Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform in Metabolomics for Assessment of Lifestyle Interventions in Human Health

Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform in Metabolomics for Assessment of Lifestyle Interventions in Human Health

By
NAOMI L. KUEHNBAUM

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 (2014) McMaster University
(Chemistry) Hamilton, Ontario

TITLE:	Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform in Metabolomics for Assessment of Lifestyle Interventions in Human Health
AUTHOR:	Naomi L. Kuehnbaum, H.B.Sc. (University of Toronto)
SUPERVISOR:	Dr. Philip Britz-McKibbin
NUMBER OF PAGES:	xxxii, 253

Abstract

Research in this thesis has focused on development and application of novel methodologies that enhance sample throughput and data fidelity when performing untargeted metabolome profiling by multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS). Metabolomics is a valuable tool in functional genomics research to investigate underlying molecular mechanisms associated with human health since metabolites are "real-world" end-products of gene expression. CE-MS is well-suited for metabolomics because it is a high efficiency microseparation technique that can be used to resolve complex mixtures of polar metabolites in human biofluids without complicated sample workup. In this thesis, a novel CE-MS assay for estrogens and their intact ionic conjugates has been described (*Chapter II*) to expand metabolome coverage that enables resolution of positional isomers with high selectivity. This is critical for better understanding of underlying perturbations in estrogen metabolism since the biological activity of estrogens are dependent on specific primary and secondary metabolic transformations. MSI-CE-MS has been introduced as a high-throughput approach for large-scale metabolomic studies based on serial injection of multiple segments of sample within a single fused-silica capillary (*Chapter III*). It reduces analysis times while increasing data quality and confidence in peak assignment together with better quality assurance. An accelerated workflow for metabolomics has also been developed when using MSI-CE-MS, where a dilution trend filter is used as a primary screen to authenticate reproducible sample-derived metabolites from a pooled sample while eliminating spurious artifact and background signals. In this way, complicated time alignment and peak picking algorithms are avoided when processing data in metabolomics to reduce false discoveries. This strategy was subsequently used in two metabolomics applications (*Chapters IV* and *V*) to identify plasma markers associated with strenuous exercise and adaptive training responses following a six-week high intensity interval training. The impact of exercise intervention to improve the glucose tolerance of a cohort of overweight/obese yet non-diabetic women was investigated on an individual level when using a cross-over design. Personalized interventions are critical in designing more effective therapies to prevent metabolic diseases due to inter-subject variations in treatment responses, including potential adverse effects. MSI-CE-MS offers a revolutionary approach for biomarker discovery in metabolomics with high sample throughput and high data fidelity, which is critical for validation of safe yet effective lifestyle interventions that promote human health and reduce risk for chronic diseases.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Philip Britz-McKibbin for his role in shaping my graduate school experience and introducing me to academic research. His eternal optimism helped guide me through the “valleys” that seem to last forever on the winding path that is research, and his enthusiasm for making a difference to people in our community inspires me to strive to produce quality work that impacts others. “PBM>” group members past (Jennilee Gavina, Richard Lee, Lisa D’Agostino) and present (Meera Shanmuganathan, Karen Lam, Alicia DiBattista, Nadine Wellington, Adriana De Macedo, Michelle Saoi) are friends, late-night lab dinner partners, hiking buddies, and biology tutors.

My committee members, Dr. Johannes Zeidler and Dr. Jose Moran-Mirabel have offered insight in areas beyond my expertise and outside opinions when I needed to step back and re-evaluate my work. The late Dr. Brian McCarry was a vibrant, animated scientist with a passion for chemistry and mass spectrometry. Dr. Kirk Green and Tadek Olech in the mass spectrometry facility were always willing to offer tech support, strong arms, and expertise in antiquated ion traps, saving me many hours of floundering and teaching me more about the inner workings of mass spectrometers than I could have imagined. Dr. Willie Leigh has been a steadfast figure throughout my Ph.D., demonstrating to me endless enthusiasm for chemistry and showing me that that teaching the new generation of chemists can (and should be) fun, both for them and for me. This work would not have been possible without my collaborators, specifically Dr. Martin Gibala, Dr. Stuart Phillips, Jenna Gillen, Mike Percival, Amy Hector, and George Marcotte. Their organization of studies, phlebotomist training, and early morning coffee runs made my part of the projects straightforward and enjoyable.

My family has supported me through my seemingly never-ending education. I’m my mom’s “micro-millilitre girl” and my dad inspired a love of science in me at a young age. I thank my parents for grounding me at the age of 7 after finding me mixing household items in a “potion” because they taught me the potential hazards of mixing chemicals. My favourite sister Erica, favourite sister Heidi, and favourite brother Graham have been a great support to me, after having learned that as long as they didn’t ask me when I would be done with school, I would remain in a pleasant mood. Lastly, I thank my husband Adam. Living apart hasn’t been easy but it has given me space to work late in the lab without feeling guilty. I thank him for his continuous encouragement and unending support of my seemingly unending schooling.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

δ_{mass}	=	mass error
ε	=	dielectric constant
ζ	=	zeta potential
η	=	buffer viscosity
μ_{eo}	=	electroosmotic flow
μ_{ep}	=	electrophoretic mobility
μ_{ep}^{a}	=	apparent electrophoretic mobility
2D	=	2-dimensional
2hPG	=	2 hour postload plasma glucose
3-MeHis	=	3-methyl-*L*-histidine
A1C	=	glycated hemoglobin
A3G	=	androsterone 3-glucuronide
A3S	=	androsterone 3-sulfate
AA	=	ammonium acetate
AAA	=	aromatic amino acid
AC	=	ammonium carbonate
ADMA	=	asymmetric dimethylarginine
ANOVA	=	analysis of variance
ANP	=	aqueous normal phase
APCI	=	atmospheric pressure chemical ionization
APPI	=	atmospheric pressure photoionization
Asp	=	*L*-aspartic acid
ATP	=	adenosine triphosphate
AUC	=	area under the curve
BCAA	=	branched chain amino acid
BEH	=	bridged ethylene hybrid
Bet	=	betaine
BGE	=	background electrolyte
BMI	=	body mass index
BSTFA	=	bis(trimethylsilyl)trifluoroacetamide
C0	=	*L*-carnitine
C2	=	*O*-acetyl-*L*-camitine
C18	=	octadecylsilane
CAD	=	coronary artery disease
CE	=	capillary electrophoresis
CEC	=	capillary electrochromatography
CE-TOF-MS	=	capillary electrophoresis-time of flight-mass spectrometry
CFI	=	Canada Foundation for Innovation

CI	=	chemical ionization
Cit	=	*L*-citrulline
Cl-Tyr	=	3-chloro-*L*-tyrosine
CMIT	=	continuous moderate intensity training
Cret	=	creatine
Cretn	=	creatinine
CTN	=	creatine
CV	=	control variable
CV	=	coefficient of variation
CV	=	control variable
CYP	=	cytochrome P450
Cys	=	cysteine
CysGly-CysSS	=	cysteinylglycine-cysteine disulfide
CysSS	=	cysteine disulfide
DA	=	dopamine
DESI	=	desorption electrospray ionization
DHEAS	=	dehydroepiandrosterone 3-sulfate
DI	=	direct infusion
DMG	=	dimethyl glycine
DXA	=	dual x-ray absorptiometry
E	=	applied electric field
E1	=	estrone
E2	=	estradiol
E3	=	estriol
ECNI	=	electron capture negative ionization
EDTA	=	ethylenediaminetetraacetic acid
EI	=	electron ionization
EIE	=	extracted ion electropherogram
EOF	=	electroosmotic flow
ESI	=	electrospray ionization
ESI-MS	=	electrospray ionization-mass spectrometry
FA	=	formic acid
FG	=	fasting glucose
FID	=	flame ionization detection
FPG	=	fasting plasma glucose
Fru	=	fructose
FTICR	=	Fourier transform ion cyclotron resonance
FTIR	=	Fourier transform infrared spectroscopy
GABA	=	gamma-aminobutyric acid
Gal	=	galactose

GalNAc	=	*n*-acetyl galactosamine
GC	=	gas chromatography
GC-MS	=	gas chromatography-mass spectrometry
Gen7G	=	genistein 7-glucuronide
GlcNAc	=	*n*-acetyl glucosamine
Glu	=	glucose
GlyCys-SS	=	cystinyl glycine disulfide
GSH	=	oxidized glutathione
GSH-Cys-SS	=	glutathionylcysteine-cysteine disulfide
GSH-Cys-SS	=	oxidized glutathionylcysteine mixed disulphide
GSSG	=	glutathione disulfide
GWAS	=	genome wide association studies
H	=	plate height
H_2O	=	water
HA	=	hippuric acid
HCA	=	hierarchical cluster analysis
HIIT	=	high intensity interval training
HILIC	=	hydrophilic interaction liquid chromatography
HIT	=	high intensity interval training
HMDB	=	Human Metabolome Database
HP-921	=	hexakis(2,2,3,3-tetrafluoro-propoxy)phosphazine
HPLC	=	high performance liquid chromatography
HR_{max}	=	maximal heart rate
HS	=	head space
HyX	=	hypoxanthine
Ile	=	isoleucine
IS	=	internal standard
L-Ala	=	*L*-alanine
L_c	=	capillary length
LC-MS	=	liquid chromatography-mass spectrometry
L_d	=	capillary length to the detector
Leu	=	leucine
LOD	=	limit of detection
L-Pro	=	*L*-proline
LTQ	=	linear ion trap quadrupole
Lys	=	lysine
m/z	=	mass to charge ratio
Man	=	mannose
MannNAc	=	*n*-acetyl mannosamine
MCF	=	methylchloroformate

MEKC	=	micellar electrokinetic chromatography
MeOH	=	methanol
Met	=	methionine
MFE	=	molecular feature extractor
MICZE	=	multiple injection capillary zone electrophoresis
MLR	=	multiple linear regression
MNA	=	*N*-methylnicotinamide
MRM	=	multiple reaction monitoring
MS	=	mass spectrometry
MS/MS	=	tandem mass spectrometry
MSI-CE-MS	=	multisegment injection-capillary electrophoresis-mass spectrometry
MS^n	=	multi stage mass spectrometry
MSTFA	=	methyl(trimethylsilyl)trifluoroacetamide
MT	=	melatonin
MTBSTFA	=	tert-butyl(trimethylsilyl)trifluoroacetamide
MV	=	molecular volume
MWCO	=	molecular weight cut-off
N	=	plate number
NH_4Ac	=	ammonium acetate
NICI	=	negative ion chemical ionization
NMR	=	nuclear magnetic resonance
NSERC	=	National Science and Engineering Research Council of Canada
O	=	olfactometry
ODS	=	octadecylsilane
OGI	=	Ontario Genomics Institute
OGTT	=	oral glucose tolerance test
Orn	=	*L*-ornithine
PAG	=	phenylacetylglutamine
PCA	=	principle component analysis
PDMS	=	polydimethylsiloxane
PFPP	=	pentafluorophenylpropyl
PG	=	plasma glucose
Phe	=	phenylalanine
pK_a	=	acid dissociation constant
PKU	=	phenylketonuria
PLS-DA	=	partial least-squares-discriminant analysis
ppm	=	parts per million
psi	=	pounds per square inch
QC	=	quality control

Q_{eff}	=	effective charge
QqQ	=	triple quadrupole
R^2	=	coefficient of determination
R_H	=	hydrodynamic radius
RMT	=	relative migration time
ROC	=	receiver operating characteristic
ROCCET	=	receiver operating characteristic curve explorer & tester
RPA	=	relative peak area
RPE	=	rating of perceived exertion
rpm	=	rotations per minute
RRR	=	relative response ratio
RSD	=	relative standard deviation
S/N	=	signal to noise ratio
SDMA	=	symmetric dimethyl arginine
SPE	=	solid phase extraction
SPME	=	solid phase microextraction
SRM	=	selected reaction monitoring
T3	=	trifunctional
TCA	=	tricarboxylic acid
t_{eo}	=	migration time of electroosmotic flow
TLC	=	thin layer chromatography
TML	=	trimethyllysine
TOF	=	time of flight
TOF-MS	=	time of flight-mass spectrometry
Trp	=	tryptophan
Tyr	=	tyrosine
UA	=	uric acid
UHPLC-MS	=	ultrahigh pressure liquid chromatography-mass spectrometry
V	=	voltage
v	=	electrophoretic velocity
v	=	volume
Val	=	valine
V_{cap}	=	capillary voltage
VIP	=	variable importance in the projection
VO_{2max}	=	maximal oxygen uptake/maximal aerobic capacity
W_{max}	=	maximum workload
z_{eff}	=	effective charge

Declaration of Academic Achievement

The following material has been previously published and is reprinted with written permission:

Chapter I: Reprinted (adapted) with permission from *Chem. Rev.*, **113**, N.L. Kuehnbaum and P. Britz-McKibbin, *New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era*, 2437-2468. Copyright (2013) American Chemical Society.

Chapter II: Reprinted (adapted) with permission from *Anal. Chem.*, **83**, N.L. Kuehnbaum and P. Britz-McKibbin, *Comprehensive Profiling of Free and Conjugated Estrogens by Capillary Electrophoresis-Time of Flight-Mass Spectrometry*, 8063-8068. Copyright (2011) American Chemical Society.

Chapter III: Reprinted (adapted) with permission from *Anal. Chem.*, **85**, N.L. Kuehnbaum, A Kormendi, and P. Britz-McKibbin, *Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-throughput Platform for Metabolomics with High Data Fidelity*, 10664-10669. Copyright (2013) American Chemical Society.

Chapter I

Introduction to Metabolomics, Capillary Electrophoresis-Mass Spectrometry and Lifestyle Interventions for Diabetes Prevention

I. Introduction to Metabolomics, Capillary Electrophoresis-Mass Spectrometry and Lifestyle Interventions for Diabetes Prevention

1.1 Metabolomics

1.1.1 Introduction to Metabolomics

Metabolomics offers a revolutionary framework for phenotyping individuals at a molecular level that is needed for new breakthroughs in cell biology and personalized medicine.[1–3] Major interest in metabolomics is rapidly expanding in diverse research fields ranging from biotechnology and agriculture,[4] drug development,[5] nutrition,[6] food safety/quality,[7] clinical diagnostics,[8] to environmental chemistry and toxicology.[9] Metabolomics (or metabonomics) is defined as the comprehensive analysis of all low molecular weight metabolites (< 1500 Da) in a sample, such as a cell, biofluid or organism. Although several related "omics" sub-disciplines have been recently coined in the literature,[10-11] they all share the common motif of untargeted analysis of low molecular weight compounds using new advances in nuclear magnetic resonance (NMR) and mass spectrometry (MS). However, direct coupling of high resolution separation techniques to NMR[12,13] and notably MS-based metabolomics[14–16] greatly improves analytical performance when analyzing complex sample mixtures while providing complementary qualitative information for metabolite identification. For instance, known metabolites in human serum comprise greater than 4,000 chemically diverse compounds present at concentration levels that span over nine orders of magnitude,[17] which further expands to over 40,000 annotated compounds when including exogenous metabolites derived from foods, microbiota, drugs and synthetic chemicals.[18] Similar chemical diversity also exists in plants as sessile organisms given their complex secondary metabolism.[19] Thus, the total number of detectable yet identifiable compounds is extensive, which necessitates the use of complementary analytical platforms for unbiased screening of the metabolome.[20]

Since metabolites represent molecular endpoints of gene expression and cell activity, metabolomics offers a holistic approach for understanding the phenotype of an organism. In this context, it plays a vital role in systems biology for elucidating the function of unknown genes[21] and enzymes[22] while revealing molecular mechanisms underlying drug action,[23,24] dietary changes,[25,26] exercise intervention[27,28] and psychosocial impacts of environment.[29] Unlike other "omics" domains, the timescale of the metabolome as measured by isotope tracer

experiments is fast within milliseconds,[30] where subtle changes in gene expression and/or protein activity can lead to large amplifications in metabolite concentrations in order to maintain stability of highly interconnected metabolic networks.[31] There is growing recognition that living systems function as dynamic ecosystems or "superorganisms"[32] underscoring the complex interactions of host, diet and gut microflora in human health.[25,33–35] Thus, metabolomics offers unique insight into the functional impact of gene expression[36] complementary to large-scale gene sequencing efforts[37] and genome wide association studies (GWAS) that neglect the impact of environment in the etiology of human diseases.[38] For instance, most genetic mutations are not directly associated with clinical phenotype, whereas aberrant metabolism is strongly linked to disease severity and intoxication as in classical phenylketonuria (PKU).[39] Recent cytochrome P450 (CYP) genotyping initiatives for predicting drug toxicity among individuals have been largely unsuccessful since the expression and functional activity of CYP isoforms is dependent on various confounding non-genetic variables, such as occupational exposure.[40] Indeed, the majority of clinical diagnostic markers, pharmacological agents and environmental pollutants are primarily small molecules. However, there is currently no single approach amenable to comprehensive metabolite profiling due to the chemical diversity and wide dynamic range of the metabolome that remains largely uncharacterized in complex biological samples.[41]

1.1.2 The Advent of Metabolomics

Major international efforts at sequencing the genomes of organisms from the mid-1990s and early 2000s[42–44] sparked renewed interest in metabolite profiling as a way to decipher genes of unknown function or mutants that possess "silent" phenotypes.[45,46] The initial euphora derived from high-throughput gene sequencing projects soon led to the realization that gene regulation and interactions on a "systems-level" were critical for functional understanding of cell activity as a whole.[47,48] Despite earlier developments in DNA microarray and proteomic approaches in post-genomic research,[49] the first published references to "metabolome" did not occur until 1998 by Tweeddale *et al.*[50] and Oliver *et al.*[51] that involved comparative studies of wild-type and gene-deleted mutant strains of bacteria and yeast incubated under different growth conditions using Fourier-transformed infra-red (FTIR) spectroscopy and 2D TLC separations, respectively. The Ferenci group[52] later reported the impact of oxidative stress in bacteria based

on measured changes in ^{14}C-glucose-labeled metabolites in *E. coli* extracts when using 2D TLC with phosphorimaging in conjunction with an amino acid analyzer. Although their work revealed increased intra-cellular levels of valine and glutathione in paraquat-treated bacteria relative to controls, a major fraction of metabolites were unresolved, whereas two spots that decreased upon paraquat exposure remained unidentified.[52] In 1999, Nicholson *et al.*[53] introduced the term "metabonomics" when assessing the differential responses of rats exposed to various model toxins based on time-dependent metabolite changes in urine when using ^{1}H-NMR together with multivariate statistical analysis. Several follow-up studies were published in 2000 on the use of NMR-based metabonomics for evaluating drug-induced toxicity in model animal organisms,[54–56] where changes in urinary metabolic patterns represent a metabolic phenotype or "metabotype" that is characteristic for different rodent species[57] and genetic strains of mice.[58] However, only a modest number of metabolites were quantified and identified in urine, primarily abundant osmolytes, amino acids and organic acids. Subsequent work by Fiehn *et al.*[59] introduced GC-MS-based metabolomics for plant functional genomic studies based on quantitative analysis of 326 distinct metabolites detected in leaf extracts derived from four genotypes of *Arabidopsis thaliana*. Their work demonstrated that single locus mutants were distinguished from wild type plants due their characteristic metabolic clustering after principal component analysis (PCA). In this case, identification of about half of all detected molecular features was achieved by comparison of relative retention times in conjunction with EI-MS spectra from a chemical library.[59]

The early adoption of NMR in metabolomics stems from its excellent robustness and reproducibility[60,61] that is convenient for rapid screening of biofluids with minimal sample handling[62] as well as direct analysis of nanolitre tissue biopsies.[63] Nevertheless, the limited sensitivity of ^{1}H-NMR enables characterization of only a few dozen metabolites in most biological samples with up to 50 metabolites with concentrations above 10 μM recently identified in human serum.[17] In addition, stable isotope tracer experiments[64] or chemical derivatization[65] are needed for acquiring spectra for other lower abundance nuclei, such as ^{13}C. The use of 2D NMR pulse sequence methods can increase spectral peak resolution however at the cost of reduced sensitivity and longer acquisition times.[66] In contrast, direct infusion-MS offers a sensitive yet high-throughput platform for metabolic phenotyping[67,68] that can measure hundreds of molecular features without separation, where accurate mass and relative isotope abundance measurements allow for determination of elemental composition.[69] However, it suffers from poor accuracy and precision due to matrix-induced signal

suppression or enhancement that prevents reliable quantification when comparing different specimens without stable-isotope internal standards.[70] Moreover, a search query for a putative elemental formula from public databases results in hundreds of potential chemical structures and their isomers that is inadequate for structural elucidation of unknown metabolites.[71]

Thus, there is a clear need to expand metabolome coverage with greater sensitivity and selectivity while ensuring robust analytical performance to avoid false discoveries. This has led to expanding interest in hyphenated separation techniques in metabolomics first reported in 2002, such as hydrophilic interaction liquid chromatography (HILIC-MS),[72] RP-LC-MS,[73] and CE-MS[74] that are ideal for resolution of polar, non-polar and ionic classes of metabolites, respectively. Despite earlier divergence in terminologies used in the literature, metabolomics and metabonomics[75] are now considered interchangeable terms and in essence refer to the same paradigm of comprehensive identification and quantification of "all" metabolites in a given specimen,[76] although the terms are usually reserved to describe biologically significant molecular features identified across different samples/groups after rigorous data pre-processing and multivariate analysis. This process is distinct from classical metabolite profiling involving quantitative analysis of targeted compound classes[77] or metabolite fingerprinting, where rapid pattern recognition/classification of samples is desired without full resolution or identification of specific molecular features.[78] An "expanded" targeted metabolomics approach[79] has also been proposed for quantitative analysis of hundreds of known standard metabolites often across multiple analytical platforms.[20,80,81] However, a unique feature of metabolomics is related to "hypothesis-free" testing for discovery of unknown and/or poorly characterized metabolites of biological or clinical significance; indeed, this "untargeted" approach is critical for novel insights into complex mechanisms of drug/toxin action and/or disease pathogenesis on a systemic level.[23,82]

To date, there have been over 2,600 total publications in metabolomics (search query with *metabolom* or metabonom** in title from 1998-2012) that has increased rapidly over the past five years with about 54% of all reports utilizing one or more separation techniques based on LC (32%), GC (17%) and/or CE (5%) as depicted in **Figure 1.1**. Thus, NMR still remains a major platform for metabolomics/metabonomics despite its lower sensitivity, limited metabolome coverage and higher infrastructure costs. Due to the lack of a truly universal method for metabolomics, recent large-scale investigations compile data across multiple instrumental platforms and/or multi-dimensional separation systems.[83]

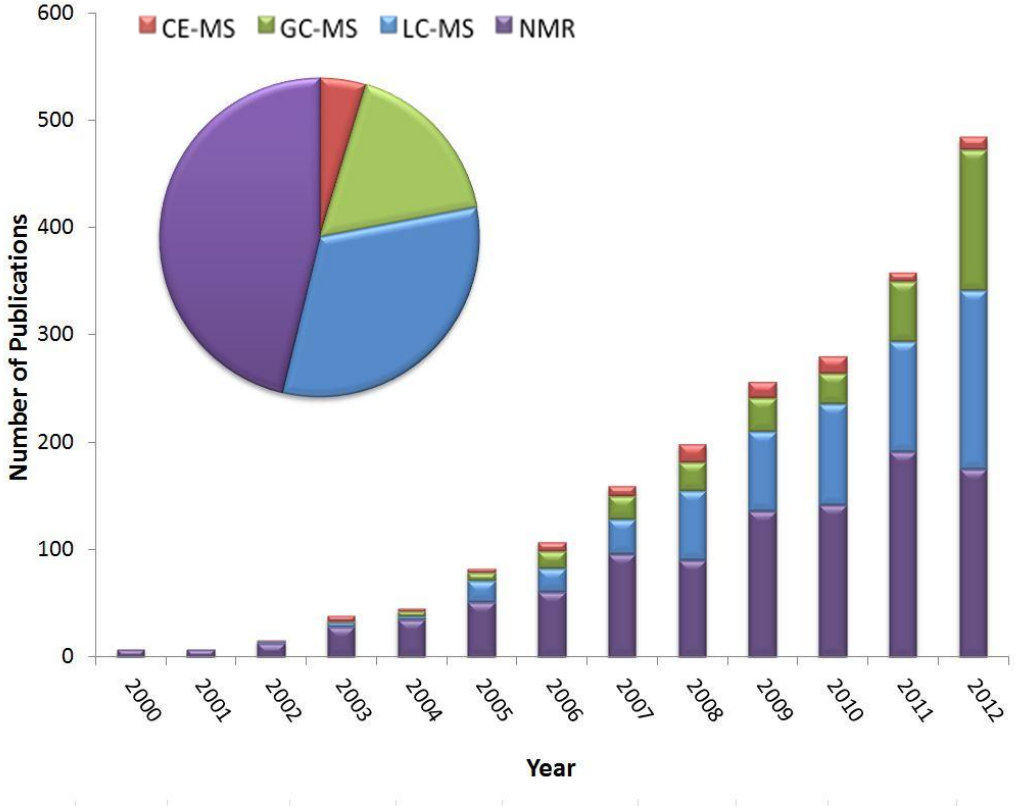


Figure 1.1. Literature survey of published research articles in metabolomics using a search query with keywords "*metabolom* or metabonom*"* in title from 1998-2012 via Web of Knowledge v. 5.8 excluding reviews and abstract proceedings. The pie chart insert depicts the average fraction of articles based on metabolomics using NMR (46%), LC-MS (32%), GC-MS (17%) and CE-MS (5%) over the past five years (2008-2012). Overall, about 54% of all reports utilize one or more separation methods that is critical for expanding the coverage of the metabolome. There is growing interest in using complementary separation modes or multi-dimensional separations, as well as compiling data across multiple analytical platforms for expanded metabolite profiling.

Despite several critical reviews dedicated to instrumental developments in NMR and/or MS-based metabolomics,[16,84–86] there have been few systematic reviews of separation science[87–89] with most reports restricted to specific separation platforms based on GC-MS,[90] LC-MS,[91–94] or CE-MS.[95,96] Thus, after 60 years since Williams' pioneering work with 2D TLC for evaluating biochemical individuality[97] and 40 years since Horning's classic work on metabolite profiling of human biological fluids by GC-MS,[98] new advances in separation science continue to play vital roles in recent metabolomics initiatives.

1.2 Separation Science in Metabolomics

1.2.1 Origin of Metabolomics and its Link to Separation Science

The origin of metabolomics can be traced to early technology developments in separation science and clinical medicine early last century, as well as to the hypothesis of "biochemical individuality" that linked metabolite patterns present in biofluids with phenotypic traits associated with human health and disease susceptibility.[97,99] Williams and co-workers[100] were early practitioners of metabolomics using paper chromatography to analyze individual metabolic patterns or "metabolic personality" in biological fluids derived from human subjects and animal models. His breakthrough work countered genetic fatalism by demonstrating that dietary changes and nutritional supplementation can treat human behavioral disorders of obscure etiology, including alcoholism and schizophrenia. The advent of paper and notably 2D thin layer chromatography (TLC)[101,102] in the 1940s and 1950s further improved peak capacity for resolution of hundreds of "spots" from complex biofluids that were detected by their intrinsic UV absorbance and/or reaction with specific colorimetric reagents.[103,104] However, the low separation efficiency, long elution times, poor automation and lack of quantitative analyses limited the applicability of planar chromatography for routine metabolite profiling applications.[77]

Higher resolution separation methods for quantitative metabolite profiling awaited later advances in chromatography,[105] including column technology, bench-top instrumentation and computer data processing and storage capabilities. For instance, Hamilton[106] reported that up to 95 ninhydrin-positive compounds were resolved in human urine after separation on a synthetic anion-exchange resin with post-column chemical derivatization when using a commercial amino acid analyzer. Similarly, Pitt *et al.*[107] demonstrated that up to 150 UV-absorbing metabolites were detected in human urine when using an ion-exchange resin with an optimized buffer gradient elution performed over 20 hrs. New advances in high performance liquid chromatography (HPLC) awaited the development of uniformly packed columns prepared from porous microparticles (*i.e.*, type B silica < 10 μm i.d.) modified with octadecyl silica (*i.e.*, reversed-phase, RP).[108,109] This greatly enhances separation efficiency and sample throughput with over 100 metabolites resolved from an acidified urine extract by RP-HPLC under 30 min when using a solvent gradient elution program.[110] Separation selectivity was further expanded when using ion pairing reagents in the mobile phase to increase the retention behavior of polar metabolites.[111–113] Since the introduction of

partition column chromatography, rapid developments were also taking place in gas chromatography (GC) that coincided with the development of sensitive post-column detectors for gas effluents based on flame ionization detection (FID),[114] electron capture,[115] and electron ionization (EI)-MS.[116–118] Pauling *et al.*[119] demonstrated that over 250 peaks from a head-space analysis of human urine vapor or breath condensates were detected by GC-FID whose composition was reproducible within a few days provided that standardized methods were used to collect samples from individual subjects. The first commercial GC-MS instruments with temperature programming together with reactive silylating reagents further enhanced method performance for characterization of multi-component mixtures in biological samples.[120,121] Untargeted metabolite profiling soon became a reality with GC-MS instruments configured with low resolution quadrupole mass filters that allowed for repetitive mass scanning over the full mass range during separation with improved resolution of co-eluting peaks.[98,105,122,123] Also, the use of stable isotopes as internal standards in GC-MS provided exceptional precision and accuracy for quantitative analysis of metabolites when using selective reaction monitoring,[124] whereas retention index markers based on co-injection of a homologous series of *n*-alkane standards corrected for variation in retention times.[125] The use of glass-based capillary columns to reduce axial diffusion with the subsequent availability of flexible open tubular fused-silica capillaries soon transformed GC-MS into the "gold standard" for metabolite profiling.[126–129]

Although paper and gel electrophoresis have been important separation techniques for preparative and qualitative analysis of charged amino acids, proteins and nucleic acids since the seminal work of Tiselius in 1930,[130] it was not until narrow bore tubes[131] and subsequently fused-silica capillaries[132,133] were introduced as effective heat dissipation media to allow for fast separations under a high voltage by capillary electrophoresis (CE). High separation efficiencies are realized due to the near flat profile of the electroosmotic flow (EOF) relative to pressure-driven laminar flow[134] and the lack of mass transfer kinetics when performing electrophoresis in free solution. Selectivity in CE is readily extended to include neutral solutes[135] and their enantiomers[136] when using charged micelles and chiral selectors as additives in the background electrolyte (BGE), respectively.

Despite earlier attempts at direct coupling of LC with MS via atmospheric pressure ionization[137–141] the first report of direct infusion electrospray ionization (ESI) of intact proteins by Fenn and co-workers[142] revolutionized metabolite

profiling by providing a "soft" ionization approach to desolvate ions prior to their extraction into the low pressure region of the mass analyzer. The introduction of commercial bench-top tandem mass spectrometer instruments in the 1990s transformed metabolite profiling in the clinical laboratory by enabling multiplexed and high-throughput analysis by direct infusion ESI-MS/MS that is rapidly replacing traditional biochemical assays.[143,144] The phenomenal explosion of LC-MS during the past two decades stems from the ability to perform selective yet sensitive analyses of polar and thermally unstable metabolites not amenable by GC-MS without complex sample pretreatment.[145] However, direct coupling of CE with ESI-MS has proven to be more challenging relative to HPLC due to the high electric field and the low flow rate of the EOF that hinder formation of a stable microspray.[146,147] Although NMR has long been the premier technique for structural elucidation of small molecules with early studies demonstrating its potential for multi-component analysis in biological fluids,[148,149] direct coupling to HPLC[150] and CE[151] still remains non-trivial when acquiring high quality spectra during sample elution. Due to fundamental limitations in the selectivity and sensitivity of single instrumental platforms, there is growing recognition of the need for orthogonal separation techniques for expanded coverage of the metabolome,[17,20,152–156] including integrated total analysis systems based on LC-solid-phase extraction (SPE)-NMR-MS.[157] Given the chemical diversity and dynamic range of the metabolome that remains largely uncharacterized, separation techniques based on chromatography and electrophoresis offer distinct advantages for different classes of metabolites as summarized in **Table A1.1**.

1.2.2 Improved Column Technology in Chromatography

Modern separation science has evolved rapidly since Mikhail Tswett's seminal report on chromatography, which described the extraction, separation and recovery of chloroplasts from plant extracts.[158] A common principle underlying all separation techniques based on differential solute migration is that the rate of generation of variance determines overall peak width, which impacts separation efficiency (plate height, *H*), resolution and peak capacity.[159] The mass balance equation can be used to describe the movement of a solute as a boundary peak through a tube, where both separative and non-separative processes influence its average linear velocity and rate of band broadening.[160] Conditions that maximize separative transport while minimizing band dispersion enable higher separation efficiency that is required for resolving complex sample mixtures prior to

detection. The advent of comprehensive two-dimensional separations further enhances separation performance via direct coupling of two columns in series that have complementary retention mechanisms, since the overall peak capacity of the system is the product of the peak capacity for each dimension under ideal conditions.[161] In chromatography, separation is driven thermodynamically by the distribution equilibrium of a solute between a mobile phase and an immiscible stationary phase, whereas the extent of band broadening is dependent on diffusion and mass transfer kinetics.[162] Selectivity in chromatography is determined by differences in retention factor due to favorable changes in the enthalpy and/or entropy of a solute upon its reversible transfer from mobile phase to stationary phase. The major modes of separation in chromatography thus exploit specific types of non-covalent intermolecular interactions to retain solutes based on adsorption, partitioning and/or ion-exchange mechanisms; non-retained solutes co-elute with the mobile phase void volume that defines the hold-up time of a chromatographic system. Retention behavior is most readily optimized by changes in the solvent composition of the mobile phase in LC when using covalently-modified silica particles,[163] whereas temperature programming is used in GC to enhance solute volatility and desorption in wall-coated open tubular capillaries.[164] Gradient elution programs are frequently used in order to achieve faster elution times relevant for optimal resolution of complex sample mixtures; however, sufficient time for re-equilibration of the column is required for reproducible retention times.[165,166] In general, GC is the method a choice for volatile, non-polar and thermally stable metabolites due to its high separation efficiency, reproducible retention times and extensive EI-MS database library.[90] Selectivity in GC is determined primarily by the polarity of the stationary phase since an inert gas (*e.g.,* He) serves to passively transport sample through the capillary, where non-polar analytes elute as a function of their volatility and molecular weight. The advent of immobilized/cross-linked ionic liquid stationary phases for GC-MS allows for unique dual retention behavior for polar and non-polar classes of metabolites with higher thermal stability (> 350°C) and minimal column bleed that has yet to be fully explored in metabolomics.[167,168] Further improvements in thermally stable/ultra-low bleed stationary phases in conjunction with commercial two-dimensional gas chromatography-time of flight-mass spectrometry (2D GC x GC-TOF-MS) instruments with cryogenic modulation[169] to minimize off-column peak broadening[170] herald separations with unprecedented peak capacity for metabolomics.

Recent developments in LC column technology over the past decade have greatly improved separation efficiency while expanding selectivity, which is now increasingly dominated by columns packed with fused-core and sub-2μm porous particles for ultra-fast separations.[171] The ratio of the hold-up time to column efficiency (*i.e.,* Poppe plot) can be used to assess the resolving power of a separation system, which is favored when using columns that possess high porosity and low plate heights while operating under high linear velocities.[172] The introduction of monolithic columns based on a continuous porous silica rod allows for fast separations due to its high porosity with plate heights comparable to conventional 5 μm porous particles that is achieved at much lower back pressures.[173–175] However, separation efficiency in monolithic columns is compromised by Eddy diffusion due to their radial heterogeneity in wide bore columns, which can be improved when using long and narrow capillary formats at low linear velocities.[176,177] Alternatively, the reduction in particle size within packed columns is an effective way to enhance separation efficiency[178] by minimizing variance contributed by Eddy diffusion and mass transfer kinetics, which also increases total number of metabolites detected due to their elution as sharper peaks with higher signal/noise.[179] Since pressure drop across packed columns is inversely dependent on the square of particle size, specialized equipment is required for operating under optimum flow rates that exceed conventional HPLC systems with particles having average diameter ranging from 1.0 to 1.7 μm.[180–182] For instance, commercial instruments based on ultra high pressure or very high pressure liquid chromatography (UHPLC or vHPLC) generate back pressures greater than 700 bar or 10,000 psi when packed with 1.5 μm porous particles that result in a two-fold enhancement in peak efficiency relative to 3 μm particles with a concomitant reduction in analysis time under optimum linear flow rates.[183] For instance, UHPLC using a C18 stationary phase with gradient elution provides higher resolution and improved sensitivity relative to conventional HPLC along with a three-fold shorter run time as shown in **Figure 1.2**.[184]

The shorter run times in UHPLC allows for the implementation of two separate runs optimized for positive and negative ions in less time than a single HPLC separation while also allowing for the identification of more than twice as many plasma metabolites with greater precision.[184] However, separation performance is ultimately limited by resistive heating at very high pressures and extra-column band broadening notably when using micro- or nanoflow UHPLC systems.[185] Operating columns at elevated temperatures (> 40°C) offers a simple

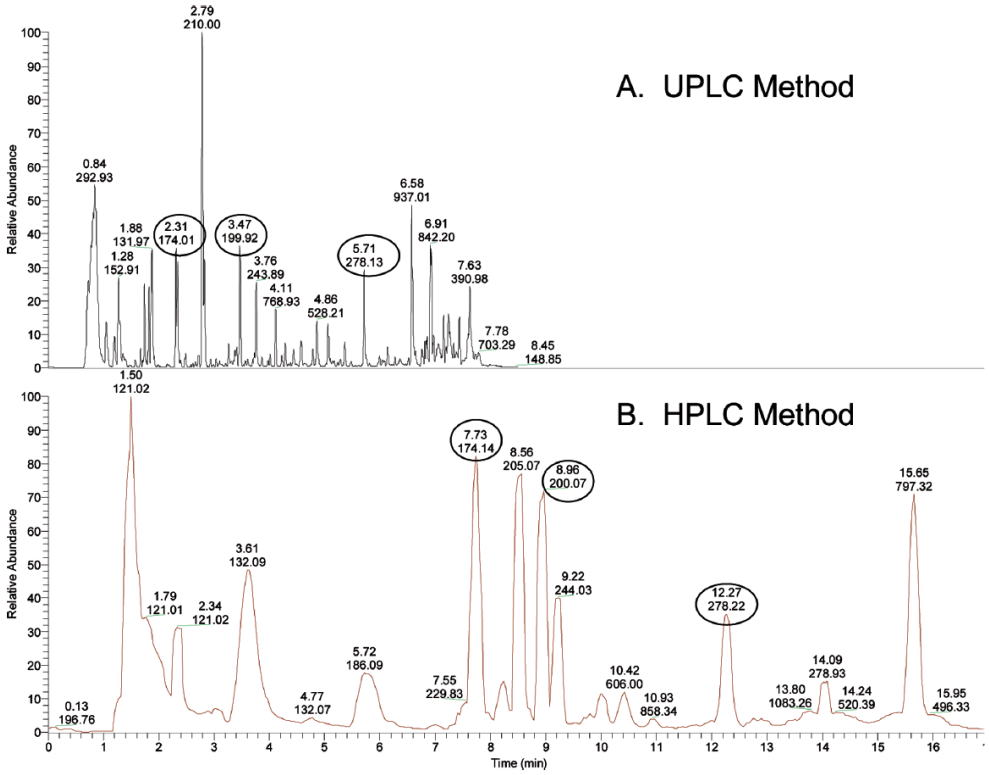


Figure 1.2. Comparison of base peak LC-MS chromatograms based on untargeted metabolite profiling of methanol-extracts of human plasma reconstituted in 0.1% formic acid (UPLC or UHPLC) and 10% methanol (HPLC) that highlights the increase in resolution power and sensitivity along with three-fold reduction in run time when analyzing the same sample under positive ion mode. LC conditions were (a) 2.1 mm × 100 mm BEH C18 1.7 µm particle columns heated to 40 °C with a gradient-elution at 350 µL/min or (b) HPLC-MS method using a 3 µm particle 2.1 mm × 100 mm Aquasil C18 column with gradient-elution at 200 µL/min. The three major peaks circled represent common ions detected in both LC-MS conditions as a reference for comparison of separation performance.[184]

way for achieving faster separations at lower back pressures due to the lower viscosity of the mobile phase that can enhance peak capacity up to a 1000 with plate numbers in excess of 100,000.[186–188] Nevertheless, most siloxane-based materials have limited pH stability (pH ≈ 6-8) at high temperatures, whereas residual silanol residues can adsorb basic solutes resulting in band broadening and peak asymmetry.[189] However, bridged ethylene hybrid (BEH) modified silica

stationary phases reduce the number of residual silanol moieties that contribute to undesirable ion-exchange processes while also improving the pH/thermal stability of the column.[190] The recent development of "superficially porous" sub-3 μm fused-core particles provide highly efficient columns without high back pressures that achieve similar separation performance relative to sub-2 μm porous particles while using standard HPLC instrumentation.[191] In this case, a 1.7 μm solid silica particle is used as support for growing an outer thin porous shell (0.5 μm) resulting in greater mass transfer kinetics even at high linear flow velocities.[192]

Since the metabolome is composed a diverse array of compounds with different physicochemical properties, no single retention mechanism is adequate to resolve complex sample mixtures that vary widely in their polarity, charge and stability.[94] In most cases, orthogonal separation modes are used in parallel to analyze the same sample as a way to expand metabolome coverage for both non-polar and polar metabolites,[193] such as HILIC and RP-LC.[153,194] Alternatively, there is growing interest in developing "mixed-mode" retention mechanisms based on dual modes[195] or tri-modal stationary phases within a single column that exploit RP, cation-exchange and/or anion-exchange retention mechanisms;[196] however, method optimization is challenging given the large number of experimental variables that can impact separation performance. To date, RP-LC is by far the most widely used mode of chromatography in instrumental analysis, including metabolomic studies reported to date (> 90%) when using silica-based octadecylsilane (ODS, C18) bonded-phase columns in conjunction with an aqueous mobile phase with methanol and/or acetonitrile as organic modifiers.[197] Solvophobic theory[198] and quantitative structure-retention relationships[199] have been developed to accurately model the retention behavior of non-polar and semi-polar solutes based on their physicochemical properties when using type-B alkylsilica stationary phases. Indeed, five different types of intermolecular interactions fully account for solute retention selectivity in RP-LC, namely hydrophobic, steric hindrance, hydrogen bonding acceptor or donor and Coulombic interactions (*e.g.*, cation-exchange capacity).[200,201] Recent computer simulations of retention mechanisms in RP-LC suggest multiple sorption sites involving both adsorption (at interface) and partition (within bonded phase) interactions at the molecular level that is solute dependent.[202] Indeed, the extensive use of RP-LC stems from its reproducible and predictable retention times, wide applicability to various classes of metabolites/xenobiotics and mobile phase compatibility for direct coupling to ESI-MS.

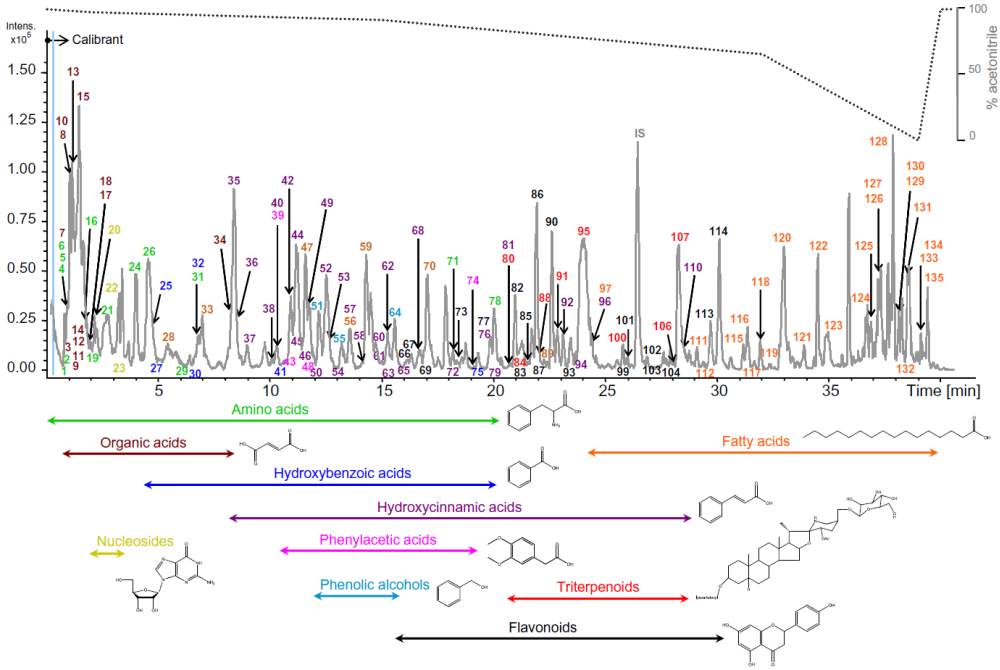


Figure 1.3. Base peak chromatogram highlighting the chemical diversity of semi-polar phytochemicals derived from methanol extracts of tomato (*Daniela*) for comprehensive metabolite profiling of different tomato cultivars when using UHPLC-TOF-MS with an optimized solvent gradient under negative ion mode detection. Over 135 different compounds, including 21 unique phytochemicals reported for the first time in tomato fruit were assigned based on their accurate mass, retention time, as well as UV absorbance and MS/MS spectral matching. Separations were performed on a C18 analytical column (4.6 × 150 mm, 1.8 µm particle size) at 37 °C.[203]

Figure 1.3 highlights the broad range of semi-polar primary (*i.e.*, organic acids) and secondary metabolites (*e.g.*, flavonoids) from methanol extracts of tomato when using RP-LC that can resolve over 135 different compounds when using an optimized solvent gradient elution program.[203] In this case, UHPLC-TOF/MS with negative ion mode enables differentiation of three major cultivars of tomatoes based on their characteristic profile of phytochemicals, notably biological active fractions of phenolic compounds. However, a major limitation of RP-LC is the poor retention of hydrophilic/ionic metabolites that prevents their reliable quantification and identification due to their co-elution within the void volume. In this case, chemical derivatization can be implemented as a strategy to improve the retention of hydrophilic metabolites (*e.g.*, amino acids, amines) when using RP-LC that also enhances their ionization efficiency with lower detection limits.[204,205] A simple "label-free" approach to enhance the retention of polar/charged solutes in RP-LC is the use of ion pairing reagents[113] in the mobile

phase, which also provides pH control while improving peak shape caused by unwanted silica interactions. However, ion pairing RP-LC (and ion-exchange LC) is not readily compatible with ESI-MS due to signal suppression and/or ion source contamination by involatile electrolytes in the mobile phase, such as trifluoroacetate under negative ion mode.[206–208] For this reason, the application of ion pair RP-LC[21,208] for separation of polar metabolites in biological samples is not widely used in metabolomics (< 1% of total LC-MS papers) since separations are constrained by the use of low concentrations of ion pair reagents together with nanoflow ESI interfaces for adequate sensitivity.[209]

HILIC represents a complementary normal-phase LC separation mode for resolving polar/ionic metabolites and their isomers that are poorly retained by RP-LC when using high amounts of organic modifiers (*e.g.,* acetonitrile) containing 10-40% water.[210] Under these conditions, retention of hydrophilic solutes occurs via partitioning onto a hydrated surface of the stationary phase relative to an organic-enriched mobile phase (*i.e.,* biphasic system) with gradient elution performed by increasing the water content. In contrast to conventional normal phase-LC, HILIC is more compatible for polar metabolite analysis when coupled to ESI-MS due to the high content of organic modifier in the mobile phase that leads to improved ionization efficiency and solvent evaporation.[211] A plethora of polar stationary phases have been developed for HILIC ranging from unmodified silica to bonded-phase particles with non-ionic (*e.g.,* polyol), ionic (*e.g.,* amine) or zwitter-ionic (*e.g.,* sulfoalkylbetaine) functional groups, which can result in complex and poorly defined retention mechanisms based on adsorption, partitioning, hydrogen bonding and/or ion-exchange.[212,213] Moreover, HILIC columns exhibit increased retention time drifts and require longer re-equilibration times relative to RP-LC while prone to poor separation performance for some classes of ionic metabolites, such as estrogen conjugates[214] and nucleotides.[215] Alternatively, aqueous normal phase (ANP)-LC based on silica hydride-based stationary phases (*i.e.,* type-C silica columns) offer greater stability since hydride groups replace the majority of surface-active silanol moieties without endcapping, where hydrophilic partitioning occurs via a gradient elution program over a wide range of water content.[216,217] Recent work has demonstrated that ANP-LC-MS offers good retention time repeatability (CV < 1%) with a similar total number of detectable features in human urine as compared to RP-LC that would have otherwise largely eluted in the void volume.[218] Recently, a systematic comparison of the separation performance of RP (C18), ANP (silica hydride) and HILIC (zwitter-ionic) MS-based analysis of human urine demonstrated that HILIC

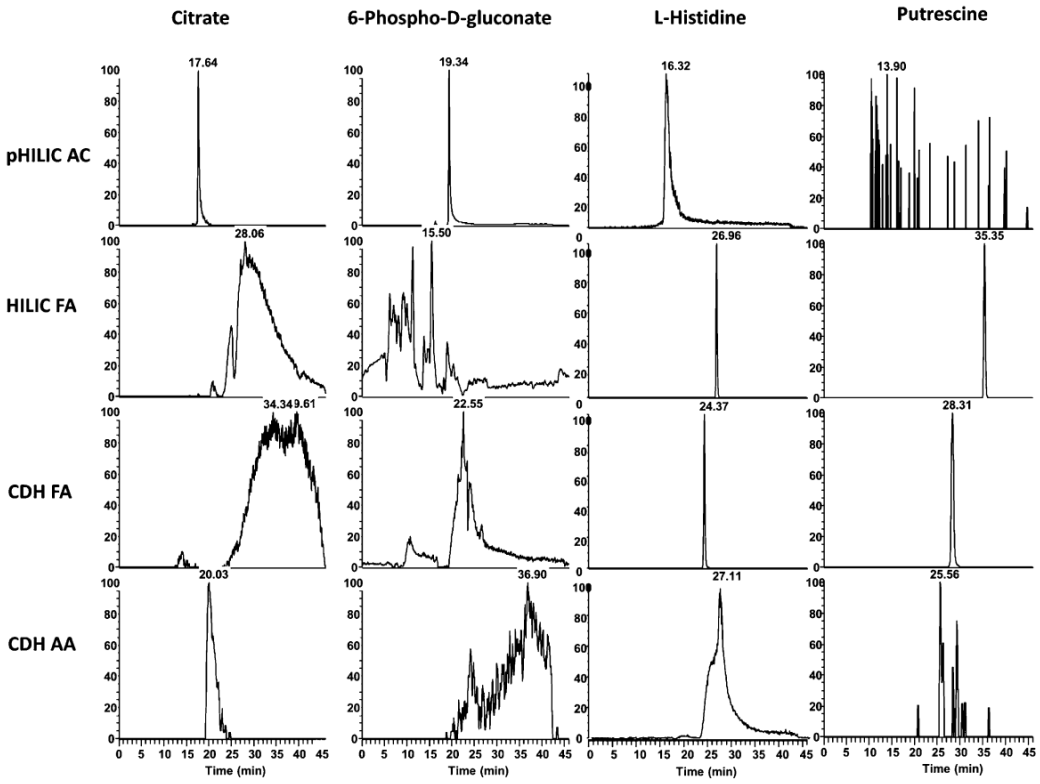


Figure 1.4. Comparison of the separation performance of HILIC-MS with two different types of columns using acidic (formic acid, FA) and basic (ammonium acetate/carbonate, AA/AC) gradient elution conditions for model polar/ionic metabolites unretained by RP-LC. The two zwitter-ionic HILIC columns were found to offer better separation performance for urinary metabolites and their isomers in comparison to ANP (type-C silica hydride) and RP-LC (C18). The three LC-MS columns used for polar metabolite separations are (a) ZIC-HILIC: 150×4.6 mm 5 µm Merck, (b) ZIC-pHILIC: 150×4.6 mm 5 µm Merck and (c) CDH: Cogent Diamond Hydride, 150×4.6 mm 4 µm MicroSolv. Due to their multi-valent charge status that is pH dependent, deleterious ionic interactions occur with both stationary phases that contribute to peak tailing and poor separation efficiency notably for sugar phosphates, tricarboxylic acids, basic amino acids and polyamines.[219]

enabled detection of the greatest fraction of significant features (70%) relative to ANP (53%) and RP (25%) with 12% of total features detected across all three platforms; however, both HILIC and ANP required the application of two different elution conditions under acidic (pH 2.8) and alkaline (pH 9.2) conditions to expand metabolome coverage as a result of the severe peak skewing reported for several classes of ionic metabolites as shown in **Figure 1.4**.[219] Despite these drawbacks, HILIC remains the most common separation mode used in LC-MS for polar metabolites, which is increasingly being used together with RP-LC for comprehensive metabolomic studies.

1.2.3 Electrophoresis for Tunable Selectivity

Separation in electrophoresis is driven electrokinetically under an electric field based on differences in electrophoretic mobility in free solution that reflects the effective charge density of an ion.[220] The ideal state (limiting) mobility of an ion and its thermodynamic weak acid dissociation constant (pK_a) determine the effective charge density of a weakly ionic species at a given buffer pH condition.[221] Due the lack of pressure-driven laminar flow, Eddy diffusion and mass transfer kinetics prevalent in packed columns, separation efficiency in CE is typically an order of magnitude greater than conventional HPLC; however other band broadening sources need to be minimized during separation optimization, such as Joule heating and electrokinetic dispersion.[222] As separation efficiency in CE is directly dependent on the magnitude of the electric field, effective heat dissipation is required when using commercial instruments that permit the application of 30 kV, whereas voltages up to 120 kV have been reported with CE systems with electrical shielding generating plate numbers up to 6 million.[223] Separation optimization in CE is dependent primarily on aqueous buffer conditions (*e.g.*, pH, electrolyte, ionic strength, organic modifier),[224] which can be performed *in silico* for rapid separation optimization based on intrinsic properties of solutes and the buffer system. Unlike LC, separations in CE are typically performed isocratically using an aqueous buffer system, however non-aqueous solutions (*e.g.*, acetonitrile) can also be used as a way to enhance solubility of non-polar metabolites, as well as alter selectivity and improve compatibility to ESI-MS.[225] The latter detector interface necessarily restricts the use of MEKC since nonvolatile micelles in the BGE contribute to ion suppression unless using partial filling methods[226] or alternative ion sources that better tolerate matrix effects, such as atmospheric photoionization (APPI).[227] A unique way to enhance resolution and separation efficiency in CE is the use of "flow counterbalance" mode operation via application of a pressure-induced[228] or EOF counterflow[229] to balance or retard electrokinetic migration in the capillary, which can be used to resolve minor differences in solute mobility, including enantiomers and isotopomers.

