bound Thiols

In this chapter, the development of a method for the discovery of novel thiols covalently bound to plasma protein is described when using CE-MS. As the work was accomplished in three stages, each stage of the study will be discussed separately in detail.

Section 2.1 introduces a simple, selective, and sensitive approach using chemical derivatization by maleimide to enhance the detection of labile and weakly ionizable thiols by CE-MS. Section 2.2 applies the novel thiol-labeling strategy towards the analysis of protein-bound thiols in human plasma, following optimization of a protein preparation protocol with thiol-selective reduction. Section 2.3 presents a data workflow for high-throughput screening of protein-bound thiols using MSI-CE-MS for the discovery of novel thiols bound to plasma protein, which may have utility as circulatory biomarkers relevant to human health and disease risk assessment.

2.1 Thiol Detection Enhancement by Chemical Derivatization with Maleimide

The work described in this section presents the application of *N*-*tert*-butylmaleimide (NTBM) as an improved derivatization reagent for thiol analysis by CE-MS that overcomes several limitations imposed by the cationic maleimide, *N*-[2-(trimethylammonium)-ethyl]maleimide (NTAM), which was previously explored by D'Agostino et al. (2011) for determination of plasma thiol redox status.

2.1.1 Introduction

Biological thiols are an important class of low-molecular-weight (LMW) sulfur-containing metabolites, such as cysteine (Cys), cysteinylglycine (CysGly), homocysteine (Hcy), γ-glutamyl-cysteine (γ-GluCys), and glutathione (GSH), which have essential roles in antioxidant defense and maintaining cellular homeostasis via redox regulation of cellular responses and metabolic pathways in biological fluids and tissues (Jones and Liang, 2009). Cys represents the most abundant extra-cellular thiol in plasma with various roles associated with metabolism, detoxification, catalysis and protein synthesis (Giles et al., 2003), whereas CysGly is the second most abundant thiol in plasma, which is formed from the enzymatic breakdown of GSH (Bald et al., 2004). Hcy is formed as an intermediate during the conversion of methionine (Met) to Cys (Brosnan and Brosnan, 2006), where elevations in total plasma concentrations represent a risk factor for stroke and cardiovascular disease (Hultdin et al., 2011; Rueda-Clausen et al., 2012; Shi et al., 2015). γ-GluCys is the biogenic precursor to GSH, which was recently found in mitochondria to respond to oxidative stress independently of intracellular GSH concentration by serving as an enzymatic co-factor

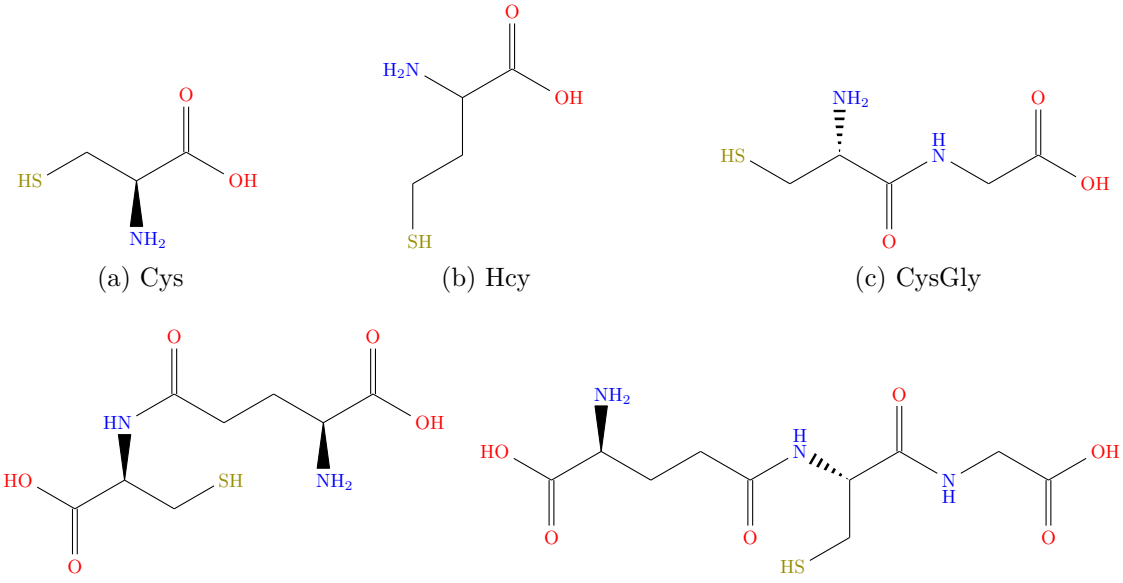


Figure 2.1: Structures of reduced (a) cysteine (Cys), (b) homocysteine (Hcy), (c) cysteinylglycine (CysGly), (d) γ-glutamyl-cysteine (γ-GluCys), and (e) glutathione (GSH).

for glutathione peroxidase-1 (Quintana-Cabrera et al., 2012). Finally, GSH, a tripeptide consisting of Cys, glutamic acid (Glu), and glycine (Gly), is the most abundant intracellular thiol present at millimolar concentrations that serves as the primary antioxidant and redox buffer to protect cells from oxidative injury (Carlucci and Tabucchi, 2009). The structures of these five thiols are shown in Figure 2.1.

Perturbations in the ratio of reduced thiols to their corresponding oxidized disulfides have long been associated with aging and disease development (Go and Jones, 2017) as a result of changes induced on protein conformation, enzyme activity, and protein interactions with deoxyribonucleic acid (DNA) (Isokawa et al., 2014). Due to their biological significance, reliable methods for accurate measurement of thiol redox status are critical to progress in health status and disease risk assessment (Toyo'oka, 2009). However, as LMW reduced thiols altogether only constitute 12–20 µM in plasma (Turell et al., 2013), and are chemically unstable due to their highly reactive sulfhydryl (–SH) group, which is prone to artifactual oxidation and thiol-disulfide exchange

reactions during sample preparation (Huang et al., 2011c), accurate determination of plasma thiol redox status is challenging. To avoid oxidation artifacts, –SH groups need to be stabilized rapidly, quantitatively, and irreversibly during early stages of sample handling (Zhu et al., 2008). In this context, thiol-specific alkylating reagents typically rely on nucleophilic substitution (e.g., monobromobimane, iodoacetamide, and halogenobenzofurazans) (Winther and Thorpe, 2014), thiol-exchange reaction (e.g., 5,5-dithio-bis2-nitrobenzoic acid (DTNB), also known as Ellman's reagent) (Aitken and Learmonth, 2009) or Michael-addition reaction via conjugation at C=C bonds (e.g., maleimide derivatives) (Higashi et al., 2003; Kand'ár et al., 2007). The development of alkylating probes with fluorogenic moieties allows for fluorescence (FL) detection by LC or CE, and is the most commonly used method of thiol analysis due to its high sensitivity (Lačná et al., 2017). While thiol detection limits with LC-FL generally range from 0.1–0.5 µM (Isokawa et al., 2014), recent LC methods have incorporated the use of novel synthesized complex derivatization reagents (Guo et al., 2013), and performed thiol-enrichment via SPE (Huang et al., 2011b) and SPE modified with multi-walled carbon nanotubes (MWCNTs) (Huang et al., 2011a) to further enhance the detection sensitivity for known thiols with low nanomolar detection limits, such as free plasma Hcy. As well, nanomolar detection limits are typically achieved by CE coupled with laser-induced fluorescence detection, while some cases have even reported picomolar detection limits (Lačná et al., 2017).

With advances in MS providing greater chemical specificity and sensitivity, MS-based methods with ESI have also become more attractive for the analysis of biological thiols. In this case, detection sensitivity is further enhanced with the incorporation of derivatization reagents that, not only stabilize the thiol, but also improves analyte ionization and increases analyte molecular weight for improved detection outside the low m/z region where matrix interferences are prevalent (Deng et al., 2012).

Improved analyte ionization efficiency with derivatization has been achieved by the introduction of a permanently charged/ionizable functional group and/or a moiety with non-polar character, which enhances solute desorption efficiency during spray formation (D'Agostino et al., 2011; Iwasaki et al., 2011; Xu et al., 2011). While MS methods typically implement thiol derivatization using conventional and commercially available reagents (Liem-Nguyen et al., 2015; Solecka et al., 2016; Sun et al., 2016), a few studies have synthesized novel in-house reagents (Gori et al., 2014; Huang et al., 2011c; Liu et al., 2014; Seiwert et al., 2008; Wagner et al., 2015), such as *N*-(2-ferrocene-ethyl)maleimide (FEM), ω-bromoacetonylquinolinium bromide (BQB), and N-benzoyloxysuccinimide to obtain further improved thiol sensitivity. Recently, BQB and N-benzoyloxysuccinimide reagents have also been used in stable isotope labeling strategies for profiling of thiols in biological samples (Liu et al., 2016; Wagner et al., 2015).

Despite the low nanomolar thiol detection limits that can be obtained with FL- or MS-based methods, some disadvantages of these techniques include long derivatization reaction times, or the use of complex derivatization reagents that are not commercially available and therefore, must be synthesized in-house. In a recent study, the plasma thiol redox status was determined using CE-MS with pre-column derivatization by a readily available cationic maleimide, *N*-[2-(trimethylammonium)-ethyl]maleimide (NTAM). Although the method achieved nanomolar detection limits when used in conjunction with on-line sample preconcentration, it suffered from complicated sample workup due to poor long-term reagent stability in aqueous solutions and an additional quenching reaction step needed to eliminate background interference from excess reagent (D'Agostino et al., 2011). In this work, an alternative maleimide for thiol labeling was developed to address the major limitations imposed by NTAM. Using *N-tert*-butylmaleimide (NTBM) as the alternative derivatization reagent, this section

presents an improved method for the selective analysis of labile and low abundance reduced thiols with lower detection limits, by using a simple, yet cost-effective labeling approach with CE-MS.

2.1.2 Materials and Methods

Chemicals and Reagents

Deionized water (DI H_2O) generated with a Thermo Scientific Barnstead™ EASY-pure™ II LF ultrapure water system (Cole Parmer, Vernon Hills, IL, USA) and HPLC-grade acetonitrile (Honeywell, Muskegon, MI, USA) were used in the preparation of solutions and working mixtures. Ultra LC-MS-grade water and methanol (MeOH) (Caledon, Georgetown, ON, Canada) were used in the preparation of background electrolyte (BGE) and sheath liquid, respectively. L-cysteine (Cys), DL-homocysteine (Hcy), reduced glutathione (GSH), cysteinylglycine (CysGly), γ-L-glutamyl-L-cysteine (γ-GluCys), 2-mercaptoethane-sulfonate (MESNA), *N*-*tert*-butylmaleimide (NTBM), 3-chloro-L-tyrosine (3-Cl-Tyr), formic acid, ascorbic acid, ammonium acetate (NH_4Ac), sodium hydroxide (NaOH), and glacial acetic acid were obtained from Sigma-Aldrich (St. Louis, MO, USA). *N*-[2-(trimethylammonium)-ethyl]maleimide chloride (NTAM) was purchased from Toronto Research Chemicals Inc. (North York, ON, Canada).

Preparation of Standards and Stock Solutions

Stock solutions and buffers were degassed and stored at 4 °C. Stock solutions (10 mM) of Cys, Hcy, GSH, CysGly, and γ-GluCys were prepared in 0.05% formic acid with 1 mM ascorbic acid as 1 mL solutions; 1 mL each of 2 M acetic acid and 10 mM of 3-Cl-Tyr internal standard (IS) solutions, and 50 mL of 1 M NaOH solution, were prepared in DI H_2O. Formic acid solution (1 M, pH 1.8) was prepared with 15% v/v

acetonitrile (ACN) in a 50 mL volume as BGE and 50 mL of 400 mM NH_4Ac solution was adjusted to pH 5 with 2 M acetic acid solution. NTBM (50 mM) was prepared in ACN as a 1 mL solution. MESNA (50 mM) and NTAM (50 mM) were prepared as 1 mL solutions in 0.1% formic acid. NTAM and thiol stocks were used within a week. MESNA was prepared fresh before each use.

Derivatization of Thiols Standards with Maleimide

Chemical derivatization of thiol standards was performed separately by the addition of 200 µM Hcy, Cys, GSH, CysGly, or γ-GluCys into 20 mM maleimide solution (NTAM or NTBM) using freshly prepared standards stored at 4 °C in 200 mM NH_4Ac, pH 5. Derivatization with NTAM required an additional quenching step to eliminate background interference from excess unreacted cationic maleimide by adding MESNA as a strongly acidic thiol in at least a 1.25-fold excess over the NTAM concentration, which forms a zwitterionic neutral adduct. NTBM derivatization did not require quenching since the neutral maleimide co-migrated with the EOF by CE-MS. For investigation of the kinetics between reduced thiols and NTBM, thiol-NTBM adducts were analyzed by CE-MS approximately 2, 30, 60, 90, and 120 min following the reaction and, in a separate stability study, NTBM adducts were measured 1 h, 24 h, 48 h, and 7 d following the reaction.

Calibration solutions for thiol-NTBM adducts were prepared in triplicate with 100-fold excess of NTBM at the following concentrations: 0.5, 1, 5, 10, 50, 100, and 200 µM for Cys and 0.1, 0.5, 1, 2, 5, 10, and 20 µM for Hcy, GSH, CysGly, and γ-GluCys. The integrated peak areas of NTBM-adducts were normalized to that of the IS, 25 µM 3-Cl-Tyr, and were subsequently plotted against concentration, in which the curves were fitted by least squares linear regression analysis.

Capillary Electrophoresis Mass Spectrometry (CE-MS)

All CE separations were performed on an Agilent CE system equipped with an XCT three-dimensional (3D) ion-trap (IT) mass spectrometer using an applied voltage of 30 kV at 25 °C, an Agilent 1100 series isocratic pump and a G16107 CE-MS sprayer kit (Agilent Technologies Inc., Mississauga, ON, Canada). Uncoated 85 cm fused silica capillaries (Polymicro Technologies, AZ, USA) with 50 µm internal diameter were used. Capillaries were preconditioned by rinsing with MeOH, 1 M NaOH, and DI H_2O for 10 minutes each (∼5 capillary volumes), respectively, prior to rinsing with BGE for 15 minutes (∼10 capillary volumes) at 950 mbar. Between runs, capillaries were rinsed with BGE for 5 minutes (∼3 capillary volumes) at 950 mbar. Single sample injections were performed by hydrodynamic injection of the sample for 5 s at 50 mbar, followed by 60 s injection of BGE at 50 mbar. On the other hand, CE-MS with on-line sample preconcentration via transient isotachophoresis was performed by injecting the sample prepared in a matrix of 200 mM NH_4Ac for 90–180 s at 50 mbar followed by 60 s injection of BGE at 50 mbar as previously demonstrated by D'Agostino et al. (2011). Sample run-time was set to 30 minutes.

The 3D-IT mass spectrometer was operated in positive-ion mode over a mass scan range of 100–500 m/z with maximum acquisition time of 200 ms with 10 average scans using a smart target of 250 000 ions. Nitrogen (N_2) was used as the nebulizing gas at 10 psi and drying gas at 4 L/min using a temperature of 300 °C in the ion source, where the electrospray ionization cone voltage was 4 kV. The coaxial sheath liquid flow rate was 10 µL/min, using 60% v/v MeOH with 0.1% formic acid via a 100:1 splitter. To minimize trapping efficiency bias for ions with different m/z, the target mass was adjusted to the average m/z of the thiol-maleimide adducts of interest in the sample. This allowed for a direct comparison of the ionization efficiency between NTAM- and

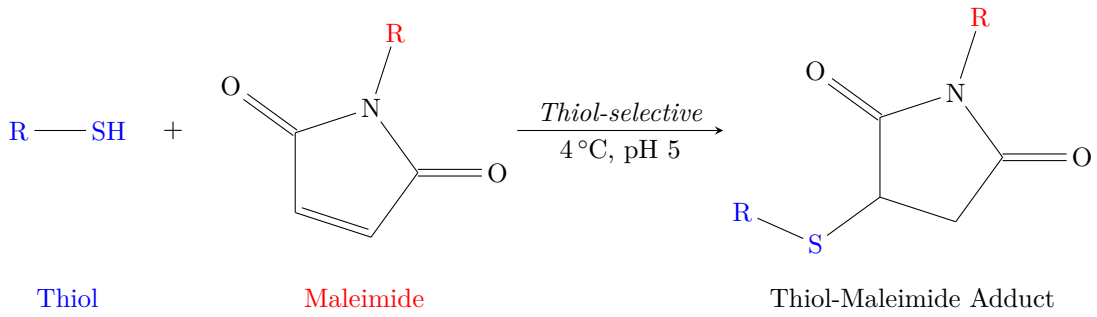


Figure 2.2: Chemical derivatization of free reduced thiols with *N*-substituted maleimides.

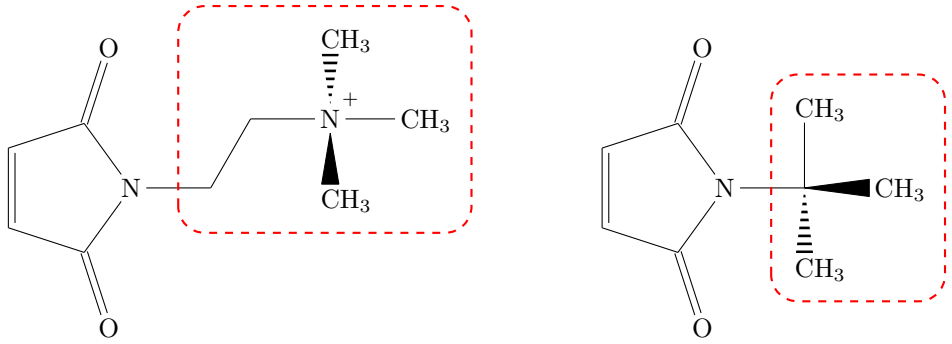
NTBM-labeled thiols. Target m/z values for NTAM and NTBM mixtures were set at 289 and 313, respectively.

2.1.3 Results & Discussion

Thiol Labeling by Maleimide and Signal Enhancement with NTBM

Maleimides react selectively towards sulfhydryls under weakly acidic pH conditions based on the nucleophilic addition of the sulfhydryl moiety to form a stable thioether adduct (Yin et al., 2009) (Figure 2.2). Weakly acidic conditions and low temperatures are favoured for thiol derivatization since auto-oxidation of thiols occurs at room temperature and may occur at pH values greater than 7.0 (Camera and Picardo, 2002; Hansen and Winther, 2009), considering that thiol pK_a values range from 8–10 (Tajc et al., 2004). In addition to stabilizing reduced thiols, formation of the thioether adduct allows for greater ionization efficiency and thus, an enhancement in the detection of small and polar compounds, such as thiols, by ESI-MS.

Recently, a cationic maleimide, NTAM, was reported to be optimal for the nanomolar analysis of thiols in plasma (D'Agostino et al., 2011) as it provided improved resolution and an overall 15-fold enhancement in peak height relative to those of the native reduced thiols (Cys, Hcy, CysGly, γ-GluCys, and GSH). Possessing a



(a) NTAM (b) NTBM

Figure 2.3: Structures of (a) *N*-[2-(trimethylammonium)-ethyl]maleimide and (b) *N*-*tert*-butylmaleimide.

permanently charged quaternary ammonium moiety independent of the pH of the CE background electrolyte (refer to Figure 2.3), NTAM conferred greater positive electrophoretic mobility to its thiol adducts relative to the unlabeled thiols, thereby, leading to shorter analysis times. On its own, NTAM, also migrated with a high positive electrophoretic mobility in CE, such that when present in large excess, unquenched label is detected as an intense broad peak with a peak width of at least 0.6 min – a large time window in which signal suppression occurred for those analytes migrating within the same time frame. To eliminate background interferences, a thiol possessing a negatively-charged sulfonate group, MESNA, was added in slight excess of the unreacted cationic NTAM in order to convert the NTAM to an electrically neutral adduct. As a result, the neutral NTAM-MESNA adduct co-migrated with the EOF, and away from the migration window of the positively-charged analytes. When applied to plasma analysis, however, an additional final quenching step after the addition of MESNA was required to prevent excess MESNA from cross-reacting with endogenous oxidized disulfides in the sample. The quench was accomplished by the addition of a neutral maleimide, *N*-methylmaleimide.

Given the complicated sample workup when using NTAM, NTBM was considered

and investigated as an alternative to simplify the thiol-labeling procedure for several reasons:

1. the lack of a positively charged moiety, eliminates the need for quenching, since the neutral NTBM would co-migrate with the EOF;
2. NTBM possesses a bulky and hydrophobic tertiary butyl moiety (Figure 2.3), which is expected to enhance ion surface activity, thus, improving ionization efficiency in ESI;
3. NTBM is cost-effective;
4. NTBM is soluble in acetonitrile (ACN), which allows for the reagent solution to remain stable over time and not be subject to hydrolysis.

Signal Enhancement with NTBM

Comparison of equimolar 10 µM NTAM- and NTBM-labeled adducts of Hcy, Cys, and GSH, as shown in the extracted ion electropherogram (EIE) of Figure 2.4, reveal a greater overall ion response (i.e., relative peak area (RPA)) in NTBM-adducts relative to NTAM by an overall factor of 2.5, specifically, with a 3.2-fold enhancement in Hcy, 2.8-fold enhancement in GSH, and a modest enhancement in Cys of 1.4-fold (Table 2.1). As expected, thiol-NTBM adducts displayed longer migration times due to the combination of a bulky/neutral moiety experiencing greater hydrodynamic friction during electromigration in free solution by CE, and the lack of a permanent positive charge, which significantly enhanced the electrophoretic mobility for thiol-NTAM adducts as reflected by their shorter migration times.

When relative ion responses of the NTBM-adducts were compared to those of the 5 unlabeled free reduced thiols (Figure 2.5), an overall average 23-fold enhancement

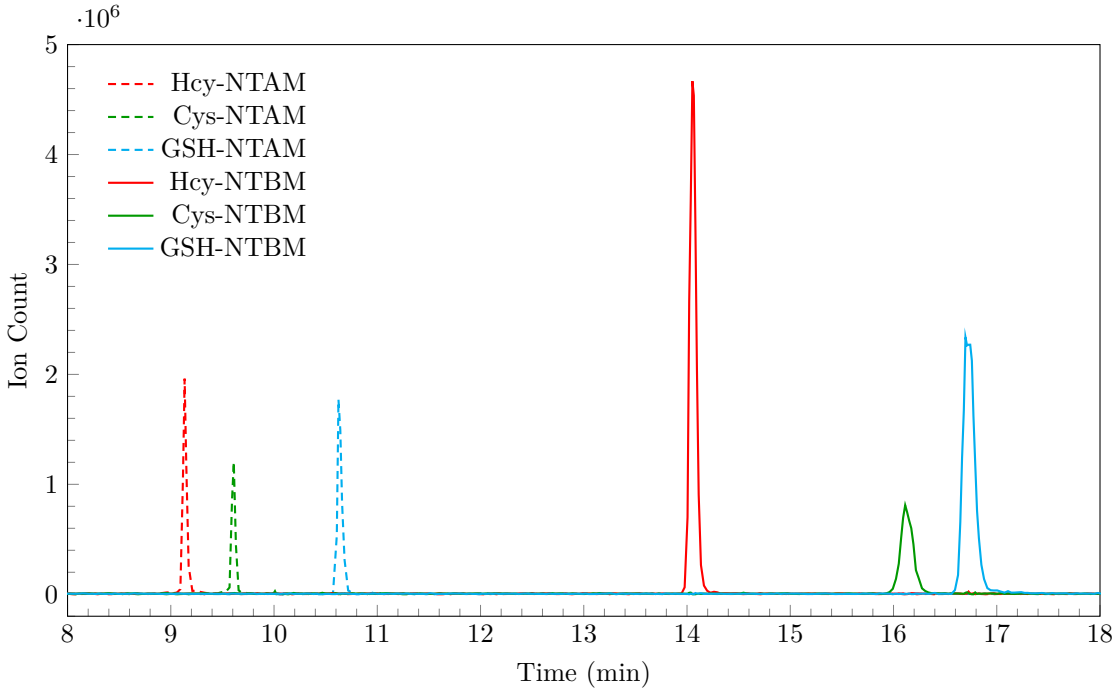


Figure 2.4: An EIE acquired with IT-MS of equimolar thiol-maleimide thioether adducts reveals greater ionization efficiency with derivatization by NTBM relative to NTAM.

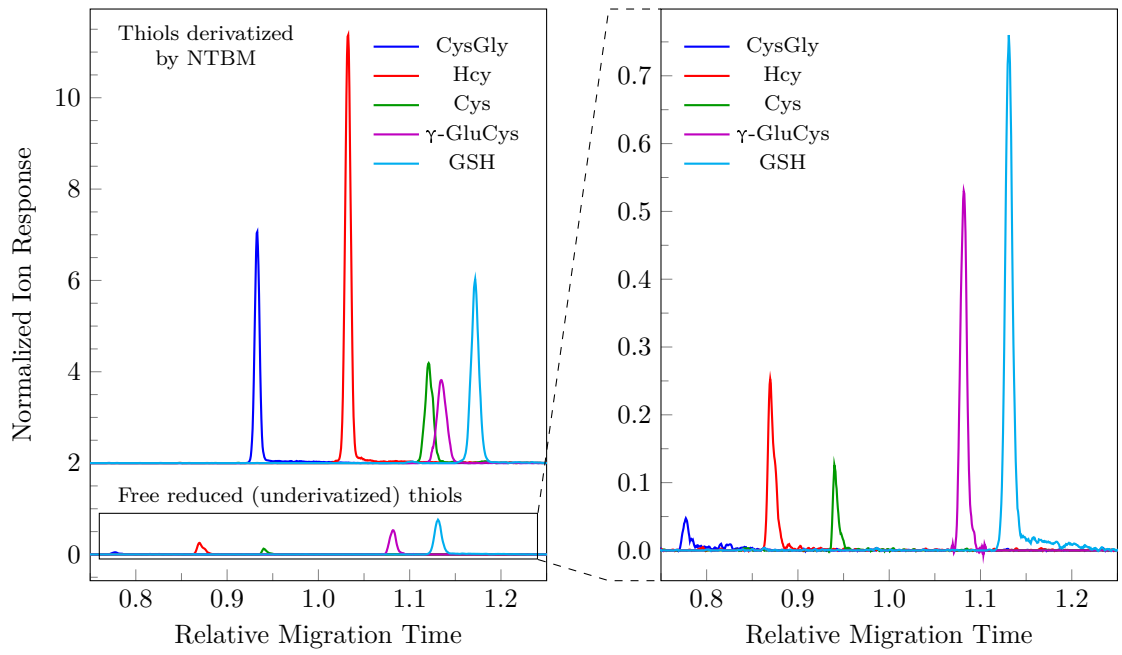


Figure 2.5: EIE comparison of equimolar unlabeled free reduced thiols and NTBM-labeled thiols, resulting in signal enhancement ranging from 4- to 50-fold.

Table 2.1: Enhancement in relative ion response (normalized to internal standard) and intensity of NTBM-adducts relative to NTAM-adducts.

Thiol	Relative Ion Response	Relative Intensity
Hcy	3.2 ± 0.7	2.3 ± 0.7
Cys	1.4 ± 0.2	0.7 ± 0.2
GSH	2.8 ± 0.4	1.4 ± 0.3
Average	2.5	1.4

Table 2.2: Summary of relative ion response and MV of free reduced thiols and thiol-NTBM adducts.

Thiol	Relative Ion Response			Molecular Volume ($Å^3$)*		
	Free thiol	Thiol-NTBM	Fold-change	Free thiol	Thiol-NTBM	Fold-change
Cys	0.12 ± 0.02	3.3 ± 0.5	28 ± 6	102.22	242.41	2.37
Hcy	0.33 ± 0.06	9.2 ± 1.6	28 ± 7	119.02	259.21	2.18
CysGly	0.07 ± 0.01	3.4 ± 0.5	50 ± 10	150.40	290.60	1.93
γ-GluCys	0.70 ± 0.08	2.7 ± 0.6	3.9 ± 1.0	211.03	351.23	1.66
GSH	1.1 ± 0.2	6.0 ± 1.4	5.5 ± 1.5	259.22	399.41	1.54

* Values for molecular volume were calculated in silico using Molinspiration (Jarrahpour et al., 2012, www.molinspiration.com), which were obtained from fitting the sum of fragment contributions to "real" 3D volume from a training set of 12 000 mostly drug-like molecules. Semi-empirical Austin Model 1 (AM1) method was used for optimization of 3D molecular geometries.

in concentration sensitivity was achieved. Notably, CysGly exhibited the greatest enhancement when derivatized with NTBM, showing approximately 50-fold enhancement in the normalized ion response relative to its native thiol (Table 2.2). The response enhancement in other thiols ranged from 4- to 28-fold with γ-GluCys and GSH exhibiting the lowest fold changes of 4 and 5.5, respectively.

The trend in ion response enhancement for the various adducts may be explained by the respective pK_a values of the –SH group in the free reduced thiols, since their acidity determines the nucleophilicity of the thiolate anion to participate in the Michael addition with the maleimide reagent at a given pH, according to the Henderson-Hasselbalch equation, which may be rewritten as Equation (2.1) (Iciek et al., 2004):

$$\frac{[RS]^-}{[RSH]} = 10^{pH-pK_a} \tag{2.1}$$

Table 2.3: Summary of pK_a, m/z, and RMT of free reduced thiols and thiol-NTBM adducts.

Thiol	pK_a (–SH)*	m/z		RMT	
		Free thiol	Thiol-NTBM	Free thiol	Thiol-NTBM
Cys	8.15	122.027	275.106	0.936 ± 0.005	1.121 ± 0.010
Hcy	8.87	136.043	289.122	0.870 ± 0.002	1.033 ± 0.002
CysGly	6.36	179.049	332.128	0.777 ± 0.005	0.933 ± 0.008
γ-GluCys	9.70	251.070	404.149	1.080 ± 0.003	1.135 ± 0.006
GSH	9.20	308.091	461.170	1.131 ± 0.004	1.172 ± 0.007

* pK_a values were obtained from Chu et al. (2017); Iciek et al. (2004); Stark et al. (1989).

In this case, when pre-column thiol derivatization conditions were performed with NTBM at pH 5 (in order to stabilize free thiols thus preventing their oxidation), a strong negative inverse correlation of -0.94 was observed (Figure 2.6b) when correlating measured average ion response enhancement with pK_a of the –SH group, which are listed for the various thiols in Tables 2.2 and 2.3, respectively. The inverse correlation is in agreement with the expectation that thiolate anions with higher pK_a values will have lower reactivity due to the thiol group being less likely to be deprotonated and therefore, less readily available for reaction with maleimide. In addition to the pK_a, the ion response enhancement of the thiols was also found to be dependent on the fold-change increase of the MV following derivatization, in which a moderate positive correlation of 0.55 was observed (Figure 2.6a). The positive correlation between MV fold-change and ion response enhancement is in agreement with earlier work by D'Agostino et al. (2011) and is consistent with ESI-MS models (Chalcraft et al., 2009; Oss et al., 2010), which rationalize that the increase in ion response is a result of increased ionization efficiency owing to larger MV. By applying multiple linear regression, the strength of the effects of pK_a and MV fold-change on ion response enhancement of the thiols were determined. A nearly significant two-factor regression model was found ($F(2, 2) = 14.965$, $p = 0.06$) with an R^2 of 0.937, in which the ion response enhancement of thiols derivatized by NTBM could be predicted by

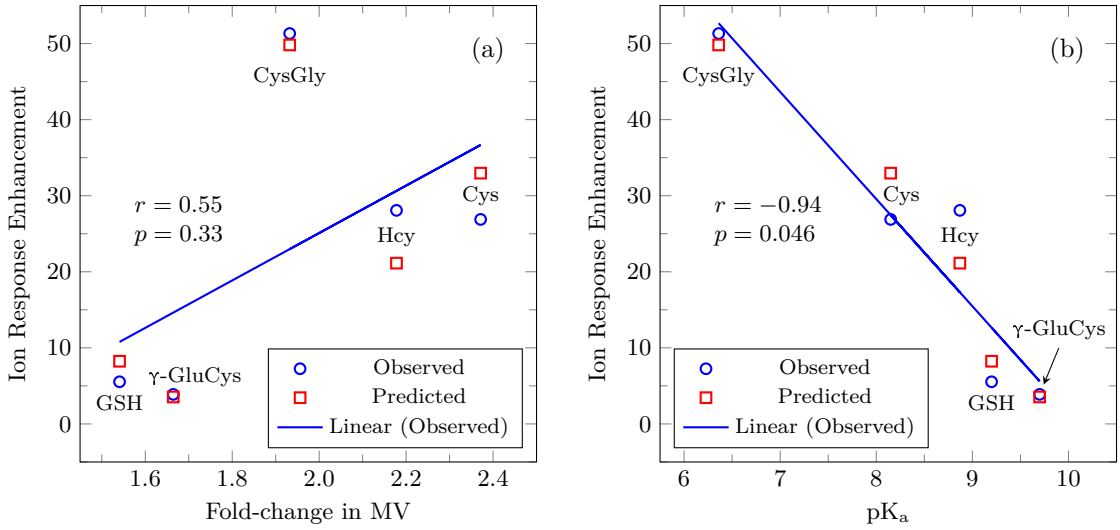


Figure 2.6: Correlation plots between the ion response enhancement (blue – observed; red – predicted from multiple linear regression modeling) and (a) fold-change in MV or (b) pK_a of sulfhydryl moiety, following maleimide derivatization for five biological thiols with NTBM. Pearson correlation coefficients (r) and p-values are indicated with respect to the observed ion response enhancement.

Equation (2.2):

$$\text{Ion response enhancement} = 104.645 + 13.631\ (\text{MV fold-change}) - 12.762\ (\text{thiol } pK_a) \tag{2.2}$$

Thiol pK_a was a significant predictor ($p = 0.046$) while MV fold-change was found to be non-significant ($p = 0.33$) in the prediction of ion response enhancement. The correlation plots between fold-change in MV, or pK_a of the thiol moiety, on the measured and predicted ion response enhancement are shown in Figure 2.6. Molecular volume data for reduced thiols and the corresponding NTBM-adducts, are summarized in Table 2.2, in which Cys-NTBM and Hcy-NTBM show the greatest increase fold-change in MV relative to the reduced thiols.

Reaction Time of NTBM

The reaction time of the derivatization reaction was investigated with NTBM being present at a 100-fold excess relative to total concentration of reduced thiols. The reaction was performed on ice (4 °C) to simulate conditions that would be required for the derivatization of plasma samples in order to minimize auto-oxidation artifacts during sample processing. Plots of concentration of Cys-, Hcy-, and GSH-NTBM adducts versus time (t = 1–200 min) in Figure 2.7 indicate that derivatization with 100-fold excess of NTBM is complete for the 3 thiols after approximately 1 h on ice. Given that the processing time for protein and plasma samples following derivatization takes at least 1.5 h, it can be assumed that the labeling of protein-bound or plasma thiols by NTBM will have reached completion by the time the sample is ready for analysis.

Stability of NTBM Reagent and Thiol Adducts

The short-term stabilities of Cys-, CysGly-, γ-GluCys-, Hcy-, and GSH-NTBM adducts were monitored at 1 h, 24 h, 48 h, and 7 d following derivatization as a mixture. In Table 2.4, the percent coefficient of variation (%CV) of the ion responses for the 5 thiol-NTBM adducts were all found to be within 5% when measured over a 1-week period. The %CV is sufficiently low such that the values are within the normal variation of instrumental performance. Thus, the results confirm that the thiol-NTBM adducts are stable for at least one week following their formation when stored at 4 °C and therefore, are expected to be stable for a much longer time period especially when stored under conditions that are typical for biological samples (i.e., −80 °C).

Since the work of D'Agostino et al. (2011) had previously found that certain maleimide reagents, such as NTAM, were susceptible to hydrolysis, thereby requiring

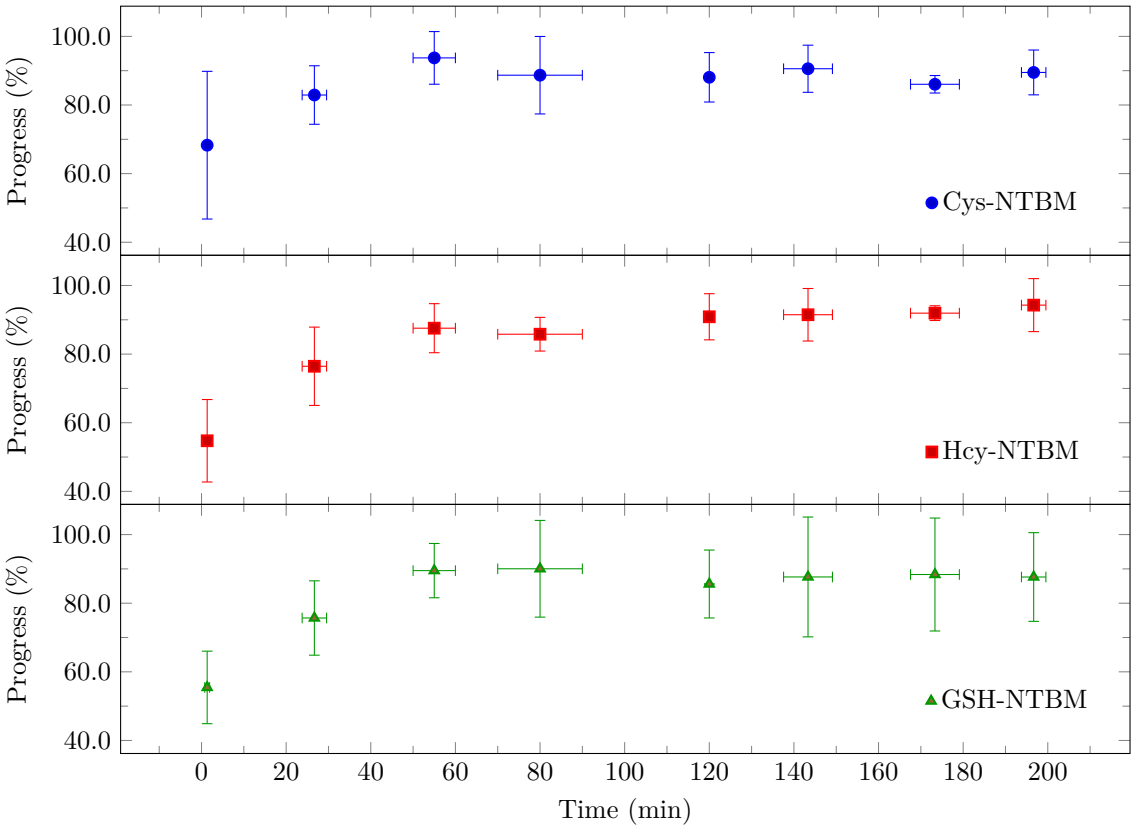


Figure 2.7: Normalized responses of thiol-NTBM products over time from the reaction between reduced thiols and 100-fold excess of NTBM at 4 °C in pH 5 buffer.

Table 2.4: Relative ion responses of five thiol-NTBM adducts measured within a week to assess short-term stability in a mixture stored at 4 °C.

Thiol-Adduct	1 h	24 h	48 h	7 d	%CV
CysGly-NTBM	3.38	3.53	3.42	3.40	2.0
Hcy-NTBM	3.27	3.33	3.30	3.23	1.3
Cys-NTBM	3.14	3.33	3.35	3.12	3.7
γ-GluCys-NTBM	2.65	2.77	2.76	2.67	2.2
GSH-NTBM	2.05	2.19	2.10	2.06	3.0

stock solutions to be used within 24 h, the stability of NTBM reagent in ACN was assessed – initially by UV-visible (VIS) absorption measurements made over 36 days. Figure 2.8 shows an overlay of the UV-VIS spectra collected for NTBM reagent from day 1 to day 36, which reveals that NTBM reagent is stable in ACN within this period. In response to later work (discussed next in Section 2.2.3) that favoured the use of methanol (MeOH) over ACN for the preparation of plasma samples, the preparation of NTBM reagent in methanol was also considered as an alternative. Following a 9-month storage period of NTBM reagent in methanol at 4 °C, its performance on derivatizing a thiol mixture was compared to that of a fresh stock of NTBM prepared in ACN. Comparison of the relative ion responses for the five thiol-adducts prepared from both NTBM stocks revealed an overall average %CV of 5.1, as shown in Table 2.5. The results are indicative of the stability of NTBM in MeOH after a 9-month storage period, since its thiol-labeling performance is comparable to that of NTBM freshly prepared in ACN, with the observed variation in ion responses being well within the largest variation typically exhibited with independent technical replicates, 24%. In order to maintain the long-term stability of NTBM reagent, it was important to use dry solvents for reagent preparation since water content in ACN and MeOH, which are both hygroscopic, could compromise and lead to hydrolysis of the maleimide. Thus, in addition to the formation of stable thiol adducts and its compatibility in both MeOH and ACN solvents, NTBM has been found to be a much more stable and user-friendly reagent than NTAM.

2.1.4 Conclusions

In summary, NTBM was shown to be an improved thiol-derivatization reagent, which overcame obstacles encountered with the use of NTAM in CE-ESI-MS. Due to the lack

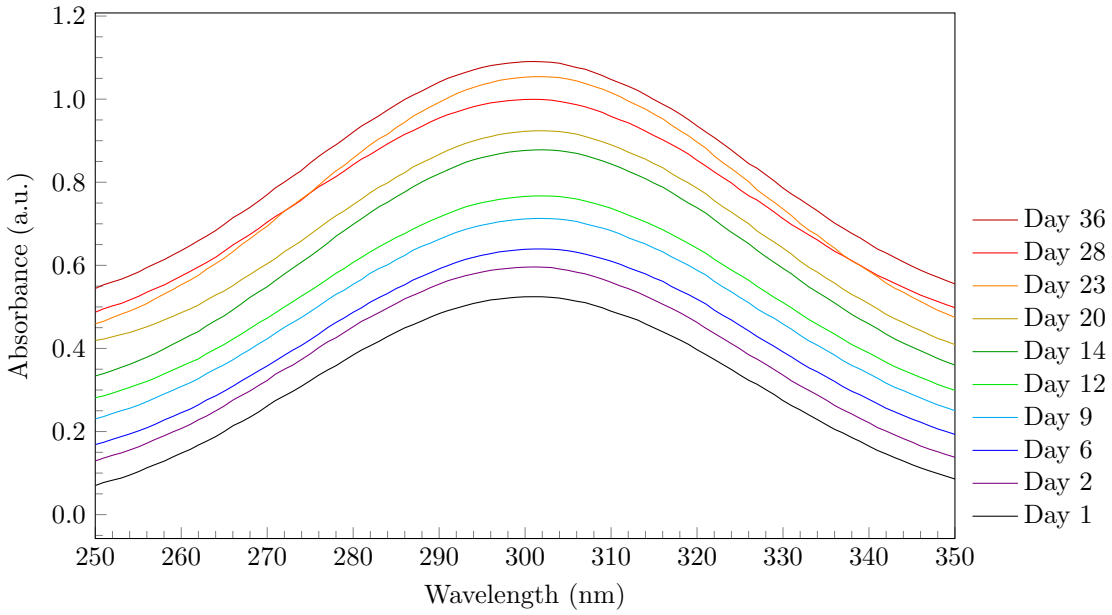


Figure 2.8: Overlay of UV-VIS spectra collected for NTBM reagent in ACN to assess its short-term stability within a 36-day period. For clarity, the spectra have been vertically offset in chronological order as listed in the legend. No spectral changes are observed and are indicative of the stability of NTBM within this short time period.

Table 2.5: Comparison of ion responses of thiol-NTBM adducts prepared when using a 9-month-old stock of NTBM in MeOH and a fresh stock of NTBM in ACN, respectively, to assess the long-term stability of NTBM reagent.

Thiol-Adduct	NTBM in MeOH Stored for 9 months	Fresh NTBM in ACN	%CV
CysGly-NTBM	2.62	2.69	1.7
Hcy-NTBM	2.37	2.57	5.6
Cys-NTBM	2.52	2.39	3.9
γ-GluCys-NTBM	1.67	1.97	11.6
GSH-NTBM	1.38	1.44	2.9
Average			5.1

of a permanent positively charged moiety, NTBM eliminates the need for additional quenching steps to remove background interference, which greatly simplifies the derivatization procedure to a single step. NTBM shows greater enhancement in the detection of thiols by 2.5-fold relative to NTAM, and 23-fold relative to unlabeled reduced thiols. When derivatization with NTBM is carried out on ice, one hour is needed for the reaction to reach completion, whereas recent novel labeling reagents generally require one hour of reaction time at elevated temperatures. The thiol adducts of NTBM were found to be stable within a week when stored at 4 °C and are expected to be stable for longer time periods when stored at −80 °C. The stability of NTBM adducts is advantageous as it allows for the reduced thiol fraction in samples to be captured at the time of sample processing and analyzed at a later date for assessment of redox status in samples when measured together with oxidized disulfides. Furthermore, the stability of NTBM reagent itself is an added benefit since it can greatly reduce the frequency of reagent preparation, as well as minimize costs associated with the elimination of unused excess reagent, which can instead be saved for future use. Overall, the simplified derivatization procedure, enhanced detection of thiols, low cost, commercial availability, and stability of the derivatized thiol-adducts make NTBM a practical and effective alternative to NTAM for thiol analysis by CE-ESI-MS.

2.2 Characterization of Protein-Bound Thiols in Plasma

This section describes the development, optimization and method validation of a protocol for the extraction and quantification of protein-bound thiols in human plasma using NTBM as the thiol-derivatizing agent.

2.2.1 Introduction

In plasma, human serum albumin (HSA) represents the most abundant protein, comprising 50% of total plasma proteins (Borowczyk et al., 2015). Aside from important physiological roles in buffering of plasma pH, regulating osmotic pressure, transporting endogenous and exogenous ligands (including metals, fatty acids, cholesterol, and hormones), HSA also functions as an extracellular antioxidant by scavenging reactive oxygen species (ROS) or reactive nitrogen species (RNS) via oxidation of the only free Cys residue (Cys34) (Ogasawara et al., 2007). Being the only thiol residue not involved in intra-protein disulfide bonding and having a pK_a of ∼5 (Borowczyk et al., 2015), which is lower than most plasma thiols (pK_a 8–9), Cys34 acts as a major redox-reactive probe that is capable of undergoing various post-translational oxidative modifications (Giles et al., 2003). Due to the naturally high abundance of HSA, Cys34 accounts for approximately 80% of the total redox-sensitive thiols in plasma with the circulating concentrations ranging from 0.6 to 0.75 mM (Borowczyk et al., 2015; Jovanović et al., 2013), in which 70–80% of total Cys34 exists in the free sulfhydryl form in healthy adults (Oettl and Stauber, 2007).

Among the various reversible and irreversible oxidative Cys modifications that can occur in HSA (as well as in other peptides and proteins), including inter-protein disulfide bonding, sulfenation, sulfination, sulfonation, nitrosation, and sulfenamidation (Rudyk and Eaton, 2014), reversible protein thiolation is the predominant form of biological oxidation with minimal redox change (Di Simplicio et al., 2005). Protein thiolation is the non-enzymatic formation of a disulfide between a protein and a LMW thiol, such as Cys and GSH (Di Simplicio et al., 2005), to either modulate protein function and/or protect protein structure by preventing irreversible oxidative modifications (Dalle-Donne et al., 2007). In healthy human plasma, the ratio of

protein-bound thiol to unbound free thiol can vary greatly between different thiols, due to various complex interactions, as well as differences in binding characteristics (Wiley et al., 1988), such that, for example, the ratio is 60:40 for Cys and 95:5 for Hcy (Bald et al., 2004).

While many HPLC-UV and HPLC-FL methods have been developed for the determination of protein-bound thiols in human plasma (Andersson et al., 1993; Bald et al., 2004; Borowczyk et al., 2015; Giustarini et al., 2005; Mansoor et al., 1992), there are often inconsistent quantitative results in reported protein-bound thiol concentrations when comparing various methods. Furthermore, these methods have generally been limited to the determination of Cys, Hcy, GSH, and CysGly, and have not included γ-GluCys (an important precursor to the synthesis of GSH). The objective of the present study is to develop a sensitive and reliable method that will allow for the measurement of low-abundance protein-bound thiols in human plasma, which may later be applied towards the untargeted profiling of Cys34 disulfides (Section 2.3), while attempting to address discrepancies in reported concentration levels of protein-bound thiols between different methods. The method is based on the derivatization of protein-bound thiols using NTBM, which first requires experimental optimization of protein precipitation and extraction of plasma metabolites, prior to analysis by CE-MS.

2.2.2 Materials and Methods

Chemicals and Reagents

In addition to the chemicals and preparation of standard solutions as described previously in Section 2.1.2, human plasma samples were obtained from the work of D'Agostino et al. (2011). As well, DL-dithiothreitol (DTT), tris(2-carboxyethyl)-

phosphine hydrochloride (TCEP), lyophilized (freeze-dried) pooled human plasma, and metabolite reference standards were purchased Sigma-Aldrich (St. Louis, MO, USA). Pooled human plasma was reconstituted to 1 mL in DI H_2O and stored as 50 µL aliquots at −80 °C until sample processing. DTT and TCEP were prepared in DI H_2O as 5 mM solutions (1 mL each), and prepared fresh before each use.

Sample Workup of Human Plasma for Protein-Thiol Reduction and Chemical Derivatization

Two organic solvents, methanol and acetonitrile, were compared in their efficacy of both protein precipitation and metabolite extraction in plasma. Four volumes (200 µL) of ice-cold solvent were added to one volume of plasma (50 µL). Adapting a method previously described for the preparation of serum (Want et al., 2006), plasma samples were left to precipitate at −20 °C for one hour following brief vortexing. Samples were subsequently centrifuged for 10 min at 13 000 *g* and the supernatants were transferred to 3 kDa cut-off Nanosep centrifugal devices (Pall Life Sciences, Washington, NY, USA), where they were ultra-filtered for 15 min at 13 000 *g*. Plasma filtrate (150 µL) was transferred to a new vial to be evaporated to dryness under vacuum at room temperature using a centrifugal evaporator (Eppendorf Vacufuge plus, AG22331, Hamburg, Germany) for 30 min before being reconstituted to 30 µL in 200 mM NH_4Ac (pH 5) aqueous buffer solution containing the IS, 25 µM 3-Cl-Tyr. The sample was stored at −80 °C until analysis. Solvent blanks for each extraction method were prepared as controls to ensure that detected features in the plasma did not originate from the extraction solvents or use of the centrifugal devices.

Upon the removal of plasma supernatant following protein precipitation, protein pellets from both methanol and acetonitrile extractions were washed with 3 different solutions (Ultra LC-MS-grade water, Ultra LC-MS-grade methanol, or 50:50 metha-

nol:water solution) in order to determine the optimal solution that would minimize the amount of residual plasma carryover, while minimizing the amount of protein lost in the procedure as explained below. The protein-washing protocol involved the addition of 500 µL of ice-cold wash solution to the pellet followed by brief vortexing, centrifugation for 10 min at 13 000 *g*, then careful removal of the supernatant to minimize removal of protein. The washing protocol was repeated three times to ensure thorough removal of residual metabolites in plasma. Reductive cleavage of disulfide bonds in the isolated protein pellet was performed by treatment with 100 µL of 5 mM DTT or TCEP for 15 min with sonication. LMW thiols were subsequently derivatized by the addition of 100 µL of 25 mM NTBM then ultra-filtered for 15 min at 13 000 *g* in 3 kDa cut-off Nanosep centrifugal devices. Filtrate (150 µL) was dried under vacuum at room temperature for 1 h prior to being resuspended in 30 µL of 200 mM NH_4Ac (pH 5) aqueous buffer solution containing the IS, 25 µM 3-Cl-Tyr. The sample was stored at −80 °C until analysis.

Capillary Electrophoresis Mass Spectrometry (CE-MS)

CE-MS separations were performed on an Agilent G7100A CE system interfaced with an orthogonal Agilent coaxial sheath liquid Jet Stream ESI source to an Agilent 6230 TOF mass spectrometer, in which a voltage of 30 kV at 25 °C was applied to 110 cm 50 µm internal diameter uncoated fused silica capillaries (Polymicro Technologies, AZ, USA), which were preconditioned similarly as described earlier in Section 2.1.2 with the exception of additional 5 min rinsing time to all capillary rinses, relative to rinse times on an 85 cm capillary. For high-throughput profiling of protein-bound thiols, MSI-CE-MS (Kuehnbaum et al., 2013) was used in this work, which enables the simultaneous analysis of seven samples within a single run. Sample run-time was set to 45 min. The TOF mass spectrometer was operated in positive-ion mode

scanning *m/z* 50–1700 with an acquisition rate of 2 Hz and acquisition time of 500 ms. The temperature of the nitrogen nebulizing gas was 300 °C at 10 psi with drying gas at 8 L/min. Sheath gas was at 3.5 L/min at 195 °C. Both the nozzle and capillary voltages (V_{cap}) were set at 2 kV, while the MS fragmentor, skimmer, and octopole radio-frequency voltages were at 120, 65, and 750 V, respectively.

2.2.3 Results & Discussion

Comparison of Plasma Extraction and Protein Precipitation Efficacy by Methanol and Acetonitrile

Efficient protein precipitation and plasma metabolite extraction were optimized in this work by first comparing the performance of ACN and MeOH in protein denaturation, in which the volume ratio of organic solvent to plasma was 4:1 (Nirungsan and Thongnopnua, 2006; Pucci et al., 2003; Yang et al., 2013). Signal responses for 35 plasma metabolite features were obtained for 3 technical replicates and compared between the 2 methods as shown in Table 2.6. Overall, it was found that MeOH protein precipitation resulted in an average signal that was 3-fold greater in both peak heights (RPH) and peak areas (RPA) relative to the metabolite responses from precipitation with ACN. The average percent relative standard deviation (%RSD) in ion response was found to be 7.7% and 15.0% for MeOH and ACN, respectively, in which the %RSD values ranged from 1.5 to 26% for MeOH, and 5.4 to 29% for ACN. The average signal responses of each metabolite measured from both methods, along with the fold-change increase in signal with MeOH protein precipitation relative to ACN, are tabulated in Table 2.6. The results reveal that greater sensitivity in signal responses from plasma extracts are obtained with MeOH extraction. Importantly, upon handling of protein pellets following removal of deproteinized plasma extracts,

Table 2.6: Average relative peak areas (RPAs), relative peak heights (RPHs), and fold-change in response (MeOH vs. ACN) of 35 plasma metabolite features measured from 3 replicates of MeOH and ACN extractions.

Metabolite	m/z	RMT	Methanol Extractions ($n = 3$)				Acetonitrile Extractions ($n = 3$)				MeOH / ACN	
			Average RPA	%RSD	Average RPH	%RSD	Average RPA	%RSD	Average RPH	%RSD	Average RPA	Average RPH
Serine	106.0499	0.844	1.584	11	2.297	12	0.406	5.4	0.552	10	3.899	4.165
Proline	116.0706	0.905	8.253	2.4	9.723	3.9	3.078	15	3.925	12	2.681	2.477
Guanidoacetic acid	118.0611	0.698	0.068	10	0.124	6.5	0.020	15	0.036	14	3.402	3.463
Valine	118.0863	0.832	7.944	4.2	9.877	3.8	3.124	11	4.234	13	2.542	2.333
Betaine	118.0863	0.954	3.377	6.8	3.825	2.0	1.674	6.7	1.760	5.6	2.018	2.173
Threonine	120.0655	0.886	3.527	3.4	4.505	4.9	0.931	14	1.186	19	3.790	3.798
Cysteine	122.0270	0.935	0.026	11	0.027	10	< LOD		< LOD			
Hydroxyproline	132.0655	1.021	1.174	5.1	1.265	3.8	0.317	11	0.324	15	3.708	3.901
Creatine	132.0768	0.745	1.202	5.1	2.072	4.7	0.390	7.4	0.601	13	3.078	3.446
Isoleucine	132.1019	0.846	4.068	7.5	6.761	3.5	1.952	18	2.872	13	2.084	2.354
Leucine	132.1019	0.856	10.007	3.4	10.340	5.2	4.718	8.8	5.644	10	2.121	1.832
Asparagine	133.0608	0.886	1.268	5.3	1.691	5.1	0.305	15	0.388	19	4.150	4.358
Ornithine	133.0972	0.582	1.801	8.3	3.579	10	0.463	17	0.875	15	3.890	4.090
Hypoxanthine	137.0458	1.087	0.886	3.9	0.991	2.6	0.546	13	0.596	11	1.623	1.662
Deoxycarnitine	146.1181	0.681	0.083	19	0.156	11	0.050	13	0.077	17	1.674	2.034
Glutamine	147.0764	0.909	12.966	2.2	11.573	6.9	4.693	15	5.515	15	2.763	2.098
Lysine	147.1128	0.584	4.750	13	8.491	10	1.325	17	2.500	14	3.585	3.397
Glutamic acid	148.0604	0.924	3.676	10	4.180	8.6	0.587	20	0.675	21	6.258	6.192
Methionine	150.0583	0.895	1.064	3.4	1.421	3.0	0.406	15	0.526	16	2.623	2.703
Histidine	156.0768	0.623	3.780	12	6.459	11	0.964	21	1.737	17	3.921	3.718
Free carnitine	162.1125	0.719	3.685	6.2	5.343	8.5	1.263	14	1.820	12	2.918	2.935
Hydroxylysine	163.1078	0.609	0.023	17	0.047	11	0.005	22	0.012	26	4.801	3.944
Phenylalanine	166.0863	0.923	6.146	1.8	6.586	2.6	3.251	14	3.698	8.9	1.891	1.781
3-Methylhistidine	170.0924	0.637	0.416	9.0	0.751	16	0.123	9.3	0.203	8.5	3.373	3.701
Arginine	175.1190	0.605	3.246	2.1	6.471	11	1.055	18	1.939	17	3.078	3.337
Citrulline	176.1030	0.936	0.988	2.5	1.131	3.8	0.210	16	0.234	15	4.700	4.828
Tyrosine	182.0812	0.956	4.401	2.3	4.790	0.8	1.861	16	2.030	13	2.365	2.360
Trimethyllysine	189.1598	0.606	0.051	26	0.100	29	0.013	29	0.023	17	3.896	4.370
Asymmetric dimethylarginine	203.1503	0.647	0.087	3.8	0.162	18	0.025	10	0.046	11	3.436	3.513
Symmetric dimethylarginine	203.1503	0.658	0.067	15	0.120	13	0.020	25	0.036	23	3.292	3.296
Acetylcarnitine	204.1230	0.761	0.917	15	1.350	18	0.639	8.6	0.898	1.4	1.435	1.503
Tryptophan	205.0972	0.924	3.026	1.5	3.286	3.3	1.758	15	2.024	13	1.722	1.624
Kynurenine	209.0921	0.876	0.117	4.9	0.150	7.6	0.070	20	0.090	21	1.669	1.654
Propionylcarnitine	218.1387	0.782	0.051	4.7	0.064	11	0.038	13	0.049	17	1.363	1.302
Butyrylcarnitine	232.1543	0.798	0.067	7.2	0.048	0.7	0.028	19	0.022	16	2.410	2.185
		Average	2.584	7.71	3.279	8.1	1.013	14.8	1.319	14.4	3.028	3.029
		Min.	0.023	1.5	0.027	0.7	0.005	4.1	0.012	1.4	1.363	1.302
		Max.	12.966	26	11.573	29	4.718	29	5.644	26	6.258	6.192

differences in the consistency of pellets were observed between the two organic solvent extraction methods. Protein pellets from ACN extractions were soft and semi-solid, which made it extremely difficult for pellets to be transferred quantitatively between vials during washing. In contrast, pellets from MeOH extractions were firm and intact, thereby, allowing for much greater ease and completeness of transfer between vials, resulting in higher recoveries of the protein following washing. Based on the superior performance of MeOH in terms of both hydrophilic metabolite signal response from plasma and extraction recovery of protein, MeOH was selected as the preferred extraction solvent to be used in subsequent protein-bound thiol analysis.

Eliminating Plasma Carryover in Protein Pellet

Prior to the analysis of protein fractions, it is important that the protein pellets are sufficiently rinsed in order to minimize carry-over from residual plasma which may contribute artifacts to the protein analysis. Early trials on the development of the protein workup protocol found that residual plasma metabolites were appreciably detected from the analysis of protein fractions that were only rinsed once with water. In an attempt to improve the protocol for a more thorough rinse, rinsing of the pellet with water had been repeated three times. However, it was found that the intact protein pellet gradually disintegrated and became more solubilized with each water rinse, leading to poor recoveries due to difficulty in quantitatively isolating the softened pellet. To keep the pellet intact and prevent its solubilization during washing, 100% MeOH was considered next as the rinsing solvent. Although the pellet had remained firm and intact throughout the rinse, several plasma metabolites were still considerably detected. Given the insolubility of plasma metabolites in MeOH, it was not surprising to see that glutamine, one of the most abundant amino acids in plasma, was substantially detected from the workup of the MeOH-rinsed protein pellet,

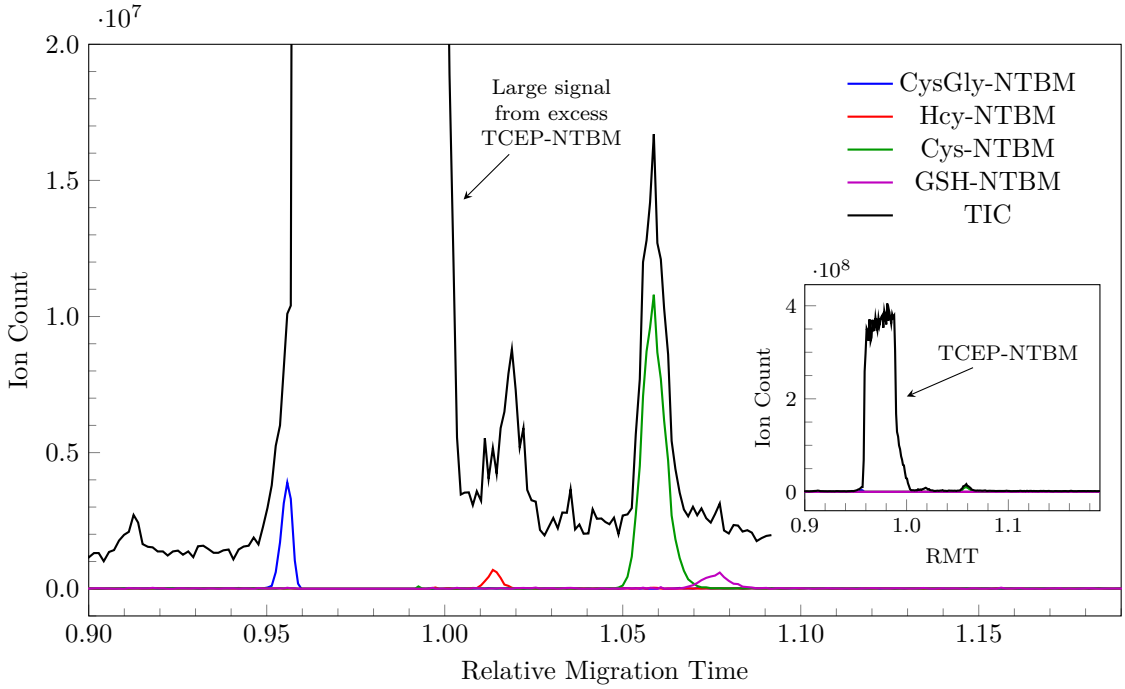


Figure 2.9: EIE overlay of the large interfering background peak from cationic fast-migrating TCEP-NTBM adduct masking low-abundance signals of thiol-NTBM adducts. (Inset) Zoomed-out TIC to show overall magnitude of ion count from TCEP-NTBM adduct.

such that its response was 50% of the measured signal from plasma. To compromise between maintaining an intact protein pellet while allowing for sufficient solubility of residual plasma metabolites in the rinse solution for their removal, 50% MeOH solution in water was tested for its efficiency in the pellet-rinsing protocol, in which it was confirmed that residual plasma carryover was best eliminated from the protein pellet. From these results, 50% ice-cold MeOH in water was confirmed as the optimal wash solution for the rinsing of protein pellets prior to their derivatization.

Reducing Protein Disulfide Bonds: DTT vs. TCEP

To quantify protein-bound thiols in plasma, reduction of protein thiols are typically carried out using DTT or TCEP prior to labeling (Hermanson, 2013). Earlier work performed on IT-MS for the analysis of NTBM-labeled protein-bound thiols reduced

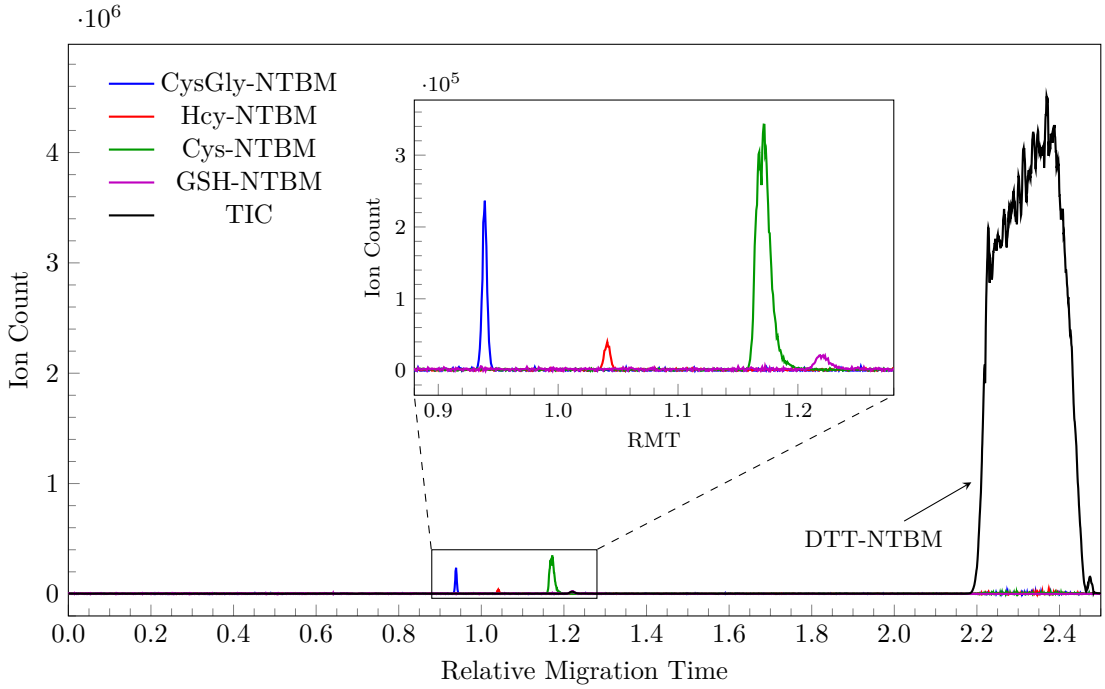


Figure 2.10: EIE overlay of low-abundance thiol-NTBM adducts (inset) and TIC showing neutral DTT-NTBM migrating late as an EOF marker, which does not interfere with measurement of thiol-NTBM adducts.

with TCEP have shown the presence of an early migrating massive background peak (Figure 2.9), whose m/z value corresponds to a singly and positively-charged TCEP-NTBM adduct (m/z 404.1). With a peak width spanning nearly a minute, the presence of an enormous background peak with a high positive mobility is highly undesirable, as it is likely to interfere and cause ion suppression to the analysis of thiol-NTBM adducts or other analytes of interest that may be migrating within the large 1-minute peak width window. As a result of the risk of possible interference, TCEP was not used for the reduction of protein-bound thiols prior to maleimide derivatization, despite it being known to be a stronger and faster reducing agent than DTT at pH < 8 (Han and Han, 1994). For this reason, DTT was considered for the reductive cleavage of protein-bound thiols. With DTT possessing two thiol groups that are reactive towards the maleimide, the addition of NTBM was ensured to be present in at least a 2-fold

Table 2.7: Validation criteria for the analysis of NTBM-derivatized thiols.

Thiol-NTBM Adduct	m/z : RMT (MH^+)	Linear Range (µM)	Sensitivity (μM^{-1})	Linearity (R^2)	LOD (SNR = 3) (µM)
Cys-NTBM	275.106 : 1.121	0.5–200	0.323 ± 0.007	0.9963	0.06
Hcy-NTBM	289.122 : 1.033	0.2–20	0.950 ± 0.057	0.9985	0.02
CysGly-NTBM	332.128 : 0.933	0.1–20	0.362 ± 0.028	0.9963	0.02
γ-GluCys-NTBM	404.149 : 1.135	0.1–20	0.274 ± 0.002	0.9920	0.04
GSH-NTBM	461.170 : 1.172	0.1–20	0.615 ± 0.058	0.9881	0.03

molar excess of the amount of DTT added plus the expected amount of reduced thiols combined. In this case, a 5-fold excess of NTBM relative to DTT was added to the protein sample. The resulting EIE, as shown in Figure 2.10, reveals a large peak at the EOF, which is believed to be the neutral DTT-NTBM adduct in excess. Therefore, the results indicate that the use of DTT as a reducing agent does not pose any risk of interference to the CE-MS analysis of protein-bound thiols as maleimide adducts. Figure 2.11 illustrates the optimized protocol for preparation and derivatization of the protein fraction.

Method Validation

External calibration curves measured in triplicate for the NTBM-derivatized thiols were acquired over a concentration range that was selected based on previously reported protein-bound thiol concentrations (Borowczyk et al., 2015). Linearity and sensitivity of the method were assessed using six to seven calibration points and are summarized in Table 2.7 along with the detection limits for each thiol. Instrument precision was determined through replicate intra- and inter-day measurements ($n = 3$), in which overall RSDs ranged from 0.9–10%.

To assess the method reproducibility of the protein-bound thiol assay, six replicate preparations of protein samples from the same batch of plasma were performed and analyzed. RSDs ranged from 6.5–28%, in which the quantification of protein-bound

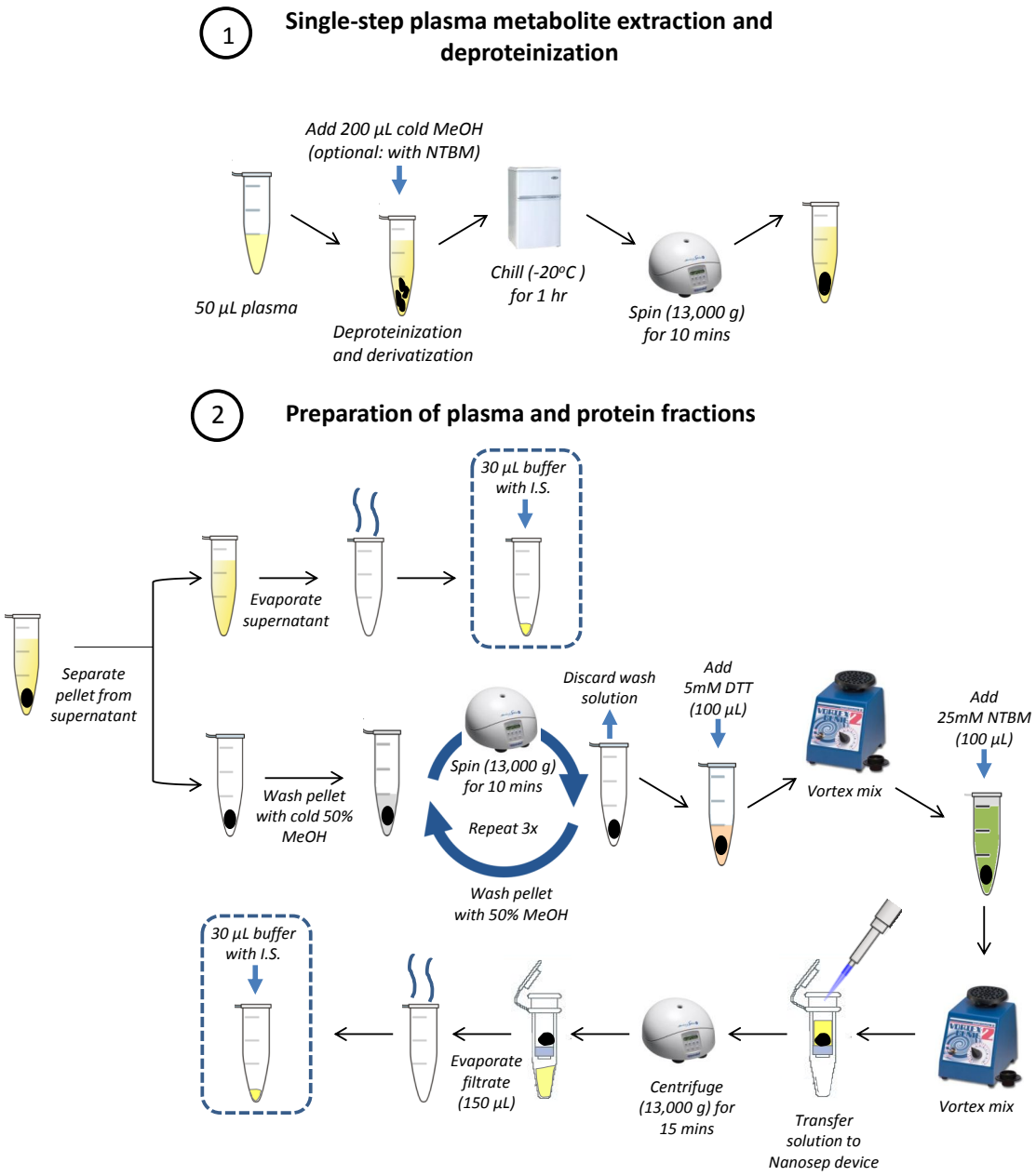


Figure 2.11: Illustration of the protocol for plasma and protein preparation.

Table 2.8: Average concentrations of protein-bound thiols in pooled human plasma.

Thiol-NTBM Adduct	Concentration (µM)*	Relative Abundance
Cys-NTBM	92 ± 6	1.000
Hcy-NTBM	0.95 ± 0.13	0.010
CysGly-NTBM	11 ± 1	0.120
γ-GluCys-NTBM	0.46 ± 0.13	0.005
GSH-NTBM	0.36 ± 0.09	0.004

* Values are expressed as mean ± standard deviation.

Table 2.9: Comparison of mass LOD between various assays for LMW thiols.

	Analytical Platform	Volume* (nL)	Cys (fmol)	Hcy (fmol)	CysGly (fmol)	γ-GluCys (fmol)	GSH (fmol)
NTBM derivatization	CE-MS	3.5	0.21	0.07	0.07	0.14	0.11
D'Agostino et al. (2011)	CE-MS	63	0.50	0.63	0.63	0.63	0.38
Borowczyk et al. (2015)	HPLC-UV	2000	160	300	200		300
Isokawa et al. (2013)	HPLC-FL	5000	7.5	4.0	1.5	17	4.0
Zhang et al. (2014)	HPLC-UV	20 000	800	1600	1200		1200

* Sample injection volume.

γ-GluCys and GSH exhibited the greatest variability with RSDs of 28% and 25%, respectively. This is not unreasonable given the low concentrations of those thiols. Average protein-bound thiol concentrations of the six replicates are shown in Table 2.8. Comparison of the absolute mass limit of detection (LOD) to those of previous methods, which were mainly based on HPLC with UV or FL detection, reveals enhancements in thiol detection limits of up to four orders of magnitude when taking into account the extremely low sample injection volume of CE-MS, in conjunction with derivatization by NTBM, as seen in Table 2.9. While the HPLC methods compared have used injection volumes of 2–20 µL, CE injects only 3.5 nL per sample, which offers the potential to further lower detection limits upon application of online sample preconcentration, as demonstrated previously in the work of D'Agostino et al. (2011).

Table 2.10: Comparison of mean concentrations of protein-bound thiols between different methods.

	n	Mean ± Standard Deviation Concentration (µM)				
		Cys	Hcy	CysGly	γ-GluCys	GSH
NTBM derivatization	6	92 ± 6	0.95 ± 0.13	11.4 ± 1.0	0.46 ± 0.13	0.36 ± 0.09
Borowczyk et al. (2015)	20	158 ± 26	3.4 ± 0.8	21.6 ± 4.6		2.8 ± 1.1
Giustarini et al. (2005)	15	147 ± 24	7.5 ± 2.8	11.3 ± 1.8		1.7 ± 0.7
Bald et al. (2004)	8	229 ± 96	8.7 ± 2.8	17.3 ± 6.6		6.7 ± 4.7
Andersson et al. (1993)	10	146 ± 29	7.3 ± 2.1	20 ± 11	1.2 ± 0.8	0.7 ± 3.0
Mansoor et al. (1992)	10	155 ± 18	9.8 ± 2.9	15.1 ± 3.2		1.3 ± 0.2
%CV		28	54	27	63	103

Pitfalls of Direct Protein-Bound Thiol Quantification

A challenge encountered with the derivatization of thiols in protein samples is the inability to correct for the amount of protein that is physically lost to repeated rinsing of the pellet and removal of the wash supernatant. Unless the combined thiol content in plasma and protein, and the total thiol content contributed by the plasma alone, are analyzed independently of the isolated protein fraction, the absolute recovery of the protein isolation protocol cannot be directly quantified, which is believed to be a reason leading to inconsistencies in reported values of absolute quantification of protein-bound thiols. Although several recent reports have developed chromatographic methods for the speciation analysis of protein-bound thiols in plasma (Bald et al., 2004; Borowczyk et al., 2015; Giustarini et al., 2005), recovery and precision relating specifically to the protein isolation procedure have not been addressed. Comparison of the measured protein-bound thiols using derivatization by NTBM with other methods reveals that concentrations in previous reports may be anywhere from being 1- to 18.5-fold greater in magnitude (Table 2.10). Though it appears at first glance that measurements of protein-bound thiols from other methods are markedly greater, closer examination of the protocols described reveals several deficiencies.

In the work of Borowczyk et al. (2015) based on HPLC with UV detection,

which show 2- to 8-fold greater protein-bound thiol concentrations compared to the current work, it is noticed that there is no mention of protein being washed prior to derivatization with thiol-labeling reagent. Not only would the lack of washing eliminate any losses of protein, but the protein-bound thiol measurements would be likely to be artificially inflated due to contamination by plasma thiols. Furthermore, no measurements of precision for the protein-bound thiol concentrations quantified in 20 different plasma samples were presented by the authors, indicating that method reproducibility was not addressed in the study. Lastly, detailed discussion of the method validation was presented solely for the determination of total thiols (Cys, CysGly, Hcy, and GSH) present in plasma, and not for the determination of thiols specifically bound to protein. As such, method reproducibility and recovery of protein-bound thiol determination were not addressed and it is very likely that the thiol measurements are overestimated. While the methods presented by Bald et al. (2004) and Giustarini et al. (2005) both incorporate duplicate washing of the protein pellets, either with 1.5% (w/v) trichloroacetic acid or water, respectively, recovery of thiols from the protein preparation protocol is not addressed in either.

The work of Bald et al. (2004) reported results for total, reduced, total free and protein-bound Cys, CysGly, Hcy, and GSH, in which subtraction of the measured plasma-derived total free amount from the measured total thiol content in protein and plasma theoretically yield the contribution of thiol from the protein fraction alone. The theoretical value of protein-bound thiol could then, in turn, be used to calculate recovery for the empirical determination of protein-bound thiols. However, in order for the recovery calculation to be valid, the determination of protein-bound thiols must be done independently of the analysis for total free thiols in plasma, especially since the two values are mathematical complements of each other. Independent analyses would prevent the technical error originating from the protocol for plasma free thiol

Table 2.11: Comparison of percent protein-bound thiol composition between different methods, excluding γ-GluCys.

	Analytical Platform	Mean ± Standard Deviation Composition (%)			
		Cys	Hcy	CysGly	GSH
NTBM derivatization	CE-MS	88 ± 8	0.91 ± 0.13	10.5 ± 1.1	0.34 ± 0.09
Borowczyk et al. (2015)	HPLC-UV	85 ± 3	1.8 ± 0.3	11.6 ± 2.9	1.5 ± 0.6
Giustarini et al. (2005)	HPLC-FL	88 ± 19	4.5 ± 1.7	6.7 ± 1.1	1.0 ± 0.4
Bald et al. (2004)	HPLC-UV	88 ± 4	3.3 ± 1.4	6.7 ± 1.7	2.5 ± 1.7
Andersson et al. (1993)	HPLC-UV	84 ± 22	4.2 ± 1.7	11.5 ± 6.6	0.4 ± 2.0
Mansoor et al. (1992)	HPLC-FL	86 ± 13	5.4 ± 1.7	8.3 ± 1.9	0.7 ± 0.1

determination from being carried forward and directly impacting the subsequent determination of the thiols in the protein-bound fraction within the same trial. In the case of the work by Bald et al. (2004), analysis of both the total plasma thiol content and the total protein-bound thiols for each sample were conducted within the same trial, which precludes the independence of the two measurements and thus, prevents a proper assessment of the recovery of the protein-bound thiol determination protocol.

Besides the lack of inadequate method validation, and direct methods for comparing absolute quantification of protein-bound thiols between methods, the absence of subject cohort information in most method validation studies also further makes it difficult to identify inter-method differences in measurements as arising from biological variation due to differences in subject cohort characteristics, or originating from method bias. As a result, large %CV values are observed, ranging from 27–103% between methods, as shown in Table 2.10. Since the amount of protein lost cannot be accurately determined, quantification of protein-bound thiol concentrations should be instead calculated relative to the sum of measured protein-bound thiol concentrations, such that the percent composition of protein-bound thiols is reported. Relative quantification would not be affected by variations in the absolute amount of recovered protein, thereby, making it a much more reliable and robust metric for the inter-method comparison of protein-bound thiols. Table 2.11 summarizes the percent protein-bound thiol

composition measured by different methods. Since γ-GluCys was not detected by all of the methods compared in Table 2.11, it was excluded from the calculation of total protein-bound thiols. Compared to Table 2.10, there is improved agreement between the methods, such that no outliers are identified in the box-and-whisker plots (Figure 2.12) for the relative quantifications of protein-bound thiols determined by six different methods. Based on the improved comparability between methods, relative quantification with respect to the sum of measured protein thiols has been shown to provide much greater robustness in its performance relative to absolute quantification, and is therefore, strongly recommended for reporting purposes in future studies involving the determination of protein-bound thiols.

2.2.4 Conclusions

A reliable CE-MS method for the determination of protein-bound thiols in plasma using was developed using NTBM as the thiol-selective labeling reagent following DTT reduction of washed plasma protein retentate. Due to the unexpected cross-reactivity between TCEP and NTBM, DTT was the preferred reducing agent for the current study due to its formation of a neutral adduct with NTBM. Overall, technical precision in the preparation protocol for six replicates was found to be acceptable (RSD $< 15\%$) for the abundant protein-bound thiols, Cys, CysGly, and Hcy, as well as for the low-abundance thiols, GSH, and γ-GluCys (RSD $< 30\%$).

Relative quantification of protein-bound thiols in human plasma allows for more reliable reporting as it is not sensitive to variations in the sample preparation protocol when comparing different methods and is therefore, more robust. Based on the results from this work in which protein was extracted from pooled plasma, relative quantification of protein-bound thiols in human plasma revealed that Cys (88%)

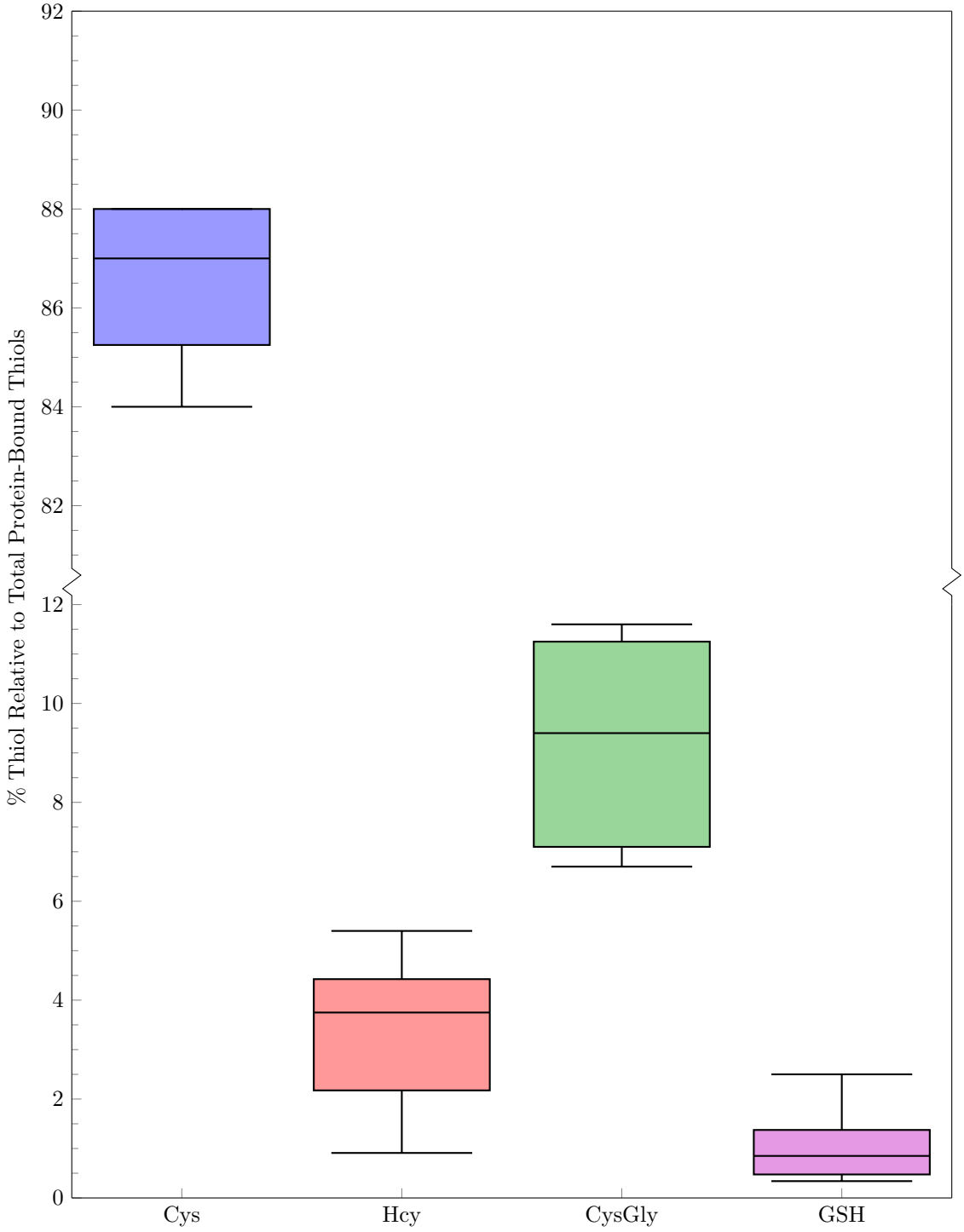


Figure 2.12: Box-and-whisker plots of the distribution of average relative quantifications of protein-bound thiols as determined by 6 methods. No outliers were detected.

and CysGly (10.5%) were the most abundant thiols, followed by Hcy (0.9%), γ-GluCys (0.4%) and GSH (0.3%). While the current investigation was focused on the optimization of protein sample preparation and analysis of protein-bound thiols by derivatization with NTBM, future studies could consider the addition of NTBM in the methanol solvent used for the initial deproteinization and extraction of metabolites from plasma. This would easily allow for both plasma deproteinization and NTBM-derivatization of free plasma thiols to occur simultaneously within a single step upon the addition of NTBM solution. Overall, the high sensitivity and precision of the method presented here makes it a promising tool for future metabolomic studies of protein-bound thiols in plasma, not only for the improved assessment of redox status that is relevant for assessment of physiological and pathological states, but also for the discovery of potentially unknown thiols that may be derived from human exposures, including diet and drug intake.

2.3 A Novel Approach for the Discovery of Unknown Protein-Bound Thiols in Plasma

Using the NTBM thiol-labeling procedure on plasma protein developed in Section 2.2, this section describes the application of a recently developed high-throughput MSI technique for CE (Kuehnbaum et al., 2013) towards the potential discovery of novel protein-bound thiols in plasma.

2.3.1 Introduction

Despite recent studies that have developed untargeted methods to analyze covalent modifications to HSA-Cys34 arising from environmental exposures (i.e., 'Cys34 ad-

ductomics') (Grigoryan et al., 2016; Rappaport et al., 2012), few studies to date have considered performing untargeted characterization of protein-bound thiols, which so far, are understood to be associated with alterations in serum redox state (Nagumo et al., 2014; Oettl and Marsche, 2010). Unlike irreversible covalent Cys34 adducts which comprise 4% of circulating HSA (Aldini et al., 2008) and require digestion with trypsin for quantification, formation of reversible Cys34 disulfides represent approximately 25–40% of HSA (Carballal et al., 2003; Lepedda et al., 2014; Ogasawara et al., 2007), and are readily reduced with DTT.

Thus far, analysis of thiols in biological samples has mainly been performed in a targeted manner for the determination of known endogenous thiols (e.g., Cys, Hcy, and GSH), thiol drugs (e.g., Cysteamine, Captopril, Tiopronin, and Penicillamine) and expected thiol-drug adducts (Albert et al., 2012; Isokawa et al., 2014; Kuśmierek et al., 2009, 2011). This has allowed for speed and efficiency of quantitative analyses, however, metabolite coverage is limited and as such, does not permit the discovery of potentially novel metabolites or pathways that may be linked to environmental exposures (including drugs and dietary intake), or associated with unique phenotypes, which could be important to understanding the pathology of disease (Tautenhahn et al., 2012; Vinayavekhin and Saghatelian, 2010). While many thiol screening assays have been employed, these methods have generally been applied towards the compositional analysis of petroleum distillates (Thomson et al., 1997), air samples (Bramanti et al., 2006), and in foods and beverages for aroma and flavour enhancement, including meat, cheese, wine, beer, coffee, tea, olive oil, and fruits (Gascó and Barrera, 1972; Sourabié et al., 2008, 2011; Vermeulen et al., 2005). Due to the complexity of the matrix and/or low concentration of thiols in these sample types, many of these methods have relied on a combination of derivatization and extraction techniques to selectively concentrate thiols prior to analysis. However, despite the variety of thiol screening assays reported,

the actual screening of unknown thiols from chromatographic data acquired from conventional analytical methods continues to be a major bottleneck in data processing.

Recently, a novel approach for thiol screening in wine based on differential analysis was presented (Inoue et al., 2013). Developed using ultra-performance liquid chromatography (UPLC) with FL and ESI-TOF mass spectrometric detections, the screening assay consisted of monitoring (thiol) peaks from underivatized samples, which would be decreased upon derivatization. New peaks which appeared in the analysis of derivatized samples at a larger m/z, whose mass difference is equal to the mass of the derivatization reagent, therefore, corresponded to derivatized thiols. Although this method is promising and was able to detect the presence of glutathione in rice wine for the first time, the method requires underivatized and derivatized preparations of each sample to be run separately, which is low-throughput and time-consuming, depending on the number of samples. Secondly, with this method, differential analysis is only limited to those thiols that are detectable without derivatization, thereby, excluding those low-abundance and/or low sensitivity thiols that require derivatization in order to be detected. Since it is expected that undiscovered protein-bound thiols in plasma are present at low levels, such that they are currently not amenable to detection without prior derivatization, extraction or pre-concentration, the method of differential analysis is not applicable to the screening of biological thiols.

By applying the method described in the previous section (2.2), which was shown to allow for the characterization of both high- and low-abundance protein-bound thiols in plasma, the screening of unknown and novel thiols may be achieved in a high-throughput manner, with the incorporation of the MSI-CE-MS strategy (Kuehnbaum et al., 2013). Possessing a half-life of three weeks (Andersen et al., 2014), HSA may be thought of as a long-lived redox probe, whose reactive Cys34 residue is sensitive to thiol-disulfide exchange interactions while circulating at high concentrations. Thus,

effective enrichment of known and unknown thiols can be simply achieved by isolation of plasma protein following which thiol-specific reduction, chemical derivatization by maleimide, and MS/MS would allow for subsequent detection and characterization. In this section, a novel strategy for the screening and discovery of unknown protein-bound thiols in plasma by CE is presented using thiol-specific labeling by NTBM and by adaptation of the MSI-CE-MS technique.

2.3.2 Materials and Methods

Chemicals and Reagents

Chemicals, reagents, and samples were obtained and prepared as described previously in Section 2.2.2. Protein was precipitated, washed, reduced and derivatized as per the protocol established in Section 2.2.2, in which methanol and 50% methanol were selected as the preferred solvents for metabolite extraction from plasma and the washing of protein pellets, respectively, based on the results shown in Section 2.2.3. Reduction of disulfide bonds and thiol derivatization were performed using DTT and NTBM, respectively, and the resulting protein solution was prepared for sample analysis. Filtered matrix blank samples were prepared by combining all reagents, excluding the protein sample, and filtering the solution through 3 kDa cut-off Nanosep centrifugal devices to account for the presence of peaks that originate from the sample matrix. Negative control samples were prepared by the addition of protein with only DTT or NTBM reagent to identify peaks corresponding to free reduced thiols or correct for reagent peaks.

Capillary Electrophoresis Mass Spectrometry (CE-MS)

Optimized protein sample preparations were analyzed on an Agilent G7100 CE system interfaced with an orthogonal Agilent coaxial sheath liquid Jet Stream ESI source to an Agilent 6550 iFunnel quadrupole-time-of-flight (Q-TOF) mass spectrometer. Uncoated 120 cm fused silica capillaries (50 µm internal diameter) were preconditioned as described earlier (Section 2.2.2). Separations were performed at 30 kV and 25 °C using the MSI-CE-MS strategy (Kuehnbaum et al., 2013), which allowed for high-throughput screening of protein-bound thiols while temporally encoding mass spectral information for signal pattern recognition via combinations of injection configurations specified by the user. Acquisition was operated in positive-ion mode scanning *m/z* 50–1700 at a rate of 1 spectra/s. The temperature of the nitrogen nebulizing gas was 200 °C at 8 psi with drying gas at 16 L/min. Sheath gas was delivered at 3.5 L/min at 199 °C. Both the nozzle and capillary voltages were set at 2 kV, while the MS fragmentor, skimmer, and octopole radio-frequency voltages were at 380, 65, and 750 V, respectively. MS/MS experiments were performed at 3 collisional energies (10, 20, and 40 V) in a continuous alternating scan mode on known thiol-NTBM adducts. Mass scan range was 20–500 *m/z* at a rate of 1 spectra/s with medium isolation width of 4 *m/z*.

Untargeted Feature Extraction of Novel Protein-Bound Thiols

To aid in the novel discovery of protein-bound thiols by Q-TOF-MS, untargeted feature selection was performed on NTBM-derivatized protein samples using the Molecular Feature Extractor (MFE) (MassHunter Qualitative Analysis, Agilent Technologies Inc.) algorithm, in which a feature was defined as having a peak height of at least 300 counts, a minimum signal-to-noise ratio (SNR) of 10 in the EIE, and a minimum

overall peak quality score of 75, which is calculated based on the match of the measured monoisotopic *m/z* compared to the predicted monoisotopic *m/z* of the most likely molecular formula, and the match of the measured isotope abundance and isotope spacing relative to the predicted isotopic distribution and spacing of the proposed formula. Following the generation of the list of extracted features by MFE, the Molecular Formula Generation (MFG) algorithm in the Agilent MassHunter software was then used to generate molecular formulae from the exact mass of the monoisotopic peak based on the measured isotope spacing and isotope abundances. The mass errors and overall scores of the generated formulae are given in Table 2.12. Subsequently, screening of thiol-NTBM adducts was performed manually to select only those features which satisfied all of the following criteria: 1) Must be absent in the negative controls and matrix blank samples; 2) Must contain at least one sulfur atom; 3) Must at least be the mass of the NTBM adduct (i.e., mass = 153.079); and 4) Empirical formula must at least contain the formula of the NTBM adduct (i.e., $C_8H_{11}NO_2$).

Automated Assignment of MS/MS Spectra

Interpretation of MS/MS spectra on known thiol standards was performed using the Agilent MassHunter Molecular Structure Correlator (MSC) software, in which Molfiles of the known structures were inputted prior to running the program. MSC operates based on correlating accurate mass MS/MS fragment ions with user-defined or MSC-calculated probable molecular formulae and structures by implementation of a systematic bond-breaking approach described by Hill and Mortishire-Smith (2005). Proposed molecular formulae were then searched against structures from the integrated ChemSpider (Pence and Williams, 2010, http://www.chemspider.com) and PubChem (Kim et al., 2016, http://pubchem.ncbi.nlm.nih.gov) databases, which were appropriate in this case since both databases contain synthetic, in addition to

naturally occurring, compounds among which the novel thiol-NTBM adducts may be found. Following the database search by MSC, proposed molecular structures were assigned an overall correlation score to describe the degree of match with the MS/MS spectrum. The overall correlation score consists of individual scores for each fragment ion signal, which considers multiple substructures candidates that are each assigned a "penalty" based on the number and type of bonds broken to generate the substructure. Bonds that require more energy to break are more unlikely to occur and are, therefore, assigned a higher penalty and a lower score.

2.3.3 Results & Discussion

Signal Pattern Recognition of Unknown Thiols by MSI-CE-MS

To increase sample throughput, MSI-CE-MS, a technique for the simultaneous analysis of multiple samples previously demonstrated by Kuehnbaum et al. (2013), was implemented for the screening of novel thiols. Capable of analyzing seven samples simultaneously while maintaining sufficient peak resolution between injected samples with high reproducibility, MSI-CE-MS is a much higher throughput approach than conventional methods that analyze a single sample at a time. By implementing MSI-CE-MS, analysis of derivatized protein samples along with negative control samples, and a filtered matrix blank were acquired within a single run. Since derivatization of protein-bound thiols requires the presence of both the reducing (DTT) and derivatization (NTBM) agents, the negative control samples consisted of protein samples with only either DTT or NTBM present, while the matrix blank consisted of all the reagents, except for the protein. In order to distinguish between features or signals originating from different sample types within the same run, a customized injection and dilution sequence was applied such that each sample type would have a characte-

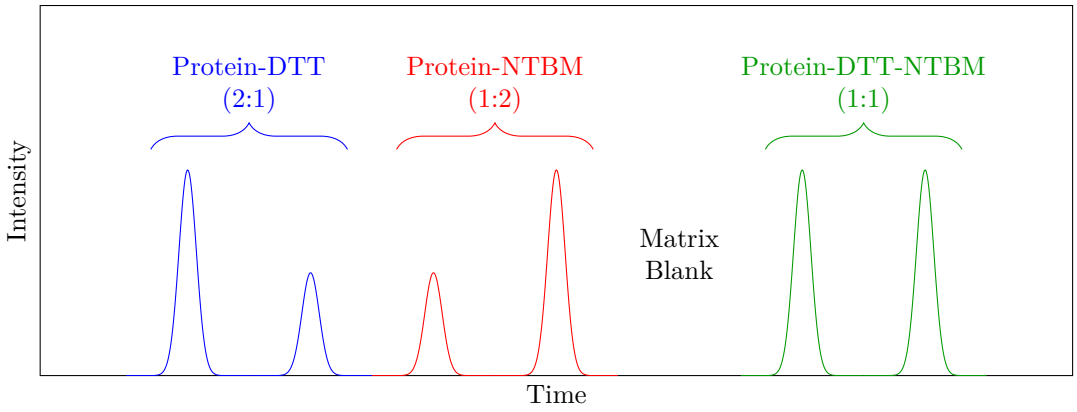


Figure 2.13: Depiction of sample configuration used in the current study for signal pattern recognition by MSI-CE-MS for the discrimination of signals originating from different samples.

ristic signal pattern. The success of this strategy has recently been demonstrated in the work of DiBattista et al. (2017) and allows for greater visual clarity and ease of identifying signals through the pattern-based encoding of signals. In this particular case involving a matrix blank and three different samples, a signal dilution pattern of 1:2, 2:1, and 1:1 was implemented to encode for the protein+DTT, the protein+NTBM, and the protein+DTT+NTBM samples, respectively. This was created using the following 7-injection sequence (depicted in Figure 2.13): protein+DTT, protein+DTT (2× dilution), protein+NTBM (2× dilution), protein+NTBM, filtered matrix blank, protein+DTT+NTBM, and protein+DTT+NTBM (duplicate injection).

Detection of Novel Unknown Compounds and Screening of Candidate Molecular Formulae

Manual filtering of the MFE results for signals that exhibited only the 1:1 pattern, which is assumed to correspond to derivatized protein-bound thiols, revealed 11 distinctive molecular features compiled from the analyses of pooled human plasma and plasma from a healthy male volunteer, respectively. Prospective formulae or candidate elemental compositions for each m/z was limited to a mass deviation of

within 5 ppm of the measured accurate mass. The elemental compositions of both the thiol-NTBM adducts and the corresponding native thiols (i.e., without NTBM adduct) were evaluated for their plausibility based on a set of rules defined by Watson (2013). In this case, rules that were applied to the current study include the following:

1. Formulae containing any number of carbon, hydrogen, sulphur, and oxygen atoms cannot have an even protonated molecular weight;

2. Formulae with odd numbers of nitrogen have even protonated molecular weights and vice versa;

3. Formulae with greater than seven nitrogen atoms are rare unless peptides are being considered;

4. It is rare for the number of nitrogen atoms to exceed the number of carbons or the number of oxygen atoms to be greater than the number of carbon atoms plus one – thus, it is unlikely for the total of nitrogen and oxygen atoms to exceed the number of carbon atoms;

5. It is unusual to find more than two sulphur atoms in a formula.

Following screening of the candidate elemental formulae, the elemental composition corresponding to the NTBM adduct was subtracted from the candidate formula of each unknown derivatized thiol-NTBM adduct in order to calculate the 'expected' chemical formula of the unknown native thiol. The resulting 'expected' formulae of the unknown thiols were then screened and evaluated as per rules 3–5 stated above.

The list of the 11 molecular features postulated as putative thiol-NTBM adducts based on the 1:1 signal patterning, as shown in Figure 2.14, are summarized in Table 2.12 along with their respective m/z, relative migration time (RMT), SNR,

Table 2.12: Summary of detected derivatized protein-bound thiols.

Derivatized Thiols			Proposed Molecule			Expected Native Thiol	
m/z : RMT (Q)*	SNR†	Formulae	Mass Error (ppm)‡	Score§	Formulae♯	ID
275.106 : 1.177 (+)	900	$C_{11}H_{18}N_2O_4S$	−0.52	98.58	$C_3H_7NO_2S$	Cys
289.122 : 1.045 (+)	55	$C_{12}H_{20}N_2O_4S$	−0.36	99.30	$C_4H_9NO_2S$	Hcy
293.117 : 1.240 (+)	50	$C_{11}H_{20}N_2O_5S$	−0.20	95.94	$C_3H_9NO_3S$	
307.132 : 1.200 (+)	25	$C_{12}H_{22}N_2O_5S$	−0.01	97.21	$C_4H_{11}NO_3S$	
307.132 : 1.272 (+)	10	$C_{12}H_{22}N_2O_5S$	−0.01	97.21	$C_4H_{11}NO_3S$	
332.128 : 0.928 (+)	145	$C_{13}H_{21}N_3O_5S$	0.37	99.30	$C_5H_{10}NO_3S$	CysGly
367.153 : 1.255 (+)	15	$C_{14}H_{26}N_2O_7S$	0.19	95.43	$C_6H_{15}NO_5S$	
367.153 : 1.329 (+)	10	$C_{14}H_{26}N_2O_7S$	0.19	95.43	$C_6H_{15}NO_5S$	
404.149 : 1.188 (+)	20	$C_{16}H_{25}N_3O_7S$	1.17	96.57	$C_8H_{14}N_2O_5S$	γ-GluCys
461.170 : 1.222 (+)	10	$C_{18}H_{28}N_4O_8S$	−0.50	99.31	$C_{10}H_{17}N_3O_6S$	GSH
573.357 : 0.895 (+)	15	$C_{30}H_{48}N_6O_3S$	−0.48	96.15	$C_{24}H_{37}N_5OS$	

* Q = charge state. Here, the charge state '+' represents $[M + H]^+$.
† The SNR indicated is the higher of the values obtained from measurements on two sets of plasma samples (pooled human plasma purchased from Sigma-Aldrich or plasma from a healthy male volunteer).
‡ Mass accuracy is calculated as ppm error where calculated m/z is subtracted from the experimental m/z, divided by the experimental m/z, and multiplied by 10^6.
§ Score is calculated based on the degree of matching to the mass, isotope abundance, and isotope spacing of the proposed formula after applying the corresponding weighting factors of 100, 60, and 50%, respectively.
♯ The expected native thiol formula is calculated by subtraction of NTBM adduct (i.e., $C_8H_{11}NO_2$) from the proposed formula.

Table 2.13: Summary of underivatized native protein-bound thiols.

Native Thiol		Mass Error (ppm)†	Score‡	Proposed Thiol	
m/z : RMT (Q)*	SNR			Formulae	ID
122.027 : 0.944 (+)	15	−3.70	91.09	$C_3H_7NO_2S$	Cys
134.027 : 1.187 (+)	145	−1.08	82.20	$C_4H_7NO_2S$	
206.048 : 1.349 (+)	15	0.07	94.12	$C_7H_{11}NO_4S$	

* Q = charge state. Here, the charge state '+' represents $[M + H]^+$.
† Mass accuracy is calculated as ppm error where calculated m/z is subtracted from the experimental m/z, divided by the experimental m/z, and multiplied by 10^6.
‡ Score is calculated based on the degree of matching to the mass, isotope abundance, and isotope spacing of the proposed formula after applying the corresponding weighting factors of 100, 60, and 50%, respectively.

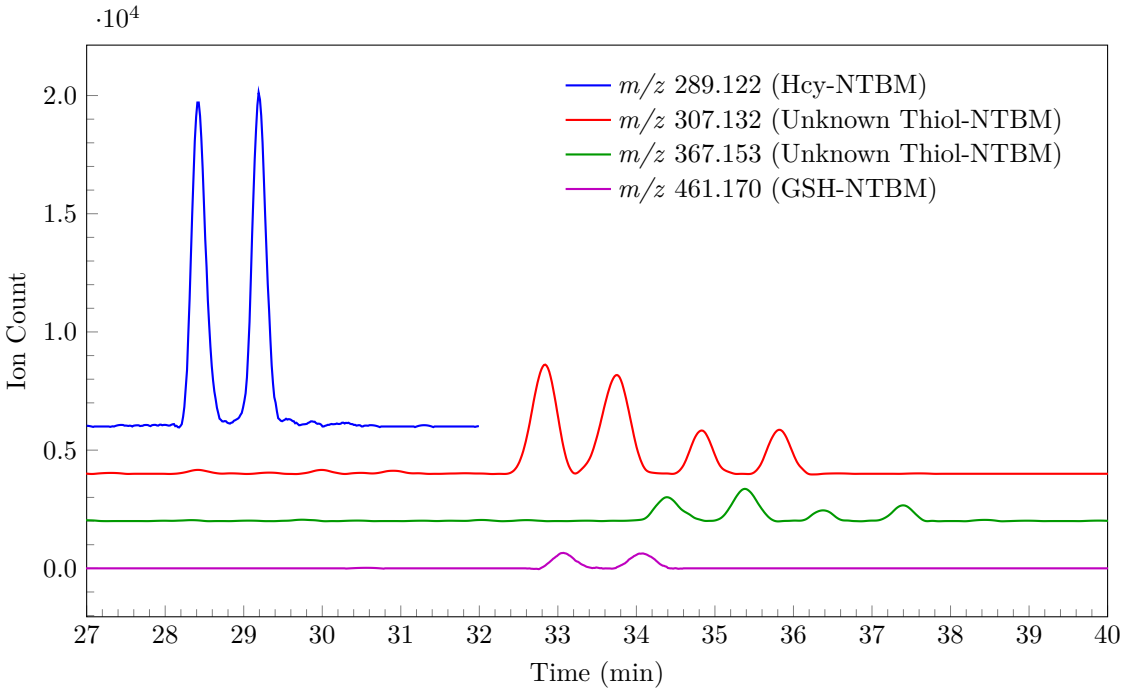


Figure 2.14: EIE overlay of four of the detected protein-bound thiols derivatized with NTBM, which show the 1:1 signal patterning.

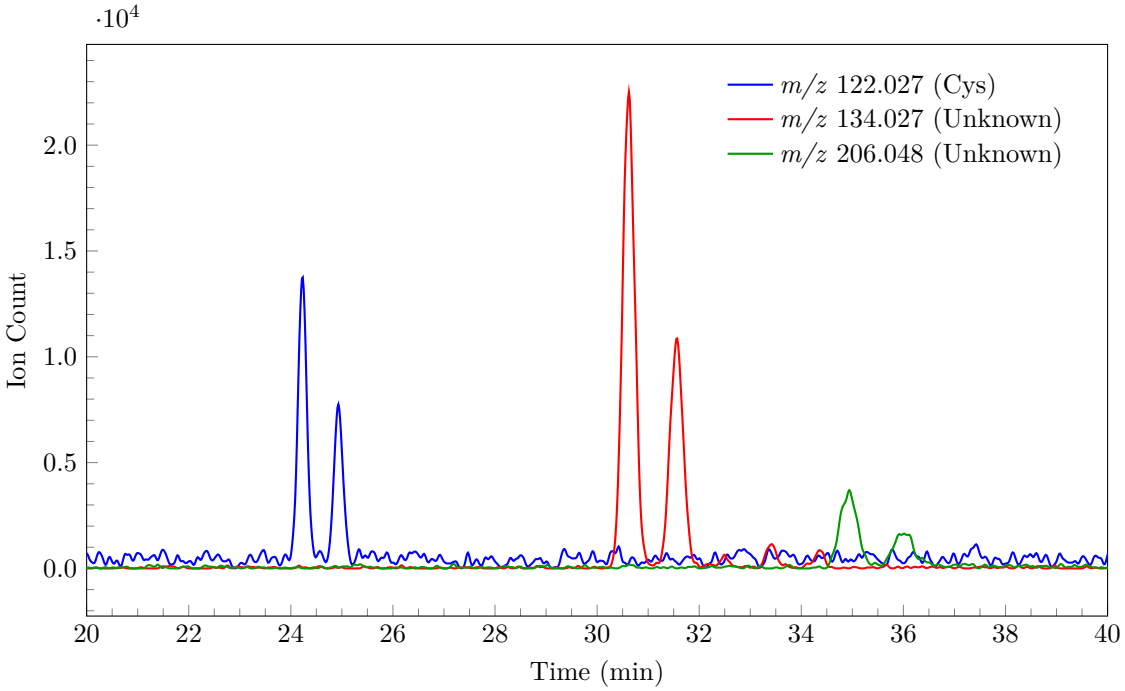


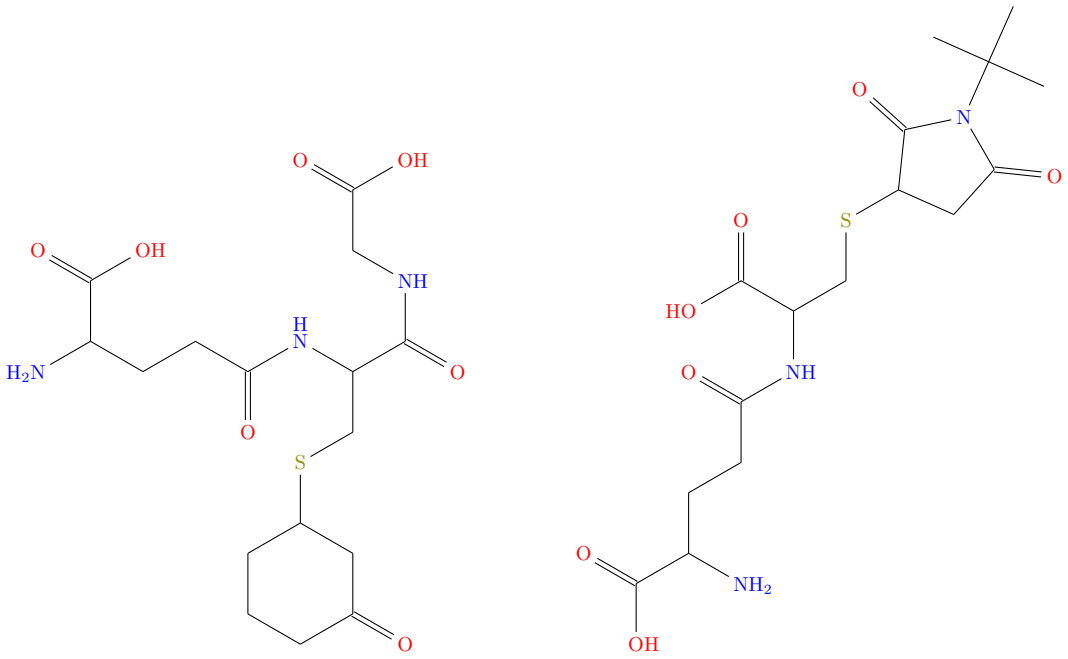
Figure 2.15: EIE overlay of detected underivatized thiols deconjugated from plasma protein, which exhibit the 2:1 dilution pattern.

mass error, formula score, and most likely molecular formulae proposed for both the thiol-NTBM adduct and its corresponding native thiol. It is interesting to note that isomers are possibly detected for unknown thiol-NTBM adducts, *m/z* 307.132 and *m/z* 367.153, in which two sets of 1:1 peaks are observed in both EIEs (Figure 2.14). Formulae for the thiol-NTBM adducts were only indicated in the table if the resulting expected formula of the native thiol was found to be plausible. Included in the list of features are the five known compounds corresponding to the NTBM adducts of Cys, Hcy, GSH, CysGly, and γ-GluCys, while the remaining six are believed to be NTBM adducts of unknown protein-bound thiols. When filtering for peak signals exhibiting only the 2:1 dilution pattern, which correspond to the sample containing underivatized thiols cleaved from protein by reduction with DTT, 3 features were detected, in which Cys (reduced) is the only known feature (Table 2.13 and Figure 2.15). In an attempt to identify the derivatized form of the two unknown reduced thiols, the mass of the NTBM adduct ($C_8H_{11}NO_2$ = 153.079) plus the measured *m/z* of the native thiol was extracted, however, NTBM adducts were not detected. Since the two underivatized (free) unknown thiols were found to have later migration times compared to reduced free Cys, it is possible that the 45 min CE runtime was not sufficiently long enough to detect the migration of the corresponding thiol-NTBM adducts for the unknown thiols given that Cys-NTBM already migrates at 22–27 min under the operating conditions specified earlier in a 120 cm capillary. While the search of unknown 'expected' native thiol molecular formulae in both metabolite and synthetic compound databases, including Human Metabolome Database (HMDB) (Wishart et al., 2018) and Metlin (Guijas et al., 2018), have largely resulted in either no hits or unlikely structures containing alkyl sulfides, dialkyl disulfides, sulfonamides, sulfones, sulfoxides, sulfur heterocycles, or sulfonic acid, the molecular formula of *m/z* 206.048 has so far been found to correspond to a drug or synthetic compound

2-(diacetylamino)-3-sulfanylpropanoic acid (or its stereoisomer, N,N-diacetylcysteine). Although it is unlikely for this compound to be naturally found in the protein of human plasma, the presence of a thiol product of drug metabolism cannot be ruled out unless structural identification by MS/MS has been performed with comparison to authentic standards. Due to insufficient information to prove or disprove the proposed molecular formulae for this proof-of-concept study, the detected compounds in the NTBM-derivatized protein sample currently remain unknown and warrants further investigation. Nonetheless, the signal-encoding strategy employed in this study has been demonstrated to be successful in the discovery of novel features by integrating the analyses of negative controls and the method blank within the same run.

Investigating MS/MS Fragmentation Patterns of Known Thiol-NTBM Adducts

While this work presents a proof-of-concept high-throughput CE-Q-TOF-MS strategy for the discovery of novel protein-bound thiols, future work on this study would involve structural elucidation of the unknown compounds by MS/MS. As a first step, the ability of the MSC software to properly identify fragments of thiols-NTBM was validated through MS/MS experiments on only the five known thiol-NTBM adducts. Although MSC was capable of generating the correct formulae for the various thiol-NTBM adducts, it was not able to generate or produce the correct structures based on the structures currently available in the PubChem (Kim et al., 2016) and ChemSpider (Pence and Williams, 2010) databases. Besides, it is not expected that such public databases would contain comprehensive listing of thiols that are derivatized by synthetic compounds. For example, in the case of γ-GluCys-NTBM, the structure was proposed by MSC to be γ-Glutamyl-S-(3-oxocyclohexyl)cysteinylglycine (ChemSpider #23241467) (Figure 2.16) as it was top-ranked among the generated list



(a) γ-Glutamyl-S-(3-oxocyclohexyl)cysteinyl-glycine (b) γ-GluCys-NTBM adduct

Figure 2.16: (a) Proposed candidate structure, γ-glutamyl-S-(3-oxocyclohexyl)cysteinylglycine (ChemSpider #23241467), by MSC based on MS/MS data of (b) γ-GluCys-NTBM adduct.

of 28 structural isomer candidates with a compatibility score of 79.89%. However, when the known structure of the thiol-NTBM adduct was inputted in the software, the compatibility score was calculated by MSC to be 91.23%. Therefore, when provided with known structures, MSC was able to propose plausible fragment structures for the majority of major ions observed in the MS/MS spectra with compatibility scores greater than 90%, such that on average, 85% of the overall fragment ion intensity for each spectra were plausibly explained.

Given the limitations of current software algorithms and databases to properly identify unknown derivatized structures, it is important to establish or identify known fragmentation patterns of five thiol-NTBM standards (Figure 2.17) that may aid in future identification and structural elucidation of unknown compounds. Thus, the 5 thiol-NTBM standards and their fragmentation patterns were investigated by

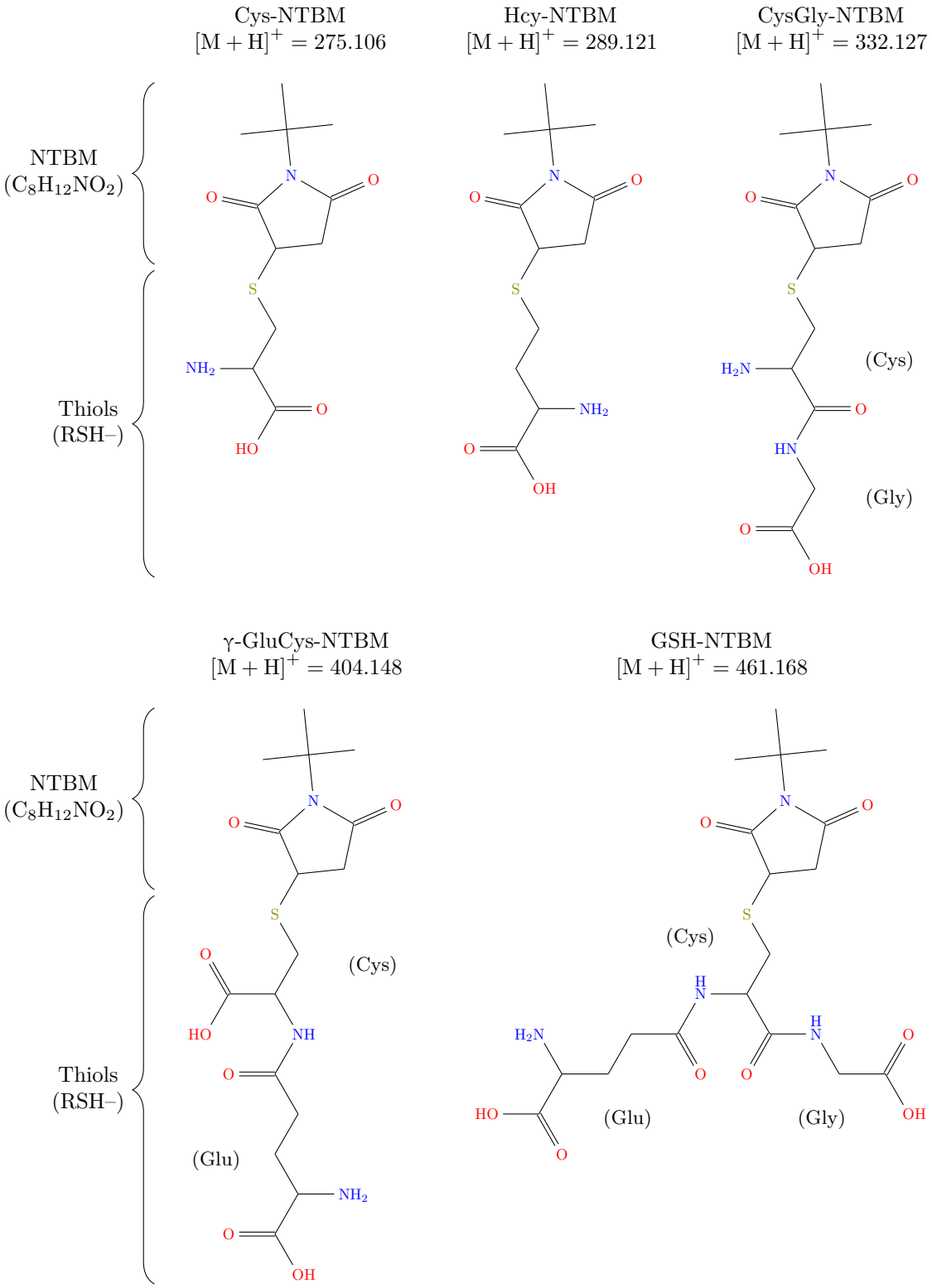


Figure 2.17: Structures of Cys, Hcy, CysGly, γ-GluCys, and GSH derivatized by NTBM.

Table 2.14: Summary of relative ion abundances (%) in tandem mass spectra of five thiol-NTBM standards at CID = 20 and 40 V.

Fragment Ion	m/z	Cys		CysGly		γ-GluCys		GSH		Hcy	
		20 V	40 V	20 V	40 V	20 V	40 V	20 V	40 V	20 V	40 V
$[C_{18}H_{29}N_4O_8S]^+$	461.168							15*			
$[C_{16}H_{26}N_3O_7S]^+$	404.148					10*					
$[C_{16}H_{23}N_2O_7S]^+$	387.122					5					
$[C_{16}H_{24}N_3O_6S]^+$	386.137							20			
$[C_{12}H_{18}N_3O_7S]^+$	348.085					2					
$[C_{13}H_{22}N_3O_5S]^+$	332.127			5*				100			
$[C_{12}H_{15}N_2O_7S]^+$	331.060					2					
$[C_{12}H_{16}N_3O_6S]^+$	330.075								15		
$[C_{12}H_{21}N_2O_4S]^+$	289.121									5*	
$[C_{11}H_{13}N_2O_5S]^+$	285.054					5					
$[C_{11}H_{14}N_3O_4S]^+$	284.070								10		
$[C_9H_{14}N_3O_5S]^+$	276.065			40				60	2		
$[C_{11}H_{19}N_2O_4S]^+$	275.106	5*				50					
$[C_9H_{11}N_2O_5S]^+$	259.038			65				30	10		
$[C_{11}H_{16}NO_4S]^+$	258.080					3					
$[C_{10}H_{11}N_2O_3S]^+$	239.050						15				
$[C_8H_{13}N_2O_4S]^+$	233.059									20	
$[C_{10}H_{17}N_2O_2S]^+$	229.100			5				30			
$[C_7H_{11}N_2O_4S]^+$	219.044	50				100	20				
$[C_7H_8NO_4S]^+$	202.017	100				20	65				
$[C_7H_{11}N_2O_2S]^+$	187.054									20	
$[C_7H_6NO_3S]^+$	184.007	30	2	25	15	2	50		35		
$[C_6H_8NO_3S]^+$	174.022						15				
$[C_6H_9N_2O_2S]^+$	173.038	100		95	15	20	90	15	100		
$[C_6H_6NO_2S]^+$	156.012	30	25	5	15		40		10		
$[C_6H_8NOS]^+$	142.033						5				
$[C_4H_8NO_2S]^+$	134.027									5	
$[C_5H_8NO_3]^+$	130.050			5		5	15	5	10		
$[C_4H_4NO_2S]^+$	129.996	30	70		50		40				
$[C_4H_4NOS]^+$	114.001		5								
$[C_5H_5OS]^+$	113.006		50		40		15		5		
$[C_4H_4NO_2]^+$	98.024	10	100		15		25		5	5	
$[C_3H_6NO_2]^+$	88.040			5							
$[C_3H_6NS]^+$	88.022										5
$[C_3H_3OS]^+$	86.989		45								
$[C_4H_5S]^+$	85.011		30		15		10				
$[C_4H_6NO]^+$	84.044					5	100	5	25		
$[C_4H_2NO]^+$	80.014		10								
$[C_2H_6NS]^+$	76.022		15		15						
$[C_2H_4NS]^+$	74.006		30								
$[C_3H_4NO]^+$	70.029		15								
$[C_2H_3S]^+$	58.995		15								
$[C_4H_9]^+ = [\text{t-butyl}]^+$	57.070	10	30		10		5				
$[C_3H_6N]^+$	56.050						5			100	85
$[C_2H_8N]^+$	46.065					5	10				
$[C_2H_6N]^+$	44.049	30	70	10	90		55	5	45		

* Denotes molecular and precursor ion.

acquiring their tandem mass spectra at 3 different energies (10, 20, and 40 V) for collision-induced dissociation (CID). Since the fragments in the spectra acquired at 10 V were all present in the spectra of 20 V, but with lower intensity in 10 V, the relative ion abundances in 10 V spectra have been omitted from Table 2.14, which summarizes the fragment ions detected along with their relative abundances (%) in the tandem mass spectra for the five thiol-NTBM standards at 20 and 40 V.

In Table 2.14, much overlap in the fragmentation of Cys, CysGly, γ-GluCys, and GSH thiol derivatives is observed. This is not surprising since the Cys residue is present in all four thiols, the Gly peptide is present in CysGly and GSH, and the Glu peptide is present in γ-GluCys and GSH. Out of the 5 thiols studied, Hcy-NTBM has the fewest fragments in both 20 V and 40 V spectra and is the only thiol adduct whose product ions are nearly all unique. Based on the masses lost from the various precursor molecular ions, $[M + H]^+$, it is noted that fragmentation of the thiol-NTBM adducts at low energy (20 V) frequently results in the loss of the tertiary butyl functional group (–$C(CH_3)_3$); losses of small neutral molecules such as ammonia (NH_3), carbon monoxide (CO), and H_2O from side chain or terminal –NH_2 and –COOH groups; and losses of individual intact amino acid residues due to fragmentation along the backbone of amide bonds. These sites of fragmentation at low energy are commonly seen in fragmentations of peptides (Paizs and Suhai, 2005) and are shown in Figure 2.18, along with the corresponding tandem mass spectrum.

Loss and Rearrangement of Tertiary Butyl Moiety

In the 20 V fragmentation spectra of Cys-, CysGly-, and Hcy-NTBM, the largest product ion is typically observed to be associated with the neutral loss of the tertiary butyl group (56.062 Da). Previous studies have shown that fragmentation occurs preferentially at branched carbon atoms because of the stability of the positively-

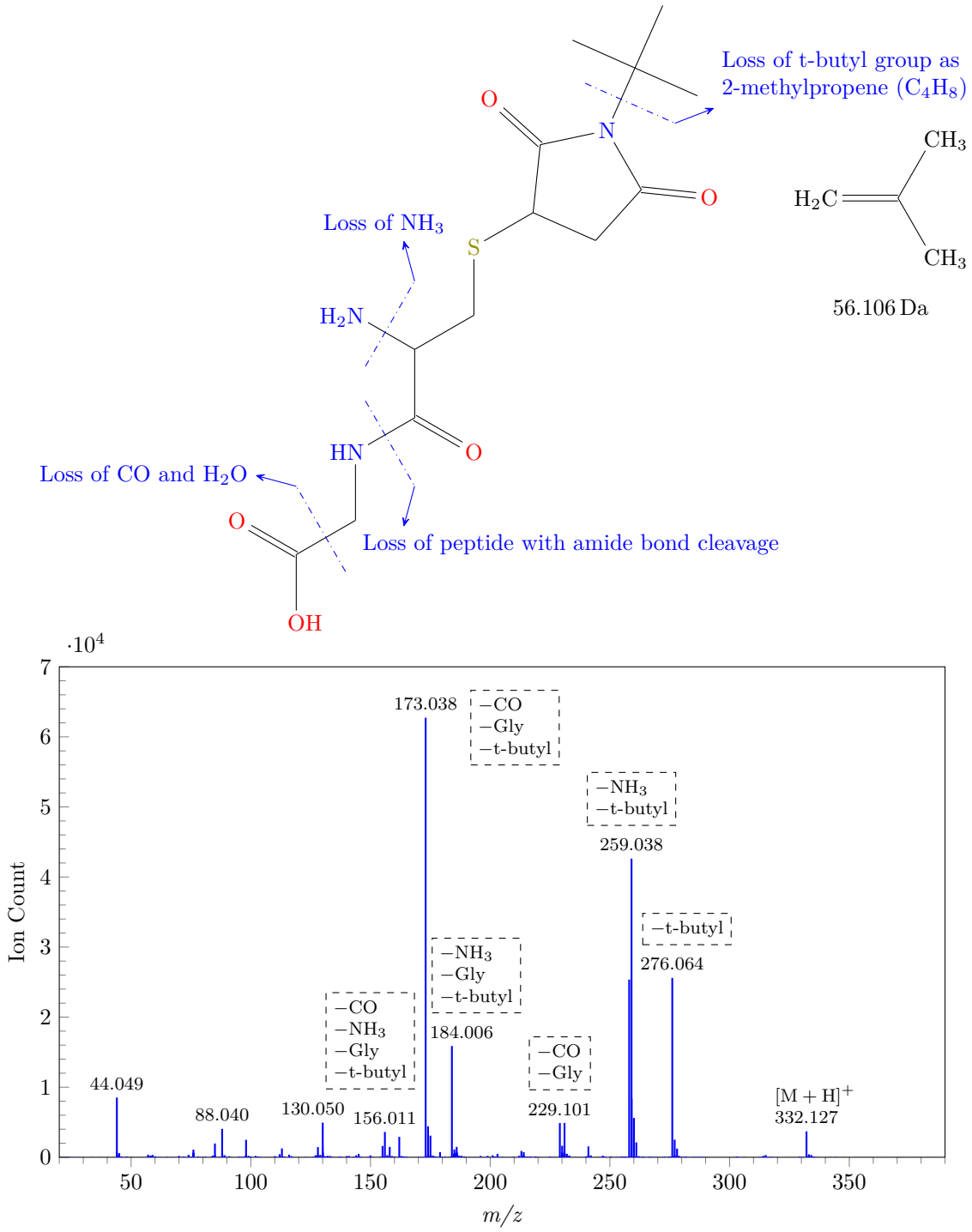
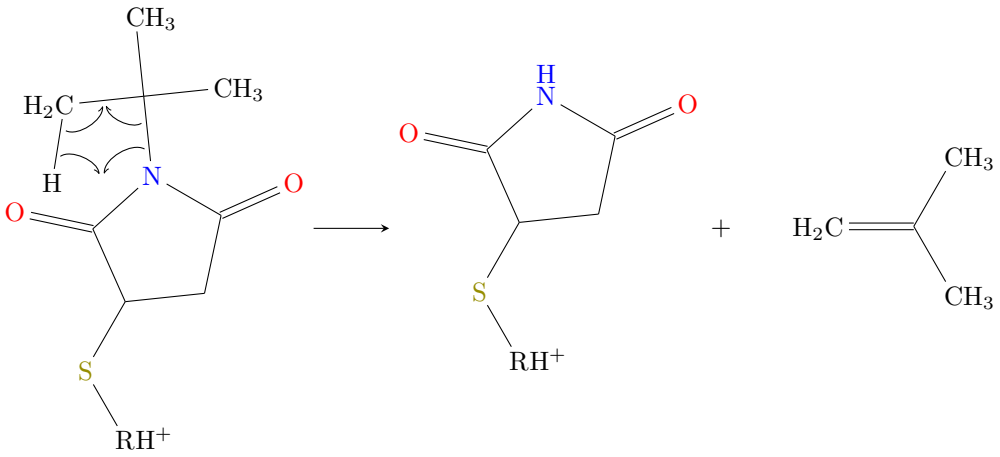


Figure 2.18: Interpretation of MS/MS spectra for CysGly-NTBM adduct. (Top) Common low-energy (20 V) MS/MS fragmentation sites of thiol-NTBM adducts, as indicated on the structure of CysGly-NTBM, as an example. (Bottom) MS/MS spectrum of CysGly-NTBM at 20 V with textboxes indicating functional groups lost from the precursor molecular ion.



Thiol-NTBM adduct $[M + H]^+$ | Thiol-maleimide adduct $[M + H - C_4H_8]^+$ | 2-Methylpropene (56.062 Da)

Figure 2.19: Proposed mechanism for the in-source elimination of C_4H_9 from NTBM adduct as C_4H_8 in order to rationalize the neutral loss of 56.062 Da.

charged carbocation such that tertiary > secondary > primary (Sharma, 2002), which results in a strong peak corresponding to the carbocation. However, in the case of thiol-NTBM adducts, while fragmentation does occur at the tertiary carbon, the ion abundance of the even-electron (EE^+) carbocation peak at *m/z* 57.070 is at best only 20% of its complementary product ion, which is indicative of its lesser stability relative to its counterpart produced in a competing pathway. Furthermore, it is unexpected to see that the mass difference between the precursor and the largest product ion is 56.062 Da, which corresponds to the neutral loss of C_4H_8, since there are two less hydrogen atoms in the formula than the expected neutral loss of the tertiary butyl group as C_4H_{10} ($HC(CH_3)_3$). Because the 56.062 Da loss from the molecular ion, $[M + H]^+$, was found to be common to the 20 V fragmentation of Cys-, CysGly-, and Hcy-NTBM, it is deduced that the C_4H_8 loss must originate from the loss of the tertiary butyl moiety, such that *m/z* 275.106→219.044, *m/z* 332.127→276.065, *m/z* 289.121→233.059 for Cys-, CysGly-, and Hcy-NTBM, respectively, since the loss of C_4H_8 from other moieties would require unique substantial rearrangements from

each thiol-NTBM adduct, which is highly unlikely. Elimination of the tertiary butyl group as C_4H_8 may be explained by the occurrence of a hydrogen rearrangement, whose initiation requires the formation of a radical (McLafferty and Turecek, 1993). Although radical formation in ESI is rare and considered forbidden according to the "even-electron rule", which states that EE^+ ions preferentially decompose to an EE^+ ion and neutral molecule (McLafferty and Turecek, 1993) due to separation of an electron pair being energetically unfavourable, several recent reports have demonstrated exceptions to this rule, in which odd-electron species were shown to be formed from the fragmentation of EE^+ ions produced in electrospray ionization, chemical ionization, or electron ionization (Chen et al., 2008). While these studies report the elimination of radical losses, it is proposed here, however, that radical formation is transient and initiates a concerted H rearrangement, in which the net outcome is H being transferred to a different atom, still resulting in the formation of an EE^+ ion and the elimination of a neutral compound. Specifically, the transient radical formed at the tertiary butyl carbon from homolytic cleavage from nitrogen forms a new bond with an adjacent carbon from a methyl group, which occurs with the concomitant homolytic cleavage of a C–H bond from $–CH_3$. As H becomes released in the process, it forms a new bond with the radical nitrogen to form N–H, due to proximity arising from steric factors. As a result, the tertiary butyl moiety (C_4H_9) is converted to 2-methylpropene (C_4H_8), which is eliminated as a neutral compound. The proposed mechanism of the elimination of tertiary butyl group as C_4H_8 is shown in Figure 2.19. Since the tertiary butyl group was observed to be lost from all five known thiol-NTBM adducts, a MS/MS strategy for confirming unknown thiols in future work could involve performing a neutral loss scan for each unknown derivatized thiol. By specifying the neutral loss to be 56.062 Da, only those compounds that experience a fragment loss corresponding to 56.062 Da would be detected. Structural

elucidation for identification of unknown thiols could then be subsequently performed via MS/MS product ion scans.

2.3.4 Conclusions

For the first time, an original high-throughput and comprehensive CE-MS approach for the screening and discovery of unknown biological thiols, specifically protein-bound thiols in plasma, was presented. MSI-CE-MS in conjunction with temporal signal pattern recognition allowed for the streamlined screening of thiols via the simultaneous analysis of underivatized and NTBM-derivatized thiols along with negative controls and a blank sample to ensure high data fidelity. Through this method, 11 thiols were found, of which 6 are unknown. The robustness and ease of operation make the method amenable and easily adaptable to other sample types for screening of thiols or other compound classes.

Furthermore, MS/MS experiments on five known thiol-NTBM derivatives were performed for the first time, in order to establish trends in fragmentation pathways that may be helpful to the structural elucidation of unknowns in future studies. From the MS/MS experiments, it was observed that low-energy CID at 20 V resulted in charge-driven fragmentations dominated by large and intact peptide fragments with the loss of side chains as neutral species, thereby providing sequence-related information, while higher energy CID at 40 V led to more complex spectra owing to multiple charge-remote cleavages within peptide moieties and the maleimide ring, which provided additional information on the structural features within the intact peptide fragments produced by low-energy CID.

Since this work has revealed the presence of two late-migrating unknown free thiols following reduction by DTT, for which the migration of their corresponding

thiol-NTBM adducts are believed to occur beyond the CE runtime used in the present study, future studies may consider performing single injection CE runs for NTBM-derivatized protein-bound thiols on shorter capillary lengths. This will allow for the detection of slow-migrating unknown compounds, without having to further extend the 45 min runtime that is currently applied for long capillaries. Additionally, as the majority of novel thiol-NTBM adducts detected by MSI-CE-MS appear to be low in abundance, on-line sample preconcentration studies by CE can be applied in future work for the enhanced detection, MS/MS characterization, and confirmation of these novel thiol compounds.

Although it was seen here that the combination of both low- and high-energy CID mass spectra contain a wealth of structural information in the characterization of the known thiols, in which identified fragmentation patterns may be applied towards the structural elucidation of novel unknown thiols, it is important to acknowledge that metabolite identification by MS is a significant analytical challenge. Even though exact mass measurements are of great utility in the automated assignment of product ion spectra, results of in silico structural identification should never be taken at face value without careful consideration, since the number of possible structures for a given formula could range from tens to hundreds, and the correlation scores of candidate structures are not sufficiently discriminatory to provide an unambiguous identity of a true unknown. Integrated analytical approaches, such as the combination of MS with NMR, is a powerful strategy that could aid in the elucidation of specific substructures of unknown thiols in the future. However, the excellent sensitivity and selectivity offered by MS make MS/MS a powerful technique of first choice for structure elucidation, especially in the discovery of novel metabolites and in drug development.

Chapter 3

Differential Metabolomics of Obesity and its Subtypes

The work presented in this chapter is a two-part collaborative targeted metabolomic investigation to understand metabolic differences between lean healthy, healthy obese, and unhealthy obese subjects, which are responsible for differences in cardiometabolic risk between individuals. Understanding these differences may lead to improved strategies for the prevention and treatment of obesity-related diseases. Results of targeted amino acid analysis by CE-MS on serum and plasma samples were an important contribution in both studies described in this chapter which explored: 1) metabolic differences in amino acid homeostasis of serum and adipose tissue between subject groups, and 2) differences in postprandial metabolism of plasma amino and fatty acids between subject groups following consumption of a standardized high-caloric meal, respectively. The findings from each section have been separately published in two journal articles:

Section 3.1: Badoud, F.[a], Lam, K. P.[b], DiBattista, A.[b], Perreault, M.[a], Zulyniak, M. A.[a], Cattrysse, B.[a], Stephenson, S.[a], Britz-McKibbin, P.[b], and Mutch, D. M.[a] (2014). Serum and Adipose Tissue Amino Acid Homeostasis in the Metabolically Healthy Obese. *Journal of Proteome Research,* 13(7):3455–3466. Reprinted with permission: © 2014 American Chemical Society.

Section 3.2: Badoud, F.[a,*], Lam, K. P.[b,*], Perreault, M.[a], Zulyniak, M. A.[a], Britz-McKibbin, P.[b], and Mutch, D. M.[a] (2015a). Metabolomics Reveals Metabolically Healthy and Unhealthy Obese Individuals Differ in their Response to a Caloric Challenge. *PLOS ONE,* 10(8):e0134613. Reprinted with permission under the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/legalcode).

3.1 Serum and Adipose Tissue Amino Acid Homeostasis in the Metabolically Healthy Obese

Author Contributions

FB, KPL, PBM, and DMM conceived and designed the study. Clinical trial, blood collection, and classification of subjects based on BMI, fat mass %, and metabolic status measurements were conducted by MP, MAZ, FB, and DMM. SS collected adipose tissue biopsies. Experiments on serum samples were performed by FB (GC-MS), KPL (CE-MS), and FB and BC (Microarray analysis). Calibration data was generated by FB and KPL, with assistance from AD. Data analysis was performed

[a]Department of Human Health and Nutritional Sciences, University of Guelph, Guelph, ON, Canada
[b]Department of Chemistry and Chemical Biology, McMaster University, Hamilton, ON, Canada
*Equal contributing first author.

by FB and KPL. Data interpretation was performed by FB and KPL. Reagents and materials were contributed by FB, KPL, PBM, and DMM. FB, KPL, and DMM wrote the manuscript. FB, KPL, PBM, and DMM edited the manuscript. All authors approved the final version of the manuscript.

Abstract

A subgroup of obese individuals, referred to as metabolically healthy obese (MHO), have preserved insulin sensitivity and a normal lipid profile despite being obese. The molecular basis for this improved cardiometabolic profile remains unclear. Our objective was to integrate metabolite and gene expression profiling to elucidate the molecular distinctions between MHO and metabolically unhealthy obese (MUO) phenotypes. A subset of individuals were selected from the Diabetes Risk Assessment study and classified into three groups using anthropometric and clinical measurements: lean healthy (LH), MHO, and MUO. Serum metabolites were profiled using gas chromatography coupled to mass spectrometry. Multivariate data analysis uncovered metabolites that differed between groups, and these were subsequently validated by capillary electrophoresis coupled to mass spectrometry. Subcutaneous adipose tissue (SAT) gene expression profiling using microarrays was performed in parallel. Amino acids were the most relevant class of metabolites distinguishing MHO from MUO individuals. Serum levels of glutamic acid, valine, and isoleucine were positively associated (i.e., LH < MHO < MUO) with homeostatic model assessment of insulin resistance (HOMA-IR) and glycated hemoglobin (HbA1c) values, while leucine was only correlated with HOMA-IR. The glutamine-to-glutamic acid ratio and glycine were inversely correlated (i.e., LH > MHO > MUO) with HbA1c values. Concomitantly, SAT gene expression profiling revealed that genes related to branched-chain amino

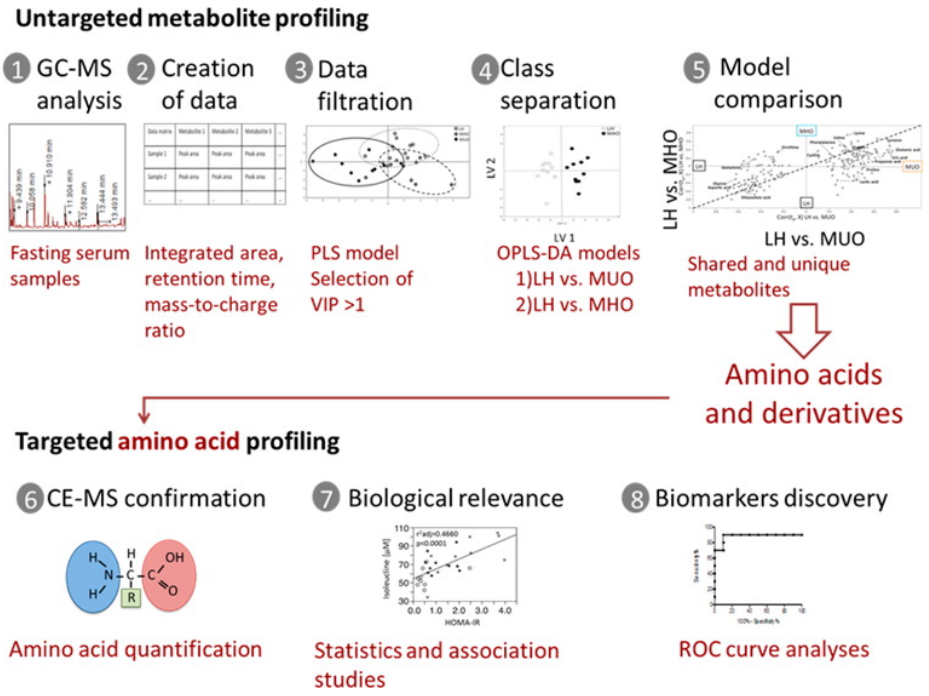


Figure 3.1: Workflow for untargeted and targeted metabolite profiling.

acid catabolism and the tricarboxylic acid cycle were less down-regulated in MHO individuals compared to MUO individuals. Together, this integrated analysis revealed that MHO individuals have an intermediate amino acid homeostasis compared to LH and MUO individuals.

3.1.1 Introduction

Obesity is widely recognized as a primary risk factor for the development of type 2 diabetes (T2D), hypertension, and cardiovascular disease (Haslam and James, 2005; Pedersen, 2013). However, the recent finding that a subgroup of obese individuals, referred to as MHO, may be protected from the downstream cardiometabolic complications typically associated with a high BMI has generated considerable interest (Primeau et al., 2011). MHO individuals are characterized by reduced cardiometabolic risk, as reflected by preserved insulin sensitivity, a reduced inflammatory status, and

a normal circulating lipid profile (Primeau et al., 2011). Importantly, recent studies estimate that the prevalence of this phenotype may be as high as 30% of the obese population (Pataky et al., 2010; Velho et al., 2010). Weight loss interventions are promoted to reduce cardiometabolic risk; however, reports suggest MHO individuals may respond to these interventions differently than their MUO counterparts (Karelis et al., 2008; Shin et al., 2006). As such, it is crucial to elucidate the metabolic and molecular basis underlying the MHO phenotype as this may lead to more tailored health care strategies.

Much of our current knowledge regarding the MHO phenotype stems from retrospective analyses of large cohort studies. While these analyses have provided substantial population-based evidence highlighting the existence and prevalence of MHO (Roberson et al., 2014; Romero et al., 2012; van Vliet-Ostaptchouk et al., 2014), the molecular and metabolic distinctions between MHO and MUO remain largely unknown. Evidence demonstrates that metabolite profiling (i.e., metabolomics) is particularly well suited to identify subtle differences in metabolism between distinct groups of individuals (Zhang et al., 2013). To date, most metabolomic studies have compared lean and obese individuals rather than distinct groups of obese individuals. These past studies have revealed that amino acid metabolism is altered with obesity and T2D (Adams, 2011; Kim et al., 2010; Morris et al., 2012; Newgard et al., 2009; Xie et al., 2012). To our knowledge, only a single study has examined metabolite profiles in obese individuals classified by cardiometabolic parameters (i.e., fasting glucose, blood pressure, blood lipids, etc.) as either metabolically healthy or metabolically unhealthy (Batch et al., 2013). The authors of this study reported that a cluster of metabolites composed of amino acids and acylcarnitines reflected metabolic wellness (Batch et al., 2013). Together, these studies suggest a potential relationship between amino acid metabolism, obesity, and metabolic health.

Although studying blood metabolites can uncover novel markers of potential clinical relevance, it is difficult to determine the originating tissues underlying changes in circulating metabolite concentrations. Little research has examined differences in tissue function between MHO and MUO, but past work suggests that studying adipose tissue will provide important insights to better understand the MHO phenotype (Naukkarinen et al., 2014). Indeed, adipose tissue has been shown to play a central role in amino acid metabolism (Lackey et al., 2013), and recent work reported that the expression of genes related to amino acid metabolism is reduced with obesity (Lackey et al., 2013; Pietiläinen et al., 2008). However, it is currently unknown if adipose tissue gene expression differs between MHO and MUO individuals.

The goal of this work was to conduct an integrative analysis combining serum metabolomics and SAT transcriptomics to elucidate the molecular basis for the differences in metabolic profiles observed between MHO and MUO individuals. Serum metabolites were initially profiled using an untargeted GC-MS approach. Metabolites of interest were then validated using a quantitative method based on CE-MS and subsequently integrated with SAT gene expression profiles. This integrative approach highlighted the relevance of amino acid metabolism for cardiometabolic health and has led to the identification of potential biomarkers of metabolic wellness.

3.1.2 Materials and Methods

Study Design

Serum and SAT samples were obtained from individuals participating in the Diabetes Risk Assessment (DRA) study (Clinicaltrials.gov identification (ID) #NCT01884714) (Perreault et al., 2014b). Persons were screened over the phone and excluded if they met any one of the following criteria: (1) below 35 or above 70 years of age; (2)

diagnosed with an acute or chronic autoimmune inflammatory disease, infectious disease, viral infection, and/or cancer; or (3) regular alcohol consumption exceeding 2 drinks/d (1 drink = 10 g alcohol). All participants signed a consent form, and the research protocol was approved by the University of Guelph Human Research Ethics Board (REB#10AP033).

Subject Classification

Thirty participants were classified into three distinct groups: LH ($n = 10$), MHO ($n = 10$), and MUO ($n = 10$) based on their adiposity (BMI, fat mass %) and metabolic status (blood lipids, glucose, and insulin), as previously described (Perreault et al., 2014b). Adiposity status was determined using the revised BMI cutoffs proposed by Shah and Braverman (2012), where lean was considered $<28\,kg/m^2$ for males and $<24\,kg/m^2$ for females, and obese was considered $\geq 28\,kg/m^2$ for males and $\geq 24\,kg/m^2$ for females. An individual was considered metabolically healthy if three or more of the following criteria were met: high-density lipoprotein cholesterol (HDL-c) > 1.0 mmol/L for males and >1.3 mmol/L for females; triglycerides (TG) < 1.7 mmol/L without use of lipid-lowering drugs; total cholesterol (Total-c) < 5.2 mmol/L; low-density lipoprotein cholesterol (LDL-c) < 2.6 mmol/L; and HOMA-IR < 1.95 without use of antidiabetic drugs. Each group was composed of seven women and three men. LH, MHO, and MUO groups were matched for age, while MHO and MUO groups were also matched for BMI, waist-to-hip ratio, and body fat %. None of the participants were taking antidiabetic medications.

Gas Chromatography Coupled to Mass Spectrometry (GC-MS) Analysis

Chemicals and Reagents. Heptadecanoic acid (≥98%), methoxyamine hydrochloride, pyridine, *N*-methyl-trimethylsilyl-trifluoroacetamide (MSTFA) with 1% trimet-

hylsilyl chloride (TMCS), chloroform (≥99%), and methanol of analytical grade were obtained from Sigma-Aldrich (St. Louis, MO).

GC-MS Procedure. Fasting serum samples were collected from all participants following an overnight fast (∼12 h). Ten microliters of the internal standard heptadecanoic acid at 1 mg/mL in methanol was added to a 100 µL serum aliquot, and then mixed with 300 µL chloroform/methanol (1:3, v/v). The mixture was shaken at 60 rpm for 10 min at room temperature and then placed at −20 °C for 10 min. Samples were then centrifuged at 15 400 *g* for 10 min at 4 °C, and the supernatant was transferred to a clean glass vial before being evaporated to dryness using a centrifugal evaporator (Savant SpeedVac SVC 100H, Thermo Scientific, San Jose, CA). Oximation of the analytes was performed by adding 80 µL of a 15 mg/mL methoxyamine solution in pyridine. Samples were immediately sealed and shaken at 60 rpm for 90 min at 30 °C. Next, 80 µL of MSTFA containing 1% TMCS was added for derivatization, and the mixture was incubated at 70 °C for 60 min. Samples were then centrifuged at 15 400 *g* for 10 min at 4 °C and the supernatant was injected in the GC-MS system after leaving the samples 2 h at room temperature.

A Bruker Daltonics Scion TQ GC-MS/MS (Bruker Daltonics, Fremont, CA) system with an Agilent column HP-1 100% dimethylpolysiloxane (30 m × 0.2 mm I.D., 0.11 µm) was used for untargeted metabolite profiling. Samples were injected using a 1:10 split ratio, and the injector temperature was set at 280 °C. Helium flow was maintained at 1 mL/min. Column temperature was first set at 80 °C for 2 min, then increased to 230 °C at 10 °C/min, then increased to 310 °C at 40 °C/min, and finally maintained at 310 °C for 3 min. The transfer line was set at 280 °C, and the ionization source at 230 °C. MS acquisition was set from 50 to 600 *m/z* in scan mode with a scan time of 200 ms. GC-MS data was processed with MS Workstation 8 software (Bruker

Daltonics).

GC-MS Data Processing. GC-MS total ion chromatogram (TIC) data was converted into xml file format using mzXML conversion files (Bruker Daltonics), and subsequently processed with XCMS software (version 2.9.1 in R, http://www.r-project.org) (Smith et al., 2006). A list of 608 variables (i.e., consisting of ion, retention time, and an integrated area) was obtained, and data was normalized by dividing the integrated peak areas to the sum of all of the peak areas from the TIC. Peak identification was performed using NIST 08 mass spectral libraries (match factor $> 60\%$, probability score $> 20\%$). PLS analysis was first performed using SIMCA (V13 Demo Umetrics, Umeå, Sweden) to evaluate the intrinsic variation of the variables within the three groups of individuals. The initial data set was subsequently filtered by selecting only features with a VIP value greater than or equal to 1. After filtering, 211 variables were investigated further. To increase the visualization and interpretation of the data set, two independent OPLS-DA models were built in comparison to the common LH group. This approach was deemed more appropriate than building a single model on three classes that will create a reference point, to which all plots and graphs would be anchored and thus reducing the interpretation of the model (Wiklund et al., 2008). The data were pareto-scaled and mean centered, and the predictive ability of the models was verified using leave-one-out cross-validation. Two OPLS-DA models with two-components, one predictive (first latent variable; LV1) and one orthogonal (orthogonal component; LV2), were obtained for the models LH versus MUO and LH versus MHO (Figure 3.2A and B, respectively). Since both models used the LH group as a reference, a shared and unique structures (SUS)-plot was built from the correlation vector $\mathrm{Corr}(t_p,X)$ from the predictive component (LV1) of each model (Figure 3.2C), as described elsewhere (Boccard et al., 2011).

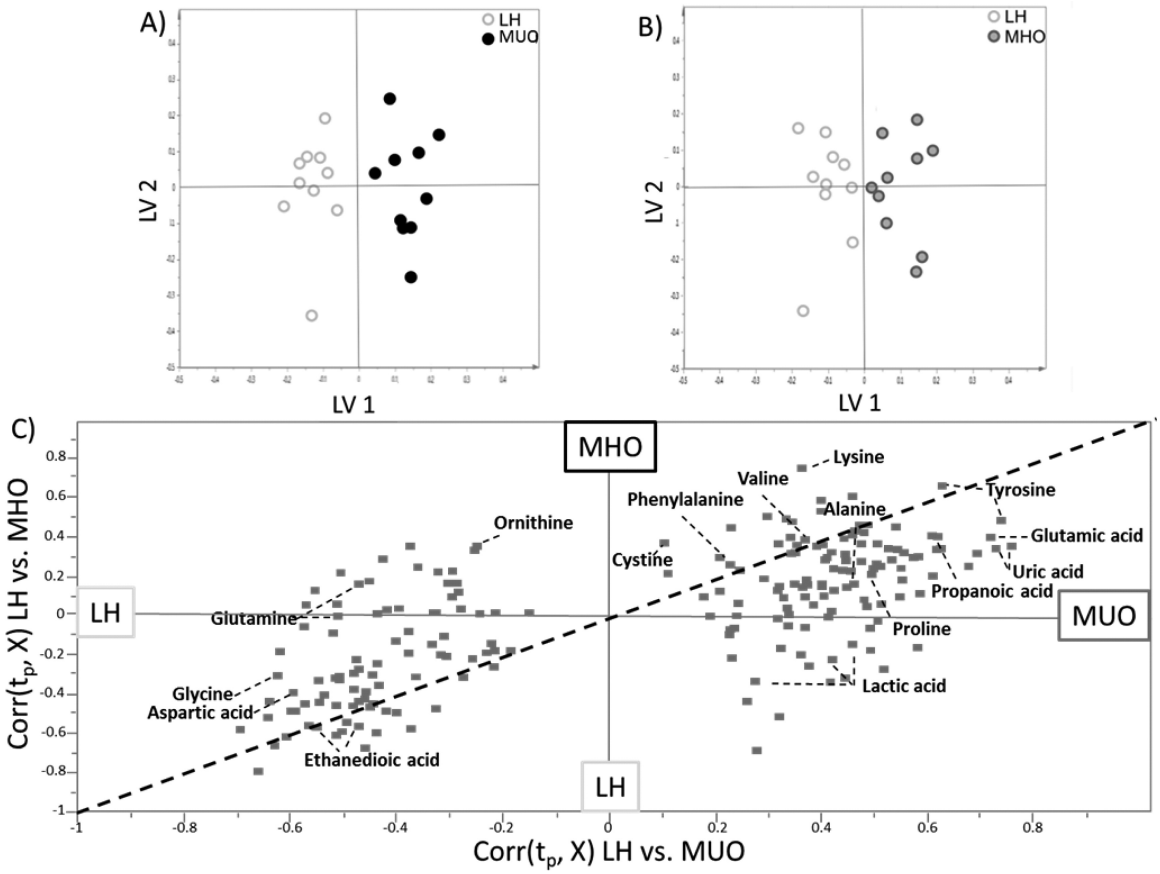


Figure 3.2: Multivariate analyses of metabolite data from GC-MS. (A) OPLS-DA scatter plot for the LH versus MUO phenotypes (R^2X= 0.46, R^2Y= 0.89, Q^2cum= 0.67; CV-ANOVA $p = 0.001$). (B) OPLS-DA scatter plot for the LH versus MHO phenotypes (R^2X= 0.41, R^2Y= 0.786, Q^2cum= 0.47; CV-ANOVA $p = 0.038$). LV1 and LV2, represent the predictive and orthogonal components, respectively. (C) Shared and Unique Structure (SUS) plot representation combining the correlation vector ($Corr(t_p,X)$) from the predictive component (LV1) of both OPLS-DA scatter plots.

Capillary Electrophoresis Coupled to Mass Spectrometry (CE-MS) Analysis

Chemicals and Reagents. Ammonium acetate, acetic acid, formic acid, 3-chloro-l-tyrosine, and other metabolite reference standards were obtained from Sigma-Aldrich (St. Louis, MO). HPLC-grade acetonitrile and methanol were obtained from Honeywell (Muskegon, MI) and Caledon (Georgetown, ON, Canada), respectively. Buffers and stock solutions were prepared in deionized water from a Thermo Scientific Barnstead EasyPure II LF ultrapure water system (Vernon Hills, IL). All metabolite stocks were prepared in water and stored at 4 °C.

CE-MS Procedure. Serum (50 µL) was diluted 4-fold with pH 5.0 ammonium acetate buffer to a final concentration of 200 mM ammonium acetate and 25 µM of internal standard (IS), 3-chloro-L-tyrosine. Deproteinization was performed by ultrafiltration using a 3 kDa MWCO Nanosep centrifugal device (Pall Life Sciences, Washington, NY) at 13 000 *g* for 20 min. A QC sample was obtained by pooling 5 µL of each of the 30 serum samples collected, which would contain all of the analytes to be profiled. To monitor variation in the instrumental performance and assess instrumental precision throughout the duration of the analysis, the reproducibility of the QC sample was assessed by incorporating the QC sample as the seventh segment in the MSI-CE-MS method in all of the sample runs, as previously described by Kuehnbaum et al. (2013). Targeted metabolite analysis was performed using an Agilent G7100A CE system (Agilent Technologies Inc., Mississauga, ON, Canada) interfaced with a coaxial sheath liquid Jetstream electrospray ionization source with heated nitrogen gas coupled to an Agilent 6230 TOF-MS, as detailed elsewhere (Kuehnbaum et al., 2013). An external standard calibration with seven levels of concentration (200, 100, 50, 25, 5, 2, and 0.5 µM) containing the IS was obtained by serial dilution of reference

standards in 200 mM ammonium acetate buffer at pH 5. The calibration curve was performed in triplicate to quantify the targeted metabolites of interest. This method allowed for quantitative measurement of the concentration of 26 amino acids and derivatives (Table 3.1).

Statistical Analysis of Metabolite Data Sets

Statistical analyses were performed using GraphPad Prism 5 software (La Jolla, CA). Nonparametric ANOVA Kruskal–Wallis tests were used to determine if the 26 amino acids and derivatives were statistically different between the three groups ($p < 0.05$). When significance was observed, a posthoc nonparametric Mann–Whitney test was used for pairwise analyses ($p < 0.05$). Receiver operating characteristic (ROC) curves were generated using GraphPad Prism 5 software and obtained by plotting the true positive rate (sensitivity) and false positive rate (1−specificity) at different thresholds. Separate regressions were fitted for each combination of trait (BMI, fat mass %, HbA1c and HOMA-IR) and amino acids and derivatives using JMP Genomics v5.1 (SAS Institute, Cary, NC).

Subcutaneous Adipose Tissue Gene Expression Analysis

RNA Extraction. Adipose tissue samples were obtained from the periumbilical region under local anesthesia after an overnight fast (∼12 h). Samples (50–100 mg) were homogenized in 1 mL of TRIzol reagent (Life Technologies, Carlsbad, CA) using a tissue homogenizer. Following tissue lysis and addition of chloroform, samples were centrifuged at 12 000 *g* for 15 min at 4 °C. The upper aqueous phase containing RNA was purified using the Qiagen RNeasy Mini Kit (Qiagen, Mississauga, ON, Canada), according to the manufacturer's instructions. Extracted RNA was quantified with a Nanodrop 8000 instrument (Thermo Scientific, Wilmington, DE, USA), and RNA

Table 3.1: Mean circulating serum metabolite concentrations analyzed by CE-MS*.