Hybrid separation techniques that utilize both electric field (*i.e.*, mobility) and equilibria (*i.e.*, thermodynamic) processes for separative transport, such as capillary electrochromatography (CEC),[134,230] MEKC,[231] and dynamic complexation CE[232] offer unique selectivity for the resolution of a wide range of analytes, including neutral/non-polar and polar/ionic metabolites. In the case of CEC, various approaches have been used to introduce RP partitioning during

solute electromigration, including open tubular, monolithic and porous particle column formats.[233] For instance, ultra-high separation efficiency can be realized in CEC when using uniformly packed sub-micron silica colloidal crystals packed in narrow bore capillaries without massive back-pressure or deleterious resistive heating since fluid transport is driven by the EOF.[234] However, monolithic columns are generally favored in CEC since they do not require complicated packing procedures or the use of frits to restrict electromigration of porous particles since sol-gel formation occurs in-situ with covalent attachment of a porous silica rod onto the capillary surface.[235] Resolution of neutral polyol enantiomers in CE when using charged arylboronic acids as selective sugar-binding additives in the BGE can be achieved electrokinetically based on differences in their complex mobility even in cases when their binding constants are similar.[236,237] When using soluble additives in CE, the retention factor of a solute is defined as the product of partition coefficient (or binding constant) and additive concentration,[238] which allows for independent control of a moving yet discrete "pseudo-stationary phase"[239] in free solution without surface immobilization as required in LC/GC columns. Furthermore, multiple additives with distinct binding mechanisms and/or charge states can be used simultaneously for high resolution chiral separations, such as anionic micelles and neutral cyclodextrins.[240]

There is growing interest in CE-based separations in the pharmaceutical industry that demands robust analytical performance,[241] notably when performing chiral separations, glycan analysis and quality control of biopharmaceuticals. Indeed, capillary gel electrophoresis was a key instrumental platform in multiplexed gene sequencers that contributed to the early completion of the Human Genome Project.[242] However, there are common misconceptions regarding the suitability of CE as a robust platform in metabolomics notably in the context of its poor migration time reproducibility. Since bulk fluid transport is distinct from pressure-driven flow in chromatography, optimized software tools need to be developed for data processing to correct for migration time drifts in CE, which is dependent on both the magnitude and direction of the effective electrophoretic mobility of ion and the EOF.[243] Indeed, the major advantages of the EOF arise from its near flat-like profile that minimizes solute axial diffusion during electromigration that serves as a natural electrokinetic pumping mechanism. However, since the EOF is dependent on the surface properties of the capillary and the electrolyte properties of the BGE, absolute migration times are more variable in CE relative to retention times in RP-LC or GC that are regulated by an external peristaltic pump or pressure regulator;[244] this issue can be readily

addressed by the use of a neutral EOF marker during separation, which allows for precise determination of electrophoretic mobility (*CV* < 1%) when using commercial instruments with thermostatic control of the capillary. Indeed, mobility shift measurements are widely used in CE for reliable determination of fundamental thermodynamic parameters in free solution, including pK_a of weakly ionic drugs,[245] dissociation constants for affinity interactions[246] and free energy changes for protein unfolding.[247] Alternatively, an internal standard can be used to determine relative migration times,[248] a migration time normalization can be performed using two or more internal standards[249] or dynamically coated capillaries can be used to stabilize and/or suppress the EOF with good inter-day precision (*CV* < 2%) for migration times when analyzing diluted human urine samples.[250] Recently, a temperature-corrected mobility scale was introduced for data transformation to improve the long-term reproducibility of CE-MS that accounts for multiple sources of migration time variance, including EOF, Joule heating, thermostatic control and suction effects by the nebulizer gas flow for spray formation.[251]

In general, CE-MS is ideal for the analysis of weakly/strongly ionic metabolites that constitute the major fraction (> 60%) of known metabolites or degradation products in primary metabolism,[252] as well as secondary metabolites and their stereoisomers that are weakly retained or have poor separation efficiency in RP-LC or HILIC.[253] In comparison to GC-MS, metabolite profiling of cell extracts can be performed by CE-MS with minimal sample pretreatment that increases sample throughput and overall reliability.[254] **Figure 1.5** highlights the selectivity of CE-MS for resolving a wide range of polar metabolites and their isomers under a single elution condition, including polar amino acids and long-chain acylcarnitines that represent clinically relevant biomarkers for early detection of in-born errors of metabolism.[255] In this case, the addition of acetonitrile as organic modifier in the BGE is needed to ensure adequate solubilization of surface-active acylcarnitines without compromising resolution of amino acids, such as three Leu stereoisomers. However, caution is warranted when adopting CE-MS protocols based on coated capillaries under reversed polarity, where the sprayer serves as the anode to reverse the EOF.[256] A recent cross-platform comparison of GC-MS, LC-MS (*i.e.*, ion pair-RP/HILIC) and CE-MS for analysis of polar metabolites of central energy metabolism derived from *E. coli* extracts concluded that CE-MS lacked overall robustness.[257] Regrettably, this result was impacted by a flawed CE-MS method for anions that was subsequently discovered to cause premature column clogging at the capillary

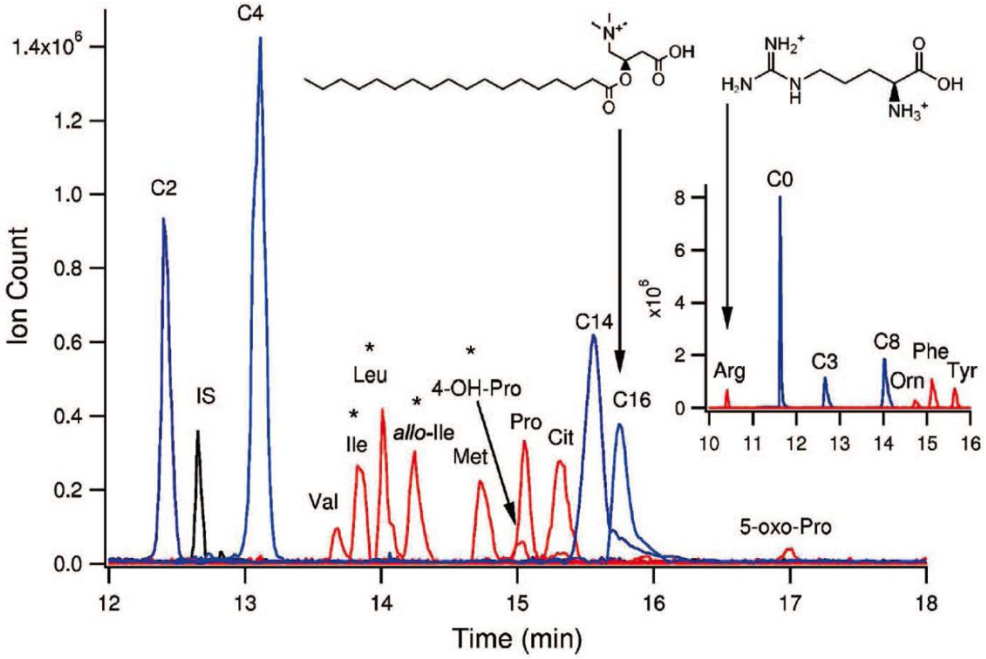


Figure 1.5. Extracted ion electropherogram depicting the separation of 20 different amino acids and acylcarnitines from dried blood spot extracts that serve as diagnostic markers for in-born errors of metabolism in newborn screening programs. Note that metabolites of widely different polarity can be analyzed simultaneously by CE-ESI-MS under a single elution condition while providing resolution of relevant isomeric and isobaric ions (*), including cationic amino acids (Lys) and long-chain acylcarnitines (C16). Conditions used were 1 M formic acid, 15% v acetonitrile, pH 1.8 as BGE with a bare fused-silica capillary and a coaxial sheath liquid interface under positive ion mode.[255]

outlet[256] since its first introduction in 2002.[74] Thus, similar to several major GC-MS,[258] LC-MS,[259] and NMR[260] initiatives, standardized CE-MS protocols require validation across multiple laboratories in order to properly evaluate method robustness. To date, metabolic profiling of cations by CE-ESI-MS under strongly acidic conditions with positive ion mode[261] is a reliable format consistently adopted by several research groups, which allows for quantitative yet reproducible analysis of cationic metabolites in various biological samples.[262] However, further development is needed in CE-MS for reliable yet sensitive analysis of anionic metabolites under alkaline conditions that offers excellent selectivity for polyprotic acids/strongly ionic metabolites that are not well served by LC separation modes based on HILIC or ANP.[253]

Figure 1.6 demonstrates excellent selectivity for a wide variety of anionic metabolites with negative ion mode when using a platinum-based electrode

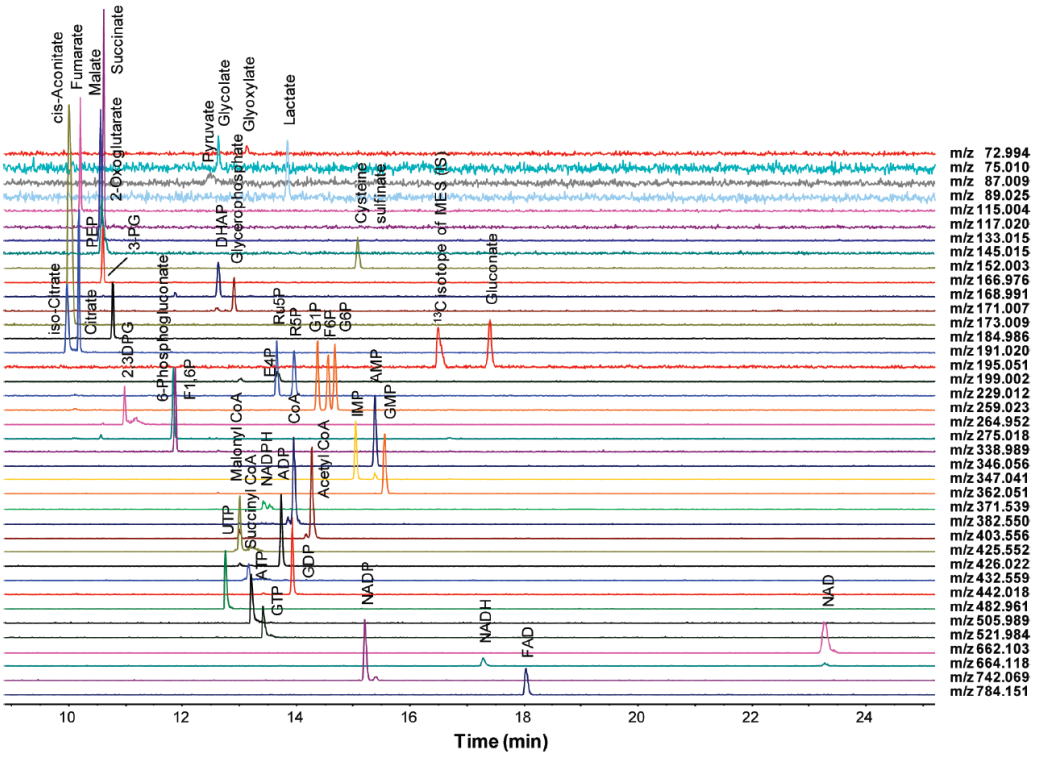


Figure 1.6. Anionic metabolite analysis by CE-MS under negative ion mode using a platinum needle with a COSMO(+) cationic polymer-coated capillary using 50 mM ammonium acetate, pH 8.5. Selected extracted electropherograms for a standard mixture of anionic metabolites (10 µM) relevant to glycolysis, pentose phosphate and the TCA pathways in rat liver extracts, including sugar phosphates, organic acids, nucleotides and redox cofactors. The use of a platinum electrode instead of a stainless steel sprayer improves sensitivity, migration time precision and overall robustness while avoiding capillary outlet clogging at anode under reversed polarity conditions.[256]

sprayer, including nucleotides, organic acids, sugar phosphates and redox-active cofactors[256] In this case, a platinum sprayer was needed to extend capillary lifetime due to incidental corrosion of conventional stainless steel coaxial sprayers when operating under reversed polarity conditions using a polymer coated capillary to reverse the EOF.[263] Recent developments in low flow sprayer designs for CE-MS with improved sensitivity[264–266] require further evaluation under negative ion mode,[263,267–269] which is prone to corona discharge with higher background noise and poor spray stability.[270] Despite the advent of two-dimensional gel electrophoresis as the preeminent preparative tool in biochemistry and proteomics,[49,271,272] there have been few precedents of on-line two-dimensional CE x CE[273,274] or LC x CE[275] systems for metabolomics without off-line fractionation prior to ESI-MS.[276]

1.2.4 New Advances in Separation Science for Metabolomics

Separation science plays several critical roles in metabolomics for improving the performance of MS and NMR-based detection methods needed for reliable quantitative and improved identification of metabolites in complex sample mixtures.[277] Separation efficiency (plate height, *H* or plate number, *N*) and peak capacity are two important parameters for determining the performance of a separation in terms of the total number of unique compounds that are baseline resolved prior to detection. Higher peak capacity within a wider separation window minimizes co-elution issues that can complicate relative quantification and/or spectral matching for unknown identification. For instance, the use of shallow gradient elution programs with ultra-high resolution RP-UHPLC separations can expand peak capacity up to 1000 at the cost of longer analysis times that exceed one hour when including column re-equilibration.[187] Alternatively, the application of orthogonal RP and HILIC separation modes in parallel using optimized solvent elution programs (< 15 min) enables a wider range of compounds that can be detected by ESI-MS for global metabolic profiling.[278] Moreover, optimized 2D GC x GC-TOF-MS using a polar and apolar column connected in series allows for a major boost in peak capacity to over 9,000 resolved features within a short total analysis time.[279]

Figure 1.7 highlights that the use of a wider bore/thicker non-polar stationary phase film as the second dimension column enhances separation performance by increasing sample loadability that is important when quantifying low abundance metabolites from major components in complex biological samples. However, due to the larger retention time variability notably for HILIC, CE and 2D GC x GC separations,[280] additional time resources are needed for data alignment, peak picking and normalization that complicates data pre-processing and statistical analyses in metabolomics.[281,282] Therefore, there is a fundamental trade-off between comprehensive metabolite profiling when using hyphenated separation platforms for deeper metabolome coverage and improved data quality in comparison to high throughput screening approaches based on direct infusion (DI)-ESI-MS,[283] desorption electrospray ionization (DESI)-MS,[284] or ^{1}H-NMR.[285] For instance, a comparison of the performance of DI-FT-ICR-MS relative to LC-ESI-TOF/MS for the discovery of putative serum markers of kidney cancer revealed that twice the number of significant molecular features were detected by LC-MS, however this required about a twenty-fold increase in total analysis time.[286] Alternatively, a two-stage strategy can be adapted based on rapid primary screening by DI-ESI-MS followed by confirmatory testing by hyphenated

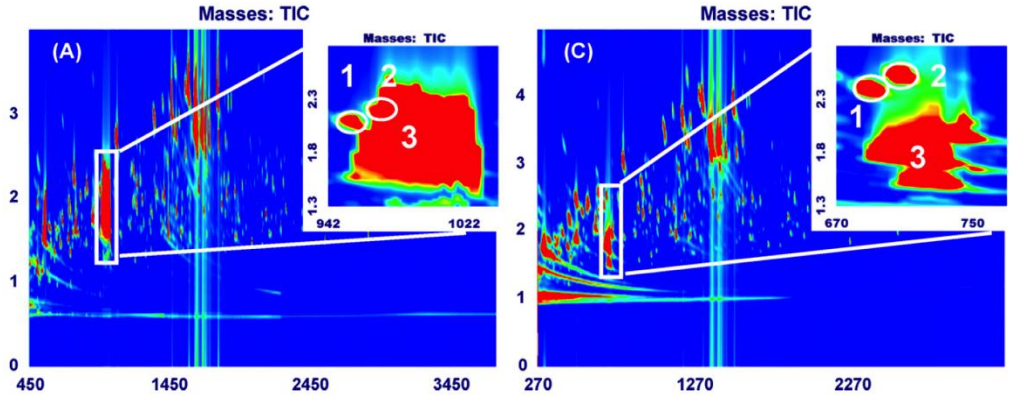


Figure 1.7. Improved separation performance in 2D GC x GC-TOF/MS when using a polar/non-polar column configuration to enhance peak capacity (> 9,000) with greater loadability and separation efficiency, where a wider bore and thicker wall coated non-polar stationary phase is used as the column in the second dimension. Full scan chromatograms of methanol extracted fetal bovine serum samples with conventional setup A (2D: 1m x 0.1 mm, 0.1 µm, BPX5 column) and high loadability setup C (2D: 2m x 0.32 mm, 0.25 µm, BPX5 column) where the first column is polar consisting of 30 m x 0.25 mm, 0.25 µm, BPX50. Enlargements (100 s x 1.5 s) highlight better resolution of trimethylsilylated amino acids in the presence of excess urea from plasma when using the loadable column configuration, namely serine (1), threonine (2) and urea (3).[279]

separation techniques that possess greater selectivity for confirmatory testing similar to expanded newborn screening initiatives.[255] However, this strategy has yet to be adapted in metabolomics research as it risks false-positives due to a plethora of artifact signals when using DI-ESI-MS as the primary screen, as well as false negatives due to matrix-induced ion suppression effects.[287] For this reason, a compromise between sample throughput, metabolome coverage and peak capacity/resolution of a separation must be reached without sacrificing data quality when performing large-scale metabolomics investigations for biomarker discovery.[278]

One major contribution of separations is the ability to resolve isobaric compounds since they are not readily distinguished by high resolution MS or MS/MS, including diastereoisomers[253,288] and enantiomers[289,290] For instance, stereoselective separations are critical for second-tier testing of genetic diseases with improved sensitivity and selectivity relative to high-throughput MS/MS screening, such as plasma *allo*-Ile that is a pathognomonic marker of maple syrup urine disease.[291] Differential metabolic profiling of cellular responses to racemic drugs or single enantiomers also highlights the importance of drug stereochemistry on biological activity that can be resolved by appropriate separation techniques.[292] In general, chiral resolution is realized by pre-column

chemical derivatization to form diastereomeric adducts that can be resolved by differences in their retention factor and/or mobility or the use of chiral-based stationary phases or chiral additives in the BGE/mobile phase.[293] Indeed, there are few reported studies to date that have performed comprehensive metabolite profiling that is also enantioselective for resolution of optically-active classes of metabolites. For example, chiral separations contributed to improved discrimination of liver cirrhosis patients[290] and bioequivalence of transgenic crops[294] when measuring the enantiomeric excess ratio of *D*-amino acids. Another important function of separations notably when using ESI-MS is improved detection of low abundance metabolites that is compromised by co-elution of abundant co-ions and involatile salts in a biological sample which lowers sensitivity.[206] This also impacts the quantitative performance of MS-based metabolomics when measuring relative signal responses with acceptable accuracy that is independent of sample matrix. Given the wide disparity in solute ionization efficiency in ESI-MS, reliable quantification of metabolites in complex biological samples is only feasible if separations are optimized to avoid co-elution, which is critical in metabolomics since purified reference standards and/or stable-isotope internal standards are often unavailable.[262] For instance, post-column infusions experiments of a calibrator ion can be used in LC-MS to monitor potential ion suppression or enhancement effects during elution,[295] where ESI is generally found to be more susceptible to signal suppression relative to atmospheric pressure chemical ionization (APCI) or APPI even when using UHPLC methods.[296,297] One way to improve method robustness and sensitivity in metabolomics is to integrate an effective sample enrichment and cleanup protocol together with separation prior to ionization, such as SPE-LC-MS.[298] For instance, an optimized mixed-mode solid-phase microextraction (SPME) procedure can be used to enhance separation performance for analysis of both hydrophobic and hydrophilic metabolites with reduced ionization suppression in LC-MS.[299] However, since extraction is not exhaustive and depends on sorbent material while reflecting the free circulating levels metabolites in biological fluids, the number of compounds detectable by LC-MS is reduced when compared to solvent-based extraction methods that denature protein during sample preprocessing to remove protein-bound metabolites. Also, in most cases SPE/SPME extraction steps in metabolomics are performed off-line prior to analysis and require long extraction or desorption times that prevents absolute quantification since recovery is highly solute dependent.[300] Other sample clean-up methods offer better extraction efficiency, but still are limited by complicated sample handling that contributes to higher process variability, including single-

drop microextraction[301] and stir bar exhaustive extraction.[302] On-line sample preconcentration together with desalting can be integrated during separation prior to ESI-MS when using a discontinuous electrolyte system in CE-MS as highlighted in **Figure 1.8**. Overall, about a 50-fold enhancement in concentration sensitivity can be realized in CE-MS for cationic metabolites (*e.g.,* Trp) under positive ion mode without ionization suppression or extra-column band broadening.[303] Although ion mobility-ESI-MS provides additional selectivity to resolve isomeric metabolites prior to mass analysis due to the characteristic mobility drift time of an ion, it does not offer preconcentration and/or desalting benefits similar to CE since electrophoretic separations occur in the gas-phase after ionization.[304] In addition to resolving isobars/isomers, reducing matrix interferences that modulate signal response, boosting concentration sensitivity, as well as improving quantitative reliability, high resolution separations also improve mass spectral matching to reference libraries while also providing qualitative information for supporting unknown metabolite identification that is complementary to MS.

In most cases, the majority of detectable signals in MS-based metabolomics studies are unknown, where a minimum of two independent parameters are necessary to describe a molecular feature, namely its accurate mass (*m/z*) and (corrected) retention time, migration time or retention index.[305] Although ultra-high resolution MS instruments (*e.g.,* FT-ICR-MS, Orbitrap, etc.) may correctly assign a putative empirical formula derived from a molecular feature, this is subject to a high false-positive rate when adequate confidence intervals are not properly used to ensure annotation quality.[306–308] Moreover, the confidence of annotation decreases with lower signal/noise ratio (*e.g., S/N* < 20),[309] such that low abundance ions typically remain uncharacterized due to their poor mass accuracy (> 10 ppm) and/or undetectable isotope abundance patterns that are needed for discrimination among numerous molecular formulae.[69,71,309] Even in the case of a defined elemental formula, such as phenylalanine ($C_9H_{11}NO_2$), there are over 6 million putative chemical structures as determined by Molgen Online[184] with 1,974 known structures with the same accurate mass archived in PubChem (search query performed on 11/2012). However, a greater chance for correct spectral assignment is achieved after pre-selection of "authentic" molecular features with robust data filtering[310] or subtracting a process blank to all samples in order to remove confounding background ions.[184] As GC-EI-MS provides reproducible mass spectra that are largely independent of matrix-effects and vendor-specific MS instruments, this allows for the

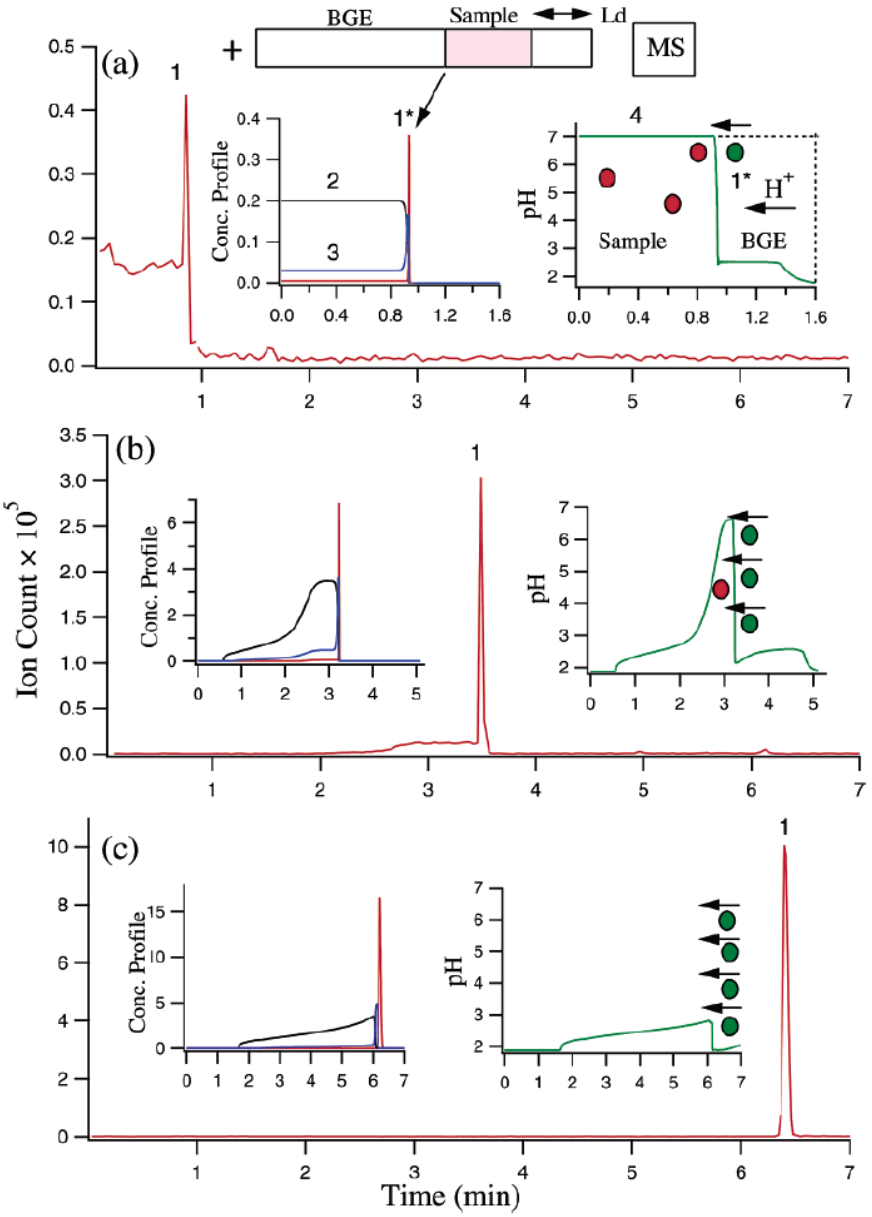


Figure 1.8. Time-resolved electropherograms and computer simulations (as insets) highlighting the dynamics of electrokinetic focusing of a long/dilute sample plug of Trp in CE, where the capillary functions as a preconcentrator, desalter and separator prior to ESI-MS. A long sample plug (12.5 cm) of dilute Trp prepared in 200 mM ammonium acetate, pH 7.0 with 15 mM NaCl is placed at increasing distances from the capillary outlet (a) 7.5 cm, (b) 25 cm and (c) 40 cm after it is injected within a discontinuous electrolyte system that is filled with 1 M formic acid, pH 1.8 as the BGE. About a 50-fold enhancement in concentration sensitivity is realized without signal suppression. Analyte peak numbering corresponds to concentration profiles of (1) Trp (experimental), (1*) Trp (predicted), (2) NH_4^+, (3) Na^+ and (4) pH.[303]

development of universal database libraries for mass spectral searches, such as NIST11.[311] In most cases, customized "in-house" mass spectral libraries are required for improving the confidence of peak assignments in LC-MS metabolomic studies due to the greater variability in ion fragmentation processes when using ESI-MS/MS.[184,312,313] Ultimately, structural elucidation from a molecular formula requires additional experimentation derived from collisional-induced dissociation (MS/MS), multi-stage MS (MS^n) or EI-MS spectra that are consistent with reference spectral libraries acquired under standardized conditions; however, the quality of mass spectral matching is highly dependent on the specific database that can result in different search query results (*e.g.*, PubChem, HMDB, Massbank, METLIN etc.).[18,314,315] One of the main functions of separations is the resolution of multi-component mixtures prior to ionization in order facilitate mass spectral matching with pure compound reference libraries. Indeed, it was long recognized that the inclusion of both mass spectral and retention time index criteria within customized GC-MS databases allows for greater confidence in identifying biologically significant molecular features.[316] However, without reference spectra or purified chemical standards, de novo structural elucidation remains tenuous despite the use of computer-aided structural elucidation strategies, such as *in silico* mass fragmentation prediction.[317–321]

The advent of computer-assisted design tools for separation optimization in GC/LC[199,322] and CE[323,324] offers an orthogonal approach for metabolite identification based on accurate prediction of retention or migration times *in silico*. CE is particularly amenable for predictive modeling of polar metabolites based on intrinsic physicochemical parameters of an ion[303] since separations are performed under defined isocratic conditions using a homogenous aqueous buffer in a fused-silica capillary. Indeed, fundamental electrokinetic, thermodynamic and molecular properties of an ion not only permit modeling of solute migration behavior in CE, but can also be used for virtual quantification (*i.e.*, relative response factor) of polar metabolites by ESI-MS given their wide disparity in ionization efficiency.[262] For instance, Sugimoto *et al.*[249] demonstrated that molecular descriptors can accurately predict migration times of cationic metabolites in CE-TOF/MS when using support vector regression, which increases the assigned molecular features for urinary metabolites from 39% (*i.e.*, standard mass spectral library match alone) to over 52% when using a *m/z* and migration time tolerance of 15 ppm and 1.5 min, respectively. Similar approaches have been reported for prediction of retention times when using RP-LC-MS,[325,326] HILIC-MS,[327] and GC-MS[71] based on *in silico* molecular descriptors to improve metabolite identification, which can reduce false negative assignments by about

40% as compared to accurate mass alone, including isomer discrimination.[327] However, model robustness is dependent on the selection of an appropriate training/validation set that encompass a diverse classes of metabolites while implementing measures to minimize major drifts in retention times and non-ideal interactions that contribute to bias exceeding 35% in HILIC.[327] Another major challenge in predictive modeling of retention times in chromatography is the non-linear change in isocratic retention factor that is solute dependent and not fully corrected by retention indexing or time warping algorithms for retention time normalization, which assume similar retention behavior of neighboring eluting internal standards.[328] This premise fundamentally prevents accurate modeling of analyte retention behavior in LC-MS due to the wide range of solvent compositions and/or gradient elution programs used in separations for metabolomics.

Recently, Boswell *et al.*[328,329] introduced a novel retention projection method for compound identification in LC-MS based on their measured isocratic retention versus eluent composition relationships that can tolerate different solvent gradients, flow rates and instrument platforms. The inclusion of predicted retention times with good accuracy (< 1%) substantially improves compound identification since retention times in RP-LC are highly orthogonal to accurate mass[328] as depicted in **Figure 1.9** for over 7,307 model compounds derived from the KEGG database. Interestingly, not only does retention time data with accurate mass double the total number of compounds identified by FT-ICR-MS, but it also enhances the performance of nominal mass resolution quadrupole mass filters, such that a larger fraction of metabolites can be identified relative to accurate mass alone provided that retention time accuracy is within 0.25%.[328,329] A metabolomics workflow for identification of unknown metabolites was recently introduced for GC-EI-TOF-MS based on unbiased query searches of known chemical structures deposited in the PubChem database,[71] where in silico retention index and mass spectra matching allowed for correct structure assignment in 73% of cases among putative compounds within the top 5 hits. Moreover, the integration of both Kovats retention index estimation with EI-MS fragmentation prediction software filtered out 89% of all potential isomeric structures not feasible by elemental formula alone. However, retention index predictions in GC using NIST MS software[330] is not sufficiently robust for prediction of polar metabolites labeled with multiple trimethylsilyl groups since steric effects are not included in retention modeling,[331] which limits its applicability for de novo structural elucidation of unknown molecular features.[71] Similarly, a dual *in silico* retention index and mass spectra model was reported for

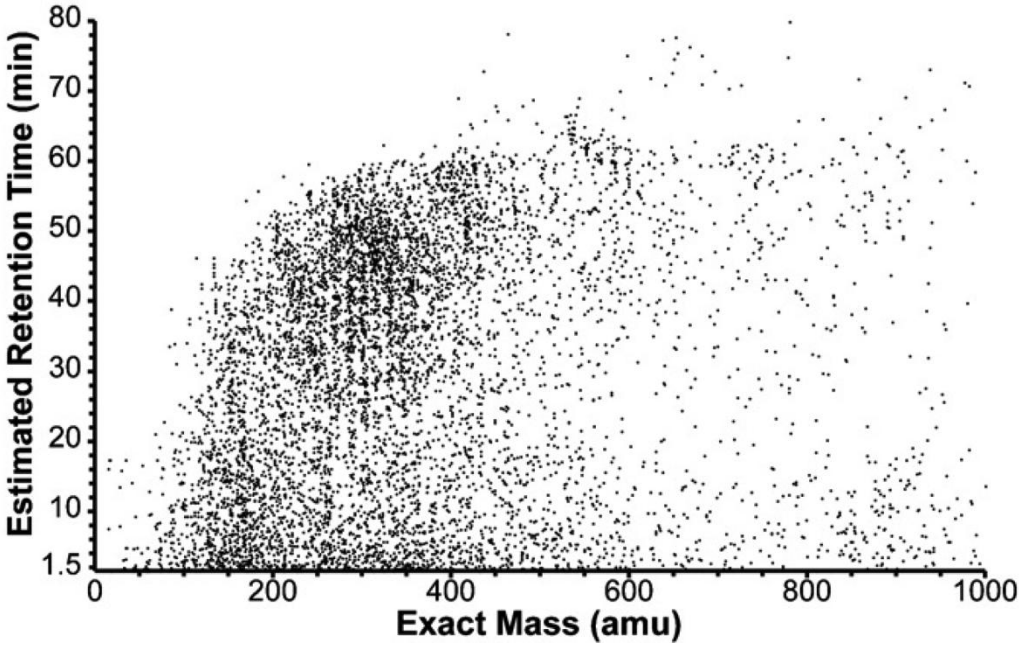


Figure 1.9. Predicted RP-LC retention time and exact mass of 7307 compounds from the KEGG metabolite database, which highlights their orthogonal relationship for metabolite identification. Simulated compound identification revealed that a two-fold increase in correct structural assignment (70% of total compounds, $P = 0.01$) is realized when accurate mass (1 ppm, 1×10^6 mass resolution) is supplemented with accurate retention time (0.25% error). However, a major improvement in compound identification was also found when combining nominal mass resolution with predicted retention times from only 1% to about 44% of total compounds, which is greater than FT-ICR-MS alone.[328]

LC-MS to rank lead candidates from PubChem search queries for subsequent screening using multiple filters, including accurate mass, retention index, precursor ion stability and CID fragmentation spectral matching.[332,333] New computational approaches for structural elucidation of unknown metabolites are urgently needed for ranking lead candidates based on a consensus scoring function since a majority of annotated features remain largely unidentified in metabolomics.[334]

1.2.5 The Volatile Metabolome by GC-MS

GC-MS represents an established separation platform in metabolomics that is the method of choice for analysis of volatile/non-polar metabolites; however chemical derivatization is required for polar metabolites to improve their volatility, thermal stability, detectability and/or retention while minimizing deleterious column

adsorption. Indeed, metabolome coverage by GC-MS is limited by the intrinsic thermal stability of the stationary phase and metabolite as their trimethylsilyl derivative[335] that also excludes high molecular weight compounds with low vapor pressure unless using high temperature GC-based columns.[336] Plant sciences still dominate the use of GC-MS as a central platform in metabolomics since its first introduction in 2000[59] with growing interest in food science, microbiology and biomedical applications.[334] A survey of GC-MS and 2D GC x GC-MS metabolomic studies recently published indicates that overall methodologies are fairly standardized and differ primarily in terms of sample pretreatment protocol as related to sample extraction and chemical derivatization for specimens ranging from biofluids,[337,338] cells,[339,340] tissues,[341,342] and whole organisms.[343,344] In general, most methods use narrow bore fused-silica capillaries with wall-coated 5%(diphenyl)polydimethylsiloxane (PDMS) as a generic non-polar stationary phase and He as an inert gas with elution gradient programs performed with a maximum temperature under 320°C within 50 min to reduce column bleed; however, some reports also utilize more polar stationary phases for improved selectivity in GC-MS and/or higher peak capacity in 2D GC x GC-MS, such as 50%(diphenyl)-PDMS[345] and 14%(cyanopropylphenyl)-PDMS[346] bonded phase columns. If sensitivity is an issue, then splitless injection of the sample is performed, but it can result in column overloading, poor separation efficiency and reduced column lifespan when exposed to excess residual silylating reagents. For these reasons, splitless injection is often used in GC at a high temperature to induce flash evaporation, whereas the initial temperature of oven is set low to re-condense and preconcentrate the sample zone at the column head; however, this risks in-situ thermal degradation of labile metabolites in the injector, notably partially silylated derivatives. A similar type of analyte focusing mechanism is applied in 2D GC x GC to reduce extra-column band broadening and enhance sensitivity during thermal modulation of eluting bands at the end of the first column with their subsequent transfer into a much shorter second column having complementary selectivity.[170] Despite the advent of commercial 2D GC x GC-MS instrumentation, time-consuming data processing and peak picking procedures still hamper large-scale metabolomics studies.[347] Clean injection port liners are critical in GC-MS due to surface adsorption and carry-over contamination by sample matrix and/or excess reagents, which can also potentially catalyze chemical reactions at elevated temperatures leading to sampling bias and unanticipated artifacts.[348] In most cases, EI-MS (70 eV) with positive ion mode is used as a universal ion source preferably with TOF-MS, which allows for mass spectral matching with known compounds within customized libraries or public

databases. Heuristic filtering of accurate mass measurements (< 3 ppm) with good isotope precision (< 5%) of a molecular ion (*e.g.*, radical cation) in EI-MS allows for correct elemental formula assignment in most cases,[306] whereas a tentative chemical structure can be assigned based on its characteristic fragmentation pattern and retention index with subsequent verification using a reference standard.[313] However, extensive fragmentation hampers reliable elemental composition in EI-MS unless complemented by chemical ionization (CI) that is needed for generating more intense molecular ions and isotope signatures.[349] In addition, other more sensitive ionization modes are also available in GC-MS but rarely used in metabolomic studies to date, including negative ion chemical ionization (NICI)[350] or electron capture negative ion (ECNI)[351] detection that offer lower detection limits for metabolites as their perfluorobenzyl derivatives.

Trimethylsilylation remains the most widely used labeling procedure in GC-MS-based metabolomics given its general applicability to various classes of polar metabolites.[352] After sample extraction, a two-step derivatization reaction is performed involving methoximation in pyridine to prevent reversible cyclization of reducing sugars followed by silylation using bis(trimethylsilyl)trifluoroacetamide (BSTFA) or preferably *N*-methyl(trimethylsilyl)-trifluoroacetamide (MSTFA) catalyzed in the presence of 1% trimethylchlorosilane.[90] Although chemical derivatization is typically performed off-line using an autosampler vial, in-liner derivatization of volume-restricted biological samples or single cells can also be achieved by GC-MS using a direct thermal desorption interface inside the injector port.[353] However, the optimum derivatization temperature of the injector port must be kept sufficiently low (70°C) in order to prevent degradation of labile metabolites that restricts thermal focusing of sample at the column head. Despite its broad reactivity with various classes of functional groups with good analytical performance for "class-1" metabolites (*e.g.*, sugars, fatty acids, organic acids), silylation is prone to numerous artifacts that leads to poor precision and lower recoveries in the case of "class-2" (*e.g.*, amines, phosphates) and notably for "class-3" (*e.g.*, amide, thiols) compounds.[354] Moreover, silylation is associated with multiple and/or incomplete derivatives for polar/ionic metabolites that vary significantly in their reactivity and stability, which complicates absolute quantification with reduced sensitivity.[355] Excess silyl reagent and high temperatures are required to enhance reaction kinetics under anhydrous conditions that require stringent quality control measures during all stages of sample handling, including stable-isotope internal standards, blank controls and frequent injector liner replacements.[90,345] Sample clean-up protocols to remove residual silylating reagents after derivatization using

silica is ineffective due to co-extraction of most derivatized metabolites from sample.[356] Alternatively, bulkier reagents can be used to impart greater stability, such as *tert*-butyl(dimethylsilyl)trifluoroacetamide (MTBSTFA) derivatives that are 10^4 times more hydrolytically stable than trimethylsilyl protecting groups.[357] Reaction conditions and GC-MS protocols can be systemically optimized when using a design of experiments to elucidate major factors impacting derivatization efficiency for rigorous method validation.[358] In addition, stable isotope labeling using MTBSTFA together with its heavier deuterated (d_6) analog offers a novel approach for differential metabolomics that corrects for bias and the high variability associated with silylation.[359] In this case, relative quantification of metabolites in human serum extracts can improve precision and accuracy based on ratiometric measurements of each encoded metabolite isotopomer pair. Similarly, stable isotope labeling using a mixture of light and heavy (d_9) isotopomers of MSTFA improves relative quantification of polar metabolites in human urine and plasma extracts by GC-MS.[360] However, multiple silylated adducts with polyprotic metabolites along with resolution of deuterated isotopomers by GC complicates data processing, although it does not impact quantitative performance since EI-MS is not subject to sample matrix effects unlike ESI-MS that requires co-elution of isotopomers.[359] Due to the long reaction times and elevated temperatures that are required for oximation/silylation, microwave-assisted derivatization protocols allow for rapid silylation with total reaction times under 5 min.[361–363] Despite the continued use of classical trimethylsilylation in GC-MS-based metabolomics in large-scale inter-laboratory studies[258,345] on-going controversies raise concerns regarding its suitability for metabolite profiling due to its poor analytical performance for amino acids in comparison to LC-fluorescence.[364] Indeed, recent studies have demonstrated the advantages of pre-column derivatization based on alkylation using methylchloroformate (MCF)[346,365] or ethylchloroformate[366,367] that converts amino and non-amino organic acids into volatile carbamate and ester derivatives, respectively. This chemical derivatization occurs rapidly in biological samples under aqueous alkaline conditions without heating, prior extraction or stringent quality controls as needed for silylation. Similar to esterification derivatization of plasma fatty acids,[368] alkylation reactions are performed in aqueous solution with high alcohol content for improved solubility and stabilization of long-chain carboxylic acids, where ethyl esters of short-chain organic acids exhibit better retention in GC.[367] For instance, microwave-assisted oximation together with MCF results in formation of stable derivatives that are less susceptible to sample matrix effects allowing for reliable GC-MS analyses while using a single internal

standard without quartz liner replacement over several months.[369] Also, highly fluorinated alkyl chloroformates with encoded chlorine tags represent a novel class of label that greatly enhances sensitivity when using GC-MS with ECNI detection.[370,371] However, chloroformate reagents are more susceptible to fragmentation in EI-MS or ECNI due to their weak carbamate bond that may lack a detectable radical cation without CI mass spectra validation.[366] Thus, the development of improved extraction and derivatization protocols are still needed for GC-MS in order to avoid the large sources of bias and process variability associated with classical trimethylsilylation[372] that ultimately limits the discrimination power of metabolomics for molecular phenotyping biological samples.[369]

Head-space (HS) sampling of the volatile metabolome represents one of the most unique features of GC-MS that avoids many of the technical hurdles of sample processing of polar metabolites in biological samples. SPME is frequently used for direct sampling of the volatile fraction of organic compounds distributed above a solid or liquid specimen using a fused-silica fibre coated with a specific sorbent that integrates sampling, extraction, concentration and sample desorption in a solvent-free system.[373–375] In addition, in-fiber chemical derivatization can also be integrated together with HS-SPME for single-step sample workup prior to GC-MS.[376] Sampling bias is inherent with SPME in metabolomics due to differences in sorption kinetics and metabolite volatility that is also matrix dependent; however relative quantification with good precision is feasible when performed under standardized conditions. Due to the limited loading capacity of SPME, improved sample enrichment and lower detection limits can be acheived using headspace stir bar sorptive extraction methods.[377,378] Given the importance of odour and fragrant compounds in food quality, agriculture and plant sciences, high throughput screening of volatiles from tomatoes was demonstrated by HS-SPME-GC-MS[379] that detected 322 distinct plant-derived compounds after subtraction of SPME fiber blanks. Several major classes of volatile organics were identified, including phenolics, lipid derivatives, terpenoids and carotenoids that allows for the discrimination of ripe fruits among 94 different tomato genotypes using unsupervised multivariate data analysis.[379] Similarly, untargeted metabolite profiling of volatile organics from four apple varieties was investigated by GC-TOF-MS, where an experimental design was used to optimize HS-SPME conditions to maximize total signal responses, including pre-heating time, extraction time and temperature.[380] A comparative analysis of apple extracts after freeze drying was found to reduce levels of volatiles with greater variability relative to freshly processed samples highlighting the impact of sample collection

and storage. In this case, differentiation of the four apples varieties was achieved by multivariate data analysis based on 10 key volatile alcohols and ester derivatives, whereas residual levels of the pre-/post-harvest scald inhibitor, *N*-phenylaniline was measured in a specific apple variety prone to oxidation/discoloration during storage.[380]

There is growing interest in using HS-SPME-GC-MS for comprehensive analysis of volatile organics in biological samples as applied to medical diagnostics in order to assess the metabolic condition and health status of an individual[381] as first reported by Pauling's pioneering work in the early 1970s.[119] Recently, differential metabolite profiling of volatile metabolites from skin was developed for early detection and diagnosis of melanoma relative to normal skin and benign naevus (mole) samples from a 3-mm punch biopsy, as well as air blank samples.[82] Three volatile hydrocarbons were significantly elevated in both fresh and frozen melanoma specimens as secondary metabolites of membrane lipid peroxidation/oxidation stress, namely 4-methyldecane, dodecane and undecane, whereas pyridine and 3-hexanol were expressed from naevi samples. Since this is a non-invasive method that preserves skin morphology, the same specimens used for volatile collection by HS-SPME were subsequently analyzed by histopathological staining within 3 hrs of the original biopsy collection.[82] The discovery of putative markers for non-invasive diagnostics was recently reported by Silva *et al.*,[382] which detected up to 82 different metabolites in the headspace of urine samples for three different cancer types (*i.e.* leukemia, colorectal, lymphoma). A comparison of different sorbent materials and experimental conditions used in HS-SPME was performed to maximize extraction efficiency. **Figure 1.10** highlights differential volatile urinary metabolite signatures measured in three representative urine samples from cancer patients relative to healthy controls by GC-MS. Optimal extraction was performed using a carboxen-polydimethylsiloxane sorbent in HS-SPME, where benzene derivatives, terpenoids and phenols were significantly elevated in the oncological group, including lower levels of dimethyl disulfide;[382] however sample sizes were small and not adequately age or gender matched for this pilot study since volatile compounds from urine have been demonstrated to be gender-specific and individual dependent.[383] Moreover, volatile organics derived from biological samples are known to be dependent on both genetic and environmental factors, such as diet, lifestyle and microbiota that contribute to significant inter-individual variability. Indeed, greater volatile metabolome complexity has been reported in

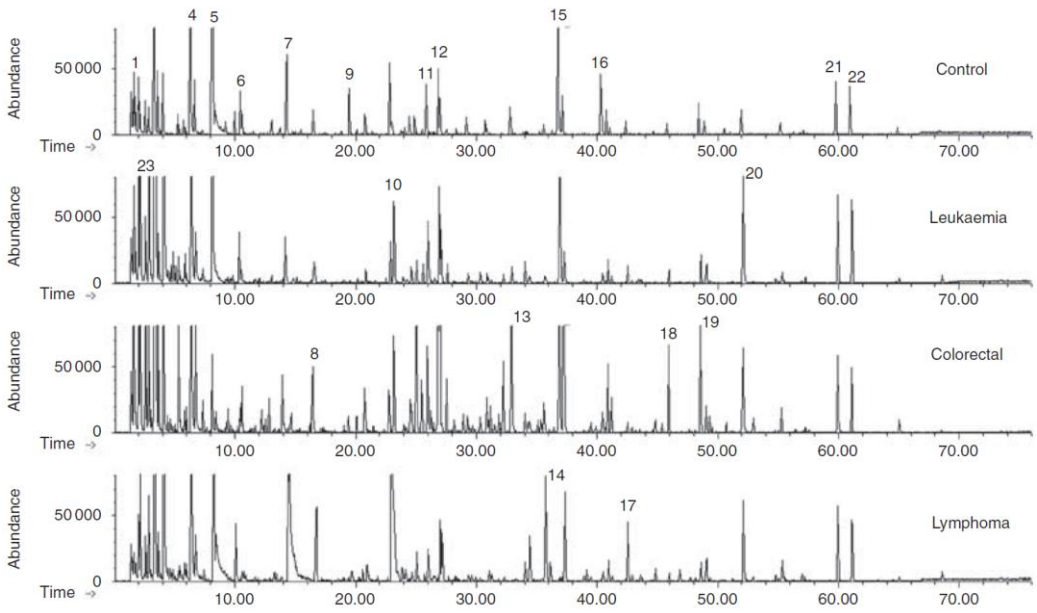


Figure 1.10. Total ion chromatograms by GC-MS comparing metabolomic profiles of volatile urinary metabolites from cancer patients (leukaemia, colorectal and lymphoma) contrasted with a healthy volunteer (control group) after sampling by HS-SPME for 60 min at 50°C for pH-adjusted urine samples with 17% w/v sodium chloride under agitation. A significant increase in the peak area of 2-methyl-3-phenyl-2-propenal, *p*-cymene, anisole, 4-methyl-phenol and 1,2-dihydro-1,1,6-trimethyl-naphthalene along with lower levels of dimethyl disulfide were detected in cancer patients.[382]

human sweat in comparison to urine or saliva.[377] Thus, standardized conditions are critical for ensuring reproducible volatile extraction using HS-SPME without background contamination, including high purity chemical reagents due to the high amounts of salt added to sample for increasing metabolite volatility.[383]Due to difficulty in collection of breath samples for routine clinical diagnostics,[384] metabolomic studies of volatile metabolites from urine were also conducted in a mouse model of lung cancer.[385] In order to validate that volatile odourants represent authentic markers of lung cancer, sensor mice were trained to discriminate among tumour bearing and non-tumour control mice based on their behavioral response similar to canine scent detection approaches reported for the diagnosis of lung and breast cancer.[386] In order to better correlate perceived odor quality/intensity in relation to its chemical composition, GC-MS together with on-line olfactometry[387,388] offers a unique approach to understand food quality traits associated with human sensory responses (*i.e.,* sensometabolome)[389] using a trained panel of subjects who function as detectors for odour-active components.

The chemical-sensory characterization of offensive or unpalatable odours can also be explored for monitoring livestock environments that impact air quality on local and regional scales, including human/animal health within confined feeding operations.[390] For instance, untargeted metabolite profiling of dairy manure odour was performed by HS-SPME-GC-MS with simultaneous olfactometry (O),[391] which revealed over 86 different types of volatile organics, including 17 compounds only detected by smell that were present below the detection limit of either FID or EI-MS. Dimethyltrisulfide was identified above its odor detection threshold, which is a volatile pheromone produced by the fly-attracting plant, dead-horse arum florets[392] and a characteristic odorant emitted by fungating cancer wounds.[393] Improved resolution and detection of early eluting volatile organics was achieved by multidimensional GC-MS with heart cutting thereby allowing for a shorter extraction time of 30 min by HS-SPME, which was needed to reduce bias caused by microbial action during sampling.[391] Thus, GC-MS-O allows for improved characterization of complex mixtures of low abundance odor-active compounds, including sulfur-containing compounds, fatty acids and phenol/indole derivatives.[391]

1.2.6 The Non-Volatile Metabolome by LC-MS

LC-MS has undergone the greatest expansion within metabolomics since the first publications in 2002[72,73] due to its wide selectivity for resolving various classes of involatile and/or thermally unstable metabolites. Due to recent advances in column technology, UHPLC now offers analogous separation efficiency as GC based on solvent and/or temperature programming for ultra-fast separations with high peak capacity.[187] For instance, a major fraction of known metabolites within the human serum metabolome (> 70%) comprise non-polar lipid classes (*e.g.*, phospholipids, glycerolipids and steroids) which are well-suited to RP-LC-MS.[17] For this reason, thousands of molecular features are detected in metabolomic studies from methanol extracts of plasma when using RP-LC-MS under both positive and negative ion mode ESI.[394] In addition, greater metabolome coverage can also be acheived by LC-MS when using complementary ion sources that have greater response towards non-polar metabolites, such as APCI.[33,395] Hundreds of different commercial stationary phases exist in LC, where proprietary column packings are characterized based on their hydrophobicity and silanophilic activity at pH 7 using model non-polar and polar solutes;[396] indeed, similar C18 packed columns made by various manufacturers often exhibit different hydrophobicity

properties under analogous elution conditions. Thus, reproducible analytical performance using LC-MS in metabolomics can only be achieved if using batches of columns from the same manufacturer while implementing preventative maintenance protocols (*e.g.,* cleaning, reconditioning or calibration) between blocks of runs when analyzing complex biological matrices.[394] Moreover, quality control samples that are intermittently injected throughout an analysis block is an effective approach for signal drift correction in LC-MS that minimizes intra-laboratory and inter-laboratory variation with CV < 20%.[259,397] Indeed, the temporal stability of both measured signal and retention time in LC-MS are critical for relative quantification in major metabolomics projects due to long-term changes in ion source, column properties or separation conditions. Similar to GC-MS, rigorous optimization of pre-analytical steps prior to LC-MS, including sample collection/storage[398] and sample pretreatment/cleanup,[399] as well as default software parameters used in data alignment and peak picking are required to ensure peak quality as a way to reduce false discovery rate.[400] Also, since the eluent composition and intrinsic solute properties determine ionization efficiency in ESI-MS,[296] careful optimization of separation conditions in LC-MS with sufficient column re-equilibration time after gradient elution is needed when stable isotope internal standards are unavailable.

A survey of major LC-MS configurations in metabolomics indicates that RP-HPLC using standard 3 μm and increasingly UHPLC with 1.7 μm BEH particles typically use C18 bonded-phase columns under elevated temperatures (> 40°C) for improved stability, performance and reproducibility. However, most LC-MS methods differ widely in terms of sample pretreatment and optimum solvent gradient elution program reflecting a compromise between sample throughput and peak capacity requirements. Although C18 remains by far the generic RP stationary phase of choice in UHPLC-MS, improved performance has been reported with C8 columns when analyzing plasma extracts in terms of reduced injection carry-over of strongly retained lipids, thus allowing for shorter wash times between runs without baseline drifts as evaluated by post-blank samples.[401] Unlike most LC-MS studies in metabolomics that use standard narrow bore (2.1 mm i.d.) columns with relatively high flow rates (300-500 μL/min), capillary LC columns (150 μm i.d.) packed with 1.7 μm trifunctional (T3) C18 particles provide better retention of polar metabolites with higher peak capacity and greater sensitivity for plasma lipids when coupled to nanoflow ESI-MS.[402] Alternatively, perfluorinated stationary phases offer extended selectivity for polar metabolites relative to alkyl and phenyl-based RP columns since they exhibit

mixed-mode retention mechanisms with greater solute retention at elevated temperatures and higher organic solvent content.[403] For instance, pentafluorophenylpropyl (PFPP) columns have been shown to provide higher separation efficiency with better resolution for polar metabolites and their stereoisomers, notably organic acids and amino acids relative to amino-based HILIC or C18 RP columns.[215] Superheated water at elevated temperatures (> 100°C) lowers the viscosity of mobile phase when using UHPLC that also improves solvent eluotropic strength without organic modifier resulting in shorter analysis times, lower back pressure and higher selectivity. Indeed, an optimized thermal and solvent gradient program in UHPLC-MS allows for fast yet high resolution separations when analyzing diluted urine samples relative to a conventional isothermal gradient elution at elevated temperature.[404] However, poor separation performance was found when analyzing plasma extracts due to the higher content of lipids that are not adequately soluble in a primarily aqueous mobile phase (0-10% acetonitrile). High temperature UHPLC is also a key strategy for comprehensive 2D LC x LC separations in order to minimize under-sampling of eluting first dimension peaks to the second column with ultrafast elution programs.[405,406] Although 2D LC x LC separations using a switching valve has been developed in proteomics as a way to greatly enhance resolving power,[275] "off-line" fraction collection methods based on LC x LC-MS have only been reported in metabolomics[407] due to technical hurdles associated with elution solvent compatibility, such as strong cation-exchange/HILIC,[408] ion pair RP/RP,[409] and RP/HILIC.[410]

LC-ESI-TOF-MS is widely used in metabolomics studies given its fast duty cycles for sharp eluting peaks in UHPLC, whereas other mass analyzers have been used to reduce isobaric interferences, including ultra-high mass resolution FT-ICR-MS[411] and orbitrap instruments.[412] Since ESI-MS rarely leads to extensive in-source fragmentation, Q-TOF in conjunction with CID experiments is required for structural elucidation of unknown molecular features that may serve as putative markers of disease progression;[413] however, unlike EI-MS, spectral matching by MS/MS from archived database entries or via in silico modeling are non-trivial given complex fragmentation reaction pathways that are dependent on both metabolite structure and collisional energy selected.[321] In most cases, only a small fraction of metabolites in LC-MS (< 20-30%) can be unambiguously identified given limited access to authentic commercial standards.[414] Indeed, various salt adducts and background artifacts are prevalent in ESI-MS that complicates data processing if not adequately filtered based on peaks that satisfy minimum detection thresholds (S/N > 10) within a defined retention

window while comprising reproducible signals measured in pooled samples (CV < 15%).[310] Given the large number of unknown molecular features that are not present in current databases, open access repositories for metabolomics studies and meta-data are required for dissemination, validation and future identification of novel metabolites, such as MetaboLights.[415] However, the qualitative and quantitative performance of LC-MS is highly dependent on optimization of separation conditions that generate high peak capacity. In order to enhance resolution of polar/ionic metabolites that are not well retained by conventional RP-LC, complementary retention mechanisms can be used in parallel for the same sample, notably HILIC[153,416] or ANP.[217,417] Indeed, two chromatographic systems also permit greater confidence in peak assignment if the ion is detected under both separation and/or ionization modes while also allowing for detection of weakly retained compounds that co-elute in the void front in one of the separation modes.[414] For instance, a greater coverage of urinary metabolites can be analyzed when using HILIC-MS in positive/negative ESI that is largely complementary to RP-LC/MS since urine is largely comprised of polar/ionic metabolites.[219] Similarly, metabolomics studies performed on several different specimens derived from the same organism further enhances confidence of the biological significance of putative disease markers when using complementary C18 and fluorinated-based stationary phases in LC-MS.[418] However, in many cases, unique metabolic fingerprints are characteristic to specific organs or biofluids as shown by UHPLC-MS when using a PFPP column that optimally resolved up to 112 unique hydrophilic metabolites within 8 min that was not feasible by RP (C18) and HILIC (amino) modes of separation.[419] The application of a complementary zwitter-ionic HILIC and C18 RP-LC strategy for comprehensive profiling of the tear metabolome that lead to the identification of 60 diverse classes of metabolites with a majority of compounds reported for the first time.[420] Indeed, expanded coverage of the metabolome in human serum using both HILIC and RP-LC datasets allows for better classification and staging of disease stage relative to healthy controls, such as renal cell carcinoma.[421] Alternatively, RP-LC in conjunction with ANP-LC can provide improved coverage of the metabolome with excellent reproducibility, which is important for revealing unanticipated perturbations of xanthine oxidreductase inactivation in mice models induced by gene deletion or pharmacological inhibition, such as treatment with allopurinol.[23] **Figure 1.11** highlights the technical reproducibility (CV = 6.5%, n = 56) and separation efficiency of polar metabolites in murine plasma when using ANP-LC-MS, including various classes of cationic and anionic metabolites that was shown to have better retention time precision while tolerating a wider solvent elution

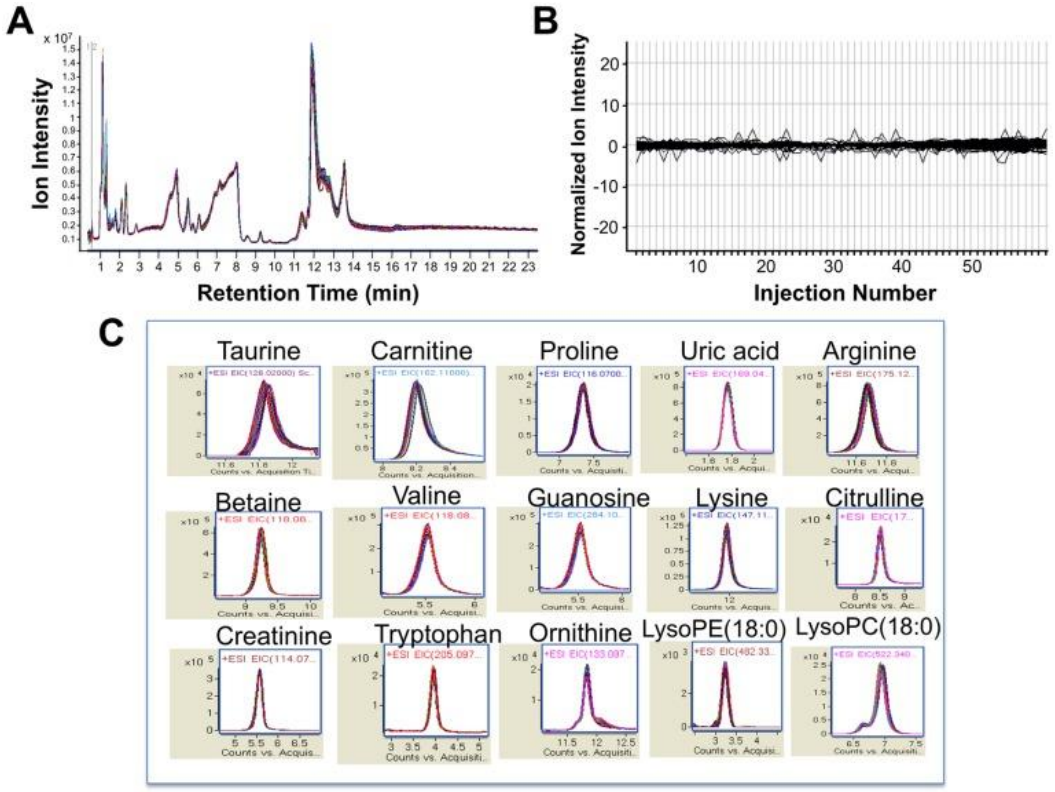


Figure 1.11. Technical reproducibility of plasma metabolite profiling by ANP-LC-MS. (A) Overlay of chromatograms acquired for 56 repeated analyses of a plasma specimen. . (B) Profile plot overlay of normalized extracted ion intensities for 374 distinct metabolites, quantified as a function of injection number for the 56 repeated plasma analyses depicted in A. Results depict flat run-to-run variation (CV = 6.5%) in the levels of the 374 metabolites with repeated analysis. (C) Peak overlay of 56 repeated assessments of extracted ion counts for various classes of polar metabolites in plasma demonstrating long-term method reproducibility.[23]

gradient range relative to HILIC.[23] Recently, a novel strategy for comprehensive metabolite profiling using aqueous and organic fractions of liver extracts by UHPLC-MS was reported based on three different stationary phases, namely C8, C18 and HILIC (amide) for non-polar lipids (organic), semi-polar (aqueous) and polar (aqueous) metabolites, respectively.[422] In this case, a differential metabolomics study was performed for identification of putative steatotic (fatty) markers of impaired liver function, which revealed systemic dysregulation in lipid, bile acid and amino acid metabolism.