Metabolites	Mean ± SEM			ANOVA (p-value)	Post-Hoc Mann–Whitney Groups Comparison		
	LH (µM)	MHO (µM)	MUO (µM)		LH vs. MUO (p-value)	LH vs. MHO (p-value)	MHO vs. MUO (p-value)
Glycine	221 ± 16	208 ± 17	169 ± 10	0.048	0.018	0.352	0.123
Alanine	336 ± 22	423 ± 23	399 ± 32	0.068			
Valine	206 ± 12	266 ± 10	278 ± 26	0.014	0.043	0.002	1.000
Isoleucine	55.9 ± 3.3	70.5 ± 3.9	78.4 ± 7.1	0.006	0.005	0.015	0.143
Leucine	107 ± 6	134 ± 8	138 ± 1	0.042	0.035	0.029	0.853
Ornithine	53.9 ± 2.8	71.5 ± 5.9	57.1 ± 5.6	0.107			
Arginine	111 ± 7	131 ± 6	119 ± 7	0.146			
Citrulline	32.6 ± 2.8	37.1 ± 2.5	34.2 ± 1.9	0.336			
Aspartic acid	12.8 ± 0.4	13.4 ± 0.4	14.40 ± 0.57	0.155			
Glutamic acid	30.1 ± 3.2	43.5 ± 4.4	57.2 ± 6.5	0.006	0.004	0.018	0.165
Glutamine	640 ± 25	668 ± 14	604 ± 22	0.159			
Glutamine-to-glutamic acid ratio	22.7 ± 1.7	16.7 ± 1.6	12.3 ± 1.8	0.005	0.004	0.029	0.075
Phenylalanine	51.5 ± 2.0	67.5 ± 4.2	65.1 ± 4.1	0.012	0.018	0.005	0.739
Tyrosine	58.2 ± 3.4	78.3 ± 4.4	77.9 ± 5.7	0.003	0.007	0.001	0.912
Cystine	62.6 ± 2.9	87.2 ± 8.8	74.6 ± 5.6	0.033	0.075	0.015	0.352
Methionine	22.0 ± 0.7	26.2 ± 1.0	24.6 ± 1.4	0.013	0.052	0.004	0.315
Histidine	92.2 ± 4.0	102.0 ± 4.3	94.4 ± 4.8	0.303			
Tryptophan	57.2 ± 3.1	64.6 ± 3.3	62.1 ± 3.7	0.289			
Lysine	164 ± 12	213 ± 12	174 ± 10	0.039	0.63	0.023	0.043
Hypoxanthine	7.13 ± 0.67	6.75 ± 0.57	7.39 ± 0.71	0.849			
Serine	160.0 ± 6.8	163 ± 12	138.600 ± 6.202	0.089			
Threonine	121.0 ± 8.7	139.0 ± 9.4	120.0 ± 8.9	0.436			
Proline	142 ± 10	182.0 ± 9.6	183 ± 18	0.042	0.105	0.009	0.796
Hydroxyproline	15.8000 ± 0.2405	18.9 ± 0.6	15.9 ± 0.9	0.002	0.529	0.001	0.009
Carnitine	37.2 ± 2.5	61.0 ± 5.6	49.6 ± 5.0	0.005	0.089	0.001	0.165
Acetylcarnitine	26.5 ± 2.2	27.1 ± 3.2	25.5 ± 3.6	0.946			
Asparagine	77.6 ± 3.8	83.4 ± 2.5	79.0 ± 3.2	0.443			
Asparagine-to-aspartic acid ratio	6.09 ± 0.34	6.33 ± 0.35	5.57 ± 0.32	0.284			

* Data represented as mean ± SEM. LH, lean healthy; MHO, metabolically healthy obese; MUO, metabolically unhealthy obese. A nonparametric one-way ANOVA followed by a post-hoc Mann-Whitney test was used to determine the significance ($p < 0.05$).

quality was verified using the Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA). Only samples having a RNA integrity number greater than 8 were used for microarray analyses.

Microarray Analysis and Data Treatment. Total RNA (100 ng) was used to synthesize cDNA and then cRNA, according to standard Affymetrix protocols. Second cycle cDNA was synthesized, fragmented, biotin labeled, and hybridized to Affymetrix Human Gene 2.1 ST array strips (Affymetrix Inc., Fremont, CA) according to the manufacturer's guidelines. Strips were washed, stained, and scanned on the GeneAtlas platform. Global gene expression was performed on a subset of SAT samples from the same individuals used for serum metabolomic analyses. Specifically, we analyzed SAT from seven LH individuals, eight MHO individuals, and eight MUO individuals; with each group having two men. Data analysis was performed using the Expression Console and Transcriptome Analysis Console from Affymetrix (Affymetrix Inc., Fremont, CA). The robust multiarray average method was used for normalization of the microarray data. An ANOVA, which was adjusted for multiple testing using a false-discovery rate, was performed to identify differentially expressed genes. The LH group was used as the reference for pairwise comparisons between LH versus MHO and LH versus MUO. After correction for multiple testing, 2940 and 1654 genes were significantly and differentially expressed ($p < 0.05$) in the LH versus MUO and LH versus MHO models, respectively. Biological pathway analysis was performed using FunNet software (http://www.funnet.info) to identify pathways that were differentially regulated from within the lists of differentially expressed genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) biological pathways that are over-represented in gene expression data lists were identified by FunNet using Fisher's exact test.

3.1.3 Results

Untargeted Metabolite Profiling Reveals Amino Acids as Metabolites of Interest

We first performed a comprehensive and untargeted metabolite analysis of serum collected from LH, MHO, and MUO individuals using GC-MS. Multivariate data analyses and discriminant modeling were then used to highlight the most pertinent metabolites distinguishing the three groups of individuals. Using an OPLS-DA, we obtained a good fit for the separation of LH versus MUO (R^2X= 0.46, R^2Y= 0.89, Q^2cum= 0.67, cross-validated residuals analysis of variance (CV-ANOVA) $p = 0.001$) and LH versus MHO (R^2X= 0.41, R^2Y= 0.78, Q^2cum= 0.47, CV-ANOVA $p = 0.038$) (Figure 3.2A and B). Relevant metabolites distinguishing the three subject groups were subsequently identified with a SUS-plot analysis (Figure 3.2C). This included metabolites shared by both the MHO and MUO groups in comparison to the LH group, that is, metabolites located near the diagonal (e.g., tyrosine, glutamic acid, uric acid, glycine, aspartic acid, and proline), as well as metabolites unique to either the MHO or MUO groups, that is, metabolites located away from the diagonal (e.g., glutamine, lysine, and ornithine). As the majority of the shared and unique metabolites corresponded to amino acids and derivatives, we next performed a targeted analysis to quantify these specific metabolites of interest.

Quantitative Analysis of Amino Acids and Derivatives

A quantitative and targeted approach using CE-MS was implemented in order to confirm the metabolites of interest highlighted by the SUS-plot analysis (Figure 3.2C). Importantly, the CE-MS platform generated reliable and useful data, as indicated by an analytical variability of <12% in QC samples. Moreover, this targeted analysis

also included other essential and nonessential amino acids, as well as carnitine and acetylcarnitine (Table 3.1) in addition to the metabolites indicated in the SUS-plot (Figure 3.2C).

Distinct serum amino acid profiles were detected in the three groups of subjects (Table 3.1). MHO individuals had intermediate glycine levels that were not significantly different compared to LH and MUO individuals, while glycine levels were significantly reduced in MUO compared to LH individuals (−23.9%). MHO individuals had higher levels of valine, isoleucine, and leucine compared to LH individuals, with increases of 29.3%, 25.9%, and 25.4% observed for these three branched-chain amino acids (BCAAs), respectively. Although not statistically different from BCAA levels in MHO individuals, MUO individuals showed even higher levels of valine, isoleucine, and leucine compared to LH individuals, with increases of 35.0%, 40.1%, and 29.1% observed, respectively. The aromatic amino acids tyrosine and phenylalanine were also significantly elevated in MHO and MUO individuals compared to the LH individuals. Specifically, tyrosine and phenylalanine levels were increased by 31.0% and 34.3% in MHO individuals, respectively, and by 26.4% and 33.7% in MUO individuals, respectively, in comparison to LH individuals.

Serum glutamic acid levels showed an increasing trend, with LH individuals having the lowest level and MUO individuals having the highest level. Specifically, MHO and MUO individuals showed a 44.5% and 89.8% increase in glutamic acid levels compared to LH individuals; however, the concentrations of glutamic acid did not differ significantly between the two obese groups. Glutamine did not differ between the three groups. Because glutamine is the precursor of glutamic acid, we also calculated the glutamine-to-glutamic acid ratio for all individuals. LH individuals had the highest glutamine-to-glutamic acid ratio, while MUO had the lowest ratio. Interestingly, the glutamine-to-glutamic acid ratio was significantly decreased by 26.5%

and 46.0% in MHO and MUO individuals compared to LH individuals, respectively. Furthermore, a trend was observed for a lower ratio in MUO compared to MHO individuals ($p = 0.075$).

Higher concentrations of sulfur-containing amino acids cystine and methionine were found in MHO (39.4% and 19.2%, respectively) compared to LH individuals, while MUO individuals also tended to have higher levels of cystine (19.3%) and methionine (11.9%) compared to LH individuals. However, no significant difference was observed between MHO and MUO individuals. Proline and carnitine showed a similar profile to sulfur-containing amino acids, with significantly increased levels detected in MHO individuals (28.1% and 64.2%, respectively) compared to LH individuals; however, the increases seen in MUO individuals (28.6% and 33.3%, respectively) were not statistically different from LH individuals. Lysine and hydroxyproline showed a distinct profile, with significantly lower levels seen in LH (−29.5% and −19.6%, respectively) and MUO individuals (−22.6% and −18.5%, respectively) compared to MHO individuals. No differences were observed for lysine or hydroxyproline between MUO and LH individuals.

Aspartic acid, asparagine, and the asparagine-to-aspartic acid ratio were not significantly different among the three groups, nor were amino acids involved in the urea cycle (e.g., arginine, citrulline, and ornithine). The initial untargeted GC-MS analysis found that uric acid differed between the three groups; however, this metabolite was not amenable to the CE-MS method, and hypoxanthine, which is a biosynthetic precursor for uric acid, was examined instead. The trend seen for uric acid with GC-MS (i.e., MUO > MHO > LH; data not shown) was not reflected by hypoxanthine levels. Finally, the levels of alanine, histidine, tryptophan, serine, threonine, and acetylcarnitine were not significantly different between the three groups.

Table 3.2: Associations of select metabolites with adiposity and insulin sensitivity traits*.

	BMI		Fat Mass %		HOMA-IR		HbA1c	
	r	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value
Glycine	−0.3797	0.0207	NS	NS	NS	NS	−0.3378	0.0345
Valine	0.5945	0.0002	0.2638	0.0344	0.4094	0.0074	0.3022	0.0242
Isoleucine	0.6513	0.0002	0.3924	0.0404	0.6826	<0.0001	NS	NS
Leucine	0.5941	0.0011	NS	NS	0.5531	0.003	NS	NS
Glutamic acid	0.4354	0.0086	NS	NS	0.3194	0.0378	0.3240	0.0365
Glutamic acid-to-glutamine	−0.4022	0.0097	NS	NS	NS	NS	−0.4148	0.0081
Phenylalanine	0.5986	0.0003	0.5213	0.0018	NS	NS	0.4537	0.0061
Tyrosine	0.6705	0.0002	0.5706	0.0034	NS	NS	0.4714	0.0265
Cystine	NS	NS	0.5690	0.0362	NS	NS	0.5734	0.0324
Methionine	0.3625	0.0436	NS	NS	NS	NS	NS	NS
Proline	0.5044	0.0018	NS	NS	0.6589	<0.0001	NS	NS
Hydroxyproline	0.4553	0.0066	0.4207	0.0112	NS	NS	NS	NS
Carnitine	0.5282	0.0117	0.4718	0.0354	NS	NS	0.5010	0.0204

* Separate regressions were fitted for each combination of trait (body mass index (BMI), fat mass %, and homeostatic model assessment of insulin resistance (HOMA-IR) and glycated hemoglobin (HbA1c)) with amino acids and derivatives, adjusted for age and sex. NS, nonsignificant.

Serum Amino Acids Associate with Adiposity and Insulin Sensitivity

We next examined whether the aforementioned amino acids were associated with measures of adiposity (i.e., BMI and fat mass %) and estimates of insulin sensitivity (i.e., HOMA-IR and HbA1c) in our cohort of individuals. A number of metabolites were significantly associated with the aforementioned measures and are presented in Table 3.2.

It is noteworthy that the positive correlations between BCAAs and glutamic acid with HOMA-IR and HbA1c could be visually stratified according to the three groups (LH < MHO < MUO). Specifically, increases in valine, leucine, isoleucine, and glutamic acid concentrations were in parallel to increases in HOMA-IR values, while increases in glutamic acid, valine, and isoleucine concentrations were in parallel to increases in HbA1c values (Figures 3.3 and 3.4). In contrast, the relationship between glycine and the glutamine-to-glutamic acid ratio with HbA1c values decreased from LH to MHO to MUO individuals (Figure 3.4). Taken together, the associations between amino acids (e.g., BCAAs, glutamic acid and the glutamine-to-glutamic acid ratio) and markers of insulin sensitivity consistently demonstrated that MHO individuals

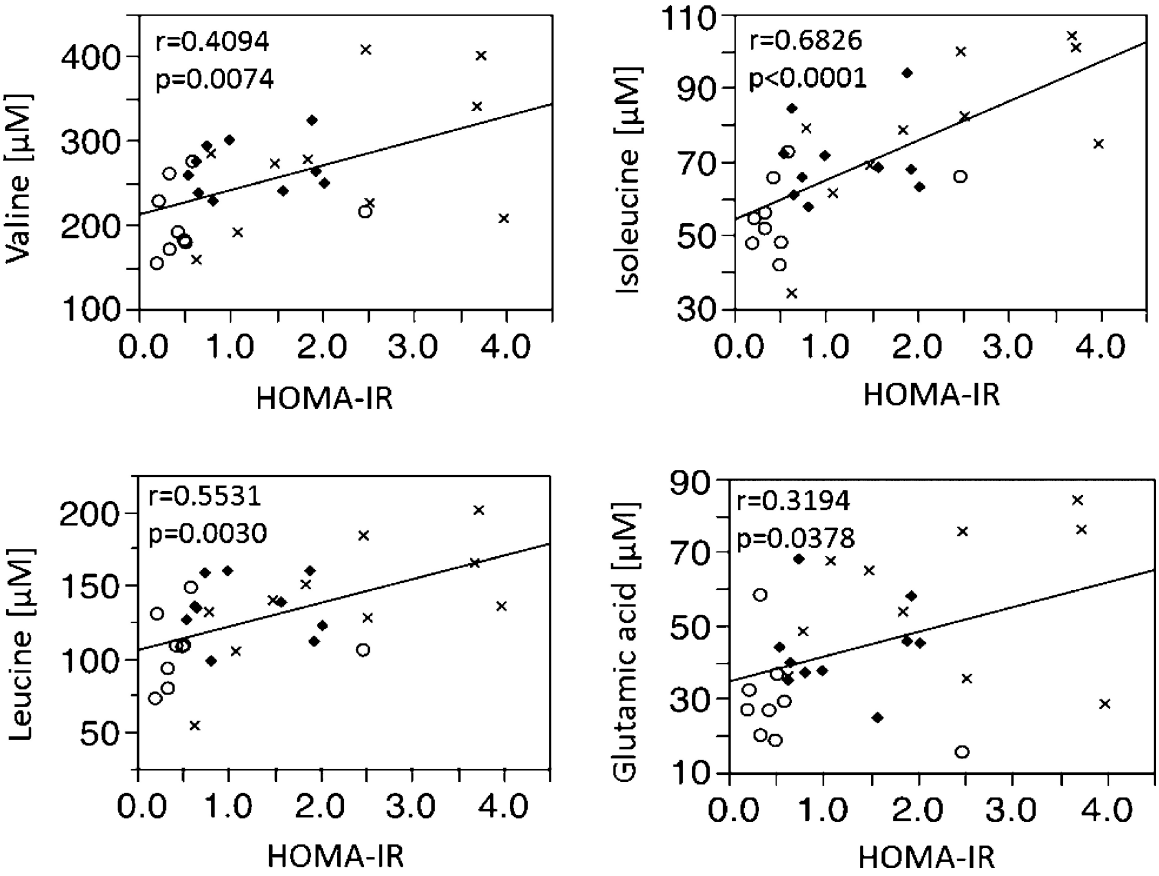


Figure 3.3: Relationship between specific amino acids and HOMA-IR. Metabolite values corresponded to quantitative data obtained by CE-MS analysis. (○) Lean healthy (LH) individuals, (◆) metabolically healthy obese (MHO) individuals, and (×) metabolically unhealthy obese (MUO) individuals.

have an intermediate phenotype compared to LH and MUO individuals.

Amino Acids as Potential Biomarkers of the MHO Phenotype

When considered individually, many of the aforementioned amino acids were not significantly different between MHO and MUO individuals; however, considering small subsets of amino acids simultaneously showed their potential to serve as biomarkers of the MHO phenotype. ROC curve analyses were used to evaluate if metabolites and/or metabolite ratios could be used to distinguish MHO from MUO individuals (Figure 3.5). Based on this analysis, we identified several potential biomarkers that

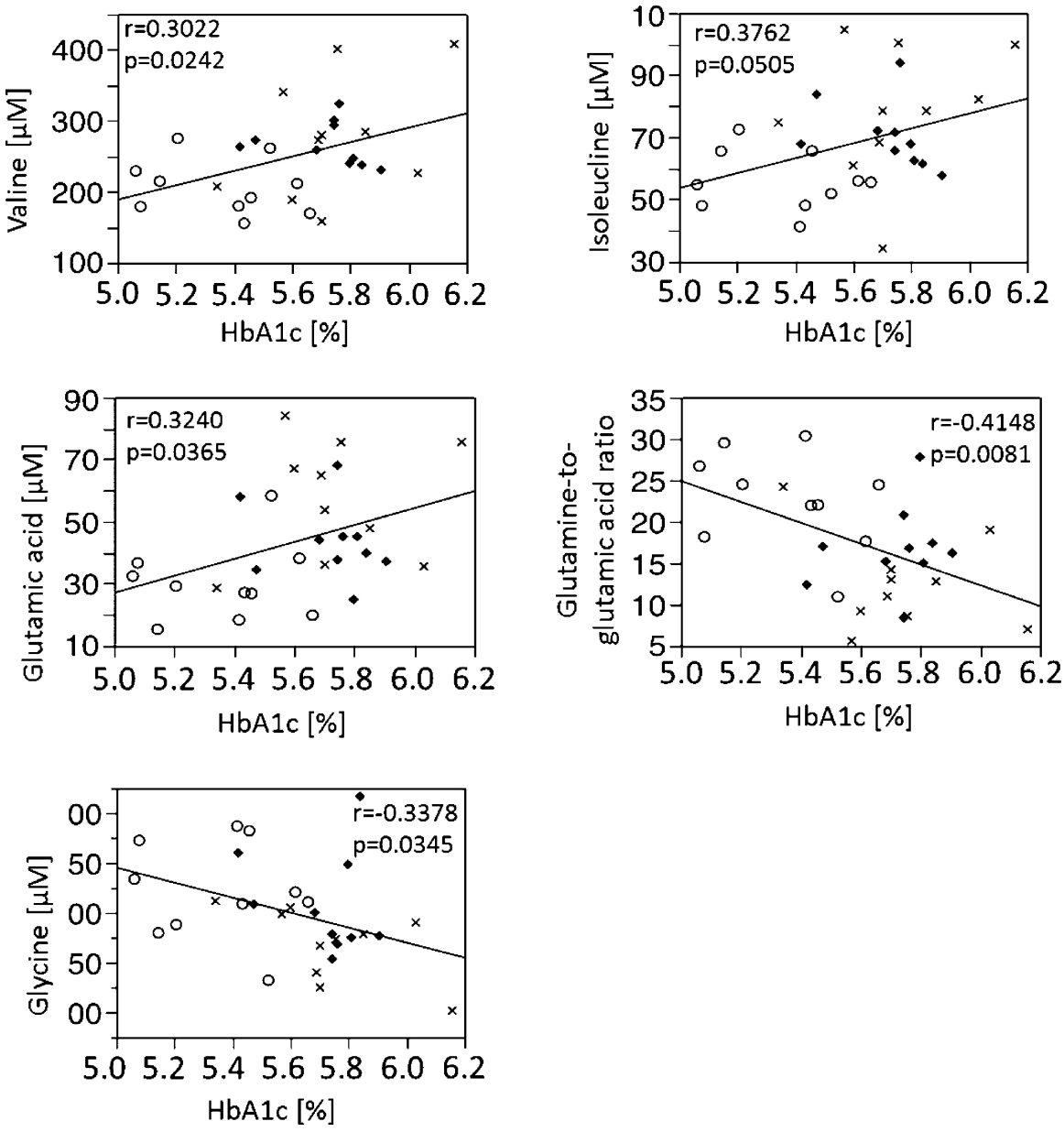


Figure 3.4: Relationship between specific amino acids and HbA1c. Metabolite values corresponded to quantitative data obtained by CE-MS analysis. (○) Lean healthy (LH) individuals, (◆) metabolically healthy obese (MHO) individuals, and (×) metabolically unhealthy obese (MUO) individuals.

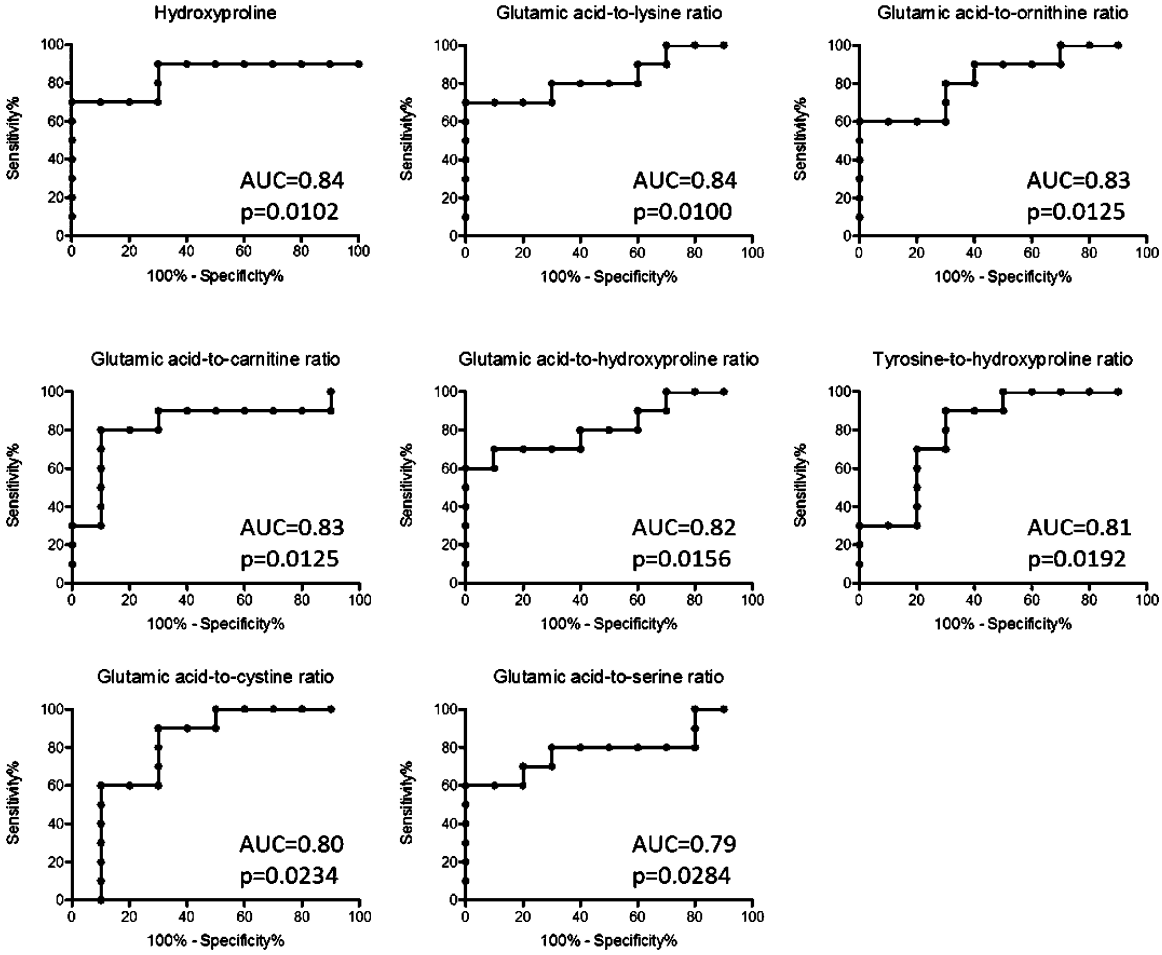


Figure 3.5: ROC curves for the significant metabolites measured by CE-MS. Metabolites and metabolite ratios that best differentiate MHO from MUO. AUC = area under the curve. A $p < 0.05$ indicated statistical significance.

could be used to differentiate MHO and MUO individuals, including the glutamic acid-to-lysine ratio (area under the curve (AUC) = 0.84; $p = 0.010$), the glutamic acid-to-ornithine ratio (AUC = 0.83; $p = 0.0125$), the glutamic acid-to-carnitine ratio (AUC = 0.83; $p = 0.0125$), the glutamic acid-to-hydroxyproline ratio (AUC = 0.82; $p = 0.0156$), the tyrosine-to-hydroxyproline ratio (AUC = 0.81; $p = 0.0192$), the glutamic acid-to-cystine ratio (AUC = 0.80; $p = 0.0234$), and the glutamic acid-to-serine ratio (AUC = 0.79; $p = 0.0284$), as well as hydroxyproline alone (AUC = 0.84; $p = 0.0102$).

Branched-Chain Amino Acid Catabolism and Energy Metabolism in Subcutaneous Adipose Tissue

Global gene expression analysis of SAT revealed that *branched-chain amino acid (BCAA) degradation* and *tricarboxylic acid (TCA) cycle* pathways were down-regulated in MHO and MUO individuals compared to LH individuals. The following genes from the *TCA cycle* pathway were decreased in both MHO and MUO individuals: *PCK1* and *PCK2*, which encode the cataplerotic enzymes that convert oxaloacetate to phosphoenolpyruvate, as well as *CS*, *DLD*, *DLST*, *FH*, *IDH3B*, *MDH2*, *PC*, *PDHB*, *SUCLA2*, and *SUCLG2* (Figure 3.6, Table 3.3). In contrast, only MUO individuals showed significant ($p < 0.05$) decreases in the expression of *ACO1*, *ACO2*, *IDH2*, *PDHA1*, *SDHB*, *SDHD*, and *SUCLG1* compared to LH individuals, suggesting that these pathways are more strongly compromised in MUO individuals compared to MHO individuals. Most of the genes in the *BCAA degradation* pathway were down-regulated in both groups of obese individuals compared to LH individuals (Table 3.4). The cytosolic transcript *BCAT1* was up-regulated in both MHO and MUO groups compared to the LH group. While *BCKDHA* was similarly decreased in both obese groups, the absolute expression of mitochondrial genes involved in the two first reactions of BCAA catabolism (i.e., *BCAT2*, *BCKDHB*, and DBT) were decreased significantly in MUO compared to MHO individuals (Table 3.4). Together, this gene expression profile suggests that MUO individuals may have a more significant impairment of SAT energy metabolism compared to MHO individuals.

3.1.4 Discussion

Obesity is considered a major risk factor for the development of T2D and cardiovascular disease. Changes in lifestyle (e.g., dietary habits, physical activity) can

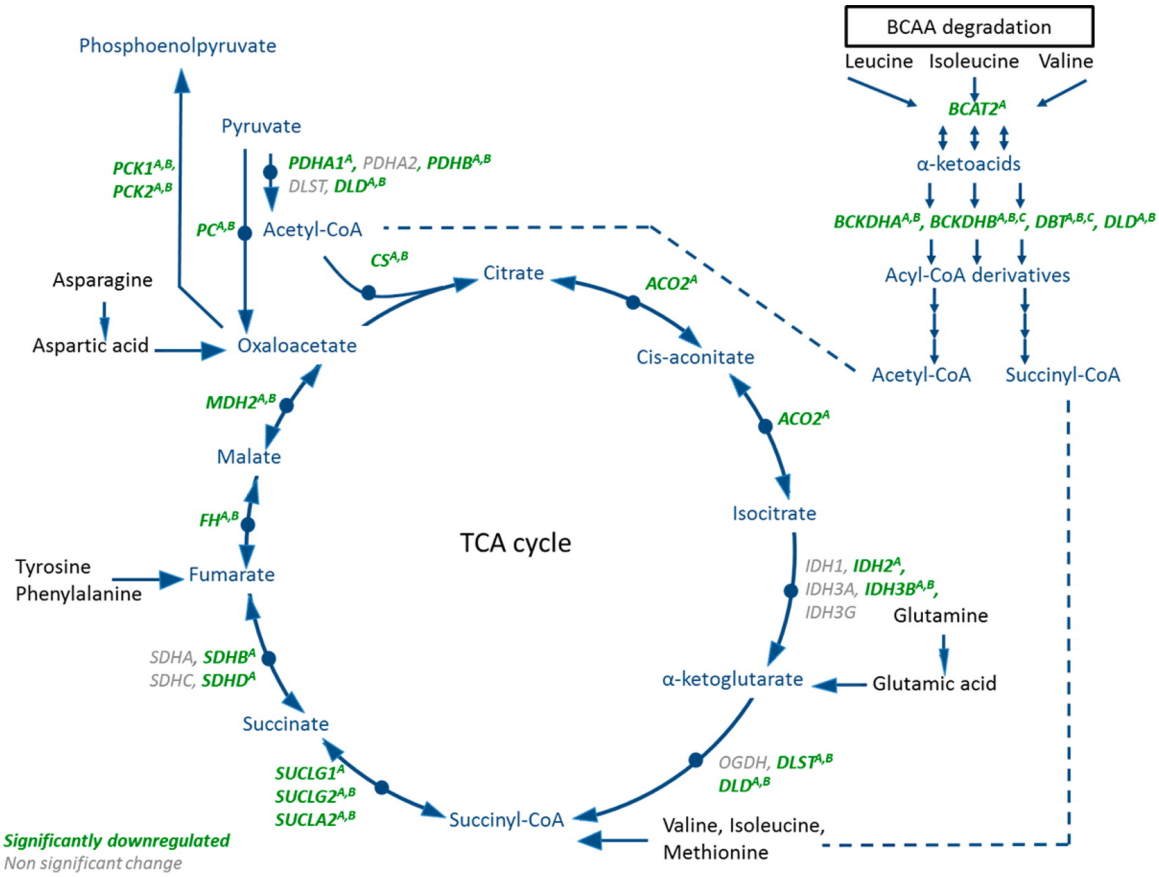


Figure 3.6: Schematic representation of tricarboxylic acid (TCA) cycle in subcutaneous adipose tissue (adapted from KEGG pathway). The superscript "A" indicates genes significantly down-regulated in MUO compared to LH ($p < 0.05$). The superscript "B" indicates genes significantly down-regulated in MHO compared to LH ($p < 0.05$). The superscript "C" indicates genes significantly down-regulated in MUO compared to MHO ($p < 0.05$). For corresponding fold changes and gene names, see Table 3.3.

Table 3.3: Subcutaneous adipose tissue gene expression related to the TCA cycle pathway (KEGG) in MHO and MUO individuals.

Gene Symbol	Gene Name	ID	Fold Change		Post-Hoc Significance
			MUO/LH	MHO/LH	
ACLY	ATP citrate lyase	47	0.94	0.94	
ACO1	Aconitase 1, soluble	48	0.98	0.99	A
ACO2	Aconitase 2, mitochondrial	50	0.96	0.98	A
CS	Citrate synthase	1431	0.95	0.96	A, B
DLD	Dihydrolipoamide dehydrogenase	1738	0.96	0.96	A, B
DLST	Dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex)	1743	0.95	0.95	A, B
FH	Fumarate hydratase	2271	0.95	0.96	A, B
IDH1	Isocitrate dehydrogenase 1 (NADP+), soluble	3417	0.99	1.02	
IDH2	Isocitrate dehydrogenase 2 (NADP+), mitochondrial	3418	0.95	0.98	A
IDH3A	Isocitrate dehydrogenase 3 (NAD+) alpha	3419	0.96	0.98	
IDH3B	Isocitrate dehydrogenase 3 (NAD+) beta	3420	0.96	0.97	A, B
IDH3G	Isocitrate dehydrogenase 3 (NAD+) gamma	3421	0.99	1.00	
MDH1	Malate dehydrogenase 1, NAD (soluble)	4190	0.98	1.00	
MDH2	Malate dehydrogenase 2, NAD (mitochondrial)	4191	0.96	0.96	A, B
OGDH	Oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide)	4967	0.98	0.99	
PC	Pyruvate carboxylase	5091	0.90	0.93	A, B
PCK1	Phosphoenolpyruvate carboxykinase 1	5105	0.83	0.85	A, B
PCK2	Phosphoenolpyruvate carboxykinase 2	5106	0.89	0.91	A, B
PDHA1	Pyruvate dehydrogenase (lipoamide) alpha 1	5160	0.94	0.97	A
PDHA2	Pyruvate dehydrogenase (lipoamide) alpha 2	5161	0.96	0.96	
PDHB	Pyruvate dehydrogenase (lipoamide) beta	5162	0.98	0.97	A, B
SDHAF1	Succinate dehydrogenase complex assembly factor 1	644096	0.98	0.96	
SDHAF2	Succinate dehydrogenase complex assembly factor 2	54949	0.99	1.00	
SDHB	Succinate dehydrogenase complex, subunit B, iron sulfur (lp)	6390	0.94	0.97	A
SDHC	Succinate dehydrogenase complex, subunit C, integral membrane protein, 15 kDa	6391	0.98	0.99	
SDHD	Succinate dehydrogenase complex, subunit D, integral membrane protein	6392	0.98	0.99	A
SUCLA2	Succinate-CoA ligase, ADP-forming, beta subunit	8803	0.95	0.97	A, B
SUCLG1	Succinate-CoA ligase, alpha subunit	8802	0.95	0.98	A
SUCLG2	Succinate-CoA ligase, GDP-forming, beta subunit	8801	0.95	0.96	A, B

A Significantly decreased in MUO in comparison to LH ($p < 0.05$)
B Significantly decreased in MHO in comparison to LH ($p < 0.05$)

Table 3.4: Subcutaneous adipose tissue gene expression related to the BCAA degradation pathway (KEGG) in MHO and MUO individuals.

Gene Symbol	Gene Name	ID	Fold Change		Post-Hoc Significance
			MUO/LH	MHO/LH	
ABAT	4-aminobutyrate aminotransferase	18	0.98	0.98	
ACAA1	Acetyl-CoA acyltransferase 1	30	0.99	0.98	
ACAA2	Acetyl-CoA acyltransferase 2	10 449	1.00	1.01	
ACAD8	Acyl-CoA dehydrogenase family, member 8	27 034	0.95	0.96	A, B
ACADM	Acyl-CoA dehydrogenase, C-4 to C-12 straight chain	34	0.94	0.95	A, B
ACADS	Acyl-CoA dehydrogenase, C-2 to C-3 short chain	35	0.94	0.96	A, B
ACADSB	Acyl-CoA dehydrogenase, short/branched chain	36	0.93	0.96	A, B, C
ACAT1	Acetyl-CoA acetyltransferase 1	38	0.93	0.95	A, B
ACAT2	Acetyl-CoA acetyltransferase 2	39	0.99	0.97	
ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)	217	0.98	0.99	A
ALDH6A1	Aldehyde dehydrogenase 6 family, member A1	4329	0.88	0.91	A, B
AOX1	Aldehyde oxidase 1	316	1.09	1.10	E
AUH	AU RNA binding protein/enoyl-CoA hydratase	549	0.96	0.97	A, B
BCAT1	Branched chain amino-acid transaminase 1, cytosolic	586	1.16	1.18	D, E
BCAT2	Branched chain amino-acid transaminase 2, mytochondrial	587	0.93	0.95	A
BCKDHA	Branched chain keto acid dehydrogenase E1, alpha polypeptide	593	0.94	0.95	A, B
BCKDHB	Branched chain keto acid dehydrogenase E1, beta polypeptide	594	0.91	0.94	A, B, C
BCKDK	Branched chain ketoacid dehydrogenase kinase	10 295	0.98	0.98	
DBT	Dihydrolipoamide branched chain transacylase E2	1629	0.93	0.95	A, B, C
DLD	Dihydrolipoamide dehydrogenase	1738	0.96	0.96	A, B
ECHS1	Enoyl CoA hydratase, short chain, 1, mitochondrial	1892	0.94	0.95	A, B
EHHADH	Enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase	1962	0.96	0.98	
HIBADH	3-hydroxyisobutyrate dehydrogenase	11 112	0.94	0.96	A, B
HIBCH	3-hydroxyisobutyryl-CoA hydrolase	26 275	0.97	0.96	B
HMGCL	3-hydroxymethyl-3-methylglutaryl-CoA lyase	3155	0.98	0.99	A, B
HMGCS1	3-hydroxymethyl-3-methylglutaryl-CoA synthase 1 (soluble)	3157	0.99	0.96	
HMGCS2	3-hydroxymethyl-3-methylglutaryl-CoA synthase 2 (mitochondrial)	3158	0.98	0.98	
HSD17B10	Hydroxysteroid (17-beta) dehydrogenase 10	3028	0.96	0.98	A
IL4I1	Interleukin 4 induced 1	259 307	1.02	1.01	
IVD	Isovaleryl-CoA dehydrogenase	3712	0.96	0.96	A, B
MCCC1	Methylcrotonoyl-CoA carboxylase 1 (alpha)	56 922	0.92	0.96	A, B, C
MCCC2	Methylcrotonoyl-CoA carboxylase 2 (beta)	64 087	0.95	0.97	A, B, C
MCEE	Methylmalonyl CoA epimerase	84 693	0.96	0.96	A, B
MUT	Methylmalonyl CoA mutase	4594	0.94	0.95	A, B
OXCT1	3-oxoacid CoA transferase 1	5019	0.95	0.97	A, B
OXCT2	3-oxoacid CoA transferase 2	64 064	1.04	1.03	
PCCA	Propionyl CoA carboxylase, alpha polypeptide	5095	0.93	0.94	A, B

A: Significantly decreased in MUO in comparison to LH ($p < 0.05$); B: Significantly decreased in MHO in comparison to LH ($p < 0.05$); C: Significantly decreased in comparison to LH and significantly decreased in MUO compared to MHO ($p < 0.05$); D: Significantly increased in MUO in comparison to LH ($p < 0.05$); E: Significantly increased in MHO in comparison to LH ($p < 0.05$)

help prevent the development of these obesity-related complications; however, it is currently unknown if all obese individuals will respond similarly to a given lifestyle intervention. The few studies that have examined the effects of lifestyle interventions in MHO and MUO individuals report conflicting findings, where in some cases the two groups responded similarly (Dalzill et al., 2014; Liu et al., 2013) and in other cases the two groups responded differently (Karelis et al., 2008; Shin et al., 2006). Given these conflicting findings, an improved understanding of the MHO phenotype is necessary before considering therapeutic and lifestyle management strategies that are personalized for distinct subgroups of obese individuals. The integrative metabolomic and transcriptomics approach presented here aimed to improve our understanding of the MHO phenotype in relation to the MUO phenotype. An untargeted metabolomic approach combined with multivariate data analysis and discriminant modeling highlighted the central role of amino acids in relation to obesity, and their association with the MHO phenotype. A targeted CE-MS method was subsequently used to confirm changes in amino acids and derivatives, thus reinforcing the strong association between this subset of metabolites with obesity and cardiometabolic risk factors. Integrating serum metabolite data with SAT gene expression suggested that energy and BCAA metabolism were more highly compromised in MUO individuals compared to MHO individuals.

In the present study, serum samples were obtained from well-characterized, age- and sex-matched groups of LH, MHO, and MUO individuals, as previously described (Perreault et al., 2014b). A comprehensive classification approach encompassing both anthropometric and clinical measurements was previously used to classify our study participants into these groups, and complemented by screening circulating inflammatory markers and serum fatty acids, as described elsewhere (Perreault et al., 2014b). Briefly, MHO individuals had a lipid profile similar to LH individuals, as

well as having lower levels of circulating inflammatory markers and increased insulin sensitivity compared to their MUO counterparts. Importantly, the differences between MHO and MUO individuals used in the present study agree with previous independent reports that MHO individuals are more insulin sensitive, have a more favorable lipid profile, and have lower levels of circulating inflammatory markers (Karelis et al., 2004b; Phillips and Perry, 2013). Metabolites are the end points of gene and protein function, and thus provide important insights regarding metabolic flux through key biological pathways (Ordovás Muñoz, 2013). Previous metabolomic studies have revealed a central role of amino acid metabolism with obesity and T2D, where specific amino acids are either increased or decreased with an unhealthy phenotype (Adams, 2011; Fiehn et al., 2010; Floegel et al., 2013; Newgard et al., 2009; Wang-Sattler et al., 2012). Our results suggest that studying amino acid and related metabolites is highly relevant to improve our understanding of the MHO phenotype. It is important to stress that the results presented here reflect changes in amino acid metabolism between the three groups rather than differences in dietary habits. This is supported by two key pieces of information. First, all study participants were fasted overnight ($\sim$12 h) prior to blood collection, thereby removing the influence of acute dietary influences. Second, a recent study by Phillips et al. (2013) found no evidence that dietary macronutrient composition differed between MHO and MUO individuals, suggesting that dietary habits do not explain differences between these two groups of obese individuals. Therefore, the amino acid signature reported here provides insight regarding differences in amino acid homeostasis between MHO and MUO individuals (Adams, 2011). Our results revealed distinct patterns of amino acids between LH, MHO, and MUO groups. Generally, most individual amino acids did not differ significantly between MHO and MUO individuals, and the specific increases or decreases in amino acid levels corroborated findings previously reported with obesity

and T2D (Adams, 2011; Kim et al., 2010; Morris et al., 2012; Newgard et al., 2009; Xie et al., 2012). However, we observed that several amino acids reflected the intermediate cardiometabolic profile that seems to be characteristic of the MHO phenotype.

For example, we observed a significant reduction of glycine levels in the MUO group that is consistent with previous reports (Fiehn et al., 2010; Floegel et al., 2013; Newgard et al., 2009; Oberbach et al., 2011; Wang-Sattler et al., 2012). It was hypothesized that decreases in glycine levels in obese individuals may reflect an increased utilization of this amino acid in gluconeogenic pathways that are typically up-regulated with insulin resistance (Floegel et al., 2013; Wang-Sattler et al., 2012). It was interesting to see that the concentration of glycine was intermediate in MHO individuals compared to LH and MUO individuals (i.e., LH > MHO > MUO), suggesting the higher insulin sensitivity in MHO may stem from a lower utilization of gluconeogenic pathways. The concentrations of glutamic acid were 44.5% and 89.8% higher in MHO and MUO individuals, respectively, compared to LH individuals, while the concentration of glutamine was not different within the three groups. A strength of our CE-MS methodology is that glutamine and glutamic acid were readily distinguished (Kuehnbaum and Britz-McKibbin, 2013), allowing us to calculate the glutamine-to-glutamic acid ratio. When considered simultaneously, this ratio was significantly reduced in MHO (−26.5%) and MUO (−46.0%) individuals compared to LH individuals. Further, the difference in the glutamine-to-glutamic acid ratio observed between MHO and MUO individuals showed a trend to be lower in MUO compared to MHO individuals ($p = 0.075$), suggesting that this ratio may serve as an indicator for reduced metabolic wellness. Interestingly, our results additionally show that glutamic acid was positively associated with BMI, HOMA-IR, and HbA1c (Figures 3.3 and 3.4), whereas the glutamine-to-glutamic acid ratio was inversely associated with HbA1c (Figure 3.4). It is noteworthy that the correlation between

glutamic acid and HOMA-IR, as well as the correlation between the glutamine-to-glutamic acid ratio and HbA1c, visually stratified the three phenotypes, with the MHO group being intermediate to LH and MUO individuals. Cheng et al. (2012) previously reported that a low glutamine-to-glutamic acid ratio was associated with higher risk of T2D in patients with metabolic syndrome. In another study, it was hypothesized that balanced levels of glutamine and glutamic acid could be important for glucose homeostasis (Xu et al., 2013). Indeed, glutamine is thought to have a beneficial effect on cardiometabolic risk via several potential mechanisms, such as increased secretion of glucagon-like peptide 1, increased externalization of glucose transporter type 4, and/or increased adipose tissue insulin sensitivity (Cheng et al., 2012). In contrast, increases in glutamic acid levels may promote metabolic complications by stimulating glucagon release from pancreatic α-cells (Cabrera et al., 2008) and increasing transamination of pyruvate to alanine; which together favors an increase in gluconeogenesis (Cheng et al., 2012). The higher absolute levels of serum glutamic acid and the decreased glutamine-to-glutamic acid ratio in MUO and MHO compared to LH individuals corroborates previous reports (Cheng et al., 2012; Newgard et al., 2009) and aligns with the prior observation that MHO individuals have an improved glucose homeostasis compared to MUO individuals.

Serum levels of the three BCAAs (i.e., valine, isoleucine, and leucine) were significantly increased in both obese phenotypes compared to LH individuals, but did not differ significantly between MHO and MUO groups. Interestingly, Batch et al. (2013) identified BCAAs and related metabolites (which included aromatic amino acids, methionine, alanine, and histidine) as being excellent biomarkers of metabolic wellness, independent of BMI. Despite failing to detect a significant difference in BCAA levels between MHO and MUO individuals, our results showed that the trend for increased insulin sensitivity previously reported for these MHO individuals (i.e.,

reduced fasting insulin and HOMA-IR values) aligns with their lower absolute concentrations of BCAAs compared to MUO individuals (Table 3.1) (Perreault et al., 2014b). The relationship between BCAAs and insulin sensitivity was further supported by the positive associations observed between the three individual BCAAs and HOMA-IR, and to a lesser extent with HbA1c (Table 3.2). While these associations need to be validated in a larger cohort, the trends observed corroborate previous research (Newgard et al., 2009; Pietiläinen et al., 2008). It was also reported in other studies that baseline plasma BCAAs associated with HOMA-IR values at a 6 year follow-up, leading to the suggestion that these amino acids may serve as early indicators for the development of insulin resistance (McCormack et al., 2013; Würtz et al., 2013). From an analytical perspective, it is noteworthy that the CE-MS method provided excellent resolution of serum amino acids, thus enabling the separation of leucine and isoleucine. This is relevant given that the majority of previous studies have been unable to distinguish these two amino acids (Batch et al., 2013; Mihalik et al., 2012; Newgard et al., 2009). Of the three BCAA metabolites, isoleucine showed the most important changes between the three groups compared to valine and leucine. In MUO individuals, isoleucine levels were increased ∼40% compared to LH individuals, while the mean isoleucine concentration in MHO individuals was increased ∼26%. The increase of isoleucine levels coincided with reductions in markers of insulin sensitivity, as reflected by the progressive association seen between the three groups with HOMA-IR and HbA1c (Figures 3.3 and 3.4).

In concordance with changes in serum amino acid profiles, we also found a role for SAT gene expression in BCAA catabolism. Transcriptomic profiling in adipose tissue showed a reduced expression of numerous genes involved in BCAA catabolism (*ACAD8*, *ACADM*, *ACADS*, *ACADSB*, *ACAT1*, *ALDH6A1*, *AUH*, *BCAT2*, *BCKDHA*, *BCKDHB*, *DBT*, *DLD*, *ECHS1*, *HIBADH*, *HMGCL*, *IVD*, *MCCC1*, *MCCC2*, *MCEE*,

MUT, *OXCT1*, and *PCCA*) in both MHO and MUO individuals (Figure 3.6 and Table 3.3), while *BCAT1* was up-regulated in both MHO and MUO individuals compared to LH individuals. We hypothesize that the differential regulation between *BCAT1* and *BCAT2* genes may reflect an attempt by SAT to compensate for the lack of mitochondrial TCA cycle intermediates by increasing a key cytosolic gene (i.e., *BCAT1*). Indeed, up-regulation of *BCAT1* is consistent with recent findings and previous reports highlighting an association with obesity and T2D (Chen et al., 2013; Soronen et al., 2012). Obesity is linked to mitochondrial alterations and the reduced expression of mitochondrial BCAA catabolic genes seen in our study aligns with previous reports showing altered BCAA catabolism with obesity in two independent populations (Lackey et al., 2013; Pietiläinen et al., 2008). However, we found that more genes related to BCAA catabolism were significantly altered in SAT from MUO individuals in comparison to MHO individuals. For example, *BCAT2* gene expression, which controls the first step in BCAA catabolism, was only reduced in MUO individuals, while genes controlling the irreversible conversion of α-ketoacids to acyl-coenzyme A (CoA) derivatives (e.g., *BCKDHA*, *BCKDHB*, and *DBT*) were more strongly reduced in MUO individuals compared to MHO individuals. Taken together, SAT gene expression suggests an intermediate BCAA catabolism profile in MHO individuals compared to LH and MUO individuals.

Interestingly, our findings suggest that MHO and MUO individuals may differ with regard to anaplerosis, that is, the availability of metabolites that enter the TCA cycle. Indeed, it was previously hypothesized that alterations in serum amino acid levels may reflect an anaplerotic stress due to reduced TCA cycle capacity in adipose tissue from obese and T2D individuals (Adams, 2011; Fiehn et al., 2010). Several studies have previously reported reduced TCA cycle activity in muscle and adipose tissue from insulin-resistant and T2D individuals (Qatanani et al., 2013; Ritov et al.,

2010). The TCA cycle is central to energy metabolism in tissues and its homeostasis reflects a balance between intermediates that leave the cycle to be converted into glucose, fatty acids, and nonessential amino acids (cataplerosis) and replenishment of these intermediates (anaplerosis) (Owen et al., 2002). Amino acid catabolism plays a major role in the production of TCA cycle intermediates (Figure 3.6). For example, glutamine is deaminated into glutamic acid, which can then be converted into α-ketoglutarate; tyrosine and phenylalanine can be converted into fumarate; valine, isoleucine, and methionine can be converted into succinyl-CoA; and asparagine can be deaminated into aspartate, which can then be converted into oxaloacetate (Owen et al., 2002). The higher serum levels of BCAAs, aromatic amino acids, glutamic acid, and methionine in both obese groups compared to LH individuals may reflect a reduced use of these metabolites for TCA cycle replenishment. This notion was previously supported in mouse and human studies showing that obesity is associated with a coordinated reduction in SAT energy metabolism that is reflected at the gene, protein, and metabolite levels (Adams, 2011; Flachs et al., 2013; Lackey et al., 2013; Naukkarinen et al., 2014). When considering BCAA and glutamic acid levels in MHO and MUO individuals compared to LH individuals, it is tempting to speculate that serum amino acid profiles may be indicative of reduced TCA cycle activity in SAT. Further, our gene expression data suggests a stronger impairment in mitochondrial TCA cycle activity in MUO compared to MHO individuals (Table 3.3 and Figure 3.6). Indeed, the anaplerotic stress suggested by serum metabolite profiles was mirrored by a decrease in the expression of numerous genes associated with the TCA cycle, including *ACO2*, *CS*, *DLD*, *DLST*, *FH*, *IDH2*, *IDH3B*, *MDH2*, *PC*, *PDHA1*, *PDHB*, *SUCLA2*, *SDHB*, *SDHD*, *SUCLG1*, and *SUCLG2* (Figure 3.6). In particular, the primary anaplerotic reaction driven by pyruvate carboxylase (*PC*) was down-regulated in both obese groups. This decrease suggests a reduced production of oxaloacetate

from pyruvate which would ultimately compromise TCA cycle activity (Gaster et al., 2012). Furthermore, genes involved in the regulation of several major steps in the TCA cycle (i.e., *ACO2*, *IDH2*, *SUCLG1*, *SDHB*, and *SDHD*) (Figure 3.6) were only down-regulated in MUO individuals compared to LH individuals, and not MHO individuals. Altogether, our data supports an important role for SAT in whole-body amino acid homeostasis; however, the cause of the transcriptional changes leading to differences in SAT gene expression and insulin resistance between MHO and MUO individuals remains to be elucidated.

Several limitations of the study should be addressed. First, we acknowledge that statistical power would be increased with a larger sample size; however, the small sample size in the present study is compensated by the extensive characterization of our study participants, which included body composition analyses, blood clinical measurements, inflammatory marker analyses, and serum fatty acid profiling. Nevertheless, future studies using a larger sample size and longitudinal approaches are a logical continuation to this line of investigation. Second, the use of two different analytical platforms to study serum metabolites prevented us from quantifying all amino acids derivatives. For example, it was not possible to quantify uric acid directly by CE-MS; therefore, its precursor hypoxanthine was measured instead. However, the high level of concordance between our GC-MS and CE-MS data sets adds considerable strength to the study conclusions. Finally, ROC curve analyses identified several amino acids that may have the potential to be used as a diagnostic tool to assess metabolic wellness; however, the clinical relevance of these amino acid biomarkers remains to be independently validated.

3.1.5 Conclusions

In conclusion, metabolite profiling revealed subtle differences in amino acid homeostasis between the two groups of obese individuals. Interestingly, the serum amino acid profile in MHO individuals reflected their improved insulin sensitivity compared to MUO individuals. Furthermore, the circulating amino acid profile in MHO individuals aligned with gene expression profiles in SAT, which indicated that MHO individuals had a smaller number of genes and/or smaller changes in the expression levels of genes associated with mitochondrial BCAA catabolism and the TCA cycle in comparison to MUO individuals, thus highlighting the contribution of adipose tissue in the development of MHO and MUO phenotypes. Taken together, this study has shed novel insight into the molecular and metabolic basis of the MHO phenotype and highlights the important link between amino acid homeostasis and cardiometabolic risk.

3.2 Metabolomics Reveals Metabolically Healthy and Unhealthy Obese Individuals Differ in their Response to a Caloric Challenge

Author Contributions

FB, KPL, PBM, and DMM conceived and designed the study. Clinical trial, blood collection, and classification of subjects based on BMI, fat mass %, and metabolic status measurements were conducted by MP, MAZ, FB, and DMM. Fatty acid profiling on serum samples and the homogenized meal was performed by FB (GC). Amino acid profiling on plasma samples and the homogenized meal was performed by KPL

(CE-MS). Calibration data was generated by FB and KPL. Data analysis and data interpretation was performed by FB and KPL. Reagents and materials were contributed by FB, KPL, PBM, and DMM. FB, KPL, and DMM wrote the manuscript. FB, KPL, PBM, and DMM edited the manuscript. All authors approved the final version of the manuscript.

Abstract

Objective

To determine if MHO individuals have a different metabolic response to a standardized diet compared to LH and MUO individuals.

Methods

Thirty adults (35–70 yr) were classified as LH, MHO, and MUO according to anthropometric and clinical measurements. Participants consumed a standardized high calorie meal (∼1330 kcal). Blood glucose and insulin were measured at fasting, and 15, 30, 60, 90 and 120 min postprandially. Additional blood samples were collected for the targeted analysis of amino acids (AAs) and derivatives, and fatty acids (FAs).

Results

The postprandial response (i.e., area under the curve, AUC) for serum glucose and insulin were similar between MHO and LH individuals, and significantly lower than MUO individuals ($p < 0.05$). Minor differences were found in postprandial responses for AAs between MHO and MUO individuals, while three polyunsaturated FAs (linoleic acid, γ-linolenic acid, arachidonic acid) showed smaller changes in serum after the meal in MHO individuals compared to MUO. Fasting levels for various AAs (notably

BCAAs) and FAs (e.g., saturated myristic and palmitic acids) were found to correlate with glucose and insulin AUC.

Conclusion

MHO individuals show preserved insulin sensitivity and a greater ability to adapt to a caloric challenge compared to MUO individuals.

3.2.1 Introduction

The widespread availability of foods rich in refined carbohydrates and fats is a major contributor to the obesity epidemic (Hruby and Hu, 2015). As most of an individual's day is spent in a postprandial state (van Dijk et al., 2009), studying a person's response to a diet provides valuable insight into metabolic function. Further, the dynamic adaptation to a caloric challenge is highly informative. For example, postprandial increases in TG are considered a risk factor for cardiovascular diseases (CVDs) (Alcala-Diaz et al., 2014; Perez-Martinez et al., 2014), while high postprandial glucose and insulin levels indicate a risk for T2D (Weyer et al., 1999). Therefore, evaluating metabolic responses following a standardized meal challenge can provide a better understanding of the link between foods and metabolism compared to fasted markers, and has the potential to reveal differences in metabolic adaptability between individuals (Moazzami et al., 2014).

Recently, a subgroup of obese individuals, commonly referred to as metabolically healthy obese (MHO), was identified as being protected from the usual downstream complications associated with obesity, such as T2D and CVD. Studying this particular phenotype has revealed that MHO exhibit distinct molecular and metabolic characteristics compared to metabolically unhealthy obese (MUO) individuals (Badoud

et al., 2015b; Samocha-Bonet et al., 2014). Previous reports have shown that the response to an oral glucose tolerance test (OGTT) (Kantartzis et al., 2011; Marini et al., 2014; Naukkarinen et al., 2014; Succurro et al., 2008) or an oral fat load differs between individuals varying in cardiometabolic risk (Perez-Martinez et al., 2014; van Dijk et al., 2009). For example, a recent study reported that MHO co-twins had lower AUCs for insulin and glucose following an OGTT compared to MUO co-twins (Naukkarinen et al., 2014). Perez-Martinez et al. (2014) showed that after an oral fat load, postprandial TG metabolism (i.e., the AUC for TG) and inflammatory status were lower in MHO individuals compared to their MUO counterparts, thus showing the greater ability of MHO to adapt to a caloric challenge. Therefore, underlying differences in response to a dietary challenge is of high interest in order to better understand mechanisms linked to cardiometabolic health.

This study is the first to examine the glycemic and insulinemic responses to a standardized high-calorie meal in lean healthy (LH), MHO, and MUO individuals. We previously reported that circulating AA and FA profiles differed between MHO and MUO individuals in a fasted state (Badoud et al., 2014; Perreault et al., 2014b). Therefore, the present study also used a targeted metabolomics approach to evaluate if the metabolic adaptation to this challenge had a different effect on AA and FA profiles in MHO and MUO individuals. Finally, we hypothesized that studying AAs and FAs could lead to the discovery of metabolites that may potentially serve as predictors for a person's postprandial glucose and/or insulin response to a caloric challenge. Together this report highlights the greater ability of MHO individuals to adapt to a high calorie meal compared to their MUO counterparts as well as their preserved insulin sensitivity.

3.2.2 Materials and Methods

Subjects and Study Design

Serum and plasma samples were obtained from individuals participating in the Diabetes Risk Assessment (DRA) study (Clinicaltrials.gov ID #NCT01884714). Persons were screened over the phone and excluded if they met any one of the following criteria: 1) aged < 35 or > 70 years; 2) diagnosed with an acute or chronic autoimmune inflammatory disease, infectious disease, viral infection, and/or cancer; or 3) regular alcohol consumption exceeding 2 drinks/day (1 drink = 10 g alcohol). None of the participants were taking antidiabetic medications. All participants signed a consent form and the research protocol was approved by the University of Guelph Research Ethics Board (REB#10AP033).

Subject Classification

Thirty participants were classified into 3 distinct groups: LH ($n = 10$), MHO ($n = 10$), and MUO ($n = 10$) based on their adiposity and metabolic status, as previously described (Perreault et al., 2014b). Briefly, adiposity status was determined using the revised BMI cut-offs proposed by Shah and Braverman (2012), where lean was considered $<28\,kg/m^2$ for males and $<24\,kg/m^2$ for females, and obese was considered $\geq 28\,kg\,m^{-2}$ for males and $\geq 24\,kg\,m^{-2}$ for females. Metabolic status was determined using criteria adapted from that originally proposed by Karelis et al. (2004a) in order to account for sex-specific differences and medication. An individual was considered "metabolically healthy" if 3 or more of the following criteria were met: HDL-c > 1.0 mmol/L for males and >1.3 mmol/L for females; TG < 1.7 mmol/L without use of lipid-lowering drugs; Total-c < 5.2 mmol/L; LDL-c < 2.6 mmol/L; and HOMA-IR < 1.95 without use of anti-diabetic drugs. Each group was comprised of 7

women and 3 men. Complete details regarding the methodologies used to measure biochemical parameters and percentage body fat can be found in (Perreault et al., 2014b). LH, MHO, and MUO groups were matched for age, while the MHO and MUO groups were matched for BMI and percentage body fat (Table 3.6).

Caloric Challenge

A high calorie meal (fast-food breakfast representative of the Western diet) was given to each participant following a 12 h overnight fast. The meal consisted of 2 sausage egg English muffins, 1 apple turnover, and ∼370 mL of concentrated orange juice, and provided a total calorie intake of ∼1330 kcal (i.e., 66 g of fat, 141 g of carbohydrates, 5 g of fibre and 42 g of proteins). The standardized meal was eaten within 20 min. Blood was collected for glucose and insulin measurements at fasting and postprandially at 15, 30, 60, 90 and 120 min. Additional blood samples were collected at fasting and at the 120 min postprandial time point (T120) for targeted analyses of AAs and derivatives, as well as FAs.

To limit potential confounding lifestyle factors, participants were asked to avoid rigorous exercise, over-the-counter medication, dietary supplements, vitamins, and herbal supplements for 48 h prior to the study visit. Furthermore, participants fasted for at least 12 h after consuming a standardized single-serving frozen dinner meal the night preceding the study visit, and 1 dinner roll (whole wheat or white), 1 vegetable side of their choice (corn, peas, broccoli, carrots, squash, zucchini, or green beans), 1 fruit (apple, orange, banana, peach, grapes, or melon), and 500 mL bottled water.

Blood Metabolite Profiling

Amino Acid Profiling. Plasma samples (50 µL) were analysed using CE-MS, as previously described (Badoud et al., 2014). Peak areas and migration times were

normalized relative to the IS, 3-chloro-L-tyrosine, and data were reported as absolute level (µM) and percent postprandial change (%PP) following the standardized meal intake, where %PP = ((PP value − fasting value)/fasting value) × 100.

Fatty Acid Profiling. FAs were profiled in serum samples using GC, as previously described (Perreault et al., 2014b). Briefly, 10 µL of a 1 µg/µL C17:0 internal standard was added to 100 µL of serum. FAs were extracted with chloroform:methanol (2:1, v/v) and methylated at 100 °C for 1.5 h. All samples were analyzed on an Agilent DB-FFAP column (15 m × 0.1 mm internal diameter; 0.1 µm), using an Agilent Technologies 7890A GC system (Agilent Technologies, Mississauga, ON, Canada) with flame ionization detector. Peaks were identified by comparison to a panel of 49 FA methyl ester standards suspended in hexane (ranging from C8:0 to C24:1n-9). Relative FA values were calculated as a % of total peak area and data were reported as %PP (as described above) using relative FA values.

Standardized Meal Composition Analysis

Protein Hydrolysis for Amino Acid Composition. Acid hydrolysis was performed on the homogenized meal to determine the AA composition. Briefly, the meal was homogenized in a standard blender and a 50 mg sample was placed in a glass tube to which 350 µL of 6 M hydrochloric acid (HCl) was added ($n = 3$). Samples were maintained at 100 °C for 24 h under nitrogen, prior to being cooled to room temperature. The acid hydrolysate containing the meal sample was evaporated to dryness, then neutralized and reconstituted in 2.1 mL of 1 M ammonium hydroxide (NH_4OH). Ultrafiltration was performed using 3 kDa MWCO NanoSep centrifugal devices (Pall Life Sciences, Washington, NY, USA) at 13 000 g for 20 min. Prior to CE-MS analysis, 15 µL of the filtered hydrolysate was diluted to 50 µL in NH_4Ac buffer

(pH of 5) with a final concentration of 200 mM NH_4Ac containing 25 µM of IS. AAs in the meal were expressed as relative abundances. Notably, during this procedure certain acid-labile amino acids have poor recoveries after protein digestion, including tryptophan and cysteine/cystine, while glutamine and asparagine are hydrolysed into glutamic acid and aspartic acid, respectively.

Fatty Acid Composition. An aliquot of 10 mg of the homogenized meal was used to determine FA composition by GC. 2.5 mL of 0.1 M potassium chloride (KCl) was added to 10 mg of the homogenized meal. The homogenate was transferred into 10 mL chloroform:methanol (2:1, v/v), and analyzed by GC using the same procedure as for the serum FA analysis. FAs were expressed as relative abundances in the meal.

Statistical Analyses

Statistical analyses were performed using GraphPad Prism 6 software (La Jolla, CA, USA). Non-parametric ANOVA Kruskal-Wallis tests were used to determine if the metabolites were statistically different between the three groups ($p < 0.05$). When significance was observed, a post-hoc non-parametric Mann-Whitney test was used for pairwise analyses ($p < 0.05$). AUC_{total} was calculated using the trapezoid rule, with a baseline of 0 over the range of fasting to T120 post meal intake. Separate regressions were fitted for all measured metabolites and insulin sensitivity indices using JMP 11 (SAS, Institute, Cary, NC), adjusted for sex, age and BMI. Significance was assessed after Bonferroni correction for multiple testing ($p \leq 0.01$).

3.2.3 Results

Glycemic and Insulinemic Responses

Glucose and insulin AUC_{total} differed significantly between the three groups (Figure 3.7a and b). The AUC_{total} for glucose and insulin (Figure 3.7c and d) were not different between LH and MHO individuals ($p = 0.40$ and $p = 0.31$, respectively), but both values were significantly higher in MUO individuals ($p = 0.02$ between LH and MUO and $p = 0.01$ between MHO and MUO for glucose; and $p < 0.01$ between LH and MUO and $p = 0.01$ between MHO and MUO for insulin).

Plasma Amino Acid and Derivatives Response

The levels of AAs and their derivatives were profiled in fasting and T120 plasma samples in the three distinct groups. A dataset comprising individual levels and ratios for the 39 AAs and their derivatives measured in each group are provided in Tables 3.7 and 3.8.

We first determined the relative abundance of the detected AAs in the meal, and then overlaid this with the corresponding %PP change in plasma AAs for LH, MHO and MUO groups (Figure 3.8a). Branched-chain AAs (BCAA: leucine, isoleucine, and valine), glutamic acid + glutamine (i.e., glutamine is converted to glutamic acid during the acid hydrolysis procedure), and proline were the most abundant AAs found in the meal. Most of the AAs and derivatives were similarly increased within the 3 groups at T120 following the meal (e.g., alanine, proline), but several showed a reduction (e.g., acetylcarnitine, citrulline) (Tables 3.7 and 3.8). When examining the %PP change after the consumption of the standardized meal (Figure 3.8a and Figure 3.9), four metabolites (i.e., asparagine, cystine, glutamine, and serine) and the carnitine-to-acetylcarnitine ratio were significantly different between the 3 groups

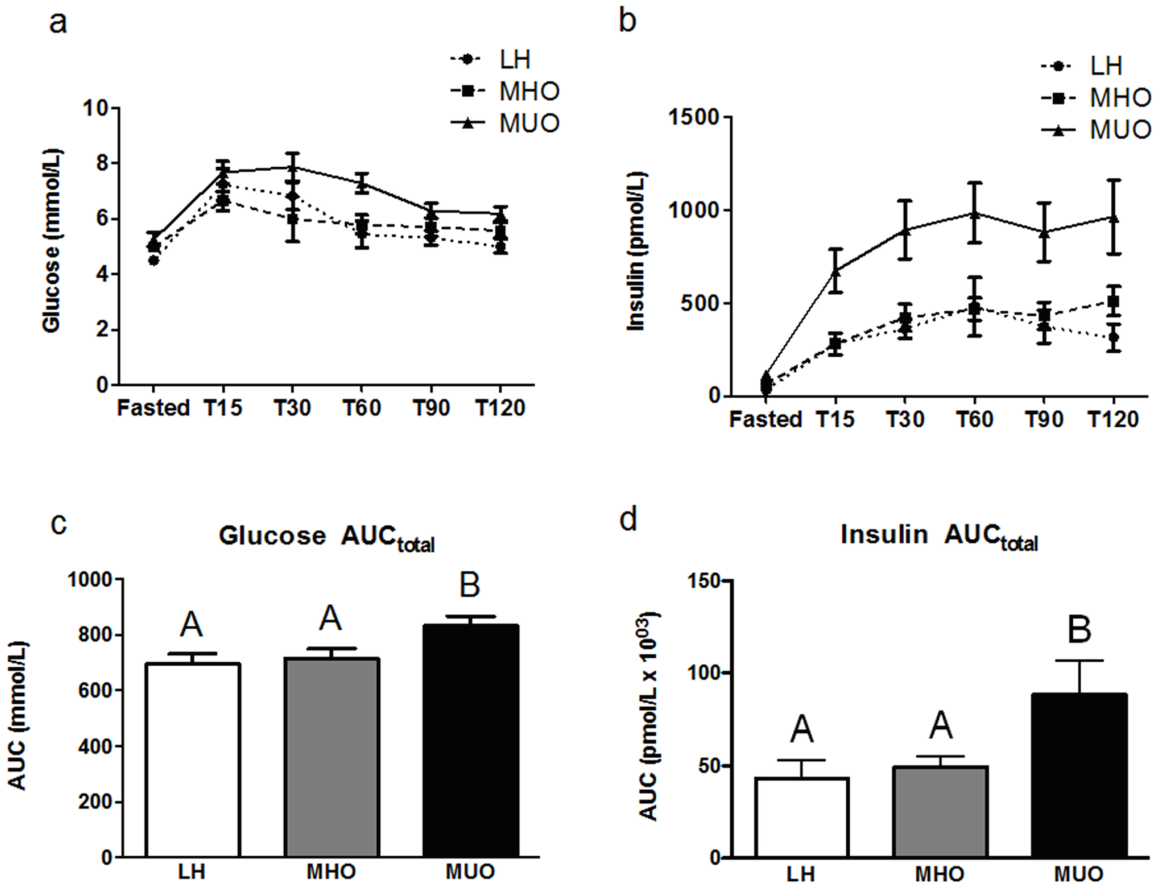


Figure 3.7: Postprandial serum glucose (a) and insulin (b) responses (mean ± SEM) following consumption of the standardized meal in lean healthy (LH; circle, $n = 10$), metabolically healthy obese (MHO; square, $n = 10$), and metabolically unhealthy obese (MUO; triangle, $n = 10$) individuals. Glucose (c) and insulin (d) Area Under the Curve (AUC_{total}), where white bars = LH; grey bars = MHO; and black bars = MUO. A non-parametric ANOVA Kruskal-Wallis test followed by a post-hoc Mann-Whitney test was used to determine differences between groups. Bars not sharing the same letter are statistically different ($p < 0.05$).

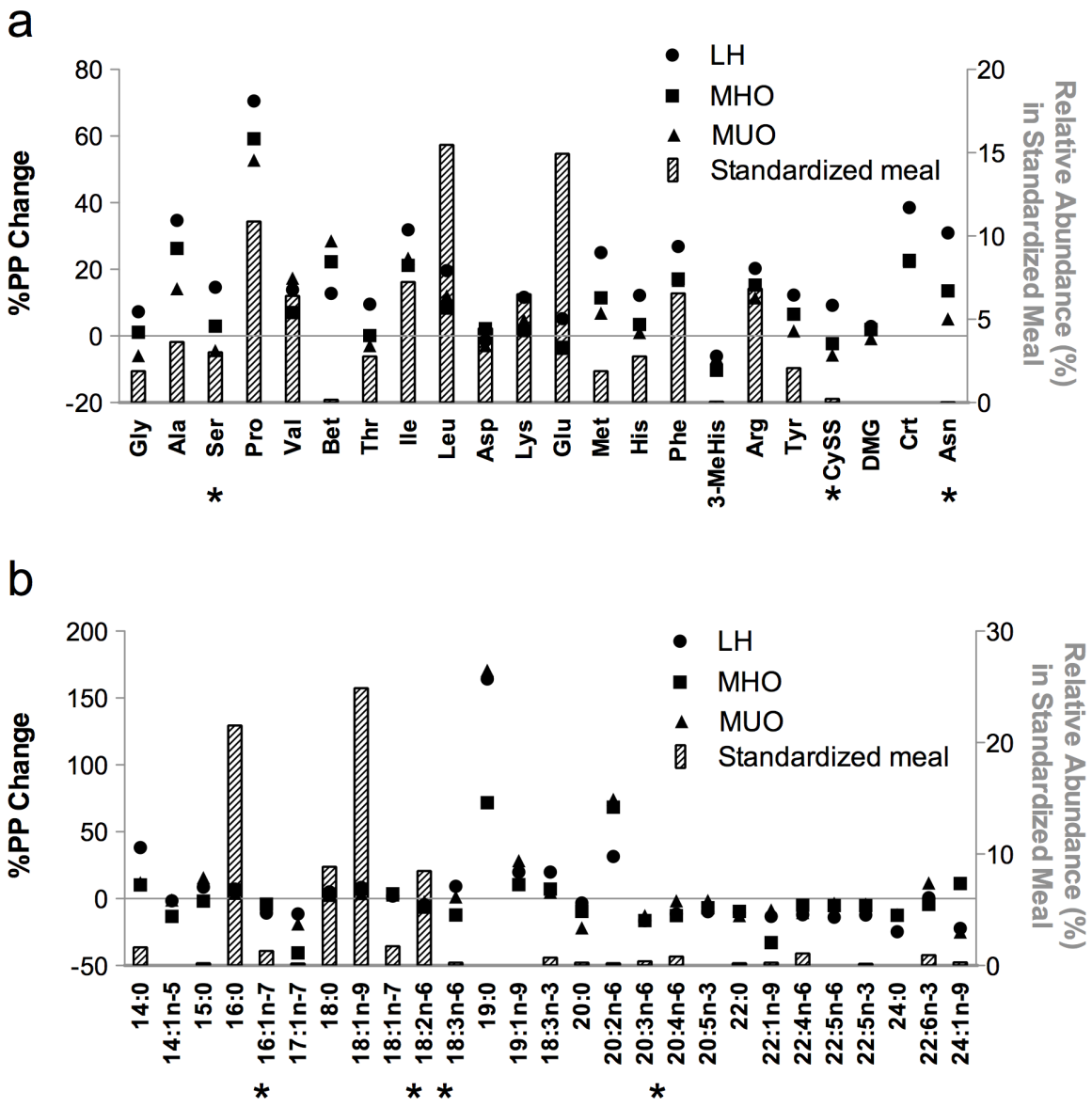


Figure 3.8: Amino acid (a) and fatty acid (b) composition of the standardized meal (grey bars, plotted on right y-axis). The corresponding %PP is plotted on the left y-axis for lean healthy (LH; circle, $n = 10$), metabolically healthy obese (MHO; square, $n = 10$) and metabolically unhealthy obese (MUO; triangle, $n = 10$) in plasma samples. Significant %PP (indicated by *) between the three groups were identified using a non-parametric ANOVA Kruskal-Wallis ($p < 0.05$) test followed by a post-hoc Mann-Whitney test ($p < 0.05$). Gly = glycine, Ala = alanine, Ser = serine, Pro = proline, Val = valine, Bet = betaine, Thr = threonine, Ile = isoleucine, Leu = leucine, Asp = aspartic acid, Lys = lysine, Glu = glutamic acid, Met = methionine, His = histidine, Phe = phenylalanine, 3-MeHis = 3-methylhistidine, Arg = arginine, Tyr = tyrosine, CySS = cystine, DMG = dimethylglycine, Crt = creatine, Asn = asparagine.

(ANOVA Kruskal-Wallis $p < 0.05$, Table 3.8). For asparagine and glutamine, the %PP change between fasting and T120 was intermediate for MHO individuals compared to MUO and LH, with levels in MUO being significantly lower compared to that of LH individuals. Cystine and serine revealed a significantly higher %PP change in the LH group relative to the obese groups, with no distinction between MHO and MUO. Finally, the carnitine-to-acetylcarnitine ratio was significantly increased for each group, where the %PP change was higher in LH compared to that of MHO individuals, while being intermediate for MUO individuals.

Serum Fatty Acid Response

The standardized meal provided the most abundant dietary FAs, as expected (Figure 3.8b). Indeed, the meal contained ∼38% saturated fatty acid (SFA), of which palmitic acid (16:0) and stearic acid (18:0) were the most abundant, ∼29% monounsaturated fatty acid (MUFA) of which oleic acid (18:1n-9) was the most abundant, and ∼17% polyunsaturated fatty acid (PUFA), which consisted predominantly of linoleic acid (18:2n-6). In parallel, we profiled serum FAs in fasting and postprandial T120 plasma samples. Twenty-seven FAs were consistently detected in all samples and either increased or decreased similarly following the caloric challenge in the LH, MHO, and MUO groups (Tables 3.9 and 3.10). We also calculated the %PP change between the two time points. As shown in Figure 3.8b, which illustrates the abundance of FAs in the meal relative to the %PP between fasting and T120, four FAs (palmitoleic acid (16:1n-7), 18:2n-6, γ-linolenic acid (18:3n-6), and arachidonic acid (20:4n-6)) revealed significantly distinct %PP changes (Figure 3.10) between the three groups (ANOVA Kruskal-Wallis $p < 0.05$, see Table 3.10). The %PP for 16:1n-7 was significantly lower in MHO and MUO compared to LH, while the %PP for 18:2n-6 and 20:4n-6 was significantly lower in MHO compared to MUO, but was intermediate for LH. In

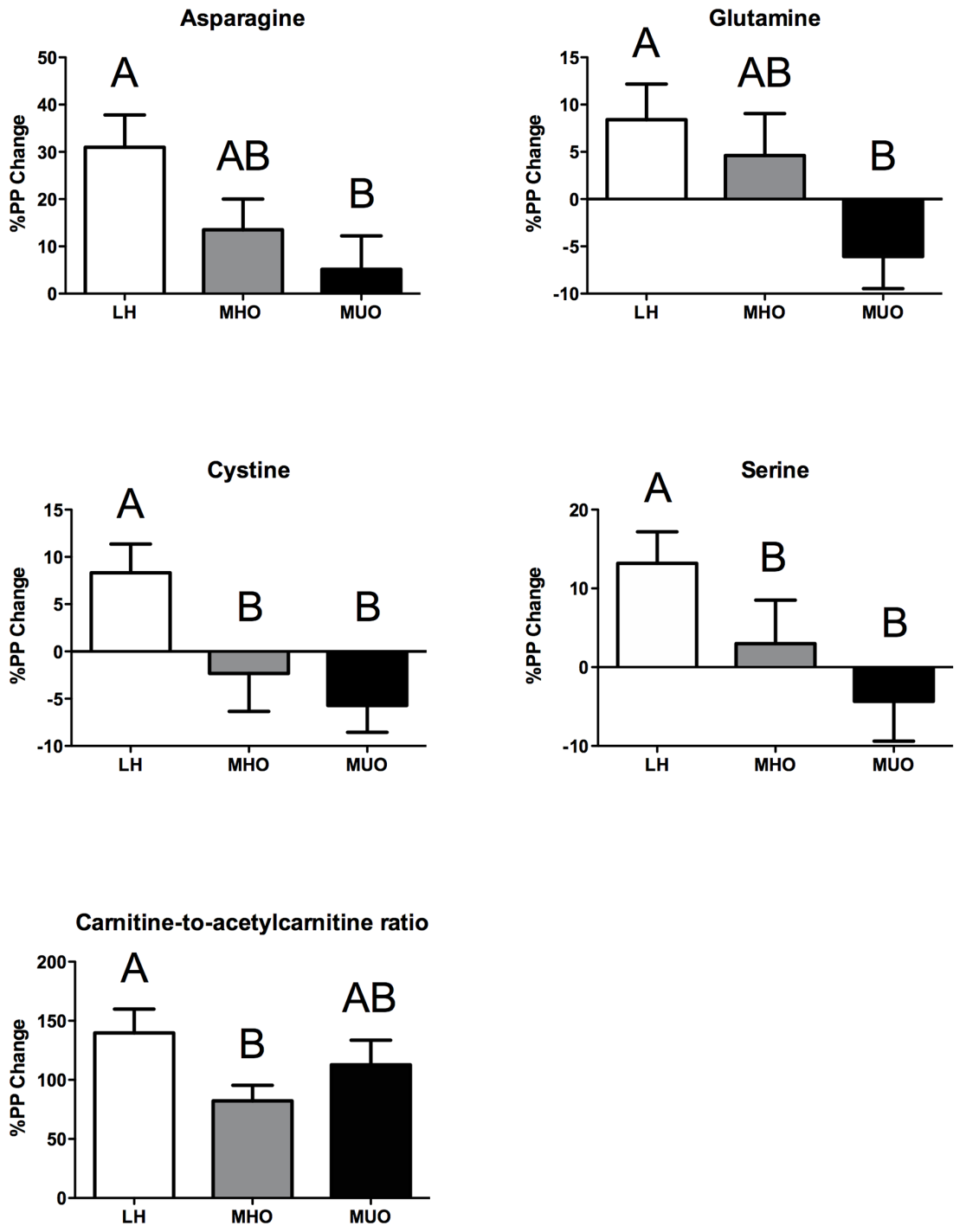


Figure 3.9: Mean %PP of plasma amino acid and derivatives measured by CE-MS for lean healthy (LH, $n = 10$, white bars), metabolically healthy obese (MHO, $n = 10$, grey bars) and metabolically unhealthy obese (MUO, $n = 10$, black bars). Significant %PP amino acids between the three groups were identified using a non-parametric ANOVA Kruskal-Wallis ($p < 0.05$) test followed by a post-hoc Mann-Whitney test ($p < 0.05$). Data is presented as mean %PP $\pm$ SEM.