Although a mixed-mode SPME coating fiber offers a sample cleanup method for biological fluids to reduce signal suppression in LC-MS,[299] the depth of the metabolome is limited by the intrinsic physicochemical properties of a

solute since small and hydrophilic metabolites have poor ionization efficiency and high detection limits.[262] Although LC-MS was originally introduced as a “label-free" approach for metabolite profiling, various functional group-specific chemical labelling strategies have been introduced to improve analytical performance,[205] including concentration sensitivity, separation selectivity, quantitative reliability and qualitative identification.[423] In most cases, a reagent that possesses an ionizable charged group (*e.g.*, quaternary ammonium) that is also bulky, hydrophobic and/or contain an isotopic signature (*e.g.*, Br, Cl ^{13}C, ^{2}H etc.) is preferred, where reactions can occur in free solution,[204] on a solid support resin for facile sample workup,[424] or dynamically during elution by formation of an ion pair complex.[425] Although chemical derivatization is often used to introduce a positive charge to weakly ionizable/neutral lipids,[426] it can also be applied to reverse the charge of acidic metabolites upon adduct formation that greatly enhances detection sensitivity over 2,500-fold when using positive ion mode, such as fatty acids.[427]

Figure 1.12 highlights the use of differential $^{12}C/^{13}C$ isotope dansylation labeling for relative quantification of amines, amino acids and phenols in human urine by LC-FTICR-MS under positive ion mode that results in a signal enhancement ranging from 1 to 3 orders of magnitude while improving retention of polar amino acids by RP-LC, including resolution of Leu and Ile isomers.[428] Since identical sample aliquots are derivatized with light and heavy (2 x ^{13}C) analogs of dansyl chloride and subsequently mixed together prior to LC-MS, co-elution of a pair of isotopomers offset by a *m/z* of 2 allows for reliable relative quantification without the need for specific stable-isotope internal standards. Absolute quantification of dansyl chloride-reactive metabolites is also feasible when isotopic encoding (^{13}C) a pooled reference sample relative to each individual sample, which allows for detection of 1,058 pairs of putative metabolites in saliva samples when using UHPLC-MS.[411] However, de novo identification of unknown metabolites remains a major obstacle since comparison of retention times and MS/MS spectra cannot be compared with archived public databases based on underivatized metabolites. Moreover, an array of selective chemical labels are ultimately needed for reacting with specific functional groups in order to extend coverage of the metabolome.[424]

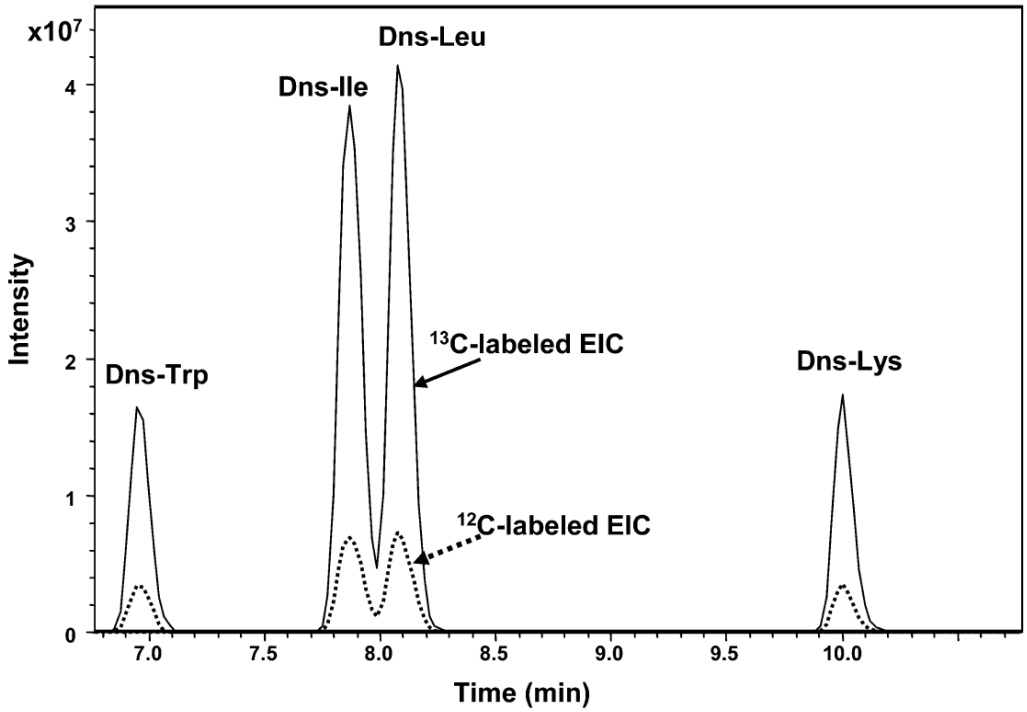


Figure 1.12. Extracted ion chromatograms of 1:5 of 12C-/13C-dansylated Trp, Ile, Leu, and Lys obtained by fast gradient RP-UHPLC with FT-ICR-MS. Pre-column chemical derivatization of amines, amino acids and phenols was performed using light and heavy isotopes of dansyl chloride for 60 min at 60°C with reaction quenched by methylamine prior to mixing of samples together prior to MS analysis**.** Dansylation labeling enhances solute ionization efficiency and improves retention of polar metabolites while also allowing for relative quantification of co-eluting isotope pairs without bias due to matrix-induced signal suppression.[428]

1.2.7 The Ionic Metabolome by CE-MS

CE-MS represents a promising platform in metabolomics that is applicable for comprehensive profiling of polar/ionic and labile metabolites that have poor retention in RP-LC or require complicated sample handling prior to GC-MS.[429] Moreover, high efficiency separations of multi-valent charged ions and their stereoisomers can be realized by CE that are difficult to resolve by HILIC or ANP without deleterious band broadening.[219] For instance, over 1692 molecular features are detected by CE-MS from cell extracts of *B. subtilus* comprising charged and hydrophilic metabolites that can be used for differentiation of bacterial sporulation.[430] Direct analysis of strongly ionic and secondary metabolites are also ideal for CE-MS without enzyme deconjugation, including *S*-glutathione, *N*-glycine, *O*-glucuronide and *O*-sulfate conjugates, as well as their

positional isomers.[253,431] Recently, a comparison of CE-MS with RP-UPLC-MS for comprehensive profiling of human urine under positive ion mode revealed that both techniques are largely complementary with respect to metabolome coverage with CE-MS preferentially detecting the majority of low molecular weight (m/z < 150) and ionic metabolite co-eluting in the void volume;[432] however, CE-MS detected 10-fold fewer significant molecular features in urine for subject gender classification relative to UPLC-MS due to sensitivity limitations related to the much lower sample volumes injected (10 nL versus 30 μL). Indeed, CE as a microseparation technique lends itself well to metabolomics studies of volume-restricted specimens and single cell analyses needed for understanding accelerated aging processes in mice models[433] and cell heterogeneity in freshly isolated and cultured neurons,[434] respectively. Nevertheless, migration time reproducibility, concentration sensitivity and method robustness remain key obstacles in CE-MS that has hindered its adaptation by the broader metabolomics community. This issue is partially explained by few commercial CE-MS vendors that offer full service support and the lack of customized software tools developed for processing electrophoretic data, which is required for correction of migration times with good precision.[243,281]

1.3 Capillary Electrophoresis Fundamentals

CE is a versatile separation technique with high separation efficiency (N > 100,000) that is appropriate for the separation of polar and ionic metabolites, where the capillary can also function as an on-line preconcentrator, desalter and microreactor.[435] ESI-MS is suitable for the analysis of large non-volatile molecules directly from the liquid phase.[262] CE-ESI-MS is a powerful hyphenated technique which separates analytes based on their electrophoretic mobilities and allows information about molecular mass and structure to be obtained.

1.3.1 Setup and Apparatus

As shown in **Figure 1.13,** a narrow bore (~25-75 μm ID) fused silica capillary (~50-120 cm length) connects a buffer vial to the ionization source and MS. The capillary is filled with a volatile buffer that is compatible with the MS and the sample is injected hydrodynamically or electrokinetically via the application of

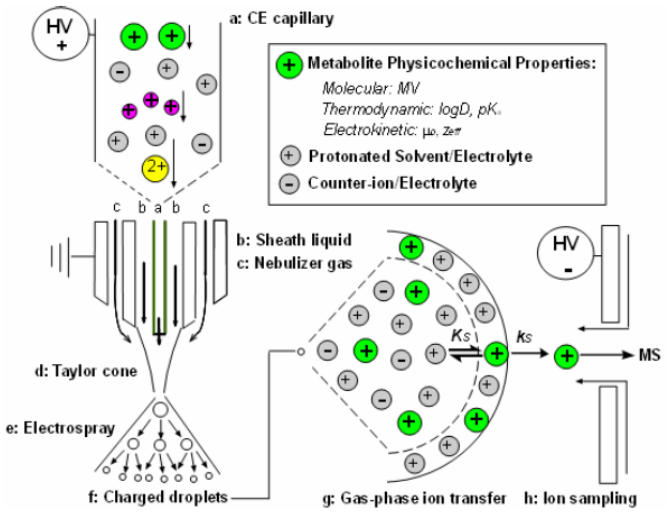


Figure 1.13. General setup of a CE-ESI-MS system showing separation in the CE capillary based on size-to-charge ratio and coaxial sheath liquid interface for electrospray ionization prior to sampling into the MS.[262]

pressure or voltage, respectively. A high voltage (up to 30 kV) is applied to the capillary and the components of the sample are separated based on their differential migration in the subsequent electric field. The narrow capillaries with high surface area-to-volume ratio allow for efficient heat dissipation and reduction in Joule heating (which can lead to band broadening) and permit the application of high voltages for fast separations.

1.3.2 Electrokinetic Parameters of Separation

There are two parameters in CE which contribute to the overall migration of analytes under an applied electric field, namely the electroosmotic flow (EOF) and electrophoretic mobility (μ_{ep}).

1.3.2.1 Electroosmotic Flow

The EOF is a flow that results from the inner charged surface of the uncoated fused silica capillary wall. The fused silica surface contains weakly acidic silanol groups (pK_a ~ 6.7), which is ionized to varying degrees depending on the buffer

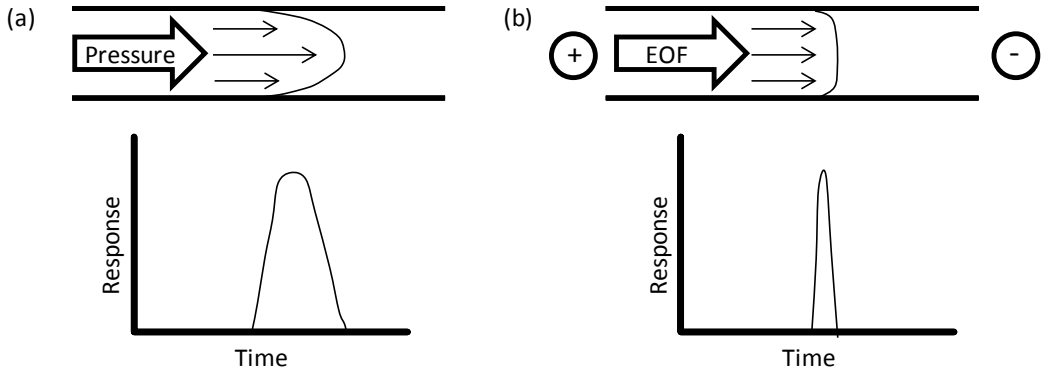


Figure 1.14. Comparison of fluid transport in (a) pressure-driven pumping system resulting in a parabolic flow profile and (b) an electrokinetic flow in CE resulting in a flat laminar flow profile with reduced band dispersion and higher separation efficiency.

pH, resulting in a negatively charged surface. An electric double layer is formed at this surface in accordance with the Debye-Hückel model. This double layer consists of the Stern layer, a rigid layer of cations adsorbed to the anionic surface, and a diffuse layer of mobile cations and anions, separated by the Stern plane. The rest of the bulk solution is separated from the double layer by the slipping plane, and the potential at this plane is the zeta potential (ζ). The charge density on the surface and the sign of the charge determines the magnitude and sign of ζ and the resulting EOF. At high pH the surface is primarily negatively charged, resulting in a large ζ and strong EOF, whereas at low pH the silanol groups are protonated and the EOF is effectively suppressed. Changing the charge of the surface by the use of a cationic coating will result in a reversed EOF, which is not generally useful for coupling with MS where a forward flow of analytes is required. The direction and magnitude of the EOF (μ_{eo}) is described by the following equation:[436]

$$\mu_{eo} = \frac{\varepsilon\zeta}{4\pi\eta} \qquad (1)$$

where ε is the dielectric constant, ζ is the zeta potential, and η is the buffer viscosity.[437] The EOF acts as an efficient mechanism for the transport of buffer and analytes in CE without external pumps. In addition, there is decreased band broadening because of the flat laminar flow profile compared to pump-driven LC separations where the parabolic flow profile increases longitudinal diffusion as shown in **Figure 1.14**.

The EOF can also be described by the following equation based on experimentally determined parameters:

$$\mu_{eo} = \frac{L_c L_d}{V t_{eo}} \quad (2)$$

Where L_c is the total length of the capillary, L_d is the length of the capillary to the detector ($L_c = L_d$ in CE-MS), V is the applied voltage, and t_{eo} is the migration time of the EOF as measured by monitoring a neutral EOF marker (*e.g.* caffeine, melatonin).

1.3.2.2 Electrophoretic Mobility

The second electrokinetic parameter involved in CE is the electrophoretic mobility (μ_{ep}) which describes the movement of solute ions in an applied electric field and is dependent on their physicochemical properties. The electrophoretic mobility is directly proportional to its charge (Q_{eff}) to hydrodynamic radius (R_H) ratio when assuming a uniformly charged spherical ion and it is described by the following equation:[436]

$$\mu_{ep} = \frac{v}{E} = \frac{Q_{eff}}{6\pi\eta R_H} \quad (3)$$

where v is the measured electrophoretic velocity, E is the applied electric field, Q_{eff} is the effective charge, and R_H is the hydrodynamic radius of an ion.

1.3.2.3 Apparent Electrophoretic Mobility

The apparent electrophoretic mobility of an analyte (μ_{ep}^{A}) is the vector sum of the EOF (μ_{eo}) and the analyte's electrophoretic mobility (μ_{ep}):

$$\mu_{ep}^{A} = \mu_{eo} + \mu_{ep} \quad (4)$$

The combination of these two electrokinetic forces allow for resolution of ions in free solution in the capillary. Small cationic species (*e.g.*, Na^+, K^+) have high positive electrophoretic mobilities and migrate with a high μ_{ep} towards the detector and effectively desalts the sample online without complicated offline sample workup. Typically, a low pH buffer (pH < 3) is used only for the analysis

of cations because the suppressed EOF allows time for effective separation of the cations in the capillary and strong anions (*e.g.*, Cl^-) with a high negative electrophoretic mobility will migrate out of the capillary and into the inlet vial. At low buffer pH, weakly anionic species will have very long migration times and suffer from band broadening. A high pH buffer (pH > 8) is generally used for the analysis of anions, where the EOF is not suppressed. Cations can also be analyzed under high buffer pH conditions but may not be adequately resolved from each other due to their fast μ_{ep} and correspondingly short migration times.

1.3.3 CE-MS Interfaces

Although CE enables high separation efficiency, it also has poor concentration sensitivity compared to LC techniques because of small sample injection volumes and small capillary inner diameters (25 – 75 μm) resulting in low flow of analytes out of the capillary. Since the first CE-ESI-MS interface was introduced in 1987,[141] a wide variety designs have been proposed to increase the concentration sensitivity and interface robustness. Since both CE and MS require electrical contact of the liquid (BGE and/or makeup flow) with an electrode at the capillary outlet for proper functioning of the CE separation and electrospray, interfaces must be designed to accommodate this. Additionally, ESI requires a certain flow rate to maintain a stable Taylor cone, which restricts the use of low-flow interfaces that are proposed to minimize dilution effects.

1.3.3.1 Sheath-Flow Interfaces

The most widely used, commercially available interfaces are sheath flow interfaces that use a make-up flow at relatively high flow rates that mixes with the CE effluent at the exit of the CE capillary prior to nebulization in ESI. Although this make-up flow provides the appropriate electrical contact between the electrode and the BGE and makes the resulting mixture appropriate for ionization of analytes, it significantly dilutes the sample and suction may be created, causing peak broadening and decreased resolution. Overall, sheath-flow interfaces are generally robust and suitable for commercialization, which has led to their prominence in CE-MS.

The most commonly used sheath-flow interface is the coaxial sheath flow interface. The flow rate out of the CE capillary is on the order of 10-100 nL/min,

which is not sufficient for creating a stable electrospray, so a constant flow of sheath liquid at a flow rate of ~ 5–20 μL/min is added to the CE flow using a coaxial sheath liquid interface.[146,438] Together, this flow leaves the capillary and makes up the Taylor cone at the exit of the electrospray tip. The nebulizer gas flows outside of the Taylor cone and assists in droplet formation and droplet desolvation while also influencing the stability of the electrospray.[31] The charged droplets are formed in the electrospray process and from these droplets, gas-phase ions are desorbed that are then sampled into the mass spectrometer and analyzed based on their *m/z* ratios. An advantage of this interface is that the separation in the CE capillary is undisturbed by the additional flow of liquid until the analytes exit the capillary. However, the flow can create suction that causes a parabolic flow profile, reduce separation efficiency, and increase peak broadening, and the high flow rates that are required to create a stable electrospray cause dilution of the analytes.[146] As a result, CE-MS suffers from poor sensitivity when compared to LC-MS due to small sample injections and post-capillary dilution effects when using the traditional coaxial sheath liquid interface.[439] To eliminate these dilution effects, various sheathless CE-MS interface designs have been developed to increase concentration sensitivity via nanospray formation over the past twenty years.[440]

1.3.3.2 Sheathless Interfaces

The first CE-ESI-MS interface, proposed in 1987, was a sheathless interface which employed a metal sheath around the fused silica capillary to replace the electrode, however its stability was poor.[141] In recent years, several sheathless interface designs have been proposed to increase concentration sensitivity and reduce dilution effects compared to the commonly used, stable, coaxial sheath-flow interface. These interfaces typically aim to improve over past designs by changing the shape of the emitter tip, the mechanism of electrical contact, and/or the physical assembly of the interface.[146] Changing the shape of the emitter tip affects the stability of the spray and stable ESI at the flow rates of the CE effluent can be achieved by narrowing the fused silica capillary. This can be accomplished by heating and stretching to a fine tip, after which hydrofluoric acid can be used for etching and refining.[441] Alternatively, the capillary can be mechanically ground, where the inner diameter and characteristics of the capillary are retained and no additional backpressure is caused in the CE capillary.[442] Although these different tip geometries are most commonly made using fused silica and glass, shaped stainless steel emitters have also been used as two piece interfaces and

offer increased robustness and ease of making electrical contact with the solution.[146] Overall, metal emitter tips are robust compared to etched or drilled fused silica capillary tips and offer mechanical stability that is unmatched by other sheathless interfaces. However, there remain significant challenges surrounding alignment of the silica capillary and metal tip, as well as electrolysis reactions inside the tip that have prevented their widespread use and commercialization.[146]

1.3.4 Capillary Coatings

Separation of cations and anions are typically performed using bare or surface coated fused-silica capillaries under strongly acidic (pH < 2) and alkaline (pH > 8) conditions, respectively. The main benefits of operating under extreme pH conditions is that it ensures that weakly ionic metabolites are adequately ionized while also generating a more stable EOF that is more variable near neutral pH as a result of the weak acid ionization of surface silanol groups ($pK_a \approx 6.5$). Indeed, long-term variations in EOF in fused-silica capillaries are highly dependent on electrolyte conditions and applied voltage, as well as surface properties of the capillary wall that alters with time due to cation adsorption[443] if proper rinsing procedures are not implemented. Although dynamically coated capillaries based on a triple layer of charged polymers improve long-term migration time reproducibility (*CV* < 1%) in high salt urine samples,[444] extra conditioning steps are needed for good performance, whereas surface adsorption of multi-valent anions can limit its applicability.[430] In most cases, bare fused-silica capillaries with 50 μm i.d. are widely used in CE-MS, however narrower capillaries (*e.g.*, 5-25 μm) have also been shown to improve separation efficiency by minimizing Joule heating and siphoning effects for ultra-fast separations (< 1 min).[445] Alternative capillary surfaces have also been explored in CE-MS, including polymer-based[446] and sulfonic acid modified fused-silica capillaries[447] in order to better control the EOF, as well as reduce ion adsorption. Soga and co-workers[430] first introduced three different operating conditions in CE-MS-based metabolomics for cations, anions and multivalent anions in order to maximize metabolome coverage; the first method is widely adopted by several other research groups using strongly acidic formic acid (5% *v*) or acetic acid (10% *v*) as the BGE since it allows for excellent analytical performance for a wide variety of cationic metabolites under suppressed EOF conditions with positive ion mode.[448–451] However, the two other methods for anion metabolite profiling that rely on a cationic coated polymer capillary (*e.g.*, SMILE (+)) or neutral

polydimethylsiloxane-coated GC-capillary under reversed polarity and negative ion mode[430] have not gained wider acceptance due to inherent flaws impacting method robustness and sample throughput.[257] Both types of coated capillaries are susceptible to hydrolysis with limited pH stability[452] that can lead to column bleed, ion adsorption and ion source contamination, whereas CE-MS under reversed polarity results in corrosion of the stainless steel sprayer.[256] Although more stable coated capillary formats with platinum-based sprayer assemblies are recently available,[256] this is ultimately unnecessary since anions (*e.g.*, organic acids, sugar phosphates, redox cofactors, nucleotides) can be readily analyzed using bare fused-silica capillaries under normal polarity with pressure-assistance flow[453] or methanol as an organic additive.[454] Although ammonium acetate or bicarbonate are widely used volatile buffers for analysis of anionic metabolites using alkaline conditions in CE-MS, improved sensitivity can be realized by using amine-based buffers as electrolytes in the BGE to reduce co-ion suppression, which contributes to a two-fold increase in the total number of molecular features detected in human urine.[263] Thus, caution is warranted against the use of polymer-coated capillaries under reversed polarity for anionic metabolite profiling by CE-MS given its many technical problems.[455]

1.3.5 Ionization Sources

For untargeted metabolomics, one or more ionization technique could be used to obtain complementary information. The most commonly used ionization technique used for CE-MS based metabolomics studies is electrospray ionization (ESI) because the CE-MS interface is commercially available and robust, while ionizing a wide range of compounds. Other less commonly used methods of ionization that can be coupled to CE are atmospheric pressure photoionization (APPI) and atmospheric pressure chemical ionization (APCI). Various different ESI interface designs are available, differing in their orientation of the sprayer, sprayer design, and the use of heated gas flow to enhance desolvation and ionization. Co-migrating compounds can cause substantial ion suppression in ESI and remains one of the major drawbacks to using ESI for reliable quantitative analysis. The use of an internal standard, ideally an isotope-labelled version of the analyte, allows for quantification that accounts for this ion suppression.

1.3.6 Mass Analyzers

Although triple quadrupole (QqQ) mass spectrometers are currently the standard for targeted quantitative analyses due to their robustness, cost, and sensitivity in selected reaction monitoring (SRM) and multiple reaction monitoring (MRM) mode, other mass spectrometers offer different characteristics and may be better suited to untargeted metabolomics. Ion traps are not often used for untargeted metabolomics due to slow acquisition times and limited mass range, but linear ion traps (LTQ), time of flight (TOF), Fourier transform ion cyclotron resonance (FT-ICR), and Orbitrap mass spectrometers can be well suited to metabolomic applications.

In metabolomics applications, the two parameters that are important to consider when selecting a mass spectrometer are duty cycle, and resolving power. Duty cycle is the time required to acquire data for a single experiment, which is important in order to obtain an adequate number of data points across a peak. The required duty cycle is dependent on the separation technique that is being used because different peak widths are associated with different techniques. In CE the peak widths are generally less than 12 s and in this case a fast acquiring mass spectrometer with a short duty cycle (< 1 s) is required in order to obtain the minimum requirement of 12 points over the peak.[456] Ion traps have a long duty cycle and are not well suited for CE, however, LTQ mass spectrometers offer similar sensitivities to QqQ and can operate in SRM mode. TOF, FTICR, and Orbitrap mass spectrometers have a fast duty cycle that is suitable for CE. For FTICR and Orbitrap instruments, the duty cycle is correlated to the resolving power of the instrument where a resolving power of 10,000 can be achieved at a duty cycle of 100 ms but a longer duty cycle (1-1.5 s) is required to obtain a resolving power of 100,000. The resolving power is not dependent on the duty cycle in a TOF, where a typical resolving power of 20,000-30,000 can be achieved at a very short duty cycle (10 ms).[456] The resolving power of an instrument is important because it permits differentiation between comigrating compounds with similar *m/*z. In metabolomics, where many analytes are being ionized at any given time, it is critical to be able to resolve analytes with high mass accuracy (< 5 ppm) for identification of unknowns and TOF, FT-ICR, and Orbitrap are well suited for because they offer this high resolution in full scan and/or MS/MS mode. Of these three fast duty cycle, high resolution mass spectrometers, TOF is the least expensive and most commonly used for metabolomics applications.

1.4 Sample Preparation

Common sample preparation protocols used in metabolomics prior to separation[457] include solvent extraction/precipitation, ultrafiltration/centrifugation, solid-phase extraction and chemical derivatization, which can lead to wide variations in metabolite recovery with distinct metabolic profiles measured for the same sample.[451] Thus, optimization of sample pretreatment is critical for ensuring stable yet representative metabolite concentration levels derived from the original specimen while also being compatible to a separation method and increasingly across multiple instrumental platforms.[458–460] A design of experiments can be used to optimize solvent extraction and chemical derivatization conditions needed to understand significant variables and their interactions that maximize recovery of metabolites.[461] All pre-analytical steps during sample pretreatment need to be investigated when using stable-isotope internal standards and/or pooled samples as quality controls to properly evaluate absolute metabolite recovery, process efficiency and residual enzyme activity with acceptable precision for various specimens.[462–465] Seemingly minor details regarding sample handling can result in significant changes in measured metabolite profiles, such as the impact of preservative agents in urine specimens,[466] delayed storage of cerebrospinal fluid samples,[467] slow deproteinization of plasma,[468] flash quenching of metabolism and extraction from cell cultures,[340] and storage temperature, pre-fractionation and repeated freeze-thaw cycles on serum lipids.[469] Indeed, the use of different solvent extraction methods can contribute to conflicting biological interpretations of intra-cellular metabolite profiles that raises questions on whether most published metabolomic studies accurately reflect the metabolic status of a sample without appropriate method validation.[470] Therefore, there is urgent need for the use of standardized protocols for sample collection, storage and processing[471] that is critical for large-scale prospective studies in metabolomics.[394,472] However, certain labile classes of metabolites remain difficult to quantify due to their inherent volatility, limited thermal or pH stability, photosensitivity or susceptibility to hydrolysis and/or oxidation, such as reduced thiols, unsaturated lipids and redox-active cofactors. Although the efficacy of liquid extraction methods are often dependent on solvent polarity and pH conditions,[463,473] combining a two-step series of extractions[474] or an optimized extraction using a single miscible solvent mixture[459] enables greater sample throughput and broader metabolite coverage with acceptable reproducibility. Long-established protocols developed for targeted metabolite profiling by GC-MS can also lead to unanticipated artifacts, such as urease treatment for sample cleanup of urine[153] and methoximation/silylation derivatization of polar metabolites.[475]

Similar to recent developments in GC-MS and LC-MS-based metabolomics, greater attention is required for optimization of sample pretreatment[451] and separation/ionization conditions in CE-MS[476] to enhance analytical performance without bias. Although separation of neutral steroids is feasible by MEKC or CEC, direct coupling to ESI-MS limits selectivity by restricting use of nonvolatile surfactants or additives due to signal suppression. To date, CE-ESI-TOF-MS using a coaxial sheath liquid interface with aqueous volatile buffers remains the most common instrumental format for routine metabolomics studies. Leon *et al.*[477] compared the metabolome profiles between three isogenic wild-lines relative to three transgenic strains of maize seed using CE-FT-ICR-MS after pressurized liquid extraction in order to evaluate their bioequivalence. In this case, genetically modified maize strains were differentiated from their parental wild lines when using partial least squares analysis, however several annotated features only expressed in genetically modified strains were not identified based on their empirical formula with the exception of *L*-carnitine, apigeninidin and 5,6-dihydroxyindole. One way to improve concentration sensitivity in CE-MS for anionic metabolite profiling is the use of chemical derivatization, such as *N*-butyl-4-aminomethylpyridium that contains a pre-charged quaternary amine, deuterium label (^{2}H x 9) and a reactive primary amine for labeling to carboxylic acids after activation.[478] In this case, CE separations can be performed under strongly acidic conditions and normal polarity when using positive ion mode ESI-MS since carboxylic acid derivatives possess a net positive mobility. Unlike LC-MS, deuterium-encoded isotopomers (d_8) were found to co-elute with light form derivatives in CE-MS thereby allowing for relative quantification of urine samples without bias, which also aided unknown structural identification by characteristic MS/MS transitions.[478]

Reduced thiols are extremely difficult to measure quantitatively due to their susceptibility to oxidation, thiol-disulfide exchange and poor ionization efficiency that requires careful optimization of all pre-analytical and analytical steps, including sample collection/storage, sample pretreatment and electrophoretic separation conditions.[431] Indeed, CE-MS allows for reliable quantification of reduced thiols and their oxidized disulfides from red blood cell lysates that is needed for assessment of the synergistic interactions of nutrition and exercise in human health.[28] For instance, time-resolved metabolomic studies revealed that over a 50-fold increase in intra-cellular oxidized glutathione was measured during the onset of strenuous exercise that was regulated back to pre-exercise levels upon recovery. High dose oral *N*-acetylcysteine pretreatment was

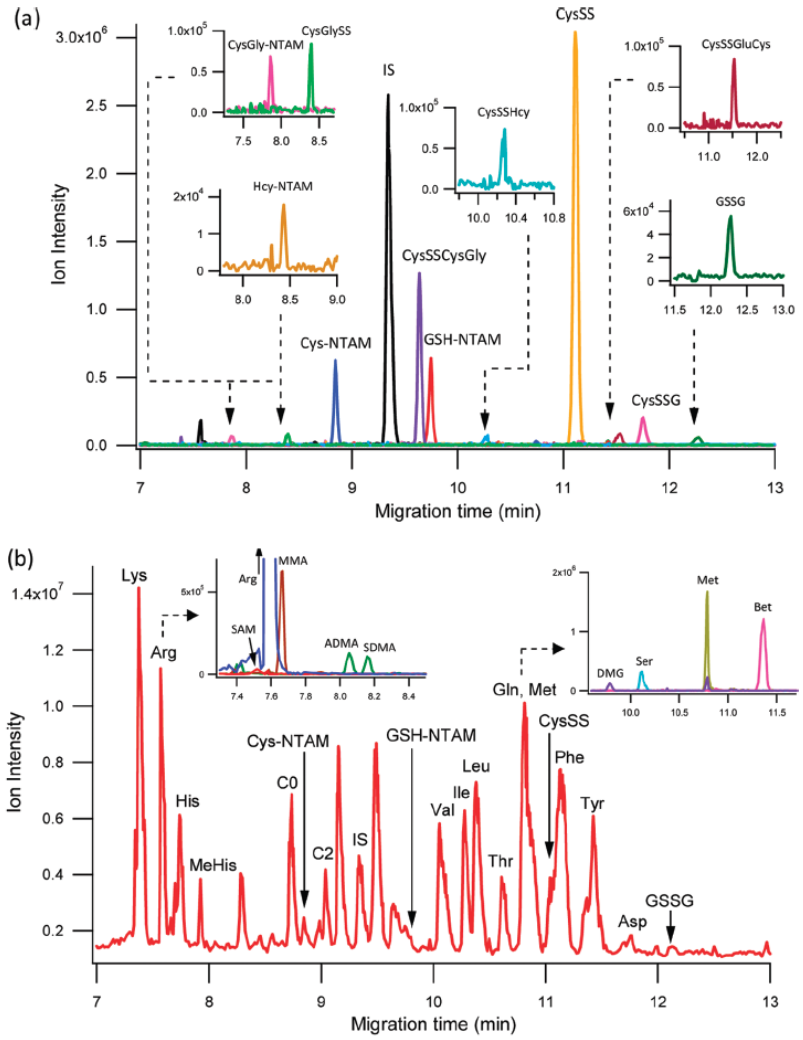


Figure 1.15. (a) Series of extracted ion electropherograms demonstrating comprehensive assessment of plasma thiol redox status by CE-MS in conjunction with thiol-selective maleimide (NTAM) derivatization for simultaneous analysis of labile reduced thiols, as well as intact symmetric and mixed oxidized disulfides over a wide dynamic range. (b) Total ion electropherogram for the same specimen that highlights integration of untargeted metabolite profiling for deeper insight into metabolic pathways modulated by thiol dysregulation.[479]

found to downregulate resting metabolism while modulating metabolic responses to exercise-induced oxidative stress that was associated with faster recovery.[28] Since oxidative stress is associated with aging and many chronic/acute diseases,[480] a new strategy for comprehensive plasma thiol redox status was introduced based on thiol-specific maleimide labeling for simultaneous determination of nanomolar levels of reduced thiols and free oxidized thiols as their intact symmetric and

mixed disulfides.[479] **Figure 1.15** highlights that artifact-free analysis of various plasma thiols can be achieved over three orders of dynamic range in CE-MS, where the use of a cationic maleimide label was found to stabilize reduced thiols as their thioether adducts while boosting sensitivity and shortening migration times. Overall, maleimide labeling was not found to change the responses of other detectable plasma metabolites, where a multivariate model was used to predict relative response factors for model thioether adducts with various maleimide substituents.[479]

A study exploring the correlation between sensory evaluation scores of Japanese sake to their metabolome using CE-MS was reported based on four trained panelists who graded samples based on their sweetness, sourness, bitterness and zatsumi, an unpleasant not clear flavor.[481] A predictive "electric tongue" model for estimating the taste and quality of different sake brands was developed using support vector regression of metabolomics datasets with cross-validation and external validation, which demonstrated good correlation notably with the undesirable zatsumi flavor score. However, only 14% of detected peaks were identified in this study, whereas the overall significance of altered levels of individual metabolites with respect to sensory responses is not well understood.[481] Recently, an extensive metabolomics study for the discovery of putative markers of liver disease among 248 serum samples derived from seven sub-classes of liver disorders and healthy controls.[449] CE-TOF-MS-based metabolomics revealed that γ-glutamyl dipeptides were found to be highly elevated in all liver diseases relative to controls, which was associated with deleterious oxidative stress due to depletion in reduced glutathione. In this case, a specific set of γ-glutamyl dipeptides, transaminases and methionine sulfoxide were collectively used to discriminate among multiple liver disorders when using receiver-operating characteristic curves as a way to provide a reliable diagnostic test without invasive liver biopsies.[449] A similar approach was reported for the discrimination of different types of cancer using non-invasive saliva sampling with CE-MS from 211 different patients, including oral, pancreatic and breast cancer patients, as well as periodontal disease and healthy controls.[482] However, since salivary metabolite profiles are highly dependent on various physiological and lifestyle factors (*e.g.*, smoking, time of day, age)[483] standardized sampling protocols and carefully designed cohort groups are needed to minimize innate biological variation notably when collected from different hospital settings or institutes.

1.5 Data Quality in Metabolomics

1.5.1 Sources of Bias and False Discovery in Metabolomics

The major aim of metabolomics is to identify statistically different concentration levels of metabolites between two or more groups of defined samples (*e.g.*, placebo/treated; wild-type/mutant etc.) based on relative quantification of their average signal intensity when using NMR and/or MS in conjunction with optimized separation systems. However, the overall significance of putative "marker" compounds will ultimately depend on the underlying stability of metabolite concentration levels in a given sample and the precision associated with the measurement. In many ways, metabolomic studies are confronted with the same problems that have plagued the failure of protein cancer biomarker discoveries to date.[484] Sources of bias are attributed to either fraud, false discoveries due to improper validation or true discoveries that ultimately lack adequate biological significance or clinical value to patient care in the context of diagnosis, prognosis, sub-class stratification or therapeutic monitoring.[485] Bias is minimized with a carefully planned study design while implementing a fully validated metabolomics workflow (*i.e.*, pre-analytical, analytical, post-analytical and bioinformatic) that depends on knowledge of the chemical properties and biological functions of a marker compound, including its natural variation in sample and intrinsic chemical stability. Indeed, there have been few reported metabolomic studies that have translated biomarker discovery into clinical practice with the probable exception of sarcosine as a prognostic marker of prostate cancer[486] that is still being evaluated for its appropriate clinical utility.[487]

Early studies published in metabolomics have recognized the impact of diurnal cycle, diet, gender and oestrus cycle on measured metabolite profiles in laboratory rats,[488–490] which were long known as "extrinsic" factors that impact metabolite profiling of biological fluids by GC-MS.[119,491] For instance, recent studies have demonstrated that about 15% of identified metabolites in human plasma (*e.g.*, fatty acids) and saliva (*e.g.*, amino acids) are under circadian control.[492] The biological variance in the metabolome of yeast can vary from about 10% CV (n=12) when cultured under defined conditions, but can exceed 40% CV for more complex organisms depending on food source, such as *C. elegans*.[493] Although different in-bred genetic strains of mice have distinct microflora populations reflected within the host metabolome, these variations are absent in the case of genetically distinct mice gestated by the same mother or isogenic mice relocated to different laboratories, highlighting the strong influence

of environment on metabolic profile.[494] Although the ability to implement standardized conditions in metabolomics is more feasible with laboratory animals, the innate biological variability of metabolism places major constraints when performing observational studies involving humans notably when specimens are only collected at a single time frame as a "metabolic snapshot", such as a morning fasting blood sample.[28] For instance, a longitudinal study of biofluids collected from healthy post-menopausal female twins over several months demonstrated that endogenous metabolite concentration variation have a stable familial and individual-environmental component with an average variance of 47% in urine that increased to over 60% in plasma for major compounds detected by ^{1}H-NMR.[495] Similarly, LC-MS-based metabolomics studies confirmed biological variability ranging from a median variance of 35% in human cerebrospinal fluid to over 46% in the case of human plasma when using an optimized methanol extraction protocol for sample deproteinization prior to LC-MS.[496] Thus, metabolomic studies based on drug or lifestyle interventions suffer from significant intra-individual (*e.g.,* diurnal cycle, menstrual cycle etc.) and inter-individual variances (*e.g.,* age, ethnicity etc.) as well as less well characterized effects, such as time of last meal, hydration status and non-compliance with study protocols.[497] In contrast to large-scale randomized controlled trials, cross-over interventions offer a powerful approach in metabolomics[498] for measuring authentic treatment effects on an individual level instead of a group level since the same subject serves as their own control.[499] Similarly, time-resolved metabolomics involving stressor/challenge experiments with multiple specimens collected from the same individual over a short time duration can distinguish subtle pre-symptomatic phenotypes not evident by steady-state experiments due to high biological variance.[500] Since acceptable technical/process (CV < 15%) and inter-laboratory variance (CV < 20%)[259] is often less than "normal" biological variance under standardized conditions (CV < 60%), a minimum threshold change of about a 2-fold for pair-wise comparison of relative responses in average metabolite signals is considered statistically significant when applying Bonferroni correction to reduce the high false discovery rate with multiple hypothesis testing.[501] False-discoveries in metabolomics often stem from inadequate sample size or group class selection that contributes to data overfitting when using multivariate statistics and machine learning. In this case, stringent testing of predictive accuracy, such as cross-validation of a training set with independent testing of a hold-out set can be performed to demonstrate biological significance.[502]

1.5.2 Quality Assurance in Data Processing

Given the recent expansion of metabolomics across various research fields that employ different instrumental techniques, a consensus on a minimum set of reporting standards to enhance data quality is required while providing meta-data information to promote transparency, best practices and data exchange.[305,503] For instance, quality assurance practices in metabolomics typically incorporate the use of a series of internal and/or recovery standards to evaluate variation in sample processing that also improves data alignment,[90] comparison of different data transformation methods on biological information content,[504] intermittent injection of a pooled sample as quality control for assessing long-term signal drift,[505] a blank subtraction and/or a dilution trend filter to reject a majority of "spurious" peaks in ESI-MS.[310] A variety of proprietary, open-source and web-based software packages[281] have been developed for data pre-processing of high resolution MS-based metabolomics data sets, including file conversion, mass filtering, retention time alignment, peak picking, data normalization/transformation together with integrated bioinformatics tools for statistical analysis and visualization;[243,506–511] however, systematic bias can result during peak picking/integration of significant molecular features due to deconvolution algorithms not optimized for specific instrumental platforms.[153,512] Indeed, there are on-going efforts to enhance data pre-processing strategies[513] in order to improve the quality of peak selection based on a "reliability index" of significant features derived from a dilution series[400] instead of maximizing the total number of detectable features. This process is essential to down-stream efforts dedicated to data analysis[514] and structural identification[321] that constitute two major bottlenecks in a metabolomics workflow. Biological interpretation of complex metabolic networks on a systems level also remains a major obstacle in metabolomics[515] that requires integration of metabolic flux studies[516] and "multi-omic" data approaches to unravel poorly understood genotype-phenotype interactions.[517,518]

1.6 Lifestyle Intervention for Prevention of Diabetes

Diabetes mellitus is a metabolic disorder characterized by hyperglycemia caused by defective insulin secretion and/or defective insulin action.[519] Type 1 diabetes is a chronic autoimmune condition that primarily results from destruction of pancreatic β-cells and failure of the body to produce insulin. Effective, safe

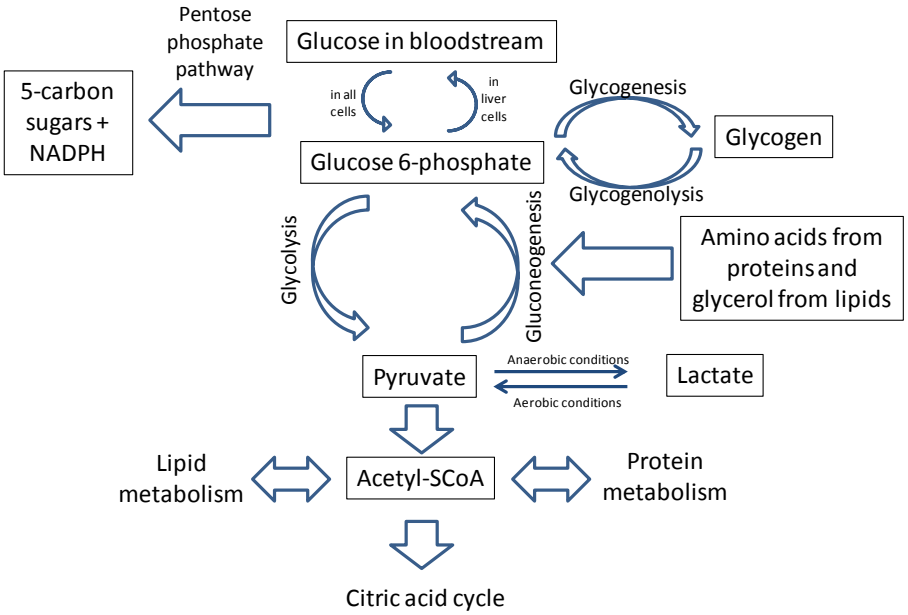


Figure 1.16. Metabolism pathways for glucose upon entry into cells from the bloodstream.[521]

interventions for prevention of type 1 diabetes have not yet been developed.[520] Type 2 diabetes is caused by both insulin resistance and insulin deficiency. Insulin resistance occurs when circulating insulin is less effective at regulating sugar (including glucose) metabolism. Insulin resistance causes a reduction in glucose uptake into muscle and fat cells and results in persistently elevated blood glucose levels.[522] Diabetes is a rapidly growing problem worldwide, and as of 2009 it has affected an estimated 6.8% of Canadians.[523,524] Corresponding with its prevalence, the direct and indirect costs associated with diabetes have also increased dramatically. Although type 1 diabetes has no successful prevention protocol to date and remains an area of intense research, intervention for type 2 diabetes has been shown to be effective at decreasing the rates of cardiovascular disease and renal failure.[525]

1.6.1 Diagnostic Criteria and OGTTs

Diagnosis of diabetes as defined by the Canadian Diabetes Association (2013) relies on meeting one of the four following criteria: a fasting plasma glucose (FPG) level of ≥ 7.0 mM; glycated hemoglobin (A1C) ≥ 6.5%; a 2-hour postload

plasma glucose (2hPG) value ≥ 11.1 mM; or a random plasma glucose (PG) ≥ 11.1 at any time of day when symptoms of diabetes are present.[526] The 2hPG is determined by the administration of an oral glucose tolerance test (OGTT). This involves the administration of a high dose of glucose (75 g) after an 8-hour fasting period, after which the targeted analytes of interest are monitored in plasma or serum. The initial fasting glucose level, in combination with the PG level 2 hours after the glucose administration are both used for diagnosis of diabetes and insulin resistance.[519] Although the most commonly targeted analytes measured in an OGTT are glucose and insulin in order to diagnose diabetes and insulin sensitivity, there are other physiological and metabolic changes during an OGTT that cause other metabolite concentrations to change.[527] Glucose metabolism begins with the immediate metabolism of glucose to glucose 6-phosphate upon entry to a cell from the blood stream, trapping it inside the cell.[521] Excess glucose is converted to other storage pathways including glycogen by the glycogenesis pathway or fatty acids via the acetyl-SCoA pathway as shown in **Figure 1.16.**

Several classes of compounds have been identified that have time-dependent patterns in OGTTs and may offer insight into glucose metabolism and the effect of fasting OGTTs on an individual. A range of metabolites are affected during an OGTT, as shown in **Figure 1.17**. Several metabolites, along with glucose, are elevated during an OGTT including lactic acid,[528] hippuric acid,[528] free carnitine,[500] mannose,[529] and lysophosphatidylcholines.[527] A variety of classes of metabolites are decreased during an OGTT, notably free fatty acids,[527] acylcarnitines,[500] branched chain amino acids (BCAAs),[528] and aromatic amino acids (AAAs).[528] Several individual metabolites have also been noted to be attenuated during an OGTT, specifically hypoxanthine, glycerol, and β-hydrobutyric acid.[528] Glutathione follows a dynamic profile where it is elevated and then decreased over the course of a 2 hour OGTT.[530] It has also been noted that insulin resistance is correlated with blunted alterations in several metabolites, including glycerol, leucine, and isoleucine.[528]

1.6.2 Interventions for Disease Prevention

Early interventions that include lifestyle changes (*i.e.* increased activity, modified diet) to reduce body weight, sometimes combined with pharmacotherapy, have been shown to be effective at preventing the progression of prediabetes to type 2 diabetes. Prevention of this progression may result in lower rates of associated

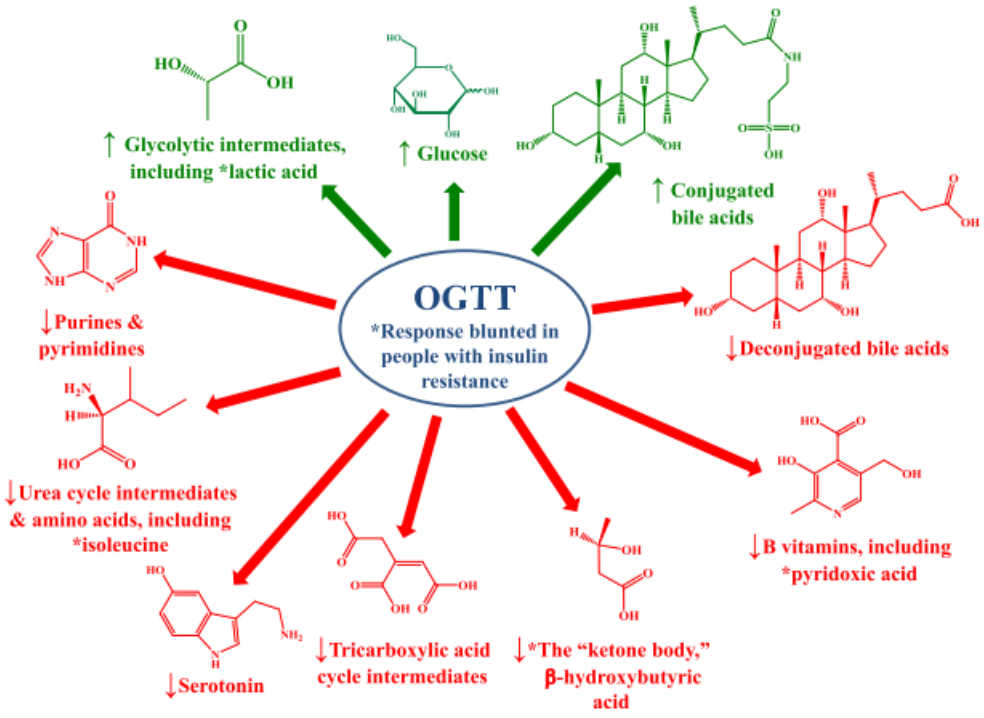


Figure 1.17. Metabolites and classes of compounds that are altered during an OGTT, both elevated (green) and decreased (red). Four metabolites (*) demonstrated attenuated responses in subjects with insulin resistance.[531]

health problems, including cardiovascular disease, renal failure, blindness, and premature mortality.[520] Various historical, physical, and biochemical variables have been associated with development of type 2 diabetes including older age, obesity, physical inactivity, certain ethnic backgrounds, history of gestational diabetes mellitus, and overt coronary artery disease.[519] These variables can be used to target high-risk individuals and population subgroups in order to more efficiently promote physical activity and healthy eating to prevent diabetes. Lifestyle intervention for type 2 diabetes typically involve a program with increased physical activity and modified diet to reduce excess body mass. Insulin resistance is strongly correlated to obesity and is a defining factor in type 2 diabetes. Reduced muscle mitochondrial capacity, impaired glucose transport, and intramyocellular lipid accumulation have all been associated with insulin resistance in sedentary populations.[532–535] Interventions involving regular exercise training are not simply about "losing weight" to reduce risk factors associated with obesity, but also to promote the proper functioning of mechanisms such as

blood glucose homeostasis, insulin sensitivity, and skeletal muscle metabolic health. Numerous studies have shown that exercise increases insulin action in skeletal muscle in obese and insulin resistant people and slows the progression of diabetes.[536] Overall, it has been well-established that changes in lifestyle via modification of diet and increased physical activity in order to reduce body weight are a significant factor in reducing risk and symptoms of diabetes, however, there is not necessarily a "one size fits all" intervention.[537–540]

1.6.3 High-Intensity Interval Training Intervention

High-intensity interval training (HIT) is a physical exercise protocol characterized by short, intense bursts of activity interspersed with periods of rest or low-intensity exercise. The types of exercise, number and duration of intervals, and intensity can be varied for different training strategies and/or goals. Recent work indicates that HIT can induce metabolic adaptations similar to those from traditional continuous moderate-intensity training (CMIT).[541–543] When HIT is compared to CMIT on a "matched work" basis where energy expenditure by the individual is equivalent in each case, as in these studies, HIT is generally equivalent or superior to traditional CMIT for improving a range of physiological and health-related markers.[544] Gibala *et al.* have proposed a HIT model consisting of 10 x 60s constant intensity bouts of high intensity exercise at a load that elicits ~90% of maximal heart rate, interspersed with 10 x 60s low intensity recovery periods for a total of 10 min of exercise in a 20 min session. The low-volume aspect of this exercise protocol is advantageous as it only requires approximately one hour per week (3 x 20 min sessions) compared to traditional endurance training which requires approximately 3-5 hours per week. This low-volume feature combined with the physiological and metabolic improvements associated with HIT suggest that it may be an ideal strategy for people who are overweight and/or demonstrate insulin resistance.

To date, most studies investigating the effect of HIT on cardiovascular disease and metabolic disorders have utilized high-volume exercise protocols that are similar in workload to traditional high-volume endurance training. There are few studies that have focused on low-volume HIT interventions for diabetes-related diseases, however insulin sensitivity has shown to be improved over a 2-week low-volume HIT protocol in sedentary adults.[545] Another 2-week low-volume HIT study demonstrated increased concentrations of skeletal muscle protein GLUT4 compared to high-volume endurance training, which may benefit

glucose transport capacity and improve insulin sensitivity.[546] HIT has also been shown to reduce postprandial hyperglycemia (cumulative AUC of post-meal glucose response) and overall 24 h hyperglycemia (time spent > 10.0 mM) after one HIT session, with each subject acting as their own control compared to a non-exercise control day.[547] In a longer, 12-week study (n = 22), subjects with coronary artery disease (CAD) showed similar improvements in fitness and brachial artery flow-mediated dilation in response to both low-volume HIT and high-volume CMIT, which may be useful to diabetic patients with CAD.[548]

Traditional parameters that are measured in HIT studies include physiological parameters such as body composition (muscle/fat/bone), and BMI, fitness parameters such as maximal oxygen uptake and heart rate, along with parameters specific to the study such as heart function parameters related to cardiovascular disease, targeted analysis of glucose and insulin, and/or mitochondrial enzyme activity involved in muscle protein synthesis. To date, no metabolomic studies of HIT interventions have been used to study either the effect of acute HIT sessions on metabolic profile or the effect of a HIT intervention on an individual's response to a stressor. In this thesis, metabolomics will be used to more deeply explore HIT to group participants into categories (*e.g.*, good responders, non-responders, poor responders) in an effort to lead to more personalized exercise interventions for improving health and physiological outcomes for overweight/obese women at risk for developing metabolic diseases, such as diabetes.

1.7 Research Objectives

Metabolomics offers valuable insights into understanding the phenotype of an organism and mechanisms of lifestyle or drug interventions via untargeted metabolite profiling for discovery of specific metabolites of clinical or biological significance. Metabolomics can also be applied to monitor the effects of acute stressors, such as ingestion of high dose glucose, as well as evaluate differential treatment responses to therapy, including low volume high intensity interval training. Since metabolites represent functional endpoints of gene expression and cell activity, they can be used to classify individuals into groups based on their characteristic metabolic response to a stimulus/stressor in order to develop more efficacious yet customized interventions without adverse effects (*i.e.* personalized medicine). CE-ESI-MS is a powerful, high efficiency technique where the CE and TOF-MS offer complementary separations, based on charge-to-size and mass-to-

charge ratios, respectively. It is appropriate for the separation of polar, ionic metabolites, and non-volatile molecules directly from the liquid phase, and is well-suited for metabolomics applications because of the wide variety of compounds that can be analyzed simultaneously without complicated sample workup of human biological fluids. In addition, CE is well-suited for the high efficiency separation of a metabolites that differ greatly in their polarity, such as weakly acidic estrogens and their intact ionic conjugates, which have poor retention by other separation techniques (*e.g.*, reversed-phase LC) unless compounds are enzymatically deconjugated prior to analysis. However, the latter process reduces biological insight due to the critical importance of secondary metabolism (*i.e.* sulfation and glucuronidation) in modulating the activity of potent estrogen hormones. Moreover, resolution of stereoisomers by CE provides greater selectivity since certain positional isomers of estrogen metabolites have genotoxic properties that can increases risk for breast cancer. However, a common limitation to all MS-based separation platforms is their low sample throughput due to long analysis times (> 20 min) associated with solute elution and column re-equilibration that is further hindered by complicated data pre-processing in large-scale metabolomic studies.