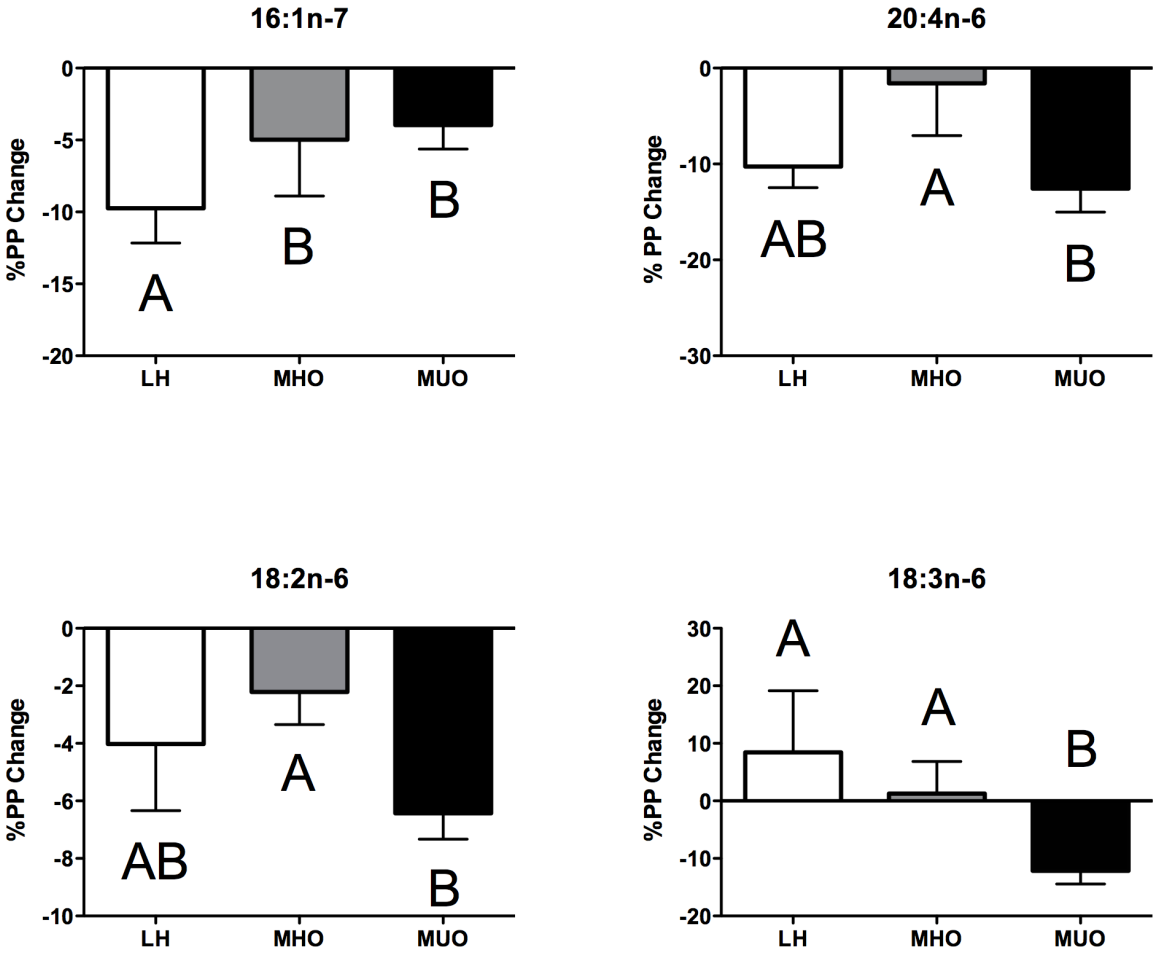


Figure 3.10: Mean %PP of serum fatty acids for lean healthy (LH, $n = 10$, white bars), metabolically healthy obese (MHO, $n = 10$, grey bars) and metabolically unhealthy obese (MUO, $n = 10$, black bars). Significant %PP between the three groups were identified using a non-parametric ANOVA Kruskal-Wallis test ($p < 0.05$) followed by a post-hoc Mann-Whitney test ($p < 0.05$). Data is presented as mean %PP ± SEM.

contrast, the %PP for 18:3n-6 was higher for LH and MHO in comparison with the MUO group.

Baseline Metabolites Associate with Measures of Insulin Sensitivity

We next examined the associations between all measured metabolites with markers of insulin sensitivity at both fasting (i.e., HOMA-IR, fasting glucose and insulin) and postprandially (i.e., glucose and insulin AUCs) to evaluate if metabolites correlated

with indices of insulin sensitivity and if they could be used to potentially predict glucose and insulin responses (Table 3.5). None of the AAs that showed a significantly different %PP change between the 3 groups correlated with markers of insulin sensitivity. However, we observed positive associations between other AAs and their derivatives with insulin sensitivity indices. Proline and leucine were positively associated with HOMA-IR; creatine and proline with fasting glucose; and proline and leucine with fasting insulin. Moreover, fasting isoleucine levels correlated positively with insulin AUC.

18:2n-6 (whose %PP change differed significantly between the 3 groups) showed a significant inverse relationship with fasting glucose. For the other measured FAs, the associations between SFAs and markers of glucose homeostasis tended to separate into two groups. On the one hand, 14 and 16-carbon SFAs (i.e., myristic acid (14:0) and 16:0) were positively correlated with HOMA-IR, fasting insulin, and glucose AUC, while 14:0 was also positively associated with fasting glucose and 16:0 with insulin AUC. On the other hand, the longer chain SFA 18:0 was inversely associated with fasting glucose. The MUFA myristoleic acid (14:1n-5) was positively correlated with HOMA-IR, fasting glucose and glucose AUC; and 18:1n-9 was positively associated with fasting glucose.

3.2.4 Discussion

General Summary

We have previously reported that during catabolic conditions (i.e., fasting), the MHO group differed from LH and MUO individuals. Specifically, circulating AA, FA, and inflammatory marker profiles in MHO individuals were intermediate to LH and MUO individuals (Badoud et al., 2014; Shah and Braverman, 2012). Building on these

Table 3.5: Associations between amino acids and derivatives, and fatty acids, with fasting and postprandial indices of insulin sensitivity.

Metabolites		Baseline Indices									Postprandial Indices					
		HOMA-IR			Fasting Glucose			Fasting Insulin			AUC Glucose			AUC Insulin		
		r^*	p-value†	β	r^*	p-value†	β	r^*	p-value†	β	r^*	p-value†	β	r^*	p-value†	β
Amino Acid and Derivatives																
Creatine (µM)					0.464	<0.01	0.533									
Proline (µM)		0.645	<0.01	0.658	0.609	<0.01	0.490	0.630	<0.01	0.648						
Isoleucine (µM)														0.730	<0.01	0.460
Leucine (µM)		0.761	0.01	0.368				0.760	0.01	0.388						
Class	Fatty Acids															
SFA	Myristic acid (14:0, %)	0.516	0.01	0.549	0.520	0.01	0.450	0.489	0.01	0.524	0.527	<0.01	0.458			
	Palmitic acid (16:0, %)	0.681	<0.01	0.662				0.675	<0.01	0.662	0.729	<0.01	0.523	0.700	<0.01	0.623
	Stearic acid (18:0, %)				−0.536	<0.01	−0.455									
MUFA	Myristoleic acid (14:1n-5, %)	0.433	0.01	0.617	0.671	<0.01	0.687				0.486	<0.01	−0.407			
	Oleic acid (18:1n-9, %)				0.644	<0.01	0.458									
PUFA	Linoleic acid (18:2n-6, %)				−0.629	<0.01	−0.433									

* Linear regressions were adjusted for age, sex, and BMI.
† Significance was set at $p \leq 0.01$ using Bonferroni correction for multiple testing.

previous observations, the present study investigated for the first time the response to a high calorie Western meal within these distinct groups of individuals. This is highly relevant as this corresponds to a real-world meal typical of the Western diet, as opposed to clinical tolerance tests. Importantly, study participants received standardized lifestyle advice 48 h prior to the caloric challenge in order to minimize inter-individual lifestyle differences, which was previously shown to adequately normalize the human metabolome to reduce variations between people (Winnike et al., 2009). We acknowledge that the caloric challenge imposed with the fast-food meal may be significantly greater for lean individuals compared to obese individuals when considering their habitual dietary habits; however, imposing the same meal challenge in all subjects was essential in order to emphasize the differences in the responses between the MHO and MUO groups in comparison to LH individuals. Nevertheless, future studies could account for habitual caloric intake and standardize the increase in calories in accordance to each participant's daily caloric intake. Our results showed that the response to the high calorie meal differed between the three groups. We found that MHO individuals, similar to LH individuals, had a preserved glycemic and insulinemic postprandial response compared to their MUO counterparts despite having the same BMI and body fat %. Moreover, we observed distinct AA and FA profiles following the caloric challenge, and identified metabolites that significantly correlated with an individual's glycemic and insulinemic response.

Amino acids and their Derivatives

We examined AAs and AA-derivatives at fasting and postprandially to evaluate the %PP change within LH, MHO, and MUO groups. Analyzing the high calorie meal by CE-MS showed that it contained detectable levels for the majority of AAs. Correspondingly, most of the metabolites experienced similar increases or decreases

following the meal between the three groups (Tables 3.7 and 3.8). This aligns with previous studies showing that AA homeostasis was perturbed in anabolic states (i.e., postprandially) following a caloric challenge (Bondia-Pons et al., 2014; Bos et al., 2003; Krug et al., 2012; Moazzami et al., 2014; Pellis et al., 2012). For example, concordant with our observations, BCAAs were increased following the ingestion of different breads in healthy postmenopausal women (Moazzami et al., 2014), and following a standardized high calorie Big Mac meal in discordant obese and lean twins (Bondia-Pons et al., 2014). Here, our goal was to determine if postprandial alterations in the AA profile differed between individuals varying in cardiometabolic risk.

When studying the %PP change, several AAs significantly differed between the 3 groups (Figure 3.9). This included serine and cystine, which showed a similar reduced %PP change in MHO and MUO individuals compared to LH. This is interesting, as total cyst(e)ine levels were shown to be a predictor of obesity and insulin resistance in both children and adults (Elshorbagy et al., 2009, 2012). In contrast, asparagine and glutamine levels had an intermediate %PP change in MHO relative to those of LH and MUO individuals. Asparagine and glutamine are precursors to aspartic acid and glutamic acid, respectively, via deamination, which are subsequently used to replenish intermediates for the TCA cycle. This is notable as the TCA cycle was shown to reach its optimal capacity early in the postprandial phase (Pellis et al., 2012); therefore, the different groups of individuals may potentially differ with regards to metabolite flux entering and/or leaving the TCA cycle. Interestingly, we previously showed that the TCA cycle was impaired in subcutaneous adipose tissue from obese individuals, but to a lesser extent in MHO relative to MUO individuals (Badoud et al., 2014). We and others have also shown that the decreased expression of a number of genes associated with the TCA cycle and BCAA catabolism in adipose tissue from obese individuals was associated with higher serum BCAA levels and insulin resistance (Badoud et al.,

2014; Lackey et al., 2013; Pietiläinen et al., 2008). As such, it is tempting to speculate that the variable circulating levels of the aforementioned AAs, for which the %PP change between the three groups followed the general pattern of LH > MHO > MUO at T120, may stem from differences in TCA cycle functionality.

We also observed that the carnitine-to-acetylcarnitine ratio differed significantly between LH and MHO individuals, while found at intermediate levels in MUO individuals. Interestingly, we observed a trend for a higher %PP change in this ratio in LH individuals compared to the obese groups (absolute value, Figure 3.9 and Table 3.8). This is intriguing given that carnitine and acylcarnitines, and their ratio, were suggested to indicate a switch from a catabolic (i.e., fasting) state to an anabolic state (i.e., postprandial) following a number of different caloric challenges (e.g., a standard liquid diet, an OGTT, an oral lipid tolerance test (Krug et al., 2012)), and also for predicting OGTT changes after high-intensity interval training (Kuehnbaum et al., 2015). During fasting, reduced carnitine levels in blood indicate an increased cellular uptake, while the release of acetyl- and acyl-CoA into blood due to increased FA β-oxidation is buffered by increased acylcarnitines. Therefore, the catabolic state is reflected by an increase in acylcarnitines and the anabolic status is reflected by a decrease in acylcarnitines (Krug et al., 2012). Our results are consistent with this phenomenon, and revealed the highest %PP change in the ratio for the LH group (Figure 3.9), suggesting their heightened ability to metabolically respond to the caloric challenge.

As it has been shown that AAs, and particularly the most potent insulinogenic BCAAs (i.e., leucine and isoleucine), correlate with indices of insulin sensitivity and could predict T2D (Moazzami et al., 2014; Nakamura et al., 2014; Palmer et al., 2015), we evaluated associations between fasting AAs and parameters of insulin sensitivity before and after the caloric challenge. Fasting levels for isoleucine correlated

significantly with both fasting insulin levels and insulin AUC (Table 3.5); while leucine associated with HOMA-IR and fasting insulin. These results are in agreement with previous works suggesting that BCAAs could be used as predictors of insulin resistance and as biomarkers of T2D development (McCormack et al., 2013; Wang et al., 2011; Würtz et al., 2013).

Fatty Acids

In parallel to changes in AA homeostasis, we also observed modifications in the FA profile after consumption of the standardized meal. Most of the FAs increased or decreased in a similar way between fasting and T120 in the three groups of individuals (Tables 3.9 and 3.10). However, we found that four FAs (16:1n-7, 18:2n-6, 18:3n-6, and 20:4n-6) had significantly different %PP changes (Figure 3.10). Interestingly, 16:1n-7 showed the same trend in MHO and MUO individuals in comparison to LH, while the other FAs revealed %PP changes that were similar between LH and MHO groups.

Recent evidence has shown relationships between FAs, insulin sensitivity indices, and T2D risk (Rhee et al., 2011); therefore, we evaluated if FAs were also associated with parameters of insulin sensitivity in our study. Of the four FAs mentioned above, we observed a positive correlation between fasting 16:1n-7 levels and glucose AUC, and inverse correlations between 18:2n-6 and fasting glucose. We also extended our association analyses to include all measured FAs. This allowed the link between the profiled FAs and indices of insulin sensitivity at fasting, as well as their potential association with glycemic and insulinemic postprandial responses, to be investigated. We found that fasting levels of SFAs 14:0 and 16:0 correlated positively with all parameters of fasting and postprandial insulin sensitivity, while longer-chain 18:0 was inversely correlated with fasting glucose. Moreover, we also observed a trend for inverse associations between 18:0 ($r = -0.492$, $p = 0.02$) and 20:0 ($r = -0.533$,

$p = 0.03$) with glucose AUC (data not shown). This is relevant given that, according to our observations, shorter chain SFAs (14:0 and 16:0) were previously linked with an unhealthy cardiometabolic profile, compared to longer chain SFAs (18:0, 22:0 and 24:0) (Komatsu and Sharp, 1998; Perreault et al., 2014a; Ralston et al., 2013). This was further reinforced by the fact that the shorter chain SFAs associated with HOMA-IR, while 18:0 and 20:0 did not. Our results also revealed inverse correlations between 18:2n-6 PUFAs and fasting glucose that supports previous data indicating that this n-6 PUFA is inversely associated with T2D risk (Hodge et al., 2007; Patel et al., 2010).

Study Limitations

The T120 time point is optimal for assessing glucose and insulin responses, and although differences in TG and free FAs were reported by van Dijk et al. (2009) in the 2 h postprandial period, we acknowledge that longer follow-up times and additional time points would generate further information regarding lipid responses following the standardized meal. This will help provide insights into inter-individual variations in nutrient absorption and gut microbial co-metabolism (Heymsfield and Pietrobelli, 2011), as well as the association between metabolites and overall cardiometabolic risk (notably when expanding metabolomic coverage to include organic acids and lipids). While our sample size may be construed as small, the robust clinical characterization of our study participants, the use of a standardized caloric challenge, and the fact that each individual serves as their own control ensures a high degree of confidence in our results.

3.2.5 Conclusions

In conclusion, the present study showed that the glycemic and insulinemic postprandial responses were significantly different between individuals varying in cardiometabolic risk. MHO had a greater ability to adapt to the caloric challenge compared to their MUO counterparts, thereby highlighting their preserved insulin sensitivity. The targeted metabolomic and FA profiling approaches revealed that several metabolites differed significantly after the challenge. Additionally, we identified metabolites at baseline that should be further studied for their potential to predict an individual's postprandial response and cardiometabolic risk, independent of BMI. Indeed, the positive correlations seen between fasting levels of isoleucine and both fasting insulin levels and insulin AUC, as well as the positive associations seen between leucine and both HOMA-IR and fasting insulin, show the high potential of BCAA to identify "at risk" obese individuals. Further, the fasting levels of 14:0, 16:0, and 18:0 show promise as distinct markers of fasting and/or postprandial insulin sensitivity. This highlights the added value of postprandial measurements and underscores the importance to identify "at risk" obese individuals that could benefit from tailored diet interventions.

3.2.6 Supporting Information

Study Population Characteristics (Table 3.6)

Data represented as mean ± standard error of the mean (SEM). LH, lean healthy; MHO, metabolically healthy obese; MUO, metabolically unhealthy obese; BMI, body mass index; BP, blood pressure; Total-c, total cholesterol; LDL-c, low-density lipoprotein cholesterol; HDL-c, high-density lipoprotein cholesterol; TG, triglycerides; HbA1c, glycated hemoglobin; HOMA-IR, homeostatic model assessment of insulin resistance; HOMA%B, homeostatic model assessment for β-cell function. A non-parametric

ANOVA Kruskal-Wallis followed by a post-hoc Mann-Whitney test was used to determine the significance between groups ($p < 0.05$). Adapted from Perreault et al. (2014b).

Mean Circulating Concentrations of Amino Acid and Derivatives at Fasting and T120 min Time Points (Tables 3.7 and 3.8)

Data represented as mean concentration ± SEM. LH, lean healthy; MHO, metabolically healthy obese; MUO, metabolically unhealthy obese. A non-parametric ANOVA Kruskal-Wallis followed by a post-hoc Mann-Whitney test was used to determine significance ($p < 0.05$). Significant percent postprandial change (%PP) changes are indicated in bold.

Mean Circulating Concentrations of Fatty Acids at Fasting and T120 min Time Points (Tables 3.9 and 3.10)

Data presented as mean relative percentage ± SEM. LH, lean healthy; MHO, metabolically healthy obese; MUO, metabolically unhealthy obese. A non-parametric ANOVA Kruskal-Wallis followed by a post-hoc Mann-Whitney test was used to determine significance ($p < 0.05$). Significant percent postprandial change (%PP) changes are indicated in bold.

3.2.7 Acknowledgments

The authors express their gratitude to Prof. Anthony Clarke for his help with the hydrolysis of proteins from the meal, and to Drs. Susan Stephenson, Andrea Buchholz, and Alison Duncan, as well as Ms. Alicia DiBattista and Ms. Michelle Saoi, for their conceptual contributions to this work. The authors thank all of the study

Table 3.6: Study population characteristics.

Parameters	Mean ± SEM			ANOVA (p-value)	Post-Hoc Mann-Whitney Group Comparison (p-value)		
	LH	MHO	MUO		LH vs. MHO	LH vs. MUO	MHO vs. MUO
Anthropometric measurements							
Number of subjects	10	10	10				
Number of men/women	3/7	3/7	3/7				
Age (yr)	51 ± 3	50 ± 4	48 ± 2	0.8418			
Weight (kg)	61.9 ± 2.8	86.2 ± 3.4	92.7 ± 6.2	0.0004	0.0001	0.0003	0.5288
Height (cm)	167 ± 3	168 ± 3	167 ± 3	0.9299			
BMI (kg/m^2)	22.1 ± 0.6	30.6 ± 1.1	33.0 ± 1.9	<0.0001	<0.0001	<0.0001	0.4813
Waist circumference (cm)	77 ± 3	98 ± 3	104 ± 5	0.0002	0.0007	0.0003	0.4359
Hip circumference (cm)	96 ± 1	109 ± 2	113 ± 4	0.0002	0.0003	0.0006	0.5960
Waist-to-hip ratio	0.80 ± 0.02	0.90 ± 0.03	0.92 ± 0.02	0.0075	0.0172	0.0046	0.5787
Fat mass (%)	27.4 ± 2.7	39.8 ± 2.4	39.3 ± 2.4	0.0086	0.0039	0.0115	0.9397
Fat mass (kg)	16.7 ± 1.6	34.1 ± 2.4	36.4 ± 3.4	<0.0001	<0.0001	<0.0001	0.7959
Lean mass (%)	72.6 ± 2.7	60.2 ± 2.4	60.8 ± 2.4	0.0086	0.0039	0.0115	0.9397
Lean mass (kg)	45.1 ± 3.1	52.0 ± 3.2	56.2 ± 4.4	0.1071			
Clinical measurements							
Systolic BP (mmHg)	118 ± 4	128 ± 5	128 ± 4	0.1063			
Diastolic BP (mmHg)	75 ± 3	82 ± 2	82 ± 2	0.0664			
Total-c (mmol/L)	4.43 ± 0.30	4.26 ± 0.32	5.34 ± 0.23	0.0169	0.7054	0.0355	0.0073
LDL-c (mmol/L)	2.52 ± 0.25	2.39 ± 0.30	3.27 ± 0.19	0.0401	0.6842	0.0433	0.0232
HDL-c (mmol/L)	1.57 ± 0.08	1.17 ± 0.12	1.04 ± 0.05	0.0018	0.0256	0.0004	0.3634
Total-c/HDL ratio	2.85 ± 0.16	3.81 ± 0.22	5.17 ± 0.18	<0.0001	0.0073	0.0002	0.0010
TG (mmol/L)	0.77 ± 0.05	1.54 ± 0.33	2.26 ± 0.23	0.0004	0.0311	0.0002	0.0524
Fasting glucose (mmol/L)	4.5 ± 0.2	5.0 ± 0.1	5.3 ± 0.2	0.0243	0.0335	0.0171	0.5178
Fasting insulin (pmol/L)	37 ± 17	64 ± 9	118 ± 22	0.0025	0.0129	0.0030	0.0887
HbA1c (%)	5.37 ± 0.07	5.72 ± 0.08	5.73 ± 0.08	0.0051	0.0055	0.0066	0.8775
HOMA-IR	0.65 ± 0.28	1.18 ± 0.16	2.19 ± 0.42	0.0024	0.0143	0.0015	0.0892
HOMA%B	94 ± 33	106 ± 10	140 ± 18	0.0130	0.0337	0.0076	0.2176

Table 3.7: Mean circulating concentrations of amino acid and derivatives at fasting and T120 min time points.

Plasma Amino Acid and Derivatives	Fasting (mean ± SEM)			Postprandial (mean ± SEM)		
	LH	MHO	MUO	LH	MHO	MUO
Alanine (µM)	339 ± 30	446 ± 24	413 ± 34	451 ± 39	557 ± 31	456 ± 28
Arginine (µM)	82.7 ± 4.8	95.6 ± 6.0	88.5 ± 9.7	100 ± 8	108 ± 5	95.2 ± 8.2
Asparagine (µM)	56.9 ± 3.2	56.9 ± 3.1	56.2 ± 4.9	75.0 ± 6.7	64.2 ± 4.7	56.9 ± 3.7
Aspartic acid (µM)	9.97 ± 0.22	9.82 ± 0.16	10.8 ± 0.4	9.81 ± 0.19	10.0 ± 0.2	10.3 ± 0.3
Asparagine-to-aspartic acid ratio	5.73 ± 0.35	5.81 ± 0.32	5.28 ± 0.48	7.69 ± 0.74	6.46 ± 0.50	5.57 ± 0.46
Betaine (µM)	51.0 ± 4.0	44.4 ± 4.7	38.5 ± 3.4	57.1 ± 4.1	52.4 ± 4.0	45.8 ± 2.8
Carnitine (µM)	42.1 ± 4.9	72.5 ± 8.7	48.6 ± 4.3	49.9 ± 4.8	71.4 ± 6.1	54.5 ± 3.7
Acetylcarnitine (µM)	37.2 ± 2.3	40.6 ± 4.4	33.2 ± 3.7	18.0 ± 1.6	22.4 ± 1.6	17.9 ± 1.5
Carnitine-to-acetylcarnitine ratio	1.14 ± 0.13	1.79 ± 0.09	1.60 ± 0.19	2.81 ± 0.21	3.18 ± 0.14	3.18 ± 0.26
Propylcarnitine (µM)	0.65 ± 0.07	0.84 ± 0.08	0.72 ± 0.07	0.73 ± 0.08	1.00 ± 0.07	0.84 ± 0.06
Citrulline (µM)	27.6 ± 2.0	30.7 ± 1.6	29.3 ± 2.0	22.0 ± 1.5	24.7 ± 1.5	22.4 ± 1.2
Creatinine (µM)	64.0 ± 4.3	73.6 ± 5.9	64.0 ± 4.5	66.2 ± 5.8	70.0 ± 5.1	63.3 ± 3.9
Creatine (µM)	32.7 ± 3.8	49.4 ± 5.5	42.5 ± 6.2	45.1 ± 5.0	60.2 ± 7.0	48.5 ± 4.6
Cystine (µM)	46.0 ± 2.3	61.5 ± 5.5	52.5 ± 3.0	49.9 ± 2.1	58.6 ± 3.1	49.5 ± 3.2
Cysteine-glutathione disulfide (µM)	9.3 ± 1.1	8.06 ± 0.62	7.97 ± 0.77	7.76 ± 0.40	7.52 ± 0.35	6.86 ± 0.21
Deoxycarnitine (µM)	1.37 ± 0.14	1.83 ± 0.16	1.27 ± 0.13	1.28 ± 0.14	1.73 ± 0.17	1.23 ± 0.13
Dimethylglycine (µM)	8.44 ± 0.25	8.17 ± 0.16	9.01 ± 0.29	8.67 ± 0.27	8.33 ± 0.21	8.89 ± 0.21
γ-Amino-butyric acid (µM)	6.70 ± 0.13	6.34 ± 0.09	6.76 ± 0.23	6.35 ± 0.14	6.16 ± 0.07	6.40 ± 0.20
Glutamine (µM)	537 ± 16	545 ± 20	510 ± 30	587 ± 27	570 ± 31	473 ± 20
Glutamic acid (µM)	33.7 ± 4.1	40.4 ± 3.9	51.5 ± 5.0	34.5 ± 3.7	36.2 ± 2.4	49.3 ± 7.5
Gutamine-to-glutamic acid ratio	17.5 ± 1.8	14.7 ± 1.6	11.2 ± 1.6	18.9 ± 2.4	16.5 ± 1.7	11.4 ± 1.4
Glycine (µM)	209 ± 18	204 ± 18	165 ± 14	226 ± 21	207 ± 21	153 ± 13
Histidine (µM)	85.1 ± 3.4	95.8 ± 5.8	90.9 ± 5.3	95.3 ± 3.9	98.2 ± 6.4	90.5 ± 3.8
Hydroxyproline (µM)	15.1 ± 0.5	18.0 ± 0.6	15.9 ± 0.9	17.9 ± 0.6	19.50 ± 0.47	17.5 ± 0.8
Hydroxyxanthine (µM)	4.14 ± 0.93	3.97 ± 0.29	3.84 ± 0.43	2.79 ± 0.35	2.54 ± 0.20	2.79 ± 0.19
Isoleucine (µM)	49.2 ± 2.6	59.2 ± 3.3	62.2 ± 4.4	64.6 ± 3.4	70.4 ± 3.5	74.7 ± 4.7
Kynurenic acid (µM)	3.03 ± 0.08	2.99 ± 0.08	3.17 ± 0.05	2.98 ± 0.10	2.95 ± 0.10	3.13 ± 0.05
Leucine (µM)	96.5 ± 6.0	117.0 ± 7.0	121 ± 12	114 ± 7	125 ± 6	129 ± 9
Lysine (µM)	129 ± 7	179 ± 11	151 ± 13	144 ± 7	179 ± 8	158 ± 13
Methionine (µM)	24.4 ± 1.0	27.8 ± 1.6	27.1 ± 1.9	30.5 ± 2.2	30.5 ± 1.3	28.4 ± 1.6
3-Methylhistidine (µM)	21.9 ± 1.3	29.7 ± 2.7	22.3 ± 1.7	20.5 ± 1.1	26.1 ± 2.1	20.9 ± 1.7
Ornithine (µM)	43.3 ± 2.1	56.0 ± 4.9	50.2 ± 4.8	52.2 ± 2.6	63.5 ± 4.4	56.9 ± 4.8
Phenylalanine (µM)	52.6 ± 1.7	60.9 ± 3.3	58.8 ± 3.4	66.6 ± 2.7	70.3 ± 2.1	67.5 ± 2.1
Proline (µM)	130 ± 11	163 ± 10	153 ± 15	215 ± 12	254 ± 11	226 ± 15
Serine (µM)	120 ± 6	123 ± 12	103 ± 7	139 ± 10	124 ± 11	96.7 ± 5.8
Threonine (µM)	133 ± 11	147 ± 11	123 ± 11	147 ± 15	146 ± 11	117 ± 8
Tryptophan (µM)	53.7 ± 2.6	58.7 ± 2.9	57.1 ± 3.4	59.0 ± 3.2	61.9 ± 2.8	57.4 ± 2.6
Tyrosine (µM)	56.4 ± 3.3	71.3 ± 3.8	73.7 ± 7.3	62.9 ± 3.6	77.3 ± 2.5	74.9 ± 5.3
Valine (µM)	210 ± 16	267 ± 15	256 ± 22	238 ± 17	284 ± 18	296 ± 28

Table 3.8: Mean circulating concentrations of amino acid and derivatives at fasting and T120 min time points (continued).

Plasma Amino Acid and Derivatives	% Postprandial Change (%PP)			ANOVA (p-value)	Post-Hoc Mann-Whitney Group Comparison (p-value)		
	LH	MHO	MUO		LH vs. MUO	LH vs. MHO	MHO vs. MUO
Alanine	34.7 ± 8.1	26.2 ± 7.3	14.2 ± 8.1	0.19			
Arginine	20.3 ± 3.1	15.3 ± 5.4	11.4 ± 6.6	0.21			
Asparagine	30.9 ± 6.8	13.5 ± 6.5	5.2 ± 7.1	0.03	0.01	0.06	0.29
Aspartic acid	−1.4 ± 1.7	2.2 ± 1.6	−2.9 ± 3.2	0.36			
Asparagine-to-aspartic acid ratio	33.5 ± 8.0	11.9 ± 7.6	8.3 ± 6.2	0.07			
Betaine	12.8 ± 3.5	22.3 ± 6.9	29 ± 16	0.54			
Carnitine	21.6 ± 6.4	2.3 ± 4.9	14.4 ± 4.0	0.11			
Acetylcarnitine	−51.7 ± 2.3	−41.8 ± 4.1	−43.4 ± 3.5	0.12			
Carnitine-to-acetylcarnitine ratio	155 ± 15	82 ± 13	113 ± 21	0.02	0.06	0.01	0.26
Propylcarnitine	15 ± 10	23.0 ± 6.2	21.7 ± 6.2	0.34			
Citrulline	−19.8 ± 2.4	−18.5 ± 4.6	−21.9 ± 4.1	0.85			
Creatinine	2.6 ± 4.1	−3.6 ± 3.9	0.3 ± 4.3	0.49			
Creatine	38.6 ± 3.7	22.7 ± 4.3	22.4 ± 8.7	0.06			
Cystine	9.2 ± 3.2	−2.3 ± 4.0	−5.7 ± 2.8	0.02	<0.01	0.04	0.52
Cysteine-glutathione disulfide	−11.2 ± 7.7	−4.0 ± 5.1	−9.0 ± 5.8	0.46			
Deoxycarnitine	−6.9 ± 3.8	−5.6 ± 4.2	−2.7 ± 4.5	0.74			
Dimethylglycine	2.8 ± 2.2	1.9 ± 1.3	−0.8 ± 2.5	0.68			
γ-Amino-butyric acid	−5.1 ± 1.4	−2.7 ± 1.2	−5.2 ± 1.3	0.38			
Glutamine	9.3 ± 4.1	4.6 ± 4.4	−6.1 ± 3.4	0.03	0.01	0.46	0.07
Glutamic acid	5.2 ± 9.1	−3.4 ± 9.6	−3.7 ± 9.3	0.83			
Gutamine-to-glutamic acid ratio	11 ± 11	21 ± 15	9 ± 15	0.77			
Glycine	7.3 ± 4.2	1.1 ± 4.0	−5.8 ± 4.3	0.12			
Histidine	12.2 ± 2.1	3.4 ± 5.3	1.1 ± 3.7	0.06			
Hydroxyproline	18.9 ± 3.3	8.8 ± 3.2	11.6 ± 4.1	0.13			
Hydroxyxanthine	−22 ± 11	−34.7 ± 4.2	−22.4 ± 6.6	0.18			
Isoleucine	31.9 ± 3.8	21.2 ± 6.5	23.3 ± 8.0	0.34			
Kynurenic acid	−1.5 ± 2.0	−1.6 ± 2.2	−0.9 ± 1.5	0.85			
Leucine	19.6 ± 2.4	8.4 ± 5.7	12.1 ± 9.7	0.25			
Lysine	11.6 ± 3.4	1.6 ± 3.3	4.8 ± 3.8	0.15			
Methionine	25.0 ± 7.7	11.5 ± 4.5	6.8 ± 5.9	0.23			
3-Methylhistidine	−6.0 ± 1.7	−10.3 ± 3.9	−6.4 ± 1.5	0.90			
Ornithine	21.4 ± 5.9	15.8 ± 5.7	15.3 ± 4.4	0.73			
Phenylalanine	26.9 ± 4.5	17.1 ± 4.6	16.7 ± 5.1	0.30			
Proline	71 ± 11	59.2 ± 8.7	52.8 ± 8.6	0.54			
Serine	14.6 ± 4.2	3.0 ± 5.5	−4.3 ± 5.0	0.03	0.03	0.02	0.36
Threonine	9.5 ± 3.4	0.1 ± 4.4	−2.9 ± 4.4	0.10			
Tryptophan	9.9 ± 3.0	6.5 ± 4.4	1.6 ± 2.9	0.27			
Tyrosine	12.3 ± 4.1	9.9 ± 4.2	4.3 ± 6.0	0.80			
Valine	13.8 ± 3.1	7.0 ± 4.3	17.4 ± 8.9	0.77			

Table 3.9: Mean circulating concentrations of fatty acids at fasting and T120 min time points.

Serum fatty acids (%)	Fasting (mean ± SEM)			Postprandial (mean ± SEM)		
	LH	MHO	MUO	LH	MHO	MUO
Myristic acid (14:0)	0.62 ± 0.04	0.84 ± 0.08	1.20 ± 0.11	0.84 ± 0.04	0.93 ± 0.09	1.30 ± 0.10
Pentadecanoic acid (15:0)	0.23 ± 0.01	0.20 ± 0.01	0.23 ± 0.01	0.25 ± 0.01	0.24 ± 0.02	0.23 ± 0.01
Palmitic acid (16:0)	20.2 ± 0.4	21.2 ± 0.6	22.20 ± 0.64	21.7 ± 0.4	22.0 ± 0.5	23.7 ± 0.7
Stearic acid (18:0)	7.71 ± 0.18	7.50 ± 0.27	6.73 ± 0.18	8.02 ± 0.11	7.64 ± 0.20	6.88 ± 0.16
Nonadecanoic acid (19:0)	0.05 ± 0.01	0.07 ± 0.01	0.05 ± 0.01	0.17 ± 0.04	0.13 ± 0.02	0.08 ± 0.01
Arachidic acid (20:0)	0.29 ± 0.03	0.20 ± 0.02	0.19 ± 0.01	0.26 ± 0.04	0.16 ± 0.02	0.17 ± 0.01
Behenic acid (22:0)	0.46 ± 0.03	0.39 ± 0.05	0.35 ± 0.04	0.38 ± 0.02	0.31 ± 0.07	0.29 ± 0.04
Lignoceric acid (24:0)	2.65 ± 0.31	2.37 ± 0.31	1.80 ± 0.10	1.75 ± 0.28	1.85 ± 0.27	1.58 ± 0.10
Myristoleic acid (14:1n-5)	0.06 ± 0.01	0.08 ± 0.01	0.11 ± 0.01	0.05 ± 0.01	0.08 ± 0.01	0.08 ± 0.01
Palmitoleic acid (16:1n-7)	1.57 ± 0.13	2.16 ± 0.25	2.33 ± 0.14	1.42 ± 0.11	2.02 ± 0.22	2.23 ± 0.10
Heptadecanoic acid (17:1n-7)	0.14 ± 0.01	0.19 ± 0.02	0.20 ± 0.01	0.12 ± 0.02	0.11 ± 0.02	0.15 ± 0.03
Vaccenic acid (18:1n-7)	1.81 ± 0.07	1.88 ± 0.06	1.77 ± 0.04	1.89 ± 0.06	1.93 ± 0.07	1.83 ± 0.05
Oleic acid (18:1n-9)	19.6 ± 0.5	21.7 ± 1.2	23.7 ± 0.4	21.3 ± 0.7	22.50 ± 1.29	25.4 ± 0.5
cis-Nonadecanoic acid (19:1n-9)	0.28 ± 0.05	0.27 ± 0.04	0.15 ± 0.01	0.27 ± 0.07	0.25 ± 0.06	0.16 ± 0.03
Erucic acid (22:1n-9)	0.56 ± 0.06	0.58 ± 0.07	0.36 ± 0.03	0.53 ± 0.17	0.46 ± 0.14	0.25 ± 0.08
Nervonic acid (24:1n-9)	0.07 ± 0.01	0.06 ± 0.01	0.04 ± 0.01	0.05 ± 0.01	0.04 ± 0.01	0.04 ± 0.01
Linoleic acid (18:2n-6)	30.4 ± 0.8	25.7 ± 0.9	26.5 ± 1.0	28.7 ± 0.4	25.1 ± 1.0	24.8 ± 1.0
γ-linolenic acid (18:3n-6)	0.36 ± 0.04	0.60 ± 0.06	0.55 ± 0.04	0.40 ± 0.03	0.61 ± 0.07	0.49 ± 0.04
α-linolenic acid (18:3n-3)	0.70 ± 0.05	0.71 ± 0.05	0.95 ± 0.05	0.82 ± 0.04	0.74 ± 0.06	1.01 ± 0.05
Eicosadienoic acid (20:2n-6)	0.16 ± 0.02	0.13 ± 0.01	0.12 ± 0.01	0.16 ± 0.02	0.19 ± 0.03	0.20 ± 0.02
Dihomo-γ-linolenic acid (20:3n-6)	1.47 ± 0.14	1.79 ± 0.09	1.64 ± 0.11	1.30 ± 0.10	1.57 ± 0.10	1.37 ± 0.10
Arachidonic acid (20:4n-6)	7.38 ± 0.52	8.31 ± 0.66	6.07 ± 0.36	6.72 ± 0.40	8.12 ± 0.66	5.31 ± 0.30
Eicosapentaenoic acid (20:5n-3)	0.95 ± 0.16	0.91 ± 0.09	0.85 ± 0.11	0.73 ± 0.11	0.90 ± 0.10	0.78 ± 0.09
Adrenic acid (22:4n-6)	0.22 ± 0.02	0.20 ± 0.02	0.15 ± 0.02	0.18 ± 0.01	0.17 ± 0.01	0.13 ± 0.01
Docosapentaenoic acid (22:5n-3)	0.69 ± 0.03	0.64 ± 0.05	0.59 ± 0.02	0.61 ± 0.03	0.60 ± 0.04	0.56 ± 0.03
Docosapentaenoic acid (22:5n-6)	0.21 ± 0.01	0.20 ± 0.01	0.18 ± 0.01	0.17 ± 0.01	0.19 ± 0.01	0.17 ± 0.01
Docosahexaenoic acid (22:6n-3)	0.77 ± 0.04	0.66 ± 0.07	0.52 ± 0.06	0.77 ± 0.05	0.71 ± 0.07	0.50 ± 0.06

Table 3.10: Mean circulating concentrations of fatty acids at fasting and T120 min time points (continued).

Serum fatty acids (%)	% Postprandial Change (%PP)						ANOVA (p-value)	Post-Hoc Mann-Whitney Group Comparison (p-values)		
	LH		MHO		MUO			LH vs. MUO	LH vs. MHO	MHO vs. MUO
Myristic acid (14:0)	38	± 12	12	± 6	10	± 4	0.15			
Pentadecanoic acid (15:0)	8	± 5	15	± 11	−1	± 4	0.52			
Palmitic acid (16:0)	7	± 1	3	± 1	6	± 1	0.11			
Stearic acid (18:0)	5	± 1	2	± 2	2	± 2	0.55			
Nonadecanoic acid (19:0)	164	± 60	170	± 58	71	± 31	0.36			
Arachidic acid (20:0)	−3	± 21	−21	± 6	−9	± 7	0.48			
Behenic acid (22:0)	−10	± 7	−12	± 15	−9	± 10	0.96			
Lignoceric acid (24:0)	−24	± 10	−12	± 11	−12	± 2	0.49			
Myristoleic acid (14:1n-5)	−1	± 8	−1	± 7	−13	± 10	0.38			
Palmitoleic acid (16:1n-7)	−11	± 2	−5	± 3	−3	± 1	0.04	0.04	0.02	0.73
Heptadecanoic acid (17:1n-7)	−11	± 14	−18	± 15	−40	± 12	0.22			
Vaccenic acid (18:1n-7)	1	± 1	2	± 2	3	± 1	0.46			
Oleic acid (18:1n-9)	8	± 2	3	± 2	7	± 1	0.79			
cis-Nonadecanoic acid (19:1n-9)	19	± 25	28	± 43	10	± 25	0.84			
Erucic acid (22:1n-9)	−13	± 19	−8	± 28	−32	± 18	0.84			
Nervonic acid (24:1n-9)	−22	± 11	−24	± 8	11	± 32	0.94			
Linoleic acid (18:2n-6)	−4.4	± 2.0	−2.2	± 1.0	−6	± 1	0.04	0.66	0.13	0.01
γ-linolenic acid (18:3n-6)	9	± 11	1	± 5	−12	± 2	0.04	0.02	0.90	0.04
α-linolenic acid (18:3n-3)	19	± 9	4.8	± 4.0	7	± 1	0.78			
Eicosadienoic acid (20:2n-6)	31	± 15	74	± 43	68	± 29	0.74			
Dihomo-γ-linolenic acid (20:3n-6)	−16	± 2	−12	± 2	−16	± 2	0.21			
Arachidonic acid (20:4n-6)	−11	± 2	−1	± 5	−12	± 2	0.04	0.60	0.06	0.02
Eicosapentaenoic acid (20:5n-3)	−9	± 5	−1	± 3	−6	± 3	0.21			
Adrenic acid (22:4n-6)	−11	± 10	−10	± 9	−4	± 9	0.83			
Docosapentaenoic acid (22:5n-3)	−12	± 2	−4	± 2	−5	± 4	0.11			
Docosapentaenoic acid (22:5n-6)	−13	± 5	−3	± 6	−5	± 6	0.49			
Docosahexaenoic acid (22:6n-3)	1	± 7	11	± 10	−4	± 6	0.46			

participants, as well as the phlebotomists in the Human Nutraceutical Research Unit at the University of Guelph.

Chapter 4

Metabolomic Assessment of Treatment Responses to Protein Supplementation during Caloric Restriction

Protein supplementation has been shown in recent studies to attenuate the loss of muscle mass during weight-loss interventions. In a study by Hector et al. (2015), it has been shown that whey protein is more effective than soy in stimulating myofibrillar protein synthesis in a 14-day intervention with a controlled hypoenergetic diet. Unexpectedly, however, no significant difference was found in loss of fat mass between groups consuming carbohydrate, soy, and whey supplementation, respectively. Using CE-MSI-Q-TOF-MS, the objective of the current study was to investigate differences in the impact of protein supplementation during short-term caloric restriction by measuring changes in the urine metabolome involving a cohort of healthy yet overweight/obese individuals following a 2-week intervention where participants were

supplemented with whey or soy protein, or maltodextrin (carbohydrate control).

4.1 Introduction

Obesity is well-known to be a major risk factor for T2D, heart disease, and stroke (Eckel et al., 2004). Due to its increasing prevalence worldwide and its burden on economy and society, obesity is a global health concern which has made its treatment and prevention the focus of many public health efforts around the world (Hu et al., 2011; Institute of Medicine, 2012, 2014; Joint WHO/FAO Expert Consultation, 2003; Moffat and Thrasher, 2016; Newman et al., 2016). Although various weight-loss treatment options are available, including lifestyle changes, invasive bariatric surgery, and usage of weight-loss drugs associated with severe side-effects, energy-restricted dietary interventions are typically the first approach taken towards the treatment of obesity by controlling energy intake (Langeveld and DeVries, 2015).

Many studies of various dietary interventions have been shown to aid in weight loss while improving cardiometabolic risk factors (Atallah et al., 2014; Bazzano et al., 2014; Gardner et al., 2016; Johnston et al., 2014; Sacks et al., 2009; Shai et al., 2008; Tay et al., 2014). However, during diet-induced weight-loss periods, it is well known that the reduction in body weight is accompanied by loss of strength and muscle mass (Ballor et al., 1988). The loss of lean mass is undesirable, particularly in active and in elderly individuals, since it may compromise the metabolic benefits gained from the loss of excess body fat. Dietary protein intake has been shown to stimulate muscle protein synthesis (Koopman et al., 2007; Moore et al., 2009; Paddon-Jones and Rasmussen, 2009) and several studies have supported that the loss of muscle mass can be mitigated by increased protein consumption during energy restriction (Leidy et al., 2007; Pasiakos et al., 2013; Soenen et al., 2013). Furthermore, diets with higher

protein content (> 20% energy intake) have resulted in greater weight loss (Clifton et al., 2008; Due et al., 2004; Johnstone et al., 2008; Weigle et al., 2005), due to the greater thermogenic effect of protein than fat and carbohydrate (Halton and Hu, 2004; Matarese and Pories, 2014), and the longer duration of satiety, helping to suppress appetite (Astrup, 2005; Johnstone et al., 2008; Liu et al., 2015; Weigle et al., 2005; Westerterp-Plantenga et al., 2009). While some studies were unable to confirm the contribution of increased protein intake to the preservation of lean body mass (Backx et al., 2016; Mojtahedi et al., 2011; Parker et al., 2002), other studies have shown that differential outcomes on body compositional changes, energy expenditure, or cellular processes can result from differences in the source of protein ingested (Baer et al., 2011; Belobrajdic et al., 2004; Hector et al., 2015; Mikkelsen et al., 2000), which may be marine-, plant-, dairy-, or animal-based, suggesting that isoenergetic diets are not necessarily isometabolic (Matarese and Pories, 2014).

Recently, the effects of soy and whey protein supplementations on changes in body weight, composition, and rates of myofibrillar protein synthesis (MPS) have been compared and investigated (Baer et al., 2011; Hector et al., 2015). In the study of Baer et al., it was found that postprandial body weight and fat mass of healthy free-living overweight and obese individuals consuming whey protein were significantly lower when compared to the control group after 23 weeks of supplementation, while no differences were observed between soy and control groups. Meanwhile, Hector et al. concluded that the reduction in the postprandial rate of MPS was more attenuated with whey supplementation, than with soy or the control groups, in overweight and obese adults after a 14-day energy restriction, suggesting that whey confers greater ability than soy towards the maintenance of muscle mass.

Despite both studies revealing whey supplementation yielding greater positive outcomes associated with loss of fat mass and preservation of lean mass relative to soy,

the metabolic processes responsible for the differential responses between the different protein sources have not been reported and therefore, were investigated in the present study. Metabolomic analysis of urine has proven to be a useful tool for the study and understanding of diet-related influences (Hjerpsted et al., 2014; Ross et al., 2013; Vázquez-Fresno et al., 2015), including the intake of high-protein diets in rats (Mu et al., 2015) and in humans (Rasmussen et al., 2012), in which the latter two studies assessed the metabolic effect of different levels of protein content (i.e., high vs. low), but not between different protein sources. Through untargeted metabolomic analysis of urine samples collected from the study of Hector et al. (2015), the current study aims to gain insight into the underlying differences in metabolism between different sources of protein, namely soy and whey supplementations, in order to understand why whey protein might exhibit differences versus soy protein and control in weight loss and conservation of MPS. The use of untargeted metabolomics allows for a more comprehensive approach than targeted analysis such that both known and unknown metabolites affected by the dietary intervention are measured to better distinguish metabolic profiles between diet treatments.

4.2 Materials and Methods

4.2.1 Subject Cohort

A total of 50 participants were recruited through poster and newspaper advertisements and were screened for meeting the following inclusion criteria as previously described (Hector et al., 2015): 1) BMI of 28–50 kg/m^2; 2) 35–65 years old; 3) non-smoker; 4) non-diabetic; and 5) otherwise healthy based on responses to a standard medical screening questionnaire. Nine subjects declined participation before the trial started and one

subject dropped out during the trial for personal reasons. All participants signed a consent form, which informed them of the purpose of the study, the experimental procedures, and potential risks. The study was approved by the Hamilton Health Sciences Research Ethics Board (now known as Hamilton Integrated Research Ethics Board, HiREB) and research protocol was in accordance with standards set by the Canadian Tri-Council Policy Statement (Canadian Institutes of Health Research et al., 2010) on the use of human subjects in research. None of the participants were enrolled in a weight-loss or exercise program at the time of recruitment. The basic characteristics of the participants are outlined in Table 4.6 (Supporting Information). To minimize the effect of confounding variables, participants were matched by age, sex, and BMI in all groups.

4.2.2 Dietary Intervention and Study Design

As described in Hector et al. (2015), standardized pre-packaged hypoenergetic weight-loss diets, with twice daily (mid-morning and mid-afternoon) supplementation of whey protein (WHY, 27 g, IsoChill 8000), soy protein (SOY, 26 g, SoyPro950M), or an isoenergetic amount of carbohydrate (CHO, 25 g, GlobePlus) in the form of flavoured beverages (i.e., supplements mixed with equal amounts of acesulfame potassium to enhance palatability, in addition to 15 g of Nesquik®) were randomly assigned to participants in a double-blind procedure, to be consumed over a 14 day period. Based on the estimated energy requirements of each participant, the meals and supplements were designed to provide a caloric deficit of −750 kcal/day. Energy requirements were estimated by a registered dietician from baseline height and body mass measurements using the Mifflin-St. Jeor equation (Frankenfield et al., 2005), along with an appropriate activity factor, calculated based on the participant's activity log. Prior to the start of

the study, participants completed a 3-day food journal (2 weekdays and 1 weekend day) to provide an estimate of their regular dietary intake. Three days prior to the study, participants consumed pre-packaged weight maintenance meals (Copper County Foods) that supplied 100% of the estimated energy requirements, at a protein intake rate of 1 g/kg/day. Participants underwent an overnight fast prior to consuming Day 1 of the standardized breakfast and were instructed not to consume any vitamin or mineral supplements, nor alcohol, throughout the duration of the study. However, participants were asked to maintain their physical activity level throughout the entire study duration.

4.2.3 Urine Collection and Sample Preparation

Single-spot mid-stream urine samples from the subjects' first urinary void of the day were obtained from 40 participants (19 males, 21 females) at two time points: 1) the morning prior to consumption of the Day 1 standardized breakfast (Pre) and 2) the morning following the 14-day weight loss intervention period (Post). Urine samples were collected in 3-litre sterile polypropylene urine collection containers (Fisher Scientific) and participants were instructed to store the samples at 4 °C until the samples could be transferred to a −80 °C freezer. The time between sample collection and storage in a −80 °C freezer was typically within a 3-hour period. Frozen urine samples were thawed at room temperature, aliquoted, then re-frozen at −80 °C until analysis.

4.2.4 Chemicals and Reagents

Ultra LC-MS-grade water (H_2O), acetonitrile (ACN), and methanol (MeOH) (Caledon, Georgetown, ON, Canada), and HPLC-grade ammonium hydroxide solution

(Sigma-Aldrich, St. Louis, MO, USA) were used in the preparation of background electrolyte (BGE) solutions and sheath liquid. 3-Chloro-L-tyrosine (3-Cl-Tyr), sodium-2-naphthalene sulfonate (NMS), formic acid, ammonium acetate (NH_4Ac), sodium hydroxide (NaOH), and metabolite reference standards were purchased from Sigma-Aldrich (St. Louis, MO, USA). L-homocitrulline (Hci) was purchased from Toronto Research Chemicals Inc. (Toronto, ON, Canada). Ultra LC-MS-grade water was used in the dilution of samples and preparation of standard stock solutions.

4.2.5 Sample Preparation

Prior to analysis by CE-MS, urine samples were thawed on ice and prepared as follows: 10 µL urine samples were diluted 100-fold in H_2O with 10 µM 3-Cl-Tyr as internal standard for positive-mode analysis, and diluted 10-fold in H_2O with 10 µM NMS as internal standard for negative-mode analysis. QC samples for assessing instrument stability and analytical reproducibility were prepared by pooling equal aliquots of all urine samples. Pooled subgroup samples for purposes of preliminary screening of group-specific features were prepared by pooling together equal aliquots of samples belonging to the same group based on the supplement type (WHY, SOY, or CHO) and time point (Pre or Post). Altogether, six subgroup samples were prepared: Pre-WHY, Pre-SOY, Pre-CHO, Post-WHY, Post-SOY, and Post-CHO.

4.2.6 Capillary Electrophoresis Mass Spectrometry (CE-MS)

Sample analyses were performed on an Agilent 7100 series CE system interfaced with an orthogonal Agilent coaxial sheath liquid Jet Stream ESI source to an Agilent 6550 iFunnel Q-TOF mass spectrometer. Acquisition was operated in positive-ion and negative-ion modes separately scanning m/z 50–1700 at a rate of 1 spectra/s. The

temperature of the nitrogen nebulizing gas was 200 °C at 8 psi with drying gas at 16 L/min. Sheath gas was delivered at 3.5 L/min at 199 °C. Both the nozzle voltage and the V_{cap} were set to 2 kV, while the MS fragmentor, skimmer, and octopole radio-frequency voltages were set to 380 V, 65 V, and 750 V, respectively. MS/MS experiments were performed at 3 collisional energies (10 V, 20 V, and 40 V) in a continuous alternating scan mode. The mass scan range was 20–500 m/z at a rate of 1 spectra/s with medium isolation width of 4 m/z.

CE separations were performed on uncoated fused-silica capillaries (Polymicro Technologies, AZ, USA) with 50 µM internal diameter at 30 kV and 21 °C. Capillary lengths used were 110 cm and 135 cm for negative and positive mode analyses, respectively. Capillaries were preconditioned by rinsing with 5 capillary volumes of MeOH, 1 M NaOH, and H_2O, respectively, prior to rinsing with 10 capillary volumes of BGE at 950 mbar. The BGE consisted of 1 M formic acid, pH 1.8, in 15% v/v ACN for positive-ion mode, and 30 mM NH_4Ac, pH 10.0, for negative-ion mode. Sample injections were performed using MSI-CE-MS developed by Kuehnbaum et al. (2013), in which seven sample injections (50 mbar for 5 s) are separated by BGE spacer injections (50 mbar for 40 s). Between sample runs, capillaries were rinsed with 3 capillary volumes of BGE at 950 mbar. Sample run-time was set to 45 min.

Sheath liquid consisting of 60:40 MeOH:H_2O containing 0.1% formic acid was delivered at a flow rate of 10 µL/min via a 100:1 splitter. Three Agilent reference mass standards (purine, hexamethoxyphosphazine (HP-0321), and hexakis(2,2,3,3-tetrafluoro-propoxy)phosphazine (HP-0921)) with m/z of 121.050873, 322.048121, and 922.009798, respectively, were spiked at levels of 2 µL into 250 mL sheath liquid for real-time internal mass calibration.

4.2.7 Method Workflow

In order to ensure quality control throughout the duration of sample analyses, a unique analytical workflow based on the MSI technique was designed to allow for the assessment of instrument stability, correction of system drift, and elimination of features exhibiting poor reproducibility. As shown in Figure 4.1, the method workflow consisted of three different run types and configurations.

Subgroup Comparison Run

The first run in the method workflow consisted of the analysis of the six pooled subgroup samples, along with the incorporation of a sample blank, in order to screen for compounds exhibiting distinct differences between sample subgroup types as potential markers of interest. This was performed in triplicate, in which Pre and Post pooled samples of each subgroup sample type were analyzed in a paired fashion (i.e., Pre-SOY, Post-SOY, Pre-CHO, Post-CHO, Pre-WHY, and Post-WHY), along with an injection of a sample blank as a control, which consisted of water stored and sampled from a sterile urine container identical to the ones used in the current study for storage of urine samples. Any features that were found in the blank injection were to be removed from the list of metabolite features to be analyzed. To eliminate the possibility of unusual differences between signals being attributed to injection bias, a subgroup run with the reversed sequence of injections was performed.

QC Run and Maintaining Quality Control

The second run in the workflow is a QC run, in which a fresh pooled QC sample is injected seven times with BGE as a spacer segment between sample injections. This run is repeated once for every 8 hours of acquisition using a fresh QC sample. The

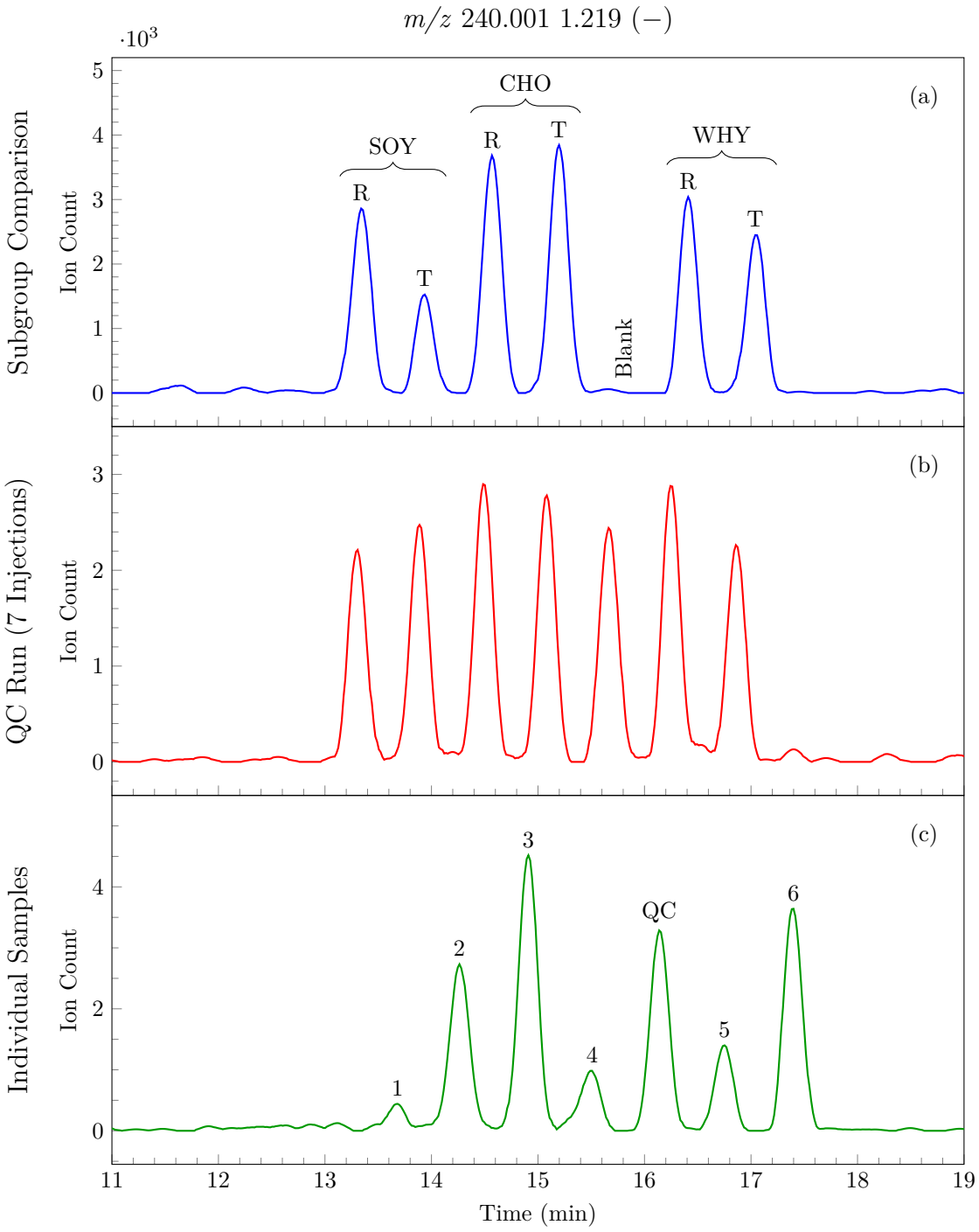


Figure 4.1: Examples of various MSI configurations used in the analytical workflow. (a) Subgroup Pre (R) and Post (T) comparison run: Rapidly screen for features that are differentially expressed between sample subgroups; (b) QC run: Correct for variations in injection volume and ionization processes; and (c) Individual sample analysis: Analyze 6 samples simultaneously along with a randomized injection of a QC sample, which corrects for instrument drift and allows for urine normalization.

purpose of the QC run is to screen for those features whose seven injections vary within acceptable technical precision limits (as determined by the overall average coefficient of variation (CV) of all detected features), in order for significantly different signal variations between samples to be confidently attributed to real biological variation, and not to variations arising from instrumental drift or ionization suppression effects.

Sample Analysis Run

Individual samples (80 samples, derived from 40 participants × 2 time points) were analyzed in randomized order using the typical 7-injection configuration of MSI-CE-MS, in which 6 samples and a pooled QC sample were analyzed simultaneously per sample run. In order to correct for system measurement bias that may exist for specific injection positions, the QC sample position was randomized through positions 1 to 7 within each run and served a three-fold purpose allowing for 1) assessment of analytical reproducibility; 2) correction of system drift; and 3) urine sample normalization.

4.2.8 Untargeted Feature Extraction

Untargeted feature extraction was performed on pooled subgroup samples acquired in both positive and negative ion modes using the MFE algorithm in Mass Hunter Qualitative Analysis (Agilent Technologies Inc.) to compile a list of analytes with their corresponding m/z values to be assessed for reproducibility within the QC run, and subsequently extracted in the analysis of the individual samples. In MFE, a feature was defined as a mass spectral peak with a SNR of at least 10 and a peak height of at least 100 counts.

Interpretation and assignment of MS/MS spectra were guided by MassHunter MSC software (Agilent Technologies Inc.), in which accurate mass MS/MS fragment ions

were correlated to probable molecular formulas and structures. Proposed molecular formulas were searched against referenced online structures in PubChem (Kim et al., 2016), which were assigned correlation scores to describe the degree of match with the MS/MS spectrum. Identities of those candidate structures with high correlation scores, for which standards were readily available, were confirmed by either spiking samples with the authentic standards and/or comparison of the MS/MS spectra of the sample to those of the reference.

4.2.9 Urine Sample Normalization & Correction of System Drift

In order to correct for differences in urine concentration between samples that would otherwise lead to unwanted sample-to-sample variation, urine samples were normalized using three different methods, which were compared in this study. Urine metabolite levels were normalized relative to measured creatinine levels and osmolality measurements, in addition to normalization using the Probabilistic Quotient Normalization (PQN) method (Dieterle et al., 2006). Osmolality measurements were made using The Advanced Micro-Osmometer 3300 (Advanced Instruments Inc., MA, USA). For the PQN method, a pooled QC urine sample, which was analyzed simultaneously with six samples in each run using the MSI-CE-MS technique, was used as the reference sample. PQN was performed by first dividing the signal of each metabolite in a sample by the corresponding signal in the QC sample, which was analyzed in the same run as the sample of interest. From the resulting distribution of quotients, the median value of the distribution was taken as the correction factor, also known as the probabilistic quotient (PQ). All of the metabolite signals in the sample were then normalized by dividing the signals by the PQ. Since urine samples were analyzed in both positive

and negative ion modes, the average values of creatinine and the PQ determined from the two modes were used for normalization purposes.

In addition to the pooled QC sample allowing for assessment of system stability and PQN of urine concentration over different sample runs, correction of system drift was also achieved with the presence of the QC samples in each sample run. All of this was accomplished in a single step with PQN. While the concept of normalization to a reference or QC sample for correction of instrument drift has previously been shown to be possible (van der Greef et al., 2007) by fitting a low-order nonlinear locally weighted smoothed scatterplot (LOESS) to QC data in the order of their injection, which is then subsequently used as a correction curve to which the total data set is normalized, normalization to QC samples in this manner requires an additional step prior to the normalization of urine concentration. With MSI, however, the within-run QC sample served as a single reference sample, thereby, allowing for the simultaneous normalization of urine concentration and correction of system drift to be combined in one step.

4.2.10 Statistical Analyses

Initial Data Exploration

Following the elimination of irreproducible features and normalization by PQN, multivariate analysis by principle component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were performed on auto-scaled data transformed by the generalized logarithm (glog) using the web-based MetaboAnalyst 3.0 (Xia and Wishart, 2016) for initial data exploration and identification of discriminatory urinary metabolites associated with dietary intervention. PCA was used to examine the clustering of the QC samples relative to the individual Pre and Post samples, in order

to assess overall instrument stability and method robustness, while multi-level PLS-DA using fold-change values (i.e., Post/Pre ratio) was used to investigate potential markers discriminating between supplementation types. Variable importance to the projection (VIP) and non-parametric Kruskal-Wallis tests were used preliminarily to highlight any features that exhibited strong Post/Pre differences between supplementation types, before checking whether the dataset satisfied the multiple assumptions that are required in parametric statistical testing by one-way ANOVA.

Statistical Data Filtering and Analysis

In order to identify metabolites that were consistently ranked high, regardless of transformation and scaling procedures, PLS-DA and random forest (RF) analysis using MetaboAnalyst were conducted on Post/Pre ratios from various data sets (with/without transformation by the glog function and with/without data scaling (Pareto scaling)). Metabolites initially considered important possessed PLS-DA VIP values greater than 1.0 in the first two components, and ranked within the top 5 with respect to PLS-DA coefficient scores. For RF analysis, a combination of 7 predictors in each node with 500 trees grown was selected, and the top 10 ranking metabolites, based on their contributions to classification accuracy, were also considered important. Features which satisfied the above criteria were tabulated to generate a pre-filtered list of 67 metabolites. Since extreme skewness and kurtosis were observed in more than 70% of the metabolites in non-transformed data, it was, thus, deemed necessary that glog transformation be performed, which, thereby, reduced the observed skewness to only 20% of the metabolites. Following glog transformation, compounds that violated assumptions of normality and/or multicollinearity, possessed VIP scores less than 1.5, or were known to be exogenous dietary markers, were excluded from analysis, which resulted in a refined list of 20 metabolites. The order of the various data filtering

procedures leading up to statistical analysis is illustrated in Figure 4.2.

In order to reduce the number of remaining metabolites to a size smaller than the number of subjects present in each treatment group, which is a requirement of one-way multivariate analysis of variance (MANOVA), McCabe analysis (Huberty and Olejnik, 2006; McCabe, 1975) was performed using an algorithm written in GNU Octave. The McCabe algorithm computed Wilks' Λ for all possible combinations of the top 20 metabolites in subset sizes ranging from 1 to 11, in which the Wilks' Λ statistic represents the proportion of unexplained variance. The algorithm only outputted the top 10 best metabolite combinations exhibiting the lowest Wilks' Λ for each subset size. Since the majority of metabolites were unknown, the choice of subset used in the one-way MANOVA was determined based on three criteria: 1) the proportion of known or putatively identified metabolites present in the combination; 2) the a priori relevance of some known metabolites in the discrimination of the groups; and 3) the optimal balance between a low Wilks' Λ and the lowest number of metabolite variables, which was determined from the plot of lowest Wilks' Λ vs. subset size (Figure 4.17, Supporting Information), in order to achieve parsimony, or "simplicity of description" (Huberty and Olejnik, 2006). Prior to performing a one-way MANOVA on the selected subset of metabolites, using SPSS Inc.'s Predictive Analytics Software (PASW Statistics 18), assumptions of independence of observations, linearity, multivariate outliers, and homogeneity of variance-covariance matrices, were checked.

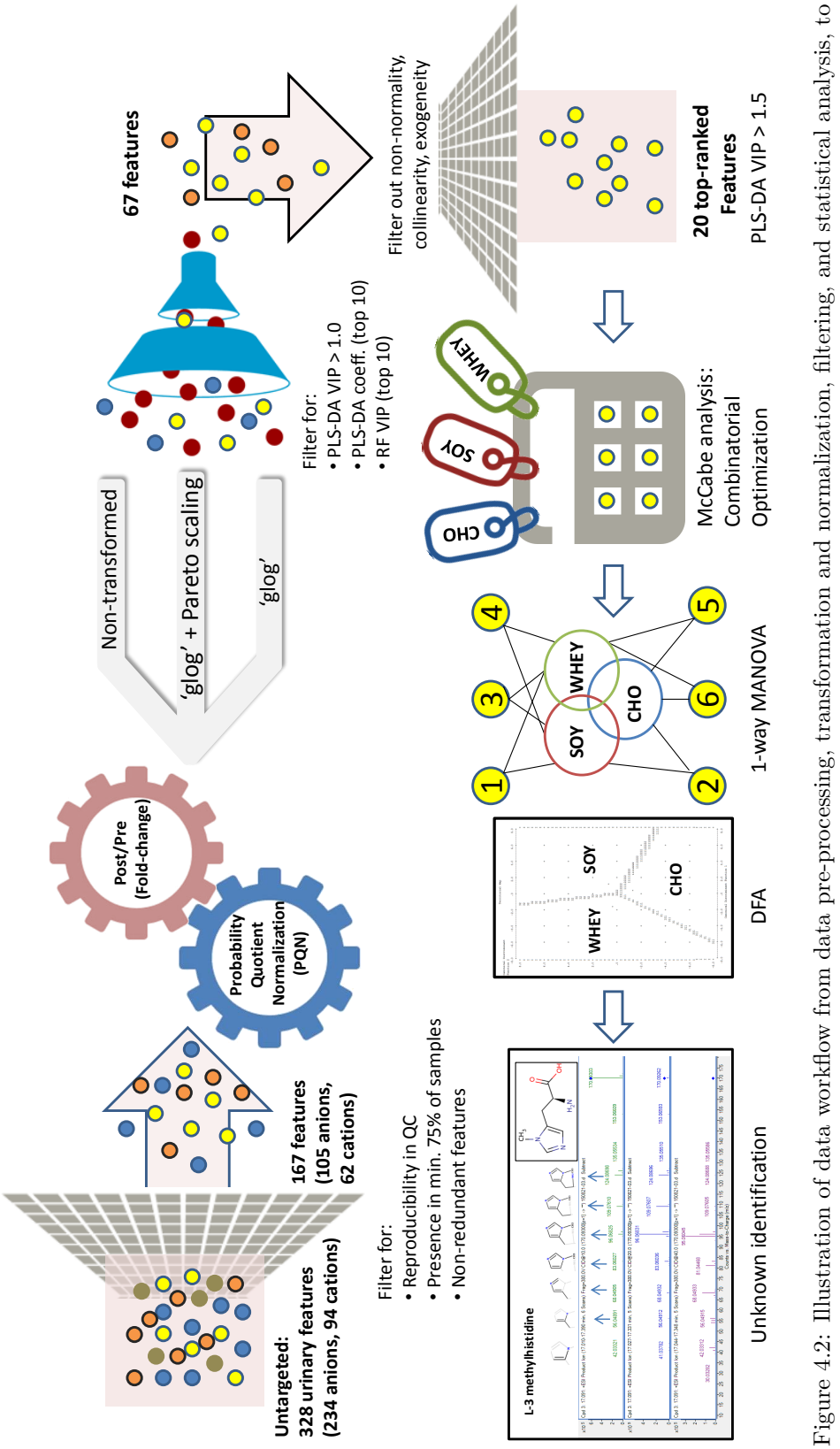


Figure 4.2: Illustration of data workflow from data pre-processing, transformation and normalization, filtering, and statistical analysis, to unknown identification.

4.3 Results

4.3.1 Evaluation of Extracted Features & Establishing Technical Precision Limits

Upon untargeted feature extraction from the subgroup comparison, a combined list of 328 cationic and anionic urinary features with unique m/z : RMT was compiled. The list of 328 features was then extracted in the QC run, which consisted of 7 identical injections of a QC sample. Features that were redundant (i.e., peaks of different m/z having identical RMT), due to the formation of salt adducts or mass fragments, as well as those features that were found to exhibit poor reproducibility and/or poor resolution, were eliminated. While variations in injection volume are random and typically vary only within 10% within the same run, signal variations due to differences in ionization processes, on the other hand, can be relatively large. These differences arise due to competitive ionization and suppression effects in the ion source, whose time of occurrence is dependent on the temporal migration of the abundant ionization-suppressing species during the electrophoretic separation. This can impose a systematic positional bias on signals such that certain injection positions in a particular electrophoretic migration time window are more susceptible to ionization suppression effects, thereby, leading to markedly decreased signal responses in specific injection positions and resulting in an overall lack of signal uniformity across seven injections of the same sample. In addition to ionization suppression effects, co-migration of isobaric compounds is another contributing factor that can negatively impact the reproducibility of peaks, resulting in unresolved peaks within each of the seven sample injections. Therefore, although the QC run consists of seven identical injections, the signals arising from each injection are subject to variations

in injection volume and ionization processes, as well as co-migration effects. Since the instrumental precision of identical injections in the absence of co-migration and suppression effects was previously found to vary within 10%, a 2-fold greater threshold tolerance of 20% was, thus, established to be the cut-off limit for the first-pass filtering of reproducible features. The filtered features were found to exhibit an overall CV of 8.1% within the single QC run. A total of 163 features from both positive- and negative-ion modes was compiled following the initial QC screening procedure.

4.3.2 Assessment of Subgroup Comparison Run

Consisting of paired injections of the pooled Pre and Post samples for each of the three subgroup types (CHO, SOY, WHY), along with the injection of a sample blank, the subgroup comparison run served as a quick screen among the pre-screened metabolites for interesting features exhibiting varied responses either between subgroup types and/or between the Pre and Post states. Of the features extracted in the subgroup and reversed subgroup runs, no single feature was found to exhibit three unique trends for each of the CHO, SOY, and WHY subgroups within an electropherogram. Usually, either all three dietary interventions exhibited a consistent Pre-to-Post trend or two dietary groups exhibited a common trend with the third group exhibiting a distinct trend (Figure 4.3a,b). Though the subgroup run was useful as a preliminary screen for any potentially interesting features differing in their responses between the three dietary groups, it was later found that many of the features that had displayed different trends in the subgroup run were, in fact, not found to be statistically significant following individual sample analysis. Conversely, features which were later found to be significant in group discrimination did not necessarily reveal any evident or obvious contrasts between groups in the initial subgroup run. For this reason, the

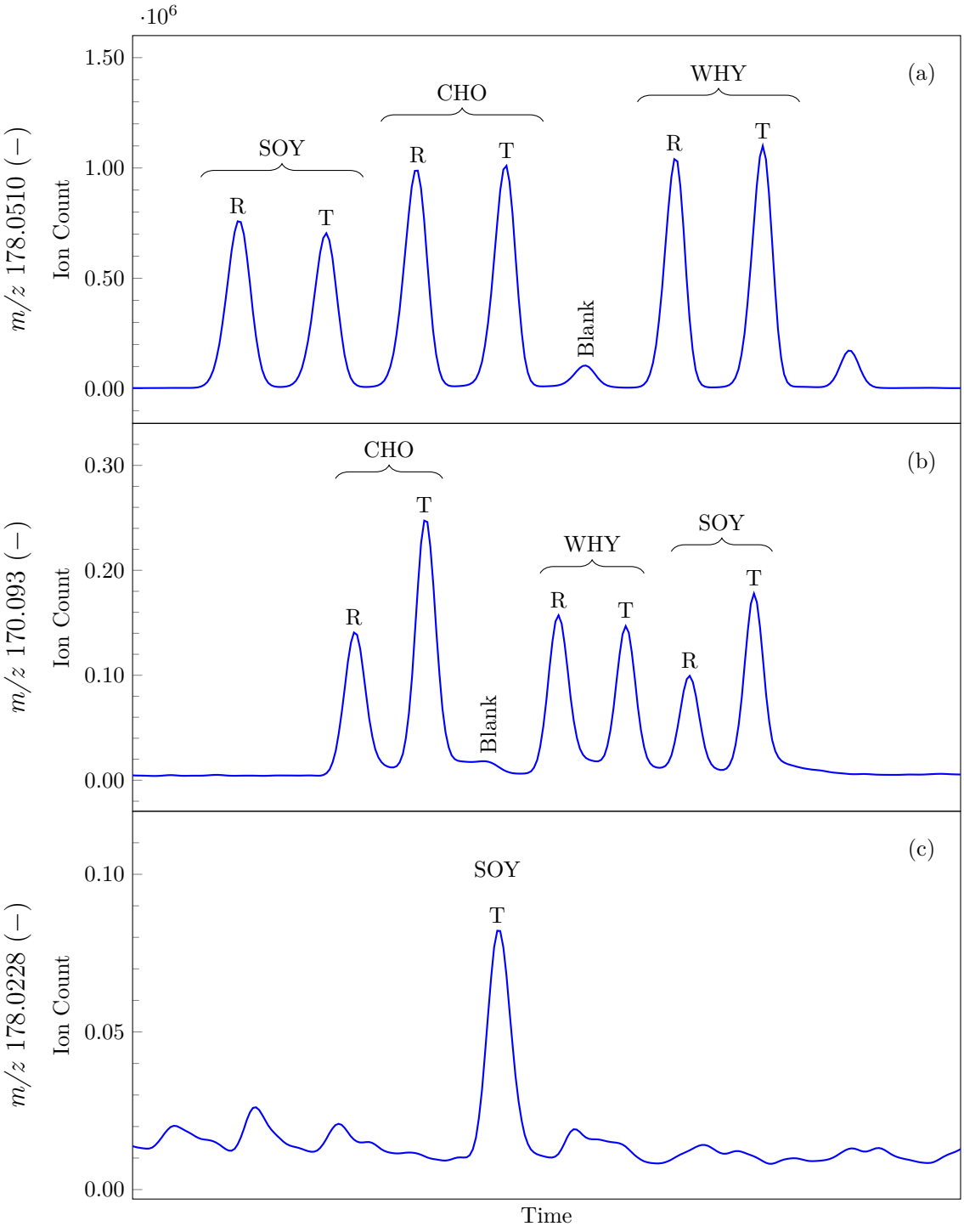


Figure 4.3: Examples of trends observed in subgroup comparison runs: (a) All three groups exhibit a common Pre (R)-to-Post (T) trend; (b) Only two of the three groups exhibit a common Pre-to-Post trend; (c) Observation of a single peak (i.e., feature is exclusive to a single subgroup).

analysis of the subgroup run was found to be misleading. Two reasons to explain this discrepancy in results between the subgroup run and individual sample analysis are: 1) the sensitivity of the pooled subgroup samples to outlier samples, which was sufficient to skew the overall pooled response of the subgroup sample; and 2) the fact that the pooled subgroup samples do not take into account the variations in dilution and hydration status of the constituent individual samples. However, in the case of four metabolite features, which were noted to have been eliminated due to low signal or poor reproducibility in the QC run, yet exhibited a measurable signal in the initial feature extraction step from the subgroup run, it was unexpectedly observed that these four features had revealed only a single peak within the six-sample subgroup comparison run (Figure 4.3c). Upon extraction in individual samples, it was found that the four peaks were mainly found in Post-SOY individuals. As such, the four unique features were labelled as metabolites specific to the SOY diet intervention, and were investigated separately from the pre-screened list of 163 features, which will be discussed in more detail in Section 4.4.3. Although the subgroup run had a tendency to produce false leads, it was nevertheless found to be a useful tool in providing a rapid overview of potential markers of interest to be further investigated, including drawing attention to the presence of soy-specific metabolites, which otherwise may have been overlooked in the analysis of individual samples.

4.3.3 Individual Sample Analysis & Establishing Between-Run QC Cut-Offs

Analysis of the QC replicates between sample runs provided a measure of the analytical reproducibility of the method. Since 14 sample runs were required to analyze 80 samples, therefore, a total of 14 QC measurements were acquired. Analysis of the

pre-screened list of features ($n = 163$) on each of the 14 QC injections resulted in an overall average CV of 13.6%. Therefore, an approximately 2-fold greater CV cut-off of 30% was set to be an acceptable threshold limit, beyond which metabolite signals in the QC sample were considered to have poor reproducibility and were eliminated from the list of features to be extracted on individual samples. This was done in order to reduce the chance of false discovery and highlights the need for technical variation to be lower than biological variation in order to increase the chances of true discovery (Dunn et al., 2008). However, due to the rigorous first-pass filtering in the QC run, it was found that no features had exceeded the 30% CV cut-off in this step. Therefore, no additional features were removed from the list of pre-filtered 163 features.

4.3.4 Characterization of Urinary Metabolites

The list of carefully screened 163 metabolite features, along with the four sporadic features that were found to be exclusive to the subjects consuming the SOY diet, resulted in the detection of a total of 167 unique features (57 cations, 110 anions). These metabolites are denoted by their characteristic m/z : RMT and are summarized in Table 4.7 (Supporting Information). Based on the accurate mass, isotopic distribution, and rules for filtering viable candidates of elemental composition (Watson, 2013), the most probable chemical formula (within 3 ppm mass error), as calculated by the Agilent MassHunter software, is indicated for each feature. Given the large number of metabolites detected, it was not feasible to acquire the MS/MS spectra for each feature for purposes of identification and comparison to known reference standards, especially when only a small handful of features were expected to be relevant to the current investigation. Therefore, only those features for which commercial standards had previously been characterized in-house with known m/z : RMT, as well as a few of the

features that were later found to be significant to the study, were conclusively identified. Although the majority of the metabolites remain unidentified, these unknown features were putatively annotated based on the availability of comprehensive public databases, such as the HMDB (Wishart et al., 2018, http://www.hmdb.ca) and the METLIN Metabolomics Database (Guijas et al., 2018, https://metlin.scripps.edu), which permit screening of candidate molecular formulae for a given accurate mass and searching of experimental or predicted fragmentation mass spectra.

In 2007, the Metabolomics Standards Initiative defined four levels of confidence in metabolite identification, depending on the amount of known information on the metabolite of interest and degree of matching to spectral and physicochemical properties of a commercial standard (Sumner et al., 2007). According to this system of classification, the matching of two or more orthogonal properties (such as m/z, RMT, and MS/MS spectra) to an authentic chemical standard, measured under identical analytical conditions, results in a confident (level 1) identification. When a chemical standard is not available, the matching of properties to data acquired from different laboratories using different analytical platforms is considered to be a putative (level 2) annotation. In cases where only in silico predicted MS/MS spectra are available in public databases, any spectral similarity can only at best yield a putative annotation regarding the chemical class of the compound (level 3). Finally, for features that cannot be identified or classified due to lack of any experimental or predicted spectral data, they are deemed unknown compounds (level 4). While this classification system provides a clear definition of the various levels of metabolite identification and is an appropriate starting point for most untargeted metabolomic investigations, the criteria for putative annotations (levels 2 and 3) were found to be unnecessarily stringent, due to its strong reliance on the acquisition of MS/MS spectra (which is not realistic for the majority of features detected in untargeted

metabolomic studies) and its failure to consider sources of information other than physicochemical properties, which may be highly valuable and beneficial to inferring a metabolite's identity. As a result, faithfully following these rigid definitions for putative annotations can greatly limit the number of unknown metabolites that can be annotated by other means, thereby, preventing the use of many unknown compounds as tentative markers to explain and understand fundamental metabolic processes. Over time, newer classification criteria have evolved to refine the mechanism for reporting confidence in metabolite identification, including a quantitative scoring system (Sumner et al., 2014) and the addition of more classification levels (Schymanski et al., 2014) to accommodate cases that fall "in between" levels established by the Metabolomics Standards Initiative. Additionally, alternative approaches that rely on other information besides direct comparison to reference mass spectra for inferring metabolite identity have also been explored. For example, the work of Krumsiek et al. (2012) combined the knowledge of metabolic networks and pathways, with the use of statistical association analysis and Gaussian graphical modeling, based on partial correlation coefficients, to successfully infer the biochemical identities of unknown metabolites from high-throughput metabolomics data alone. Meanwhile, the work of Zhou et al. (2013) explored a computational pipeline that incorporated biochemical pathways to guide the metabolite identification process, which resulted in much more relevant metabolite annotations compared to the conventional method of manually verifying metabolite identities through comparison to multiple authentic compounds. In a similar manner to the above methods, the current study incorporated the results from correlation analysis and the knowledge of metabolic pathways in order to infer the biochemical identity of many unknown metabolites with known accurate m/z. To allow for more pertinent annotations of unknown compounds, which will not only guide future experiments to verify their identities, but also facilitate the understanding

of their biological roles and the metabolic processes involved, the classification system of the Metabolomics Standards Initiative has been modified and tailored for the characterization of the metabolites in the current investigation, as follows:

Level 1: Confirmed identity – based on comparison to two or more orthogonal properties of authentic standards (e.g., *m/z*, RMT, MS/MS spectra)

Level 2: Putative identity – deduced from the low number (< 5) of probable candidate molecules/structures based on the context of biological samples being studied and relevance to expected metabolic pathway(s) under investigation, along with one or both of the following additional criteria: unambiguous matching of MS/MS spectra from literature/database; significant statistical association with metabolite(s) of known biological role to provide evidence for metabolic and biological context, thereby, increasing confidence in annotation and characterization of metabolite

Level 3: Putatively annotated compound – deduced from the low number (< 5) of probable candidate molecules/structures based on the context of biological samples being studied and relevance to expected metabolic pathway(s) under investigation

Level 4: Putatively annotated compound class – similarity to predicted in silico spectra; spectral similarity to measured compounds of a class but unable to distinguish between positional isomers; insufficient information is available to confirm one exact structure

Level 5: Unknown compound – molecular formula with unknown structure due to: a) no available candidate molecule/structure for the given accurate mass;

or b) availability of tens to hundreds of candidate molecules for a single accurate mass

4.3.5 Comparison of Urine Normalization by Creatinine, Osmolality, & PQN

Prior to the multivariate statistical analysis of individual samples, normalization of urine was required. Due to the high day-to-day variability in urine volume (600–2500 mL/day) (Ryan et al., 2011) within the same individual and between individuals as a result of differences in hydration status or other physiological factors, normalization of urine is necessary to correct for variation in individual urine output and adjust for urinary metabolite concentrations. Various post-sample normalization methods exist thereby eliminating the need for measurement of the renal elimination rate or the absolute urine volume (Warrack et al., 2009). Creatinine (Crn) and osmolality measurements are most commonly used for urine normalization. Although it is generally accepted that the excretion of endogenous urinary Crn is relatively constant in the absence of kidney disease (Saude and Sykes, 2007), several studies have shown that urine Crn levels in humans are susceptible to variations of as much as 4–5 fold change over a 30-day period due to such factors as meat and creatine intake, sleep deprivation, time of day, level of physical exercise, mental state and other metabolomic responses (Gray et al., 1990; Miller et al., 2004; Shockcor and Holmes, 2002). In a study of the effect of a high-protein diet on urinary metabolic profiles in human adults, Rasmussen et al. (2012) found that urinary Crn levels increased while Reimer et al. (2012) also reported that Crn levels change with dietary intervention. As such, normalization to Crn cannot be assumed to be appropriate for normalization of all metabolites. On the other hand, normalization by osmolality, which is based on the

premise that the concentration of solutes is representative of the total endogenous metabolic output, has been reported to offer advantages over normalization by Crn. For example, osmolality is not as easily influenced as Crn levels by such factors as diurnal rhythms, diet, activity, stress, or health (Chadha et al., 2001). However, osmolality measurements may not always be accessible, and in such cases, an alternative normalization strategy is needed.

Recently, Dieterle et al. (2006) introduced a new statistical normalization method in response to the technical challenges hampering the measurement of urinary Crn by NMR. Referred to as probabilistic quotient normalization, or PQN, the approach assumes that the intensity of the majority of signals or metabolites is a function of dilution only. To use this approach, determination of the most probable 'dilution' quotient requires a reference sample to which all metabolites in the samples of interest are compared. By dividing the signal of each metabolite in a sample by the corresponding signal in the reference sample, a distribution of dilution quotients is obtained. The most probabilistic quotient for the sample is then estimated to be the median of this distribution, which then serves as the correction factor to which all of the metabolites in the sample are normalized. To compare the effects of urine normalization by PQN with those of Crn and osmolality in all 80 samples, correlation plots were examined (Figure 4.15, Supporting Information). Both plots reveal a strong and positive correlation between PQ and osmolality ($r = 0.844$, $p < 0.05$), and between PQ and Crn ($r = 0.934$, $p < 0.01$) measurements. These observations indicate that the PQN method is in good agreement with well-established conventional urine normalization techniques, such as those involving Crn and osmolality.

The distribution of the dilution quotients for a sample analyzed in both positive and negative ion modes may be assessed by constructing a frequency histogram. Figure 4.4 shows frequency histograms for each of the distributions of dilution quotients calculated

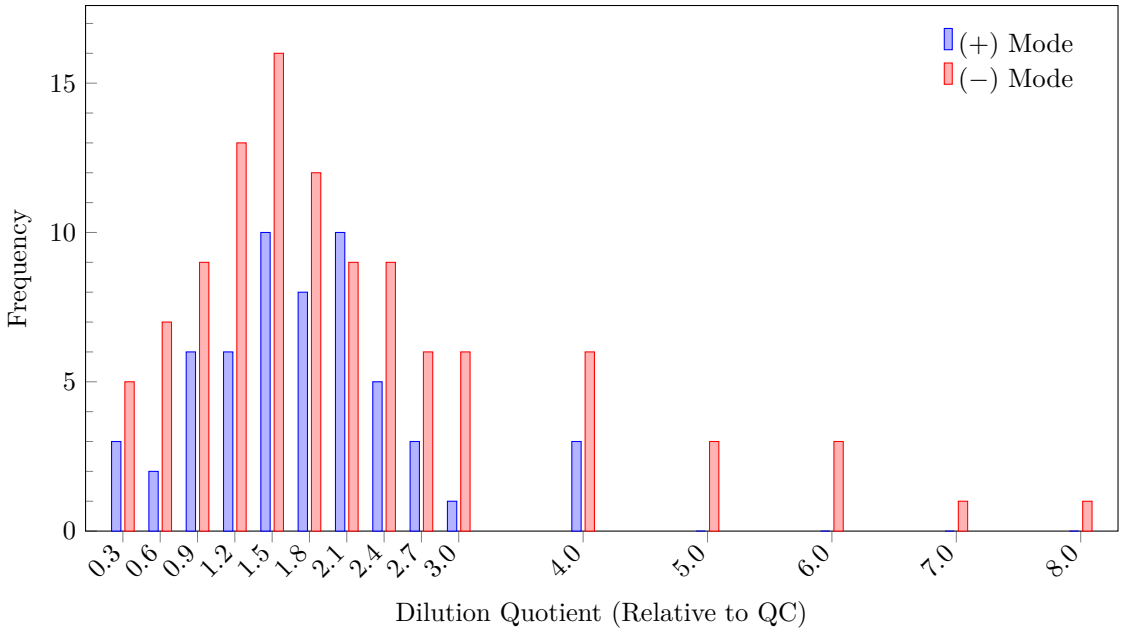


Figure 4.4: Frequency histogram of the dilution quotients for a urine sample analyzed in (red) positive-ion mode ($n = 57$ metabolites) and (blue) negative-ion mode ($n = 106$ metabolites). In this particular case, the median of both distributions are approximately equal. The average probabilistic quotient for the sample was determined to be 1.56 (overall interquartile range: 1.12–2.20).

for cationic and anionic metabolites in a particular sample, revealing in each case that the most probable quotient coincides with the median of the distribution. The median is a robust estimate of the PQN, which is not influenced by outliers, as supported by Dieterle et al. (2006). Furthermore, by plotting the PQ values measured from positive- and negative-ion modes against each other, as shown in Figure 4.16 (Supporting Information), a strong positive correlation ($r = 0.896$, $p < 0.05$) is revealed. These findings suggest that PQN is robust in its application, whose determination is fairly reproducible. While normalization by Crn and osmolality measurements may often be subject to technical challenges and/or metabolic perturbations, determination of the PQ is facile and straightforward, regardless of the mode of analysis. For these reasons, PQN was selected as the method of normalization prior to the statistical analyses of urine samples in this study.

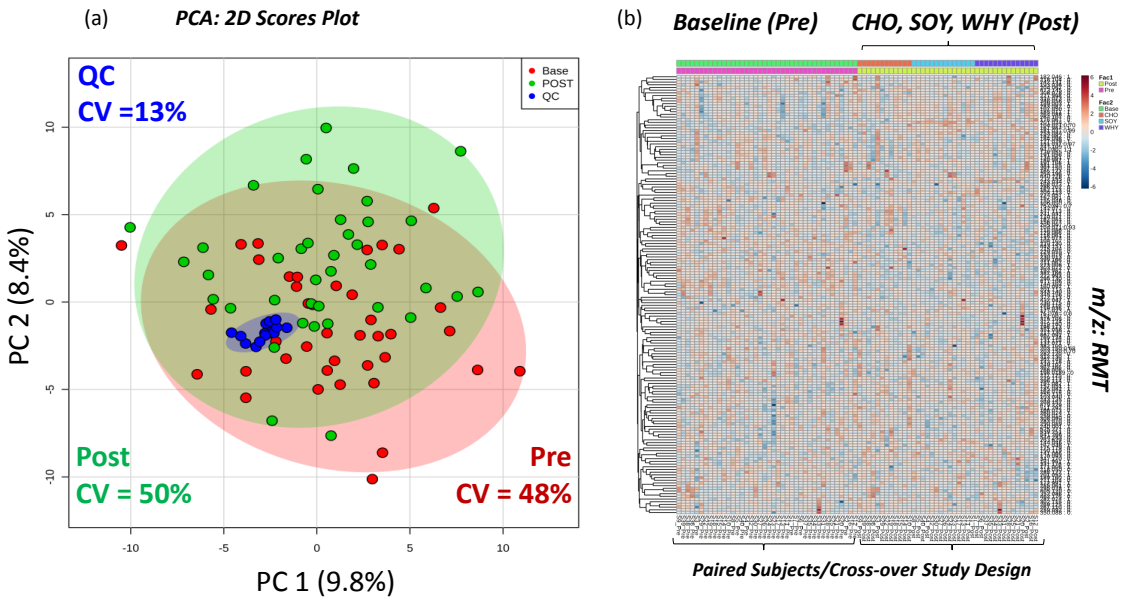


Figure 4.5: (a) PCA 2D scores plot of PQ normalized urine metabolite data Pre- and Post-intervention from 40 subjects highlighting tight clustering of 14 QC replicates (average CV = 13.6%) relative to the large biological variability between subjects, demonstrating good system stability and reproducibility of the MSI-CE-MS technique; (b) 2D heat map with HCA providing an overview of the urine metabolite data set of 40 subjects classified based on pre- or post-intervention and supplementation type.

4.3.6 Preliminary Data Exploration

Figure 4.5a shows a PCA two-dimensional (2D) scores plot, which highlights the tight clustering of 14 QC replicates relative to the large biological variability in the PQ-normalized urinary metabolites derived from the obese subjects before and after the dietary intervention, demonstrating good system stability and reproducibility of the MSI-CE-MS technique. Figure 4.5b is a 2D heat map summarizing the overall data structure of 163 metabolites, measured before and after the dietary intervention, by hierarchical cluster analysis (HCA). To investigate preliminary postprandial differences between the three diets, multi-level PLS-DA was performed on metabolite fold-change values in the postprandial state relative to the pre-intervention state. Despite poor accuracy and robustness of the multi-level PLS-DA model (accuracy = 0.375;

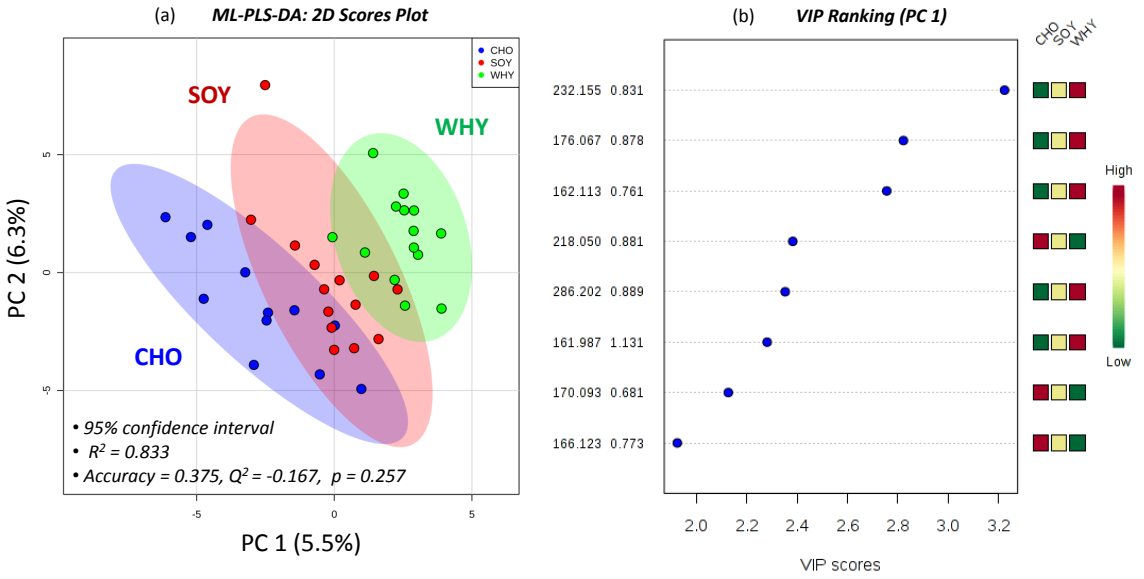


Figure 4.6: (a) Multi-level PLS-DA 2D scores plot and (b) VIP ranking of postprandial fold-change urinary metabolite responses ($n = 163$) in 3 dietary interventions relative to the pre-intervention state of 40 subjects.

$Q^2 = -0.167$), the multi-level PLS-DA scores plot (Figure 4.6a), nonetheless, reveals a directional trend among the 3 diets along the PC-1 axis such that the responses from the SOY intervention are generally observed as an intermediate state between the responses from the CHO and WHY diets. This directional trend was observed in the top 8 metabolites of the VIP ranking plot (Figure 4.6b). Non-parametric Kruskal-Wallis test was performed on all 163 metabolites, from which the top 20 metabolites based on univariate significance are shown in Table 4.1. Performing the Benjamini-Hochberg Procedure on these top 20 metabolites revealed two significant features, a bile acid-glycine-sulphate conjugate (BGS) and isobutyrylcarnitine (iso-C4).

Table 4.1: Top 20 metabolites based on non-parametric Kruskal-Wallis testing.

m/z : RMT (Q)	Compound	Significance (p-value)
263.629 : 0.963 (−2)	Bile acid glycine sulfate conjugate*	0.001
232.155 : 0.831 (+)	Isobutyrylcarnitine*†	0.004
176.067 : 0.878 (+)	Guanidinosuccinic acid†	0.010
168.023 : 1.037 (−)	Uric acid (isotope peak)	0.017
230.115 : 0.851 (−)	Unknown	0.017
240.001 : 1.219 (−)	3-Mercaptolactate cysteine disulfide	0.033
190.119 : 0.955 (+)	Homocitrulline†	0.037
162.113 : 0.761 (+)	Free carnitine†	0.038
218.050 : 0.881 (−)	Unknown	0.039
170.093 : 0.681 (+)	3-Methylhistidine†	0.046
156.077 : 0.667 (+)	Histidine†	0.055
161.987 : 1.131 (−)	Acesulfame†	0.063
201.052 : 1.184 (−)	Unknown	0.073
110.071 : 0.667 (+)	Unknown	0.079
147.113 : 0.629 (+)	Lysine	0.081
286.202 : 0.889 (+)	2-Octenoylcarnitine	0.085
308.099 : 0.817 (−)	N-Acetylneuraminic acid	0.104
218.104 : 0.857 (−)	Pantothenic acid	0.111
141.066 : 0.737 (+)	Unknown	0.111
138.055 : 0.906 (+)	Unknown	0.131

* Compounds found to be significant after correction by Benjamini-Hochberg procedure ($n = 20$).
† Metabolites validated by reference standards or tandem mass spectra. Others were putatively annotated.

4.3.7 Feature Selection by McCabe Analysis for One-Way MANOVA

The plot of lowest Wilks' Λ vs. subset size (Figure 4.17, Supporting Information) revealed that the optimal balance between a small number of metabolite variables and high discriminatory ability was achieved with a subset size of 6 (Wilks' Λ = 0.184; between-groups effect, $p = 4.9 \times 10^{-7}$). For this size, one combination of metabolites was found to consist entirely of known (or putatively known) features: BGS, guanidinosuccinic acid (GSA), 3-mercaptolactate-cysteine-disulfide (MLCD), homocitrulline (Hci), iso-C4, and 3-methylhistidine (3-MeHis). In this case, it was preferable to select the combination with as many known metabolites as possible, since the inclusion of uncharacterized unknown compounds, or "dark matter" (da Silva et al., 2015), could result in the potential loss of metabolic information (Dias et al., 2016). As the dataset was found to satisfy the assumptions of the parametric statistical tests, one-way MANOVA was subsequently performed.

Results of the one-way MANOVA revealed that there were statistically significant mean fold-change differences between the 3 treatment groups with respect to the combination of the 6 metabolites: $F(12, 64) = 6.789$, $p = 1.05 \times 10^{-7}$; Wilks' $\Lambda = 0.194$. The effect size based on Wilks' Λ, 0.560, was indicative of a large difference between the groups. Given the significance of the MANOVA, the univariate main effects were assessed via a follow-up ANOVA. Since Levene's test for homogeneity of variance indicated that all metabolites, except for Hci, had the same variance between groups, a more conservative critical level of 0.025 was set for determining the univariate significance of Hci (Tabachnick and Fidell, 2007). Controlling for the FDR of 0.05 using the Benjamini-Hochberg procedure, all metabolites, except for Hci, yielded significant univariate effects. The liberal Games-Howell post-hoc tests revealed several

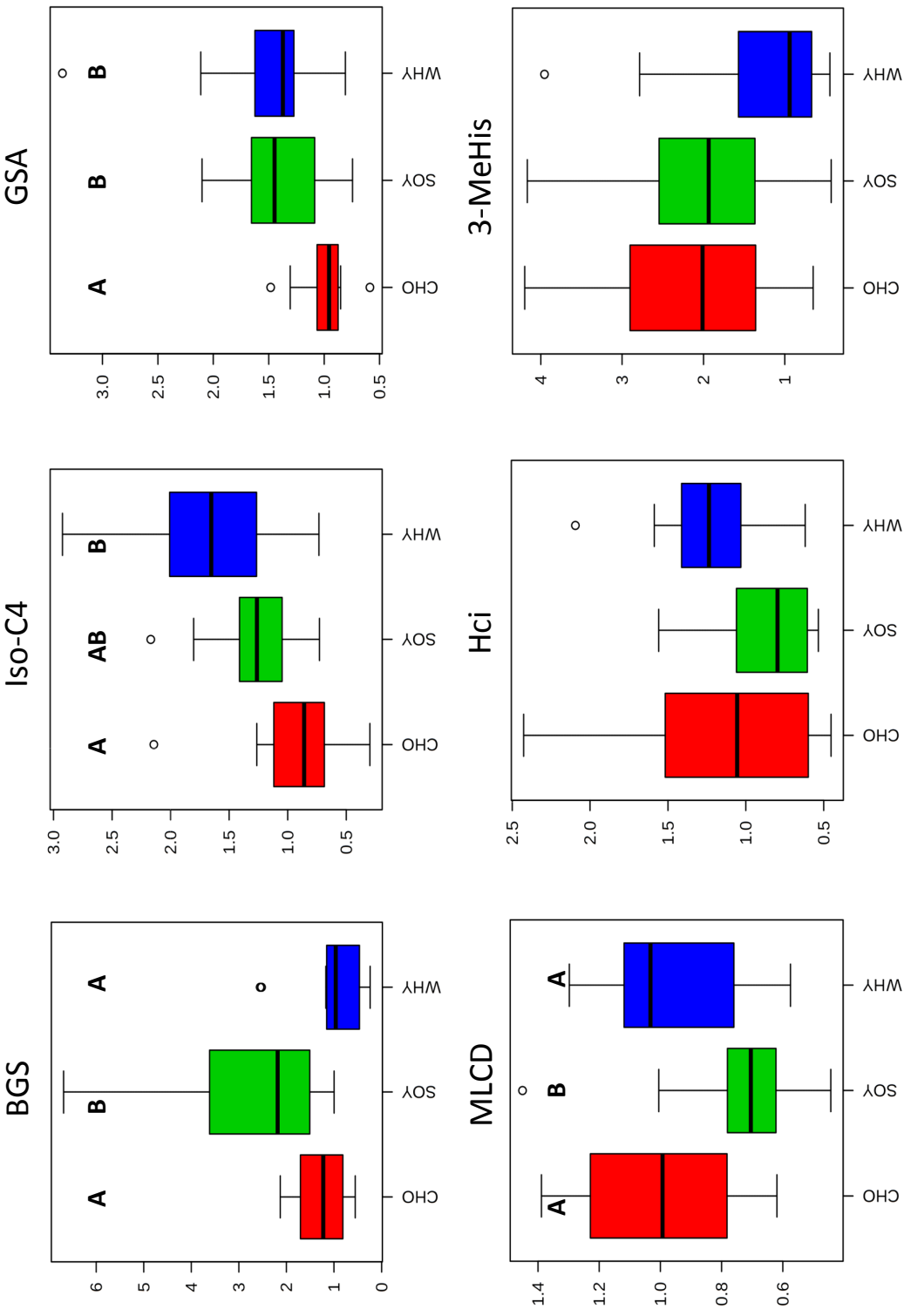


Figure 4.7: Box and whiskers plots for the top 6 metabolites (unknown BGS, iso-C4, GSA, MLCD, Hci, and 3-MeHis) in discriminating between SOY, WHY, and CHO treatment groups. Box and whiskers not sharing the same letter are statistically different ($p < 0.05$).

Table 4.2: Top-ranking urinary metabolites with differential excretion between SOY, WHY, and CHO groups after dietary intervention.

Urinary metabolite		$F_{2,37}$	Effect Size	Significance (p-value)	FDR (q-value)	Significant Post-Hoc Test Differences*	%CV in QC†
m/z : RMT (Q)	ID						
263.629 : 0.963 (−2)	BGS	11.139	0.376	1.63×10^{-4}‡	9.78×10^{-4}§	SOY vs. CHO, $p = 0.006$ SOY vs. WHY, $p = 0.001$	13.1%
232.155 : 0.831 (+)	iso-C4♯	7.383	0.285	2.00×10^{-3}‡	6.00×10^{-3}§	WHY vs. CHO, $p = 0.008$	8.6%
176.067 : 0.878 (+)	GSA	6.090	0.248	5.17×10^{-3}‡	1.03×10^{-2}§	SOY vs. CHO, $p = 0.015$ WHY vs. CHO, $p = 0.006$	7.6%
240.001 : 1.219 (−)	MLCD	4.510	0.196	1.77×10^{-2}	2.66×10^{-2}§	SOY vs. CHO, $p = 0.042$ SOY vs. WHY, $p = 0.047$	8.7%
190.119 : 0.955 (+)	Hci♯♭	3.619	0.164	3.67×10^{-2}	4.40×10^{-2}§		20.2%
170.093 : 0.681 (+)	3-MeHis♯	3.285	0.151	4.86×10^{-2}	6.94×10^{-2}		6.8%

* Significant post-hoc Games-Howell tests.
† Variation of the metabolite relative peak areas in QC samples ($n = 14$) expressed as percent coefficient of variation (%CV).
‡ Metabolites significantly different after Bonferroni correction ($p < 0.008$) for $n = 6$.
§ Metabolites significantly different based on Benjamini-Hochberg procedure to control for false discovery rate ($q < 0.05$) for $n = 6$.
♯ Metabolites validated by reference standards. Others were putatively annotated.
♭ Metabolite violates Levene's Test – significance is assessed at more stringent level, $\alpha/2$.

important differences between groups. The SOY group exhibited significantly greater urinary fold-change levels in BGS than in both the CHO and WHY groups. Conversely, the SOY group also exhibited the lowest fold-change of MLCD than in both the WHY and CHO groups. Meanwhile, the WHY group exhibited significantly greater fold-change in iso-C4 than that by the CHO group. Lastly, SOY and WHY groups both had significantly greater excretion of GSA than in the CHO group. Figure 4.7 shows the box and whiskers plots, which illustrate the differences between treatment groups, for the six metabolites. Table 4.2 summarizes the between-subjects effects of the six metabolites in MANOVA, including the effect size, univariate significance, FDR, significant post-hoc test differences, and the technical reproducibility in QC samples.

4.3.8 Discriminant Function Analysis (DFA)

Although the MANOVA and post-hoc tests allowed us to identify metabolites for which significant group differences existed, a discriminant function analysis (DFA) was subsequently performed to fully explore and interpret the underlying multivariate nature of the six metabolites that best separated the groups. Since the study involved three treatment groups, two discriminant functions (DFs) were created. Both DFs were found to be statistically significant in discriminating between the groups, $\chi_1^2(12) = 56.655$, $p = 9.12 \times 10^{-8}$ and $\chi_2^2(5) = 20.191$, $p = 1.15 \times 10^{-3}$, respectively. The proportion of explained variance by DFs 1 and 2 were 70.2% and 29.8%, respectively, and the canonical correlations (or effect sizes), were 0.808 and 0.666, respectively, which are indicative of both discriminant functions being strongly associated to group membership. Therefore, the DFA indicates that two underlying dimensions contribute significantly to the group separation.

Table 4.3 summarizes the individual metabolite fold-change means and standard

Table 4.3: Fold-change means, standard deviations, and structure matrix correlation coefficients for top-ranking urinary metabolites.

m/z : RMT (Q)	Group	Fold-change*		N	Structure Matrix	
		glog	Linear		Function 1	Function 2
170.093 : 0.681 (+)	CHO	0.88 ± 0.83	2.11 ± 1.07	12	$0.259^{\dagger}$	-0.255
	SOY	0.82 ± 0.93	2.09 ± 1.18	14		
	WHY	0.09 ± 0.90	1.31 ± 0.99	14		
176.067 : 0.878 (+)	CHO	-0.04 ± 0.33	1.00 ± 0.23	12	-0.071	$0.634^{\dagger}$
	SOY	0.43 ± 0.45	1.41 ± 0.42	14		
	WHY	0.53 ± 0.50	1.54 ± 0.63	14		
190.119 : 0.955 (+)	CHO	-0.04 ± 0.81	1.11 ± 0.61	12	$-0.323^{\dagger}$	0.020
	SOY	-0.32 ± 0.49	0.85 ± 0.31	14		
	WHY	0.27 ± 0.43	1.26 ± 0.36	14		
232.155 : 0.831 (+)	CHO	-0.24 ± 0.70	0.94 ± 0.46	12	-0.182	$0.651^{\dagger}$
	SOY	0.30 ± 0.40	1.28 ± 0.37	14		
	WHY	0.63 ± 0.60	1.68 ± 0.66	14		
240.001 : 1.219 (−)	CHO	-0.05 ± 0.39	1.00 ± 0.26	12	$-0.307^{\dagger}$	-0.290
	SOY	-0.47 ± 0.44	0.75 ± 0.25	14		
	WHY	-0.08 ± 0.39	0.98 ± 0.24	14		
263.629 : 0.963 (−)	CHO	0.23 ± 0.68	1.29 ± 0.55	12	$0.556^{\dagger}$	0.164
	SOY	1.27 ± 0.86	2.87 ± 1.82	14		
	WHY	-0.28 ± 1.04	1.03 ± 0.71	14		

* Values are expressed as mean ± standard deviation.
† Largest absolute correlation between variable and discriminant function.

deviations, as well as the structure matrix correlation coefficients between each metabolite and the two multivariate DFs. The highest absolute value correlation coefficient of each metabolite, marked by an asterisk (*), indicates the DF to which the metabolite contributes the most to group separation. Results suggest that high fold-change excretions of BGS and 3-MeHis, in combination with low fold-change excretions of MLCD and Hci, lead to high scores on DF 1. In contrast, high scores on DF 2 were associated with high fold-changes in iso-C4 and GSA.

From the DFA scatter plot (Figure 4.8), it is observed that DF 1 separates WHY from the CHO and SOY groups, while DF 2 separates SOY and WHY groups from CHO. Conceptually, DF 1 can, therefore, be viewed as the tendency towards exhibiting

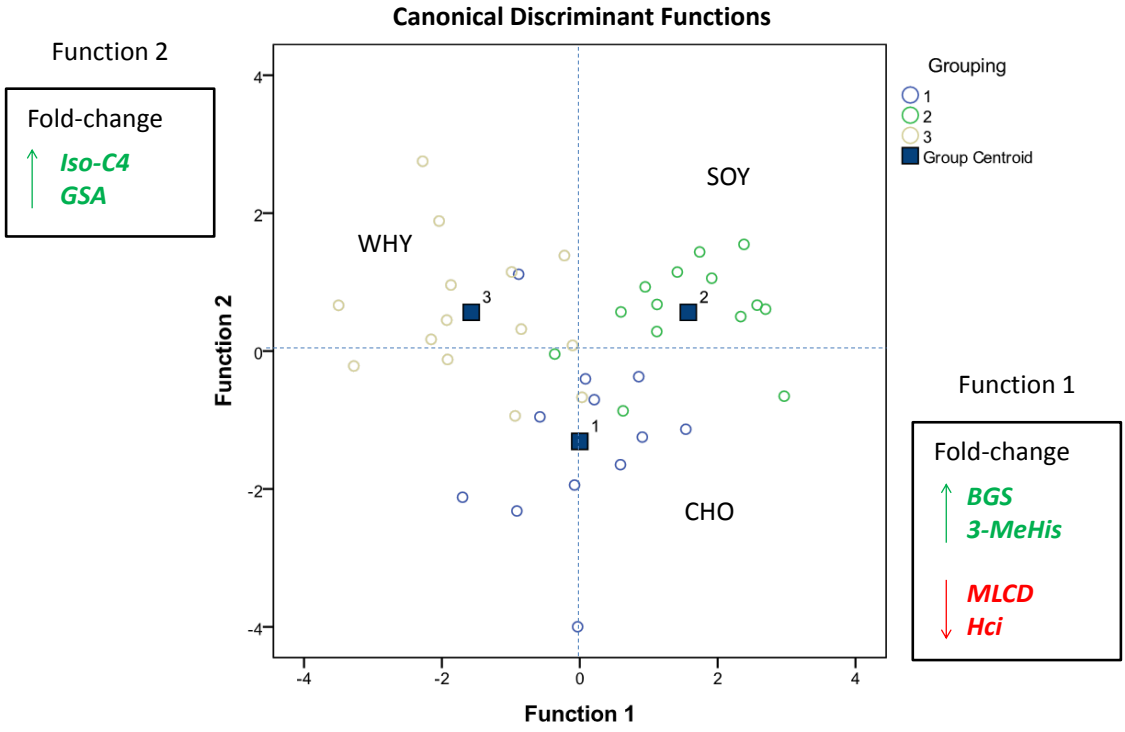


Figure 4.8: Scatter plot from DFA illustrating the discrimination of the three treatment group centroids by DFs 1 and 2. DF 1 separates WHY from CHO and SOY, while DF 2 separates WHY and SOY from CHO.

a urinary metabolic profile reflective of the SOY dietary intervention, while DF 2 corresponds to the tendency towards urinary metabolite profiles of subjects consuming protein supplementation (SOY or WHY). Combining these interpretations, the DFA reveals that the urinary profile from the SOY treatment corresponds to greater fold-change excretion of BGS and 3-MeHis, and low fold-changes in MLCD and Hci, while protein-supplemented urinary profiles are collectively discriminated from the control group based on high fold-change excretions in iso-C4 and GSA.

Owing to the low Wilks' Λ value, the prediction of group membership by the derived functions was excellent, with 82.5% of the subjects correctly classified into the 3 treatment groups by discriminant analysis (refer to Table 4.8, Supporting Information). With the leave-one-out cross-validation procedure, 77.5% of the cross-

validation classifications were correct. Since the likelihood of correct classification by chance alone is 1 in 3 (i.e., 33.3%) in this case, the cross-validation classification accuracy by DFA was found to be statistically significant at $p < 0.05$. Compared to the univariate approach, which at best would provide an effect size of 0.376 just based on BGS alone, which is statistically the most significant metabolite, DFA provided maximal discrimination between the 3 groups based on multivariate linear combinations of metabolites, producing an effect size of 0.652. Therefore, the multivariate effect is greater than the strongest univariate effect by 27.6 percentage points in terms of the variance explained by the group membership.

4.3.9 Correlation Studies

Correlation studies were performed between the 6 significant metabolites against a list of 27 known or putatively identified compounds involved in AA metabolism, in order to identify additional relevant biosynthetic and catabolic pathways in protein metabolism that may have been strongly influenced by the dietary intervention. After controlling for the FDR when data-mining for multiple correlations ($n = 177$), only 5 significant correlations involving 3 of the top urinary markers (iso-C4, GSA, and 3-MeHis) were found (Table 4.4). Additionally, 3-MeHis had the strongest correlations with 2 metabolites (aspartyl-histidine (Asp-His) and trimethyllysine (TML)), at $r = 0.600$ and $r = 0.589$, respectively, which were both found to be significant even with the more stringent Bonferroni correction. GSA was significantly correlated with symmetric dimethylarginine (SDMA) ($r = 0.519$), while iso-C4 was significantly correlated to free carnitine (C0) ($r = 0.498$) and alanine (Ala) ($r = 0.488$). In all of the five cases, the metabolites were positively correlated.

Table 4.4: Significant metabolite correlations associated with top-ranked urinary markers.

Urinary marker		Correlated metabolite		Pearson corr., r	Significance (p-value)	FDR (q-value)
m/z : RMT (Q)	ID*	m/z : RMT (Q)	ID*			
170.093 : 0.681 (+)	3-MeHis†	271.104 : 0.768 (+)	Asp-His	0.600	4.34×10^{-5}‡	5.63×10^{-3}§
		189.160 : 0.649 (+)	TML†	0.589	6.36×10^{-5}‡	5.63×10^{-3}§
176.067 : 0.878 (+)	GSA†	203.150 : 0.700 (+)	SDMA†	0.519	6.01×10^{-4}	3.54×10^{-2}§
232.155 : 0.831 (+)	iso-C4†	162.113 : 0.761 (+)	C0†	0.498	1.00×10^{-3}	3.54×10^{-2}§
		90.055 : 0.818 (+)	Ala†	0.488	1.00×10^{-3}	3.54×10^{-2}§

* 3-MeHis = 3-methylhistidine; GSA = guanidinosuccinic acid; iso-C4 = isobutyrylcarnitine; Asp-His = aspartylhistidine; TML = trimethyllysine; SDMA = symmetric dimethylarginine; C0 = free carnitine; Ala = alanine.
† Metabolites validated by reference standards or tandem mass spectra. Others were putatively annotated.
‡ Significant correlations after Bonferroni correction, $p < 2.82 \times 10^{-4}$ (177 correlations at $\alpha = 0.05$).
§ Significant correlations after Benjamini-Hochberg procedure to control for false discovery rate ($q < 0.05$) for $n = 177$.

4.4 Discussion

4.4.1 Significance of Discriminatory Metabolites

Prior to performing the metabolomic analysis, it is important to acknowledge the compositional differences between dietary soy and whey protein sources. Whey is a liquid by-product of cheese manufacturing, which comprises 20% of the protein in cow's milk (Hoffman and Falvo, 2004), while soy protein is plant-based and typically contains high levels of isoflavones and other biologically active phytochemicals. Although soy and whey protein both contain a high concentration of all essential amino acids, whey protein is a richer source of BCAA than soy (Morifuji et al., 2009), namely, isoleucine, leucine, and valine (Table 4.9, Supporting Information), in which their supplementation has been shown to promote growth of muscle tissue and reduce muscle degradation (Maki et al., 2012; Monirujjaman and Ferdouse, 2014). During periods of brief starvation, caloric restriction or exercise, catabolism of BCAA has been shown to occur (Biolo et al., 2007; Shimomura et al., 2004; Vazquez et al., 1985) – a process which is associated with both muscle proteolysis (Paul and Adibi, 1980) and increased fatty acid oxidation (Kainulainen et al., 2013). Findings from Hector et al.

(2015) had previously shown that, despite a lack of significant differences in body composition changes following a 2-week hypoenergetic dietary intervention between the CHO, SOY, and WHY groups, the WHY group was able to best maintain their postprandial MPS rates relative to the pre-intervention state with a reduction of only 9%, while the postprandial rates in the SOY and CHO groups decreased by 28% and 31%, respectively. Though the findings were consistent with higher levels of leucine present in whey protein conferring a protective effect against muscle loss during caloric restriction, it was unclear why the MPS rate reduction between soy and whey groups were considerably different, by at least 3-fold, given that BCAA content is only approximately 1.25-fold greater in whey than in soy. In order to better understand the underlying metabolic differences contributing to this result, urinary metabolite profiles were analyzed in the current study as a non-invasive alternative to blood samples. Despite being a waste product, urine is often overlooked as a metabolite-rich biofluid with diverse chemical complexity that includes breakdown products from a variety of metabolic processes, ranging from endogenous processes, the consumption of food, drink or drugs, exposure to environmental contaminants, to by-products generated by intestinal bacteria (Bouatra et al., 2013). Currently, approximately 3100 metabolites found in human urine are catalogued in the Urine Metabolome Database (UMDB) (Bouatra et al., 2013, http://www.urinemetabolome.ca) covering such chemical classes as amino acids, organic acids, carbohydrates, lipids, aromatic heterocyclic compounds, aliphatic cyclic compounds, alkaloids, nucleosides and polyketides. Out of the 163 urinary features detected by the current study, 6 metabolites are highlighted here as being able to collectively provide significant discrimination between the 3 dietary treatment groups, which may offer some insight into the differences in energy and muscle metabolism observed previously between the two protein supplementation types (Hector et al., 2015).

Isobutyrylcarnitine (iso-C4)

Among the list of significant metabolites, iso-C4, an acylcarnitine, was conclusively identified as one of the top 2 ranked metabolites in this study, in which an increasing trend in excretion was observed from SOY to WHY, while the CHO group did not show any fold-change. Unlike most acylcarnitines that are derived from fatty acid oxidation, iso-C4, in contrast, is generated from the catabolism of the branched-chain amino acid, valine, via the intermediate isobutyryl-coenzyme A (isobutyryl-CoA) (Wanders et al., 2012) and enzymatic action of mitochondrial carnitine palmitoyltransferase I (carnitine acyltransferase I, CPT1), which exchanges the CoA moiety for L-carnitine (Schooneman et al., 2013). As an acylcarnitine, iso-C4 readily passes into the mitochondria to be oxidized and used as an energy substrate. Depending on diet, fed/fasting status and/or health status, the mitrochondrial efflux of acylcarnitines into the plasma can be modulated, which in turn, would reflect either normal variations in the flux of the fatty acid oxidation pathway or impaired metabolism (Costa et al., 1999; Huffman et al., 2012; Schooneman et al., 2013). It is interesting to note that while Costa et al. (1999) observed a general increase in plasma acylcarnitines associated with fatty acid metabolism following a 20-hour fast in children, the short-chain acylcarnitines derived from BCAA, including iso-C4, remained unchanged. This observation seems to support that normal levels of iso-C4 are generally not modulated by shifts in fatty acid oxidation that may occur as a result of fasting or fat loading, but instead are influenced due to increased substrate availability for its formation. Although the role of acylcarnitine efflux into plasma is not completely clear, it has been suggested that CoA trapping is prevented to allow for equilibrium-dependent mitochondrial CoA metabolic activities to remain unhindered (Schooneman et al., 2013), as well as to maintain the acyl-CoA/free CoA ratio, which is regulated by L-carnitine (Möder et al.,

2003). Following export to the plasma, acylcarnitines are then excreted in urine or bile as a possible detoxification mechanism (Schooneman et al., 2013) and supports, once again, the maintenance of homeostatic metabolite ratios, since the ratio of urinary acylcarnitines/carnitine in healthy individuals is 2.0 ± 1.1 (Chalmers et al., 1984). In humans, iso-C4 is the second most excreted short chain acylcarnitine (19.7%) after acetylcarnitine (70.7%) (Schmidt-Sommerfeld et al., 1989) with normal urinary levels of $7.1 \pm 3.5\,\mu mol/g$ Crn (Maeda et al., 2007). As with plasma levels, its excretion in urine has also been shown to be influenced by age, health status, and diet (de Sousa et al., 1990; Möder et al., 2003), such that the consumption of valine-rich foods has been found to result in higher urinary levels of iso-C4 (Abe et al., 2017; Carroll et al., 1981). Therefore, in this study, the greater urinary excretion of iso-C4 observed in WHY subjects, as shown in Figure 4.7, appears to be consistent with the increased availability of valine as a substrate in whey protein over soy (Table 4.9, Supporting Information).

3-Methylhistidine (3-MeHis)

Despite the lack of statistically significant differences in the fold-change of 3-MeHis excretion between treatment groups, 3-MeHis was still identified as an important urinary metabolite which contributed to group discrimination in DFA. Often suggested as an index of rate of muscle myofibrillar protein catabolism (Chinkes, 2005; McKeran et al., 1979; Plaizier et al., 2000), 3-MeHis results from the post-translational methylation of certain histidine residues present in actin and myosin from all muscles and white muscle fibres (McKeran et al., 1979), which constitutes over 90% of the protein-bound 3-MeHis, as demonstrated in rat skeletal muscle (Haverberg et al., 1975). Released upon degradation of muscle protein, 3-MeHis is excreted unmetabolized in human urine (Long et al., 1975). In support of the aforementioned protective effect conferred

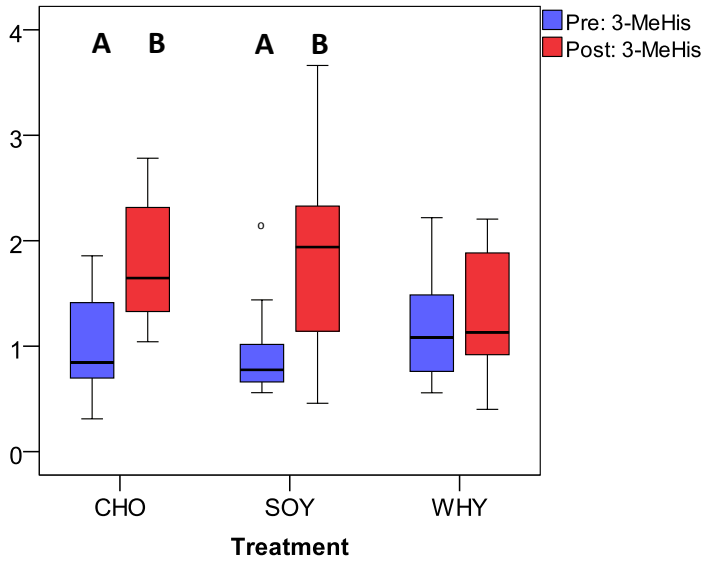


Figure 4.9: Box and whiskers plots for the Pre- and Post- urinary levels of 3-MeHis for CHO, SOY, and WHY treatment groups. Box and whiskers not sharing the same letter are statistically different ($p < 0.05$).

by BCAA against muscle damage and breakdown (Maki et al., 2012; Shimomura et al., 2004) and consistent with the maintenance of postprandial MPS rates by whey supplementation as reported by Hector et al. (2015), the WHY group had shown no significant loss of 3-MeHis in urine, with an average fold-change of only 1.3 between Pre- and Post-intervention, while in contrast, the CHO and SOY groups revealed significantly increased elimination ($p < 0.05$) of 3-MeHis by 2-fold following intervention, as shown in Figure 4.9. It is interesting to note that these results were also reflected in the pooled subgroup analysis of 3-MeHis as seen in Figure 4.3b. These results imply that muscle breakdown may have been elevated in CHO and SOY groups following caloric restriction, while the absence of change in urinary levels of 3-MeHis in WHY subjects indicated that muscle proteolysis was either unchanged or possibly reduced over the weight-loss period. Therefore, the results show that whey supplementation provides increased protection against the effects of muscle catabolism during caloric restriction and weight loss.

Unknown Bile Acid-Glycine-Sulphate Conjugate (BGS)

Ranked as the top discriminatory urinary metabolite in the current study, the $[M-2H]^{2-}$ unknown compound, with m/z 263.629 : 0.963, has only been associated with the metabolite class of BGS for which the identity of the bile acid may be either chenodeoxycholic acid, deoxycholic acid or ursodeoxycholic acid. Tandem mass spectra of this compound revealed characteristic peaks corresponding to the sulfate moiety (m/z 97) and deprotonated glycine molecule (m/z 74) (Figure 4.10). Bile acids play important roles in fat digestion, modulation of gut microbiota, and in the homeostasis of glucose, lipid, and cholesterol through generation of bile flow, which aids in the excretion of toxins and drugs, and re-absorption of vitamins and endogenous compounds (Hylemon et al., 2009; Liao et al., 2016; Trauner et al., 2010). Following the intervention, the SOY group showed the greatest urinary fold-change levels of this compound, which led to significant differences when compared to both the CHO and WHY groups. Literature has previously reported that increased bile excretion is an indication of disruption in micelle formation, which lowers the solubilization and absorption of cholesterol, thereby, leading to a cholesterol-lowering effect (Taylor, 2006; Zhang et al., 2012b). In particular, the high surface activity of amphiphilic saponin molecules and the presence of hydrophobic amino acids, composing the insoluble high molecular-weight peptides fractions in soy protein, have been found to increase bile excretion and thus, contribute to the hypocholesterolemic effects observed with dietary soy protein in animal studies (Greaves et al., 2000; Nagaoka et al., 2010). Furthermore, soy protein hydrolysate has been shown in vitro to decrease the micellar solubility of cholesterol (Nagaoka et al., 1999). Due to formation of large mixed micelles of bile acid and saponin molecules forming aggregates greater than 10^6 Da, the bile acid is not available for absorption (Sidhu and Oakenfull, 1986). The suppression of the

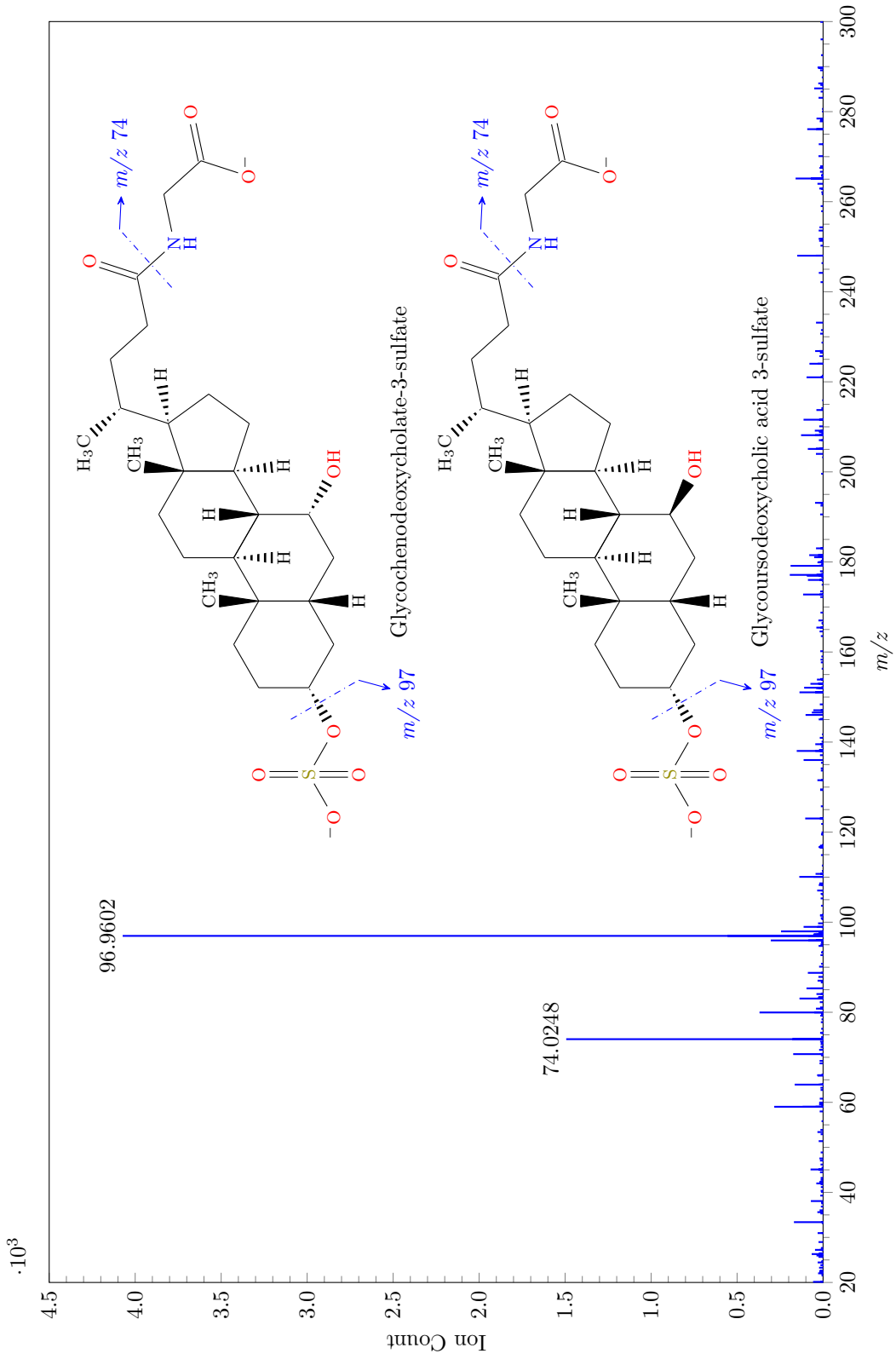


Figure 4.10: MS/MS spectrum (CID = 20 V, negative mode) of an unknown doubly-charged anionic compound, m/z 263.629 : 0.963, believed to belong to the class of bile acid-glycine-sulphate compounds due to the presence of the MS peaks corresponding to sulfate (m/z 97) and deprotonated glycine (m/z 74) moieties.

intestinal absorption of bile acid and cholesterol, thus, enhances their elimination and excretion (Jahan-Mihan et al., 2011). Therefore, the findings of increased bile excretion observed in the SOY subjects were consistent with previous reports of soy protein possessing cholesterol-lowering properties.

Guanidinosuccinic Acid (GSA)

A normal component of urine, GSA was found to be a significant metabolite in the current study. The fold-change excretion was found to be similar between SOY and WHY groups, in which both groups showed significantly increased excretion from the CHO group. GSA is an uremic toxin whose excretion occurs at a relatively constant rate, with a mean daily output of 10 mg in healthy individuals, and is elevated with increased protein intake (Cohen, 1970). GSA derives its guanidino group from arginine through the urea cycle, as well as from the transamidination from arginine to aspartic acid in the liver (Cohen, 1970; Perez et al., 1976). In another study, GSA has also been proposed to be produced from the action of the hydroxyl radical on the arginine-precursor, argininosuccinic acid (ASA) (Aoyagi et al., 2001). Owing to the precursory role of arginine in the synthesis of GSA, it was initially unexpected to see that the fold-change excretion of GSA were similar in both SOY and WHY groups, since arginine content is about 3 times greater in soy protein than in whey. On the other hand, the increased excretion of GSA in the SOY and WHY subjects was consistent with the greater protein intake by those groups. Given that GSA is a derivative of aspartic acid, and that aspartic acid is a precursor of arginine in the urea cycle, the similarity in the pattern of GSA excretion in SOY and WHY subjects was in fact, in agreement with the aspartic acid content being similar in both soy and whey protein (Table 4.9, Supporting Information). This observation confirms earlier work by Cohen (1970), which found that excess loading of arginine precursors increased GSA output.

Comparison of the tandem mass spectra with reference spectra of GSA from HMDB (Wishart et al., 2018) confirmed the identity of GSA (Figure 4.11).

Homocitrulline (Hci)

An analogue of citrulline, Hci is a normal metabolite of lysine, which is an essential amino acid. It is formed from the reaction of nitric oxide synthases (NOS) on homoarginine (Rodionov et al., 2016), which is synthesized from lysine via a three-step process by urea cycle enzymes (Rodionov et al., 2016) or by arginine:glycine amidinotransferase (AGAT) (Rodionov et al., 2016). The work of Paik et al. (1977) has shown that the formation of Hci results from the action of ornithine transcarbamylase (OTC) on lysine in the urea cycle. Under caloric restriction, mice have been found to exhibit increased levels of mitochondrial deacetylase protein, Sirt3, which deacetylate and stimulate OTC activity in the urea cycle (Hallows et al., 2011), and promotes amino acid catabolism and beta-oxidation. Interestingly, Hci can also be derived from the reaction between cyanate and lysine residues through the heating of milk (Metwalli et al., 1998). Therefore, its prevalence in the urine of infants and children has been attributed to milk consumption (Gerritsen et al., 1963). As well, its elevated levels in urine have been reported with lysine loading in healthy individuals (Ryan and Wells, 1964) and also with increased protein intake in patients with hyperornithinemia, hyperammonemia, and homocitrullinuria (Fell et al., 1974; Gatfield et al., 1975). Fell et al. found that supplementary ornithine or arginine was found to lower excretion of Hci. Furthermore, with conditions of caloric or nitrogen deficiency generally resulting in increased endogenous protein metabolism, Hci becomes released in the form of ε-amino-carbamoyl-lysine residues from the catabolism of proteins that have been post-translationally modified non-enzymatically via irreversible carbamoylation of lysine residues in proteins by urea-derived isocyanic acid, the active form of cyanate

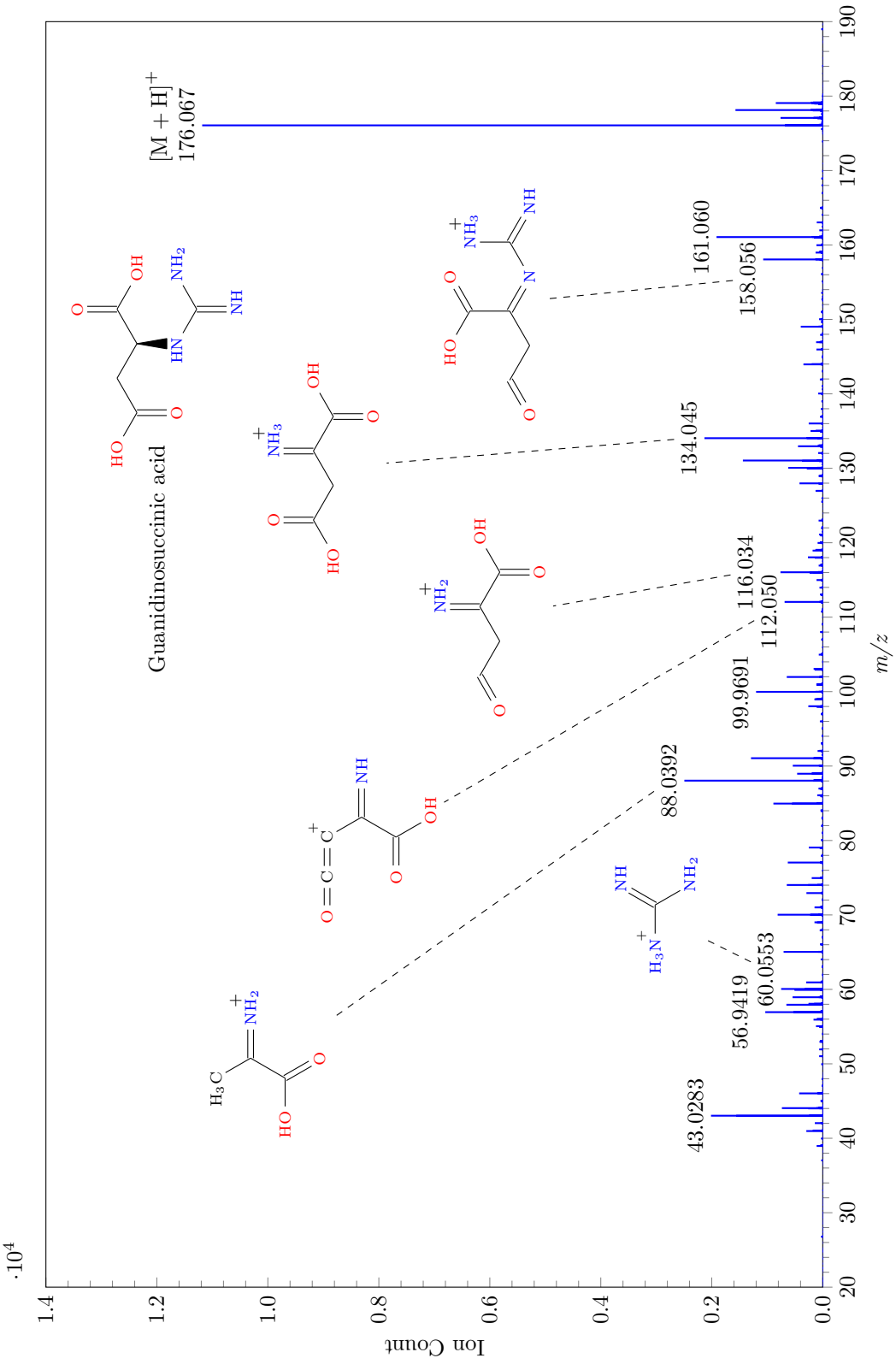


Figure 4.11: MS/MS spectrum (CID = 10 V, positive mode) of GSA, m/z 176.067 : 0.878, annotated after reference spectra from HMDB (Wishart et al., 2018).

(Kraus and Kraus Jr., 2001).

In the current study, greater fold-change excretion of Hci was found with WHY subjects relative to SOY, which is consistent with the presence of higher content of lysine in whey protein than in soy. Following this trend, the wide range of Hci excretion observed in CHO subjects was highly unusual since CHO supplement was expected to contain no lysine content and yet, the range of excretion was seen to overlap the ranges observed in both SOY and WHY. However, when considering the relative deficiency of dietary lysine imposed on the CHO subjects compared to the SOY and WHY groups, it may be speculated that the Hci excreted by the CHO group is primarily the consequence of protein catabolism rather than the metabolism of dietary lysine, such that the urinary Hci originates mainly from the breakdown of carbamoylated proteins. Similar to 3-MeHis excretion being an indicator of muscle protein breakdown, it therefore follows that Hci excretion by the CHO group, in particular, could serve as a marker of general protein degradation since carbamoylation can occur on all free amino groups in a multitude of peptides and proteins (Delanghe et al., 2017). A process which occurs throughout the lifespan of the protein, carbamoylation is irreversible and thus, results in an accumulation of protein-bound Hci, which can be linked with age, health status, and various diseases and disorders, including renal failure, immune system dysfunction, and atherosclerosis (Delanghe et al., 2017; Kraus and Kraus Jr., 2001). Therefore, the large variation of Hci excretion observed, particularly in the CHO group, may be explained by the highly variable nature of protein carbamoylation.

3-Mercaptolactate-Cysteine-Disulfide (MLCD)

First discovered in 1968 by Ubuka et al. in the urine from healthy individuals as well as in a patient following ingestion of cysteine, MLCD was tentatively identified as a significant urinary metabolite in the current study. Based on the findings

by Yuasa et al. (1990), administration of either L-cysteine or L-cystine resulted in increased excretion of MLCD, as well as that of taurine and free sulfate, which are two major sulfur-containing metabolites. These findings were indicative of cysteine metabolism through the transamination pathway (3-mercaptopyruvate pathway) and the maintenance of sulfur equilibrium in humans. Interestingly, it was shown that loading of the sulfur-containing amino acid, L-methionine, did not increase excretion of MLCD, which suggested that methionine metabolism does not occur through cysteine (Yuasa et al., 1990). Ubuka et al. (1992) found that MLCD formation is favoured under conditions of low activity of 3-mercaptopyruvate sulfurtransferase (MST) activity and high activity of lactate dehydrogenase, which generates 3-mercaptolactate from the reduction of 3-mercaptopyruvate upon the reversible transamination of cysteine via the catalytic activity of aspartate or cysteine aminotransferase (CAT) in the presence of α-ketoglutarate. Mercaptopyruvate then reacts subsequently with cysteine to form the mixed disulfide of mercaptolactate and cysteine (Nagahara and Sawada, 2006), also known as MLCD, which is then excreted. In patients with mercaptolactate cysteine disulfiduria, where MST activity is deficient, MLCD is excreted in excessive quantities (Crawhall et al., 1968).

Consistent with the higher cysteine/cystine content in whey protein than in soy, results from the current study have shown that WHY subjects excrete significantly higher levels of MLCD compared to those of SOY subjects. However, unexpectedly, the MLCD excretion by the CHO group was also found to be significantly higher than that of the SOY group. This observation was unusual given that the controls were on a maltodextrin supplement, which contains no cysteine/cystine content, and was, therefore, expected to exhibit the lowest MLCD excretion levels. As well, it was surprising to see that the MLCD excretion between the CHO and WHY subjects were quite similar to each other. A possible mechanism for the increased excretion of MLCD

by the CHO group may be explained in part by the findings of Mårtensson (1982), who found that short-term fasting in healthy subjects resulted in increased excretion of mercaptolactate, the precursor to MLCD. According to Mårtensson, this result suggested that the metabolite was derived from endogenous sulfur amino acids (SAAs), which are liberated during protein catabolism during fasting. Given this information, it may be speculated that protein catabolism may have been triggered in the CHO subjects as a result of insufficient SAA intake. With increased levels of endogenous SAA arising from protein degradation, production of MLCD would be enhanced. An alternative mechanism to support the increased excretion of MLCD by the CHO group may be the decreased activity of the enzyme cysteine dioxygenase, which has been reported in SAA-depleted rats (Mårtensson, 1982). Lowered activity of this enzyme favours the transamination of cysteine and formation of MLCD, as opposed to the oxidative pathway leading to the production of taurine and inorganic sulfate (Stipanuk and Ueki, 2011). Akin to the findings observed with Hci excretion, the CHO group exhibited the greatest range of variation in MLCD excretion compared to the SOY and WHY groups. In contrast to the trends observed in the SOY and WHY groups, where MLCD excretion is believed to be dependent on the intake of dietary SAA, excretion levels in the CHO group may be explained by the inter-individual variations in metabolic processes associated with protein catabolism and alternate enzymatic pathways triggered by dietary SAA deficiency.

4.4.2 Interpretation of Correlations

Considering all the correlated metabolites together shown earlier in Table 4.4, it is of interest to note that C0 and TML are involved in the lysine degradation pathway, such that both are by-products of lysine metabolism, which ultimately leads to fatty

acid catabolism. In the case of C0, which is a key component of skeletal muscle and important in regulating energy metabolism, its strong positive correlation with the increasing trend of iso-C4 excretion from CHO to SOY to WHY may be justified by the fact that C0 can be synthesized from Lys and Met (Rebouche, 2004; Rebouche et al., 1989), both of which are present in increasing amounts from CHO to SOY to WHY. Thus, the excretion of C0 appears to be closely linked to the amount of the dietary amino acid precursors in the different supplement types. TML is a derivative of lysine, however, as it was found to be significantly correlated with 3-MeHis, its excretion by the subjects in the current study is most likely the result of proteins containing TML residues being catabolized due to insufficient dietary intake, which is the first step in carnitine biosynthesis (Servillo et al., 2014). Tentatively identified Asp-His was also found to be correlated strongly with 3-MeHis. However, as it has not yet been found in humans according to the HMDB (Wishart et al., 2018), the tentative identity of Asp-His in this study is so far only supported by its potential association with protein degradation, which is described by HMDB to be "an incomplete breakdown of protein digestion or protein catabolism." Due to the currently limited information on this metabolite, additional details on how the compound is linked to the study are unable to be provided. Further work would be needed to confirm its identity.

Ala is a non-essential amino acid that was strongly correlated with iso-C4. It is a major amino acid that originates from the breakdown of muscle and plays an important role in the glucose-alanine cycle, in which amino acids are degraded by muscle protein in order to allow for gluconeogenesis and muscle contraction (Felig, 1973; Felig et al., 1970). During brief periods of fasting, it has been shown that the breakdown of skeletal muscle is triggered to provide energy through the glucose-alanine cycle (Pozefsky et al., 1976). In the process, Ala is predominantly formed and exported by muscle, serving to shuttle pyruvate to the liver for glucose production (Pozefsky et al., 1976). While

the excretion of Ala may arise from the breakdown of muscle protein, the correlation of Ala excretion pattern with iso-C4, which is increased from CHO to SOY to WHY, is likely the result of increased dietary content of Ala provided by the SOY and WHY supplements, in which greater content of Ala was expected to be present in WHY than in SOY (Table 4.9, Supporting Information). The dependence of the excretion pattern of Ala on the amino acid composition of the dietary supplements is, thus, similar to the manner in which excretion of iso-C4 is believed to be dependent on the dietary content of its precursor, Val, in the protein supplements. Furthermore, this is consistent with early reports, in which increased excretion of individual amino acids has been found with increased protein intake (Eckhardt and Davidson, 1949). The correlation between Ala and iso-C4 excretion is, therefore, explained by the similarity in the relative ratios in which Ala and Val are expected to be found in CHO, SOY, and WHY supplements, respectively.

Lastly, SDMA was found to be strongly correlated to GSA. SDMA and GSA are, in fact, both well-known uremic toxins derived from urea, and therefore, contain the strongly basic guanidine group, $HNC(NH_2)_2$ (Vanholder and De Smet, 1999). Produced during turnover of proteins containing arginine that have been post-translationally methylated (Fickling et al., 1993) by the enzymatic action of type II protein-arginine methyltransferases (PRMTs) (Bedford and Clarke, 2009), SDMA is released into the cytoplasm before being excreted almost entirely (Franceschelli et al., 2013). Although SDMA is the biologically inactive stereoisomer of asymmetric dimethylarginine (ADMA), which is the endogenous inhibitor of NOS (Päivä et al., 2004), SDMA has recently been found to be linked to kidney function and correlated closely with glomerular filtration rate (GFR), since it is strictly eliminated via renal excretion (Bode-Böger et al., 2006). Its strong correlation with GSA, and greater excretion by participants consuming soy or whey protein, is consistent with the work of Juraschek

et al. (2013), which has shown that high-protein diets lead to increased estimated GFR without being able to conclude on the long-term effects of high-protein diets on kidney function. Thus, while GSA was aforementioned to be a direct dietary marker of protein intake, it appears that SDMA may also indirectly reflect fluctuations in protein intake in the current study owing to its strong dependence on the GFR, which in this case, was primarily determined by levels of protein intake in participants which were overweight to obese, but otherwise, healthy individuals.

4.4.3 Sporadic Features and Exogenous Dietary Non-Nutrients

As the main goal of this study was to elucidate differences in energy metabolism and muscle synthesis that were brought on by differences in dietary supplementation type (namely carbohydrate, soy protein, or whey protein supplements) during a hypo-caloric intervention period, the current investigation was limited to endogenous compounds and imposed strict criteria throughout the data screening procedures in order to generate robust discriminatory markers of high quality. In the process, compounds which were known or tentatively identified to be exogenous in nature, including caffeine metabolites and artificial sweeteners, were considered to be confounding variables and therefore, excluded from the analyses. Furthermore, compounds which did not meet the criterion of being present in at least 75% of the samples were also excluded, such that only those features which were reliably detected would be retained. However, the dangers of imposing these well-intentioned stringent limitations are that the presence of any important effects or potential interactions between exogenous and endogenous compounds will be overlooked. As well, the exclusion of low-frequency features, which may include distinctive markers of some uncommon yet relevant characteristic(s)

inherent to certain samples, may result in loss of valuable information to the study. By failing to acknowledge these drawbacks resulting from the application of rigorous data-screening measures, full understanding of the system under investigation may be hindered, and possibly lead to flawed and incomplete explanations of the underlying mechanisms at play. Therefore, in order to fully appreciate and recognize the potential value of exogenous and/or sporadic features in the current study, these compounds were investigated separately post-hoc and are briefly discussed below.

Sporadic Features: Soy-Specific Markers

In addition to the soy-specific marker shown earlier in Section 4.3.3, three additional soy-specific features were extracted with the help of the MFE algorithm. Similar to *m/z* 201.0228 : 1.008 (−), these three metabolites were only sporadically detected in the sample runs, being mostly detected in the Post-SOY samples, and exhibited low and irreproducible signals in the pooled QC sample. The large contrasts in SNR between the pooled QC sample and the pooled Post-SOY sample for each of the four metabolite measurements indicated that the metabolites were strongly specific to the urine of the soy participants post-intervention (Table 4.5). However, it is also notable that the soy-specific metabolites were not detected in all participants of the SOY group following dietary intervention. In the case of *m/z* 417.1191 (−), the frequency of detection was only 6 out of 14 (43%), while for *m/z* 201.0228 (−) and *m/z* 431.0984 (−), detection rates in SOY subjects were 86% and 79%, respectively. Meanwhile, *m/z* 353.0337 (−) was the only soy-specific marker whose detection rate was 100%. Differences in the detection frequency of these features seemed to indicate variations in the metabolic production of the compounds between individuals.

Aside from being rich in protein, vitamins, and minerals, soy-based foods provide the richest sources of isoflavones in the human diet, which is a subclass of flavonoids,

Table 4.5: Summary of soy-specific markers.

Soy-specific marker		Average SNR		Samples detected	
m/z : RMT (Q)	Putative ID	Pooled QC ($n = 14$)	SOY	# out of 14	%
201.0228 : 1.008 (−)	4-EPS*†	3.5	17.7	12	85.7
353.0337 : 1.208 (−)	6'-OH-O-DMA-sulfate‡	8.9	40.8	14	100.0
417.1191 : 0.853 (−)	equol 7-O-GLU§	12.3	107.0	6	42.9
431.0984 : 0.872 (−)	DHD-GLU♯	4.1	22.4	11	78.6

* Metabolite validated by reference tandem mass spectra.
† 4-EPS = 4-ethylphenyl sulfate.
‡ 6'-OH-O-DMA-sulfate = 6'-Hydroxy-O-desmethylangolensin-sulfate.
§ equol 7-O-GLU = equol 7-O-glucuronide.
♯ DHD-GLU = dihydrodaidzein glucuronide.

a diverse group of naturally-occurring polyphenolic plant metabolites (Bhagwat et al., 2008). Exhibiting estrogenic activity in mammals and conferring antioxidant and anti-inflammatory properties (Yu et al., 2016), isoflavones are the most well-known class of phytoestrogens that are present as glucosides (i.e., bound to a sugar molecule) in soy (Day et al., 1998). Upon digestion and fermentation following ingestion, hydrolysis of the isoflavone glucosides occurs by the action of intestinal β-glucosidases (Day et al., 1998), thereby, releasing the sugar molecule, and resulting in isoflavone aglycones, for which daidzein, genistein, and glycitein are the major species (Wang et al., 2013). From this point, the absorption and metabolic biotransformative fates of the aglycones vary with the isoflavone species present and depend on the colonic microflora composition of the individual, such that in the absence of microflora, isoflavones are unabsorbed (Chang and Choue, 2013). Once absorbed, isoflavones typically become conjugated with glucuronic or sulfuric acid in the bowels or liver (Chang and Choue, 2013).

Based on the accurate mass of the unknown compounds, three of the compounds have been putatively annotated as metabolites of daidzein. By the enzymatic action of gut microbiota, daidzein may be transformed into different metabolites, such as dihydrodaidzein (DHD), *O*-desmethylangolensin (O-DMA) and equol (7-hydroxyisoflavan) (Chang and Choue, 2013; Minamida et al., 2006) (Figure 4.12). Due to differences in

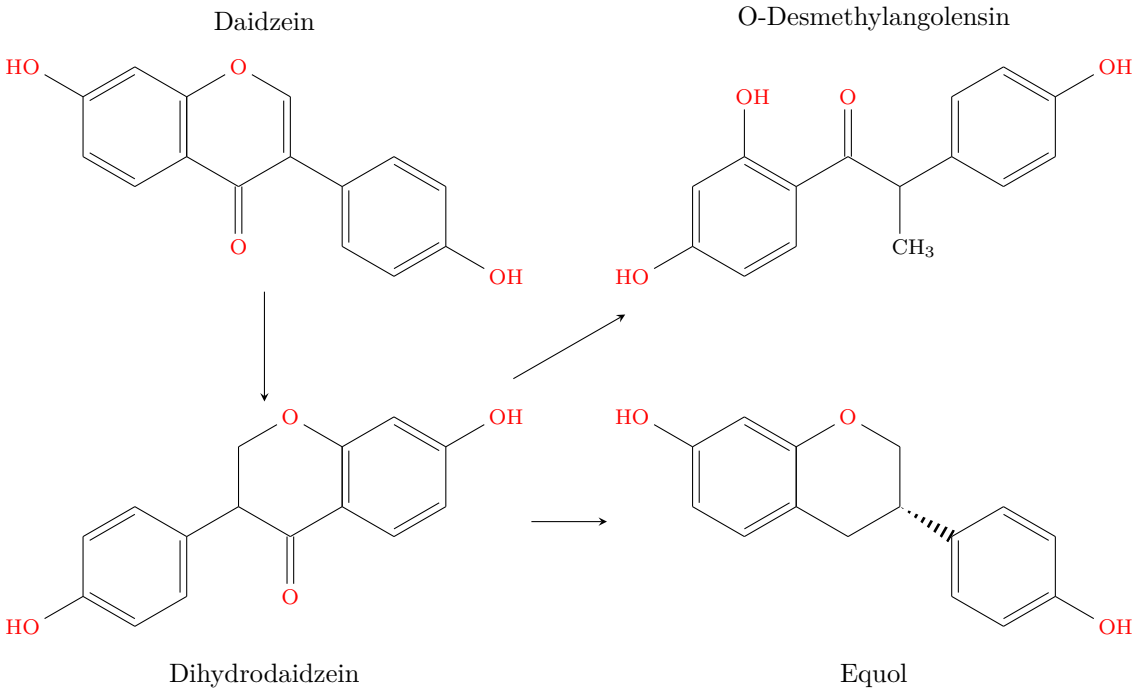


Figure 4.12: Chemical structures of daidzein metabolites.

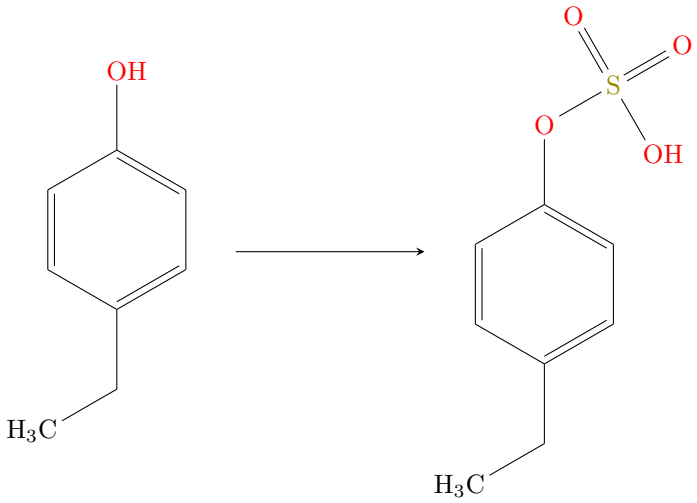


Figure 4.13: Intestinal sulfation of a bacterial metabolite, 4-EP, to 4-EPS is believed to occur via sulfotransferase 1A3.

microbiota composition between individuals, it has been reported that 80–90% of the adult population produce O-DMA, while only 25–30% of the Western population are equol-producers (Atkinson et al., 2008; Setchell and Clerici, 2010). In the current study, 3 of the soy-specific metabolites were tentatively identified to be O-DMA, DHD, and equol metabolites, based on their detection in 100%, 79%, and 43% of the participants, respectively. Taking into account the accurate mass, it is most likely that the sulfated form of O-DMA and the glucuronidated forms of equol and DHD were detected. Although the values for O-DMA and equol both appear to be slightly elevated compared to those previously reported in literature, this may be an overestimation due to the small sample size of only 14 adults in the SOY treatment group, or alternatively, could be the result of increased instrumental sensitivity inherent to the current method. The divergence of DHD levels from O-DMA and equol levels in human urine (Setchell and Clerici, 2010) has been suggested to be indicative of multiple bacteria being responsible for producing equol, such that the list of colonic microflora involved in the biotransformation of daidzein or its precursor glucoside, daidzin, is ever-growing (Setchell and Clerici, 2010). Given this information, the variations in prevalence rate for these urinary soy-specific metabolites would be in close agreement with those reported in literature and support the compound identities as being intestinal bacterial metabolic products of daidzein, for which colonic degradation and absorption rates for different metabolite species have been shown to vary between individuals, depending on the microbiota environment.