In this thesis, multisegment injection capillary electrophoresis-mass spectrometry (MSI-CE-MS) is introduced as a novel approach for multiplexed analysis for metabolomics and biomarker discovery. MSI-CE-MS allows for higher throughput analysis via the injection of multiple samples in series while providing better data quality because of the reduced time for ionization between samples, built-in QCs samples, and encoded information based on injection sequence and dilution patterning to rigorously filter raw data. In MSI-CE-MS, isomer and metabolite resolution is maintained while enhancing sample throughput by up to one order of magnitude. The seven sample segment sequence allows for various different configurations to be creatively designed to impart high data fidelity and reduce false discoveries by eliminating artifact and spurious background ion peaks that can lead to data over-fitting. Additionally, the seven-segment injection sequence can be used to evaluate differential treatment outcomes among heterogeneous subject populations, whose responses to a stressor (*i.e.* acute glucose consumption, strenuous exercise) are highly variable. With each subject acting as their own control, time-dependent samples can be analyzed within a single run, increasing confidence in data quality due to reduction in analysis time and instrumental variation.

MSI-CE-MS is used to investigate whether a 6-week exercise training protocol can improve the metabolic status of overweight/obese yet non-diabetic women (beyond glucose tolerance and insulin resistance) as evidence for diabetes prevention. Diabetes is a rapidly growing problem in North America, and exercise interventions have been shown to be effective at prevention, although response to exercise varies dramatically between subjects. Ideally, personalized interventions for disease an health disorders such as diabetes and obesity can be developed based on individual metabolic profile, wherein individual response to exercise can be accounted for based on a simple blood test involving a panel of prognostic markers that predict gains in aerobic endurance capacity and/or increases in insulin sensitivity while also identifying non-responsive subjects or individuals at risk for adverse effects to exercise training. HIT is a practical yet effective lifestyle intervention for prevention of diabetes and metabolomics is a valuable tool that can be used to probe underlying metabolic profiles and classify adaptive training response to therapy for sub-groups of subjects. In this thesis, both plasma samples from HIT exercise sessions as well as plasma samples from diagnostic OGTTs before and after the 6 week exercise intervention are used to investigate underlying metabolic phenotypes (*i.e.,* metabotypes) of subjects associated with physiological changes and fitness improvement in an effort towards the design of more personalized interventions for at-risk individuals.

The work in this thesis focuses on overcoming analytical challenges through development of novel analytical and bioinformatic methods in metabolomics related to promotion of human health and primary disease prevention. The major research objectives in this work are (a) development of a novel assay for comprehensive profiling of estrogen conjugates with improved metabolome coverage, (b) development of a high-throughput method for large-scale metabolomics studies to maintain high data quality with decreased analysis times, (c) an untargeted metabolomics investigation of adaptive training responses to a 6-week HIT intervention in overweight/obese women, and (d) evaluation of the efficacy of HIT to modulate metabolic responses to standardized oral glucose tolerance tests as evidence for diabetes prevention in healthy yet at risk individuals. The ultimate goal of this thesis is to overcome existing analytical challenges in metabolomics to improve metabolome coverage, throughput, and data quality and apply these methods to current metabolomic challenges to provide insight into the efficacy and mechanisms of lifestyle interventions that elicit largely positive yet highly variable physiological responses within a heterogeneous cohort despite similarity in overall body composition ($BMI > 25$). Moreover, this work also proposes an accelerated metabolomics workflow for

biomarker discovery that takes advantage of the increased sample throughput and higher information content derived from multiplexed CE-based separations.

1.8 References

(1) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263.

(2) Veenstra, T. D. *Genome Med.* **2012**, *4*, 40.

(3) Johnson, C. H.; Gonzalez, F. J. *J. Cell. Physiol.* **2012**, *227*, 2975.

(4) Zhao, L.; Nicholson, J. K.; Lu, A.; Wang, Z.; Tang, H.; Holmes, E.; Shen, J.; Zhang, X.; Li, J. V; Lindon, J. C. *J. Proteome Res.* **2012**, *11*, 3509.

(5) Saito, K.; Matsuda, F. *Ann. Rev. Plant Biol.* **2010**, *61*, 463.

(6) Jones, D. P.; Park, Y.; Ziegler, T. R. *Annu. Rev. Nutr.* **2012**, *32*, 183.

(7) Antignac, J.-P.; Courant, F.; Pinel, G.; Bichon, E.; Monteau, F.; Elliott, C.; Le Bizec, B. *TrAC, Trends Anal. Chem.* **2011**, *30*, 292.

(8) D'Alessandro, A.; Giardina, B.; Gevi, F.; Timperio, A. M.; Zolla, L. *Blood Transfus.* **2012**, *10 Suppl 2*, s19.

(9) Viant, M. R.; Sommer, U. *Metabolomics* **2013**, *9*, S144.

(10) Han, X.; Gross, R. W. *J. Lipid Res.* **2003**, *44*, 1071.

(11) Marshall, A. G.; Rodgers, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18090.

(12) Jones, C. J.; Larive, C. K. *Anal. Bioanal. Chem.* **2012**, *402*, 61.

(13) Agnolet, S.; Jaroszewski, J. W.; Verpoorte, R.; Staerk, D. *Metabolomics* **2010**, *6*, 292.

(14) Lei, Z.; Huhman, D.; Sumner, L. *J. Biol. Chem.* **2011**, *286*, 25435.

(15) Metz, T. O.; Zhang, Q.; Page, J. S.; Shen, Y.; Callister, S. J.; Jacobs, J. M.; Smith, R. D. *Biomark. Med.* **2007**, *1*, 159.

(16) Dettmer, K.; Aronov, P.; Hammock, B. *Mass Spectrom. Rev.* **2007**, *26*, 51.

(17) Psychogios, N.; Hau, D.; Peng, J.; Guo, A.; 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.; Smith, S.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.; Goodfriend, T.; Wishart, D. *PLoS One* **2011**, *6*, e16957.

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

(19) Kind, T.; Scholtz, M.; Fiehn, O. *PLoS One* **2009**, *4*, e5440.

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

(21) Lu, W.; Clasquin, M. F.; Melamud, E.; Amador-Noguez, D.; Caudy, A.; Rabinowitz, J. D. *Anal. Chem.* **2010**, *82*, 3212.

(22) Saito, N.; Robert, M.; Kitamura, S.; Baran, R.; Soga, T.; Mori, H.; Nishioka, T.; Tomita, M. *J. Proteome Res.* **2006**, *5*, 1979.

(23) Chen, Q.; Park, H.-C.; Goligorsky, M. S.; Fischer, S. M.; Gross, S. S. *PLoS One* **2012**, *7*, e37149.

(24) Trupp, M.; Zhu, H.; Wikoff, W.; Baillie, R.; Zeng, Z.; Karp, P.; Fiehn, O.; Krauss, R.; Kaddurah-Daouk, R. *PLoS One* **2012**, *7*, e38386.

(25) Van Duynhoven, J.; Vaughan, E.; Jacobs, D.; Kemperman, R.; van Velzen, E.; Gross, G.; Roger, L.; Possemiers, S.; Smilde, A.; Doré, J.; Westerhuis, J.; Van de Wiele, T. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108 Suppl.*, 4531.

(26) Van Velzen, E.; Westerhuis, J.; van Duynhoven, J.; van Dorsten, F.; Grün, C.; Jacobs, D.; Duchateau, G.; Vis, D.; Smilde, A. *J. Proteome Res.* **2009**, *8*, 3317.

(27) Brugnara, L.; Vinaixa, M.; Murillo, S.; Samino, S.; Rodriguez, M. A.; Beltran, A.; Lerin, C.; Davison, G.; Correig, X.; Novials, A. *PLoS One* **2012**, *7*, e40600.

(28) Lee, R.; West, D.; Phillips, S.; Britz-McKibbin, P. *Anal. Chem.* **2010**, *82*, 2959.

(29) Capriotti, A. L.; Caruso, G.; Cavaliere, C.; Foglia, P.; Bizzarri, M.; Laganà, A. *Chromatographia* **2011**, *73*, 45.

(30) Weckwerth, W.; Morgenthal, K. *Drug Discov. Today* **2005**, *10*, 1551.

(31) Nielsen, J.; Oliver, S. *Trends Biotechnol.* **2005**, *23*, 544.

(32) Goodacre, R. *J. Nutr.* **2007**, *137*, 259S.

(33) Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3698.

(34) Holmes, E.; Li, J.; Athanasiou, T.; Ashrafian, H.; Nicholson, J. *Trends Microbiol.* **2011**, *19*, 349.

(35) Wang, Z.; Klipfell, E.; Bennett, B. J.; Koeth, R.; Levison, B. S.; DuGar, B.; Feldstein, A. E.; Britt, E. B.; Fu, X.; Chung, Y.-M.; Wu, Y.; Schauer, P.; Smith, J. D.; Allayee, H.; Tang, W. H. W.; DiDonato, J. A.; Lusis, A. J.; Hazen, S. L. *Nature* **2011**, *472*, 57.

(36) Lewis, N.; Nagarajan, H.; Palsson, B. *Nat. Rev. Microbiol.* **2012**, *10*, 291.

(37) Levy, S.; Sutton, G.; Ng, P. C.; Feuk, L.; Halpern, A. L.; Walenz, B. P.; Axelrod, N.; Huang, J.; Kirkness, E. F.; Denisov, G.; Lin, Y.; MacDonald, J. R.; Pang, A. W. C.; Shago, M.; Stockwell, T. B.; Tsiamouri, A.; Bafna, V.; Bansal, V.; Kravitz, S. A.; Busam, D. A.; Beeson, K. Y.; McIntosh, T. C.; Remington, K. A.; Abril, J. F.; Gill, J.; Borman, J.; Rogers, Y.-H. Y.-H.; Frazier, M. E.; Scherer, S. W.; Strausberg, R. L.; Venter, J. C. *PLoS Biol.* **2007**, *5*, e254.

(38) Adamski, J. *Genome Med.* **2012**, *4*, 34.

(39) Chace, D.; Sherwin, J.; Hillman, S.; Lorey, F.; Cunningham, G. *Clin. Chem.* **1998**, *44*, 2405.

(40) Fleeman, N.; Dundar, Y.; Dickson, R.; Jorgensen, A.; Pushpakom, S.; McLeod, C.; Pirmohamed, M.; Walley, T. *Pharmacogenomics* **2011**, *11*, 1.

(41) Bowen, B.; Northen, T. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 1471.

(42) Fleischmann, R. D.; Adams, M. D.; White, O.; Clayton, R. A.; Kirkness, E. F.; Kerlavage, A. R.; Bult, C. J.; Tomb, J. F.; Dougherty, B. A.; Merrick, J. M. *Science.* **1995**, *269*, 496.

(43) Goffeau, A. A.; Barrell, B. G.; Bussey, H.; Davis, R. W.; Dujon, B.; Feldmann, H.; Hoheisel, J. D.; Jacq, C.; Johnston, M.; Louis, E. J.; Mewes, H. W.; Murakami, Y.; Tettelin, H.; Oliver, S. G. *Science.* **1996**, *274*, 546.

(44) Venter, J. C. *Science.* **2001**, *291*, 1304.

(45) Raamsdonk, L. M.; Teusink, B.; Broadhurst, D.; Zhang, N.; Hayes, A.; Walsh, M. C.; Berden, J. A.; Brindle, K. M.; Kell, D. B.; Rowland, J. J.; Westerhoff, H. V; Dam, K.; Oliver, S. G.; van Dam, K. *Nat. Biotech.* **2001**, *19*, 45.

(46) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155.

(47) Dolinski, K.; Botstein, D. *Genome Res.* **2005**, *15*, 1611.

(48) Sauer, U.; Heinemann, M.; Zamboni, N. *Science.* **2007**, *316*, 550.

(49) Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837.

(50) Tweeddale, H.; Notley-McRobb, L.; Ferenci, T. J. *J. Bacteriol.* **1998**, *180*, 5109.

(51) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. *Trends Biotechnol.* **1998**, *16*, 373–378.

(52) Tweeddale, H.; Notley-McRobb, L.; Ferenci, T. *Redox Rep.* **1999**, *4*, 237.

(53) Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica* **1999**, *29*, 1181.

(54) Warne, M. A.; Lenz, E. M.; Osborn, D.; Weeks, J. M.; Nicholson, J. K. *Biomarkers* **2000**, *5*, 56.

(55) Robertson, D. G.; Reily, M. D.; Sigler, R. E.; Wells, D. F.; Paterson, D. A.; Braden, T. K. *Toxicol. Sci.* **2000**, *57*, 326.

(56) Nicholls, A. W.; Nicholson, J. K.; Haselden, J. N.; Waterfield, C. J. *Biomarkers* **2000**, *5*, 410.

(57) Griffin, J. L.; Walker, L. A.; Garrod, S.; Holmes, E.; Shore, R. F.; Nicholson, J. K. *Comp. Biochem. Physiol., Part B Biochem. Mol. Biol.* **2000**, *127*, 357.

(58) Gavaghan, C. L.; Holmes, E.; Lenz, E.; Wilson, I. D.; Nicholson, J. K. *FEBS Lett.* **2000**, *484*, 169.

(59) Fiehn, O.; Kopka, J.; Dörmann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. *Nat. Biotech.* **2000**, *18*, 1157.

(60) Potts, B. C.; Deese, A. J.; Stevens, G. J.; Reily, M. D.; Robertson, D. G.; Theiss, J. *J. Pharm. Biomed. Anal.* **2001**, *26*, 463.

(61) Keun, H. C.; Ebbels, T. M. D.; Antti, H.; Bollard, M. E.; Beckonert, O.; Schlotterbeck, G.; Senn, H.; Niederhauser, U.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Chem. Res. Toxicol.* **2002**, *15*, 1380.

(62) Brindle, J. T.; Antti, H.; Holmes, E.; Tranter, G.; Nicholson, J. K.; Bethell, H. W.; Clarke, S.; Schofield, P. M.; McKilligin, E.; Mosedale, D. E.; Grainger, D. J. *Nat. Med.* **2002**, *8*, 1439.

(63) Wong, A.; Jiménez, B.; Li, X.; Holmes, E.; Nicholson, J. K.; Lindon, J. C.; Sakellariou, D. *Anal. Chem.* **2012**, *84*, 3843.

(64) Fan, T. W.; Lane, A. N.; Higashi, R. M. *Russ. J. Plant Physiol.* **2003**, *50*, 879.

(65) Shaniah, N.; Desilva, M. A.; Gowda, G. A. N.; Raferty, M. A.; Hainline, B. E.; Raferty, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11540.

(66) Viant, M. R. *Biochem. Biophys. Res. Commun.* **2003**, *310*, 943.

(67) Aharoni, A.; Ric de Vos, C.; Verhoeven, H.; Maliepaard, C.; Kruppa, G.; Bino, R.; Goodenowe, D. *Omics* **2002**, *6*, 217.

(68) Goodacre, R.; Vaidyanathan, S.; Bianchi, G.; Kell, D. B. *Analyst* **2002**, *127*, 1457.

(69) Kind, T.; Fiehn, O. *BMC Bioinf.* **2006**, *7*, 234.

(70) Han, J.; Danell, R. M.; Patel, J. R.; Gumerov, D. R.; Scarlett, C. O.; Speir, J. P.; Parker, C. E.; Rusyn, I.; Zeisel, S.; Borchers, C. H. *Metabolomics* **2008**, *4*, 128–140.

(71) Kumari, S.; Stevens, D.; Kind, T.; Denkert, C.; Fiehn, O. *Anal. Chem.* **2011**, *83*, 5895–5902.

(72) Tolstikov, V.; Fiehn, O. *Anal. Biochem.* **2002**, *301*, 298.

(73) Plumb, R. S.; Stumpf, C. L.; Gorenstein, M. V; Castro-Perez, J. M.; Dear, G. J.; Anthony, M.; Sweatman, B. C.; Connor, S. C.; Haselden, J. N. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1991.

(74) Soga, T.; Ueno, Y.; Naraoka, H.; Ohashi, Y.; Tomita, M.; Nishioka, T. *Anal. Chem.* **2002**, *74*, 2233.

(75) Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E. M. *Nat. Rev. Drug Discov.* **2002**, *1*, 153.

(76) Weckwerth, W.; Fiehn, O. *Curr. Opin. Biotechnol.* **2002**, *13*, 156.

(77) Robinson, A. B.; Robinson, N. E. *Methods Mol. Biol.* **2011**, *708*, 1.

(78) Beckmann, M.; Parker, D.; Enot, D. P.; Duval, E.; Draper, J. *Nat. Protoc.* **2008**, *3*, 486–504.

(79) Lu, W.; Bennett, B. D.; Rabinowitz, J. D. *J. Chromatogr. B.* **2008**, *871*, 236.

(80) He, Y.; Yu, Z.; Giegling, I.; Xie, L.; Hartmann, A. M.; Prehn, C.; Adamski, J.; Kahn, R.; Li, Y.; Illig, T.; Wang-Sattler, R.; Rujescu, D. *Transl. Psychiatry* **2012**, *2*, e149.

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

(82) Abaffy, T.; Duncan, R.; Riemer, D. D.; Tietje, O.; Elgart, G.; Milikowski, C.; DeFazio, R. A. *PLoS One* **2010**, *5*, e13813.

(83) Mondello, L.; Tranchida, P. Q.; Dugo, P.; Dugo, G. *Mass Spectrom. Rev.* **2008**, *27*, 101.

(84) Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. *Analyst* **2012**, *137*, 293.

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

(86) Bedair, M.; Sumner, L. W. *TrAC, Trends Anal. Chem.* **2008**, *27*, 238.

(87) Patti, G. J. *J. Sep. Sci.* **2011**, *34*, 3460.

(88) Issaq, H. J.; Abbott, E.; Veenstra, T. D. *J. Sep. Sci.* **2008**, *31*, 1936.

(89) Novotny, M. V; Soini, H. A.; Mechref, Y. *J. Chromatogr. B* **2008**, *866*, 26.

(90) Koek, M. M.; Jellema, R. H.; van der Greef, J.; Tas, A. C.; Hankemeier, T. *Metabolomics* **2011**, *7*, 30.

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

(92) Wang, X.; Sun, H.; Zhang, A.; Wang, P.; Han, Y. *J. Sep. Sci.* **2011**, *34*, 3451.

(93) Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D.; Theodoridis, G. *J. Sep. Sci.* **2010**, *33*, 716–727.

(94) Pham-Tuan, H.; Kaskavelis, L.; Daykin, C.; Janssen, H. *J. Chromatogr. B* **2003**, *789*, 283.

(95) Britz-McKibbin, P. *Methods Mol. Biol.* **2011**, *708*, 229.

(96) Ramautar, R.; Mayboroda, O. A.; Somsen, G. W.; de Jong, G. J. *Electrophoresis* **2011**, *32*, 52.

(97) Williams, R. J. *Chem. Eng. News* **1947**, *25*, 1112.

(98) Horning, E.; Horning, M. G. *J. Chromatogr. Sci.* **1971**, *9*, 129.

(99) Garrod, A. E. *Lancet* **1902**, *2*, 1616.

(100) Williams, R. J.; Berry, L. J.; Beerstecher, E. J. *Proc. Natl. Acad. Sci. U.S.A.* **1949**, *35*, 265.

(101) Shostenkot, Y. V.; Georgievskii, V. P.; Levin, M. G. *Anal. Chem.* **2000**, *55*, 905.

(102) Zakaria, M.; Gonnord, M.-F.; Guiochon, G. *J. Chromatogr.* **1983**, *27*, 127.

(103) Williams, R. J.; Kirby, H. *Science.* **1948**, *107*, 481.

(104) Armstrong, M. D.; Shaw, K. F.; Wall, P. E. *J. Biol. Chem.* **1956**, *218*, 293.

(105) Gates, S. C.; Sweeley, C. C. *Clin. Chem.* **1978**, *24*, 1663.

(106) Hamilton, P. B. *Anal. Chem.* **1963**, *35*, 2055.

(107) Pitt, W. W. J.; Scott, S. D.; Johnson, W. F.; Jones, G. J. *Clin. Chem.* **1970**, *16*, 657.

(108) Majors, R. E. *Anal. Chem.* **1973**, *45*, 755.

(109) Horváth, C. S.; Lipsky, S. R. *J. Chromat. Sci.* **1969**, *7*, 109.

(110) Molnar, I.; Horvath, C. *J. Chromatogr.* **1977**, *143*, 391.

(111) Persson, B. A.; Karger, B. L. *J. Chromatogr. Sci.* **1974**, *12*, 521.

(112) Wittmer, D. P.; Nuessle, N. O.; Haney, W. G. J. *Anal. Chem.* **1975**, *47*, 1422.

(113) Horvath, C.; Melander, W.; Molnar, I.; Molnar, P. *Anal. Chem.* **1977**, *49*, 2295.

(114) McWilliam, I. G.; Dewar, R. A. *Nature* **1958**, *181*, 760.

(115) Lovelock, J. E.; Lipsky, S. R. *J. Am. Chem. Soc.* **1960**, *82*.

(116) Gohlke, R. S. *Anal. Chem.* **1959**, *31*, 535.

(117) Ryhage, R. *Anal. Chem.* **1964**, *36*, 759.

(118) Watson, J. T.; Bieman, K. *Anal. Chem.* **1964**, *36*, 1135.

(119) Pauling, L.; Robinson, A. R.; Teranishi, R.; Cary, P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 2374.

(120) Dalgliesh, C. E.; Horning, E. C.; Horning, M. G.; Knox, K. L.; Yarger, K. *Biochem. J.* **1966**, *101*, 792.

(121) Stalling, D. L.; Gehrke, C. W.; Zumwalt, R. W. *Biochem. Biophys. Res. Comm.* **1968**, *31*, 616.

(122) Horning, E.; Horning, M. G. *Methods Med. Res.* **1970**, *12*, 369.

(123) Jellum, E.; Stokke, O.; Eldjarn, L. *Clin. Chem.* **1972**, *18*, 800.

(124) Gaffney, T. E.; Hammar, G. C.; Holmstedt, B.; McMahon, R. E. *Anal. Chem.* **1971**, *43*, 307.

(125) Kovats, E. *Helv. Chim. Acta* **1958**, *41*, 1915.

(126) Sauter, A. D.; Betowski, L. D.; Smith, T. R.; Strickler, V. A.; Beimer, R. G.; Colby, B. N.; Wilkinson, J. E. *J. High Res. Chromatogr. Chromatogr. Comm.* **1981**, *4*, 366.

(127) Settlage, J.; Jaeger, H. *J. Chromatogr. Sci.* **1984**, *22*, 192.

(128) Chamberlin, B. A.; Sweeley, C. C. *Clin. Chem.* **1987**, *33*, 572.

(129) Kuhara, T; Shinka, T.; Inoue, Y.; Ohse, M.; Zhen-wei, X.; Yoshida, I.; Inokuchi, T.; Yamaguchi, S.; Takayanagi, M, Matsumoto, I. *J. Chromatogr. B. Biomed. Sci. Appl.* **1999**, *731*, 141.

(130) A. Tiselius. *Ph.D. Thesis. The moving boundary method of studying the electrophoresis of proteins*; Almqvist & Wiksell: University of Uppsala, Sweden, 1930.

(131) Mikkers, F. E. P.; Everaerts, F. M.; Verheggen, T. P. E. M. *J. Chromatogr.* **1979**, *169*, 11.

(132) Jorgenson, J. W.; Lukacs, K. D. *Clin. Chem.* **1981**, *53*, 1298.

(133) Jorgenson, J. W.; Lukacs, K. D. *Clin. Chem.* **1981**, *27*, 1551.

(134) Pretorius, V.; Hopkins, B. J.; Schieke, J. D. *J. Chromatogr. A* **1974**, *99*, 23.

(135) Terabe, S.; Otsuka, K.; Ichikawa, K.; Tsuchiya, A.; Ando, T. *Anal. Chem.* **1984**, *56*, 111.

(136) Terabe, S.; Ozaki, H.; Otsuka, K.; Ando, T. *J. Chromatogr. A* **1985**, *332*, 211.

(137) Horning, E.; Carroll, D.; Dzidic, I.; Haegele, K.; Horning, M.; Stillwell, R. *J. Chromatogr. Sci.* **1974**, *12*, 725.

(138) Willoughby, R. C.; Browner, R. F. *Anal. Chem.* **1984**, *56*, 2626.

(139) Vestal, M. L.; Fergusson, G. J. *Anal. Chem.* **1985**, *57*, 2378.

(140) Caprioli, M.; Fan, T.; Cottrell, J. S. *Anal. Chem.* **1986**, *58*, 2949.

(141) Bruins, A. P.; Covey, T. R.; Henion, J. D. *Anal. Chem.* **1987**, *59*, 2642.

(142) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science.* **1989**, *246*, 64.

(143) Rashed, M. S.; Bucknall, M. P.; Little, D.; Awad, A.; Jacob, M.; Alamoudi, M.; Alwattar, M.; Ozand, P. T. *Clin. Biochem. Rev.* **1997**, *43*, 1129.

(144) Chace, D. H.; Kalas, T. A.; Naylor, E. W. *Clin. Chem.* **2003**, *49*, 1797.

(145) Shushan, B. *Mass Spectrom. Rev.* **2010**, *29*, 930.

(146) Maxwell, E. J.; Chen, D. D. *Anal. Chim. Acta* **2008**, *627*, 25–33.

(147) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1988**, *60*, 1948.

(148) Nicholson, J. K.; Buckingham, M. J.; Sadler, P. J. *Biochem. J.* **1983**, *211*, 605.

(149) Bales, J. R.; Higham, D. P.; Howe, I.; Nicholson, J. K.; Sadler, P. J. *Clin. Chem.* **1994**, *30*, 426.

(150) Cloarec, O.; Campbell, A.; Tseng, L. H.; Braumann, U.; Spraul, M.; Scarfe, G.; Weaver, R.; Nicholson, J. K. *Anal. Chem.* **2007**, *79*, 3304.

(151) Jayawickrama, D.; Sweedler, J. *J Chromatogr. A.* **2003**, *1000*, 819.

(152) Williams, R.; Lenz, E.; Wilson, A.; Granger, J.; Wilson, I.; Major, H.; Stumpf, C.; Plumb, R. *Mol. BioSyst.* **2006**, *2*, 174.

(153) Kind, T.; Tolstikov, V.; Fiehn, O.; Weiss, R. H. *Anal. Biochem.* **2007**, *363*, 185.

(154) Ong, E. S.; Chor, C. F.; Zou, L.; Ong, C. N. *Mol. BioSyst.* **2009**, *5*, 288.

(155) Nevedomskaya, E.; Mayborada, O. A.; Deelder, A. M. *Mol. BioSyst.* **2011**, *7*, 3214.

(156) Kusano, M.; Redestig, H.; Hirai, T.; Oikawa, A.; Matsuda, F.; Fukushima, A.; Arita, M.; Watanabe, S.; Yano, M.; Hiwasa-Tanase, K.; Ezura, H.; Saito, K. *PLoS One* **2011**, *6*, e16989.

(157) Schlotterbeck, G.; Ceccarelli, S. M. *Bioanalysis* **2009**, *1*, 549.

(158) Strain, H. H. *J. Chem. Educ.* **1955**, *32*, 601.

(159) Giddings, J. C. *Sep. Sci.* **1969**, *4*, 181.

(160) Bowser, M. T.; Bebault, G. M.; Peng, X.; Chen, D. D. Y. *Electrophoresis* **1997**, *18*, 2928.

(161) Potts, L.; Stoll, D.; Li, X.; Carr, P. *J. Chromatogr. A* **2010**, *1217*, 5700.

(162) Van Deemter, J. J.; Zuiderweg, F. J.; Klinkenberg, A. *Chem. Eng. Sci.* **1956**, *5*, 271.

(163) Miyabe, K.; Okada, A. *Analyst* **2002**, *127*, 1420.

(164) Castello, G.; Moretti, P.; Vezzani, S. *J. Chromatogr. A* **2009**, *1216*, 1607.

(165) Schellinger, A. P.; Stoll, D. R.; Carr, P. W. *J. Chromatogr. A* **2005**, *1064*, 143.

(166) Gika, H. G.; Macpherson, E.; Theodoridis, G. A.; Wilson, I. D. *J. Chromatogr. B* **2008**, *871*, 299.

(167) Anderson, J. L.; Armstrong, D. W. *Anal. Chem.* **2005**, *77*, 6453.

(168) Poole, C. F.; Poole, S. K. *J. Sep. Sci.* **2011**, *34*, 888.

(169) Rocha, S. M.; Caldeira, M.; Carrola, J.; Santos, M.; Cruz, N.; Duarte, I. F. *J. Chromatogr. A* **2012**, *1252*, 155.

(170) Wilson, R. B.; Siegler, W. C.; Hoggard, J. C.; Fitz, B. D.; Nadeau, J. S.; Synovec, R. E. *J. Chromatogr. A* **2011**, *1218*, 3130.

(171) Fekete, S.; Oláh, E.; Fekete, J. *J. Chromatogr. A* **2012**, *1228*, 57.

(172) Gritti, F.; Guiochon, G. *J. Chromatogr. A* **2012**, *1228*, 2.

(173) Minakuchi, H.; Nakanishi, K.; Soga, N.; Ishizuka, N.; Tanaka, N. *Anal. Chem.* **1996**, *68*, 3498.

(174) Tanaka, N.; Kobayashi, H.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Hosoya, K.; Ikegami, T. *J. Chromatogr. A* **2002**, *965*, 35.

(175) Kobayashi, H.; Ikegami, T.; Kimura, H.; Hara, T.; Tokuda, D.; Tanaka, N. *Anal. Sci.* **2006**, *22*, 491.

(176) Ikegami, T.; Dicks, E.; Kobayashi, H.; Morisaka, H.; Tokuda, D.; Cabrera, K.; Hosoya, K.; Tanaka, N. *J. Sep. Sci.* **2004**, *27*, 1292.

(177) Guiochon, G. *J. Chromatogr. A* **2007**, *1168*, 101.

(178) Unger, K. K.; Jilge, O.; Kinkel, J. N.; Hearn, M. T. W. *J. Chromatogr. A* **1986**, *359*, 61.

(179) Wilson, I. D.; Nicholson, J. K.; Castro-Perez, J.; Granger, J. H.; Johnson, K. A.; Smith, B. W.; Plumb, R. S. *J. Proteome Res.* **2005**, *4*, 591–598.

(180) MacNair, J. E.; Lewis, K. C.; Jorgenson, J. W. *Anal. Chem.* **1997**, *69*, 983.

(181) MacNair, J. E.; Patel, K. D.; Jorgenson, J. W. *Anal. Chem.* **1999**, *71*, 700.

(182) Mellors, J. S.; Jorgenson, J. W. *Anal. Chem.* **2004**, *76*, 5441.

(183) Anspach, J. A.; Maloney, T. D.; Colón, L. A. *J. Sep. Sci.* **2007**, *31*, 1207.

(184) Evans, A. M.; DeHaven, C. D.; Barrett, T.; Mitchell, M.; Milgram, E. *Anal. Chem.* **2009**, *81*, 6656–6667.

(185) Uehara, T.; Yokoi, A.; Aoshima, K.; Tanaka, S.; Kadowaki, T.; Tanaka, M.; Oda, Y. *Anal. Chem.* **2009**, *81*, 3836.

(186) Greibrokk, T.; Andersen, T. *J. Chromatogr. A* **2003**, *1000*, 743.

(187) Plumb, R. S.; Rainville, P.; Smith, B. W.; Johnson, K. A.; Castro-Perez, J.; Wilson, I. D.; Nicholson, J. K. *Anal. Chem.* **2006**, *78*, 7278.

(188) Plumb, R.; Mazzeo, J. R.; Grumbach, E. S.; Rainville, P.; Jones, M.; Wheat, T.; Neue, U. D.; Smith, B.; Johnson, K. A. *J. Sep. Sci.* **2007**, *31*, 1158.

(189) Kirkland, J. J.; Henderson, J. W.; DeStefano, J. J.; van Straten, M. A.; Claessens, H. A. *J. Chromatogr. A* **1997**, *762*, 97.

(190) Hupp, A. M.; McGuffin, V. L. *J. Chromatogr. A* **2010**, *1217*, 6241.

(191) Cunliffe, J. M.; Maloney, T. D. *J. Sep. Sci.* **2007**, *30*, 3104.

(192) Brice, R.; Zhang, X.; Colón, L. *J. Sep. Sci.* **2009**, *32*, 2723.

(193) Wilson, I.; Plumb, R.; Granger, J.; Major, H.; Williams, R.; Lenz, E. *J. Chromatogr. B* **2005**, *817*, 67.

(194) Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. *J. Sep. Sci.* **2008**, *31*, 1598.

(195) Zhang, Y.; Carr, P. *J. Chromatogr. A* **2011**, *1218*, 763.

(196) Liu, X.; Pohl, C.; Woodruff, A.; Chen, J. *J. Chromatogr. A* **2011**, *1218*, 3407.

(197) Dorsey, J. G.; Dill, K. A. *Chem. Rev.* **1989**, *89*, 331.

(198) Vailaya, A.; Horváth, C. S. *J. Phys. Chem. B* **1997**, *101*, 5875.

(199) Héberger, K. *J. Chromatogr. A* **2007**, *1158*, 273.

(200) Snyder, L.; Dolan, J.; Carr, P. *J. Chromatogr. A* **2004**, *1060*, 77.

(201) Zhang, Y.; Carr, P. W. *J. Chromatogr. A* **2009**, *1216*, 6685.

(202) Rafferty, J.; Zhang, L.; Siepmannm, J.; Schure, M. *Anal. Chem.* **2007**, *79*, 6551.

(203) Gómez-Romero, M.; Segura-Carretero, A.; Fernández-Gutiérrez, A. *Phytochemistry* **2010**, *71*, 1848.

(204) Shortreed, M.; Lamos, S.; Frey, B.; Phillips, M.; Patel, M.; Belshaw, P.; Smith, L. *Anal. Chem.* **2006**, *78*, 6398.

(205) Santa, T. *Biomed. Chromatogr.* **2011**, *25*, 1.

(206) Annesley, T. M. *Clin. Chem.* **2003**, *49*, 1041.

(207) Mallet, C. R.; Lu, Z.; Mazzeo, J. R. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 49.

(208) Coulier, L.; Bas, R.; Jespersen, S.; Verheij, E.; van der Werf, M. J.; Hankemeier, T. *Anal. Chem.* **2006**, *78*, 6573.

(209) Kiefer, P.; Delmotte, N.; Vorholt, J. A. *Anal. Chem.* **2011**, *83*, 850.

(210) Alpert, A. J. *J. Chromatogr. A* **1990**, *499*, 177.

(211) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. *Mass Spectrom. Rev.* **2010**, *29*, 671.

(212) Jandera, P. *J. Sep. Sci.* **2008**, *31*, 1421.

(213) Rappold, B. A.; Grant, R. P. *J. Sep. Sci.* **2011**, *34*, 3527.

(214) Nguyen, H. P.; Yang, S. H.; Wigginton, J. G.; Simpkins, J. W.; Schug, K. A. *J. Sep. Sci.* **2010**, *33*, 793.

(215) Yang, S.; Sadilek, M.; Lidstrom, M. E. *J. Chromatogr. A* **2010**, *1217*, 7401.

(216) Pesek, J.; Matyska, M. *J. Sep. Sci.* **2005**, *28*, 1845.

(217) Pesek, J.; Matyska, M.; Loo, J.; Fischer, S.; Sana, T. *J. Sep. Sci.* **2009**, *32*, 2200.

(218) Callahan, D.; De Souza, D.; Bacic, A.; Roessner, U. *J. Sep. Sci.* **2009**, *32*, 2273.

(219) Zhang, T.; Creek, D.; Barrett, M.; Blackburn, G.; Watson, D. *Anal. Chem.* **2012**, *84*, 1994.

(220) Frost, N. W.; Jing, M.; Bowser, M. T. *Anal. Chem.* **2010**, *82*, 4682.

(221) Zusková, I.; Novotná, A.; Vceláková, K.; Gas, B. *J. Chromatogr. B* **2006**, *841*, 129.

(222) Gas, B.; Kenndler, E. *Electrophoresis* **2000**, *21*, 3888.

(223) Hutterer, K. M.; Jorgenson, J. W. *Anal. Chem.* **1999**, *71*, 1293–1297.

(224) Beckers, J. L.; Bocek, P. *Electrophoresis* **2003**, *24*, 518.

(225) Geiser, L.; Veuthey, J. L. *Electrophoresis* **2009**, *30*, 36.

(226) Amundsen, L. K.; Kokkonen, J. T.; Sirén, H. *J. Sep. Sci.* **2008**, *31*, 803.

(227) Hommerson, P.; Khan, A.; de Jong, G.; Somsen, G. *J. Chromatogr. A* **2008**, *1204*, 197.

(228) Henley, W.; Wilburn, R.; Crouch, A.; Jorgenson, J. *Anal. Chem.* **2005**, *77*, 7024.

(229) Yeung, K.; Lucy, C. *Electrophoresis* **1999**, *20*, 2554.

(230) Knox, J. H.; Grant, I. H. *Chromatographia* **1987**, *24*, 135.

(231) Terabe, S. *Anal. Chem.* **2004**, *76*, 241A.

(232) Bowser, M.; Chen, D. *Electrophoresis* **1998**, *19*, 383.

(233) Klampfl C.W. *J. Chromatogr. A* **2004**, *1044*, 131.

(234) Malkin, D. S.; Wei, B.; Fogiel, A. J.; Staats, S. L.; Wirth, M. J. *Anal. Chem.* **2010**, *82*, 2175.

(235) Svec, F. *Electrophoresis* **2009**, *30*, S68.

(236) Kaiser, C.; Segui-Lines, G.; D'Amaral, J. C.; Ptolemy, A. S.; Britz-McKibbin, P. *Chem. Comm.* **2008**, 338.

(237) Fei, F.; Britz-McKibbin, P. *Anal. Bioanal. Chem.* **2010**, *398*, 1349.

(238) Britz-McKibbin, P; Chen, D. *J. Chromatogr. A* **1997**, *781*, 23.

(239) Terabe, S.; Otsuka, K.; Ando, T. *Anal. Chem.* **1985**, *57*, 834.

(240) Nishi, H.; Fukuyama, T.; Terabe, S. *J. Chromatogr. A* **1991**, *553*, 503.

(241) Espada, A.; Molina-Martin, M. *Drug Discov. Today* **2012**, *17*, 396.

(242) Dovichi, N. J.; Zhang, J. *Angew. Chem. Int. Ed.* **2000**, *39*, 4463–4468.

(243) Sugimoto, M.; Hirayama, A.; Ishikawa, T.; Robert, M.; Baran, R.; Uehara, K.; Kawai, K.; Soga, T.; Tomita, M. *Metabolomics* **2010**, *6*, 27–41.

(244) Jumppanen, J. H.; Riekkola, M.-L. *Anal. Chem.* **1995**, *67*, 1060.

(245) Shalaeva, M.; Kenseth, J.; Lombardo, F.; Bastin, A. *J. Pharm. Sci.* **2008**, *97*, 2581.

(246) Colton, I.; Carbeck, J.; Rao, J.; Whitesides, G. M. *Electrophoresis* **1998**, *19*, 367.

(247) Gavina, J.; Mazhab-Jafari, M.; Melacini, G.; Britz-McKibbin, P. *Biochemistry* **2009**, *48*, 223.

(248) Yang, J.; Bose, S.; Hage, D. S. *J. Chromatogr. A* **1996**, *735*, 209.

(249) Sugimoto, M.; Hirayama, A.; Robert, M.; Abe, S.; Soga, T.; Tomita, M. *Electrophoresis* **2010**, *31*, 2311–2318.

(250) Ramautar, R.; Mayboroda, O. A.; Derks, R. J.; van Nieuwkoop, C.; van Dissel, J. T.; Somsen, G. W.; Deelder, A. M.; de Jong, G. J. *Electrophoresis* **2008**, *29*, 2714.

(251) Petersen, N. J.; Hansen, S. H. *Electrophoresis* **2012**, *33*, 1021.

(252) Ohashi, Y.; Hirayama, A.; Ishikawa, T.; Nakamura, S.; Shimizu, K.; Ueno, Y.; Tomita, M.; Soga, T. *Mol. BioSyst.* **2008**, *4*, 135.

(253) Kuehnbaum, N. L.; Britz-McKibbin, P. *Anal. Chem.* **2011**, *83*, 8063.

(254) Williams, B.; Cameron, C.; Workman, R.; Broeckling, C.; Sumner, L.; Smith, J. *Electrophoresis* **2007**, *28*, 1371.

(255) Chalcraft, K. R.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 307–314.

(256) Soga, T.; Igarashi, K.; Itoh, C.; Mizobuchi, K.; Zimmermann, H.; Tomita, M. *Anal. Chem.* **2009**, *81*, 6165.

(257) Büscher, J. M.; Czernik, D.; Ewald, J. C.; Sauer, U.; Zamboni, N. *Anal. Chem.* **2009**, *81*, 2135.

(258) Allwood, J. W.; Erban, A.; de Koning, S.; Dunn, W. B.; Luedemann, A.; Lommen, A.; Kay, L.; Löscher, R.; Kopka, J.; Goodacre, R. *Metabolomics* **2009**, *5*, 479.

(259) Benton, H.; Want, E.; Keun, H.; Amberg, A.; Plumb, R.; Goldfain-Blanc, F.; Walther, B.; Reily, M.; Lindon, J.; Holmes, E.; Nicholson, J.; Ebbels, T. *Anal. Chem.* **2012**, *84*, 2424.

(260) Ward, J. L.; Baker, J. M.; Miller, S. J.; Deborde, C.; Maucourt, M.; Biais, B.; Rolin, D.; Moing, A.; Moco, S.; Vervoort, J.; Lommen, A.; Schäfer, H.; Humpfer, E.; Beale, M. H. *Metabolomics* **2010**, *6*, 263.

(261) Soga, T.; Heiger, D. N. *Anal. Chem.* **2000**, *72*, 1236.

(262) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 2506–2515.

(263) Kok, M. G. M.; de Jong, G. J.; Somsen, G. W. *Electrophoresis* **2011**, *32*, 3016.

(264) Hirayama, A.; Tomita, M.; Soga, T. *Analyst* **2012**, *137*, 5026.

(265) Ramautar, R.; Busnel, J.-M.; Deelder, A. M.; Mayboroda, O. A. *Anal. Chem.* **2012**, *84*, 885–892.

(266) Maxwell, E. J.; Zhong, X.; Chen, D. D. *Anal. Chem.* **2010**, *82*, 8377–8381.

(267) Wu, Z.; Gao, W.; Phelps, M.; Wu, D.; Miller, D.; Dalton, J. *Anal. Chem.* **2004**, *76*, 839.

(268) Yanes, O.; Tautenhahn, R.; Patti, G.; Siuzdak, G. *Anal. Chem.* **2011**, *83*, 2152.

(269) Zhang, X.; Clausen, M.; Zhao, X.; Zheng, H.; Bertram, H. *Anal. Chem.* **2012**, *84*, 7785.

(270) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2002**, *20*, 362.

(271) O'Farrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007.

(272) Rabilloud, T.; Chevallet, M.; Luche, S.; Lelong, C. *J. Proteomics* **2010**, *73*, 2064.

(273) Dickerson, J.; Ramsay, L.; Dada, O.; Cermak, N.; Dovichi, N. *Electrophoresis* **2010**, *31*, 2650.

(274) Wang, T.; Ma, J.; Wu, S.; Sun, L.; Yuan, H.; Zhang, L.; Liang, Z.; Zhang, Y. *J. Chromatogr. B* **2011**, *879*, 804.

(275) Evans, C.; Jorgenson, J. *Anal. Bioanal. Chem.* **2004**, *378*, 1952.

(276) Jia, L.; Tanaka, N.; Terabe, S. *Electrophoresis* **2005**, *26*, 3468.

(277) Metz, T. O.; Page, J. S.; Baker, E. S.; K.Tang; Ding, J.; Shen, Y.; Smith, R. D. *Trends Anal. Chem.* **2008**, *27*, 205.

(278) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1005.

(279) Koek, M. M.; Muilwijk, B.; van Stee, L. L. P.; Hankemeier, T. *J. Chromatogr. A* **2008**, *1186*, 420.

(280) Koek, M.; van der Kloet, F.; Kleemann, R.; Kooistra, T.; Verheij, E.; Hankemeier, T. *Metabolomics* **2011**, *7*, 1.

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

(282) Aberg, K. M.; Alm, E.; Torgrip, R. J. *Anal. Bioanal. Chem.* **2009**, *394*, 151.

(283) Junot, C.; Madalinski, G.; Tabet, J.; Ezan, E. *Analyst* **2010**, *135*, 2203.

(284) Chen, H.; Pan, Z.; Talaty, N. N.; Cooks, R. G.; Raftery, D. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1577.

(285) Dumas, M.-E.; Maibaum, E.; Teague, C.; Ueshima, H.; Zhou, B.; Lindon, J.; Nicholson, J.; Stamler, J.; Elliott, P.; Chan, Q.; Holmes, E. *Anal. Chem.* **2006**, *78*, 2199.

(286) Lin, L.; Yu, Q.; Yan, X.; Hang, W.; Zheng, J.; Xing, J.; Huang, B. *Analyst* **2010**, *135*, 2970.

(287) Stahnke, H.; Kittlaus, S.; Kempe, G.; Hemmerling, C.; Alder, L. *J. Mass Spectrom.* **2012**, *47*, 875–884.

(288) Hinterwirth, H.; Lämmerhofer, M.; Preinerstorfer, B.; Gargano, A.; Reischl, R.; Bicker, W.; Trapp, O.; Brecker, L.; Lindner, W. *J. Sep. Sci.* **2010**, *33*, 3273.

(289) Giuffrida, A.; León, C.; García-Cañas, V.; Cucinotta, V.; Cifuentes, A. *Electrophoresis* **2009**, *30*, 1734.

(290) Waldhier, M. C.; Almstetter, M. F.; Nürnberger, N.; Gruber, M. A.; Dettmer, K.; Oefner, P. J. *J. Chromatogr. A* **2011**, *1218*, 4537.

(291) Oglesbee, D.; Sanders, K.; Lacey, J.; Magera, M.; Casetta, B.; Strauss, K.; Tortorelli, S.; Rinaldo, P.; Matern, D. *Clin. Chem.* **2008**, *54*, 542.

(292) Wang, M.; Bai, J.; Chen, W.; Ching, C. *PLoS One* **2010**, *5*, e15441.

(293) Stalcup, A. M. *Annu. Rev. Anal. Chem.* **2010**, *3*, 341.

(294) Herrero, M.; Ibáñez, E.; Martín-Alvarez, P.; Cifuentes, A. *Anal. Chem.* **2007**, *79*, 5071.

(295) King, R.; Bonfiglio, R.; Fernandez-Metzler, C.; Miller-Stein, C.; Olah, T. *J. Am. Soc. Mass. Spectrom.* **2000**, *11*, 942.

(296) Kostiainen, R.; Kauppila, T. J. *J. Chromatogr. A* **2009**, *1216*, 685.

(297) Remane, D.; Meyer, M.; Wissenbach, D.; Maurer, H. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 3103.

(298) León-González, Z.; Ferreiro-Vera, C.; Priego-Capote, F.; de Castro, M. D. L. *J. Chromatogr. A* **2011**, *1218*, 3013.

(299) Vuckovic, D.; Pawliszyn, J. *Anal. Chem.* **2011**, *83*, 1944.

(300) Risticevic, S.; DeEll, J.; Pawliszyn, J. *J. Chromatogr. A* **2012**, *1251*, 208.

(301) Jain, A.; Verma, K. K. *Anal. Chim. Acta* **2011**, *706*, 37.

(302) Lancas, F.; Queiroz, M.; Grossi, P.; Olivares, I. *J. Sep. Sci.* **2009**, *32*, 813.

(303) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403–415.

(304) Dwivedi, P.; Schultz, A. J.; Hill, H. H. *Int. J. Mass Spectrom.* **2010**, *298*, 78.

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

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

(307) Matsuda, F.; Shinbo, Y.; Oikawa, A.; Hirai, M.; Fiehn, O.; Kanaya, S.; Saito, K. *PLoS One* **2009**, *7*, e7490.

(308) Kilgour, D. P.; Mackay, C. L.; Langridge-Smith, P. R.; O'Connor, P. B. *Anal. Chem.* **2012**, *84*, 7431.

(309) Weber, R. J. M.; Southam, A. D.; Sommer, U.; Viant, M. R. *Anal. Chem.* **2011**, *83*, 3737.

(310) Jankevics, A.; Merlo, M.; de Vries, M.; Vonk, R.; Takano, E.; Breitling, R. *Metabolomics* **2012**, *8*, 29–36.

(311) Stein, S. *Anal. Chem.* **2012**, *84*, 7274.

(312) Courant, F.; Royer, A. L.; Chéreau, S.; Morvan, M. L.; Monteau, F.; Antignac, J. P.; Le Bizec, B. *Analyst* **2012**, *137*, 4958.

(313) Kind, T.; Wohlgemuth, G.; Lee, D. Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O. *Anal. Chem.* **2009**, *81*, 10038.

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

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

(316) Wagner, C.; Sefkow, M.; Kopka, J. *Phytochemistry* **2003**, *62*, 887.

(317) Neumann, S.; Böcker, S. *Anal. Bioanal. Chem.* **2010**, *398*, 2779.

(318) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. *BMC Bioinformatics* **2010**, *22*, 148.

(319) Wishart, D. S. *Bioanalysis* **2011**, *3*, 1769.

(320) Hill, D. W.; Kertesz, T. M.; Fontaine, D.; Friedman, R.; Grant, D. F. *Anal. Chem.* **2008**, *80*, 5574.

(321) Kind, T.; Fiehn, O. *Bioanal. Rev.* **2010**, *2*, 23.

(322) Molnar, I. *J. Chromatogr. A* **2002**, *965*, 175.

(323) Hruska, V.; Jaros, M.; Gas, B. *Electrophoresis* **2006**, *27*, 984.

(324) Thormann, W.; Breadmore, M. C.; Caslavska, J.; Mosher, R. A. *Electrophoresis* **2010**, *31*, 726.

(325) Hagiwara, T.; Saito, S.; Ujiie, Y.; Imai, K.; Kakuta, M.; Kadota, K.; Terada, T.; Sumikoshi, K.; Shimizu, K.; Nishi, T. *Bioinformation* **2010**, *5*, 255.

(326) Albaugh, D.; Hall, L. L.; Hill, D.; Kertesz, T.; Parham, M.; Grant, D. *J. Chem. Inf. Model* **2009**, *49*, 788.

(327) Creek, D.; Jankevics, A.; Breitling, R.; Watson, D.; Barrett, M.; Burgess, K. *Anal. Chem.* **2011**, *83*, 8703.

(328) Boswell, P. G.; Schellenberg, J. R.; Carr, P. W.; Cohen, J. D.; Hegeman, A. D. *J. Chromatogr. A* **2011**, *1218*, 6732–6741.

(329) Boswell, P. G.; Schellenberg, J. R.; Carr, P. W.; Cohen, J. D.; Hegeman, A. D. *J. Chromatogr. A* **2011**, *1218*, 6742.

(330) Stein, S.; Babushok, V.; Brown, R.; Linstrom, P. *J. Chem. Inf. Model.* **2007**, *47*, 975.

(331) Schymanski, E. L.; Gallampois, C. M.; Krauss, M.; Meringer, M.; Neumann, S.; Schulze, T.; Wolf, S.; Brack, W. *Anal. Chem.* **2012**, *84*, 3287.

(332) Hall, L. M. L. H.; Kertesz, T. M.; Hill, D. W.; Sharp, T. R.; Oblak, E. Z.; Dong, Y. W.; Wishart, D. S.; Chen, M.-H.; Grant, D. F. *J. Chem. Inf. Model.* **2012**, *52*, 1222.

(333) Menikarachchi, L.; Cawley, S.; Hill, D.; Hall, L.; Lai, S.; Wilder, J.; Grant, D. *Anal. Chem.* **2012**, *84*, 9388.

(334) Fiehn, O. *TrAC, Trends Anal. Chem.* **2008**, *27*, 261.

(335) Kaal, E.; Janssen, H.-G. *J. Chromatogr. A* **2008**, *1184*, 43.

(336) Jung, H.-J. H.; Lee, W. W.-Y.; Yoo, Y. S. Y.; Chung, B. C. B.; Choi, M. H. M. *Clin. Chim. Acta* **2010**, *411*, 818.

(337) Chan, E. C. Y.; Pasikanti, K. K.; Nicholson, J. K. *Nat. Protoc.* **2011**, *6*, 1483.

(338) Dettmer, K.; Almstetter, M. F.; Appel, I. J.; Nürnberger, N.; Schlamberger, G.; Gronwald, W.; Meyer, H. H. D.; Oefner, P. J. *Electrophoresis* **2010**, *31*, 2365.

(339) Mohler, R. E.; Tu, B. P.; Dombek, K. M.; Hoggard, J. C.; Young, E. T.; Synovec, R. E. *J. Chromatogr. A* **2008**, *1186*, 401.

(340) Sellick, C. A.; Hansen, R.; Stephens, G. M.; Goodacre, R.; Dickson, A. J. *Nat. Protoc.* **2011**, *6*, 1241.

(341) Mal, M.; Koh, P. K.; Cheah, P. Y.; Chan, E. C. Y. *Anal. Bioanal. Chem.* **2012**, *403*, 483.

(342) Warren, C. R.; Aranda, I.; Cano, F. J. *Metabolomics* **2011**, *8*, 186.

(343) Butler, J.; Mishur, R.; Bokov, A.; Hakala, K.; Weintraub, S.; Rea, S. *PLoS One* **2012**, *7*, e46140.

(344) McKelvie, J. R.; Yuk, J.; Xu, Y.; Simpson, A. J.; Simpson, M. J. *Metabolomics* **2009**, *5*, 84.

(345) Begley, P.; Francis-McIntyre, S.; Dunn, W. B.; Broadhurst, D. I.; Halsall, A.; Tseng, A.; Knowles, J.; Goodacre, R.; Kell, D. B. *Anal. Chem.* **2009**, *81*, 7038.

(346) Smart, K. F.; Aggio, R. B. M.; van Houtte, J. R.; Villas-Bôas, S. G. *Nat. Protoc.* **2010**, *5*, 1709.

(347) Castillo, S.; Mattila, I.; Miettinen, J.; Orešič, M.; Hyötyläinen, T. *Anal. Chem.* **2011**, *83*, 3058.

(348) Sverko, E.; Tomy, G. T.; Marvin, C. H.; Zaruk, D.; Reiner, E.; Helm, P. A.; Hill, B.; McCarry, B. E. *Environ. Sci. Technol.* **2008**, *42*, 361.

(349) Abate, S.; Ahn, Y. G.; Kind, T.; Cataldi, T. R. I.; Fiehn, O. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1172.

(350) Nakamura, S.; Sian, T. H.; Daishima, S. *J. Chromatogr. A* **2001**, *19*, 275.

(351) Rosenfelder, N.; Vetter, W. *Mass Spectrom.* **2009**, *23*, 3807.

(352) Fiehn, O.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Anal. Chem.* **2000**, *72*, 3573.

(353) Koek, M. M.; Bakels, F.; Engel, W.; van den Maagdenberg, A.; Ferrari, M. D.; Coulier, L.; Hankemeier, T. *Anal. Chem.* **2010**, *82*, 156.

(354) Koek, M. M.; Muilwijk, B.; van der Werf, M. J.; Hankemeier, T. *Anal. Chem.* **2006**, *8*, 1272.

(355) Little, J. L. *J. Chromatogr. A* **1999**, *84*, 1.

(356) Gu, Q.; David, F.; Lynen, F.; Rumpel, K.; Dugardeyn, J.; Van Der Straeten, D.; Xu, G.; Sandra, P. *J. Chromatogr. A* **2011**, *1218*, 3247.

(357) Corey, E. J.; Venkateswarlu, A. *J. Am. Chem. Soc.* **1972**, *94*, 6190.

(358) Danielsson, A. P. H.; Moritz, T.; Mulder, H.; Spégel, P. *Metabolomics* **2012**, *8*, 50.

(359) Huang, X.; Regnier, F. E. *Anal. Chem.* **2008**, *80*, 107.

(360) Lien, S. K.; Kvitvang, H. F.; Bruheim, P. *J. Chromatogr. A* **2012**, *1274*, 118.

(361) Hong, Z.; Lin, Z.; Liu, Y.; Tan, G.; Lou, Z.; Zhu, Z.; Chai, Y.; Fan, G.; Zhang, J.; Zhang, L. *J. Chromatogr. A* **2012**, *1254*, 14.

(362) Liebeke, M.; Wunder, A.; Lalk, M. *Anal. Biochem.* **2010**, *401*, 312.

(363) Kouremenos, K. A.; Harynuk, J. J.; Winniford, W. L.; Morrison, P. D.; Marriott, P. J. *J. Chromatogr. B* **2010**, *878*, 1761.

(364) Noctor, G.; Bergot, G.; Mauve, C.; Thominet, D.; Lelarge-Trouverie, C.; Priou, J. *Metabolomics* **2005**, *3*, 161.

(365) Villas-Bôas, S. G.; Delicado, D. G.; Akesson, M.; Nielsen, J. *Anal. Biochem.* **2003**, *322*, 134.

(366) Gao, X.; Pujos-Guillot, E.; Martin, J.-F.; Galan, P.; Juste, C.; Jia, W.; Sebedio, J.-L. *Anal. Biochem.* **2009**, *393*, 163.

(367) Tao, X.; Liu, Y.; Wang, Y.; Qiu, Y.; Lin, J.; Zhao, A.; Su, M.; Jia, W. *Anal. Bioanal. Chem.* **2008**, *391*, 2881.

(368) Han, L.-D.; Xia, J.-F.; Liang, Q.-L.; Wang, Y.-M. Y.; Hu, P.; Li, P.; Luo, G.-A. *Anal. Chim. Acta* **2011**, *689*, 85.

(369) Villas-Bôas, S. G.; Smart, K. F.; Sivakumaran, S.; Lane, G. A. *Metabolites* **2011**, *1*, 3.

(370) Vincenti, M.; Fasano, F.; Valsania, M. C.; Guarda, P.; Richardson, S. D. *Anal. Bioanal. Chem.* **2010**, *397*, 43.

(371) Hušek, P.; Šimek, P.; Hartvich, P.; Zahradníčková, H. *J. Chromatogr. A* **2008**, *1186*, 391.

(372) Kanani, H.; Chrysanthopoulos, P. K.; Klapa, M. I. *J. Chromatogr. B* **2008**, *871*, 191.

(373) Zhang, Z.; Pawliszyn, J. *Anal. Chem.* **1993**, *65*, 1843.

(374) Musteata, F. M.; Pawliszyn, J. *TrAC Trends Anal. Chem.* **2007**, *26*, 36.

(375) Vas, G.; Vekey, K. *J. Mass Spectrom.* **2004**, *39*, 233.

(376) Llop, A.; Pocurull, E.; Borrull, F. *J. Sep. Sci.* **2011**, *34*, 1531.

(377) Penn, D. J.; Oberzaucher, E.; Grammer, K.; Fischer, G.; Soini, H. A.; Wiesler, D.; Novotny, M. V.; Dixon, S. J.; Xu, Y.; Brereton, R. G. *J. R. Soc. Interface* **2007**, *4*, 331.

(378) Kaur, N.; Cabral, J. L.; Morin, A.; Waldron, K. C. *J. Chromatogr. A* **2011**, *1218*, 324.

(379) Tikunov, Y.; Lommen, A.; de Vos, C. H.; Verhoeven, H. A.; Bino, R. J.; Hall, R. D.; Bovy, A. G. *Plant Physiol.* **2005**, *139*, 1125.

(380) Aprea, E.; Gika, H.; Carlin, S.; Theodoridis, G.; Vrhovsek, U.; Mattivi, F. *J. Chromatogr.* **2011**, *1218*, 4517.

(381) Shirasu, M.; Touhara, K. *J. Biochem.* **2011**, *150*, 257.

(382) Silva, C. L.; Passos, M.; Câmara, J. S. *Br. J. Cancer* **2011**, *105*, 1894.

(383) Zhang, S.; Liu, L.; Steffen, D.; Ye, T.; Raftery, D. *Metabolomics* **2011**, *8*, 323.

(384) Phillips, M.; Cataneo, R. N.; Ditkoff, B. A.; Fisher, P.; Greenberg, J.; Gunawardena, R.; Kwon, C. S.; Tietje, O.; Wong, C. *Breast Cancer Res.* **2006**, *99*, 19.

(385) Matsumura, K.; Opiekun, M.; Oka, H.; Vachani, A.; Albelda, S.; Yamazaki, K.; Beauchamp, G. *PLoS One* **2010**, *5*, e8819.

(386) McCulloch, M.; Jezierski, T.; Broffman, M.; Hubbard, A.; Turner, K. *Integr. Cancer Ther.* **2006**, *5*, 30.

(387) Delahunty, C.; Eyres, G.; Dufour, J. *J. Chromatogr. A* **2008**, *1186*, 123.

(388) d'Acampora Zellner, B.; Dugo, P.; Dugo, G.; Mondello, L.; Delahunty, C.; Eyres, G.; Dufour, J. *J. Chromatogr. A* **2008**, *1186*, 123.

(389) Hufnagel, J. C.; Hofmann, T. *J. Agric. Food Chem.* **2008**, *56*, 9190.

(390) Lo, Y.; Koziel, J.; Cai, L.; Hoff, S.; Jenks, W.; Xin, H. *J. Environ. Qual.* **2008**, *37*, 521.

(391) Laor, Y; Koziel, J.; Cai, L.; Ravid, U. *J. Air Waste Manag. Assoc.* **2008**, *58*, 1187.

(392) Stensmyr, M. C.; Urru, I.; Collu, I.; Celander, M.; Hansson, B. S.; Angioy, A.-M. *Nature* **2002**, *420*, 625.

(393) Shirasu, M.; Nagai, S.; Hayashi, R.; Ochiai, A.; Touhara, K. *Biosci., Biotechnol., Biochem.* **2009**, *73*, 2117.

(394) Zelena, E.; Dunn, W. B. W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. K. M.; Begley, P.; O'Hagan, S.; Knowles, J. J. D.; Halsall, A. A.; HUSERMET Consortium, Wilson, I.; Kell, D. B. D. B.; Wilson, I. D. *Anal. Chem.* **2009**, *81*, 1357.

(395) Nordstrom, A.; Want, E.; Northen, T.; Lehtio, J.; Siuzdak, G. *Anal. Chem.* **2008**, *80*, 421.

(396) Neue, U. D.; Tran, K. V.; Iraneta, P. C.; Alden, B. A. *J. Sep. Sci.* **2003**, *26*, 174.

(397) Kamleh, M. A.; Ebbels, T. M.; Spagou, K.; Masson, P.; Want, E. J. *Anal. Chem.* **2012**, *84*, 2670–2677.

(398) Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. *J. Chromatogr. A* **2008**, *1189*, 314.

(399) t'Kindt, R.; De Veylder, L.; Storme, M.; Deforce, D.; Van Bocxlaer, J. *J. Chromatogr. B* **2008**, *871*, 37.

(400) Eliasson, M; Rännar, S.; Madsen, R.; Donten, M.; Marsden-Edwards, E.; Moritz, T.; Shockcor, J.; Johansson, E.; Trygg, J. *Anal. Chem.* **2012**, *84*, 6869.

(401) Bruce, S. J.; Jonsson, P.; Antti, H.; Cloarec, O.; Trygg, J.; Marklund, S. L.; Moritz, T. *Anal. Biochem.* **2008**, *372*, 237.

(402) Gao, X.; Zhang, Q.; Meng, D.; Isaac, G.; Zhao, R.; Fillmore, T. L.; Chu, R. K.; Zhou, J.; Tang, K.; Hu, Z.; Moore, R. J.; Smith, R. D.; Katze, M. G.; Metz, T. O. *Anal. Bioanal. Chem.* **2012**, *402*, 2923.

(403) Bell, D. S.; Cramer, H. M.; Jones, A. D. *J. Chromatogr. A* **2005**, *1095*, 113.

(404) Gika, H. G.; Theodoridis, G.; Extance, J.; Edge, A. M.; Wilson, I. D. *J. Chromatogr. B* **2008**, *871*, 279–287.

(405) Stoll, D.; Wang, X.; Carr, P. *Anal. Chem.* **2008**, *80*, 268.

(406) D'Attoma, A.; Grivel, C.; Heinisch, S. *J. Chromatogr. A* **2012**, *1262*, 148.

(407) Stoll, D. R. *Anal. Bioanal. Chem.* **2010**, *397*, 979.

(408) Fairchild, J. N.; Horvath, K.; Gooding, J. R.; Campagna, S. R.; Guiochon, G. *J. Chromatogr. A* **2010**, *1217*, 8161.

(409) Guo, K.; Peng, J.; Zhou, R.; Li, L. *J. Chromatogr. A* **2011**, *1218*, 3689.

(410) Liu, Y.; Xue, X.; Guo, Z.; Xu, Q.; Zhang, F.; Liang, X. *J. Chromatogr. A* **2008**, *1208*, 133.

(411) Zheng, J.; Dixon, R. A.; Li, L. *Anal. Chem.* **2012**, *84*, 10802–10811.

(412) Dunn, W. B.; Broadhurst, D.; Brown, M.; Baker, P. N.; Redman, C. W. G.; Kenny, L. C.; Kell, D. B. *J. Chromatogr. B* **2008**, *871*, 288.

(413) Ciborowski, M.; Teul, J.; Martin-Ventura, J. L.; Egido, J.; Barbas, C. *PLoS One* **2012**, *7*, e31982.

(414) Roux, A.; Xu, Y.; Heilier, J.; Olivier, M.; Ezan, E.; Tabet, J.; Junot, C. *Anal. Chem.* **2012**, *84*, 6429.

(415) Haug, K.; Salek, R.; Conesa, P.; Hastings, J.; de Matos, P.; Rijnbeek, M.; Mahendraker, T.; Williams, M.; Neumann, S.; Rocca-Serra, P.; Maguire, E.; González-Beltrán, A.; Sansone, S.; Griffin, J.; Steinbeck, C. *Nucleic Acids Res.* **2012**, *47*, D781.

(416) Cubbon, S.; Bradbury, T.; Wilson, J.; Thomas-Oates, J. *Anal. Chem.* **2007**, *79*, 8911.

(417) De Carvalho, L.; Zhao, H.; Dickinson, C.; Arango, N.; Lima, C.; Fischer, S.; Ouerfelli, O.; Nathan, C.; Rhee, K. *Chem. Biol.* **2010**, *17*, 323.

(418) Tsutsui, H.; Maeda, T.; Min, J. Z.; Inagaki, S.; Higashi, T.; Kagawa, Y.; Toyo'oka, T. *Clin. Chim. Acta* **2011**, *412*, 861.

(419) Lv, H.; Palacios, G.; Hartil, K.; Kurland, I. J. *J. Proteome Res.* **2011**, *10*, 2104.

(420) Chen, L.; Zhou, L.; Chan, E. C. Y.; Neo, J.; Beuerman, R. W. *J. Proteome Res.* **2011**, *10*, 4876.

(421) Lin, L.; Huang, Z.; Gao, Y.; Yan, X.; Xing, J.; Hang, W. *J. Proteome Res.* **2011**, *10*, 1396.

(422) García-Cañaveras, J.; Donato, M.; Castell, J.; Lahoz, A. *J. Proteome Res.* **2011**, *10*, 4825.

(423) O'Maille, G.; Go, E. P.; Hoang, L.; Want, E. J.; Smith, C.; O'Maille, P.; Nordstrom, A.; Morita, H.; Qin, C.; Uritboonthai, W.; Apon, J.; Moore, R.; Garrett, J.; Siuzdak, G. *Spectroscopy* **2008**, *22*, 327.

(424) Carlson, E.; Cravatt, B. *Nat. Methods* **2007**, *4*, 429.

(425) Soukup-Hein, RJ; Remsburg, J.; Dasgupta, P.; Armstrong, D. *Anal. Chem.* **2007**, *79*, 7346.

(426) Woo, H. K.; Go, E. P.; Hoang, L.; Trauger, S. A.; Bowen, B.; Siuzdak, G.; Northen, T. R. *Rapid Comm. Mass Spectrom.* **2009**, *23*, 1849.

(427) Yang, W.; Adamec, J.; Regnier, F. *Anal. Chem.* **2007**, *79*, 5150.

(428) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919.

(429) Barbas, C.; Moraes, E. P.; Villaseñor, A. *J. Pharm. BIomed. Anal.* **2011**, *55*, 823.

(430) Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. *J. Proteome Res.* **2003**, *2*, 488.

(431) Lee, R.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 7047.

(432) Ramautar, R.; Nevedomskaya, E.; Mayboroda, O.; Deelder, A.; Wilson, I.; Gika, H.; Theodoridis, G.; Somsen, G.; de Jong, G. *Mol. BioSyst.* **2011**, *7*, 194–199.

(433) Nevedomskaya, E.; Ramautar, R.; Derks, R.; Westbroek, I.; Zondag, G.; Pluijm, I.; Deelder, A. M.; Mayboroda, O. A.; van der Pluijm, I. *J. Proteome Res.* **2010**, *9*, 4869–4874.

(434) Nemes, P.; Knolhoff, A.; Rubakhin, S.; Sweedler, J. *ACS Chem. Neurosci.* **2012**, *3*, 782.

(435) Ptolemy, A. S.; Britz-McKibbin, P. *J. Chromatogr. A* **2006**, *1106*, 7–18.

(436) Copper, C. L. **1998**, *75*, 343–347.

(437) Schmitt-Kopplin, P.; Frommberger, M. *Electrophoresis* **2003**, *24*, 3837–3867.

(438) Zheng, J.; Jann, M. W.; Hon, Y. Y.; Shamsi, S. A. *Electrophoresis* **2004**, *25*, 2033–2043.

(439) Mokaddem, M.; Gareil, P.; Belgaied, J.-E.; Varenne, A. *Electrophoresis* **2009**, *30*, 1692.

(440) Zamfir, A. D. *J. Chromatogr. A* **2007**, *1159*, 2.

(441) Moini, M. *Anal. Chem.* **2007**, *79*, 4241–4246.

(442) Fanali, S.; D'Orazio, G.; Foret, F.; Kleparnik, K.; Aturki, Z. *Electrophoresis* **2006**, *27*, 4666–4673.

(443) Bello, M.; Capelli, L.; Righetti, P. *J. Chromatogr. A* **1994**, *684*, 311.

(444) Ramautar, R.; Toraño, J. S.; Somsen, G. W.; de Jong, G. J. *Electrophoresis* **2010**, *31*, 2319.

(445) Grundmann, M.; Matysik, F. *Anal. Bioanal. Chem.* **2011**, *400*, 269.

(446) Tanaka, Y.; Higashi, T.; Rakwal, R.; Wakida, S.; Iwahashi, H. *Electrophoresis* **2008**, *29*, 2016.

(447) Harada, K.; Ohyama, Y.; Tabushi, T.; Kobayashi, A.; Fukusaki, E. *J. Biosci. Bioeng.* **2008**, *105*, 249.

(448) Lee, R.; Britz-McKibbin, P. *Electrophoresis* **2010**, *31*, 2328.

(449) Soga, T.; Sugimoto, M.; Honma, M.; Mori, M.; Igarashi, K.; Kashikura, K.; Ikeda, S.; Hirayama, A.; Yamamoto, T.; Yoshida, H.; Otsuka, M.; Tsuji, S.;

Yatomi, Y.; Sakuragawa, T.; Watanabe, H.; Nihei, K.; Saito, T.; Kawata, S.; Suzuki, H.; Tomita, M.; Suematsu, M. *J. Hepatol.* **2011**, *55*, 896.

(450) Ramautar, R.; Shyti, R.; Schoenmaker, B.; de Groote, L.; Derks, R. J.; Ferrari, M. D.; van den Maagdenberg, A. M.; Deelder, A. M.; Mayboroda, O. A. *Anal. Bioanal. Chem.* **2012**, *404*, 2895.

(451) Simó, C.; Ibáñez, C.; Gómez-Martínez, A.; Ferragut, J. A.; Cifuentes, A. *Electrophoresis* **2011**, *32*, 1765.

(452) Soga, T.; Ishikawa, T.; Igarashi, S.; Sugawara, K.; Kakazu, Y.; Tomita, M. *J. Chromatogr. A* **2007**, *1159*, 125.

(453) Harada, K.; Fukusaki, E.; Kobayashi, A. *J. Biosci. Bioeng.* **2006**, *101*, 403.

(454) Sato, S.; Yanagisawa, S. *Metabolomics* **2010**, *6*, 529.

(455) Timischl, B.; Dettmer, K.; Kaspar, H.; Thieme, M.; Oefner, P. J. *Electrophoresis* **2008**, *29*, 2203.

(456) Hopfgartner, G.; Varesio, E. In *Metabolomics in Practice*; Lammerhofer, M.; Weckwerth, W., Eds.; Wiley-VCH: Weinheim, Germany, 2013; pp. 21–37.

(457) Vuckovic, D. *Anal. Bioanal. Chem.* **2012**, *403*, 1523.

(458) Geier, F. M.; Want, E. J.; Leroi, A. M.; Bundy, J. G. *Anal. Chem.* **2011**, *83*, 3730.

(459) Saric, J.; Want, E.; Duthaler, U.; Lewis, M.; Keiser, J.; Shockcor, J.; Ross, G.; Nicholson, J.; Holmes, E.; Tavares, M. *Anal. Chem.* **2012**, *84*, 6963–6972.

(460) Beltran, A.; Suarez, M.; Rodríguez, M. A.; Vinaixa, M.; Samino, S.; Arola, L.; Correig, X.; Yanes, O. *Anal. Chem.* **2012**, *84*, 5838.

(461) Gullberg, J.; Jonsson, P.; Nordström, A.; Sjöström, M.; Moritz, T. *Anal. Biochem.* **2004**, *331*, 283.

(462) T'Kindt, R.; Morreel, K.; Deforce, D.; Boerjan, W.; Van Bocxlaer, J. *J. Chromatogr. B* **2009**, *877*, 3572.

(463) Canelas, A.; ten Pierick, A.; Ras, C.; Seifar, R.; van Dam, J.; van Gulik, W.; Heijnen, J. *Anal. Chem.* **2009**, *81*, 7379.

(464) Shin, M. H.; Lee do, Y.; Liu, K. H.; Fiehn, O.; Kim, K. H. *Anal. Chem.* **2010**, *82*, 6660.

(465) El Rammouz, R.; Létisse, F.; Durand, S.; Portais, J.-C.; Moussa, Z. W.; Fernandez, X. *Anal. Biochem.* **2010**, *398*, 169.

(466) Smith, L.; Maher, A.; Want, E.; Elliott, P.; Stamler, J.; Hawkes, G.; Holmes, E.; Lindon, J.; Nicholson, J. K. *Anal. Chem.* **2009**, *81*, 4847.

(467) Rosenling, T.; Stoop, M.; Smolinska, A.; Muilwijk, B.; Coulier, L.; Shi, S.; Dane, A.; Christin, C.; Suits, F.; Horvatovich, P.; Wijmenga, S.; Buydens, L.; Vreeken, R.; Hankemeier, T.; van Gool, A.; Luider, T.; Bischoff, R. *Clin. Chem.* **2011**, *57*, 1703.

(468) Boernsen, K. O.; Gatzek, S.; Imbert, G. *Anal. Chem.* **2005**, *77*, 7255.

(469) Zivkovic, A. M.; Wiest, M. M.; Nguyen, U. T.; Davis, R.; Watkins, S. M.; German, J. B. *Metabolomics* **2009**, *5*, 507.

(470) Duportet, X.; Bastos, R.; Aggio, M.; Carneiro, S.; Villas-Bôas, S. G. *Metabolomics* **2012**, *8*, 41.

(471) Elliott, P.; Peakman, T. C. *Int. J. Epidemiol.* **2008**, *37*, 234.

(472) Dunn, W. B.; Broadhurst, D.; Ellis, D. I.; Brown, M.; Halsall, A.; O'Hagan, S.; Spasic, I.; Tseng, A.; Kell, D. B. *Int. J. Epidemiol.* **2008**, *37*, i23.

(473) Want, E. J.; O'Maille, G.; Smith, C. A.; Brandon, T. R.; Uritboonthai, W.; Qin, C.; Trauger, S. A.; Siuzdak, G.; Uritboonthai, W. *Anal. Chem.* **2006**, *78*, 743.

(474) Masson, P; Alves, A.; Ebbels, T.; Nicholson, J.; Want, E. *Anal. Chem.* **2010**, *82*, 7779.

(475) Kanani, H. H.; Klapa, M. I. *Metab. Eng.* **2007**, *9*, 39.

(476) Nilsson, S.; Bylund, D.; Jörntén-Karlsson, M.; Petersson, P.; Markides, K. *Electrophoresis* **2004**, *25*, 2100.

(477) Leon, C.; Rodriguez-Meizoso, I.; Lucio, M.; Garcia-Cañas, V.; Ibañez, E.; Schmitt-Kopplin, P.; Cifuentes, A. *J. Chromatogr. A* **2009**, *1216*, 7314.

(478) Yang, W.-C.; Regnier, F. E.; Adamec, J. *Electrophoresis* **2008**, *29*, 4549.

(479) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(480) Jones, D. P. *Rejuvenation Res.* **2006**, *9*, 169.

(481) Sugimoto, M.; Koseki, T.; Hirayama, A.; Abe, S.; Sano, T.; Tomita, M.; Soga, T. *J. Agric. Food Chem.* **2010**, *58*, 374.

(482) Sugimoto, M.; Wong, D. T.; Hirayama, A.; Soga, T.; Tomita, M. *Metabolomics* **2010**, *6*, 78.

(483) Sugimoto, M.; Saruta, J.; Matsuki, C.; To, M.; Onuma, H.; Kaneko, M.; Soga, T.; Tomita, M.; Tsukinoki, K. *Metabolomics* **2013**, *9*, 454–463.

(484) Ranshoff, D. F. *Nat. Rev.* **2005**, *5*, 142.

(485) Diamandis, E. P. *BMC Med.* **2012**, *10*, 87.

(486) Sreekumar, A.; Poisson, L.; Rajendiran, T.; Khan, A.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R.; Li, Y.; Nyati, M.; Ahsan, A.; Kalyana-Sundaram, S.; Han, B.; Cao, X.; Byun, J.; Omenn, G.; Ghosh, D.; Pennathur, S.; Alexander, D.; Berger, A.; Shuster, J.; Wei, J.; Varambally, S.; Beecher, C.; Chinnaiyan, A. *Nature* **2009**, *457*, 910.

(487) Stephan, C.; Jentzmik, F.; Jung, K. *Eur. Urol.* **2010**, *58*, 19.

(488) Bollard, M.; Holmes, E.; Lindon, J.; Mitchell, S.; Branstetter, D.; Zhang, W.; Nicholson, J. K. *Anal. Biochem.* **2001**, *295*, 194.

(489) Plumb, R.; Granger, J.; Stumpf, C.; Wilson, I.; Evans, J.; Lenz, E. *Analyst* **2003**, *128*, 819.

(490) Shi, H.; Vigneau-Callahan, K. E.; Shestopalov, A. I.; Milbury, P. E.; Matson, W. R.; Kristal, B. S. *J. Nutr.* **2002**, *132*, 1039.

(491) Hutterer, F; Roboz, J.; Sarkozi, L.; Ruhig, A.; Bacchin, P. *Clin. Chem.* **1971**, *17*, 789.

(492) Dallmann, R.; Viola, A. U.; Tarokh, L.; Cajochen, C.; Brown, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *108*, 2625.

(493) Szeto, S.; Reinke, S.; Lemire, B. *J. Biomol. NMR.* **2011**, *49*, 245.

(494) Friswell, M. K.; Gika, H.; Stratford, I. J.; Theodoridis, G.; Telfer, B.; Wilson, I. D.; McBain, A. J. *PLoS One* **2010**, *5*, e8584.

(495) Nicholson, G.; Rantalainen, M.; Maher, A.; Li, J.; Malmodin, D.; Ahmadi, K.; Faber, J.; Hallgrímsdóttir, I.; Barrett, A.; Toft, H.; Krestyaninova, M.; Viksna, J.; Neogi, S.; Dumas, M.; Sarkans, U.; Silverman, B.; Donnelly, P.; Nicholson, J.; Allen, M.; Zondervan, K.; Lindon, J.; Spector, T.; McCarthy, M.; Holmes, E.; Baunsgaard, D.; Holmes, C. *Mol. Syst. Biol.* **2011**, *7*, 525.

(496) Crews, B.; Wikoff, W.; Patti, G.; Woo, H.; Kalisiak, E.; Heideker, J.; Siuzdak, G. *Anal. Chem.* **2009**, *81*, 8538.

(497) Favé, G.; Beckmann, M.; Lloyd, A. J.; Zhou, S.; Harold, G.; Lin, W.; Tailliart, K.; Xie, L.; Draper, J.; Mathers, J. C. *Metabolomics* **2011**, *7*, 469.

(498) Van Velzen, E.; Westerhuis, J.; van Duynhoven, J.; van Dorsten, F.; Hoefsloot, H.; Jacobs, D.; Smit, S.; Draijer, R.; Kroner, C.; Smilde, A. *J. Proteome Res.* **2008**, *7*, 4483.

(499) Van der Greef, J.; Hankemeier, T.; McBurney, R. N. *Pharmacogenomics* **2006**, *7*, 1087.

(500) Krug, S.; Kastenmüller, G.; Stückler, F.; Rist, M. J. M.; Skurk, T.; Sailer, M.; Raffler, J.; Römisch-Margl, W.; Adamski, J.; Prehn, C.; Frank, T.; Engel, K. K.-H.; Hofmann, T.; Luy, B.; Zimmermann, R.; Moritz, F.; Schmitt-Kopplin, P.; Krumsiek, J.; Kremer, W.; Huber, F.; Oeh, U.; Theis, F. F. J.; Szymczak, W.; Hauner, H.; Suhre, K.; Daniel, H. *FASEB J.* **2012**, *26*, 2607–2619.

(501) Patti, G. J.; Tautenhahn, R.; Siuzdak, G. *Nat. Protoc.* **2012**, *7*, 508.

(502) Broadhurst, D. I.; Kell, D. B. *Metabolomics* **2007**, *2*, 171–196.

(503) Fernie, A.; Aharoni, A.; Willmitzer, L.; Stitt, M.; Tohge, T.; Kopka, J.; Carroll, A.; Saito, K.; Fraser, P.; DeLuca, V. *Plant Cell.* **2011**, *23*, 2477.

(504) Van den Berg, R.; Hoefsloot, H.; Westerhuis, J.; Smilde, A.; van der Werf, M. *BMC Genomics* **2006**, *7*, 142.

(505) Sangster, T.; Major, H.; Plumb, R.; Wilson, A. J.; Wilson, I. D. *Analyst* **2006**, *131*, 1075–1078.

(506) Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S. *Nucleic Acids Res.* **2012**, *40*, W127–W133.

(507) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84*, 5035–5039.

(508) Wei, X.; Shi, X.; Kim, S.; Zhang, L.; Patrick, J.; Binkley, J.; McClain, C.; Zhang, X. *Anal. Chem.* **2012**, *84*, 7963.

(509) Creek, D.; Jankevics, A.; Burgess, K.; Breitling, R.; Barrett, M. *Bioinformatics* **2012**, *28*, 1048.

(510) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinf.* **2010**, *11*, 395.

(511) Melamud, E.; Vastag, L.; Rabinowitz, J. D. *Anal. Chem.* **2010**, *82*, 9818.

(512) Dragsted, L. O. *Metabolites* **2012**, *2*, 77.

(513) Hendriks, M. M. W. B.; van Eeuwijk, F. A.; Jellema, R. H.; Westerhuis, J. A.; Reijmers, T. H.; Hoefsloot, H. C. J.; Smilde, A. K. *TrAC, Trends Anal. Chem.* **2011**, *30*, 1685.

(514) Hrydziuszko, O.; Viant, M. R. *Metabolomics* **2012**, *8*, 161.

(515) Duarte, N. C.; Becker, S. A.; Jamshidi, N.; Thiele, I.; Mo, M. L.; Vo, T. D.; Srivas, R.; Palsson, B. Ø. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1777.

(516) Zhang, G.-F.; Sadhukhan, S.; Tochtrop, G. P.; Brunengraber, H. *J. Biol. Chem.* **2011**, *286*, 23631.

(517) Bordbar, A.; Mo, M. L.; Nakayasu, E. S.; Schrimpe-Rutledge, A. C.; Kim, Y.-M.; Metz, T. O.; Jones, M. B.; Frank, B. C.; Smith, R. D.; Peterson, S. N.; Hyduke, D. R.; Adkins, J. N.; Palsson, B. Ø. *Mol. Syst. Biol.* **2012**, *8*, 558.

(518) Ishii, N.; Nakahigashi, K.; Baba, T.; Robert, M.; Soga, T.; Kanai, A.; Hirasawa, T.; Naba, M.; Hirai, K.; Hoque, A.; Ho, P.; Kakazu, Y.; Sugawara, K.; Igarashi, S.; Harada, S.; Masuda, T.; Sugiyama, N.; Togashi, T.; Hasegawa, M.; Takai, Y.; Yugi, K.; Arakawa, K.; Iwata, N.; Toya, Y.; Nakayama, Y.; Nishioka, T.; Shimizu, K.; Mori, H.; Tomita, M. *Science.* **2007**, *316*, 593.

(519) Canadian Diabetes Association. *Can. J. Diabetes* **2008**, *32*.

(520) Ransom, T.; Goldenberg, R.; Mikalachki, A.; Prebtani, A. P. H.; Punthakee, Z. *Can. J. Diabetes* **2013**, *37*, S16–19.

(521) *Glucose metabolism: An Overview (Chapter 23.2)*; Pearson Education, 2010.

(522) Matthews, D. R.; Hosker, J. R.; Rudenski, A. S.; Naylor, B. A.; Treacher, D. F.; Turner, R. C.; Infirmary, R. *Diabetologia* **1985**, 412–419.

(523) Knowler, W. C.; Barrett-Connor, E.; Fowler, S. E.; Hamman, R. F.; Lachin, J. M.; Walker, E. a; Nathan, D. M. *N. Engl. J. Med.* **2002**, *346*, 393–403.

(524) Public Health Agency of Canada. *Diabetes in Canada: Facts and Figures from a Public Health Perspective*; Ottawa, 2011.

(525) Alberti, K. G. M. M.; Zimmet, P.; Shaw, J. *Diabet. Med.* **2007**, *24*, 451–463.

(526) Goldenberg, R.; Punthakee, Z. *Can. J. Diabetes* **2013**, *37*, S8–S11.

(527) Zhao, X.; Peter, A.; Fritsche, J.; Elcnerova, M.; Fritsche, A.; Häring, H.-U.; Schleicher, E. D.; Xu, G.; Lehmann, R. *Am. J. Physiol. Endocrinol. Metab.* **2009**, *296*, E384–93.

(528) Shaham, O.; Wei, R.; Wang, T. J.; Ricciardi, C.; Lewis, G. D.; Vasan, R. S.; Carr, S. a; Thadhani, R.; Gerszten, R. E.; Mootha, V. K. *Mol. Syst. Biol.* **2008**, *4*, 214.

(529) Spégel, P.; Danielsson, A. P. H.; Bacos, K.; Nagorny, C. L. F.; Moritz, T.; Mulder, H.; Filipsson, K. *Metabolomics* **2009**, *6*, 56–66.

(530) Wopereis, S.; Rubingh, C. M.; van Erk, M. J.; Verheij, E. R.; van Vliet, T.; Cnubben, N. H. P.; Smilde, A. K.; van der Greef, J.; van Ommen, B.; Hendriks, H. F. J. *PLoS One* **2009**, *4*, e4525.

(531) Bain, J. R.; Muehlbauer, M. J. *Diabetes* **2013**, *62*, 2651–2653.

(532) Razak, F.; Anand, S. S. *N. Engl. J. Med.* **2004**, *9*, 223–224.

(533) Holloszy, J. O. *J. Appl. Physiol.* **2005**, *99*, 338–343.

(534) Knapik, J. J.; Meredith, C. N.; Jones, B. H.; Suek, L.; Young, V. R.; Evans, W. J. *J. Appl. Physiol.* **1988**, *64*, 1923–1929.

(535) Petersen, K. F.; Befroy, D.; Dufour, S.; Dziura, J.; Ariyan, C.; Rothman, D. L.; DiPietro, L.; Cline, G. W.; Shulman, G. I. *Science.* **2003**, *300*, 1140–1142.

(536) Bird, S. R.; Hawley, J. A. *Maturitas* **2012**, *72*, 311 – 316.

(537) Kosaka, K.; Noda, M.; Kuzuya, T. *Diabetes Res. Clin. Pract.* **2005**, *67*, 152–162.

(538) Diabetes Prevention Research Group. *Diabetes Care* **2003**, *26*, 977–980.

(539) Chiasson, J.-L.; Josse, R. G.; Gomis, R.; Hanefeld, M.; Karasik, A.; Laakso, M. *Lancet* **2002**, *359*, 2072–2077.

(540) Tuomilehto, J.; Lindström, J.; Eriksson, J. G.; Valle, T. T.; Hämäläinen, H.; Ilanne-Parikka, P.; Keinänen-Kiukaanniemi, S.; Laakso, M.; Louheranta, A.; Rastas, M.; Salminen, V.; Uusitupa, M. *N. Engl. J. Med.* **2001**, *344*, 1343–1350.

(541) Burgomaster, K. A.; Cermak, N. M.; Phillips, S. M.; Benton, C. R.; Bonen, A.; Gibala, M. J. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2007**, *292*, R1970–6.

(542) Burgomaster, K. A.; Howarth, K. R.; Phillips, S. M.; Rakobowchuk, M.; Macdonald, M. J.; McGee, S. L.; Gibala, M. J. *J. Physiol.* **2008**, *586*, 151–160.

(543) Gibala, M. J.; Little, J. P.; van Essen, M.; Wilkin, G. P.; Burgomaster, K. A.; Safdar, A.; Raha, S.; Tarnopolsky, M. A. *J. Physiol.* **2006**, *575*, 901–911.

(544) Gibala, M. J.; Little, J. P.; MacDonald, M. J.; Hawley, J. A. *J. Physiol.* **2012**, *590*, 1077–1084.

(545) Hood, M. S.; Little, J. P.; Tarnopolsky, M. A.; Myslik, F.; Gibala, M. J. *Med. Sci. Sports Exerc.* **2011**, *43*, 1849–1856.

(546) Little, J. P.; Gillen, J. B.; Percival, M. E.; Safdar, A.; Tarnopolsky, M. A.; Punthakee, Z.; Jung, M. E.; Gibala, M. J. *J. Appl. Physiol.* **2011**, *111*, 1554–1560.

(547) Gillen, J. B.; Little, J. P.; Punthakee, Z.; Tarnopolsky, M. A.; Riddell, M. C.; Gibala, M. J. *Diabetes. Obes. Metab.* **2012**, *14*, 575–577.

(548) Currie, K. D.; Dubberley, J. B.; McKelvie, R. S.; MacDonald, M. J. *Med. Sci. Sports Exerc.* **2013**, *45*, 1436–1442.

1.9 Appendix A.

Table A1.1. Summary of major separation platforms and recent advances for comprehensive metabolite profiling.

Separation platform	*Metabolite compatibility*	*Advantages*	*Limitations*	*Recent advances*
GC-MS	Volatile, non-polar, thermally stable metabolites; chemical derivatization required for polar metabolites; ideal with headspace sampling with SPME	High separation efficiency; robust/mature platform with EI/CI-MS; precise retention indices; extensive NIST database	Complex sample pretreatment; stringent QC/validation; not applicable to labile metabolites; artifact-prone silylation	2D GCxGC for higher peak capacity; high T ionic liquid columns; rapid alkylation in aqueous conditions
LC-MS	Involatile and thermally unstable metabolites; label-free approach for non-polar/semi-polar metabolites; ideal for involatile, non-polar and semi-polar metabolites	Wide selectivity with RP, HILIC and ANP; high separation efficiency with UHPLC; robust performance of RP-LC with broad metabolome coverage	Variable retention times and complex mechanisms for HILIC/ANP; susceptible to ion suppression with ESI-MS; limited MS/MS databases for identification	Sub-2 μm or core-fused particles; elevated T/thermal elution; chemical derivatization for improved performance
CE-MS	Weakly ionic/strongly ionic and labile metabolites; label-free approach for charged metabolites and peptides; ideal for multivalent ions and their isomers	High separation efficiency; on-line sample preconcentration/ desalting; minimal sample handling; predictive modeling of migration times	Migration time variability; poor concentration sensitivity; lack of inter-lab validation for ensuring robustness; few commercial CE-MS vendors/support	Sheathless ion source; capillary coating or data transformation for stable runs; chemical derivatization for improved performance

Chapter II

Comprehensive Profiling of Free and Conjugated Estrogens by Capillary Electrophoresis-Time of Flight-Mass Spectrometry

II. Comprehensive Profiling of Free and Conjugated Estrogens by Capillary Electrophoresis-Time of Flight-Mass Spectrometry

2.1 Abstract

The biological activity of estrogens is tightly regulated by regioselective phase I/II metabolic transformations that are critical to human health. Current methods for analysis of urinary estrogens are limited by complicated sample pre-treatment and/or inadequate specificity for free estrogens and their glucuronide/sulfate conjugates that vary widely in their intrinsic polarity. In this work, direct speciation of intact estrogen conjugates and their regioisomers is demonstrated using capillary electrophoresis-time of flight/mass spectrometry (CE-TOF/MS) when using an alkaline buffer system with negative ion mode detection. This method allows for resolution of weakly acidic native estrogens, anionic estrogen conjugates and their positional isomers without significant matrix-induced ion suppression effects in human urine. Identification of unknown estrogen metabolites using CE-TOF/MS is supported by accurate mass together with their characteristic relative migration times, which can be predicted based on two intrinsic physicochemical properties of an ion. CE-TOF/MS offers a promising strategy for comprehensive profiling of estrogens and other classes of steroid conjugates that is needed for deeper insight into the etiology and treatment of chronic disorders associated with impaired estrogen metabolism.

2.2 Introduction

Estrogens are an important class of hormone that play essential roles in the female reproductive system, stimulate secondary sexual characteristics, and regulate bone growth and maturation. Dysregulation in estrogen metabolism has been implicated in infertility, obesity and several human diseases, including post-menopausal breast cancer, prostate cancer, Alzheimer's disease and osteoporosis.[1] The biological activity of the major endogenous estrogens, estrone (E_1), estradiol (E_2), and estriol (E_3), is tightly regulated by regioselective phase I/II metabolic transformations that generate a plethora of estrogen metabolites, positional isomers and hydrophilic conjugates that vary in their distribution within specific organs, tissues and biological fluids.[2] For instance, oxidation of estrogens by human cytochrome P450 isoforms can generate 2, 4 or 16-hydroxylated estrogen isomers that differ in their potency to stimulate cell proliferation and/or elicit oxidative damage via catechol estrogen quinone formation.[3] There is growing

concern regarding the confounding impact of xenobiotics in human health due to long-term exposure to pharmacological agents, dietary phytoestrogens and endocrine-disrupting substances in the environment with estrogenic activity.[4–7] Deactivation of estrogens and their subsequent excretion in urine is primarily modulated by glucuronide and sulphate conjugation reactions that are catalyzed by UDP-glucuronosyltransferases, β-glucuronidase, sulfotransferases and steroid sulfatase in the liver and other estrogen-responsive tissues.[8] Although conjugation is considered an important detoxification pathway, it also serves as a regulatory mechanism for controlling the transport and local activity of circulating estrogens to hormone-responsive cells, such as desulfation of estrone 3-sulphate ($E_1$3S).[9]

Various strategies have been developed to overcome the major technical challenges for estrogen analysis as required for clinical, environmental, food science and epidemiological studies.[8,10,11] Commercial immunoassay kits allow for rapid and sensitive analysis of trace levels of targeted estrogens; however suffer from poor specificity with cross-reactivity to other interferences in complex biological samples that often contribute to positive bias.[12,13] Indirect immunoassays with off-line liquid extraction and/or chromatographic cleanup can improve method specificity but are time-consuming and not amenable to automation.[8] Gas chromatography-mass spectrometry (GC-MS) offers greater selectivity for native estrogens together with their qualitative identification but is not applicable to the direct analysis of intact estrogen conjugates while requiring complex sample handling, such as enzymatic deconjugation, chemical derivatization, and solid-phase extraction (SPE).[14,15] Currently, liquid chromatography-tandem mass spectrometry (LC-MS/MS) represents the gold standard for analysis of free estrogens[3,16–18] and their intact glucuronide and/or sulfate conjugates.[19–22] However, simultaneous resolution of lipophilic estrogens and their ionic conjugates by reverse-phase chromatography is challenging since it requires prefractionation,[23] multiple columns,[24] column switching,[25] two dimensional separations[26] or complex gradient elution programs.[27] Alternatively, hydrophilic interaction liquid chromatography (HILIC) offers better retention of polar metabolites with elution using a high content of acetontrile that improves sensitivity when using electrospray ionization (ESI)-MS.[28] However, free estrogens and/or strongly acidic estrogen sulfates are often poorly retained on commercial HILIC columns[29] that can undergo band broadening by complex retention mechanisms.[30] In most cases, adequate sensitivity for trace analysis of estrogens by LC-MS/MS can only be achieved when using SPE[21,31] and/or enzymatic deconjugation with chemical derivatization in order to enhance

ionization efficiency with a cationic label (*e.g.*, dansyl chloride) under positive ion mode detection.[27, 1,27,32–35] The latter strategy necessarily precludes the direct analysis of intact estrogen conjugates in complex biological samples, such as human urine.

Capillary electrophoresis-mass spectrometry (CE-MS) represents a promising platform in metabolomics research[36,37] that is particularly useful for simultaneous resolution of polar metabolites that vary widely in their intrinsic polarity using a simple aqueous buffer system.[38,39] Previous CE methods for analysis of free estrogens[40–42] and their conjugates[43] have used micellar electrokinetic chromatography with UV detection that is only directly compatible to ESI-MS when using partial filling strategies.[44] Alternatively, an alkaline buffer can be used to induce ionization of weakly acidic free estrogens ($pK_a \approx 10.4$)[45] for their resolution by CE without involatile detergent additives.[46] Recently, Cho *et al.*[47] demonstrated the analysis of urinary androgen glucuronides by CE-ESI-MS/MS; however, the method required a dynamically coated capillary under reverse polarity that suffers from poor stability due to anodic corrosion of the stainless steel needle.[48]

2.3 Experimental

2.3.1 Chemicals and Reagents

Twelve standard estrogens were purchased from Sigma-Aldrich (St. Louis, MO, USA), including five glucuronide conjugates: estrone 3-glucuronide ($E_1$3G), 17β-estradiol 3-glucuronide ($E_2$3G), β-estradiol 17-glucuronide ($E_2$17G), estriol 3-glucuronide ($E_3$3G), estriol 16-glucuronide ($E_3$16G), three sulfate conjugates, estrone 3-sulfate ($E_1$3S), β-estradiol 3-sulfate ($E_2$3S), estriol 3-sulfate ($E_3$3S), as well as four free estrogens: estrone (E_1), estradiol (E_2), 16-ethynyl estradiol (EE_2) and estriol (E_3). Stock solutions were prepared at concentrations of 3-15 mM in 1:1 MeOH:H_2O and stored at -20°C. Creatinine, uric acid, sodium chloride, melatonin, and HEPES were also purchased from Sigma-Aldrich. 10 mM stock solutions were prepared in water and stored at 4°C. A simulated urine matrix was used for preparation of estrogen standard solutions containing 15 mM sodium chloride, 1 mM creatinine, 500 μM uric acid, 150 μM melatonin (EOF marker) and 150 μM HEPES (internal standard). Ammonium bicarbonate (Sigma Aldrich Inc.) was prepared as a 500 mM stock in water and used as the background electrolyte for both sheath liquid and CE separations. HPLC-grade methanol

(Caledon Labs, Georgetown, ON, Canada) was used for preparation of all stock solutions and sheath liquid. All aqueous buffers and stock solutions were prepared with water purified using a Thermo Scientific Barnstead EasyPure II LF ultrapure water system (Cole Parmer, Vernon Hills, IL, USA).

2.3.2 Instrumentation

All CE-ESI-TOF-MS experiments were performed using an Agilent G7100A CE system (Mississauga, ON, Canada) interfaced with coaxial sheath liquid electrospray ion source to an Agilent 6224 TOF LC/MS orthogonal axis time-of-flight mass spectrometer (Mississauga, ON, Canada). Nitrogen gas was used as the nebulizer gas in ESI and the drying gas for the MS and helium gas was used as a damping and collision gas. The system software was 3D-CE ChemStation (CE) and Agilent MassHunter Workstation Data Acquisition (TOF-MS). Data processing was performed using MassHunter Qualitative and MassHunter Quantitative software. All data processing, electrophoregrams and surface response models were performed using Igor Pro 5.0 (Wavemetrics Inc., Lake Oswego, OR, USA).

2.3.3 CE-ESI-TOF-MS Conditions

An uncoated fused silica capillary (Polymicro Technologies, AZ, USA) with a 50 μm internal diameter and 85 cm length was used for all experiments and maintained at 20°C. Unless otherwise noted, as in the experimental design, all conditions were as follows: A 50 mM ammonium bicarbonate buffer adjusted to pH 9.5 with ammonium hydroxide was used as the background electrolyte. The sample was injected at 50 mbar for 5 seconds (approximately 5 nL) and the applied voltage was 20 kV. The sheath liquid was 5 mM ammonium bicarbonate in 80:20 MeOH:H_2O at a flow rate of 14 μL/min. The mass spectrometer was operated in negative-ion mode under the following conditions: ion spray voltage at -4.0 kV, drying gas temperature at 300°C, drying gas flow rate at 4 L/min and nebulizer gas flow rate at 10 L/min. The MS settings for ion extraction were fragmentor = -145 V, skimmer = -65 V and Oct 1 RF Vpp = 750 V. The mass range scanned was *m/z* 50 – 1100 with 6800 transients/scan. Purine and hexakis(2,2,3,3-tetrafluoropropoxy)phosphazine (HP-0921) were spiked into the sheath liquid at a concentration of 0.02% v and produced corresponding reference

ions at *m/z* 119.03632 and 981.9956 that were used for real-time internal mass-correction allowing for mass accuracy < 2 ppm in most cases.

2.3.4 Experimental Design

Experimental design was used to determine the optimum ESI conditions to maximize the ionization efficiency of free estrogens and their anionic conjugates. The three factors that were selected for optimization were capillary voltage (2.5 – 5.0 kV), fraction of methanol in the sheath liquid (40 – 80%), and sheath liquid flow rate (4 – 20 µL/min) since they have significant impacts on spray stability when using a coaxial sheath liquid interface that impact separation efficiency and concentration sensitivity in CE-MS due to potential suction and post-capillary dilution effects, respectively.[49] A two level (±1)-three factor (*i.e.*, 2^3) central composite design with six axial (± 1.7) and five central (0) conditions for a total of 20 experiments were performed to systematically maximize estrogen detectability.[50] The responses of E_3, $E_3$3S and $E_3$3G were measured by calculating the absolute peak areas or signal-to-noise (*S/N*) ratios of these ions in each experiment. Multiple linear regression of the data matrix was performed by Excel (Microsoft Inc., Redmond, WA, USA) and used to develop an empirical model based on three main factors and their interactions, which was refined iteratively by eliminating insignificant variables ($P < 0.05$) that had minimal effect on the overall predictive accuracy as reflected by changes in R^2 value. The optimized models were then used to graph 3D surface response plots in the experimental space for the three estriol metabolites. The optimum conditions for maximizing estrogen detection was determined to be a capillary voltage of -4.0 kV with 80% v/v methanol in the sheath liquid at a flow rate of 14 µL/min.

2.3.5 Molecular Modeling and Prediction of Relative Migration Times

Molecular volume (*MV*) was calculated as the Connolly solvent-excluded volume after minimization of energy using molecular mechanics (MM2) in Chem3D Ultra 8.0. pK_a values for estrogens were derived from literature,[45] which were used to determine z_{eff} in the background electrolyte (pH 9.5) using the Henderson-Hasselbalch equation. From these two physicochemical parameters (*MV* and z_{eff}) the relative migration time (RMT) for twelve estrogen metabolites (training set) were modeled using multiple linear regression (MLR).[38] A 12-fold cross-validation that randomly held out a estrogen metabolite (test set) was then

performed to assess overall model robustness. The predictive accuracy of the model was assessed in terms of correlation of determination (R^2) for the training set, the average correlation of determination for the test set (Q^2) and the average absolute bias of predicted RMT relative to average RMT (n=30) values measured experimentally by CE-TOF/MS.

2.3.6 Urine Collection and Sample Pretreatment

Urine samples were collected from three healthy female volunteers who were not taking any regular medication and had provided written informed consent prior to beginning the study. Urine samples were collected in sterile containers and stored at -80°C until analysis. A recent study has demonstrated that estrogen conjugates have excellent long-term stability when urine samples are stored frozen without addition of chemical preservatives.[51] Urine samples were centrifuged for 10 min at 150 *g* and supernatant urine was then analyzed directly by CE-TOF/MS after 10-fold dilution in de-ionized water. Solid-phase extraction (SPE) was also performed on the urine supernatant using Oasis HLB cartridges (Waters Inc., Milford, USA) with a vacuum manifold. The protocol for SPE of human urine was performed based on recent work reported by Qin *et al.*[29] with a few modifications to make the concentrated eluent compatible with CE-TOF/MS analysis. Before extraction, each HLB cartridge was preconditioned with 1.0 mL methanol followed by 1.0 mL water and 1.0 mL aqueous phosphoric acid (0.3%, v/v). A 5.0 mL aliquot of supernatant urine was mixed with 5.0 mL aqueous phosphoric acid (0.3%, v/v) and then loaded onto the HLB cartridge. After sample loading, the cartridge was washed with 1.0 mL water followed by 1.0 mL methanol:water:acetic acid (60:40:2 v). Elution of the analytes was performed using 1.0 mL methanol without addition of 2% ammonium hydroxide,[29] which was found to generate excessive band broadening of estrogen conjugates due to electrokinetic dispersion in CE-TOF/MS. The eluent was evaporated to dryness under a stream of nitrogen gas at 50°C and the resulting residue was reconstituted in 40 µL acetonitrile:water (75:25 v) containing 50 µM HEPES as an internal standard. By loading 5.0 mL of urine and reconstituting to 40 µL, the effective concentration of urinary estrogens was increased by about 125-fold, which has been found to offer good recoveries when using HLB columns ranging from 90-110%.[29]

2.3.7 Method Calibration and Assay Validation

Calibration standards were prepared as standards in a simulated urine matrix (15 mM sodium chloride, 1 mM creatinine, 500 µM uric acid) at six levels over a 100-fold concentration range from 0.5 to 50 µM using 50 µM HEPES as the internal standard as a way to improve precision for quantification in terms of relative peak areas (RPA) and metabolite identification in terms of relative migration times (RMT) for urinary metabolites. This is needed to correct for long-term drift in ion source and random changes in EOF when using CE-TOF/MS, respectively.[38] Ion suppression effects was investigated by comparison of the sensitivity (*i.e.*, slope of calibration curve) of three estrogen standards (*i.e.*, E_1, E_1-3G, E_1-3S) prepared in the mock urine solutions relative to spiked standards in authentic urine samples that were pooled from five pre-menopausal female volunteers and subsequently diluted 10-fold in de-ionized water. Intra-day reproducibility was determined by repeated measurements of estrogen standards in mock urine at three different concentration levels (1, 5 and 15 µM) in order to determine reproducibility of the method (n = 30). Limits of detection ($S/N \approx 3$) and limits of quantification ($S/N \approx 10$) for estrogen conjugates determined from the slope of the calibration curve and background noise measured at the specific *m/z* for each ion.

2.3.8 Ethics Approval

This study involving the collection of urine samples from female volunteers was approved by the McMaster University Research Ethics Board.

2.4 Results and Discussion

In this report, we introduce for the first time a CE-MS strategy for comprehensive profiling of free estrogens together with their intact glucuronide and sulfate conjugates using an unmodified capillary under normal polarity (**Figure 2.1**).

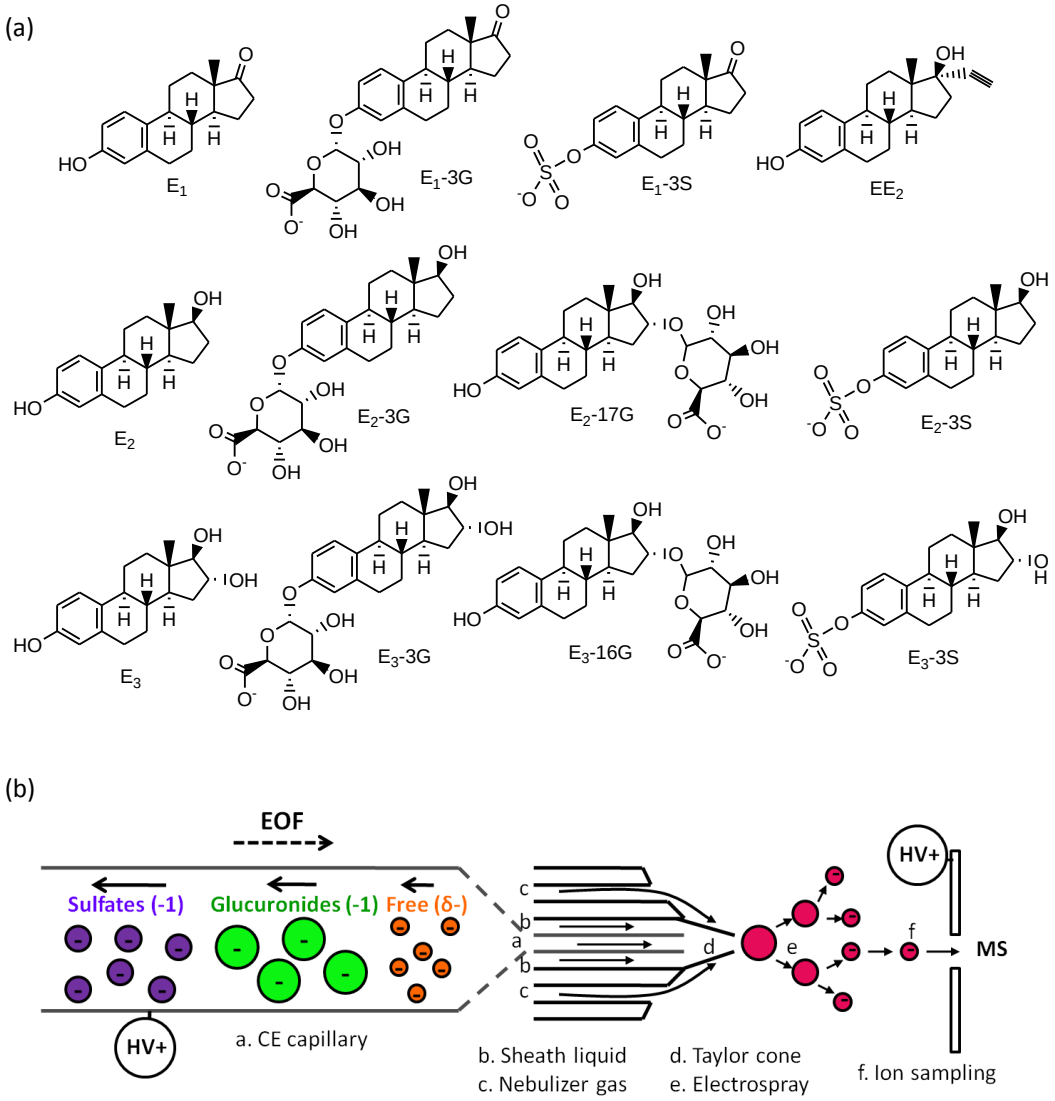
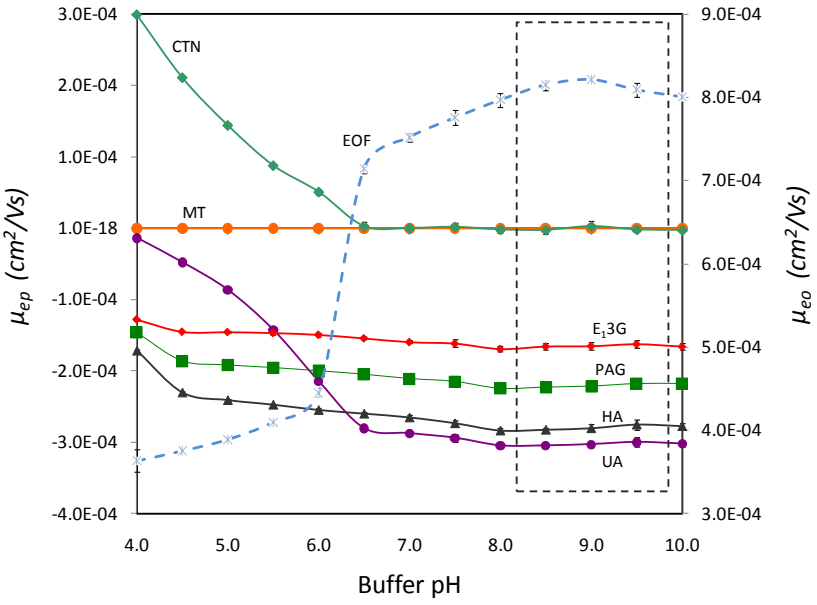


Figure 2.1. (a) 2D chemical structures of model weakly acidic estrogens and their anionic conjugates examined in this study, including estrone (E_1), estrone 3-glucuronide ($E_1$3G), estrone 3-sulfate ($E_1$3S), 17β-ethinyl 17β-estradiol (EE_2), 17β-estradiol (E_2), 17β-estradiol 3-glucuronide ($E_2$3G), β-estradiol 17-glucuronide ($E_2$17G), β-estradiol 3-sulfate ($E_2$3S), estriol (E_3), estriol 3-glucuronide ($E_3$3G), estriol 16-glucuronide ($E_3$16G) and estriol 3-sulfate ($E_3$3S). (b) Schematic depicting the resolution of anionic urinary metabolites by CE-TOF/MS based on their effective charge density, where weakly acidic estrogens and their ionic conjugates are desorbed into the gas-phase via a coaxial sheath liquid interface.

(a)



(b)

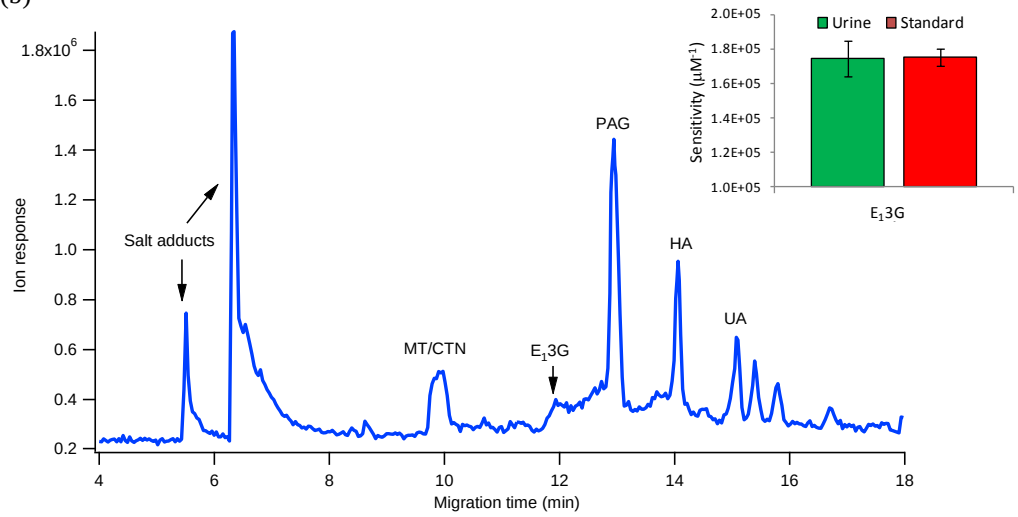


Figure 2.2(a) Overlay plot showing the pH dependent changes in the electrophoretic mobility (μ_{ep}) and electroosmotic mobility (μ_{eo}) highlighting a stable EOF and interference-free region for estrogen analysis under alkaline buffer conditions in CE and (b) representative total ion electropherogram of a 10-fold diluted female urine sample at pH 9.5 under full-scan negative ion mode detection showing resolution of estrogen conjugates from salts (cations) and major anionic urinary metabolites by CE-ESI-TOF/MS. The inset highlights the lack of ion suppression effects on apparent sensitivity for E1G when spiked in diluted pooled urine samples relative to standard solutions. Metabolite acronyms refer to MT, melatonin (EOF marker); CTN, creatinine; $E_1$3G. estrone 3-glucuronide; PAG, phenylacetylglutamine; HA, hippuric acid; and UA, uric acid. CE-ESI-TOF/MS conditions were 50 mM ammonium acetate as the background electrolyte, 85 cm total capillary length, separation voltage of 15 kV, and cone voltage of -4.0 kV.