Lastly, examination of the MS/MS spectra of the fourth soy-specific marker, m/z 201.0228 (−), and searching the literature have revealed the identity of the compound to likely be 4-ethylphenyl sulfate (4-EPS), a metabolite of gut bacteria that is structurally similar to the well-known toxin, *p*-cresol sulfate (4-methylphenyl sulfate, 4-MPS). 4-EPS has only recently been reported in human serum and blood, and was

found to have an association with the consumption of tofu and soymilk (Guertin et al., 2014; Pallister et al., 2016). Interestingly, high concentration of 4-EPS in the serum of mice has been linked to autistic anxiety-like behaviour (Hsiao et al., 2013), in which the precursor, 4-ethylphenol (4-EP), is believed to be produced by several species of *Clostridium* bacteria (Nicholson et al., 2012). The conversion of 4-EP to 4-EPS is expected to occur via the enzyme, sulfotransferase 1A3 (HMDB, Wishart et al., 2018) (Figure 4.13). 4-EP has recently been quantified in human urine and blood by GC-MS, following oral consumption of quercetin, which is a flavonoid component of soy, in which significantly elevated levels of 4-EP were found in 83% of participants (10 out of 12) (Loke et al., 2009). As well, 4-EP has been known to be a product of the major degradation pathway of genistein, another major isoflavone species in soy (Setchell and Clerici, 2010). Despite the precedence of urinary 4-EP, 4-EPS has not yet been reported in human urine. However, it has been quantified in the urine of rats with chronic renal failure (CRF) (Kikuchi et al., 2010). Comparison of the Q-TOF-MS/MS spectra of the soy-specific compound from the current work with the quadrupole IT-TOF MS/MS spectra of 4-EPS by Kikuchi et al. (2010) confirms the identity of m/z 201.0228 $(-)$ as 4-EPS, in which the presence of product ions at m/z 79.9574 and m/z 121.0660 are characteristic of the sulfate conjugate and correspond to $[SO_3]^-$ and $[M - SO_3]^-$, respectively (Figure 4.14). Additionally, the observed prevalence rate of 86% for m/z 201.0228 $(-)$ in the current work provides further evidence to support the compound identity to be 4-EPS as it is consistent with the rate of 83% previously reported for its precursor metabolite, 4-EP. Herein, this work is the first evidence of 4-EPS in human urine.

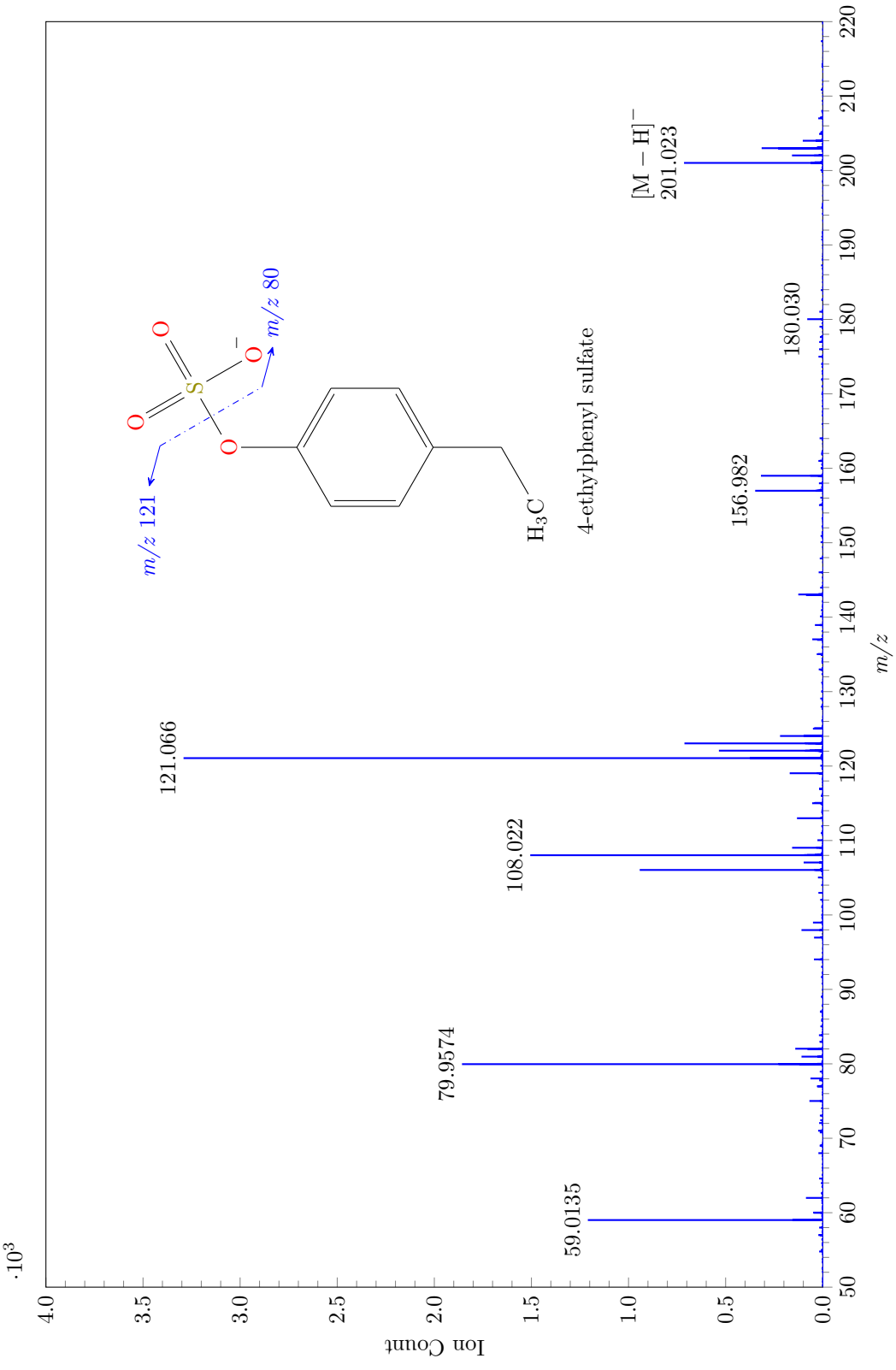


Figure 4.14: MS/MS spectrum (CID = 10 V, negative mode) of a urinary soy-marker metabolite, m/z 201.023 : 1.008, which was identified to be 4-EPS, after comparison with the MS/MS spectrum of 4-EPS by Kikuchi et al. (2010).

Exogenous Dietary Non-Nutrients

In order to avoid potential strong effects of exogenous dietary non-nutrients from confounding the desired analysis of the nutritional effects of important dietary precursors and essential proteinogenic components on muscle synthesis and energy metabolism, the current investigation was limited to endogenous metabolites whenever possible. A considerable challenge of metabolomic analysis in nutritional research, as with the current study, is that the measured effects of dietary nutrients must out-compete the presence of possibly stronger and greater number of non-nutrient effects (Gibney et al., 2005), in order to identify even the slightest variations in metabolism that are contributed by nutrients alone. The reason for this is simply due to the existence of a much higher abundance of plant-based and man-made non-nutrient molecules than there are nutrients in the food supply – therefore, it is important that non-nutrient dietary components are either controlled for, or taken into account before one compares the effects of different diets (Gibney et al., 2005). Furthermore, it should be emphasized that any significant effects of exogenous factors that may influence the metabolome should be considered (Gibney et al., 2005).

For the current study, all subjects were provided with pre-packaged meals that would result in a 750 kcal deficit, in addition to supplements (i.e., CHO, SOY, or WHY), which were prepared in-house from commercially isolated powder as artificially-sweetened cocoa beverages using Nesquik® Chocolate Powder (Nestlé, Vevey, Switzerland) by Hector et al. (personal communication, September 4, 2015). The protein beverages were prepared carefully and identically to control for any undesirable metabolic effects that may arise due to the presence of non-nutrient flavouring agents, which were intentionally added to enhance supplement palatability. However, throughout the two-week duration of the study, participants were allowed to additionally consume

fluids and beverages of their own choice, so long as they were free of alcohol, vitamin, mineral, calories, and protein (Hector et al., 2015). Because of this, several known and putatively identified common exogenous dietary non-nutrients were detected in all participants, including artificial sweeteners (acesulfame potassium, saccharin), caffeine metabolites (1-methyluric acid, 3-methyluric acid), a citrus juice metabolite (proline-betaine), and a sulfated cocoa metabolite. Examination of the univariate significance of these six dietary non-nutrients did not reveal any significant discriminating markers after controlling for the FDR, which was expected. However, when correlations were explored between the dietary non-nutrients and the six endogenous discriminatory markers, acesulfame potassium (Ace-K), which was confirmed by MS/MS, was revealed to be so strongly correlated with GSA, such that it was found to be significant, not only with the FDR method ($q = 0.035$), but also at the more conservative Bonferroni-corrected level of $p < 8 \times 10^{-3}$ ($r = 0.504$, unadjusted $p = 9.26 \times 10^{-4}$).

Ace-K was included as an additive in the preparation of carbohydrate or protein supplement beverages. As it is widely used as a popular low-calorie non-nutritive artificial sweetener in packaged food, candy and drinks, the possible consumption of Ace-K by participants from other beverages or dietary sources, besides the provided carbohydrate/protein supplement beverage from the intervention diet, was not known and could not be monitored. However, since all participants were known to be consuming supplement beverages twice daily, which is expected to possess doses of Ace-K at 18–40 mg/serving (based on acesulfame levels present in diet sodas (Franz, 2010)) and thereby, resulting in the consumption of at least 36–80 mg/day of Ace-K based on consumption of the supplement beverages alone, it may be assumed that the bulk of dietary Ace-K in the participants originated primarily from the supplement beverages themselves. Even if this assumption were to be false, any random variation

or at worst, extreme differences, in Ace-K ingestion from external dietary sources between subjects would unlikely be a factor in causing significant group differences, and even more unlikely to lead to the observation of such a highly significant correlation with GSA.

Approved to be safe for consumption by the United States Food and Drug Administration (FDA) since 1988 (Food and Drug Administration, 1988) Ace-K has increasingly been used in the modern diet as an alternative sweetener to restrict high sugar consumption, which is known universally to promote health problems, such as obesity and diabetes (Bray and Popkin, 2014). Recently, Ace-K has also received health concerns regarding its unknown long-term effects and impact on metabolism, in which studies have shown altered neurometabolic functions in mice with prolonged exposure (Cong et al., 2013), as well as increased body weight and perturbations in gut microbiota composition (Bian et al., 2017). Ace-K is not metabolized by the human body, and is rapidly excreted in the urine unchanged (Wilson et al., 1999). In light of this, the strong correlation between Ace-K and GSA, a marker of dietary protein intake, was quite perplexing and seemingly paradoxical. It would be preposterous to reason that the excretion of Ace-K and GSA were correlated on the same grounds given their contradictory roles as non-metabolite and metabolite, respectively. Keeping in mind that "correlation does not imply causation", attention was diverted to understanding the absorption and excretion kinetics of Ace-K, independently of GSA.

Studies have shown that Ace-K is rapidly absorbed into the blood, with maximum peak levels (0.2–0.3 µg/mL) in humans occurring 1–1.5 h after oral consumption of a single 30 mg dose, followed by clearance from the general circulation at a half-life of 2.5 h. Within 24 h of the initial intake, almost all (98%) Ace-K is eliminated into urine, in which greater than 99% and less than 1% is excreted in urine and feces, respectively (Magnuson et al., 2016). Low levels of Ace-K have also recently been

detected in human breast milk within 24 h of maternal consumption (Sylvetsky et al., 2015). In the current study, no participants were reported to be lactating during the dietary intervention period. Owing to the relatively rapid clearance of Ace-K from the blood, which is immediately followed by elimination into urine, it then follows that the urinary excretion of Ace-K may be dependent on the GFR, in a manner similar to that observed for the excretion of SDMA, as mentioned earlier. For example, it has been shown that children exhibit 2-fold greater plasma concentrations of an artificial sweetener, sucralose, than in adults after consumption of a single 12 fl oz can of diet soda (Sylvetsky et al., 2017). This observation was explained by the authors as being due to the lower GFR in children compared to adults, resulting in a reduced rate of excretion. By the same token, the authors speculated that artificial sweetener levels would be even higher in infants up to 2 years of age, who are expected to have a significantly lower GFR. In the absence of kidney failure and after controlling for age, GFR in healthy individuals has been shown to increase with the consumption of coffee (Herber-Gast et al., 2016; Nakajima et al., 2010) and protein (Brändle et al., 1996; Lohsiriwat, 2013), as well as vary depending on the protein type (Jones et al., 1987; Nakamura et al., 1993). Applying these observations to the current study involving healthy obese/overweight participants, in which BMI status, gender, and age were controlled for and balanced between the treatment groups, any group differences in GFR would, therefore, be predominantly influenced by the amount of dietary protein intake (0.7 g/(kg d) (CHO) vs. 1.3 g/(kg d) (SOY, WHY)) and/or protein source (SOY vs. WHY), which would overshadow the smaller contributions to GFR by inter-individual variations of unmonitored coffee consumption (whose metabolites in this case did not reveal any significant group differences, which was favourable).

Based on the information above, we can, therefore, infer that GFR is expected to be: 1) increased in the SOY and WHY groups relative to the CHO group; and 2)

similar between the SOY and WHY groups, since SDMA (an endogenous marker of estimated GFR and renal function) (Bode-Böger et al., 2006; Kielstein et al., 2006) was strongly associated with GSA, whose excretion patterns were similar in both SOY and WHY groups. Taking into account these considerations, the strong association of Ace-K with GSA can finally be explained by the influence of protein quantity in the SOY and WHY diets (as reflected by GSA excretion levels) on increasing GFR, which in turn led to the apparent proportionally enhanced elimination of Ace-K, as well as SDMA, since renal excretion is the primary route of clearance for both compounds from the human body. In other words, the correlation of GSA with Ace-K and SDMA in the current study revealed that the excretions of Ace-K and SDMA were both dependent on and sensitive to changes in the GFR, which in this case was influenced and enhanced with increased protein intake (for which GSA is a urinary marker) in SOY and WHY groups.

While SDMA is recognized as an important endogenous biomarker of kidney function in humans and also recently, in veterinary medicine for the detection and staging of chronic kidney disease (CKD) in dogs and cats (Brown, 2015; Dahlem et al., 2017; International Renal Interest Society, 2016), the study described here has revealed the potential for Ace-K to be an exogenous marker of GFR. Although the exact amounts of Ace-K ingested by the participants were not known nor controlled for, it was largely assumed that the dietary Ace-K levels would be approximately equivalent in all subjects independent of treatment groups given that the twice-daily administration of the controlled supplement beverages contained equal amounts of Ace-K across all treatment groups. Currently, there is no evidence or published reports in the literature to demonstrate Ace-K as a potential marker of kidney function or indicator of GFR. However, it appears that a pilot study at Oxford University (Oxford, UK) was recently proposed to assess the use of "Acesulfame K as an Exogenous

Marker of Glomerular Filtration Rate (GFR)" by O'Callaghan (2013). Regardless, the work discussed here is the first to present preliminary evidence of the ability of Ace-K excretion to reflect changes in GFR, which in this case, was believed to have been modulated with protein intake based on findings from earlier studies. Although GFR was never measured or assessed directly, changes in GFR were inferred based on the highly significant correlation of SDMA (a marker of GFR) with GSA (a marker of protein intake). From the inferred changes of GFR, the unexpected yet highly significant correlation between Ace-K and GSA was reasonably justified. While initially confounding and dismissed from the analysis due to its irrelevance to the study as a non-nutrient, the unusual finding of Ace-K as having potential utility as an exogenous indicator of GFR was rather accidental and serendipitous, thereby, demonstrating the importance of unbiased and thorough analytical approaches in making novel discoveries and possibly minimizing false negatives.

4.5 Conclusions

In conclusion, untargeted metabolomics in both positive- and negative-ion mode Q-TOF-MS on urine samples from participants consuming soy, whey or carbohydrate supplements during a 2-week dietary weight-loss intervention, resulted in the detection of 167 unique metabolites, including 4 soy-specific metabolites, after rigorous filtering. The application of strict filtering thresholds and parametric statistical analysis eliminated spurious and highly variable measurements, and consequently, led to the identification of several top robust and authentic discriminating metabolites, which were found to be associated with the presence of diet-derived phytochemicals and hydrophobic peptides in soy protein (i.e., BGS), as well as with altered protein catabolism (i.e., Hci, 3-MeHis), and differences in the amino acid composition (i.e., iso-C4,

MLCD, Hci) and total protein content (i.e., GSA) of the supplements. Investigation of significant correlations between the discriminatory markers and metabolites involved in protein metabolism revealed further downstream metabolic perturbations concerning the glucose-alanine cycle, carnitine biosynthesis, and glomerular filtration rate. Most importantly, the WHY group being observed to excrete the least 3-MeHis when compared to the CHO and SOY groups confirmed the findings of Hector et al. (2015), in which MPS following weight loss was found to be attenuated with whey protein supplementation. Additionally, the examination of sporadic features derived from soy and other exogenous dietary non-nutrients revealed, respectively, the first-time detection of 4-EPS in human urine, and the novel finding of the artificial sweetener food additive, Ace-K, as a potential exogenous marker of glomerular filtration rate. Study limitations include the relatively small cohort size and the short duration of the study. Further validation of the observations can be made on a larger number of human participants in the future over a longer intervention time period. Also, the collection of 24-hour urine samples, rather than single-spot, will allow for more quantitative changes in total urinary metabolites excreted in a day between subjects of different groups, as well as, minimize any confounding inter-individual differences in excretion patterns (Van Bemmel et al., 1988) and diurnal rhythms that would be more easily manifested in single-spot urine (Ji et al., 2012). Furthermore, hydrolysis of the various supplements would be helpful in future studies to link measured amino acid composition to the observed metabolic differences between treatment groups since slight discrepancies exist in reported amino acid levels between different information sources. Recommendations for future studies include limiting other dietary/beverage intakes by participants during the intervention period so that external sources of possibly confounding exogenous compounds may be minimized, in order to avoid undesired effects of non-nutrients potentially masking contributions from the nutrients

to the observed phenotype. Last, but not least, future work will be needed to confirm the identity of Asp-His in urine, and the viability of Ace-K as a robust, convenient and clinically reliable indicator of GFR. In summary, findings from this study have demonstrated that urinary metabolite differences arising from dietary nutritional intervention are dependent on the complex array of interactions between intestinal microflora, endogenous metabolites, dietary nutrients, and exogenous non-nutrient compounds. It is impossible, and erroneous, to ascribe the differences in the metabolite phenotype to a single effect, whether it be differences in amino acid composition, soy- or milk-derived peptides, or the presence of soy-specific phytochemicals. Nonetheless, the results highlight the importance of protein quality in dietary and nutritional interventions for effective loss of weight while minimizing the loss of lean muscle mass. Finally, this study demonstrates the wealth of information derived from human urine, which can be used to elucidate underlying mechanisms and reveal new insights responsible for the observed changes in the metabolic condition.

4.6 Supporting Information

Table 4.6: Baseline participant characteristics of the WHY, SOY, and CHO groups*. Adapted from Hector et al. (2015).

Group	Sex (male/female)	Age (yr)	BMI (kg/m^2)	Body fat (%)	Lean mass (kg)	Trunk fat (kg)
WHY	7/7	52 ± 2	34.7 ± 1.1	36.0 ± 1.9	60.6 ± 3.3	8.9 ± 0.8
SOY	7/7	52 ± 2	34.8 ± 1.5	35.9 ± 1.7	61.6 ± 3.9	9.8 ± 1.3
CHO	5/7	48 ± 3	36.9 ± 1.2	37.9 ± 2.2	60.4 ± 4.0	9.5 ± 1.1
p-value		0.257	0.423	0.699	0.794	0.805

* Values are means ± SEMs unless otherwise indicated.

Table 4.7: Summary of 167 metabolites detected in urine, including m/z, RMT, ionization mode, molecular formula, compound ID (confirmed or putative), classification, and identification confidence.

m/z : RMT (Q)	Formula*	Compound ID†	Classification	IDC‡
76.039 : 0.744 (+)	$C_2H_5NO_2$	Glycine (Gly)	Amino acid	1
104.071 : 0.706 (+)	$C_4H_9NO_2$	β-Aminoisobutyric acid (BAIBA)	Amino acid derivative	1
104.071 : 0.938 (+)	$C_4H_9NO_2$	Dimethylglycine (DMG)	Amino acid derivative	1
104.107 : 0.623 (+)	$C_4H_9NO_2$	γ-Aminobutyric acid (GABA)	Amino acid derivative	1
106.050 : 0.869 (+)	$C_3H_7NO_3$	Serine (Ser)	Amino acid	1
110.071 : 0.667 (+)	$C_5H_7N_3$	Unknown	Unknown	5
118.061 : 0.739 (+)	$C_3H_7N_3O_2$	Guanidinoacetic acid (GAA)	Amino acid derivative	1
118.086 : 0.962 (+)	$C_5H_{11}NO_2$	Betaine (Bet)	Quaternary ammonium salt	1
129.066 : 0.773 (+)	$C_4H_9NO_3$	Unknown	Unknown	5
131.118 : 0.755 (+)	$C_6H_{14}N_2O$	Unknown	Unknown	5
132.077 : 0.782 (+)	$C_4H_9N_3O_2$	Creatine (Crt)	Amino acid derivative	1
137.071 : 0.667 (+)	$C_7H_8N_2O$	N-Methylnicotinamide	Pyridinecarboxylic acid derivative	1
138.055 : 0.906 (+)	$C_7H_7NO_2$	Unknown	Unknown	5
141.066 : 0.737 (+)	$C_6H_8N_2O_2$	Unknown	Unknown	5
143.118 : 0.697 (+)	$C_7H_{14}N_2O$	Unknown	Unknown	5
144.102 : 0.969 (+)	$C_7H_{13}NO_2$	Proline betaine (Pro-Bet)	Amino acid derivative (citrus dietary marker)	2
147.077 : 0.926 (+)	$C_5H_{10}N_2O_3$	Glutamine (Gln)	Amino acid	1
147.113 : 0.629 (+)	$C_6H_{14}N_2O_2$	Lysine (Lys)	Amino acid	1
156.077 : 0.667 (+)	$C_6H_9N_3O_2$	Histidine (His)	Amino acid	1
157.061 : 0.786 (+)	$C_6H_8N_2O_3$	Unknown	Unknown	5
160.097 : 1.069 (+)	$C_7H_{13}NO_3$	4-Hydroxystachydrine	Amino acid derivative (citrus dietary marker)	2
162.113 : 0.761 (+)	$C_7H_{15}NO_3$	Carnitine (C0)	Quaternary ammonium salt	1
164.074 : 0.773 (+)	$C_6H_{13}NO_2S$	Unknown	Unknown	5
166.073 : 0.750 (+)	$C_6H_7N_5O$	Unknown	Unknown	5
166.123 : 0.773 (+)	$C_{10}H_{15}NO$	Unknown	Unknown	5
169.134 : 0.759 (+)	$C_9H_{16}N_2O$	Unknown	Unknown	5
170.093 : 0.681 (+)	$C_7H_{11}N_3O_2$	3-Methylhistidine (3-MeHis)	Amino acid derivative	1
175.119 : 0.648 (+)	$C_6H_{14}N_4O_2$	Arginine (Arg)	Amino acid	1
176.067 : 0.878 (+)	$C_5H_9N_3O_4$	Guanidinosuccinic acid (GSA)	Amino acid derivative	2
182.081 : 0.965 (+)	$C_9H_{11}NO_3$	Tyrosine (Tyr)	Amino acid	1
189.160 : 0.649 (+)	$C_9H_{20}N_2O_2$	Trimethyllysine (TML)	Amino acid derivative	1
190.119 : 0.955 (+)	$C_7H_{15}N_3O_3$	Homocitrulline (Hci)	Amino acid derivative	1
198.019 : 0.663 (+)	$C_4H_{11}N_3S_3$	Unknown	Unknown	5
203.150 : 0.689 (+)	$C_8H_{18}N_4O_2$	Asymmetric dimethylarginine (ADMA)	Amino acid derivative	1
203.150 : 0.700 (+)	$C_8H_{18}N_4O_2$	Symmetric dimethylarginine (SDMA)	Amino acid derivative	1
204.123 : 0.799 (+)	$C_9H_{17}NO_4$	Acetyl carnitine (C2)	Fatty acid ester	1
205.097 : 0.938 (+)	$C_{11}H_{12}N_2O_2$	Tryptophan (Trp)	Amino acid	1
217.130 : 0.872 (+)	$C_8H_{16}N_4O_3$	Unknown	Unknown	5
227.125 : 0.663 (+)	$C_8H_{14}N_6O_2$	Unknown	Unknown	5

Continued on next page

Summary of 167 metabolites detected in urine, including m/z, RMT, ionization mode, molecular formula, compound ID (confirmed or putative), classification, and identification confidence.

m/z : RMT (Q)	Formula*	Compound ID†	Classification	IDC‡
229.119 : 0.861 (+)	$C_{11}H_{12}N_6$	Unknown	Unknown	5
229.130 : 0.727 (+)	$C_9H_{16}N_4O_3$	Unknown	Unknown	5
232.155 : 0.831 (+)	$C_{11}H_{21}NO_4$	Isobutyryl carnitine (iso-C4)	Fatty acid ester	1
241.031 : 0.944 (+)	$C_6H_{12}N_2O_4S_2$	Cystine (CySS)	Cysteine conjugate	1
242.131 : 0.866 (+)	$C_9H_{17}N_6S$	Unknown	Unknown	5
249.107 : 0.663 (+)	$C_7H_{14}N_5O_5$	Unknown	Unknown	5
259.092 : 0.897 (+)	$C_9H_{16}N_5S_2$	Unknown	Unknown	5
264.196 : 0.897 (+)	$C_{16}H_{25}NO_2$	Unknown	Unknown	5
266.006 : 0.663 (+)	$C_{18}H_3NS$	Unknown	Unknown	5
269.125 : 0.929 (+)	$C_{11}H_{16}N_4O_4$	Unknown	Unknown	5
271.104 : 0.768 (+)	$C_{10}H_{14}N_4O_5$	Aspartyl-Histidine (Asp-His) or Histidnyl-Aspartate (His-Asp)	Amino acid derivative	4
282.120 : 0.877 (+)	$C_{11}H_{15}N_5O_4$	1-Methyladenosine (1-MAD)	Purine nucleoside	1
286.202 : 0.889 (+)	$C_{15}H_{27}NO_4$	2-octenoylcarnitine	Fatty acid ester	2
312.130 : 1.055 (+)	$C_{12}H_{25}NO_4S_2$	Unknown	Unknown	5
362.166 : 0.663 (+)	$C_{19}H_{25}N_2O_3S$	Unknown	Unknown	5
367.149 : 1.069 (+)	$C_{16}H_{32}NS_4$	Unknown	Unknown	5
432.047 : 1.044 (+)	$C_{14}H_9N_9O_6S$	Unknown	Unknown	5
440.228 : 1.047 (+)	$C_{29}H_{31}N_2S$	Unknown	Unknown	5
76.076 : 0.600 (−)	C_3H_9NO	Trimethylamine N-oxide (TMAO)	Aminoxide	1
94.030 : 1.111 (−)	C_5H_5NO	Unknown	Unknown	5
103.040 : 1.047 (−)	$C_4H_8O_3$	Hydroxybutyric acid	Hydroxy acid derivative	4
107.050 : 1.051 (−)	C_7H_8O	Unknown	Unknown	5
112.052 : 0.632 (−)	$C_4H_7N_3O$	Creatinine (Crn)	Amino acid derivative	1
119.036 : 1.015 (−)	$C_4H_8O_4$	Unknown	Unknown	5
121.029 : 1.022 (−)	$C_7H_6O_2$	Unknown	Benzoyl derivative	4
134.061 : 0.926 (−)	C_8H_9NO	Unknown	Unknown	5
144.046 : 0.943 (−)	C_9H_7NO	Unknown	Unknown	5
151.007 : 0.999 (−)	$C_4H_8O_4S$	Unknown	Unknown	5
153.019 : 1.137 (−)	$C_7H_6O_4$	Unknown	Hydroxybenzoic acid derivative	4
155.046 : 0.955 (−)	$C_6H_8N_2O_3$	Unknown	Unknown	5
157.051 : 1.397 (−)	$C_7H_{10}O_4$	Succinylacetone	Medium-chain keto acid derivative	3
161.987 : 1.131 (−)	$C_4H_5NO_4S$	Acesulfame	Exogenous	1
167.046 : 1.037 (−)	$C_7H_8N_2O_3$	2,3-Diaminosalicylic acid	Benzoic acid derivative	3
168.023 : 1.037 (−)	$C_5H_4N_4O_3$	Uric acid (M + 1 peak)	Purine derivative	3
171.006 : 1.260 (−)	$C_6H_8N_2S_2$	Unknown	Unknown	5
172.991 : 1.117 (−)	$C_6H_6O_4S$	Phenylsulfate	Arylsulfate	3
177.023 : 0.929 (−)	$C_6H_{10}O_4S$	Unknown	Unknown	5
181.037 : 0.979 (−)	$C_6H_6N_4O_3$	1/3/7-methyluric acid	Purine derivative	4
181.037 : 0.999 (−)	$C_6H_6N_4O_3$	1/3/7-methyluric acid	Purine derivative	4
181.991 : 1.075 (−)	$C_7H_5NO_3S$	Saccharin	Exogenous	3
182.046 : 1.025 (−)	$C_8H_9NO_4$	4-Pyridoxic acid	Pyridinecarboxylic acid derivative	3
182.082 : 0.900 (−)	$C_9H_{13}NO_3$	Unknown	Unknown	5

Continued on next page

Summary of 167 metabolites detected in urine, including m/z, RMT, ionization mode, molecular formula, compound ID (confirmed or putative), classification, and identification confidence.

m/z : RMT (Q)	Formula*	Compound ID†	Classification	IDC‡
184.098 : 0.893 (−)	$C_9H_{15}NO_3$	Unknown	Unknown	5
185.082 : 1.263 (−)	$C_9H_{14}O_4$	Unknown	Unknown	5
188.011 : 1.052 (−)	$C_7H_8O_4S$	p-Cresol sulfate (M + 1 peak)	Arylsulfate	3
188.987 : 1.301 (−)	$C_7H_{10}S_3$	Unknown	Unknown	5
189.003 : 1.051 (−)	$C_4H_4N_3O_6$	Unknown	Unknown	5
191.056 : 0.909 (−)	$C_7H_{12}O_6$	Quinic acid	Cyclic polyol	3
191.108 : 1.159 (−)	$C_{12}H_{16}O_2$	Unknown	Unknown	5
192.982 : 1.068 (−)	$C_6H_{10}OS_3$	Unknown	Unknown	5
193.036 : 0.898 (−)	$C_6H_{10}O_7$	Unknown	Glucuronic acid or derivative	4
197.082 : 1.223 (−)	$C_{10}H_{14}O_4$	Unknown	Medium-chain fatty acid	4
198.113 : 0.875 (−)	$C_{10}H_{17}NO_3$	Unknown	Unknown	5
201.023 : 1.008 (−)	$C_8H_{10}O_4S$	4-ethylphenyl sulfate§	Arylsulfate	2
201.052 : 1.184 (−)	$C_7H_{10}N_2O_5$	Unknown	Unknown	5
201.113 : 0.853 (−)	$C_{10}H_{18}O_4$	sebacic acid	Medium-chain fatty acid	3
204.067 : 0.890 (−)	$C_{11}H_{11}NO_3$	Indolelactic acid	Indolyl carboxylic acid	1
211.062 : 0.976 (−)	$C_{10}H_{12}O_5$	Unknown	Unknown	5
215.104 : 0.855 (−)	$C_9H_{16}N_2O_4$	Unknown	Unknown	5
218.050 : 0.881 (−)	$C_9H_9N_5S$	Unknown	Unknown	5
218.104 : 0.857 (−)	$C_9H_{17}NO_5$	Pantothenic acid	Amino acid derivative	3
222.993 : 0.975 (−)	$C_7H_4N_4O_3S$	Unknown	Unknown	5
224.057 : 1.161 (−)	$C_{11}H_{15}NS_2$	Unknown	Unknown	5
227.998 : 0.981 (−)	$C_9H_{11}NS_3$	Unknown	Unknown	5
229.055 : 0.932 (−)	$C_{11}H_{10}N_4S$	Unknown	Unknown	5
230.013 : 0.955 (−)	$C_8H_9NO_5S$	Unknown	Unknown	5
230.115 : 0.851 (−)	$C_9H_{17}N_3O_4$	Unknown	Unknown	5
240.001 : 1.219 (−)	$C_6H_{11}NO_5S_2$	3-Mercaptolactate-cysteine disulfide	Cysteine conjugate	2
243.034 : 0.931 (−)	$C_{10}H_{12}O_5S$	Unknown	Unknown	5
243.135 : 0.841 (−)	$C_{11}H_{20}N_2O_4$	Unknown	Unknown	5
247.109 : 0.952 (−)	$C_{13}H_{16}N_2O_3$	Unknown	Unknown	5
253.084 : 0.859 (−)	$C_{12}H_{18}N_2S_2$	Unknown	Unknown	5
255.135 : 0.839 (−)	$C_{12}H_{20}N_2O_4$	Unknown	Unknown	5
259.014 : 0.886 (−)	$C_7H_{16}O_4S_3$	Unknown	Unknown	5
260.024 : 1.205 (−)	$C_{10}H_{15}NOS_3$	Unknown	Unknown	5
263.629 : 0.963 (−)	$C_{26}H_{43}NO_8S$	Unknown bile acid glycine-sulfate conjugate $[M-2H]^{2-}$	Glycinated bile acid derivative	4
264.088 : 1.154 (−)	$C_{13}H_{15}NO_5$	Unknown	Unknown	5
264.107 : 0.848 (−)	$C_{13}H_{16}N_2O_4$	Phenylacetylglutamine (M + 1 peak)	Amino acid derivative	3
267.110 : 0.838 (−)	$C_{11}H_{16}N_4O_4$	Unknown	Unknown	5
269.152 : 0.830 (−)	$C_{18}H_{22}O_2$	Unknown	Unknown	5
273.008 : 1.326 (−)	$C_{10}H_{10}O_7S$	Unknown	Phenylsulfate derivative	4
283.070 : 0.856 (−)	$C_{12}H_{14}NO_7$	Unknown	Unknown	5
283.082 : 0.834 (−)	$C_{12}H_{18}N_3OS_2$	Unknown	Unknown	5

Continued on next page

Summary of 167 metabolites detected in urine, including m/z, RMT, ionization mode, molecular formula, compound ID (confirmed or putative), classification, and identification confidence.

m/z : RMT (Q)	Formula*	Compound ID†	Classification	IDC‡
285.061 : 0.907 (−)	$C_{11}H_{16}N_3O_2S_2$	Unknown	Unknown	5
287.024 : 1.051 (−)	$C_{11}H_{12}O_7S$	5'-(3',4'-Dihydroxyphenyl)-γ-valerolactone sulfate	Phenylsulfate derivative	3
288.122 : 1.055 (−)	$C_6H_{21}N_6O_5S$	Unknown	Unknown	5
289.116 : 0.980 (−)	$C_{11}H_{22}N_4OS_2$	Unknown	Unknown	5
290.089 : 0.822 (−)	$C_{12}H_{21}NO_3S_2$	Unknown	Unknown	5
295.028 : 1.288 (−)	$C_{12}H_6N_7OS$	Unknown	Unknown	5
295.130 : 0.827 (−)	$C_{14}H_{20}N_2O_5$	Unknown	Unknown	5
302.115 : 0.835 (−)	$C_{15}H_{17}N_3O_4$	Indoleacetyl glutamine	Amino acid derivative	3
303.072 : 0.832 (−)	$C_{11}H_{10}N_7O_4$	Unknown	Unknown	5
308.078 : 0.831 (−)	$C_7H_{15}N_7O_5S$	Unknown	Unknown	5
308.099 : 0.817 (−)	$C_{11}H_{19}NO_9$	N-Acetylneuraminic acid	Carbohydrate conjugate	3
315.108 : 0.813 (−)	$C_{13}H_{22}N_3O_2S_2$	Unknown	Unknown	5
317.124 : 0.814 (−)	$C_{13}H_{16}N_7O_3$	Unknown	Unknown	5
319.140 : 0.812 (−)	$C_{14}H_{24}O_8$	Octanoylglucuronide	Carbohydrate conjugate	3
324.073 : 0.826 (−)	$C_{21}H_{13}N_2S$	Unknown	Unknown	5
331.176 : 0.806 (−)	$C_{15}H_{22}N_7O_2$	Unknown	Unknown	5
343.140 : 0.808 (−)	$C_{23}H_{22}NS$	Unknown	Unknown	5
345.156 : 0.799 (−)	$C_9H_{26}N_6O_6S$	Unknown	Unknown	5
346.056 : 0.827 (−)	$C_{10}H_{14}N_5O_7P$	Unknown	Monosaccharide phosphate	4
347.170 : 0.797 (−)	$C_{14}H_{26}N_3O_7$	Unknown	Unknown	5
350.088 : 0.816 (−)	$C_{16}H_{17}NO_8$	Indole-3-acetic-acid-O-glucuronide	Carbohydrate conjugate	3
351.058 : 1.059 (−)	$C_{13}H_{20}O_7S_2$	Unknown	Unknown	5
352.086 : 1.000 (−)	$C_{16}H_{19}NO_6S$	4-Hydroxybenzyl isothiocyanate 4"-acetylrhamnoside	Carbohydrate conjugate	3
353.034 : 1.208 (−)	$C_{15}H_{14}O_8S$	6'-hydroxy-O-desmethylangolensin-sulfate§	Flavonoid	3
372.113 : 0.880 (−)	$C_9H_{23}N_7O_5S_2$	Unknown	Unknown	5
381.155 : 0.793 (−)	$C_{18}H_{28}N_3O_2S_2$	Unknown	Unknown	5
382.100 : 1.024 (−)	$C_{14}H_{25}NO_7S_2$	Unknown	Unknown	5
387.166 : 0.793 (−)	$C_{17}H_{22}N_7O_4$	Unknown	Unknown	5
396.114 : 0.794 (−)	$C_{12}H_{27}N_7S_4$	Unknown	Unknown	5
397.150 : 0.785 (−)	$C_{18}H_{28}N_3O_3S_2$	Unknown	Unknown	5
399.166 : 0.788 (−)	$C_{18}H_{22}N_7O_4$	Unknown	Unknown	5
405.120 : 0.796 (−)	$C_{20}H_{22}O_9$	Unknown	Phenolic glycoside	4
417.119 : 0.853 (−)	$C_{21}H_{22}O_9$	Equol 7-O-glucuronide§	Isoflavonoid o-glycoside	3
431.098 : 0.872 (−)	$C_{21}H_{20}O_{10}$	Dihydrodaidzein 7-O-glucuronide§	Isoflavonoid o-glycoside	3
443.157 : 0.968 (−)	$C_{20}H_{18}N_{11}O_2$	Unknown	Unknown	5
462.176 : 0.768 (−)	$C_{22}H_{39}O_2S_4$	Unknown	Unknown	5
467.092 : 1.180 (−)	$C_{15}H_{24}N_4O_9S_2$	Unknown	Unknown	5
473.145 : 0.837 (−)	$C_{16}H_{36}N_5OS_5$	Unknown	Unknown	5

Continued on next page

Summary of 167 metabolites detected in urine, including m/z, RMT, ionization mode, molecular formula, compound ID (confirmed or putative), classification, and identification confidence.

m/z : RMT (Q)	Formula*	Compound ID†	Classification	IDC‡
477.106 : 0.967 (−)	$C_{24}H_{22}N_4O_3S_2$	Unknown	Unknown	5
479.229 : 0.778 (−)	$C_{17}H_{38}N_9OS_3$	Unknown	Unknown	5
511.292 : 0.763 (−)	$C_{20}H_{44}N_6O_7S$	Unknown	Unknown	5
525.271 : 0.765 (−)	$C_{19}H_{36}N_{13}O_3S$	Unknown	Unknown	5
539.251 : 0.764 (−)	$C_{27}H_{30}N_{11}O_2$	Unknown	Unknown	5
541.266 : 0.761 (−)	$C_{19}H_{36}N_{13}O_4S$	Unknown	Unknown	5
544.283 : 0.759 (−)	$C_{33}H_{35}N_7O$	Unknown	Unknown	5

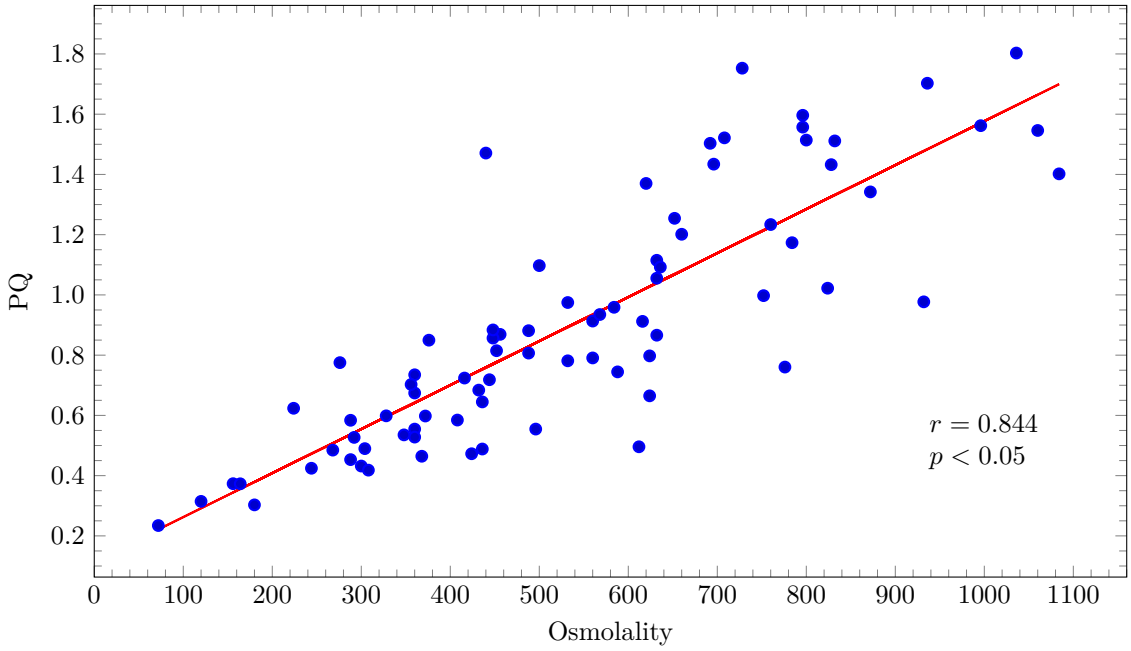
* The most probable formula is presented for tentatively identified compounds and unknowns.

† Stereochemistry (D/L) not confirmed; presumed to be the biologically active isomer (i.e., mainly L-amino acids and D-sugars).

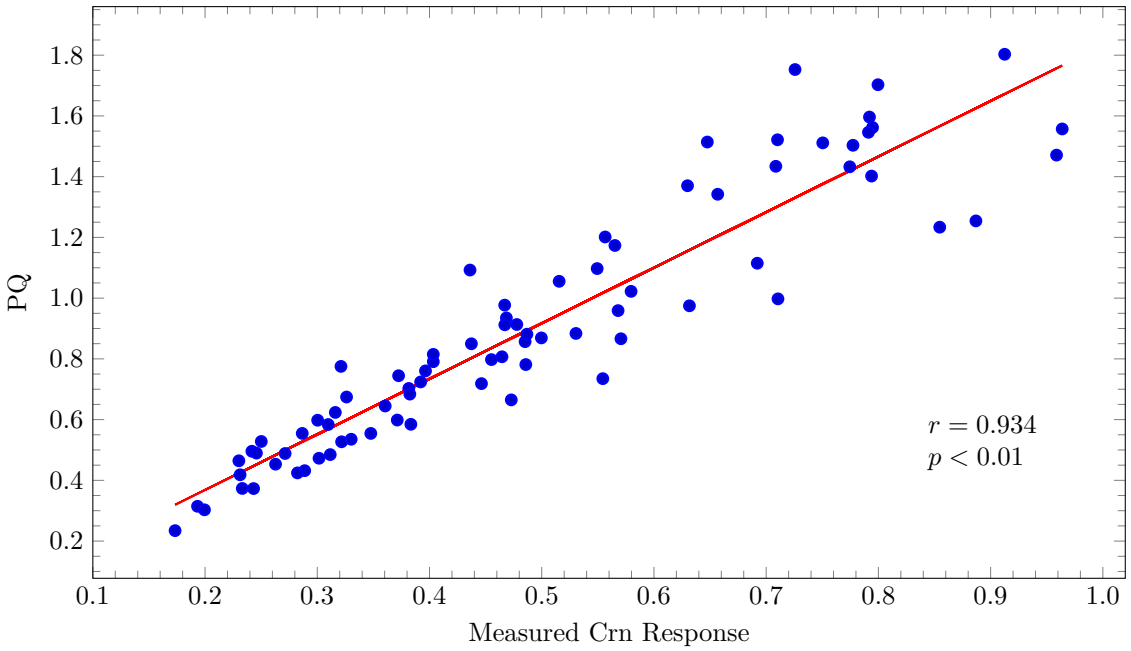
‡ Identification confidence (IDC), modified from the classification system proposed by the Metabolomics Standards Initiative:

1. Confirmed identity: based on comparison to authentic standards.
2. Putative identity: supported by unambiguous matching to MS/MS spectra in literature/database; inferred based on physicochemical properties, biological context and/or statistical associations.
3. Putatively annotated compound: deduced based on biological context of sample and relevance to expected metabolic pathways.
4. Putatively annotated compound class: supported by similarity to predicted spectra; unable to confirm exact positional isomer.
5. Unknown compound: candidate structure is not available or cannot be inferred.

§ Compound excluded from data matrix for not being detected in at least 75% of the samples.



(a) Correlation plot of average PQ vs. osmolality.



(b) Correlation plot of average PQ vs. average measured Crn response.

Figure 4.15: Correlation plots of: (a) average PQ vs. osmolality; and (b) average PQ vs. average creatinine response.

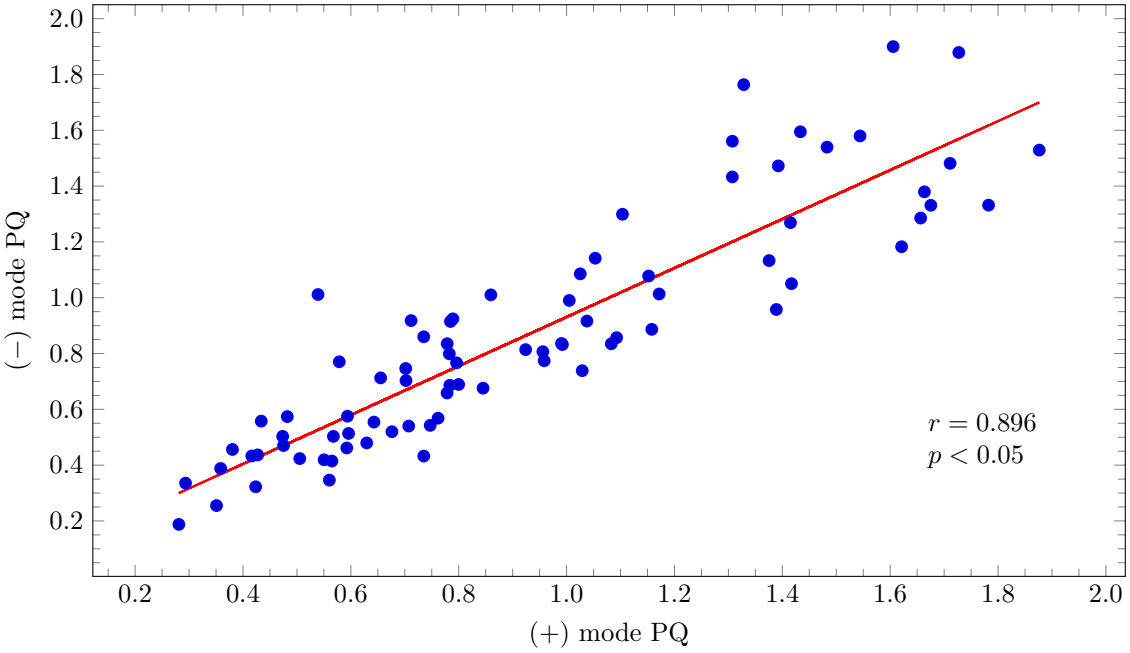


Figure 4.16: Correlation plot of PQ measured in negative-ion mode vs. PQ measured in positive-ion mode shows a strong correlation between both modes of analysis, which is indicative of the robustness of PQ determination.

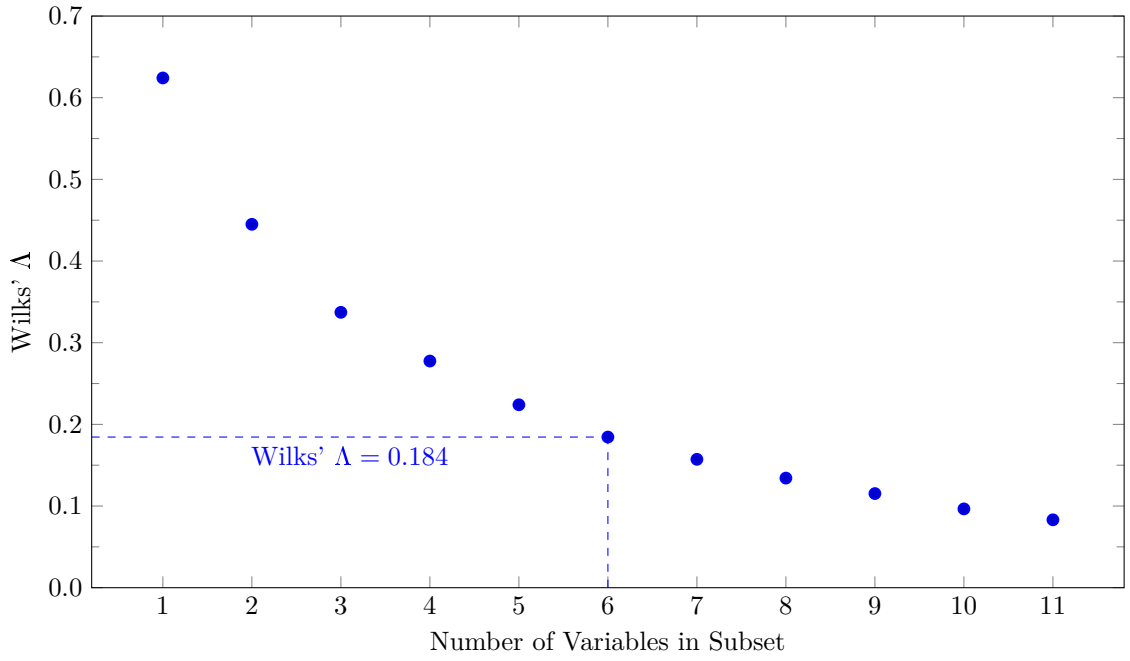


Figure 4.17: Plot of lowest Wilks' Λ vs. subset size. The optimal balance of low Wilks' Λ and low subset size was determined to be achieved with $n = 6$ variables.

Table 4.8: Classification results from DFA show that 82.5% of subjects are correctly classified into the 3 treatment groups based on fold-change levels of 6 top-ranked urinary metabolites. With leave-one-out-cross-validation procedure, 77.5% accuracy in classification is achieved.

		Grouping	Predicted Group Membership			Total
			CHO	SOY	WHY	
Original*	Count	CHO	10	1	1	12
		SOY	1	12	1	14
		WHY	3	0	11	14
	%	CHO	83.3	8.3	8.3	100.0
		SOY	7.1	85.7	7.1	100.0
		WHY	21.4	0.0	78.6	100.0
Cross-validated†‡	Count	CHO	9	2	1	12
		SOY	1	12	1	14
		WHY	3	1	10	14
	%	CHO	75.0	16.7	8.3	100.0
		SOY	7.1	85.7	7.1	100.0
		WHY	21.4	7.1	71.4	100.0

* 82.5% of original grouped cases correctly classified.
† 77.5% of cross-validated grouped cases correctly classified.
‡ Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

Table 4.9: Amino acid composition of whey and soy protein isolates (expressed per 100 g).

Amino Acid	Soy protein isolate*	Whey protein isolate†
Alanine	3.589	6.00
Arginine	6.670	2.65
Aspartic acid	10.203	9.00
Cystine	1.046	3.10
Glutamic acid	17.452	13.00
Glycine	3.603	2.35
Histidine‡	2.303	1.35
Isoleucine‡§	4.253	5.90
Leucine‡§	6.783	13.00
Lysine‡	5.327	9.15
Methionine‡	1.130	2.05
Phenylalanine‡	4.593	2.30
Proline	4.960	4.80
Serine	4.593	5.00
Threonine‡	3.137	6.25
Tryptophan‡	1.116	1.50
Tyrosine	3.222	3.15
Valine‡§	4.098	5.35

* Reference: U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory (2016).
† Reference: U.S. Dairy Export Council (2008).
‡ Essential amino acid.
§ Branched-chain amino acid (BCAA).

Chapter 5

Research Contributions and Perspectives

In summary, MSI-CE-MS offers a high-throughput analytical platform for biomarker discovery in metabolomics, which has been demonstrated in this thesis through the comprehensive targeted and untargeted analysis of biologically relevant metabolites in various biological samples, including protein-bound oxidized disulfides and phytochemicals reflecting habitual diet. Though regarded by many as still being in its infancy when compared to more well-established instrumental platforms, such as NMR, GC-MS, and LC-MS, CE-MS has steadily gained recognition in the past decade as a promising tool in the field of metabolomics due to its ability to perform high-efficiency separations of polar/ionic metabolites and their isomers in volume-limited biospecimens with low operating costs and minimal organic solvent. Motivated by the critical need to address health concerns as obesity becomes a growing epidemic worldwide, the studies presented in this thesis aim to use metabolomics to better understand and elucidate metabolic responses of obese subjects towards dietary interventions, including caloric restriction strategies that are supplemented with protein. Additionally, this

thesis was also focused on expanding metabolome coverage of labile reduced thiols that are covalently bound to plasma protein (i.e., human serum albumin) when using selective maleimide chemical derivatization in conjunction with MSI-CE-MS. Overall, major technical challenges were encountered and successfully overcome in this thesis, which contributed to the discovery of novel protein-bound thiols, as well as specific biomarkers of dietary interventions in overweight/obese subjects when performing metabolomic studies of urine and plasma.

The work in Chapter 2 presented a novel sensitive CE-MS method, which was developed for the reliable determination of protein-bound thiols in plasma using NTBM as a simple thiol-specific maleimide derivatization reagent due to its:

1. greater signal enhancement (of up to 50-fold) relative to underivatized free thiols;
2. improved simplicity in the derivatization procedure without the need for quenching excess maleimide to eliminate background interferences;
3. faster reaction kinetics, which reduces processing time; and
4. much greater long-term chemical stability.

With this optimized pre-column derivatization method, the detection sensitivity of Cys, Hcy, and GSH was 2-fold greater when derivatized by NTBM as compared to previous maleimide analogs, with an average 20-fold greater ionization response relative to native free thiols. Chemical derivatization by NTBM also allowed for the determination of protein-bound γ-GluCys, which is seldom reported in the literature due to its inherent low abundance in biological samples and inability to be detected by traditional LC with UV absorbance- or fluorescence detection. The developed thiol derivatization protocol for plasma protein yielded acceptable technical precision, such that the coefficient of variation was within 15% for abundant protein thiols (i.e.,

Cys, CysGly, and Hcy) and 30% for low-abundance species (i.e., GSH and γ-GluCys). Since protein lost during sample processing cannot be determined or accounted for, this has resulted in large discrepancies in reported absolute thiol measurements between methods in several publications (Andersson et al., 1993; Bald et al., 2004; Borowczyk et al., 2015; Giustarini et al., 2005; Mansoor et al., 1992). As a result, it is recommended that relative quantification be used as the preferred reporting metric for the determination of protein thiols. By combining multiplexed CE-MS technology with the thiol-specific derivatization method, an innovative strategy developed for the discovery of unknown thiols bound to plasma protein was also presented, in which 8 unknown thiol compounds were detected, such that 6 were detected as thiol-NTBM adducts and 2 were detected as reduced free thiols. Due to the ease with which serial sample injections in MSI-CE-MS can be designed based on the specific experimental design (e.g., differential analysis and identification of labeled protein-bond thiols), the presented multiplexed separation strategy may also be widely suitable for other screening and discovery applications in MS-based metabolomic studies. Finally, high-resolution MS/MS experiments on the five known thiol-NTBM derivatives revealed trends in fragmentation pathways which will facilitate the structural elucidation of unknown thiol-NTBM derivatives in future studies.

In Chapter 3, the differences in baseline serum and postprandial metabolic responses between individuals varying in cardiometabolic risk (i.e., LH, MHO, and MUO) were evaluated before and after a standardized high-fat and high-caloric meal using a targeted metabolomics approach. Generally, it was found that MHO individuals possess a greater adaptive response to the dietary challenge relative to MUO, wherein their postprandial alterations in amino acid profiles were intermediate to those of LH and MUO. In particular, differential postprandial changes in asparagine and glutamine between the groups, which are important precursors to TCA cycle intermediates, sug-

gest that LH, MHO, and MUO may exhibit differences in the metabolite flux entering or leaving the TCA cycle, which may have important downstream implications in the efficiency of energy metabolism. Additionally, the plasma carnitine-to-acetylcarnitine ratio, which is reflective of the current catabolic or anabolic state of an individual, was interestingly found to have the highest postprandial change in the LH group compared to both obese groups. These findings suggested that the metabolism of LH individuals were most evidently switched to an anabolic state in response to the caloric challenge. Furthermore, this study has shown correlations between fasting levels of BCAA and indices of insulin sensitivity, which is in support of previous findings from literature that indicate the potential of BCAA to be predictors of insulin resistance and development of type 2 diabetes, independently of BMI (McCormack et al., 2013; Wang et al., 2011; Würtz et al., 2013). Therefore, amino acid profiles may be useful in diabetes risk assessment, which could be advantageous in identifying "at risk" individuals who could benefit earlier from personalized treatments or interventions. Overall, the work from Chapter 3 has highlighted the ability of postprandial measurements to reveal subtle differences in metabolism that may be applied towards cardiometabolic risk assessment among metabolically distinctive sub-groups of obese subjects.

Finally, Chapter 4 analyzed the effects of protein supplementation following a short-term hypo-caloric dietary intervention, which were measured in paired urine samples (i.e., baseline and post-treatment for each subject) using a comprehensive untargeted metabolite profiling approach by MSI-CE-MS. Through this work, 167 unique urinary cationic and anionic metabolites were consistently detected from a cohort of 40 overweight/obese participants with adequate technical precision (RSD $< 30\%$) and frequency ($> 75\%$ of all urine samples) after rigorous data filtering to reject spurious signals and redundant ions (i.e., salt adducts, in-source fragments, isotopes, etc.). By applying a robust feature selection strategy in conjunction with multivariate parametric

statistical methods rarely implemented in conventional metabolomic studies, six urinary metabolites were identified as top-ranked biomarkers for discriminating between subjects consuming carbohydrate-, soy-, or whey-supplemented diets. In support of earlier research which compared the effects of soy and whey supplementation on myofibrillar protein synthesis (Baer et al., 2011; Hector et al., 2015), subjects consuming the whey diet had the lowest urinary excretion of 3-methylhistidine, which indicated that muscle loss was most attenuated with whey supplementation as compared to soy or carbohydrate control. Through the process of maintaining quality assurance and following strict criteria for rigorous metabolite screening, it was also determined that four urinary metabolites were only occasionally detected among specific subset of samples, which were later found to be dietary biomarkers specific to the soy diet. While three of these sporadic urinary metabolites are believed to be bacterially-derived metabolic products of the isoflavonoid phytoestrogen, daidzein, as supported by their respective proportion of detection in the soy subjects, the other molecular feature has tentatively been identified as 4-ethylphenyl sulfate through comparison to literature MS/MS data, which is the sulfate conjugate of a bacterial metabolic product of a different isoflavone species, genistein. Though its precursor, 4-ethylphenol, has previously been quantified in human blood and urine, 4-ethylphenyl sulfate has so far only been detected in mice, thereby, making its detection in this study the first reported occurrence in human urine. Additionally, this chapter has also unexpectedly revealed a strong correlation between the unmetabolized artificial sweetener, acesulfame potassium, and the marker of dietary protein intake, guanidinosuccinic acid. Since it is known that protein intake increases glomerular filtration rate, the correlation of acesulfame potassium with guanidinosuccinic acid is believed to be explained by its increased excretion due to increased glomerular filtration rate induced by protein intake. As a result, these findings indicate that acesulfame potassium could potentially

serve as an exogenous marker of renal function or glomerular filtration rate in humans, which has not yet been reported in the literature.

Chapter 6

Future Directions

6.1 Structural Elucidation of Unknown Protein-Bound Thiols by MS/MS Studies

The novel detection of 8 unknown thiol compounds using the MSI-CE-MS method in the work of Chapter 2 is intriguing and can potentially lead to further studies to investigate their identity and origin. By performing neutral loss scan and MS/MS experiments, comparison of the acquired spectra to MS/MS spectra of known thiol-NTBM adducts with known MS fragmentation patterns will enable structural elucidation of unknown thiol-NTBM adducts. Although recent literature in the area of thiol analysis have been focused on:

1. the determination of thiol compounds arising from the metabolism of sulfur-containing drugs, such as vicagrel, clopidogrel, tiopronin and penicillamine (Cao et al., 2018; Kavitha et al., 2018; Liu et al., 2018),

2. the discovery of novel compounds produced through the conjugation of endogenous thiols to exogenous or dietary metabolites (Nagakubo et al., 2018),

and

3. the identification of microbial-derived volatile thiols contributing to fruit aromas in the production of wine and beer (Inoue et al., 2013; Takoi et al., 2009),

no work has yet been done to investigate novel thiols bound to protein in plasma, which may possibly arise from disulfide formation with exogenous thiol compounds originating from drugs or dietary sources. Given that Nakashima et al. (2018) recently discovered that protein-thiolation could occur at multiple cysteine residues in HSA of hyperlipidemia patients in addition to the free cysteine residue at position 34 (Cys34) that is already known to be a target of oxidation, it is likely that multiple sites for thiol-binding in HSA become available under certain circumstances to allow for the additional binding of exogenous thiol compounds, which are expected to be less abundant than cysteine in the extracellular matrix. Furthermore, since the diverse multidomain structure of HSA allows for the binding and transport of various endogenous and exogenous compounds (such as metals, vitamins, hormones, bilirubin, fatty acids, amino acids, and drugs) (Turell et al., 2013), the processing and analysis of HSA, therefore, also includes the analysis of its bound substrates, peptides and proteins (i.e., "albuminome") (Gundry et al., 2007). As HSA represents the most abundant protein in plasma, the majority (70–90%) of protein-bound thiols and substrates are, thus, linked to HSA as compared to globulins (Hortin et al., 2006). With this information, the novel thiols detected in Chapter 2 are expected to be most likely associated with HSA. Once the unknown thiols have been identified, it would be of interest to compare the abundances and species of protein-bound thiols detected in the plasma between healthy subjects and those at risk for metabolic syndrome, which will provide new insights on the impact of potentially altered structure and functionality of plasma proteins on the development of risk factors and disease.

Additionally, it would be valuable to determine how the plasma thiol redox status impacts both the magnitude and type of thiolation observed in the protein fraction. To achieve this, the method developed in this thesis for the stabilization and enhanced detection of thiols may be adapted in future studies for improved simultaneous determination of free reduced thiols and oxidized disulfides for the measurement of thiol redox status in plasma, which is implicated in the process of aging and development of disease, including obesity. Currently, thiol redox status assessment in plasma remains challenging due to its susceptibility to artifactual oxidation during sample collection and processing. To minimize auto-oxidation during sample preparation, collection of blood plasma into tubes containing maleimide reagent prior to sample storage should be explored in future studies.

6.2 Further Evaluation of Postprandial Effects in Diabetes Risk Assessment

Although the results presented in Chapter 3 have shown how postprandial changes were able to differentiate between individuals of different metabolic phenotypes after a two-hour period, future studies should consider additional sampling at longer time points to measure time-based changes in the plasma amino acid profile, which may vary between the subject groups. In addition to evaluating a larger sample size and replication of the study among different cohorts for validation, follow-up studies could also aim to include metabolically unhealthy lean individuals, who, despite having normal weight and BMI, are at three-fold greater risk of cardiovascular diseases compared to lean healthy individuals and represent roughly 20% of the lean adult population (Stefan et al., 2017). Comparing their postprandial metabolism with those

of lean healthy and obese subjects may offer new evidence for the understanding and treatment of "at risk" metabolic phenotypes in general, regardless of the lean or obese state, either through the discovery of novel mechanisms or the unexpected finding of metabolic commonalities to those of unhealthy obese individuals. Furthermore, since fasting levels of BCAA were shown in this work to have the potential to identify "at risk" obese individuals through positive correlations with HOMA-IR and fasting insulin levels, future studies should strive to confirm the findings with a larger subject cohort for increased study power, along with the inclusion of unhealthy lean individuals in order to assess the sensitivity of BCAA levels in the identification of metabolically unhealthy individuals (lean or obese), independently of BMI. Findings from these studies would be highly relevant to the understanding and treatment of metabolic and cardiovascular diseases given that BCAAs have recently attracted much interest due to their association with insulin resistance and emergence as potential biomarkers of metabolic syndrome.

6.3 Long-Term Assessment of Soy or Whey Supplementation During Caloric Restriction

The identification of six top-ranking urinary metabolites in Chapter 4, which collectively discriminated between responses to soy or whey supplementation relative to controls during a two-week weight loss intervention, has demonstrated that changes resulting from nutritional interventions are complex, yet subtle – such that a single metabolite alone is not sufficient to provide group discrimination and explain differences in underlying biochemical mechanisms. In order to confirm the relevance of these six biomarkers in response to the soy and whey supplementation, future

studies should involve a longer-term dietary intervention period with a larger sample size for study replication. Due to urinary metabolite levels being susceptible to variations in estimated GFR, which can vary with age, between individuals, as well as be influenced by dietary protein intake, 24-hour urine samples should ideally be collected in follow-up studies. While the collection of single-spot urine samples in the current study allowed for easy sample handling and convenient sample collection, 24-hour urine samples will permit measurement of daily total excreted metabolite levels without being influenced by temporal differences in the estimated GFR. If necessary, correcting for variations in estimated GFR may be performed through the measurement of creatinine clearance from the 24-hour urine samples. Alternatively, if single-spot urine is still preferred for future studies, then GFR may be estimated through measurement of blood creatinine level. Since the work had also revealed the potential of Ace-K excretion to be linked to GFR or renal function, it would be of interest to confirm this finding given that Ace-K is rapidly absorbed upon ingestion and excreted unmetabolized by the kidney within 24 hours (Magnuson et al., 2016). Aside from the collection of urine samples, which are known to be more sensitive to changes in diet, lifestyle, and environmental factors, fasting blood samples could also be collected to identify those changes relevant to energy metabolism that are allowed to occur within the tightly regulated environment of homeostatic control. Finally, as the work shown here has putatively identified the presence of 4-EPS in human urine for the first time based on unambiguous matching between the experimental tandem mass spectra with data found in the literature, further work is needed to confirm the compound identity through MS/MS and migration time measurements of an authentic standard. Currently, 4-EPS is recognized as an uremic toxin, in which its detection in plasma is due its abnormal accumulation in patients with CKD or renal failure (Tanaka et al., 2015). With evidence suggesting it to be derived from

intestinal gut microbial metabolism (Kikuchi et al., 2010), 4-EPS is expected to be eliminated by the kidneys and is, therefore, an expected metabolite of human urine. However, perhaps, due to low concentrations, 4-EPS has not yet been detected in human urine. The finding of 4-EPS in this thesis as a urinary metabolite specific to the soy intervention is important to confirm because it supports previous work in the literature, which linked the presence of 4-EPS in human serum and blood to the consumption of tofu and soymilk (Guertin et al., 2014; Pallister et al., 2016), and demonstrates how diet and microflora-derived metabolites can contribute to notable differences between treatment groups in dietary intervention studies. Given the host-microbiome relationship is involved in the co-metabolism of many dietary compounds, integration of metabolomic studies with microbiome analysis should be considered in future studies to better understand dynamic interactions between gut microflora, diet, and host metabolism. Lastly, future MS/MS experiments on an unknown compound, m/z : RMT 271.104 : 0.768 (+), which was found to be significantly correlated to 3-MeHis and is tentatively annotated to be a dipeptide formed from aspartic acid and histidine, should be performed for structural identification in future studies, due to its possible involvement in the process of protein catabolism and muscle protein turnover, which may further our understanding of differences in energy metabolism arising from different protein supplementations in weight-loss interventions.

References

Abe, K., Suzuki, H., Maekawa, M., Shimada, M., Yamaguchi, H., and Mano, N. (2017). Matrix effect-corrected liquid chromatography/tandem mass-spectrometric method for determining acylcarnitines in human urine. *Clinica Chimica Acta*, 468:187–194, DOI: 10.1016/j.cca.2017.03.001.

Adams, S. H. (2011). Emerging Perspectives on Essential Amino Acid Metabolism in Obesity and the Insulin-Resistant State. *Advances in Nutrition: An International Review Journal*, 2(6):445–456, DOI: 10.3945/an.111.000737.

Aitken, A. and Learmonth, M. (2009). Estimation of Disulfide Bonds Using Ellman's Reagent. In Walker, J. M., editor, *The Protein Protocols Handbook*, Springer Protocols Handbooks, pages 1053–1055. Humana Press, Totowa, NJ, ISBN: 978-1-60327-474-6, DOI: 10.1007/978-1-59745-198-7.

Albert, A., Brauckmann, C., Blaske, F., Sperling, M., Engelhard, C., and Karst, U. (2012). Speciation analysis of the antirheumatic agent Auranofin and its thiol adducts by LC/ESI-MS and LC/ICP-MS. *Journal of Analytical Atomic Spectrometry*, 27(6):975, DOI: 10.1039/c2ja30109a.

Alcala-Diaz, J. F., Delgado-Lista, J., Perez-Martinez, P., Garcia-Rios, A., Marin, C., Quintana-Navarro, G. M., Gomez-Luna, P., Camargo, A., Almaden, Y., Caballero,

J., Tinahones, F. J., Ordovas, J. M., Perez-Jimenez, F., and Lopez-Miranda, J. (2014). Hypertriglyceridemia Influences the Degree of Postprandial Lipemic Response in Patients with Metabolic Syndrome and Coronary Artery Disease: From the Cordioprev Study. *PLoS ONE*, 9(5):e96297, DOI: 10.1371/journal.pone.0096297.

Aldini, G., Vistoli, G., Regazzoni, L., Gamberoni, L., Facino, R. M., Yamaguchi, S., Uchida, K., and Carini, M. (2008). Albumin Is the Main Nucleophilic Target of Human Plasma: A Protective Role Against Pro-atherogenic Electrophilic Reactive Carbonyl Species? *Chemical Research in Toxicology*, 21(4):824–835, DOI: 10.1021/tx700349r.

Alonso, A., Marsal, S., and Julià, A. (2015). Analytical methods in untargeted metabolomics: state of the art in 2015. *Frontiers in Bioengineering and Biotechnology*, 3:23, DOI: 10.3389/fbioe.2015.00023.

Alter, O., Brown, P. O., and Botstein, D. (2000). Singular value decomposition for genome-wide expression data processing and modeling. *Proceedings of the National Academy of Sciences*, 97(18):10101–10106, DOI: 10.1073/pnas.97.18.10101.

Andersen, C. M. and Bro, R. (2010). Variable selection in regression—a tutorial. *Journal of Chemometrics*, 24(11-12):728–737, DOI: 10.1002/cem.1360.

Andersen, M.-B. S., Rinnan, Å., Manach, C., Poulsen, S. K., Pujos-Guillot, E., Larsen, T. M., Astrup, A., and Dragsted, L. O. (2014). Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. *Journal of Proteome Research*, 13(3):1405–1418, DOI: 10.1021/pr400964s.

Andersson, A., Isaksson, A., Brattström, L., and Hultberg, B. (1993). Homocysteine and other thiols determined in plasma by HPLC and thiol-specific postcolumn

derivatization. *Clinical Chemistry*, 39(8):1590–1597, http://www.ncbi.nlm.nih.gov/pubmed/8353942.

Andreopoulos, B., An, A., Wang, X., and Schroeder, M. (2009). A roadmap of clustering algorithms: finding a match for a biomedical application. *Briefings in Bioinformatics*, 10(3):297–314, DOI: 10.1093/bib/bbn058.

Ang, J. E., Revell, V., Mann, A., Mäntele, S., Otway, D. T., Johnston, J. D., Thumser, A. E., Skene, D. J., and Raynaud, F. (2012). Identification of Human Plasma Metabolites Exhibiting Time-of-Day Variation Using an Untargeted Liquid Chromatography–Mass Spectrometry Metabolomic Approach. *Chronobiology International*, 29(7):868–881, DOI: 10.3109/07420528.2012.699122.

Aoyagi, K., Shahrzad, S., Iida, S., Tomida, C., Hirayama, A., Nagase, S., Takemura, K., Koyama, A., Ohba, S., Narita, M., and Cohen, B. D. (2001). Role of nitric oxide in the synthesis of guanidinosuccinic acid, at activator of the N-methyl-d-asparate receptor. *Kidney International*, 59:S93–S96, DOI: 10.1046/j.1523-1755.2001.59780093.x.

Astarita, G. and Langridge, J. (2013). An Emerging Role for Metabolomics in Nutrition Science. *Journal of Nutrigenetics and Nutrigenomics*, 6(4-5):181–200, DOI: 10.1159/000354403.

Astrup, A. (2005). The satiating power of protein—a key to obesity prevention? *The American Journal of Clinical Nutrition*, 82(1):1–2, http://ajcn.nutrition.org/content/82/1/1.short.

Atallah, R., Filion, K. B., Wakil, S. M., Genest, J., Joseph, L., Poirier, P., Rinfret, S., Schiffrin, E. L., and Eisenberg, M. J. (2014). Long-Term Effects of 4 Popular Diets on Weight Loss and Cardiovascular Risk Factors: A Systematic Review of

Randomized Controlled Trials. *Circulation: Cardiovascular Quality and Outcomes*, 7(6):815–827, DOI: 10.1161/CIRCOUTCOMES.113.000723.

Atkinson, C., Newton, K. M., Bowles, E. J. A., Yong, M., and Lampe, J. W. (2008). Demographic, anthropometric, and lifestyle factors and dietary intakes in relation to daidzein-metabolizing phenotypes among premenopausal women in the United States. *The American Journal of Clinical Nutrition*, 87(3):679–687, http://www.ncbi.nlm.nih.gov/pubmed/18326607.

Backx, E. M. P., Tieland, M., Borgonjen-van den Berg, K. J., Claessen, P. R., van Loon, L. J. C., and de Groot, L. C. P. G. M. (2016). Protein intake and lean body mass preservation during energy intake restriction in overweight older adults. *International Journal of Obesity*, 40(2):299–304, DOI: 10.1038/ijo.2015.182.

Badimon, L., Vilahur, G., and Padro, T. (2017). Systems biology approaches to understand the effects of nutrition and promote health. *British Journal of Clinical Pharmacology*, 83(1):38–45, DOI: 10.1111/bcp.12965.

Badoud, F., Lam, K. P., DiBattista, A., Perreault, M., Zulyniak, M. A., Cattrysse, B., Stephenson, S., Britz-McKibbin, P., and Mutch, D. M. (2014). Serum and Adipose Tissue Amino Acid Homeostasis in the Metabolically Healthy Obese. *Journal of Proteome Research*, 13(7):3455–3466, DOI: 10.1021/pr500416v.

Badoud, F., Lam, K. P., Perreault, M., Zulyniak, M. A., Britz-McKibbin, P., and Mutch, D. M. (2015a). Metabolomics Reveals Metabolically Healthy and Unhealthy Obese Individuals Differ in their Response to a Caloric Challenge. *PLOS ONE*, 10(8):e0134613, DOI: 10.1371/journal.pone.0134613.

Badoud, F., Perreault, M., Zulyniak, M. A., and Mutch, D. M. (2015b). Molecular insights into the role of white adipose tissue in metabolically unhealthy normal weight

and metabolically healthy obese individuals. *The FASEB Journal*, 29(3):748–758, DOI: 10.1096/fj.14-263913.

Baer, D. J., Stote, K. S., Paul, D. R., Harris, G. K., Rumpler, W. V., and Clevidence, B. A. (2011). Whey Protein but Not Soy Protein Supplementation Alters Body Weight and Composition in Free-Living Overweight and Obese Adults. *Journal of Nutrition*, 141(8):1489–1494, DOI: 10.3945/jn.111.139840.

Bald, E., Chwatko, G., Głowacki, R., and Kuśmierek, K. (2004). Analysis of plasma thiols by high-performance liquid chromatography with ultraviolet detection. *Journal of Chromatography A*, 1032(1-2):109–115, DOI: 10.1016/j.chroma.2003.11.030.

Ballor, D. L., Katch, V. L., Becque, M. D., and Marks, C. R. (1988). Resistance weight training during caloric restriction enhances lean body weight maintenance. *The American Journal of Clinical Nutrition*, 47(1):19–25, http://ajcn.nutrition.org/content/47/1/19.abstract.

Bally, L., Bovet, C., Nakas, C. T., Zueger, T., Prost, J.-C., Nuoffer, J.-M., Leichtle, A. B., Fiedler, G. M., and Stettler, C. (2017). A metabolomics approach to uncover effects of different exercise modalities in type 1 diabetes. *Metabolomics*, 13(7):78, DOI: 10.1007/s11306-017-1217-8.

Barnes, S., Prasain, J., and Kim, H. (2013). In Nutrition, Can We "See" What Is Good for Us? *Advances in Nutrition*, 4(3):327S–334S, DOI: 10.3945/an.112.003558.

Barri, T. and Dragsted, L. O. (2013). UPLC-ESI-QTOF/MS and multivariate data analysis for blood plasma and serum metabolomics: Effect of experimental artefacts and anticoagulant. *Analytica Chimica Acta*, 768:118–128, DOI: 10.1016/j.aca.2013.01.015.

Bartel, J., Krumsiek, J., and Theis, F. J. (2013). Statistical Methods for the Analysis of High-Throughput Metabolomics Data. *Computational and Structural Biotechnology Journal*, 4(5):e201301009, DOI: 10.5936/csbj.201301009.

Batch, B. C., Shah, S. H., Newgard, C. B., Turer, C. B., Haynes, C., Bain, J. R., Muehlbauer, M., Patel, M. J., Stevens, R. D., Appel, L. J., Newby, L. K., and Svetkey, L. P. (2013). Branched chain amino acids are novel biomarkers for discrimination of metabolic wellness. *Metabolism*, 62(7):961–969, DOI: 10.1016/j.metabol.2013.01.007.

Bazzano, L. A., Hu, T., Reynolds, K., Yao, L., Bunol, C., Liu, Y., Chen, C.-S., Klag, M. J., Whelton, P. K., and He, J. (2014). Effects of Low-Carbohydrate and Low-Fat Diets. *Annals of Internal Medicine*, 161(5):309, DOI: 10.7326/M14-0180.

Bedford, M. T. and Clarke, S. G. (2009). Protein Arginine Methylation in Mammals: Who, What, and Why. *Molecular Cell*, 33(1):1–13, DOI: 10.1016/j.molcel.2008.12.013.

Begou, O., Gika, H. G., Wilson, I. D., and Theodoridis, G. (2017). Hyphenated MS-based targeted approaches in metabolomics. *Analyst*, 142(17):3079–3100, DOI: 10.1039/C7AN00812K.

Beleggia, R., Platani, C., Papa, R., Di Chio, A., Barros, E., Mashaba, C., Wirth, J., Fammartino, A., Sautter, C., Conner, S., Rauscher, J., Stewart, D., and Cattivelli, L. (2011). Metabolomics and Food Processing: From Semolina to Pasta. *Journal of Agricultural and Food Chemistry*, 59(17):9366–9377, DOI: 10.1021/jf2022836.

Belobrajdic, D. P., McIntosh, G. H., and Owens, J. A. (2004). A High-Whey-Protein Diet Reduces Body Weight Gain and Alters Insulin Sensitivity Relative to Red Meat in Wistar Rats. *The Journal of Nutrition*, 134(6):1454–1458, http://jn.nutrition.org/content/134/6/1454.abstract.

Benito, M., Parker, J., Du, Q., Wu, J., Xiang, D., Perou, C. M., and Marron, J. S. (2004). Adjustment of systematic microarray data biases. *Bioinformatics*, 20(1):105–114, DOI: 10.1093/bioinformatics/btg385.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, http://www.jstor.org/stable/2346101.

Bhagwat, S., Haytowitz, D. B., and Holden, J. M. (2008). USDA Database for the Isoflavone Content of Selected Foods, Release 2. Technical report, United States Department of Agriculture, https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/nutrient-data-laboratory/docs/usda-database-for-the-isoflavone-content-of-selected-foods-release-20/.

Bhupathiraju, S. N. and Hu, F. B. (2017). One (small) step towards precision nutrition by use of metabolomics. *The Lancet Diabetes & Endocrinology*, 5(3):154–155, DOI: 10.1016/S2213-8587(17)30007-4.

Bian, X., Chi, L., Gao, B., Tu, P., Ru, H., and Lu, K. (2017). The artificial sweetener acesulfame potassium affects the gut microbiome and body weight gain in CD-1 mice. *PLOS ONE*, 12(6):e0178426, DOI: 10.1371/journal.pone.0178426.