Figure 2.2(a) depicts an overlay of the electrophoretic mobility (μ_{ep}) of major urinary metabolites as well as the average electroosmotic mobility (μ_{eo}) of a bare fused-silica capillary as a function of buffer pH when analyzing urine samples spiked with 10 µM estrone-3-glucuronide ($E_1$3G). A 10-fold dilution of urine specimens in deionized water prior to injection was found to provide stable separations with reproducible migration times due to lowering of the high salt matrix. All CE-TOF/MS experiments were performed under negative ion mode detection with full-scan ion monitoring (*m/z* 50-1100). Resolution of $E_1$3G from other abundant anionic metabolites (*e.g.,* hippuric acid, HA) together with a stable electroosmotic flow (EOF) is achieved under alkaline buffer conditions (pH 8.5-10). **Figure 2.2(b)** demonstrates that a low ion background is evident between the EOF and major urinary metabolites that reduces matrix-induced ion suppression of measured estrogen responses when using LC-MS/MS.[3,4,27] This feature was confirmed by comparing the slope (*i.e.,* sensitivity) of calibration curves for standard estrogen solutions (*i.e.,* E_1, $E_1$3G and $E_1$3S) with spiked samples in pooled female urine as shown in **Figure 2.3.**

Although later eluting estrogen sulfates (e.g., $E_1$3S) migrate close to phenylacetylglutamine (PAG), this was not observed to result in ion suppression under these conditions. Overall, there was a no significant difference ($P < 0.01$) in measured analytical sensitivity for estrogen conjugates in diluted urine samples as depicted in the inset of **Figure 2.2(b)**. This compares well with the performance of LC-MS with atmospheric pressure photoionization (APPI) for estrogen analysis that has better matrix tolerance than ESI at the expense of reduced sensitivity.[52] Thus, direct quantification of metabolites in complex biological samples can be realized by CE-ESI-MS when using a coaxial sheath liquid interface provided that adequate sensitivity is achieved.[53,54] In this study, free estrogens overall have about a 50-fold weaker ion response relative to the bulkier anionic estrogen conjugates under negative ion mode detection due to their lower molecular volume and reduced charge density.[54,55]

Regioselective phase I/II metabolic transformations impact the biological activity of circulating estrogens and their isomers in vivo. **Figure 2.4(a)** highlights the impact of buffer pH conditions for the selective resolution of two pairs of positional isomers of estrogen conjugates, namely, $E_2$3G/$E_2$17G and $E_3$3G/$E_3$16G. **Figure 2.4(b)** highlights the TOF/MS spectra for the deprotonated molecular ions $[M-H]^-$ as related to two pairs of unresolved estrogen conjugate

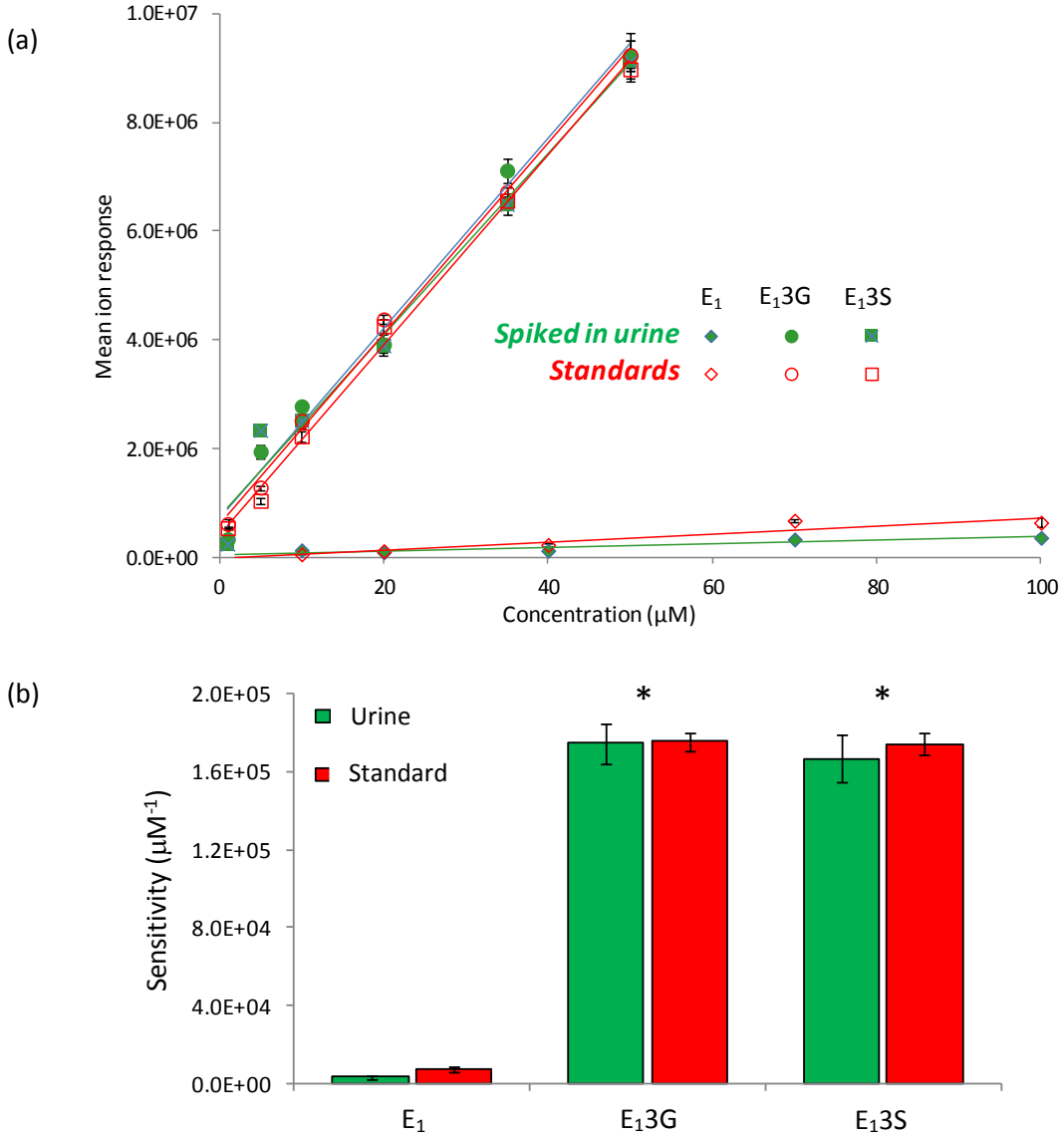


Figure 2.3. (a) Overlay of representative calibration curves for native and conjugates estrone (E_1) prepared as standard solutions in a simulated urine matrix (15 mM sodium chloride, 1 mM creatinine, 500 µM uric acid) and spiked standards in pooled female urine (n=5). (b) Bar graph comparing the apparent sensitivity for model estrogens derived from the slope of calibration curves after a 10-fold dilution in de-ionized water. Overall, no significant matrix-induced ion suppression (* $P < 0.01$) was observed for estrogen conjugates in diluted urine samples. Note the 50-fold lower ion response for free E_1 under negative ion mode detection relative to bulkier and strongly ionic estrogen conjugates.

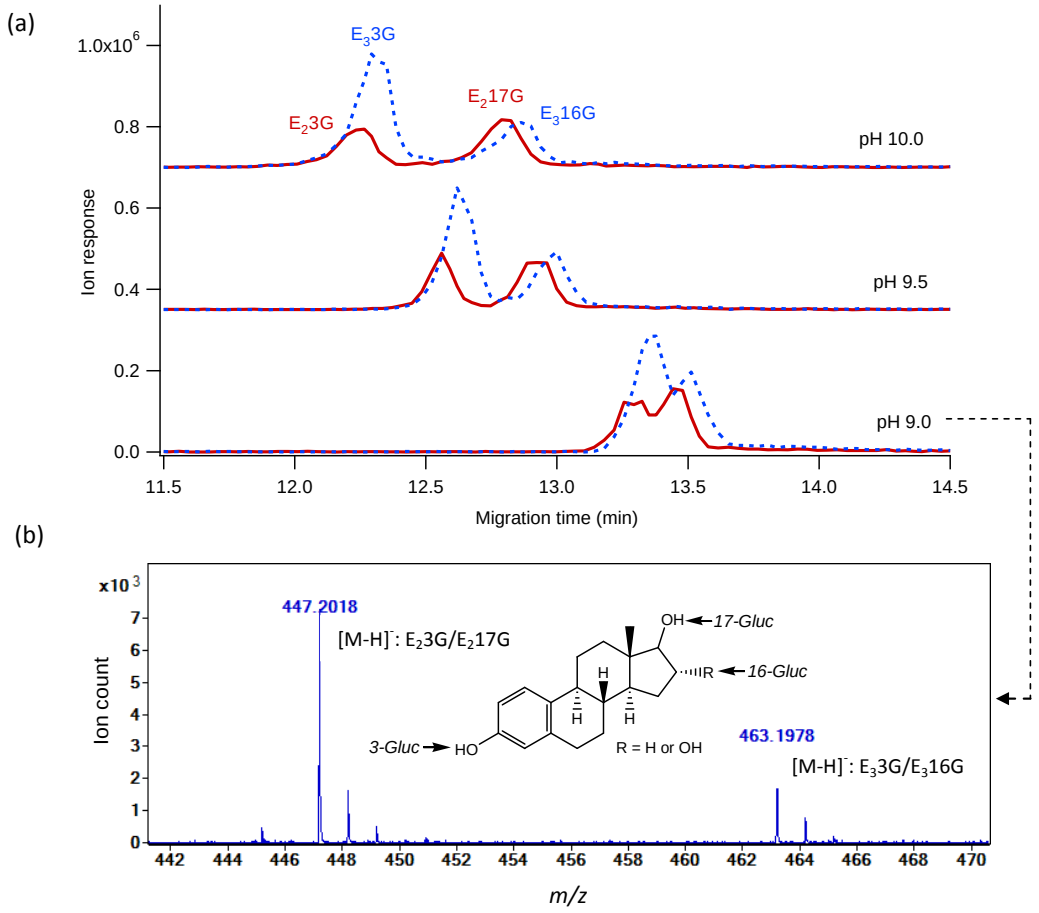


Figure 2.4. (a) Resolution of two pairs of estrogen glucuronide position isomers by CE based on optimization of buffer pH due to selective ionization of unconjugated weakly acidic phenolic moiety and (b) representative TOF-MS spectra under negative-ion mode highlighting molecular ions $[M-H]^-$ for unresolved isobaric/isomeric ions, namely $E_2$3G/$E_2$17G and $E_3$3G/$E_3$16G. The CE-ESI-TOF/MS conditions used are similar to Figure 2.2.

isomers at pH 9 with an average mass accuracy of 1.1 ppm. Since estrogen glucuronides undergo common neutral losses of 176 Da under low collision energies, identification of the site of conjugation is only feasible using MS/MS by altering their fragmentation pathways via chemical derivatization.[56] Conjugation at distal hydroxyl sites from the aromatic ring allows for their resolution by CE due to their higher negative mobility under alkaline conditions (pH > 9.5) as a result of the weak acidity of the free phenolic hydroxyl. Thus, 16/17-estrogen

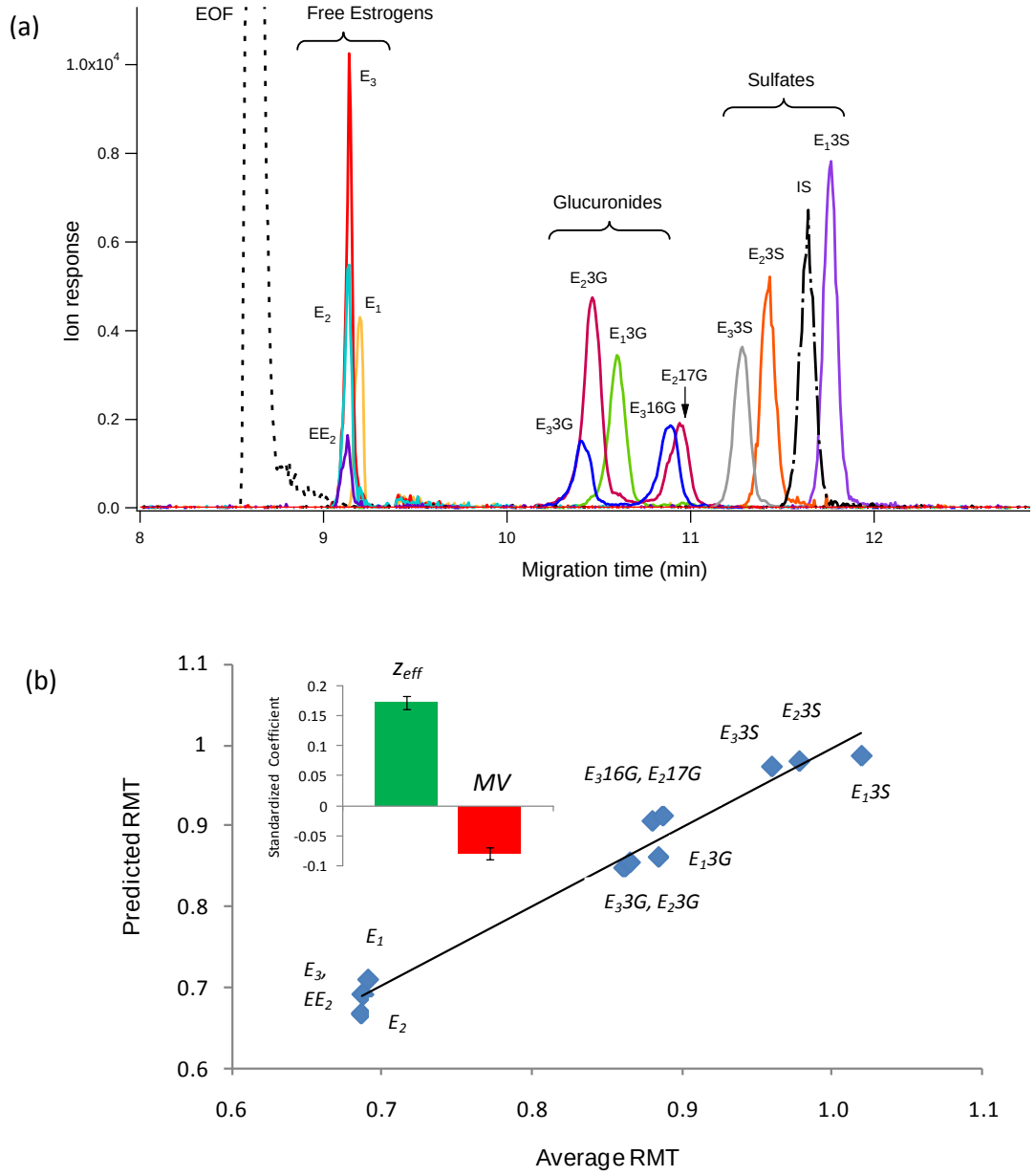


Figure 2.5. (a) Overlay of extracted ion electropherograms showing simultaneous resolution of 20 µM free estrogens and 10 µM estrogen conjugates using 50 mM ammonium bicarbonate, pH 9.5 by CE-TOF/MS. (b) Multivariate model for predicting the relative migration time (RMT) of estrogen metabolites relative to an internal standard (HEPES) based on two intrinsic physicochemical parameters of an ion, namely, effective charge (z_{eff}) and molecular volume (*MV*). Multiple linear regression of standardized variables was performed for model estrogens in order to predict their RMT (y) based on the equation: $y = (0.8407 \pm 0.0062) + (0.173 \pm 0.010)z_{eff} - (0.080 \pm 0.010)MV$, where $R^2 = 0.9750$ and $Q^2 = 0.9543$ when performing 12-fold cross-validation. CE-ESI-TOF/MS conditions are similar to Figure 2.2 except that voltage of 20 kV was used for separation.

glucuronides migrate with larger negative mobility and longer migration times than their corresponding 3-regioisomer. Although estrogen sulfate positional isomers were not examined in this study since commercial standards were not available, a similar principle governing their separation by CE is still applicable.[57] **Figure 2.5(a)** depicts an extracted ion electropherogram overlay highlighting the simultaneous analysis of free estrogens and their intact anionic conjugates by CE-TOF/MS, where their migration time increases as a function of higher charge density with free (unconjugated) estrogens < glucuronides < sulfates. Given the lack of ion suppression effects for urinary estrogen conjugates under these conditions, a single non-isotopic internal standard (HEPES) was used to improve robustness in terms of corrections for long-term drift of the ion source, as well as migration time variance due to changes in EOF.[57,58] In this case, the relative migration time (RMT) for twelve model estrogens was accurately predicted in CE when using multiple linear regression based on two intrinsic physicochemical parameters estimated in silico,[38] namely, molecular volume (*MV*) and effective charge (z_{eff}) as derived from the pK_a of the phenolic moiety as summarized in **Table 2.1**. The latter parameter is only required for weakly acidic free estrogens and their distal 16/17-conjugates since strongly acidic glucuronides and sulfates are fully ionized under the alkaline buffer conditions. Overall, an excellent agreement was found between average and predicted RMTs for 12 model estrogens with cross-validation as reflected by a R^2 = 0.9755, Q^2 = 0.9543, and a mean absolute bias of 1.9% that approaches experimental precision (*CV* < 1%) as depicted in **Figure 2.5(b).** This is significant as it not only allows for modeling of ion migration behavior,[57] but also supports TOF/MS for identification of unknown metabolites among several putative isobaric and/or isomeric candidates in cases when authentic standards are lacking and/or databases are incomplete.[59]

A major challenge remains adequate concentration sensitivity for submicromolar detection of estrogen metabolites in human urine. Various experiments were performed to maximize the detectability for estrogens by CE-TOF/MS under negative ion mode detection, including the electrolyte co-ion (*e.g.*, acetate, fluoride, hydroxide), ionic strength and pH in separation buffer and/or sheath liquid composition (data not shown). However, 50 mM ammonium bicarbonate, pH 9.5 was found to be the optimum electrolyte for resolution of estrogens while maintaining a low background noise with a stable current profile relative to other buffer systems. An advantage of the coaxial sheath liquid interface in CE-MS is the ability to independently optimize aqueous buffer conditions required for resolution in CE distinct from ionization processes in ESI-

Table 2.1. Identification of estrogens by CE-TOF/MS based on accurate mass and RMT predicted based on two intrinsic physicochemical properties.

Estrogen Metabolites	***Measured RMT[a]***	***m/z [M-H]⁻***	***- z_{eff}[b] (pH 9.5)***	***MV[c] (Å³)***	***Predicted RMT[d]***	***Bias (%)***
Estradiol (E_2)	0.687 ± 0.002	271.1704	0.11	262	0.692	+0.68
Ethynyl estradiol (EE_2)	0.687 ± 0.002	295.1704	0.21	283	0.668	-2.74
Estriol (E_3)	0.688 ± 0.001	287.1653	0.13	268	0.692	+0.52
Estrone (E_1)	0.691 ± 0.002	269.1547	0.14	257	0.710	+2.74
Estriol 3-glucuronide ($E_3$3G)	0.861 ± 0.005	463.1974	1.00	401	0.848	-1.49
Estradiol 3-glucuronide ($E_2$3G)	0.865 ± 0.001	447.2024	1.00	396	0.855	-1.16
Estriol 16-glucuronide ($E_3$16G)	0.880 ± 0.001	463.1974	1.13	395	0.906	+2.98
Estrone 3-glucuronide ($E_1$3G)	0.884 ± 0.001	445.1868	1.00	391	0.862	-2.54
Estradiol 17-glucuronide ($E_2$17G)	0.887 ± 0.002	447.2024	1.14	393	0.913	+2.90
Estriol 3-sulfate ($E_3$3S)	0.960 ± 0.001	367.1221	1.00	307	0.974	+1.49
Estradiol 3-sulfate ($E_2$3S)	0.978 ± 0.001	351.1272	1.00	302	0.980	+0.27
Estrone 3-sulfate ($E_1$3S)	1.019 ± 0.001	349.1115	1.00	297	0.987	-3.14

a. Measured RMTs for metabolites were measured in mocked urine samples (n=10), where error represents ±1σ.

b. Calculated from measured estrogen pK_a at pH 9.5, where $z_{eff} = 1/(10^{[pH-pK_a]}+1)$.

c. Chem3D Ultra 8.0 after MM2 energy minimization as Connelly Solvent-Excluded Volume.

d. Predicted RMTs determined by MLR based on two physicochemical properties (z_{eff}, MV), where model equation was $y = (0.8407 \pm 0.0062) + (0.173 \pm 0.010)z_{eff} - (0.080 \pm 0.010)MV$, where $R^2 = 0.9750$ and $Q^2 = 0.9543$ when performing 12-fold cross-validation.

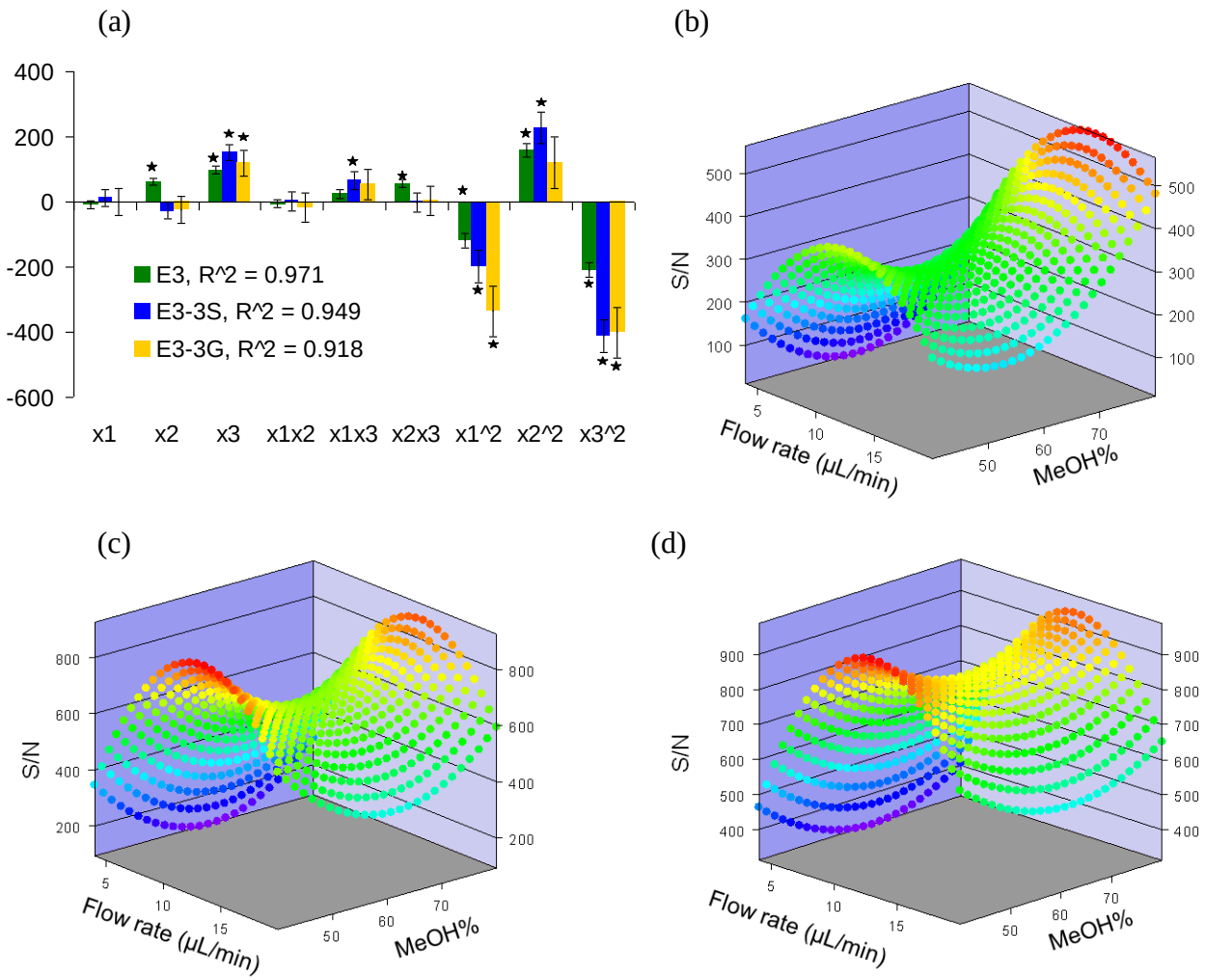


Figure 2.6.(a) Experimental design for optimization of ionization conditions for estrogens when using a coaxial sheath liquid interface in CE-ESI-TOF/MS under negative ion mode, where coefficients from multiple linear regression represent three factors (*x1* = ESI voltage, *x2* = MeOH% in the sheath liquid, *x3* = sheath liquid flow rate; ∗=significant at 95% CL) and their second-order terms. Representative bimodal surface response curves as a function of signal/noise (*S/N)* for b) E_3 (*m/z* 287.1653) c) $E_3$3S (*m/z* 367.1221) and d) $E_3$3G (*m/z* 463.1974) are depicted, where optimal conditions to enhance ionization efficiency of estrogens were a cone voltage of 4.0 kV using 80% MeOH as the sheath liquid at a flow rate of 14 µL/min.

MS. However, the sheath liquid and nebulizer spray can also contribute to deleterious peak broadening and reduced sensitivity in CE-MS due to suction and post-capillary dilution effects if not carefully investigated.[49] In this case, a face-centered central composite experimental design was performed to generate surface response curves for model estrogens as a way to systematically optimize ion responses based on three factors, including MeOH content, sheath flow rate, and capillary voltage (**Figure 2.6**). The bimodal response curves demonstrated a global maximum in terms of signal/noise (*S/N*) for both free estrogens and their anionic conjugates when using 80% MeOH/5 mM ammonium bicarbonate as the sheath liquid at a flow rate of 14 µL/min with a capillary voltage set at -4 kV. Figures of merit for the optimized CE-TOF/MS assay for estrogens are summarized in **Table 2.2**. Overall, good linearity (R^2 > 0.990) over a 100-fold concentration range with detection limits of about 0.4 µM (*S/N* ≈ 3) was achieved with excellent precision when using a single nonisotopic internal standard as reflected by an average *CV* of 11% and 0.2% (*n* = 30) for estrogen quantification and RMT, respectively. However, concentration sensitivity was found to be still inadequate for the direct detection of endogenous estrogens in urine samples by CE-TOF/MS, which has about a 20-fold higher detection limit for intact estrogen conjugates than LC-MS/MS.[25] The only major urinary steroid conjugates directly detected by CE-TOF/MS in diluted urine samples were androsterone 3-glucuronide (A3G, *m/z* 465.2500; *RMT* 0.910 ± 0.015) and androsterone 3-sulfate (A3S, *m/z* 369.1741; *RMT* 0.994 ± 0.010) as depicted in **Figure 2.7.**

Although estrogen levels are elevated during pregnancy, their concentrations levels are highly variable for women during the menstrual cycle, which are also dependent on several factors, including age, diet, health status, and drug usage. For instance, mean $E_1$3G levels in urine for menstruating females have been reported to vary from a low baseline level of about 15 nM and peak to 80 nM for a few days following luteinizing hormone surge.[60] Recent LC-MS methods for analysis of deconjugated estrogens[1] or intact estrogen conjugates[29] in human urine have required SPE and/or chemical derivatization with multiple reaction monitoring by tandem MS detection in order to achieve adequate sensitivity. However, isobaric interferences are still problematic for quantification of low abundance urinary estrogens even after SPE workup and LC-MS/MS analysis.[29] In our case, preliminary studies were also performed using SPE in order to improve concentration sensitivity prior to CE-TOF/MS, but estrogen conjugates were not detected in enriched premenopausal female urine extracts. However, other exogenous phytoestrogen and endogenous steroid conjugates

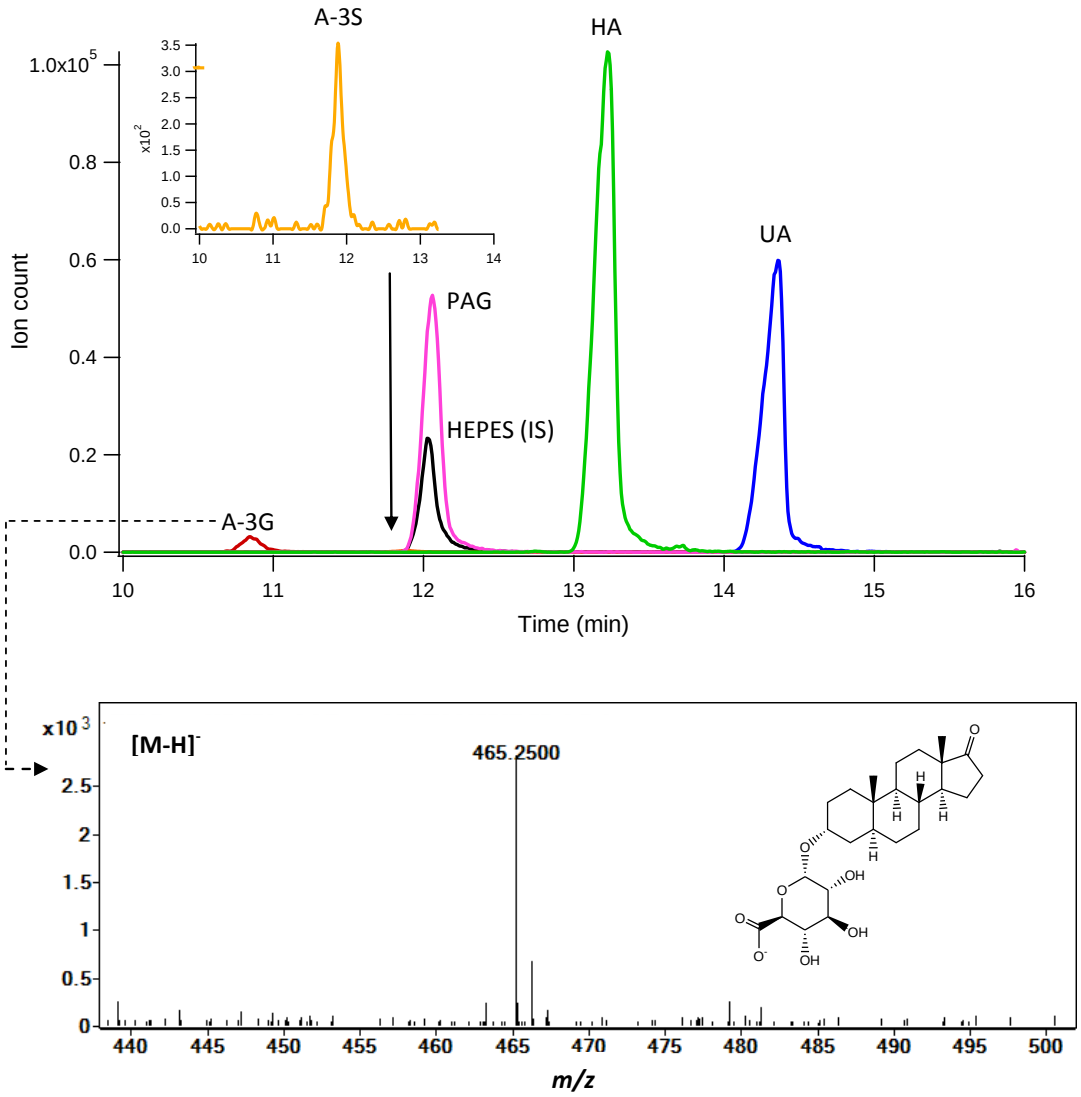


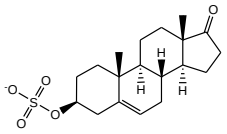
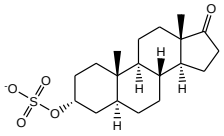
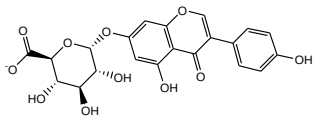
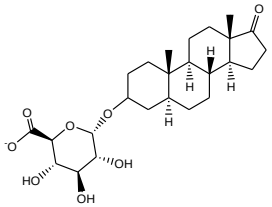
Figure 2.7. Representative extracted ion electropherogram overlay of major urinary metabolites and low abundance steroid conjugates over a 10^3 dynamic range in a pre-menopausal female urine specimen directly analyzed by CE-TOF/MS after a 10-fold dilution in de-ionized water. Analyte acronyms refer to androsterone 3-glucuronide (A-3G, *m/z* 465.2494, RMT 0.910 ± 0.015), androsterone 3-sulfate (A-3S, *m/z* 369.1741, RMT 0.994 ± 0.010), phenylacetylglutamine (PAG, *m/z* 263.1038; RMT 1.001 ± 0.012), hippuric acid (HA, *m/z* 178.0510, RMT 1.105 ± 0.010), uric acid (UA, *m/z* 167.0211, RMT 1.200 ± 0.008) and HEPES as internal standard (IS). Accurate mass and isotope ratio measurements by TOF/MS together with RMT predictions by CE can be used to confirm the identity or exclude isobaric/isomeric candidate ions selected after a database search (*e.g.*, KEGG and Human Metabolome Database). Low abundance estrogen conjugates were not reliably detected in diluted urine specimens from non-pregnant females due to inadequate sensitivity when using CE-TOF/MS in full-scan negative ion mode detection (LOD ≈ 0.4 µM).

Table 2.2. Figures of merit for estrogen conjugate analysis by CE-TOF/MS under negative ion mode detection.

Estrogen Metabolites	***Measured RMT[a]***	***m/z [M-H]⁻***	***LOD (S/N=3, μM)***	***Linearity[b] (R^2)***	***Precision RPA[c] (CV, n=30)***	***Precision RMT[c] (CV, n=30)***
Estriol 3-glucuronide ($E_3$3G)	0.861 ± 0.005	463.1974	0.17	0.991	10.9	0.09
Estradiol 3-glucuronide ($E_2$3G)	0.865 ± 0.001	447.2024	0.07	0.993	13.2	0.11
Estriol 16-glucuronide ($E_3$3G)	0.880 ± 0.001	463.1974	0.15	0.988	11.8	0.11
Estrone 3-glucuronide ($E_1$3G)	0.884 ± 0.001	445.1868	0.17	0.989	11.2	0.12
Estradiol 17-glucuronide ($E_2$17G)	0.887 ± 0.002	447.2024	0.13	0.997	10.8	0.17
Estriol 3-sulfate ($E_3$3S)	0.960 ± 0.001	367.1221	0.23	0.996	11.2	0.14
Estradiol 3-sulfate ($E_2$3S)	0.978 ± 0.001	351.1272	0.12	0.997	8.6	0.06
Estrone 3-sulfate ($E_1$3S)	1.019 ± 0.001	349.1115	0.07	0.995	10.0	0.04

a. Measured RMTs for metabolites were measured in mocked urine samples measured as ten replicates at three different concentrations (n=30), where error represents ±1σ.
b. Calibration curves derived from linear regression of average normalized ion responses (n=3) of estrogens relative to internal standard at six different concentrations over a 100-fold range from 0.5-50 μM
c. Precision was determined by performing ten replicate analysis of estrogen standards at three different concentration levels (1, 5 and 15 μM)

Table 2.3. Identification of putative urinary steroid conjugates based on accurate mass (< 3 ppm) and relative migration time (RMT) by CE-TOF/MS after 125-fold off-line preconcentration using solid-phase extraction.

Putative Urinary Metabolite	*Empirical Formula*	*2D Chemical Structure*	*m/z [M-H]⁻*	*δ_{mass} [a] (ppm)*	*Measured RMT*
Dihydroepiandrosterone 3-sulfate (DHEAS)[b]	$C_{19}H_{28}O_5S$		367.1576	-1.29	1.021 ± 0.002
Androsterone 3-sulfate[b] (A-3S)	$C_{19}H_{30}O_5S$		369.1740	0.27	1.015 ± 0.003
Genistein 7-glucuronide (Gen-7G)	$C_{21}H_{18}O_{11}$		445.0774	0.45	1.175 ± 0.017
Androsterone 3-glucuronide (A-3G)[b]	$C_{25}H_{38}O_8$		465.2506	-2.58	0.950 ± 0.002

a. Refers to mass error (δ_{mass}) defined as the relative difference between measured (TOF/MS) and theoretical m/z as derived from its putative chemical structure
b. Putative steroid conjugates identified in urine also possess low abundance/isobaric ions, namely (epi)testosterone 3-sulfate (m/z 367.1585), etiocholanolone 3-sulfate (m/z 369.1741) or etiocholanolone 3-glucuronide (m/z 465.2494)

were tentatively identified in this case, including the major soya isoflavone metabolite, genistein 7-glucuronide (Gen7G, *m/z* 445.0776; *RMT* 1.175 ± 0.017), and dehydroepiandrosterone 3-sulfate (DHEAS, *m/z* 367.1585; *RMT* 1.021 ± 0.002) as summarized in **Table 2.3.** The latter 17-ketosteroid conjugate together with A3G and A3S are forensic urine signatures of human origin,[61] whereas the isoflavone aglycone precursor of Gen7G possesses therapeutic properties for prevention of breast and prostate cancer.[62]

2.5 Conclusions

In summary, CE-TOF/MS offers a complementary platform to LC-MS for resolution of free estrogens together with their intact anionic conjugates with excellent specificity and low matrix-induced suppression effects in human urine.

The separation mechanism for estrogen metabolites by CE can be accurately modeled in terms of their intrinsic physicochemical parameters, which allows for accurate prediction of RMT to support their qualitative identification. CE-TOF/MS enables resolution of anionic estrogen conjugates and their positional isomers that is particularly useful for strongly ionic phytoestrogen diconjugates[62] and glutathionylated estrogen adducts.[63] Given the limited sensitivity of full-scan negative ion mode TOF/MS for detection of endogenous urinary estrogen metabolites, future work will focus on enhancing concentration sensitivity when using tandem MS (*e.g.,* Q-TOF), in conjunction with online sample preconcentration techniques and/or recent low flow/sheathless interfaces for CE-ESI-MS.[64,65]

2.6 Acknowledgement

P.B.M. wishes to acknowledge funding support from National Science and Engineering Research Council of Canada. N.L.K. also acknowledges support in the form of a Queen Elizabeth II Ontario Graduate Scholarship in Science and Technology. We wish to also thank Dr. Julie Marr from Agilent Technologies Inc. for helpful assistance.

2.7 References

(1) Ziegler, R. G.; Faupel-Badger, J. M.; Sue, L. Y.; Fuhrman, B. J.; Falk, R. T.; Boyd-Morin, J.; Henderson, M. K.; Hoover, R. N.; Veenstra, T. D.; Keefer, L. K.; Xu, X. *J. Steroid Biochem. Molec. Biol.* **2010**, *121*, 538–545.

(2) Taioli, E.; Im, A.; Xu, X.; Veenstra, T. D.; Ahrendt, G.; Garte, S. *Reprod. Biol. Endocrinol.* **2010**, *8*, 93.

(3) Xu, X.; Keefer, L. K.; Ziegler, R. G.; Veenstra, T. D. *Nat. Prot.* **2007**, *2*, 1350–1355.

(4) Gabet-Giraud, V.; Miege, C.; Herbreteau, B.; Hernandez-Raquet, G.; Coquery, M. *Anal. Bioanal. Chem.* **2010**, *396*, 1841–1851.

(5) Vandenberg, L. M.; Chahoud, I.; Heindel, J. J.; Padmanabhan, V. *Environ. Heal. Perspec.* **2010**, *118*, 1055–1070.

(6) Kunisue, T.; Tanabe, S.; Isobe, T.; Aldous, K. M.; Kannan, K. *J. Agric. Food Chem.* **2010**, *58*, 9838–9846.

(7) Grace, P. B.; Mistry, N. S.; Carter, M. H.; Leathem, A. J. C.; Teale, P. *J. Chromatogr. B* **2007**, *853*, 138–146.

(8) Blair, I. A. *Steroids* **2010**, *75*, 297–306.

(9) Raftogianis, R.; Creveling, C.; R.Weinshilboum; Weisz, J. *J. Natl. Cancer Inst. Monogr.* **2000**, *27*, 113–124.

(10) Briciu, R. D.; Kot-Wasik, A.; Namiesnik, J. *J. Chromatogr. Sci.* **2009**, *47*, 127–139.

(11) Giese, R. W. *J. Chromatogr. A* **2003**, *1000*, 401–412.

(12) Rothman, M. S.; Carson, N. E.; Xu, M.; Wang, C.; Swerdloff, R.; Lee, P.; Goh, V. H. H.; Ridgway, E. C.; Wierman, M. E. *Steroids* **2011**, *76*, 177–182.

(13) Giton, F.; Caron, P.; Berube, R.; Belanger, A.; Fiet, J. *Clin. Chim. Acta* **2010**, *411*, 1208–1213.

(14) Knust, U.; Strowitzki, T.; Spiegelhandler, B.; Bartsch, H.; Owen, R. W. *Rapid Comm. Mass Spectr.* **2007**, *21*, 2245–2254.

(15) Courant, F.; Antignac, J. P.; Maume, D.; Monteau, F.; Andersson, A. M.; Skakkebaek, N.; Andre, F.; Bizec, B. Le. *Anal. Chim. Acta* **2007**, *586*, 105–114.

(16) Croley, T. R.; Hughes, R. J.; Koenig, B. G.; Metcalfe, C. D.; March, R. E. *Rapid Comm. Mass Spectr.* **2000**, *14*, 1087–1093.

(17) Falk, R. T.; Xu, X.; Keefer, L.; Veenstra, T. D.; Ziegler, R. G. *Cancer Epidemiol. Biomarkers* **2008**, *17*, 3411–3418.

(18) Kushnir, M. M.; Rockwood, A. L.; Roberts, W. L.; Yue, B.; Bergquist, J.; Meickle, A. W. *Clin. Biochem.* **2011**, *44*, 77–88.

(19) Zhang, H.; Henion, J. *Anal. Chem.* **1999**, *71*, 3955–3964.

(20) Clarke, D. B.; Lolyd, A. S.; Botting, N. P.; Oldfield, M. F.; Needs, P. W.; Wiseman, H. *Anal. Biochem.* **2002**, *309*, 158–172.

(21) Isobe, T. *J. Chromatogr. A* **2003**, *984*, 195–202.

(22) Tso, J.; Aga, D. S. *Anal. Chem.* **2011**, *83*, 269–277.

(23) D'Ascenzo, G.; Corcia, A. Di; Gentili, a; Mancini, R.; Mastropasqua, R.; Nazzari, M.; Samperi, R.; Di Corcia, a. *Sci. Total Environ.* **2003**, *302*, 199–209.

(24) Caron, P.; Audet-Walsh, E.; Lepine, J.; Belanger, A.; Guillemette, C. *Anal. Chem.* **2009**, *81*, 10143–10148.

(25) Yang, Y. J.; Lee, J.; Choi, M. H.; Chung, B. C. *Biomed. Chromatogr.* **2003**, *17*, 219–225.

(26) Matejicek, D. *J. Chromatogr. A* **2011**, *1218*, 2292–2300.

(27) Tso, J.; Dutta, S.; Inamdar, S.; Aga, D. S. *J. Agric. Food Chem.* **2011**, *59*, 2212–2222.

(28) Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D. *J. Sep. Sci.* **2010**, *33*, 716–727.

(29) Qin, F.; Zhao, Y.-Y.; Sawyer, M. B.; Li, X.-F. *Anal. Chem.* **2008**, *80*, 3404–3411.

(30) Nguyen, H. P.; Yang, S. H.; Wigginton, J. G.; Simpkins, J. W.; Schug, K. A. *J. Sep. Sci.* **2010**, *33*, 793.

(31) Dutta, S.; Inamdar, S.; Tso, J.; Aga, D. S.; Sims, J. T. *J. Environ. Qual.* **2010**, *39*, 1688–1698.

(32) Tai, S. S.; Wlech, M. J. *Anal. Chem.* **2005**, *77*, 6359–6363.

(33) Lin, Y.-H.; Chen, C.-Y.; Wang, G.-S. *Rapid Comm. Mass Spectr.* **2007**, *21*, 1973–1983.

(34) Yang, W.-C.; Regnier, F. E.; Silva, D.; Adamec, J. *J. Chromatogr. B* **2008**, *870*, 233–240.

(35) Kushnir, M. M.; Rockwood, A. L.; Bergquist, J.; Varshavsky, M.; Roberts, W. L.; Yue, B.; al., et. *Am. J. Clin. Pathol.* **2008**, *129*, 530–539.

(36) Britz-McKibbin, P. *Methods Mol. Biol.* **2011**, *708*, 229.

(37) Ramautar, R.; Mayboroda, O. A.; Somsen, G. W.; de Jong, G. J. *Electrophoresis* **2011**, *32*, 52–65.

(38) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(39) Chalcraft, K. C.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 307–314.

(40) Harino, H.; Tsunoi, S.; Sato, T.; Tanaka, M. *Fresenius J. Anal. Chem.* **2001**, *369*, 546–547.

(41) Silva, C. A.; Pereira, E. A.; Micke, G. A.; Farah, J. P. S.; Tavares, M. F. M. *Electrophoresis* **2007**, *28*, 3722–3730.

(42) Liu, Y.; Jia, L. *Microchem. J.* **2008**, *89*, 72–76.

(43) Katayama, M. *Biomed. Chromatogr.* **2003**, *17*, 263–267.

(44) Siren, H.; Seppanen-Laakso, T.; Oresic, M. *J. Chromatogr. B* **2008**, *871*, 375–382.

(45) Hurwitz, A.; Liu, S. *J. Pharm. Sci.* **1977**, *66*, 624–627.

(46) Britz-McKibbin, P.; Ichihashi, T.; Tsubota, K.; Chen, D. D. Y.; Terabe, S. *J. Chromatogr. A* **2003**, *1013*, 65–76.

(47) Cho, S.-H.; Lee, J.; Choi, M. H.; Lee, W.-Y.; Chung, B. C. *Biomed. Chromatogr.* **2009**, *23*, 426–433.

(48) Soga, T.; Igarashi, K.; Itoh, C.; Mizobuchi, K.; Zimmermann, H.; Tomita, M. *Anal. Chem.* **2009**, *81*, 6165.

(49) Mokaddem, M.; Gareil, P.; Belgaied, J.-E.; Varenne, A. *Electrophoresis* **2009**, *30*, 1692.

(50) Kaiser, C.; Segui-Lines, G.; D'Amaral, J. C.; Ptolemy, A. S.; Britz-McKibbin, P. *Chem. Comm.* **2008**, 338.

(51) Fuhrman, B. J.; Xu, X.; Falk, R. T.; Hankinson, S. E.; Veenstra, T. D.; Keefer, L. K.; Ziegler, R. G. *Inter. J. Biol. Markers* **2010**, *25*, 185–194.

(52) Chen, H.-C.; Kuo, H.-W.; Ding, W.-H. *Chemosphere* **2009**, *74*, 508–514.

(53) Ramautar, R.; Toraño, J. S.; Somsen, G. W.; de Jong, G. J. *Electrophoresis* **2010**, *31*, 2319.

(54) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 2506–2515.

(55) Oss, M.; Kruve, A.; Herodes, K.; Leito, I. *Anal. Chem.* **2010**, *82*, 2865–2872.

(56) Lampinen-Salomonsson, M.; Bondesson, U.; Petersson, C.; Hedeland, M. *Rapid Comm. Mass Spectr.* **2006**, *20*, 1429–1440.

(57) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403–415.

(58) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 2506–2515.

(59) Sugimoto, M.; Hirayama, A.; Robert, M.; Abe, S.; Soga, T.; Tomita, M. *Electrophoresis* **2010**, *31*, 2311.

(60) Alliende, M. E. *Fertil. Steril.* **2002**, *78*, 90–95.

(61) Nakazono, T.; Kashimra, S.; Hayashiba, Y.; Hara, K. *J. Forensic Sci.* **2002**, *47*, 568–572.

(62) Hosoda, K.; Furuta, T.; Yokokawa, A.; Ishii, K. *Anal. Bioanal. Chem.* **2010**, *397*, 1563–1572.

(63) Peng, K.-W.; Chang, M.; Wang, Y.-T.; Wang, Z.; Qin, Z.; Bolton, J. L.; Thatcher, G. R. *J. Chem. Res. Toxicol.* **2010**, *23*, 1374–1383.

(64) Maxwell, E. J.; Zhong, X.; Zhang, H.; van Zeijl, N.; Chen, D. D. Y. *Electrophoresis* **2010**, *31*, 1130.

(65) Busnel, J.-M.; Schoenmaker, B.; Ramautar, R.; Carrasco-Pancorbo, A.; Ratnayake, C.; Feitelson, J. S.; Chapman, J. D.; Deelder, A. M.; Mayboroda, O. A. *Anal. Chem.* **2010**, *82*, 9476.

Chapter III

Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High Throughput Platform for Metabolomics with High Data Fidelity

III. Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High Throughput Platform for Metabolomics with High Data Fidelity

3.1 Abstract

A major constraint in large-scale mass spectrometry (MS)-based metabolomic initiatives is the low sample throughput associated with chromatographic or electrophoretic separations. Herein, we introduce multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) as a multiplexed separation platform for metabolomics that increases sample throughput up to one order of magnitude while improving overall data fidelity. We demonstrate that serial injection of seven or more discrete sample segments can be performed within a single capillary while maintaining isomeric resolution without ion suppression when using a high mass resolution time-of-flight-MS. Customized injection sequences can be devised to encode information temporally within a separation based on signal pattern recognition, which enables unambiguous identification and accurate quantification (mean bias < 10%) of polar metabolites in human plasma with good reproducibility ($CV \approx 10\%$, $n = 70$). False discoveries are avoided when using a rigorous dilution trend filter to reject spurious signals and background peaks that comprise the majority ($\approx$ 65%) of total detectable features. MSI-CE-MS offers an unprecedented approach to enhance sample throughput analogous to direct infusion-MS ($\approx$ 3 min/sample) while delivering far greater selectivity, quantitative performance, and data quality since the same ion from different samples migrates into the ion source within a short time interval (≈2-6 min). These distinct analytical and bioinformatic merits are achieved without column switching, isotopic labeling, hardware modifications, or costly infrastructure investments.

3.2 Introduction

Separation science plays a critical role in mass spectrometry (MS)-based metabolomics that is relevant for new advances in biomarker discovery and personalized medicine.[1] High efficiency separations improve selectivity when analyzing complex biological samples while offering a complementary approach to MS for identification of unknown metabolites.[2–4] Nevertheless, a major trade-off is the low sample throughput associated with solute elution and column pre-conditioning. High-throughput metabolome profiling can be achieved by direct-infusion[5,6] or ambient ionization[7,8] MS techniques. However, isomers/isobars are

not resolved and ion suppression reduces sensitivity with numerous artifact signals that can contribute to false discoveries.[9] Also, reliable quantification is not feasible without access to stable-isotope internal standards. As a result, large-scale metabolomic initiatives require intermittent quality controls with time alignment/peak picking algorithms to correct for system drift when comparing data derived from single[10] or multiple separation platforms.[11] Thus, there is an urgent need to enhance sample throughput in an untargeted metabolomic workflow while ensuring high data fidelity for translational research.

3.3 Experimental

3.3.1 Chemicals and Reagents

Ammonium acetate, acetic acid, formic acid, 3-chloro-*L*-tyrosine (Cl-Tyr) and all other standards of metabolites and their isomers were purchased from Sigma-Aldrich (St. Louis, MO, USA). 10 mM stock solutions of polar metabolites were prepared in water and stored at 4°C. HPLC-grade acetonitrile (Honeywell, Muskegon, MI, USA) and methanol (Caledon, Georgetown, ON, Canada) were used for preparation of background electrolyte (BGE) and sheath liquid respectively. Ammonium acetate was prepared as a 400 mM stock in water and the pH was adjusted to 5.0 using acetic acid. All aqueous buffers and stock solutions were prepared with de-ionized water purified using a Thermo Scientific Barnstead EasyPure II LF ultrapure water system (Cole Parmer, Vernon Hills, IL, USA).

3.3.2 Instrumentation and Injection Configuration

All MSI-CE-MS experiments were performed using an Agilent G7100A CE system (Agilent Technologies Inc., Mississauga, ON, Canada) interfaced with a coaxial sheath liquid Jetstream electrospray ion source with heated nitrogen gas to an Agilent 6230 TOF-MS. The system software was Agilent MassHunter Workstation Data Acquisition and data processing was performed using MassHunter Qualitative Analysis software. An uncoated fused silica capillary (Polymicro Technologies, AZ, USA) with 50 μm ID and 110 cm length was used for all experiments when using an applied voltage of 30 kV at 25°C. The BGE was 1 M formic acid with 15% *v* acetonitrile (pH 1.8). The injection configuration used in MSI-CE-MS was based on a serial hydrodynamic injection at 100 mbar

using alternating segments of sample (standard or plasma filtrate) and spacer (BGE). The total number of samples injected (1 to 10), as well as the length of sample and spacer segments (5 s to 60 s at 100 mbar) were examined during initial optimization of MSI-CE-MS. For instance, a 7-sample segment format used for untargeted metabolome profiling of human plasma comprised of the following injection sequence: (1) 5 s sample, 40 s spacer; (2) 5 s sample, 40 s spacer; (3) 5 s sample, 40 s spacer; (4) 5 s sample, 40 s spacer; (5) 5 s sample, 40 s spacer; (6) 5 s sample, 40 s spacer; (7) 5 s sample, 5 s spacer. The total time for sample introduction was about 5 min, which filled about 30% of the total capillary length. Plasma filtrate samples were diluted 4-fold in 200 mM NH_4Ac (pH 5) containing 25 µM Cl-Tyr as internal standard. Between runs the capillary was flushed for 10 min with BGE at 950 mbar. The sheath liquid was 60:40 MeOH:H_2O containing 0.1% formic acid at a flow rate of 10 µL/min via 100:1 splitter. Purine and hexakis(2,2,3,3-tetrafluoro-propoxy)phosphazine (HP-921) were added into the sheath liquid at a concentration of 0.02% v and used as reference ions at *m/z* 121.050873 and *m/z* 922.009798for real-time internal mass calibration.The TOF-MS was operated in positive-ion mode over a *m/z* 50-1700 with an acquisition rate of 2 Hz and acquisition time of 500 ms. The ESI source settings used were: Vcap = 2000 V, nozzle voltage = 2000 V, nebulizer gas = 10 psi, sheath gas = 3.5 L/min at 195 °C, drying gas 8 L/min at 300 °C, whereas MS voltage settings were fragmentor = 120 V, skimmer = 65V and Oct 1 RF = 750 V.

3.3.3 Human Plasma Filtrates and Method Validation

Pooled human female blood samples (n = 9) were collected via an in-dwelling catheter inserted in a forearm vein. Vacutainers with EDTA as anticoagulant (Becton Dickson, Franklin Lakes, NJ, USA) were used for blood collection. Blood specimens were gently mixed, placed on ice and then centrifuged at 2500 g for 5 min at 4°C to fractionate plasma from blood cells. Human plasma samples was then aliquoted and stored at -80°C. Frozen plasma aliquots were then slowly thawed/vortexed while on ice and then diluted 4-fold with ammonium acetate (pH 5) with 25 µM internal standard (Cl-Tyr). Ultrafiltration was used for plasma deproteinization using a 3 kDa MWCO Nanosep centrifugal device (Pall Life Sciences, Washington, NY, USA) at 13000 g for 15 min. A 20 µL aliquot of the pooled plasma filtrate sample was used for analysis by MSI-CE-MS. Calibration curves were performed in triplicate (n = 3) by serial dilution (100, 50, 25, 10, 5.0, 2.0, 0.5 µM) of standard mixtures of polar metabolite standards (Sigma-Aldrich

Inc., Oakville, ON, Canada) in 200 mM ammonium acetate, pH 5.0 with 25 μM Cl-Tyr as internal standard. Similarly, standard addition calibration curves were performed in triplicate (n=3) by spiking in increasing concentrations of a metabolite standard in a four-fold diluted plasma filtrate sample containing 25 μM Cl-Tyr as internal standard. Depending on the intrinsic abundance of metabolite, different concentrations were spiked into plasma in order to generate a significant response change relative to baseline levels.

3.4 Results and Discussion

Capillary electrophoresis (CE) offers a convenient approach for high throughput screening as highlighted by multiplexed gene sequencers using massively "in parallel" capillary array systems.[12] Improved sample throughput can also realized by injecting multiple sample segments "in-series" within a single capillary as used in enzyme-catalyzed reactions,[13] bioaffinity measurements,[14] online chemical derivatizations,[15] multidimensional separations,[16] and pharmaceutical analyses.[17,18] Unlike chromatographic separations that often rely on gradient elution programs, CE-MS makes use of an isocratic aqueous buffer and coaxial sheath liquid interface that provides a homogenous solution for solute electromigration with ionization under steady-state conditions.[19] The latter principle is critical for accurate quantification of ions during separation when using electrospray ionization (ESI). **Figure 3.1** depicts the configuration used for multisegment injection (MSI)-CE-MS based on hydrodynamic injection of multiple segments of sample that are spatially positioned between background electrolyte (BGE) zones (*i.e.*, spacer) within a bare fused-silica capillary. In this case, about a third of the total capillary length is loaded with alternating sample/spacer segments prior to voltage application and zonal separation when using an acidic BGE under positive ion mode ESI-MS. Due to a maximum voltage setting of 30 kV in commercial CE instruments,[20] a total capillary length of 110 cm (or 270 V/cm) is optimum for achieving high separation efficiency ($N \approx 10^5$) without deleterious band dispersion. Three parameters were optimized to maximize sample throughput while ensuring adequate peak capacity in order to resolve metabolites from adjacent sample segments, namely, sample length, BGE spacer length, and effective total capillary length. A time-of-flight (TOF)-MS serves as a high mass resolution multi-channel detector for resolving comigrating ions that may overlap between certain sample plugs based on their characteristic

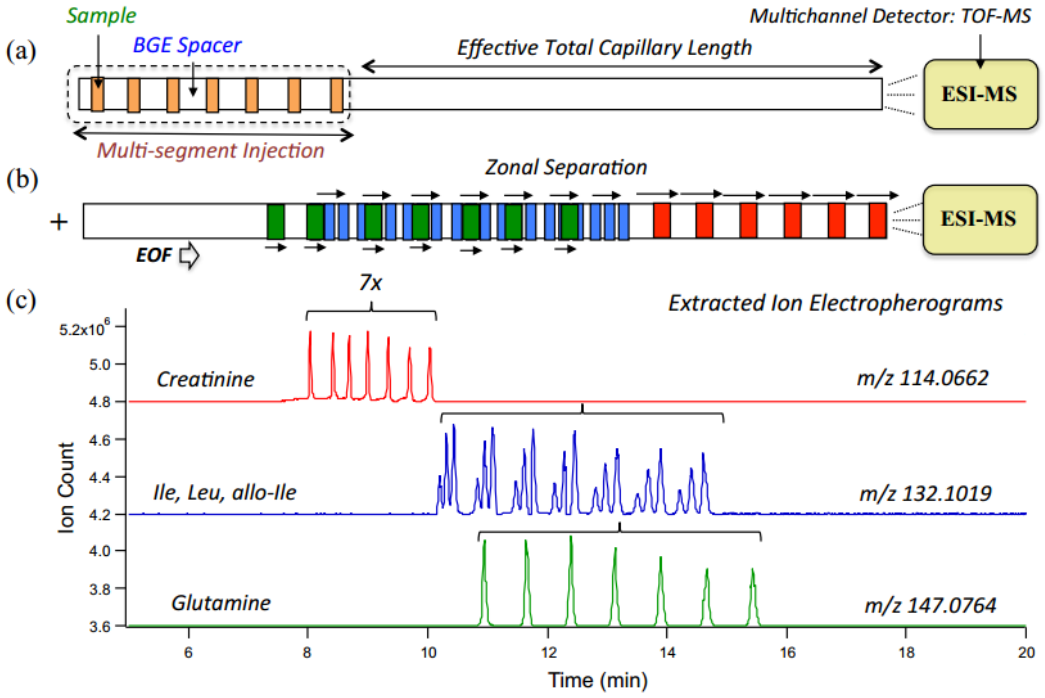


Figure 3.1. (a) Multiplexed separation of seven discrete sample segments within a single capillary by MSI-CE-MS in a single capillary, where (b) ions migrate as a series of zones in free solution prior to ionization. (c) This format enables reliable quantification of polar metabolites and their isomers in different samples since ionization occurs within a short time interval (≈ 2-6 min) under steady-state conditions when using a high mass resolution TOF-MS.

m/z (*e.g.*, Leu, Gln). Similarly, differential ion migration by CE allows for resolution of isomers/isobars prior to ionization, such as Ile, Leu, and *allo*-Ile.[21] **Figure 3.1** shows extracted ion electropherograms (EIEs) when using a 7 sample-segment format in MSI-CE-MS, where metabolite standards are detected as a series of seven equivalent bands with later migrating peaks broader due to increased longitudinal diffusion.

Figure 3.2 highlights that MSI-CE-MS is directly applicable to metabolome profiling of human plasma, where the effective separation window is defined between the salt front (*i.e.*, Na^+) and the electroosmotic flow (EOF). All plasma samples are first diluted 4-fold in ammonium acetate containing an internal standard (IS) prior to ultrafiltration to remove protein. **Figure 3.2(a)** compares base peak electropherograms when analyzing one, five or ten discrete

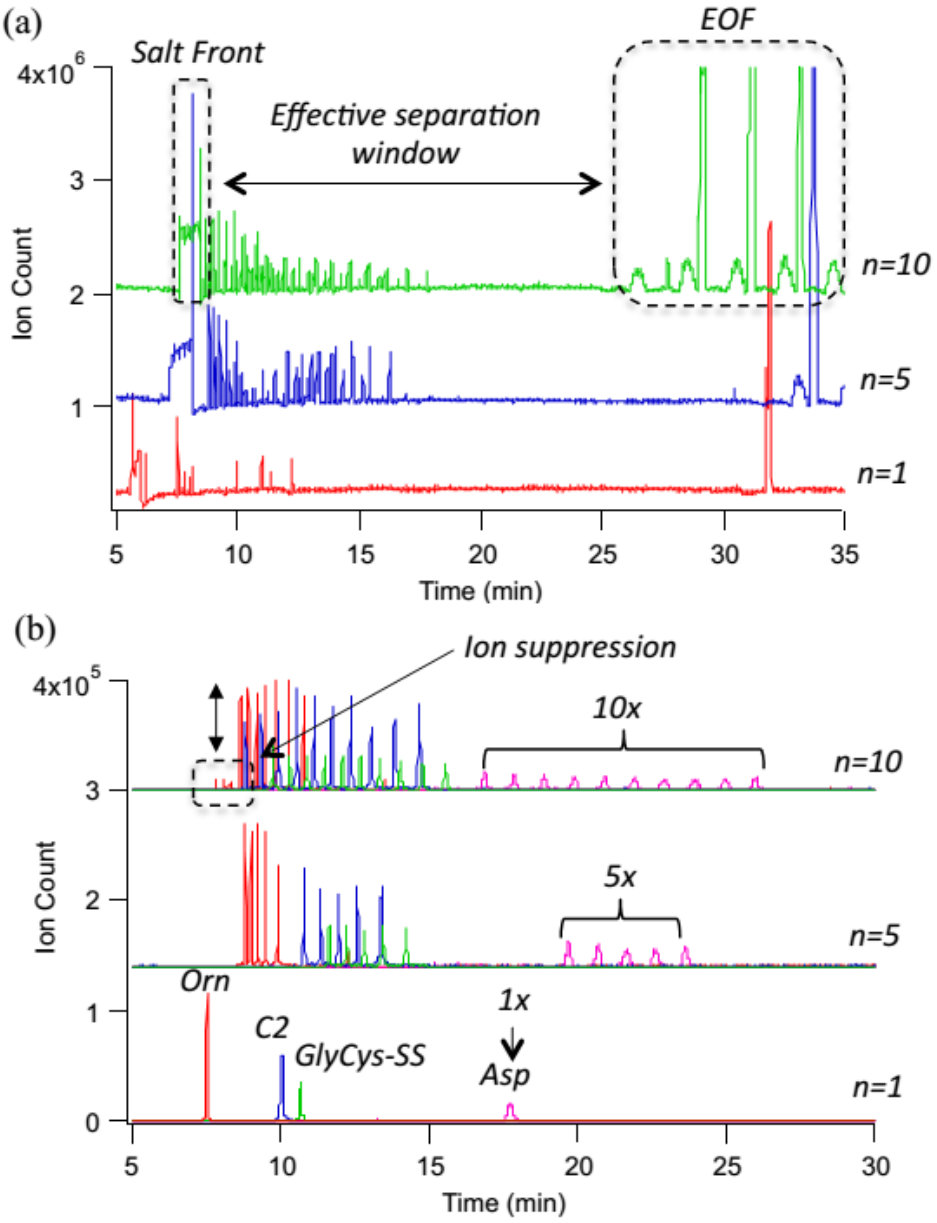


Figure 3.2. High-throughput metabolome profiling based on serial injection of 1, 5 or 10 sample segments of pooled human plasma by MSI-CE-MS highlighting (a) base peak and (b) EIEs for representative plasma metabolites, including key regions of ion suppression. The effective separation window where metabolites can be resolved and quantified without deleterious ion suppression is defined as the time interval between the salt front (*e.g.*, Na^+) and electroosmotic flow (EOF). The number of molecular features detected (> 300 counts) within the effective separation window corresponds to 96, 591 and 752 total ions when analysing 1, 5 and 10 plasma filtrate samples within the same capillary, respectively.

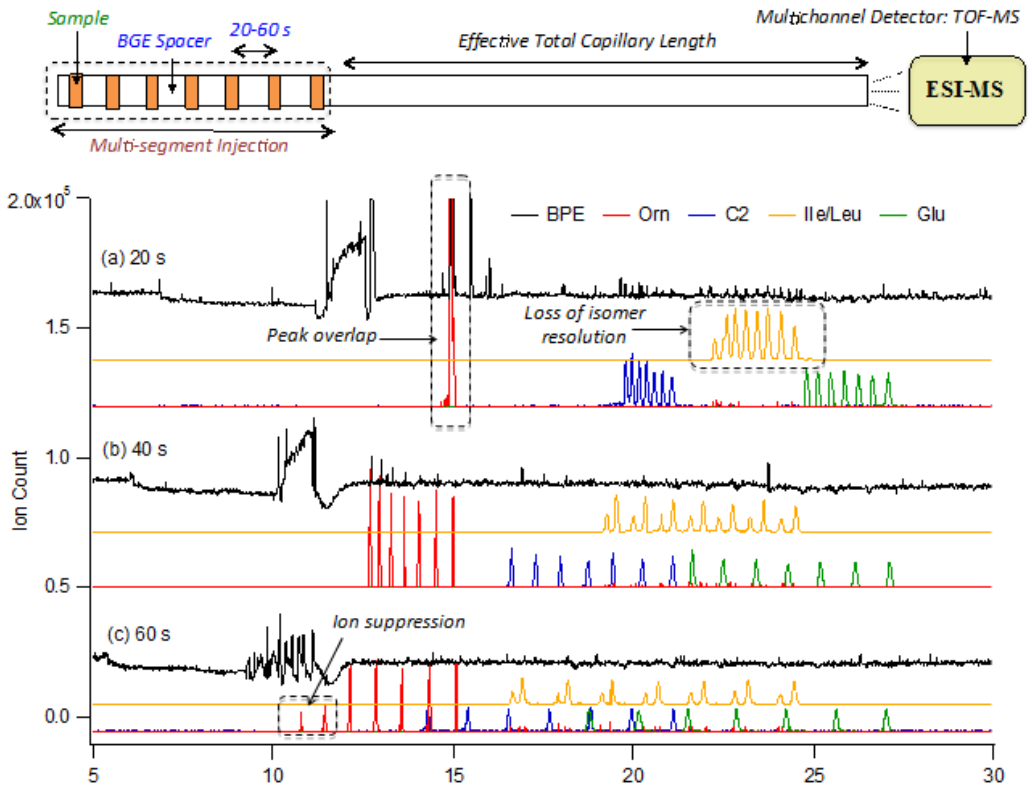


Figure 3.3. Overlay of base peak electropherograms and EIEs for selected plasma metabolites and their isomers highlighting the impact of increasing the BGE spacer length using a hydrodynamic injection (a) 20 s, (b) 40 s, or (c) 60 s at 100 mbar between all seven discrete sample segments of pooled plasma filtrate (5 s at 100 mbar). It is apparent that the use of (a) shorter spacers (20 s) results in poor resolution of early migrating metabolites with high positive mobility (*e.g.*,Orn) along with inadequate resolution of isomers due to peak overlap with adjacent sample plugs (*i.e.*, Ile/Leu), whereas (c) longer BGE spacers between sample segments leads to ion suppression of some early migration ions (*i.e.*,Orn) due to comigration with the salt front due to the reduced effective capillary length for separation. The use of (b) an intermediate spacer length (40 s) was found to offer optimum resolution of various classes of metabolites and their isomers from adjacent sample segments while avoiding ion suppression from the salt front and EOF.

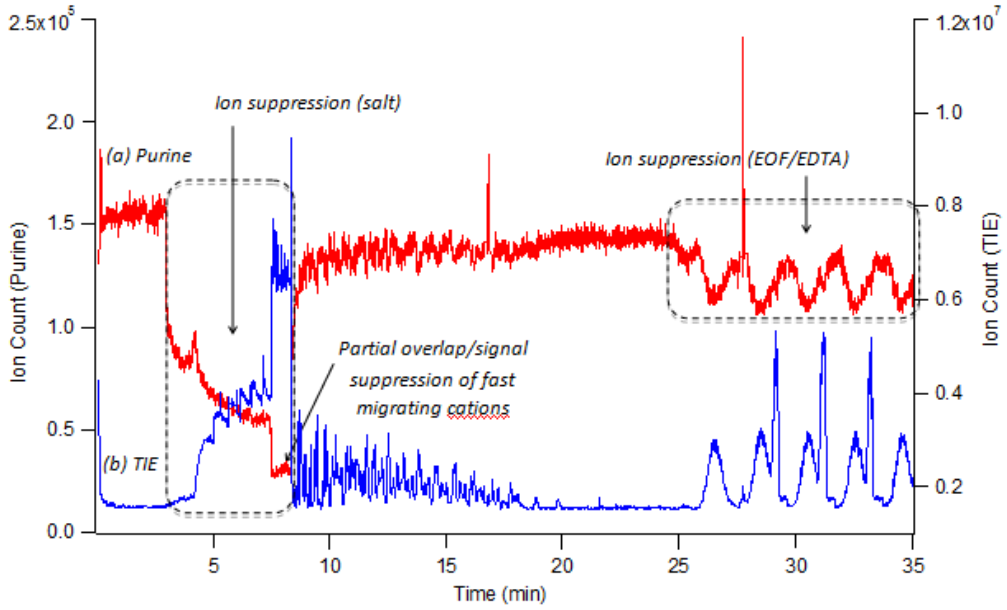


Figure 3.4. An overlay of the ion current trace for purine (*m/z* 121.0509) which is an additive in the sheath liquid (60:40 MeOH:H_2O with 0.1% v formic acid)to serve as an internal mass calibrant for the TOF-MS, and (b) total ion electropherogram when using a 10 sample segment injection of pooled plasma filtrates by MSI-CE-MS. Major signal suppression of purine occurs at two time intervals during the separation corresponding to the salt front (*e.g.*, Na^+, K^+) and EOF/EDTA regions containing abundant/involatile co-ions. Due to the reduced effective capillary length for separation, there is partial overlap of fast migrating metabolites (*e.g.*, cationic amino acids from the first three sample segments) with the salt front, which reduces the total number of molecular features detected due to partial ion suppression.

plasma filtrate samples simultaneously in a single run. Increasing the number of sample segments injected within a fixed capillary length reduces the effective peak capacity notably for early eluting ions that migrate close to the salt front. **Figure 3.2(b)** highlights EIEs for representative plasma metabolites with mobilities that span the separation window, including *L*-ornithine (Orn), *O*-acetyl-*L*-carnitine (C2), cysteinylglycine disulfide (GlyCys-SS), and *L*-aspartic acid (Asp). Sample throughput is enhanced by over one order of magnitude (≈3 min/sample) by MSI-CE-MS as demonstrated by a series of ten signal responses detected for C2, GlyCys-SS and Asp; however, the first three peaks for Orn (and other cationic amino acids) undergo ion suppression due to comigration with the salt front. Careful optimization of the serial injection configuration is required as shorter spacers result in poor resolution of metabolites and their isomers between adjacent sample segments, whereas longer spacers lead to overlap of fast migrating cations with involatile salts **(Figure 3.3**).

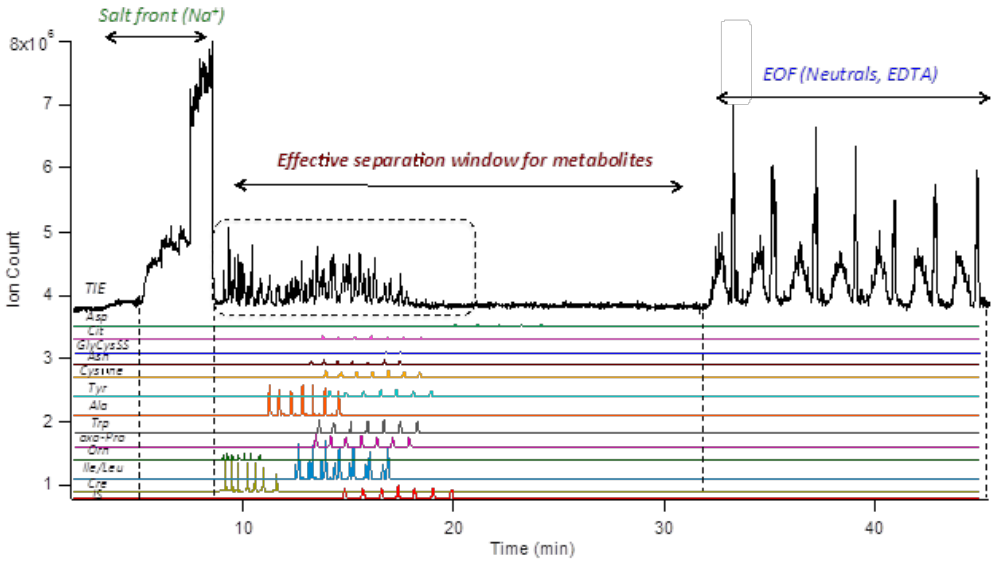


Figure 3.5. High-throughput metabolome profiling of seven discrete plasma filtrate samples by MSI-CE-MS in a electrophoretic separation, where reliable quantification of cationic metabolites and their isomers occurs within the effective separation window (≈ 9-32 min) between regions of ion suppression associated with abundant/involatile co-ions of the salt front (*e.g.*, Na^+, K^+) and EOF, where the series of seven broad and sharp peaks represent co-migrating neutral metabolites (*e.g.*, glucose) and EDTA derived from blood collection, respectively. A 7-sample segment injection format was selected as a compromise between sample throughput and metabolome coverage when analyzing human plasma, where extracted ion electropherograms (EIEs) highlight that various classes of cationic metabolites from plasma are detected as a series of 7 equivalent peaks provided they migrate between salt front and EOF. Compared to a conventional single injection format (96 molecular features), a proportional increase in detected peaks was realized when using a 7-sample segment injection format (669 molecular features) without evidence of ion suppression between salt front and EOF regions.

In this case, untargeted metabolite profiling using molecular feature extraction (> 300 counts, *m/z* 50-1700) resulted in 96, 591 and 752 total ions detected when analyzing 1, 5 and 10 plasma samples by MSI-CE-MS, respectively (**Figure 3.2(a)**). Thus, there is about a 20% loss in detectable ions with a 10-fold increase in sample throughput due to partial ion suppression as evident in current traces for purine that is used as an internal mass calibrant (**Figure 3.4**). In this case, a 7-sample segment format in MSI-CE-MS was selected as a compromise between sample throughput and metabolome coverage as reflected by 669 total ions detected in plasma (**Figure 3.5**), which confirms no loss in information content relative to a single sample injection. Hence, CE serves

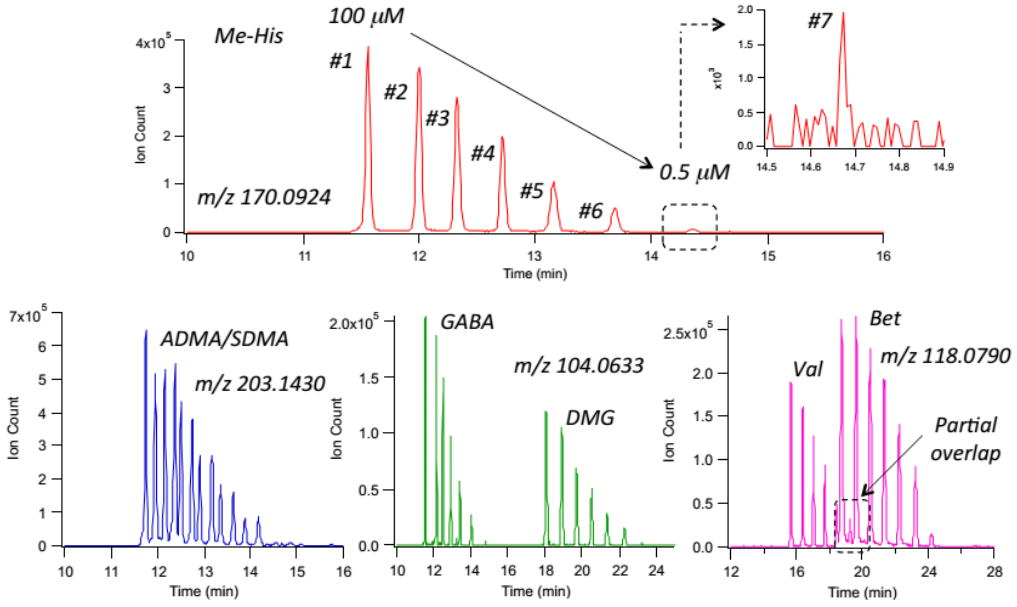


Figure 3.6. Multiplexed analysis using a 7-segment sample injection by MSI-CE-MS for single-step acquisition of external calibration curves for polar metabolites and their isomers over a 200-fold concentration range. In most cases, isobars/isomers are resolved by CE in free solution prior to ESI-MS with the exception of ions with intermediate differences in mobility, where partial peak overlap can occur between certain sample segments.

as an effective desalter to reduce ion suppression/enhancement effects originating from highly saline biofluids or reagents used during sample handling (*e.g.*, EDTA as anticoagulant for blood collection). Moreover, since metabolites from different samples enter the ion source within a short time interval (≈ 2-6 min), signal stability in ESI-MS is better than conventional CE/LC-MS methods that require several hours (≈ 3-4 h) to complete an equivalent batch of runs when analyzing "one sample at a time".

Figure 3.6 highlights the multiplexed advantage of MSI-CE-MS when acquiring an external calibration curve based on a seven-point dilution series of standards (100–0.5 µM) within a single separation, such as 3-methyl-*L*-histidine (3-MeHis), as well as pairs of isomers (*asymmetric/symmetric*-dimethyl-*L*-arginine, SDMA/ADMA) and isobaric ions (γ-aminobutyric acid, GABA, dimethylglycine, DMG; *L*-valine, Val, betaine, Bet). MSI-CE-MS retains

Table 3.1. External calibration curve data for 13 model polar metabolites with good linearity (R^2> 0.98) over a wide dynamic range (> 100-fold) with low micromolar detection limits. LOD was calculated based on the error of the calibration curve using 95% confidence interval.

Metabolite	*m/z*	*Linear range (μM)*	*Linearity (R^2)*	*LOD (μM)*	*Sensitivity (μM^{-1})*	*Concentration in pooled human plasma (μM, n=7)*
Ala	90.0550	0.50-100	0.9995	1.4	0.0168	122 ± 2
Creatinine	114.0662	0.50-100	0.9831	6.6	0.0741	14 ± 2
Creatine	132.0768	0.50-100	0.9954	3.4	0.0333	12 ± 3
Leu	132.1019	0.50-100	0.9998	0.8	0.0664	25 ± 1
Orn	133.0972	0.50-100	0.9978	2.3	0.0434	8.9 ± 2.0
HyX	137.0458	0.50-100	0.9990	2.1	0.0341	0.2 ± 0.1
Deoxycarnitine	146.1176	0.50-100	0.9917	4.6	0.1064	2.5 ± 0.7
Gln	147.0764	0.50-100	0.9998	0.8	0.0306	109 ± 6
C0	162.1125	0.50-100	0.9928	4.3	0.1035	6.3 ± 1.1
Cit	176.1030	0.50-100	0.9997	1.2	0.0368	5.5 ± 1.2
Tyr	182.0812	0.50-100	0.9991	1.9	0.0143	34 ± 2
C2	204.1230	0.50-100	0.9863	5.9	0.1238	2 ± 3
Cystine	241.0311	0.50-100	0.9997	0.9	0.0316	1.3 ± 1.1

excellent separation performance in cases when isomers have small or large differences in mobility due to their pK_a that can be predicted when using computer simulations.[22] However, isobaric ions that possess intermediate mobility differences may exhibit partial peak overlap, such as the final injected segment corresponding to a faster isobar (Val) with the first sample segment of a slower isobar (Bet). The introduction of fewer samples, use of variable spacer lengths, or an increase in capillary length can improve CE resolution; however, this reduces overall sample throughput or extends total analysis times. Apart from Val and Bet, this issue was not observed for other major isobaric ions detected in human plasma. Overall, MSI-CE-MS provides good linearity (R^2> 0.990) over a wide dynamic range (> 100-fold) with low micromolar detection limits (*LOD* ≈ 0.5-5.0 μM) for polar metabolites (**Table 3.1**) reflective on their intrinsic physicochemical properties that impact ionization efficiency.[19] However, due to the small sample volumes (≈10 nL) injected on-column along with post capillary dilution effects at the sheath liquid interface,[23] concentration sensitivity in CE-MS is about two orders of magnitude lower than LC-MS.[24] Recent developments in low-flow/sheathless interfaces[25,26] together with pre-column chemical derivatization strategies[27] offer ways to further boost sensitivity for detection of low abundance and/or weakly ionizable metabolites for expanded metabolome coverage.

MSI-CE-MS not only improves sample throughput but also offers a versatile experimental approach for conducting metabolomics with high data fidelity. **Figure 3.7(a)** depicts spike/recovery experiments for assessing accurate quantification of low micromolar levels (1.1 µM) of hypoxanthine (HyX) in pooled human plasma. In this case, a series of plasma samples are spiked with increasing concentrations of HyX (0-250 µM), which are analyzed to evaluate metabolite recovery by standard addition calibration. Due to random variations in sample volume when using hydrodynamic injections ($CV \approx 12\%$), reliable quantification in MSI-CE-MS requires normalization to its corresponding IS (25 µM 3-chloro-*L*-tyrosine, Cl-Tyr) originating from the same sample segment as a way to improve method precision in terms of relative response ratios (RRRs). Overall, good recoveries (≈95-115%) were measured for HyX and several other polar metabolites (**Table 3.2**) over a wide concentration range with excellent linearity ($R^2 > 0.990$) as highlighted in **Figure 3.7(a)**. Moreover, a correlation plot of the apparent sensitivity (μM^{-1}) derived from external (standards) and standard addition (spiked plasma) calibration curves for eight plasma metabolites confirms the lack of matrix-induced ion suppression with a slope ($m = 1.083$) and a coefficient of determination ($R^2 = 0.972$) close to unity when using a single nonisotope labeled IS (**Figure 3.8**).

MSI-CE-MS can also be configured to perform a dilution trend filter as an effective screening tool in metabolomics to reject spurious peaks and background ions that comprise the majority of molecular features detected in ESI-MS.[28] **Figure 3.7(b)** shows that replicate injection of a pooled plasma filtrate sample ($n = 3$; sample segments #1, #2, and #7) that is serially diluted (sample segments #3, #4, and #5) along with a blank as control (buffer filtrate; sample segment #6) offers an insightful way to interrogate the origin, linearity and reproducibility of signals detected in ESI-MS from a single experiment. In essence, "authentic" features for quantitation correspond to ions that satisfy three key criteria, namely, signals that are not detected in the blank yet can be measured with adequate precision ($CV < 35\%$) while exhibiting a linear response change upon dilution ($R^2 > 0.900$). In this case, low abundance metabolites present near the detection limit are excluded from subsequent data analyses due to their high variance, whereas background ions unrelated to the sample are also rejected since they contribute to data overfitting (type II error) when using multivariate statistical methods.[29]

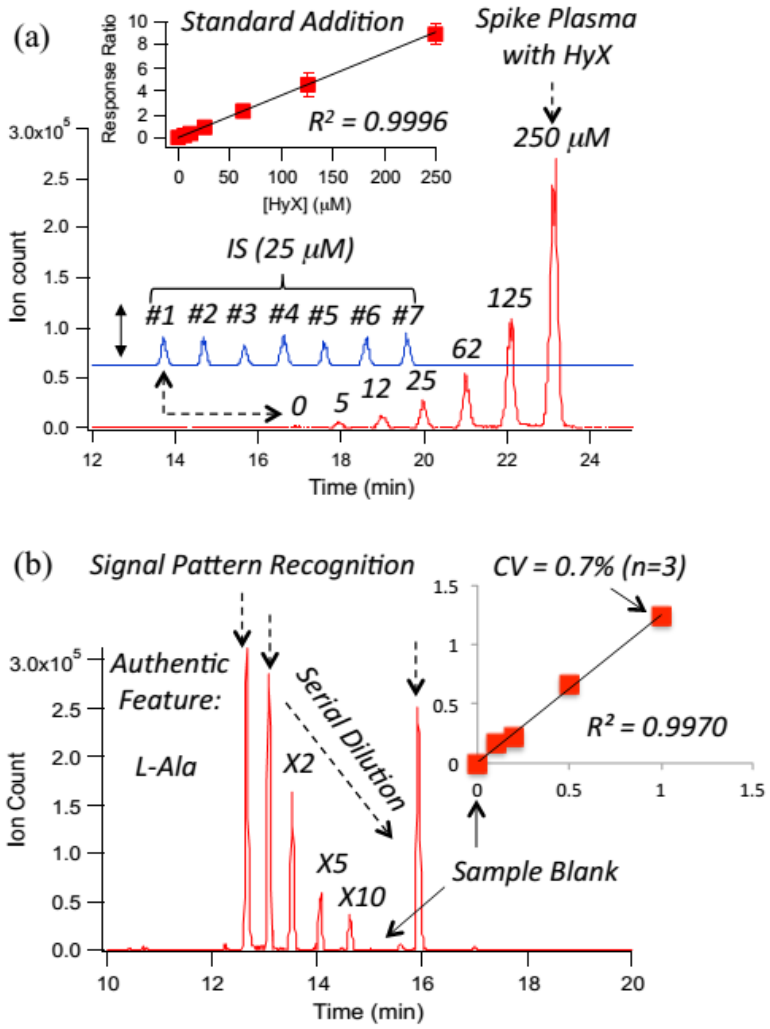


Figure 3.7. (a) Accurate quantification of plasma metabolites using spike/recovery experiments based on measured response ratio of solute (HyX) relative to internal standard (IS) from same sample segment. (b) A dilution trend filter of pooled human plasma for unambiguous identification of an authentic feature (*L*-Ala) that is reproducible and not present in the blank and exhibits a linear decrease in signal response. This process offers an effective screening tool to reject spurious signals and background ions that enhances data fidelity via signal pattern recognition without isotopic labelling.

Table 3.2. Standard addition data for quantification of polar metabolites in pooled human plasma. Method LOD was calculated based on the error of the calibration curve using 95% confidence intervals. Low abundance metabolites were spiked over a wide concentration range with an overall average recovery of 115%.

Metabolite	*m/z*	*Linear range of spike (μM)*	*Linearity (R^2)*	*LOD (μM)*	*Sensitivity ($μM^{-1}$)*	*Plasma concentration from x-intercept (μM, n=3)*	*Average % Recovery (n=6)*
Creatinine	114.0662	5.0-250	0.9996	2.6	0.0857	13 ± 4	111
Creatine	132.0768	5.0-250	0.9980	5.7	0.0456	9.4 ± 3.9	131
Orn	133.0972	5.0-250	0.9896	13	0.0446	11 ± 10	102
HyX	137.0458	5.0-250	0.9996	2.5	0.0361	1.1 ± 0.2	104
Deoxycarnitine	146.1176	5.0-250	0.9965	7.5	0.1279	9.1 ± 14	139
C0	162.1125	5.0-250	0.9997	2.3	0.1235	8.0 ± 3.6	111
Cit	176.1030	5.0-250	0.9960	7.9	0.0384	8.0 ± 3.9	105
C2	204.1230	5.0-250	0.9970	6.9	0.1265	9.2 ± 13	142
Cystine	241.0311	5.0-250	0.9991	3.7	0.0316	2.0 ± 0.4	98

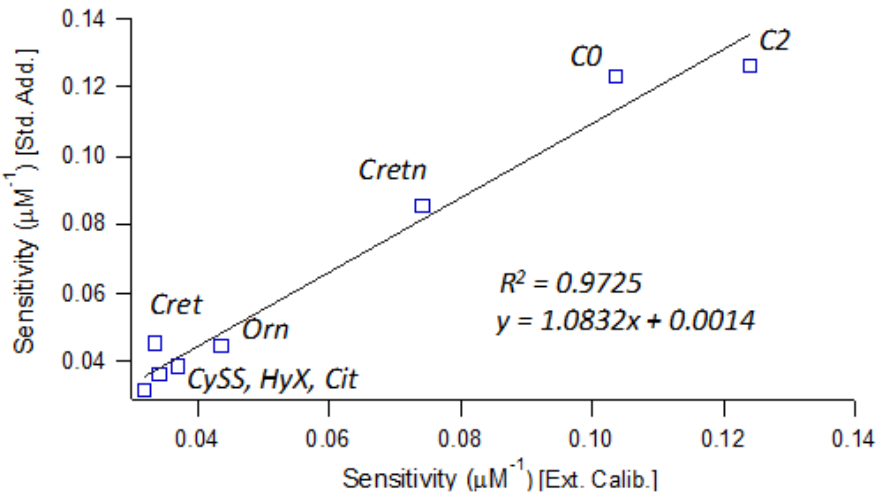


Figure 3.8. A correlation plot of the apparent sensitivity ($μM^{-1}$) derived from external (standards in ammonium acetate buffer) and standard addition (standards spiked into plasma) calibration curves for eight plasma metabolites which confirms the lack of matrix-induced signal suppression as reflected by a slope (m = 1.083) and a coefficient of determination (R^2 = 0.972) using a single non-isotope labeled IS (Cl-Tyr). High abundance polar metabolites in plasma (*e.g.*, Ala, Gln, Leu) were found to be present near the upper limit of linearity, thus spiking with standards did not generate good correspondence with external calibration curves. A wide disparity in sensitivity (response factor/ionization efficiency) was measured for polar metabolites over one order of magnitude due to differences in their physicochemical properties. Labels for metabolites correspond (from lowest to highest responder) to cystine [disulfide] (CySS), creatine (Cret), hypoxanthine (HyX), *L*-citrulline (Cit), *L*-ornithine (Orn), creatinine (Cretn), *L*-carnitine (C0) and *O*-acetyl-*L*-carnitine (C2).

Application of this rigorous dilution filter along with correction for redundant ions due to their signal multiplicity revealed 75 unique plasma metabolites as defined by their characteristic *m/z:*RMT (**Table A3.1**), where ions are excluded if they have one, two or seven signals detected. Out of 527 total molecular features, the majority (62%) correspond to spurious peaks (1 or 2 signals) or background ions (7 signals) during spray formation, such as salt adducts or solvent clusters. In contrast, authentic plasma metabolites represent only 37% of total nonredundant features with a signature pattern consisting of 3, 4, 5, 6 or more than 8 peaks depending on their ion abundance and number of detectable isobars (**Figure 3.9**). Implementation of this primary screen provides greater confidence in metabolite annotation while accelerating metabolomic workflows for biomarker discovery.[30] For instance, a peak list of authentic ions from a pooled sample can then be used to perform a "targeted" analysis for rapid screening without complicated data preprocessing typically performed on individual samples.[31] (**Table A3.2**). As a result, false discoveries are reduced,[32] which is important since major resources are devoted to qualitatively identify putative metabolites when chemical standards are unavailable. Method reproducibility was also examined for 40 plasma metabolites (**Figure 3.10** and **Table A3.3**) when analyzing ten replicate runs of pooled plasma filtrates by MSI-CE-MS using a 7 sample-segment format (n = 10 x 7 or 70 total runs within 8 h) with the same capillary. Overall, good precision was demonstrated for quantification as reflected by an average *CV* of 10% (n = 70) for RRR. Thus, MSI-CE-MS offers a novel separation format for metabolomics with a 3-fold greater sample throughput (210 samples/day) than recent ultrahigh pressure liquid chromatography (UHPLC)-MS methods.[10] Further studies are needed to evaluate the long-term performance of MSI-CE-MS when analyzing various biological samples with adequate robustness and quality assurance for large-scale metabolomic initiatives.[30]

3.5 Conclusions

In summary, this work strives to bridge the major gap that exists between sample throughput (*e.g.,* analysis speed, data preprocessing) and data fidelity (*e.g.,*selectivity, accuracy) in untargeted metabolome profiling of complex biological samples. For instance, ultrafast separations[33] and column switching methods[34] provide only modest gains in sample throughput, which limit metabolome coverage or require complicated hardware modifications unsuitable for routine analyses. In our case, signal pattern recognition based on customized

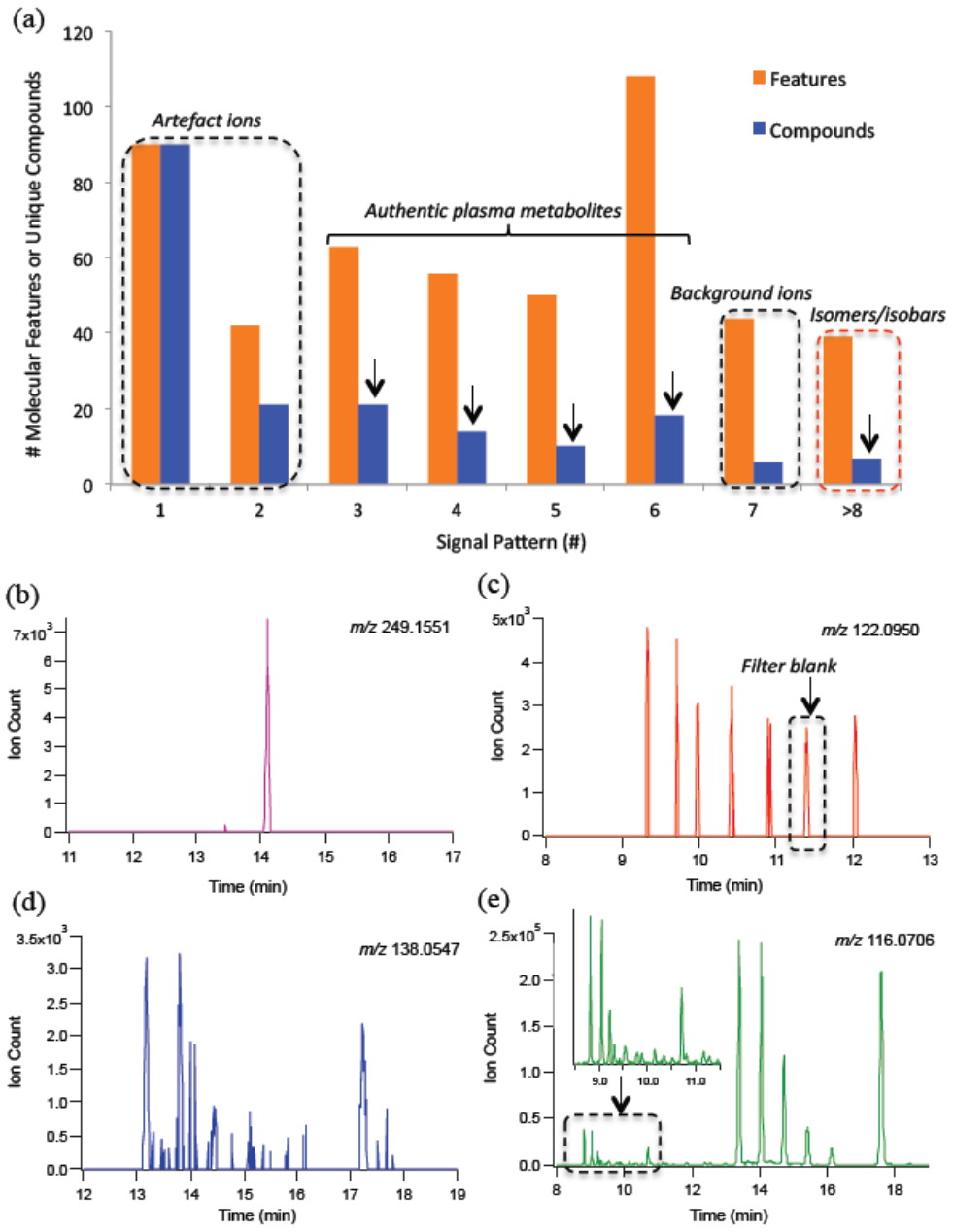


Figure 3.9. A dilution filter series experiment for assignment of "authentic" plasma metabolites in pooled human plasma by MSI-CE-MS while rejecting spurious signals, highlighting ions with (a) 1 signal (spurious peak), (b) 7 signals (background ion), (c) 3 signals (unknown metabolite near detection limit) and (d) 10 signals consisting of two isobars with 4 (unknown metabolite) and 6 features (*L*-Pro) detected. (e) Overall distribution of features detected using untargeted molecular feature extraction of pooled human plasma, where total molecular features (orange bars) detected were scaled to unique compounds (blue bars) depending on the multiplicity of signals detected for each ion based on its unique *m/z:RMT*. Overall, 75 polar metabolites (denoted by arrows) were assigned in pooled plasma based on their characteristic peak pattern (3, 4, 5, 6 or ≥7 signals).

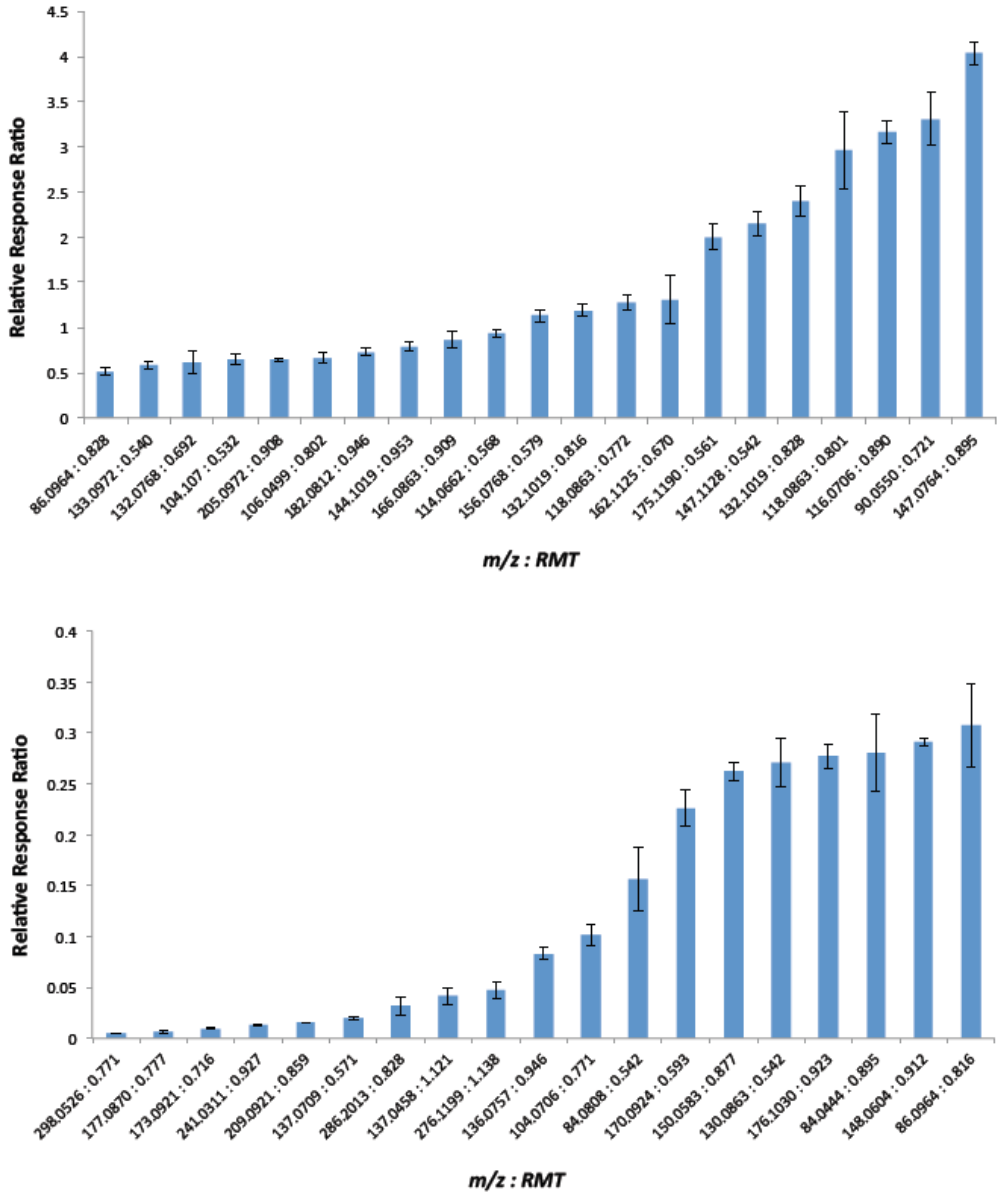


Figure 3.10. Reproducibility data based on 10 replicate analyses of pooled human plasma using a 7-sample segment injection (n = 10 x 7 = 70 runs) using the same capillary (< 8 h) highlighting 40 representative plasma metabolites annotated by their characteristic *m/z:RMT* that were selected after performing a dilution filter screen. The average coefficient of variance (*CV*) for quantification of plasma metabolites over a 900-dynamic range in relative signal response is 9.9% (median *CV* ≈ 8.5%) ranging from 3-20% (n=70) when using 3-chloro-*L*-tyrosine (Cl-Tyr, 25 µM) as a single non-isotope labelled internal standard.

serial injection sequences in MSI-CE-MS also enables annotation and identification of authentic analytical features from the background without isotopic chemical labeling[35] or tracer experiments.[36] This multiplexed separation platform offers a versatile approach for rapid screening (*e.g.,* dilution filter series), biomarker discovery (*e.g.,* healthy vs disease), and method validation (*e.g.,* recovery studies) based on relative or absolute quantification of metabolites and their isomers differentially expressed in seven or more samples. Current analytical constraints include poor concentration sensitivity, potential peak overlap due to loss in isobar/isomer resolution and its restriction to ionic metabolites. Future work will develop new strategies to improve sensitivity and metabolome coverage in MSI-CE-MS while introducing an accelerated metabolomics workflow for biomarker discovery.

3.6 Acknowledgement

This work was financially supported by the National Science and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation (CFI) and the Ontario Genomics Institute (OGI).

3.7 References

(1) Kuehnbaum, N. L.; Britz-McKibbin, P. *Chem. Rev.* **2013**, *113*, 2437–2468.

(2) Kumari, S.; Stevens, D.; Kind, T.; Denkert, C.; Fiehn, O. *Anal. Chem.* **2011**, *83*, 5895–5902.

(3) Boswell, P. G.; Schellenberg, J. R.; Carr, P. W.; Cohen, J. D.; Hegeman, A. D. *J. Chromatogr. A* **2011**, *1218*, 6732–6741.

(4) Sugimoto, M.; Hirayama, A.; Robert, M.; Abe, S.; Soga, T.; Tomita, M. *Electrophoresis* **2010**, *31*, 2311–2318.

(5) Han, J.; Danell, R. M.; Patel, J. R.; Gumerov, D. R.; Scarlett, C. O.; Speir, J. P.; Parker, C. E.; Rusyn, I.; Zeisel, S.; Borchers, C. H. *Metabolomics* **2008**, *4*, 128–140.

(6) Kirwan, J. A.; Broadhurst, D. I.; Davidson, R. L.; Viant, M. R. *Anal. Bioanal. Chem.* **2013**, *405*, 5147–5157.

(7) Rath, C. M.; Yang, J. Y.; Alexandrov, T.; Dorrestein, P. C. *J. Am. Soc. Mass. Spectrom.* **2013**, *24*, 1167–1176.

(8) Hsu, C.-C.; Elnaggar, M. S.; Peng, Y.; Fang, J.; Sanchez, L. M.; Mascuch, S. J.; Møller, K. a; Alazzeh, E. K.; Pikula, J.; Quinn, R. a; Zeng, Y.; Wolfe, B. E.; Dutton, R. J.; Gerwick, L.; Zhang, L.; Liu, X.; Månsson, M.; Dorrestein, P. C. *Anal. Chem.* **2013**, *85*, 7014–7018.

(9) Fuhrer, T.; Heer, D.; Begemann, B.; Zamboni, N. *Anal. Chem.* **2011**, *83*, 7074–7080.

(10) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2013**, *8*, 17–32.

(11) Saric, J.; Want, E.; Duthaler, U.; Lewis, M.; Keiser, J.; Shockcor, J.; Ross, G.; Nicholson, J.; Holmes, E.; Tavares, M. *Anal. Chem.* **2012**, *84*, 6963–6972.

(12) Dovichi, N. J.; Zhang, J. *Angew. Chem. Int. Ed.* **2000**, *39*, 4463–4468.

(13) Hai, X.; Yang, B.; Schepdael, A. Van. *Electrophoresis* **2012**, *33*, 211–227.

(14) Zavaleta, J.; Chinchilla, D. B.; Kaddis, C. F.; Martinez, K.; Brown, A.; Gomez, A.; Pao, A.; Ramirez, A.; Nilapwar, S.; Ladbury, J. E.; Gomez, F. A. *J. Cap. Elec Microchip Tech.* **2006**, *9*, 101–117.

(15) Ptolemy, A. S.; Britz-McKibbin, P. *J. Chromatogr. A* **2006**, *1106*, 7–18.

(16) Chambers, A.; Mellors, J.; Henley, W.; Rasmsey, J. *Anal. Chem.* **2011**, *83*, 842–849.

(17) Lodén, H.; Pettersson, C.; Arvidsson, T.; Amini, A. *J. Chromatogr. A* **2008**, *1207*, 181–185.

(18) Staub, A.; Rudaz, S.; Veuthey, J.-L.; Schappler, J. *J. Chromatogr. A* **2010**, *1217*, 8041–8047.

(19) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 2506–2515.

(20) Hutterer, K. M.; Jorgenson, J. W. *Anal. Chem.* **1999**, *71*, 1293–1297.

(21) Chalcraft, K. R.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 307–314.

(22) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403–415.

(23) Mokaddem, M.; Gareil, P.; Belgaied, J.-E.; Varenne, A. *Electrophoresis* **2008**, *29*, 1957–1964.

(24) Nevedomskaya, E.; Ramautar, R.; Derks, R.; Westbroek, I.; Zondag, G.; Pluijm, I.; Deelder, A. M.; Mayboroda, O. A.; van der Pluijm, I. *J. Proteome Res.* **2010**, *9*, 4869–4874.

(25) Ramautar, R.; Busnel, J.-M.; Deelder, A. M.; Mayboroda, O. A. *Anal. Chem.* **2012**, *84*, 885–892.

(26) Maxwell, E. J.; Zhong, X.; Chen, D. D. *Anal. Chem.* **2010**, *82*, 8377–8381.

(27) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(28) Jankevics, A.; Merlo, M.; de Vries, M.; Vonk, R.; Takano, E.; Breitling, R. *Metabolomics* **2012**, *8*, 29–36.

(29) Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S. *Nucleic Acids Res.* **2012**, *40*, W127–W133.

(30) Dunn, W.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J.; Halsall, A.; Haselden, J.; Nicholls, A.; Wilson, I.; Kell, D.; Goodacre, R. *Nat. Protoc.* **2011**, *30*, 1060–1083.

(31) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84*, 5035–5039.

(32) Broadhurst, D. I.; Kell, D. B. *Metabolomics* **2007**, *2*, 171–196.

(33) Gika, H. G.; Theodoridis, G.; Extance, J.; Edge, A. M.; Wilson, I. D. *J. Chromatogr. B* **2008**, *871*, 279–287.

(34) Livesay, E. A.; Tang, K.; Taylor, B. K.; Buschbach, M. A.; Hopkins, D. F.; LaMarche, B. L.; Zhao, R.; Shen, Y.; Orton, D. J.; Moore, R. J.; Kelly, R. T.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2008**, *80*, 294–302.

(35) Zheng, J.; Dixon, R. A.; Li, L. *Anal. Chem.* **2012**, *84*, 10802–10811.

(36) De Jong, F.; Beecher, C. *Bioanalysis* **2012**, *4*, 2303–2314.