Bieniek, J. M., Drabovich, A. P., and Lo, K. C. (2016). Seminal biomarkers for the evaluation of male infertility. *Asian Journal of Andrology*, 18(3):426–433, DOI: 10.4103/1008-682X.175781.

Biolo, G., Ciocchi, B., Stulle, M., Bosutti, A., Barazzoni, R., Zanetti, M., Antonione, R., Lebenstedt, M., Platen, P., Heer, M., and Guarnieri, G. (2007). Calorie restriction

accelerates the catabolism of lean body mass during 2 wk of bed rest. *The American Journal of Clinical Nutrition*, 86(2):366–372, http://www.ncbi.nlm.nih.gov/pubmed/17684207.

Boccard, J., Badoud, F., Grata, E., Ouertani, S., Hanafi, M., Mazerolles, G., Lantéri, P., Veuthey, J.-L., Saugy, M., and Rudaz, S. (2011). A steroidomic approach for biomarkers discovery in doping control. *Forensic Science International*, 213(1-3):85–94, DOI: 10.1016/j.forsciint.2011.07.023.

Bode-Böger, S. M., Scalera, F., Kielstein, J. T., Martens-Lobenhoffer, J., Breithardt, G., Fobker, M., and Reinecke, H. (2006). Symmetrical Dimethylarginine: A New Combined Parameter for Renal Function and Extent of Coronary Artery Disease. *Journal of the American Society of Nephrology*, 17(4):1128–1134, DOI: 10.1681/ASN.2005101119.

Boizard, F., Brunchault, V., Moulos, P., Breuil, B., Klein, J., Lounis, N., Caubet, C., Tellier, S., Bascands, J.-L., Decramer, S., Schanstra, J. P., and Buffin-Meyer, B. (2016). A capillary electrophoresis coupled to mass spectrometry pipeline for long term comparable assessment of the urinary metabolome. *Scientific Reports*, 6(1):34453, DOI: 10.1038/srep34453.

Bondia-Pons, I., Maukonen, J., Mattila, I., Rissanen, A., Saarela, M., Kaprio, J., Hakkarainen, A., Lundbom, J., Lundbom, N., Hyötyläinen, T., Pietiläinen, K. H., and Orešič, M. (2014). Metabolome and fecal microbiota in monozygotic twin pairs discordant for weight: a Big Mac challenge. *The FASEB Journal*, 28(9):4169–4179, DOI: 10.1096/fj.14-250167.

Bonvin, G., Schappler, J., and Rudaz, S. (2012). Capillary electrophoresis–electrospray ionization-mass spectrometry interfaces: Fundamental concepts

and technical developments. *Journal of Chromatography A*, 1267:17–31, DOI: 10.1016/j.chroma.2012.07.019.

Borowczyk, K., Wyszczelska-Rokiel, M., Kubalczyk, P., and Głowacki, R. (2015). Simultaneous determination of albumin and low-molecular-mass thiols in plasma by HPLC with UV detection. *Journal of Chromatography B*, 981-982:57–64, DOI: 10.1016/j.jchromb.2014.12.032.

Bos, C., Metges, C. C., Gaudichon, C., Petzke, K. J., Pueyo, M. E., Morens, C., Everwand, J., Benamouzig, R., and Tomé, D. (2003). Postprandial Kinetics of Dietary Amino Acids Are the Main Determinant of Their Metabolism after Soy or Milk Protein Ingestion in Humans. *The Journal of Nutrition*, 133(5):1308–1315, http://www.ncbi.nlm.nih.gov/pubmed/12730415.

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., and Wishart, D. S. (2013). The Human Urine Metabolome. *PLoS ONE*, 8(9):e73076, DOI: 10.1371/journal.pone.0073076.

Bramanti, E., D'Ulivo, L., Lomonte, C., Onor, M., Zamboni, R., Raspi, G., and D'Ulivo, A. (2006). Determination of hydrogen sulfide and volatile thiols in air samples by mercury probe derivatization coupled with liquid chromatography–atomic fluorescence spectrometry. *Analytica Chimica Acta*, 579(1):38–46, DOI: 10.1016/j.aca.2006.07.004.

Brändle, E., Sieberth, H. G., and Hautmann, R. E. (1996). Effect of chronic dietary protein intake on the renal function in healthy subjects. *European Journal of Clinical Nutrition*, 50(11):734–740, http://www.ncbi.nlm.nih.gov/pubmed/8933120.

Brauer, R., Leichtle, A. B., Fiedler, G. M., Thiery, J., and Ceglarek, U. (2011). Preanalytical standardization of amino acid and acylcarnitine metabolite profiling in human blood using tandem mass spectrometry. *Metabolomics*, 7(3):344–352, DOI: 10.1007/s11306-010-0256-1.

Bray, G. A. and Popkin, B. M. (2014). Dietary Sugar and Body Weight: Have We Reached a Crisis in the Epidemic of Obesity and Diabetes? *Diabetes Care*, 37(4):950–956, DOI: 10.2337/dc13-2085.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32, DOI: 10.1023/A:1010933404324.

Brennan, L. (2013). Metabolomics in nutrition research: current status and perspectives. *Biochemical Society Transactions*, 41(2):670–673, DOI: 10.1042/BST20120350.

Brennan, L. (2017). Use of metabotyping for optimal nutrition. *Current Opinion in Biotechnology*, 44:35–38, DOI: 10.1016/j.copbio.2016.10.008.

Brosnan, J. T. and Brosnan, M. E. (2006). The Sulfur-Containing Amino Acids: An Overview. *The Journal of Nutrition*, 136(6):1636S–1640S, DOI: 10.1093/jn/136.6.1636S.

Brown, S. A. (2015). Symmetric dimethylarginine (SDMA): new biomarker of renal function in cats and dogs. Technical report, International Renal Interest Society, http://www.iris-kidney.com/pdf/003-5559.001-iris-website-symmetric-dimethylarginine-pdf_220116-final.pdf.

Brunius, C., Shi, L., and Landberg, R. (2016). Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-

based within-batch signal intensity drift correction. *Metabolomics*, 12(11):173, DOI: 10.1007/s11306-016-1124-4.

Burt, T. and Nandal, S. (2016). Pharmacometabolomics in Early-Phase Clinical Development. *Clinical and Translational Science*, 9(3):128–138, DOI: 10.1111/cts.12396.

Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., and Trygg, J. (2006). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20(8-10):341–351, DOI: 10.1002/cem.1006.

Cabrera, O., Jacques-Silva, M. C., Speier, S., Yang, S.-N., Köhler, M., Fachado, A., Vieira, E., Zierath, J. R., Kibbey, R., Berman, D. M., Kenyon, N. S., Ricordi, C., Caicedo, A., and Berggren, P.-O. (2008). Glutamate Is a Positive Autocrine Signal for Glucagon Release. *Cell Metabolism*, 7(6):545–554, DOI: 10.1016/j.cmet.2008.03.004.

Caesar, L. K., Kvalheim, O. M., and Cech, N. B. (2018). Hierarchical cluster analysis of technical replicates to identify interferents in untargeted mass spectrometry metabolomics. *Analytica Chimica Acta*, 1021:69–77, DOI: 10.1016/j.aca.2018.03.013.

Cai, J. and Henion, J. (1995). Capillary electrophoresis-mass spectrometry. *Journal of Chromatography A*, 703(1-2):667–692, DOI: 10.1016/0021-9673(94)01178-H.

Cajka, T. and Fiehn, O. (2016). Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Analytical Chemistry*, 88(1):524–545, DOI: 10.1021/acs.analchem.5b04491.

Camera, E. and Picardo, M. (2002). Analytical methods to investigate glutathione and related compounds in biological and pathological processes. *Journal of Chromatography B*, 781(1-2):181–206, DOI: 10.1016/S1570-0232(02)00618-9.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada (2010). Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans.

Cao, L., Wei, T., Shi, Y., Tan, X., and Meng, J. (2018). Determination of D-penicillamine and tiopronin in human urine and serum by HPLC-FLD and CE-LIF with 1,3,5,7-tetramethyl-8-bromomethyl-difluoroboradiaza-s-indacene. *Journal of Liquid Chromatography & Related Technologies*, 41(2):58–65, DOI: 10.1080/10826076.2017.1348953.

Carballal, S., Radi, R., Kirk, M. C., Barnes, S., Freeman, B. A., and Alvarez, B. (2003). Sulfenic Acid Formation in Human Serum Albumin by Hydrogen Peroxide and Peroxynitrite. *Biochemistry*, 42(33):9906–9914, DOI: 10.1021/bi027434m.

Carlucci, F. and Tabucchi, A. (2009). Capillary electrophoresis in the evaluation of aminothiols in body fluids. *Journal of Chromatography B*, 877(28):3347–3357, DOI: 10.1016/j.jchromb.2009.07.030.

Carmen, A. and Hardiman, G., editors (2006). *Biochips as Pathways to Drug Discovery*. CRC Press, Boca Raton, FL, ISBN: 978-1-4200-1560-7.

Carroll, J. E., Brooke, M. H., Shumate, J. B., and Janes, N. J. (1981). Carnitine intake and excretion in neuromuscular diseases. *The American Journal of Clinical Nutrition*, 34(12):2693–2698, http://www.ncbi.nlm.nih.gov/pubmed/7315771.

Cazes, J., editor (2001). *Encyclopedia of Chromatography*. Marcel Dekker, Inc., New York, NY, ISBN: 0-8247-0511-4.

Chadha, V., Garg, U., and Alon, U. S. (2001). Measurement of urinary concentration:

a critical appraisal of methodologies. *Pediatric Nephrology*, 16(4):374–382, DOI: 10.1007/s004670000551.

Chalcraft, K. R., Lee, R., Mills, C., and Britz-McKibbin, P. (2009). Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Efficiency Without Chemical Standards. *Analytical Chemistry*, 81(7):2506–2515, DOI: 10.1021/ac802272u.

Chalmers, R. A., Roe, C. R., Stacey, T. E., and Hoppel, C. L. (1984). Urinary Excretion of l-Carnitine and Acylcarnitines by Patients with Disorders of Organic Acid Metabolism: Evidence for Secondary Insufficiency of l-Carnitine. *Pediatric Research*, 18(12):1325–1328, DOI: 10.1203/00006450-198412000-00021.

Chang, Y. and Choue, R. (2013). Plasma pharmacokinetics and urinary excretion of isoflavones after ingestion of soy products with different aglycone/glucoside ratios in South Korean women. *Nutrition Research and Practice*, 7(5):393–399, DOI: 10.4162/nrp.2013.7.5.393.

Chen, J., Meng, Y., Zhou, J., Zhuo, M., Ling, F., Zhang, Y., Du, H., and Wang, X. (2013). Identifying Candidate Genes for Type 2 Diabetes Mellitus and Obesity through Gene Expression Profiling in Multiple Tissues or Cells. *Journal of Diabetes Research*, 2013:1–9, DOI: 10.1155/2013/970435.

Chen, K., Rannulu, N. S., Cai, Y., Lane, P., Liebl, A. L., Rees, B. B., Corre, C., Challis, G. L., and Cole, R. B. (2008). Unusual Odd-Electron Fragments from Even-Electron Protonated Prodiginine Precursors Using Positive-Ion Electrospray Tandem Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, 19(12):1856–1866, DOI: 10.1016/j.jasms.2008.08.002.

Cheng, S., Rhee, E. P., Larson, M. G., Lewis, G. D., McCabe, E. L., Shen, D., Palma, M. J., Roberts, L. D., Dejam, A., Souza, A. L., Deik, A. A., Magnusson, M., Fox, C. S., O'Donnell, C. J., Vasan, R. S., Melander, O., Clish, C. B., Gerszten, R. E., and Wang, T. J. (2012). Metabolite Profiling Identifies Pathways Associated with Metabolic Risk in Humans. *Circulation*, 125(18):2222–2231, DOI: 10.1161/CIRCULATIONAHA.111.067827.

Chinkes, D. L. (2005). Methods for measuring tissue protein breakdown rate in vivo. *Current Opinion in Clinical Nutrition and Metabolic Care*, 8(5):534–537, DOI: 10.1097/01.mco.0000170754.25372.37.

Chong, I.-G. and Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78(1-2):103–112, DOI: 10.1016/j.chemolab.2004.12.011.

Chu, C., Stamatelatos, D., and McNeill, K. (2017). Aquatic indirect photochemical transformations of natural peptidic thiols: impact of thiol properties, solution pH, solution salinity and metal ions. *Environmental Science: Processes & Impacts*, 19(12):1518–1527, DOI: 10.1039/C7EM00324B.

Claus, S. P. (2014). Development of personalized functional foods needs metabolic profiling. *Current Opinion in Clinical Nutrition and Metabolic Care*, 17(6):567–573, DOI: 10.1097/MCO.0000000000000107.

Clifton, P. M., Keogh, J. B., and Noakes, M. (2008). Long-term effects of a high-protein weight-loss diet. *The American Journal of Clinical Nutrition*, 87(1):23–29, http://ajcn.nutrition.org/content/87/1/23.abstract.

Clish, C. B. (2015). Metabolomics: an emerging but powerful tool for precision medicine. *Molecular Case Studies*, 1(1):a000588, DOI: 10.1101/mcs.a000588.

Cohen, B. D. (1970). Guanidinosuccinic Acid in Uremia. *Archives of Internal Medicine*, 126(5):846, DOI: 10.1001/archinte.1970.00310110116020.

Cong, W.-n., Wang, R., Cai, H., Daimon, C. M., Scheibye-Knudsen, M., Bohr, V. A., Turkin, R., Wood, W. H., Becker, K. G., Moaddel, R., Maudsley, S., and Martin, B. (2013). Long-Term Artificial Sweetener Acesulfame Potassium Treatment Alters Neurometabolic Functions in C57BL/6J Mice. *PLoS ONE*, 8(8):e70257, DOI: 10.1371/journal.pone.0070257.

Costa, C. C. G., De Almeida, I. T., Jakobs, C., Poll-The, B.-T., and Duran, M. (1999). Dynamic Changes of Plasma Acylcarnitine Levels Induced by Fasting and Sunflower Oil Challenge Test in Children. *Pediatric Research*, 46(4):440–440, DOI: 10.1203/00006450-199910000-00013.

Crawhall, J. C., Parker, R., Sneddon, W., Young, E. P., Ampola, M. G., Efron, M. L., and Bixby, E. M. (1968). Beta Mercaptolactate-Cysteine Disulfide: Analog of Cystine in the Urine of a Mentally Retarded Patient. *Science*, 160(3826):419–420, http://www.jstor.org/stable/1723531.

Čuperlović-Culf, M. (2013). *NMR Metabolomics in Cancer Research*. Woodhead Publishing, Cambridge, UK, ISBN: 978-1-907568-84-8.

da Silva, R. R., Dorrestein, P. C., and Quinn, R. A. (2015). Illuminating the dark matter in metabolomics. *Proceedings of the National Academy of Sciences*, 112(41):12549–12550, DOI: 10.1073/pnas.1516878112.

D'Agostino, L. A., Lam, K. P., Lee, R., and Britz-McKibbin, P. (2011). Comprehensive Plasma Thiol Redox Status Determination for Metabolomics. *Journal of Proteome Research*, 10(2):592–603, DOI: 10.1021/pr100771g.

Dahlem, D. P., Neiger, R., Schweighauser, A., Francey, T., Yerramilli, M., Obare, E., and Steinbach, S. M. L. (2017). Plasma Symmetric Dimethylarginine Concentration in Dogs with Acute Kidney Injury and Chronic Kidney Disease. *Journal of Veterinary Internal Medicine*, 31(3):799–804, DOI: 10.1111/jvim.14694.

Dalle-Donne, I., Rossi, R., Giustarini, D., Colombo, R., and Milzani, A. (2007). S-glutathionylation in protein redox regulation. *Free Radical Biology and Medicine*, 43(6):883–898, DOI: 10.1016/j.freeradbiomed.2007.06.014.

Dallmann, R., Viola, A. U., Tarokh, L., Cajochen, C., and Brown, S. A. (2012). The human circadian metabolome. *Proceedings of the National Academy of Sciences*, 109(7):2625–2629, DOI: 10.1073/pnas.1114410109.

Dalzill, C., Nigam, A., Juneau, M., Guilbeault, V., Latour, E., Mauriège, P., and Gayda, M. (2014). Intensive Lifestyle Intervention Improves Cardiometabolic and Exercise Parameters in Metabolically Healthy Obese and Metabolically Unhealthy Obese Individuals. *Canadian Journal of Cardiology*, 30(4):434–440, DOI: 10.1016/j.cjca.2013.11.033.

Das, M. K., Arya, R., Debnath, S., Debnath, R., Lodh, A., Bishwal, S. C., Das, A., and Nanda, R. K. (2016). Global Urine Metabolomics in Patients Treated with First-Line Tuberculosis Drugs and Identification of a Novel Metabolite of Ethambutol. *Antimicrobial Agents and Chemotherapy*, 60(4):2257–2264, DOI: 10.1128/AAC.02586-15.

Day, A. J., DuPont, M. S., Ridley, S., Rhodes, M., Rhodes, M. J. C., Morgan, M. R. A., and Williamson, G. (1998). Deglycosylation of flavonoid and isoflavonoid glycosides by human small intestine and liver β-glucosidase activity. *FEBS Letters*, 436(1):71–75, DOI: 10.1016/S0014-5793(98)01101-6.

de Oliveira, M. A. L., Porto, B. L. S., Faria, I. D. L., de Oliveira, P. L., de Castro Barra, P. M., Castro, R. d. J. C., and Sato, R. T. (2014). 20 Years of Fatty Acid Analysis by Capillary Electrophoresis. *Molecules*, 19(9):14094–14113, DOI: 10.3390/molecules190914094.

de Roos, B. and Brennan, L. (2017). Personalised Interventions—A Precision Approach for the Next Generation of Dietary Intervention Studies. *Nutrients*, 9(8):847, DOI: 10.3390/nu9080847.

de Sousa, C., English, N., Stacey, T., and Chalmers, R. (1990). Measurement of L-carnitine and acylcarnitines in body fluids and tissues in children and in adults. *Clinica Chimica Acta*, 187(3):317–328, DOI: 10.1016/0009-8981(90)90117-B.

Degenhardt, F., Seifert, S., and Szymczak, S. (2017). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*, page bbx124, DOI: 10.1093/bib/bbx124.

Deidda, M., Piras, C., Bassareo, P. P., Cadeddu Dessalvi, C., and Mercuro, G. (2015). Metabolomics, a promising approach to translational research in cardiology. *IJC Metabolic & Endocrine*, 9:31–38, DOI: 10.1016/j.ijcme.2015.10.001.

Delanghe, S., Delanghe, J. R., Speeckaert, R., Van Biesen, W., and Speeckaert, M. M. (2017). Mechanisms and consequences of carbamoylation. *Nature Reviews Nephrology*, 13(9):580–593, DOI: 10.1038/nrneph.2017.103.

Delgado-Povedano, M. M., Calderón-Santiago, M., Luque de Castro, M. D., and Priego-Capote, F. (2018). Metabolomics analysis of human sweat collected after moderate exercise. *Talanta*, 177:47–65, DOI: 10.1016/j.talanta.2017.09.028.

Demetrowitsch, T. and Schwarz, K. (2014). Metabolomics: new analytical methods for metabolome research. *Ernahrungs Umschau*, 61(7):102–111.

Deng, P., Zhan, Y., Chen, X., and Zhong, D. (2012). Derivatization methods for quantitative bioanalysis by LC–MS/MS. *Bioanalysis*, 4(1):49–69, DOI: 10.4155/bio.11.298.

Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J. M., Jones, M. R., Sommer, U., Viant, M. R., and Dunn, W. B. (2016). Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics*, 12(5):93, DOI: 10.1007/s11306-016-1030-9.

Di Simplicio, P., Frosali, S., Priora, R., Summa, D., Di Simplicio, F. C., Di Giuseppe, D., and Di Stefano, A. (2005). Biochemical and Biological Aspects of Protein Thiolation in Cells and Plasma. *Antioxidants & Redox Signaling*, 7(7-8):951–963, DOI: 10.1089/ars.2005.7.951.

Di Venere, M., Viglio, S., Cagnone, M., Bardoni, A., Salvini, R., and Iadarola, P. (2018). Advances in the analysis of "less-conventional" human body fluids: An overview of the CE- and HPLC-MS applications in the years 2015-2017. *Electrophoresis*, 39(1):160–178, DOI: 10.1002/elps.201700276.

Dias, D., Jones, O., Beale, D., Boughton, B., Benheim, D., Kouremenos, K., Wolfender, J.-L., and Wishart, D. (2016). Current and Future Perspectives on the Structural Identification of Small Molecules in Biological Systems. *Metabolites*, 6(4):46, DOI: 10.3390/metabo6040046.

DiBattista, A., McIntosh, N., Lamoureux, M., Al-Dirbashi, O. Y., Chakraborty, P., and Britz-McKibbin, P. (2017). Temporal Signal Pattern Recognition in Mass

Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Analytical Chemistry*, 89(15):8112–8121, DOI: 10.1021/acs.analchem.7b01727.

Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H. (2006). Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry*, 78(13):4281–4290, DOI: 10.1021/ac051632c.

Ding, W. and Fritz, J. S. (1998). Separation of Neutral Compounds and Basic Drugs by Capillary Electrophoresis in Acidic Solution Using Laurylpoly(oxyethylene) Sulfate as an Additive. *Analytical Chemistry*, 70(9):1859–1865, DOI: 10.1021/ac970883l.

Dittmann, M. M. and Rozing, G. P. (1998). Capillary electrochromatography (CEC): methodology and applications. *Biomedical Chromatography*, 12(3):136–137, DOI: 10.1002/(SICI)1099-0801(199805/06)12:3<136::AID-BMC781>3.0.CO;2-X.

Dixon, R. A., Gang, D. R., Charlton, A. J., Fiehn, O., Kuiper, H. A., Reynolds, T. L., Tjeerdema, R. S., Jeffery, E. H., German, J. B., Ridley, W. P., and Seiber, J. N. (2006). Applications of Metabolomics in Agriculture. *Journal of Agricultural and Food Chemistry*, 54(24):8984–8994, DOI: 10.1021/jf061218t.

Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., and Alice, M. B. (1968). Molecular Beams of Macroions. *The Journal of Chemical Physics*, 49(5):2240–2249, DOI: 10.1063/1.1670391.

Dudley, E., Yousef, M., Wang, Y., and Griffiths, W. J. (2010). Targeted metabolomics and mass spectrometry. In Donev, R., editor, *Advances in Protein Chemistry and Structural Biology*, volume 80, chapter 2, pages 45–83. Academic Press, ISBN: 1876-1623, DOI: https://doi.org/10.1016/B978-0-12-381264-3.00002-3.

Dudzik, D., Barbas-Bernardos, C., García, A., and Barbas, C. (2018). Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *Journal of Pharmaceutical and Biomedical Analysis*, 147:149–173, DOI: 10.1016/j.jpba.2017.07.044.

Due, A., Toubro, S., Skov, A. R., and Astrup, A. (2004). Effect of normal-fat diets, either medium or high in protein, on body weight in overweight subjects: a randomised 1-year trial. *International Journal of Obesity*, 28(10):1283–1290, DOI: 10.1038/sj.ijo.0802767.

Duft, R. G., Castro, A., Chacon-Mikahil, M. P. T., and Cavaglieri, C. R. (2017). Metabolomics and Exercise: possibilities and perspectives. *Motriz: Revista de Educação Física*, 23(2), DOI: 10.1590/s1980-6574201700020010.

Dunn, W. B., Broadhurst, D., Brown, M., Baker, P. N., Redman, C. W., Kenny, L. C., and Kell, D. B. (2008). Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system. *Journal of Chromatography B*, 871(2):288–298, DOI: 10.1016/j.jchromb.2008.03.021.

Dunn, W. B., Erban, A., Weber, R. J. M., Creek, D. J., Brown, M., Breitling, R., Hankemeier, T., Goodacre, R., Neumann, S., Kopka, J., and Viant, M. R. (2013). Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics*, 9(S1):44–66, DOI: 10.1007/s11306-012-0434-4.

Durbin, B. P., Hardin, J. S., Hawkins, D. M., and Rocke, D. M. (2002). A variance-stabilizing transformation for gene-expression microarray data. *Bioinformatics*, 18(Suppl 1):S105–S110, DOI: 10.1093/bioinformatics/18.suppl_1.S105.

Dziubakiewicz, E. and Buszewski, B. (2013). Principles of Electromigration Techniques. In Szumski, B., Buszewski, E., and Dziubakiewicz, M., editors, *Electromigration*

Techniques: Theory and Practice, chapter 2, pages 5–26. Springer, Heidelberg, ISBN: 978-3-642-35042-9, DOI: 10.1007/978-3-642-35043-6_2.

Eckel, R. H., York, D. A., Rössner, S., Hubbard, V., Caterson, I., St. Jeor, S. T., Hayman, L. L., Mullis, R. M., and Blair, S. N. (2004). Prevention Conference VII: Obesity, a Worldwide Epidemic Related to Heart Disease and Stroke: Executive Summary. *Circulation*, 110(18):2968–2975, DOI: 10.1161/01.CIR.0000140086.88453.9A.

Eckhardt, R. D. and Davidson, C. S. (1949). Urinary Excretion of Amino Acids by a Normal Adult Receiving Diets of Varied Protein Content. *The Journal of Biological Chemistry*, 177(2):687–695, http://www.ncbi.nlm.nih.gov/pubmed/18110445.

Eichstaedt, K. E., Kovatch, K., and Maroof, D. A. (2013). A less conservative method to adjust for familywise error rate in neuropsychological research: the Holm's sequential Bonferroni procedure. *NeuroRehabilitation*, 32(3):693–696, DOI: 10.3233/NRE-130893.

Elshorbagy, A. K., Refsum, H., Smith, A. D., and Graham, I. M. (2009). The Association of Plasma Cysteine and γ-Glutamyltransferase With BMI and Obesity. *Obesity*, 17(7):1435–1440, DOI: 10.1038/oby.2008.671.

Elshorbagy, A. K., Valdivia-Garcia, M., Refsum, H., and Butte, N. (2012). The Association of Cysteine with Obesity, Inflammatory Cytokines and Insulin Resistance in Hispanic Children and Adolescents. *PLoS ONE*, 7(9):e44166, DOI: 10.1371/journal.pone.0044166.

Emwas, A.-H. M. (2015). The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research. In Bjerrum, J. T., editor, *Metabonomics: Methods and Protocols*, pages 161–193. Springer New York, New York, NY, ISBN: 978-1-4939-2377-9, DOI: 10.1007/978-1-4939-2377-9_13.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., and Wold, S., editors (2001). *Multi- and Megavariate Data Analysis: Principles and Applications*. Umetrics Academy, Umeå, ISBN: 91-973730-1-X.

Fan, Y., Murphy, T. B., Byrne, J. C., Brennan, L., Fitzpatrick, J. M., and Watson, R. W. G. (2011). Applying Random Forests To Identify Biomarker Panels in Serum 2D-DIGE Data for the Detection and Staging of Prostate Cancer. *Journal of Proteome Research*, 10(3):1361–1373, DOI: 10.1021/pr1011069.

Farrés, M., Platikanov, S., Tsakovski, S., and Tauler, R. (2015). Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics*, 29(10):528–536, DOI: 10.1002/cem.2736.

Felig, P. (1973). The glucose-alanine cycle. *Metabolism*, 22(2):179–207, DOI: 10.1016/0026-0495(73)90269-2.

Felig, P., Pozefsk, T., Marlis, E., and Cahill, G. F. (1970). Alanine: Key Role in Gluconeogenesis. *Science*, 167(3920):1003–1004, DOI: 10.1126/science.167.3920.1003.

Fell, V., Pollitt, R. J., Sampson, G. A., and Wright, T. (1974). Ornithinemia, hyperammonemia, and homocitrullinuria: A disease associated with mental retardation and possibly caused by defective mitochondrial transport. *American Journal of Diseases of Children*, 127(5):752–756, DOI: 10.1001/archpedi.1974.02110240138023.

Fickling, S. A., Leone, A. M., Nussey, S. S., Vallance, P., and Whitley, G. S. J. (1993). Synthesis of N^G, N^G Dimethylarginine by Human Endothelial Cells. *Endothelium*, 1(3):137–140, DOI: 10.3109/10623329309102688.

Fiehn, O., Garvey, W. T., Newman, J. W., Lok, K. H., Hoppel, C. L., and Adams,

S. H. (2010). Plasma Metabolomic Profiles Reflective of Glucose Homeostasis in Non-Diabetic and Type 2 Diabetic Obese African-American Women. *PLoS ONE*, 5(12):e15234, DOI: 10.1371/journal.pone.0015234.

Flachs, P., Rossmeisl, M., Kuda, O., and Kopecky, J. (2013). Stimulation of mitochondrial oxidative capacity in white fat independent of UCP1: A key to lean phenotype. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1831(5):986–1003, DOI: 10.1016/j.bbalip.2013.02.003.

Floegel, A., Stefan, N., Yu, Z., Muhlenbruch, K., Drogan, D., Joost, H.-G., Fritsche, A., Haring, H.-U., Hrabe de Angelis, M., Peters, A., Roden, M., Prehn, C., Wang-Sattler, R., Illig, T., Schulze, M. B., Adamski, J., Boeing, H., and Pischon, T. (2013). Identification of Serum Metabolites Associated With Risk of Type 2 Diabetes Using a Targeted Metabolomic Approach. *Diabetes*, 62(2):639–648, DOI: 10.2337/db12-0495.

Food and Drug Administration (1988). Food Additives Permitted For Direct Addition To Food For Human Consumption; Acesulfame Potassium. *Federal Register*, 53(145):28379–28383, https://www.fda.gov/downloads/Food/IngredientsPackagingLabeling/FoodAdditivesIngredients/UCM404328.pdf.

Franceschelli, S., Ferrone, A., Pesce, M., Riccioni, G., and Speranza, L. (2013). Biological Functional Relevance of Asymmetric Dimethylarginine (ADMA) in Cardiovascular Disease. *International Journal of Molecular Sciences*, 14(12):24412–24421, DOI: 10.3390/ijms141224412.

Frankenfield, D., Roth-Yousey, L., and Compher, C. (2005). Comparison of Predictive Equations for Resting Metabolic Rate in Healthy Nonobese and Obese Adults: A Systematic Review. *Journal of the American Dietetic Association*, 105(5):775–789, DOI: 10.1016/j.jada.2005.02.005.

Franz, M. (2010). Diet Soft Drinks. Technical report, Diabetes Self-Management, https://www.diabetesselfmanagement.com/nutrition-exercise/meal-planning/diet-soft-drinks/.

Frazier, R. A., Ames, J. M., and Nursten, H. E., editors (2000). *Capillary Electrophoresis for Food Analysis: Method Development*. Royal Society of Chemistry, Cambridge, ISBN: 978-0-85404-492-4, DOI: 10.1039/9781847550316.

Frenich, A. G., Jouan-Rimbaud, D., Massart, D. L., Kuttatharmmakul, S., Galera, M. M., and Vidal, J. L. M. (1995). Wavelength selection method for multicomponent spectrophotometric determinations using partial least squares. *Analyst*, 120(12):2787–2792, DOI: 10.1039/an9952002787.

García, A., Godzien, J., López-Gonzálvez, Á., and Barbas, C. (2017). Capillary electrophoresis mass spectrometry as a tool for untargeted metabolomics. *Bioanalysis*, 9(1):99–130, DOI: 10.4155/bio-2016-0216.

Gardner, C. D., Offringa, L. C., Hartle, J. C., Kapphahn, K., and Cherin, R. (2016). Weight Loss on Low-Fat vs. Low-Carbohydrate Diets by Insulin Resistance Status Among Overweight Adults and Adults With Obesity: A Randomized Pilot Trial. *Obesity*, 24(1):79–86, DOI: 10.1002/oby.21331.

Gascó, L. and Barrera, R. (1972). The use of derivatives for the gas-chromatographic identification of alcohols. primary and secondary amines, and thiols in food aromas. *Analytica Chimica Acta*, 61(2):253–264, DOI: 10.1016/S0003-2670(01)95057-3.

Gaster, M., Nehlin, J. O., and Minet, A. D. (2012). Impaired TCA cycle flux in mitochondria in skeletal muscle from type 2 diabetic subjects: Marker or maker of the diabetic phenotype? *Archives of Physiology and Biochemistry*, 118(3):156–189, DOI: 10.3109/13813455.2012.656653.

Gatfield, P. D., Taller, E., Wolfe, D. M., and Haust, M. D. (1975). Hyperornithinemia, hyperammonemia, and homocitrullinuria associated with decreased carbamyl phosphate synthetase I activity. *Pediatric Research*, 9(5):488–97, http://www.ncbi.nlm.nih.gov/pubmed/166348.

Gavaghan, C. L., Wilson, I. D., and Nicholson, J. K. (2002). Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. *FEBS Letters*, 530(1-3):191–196, DOI: 10.1016/S0014-5793(02)03476-2.

Gerritsen, T., Vaughn, J. G., and Waisman, H. A. (1963). The origin of homocitrulline in the urine of infants. *Archives of Biochemistry and Biophysics*, 100(2):298–301, DOI: 10.1016/0003-9861(63)90076-6.

Gibbons, H. and Brennan, L. (2017). Metabolomics as a tool in the identification of dietary biomarkers. *Proceedings of the Nutrition Society*, 76(1):42–53, DOI: 10.1017/S002966511600032X.

Gibney, M. J., Walsh, M., Brennan, L., Roche, H. M., German, B., and van Ommen, B. (2005). Metabolomics in human nutrition: opportunities and challenges. *The American Journal of Clinical Nutrition*, 82(3):497–503, DOI: 10.1093/ajcn/82.3.497.

Giles, N. M., Giles, G. I., and Jacob, C. (2003). Multiple roles of cysteine in biocatalysis. *Biochemical and Biophysical Research Communications*, 300(1):1–4, DOI: 10.1016/S0006-291X(02)02770-5.

Gillio-Meina, C., Cepinskas, G., Cecchini, E. L., and Fraser, D. D. (2013). Translational Research in Pediatrics II: Blood Collection, Processing, Shipping, and Storage. *Pediatrics*, 131(4):754–766, DOI: 10.1542/peds.2012-1181.

Girard, A. (1989). A fast 'Monte-Carlo cross-validation' procedure for large least squares problems with noisy data. *Numerische Mathematik*, 56(1):1–23, DOI: 10.1007/BF01395775.

Giustarini, D., Lorenzini, S., Rossi, R., Chindamo, D., Di Simplicio, P., and Marcolongo, R. (2005). Altered thiol pattern in plasma of subjects affected by rheumatoid arthritis. *Clinical and Experimental Rheumatology*, 23(2):205–212, http://www.ncbi.nlm.nih.gov/pubmed/15895891.

Go, Y.-M. and Jones, D. P. (2017). Redox theory of aging: implications for health and disease. *Clinical Science*, 131(14):1669–1688, DOI: 10.1042/CS20160897.

Gomez-Casati, D. F., Zanor, M. I., and Busi, M. V. (2013). Metabolomics in Plants and Humans: Applications in the Prevention and Diagnosis of Diseases. *BioMed Research International*, 2013:Article ID 792527, 1–11, DOI: 10.1155/2013/792527.

Gori, S. S., Lorkiewicz, P., Ehringer, D. S., Belshoff, A. C., Higashi, R. M., Fan, T. W.-M., and Nantz, M. H. (2014). Profiling thiol metabolites and quantification of cellular glutathione using FT-ICR-MS spectrometry. *Analytical and Bioanalytical Chemistry*, 406(18):4371–4379, DOI: 10.1007/s00216-014-7810-z.

Gowda, G. A. N. and Djukovic, D. (2014). Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. In Raftery, D., editor, *Mass Spectrometry in Metabolomics: Methods and Protocols*, pages 3–12. Springer New York, New York, NY, ISBN: 978-1-4939-1258-2, DOI: 10.1007/978-1-4939-1258-2_1.

Gowda, G. A. N., Zhang, S., Gu, H., Asiago, V., Shanaiah, N., and Raftery, D. (2008). Metabolomics-based methods for early disease diagnostics. *Expert Review of Molecular Diagnostics*, 8(5):617–633, DOI: 10.1586/14737159.8.5.617.

Gray, J., Nicholson, J. K., Creasy, D. M., and Timbrell, J. A. (1990). Studies on the relationship between acute testicular damage and urinary and plasma creatine concentration. *Archives of Toxicology*, 64(6):443–450, DOI: 10.1007/BF01977625.

Greaves, K. A., Wilson, M. D., Rudel, L. L., Williams, J. K., and Wagner, J. D. (2000). Consumption of Soy Protein Reduces Cholesterol Absorption Compared to Casein Protein Alone or Supplemented with an Isoflavone Extract or Conjugated Equine Estrogen in Ovariectomized Cynomolgus Monkeys. *The Journal of Nutrition*, 130(4):820–826, DOI: 10.1093/jn/130.4.820.

Grigoryan, H., Edmands, W., Lu, S. S., Yano, Y., Regazzoni, L., Iavarone, A. T., Williams, E. R., and Rappaport, S. M. (2016). Adductomics Pipeline for Untargeted Analysis of Modifications to Cys34 of Human Serum Albumin. *Analytical Chemistry*, 88(21):10504–10512, DOI: 10.1021/acs.analchem.6b02553.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., and Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*, 879:10–23, DOI: 10.1016/j.aca.2015.02.012.

Guasch-Ferré, M., Bhupathiraju, S. N., and Hu, F. B. (2018). Use of Metabolomics in Improving Assessment of Dietary Intake. *Clinical Chemistry*, 64(1):82–98, DOI: 10.1373/clinchem.2017.272344.

Guertin, K. A., Moore, S. C., Sampson, J. N., Huang, W.-Y., Xiao, Q., Stolzenberg-Solomon, R. Z., Sinha, R., and Cross, A. J. (2014). Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *The American Journal of Clinical Nutrition*, 100(1):208–217, DOI: 10.3945/ajcn.113.078758.

Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koellensperger, G., Huan, T., Uritboonthai, W., Aisporna, A. E., Wolan, D. W., Spilker, M. E., Benton, H. P., and Siuzdak, G. (2018). METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical Chemistry*, 90(5):3156–3164, DOI: 10.1021/acs.analchem.7b04424.

Gundry, R. L., Fu, Q., Jelinek, C. A., Van Eyk, J. E., and Cotter, R. J. (2007). Investigation of an albumin-enriched fraction of human serum and its albuminome. *Proteomics – Clinical Applications*, 1(1):73–88, DOI: 10.1002/prca.200600276.

Guo, L., Milburn, M. V., Ryals, J. A., Lonergan, S. C., Mitchell, M. W., Wulff, J. E., Alexander, D. C., Evans, A. M., Bridgewater, B., Miller, L., Gonzalez-Garay, M. L., and Caskey, C. T. (2015). Plasma metabolomic profiles enhance precision medicine for volunteers of normal health. *Proceedings of the National Academy of Sciences*, 112(35):E4901–E4910, DOI: 10.1073/pnas.1508425112.

Guo, X.-F., Zhu, H., Wang, H., and Zhang, H.-S. (2013). Determination of thiol compounds by HPLC and fluorescence detection with 1,3,5,7-tetramethyl-8-bromomethyl-difluoroboradiaza-s-indacene. *Journal of Separation Science*, 36(4):658–664, DOI: 10.1002/jssc.201200936.

Hallows, W. C., Yu, W., Smith, B. C., Devires, M. K., Ellinger, J. J., Someya, S., Shortreed, M. R., Prolla, T., Markley, J. L., Smith, L. M., Zhao, S., Guan, K.-L., and Denu, J. M. (2011). Sirt3 Promotes the Urea Cycle and Fatty Acid Oxidation during Dietary Restriction. *Molecular Cell*, 41(2):139–149, DOI: 10.1016/j.molcel.2011.01.002.

Halton, T. L. and Hu, F. B. (2004). The Effects of High Protein Diets on Thermogenesis,

Satiety and Weight Loss: A Critical Review. *Journal of the American College of Nutrition*, 23(5):373–385, DOI: 10.1080/07315724.2004.10719381.

Han, J. C. and Han, G. Y. (1994). A Procedure for Quantitative Determination of Tris(2-Carboxyethyl)phosphine, an Odorless Reducing Agent More Stable and Effective Than Dithiothreitol. *Analytical Biochemistry*, 220(1):5–10, DOI: 10.1006/abio.1994.1290.

Hancu, G., Simon, B., Rusu, A., Mircia, E., and Gyéresi, Á. (2013). Principles of Micellar Electrokinetic Capillary Chromatography Applied in Pharmaceutical Analysis. *Advanced Pharmaceutical Bulletin*, 3(1):1–8, DOI: 10.5681/apb.2013.001.

Hanhineva, K. (2015). Application of Metabolomics to Assess Effects of Controlled Dietary Interventions. *Current Nutrition Reports*, 4(4):365–376, DOI: 10.1007/s13668-015-0148-0.

Hansen, R. E. and Winther, J. R. (2009). An introduction to methods for analyzing thiols and disulfides: Reactions, reagents, and practical considerations. *Analytical Biochemistry*, 394(2):147–158, DOI: 10.1016/j.ab.2009.07.051.

Hapfelmeier, A. and Ulm, K. (2013). A new variable selection approach using Random Forests. *Computational Statistics & Data Analysis*, 60:50–69, DOI: 10.1016/j.csda.2012.09.020.

Harada, S., Hirayama, A., Chan, Q., Kurihara, A., Fukai, K., Iida, M., Kato, S., Sugiyama, D., Kuwabara, K., Takeuchi, A., Akiyama, M., Okamura, T., Ebbels, T. M. D., Elliott, P., Tomita, M., Sato, A., Suzuki, C., Sugimoto, M., Soga, T., and Takebayashi, T. (2018). Reliability of plasma polar metabolite concentrations in a large-scale cohort study using capillary electrophoresis-mass spectrometry. *PLOS ONE*, 13(1):e0191230, DOI: 10.1371/journal.pone.0191230.

Haslam, D. W. and James, W. P. T. (2005). Obesity. *The Lancet*, 366(9492):1197–1209, DOI: 10.1016/S0140-6736(05)67483-1.

Haverberg, L. N., Omstedt, P. T., Munro, H. N., and Young, V. R. (1975). Nτ-Methylhistidine content of mixed proteins in various rat tissues. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 405(1):67–71, DOI: 10.1016/0005-2795(75)90315-3.

Heather, L. C., Wang, X., West, J. A., and Griffin, J. L. (2013). A practical guide to metabolomic profiling as a discovery tool for human heart disease. *Journal of Molecular and Cellular Cardiology*, 55:2–11, DOI: 10.1016/j.yjmcc.2012.12.001.

Hector, A. J., Marcotte, G. R., Churchward-Venne, T. A., Murphy, C. H., Breen, L., von Allmen, M., Baker, S. K., and Phillips, S. M. (2015). Whey Protein Supplementation Preserves Postprandial Myofibrillar Protein Synthesis during Short-Term Energy Restriction in Overweight and Obese Adults. *Journal of Nutrition*, 145(2):246–252, DOI: 10.3945/jn.114.200832.

Herber-Gast, G.-C. M., van Essen, H., Verschuren, W. M. M., Stehouwer, C. D. A., Gansevoort, R. T., Bakker, S. J. L., and Spijkerman, A. M. W. (2016). Coffee and tea consumption in relation to estimated glomerular filtration rate: results from the population-based longitudinal Doetinchem Cohort Study. *The American Journal of Clinical Nutrition*, 103(5):1370–1377, DOI: 10.3945/ajcn.115.112755.

Hermanson, G. T. (2013). *Bioconjugate Techniques*. Academic Press, Boston, MA, 3rd edition, ISBN: 978-0-12-382239-0.

Heuberger, A. L., Lewis, M. R., Chen, M.-H., Brick, M. A., Leach, J. E., and Ryan, E. P. (2010). Metabolomic and Functional Genomic Analyses Reveal Varietal

Differences in Bioactive Compounds of Cooked Rice. *PLoS ONE*, 5(9):e12915, DOI: 10.1371/journal.pone.0012915.

Heymsfield, S. B. and Pietrobelli, A. (2011). Individual differences in apparent energy digestibility are larger than generally recognized. *The American Journal of Clinical Nutrition*, 94(6):1650–1651, DOI: 10.3945/ajcn.111.026476.

Higashi, Y., Yamashiro, M., Yamamoto, R., and Fujii, Y. (2003). HPLC Analysis of Bucillamine by Derivatization with N -(1-Pyrenyl)Maleimide in Human Blood. *Journal of Liquid Chromatography & Related Technologies*, 26(19):3265–3275, DOI: 10.1081/JLC-120025522.

Hill, A. W. and Mortishire-Smith, R. J. (2005). Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Communications in Mass Spectrometry*, 19(21):3111–3118, DOI: 10.1002/rcm.2177.

Hirabayashi, A., Sakairi, M., and Koizumi, H. (1994). Sonic Spray Ionization Method for Atmospheric Pressure Ionization Mass Spectrometry. *Analytical Chemistry*, 66(24):4557–4559, DOI: 10.1021/ac00096a025.

Hjerpsted, J. B., Ritz, C., Schou, S. S., Tholstrup, T., and Dragsted, L. O. (2014). Effect of cheese and butter intake on metabolites in urine using an untargeted metabolomics approach. *Metabolomics*, 10(6):1176–1185, DOI: 10.1007/s11306-014-0657-7.

Hodge, A. M., English, D. R., O'Dea, K., Sinclair, A. J., Makrides, M., Gibson, R. A., and Giles, G. G. (2007). Plasma phospholipid and dietary fatty acids as predictors of type 2 diabetes: interpreting the role of linoleic acid. *The American Journal of Clinical Nutrition*, 86(1):189–197, http://www.ncbi.nlm.nih.gov/pubmed/17616780.

Hoerr, V. and Vogel, H. J. (2013). Metabolomics. In Roberts, G. C. K., editor, *Encyclopedia of Biophysics*, pages 1510–1510. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-16712-6, DOI: 10.1007/978-3-642-16712-6_338.

Hoffman, J. R. and Falvo, M. J. (2004). Protein - Which is Best? *Journal of Sports Science & Medicine*, 3(3):118–30, http://www.ncbi.nlm.nih.gov/pubmed/24482589.

Holmes, E., Wilson, I. D., and Nicholson, J. K. (2008). Metabolic Phenotyping in Health and Disease. *Cell*, 134(5):714–717, DOI: 10.1016/j.cell.2008.08.026.

Hommerson, P., Khan, A. M., de Jong, G. J., and Somsen, G. W. (2011). Ionization techniques in capillary electrophoresis-mass spectrometry: Principles, design, and application. *Mass Spectrometry Reviews*, 30(6):1096–1120, DOI: 10.1002/mas.20313.

Horgan, R. P. and Kenny, L. C. (2011). 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*, 13(3):189–195, DOI: 10.1576/toag.13.3.189.27672.

Hortin, G. L., Seam, N., and Hoehn, G. T. (2006). Bound Homocysteine, Cysteine, and Cysteinylglycine Distribution between Albumin and Globulins. *Clinical Chemistry*, 52(12):2258–2264, DOI: 10.1373/clinchem.2006.074302.

Hruby, A. and Hu, F. B. (2015). The Epidemiology of Obesity: A Big Picture. *PharmacoEconomics*, 33(7):673–689, DOI: 10.1007/s40273-014-0243-x.

Hsiao, E. Y., McBride, S. W., Hsien, S., Sharon, G., Hyde, E. R., McCue, T., Codelli, J. A., Chow, J., Reisman, S. E., Petrosino, J. F., Patterson, P. H., and Mazmanian, S. K. (2013). Microbiota Modulate Behavioral and Physiological Abnormalities Associated with Neurodevelopmental Disorders. *Cell*, 155(7):1451–1463, DOI: 10.1016/j.cell.2013.11.024.

Hu, F. B., Liu, Y., and Willett, W. C. (2011). Preventing chronic diseases by promoting healthy diet and lifestyle: public policy implications for China. *Obesity Reviews*, 12(7):552–559, DOI: 10.1111/j.1467-789X.2011.00863.x.

Huang, K.-J., Han, C.-H., Han, C.-Q., Li, J., Wu, Z.-W., and Liu, Y.-M. (2011a). Determination of thiol compounds by solid-phase extraction using multi-walled carbon nanotubes as adsorbent coupled with high-performance liquid chromatography-fluorescence detection. *Microchimica Acta*, 174(3-4):421–427, DOI: 10.1007/s00604-011-0642-z.

Huang, K.-J., Han, C.-H., Li, J., Wu, Z.-W., Liu, Y.-M., and Wu, Y.-Y. (2011b). LC Determination of Thiols Derivatized with 4-(Aminosulfonyl)-7-fluoro-2,1,3-benzoxadiazole after SPE. *Chromatographia*, 74(1-2):145–150, DOI: 10.1007/s10337-011-2025-3.

Huang, Y.-Q., Ruan, G.-D., Liu, J.-Q., Gao, Q., and Feng, Y.-Q. (2011c). Use of isotope differential derivatization for simultaneous determination of thiols and oxidized thiols by liquid chromatography tandem mass spectrometry. *Analytical Biochemistry*, 416(2):159–166, DOI: 10.1016/j.ab.2011.05.020.

Hubbard, J. and Onsager, L. (1977). Dielectric dispersion and dielectric friction in electrolyte solutions. I. *The Journal of Chemical Physics*, 67(11):4850–4857, DOI: 10.1063/1.434664.

Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18(Suppl 1):S96–S104, DOI: 10.1093/bioinformatics/18.suppl_1.S96.

Huberty, C. J. and Olejnik, S. (2006). *Applied MANOVA and Discriminant Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, second edition, ISBN: 9780471789475, DOI: 10.1002/047178947X.

Huffman, K. M., Redman, L. M., Landerman, L. R., Pieper, C. F., Stevens, R. D., Muehlbauer, M. J., Wenner, B. R., Bain, J. R., Kraus, V. B., Newgard, C. B., Ravussin, E., and Kraus, W. E. (2012). Caloric Restriction Alters the Metabolic Response to a Mixed-Meal: Results from a Randomized, Controlled Trial. *PLoS ONE*, 7(4):e28190, DOI: 10.1371/journal.pone.0028190.

Hultdin, J., Van Guelpen, B., Winkvist, A., Hallmans, G., Weinehall, L., Stegmayr, B., and Nilsson, T. K. (2011). Prospective study of first stroke in relation to plasma homocysteine and MTHFR 677C>T and 1298A>C genotypes and haplotypes – evidence for an association with hemorrhagic stroke. *Clinical Chemistry and Laboratory Medicine*, 49(9):1555–1562, DOI: 10.1515/CCLM.2011.234.

Hustad, S., Eussen, S., Midttun, Ø., Ulvik, A., van de Kant, P. M., Mørkrid, L., Gislefoss, R., and Ueland, P. M. (2012). Kinetic Modeling of Storage Effects on Biomarkers Related to B Vitamin Status and One-Carbon Metabolism. *Clinical Chemistry*, 58(2):402–410, DOI: 10.1373/clinchem.2011.174490.

Huynh, R. A. and Mohan, C. (2017). Alzheimer's Disease: Biomarkers in the Genome, Blood, and Cerebrospinal Fluid. *Frontiers in Neurology*, 8:102, DOI: 10.3389/fneur.2017.00102.

Hylemon, P. B., Zhou, H., Pandak, W. M., Ren, S., Gil, G., and Dent, P. (2009). Bile acids as regulatory molecules. *Journal of Lipid Research*, 50(8):1509–1520, DOI: 10.1194/jlr.R900007-JLR200.

Iavarone, A. T. and Williams, E. R. (2003). Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *Journal of the American Chemical Society*, 125(8):2319–2327, DOI: 10.1021/ja021202t.

Iciek, M., Chwatko, G., Lorenc-Koci, E., Bald, E., and Włodek, L. (2004). Plasma levels of total, free and protein bound thiols as well as sulfane sulfur in different age groups of rats. *Acta Biochimica Polonica*, 51(3):815–824, DOI: 045103815.

Inoue, K., Nishimura, M., Tsutsui, H., Min, J. Z., Todoroki, K., Kauffmann, J.-M., and Toyo'oka, T. (2013). Foodomics Platform for the Assay of Thiols in Wines with Fluorescence Derivatization and Ultra Performance Liquid Chromatography Mass Spectrometry Using Multivariate Statistical Analysis. *Journal of Agricultural and Food Chemistry*, 61(6):1228–1234, DOI: 10.1021/jf304822t.

Institute of Medicine (2012). *Accelerating Progress in Obesity Prevention: Solving the Weight of the Nation*. The National Academies Press, Washington, DC.

Institute of Medicine (2014). *The Current State of Obesity Solutions in the United States: Workshop Summary*. The National Academies Press, Washington, DC.

International Renal Interest Society (2016). IRIS Staging of CKD. Technical report, International Renal Interest Society, http://www.iris-kidney.com/pdf/3_staging-of-ckd.pdf.

Iribarne, J. V. and Thomson, B. A. (1976). On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64(6):2287–2294, DOI: 10.1063/1.432536.

Isokawa, M., Funatsu, T., and Tsunoda, M. (2013). Fast and simultaneous analysis of biothiols by high-performance liquid chromatography with fluorescence detection

under hydrophilic interaction chromatography conditions. *The Analyst*, 138(13):3802–3808, DOI: 10.1039/c3an00527e.

Isokawa, M., Kanamori, T., Funatsu, T., and Tsunoda, M. (2014). Analytical methods involving separation techniques for determination of low-molecular-weight biothiols in human plasma and blood. *Journal of Chromatography B*, 964:103–115, DOI: 10.1016/j.jchromb.2013.12.041.

Iwasaki, Y., Nakano, Y., Mochizuki, K., Nomoto, M., Takahashi, Y., Ito, R., Saito, K., and Nakazawa, H. (2011). A new strategy for ionization enhancement by derivatization for mass spectrometry. *Journal of Chromatography B*, 879(17-18):1159–1165, DOI: 10.1016/j.jchromb.2011.02.008.

Jahan-Mihan, A., Luhovyy, B. L., El Khoury, D., and Anderson, G. H. (2011). Dietary Proteins as Determinants of Metabolic and Physiologic Functions of the Gastrointestinal Tract. *Nutrients*, 3(12):574–603, DOI: 10.3390/nu3050574.

Jarrahpour, A., Fathi, J., Mimouni, M., Hadda, T. B., Sheikh, J., Chohan, Z., and Parvez, A. (2012). Petra, Osiris and Molinspiration (POM) together as a successful support in drug design: antibacterial activity and biopharmaceutical characterization of some azo Schiff bases. *Medicinal Chemistry Research*, 21(8):1984–1990, DOI: 10.1007/s00044-011-9723-0.

Ji, C., Sykes, L., Paul, C., Dary, O., Legetic, B., Campbell, N. R. C., and Cappuccio, F. P. (2012). Systematic review of studies comparing 24-hour and spot urine collections for estimating population salt intake. *Revista Panamericana de Salud Publica*, 32(4):307–15, http://www.ncbi.nlm.nih.gov/pubmed/23299293.

Jobard, E., Trédan, O., Postoly, D., André, F., Martin, A.-L., Elena-Herrmann, B., and Boyault, S. (2016). A Systematic Evaluation of Blood Serum and Plasma

Pre-Analytics for Metabolomics Cohort Studies. *International Journal of Molecular Sciences*, 17(12):2035, DOI: 10.3390/ijms17122035.

Johansson, I. M., Huang, E. C., Henion, J. D., and Zweigenbaum, J. (1991). Capillary electrophoresis-atmospheric pressure ionization mass spectrometry for the characterization of peptides: Instrumental considerations for mass spectrometric detection. *Journal of Chromatography A*, 554(1-2):311–327, DOI: 10.1016/S0021-9673(01)88459-4.

Johnson, C. H., Athersuch, T. J., Collman, G. W., Dhungana, S., Grant, D. F., Jones, D. P., Patel, C. J., and Vasiliou, V. (2017). Yale school of public health symposium on lifetime exposures and human health: the exposome; summary and future reflections. *Human Genomics*, 11(1):32, DOI: 10.1186/s40246-017-0128-0.

Johnson, C. H. and Gonzalez, F. J. (2012). Challenges and opportunities of metabolomics. *Journal of Cellular Physiology*, 227(8):2975–2981, DOI: 10.1002/jcp.24002.

Johnson, W. E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8(1):118–127, DOI: 10.1093/biostatistics/kxj037.

Johnston, B. C., Kanters, S., Bandayrel, K., Wu, P., Naji, F., Siemieniuk, R. A., Ball, G. D. C., Busse, J. W., Thorlund, K., Guyatt, G., Jansen, J. P., and Mills, E. J. (2014). Comparison of Weight Loss Among Named Diet Programs in Overweight and Obese Adults: A Meta-analysis. *The Journal of the American Medical Association*, 312(9):923–933, DOI: 10.1001/jama.2014.10397.

Johnstone, A. M., Horgan, G. W., Murison, S. D., Bremner, D. M., and Lobley, G. E. (2008). Effects of a high-protein ketogenic diet on hunger, appetite, and weight

loss in obese men feeding ad libitum. *The American Journal of Clinical Nutrition*, 87(1):44–55, http://ajcn.nutrition.org/content/87/1/44.abstract.

Joint WHO/FAO Expert Consultation (2003). Recommendations for preventing excess weight gain and obesity. In *WHO Technical Report Series #916: Diet, Nutrition and the Prevention of Chronic Diseases*, chapter 5, pages 61–70. Geneva, Switzerland, ISBN: 924120916X.

Jones, D. P. and Liang, Y. (2009). Measuring the poise of thiol/disulfide couples in vivo. *Free Radical Biology and Medicine*, 47(10):1329–1338, DOI: 10.1016/j.freeradbiomed.2009.08.021.

Jones, M. G., Lee, K., and Swaminathan, R. (1987). The effect of dietary protein on glomerular filtration rate in normal subjects. *Clinical Nephrology*, 27(2):71–75, http://www.ncbi.nlm.nih.gov/pubmed/3829479.

Jorgenson, J. W. and Lukacs, K. D. (1981). Zone electrophoresis in open-tubular glass capillaries. *Analytical Chemistry*, 53(8):1298–1302, DOI: 10.1021/ac00231a037.

Jovanović, V. B., Penezić-Romanjuk, A. Z., Pavićević, I. D., Aćimović, J. M., and Mandić, L. M. (2013). Improving the reliability of human serum albumin-thiol group determination. *Analytical Biochemistry*, 439(1):17–22, DOI: 10.1016/j.ab.2013.03.033.

Juraschek, S. P., Appel, L. J., Anderson, C. A. M., and Miller, E. R. (2013). Effect of a High-Protein Diet on Kidney Function in Healthy Adults: Results From the OmniHeart Trial. *American Journal of Kidney Diseases*, 61(4):547–554, DOI: 10.1053/j.ajkd.2012.10.017.

Kaddurah-Daouk, R., Kristal, B. S., and Weinshilboum, R. M. (2008). Metabolomics: A Global Biochemical Approach to Drug Response and Disease. *An-*

nual Review of Pharmacology and Toxicology, 48(1):653–683, DOI: 10.1146/annurev.pharmtox.48.113006.094715.

Kainulainen, H., Hulmi, J. J., and Kujala, U. M. (2013). Potential Role of Branched-Chain Amino Acid Catabolism in Regulating Fat Oxidation. *Exercise and Sport Sciences Reviews*, 41(4):194–200, DOI: 10.1097/JES.0b013e3182a4e6b6.

Kamlage, B., Maldonado, S. G., Bethan, B., Peter, E., Schmitz, O., Liebenberg, V., and Schatz, P. (2014). Quality Markers Addressing Preanalytical Variations of Blood and Plasma Processing Identified by Broad and Targeted Metabolite Profiling. *Clinical Chemistry*, 60(2):399–412, DOI: 10.1373/clinchem.2013.211979.

Kand'ár, R., Žáková, P., Lotková, H., Kučera, O., and Červinková, Z. (2007). Determination of reduced and oxidized glutathione in biological samples using liquid chromatography with fluorimetric detection. *Journal of Pharmaceutical and Biomedical Analysis*, 43(4):1382–1387, DOI: 10.1016/j.jpba.2006.11.028.

Kantartzis, K., Machann, J., Schick, F., Rittig, K., Machicao, F., Fritsche, A., Häring, H.-U., and Stefan, N. (2011). Effects of a lifestyle intervention in metabolically benign and malign obesity. *Diabetologia*, 54(4):864–868, DOI: 10.1007/s00125-010-2006-3.

Karelis, A. D., Brochu, M., and Rabasa-Lhoret, R. (2004a). Can we identify metabolically healthy but obese individuals (MHO)? *Diabetes & Metabolism*, 30(6):569–572, DOI: 10.1016/S1262-3636(07)70156-8.

Karelis, A. D., Messier, V., Brochu, M., and Rabasa-Lhoret, R. (2008). Metabolically healthy but obese women: effect of an energy-restricted diet. *Diabetologia*, 51(9):1752–1754, DOI: 10.1007/s00125-008-1038-4.

Karelis, A. D., St-Pierre, D. H., Conus, F., Rabasa-Lhoret, R., and Poehlman, E. T.

(2004b). Metabolic and Body Composition Factors in Subgroups of Obesity: What Do We Know? *The Journal of Clinical Endocrinology & Metabolism*, 89(6):2569–2575, DOI: 10.1210/jc.2004-0165.

Kavitha, J., Kokilambigai, K. S., Seetharaman, R., Harini, R. R., Kaaviya, A. A., Sowndaravel, P., and Lakshmi, K. S. (2018). A Review of Analytical Methods for the determination of Clopidogrel in Pharmaceuticals and Biological Matrices. *Critical Reviews in Analytical Chemistry*, 48(2):119–131, DOI: 10.1080/10408347.2018.1427548.

Kebarle, P. and Verkerk, U. H. (2009). Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews*, 28(6):898–917, DOI: 10.1002/mas.20247.

Khakimov, B., Poulsen, S. K., Savorani, F., Acar, E., Gürdeniz, G., Larsen, T. M., Astrup, A., Dragsted, L. O., and Engelsen, S. B. (2016). New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC–MS Blood Plasma Metabolomics. *Journal of Proteome Research*, 15(6):1939–1954, DOI: 10.1021/acs.jproteome.6b00109.

Khamis, M. M., Adamko, D. J., and El-Aneed, A. (2017). Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrometry Reviews*, 36(2):115–134, DOI: 10.1002/mas.21455.

Kielstein, J. T., Salpeter, S. R., Bode-Boeger, S. M., Cooke, J. P., and Fliser, D. (2006). Symmetric dimethylarginine (SDMA) as endogenous marker of renal function—a meta-analysis. *Nephrology Dialysis Transplantation*, 21(9):2446–2451, DOI: 10.1093/ndt/gfl292.

Kikuchi, K., Itoh, Y., Tateoka, R., Ezawa, A., Murakami, K., and Niwa, T. (2010). Metabolomic search for uremic toxins as indicators of the effect of an oral sor-

bent AST-120 by liquid chromatography/tandem mass spectrometry. *Journal of Chromatography B*, 878(29):2997–3002, DOI: 10.1016/j.jchromb.2010.09.006.

Kim, J. Y., Park, J. Y., Kim, O. Y., Ham, B. M., Kim, H.-J., Kwon, D. Y., Jang, Y., and Lee, J. H. (2010). Metabolic Profiling of Plasma in Overweight/Obese and Lean Men using Ultra Performance Liquid Chromatography and Q-TOF Mass Spectrometry (UPLC–Q-TOF MS). *Journal of Proteome Research*, 9(9):4368–4375, DOI: 10.1021/pr100101p.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., and Bryant, S. H. (2016). PubChem Substance and Compound databases. *Nucleic Acids Research*, 44(D1):D1202–D1213, DOI: 10.1093/nar/gkv951.

Kohl, F. J., Sánchez-Hernández, L., and Neusüß, C. (2015). Capillary electrophoresis in two-dimensional separation systems: Techniques and applications. *Electrophoresis*, 36(1):144–158, DOI: 10.1002/elps.201400368.

Kok, W. T. (2000). Capillary electrophoresis: instrumentation and operation. *Chromatographia*, 51:S1–S89.

Komatsu, M. and Sharp, G. W. (1998). Palmitate and myristate selectively mimic the effect of glucose in augmenting insulin release in the absence of extracellular Ca2+. *Diabetes*, 47(3):352–357, DOI: 10.2337/diabetes.47.3.352.

Konermann, L., Ahadi, E., Rodriguez, A. D., and Vahidi, S. (2013). Unraveling the Mechanism of Electrospray Ionization. *Analytical Chemistry*, 85(1):2–9, DOI: 10.1021/ac302789c.

Koopman, R., Saris, W. H. M., Wagenmakers, A. J. M., and van Loon, L. J. C. (2007).

Nutritional Interventions to Promote Post-Exercise Muscle Protein Synthesis. *Sports Medicine*, 37(10):895–906, DOI: 10.2165/00007256-200737100-00005.

Kosmides, A. K., Kamisoglu, K., Calvano, S. E., Corbett, S. A., and Androulakis, I. P. (2013). Metabolomic Fingerprinting: Challenges and Opportunities. *Critical Reviews in Biomedical Engineering*, 41(3):205–221, DOI: 10.1615/CritRevBiomedEng.2013007736.

Kraus, L. M. and Kraus Jr., A. P. (2001). Carbamoylation of amino acids and proteins in uremia. *Kidney International*, 59(Supplement 78):S102–S107, DOI: 10.1046/j.1523-1755.2001.59780102.x.

Krug, S., Kastenmuller, G., Stuckler, F., Rist, M. J., Skurk, T., Sailer, M., Raffler, J., Romisch-Margl, W., Adamski, J., Prehn, C., Frank, T., Engel, K.-H., Hofmann, T., Luy, B., Zimmermann, R., Moritz, F., Schmitt-Kopplin, P., Krumsiek, J., Kremer, W., Huber, F., Oeh, U., Theis, F. J., Szymczak, W., Hauner, H., Suhre, K., and Daniel, H. (2012). The dynamic range of the human metabolome revealed by challenges. *The FASEB Journal*, 26(6):2607–2619, DOI: 10.1096/fj.11-198093.

Krumsiek, J., Suhre, K., Evans, A. M., Mitchell, M. W., Mohney, R. P., Milburn, M. V., Wägele, B., Römisch-Margl, W., Illig, T., Adamski, J., Gieger, C., Theis, F. J., and Kastenmüller, G. (2012). Mining the Unknown: A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. *PLoS Genetics*, 8(10):e1003005, DOI: 10.1371/journal.pgen.1003005.

Kuehnbaum, N. L. and Britz-McKibbin, P. (2013). New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era. *Chemical Reviews*, 113(4):2437–2468, DOI: 10.1021/cr300484s.

Kuehnbaum, N. L., Gillen, J. B., Gibala, M. J., and Britz-McKibbin, P. (2015). Personalized Metabolomics for Predicting Glucose Tolerance Changes in Sedentary Women After High-Intensity Interval Training. *Scientific Reports*, 4(1):6166, DOI: 10.1038/srep06166.

Kuehnbaum, N. L., Kormendi, A., and Britz-McKibbin, P. (2013). Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Analytical Chemistry*, 85(22):10664–10669, DOI: 10.1021/ac403171u.

Kuśmierek, K., Chwatko, G., Głowacki, R., and Bald, E. (2009). Determination of endogenous thiols and thiol drugs in urine by HPLC with ultraviolet detection. *Journal of Chromatography B*, 877(28):3300–3308, DOI: 10.1016/j.jchromb.2009.03.038.

Kuśmierek, K., Chwatko, G., Głowacki, R., Kubalczyk, P., and Bald, E. (2011). Ultraviolet derivatization of low-molecular-mass thiols for high performance liquid chromatography and capillary electrophoresis analysis. *Journal of Chromatography B*, 879(17-18):1290–1307, DOI: 10.1016/j.jchromb.2010.10.035.

Kvalheim, O. M., Brakstad, F., and Liang, Y. (1994). Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. *Analytical Chemistry*, 66(1):43–51, DOI: 10.1021/ac00073a010.

Lackey, D. E., Lynch, C. J., Olson, K. C., Mostaedi, R., Ali, M., Smith, W. H., Karpe, F., Humphreys, S., Bedinger, D. H., Dunn, T. N., Thomas, A. P., Oort, P. J., Kieffer, D. A., Amin, R., Bettaieb, A., Haj, F. G., Permana, P., Anthony, T. G., and Adams, S. H. (2013). Regulation of adipose branched-chain amino acid catabolism enzyme expression and cross-adipose amino acid flux in human obesity. *American*

Journal of Physiology-Endocrinology and Metabolism, 304(11):E1175–E1187, DOI: 10.1152/ajpendo.00630.2012.

Lačná, J., Foret, F., and KubáÅĹ, P. (2017). Capillary electrophoresis in the analysis of biologically important thiols. *Electrophoresis*, 38(1):203–222, DOI: 10.1002/elps.201600354.

Landers, J. P., editor (1997). *Handbook of Capillary Electrophoresis*. CRC Press, Boca Raton, FL, 2nd edition, ISBN: 0-8493-2498-X.

Langeveld, M. and DeVries, J. H. (2015). The Long-Term Effect of Energy Restricted Diets for Treating Obesity. *Obesity*, 23(8):1529–1538, DOI: 10.1002/oby.21146.

Lankadurai, B. P., Nagato, E. G., and Simpson, M. J. (2013). Environmental metabolomics: an emerging approach to study organism responses to environmental stressors. *Environmental Reviews*, 21(3):180–205, DOI: 10.1139/er-2013-0011.

Lee, E. D., Mück, W., Henion, J. D., and Covey, T. R. (1988). On-line capillary zone electrophoresis-ion spray tandem mass spectrometry for the determination of dynorphins. *Journal of Chromatography A*, 458:313–321, DOI: 10.1016/S0021-9673(00)90575-2.

Lee, E. D., Mück, W., Henion, J. D., and Covey, T. R. (1989). Liquid junction coupling for capillary zone electrophoresis/ion spray mass spectrometry. *Biological Mass Spectrometry*, 18(9):844–850, DOI: 10.1002/bms.1200180932.

Lee, R., Ptolemy, A. S., Niewczas, L., and Britz-McKibbin, P. (2007). Integrative Metabolomics for Characterizing Unknown Low-Abundance Metabolites by Capillary Electrophoresis-Mass Spectrometry with Computer Simulations. *Analytical Chemistry*, 79(2):403–415, DOI: 10.1021/ac061780i.

Lei, Z., Huhman, D. V., and Sumner, L. W. (2011). Mass Spectrometry Strategies in Metabolomics. *Journal of Biological Chemistry*, 286(29):25435–25442, DOI: 10.1074/jbc.R111.238691.

Leidy, H. J., Carnell, N. S., Mattes, R. D., and Campbell, W. W. (2007). Higher Protein Intake Preserves Lean Mass and Satiety with Weight Loss in Pre-obese and Obese Women. *Obesity*, 15(2):421–429, DOI: 10.1038/oby.2007.531.

Lepedda, A. J., Zinellu, A., Nieddu, G., De Muro, P., Carru, C., Spirito, R., Guarino, A., Piredda, F., and Formato, M. (2014). Human Serum Albumin Cys 34 Oxidative Modifications following Infiltration in the Carotid Atherosclerotic Plaque. *Oxidative Medicine and Cellular Longevity*, 2014:1–7, DOI: 10.1155/2014/690953.

Liao, W.-T., Liu, B., Chen, J., Cui, J.-H., Gao, Y.-X., Liu, F.-Y., Xu, G., Sun, B.-D., Zhang, E.-L., Yuan, Z.-B., Zhang, G., and Gao, Y.-Q. (2016). Metabolite Modulation in Human Plasma in the Early Phase of Acclimatization to Hypobaric Hypoxia. *Scientific Reports*, 6(1):22589, DOI: 10.1038/srep22589.

Liem-Nguyen, V., Bouchet, S., and Björn, E. (2015). Determination of Sub-Nanomolar Levels of Low Molecular Mass Thiols in Natural Waters by Liquid Chromatography Tandem Mass Spectrometry after Derivatization with p -(Hydroxymercuri) Benzoate and Online Preconcentration. *Analytical Chemistry*, 87(2):1089–1096, DOI: 10.1021/ac503679y.

Lin, C. Y., Wu, H., Tjeerdema, R. S., and Viant, M. R. (2007). Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics*, 3(1):55–67, DOI: 10.1007/s11306-006-0043-1.

Lindon, J. C. and Nicholson, J. K. (2008). Spectroscopic and Statistical Techniques

for Information Recovery in Metabonomics and Metabolomics. *Annual Review of Analytical Chemistry*, 1(1):45–69, DOI: 10.1146/annurev.anchem.1.031207.113026.

Liu, A. Y., Silvestre, M. P., and Poppitt, S. D. (2015). Prevention of Type 2 Diabetes through Lifestyle Modification: Is There a Role for Higher-Protein Diets? *Advances in Nutrition: An International Review Journal*, 6(6):665–673, DOI: 10.3945/an.115.008821.

Liu, C., Lu, Y., Sun, H., Yang, J., Liu, Y., Lai, X., Gong, Y., Liu, X., Li, Y., Zhang, Y., Chen, X., and Zhong, D. (2018). Development and validation of a sensitive and rapid UHPLC–MS/MS method for the simultaneous quantification of the common active and inactive metabolites of vicagrel and clopidogrel in human plasma. *Journal of Pharmaceutical and Biomedical Analysis*, 149:394–402, DOI: 10.1016/j.jpba.2017.11.019.

Liu, C. C., Zhang, J., and Dovichi, N. J. (2005). A sheath-flow nanospray interface for capillary electrophoresis/mass spectrometry. *Rapid Communications in Mass Spectrometry*, 19(2):187–192, DOI: 10.1002/rcm.1769.

Liu, L., Aa, J., Wang, G., Yan, B., Zhang, Y., Wang, X., Zhao, C., Cao, B., Shi, J., Li, M., Zheng, T., Zheng, Y., Hao, G., Zhou, F., Sun, J., and Wu, Z. (2010). Differences in metabolite profile between blood plasma and serum. *Analytical Biochemistry*, 406(2):105–112, DOI: 10.1016/j.ab.2010.07.015.

Liu, P., Huang, Y.-Q., Cai, W.-J., Yuan, B.-F., and Feng, Y.-Q. (2014). Profiling of Thiol-Containing Compounds by Stable Isotope Labeling Double Precursor Ion Scan Mass Spectrometry. *Analytical Chemistry*, 86(19):9765–9773, DOI: 10.1021/ac5023315.

Liu, P., Qi, C.-B., Zhu, Q.-F., Yuan, B.-F., and Feng, Y.-Q. (2016). Determination of thiol metabolites in human urine by stable isotope labeling in combination with pseudo-targeted mass spectrometry analysis. *Scientific Reports*, 6(1):21433, DOI: 10.1038/srep21433.

Liu, R. H., Wharton, S., Sharma, A. M., Ardern, C. I., and Kuk, J. L. (2013). Influence of a clinical lifestyle-based weight loss program on the metabolic risk profile of metabolically normal and abnormal obese adults. *Obesity*, 21(8):1533–1539, DOI: 10.1002/oby.20219.

Lloyd, A. J., Beckmann, M., Haldar, S., Seal, C., Brandt, K., and Draper, J. (2013). Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. *The American Journal of Clinical Nutrition*, 97(2):377–389, DOI: 10.3945/ajcn.112.048033.

Lohsiriwat, S. (2013). Protein Diet and Estimated Glomerular Filtration Rate. *Open Journal of Nephrology*, 03(2):97–100, DOI: 10.4236/ojneph.2013.32016.

Loke, W. M., Jenner, A. M., Proudfoot, J. M., McKinley, A. J., Hodgson, J. M., Halliwell, B., and Croft, K. D. (2009). A Metabolite Profiling Approach to Identify Biomarkers of Flavonoid Intake in Humans. *The Journal of Nutrition*, 139(12):2309–2314, DOI: 10.3945/jn.109.113613.

Long, C. L., Haverberg, L. N., Young, V. R., Kinney, J. M., Munro, H. N., and Geiger, J. W. (1975). Metabolism of 3-methylhistidine in man. *Metabolism*, 24(8):929–935, DOI: 10.1016/0026-0495(75)90084-0.

Luo, J., Schumacher, M., Scherer, A., Sanoudou, D., Megherbi, D., Davison, T., Shi, T., Tong, W., Shi, L., Hong, H., Zhao, C., Elloumi, F., Shi, W., Thomas, R., Lin, S.,

Tillinghast, G., Liu, G., Zhou, Y., Herman, D., Li, Y., Deng, Y., Fang, H., Bushel, P., Woods, M., and Zhang, J. (2010). A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. *The Pharmacogenomics Journal*, 10(4):278–291, DOI: 10.1038/tpj.2010.57.

Maeda, Y., Ito, T., Suzuki, A., Kurono, Y., Ueta, A., Yokoi, K., Sumi, S., Togari, H., and Sugiyama, N. (2007). Simultaneous quantification of acylcarnitine isomers containing dicarboxylic acylcarnitines in human serum and urine by high-performance liquid chromatography/electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21(5):799–806, DOI: 10.1002/rcm.2905.

Magnuson, B. A., Carakostas, M. C., Moore, N. H., Poulos, S. P., and Renwick, A. G. (2016). Biological fate of low-calorie sweeteners. *Nutrition Reviews*, 74(11):670–689, DOI: 10.1093/nutrit/nuw032.

Maher, A. D., Zirah, S. F. M., Holmes, E., and Nicholson, J. K. (2007). Experimental and Analytical Variation in Human Urine in 1H NMR Spectroscopy-Based Metabolic Phenotyping Studies. *Analytical Chemistry*, 79(14):5204–5211, DOI: 10.1021/ac070212f.

Maki, T., Yamamoto, D., Nakanishi, S., Iida, K., Iguchi, G., Takahashi, Y., Kaji, H., Chihara, K., and Okimura, Y. (2012). Branched-chain amino acids reduce hindlimb suspension-induced muscle atrophy and protein levels of atrogin-1 and MuRF1 in rats. *Nutrition Research*, 32(9):676–683, DOI: 10.1016/j.nutres.2012.07.005.

Mansoor, M. A., Svardal, A. M., and Ueland, P. M. (1992). Determination of the in vivo redox status of cysteine, cysteinylglycine, homocysteine, and glutathione in human plasma. *Analytical Biochemistry*, 200(2):218–229, DOI: 10.1016/0003-2697(92)90456-H.

Marini, M. A., Frontoni, S., Succurro, E., Arturi, F., Fiorentino, T. V., Sciacqua, A., Perticone, F., and Sesti, G. (2014). Differences in insulin clearance between metabolically healthy and unhealthy obese subjects. *Acta Diabetologica*, 51(2):257–261, DOI: 10.1007/s00592-013-0511-9.

Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., and Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43:34–40, DOI: 10.1016/j.copbio.2016.08.001.

Mårtensson, J. (1982). The effects of short-term fasting on the excretion of sulfur compounds in healthy subjects. *Metabolism*, 31(5):487–492, DOI: 10.1016/0026-0495(82)90239-6.

Matarese, L. E. and Pories, W. J. (2014). Adult Weight Loss Diets: Metabolic Effects and Outcomes. *Nutrition in Clinical Practice*, 29(6):759–767, DOI: 10.1177/0884533614550251.

Matysik, S., Le Roy, C. I., Liebisch, G., and Claus, S. P. (2016). Metabolomics of fecal samples: A practical consideration. *Trends in Food Science & Technology*, 57:244–255, DOI: 10.1016/j.tifs.2016.05.011.

Maxwell, E. J. and Chen, D. D. Y. (2008). Twenty years of interface development for capillary electrophoresis–electrospray ionization–mass spectrometry. *Analytica Chimica Acta*, 627(1):25–33, DOI: 10.1016/j.aca.2008.06.034.

Maxwell, E. J., Zhong, X., Zhang, H., van Zeijl, N., and Chen, D. D. Y. (2010). Decoupling CE and ESI for a more robust interface with MS. *Electrophoresis*, 31(7):1130–1137, DOI: 10.1002/elps.200900517.

McCabe, G. P. (1975). Computations for Variable Selection in Discriminant Analysis. *Technometrics*, 17(1):103–109, DOI: 10.1080/00401706.1975.10489278.

McCormack, S. E., Shaham, O., McCarthy, M. A., Deik, A. A., Wang, T. J., Gerszten, R. E., Clish, C. B., Mootha, V. K., Grinspoon, S. K., and Fleischman, A. (2013). Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. *Pediatric Obesity*, 8(1):52–61, DOI: 10.1111/j.2047-6310.2012.00087.x.

McKeran, R. O., Halliday, D., Purkiss, P., and Royston, P. (1979). 3-Methylhistidine excretion as an index of myofibrillar protein catabolism in neuromuscular disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 42(6):536–541, DOI: 10.1136/jnnp.42.6.536.

McLafferty, F. W. and Turecek, F. (1993). *Interpretation of Mass Spectra*. University Science Books, Sausalito, CA, USA, 4th edition, ISBN: 978-0-935702-25-5.

Mehmood, T., Liland, K. H., Snipen, L., and Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118:62–69, DOI: 10.1016/j.chemolab.2012.07.010.

Mehmood, T., Martens, H., Sæbø, S., Warringer, J., and Snipen, L. (2011). A Partial Least Squares based algorithm for parsimonious variable selection. *Algorithms for Molecular Biology*, 6(1):27, DOI: 10.1186/1748-7188-6-27.

Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., and Hamprecht, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, 10(1):213, DOI: 10.1186/1471-2105-10-213.

Metwalli, A. A. M., Lammers, W. L., and Van Boekel, M. A. J. S. (1998). Formation of homocitrulline during heating of milk. *The Journal of Dairy Research*, 65(4):579–589, DOI: 10.1017/S0022029998003033.

Mihalik, S. J., Michaliszyn, S. F., de las Heras, J., Bacha, F., Lee, S., Chace, D. H., DeJesus, V. R., Vockley, J., and Arslanian, S. A. (2012). Metabolomic Profiling of Fatty Acid and Amino Acid Metabolism in Youth With Obesity and Type 2 Diabetes: Evidence for enhanced mitochondrial oxidation. *Diabetes Care*, 35(3):605–611, DOI: 10.2337/DC11-1577.

Mikkelsen, P. B., Toubro, S., and Astrup, A. (2000). Effect of fat-reduced diets on 24-h energy expenditure: comparisons between animal protein, vegetable protein, and carbohydrate. *The American Journal of Clinical Nutrition*, 72(5):1135–1141, http://ajcn.nutrition.org/content/72/5/1135.abstract.

Mikkonen, J. J. W., Singh, S. P., Herrala, M., Lappalainen, R., Myllymaa, S., and Kullaa, A. M. (2016). Salivary metabolomics in the diagnosis of oral cancer and periodontal diseases. *Journal of Periodontal Research*, 51(4):431–437, DOI: 10.1111/jre.12327.

Miller, R. C., Brindle, E., Holman, D. J., Shofer, J., Klein, N. A., Soules, M. R., and O'Connor, K. A. (2004). Comparison of Specific Gravity and Creatinine for Normalizing Urinary Reproductive Hormone Concentrations. *Clinical Chemistry*, 50(5):924–932, DOI: 10.1373/clinchem.2004.032292.

Minamida, K., Tanaka, M., Abe, A., Sone, T., Tomita, F., Hara, H., and Asano, K. (2006). Production of equol from daidzein by gram-positive rod-shaped bacterium isolated from rat intestine. *Journal of Bioscience and Bioengineering*, 102(3):247–250, DOI: 10.1263/jbb.102.247.

Moazzami, A. A., Shrestha, A., Morrison, D. A., Poutanen, K., and Mykkänen, H. (2014). Metabolomics Reveals Differences in Postprandial Responses to Breads and Fasting Metabolic Characteristics Associated with Postprandial Insulin Demand in Postmenopausal Women. *The Journal of Nutrition*, 144(6):807–814, DOI: 10.3945/jn.113.188912.

Möder, M., Kießling, A., Löster, H., and Brüggemann, L. (2003). The pattern of urinary acylcarnitines determined by electrospray mass spectrometry: a new tool in the diagnosis of diabetes mellitus. *Analytical and Bioanalytical Chemistry*, 375(2):200–210, DOI: 10.1007/s00216-002-1654-7.

Moffat, T. and Thrasher, D. (2016). School meal programs and their potential to operate as school-based obesity prevention and nutrition interventions: case studies from France and Japan. *Critical Public Health*, 26(2):133–146, DOI: 10.1080/09581596.2014.957654.

Mojtahedi, M. C., Thorpe, M. P., Karampinos, D. C., Johnson, C. L., Layman, D. K., Georgiadis, J. G., and Evans, E. M. (2011). The Effects of a Higher Protein Intake During Energy Restriction on Changes in Body Composition and Physical Function in Older Women. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 66A(11):1218–1225, DOI: 10.1093/gerona/glr120.

Monirujjaman, M. and Ferdouse, A. (2014). Metabolic and Physiological Roles of Branched-Chain Amino Acids. *Advances in Molecular Biology*, 2014:1–6, DOI: 10.1155/2014/364976.

Moore, D. R., Tang, J. E., Burd, N. A., Rerecich, T., Tarnopolsky, M. A., and Phillips, S. M. (2009). Differential stimulation of myofibrillar and sarcoplasmic

protein synthesis with protein ingestion at rest and after resistance exercise. *The Journal of Physiology*, 587(4):897–904, DOI: 10.1113/jphysiol.2008.164087.

Morifuji, M., Koga, J., Kawanaka, K., and Higuchi, M. (2009). Branched-Chain Amino Acid-Containing Dipeptides, Identified from Whey Protein Hydrolysates, Stimulate Glucose Uptake Rate in L6 Myotubes and Isolated Skeletal Muscles. *Journal of Nutritional Science and Vitaminology*, 55(1):81–86, DOI: 10.3177/jnsv.55.81.

Morris, C., O'Grada, C., Ryan, M., Roche, H. M., Gibney, M. J., Gibney, E. R., and Brennan, L. (2012). The relationship between BMI and metabolomic profiles: a focus on amino acids. *Proceedings of the Nutrition Society*, 71(04):634–638, DOI: 10.1017/S0029665112000699.

Mu, C., Yang, Y., Luo, Z., and Zhu, W. (2015). Metabolomic analysis reveals distinct profiles in the plasma and urine of rats fed a high-protein diet. *Amino Acids*, 47(6):1225–1238, DOI: 10.1007/s00726-015-1949-6.

Mück, W. M. and Henion, J. D. (1989). Determination of leucine enkephalin and methionine enkephalin in equine cerebrospinal fluid by microbore high-performance liquid chromatography and capillary zone electrophoresis coupled to tandem mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications*, 495:41–59, DOI: 10.1016/S0378-4347(00)82608-0.

Muhsen Ali, A., Burleigh, M., Daskalaki, E., Zhang, T., Easton, C., and Watson, D. G. (2016). Metabolomic Profiling of Submaximal Exercise at a Standardised Relative Intensity in Healthy Adults. *Metabolites*, 6(1):9, DOI: 10.3390/metabo6010009.

Munson, P. (2001). A 'Consistency' Test for Determining the Significance of Gene Expression Changes on Replicate Samples and Two Convenient Variance-Stabilizing

Transformations. In *GeneLogic Workshop on Low Level Analysis of Affymetrix GeneChip Data*.

Nagahara, N. and Sawada, N. (2006). The Mercaptopyruvate Pathway in Cysteine Catabolism: A Physiologic Role and Related Disease of the Multifunctional 3-Mercaptopyruvate Sulfurtransferase. *Current Medicinal Chemistry*, 13(10):1219–1230, DOI: 10.2174/092986706776360914.

Nagakubo, T., Kumano, T., Hashimoto, Y., and Kobayashi, M. (2018). Hemoglobin catalyzes CoA degradation and thiol addition to flavonoids. *Scientific Reports*, 8(1):1282, DOI: 10.1038/s41598-018-19585-7.

Nagaoka, S., Miwa, K., Eto, M., Kuzuya, Y., Hori, G., and Yamamoto, K. (1999). Soy Protein Peptic Hydrolysate with Bound Phospholipids Decreases Micellar Solubility and Cholesterol Absorption in Rats and Caco-2 Cells. *The Journal of Nutrition*, 129(9):1725–1730, DOI: 10.1093/jn/129.9.1725.

Nagaoka, S., Nakamura, A., Shibata, H., and Kanamaru, Y. (2010). Soystatin (VAWWMY), a novel bile acid-binding peptide, decreased micellar solubility and inhibited cholesterol absorption in rats. *Bioscience, Biotechnology, and Biochemistry*, 74(8):1738–1741, DOI: 10.1271/bbb.100338.

Nagumo, K., Tanaka, M., Chuang, V. T. G., Setoyama, H., Watanabe, H., Yamada, N., Kubota, K., Tanaka, M., Matsushita, K., Yoshida, A., Jinnouchi, H., Anraku, M., Kadowaki, D., Ishima, Y., Sasaki, Y., Otagiri, M., and Maruyama, T. (2014). Cys34-Cysteinylated Human Serum Albumin Is a Sensitive Plasma Marker in Oxidative Stress-Related Chronic Diseases. *PLoS ONE*, 9(1):e85216, DOI: 10.1371/journal.pone.0085216.

Nakajima, K., Hirose, K., Ebata, M., Morita, K., and Munakata, H. (2010). Association between habitual coffee consumption and normal or increased estimated glomerular filtration rate in apparently healthy adults. *British Journal of Nutrition*, 103(2):149, DOI: 10.1017/S0007114509991681.

Nakamura, H., Ito, S., Ebe, N., and Shibata, A. (1993). Renal Effects of Different Types of Protein in Healthy Volunteer Subjects and Diabetic Patients. *Diabetes Care*, 16(8):1071–1075, DOI: 10.2337/diacare.16.8.1071.

Nakamura, H., Jinzu, H., Nagao, K., Noguchi, Y., Shimba, N., Miyano, H., Watanabe, T., and Iseki, K. (2014). Plasma amino acid profiles are associated with insulin, C-peptide and adiponectin levels in type 2 diabetic patients. *Nutrition & Diabetes*, 4(9):e133–e133, DOI: 10.1038/nutd.2014.32.

Nakashima, F., Shibata, T., Kamiya, K., Yoshitake, J., Kikuchi, R., Matsushita, T., Ishii, I., Giménez-Bastida, J. A., Schneider, C., and Uchida, K. (2018). Structural and functional insights into S-thiolation of human serum albumins. *Scientific Reports*, 8(1):932, DOI: 10.1038/s41598-018-19610-9.

Nakazumi, T. and Hara, Y. (2017). Effect of sieving polymer concentration on separation of 100 bp DNA Ladder by capillary gel electrophoresis. *IOP Conference Series: Materials Science and Engineering*, 242(1):012033, DOI: 10.1088/1757-899X/242/1/012033.

Narum, S. R. (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, 7(5):783–787, DOI: 10.1007/s10592-005-9056-y.

Naukkarinen, J., Heinonen, S., Hakkarainen, A., Lundbom, J., Vuolteenaho, K., Saarinen, L., Hautaniemi, S., Rodriguez, A., Frühbeck, G., Pajunen, P., Hyötyläinen, T., Orešič, M., Moilanen, E., Suomalainen, A., Lundbom, N., Kaprio, J.,

Rissanen, A., and Pietiläinen, K. H. (2014). Characterising metabolically healthy obesity in weight-discordant monozygotic twins. *Diabetologia*, 57(1):167–176, DOI: 10.1007/s00125-013-3066-y.

Newgard, C. B. (2017). Metabolomics and Metabolic Diseases: Where Do We Stand? *Cell Metabolism*, 25(1):43–56, DOI: 10.1016/j.cmet.2016.09.018.

Newgard, C. B., An, J., Bain, J. R., Muehlbauer, M. J., Stevens, R. D., Lien, L. F., Haqq, A. M., Shah, S. H., Arlotto, M., Slentz, C. A., Rochon, J., Gallup, D., Ilkayeva, O., Wenner, B. R., Yancy, W. S., Eisenson, H., Musante, G., Surwit, R. S., Millington, D. S., Butler, M. D., and Svetkey, L. P. (2009). A Branched-Chain Amino Acid-Related Metabolic Signature that Differentiates Obese and Lean Humans and Contributes to Insulin Resistance. *Cell Metabolism*, 9(4):311–326, DOI: 10.1016/j.cmet.2009.02.002.

Newman, L., Ludford, I., Williams, C., and Herriot, M. (2016). Applying Health in All Policies to obesity in South Australia. *Health Promotion International*, 31(1):44–58, DOI: 10.1093/heapro/dau064.

Nicholson, J. K., Holmes, E., Kinross, J. M., Darzi, A. W., Takats, Z., and Lindon, J. C. (2012). Metabolic phenotyping in clinical and surgical environments. *Nature*, 491(7424):384–392, DOI: 10.1038/nature11708.

Nicholson, J. K. and Lindon, J. C. (2008). Systems biology: Metabonomics. *Nature*, 455(7216):1054–1056, DOI: 10.1038/4551054a.

Nicholson, J. K., Lindon, J. C., and Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29(11):1181–1189, DOI: 10.1080/004982599238047.

Nilsson, S. L., Andersson, C., Sjöberg, P. J. R., Bylund, D., Petersson, P., Jörntén-Karlsson, M., and Markides, K. E. (2003). Phosphate buffers in capillary electrophoresis/mass spectrometry using atmospheric pressure photoionization and electrospray ionization. *Rapid Communications in Mass Spectrometry*, 17(20):2267–2272, DOI: 10.1002/rcm.1182.

Nirungsan, K. and Thongnopnua, P. (2006). Simple and rapid high-performance liquid chromatographic method for endogenous α-tocopherol determination in human plasma. *Biomedical Chromatography*, 20(8):774–781, DOI: 10.1002/bmc.596.

Nygaard, V., Rødland, E. A., and Hovig, E. (2016). Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics*, 17(1):29–39, DOI: 10.1093/biostatistics/kxv027.

Oberbach, A., Blüher, M., Wirth, H., Till, H., Kovacs, P., Kullnick, Y., Schlichting, N., Tomm, J. M., Rolle-Kampczyk, U., Murugaiyan, J., Binder, H., Dietrich, A., and von Bergen, M. (2011). Combined Proteomic and Metabolomic Profiling of Serum Reveals Association of the Complement System with Obesity and Identifies Novel Markers of Body Fat Mass Changes. *Journal of Proteome Research*, 10(10):4769–4788, DOI: 10.1021/pr2005555.

O'Callaghan, C. (2013). Pilot Study to Assess the Use of Acesulfame K as an Exogenous Marker of Glomerular Filtration Rate (GFR). Technical report, NHS Health Research Authority, London, UK, https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/acesulfame-k-as-an-exogenous-marker-of-glomerular-filtration-rate/.

Oettl, K. and Marsche, G. (2010). Redox State of Human Serum Albumin in Terms of Cysteine-34 in Health and Disease. In Cadenas, E. and Packer, L., editors, *Thiol*

Redox Transitions in Cell Signaling, Part B: Cellular Localization and Signaling, volume 474 of *Methods in Enzymology*, chapter 11, pages 181–195. Academic Press, DOI: 10.1016/S0076-6879(10)74011-8.

Oettl, K. and Stauber, R. E. (2007). Physiological and pathological changes in the redox state of human serum albumin critically influence its binding properties. *British Journal of Pharmacology*, 151(5):580–590, DOI: 10.1038/sj.bjp.0707251.

Ogasawara, Y., Mukai, Y., Togawa, T., Suzuki, T., Tanabe, S., and Ishii, K. (2007). Determination of plasma thiol bound to albumin using affinity chromatography and high-performance liquid chromatography with fluorescence detection: Ratio of cysteinyl albumin as a possible biomarker of oxidative stress. *Journal of Chromatography B*, 845(1):157–163, DOI: 10.1016/j.jchromb.2006.08.006.

O'Gorman, A. and Brennan, L. (2015). Metabolomic applications in nutritional research: a perspective. *Journal of the Science of Food and Agriculture*, 95(13):2567–2570, DOI: 10.1002/jsfa.7070.

O'Gorman, A. and Brennan, L. (2017). The role of metabolomics in determination of new dietary biomarkers. *Proceedings of the Nutrition Society*, 76(03):295–302, DOI: 10.1017/S0029665116002974.

O'Gorman, A., Gibbons, H., and Brennan, L. (2013). Metabolomics in the Identification of Biomarkers of Dietary Intake. *Computational and Structural Biotechnology Journal*, 4(5):e201301004, DOI: 10.5936/csbj.201301004.

Olivares, J. A., Nguyen, N. T., Yonker, C. R., and Smith, R. D. (1987). On-line mass spectrometric detection for capillary zone electrophoresis. *Analytical Chemistry*, 59(8):1230–1232, DOI: 10.1021/ac00135a034.

Oliver, S. G., Winson, M. K., Kell, D. B., and Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, 16(9):373–378, DOI: 10.1016/S0167-7799(98)01214-1.

Onjiko, R. M., Portero, E. P., Moody, S. A., and Nemes, P. (2017). In Situ Microprobe Single-Cell Capillary Electrophoresis Mass Spectrometry: Metabolic Reorganization in Single Differentiating Cells in the Live Vertebrate (Xenopus laevis) Embryo. *Analytical Chemistry*, 89(13):7069–7076, DOI: 10.1021/acs.analchem.7b00880.

Ordovás Muñoz, J. M. (2013). Predictors of obesity: the "power" of the omics. *Nutrición Hospitalaria*, 28 Suppl 5:63–71, DOI: 10.3305/nh.2013.28.sup5.6919.

Oriach, C. S., Robertson, R. C., Stanton, C., Cryan, J. F., and Dinan, T. G. (2016). Food for thought: The role of nutrition in the microbiota-gut–brain axis. *Clinical Nutrition Experimental*, 6:25–38, DOI: 10.1016/j.yclnex.2016.01.003.

Oshiro, T. M., Perez, P. S., and Baranauskas, J. A. (2012). How Many Trees in a Random Forest? In Perner, P., editor, *Machine Learning and Data Mining in Pattern Recognition*, pages 154–168. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-31537-4, DOI: 10.1007/978-3-642-31537-4_13.

Oss, M., Kruve, A., Herodes, K., and Leito, I. (2010). Electrospray Ionization Efficiency Scale of Organic Compounds. *Analytical Chemistry*, 82(7):2865–2872, DOI: 10.1021/ac902856t.

O'Sullivan, A., Gibney, M. J., and Brennan, L. (2011). Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. *The American Journal of Clinical Nutrition*, 93(2):314–321, DOI: 10.3945/ajcn.110.000950.

Owen, O. E., Kalhan, S. C., and Hanson, R. W. (2002). The Key Role of Anaplerosis

and Cataplerosis for Citric Acid Cycle Function. *Journal of Biological Chemistry*, 277(34):30409–30412, DOI: 10.1074/jbc.R200006200.

Paddon-Jones, D. and Rasmussen, B. B. (2009). Dietary protein recommendations and the prevention of sarcopenia. *Current Opinion in Clinical Nutrition and Metabolic Care*, 12(1):86–90, DOI: 10.1097/MCO.0b013e32831cef8b.

Paik, W. K., Pearson, E., Nochumson, S., and Kim, S. (1977). Replacement of L-ornithine with L-lysine for urea cycle enzymes. *International Journal of Biochemistry*, 8(4):317–321, DOI: 10.1016/0020-711X(77)90139-2.

Päivä, H., Lehtimäki, T., Laakso, J., Ruokonen, I., Tervonen, R., Metso, S., Nikkilä, M., Wuolijoki, E., and Laaksonen, R. (2004). Dietary composition as a determinant of plasma asymmetric dimethylarginine in subjects with mild hypercholesterolemia. *Metabolism*, 53(8):1072–1075, DOI: 10.1016/j.metabol.2003.12.028.

Paizs, B. and Suhai, S. (2005). Fragmentation pathways of protonated peptides. *Mass Spectrometry Reviews*, 24(4):508–548, DOI: 10.1002/mas.20024.

Pallister, T., Jennings, A., Mohney, R. P., Yarand, D., Mangino, M., Cassidy, A., MacGregor, A., Spector, T. D., and Menni, C. (2016). Characterizing Blood Metabolomics Profiles Associated with Self-Reported Food Intakes in Female Twins. *PLOS ONE*, 11(6):e0158568, DOI: 10.1371/journal.pone.0158568.

Palmer, N. D., Stevens, R. D., Antinozzi, P. A., Anderson, A., Bergman, R. N., Wagenknecht, L. E., Newgard, C. B., and Bowden, D. W. (2015). Metabolomic Profile Associated With Insulin Resistance and Conversion to Diabetes in the Insulin Resistance Atherosclerosis Study. *The Journal of Clinical Endocrinology & Metabolism*, 100(3):E463–E468, DOI: 10.1210/jc.2014-2357.

Parker, B., Noakes, M., Luscombe, N., and Clifton, P. (2002). Effect of a High-Protein, High-Monounsaturated Fat Weight Loss Diet on Glycemic Control and Lipid Levels in Type 2 Diabetes. *Diabetes Care*, 25(3):425–430, DOI: 10.2337/diacare.25.3.425.

Parsons, H. M., Ludwig, C., Günther, U. L., and Viant, M. R. (2007). Improved classification accuracy in 1- and 2-dimensional NMR metabolomics data using the variance stabilising generalised logarithm transformation. *BMC Bioinformatics*, 8(1):234, DOI: 10.1186/1471-2105-8-234.

Pasiakos, S. M., Cao, J. J., Margolis, L. M., Sauter, E. R., Whigham, L. D., McClung, J. P., Rood, J. C., Carbone, J. W., Combs, G. F., and Young, A. J. (2013). Effects of high-protein diets on fat-free mass and muscle protein synthesis following weight loss: a randomized controlled trial. *The FASEB Journal*, 27(9):3837–3847, DOI: 10.1096/fj.13-230227.

Pataky, Z., Bobbioni-Harsch, E., and Golay, A. (2010). Open questions about metabolically normal obesity. *International Journal of Obesity*, 34(S2):S18–S23, DOI: 10.1038/ijo.2010.235.

Patel, P. S., Sharp, S. J., Jansen, E., Luben, R. N., Khaw, K.-T., Wareham, N. J., and Forouhi, N. G. (2010). Fatty acids measured in plasma and erythrocyte-membrane phospholipids and derived by food-frequency questionnaire and the risk of new-onset type 2 diabetes: a pilot study in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk c. *The American Journal of Clinical Nutrition*, 92(5):1214–1222, DOI: 10.3945/ajcn.2010.29182.

Patterson, A. D., Bonzo, J. A., Li, F., Krausz, K. W., Eichler, G. S., Aslam, S., Tigno, X., Weinstein, J. N., Hansen, B. C., Idle, J. R., and Gonzalez, F. J. (2011). Metabolomics Reveals Attenuation of the SLC6A20 Kidney Transporter in Nonhuman

Primate and Mouse Models of Type 2 Diabetes Mellitus. *Journal of Biological Chemistry*, 286(22):19511–19522, DOI: 10.1074/jbc.M111.221739.

Patti, G. J. (2011). Separation strategies for untargeted metabolomics. *Journal of Separation Science*, 34(24):3460–3469, DOI: 10.1002/jssc.201100532.

Patti, G. J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*, 13(4):263–269, DOI: 10.1038/nrm3314.

Paul, H. S. and Adibi, S. A. (1980). Leucine Oxidation and Protein Turnover in Clofibrate-induced Muscle Protein Degradation in Rats. *Journal of Clinical Investigation*, 65(6):1285–1293, DOI: 10.1172/JCI109791.

Pauling, L., Robinson, A. B., Teranishi, R., and Cary, P. (1971). Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proceedings of the National Academy of Sciences*, 68(10):2374–2376, DOI: 10.1073/pnas.68.10.2374.

Pedersen, S. D. (2013). Metabolic complications of obesity. *Best Practice & Research Clinical Endocrinology & Metabolism*, 27(2):179–193, DOI: 10.1016/j.beem.2013.02.004.

Pêgo-Fernandes, P. M., Bibas, B. J., and Deboni, M. (2011). Obesity: the greatest epidemic of the 21st century? *Sao Paulo Medical Journal*, 129(5):283–284, DOI: 10.1590/S1516-31802011000500001.

Pellis, L., van Erk, M. J., van Ommen, B., Bakker, G. C. M., Hendriks, H. F. J., Cnubben, N. H. P., Kleemann, R., van Someren, E. P., Bobeldijk, I., Rubingh, C. M., and Wopereis, S. (2012). Plasma metabolomics and proteomics profiling

after a postprandial challenge reveal subtle diet effects on human metabolic status. *Metabolomics*, 8(2):347–359, DOI: 10.1007/s11306-011-0320-5.

Pence, H. E. and Williams, A. (2010). ChemSpider: An Online Chemical Information Resource. *Journal of Chemical Education*, 87(11):1123–1124, DOI: 10.1021/ed100697w.

Peng, B., Li, H., and Peng, X.-X. (2015). Functional metabolomics: from biomarker discovery to metabolome reprogramming. *Protein & Cell*, 6(9):628–637, DOI: 10.1007/s13238-015-0185-x.

Perez, G., Rey, A., and Schiff, E. (1976). The Biosynthesis of Guanidinosuccinic Acid by Perfused Rat Liver. *Journal of Clinical Investigation*, 57(3):807–809, DOI: 10.1172/JCI108340.

Perez-Martinez, P., Alcala-Diaz, J. F., Delgado-Lista, J., Garcia-Rios, A., Gomez-Delgado, F., Marin-Hinojosa, C., Rodriguez-Cantalejo, F., Delgado-Casado, N., Perez-Caballero, A. I., Fuentes-Jimenez, F. J., Camargo, A., Tinahones, F. J., Ordovas, J. M., Perez-Jimenez, F., and Lopez-Miranda, J. (2014). Metabolic phenotypes of obesity influence triglyceride and inflammation homoeostasis. *European Journal of Clinical Investigation*, 44(11):1053–1064, DOI: 10.1111/eci.12339.

Perreault, M., Roke, K., Badawi, A., Nielsen, D. E., Abdelmagid, S. A., El-Sohemy, A., Ma, D. W. L., and Mutch, D. M. (2014a). Plasma Levels of 14:0, 16:0, 16:1n-7, and 20:3n-6 are Positively Associated, but 18:0 and 18:2n-6 are Inversely Associated with Markers of Inflammation in Young Healthy Adults. *Lipids*, 49(3):255–263, DOI: 10.1007/s11745-013-3874-3.

Perreault, M., Zulyniak, M. A., Badoud, F., Stephenson, S., Badawi, A., Buchholz, A., and Mutch, D. M. (2014b). A Distinct Fatty Acid Profile Underlies the Redu-

ced Inflammatory State of Metabolically Healthy Obese Individuals. *PLoS ONE*, 9(2):e88539, DOI: 10.1371/journal.pone.0088539.

Phillips, C. M., Dillon, C., Harrington, J. M., McCarthy, V. J. C., Kearney, P. M., Fitzgerald, A. P., and Perry, I. J. (2013). Defining Metabolically Healthy Obesity: Role of Dietary and Lifestyle Factors. *PLoS ONE*, 8(10):e76188, DOI: 10.1371/journal.pone.0076188.

Phillips, C. M. and Perry, I. J. (2013). Does Inflammation Determine Metabolic Health Status in Obese and Nonobese Adults? *The Journal of Clinical Endocrinology & Metabolism*, 98(10):E1610–E1619, DOI: 10.1210/jc.2013-2038.

Pietiläinen, K. H., Naukkarinen, J., Rissanen, A., Saharinen, J., Ellonen, P., Keränen, H., Suomalainen, A., Götz, A., Suortti, T., Yki-Järvinen, H., Orešič, M., Kaprio, J., and Peltonen, L. (2008). Global Transcript Profiles of Fat in Monozygotic Twins Discordant for BMI: Pathways behind Acquired Obesity. *PLoS Medicine*, 5(3):e51, DOI: 10.1371/journal.pmed.0050051.

Pinto, J., Domingues, M. R. M., Galhano, E., Pita, C., Almeida, M. d. C., Carreira, I. M., and Gil, A. M. (2014). Human plasma stability during handling and storage: impact on NMR metabolomics. *Analyst*, 139(5):1168–1177, DOI: 10.1039/C3AN02188B.

Plaizier, J. C., Walton, J. P., Martin, A., Duffield, T., Bagg, R., Dick, P., and McBride, B. W. (2000). Short Communication: Effects of Monensin on 3-Methylhistidine Excretion in Transition Dairy Cows. *Journal of Dairy Science*, 83(12):2810–2812, DOI: 10.3168/jds.S0022-0302(00)75179-4.

Pleasance, S., Thibault, P., and Kelly, J. (1992). Comparison of liquid-junction and coaxial interfaces for capillary electrophoresis-mass spectrometry with application

to compounds of concern to the aquaculture industry. *Journal of Chromatography A*, 591(1-2):325–339, DOI: 10.1016/0021-9673(92)80250-X.

Poesen, R., Mutsaers, H. A. M., Windey, K., van den Broek, P. H., Verweij, V., Augustijns, P., Kuypers, D., Jansen, J., Evenepoel, P., Verbeke, K., Meijers, B., and Masereeuw, R. (2015). The Influence of Dietary Protein Intake on Mammalian Tryptophan and Phenolic Metabolites. *PLOS ONE*, 10(10):e0140820, DOI: 10.1371/journal.pone.0140820.

Pozefsky, T., Tancredi, R. G., Moxley, R. T., Dupre, J., and Tobin, J. D. (1976). Effects of brief starvation on muscle amino acid metabolism in nonobese man. *Journal of Clinical Investigation*, 57(2):444–449, DOI: 10.1172/JCI108295.

Primeau, V., Coderre, L., Karelis, A. D., Brochu, M., Lavoie, M.-E., Messier, V., Sladek, R., and Rabasa-Lhoret, R. (2011). Characterizing the profile of obese patients who are metabolically healthy. *International Journal of Obesity*, 35(7):971–981, DOI: 10.1038/ijo.2010.216.

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T. L., Smith, S. R., Bamforth, F., Greiner, R., McManus, B., Newman, J. W., Goodfriend, T., and Wishart, D. S. (2011). The Human Serum Metabolome. *PLoS ONE*, 6(2):e16957, DOI: 10.1371/journal.pone.0016957.

Pucci, V., Mandrioli, R., and Raggi, M. A. (2003). Determination of valproic acid (2-propylpentanoic acid) in human plasma by capillary electrophoresis with indirect UV detection. *Electrophoresis*, 24(1213):2076–2083, DOI: 10.1002/elps.200305405.

Pujos-Guillot, E., Hubert, J., Martin, J.-F., Lyan, B., Quintana, M., Claude, S., Chabanas, B., Rothwell, J. A., Bennetau-Pelissero, C., Scalbert, A., Comte, B., Hercberg, S., Morand, C., Galan, P., and Manach, C. (2013). Mass Spectrometry-based Metabolomics for the Discovery of Biomarkers of Fruit and Vegetable Intake: Citrus Fruit as a Case Study. *Journal of Proteome Research*, 12(4):1645–1659, DOI: 10.1021/pr300997c.

Purohit, P. V., Rocke, D. M., Viant, M. R., and Woodruff, D. L. (2004). Discrimination Models Using Variance-Stabilizing Transformation of Metabolomic NMR Data. *OMICS: A Journal of Integrative Biology*, 8(2):118–130, DOI: 10.1089/1536231041388348.

Qatanani, M., Tan, Y., Dobrin, R., Greenawalt, D. M., Hu, G., Zhao, W., Olefsky, J. M., Sears, D. D., Kaplan, L. M., and Kemp, D. M. (2013). Inverse Regulation of Inflammation and Mitochondrial Function in Adipose Tissue Defines Extreme Insulin Sensitivity in Morbidly Obese Patients. *Diabetes*, 62(3):855–863, DOI: 10.2337/db12-0399.

Quintana-Cabrera, R., Fernandez-Fernandez, S., Bobo-Jimenez, V., Escobar, J., Sastre, J., Almeida, A., and Bolaños, J. P. (2012). γ-Glutamylcysteine detoxifies reactive oxygen species by acting as glutathione peroxidase-1 cofactor. *Nature Communications*, 3(1):718, DOI: 10.1038/ncomms1722.

Ralston, J. C., Zulyniak, M. A., Nielsen, D. E., Clarke, S., Badawi, A., El-Sohemy, A., Ma, D. W. L., and Mutch, D. M. (2013). Ethnic- and sex-specific associations between plasma fatty acids and markers of insulin resistance in healthy young adults. *Nutrition & Metabolism*, 10(1):42, DOI: 10.1186/1743-7075-10-42.

Ramautar, R. (2016). Capillary Electrophoresis–Mass Spectrometry for Clini-

cal Metabolomics. In Makowski, G. S., editor, *Advances in Clinical Chemistry: Volume 74*, chapter 1, pages 1–34. Elsevier, ISBN: 978-0-12-804689-0, DOI: 10.1016/bs.acc.2015.12.002.

Ramautar, R., Somsen, G. W., and de Jong, G. J. (2013). CE-MS for metabolomics: Developments and applications in the period 2010-2012. *Electrophoresis*, 34(1):86–98, DOI: 10.1002/elps.201200390.

Ramos-Payán, M., Ocaña-Gonzalez, J. A., Fernández-Torres, R. M., Llobera, A., and Bello-López, M. Á. (2018). Recent Trends in Capillary Electrophoresis for Complex Samples Analysis: A Review. *Electrophoresis*, 39(1):111–125, DOI: 10.1002/elps.201700269.

Ramsden, J. J. (2009). Metabolomics and Metabonomics. In *Bioinformatics: An Introduction*, volume 10 of *Computational Biology*, chapter 16, pages 239–244. Springer, London, UK, second edition, ISBN: 978-1-84800-256-2, DOI: 10.1007/978-1-84800-257-9_16.

Rappaport, S. M., Li, H., Grigoryan, H., Funk, W. E., and Williams, E. R. (2012). Adductomics: Characterizing exposures to reactive electrophiles. *Toxicology Letters*, 213(1):83–90, DOI: 10.1016/j.toxlet.2011.04.002.

Rasmussen, L. G., Winning, H., Savorani, F., Toft, H., Larsen, T. M., Dragsted, L. O., Astrup, A., and Engelsen, S. B. (2012). Assessment of the Effect of High or Low Protein Diet on the Human Urine Metabolome as Measured by NMR. *Nutrients*, 4(12):112–131, DOI: 10.3390/nu4020112.

Rebouche, C. J. (2004). Kinetics, Pharmacokinetics, and Regulation of L-Carnitine and Acetyl-L-carnitine Metabolism. *Annals of the New York Academy of Sciences*, 1033(1):30–41, DOI: 10.1196/annals.1320.003.

Rebouche, C. J., Bosch, E. P., Chenard, C. A., Schabold, K. J., and Nelson, S. E. (1989). Utilization of Dietary Precursors for Carnitine Synthesis in Human Adults. *The Journal of Nutrition*, 119(12):1907–1913, DOI: 10.1093/jn/119.12.1907.

Reimer, R. A., Maurer, A. D., Eller, L. K., Hallam, M. C., Shaykhutdinov, R., Vogel, H. J., and Weljie, A. M. (2012). Satiety Hormone and Metabolomic Response to an Intermittent High Energy Diet Differs in Rats Consuming Long-Term Diets High in Protein or Prebiotic Fiber. *Journal of Proteome Research*, 11(8):4065–4074, DOI: 10.1021/pr300487s.

Rhee, E. P., Cheng, S., Larson, M. G., Walford, G. A., Lewis, G. D., McCabe, E., Yang, E., Farrell, L., Fox, C. S., O'Donnell, C. J., Carr, S. A., Vasan, R. S., Florez, J. C., Clish, C. B., Wang, T. J., and Gerszten, R. E. (2011). Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *Journal of Clinical Investigation*, 121(4):1402–1411, DOI: 10.1172/JCI44442.

Riedl, A., Gieger, C., Hauner, H., Daniel, H., and Linseisen, J. (2017). Metabotyping and its application in targeted nutrition: an overview. *British Journal of Nutrition*, 117(12):1631–1644, DOI: 10.1017/S0007114517001611.

Ritov, V. B., Menshikova, E. V., Azuma, K., Wood, R., Toledo, F. G. S., Goodpaster, B. H., Ruderman, N. B., and Kelley, D. E. (2010). Deficiency of electron transport chain in human skeletal muscle mitochondria in type 2 diabetes mellitus and obesity. *American Journal of Physiology-Endocrinology and Metabolism*, 298(1):E49–E58, DOI: 10.1152/ajpendo.00317.2009.

Roberson, L. L., Aneni, E. C., Maziak, W., Agatston, A., Feldman, T., Rouseff, M., Tran, T., Blaha, M. J., Santos, R. D., Sposito, A., Al-Mallah, M. H., Blankstein,

R., Budoff, M. J., and Nasir, K. (2014). Beyond BMI: The "Metabolically healthy obese" phenotype & its association with clinical/subclinical cardiovascular disease and all-cause mortality – a systematic review. *BMC Public Health*, 14(1):14, DOI: 10.1186/1471-2458-14-14.

Roberts, L. D., Souza, A. L., Gerszten, R. E., and Clish, C. B. (2012). Targeted Metabolomics. In *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Hoboken, NJ, USA, DOI: 10.1002/0471142727.mb3002s98.

Rocke, D. M. and Durbin, B. (2003). Approximate variance-stabilizing transformations for gene-expression microarray data. *Bioinformatics*, 19(8):966–972, DOI: 10.1093/bioinformatics/btg107.

Rodionov, R. N., Oppici, E., Martens-Lobenhoffer, J., Jarzebska, N., Brilloff, S., Burdin, D., Demyanov, A., Kolouschek, A., Leiper, J., Maas, R., Cellini, B., Weiss, N., and Bode-Böger, S. M. (2016). A Novel Pathway for Metabolism of the Cardiovascular Risk Factor Homoarginine by alanine:glyoxylate aminotransferase 2. *Scientific Reports*, 6(1):35277, DOI: 10.1038/srep35277.

Romero, C. X., Romero, T. E., Shlay, J. C., Ogden, L. G., and Dabelea, D. (2012). Changing Trends in the Prevalence and Disparities of Obesity and Other Cardiovascular Disease Risk Factors in Three Racial/Ethnic Groups of USA Adults. *Advances in Preventive Medicine*, 2012:1–8, DOI: 10.1155/2012/172423.

Ross, A. B., Pere-Trépat, E., Montoliu, I., Martin, F.-P. J., Collino, S., Moco, S., Godin, J.-P., Cléroux, M., Guy, P. A., Breton, I., Bibiloni, R., Thorimbert, A., Tavazzi, I., Tornier, L., Bebuis, A., Bruce, S. J., Beaumont, M., Fay, L.-B., and Kochhar, S. (2013). A Whole-Grain–Rich Diet Reduces Urinary Excretion of Markers

of Protein Catabolism and Gut Microbiota Metabolism in Healthy Men after One Week. *The Journal of Nutrition*, 143(6):766–773, DOI: 10.3945/jn.112.172197.

Rössner, S. (2002). Obesity: the disease of the twenty-first century. *International Journal of Obesity*, 26(S4):S2–S4, DOI: 10.1038/sj.ijo.0802209.

Rudyk, O. and Eaton, P. (2014). Biochemical methods for monitoring protein thiol redox states in biological systems. *Redox Biology*, 2(1):803–813, DOI: 10.1016/j.redox.2014.06.005.

Rueda-Clausen, C. F., Córdoba-Porras, A., Bedoya, G., Silva, F. A., Zarruk, J. G., López-Jaramillo, P., and Villa, L. A. (2012). Increased plasma levels of total homocysteine but not asymmetric dimethylarginine in Hispanic subjects with ischemic stroke FREC-VI sub-study. *European Journal of Neurology*, 19(3):417–425, DOI: 10.1111/j.1468-1331.2011.03534.x.

Russell, S., Meadows, L. A., and Russell, R. R., editors (2008). *Microarray Technology in Practice*. Academic Press, San Diego, CA, ISBN: 978-0-12-372516-5.

Ryan, D., Robards, K., Prenzler, P. D., and Kendall, M. (2011). Recent and potential developments in the analysis of urine: A review. *Analytica Chimica Acta*, 684(1-2):17–29, DOI: 10.1016/j.aca.2010.10.035.

Ryan, E. P., Heuberger, A. L., Broeckling, C. D., Borresen, E. C., Tillotson, C., and Prenni, J. E. (2013). Advances in Nutritional Metabolomics. *Current Metabolomics*, 1(2):109–120.

Ryan, W. L. and Wells, I. C. (1964). Homocitrulline and Homoarginine Synthesis from Lysine. *Science*, 144(3622):1122–1127, DOI: 10.1126/science.144.3622.1122.

Sacks, F. M., Bray, G. A., Carey, V. J., Smith, S. R., Ryan, D. H., Anton, S. D., McManus, K., Champagne, C. M., Bishop, L. M., Laranjo, N., Leboff, M. S., Rood, J. C., de Jonge, L., Greenway, F. L., Loria, C. M., Obarzanek, E., and Williamson, D. A. (2009). Comparison of Weight-Loss Diets with Different Compositions of Fat, Protein, and Carbohydrates. *New England Journal of Medicine*, 360(9):859–873, DOI: 10.1056/NEJMoa0804748.

Samocha-Bonet, D., Dixit, V. D., Kahn, C. R., Leibel, R. L., Lin, X., Nieuwdorp, M., Pietiläinen, K. H., Rabasa-Lhoret, R., Roden, M., Scherer, P. E., Klein, S., and Ravussin, E. (2014). Metabolically healthy and unhealthy obese – the 2013 Stock Conference report. *Obesity Reviews*, 15(9):697–708, DOI: 10.1111/obr.12199.

Saude, E. J. and Sykes, B. D. (2007). Urine stability for metabolomic studies: effects of preparation and storage. *Metabolomics*, 3(1):19–27, DOI: 10.1007/s11306-006-0042-2.

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., Pujos-Guillot, E., Verheij, E., Wishart, D., and Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics*, 5(4):435–458, DOI: 10.1007/s11306-009-0168-0.

Schmidt-Sommerfeld, E., Penn, D., Kerner, J., and Bieber, L. L. (1989). Analysis of acylcarnitines in normal human urine with the radioisotopic exchange-high performance liquid chromatography (HPLC) method. *Clinica Chimica Acta*, 181(2):231–238, DOI: 10.1016/0009-8981(89)90194-0.

Schooneman, M. G., Vaz, F. M., Houten, S. M., and Soeters, M. R. (2013). Acylcarnitines. *Diabetes*, 62(1):1–8, DOI: 10.2337/db12-0466.

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., and McLean, J. A. (2016). Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *Journal of The American Society for Mass Spectrometry*, 27(12):1897–1905, DOI: 10.1007/s13361-016-1469-y.

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., and Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology*, 48(4):2097–2098, DOI: 10.1021/es5002105.

Sébédio, J.-L. and Brennan, L., editors (2014). *Metabolomics as a Tool in Nutrition Research*. Woodhead Publishing, ISBN: 978-1-78242-084-2.

Seiwert, B., Hayen, H., and Karst, U. (2008). Differential labeling of free and disulfide-bound thiol functions in proteins. *Journal of the American Society for Mass Spectrometry*, 19(1):1–7, DOI: 10.1016/j.jasms.2007.10.001.

Sengupta, A., Ghosh, S., Basant, A., Malusare, S., Johri, P., Pathak, S., Sharma, S., and Sonawat, H. M. (2011). Global host metabolic response to Plasmodium vivax infection: a 1H NMR based urinary metabonomic study. *Malaria Journal*, 10(1):384, DOI: 10.1186/1475-2875-10-384.

Servillo, L., Giovane, A., Cautela, D., Castaldo, D., and Balestrieri, M. L. (2014). Where Does N^{ϵ}-Trimethyllysine for the Carnitine Biosynthesis in Mammals Come from? *PLoS ONE*, 9(1):e84589, DOI: 10.1371/journal.pone.0084589.

Setchell, K. D. R. and Clerici, C. (2010). Equol: History, Chemistry, and Formation. *The Journal of Nutrition*, 140(7):1355S–1362S, DOI: 10.3945/jn.109.119776.

Shah, N. R. and Braverman, E. R. (2012). Measuring Adiposity in Patients: The

Utility of Body Mass Index (BMI), Percent Body Fat, and Leptin. *PLoS ONE*, 7(4):e33308, DOI: 10.1371/journal.pone.0033308.

Shai, I., Schwarzfuchs, D., Henkin, Y., Shahar, D. R., Witkow, S., Greenberg, I., Golan, R., Fraser, D., Bolotin, A., Vardi, H., Tangi-Rozental, O., Zuk-Ramot, R., Sarusi, B., Brickner, D., Schwartz, Z., Sheiner, E., Marko, R., Katorza, E., Thiery, J., Fiedler, G. M., Blüher, M., Stumvoll, M., and Stampfer, M. J. (2008). Weight Loss with a Low-Carbohydrate, Mediterranean, or Low-Fat Diet. *New England Journal of Medicine*, 359(3):229–241, DOI: 10.1056/NEJMoa0708681.

Sharma, B. K. (2002). *Instrumental Methods of Chemical Analysis*. GOEL Publishing House, Meerut, India, 22nd edition.

Shi, Z., Guan, Y., Huo, Y. R., Liu, S., Zhang, M., Lu, H., Yue, W., Wang, J., and Ji, Y. (2015). Elevated Total Homocysteine Levels in Acute Ischemic Stroke Are Associated With Long-Term Mortality. *Stroke*, 46(9):2419–2425, DOI: 10.1161/STROKEAHA.115.009136.

Shimomura, Y., Murakami, T., Nakai, N., Nagasaki, M., and Harris, R. A. (2004). Exercise promotes BCAA catabolism: effects of BCAA supplementation on skeletal muscle during exercise. *The Journal of Nutrition*, 134(6 Suppl):1583S–1587S, http://jn.nutrition.org/content/134/6/1583S.full.

Shin, M.-J., Hyun, Y. J., Kim, O. Y., Kim, J. Y., Jang, Y., and Lee, J. H. (2006). Weight loss effect on inflammation and LDL oxidation in metabolically healthy but obese (MHO) individuals: low inflammation and LDL oxidation in MHO women. *International Journal of Obesity*, 30(10):1529–1534, DOI: 10.1038/sj.ijo.0803304.

Shockcor, J. P. and Holmes, E. (2002). Metabonomic Applications in Toxicity Screening

and Disease Diagnosis. *Current Topics in Medicinal Chemistry*, 2(1):35–51, DOI: 10.2174/1568026023394498.

Sidhu, G. S. and Oakenfull, D. G. (1986). A mechanism for the hypocholesterolaemic activity of saponins. *British Journal of Nutrition*, 55(3):643, DOI: 10.1079/BJN19860070.

Silvertand, L. H. H., Toraño, J. S., van Bennekom, W. P., and de Jong, G. J. (2008). Recent developments in capillary isoelectric focusing. *Journal of Chromatography A*, 1204(2):157–170, DOI: 10.1016/j.chroma.2008.05.057.

Smilowitz, J. T., Zivkovic, A. M., Wan, Y.-J. Y., Watkins, S. M., Nording, M. L., Hammock, B. D., and German, J. B. (2013). Nutritional lipidomics: Molecular metabolism, analytics, and diagnostics. *Molecular Nutrition & Food Research*, 57(8):1319–1335, DOI: 10.1002/mnfr.201200808.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry*, 78(3):779–787, DOI: 10.1021/ac051437y.

Smith, R. D., Barinaga, C. J., and Udseth, H. R. (1988a). Improved electrospray ionization interface for capillary zone electrophoresis-mass spectrometry. *Analytical Chemistry*, 60(18):1948–1952, DOI: 10.1021/ac00169a022.

Smith, R. D., Loo, J. A., Barinaga, C. J., Edmonds, C. D., and Udseth, H. R. (1989). Capillary zone electrophoresis and isotachophoresis— mass spectrometry of polypeptides and proteins based upon an electrospray ionization interface. *Journal of Chromatography A*, 480:211–232, DOI: 10.1016/S0021-9673(01)84290-4.

Smith, R. D., Olivares, J. A., Nguyen, N. T., and Udseth, H. R. (1988b). Capillary zone electrophoresis-mass spectrometry using an electrospray ionization interface. *Analytical Chemistry*, 60(5):436–441, DOI: 10.1021/ac00156a013.

Smoluchowski, M. v. (1905). Elektrosche kataphorese. *Physikalische Zeitschrift*, 6:529.

Soenen, S., Martens, E. A. P., Hochstenbach-Waelen, A., Lemmens, S. G. T., and Westerterp-Plantenga, M. S. (2013). Normal Protein Intake Is Required for Body Weight Loss and Weight Maintenance, and Elevated Protein Intake for Additional Preservation of Resting Energy Expenditure and Fat Free Mass. *Journal of Nutrition*, 143(5):591–596, DOI: 10.3945/jn.112.167593.

Soga, T., Igarashi, K., Ito, C., Mizobuchi, K., Zimmermann, H.-P., and Tomita, M. (2009). Metabolomic Profiling of Anionic Metabolites by Capillary Electrophoresis Mass Spectrometry. *Analytical Chemistry*, 81(15):6165–6174, DOI: 10.1021/ac900675k.

Solanky, K. S., Bailey, N. J. C., Holmes, E., Lindon, J. C., Davis, A. L., Mulder, T. P. J., Van Duynhoven, J. P. M., and Nicholson, J. K. (2003). NMR-Based Metabonomic Studies on the Biochemical Effects of Epicatechin in the Rat. *Journal of Agricultural and Food Chemistry*, 51(14):4139–4145, DOI: 10.1021/jf025677f.

Solecka, B. A., Weise, C., Fuchs, B., and Kannicht, C. (2016). Free thiol groups in von Willebrand factor (VWF) are required for its full function under physiological flow conditions. *Thrombosis Research*, 137:202–210, DOI: 10.1016/j.thromres.2015.10.037.

Soronen, J., Laurila, P.-P., Naukkarinen, J., Surakka, I., Ripatti, S., Jauhiainen, M., Olkkonen, V. M., and Yki-Järvinen, H. (2012). Adipose tissue gene expression analysis reveals changes in inflammatory, mitochondrial respiratory and lipid metabolic

pathways in obese insulin-resistant subjects. *BMC Medical Genomics*, 5(1):9, DOI: 10.1186/1755-8794-5-9.

Sourabié, A. M., Spinnler, H.-E., Bonnarme, P., Saint-Eve, A., and Landaud, S. (2008). Identification of a Powerful Aroma Compound in Munster and Camembert Cheeses: Ethyl 3-Mercaptopropionate. *Journal of Agricultural and Food Chemistry*, 56(12):4674–4680, DOI: 10.1021/jf800307d.

Sourabié, A. M., Spinnler, H.-E., Saint-Eve, A., Bonnarme, P., and Landaud, S. (2011). Recent Advances in Volatile Sulfur Compounds in Cheese: Thiols and Thioesters. In Qian, M. C., Fan, X., and Mahattanatawee, K., editors, *Volatile Sulfur Compounds in Food*, chapter 6, pages 119–135. American Chemical Society, Washington, D.C., DOI: 10.1021/bk-2011-1068.ch006.

Spratlin, J. L., Serkova, N. J., and Eckhardt, S. G. (2009). Clinical Applications of Metabolomics in Oncology: A Review. *Clinical Cancer Research*, 15(2):431–440, DOI: 10.1158/1078-0432.CCR-08-1059.

Stark, A.-A., Arad, A., Siskindovich, S., Pagano, D. A., and Zeiger, E. (1989). Effect of pH on mutagenesis by thiols in Salmonella typhimurium TA102. *Mutation Research/Genetic Toxicology*, 224(1):89–94, DOI: 10.1016/0165-1218(89)90007-4.

Stefan, N., Schick, F., and Häring, H.-U. (2017). Causes, Characteristics, and Consequences of Metabolically Unhealthy Normal Weight in Humans. *Cell Metabolism*, 26(2):292–300, DOI: 10.1016/j.cmet.2017.07.008.

Stern, O. (1924). Zur Theorie der Elektrolytischen Doppelschicht. *Zeitschrift für Elektrochemie und angewandte physikalische Chemie*, 30(21-22):508–516, DOI: 10.1002/bbpc.192400182.

Stipanuk, M. H. and Ueki, I. (2011). Dealing with methionine/homocysteine sulfur: cysteine metabolism to taurine and inorganic sulfur. *Journal of Inherited Metabolic Disease*, 34(1):17–32, DOI: 10.1007/s10545-009-9006-9.

Storey, J. D. and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, 100(16):9440–9445, DOI: 10.1073/pnas.1530509100.

Succurro, E., Marini, M. A., Frontoni, S., Hribal, M. L., Andreozzi, F., Lauro, R., Perticone, F., and Sesti, G. (2008). Insulin Secretion in Metabolically Obese, but Normal Weight, and in Metabolically Healthy but Obese Individuals. *Obesity*, 16(8):1881–1886, DOI: 10.1038/oby.2008.308.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W.-M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., Reily, M. D., Thaden, J. J., and Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, 3(3):211–221, DOI: 10.1007/s11306-007-0082-2.

Sumner, L. W., Lei, Z., Nikolau, B. J., Saito, K., Roessner, U., and Trengove, R. (2014). Proposed quantitative and alphanumeric metabolite identification metrics. *Metabolomics*, 10(6):1047–1049, DOI: 10.1007/s11306-014-0739-6.

Sun, Y., Yao, T., Guo, X., Peng, Y., and Zheng, J. (2016). Simultaneous assessment of endogenous thiol compounds by LC–MS/MS. *Journal of Chromatography B*, 1029-1030:213–221, DOI: 10.1016/j.jchromb.2016.06.024.

Sussulini, A., editor (2017). *Metabolomics: From Fundamentals to Clinical Applications*,

volume 965 of *Advances in Experimental Medicine and Biology*. Springer International Publishing, Cham, ISBN: 978-3-319-47655-1, DOI: 10.1007/978-3-319-47656-8.

Sylvetsky, A. C., Bauman, V., Blau, J. E., Garraffo, H. M., Walter, P. J., and Rother, K. I. (2017). Plasma concentrations of sucralose in children and adults. *Toxicological & Environmental Chemistry*, 99(3):535–542, DOI: 10.1080/02772248.2016.1234754.

Sylvetsky, A. C., Gardner, A. L., Bauman, V., Blau, J. E., Garraffo, H. M., Walter, P. J., and Rother, K. I. (2015). Nonnutritive Sweeteners in Breast Milk. *Journal of Toxicology and Environmental Health, Part A*, 78(16):1029–1032, DOI: 10.1080/15287394.2015.1053646.

Tabachnick, B. G. and Fidell, L. S. (2007). *Using Multivariate Statistics*. Allyn & Bacon/Pearson Education, Boston, MA, 5th edition.

Tajc, S. G., Tolbert, B. S., Basavappa, R., and Miller, B. L. (2004). Direct Determination of Thiol pK_a by Isothermal Titration Microcalorimetry. *Journal of the American Chemical Society*, 126(34):10508–10509, DOI: 10.1021/ja047929u.

Takada, Y., Sakairi, M., and Koizumi, H. (1995a). Atmospheric Pressure Chemical Ionization Interface for Capillary Electrophoresis/Mass Spectrometry. *Analytical Chemistry*, 67(8):1474–1476, DOI: 10.1021/ac00104a027.

Takada, Y., Sakairi, M., and Koizumi, H. (1995b). On-line combination of micellar electrokinetic chromatography and mass spectrometry using an electrospray-chemical ionization interface. *Rapid Communications in Mass Spectrometry*, 9(6):488–490, DOI: 10.1002/rcm.1290090603.

Takoi, K., Degueil, M., Shinkaruk, S., Thibon, C., Maeda, K., Ito, K., Bennetau, B., Dubourdieu, D., and Tominaga, T. (2009). Identification and Characteristics

of New Volatile Thiols Derived from the Hop (Humulus luplus L.) Cultivar Nelson Sauvin. *Journal of Agricultural and Food Chemistry*, 57(6):2493–2502, DOI: 10.1021/jf8034622.

Tanaka, H., Sirich, T. L., and Meyer, T. W. (2015). Uremic Solutes Produced by Colon Microbes. *Blood Purification*, 40(4):306–311, DOI: 10.1159/000441578.

Tautenhahn, R., Cho, K., Uritboonthai, W., Zhu, Z., Patti, G. J., and Siuzdak, G. (2012). An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology*, 30(9):826–828, DOI: 10.1038/nbt.2348.

Tay, J., Luscombe-Marsh, N. D., Thompson, C. H., Noakes, M., Buckley, J. D., Wittert, G. A., Yancy, W. S., and Brinkworth, G. D. (2014). A Very Low-Carbohydrate, Low–Saturated Fat Diet for Type 2 Diabetes Management: A Randomized Trial. *Diabetes Care*, 37(11):2909–2918, DOI: 10.2337/dc14-0845.

Taylor, S., editor (2006). *Advances in Food and Nutrition Research, Volume 51*. Academic Press, Inc., Orlando, FL, USA, ISBN: 0120164515.

Teahan, O., Gamble, S., Holmes, E., Waxman, J., Nicholson, J. K., Bevan, C., and Keun, H. C. (2006). Impact of Analytical Bias in Metabonomic Studies of Human Blood Serum and Plasma. *Analytical Chemistry*, 78(13):4307–4318, DOI: 10.1021/ac051972y.

Tebani, A., Abily-Donval, L., Afonso, C., Marret, S., and Bekri, S. (2016). Clinical Metabolomics: The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era. *International Journal of Molecular Sciences*, 17(12):1167, DOI: 10.3390/ijms17071167.

The GBD 2015 Obesity Collaborators (2017). Health Effects of Overweight and Obesity

in 195 Countries over 25 Years. *New England Journal of Medicine*, 377(1):13–27, DOI: 10.1056/NEJMoa1614362.

Thibault, P., Pleasance, S., and Laycock, M. V. (1991). Analysis of paralytic shellfish poisons by capillary electrophoresis. *Journal of Chromatography A*, 542:483–501, DOI: 10.1016/S0021-9673(01)88783-5.

Thomson, J. S., Green, J. B., and McWilliams, T. B. (1997). Determination of Sulfides and Thiols in Petroleum Distillates Using Solid-Phase Extraction and Derivatization with Pentafluorobenzoyl Chloride. *Energy & Fuels*, 11(4):909–914, DOI: 10.1021/ef960165m.

Timms, J. F., Arslan-Low, E., Gentry-Maharaj, A., Luo, Z., T'Jampens, D., Podust, V. N., Ford, J., Fung, E. T., Gammerman, A., Jacobs, I., and Menon, U. (2007). Preanalytic Influence of Sample Handling on SELDI-TOF Serum Protein Profiles. *Clinical Chemistry*, 53(4):645–656, DOI: 10.1373/clinchem.2006.080101.

Toyo'oka, T. (2009). Recent advances in separation and detection methods for thiol compounds in biological samples. *Journal of Chromatography B*, 877(28):3318–3330, DOI: 10.1016/j.jchromb.2009.03.034.

Tran, T. N., Afanador, N. L., Buydens, L. M. C., and Blanchet, L. (2014). Interpretation of variable importance in Partial Least Squares with Significance Multivariate Correlation (sMC). *Chemometrics and Intelligent Laboratory Systems*, 138:153–160, DOI: 10.1016/j.chemolab.2014.08.005.

Trauner, M., Claudel, T., Fickert, P., Moustafa, T., and Wagner, M. (2010). Bile Acids as Regulators of Hepatic Lipid and Glucose Metabolism. *Digestive Diseases*, 28(1):220–224, DOI: 10.1159/000282091.

Trezzi, J.-P., Jäger, C., Galozzi, S., Barkovits, K., Marcus, K., Mollenhauer, B., and Hiller, K. (2017). Metabolic profiling of body fluids and multivariate data analysis. *MethodsX*, 4:95–103, DOI: 10.1016/j.mex.2017.02.004.

Tuck, M. K., Chan, D. W., Chia, D., Godwin, A. K., Grizzle, W. E., Krueger, K. E., Rom, W., Sanda, M., Sorbara, L., Stass, S., Wang, W., and Brenner, D. E. (2009). Standard Operating Procedures for Serum and Plasma Collection: Early Detection Research Network Consensus Statement Standard Operating Procedure Integration Working Group. *Journal of Proteome Research*, 8(1):113–117, DOI: 10.1021/pr800545q.

Tugizimana, F., Steenkamp, P., Piater, L., and Dubery, I. (2016). A Conversation on Data Mining Strategies in LC-MS Untargeted Metabolomics: Pre-Processing and Pre-Treatment Steps. *Metabolites*, 6(4):40, DOI: 10.3390/metabo6040040.

Turell, L., Radi, R., and Alvarez, B. (2013). The thiol pool in human plasma: The central contribution of albumin to redox processes. *Free Radical Biology and Medicine*, 65:244–253, DOI: 10.1016/j.freeradbiomed.2013.05.050.

Ubuka, T., Kobayashi, K., Yao, K., Kodama, H., Fujii, K., Hirayama, K., Kuwari, T., and Mizuhara, S. (1968). S-(2-hydroxy-2-carboxyethylthio)cysteine and S-(carboxymethyltio)cysteine in human urine. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 158(3):493–495, DOI: 10.1016/0304-4165(68)90313-9.

Ubuka, T., Ohta, J., Akagi, R., Hosaki, Y., Ishimoto, Y., Kiguchi, S., Ikeda, T., and Ishino, K. (1992). Metabolism ofl-cysteine via transamination pathway (3-mercaptopyruvate pathway). *Amino Acids*, 3(3):243–252, DOI: 10.1007/BF00805999.

Udseth, H. R., Loo, J. A., and Smith, R. D. (1989). Capillary isotachophoresis/mass spectrometry. *Analytical Chemistry*, 61(3):228–232, DOI: 10.1021/ac00178a008.

U.S. Dairy Export Council (2008). Reference Manual for U.S. Whey and Lactose Products. Technical report, http://usdec.files.cms-plus.com/PDFs/2008ReferenceManuals/Whey_Lactose_Reference_Manual_Complete2_Optimized.pdf.

U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory (2016). USDA National Nutrient Database for Standard Reference, Release 28, Item 16122. Technical report, https://ndb.nal.usda.gov/ndb/foods/show/16122.

Vaher, M., Koel, M., and Kaljurand, M. (2001). Non-aqueous capillary electrophoresis in acetonitrile using lonic-liquid buffer electrolytes. *Chromatographia*, 53(S1):S302–S306, DOI: 10.1007/BF02490346.

Vaidyanathan, S., Harrigan, G. G., and Goodacre, R., editors (2005). *Metabolome Analyses: Strategies for Systems Biology*. Springer US, Boston, MA, ISBN: 978-0-387-25239-1, DOI: 10.1007/b106967.

Van Bemmel, A. L., Smeets, E. H. J., and Van Diest, R. (1988). The 24-hour pattern of urinary MHPG excretion in depressives and normals. *Acta Psychiatrica Scandinavica*, 78(3):298–303, DOI: 10.1111/j.1600-0447.1988.tb06340.x.

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(1):142, DOI: 10.1186/1471-2164-7-142.

van der Greef, J., Martin, S., Juhasz, P., Adourian, A., Plasterer, T., Verheij, E. R., and McBurney, R. N. (2007). The Art and Practice of Systems Biology in Medicine: Mapping Patterns of Relationships. *Journal of Proteome Research*, 6(4):1540–1559, DOI: 10.1021/pr0606530.

van Dijk, S. J., Feskens, E. J., Bos, M. B., Hoelen, D. W., Heijligenberg, R., Bromhaar, M. G., de Groot, L. C., de Vries, J. H., Muller, M., and Afman, L. A. (2009). A saturated fatty acid-rich diet induces an obesity-linked proinflammatory gene expression profile in adipose tissue of subjects at risk of metabolic syndrome. *The American Journal of Clinical Nutrition*, 90(6):1656–1664, DOI: 10.3945/ajcn.2009.27792.

van Vliet-Ostaptchouk, J. V., Nuotio, M.-L., Slagter, S. N., Doiron, D., Fischer, K., Foco, L., Gaye, A., Gögele, M., Heier, M., Hiekkalinna, T., Joensuu, A., Newby, C., Pang, C., Partinen, E., Reischl, E., Schwienbacher, C., Tammesoo, M.-L., Swertz, M. A., Burton, P., Ferretti, V., Fortier, I., Giepmans, L., Harris, J. R., Hillege, H. L., Holmen, J., Jula, A., Kootstra-Ros, J. E., Kvaløy, K., Holmen, T. L., Männistö, S., Metspalu, A., Midthjell, K., Murtagh, M. J., Peters, A., Pramstaller, P. P., Saaristo, T., Salomaa, V., Stolk, R. P., Uusitupa, M., van der Harst, P., van der Klauw, M. M., Waldenberger, M., Perola, M., and Wolffenbuttel, B. H. R. (2014). The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. *BMC Endocrine Disorders*, 14(1):9, DOI: 10.1186/1472-6823-14-9.

Vanholder, R. and De Smet, R. (1999). Pathophysiologic Effects of Uremic Retention Solutes. *Journal of the American Society of Nephrology*, 10(8):1815–1823, http://jasn.asnjournals.org/content/10/8/1815.short.

Vaught, J. B. (2006). Blood Collection, Shipment, Processing, and Storage. *Cancer Epidemiology Biomarkers & Prevention*, 15(9):1582–1584, DOI: 10.1158/1055-9965.EPI-06-0630.

Vazquez, J. A., Morse, E. L., and Adibi, S. A. (1985). Effect of Dietary Fat, Carbohy-

drate, and Protein on Branched-Chain Amino Acid Catabolism during Caloric Restriction. *Journal of Clinical Investigation*, 76(2):737–743, DOI: 10.1172/JCI112029.

Vázquez-Fresno, R., Llorach, R., Urpi-Sarda, M., Lupianez-Barbero, A., Estruch, R., Corella, D., Fitó, M., Arós, F., Ruiz-Canela, M., Salas-Salvadó, J., and Andres-Lacueva, C. (2015). Metabolomic Pattern Analysis after Mediterranean Diet Intervention in a Nondiabetic Population: A 1- and 3-Year Follow-up in the PREDIMED Study. *Journal of Proteome Research*, 14(1):531–540, DOI: 10.1021/pr5007894.

Velho, S., Paccaud, F., Waeber, G., Vollenweider, P., and Marques-Vidal, P. (2010). Metabolically healthy obesity: different prevalences using different criteria. *European Journal of Clinical Nutrition*, 64(10):1043–1051, DOI: 10.1038/ejcn.2010.114.

Vermeulen, C., Gijs, L., and Collin, S. (2005). Sensorial Contribution and Formation Pathways of Thiols in Foods: A Review. *Food Reviews International*, 21(1):69–137, DOI: 10.1081/FRI-200040601.

Vinaixa, M., Samino, S., Saez, I., Duran, J., Guinovart, J. J., and Yanes, O. (2012). A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites*, 2(4):775–795, DOI: 10.3390/metabo2040775.

Vinayavekhin, N. and Saghatelian, A. (2010). Untargeted Metabolomics. In *Current Protocols in Molecular Biology*, volume Chapter 30, pages Unit 30.1.1–24. John Wiley & Sons, Inc., Hoboken, NJ, USA, ISBN: 0471142727, DOI: 10.1002/0471142727.mb3001s90.

Wagner, M., Ohlund, L. B., Shiao, T. C., Vézina, A., Annabi, B., Roy, R., and Sleno, L. (2015). Isotope-labeled differential profiling of metabolites using

N -benzoyloxysuccinimide derivatization coupled to liquid chromatography/high-resolution tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 29(18):1632–1640, DOI: 10.1002/rcm.7264.

Wagner, N. D., Simpson, A. J., and Simpson, M. J. (2017). Metabolomic responses to sublethal contaminant exposure in neonate and adult Daphnia magna. *Environmental Toxicology and Chemistry*, 36(4):938–946, DOI: 10.1002/etc.3604.

Walsh, M. C., Brennan, L., Malthouse, J. P. G., Roche, H. M., and Gibney, M. J. (2006). Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition*, 84(3):531–539, DOI: 10.1093/ajcn/84.3.531.

Wanders, R. J. A., Duran, M., and Loupatty, F. J. (2012). Enzymology of the branched-chain amino acid oxidation disorders: the valine pathway. *Journal of Inherited Metabolic Disease*, 35(1):5–12, DOI: 10.1007/s10545-010-9236-x.

Wang, J. H., Byun, J., and Pennathur, S. (2010). Analytical Approaches to Metabolomics and Applications to Systems Biology. *Seminars in Nephrology*, 30(5):500–511, DOI: 10.1016/j.semnephrol.2010.07.007.

Wang, Q., Ge, X., Tian, X., Zhang, Y., Zhang, J., and Zhang, P. (2013). Soy isoflavone: The multipurpose phytochemical (Review). *Biomedical Reports*, 1(5):697–701, DOI: 10.3892/br.2013.129.

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., Lewis, G. D., Fox, C. S., Jacques, P. F., Fernandez, C., O'Donnell, C. J., Carr, S. A., Mootha, V. K., Florez, J. C., Souza, A., Melander, O., Clish, C. B., and Gerszten, R. E. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine*, 17(4):448–453, DOI: 10.1038/nm.2307.

Wang-Sattler, R., Yu, Z., Herder, C., Messias, A. C., Floegel, A., He, Y., Heim, K., Campillos, M., Holzapfel, C., Thorand, B., Grallert, H., Xu, T., Bader, E., Huth, C., Mittelstrass, K., Döring, A., Meisinger, C., Gieger, C., Prehn, C., Roemisch-Margl, W., Carstensen, M., Xie, L., Yamanaka-Okumura, H., Xing, G., Ceglarek, U., Thiery, J., Giani, G., Lickert, H., Lin, X., Li, Y., Boeing, H., Joost, H.-G., de Angelis, M. H., Rathmann, W., Suhre, K., Prokisch, H., Peters, A., Meitinger, T., Roden, M., Wichmann, H.-E., Pischon, T., Adamski, J., and Illig, T. (2012). Novel biomarkers for pre-diabetes identified by metabolomics. *Molecular Systems Biology*, 8(1):615, DOI: 10.1038/msb.2012.43.

Want, E. J., O'Maille, G., Smith, C. A., Brandon, T. R., Uritboonthai, W., Qin, C., Trauger, S. A., and Siuzdak, G. (2006). Solvent-Dependent Metabolite Distribution, Clustering, and Protein Extraction for Serum Profiling with Mass Spectrometry. *Analytical Chemistry*, 78(3):743–752, DOI: 10.1021/ac051312t.

Warrack, B. M., Hnatyshyn, S., Ott, K.-H., Reily, M. D., Sanders, M., Zhang, H., and Drexler, D. M. (2009). Normalization strategies for metabonomic analysis of urine samples. *Journal of Chromatography B*, 877(5-6):547–552, DOI: 10.1016/j.jchromb.2009.01.007.

Watson, D. G. (2013). A Rough Guide to Metabolite Identification using High Resolution Liquid Chromatography Mass Spectrometry in Metabolomic Profiling in Metazoans. *Computational and Structural Biotechnology Journal*, 4(5):e201301005, DOI: 10.5936/csbj.201301005.

Weidt, S., Haggarty, J., Kean, R., Cojocariu, C. I., Silcock, P. J., Rajendran, R., Ramage, G., and Burgess, K. E. V. (2016). A novel targeted/untargeted GC-Orbitrap

metabolomics methodology applied to Candida albicans and Staphylococcus aureus biofilms. *Metabolomics*, 12(12):189, DOI: 10.1007/s11306-016-1134-2.

Weigle, D. S., Breen, P. A., Matthys, C. C., Callahan, H. S., Meeuws, K. E., Burden, V. R., and Purnell, J. Q. (2005). A high-protein diet induces sustained reductions in appetite, ad libitum caloric intake, and body weight despite compensatory changes in diurnal plasma leptin and ghrelin concentrations. *The American Journal of Clinical Nutrition*, 82(1):41–48, http://ajcn.nutrition.org/content/82/1/41.abstract.

Westerterp-Plantenga, M. S., Nieuwenhuizen, A., Tomé, D., Soenen, S., and Westerterp, K. R. (2009). Dietary Protein, Weight Loss, and Weight Maintenance. *Annual Review of Nutrition*, 29(1):21–41, DOI: 10.1146/annurev-nutr-080508-141056.

Weyer, C., Bogardus, C., Mott, D. M., and Pratley, R. E. (1999). The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. *Journal of Clinical Investigation*, 104(6):787–794, DOI: 10.1172/JCI7231.

Whatley, H. (2001). Basic Principles and Modes of Capillary Electrophoresis. In Petersen, J. R. and Mohammad, A. A., editors, *Clinical and Forensic Applications of Capillary Electrophoresis*, chapter 2, pages 21–58. Humana Press, Totowa, NJ, ISBN: 978-0-89603-645-1, DOI: 10.1007/978-1-59259-120-6.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfries, J., Moritz, T., and Trygg, J. (2008). Visualization of GC/TOF-MS-Based Metabolomics Data for Identification of Biochemically Interesting Compounds Using OPLS Class Models. *Analytical Chemistry*, 80(1):115–122, DOI: 10.1021/ac0713510.

Wiley, V. C., Dudman, N. P. B., and Wilcken, D. E. L. (1988). Interrelations between plasma free and protein-bound homocysteine and cysteine in homocystinuria. *Metabolism*, 37(2):191–195, DOI: 10.1016/S0026-0495(98)90017-8.

Wilson, L. A., Wilkinson, K., Crews, H. M., Davies, A. M., Dick, C. S., and Dumsday, V. L. (1999). Urinary monitoring of saccharin and acesulfame-K as biomarkers of exposure to these additives. *Food Additives and Contaminants*, 16(6):227–238, DOI: 10.1080/026520399283993.

Winnike, J. H., Busby, M. G., Watkins, P. B., and O'Connell, T. M. (2009). Effects of a prolonged standardized diet on normalizing the human metabolome. *The American Journal of Clinical Nutrition*, 90(6):1496–1501, DOI: 10.3945/ajcn.2009.28234.

Winther, J. R. and Thorpe, C. (2014). Quantification of thiols and disulfides. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(2):838–846, DOI: 10.1016/j.bbagen.2013.03.031.

Wishart, D. S. (2005). Metabolomics: The Principles and Potential Applications to Transplantation. *American Journal of Transplantation*, 5(12):2814–2820, DOI: 10.1111/j.1600-6143.2005.01119.x.

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*, 15(7):473–484, DOI: 10.1038/nrd.2016.32.

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., and Scalbert, A. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research*, 46(D1):D608–D617, DOI: 10.1093/nar/gkx1089.

Wold, S., Johansson, E., and Cocchi, M. (1993). PLS—Partial Least Squares Projections to Latent Structures. In Kubinyi, H., editor, *3D QSAR in Drug Design: Theory, Methods and Applications*, pages 523–550. ESCOM Science Publishers, Leiden, ISBN: 9789072199140.

Wung, W. E. and Howell, S. B. (1980). Simultaneous liquid chromatography of 5-fluorouracil, uridine, hypoxanthine, xanthine, uric acid, allopurinol, and oxipurinol in plasma. *Clinical Chemistry*, 26(12):1704—-1708, http://clinchem.aaccjnls.org/content/26/12/1704.

Würtz, P., Soininen, P., Kangas, A. J., Rönnemaa, T., Lehtimäki, T., Kähönen, M., Viikari, J. S., Raitakari, O. T., and Ala-Korpela, M. (2013). Branched-Chain and Aromatic Amino Acids Are Predictors of Insulin Resistance in Young Adults. *Diabetes Care*, 36(3):648–655, DOI: 10.2337/dc12-0895.

Xia, J. and Wishart, D. S. (2016). Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. In *Current Protocols in Bioinformatics*, pages 14.10.1–14.10.91. John Wiley & Sons, Inc., Hoboken, NJ, USA, DOI: 10.1002/cpbi.11.

Xie, B., Waters, M. J., and Schirra, H. J. (2012). Investigating Potential Mechanisms of Obesity by Metabolomics. *Journal of Biomedicine and Biotechnology*, 2012:1–10, DOI: 10.1155/2012/805683.

Xu, F., Tavintharan, S., Sum, C. F., Woon, K., Lim, S. C., and Ong, C. N. (2013). Metabolic Signature Shift in Type 2 Diabetes Mellitus Revealed by Mass Spectrometry-based Metabolomics. *The Journal of Clinical Endocrinology & Metabolism*, 98(6):E1060–E1065, DOI: 10.1210/jc.2012-4132.

Xu, F., Zou, L., Liu, Y., Zhang, Z., and Ong, C. N. (2011). Enhancement of the capabilities of liquid chromatography-mass spectrometry with derivatization: General principles and applications. *Mass Spectrometry Reviews*, 30(6):1143–1172, DOI: 10.1002/mas.20316.

Yamashita, M. and Fenn, J. B. (1984). Electrospray ion source: another variation on the free-jet theme. *The Journal of Physical Chemistry*, 88(20):4451–4459, DOI: 10.1021/j150664a002.

Yang, W., Chen, Y., Xi, C., Zhang, R., Song, Y., Zhan, Q., Bi, X., and Abliz, Z. (2013). Liquid Chromatography–Tandem Mass Spectrometry-Based Plasma Metabonomics Delineate the Effect of Metabolites' Stability on Reliability of Potential Biomarkers. *Analytical Chemistry*, 85(5):2606–2610, DOI: 10.1021/ac303576b.

Yen, S., Bolte, E., Aucoin, M., and Allen-Vercoe, E. (2018). Metabonomic Evaluation of Fecal Water Preparation Methods: The Effects of Ultracentrifugation. *Current Metabolomics*, 6(1):57–63, DOI: 10.2174/2213235X05666161226164121.

Yin, L.-L., Chen, Z.-Z., Tong, L.-L., Xu, K.-H., and Tang, B. (2009). Progress on Fluorescent Probes for Thiols. *Chinese Journal of Analytical Chemistry*, 37(7):1073–1081, DOI: 10.1016/S1872-2040(08)60117-6.

Yin, P., Lehmann, R., and Xu, G. (2015). Effects of pre-analytical processes on blood samples used in metabolomics studies. *Analytical and Bioanalytical Chemistry*, 407(17):4879–4892, DOI: 10.1007/s00216-015-8565-x.

Yin, P., Peter, A., Franken, H., Zhao, X., Neukamm, S. S., Rosenbaum, L., Lucio, M., Zell, A., Häring, H.-U., Xu, G., and Lehmann, R. (2013). Preanalytical Aspects and Sample Quality Assessment in Metabolomics Studies of Human Blood. *Clinical Chemistry*, 59(5):833–845, DOI: 10.1373/clinchem.2012.199257.

Yu, J., Bi, X., Yu, B., and Chen, D. (2016). Isoflavones: Anti-Inflammatory Benefit and Possible Caveats. *Nutrients*, 8(12):361, DOI: 10.3390/nu8060361.

Yu, Z., Kastenmüller, G., He, Y., Belcredi, P., Möller, G., Prehn, C., Mendes, J., Wahl, S., Roemisch-Margl, W., Ceglarek, U., Polonikov, A., Dahmen, N., Prokisch, H., Xie, L., Li, Y., Wichmann, H.-E., Peters, A., Kronenberg, F., Suhre, K., Adamski, J., Illig, T., and Wang-Sattler, R. (2011). Differences between Human Plasma and Serum Metabolite Profiles. *PLoS ONE*, 6(7):e21230, DOI: 10.1371/journal.pone.0021230.

Yuasa, S., Akagi, R., Ubuka, T., Masuoka, N., and Yao, K. (1990). Excretion of 3-Mercaptolactate-Cysteine Mixed Disulfide, Sulfate and Taurine in Human Urine before and after Oral Administration of Sulfur-Containing Amino Acids. *Acta Medica Okayama*, 44(3):117–122, http://www.ncbi.nlm.nih.gov/pubmed/2382576.

Zhang, A., Sun, H., Wang, P., Han, Y., and Wang, X. (2012a). Modern analytical techniques in metabolomics analysis. *Analyst*, 137(2):293–300, DOI: 10.1039/C1AN15605E.

Zhang, A., Sun, H., and Wang, X. (2013). Power of metabolomics in biomarker discovery and mining mechanisms of obesity. *Obesity Reviews*, 14(4):344–349, DOI: 10.1111/obr.12011.

Zhang, H. and Wang, M. (2009). Search for the smallest random forest. *Statistics and Its Interface*, 2(3):381, http://www.ncbi.nlm.nih.gov/pubmed/20165560.

Zhang, H., Yokoyama, W. H., and Zhang, H. (2012b). Concentration-dependent displacement of cholesterol in micelles by hydrophobic rice bran protein hydrolysates. *Journal of the Science of Food and Agriculture*, 92(7):1395–1401, DOI: 10.1002/jsfa.4713.

Zhang, W., Hankemeier, T., and Ramautar, R. (2017). Next-generation capillary electrophoresis–mass spectrometry approaches in metabolomics. *Current Opinion in Biotechnology*, 43:1–7, DOI: 10.1016/j.copbio.2016.07.002.

Zhang, W., Li, P., Geng, Q., Duan, Y., Guo, M., and Cao, Y. (2014). Simultaneous Determination of Glutathione, Cysteine, Homocysteine, and Cysteinylglycine in Biological Fluids by Ion-Pairing High-Performance Liquid Chromatography Coupled with Precolumn Derivatization. *Journal of Agricultural and Food Chemistry*, 62(25):5845–5852, DOI: 10.1021/jf5014007.

Zhong, X., Maxwell, E. J., and Chen, D. D. Y. (2011). Mass Transport in a Micro Flow-Through Vial of a Junction-at-the-Tip Capillary Electrophoresis-Mass Spectrometry Interface. *Analytical Chemistry*, 83(12):4916–4923, DOI: 10.1021/ac200636y.

Zhou, B., Xiao, J. F., and Ressom, H. W. (2013). Prioritization of putative metabolite identifications in LC-MS/MS experiments using a computational pipeline. *Proteomics*, 13(2):248–260, DOI: 10.1002/pmic.201200306.

Zhu, P., Oe, T., and Blair, I. A. (2008). Determination of cellular redox status by stable isotope dilution liquid chromatography/mass spectrometry analysis of glutathione and glutathione disulfide. *Rapid Communications in Mass Spectrometry*, 22(4):432–440, DOI: 10.1002/rcm.3380.

Zhu, Z., Lu, J. J., and Liu, S. (2012). Protein separation by capillary gel electrophoresis: A review. *Analytica Chimica Acta*, 709:21–31, DOI: 10.1016/j.aca.2011.10.022.