3.8 Appendix A

Table A3.1. A dilution filter series by MSI-CE-MS is used to identify "authentic" metabolites from pooled human plasma and exclude artifacts/background ions generated during spray formation. This list of 75 compounds was filtered and compounds were removed if they had [a]7 peaks (a signal in the blank), [b]CV% in RRR> 30%, and/or [c]R^2 of the resulting slope of RRR vs. dilution factor < 0.9. With some overlap of exclusion criteria, a total of 21 features were later rejected as noted.

m/z : RMT	*Signal in Blank (Y/N)*	*Signal Pattern: Detectable Peaks (#)*	*Reproducibility CV% (n=3)*	*Dilution Linearity R^2*
70.0663 : 0.898	N	5	5.3	0.9891
76.0404 : 0.688	N	6	5.6	0.9944
76.0766 : 0.534	N	6	7.6	0.9898
84.0452 : 0.903[c]	N	6	4.9	0.0338
84.0816 : 0.565	N	6	18.3	0.9957
86.0972 : 0.827	N	6	9.6	0.9994
86.0972 : 0.839	N	6	1.2	0.9980
90.0556 : 0.739	N	6	4.5	0.9963
104.0710 : 0.785	N	6	11.6	0.9692
104.1073 : 0.569	N	5	28.3	0.9893
106.0499 : 0.834	N	6	9.7	0.9976
110.0714 : 0.611[c]	N	4	10.8	0.4118
114.0663 : 0.590	N	6	14.1	0.9979
116.0707 : 0.898	N	6	5.4	0.9988
118.0863 : 0.814	N	6	8.1	0.9972
118.0863 : 0.951	N	6	11.2	0.9714
119.0160 : 0.966[b]	N	5	49.3	0.9681
120.0653 : 0.876	N	6	7.3	0.9855
122.0950 : 0.613[a]	Y	7	24.5	0.9246
132.0656 : 0.880	N	6	15.2	0.9829
132.0766 : 0.719	N	6	7.3	0.9888
132.1018 : 0.827	N	6	9.2	0.9995
132.1018 : 0.839	N	6	2.7	0.9991
133.0604 : 0.880[c]	N	6	2.1	0.3220
133.0969 : 0.564	N	6	17.2	0.9957
136.0754 : 0.949	N	6	8.4	0.9959
137.0457 : 1.133	N	4	9.9	0.9885
137.0708 : 0.593	N	5	18.3	0.9951
138.0547 : 0.880	N	3	12.6	0.9656

m/z : RMT	*Signal in Blank (Y/N)*	*Signal Pattern: Detectable Peaks (#)*	*Reproducibility CV% (n=3)*	*Dilution Linearity R^2*
144.1018 : 0.956	N	6	15.1	0.9784
146.1193 : 0.653[c]	N	6	29.4	0.7853
147.0764 : 0.902	N	6	4.4	0.9985
147.1128 : 0.565	N	6	15.4	0.9994
148.0603 : 0.919	N	6	3.9	1.0000
150.0581 : 0.886	N	6	3.3	0.9990
156.0418 : 0.987[c]	N	5	13.1	0.7164
156.0766 : 0.602	N	6	14.0	0.9994
161.1283 : 0.644	N	6	12.0	0.9970
162.1121 : 0.688	N	6	19.9	0.9988
163.1437 : 0.813	N	4	12.0	0.9946
166.0862 : 0.915	N	6	7.2	1.0000
170.0366 : 1.000[a,c]	Y	7	5.6	0.1815
170.0563 : 0.837[c]	N	6	10.1	0.2433
170.0924 : 0.615	N	6	13.9	0.9991
173.0920 : 0.734	N	5	8.1	0.9367
175.1191 : 0.584	N	6	17.3	0.9995
176.1028 : 0.929	N	6	6.5	0.9991
177.0867 : 0.877	N	6	16.5	0.9936
177.1595 : 0.827	N	4	11.7	0.9346
177.1595 : 0.839	N	5	10.0	0.9879
182.0809 : 0.949	N	6	4.8	0.9973
189.1595 : 0.586	N	6	18.1	0.9718
192.1592 : 0.762	N	6	8.6	0.9858
198.1275 : 0.739[a,c]	Y	7	12.2	0.1151
204.1228 : 0.732[b]	N	6	30.6	0.9955
205.0970 : 0.914	N	6	6.1	0.9999
205.1180 : 1.137[b,c]	N	5	37.8	0.6081
209.0918 : 0.872	N	5	13.3	0.9920
214.0582 : 0.662	N	5	29.5	0.9220
218.1389 : 0.753[b]	N	6	40.7	0.9204
229.1542 : 0.796[c]	N	3	27.8	0.7853
235.0923 : 1.196[b]	N	6	36.2	0.9469
235.0923 : 1.222	N	6	30.0	0.9832
241.0308 : 0.934	N	6	8.0	0.9986
243.0525 : 3914[b]	N	5	38.5	0.9781
249.1078 : 1.407[b]	N	5	34.4	0.9862

m/z : RMT	*Signal in Blank (Y/N)*	*Signal Pattern: Detectable Peaks (#)*	*Reproducibility CV% (n=3)*	*Dilution Linearity R^2*
253.9975 : 0.999[a,b,c]	Y	7	38.6	0.8160
276.1185 : 1.124	N	4	24.1	0.9977
283.1626 : 0.953[b]	N	5	63.4	0.9883
286.2009 : 0.838	N	6	7.1	0.9544
299.1370 : 0.951[b]	N	4	70.9	0.9123
325.1519 : 0.954	N	6	14.5	0.9682
337.1084 : 0.903[b]	N	6	30.4	0.9926
477.1602 : 0.901	N	4	12.7	0.9475
733.3294 : 1.304	N	4	16.9	0.9643

Table A3.2. An authentic peak list of 51 plasma metabolites selected after rigorous dilution filter series by MSI-CE-MS (*CV*< 30%; linear dilution, R^2> 0.90; no blank signal) together with empirical formula matches with an average absolute mass error of ±2.6 ppm. This approach is used as a high quality filter for subsequent targeted analysis of individual samples in large-scale metabolomics initiatives. Both RRR and RMT for ions were normalized to the internal standard, 25 µM 3-chloro-L-tyrosine. *High mass error (> 5 ppm) is present for analytes below lowest lock mass (*m/z* 121) used for real-time internal mass calibration.

m/z : RMT	*Molecular Formula*	*Compound Name*	*RRR, CV% (n = 3)*	*Dilution Linearity R^2*	*Mass error (ppm)*
70.0663 : 0.898	C4 H7 N		5.3	0.9891	17.1*
76.0404 : 0.688	C2 H5 N O2	Glycine	5.6	0.9944	14.5*
76.0766 : 0.534	C3 H9 N O		7.6	0.9898	11.8*
84.0816 : 0.565	C5 H9 N		18.3	0.9957	9.5*
86.0972 : 0.827	C5 H11 N		9.6	0.9994	9.3*
86.0972 : 0.839	C5 H11 N		1.2	0.9980	9.3*
90.0556 : 0.739	C3 H7 N O2	Alanine	4.5	0.9963	6.7*
104.071 : 0.785	C4 H9 N O2		11.6	0.9692	3.8
104.1073 : 0.569	C5 H14 N O	Choline	28.3	0.9893	-1.9
106.0499 : 0.834	C3 H7 N O3	Serine	9.7	0.9976	0.0
114.0663 : 0.59	C4 H7 N3 O	Creatinine	14.1	0.9979	0.9
116.0707 : 0.898	C5 H9 N O2	Proline	5.4	0.9988	0.9
118.0863 : 0.814	C5 H11 N O2	Valine	8.1	0.9972	0.0
118.0863 : 0.951	C5 H11 N O2	Betaine	11.2	0.9714	0.0
120.0653 : 0.876	C4 H9 N O3	Threonine	7.3	0.9855	-1.7
132.0656 : 0.88	C5 H9 N O3	Hydroxyproline	15.2	0.9829	0.8
132.0766 : 0.719	C4 H9 N3 O2	Creatine	7.3	0.9888	-1.5
132.1018 : 0.827	C6 H13 N O2	Isoleucine	9.2	0.9995	-0.8
132.1018 : 0.839	C6 H13 N O2	Leucine	2.7	0.9991	-0.8
133.0969 : 0.564	C5 H12 N2 O2	Ornithine	17.2	0.9957	-2.3
136.0754 : 0.949	C8 H9 N O		8.4	0.9959	-2.2
137.0457 : 1.133	C5 H4 N4 O	Hypoxanthine	9.9	0.9885	-0.7
137.0708 : 0.593	C7 H8 N2 O		18.3	0.9951	-0.7
138.0547 : 0.88	C7 H7 N O2		12.6	0.9656	-2.2
144.1018 : 0.956	C7 H13 N O2		15.1	0.9784	-0.7
147.0764 : 0.902	C5 H10 N2 O3	Glutamine	4.4	0.9985	0.0
147.1128 : 0.565	C6 H14 N2 O2	Lysine	15.4	0.9994	0.0
148.0603 : 0.919	C5 H9 N O4	Glutamic acid	3.9	1.0000	-0.7
150.0581 : 0.886	C5 H11 N O2 S	Methionine	3.3	0.9990	-1.3
156.0766 : 0.602	C6 H9 N3 O2	Histidine	14.0	0.9994	-1.3
161.1283 : 0.644	C7 H16 N2 O2		12.0	0.9970	-1.2
162.1121 : 0.688	C7 H15 N O3	Carnitine	19.9	0.9988	-2.5
163.1437 : 0.813	C7 H18 N2 O2		12.0	0.9946	-2.5
166.0862 : 0.915	C9 H11 N O2	Phenylalanine	7.2	1.0000	-0.6

m/z : RMT	*Molecular Formula*	*Compound Name*	*RRR, CV% (n = 3)*	*Dilution Linearity R^2*	*Mass error (ppm)*
170.0924 : 0.615	C7 H11 N3 O2	3-Methyl histidine	13.9	0.9991	0.0
173.092 : 0.734	C7 H12 N2 O3		8.1	0.9367	-0.6
175.1191 : 0.584	C6 H14 N4 O2	Arginine	17.3	0.9995	0.6
176.1028 : 0.929	C6 H13 N3 O3	Citrulline	6.5	0.9991	-1.1
177.0867 : 0.877	C6 H12 N2 O4		16.5	0.9936	-1.7
177.1595 : 0.827	C8 H20 N2 O2		11.7	0.9346	-1.7
177.1595 : 0.839	C8 H20 N2 O2		10.0	0.9879	-1.7
182.0809 : 0.949	C9 H11 N O3	Tyrosine	4.8	0.9973	-1.6
189.1595 : 0.586	C9 H20 N2 O2		18.1	0.9718	-1.6
192.1592 : 0.762	C9 H21 N O3		8.6	0.9858	-1.0
205.097 : 0.914	C11 H12 N2 O2	Tryptophan	6.1	0.9999	-1.0
209.0918 : 0.872	C10 H12 N2 O3	Kynurenine	13.3	0.9920	-1.4
214.0582 : 0.662	C6 H7 N5 O4		29.5	0.9220	5.1
241.0308 : 0.934	C6 H12 N2 O4 S2	Cystine	8.0	0.9986	-1.2
276.1185 : 1.124	C10 H17 N3 O6		24.1	0.9977	-1.8
286.2009 : 0.838	C15 H27 N O4		7.1	0.9544	-1.4
477.1602 : 0.901	C18 H16 N14 O3		12.7	0.9475	-0.2

Table A3.3. Method reproducibility using ten replicate runs of pooled plasma filtrates by MSI-CE-MS using a 7 sample-segment format with the same capillary (n = 70 total run equivalent). Overall, the average CV% for 40 plasma metabolites in terms of relative response ratio (RRR) is 10% and for relative migration time (RMT) is 3% (n=70).

m/z: RMT	*Metabolite Name*	*RMT*	*RRR*	*CV% (RRR)*
298.0526 : 0.771	GlyCy-SS	0.771 ± 0.014	0.005 ± 0.001	13.9
177.0870 : 0.777		0.777 ± 0.214	0.006 ± 0.001	15.9
173.0921 : 0.716		0.716 ± 0.016	0.009 ± 0.001	9.1
241.0311 : 0.927	CySS	0.927 ± 0.005	0.013 ± 0.001	9.0
209.0921 : 0.859		0.859 ± 0.009	0.015 ± 0.001	8.0
137.0709 : 0.571		0.571 ± 0.020	0.019 ± 0.002	10.3
286.2013 : 0.828		0.828 ± 0.011	0.032 ± 0.009	27.6
137.0458 : 1.121		1.121 ± 0.065	0.042 ± 0.008	20.1
276.1199 : 1.138		1.138 ± 0.009	0.047 ± 0.009	18.7
136.0757 : 0.946		0.946 ± 0.004	0.083 ± 0.006	7.7
104.0706 : 0.771	GABA	0.771 ± 0.015	0.101 ± 0.010	10.3
84.0808 : 0.542		0.542 ± 0.019	0.156 ± 0.031	20.1
170.0924 : 0.593	Me-His	0.593 ± 0.018	0.226 ± 0.019	8.2
150.0583 : 0.877		0.877 ± 0.008	0.262 ± 0.009	3.4
130.0863 : 0.542	Pipecolic acid	0.542 ± 0.019	0.271 ± 0.024	8.9
176.1030 : 0.923		0.923 ± 0.005	0.277 ± 0.011	4.1
84.0444 : 0.895		0.895 ± 0.007	0.280 ± 0.038	13.7
148.0604 : 0.912	Glycine	0.912 ± 0.006	0.291 ± 0.004	1.4
86.0964 : 0.816		0.816 ± 0.012	0.307 ± 0.041	13.2
86.0964 : 0.828		0.828 ± 0.012	0.515 ± 0.034	6.7
133.0972 : 0.540	Ornithine	0.540 ± 0.019	0.581 ± 0.047	8.1
132.0768 : 0.692		0.692 ± 0.046	0.615 ± 0.119	19.3
104.1070 : 0.532	Choline	0.532 ± 0.020	0.647 ± 0.056	8.7
205.0972 : 0.908		0.908 ± 0.006	0.648 ± 0.020	3.0
106.0499 : 0.802		0.820 ± 0.012	0.669 ± 0.061	9.2
182.0812 : 0.946		0.946 ± 0.004	0.728 ± 0.035	4.8
144.1019 : 0.953		0.953 ± 0.004	0.789 ± 0.054	6.9
166.0863 : 0.909		0.673 ± 0.017	0.858 ± 0.092	10.7
114.0662 : 0.568	Creatine	0.909 ± 0.006	0.932 ± 0.035	3.7
156.0768 : 0.579	Histidine	0.568 ± 0.020	1.133 ± 0.070	6.2
132.1019 : 0.816	Isoleucine	0.579 ± 0.019	1.190 ± 0.072	6.1
118.0863 : 0.772	Betaine	0.816 ± 0.012	1.283 ± 0.081	6.3
162.1125 : 0.670	Carnitine	0.772 ± 0.204	1.313 ± 0.265	20.2
175.1190 : 0.561		0.561 ± 0.019	2.005 ± 0.147	7.4
147.1128 : 0.542	Lysine	0.542 ± 0.019	2.154 ± 0.135	6.3
132.1019 : 0.828	Leucine	0.828 ± 0.012	2.401 ± 0.165	6.9
118.0863 : 0.801	Valine	0.801 ± 0.013	2.966 ± 0.425	14.3
116.0706 : 0.890		0.890 ± 0.007	3.164 ± 0.124	3.9
90.0550 : 0.721	Alanine	0.721 ± 0.016	3.310 ± 0.294	8.9
147.0764 : 0.895		0.895 ± 0.007	4.032 ± 0.128	3.2

Chapter IV

An Accelerated Workflow in Metabolomics for Biomarker Discovery: Elucidating Adaptive Responses to High Intensity Interval Training

IV. An Accelerated Workflow in Metabolomics for Biomarker Discovery: Elucidating Adaptive Responses to High Intensity Interval Training

4.1 Abstract

Practical lifestyle modifications based on physical activity are needed to enhance public health while reducing chronic disease risk. We introduce an accelerated metabolomics workflow for biomarker discovery using multi-segment injection-capillary electrophoresis-mass spectrometry as a multiplexed separation platform to evaluate adaptive responses to exercise training on an individual level. Overweight/obese women performed a 6-week high intensity interval training (HIT) intervention using a cross-over study design. Venous blood samples were collected from each subject in their naïve and trained state when completing standardized cycling trials at the same absolute workload. Complementary statistical methods were used to classify changes in plasma metabolism associated with training status, as well as time-dependent metabolomic responses to strenuous exercise. Positive responses to HIT were associated with training-induced upregulation in plasma *L*-carnitine at rest reflecting improved muscle oxidative capacity. Attenuation in plasma hypoxanthine and higher *O*-acetyl-*L*-carnitine post-exercise also indicated lower energetic stress for trained subjects. HIT promotes fatty acid metabolism, glutathione homeostasis and aerobic ATP recycling in skeletal muscle that corresponds to gains in cardiorespiratory fitness for sedentary women despite between-subject variances.

4.2 Introduction

Human health is determined by a complex interplay of genetic and environmental factors, including diet and habitual physical activity. A global epidemic of obesity threatens to reverse overall life expectancy[1,2] while overwhelming healthcare resources since it is an independent graded risk factor of cardiometabolic disorders, including type 2 diabetes and coronary heart disease.[3,4] Systemic low-grade inflammation associated with physical inactivity and lack of adequate muscle contractions contributes to a pro-inflammatory phenotype independent from obesity.[5] In this context, physical exercise plays a key role in human health, which can be used as an integrative strategy for chronic disease prevention even without substantial weight loss.[6] High intensity interval training (HIT) is characterized by short, intense bursts of activity interspersed with periods of rest or low intensity exercise.[7] Since lack of time remains a major barrier to regular

exercise, HIT offers similar health benefits than endurance training while requiring reduced time commitments and lower training volume.[8] However, prognostic indicators are needed to predict exercise training responsiveness[9] since treatment effects are variable between-subjects with weaker responses reported for women.[10]

Metabolomics[11] is an expanding area of functional genomics research since metabolites represent real-world end-products of gene expression closely associated with clinical outcomes.[12] Untargeted metabolite profiling can decipher genes/protein of unknown function[13,14] while revealing underlying molecular mechanisms of action involving drugs,[15] environmental pollutants[16] and food products.[17] Dynamic metabolomic studies in conjunction with cross-over human trials offers a powerful way to elucidate subtle treatment effects on an individual level that are obscured by biological variance.[18] Several studies have applied metabolomics for characterizing metabolic signatures of exercise performance in human trials,[19] including the impact of dietary interventions in strenuous exercise.[20–22] Metabolomics has also been used to evaluate the efficacy of exercise training for treating patients with type 1 diabetes[23] and chronic obstructive pulmonary disease.[24] However, sample throughput and data pre-processing remain major constraints in large-scale mass spectrometry (MS)-based metabolomics that rely on chromatographic[25] or electrophoretic[26] separations. Although high throughput metabolome screening is achieved by direct infusion[27–29] or ambient ionization[30] MS, high efficiency separations are critical to enhance selectivity and improve quantitative performance while supporting metabolite identification in complex biological samples.[31]

The main objective for this study is to introduce an accelerated metabolomics workflow for the discovery of plasma markers of exercise responsiveness based on multi-segment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS).[32] Increased sample throughput and higher data quality are achieved via serial injection of seven sample segments within a fused-silica capillary that encodes information temporally prior to MS detection via signal pattern recognition. We first implemented a dilution trend filter when analyzing a pooled plasma by MSI-CE-MS as an untargeted primary screen to characterize reproducible yet sample-derived metabolites while rejecting spurious signals generated during spray formation.[33] This process enables targeted profiling of authentic plasma metabolites when analyzing individual samples without complicated peak picking and time alignment procedures that are subject to bias.[34] We then evaluated dynamic metabolomic changes to strenuous exercise

and adaptive training responses of a cohort of overweight/obese women who participated in a supervised 6-week HIT intervention.[35] MSI-CE-MS offers a multiplexed separation platform to unambiguously identify metabolic signatures of exercise training for each subject both prior to (naïve) and after (trained) HIT with quality assurance to reduce false discoveries. Positive adaptations to exercise training included improvements to fatty acid metabolism and glutathione homeostasis at rest, as well as greater aerobic ATP recycling capacity following strenuous exercise. Gains in cardiorespiratory fitness and skeletal muscle mitochondrial function derived from HIT offer a practical solution to improve public health for chronic disease prevention in high-risk populations.

4.3 Experimental

4.3.1 Cohort Selection

Nine overweight/obese but otherwise healthy women (age: 21-45 years, *BMI*: 25-36 kg/m^2) were recruited to participate in a 6-week supervised exercise training intervention as we have previously described.[35] The subjects represent a subset of a group of 16 individuals who took part in a study that evaluated the impact of pre-exercise nutritional state on adaptations to HIT.[35] In this work, blood samples at all time intervals during two HIT sessions were acquired from only 9 of the 16 subjects from the original trial. The focus of the present study was to identify plasma markers of exercise responsiveness using metabolomics when performing strenuous exercise at the same absolute workload for subjects in their naïve and trained state after 6 weeks of HIT. No effect of the nutritional intervention was observed, and all subjects in the current study were compared against themselves in a repeated measures design in which pre-exercise nutrition was standardized for a given subject. The participants were recruited through poster advertisements and the study protocol was approved by the Hamilton Integrated Research Ethics Board. Subjects were classified as "sedentary" based on self-reporting of physical activity $\leq$ 2 sessions/week of exercise with a duration $\leq$ 30 min. Participants completed a general health questionnaire to assess overall health and verify that they did not meet any of the exclusion criteria, including pre-diabetes, diabetes, hypertension or cardiovascular diseases.

4.3.2 Body Composition and Fitness Testing

Body composition analysis to determine percentages of bone, fat, and lean muscle tissue were measured by dual x-ray absorptiometry (DXA) scanning (Lunar Prodigy Advance, Madison, WI). Maximal oxygen uptake (VO_{2max}), peak power output (W_{max}), and average heart rate (HR) as indicators of cardiorespiratory fitness were determined for each participant using standardized protocols.[35] Changes in the aerobic endurance capacity for each participant from baseline was evaluated by measuring VO_{2max} and W_{max} after 6-week HIT intervention, whereas changes in average HR were assessed for individual subjects during standardized HIT trials performed in their naïve and trained state.

4.3.3 Exercise Intervention

Exercise was performed on a stationary cycle ergometer (LifeCycle C1, Life Fitness, Schiller Park, IL) while maintaining a pedal cadence of 80-100 rpm. The exercise intervention consisted of 18 sessions of HIT, performed three times a week (with a 1-2 day recovery in between) over a total of 6 weeks.[35] The HIT protocol involved intermittent cycling and consisted of 10 cycles of 60 s duration at an individualized workload designed to elicit 90% HR_{max}. Recovery intervals that involved light cycling at 50 W were performed for 1 min between high intensity bursts for a total exercise time of 25 min for each HIT trial, which included a 3 min warm-up and a 2 min cool down period. Absolute workload was increased for each subject over the course of the HIT intervention in order to maintain the training stimulus of 90% HR_{max}. At the end of each HIT session, average heart rate and rating of perceived exertion (RPE, 0-10 scale) were also recorded. All training sessions were performed in the morning between 0700–1000 h.

4.3.4 Plasma Collection and Sample Workup

During the second last HIT session (trial 17), the training workload was adjusted in order to match the initial session (trial 1). This was done in order to facilitate blood collection under standardized conditions, and permit a comparison at the same absolute workload. In this way, exercise responsiveness to HIT was evaluated for each participant in their trained relative to naïve state when using a cross-over study. Venous blood samples were collected via an indwelling catheter inserted in a forearm vein. Vacutainers (6 mL) with EDTA (10.8 mg) as

anticoagulant were used for collection (Beckton Dickson, Franklin Lakes, NJ, USA). Samples were collected prior to onset of exercise (*i.e.*, baseline, 0 min) immediately after HIT (*i.e.*, post-exercise, 20 min), and 20 min after HIT while at rest (*i.e.*, recovery, 40 min). Blood samples were placed on ice and subsequently centrifuged at 2500 *g* at 4°C for 5 min to fractionate plasma from erythrocytes. Fractionated plasma was aliquoted and frozen at -80°C until later processing and analysis. Immediately after collection, blood samples were placed on ice and subsequently centrifuged at 2500 *g* at 4°C for 5 min to fractionate plasma from erythrocytes. Fractionated plasma was aliquoted and frozen at -80°C until analysis. Frozen plasma aliquots (-80°C) were thawed on ice, then vortexed for 30 s to mix. Plasma was diluted 4-fold to a final concentration of 200 mM NH4Ac (adjusted to pH 5 with HAc) with 25 µM 3-chloro-*L*-tyrosine (Cl-Tyr) as internal standard and vortexed for 30 s. Plasma proteins were removed using a 3 kDa MWCO Nanosep centrifugal device (Pall Life Sciences, Washington, NY, USA) at 13,000 *g* for 15 min. A 20 µL aliquot of the plasma filtrate sample was subsequently used for analysis.

4.3.5 Chemicals and Reagents

Ammonium acetate, acetic acid, formic acid, and Cl-Tyr were purchased from Sigma-Aldrich (St. Louis, MO, USA). A 10 mM stock solutions of Cl-Tyr was prepared in water and stored at 4°C. HPLC-grade acetonitrile (Honeywell, Muskegon, MI, USA) and methanol (Caledon, Georgetown, ON, Canada) were used for preparation of background electrolyte (BGE) and sheath liquid respectively. Ammonium acetate was prepared as a 400 mM stock in water and the pH was adjusted to 5.0 using acetic acid. All aqueous buffers and stock solutions were prepared with deionized water purified using a Thermo Scientific Barnstead EasyPure II LF ultrapure water system (Cole Parmer, Vernon Hills, IL, USA).

4.3.6 Instrumentation and MSI-CE Configuration

All CE-TOF-MS experiments were performed using an Agilent G7100A CE system (Mississauga, ON, Canada) interfaced with a coaxial sheath liquid Jet Stream electrospray ion source with heated gas to an Agilent 6230 TOF LC/MS orthogonal axis time-of-flight mass spectrometer (TOF-MS). The TOF mass analyzer serves as an ideal multi-channel detector for MSI-CE-MS with sufficiently high mass resolution and fast data acquisition. Nitrogen gas was used

as the nebulizer gas in the ESI source and as the drying gas in the MS. An uncoated fused silica capillary (Polymicro Technologies, AZ, USA) with 50 μm ID and 110 cm length maintained at 25 °C was used for all experiments. The background electrolyte (BGE) was 1 M formic acid containing 15% v acetonitrile (pH 1.8). The applied voltage was 30 kV with a total analysis time of 35 min. In MSI-CE-MS, samples and spacers containing BGE were alternately injected in series at 100 mbar as described previously,[32] namely (1) 5 s sample, 40 s spacer; (2) 5 s sample, 40 s spacer; (3) 5 s sample, 40 s spacer; (4) 5 s sample, 40 s spacer; (5) 5 s sample, 40 s spacer; (6) 5 s sample, 40 s spacer; (7) 5 s sample, 5 s spacer. The total injection time for sample and BGE spacer segments was 280 s that is equivalent to about 30% of the total capillary length. Plasma filtrate samples were diluted 4-fold in 200 mM NH_4Ac (pH 5) with 25 μM Cl-Tyr as internal standard, which was used for determination of relative migration time (RMT) and relative peak response of metabolites. Between runs the capillary was flushed for 10 min with BGE. The sheath liquid was 60:40 MeOH:H2O containing 0.1% formic acid at a flow rate of 10 μL/min via 100:1 splitter. Purine and hexakis(2,2,3,3-tetrafluoropropoxy)phosphazine (HP-921) were spiked into the sheath liquid at a concentration of 0.02% *v* to produce corresponding reference ions at *m/z* 121.05087 and *m/z* 922.00978 for real time internal mass correction. The TOF-MS was operated in positive-ion mode for detection of cationic/zwitter-ionic metabolites over a m/z range of 50-1700 at an acquisition rate of 2 Hz and acquisition time of 500 ms with fixed voltage settings for the following: fragmentor = 120 V, skimmer = 65V and Oct 1 RF = 750 V. In the ESI source the settings were V_{cap} = 2000 V, nozzle voltage = 2000 V, nebulizer gas = 10 psi, sheath gas = 3.5 L/min at 195 °C, and drying gas = 8 L/min at 300 °C.

4.3.7 Sample Injection Configuration and Data Workflow

A seven-segment injection format was used in MSI-CE-MS for all analyses in order to enhance sample throughput without loss of information content due to ion suppression.[32] Different injection configurations can be used to encode information temporally according to the experimental design, however two specific formats were used in this work. First, a dilution trend filter based on a serial injection of a pooled plasma filtrate sample as quality control (QC) was performed at different dilutions, including a blank (*i.e.*, buffer filtrate).[32] Untargeted feature picking was performed using Molecular Feature Extractor (MFE, MassHunter Qualitative Analysis, Agilent Technologies Inc.) in order to detect reproducible yet authentic ions (*i.e.*, $[M+H]^+$ or $[M+Na]^+$) with peak

heights over 300 counts while excluding chemical and biochemical noise. MFE was conducted over a time range (≈ 8-30 min) of the separation while excluding the salt front and the electroosmotic flow (EOF) regions of the separation, where ion co-migration and signal suppression are prevalent. An authentic plasma metabolite from the dilution trend filter was annotated by its characteristic *m/z*:RMT provided the signal was detected with adequate precision (*CV* < 40%, *n*=3) and linearity (R^2 > 0.90) in at least three dilution levels with no signal in the blank sample segment.[32] This process was applied to derive a compound list of reproducible yet authentic molecular features while filtering spurious signals and background ions that comprise the majority of ion signals detected in ESI-MS.[32] A second injection configuration was also used in MSI-CE-MS when analyzing dynamic metabolome responses to HIT at three time intervals (0, 20, 40 min) for each subject either in their naïve (trial 1) or trained (trial 17) state, including a pooled plasma QC. Thus, the same metabolite from seven different samples migrate into the ion source over a short time interval (≈ 3-5 min) under stable ionization conditions with good quantitative performance.[32] In this way adaptive responses to exercise training (*i.e.*, treatment effect) can be readily measured by MSI-CE-MS for each subject in a single run using an accelerated data workflow in metabolomics for biomarker discovery.

4.3.8 Quality Assurance and Statistical Analysis

Quality assurance was achieved by inclusion of a pooled QC sample as the seventh sample segment injected during each run performed by MSI-CE-MS (*n*=9). The QC serves as an internal reference sample included within every separation in order to monitor long-term instrument bias and system drift that is similar to intermittent QCs used in large-scale metabolomic workflows.[36] Moreover, plasma metabolites that were initially identified and validated in the dilution trend filter were only included into the final data matrix if they were detected in a minimum of 75% of all samples with acceptable precision (*CV* < 40%). Multivariate analysis of *log*-transformed and auto-scaled data was performed using MetaboAnalyst 2.0,[37] including Wilcoxon-signed rank tests, principal component analysis (PCA), hierarchical cluster analysis (HCA) and partial least-squares-discriminant analysis (PLS-DA). In addition, two-way ANOVA with Bonferroni correction was performed to classify plasma metabolites associated with training status and exercise time course at baseline (0 min), post-HIT (20 min) and recovery while at rest (40 min). Receiver operating characteristic (ROC) curves were also performed using ROC Curve Explorer &

Tester (ROCCET)[38] after *log*-transformation of data for identification of single or ratiometric plasma metabolites associated with training status (naïve; trained) and exercise timing (0; 20 min). All data processing involving electropherograms was performed using Igor Pro 5.0 (Wavemetrics Inc., Lake Oswego, OR).

4.4 Results

Figure 4.1(a) depicts the blood sampling protocol used during two HIT sessions at three time intervals (*i.e.*, 0 min/baseline, 20 min/post-HIT and 40 min/recovery) for each subject. Overweight/obese women ($BMI > 25$, $n = 9$) performed 6 weeks of supervised HIT, which involved 10 X 60 s cycling efforts interspersed with 60 s of recovery, 3 times per week. The workload used in ergometer cycling was increased to maintain a maximal heart rate (HR_{max}) ≈ 90% during exercise training[35] with the exception of trial 17 (second last trial), which had the same workload as trial 1 (baseline) for each subject. This study design thus allows for evaluation of adaptive training responses on an individual level when standardized cycling trials are performed at the same absolute workload for the same subject in their naïve and trained state. **Figure 4.1(b)** highlights the underlying heterogeneity of the female cohort by principal component analysis (PCA) when using a 2D scores plot to relate subjects in terms of five variables that reflect their baseline aerobic fitness (*e.g.*, VO_{2max}, W_{max}, HR_{max}) and body composition (*e.g.*, *BMI, % fat*). **Table 4.1** also summarizes treatment outcomes after HIT, including gains in aerobic endurance capacity (in brackets). Overall, significant improvements in maximal oxygen uptake (VO_{2max}) and maximal workload (W_{max}) as well as a decrease in average *HR* during strenuous exercise were achieved for trained subjects despite differences in adiposity and baseline fitness. Although there was no significant change in total body mass or *BMI*, a modest decrease in body fat was realized after HIT notably in the abdominal and leg regions.[35]

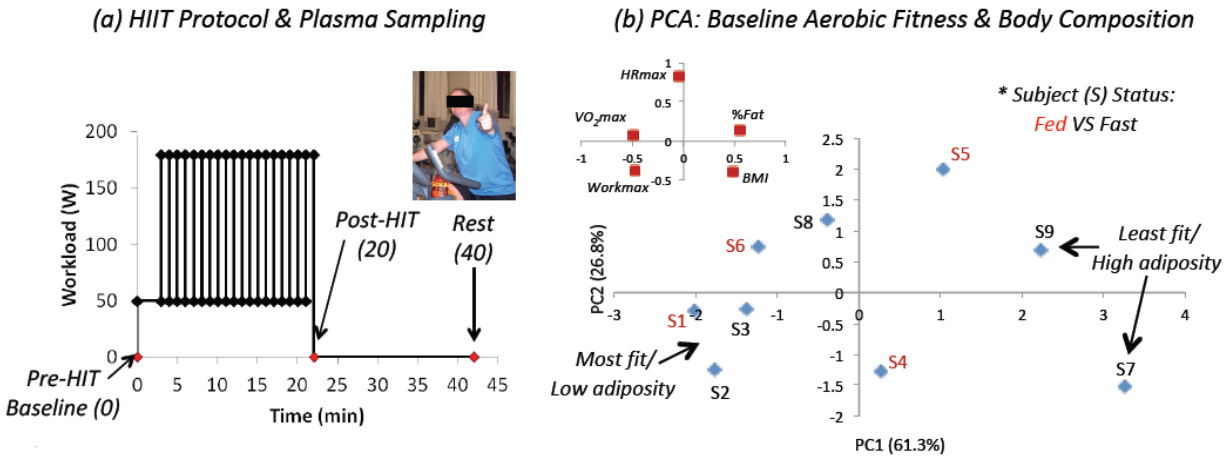


Figure 4.1. (a) Blood sampling protocol for a cross-over HIT intervention study for assessment of exercise responsiveness involving (b) a cohort of overweight/obese females (n=9, *BMI* > 25 kg/m^2) with variable body composition and aerobic fitness levels at baseline as summarized in PCA 2D scores and loadings plots. Blood samples were acquired at three time intervals for subjects in their naïve (trial 1) and trained (trial 17) state after 6 weeks of HIT. In this case, both exercise trials were performed under standardized conditions using the same absolute workload initially set for each participant.

Table 4.1. A 6-week HIT intervention trial for improving cardiorespiratory fitness and body composition of a cohort of healthy yet sedentary/obese women. Despite differences in exercise responsiveness among participants, an overall improvement in aerobic endurance capacity outcomes was measured relative to baseline (in brackets) that also contributed to a modest decrease in body fat.

Subject	***Pre-HIT status***	***Age (yrs)***	***BMI (kg/m^2)***	***Fat (%)***	***VO_2 max (mL/kg/min)***	***Workload max (W)***	***Avg. Heart Rate max (bpm)***
S1	Fed	24	25.5	33.8 (+0.2)	39.3 (-0.90)	248 (+18)	164 (-8.8)
S2	Fast	23	27.8	32.0 (-0.6)	37.5 (0.0)	257 (+32)	158 (-4.1)
S3	Fast	23	25.7	31.7 (-0.6)	26.8 (+9.2)	266 (+20)	169 (-8.1)
S4	Fed	41	27.8	40.1 (-0.3)	29.3 (+3.1)	205 (+41)	150 (-9.0)
S5	Fed	23	27.4	47.1 (-1.8)	29.8 (+11)	173 (+54)	182 (-20)
S6	Fed	33	25.6	35.1 (-1.8)	33.9 (+4.8)	228 (+17)	174 (-19)
S7	Fast	41	36.1	50.1 (-1.2)	19.4 (+3.0)	190 (+40)	156 (-12)
S8	Fast	21	25.1	37.1 (-0.4)	30.4 (+7.6)	194 (+12)	174 (-6.5)
S9	Fast	23	30.0	50.5 (-2.4)	24.1 (+5.3)	170 (+30)	170 (-10)
Median	--	*23*	*27.4*	*37.1 (-0.6)*	*29.8 (+4.8)*	*205 (+30)*	*169 (-9)*

4.4.1 Injection Configuration and Metabolomics Data Workflow

Seven discrete plasma filtrate samples are analyzed simultaneously by MSI-CE-MS to enhance sample throughput,[32] where ions from each sample migrate in free solution prior to ionization under steady-state conditions using an isocratic buffer system. **Figure 4.2(a)** highlights the specific injection configuration used in MSI-CE-MS where dynamic metabolic responses to HIT at three time intervals is evaluated for each subject in their naïve and trained states within a single run. Importantly, this approach enables direct assessment of the exercise responsiveness on an individual level (*i.e.*, treatment effect) while also including a pooled sample as a quality control (QC) in every run. **Figure 4.2(b)** depicts an extracted ion electropherogram for *L*-carnitine (C0) in seven distinct plasma filtrate samples, whereas 3-chloro-*L*-tyrosine (Cl-Tyr, 25 μM) is used as a single non-isotopic internal standard (IS) that is added to all samples in order to correct for variations in injection volume between samples. In this case, the ion response for each plasma metabolite is normalized to its corresponding IS derived from the same injection segment.[32] Thus, unambiguous evaluation of adaptive responses to HIT, as well as time course changes following strenuous exercise is realized using signal pattern recognition in MSI-CE-MS. For instance, subject 8 (S8) undergoes a 1.6-fold increase ($p < 0.01$) in baseline plasma C0 after 6 weeks of HIT, as well as a distinct time course response with lower C0 levels post-exercise (20 min) and during recovery (40 min) after exercise training. In contrast, the analogous exercise trial performed by the same subject in their naïve/untrained state resulted in an increase in plasma C0 post-exercise. Since ions from different samples enter the ion source within a short time interval (≈ 2-5 min), accurate quantification is feasible when using a single non-isotope internal standard without ion suppression or loss in separation performance.[32]

Data pre-processing also represents a major obstacle to large-scale metabolomic studies since individual samples require complicated time alignment and peak picking algorithms to identify putative ions consistently across a batch of runs.[34] MSI-CE-MS provides an alternative workflow for data processing, wherein the ability to analyze multiple samples simultaneously provides increased confidence for metabolite annotation and biomarker identification. A dilution trend filter[33] was first used to generate a list of authentic metabolites originating from plasma that increases data quality by eliminating spurious artifact peaks and background ions generated during spray formation. In this case, a pooled QC

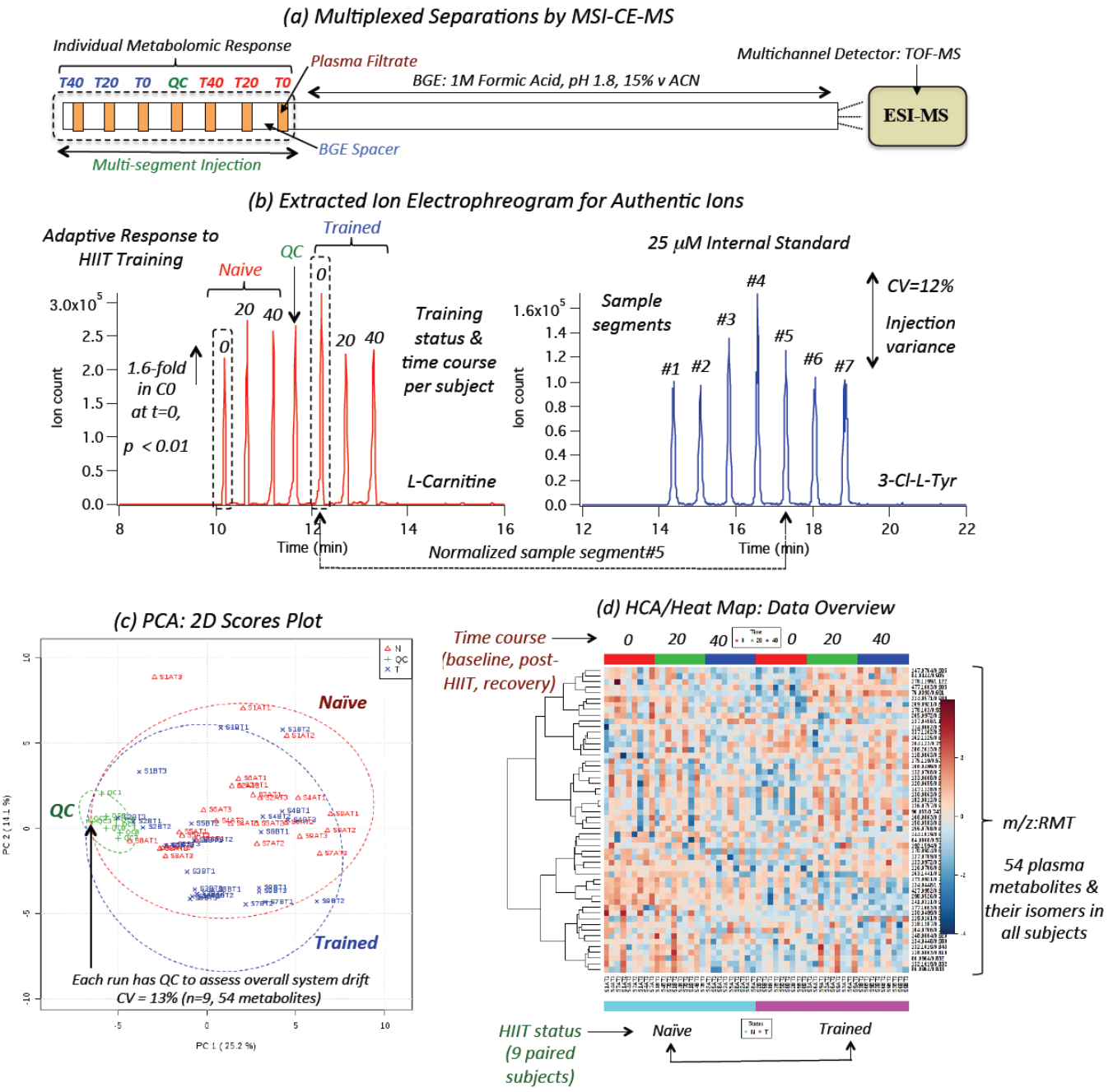


Figure 4.2. (a) High-throughput metabolomic profiling with high data fidelity by MSI-CE-MS based on serial injection of seven sample segments within a single run. (b) Extracted ion electropherograms for plasma *L*-carnitine (C0) and internal standard (IS) depicts seven discrete ion signals derived from a pair of three time-course plasma samples from an individual subject (naïve vs trained) and a pooled plasma reference as QC. (c) PCA 2D scores plot highlighting the good reproducibility of MSI-CE-MS based on tight clustering of 9 QCs relative to large biological variability in plasma metabolome derived from female subjects in their naïve and trained states. (d) 2D heat map with hierarchical cluster analysis (HCA) providing a data overview of the dynamic plasma metabotype for each subject based on their training status and exercise time course during standardized HIT trials.

sample was diluted at four levels, along with a blank (*i.e.*, buffer filtrate) that included three repeated injections of the highest concentration sample. Molecular feature extraction was performed to identify authentic ions based on their characteristic signal pattern while excluding spurious signals, background ions (in blank) or sample-derived ions detected with poor precision (*CV* > 40%) as described previously.[32] Application of this primary screen on a pooled QC generated a list of 54 polar metabolites common to all samples, which was then used for targeted analysis of individual plasma filtrate specimens as summarized in **Table A4.1**. In this way, data pre-processing is more streamlined when using a conventional targeted metabolite profiling workflow. Also, false discoveries due to data over-fitting are reduced since spurious signals detected in ESI-MS are rejected prior to statistical analysis. **Figure 4.2(c)** depicts a PCA 2D scores plot highlighting the tight clustering of 9 QCs relative to large biological variability in plasma metabolome derived from women in their naïve and trained states. Acceptable within-day reproducibility without system drift is evident as reflected by an average *CV* of 13% (n = 9) for plasma metabolites consistently detected in all QC samples over the duration of the analysis. **Figure 4.2(d)** summarizes the overall data structure when using a 2D heat map with hierarchical cluster analysis (HCA) based on 54 plasma metabolites denoted by their characteristic mass/charge:relative migration time (*m/z*:RMT) that were measured at three time intervals (0, 20, 40 min) for paired subjects in their trained and naïve states. The cross-over study design together with two independent factors evaluated in this exercise intervention (*i.e.*, training status; time course) allows for evaluation of adaptive metabolomic responses to HIT among individual participants that are associated with gains in their cardiorespiratory fitness (**Table 4.1**).

4.4.2 Plasma Markers of Adaptive Training Responses to HIT

Complementary statistical methods were used to classify plasma metabolites as putative markers of high-intensity exercise, as well as subject training status. Most metabolites were identified (> 60%) after spiking plasma samples with authentic standards, whereas unknown metabolites were annotated by their *m/z*:RMT if structures were not found after searching public databases (*e.g.*, HMDB) based on their accurate mass/empirical formula or lead candidates had inconsistent electromigration behavior (*e.g.*, cations). **Figure 4.3** highlights between-subject variance using 2-way ANOVA with Bonferroni correction (p < 0.05) to correct for multiple hypothesis testing. Two plasma metabolites are associated with training status, including C0 (p = 5.72 E-3) and an unknown

metabolite *m/z* 163.1441 (p = 2.80 E-4). In both cases, significant upregulation of circulating levels of these metabolites occur following HIT intervention irrespective of feeding status notably at rest prior to strenuous exercise. In addition, two other plasma metabolites are associated with both training status and notably exercise time course, namely hypoxanthine (HyX) and *O*-acetyl-*L*-carnitine (C2). HyX (p =8.04 E-8) and C2 (p =1.42 E-2) were significantly elevated in plasma after HIT with modest decreases during recovery (*i.e.*, 40 min) in the case of trained subjects. Interestingly, the magnitude of plasma HyX accumulation post-exercise is greatly attenuated after 6 weeks of exercise training.

Further investigation of adaptive responses to HIT when comparing plasma metabolome profiles of subjects while at rest is shown in **Figure 4.4(a)**. In order to take advantage of the cross-over study design, a 2D volcano plot derived from a Wilcoxon-signed rank test under stringent constraints ($p < 0.01$; fold-change > 1.5; significant counts 75%) demonstrates that five plasma metabolites are modulated after HIT intervention. In addition to confirming exercise-induced upregulation in plasma C0, two other unknown metabolites were also found to increase with exercise training. In addition, oxidized glutathionylcysteine mixed disulfide (GSH-Cys-SS) and an unknown metabolite decrease in plasma after low-volume HIT. A 2D scores plot from PLS-DA shown in **Figure 4.4(b)** highlights that training status is readily distinguished for the cohort (*i.e.*, naïve VS trained) with good model fitting (R^2 = 0.994) and adequate robustness (Q^2 = 0.564) after cross-validation. Similarly, C0 is by far the most significant plasma metabolite upregulated following HIT as indicated by the variable importance in the projection (VIP) ranking. Moreover, receiver operating characteristic (ROC) curves demonstrate that plasma C0 status provides good classification of training status for overweight/obese women at rest (AUC = 0.951, p = 2.94 E-4) that is further improved when using ratiometric markers depicted in **Figure 4.4(c)** with excellent specificity and sensitivity, including *m/z* 86.0964/Leu (AUC = 1.00, p = 3.89 E-5) and *m/z* 163.1141/C2 (AUC = 1.00, p = 2.92 E-4). Both unknown plasma metabolites were also identified in VIP rankings as indicative of training status when using PLS-DA.

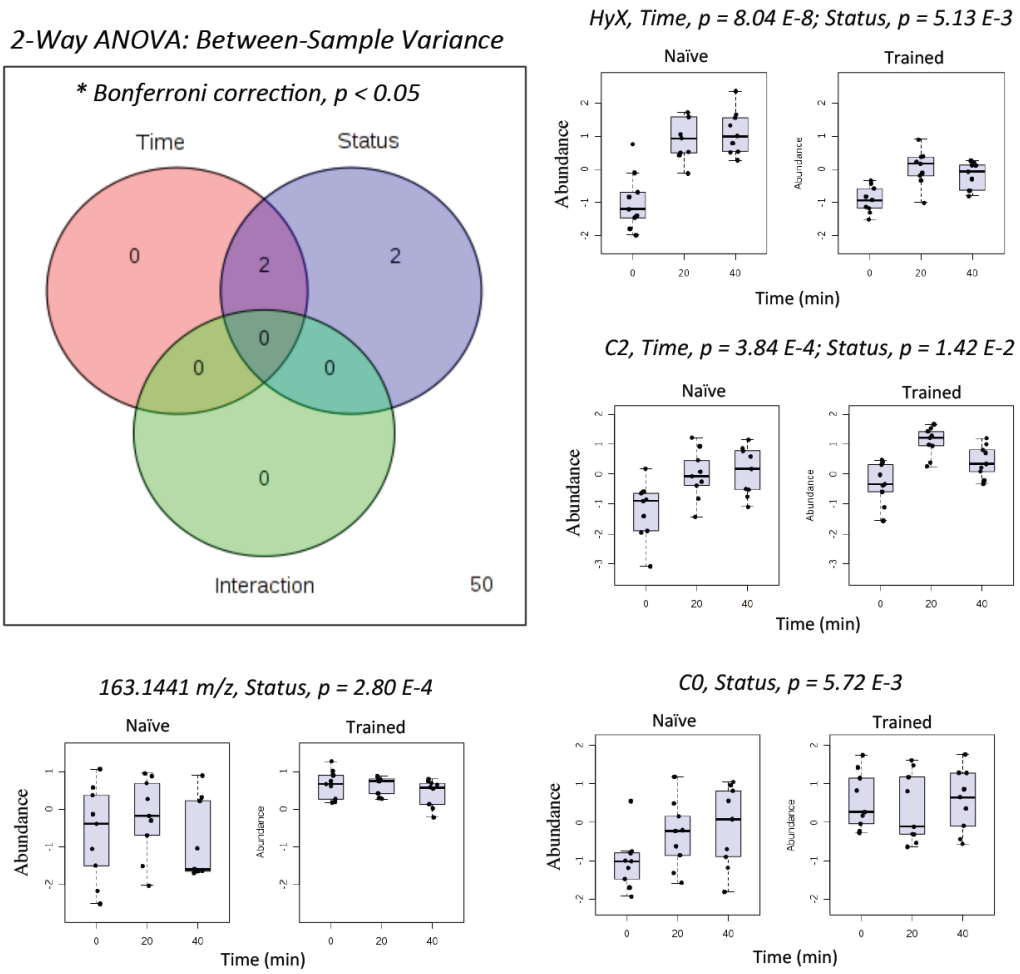


Figure 4.3. Two-way ANOVA with Bonferroni correction to reduce false discoveries for characterization of dynamic metabolomic responses to HIT training that takes advantage of time series data from the experiment at baseline (0 min), post-exercise (20 min) and recovery/rest (40 min), which confirms that both HyX and C2 are significant plasma features reflecting both time and training status among different subjects, whereas plasma C0 and unknown metabolite *m/z* 163.1441 levels are significantly elevated after 6 weeks of exercise training.

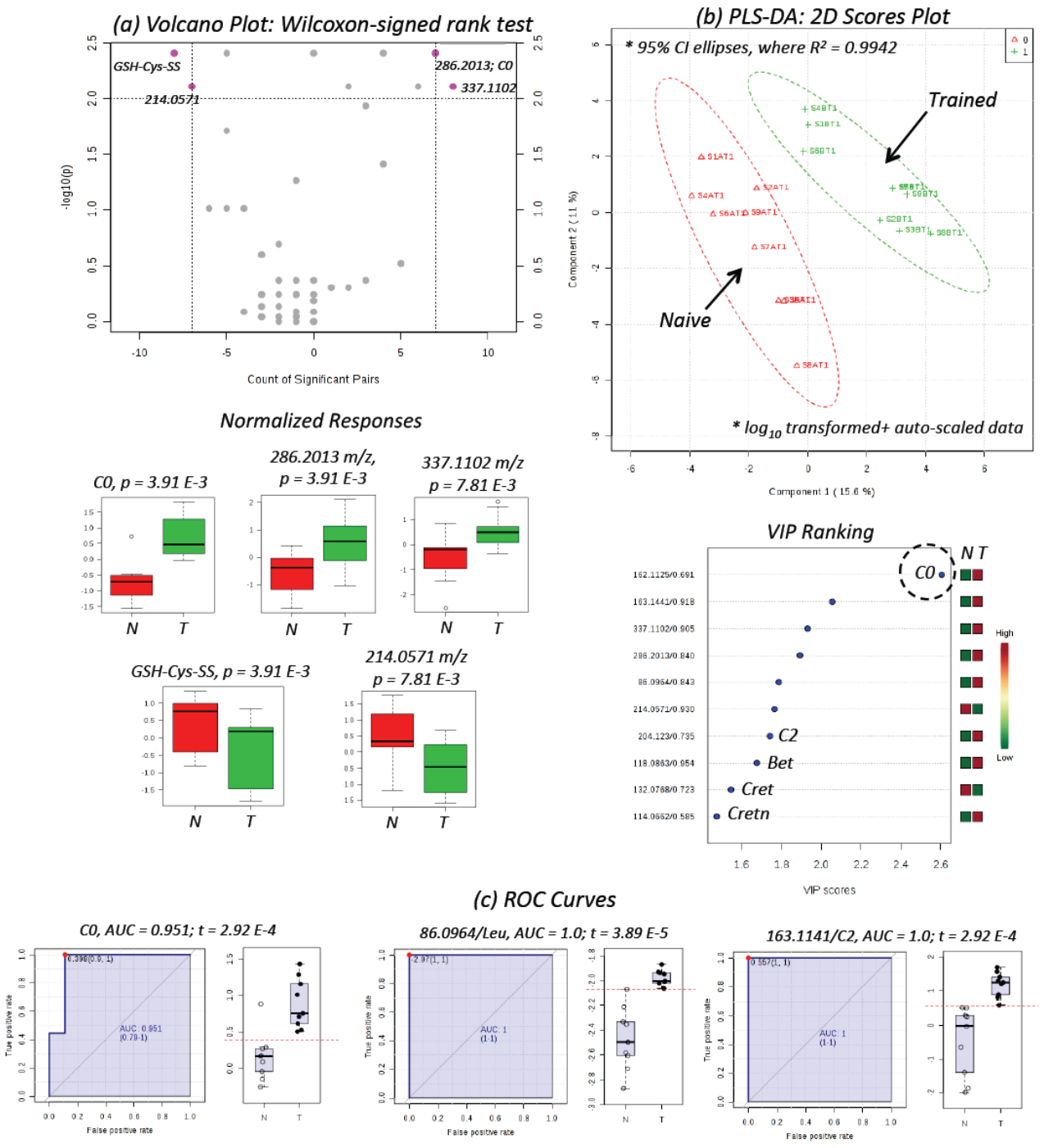


Figure 4.4. Univariate and multivariate data analysis for classification of plasma metabolites associated with adaptive responses to HIT for subjects while at rest (0 min/baseline), including (a) Wilcoxon signed rank test that takes advantage of paired subject relation, (b) PLS-DA with variable importance in projection (VIP) ranking of metabolites and (c) receiver operating characteristic (ROC) curves derived from single and ratiometric plasma metabolites associated with training status. All methods confirm that plasma *L*-carnitine (C0) levels are significantly upregulated in trained subjects after 6 week HIT intervention.

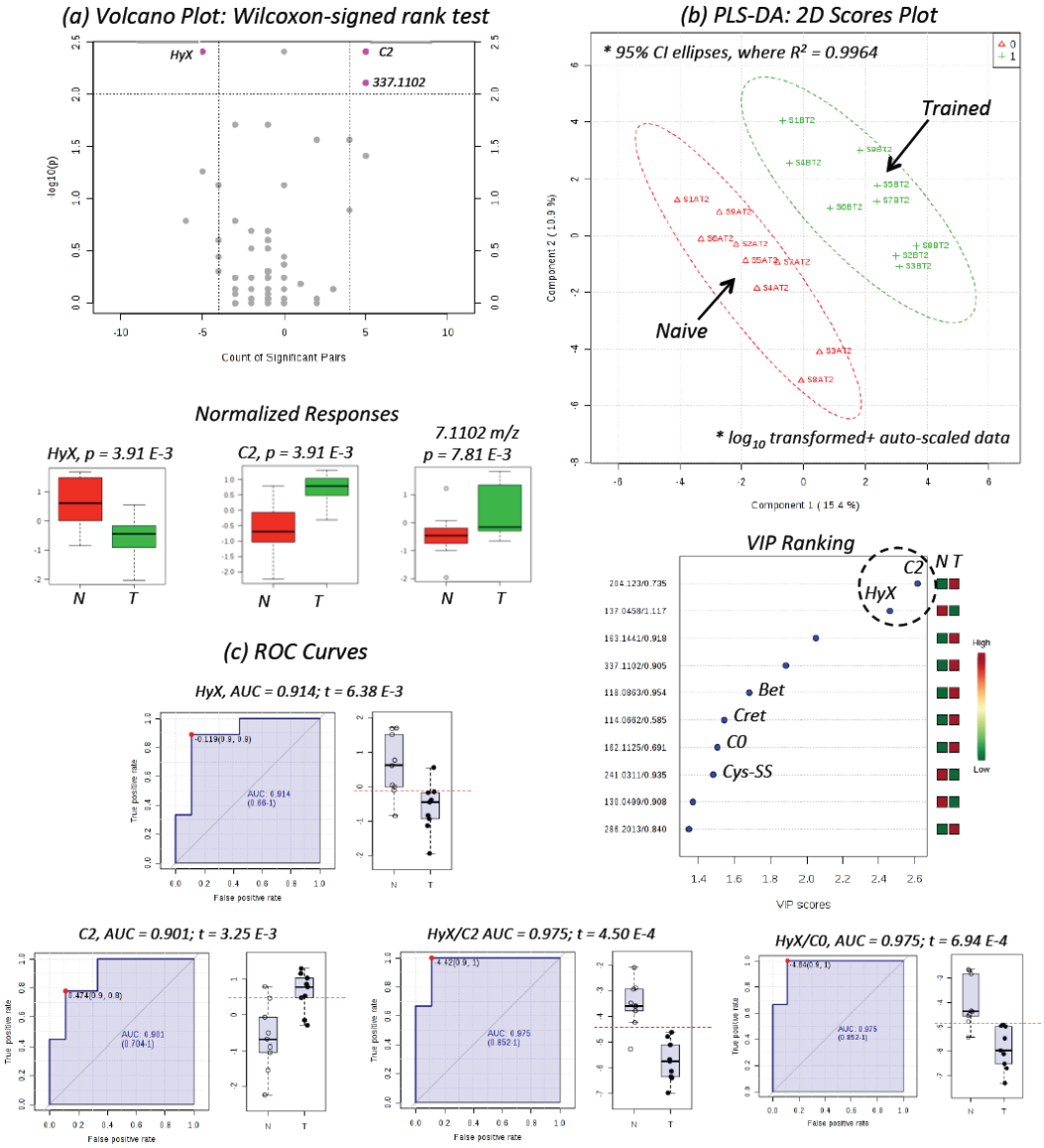


Figure 4.5. Univariate and multivariate data analysis for classification of plasma metabolites associated adaptive training responses to HIT for subjects after strenuous exercise (20 min/post-exercise), including (a) Wilcoxon signed rank test that takes advantage of paired subject relation, (b) PLS-DA with variable importance in projection (VIP) ranking of metabolites and (c) receiver operating characteristic (ROC) curves derived from single and ratiometric plasma metabolites associated with training status. All methods confirm that plasma hypoxanthine (HyX) and *O*-acetyl-*L*-carnitine (C2) levels are significantly attenuated and upregulated after HIT, respectively.

An analogous approach was also used to compare adaptive responses when examining changes in plasma metabolome for naïve and trained subjects after strenuous exercise (20 min). In this case, a 2D volcano plot based on a Wilcoxon-signed rank test reveals that two key plasma metabolites (including a third unknown metabolite in **Figure 4.4**) are significantly modulated ($p < 0.01$) after each subject completed standardized HIT sessions in their trained and naïve states as shown in **Figure 4.5(a)**. Similar trends were also found when evaluating dynamic metabolomic responses for participants during recovery (40 min) following HIT (data not shown). Consistent with 2-way ANOVA results, plasma HyX undergoes significant attenuation for trained subjects post-exercise in contrast to C2, which was elevated following strenuous exercise. The latter effect is more pronounced for subjects after HIT than plasma C2 changes measured at rest/baseline (**Figure 4.4**). Both HyX and C2 were also classified as top plasma markers indicative of training status among 54 plasma metabolites reliably measured by MSI-CE-MS based on VIP rankings from PLS-DA as shown in **Figure 4.5(b)**. In addition, ROC curves depicted in **Figure 4.5(c)** demonstrate that both HyX (*AUC* = 0.914, p = 6.38 E-3) and C2 (*AUC* = 0.901, p = 3.25 E-3) serve as promising diagnostic markers of adaptive training in women after HIT, which is further enhanced when considering ratiometric markers based on HyX/C2 (*AUC* = 0.975, p = 4.50 E-4) or HyX/C0 (*AUC* = 0.975, p = 6.94 E-4).

Figure 4.6 shows dynamic changes for HyX derived from three representative subjects in the study when using MSI-CE-MS, which underscores the large disparity in adaptive training responses between-subjects. For instance, S6 has over a 116-fold accumulation in plasma HyX post-exercise in their naïve state even 20 min after exercise is completed; however, this outcome is dramatically attenuated by almost 20-fold when performing the same standardized exercise trial following HIT. In contrast, both S2 and S9 had far less pronounced increases in HyX post-exercise in their naïve states, whereas improved training responsiveness is reflected by a faster return to baseline during recovery while at rest, notably in the case of S9. Indeed, there was a large biological variation in plasma HyX levels for naïve subjects at recovery (40 min) ranging from a mere 2-fold to a 116-fold increase after strenuous exercise. Thus, dynamic metabolomic responses to high-intensity exercise (*i.e.*, exercise-induced oxidative stress) and adaptive changes to exercise training (*i.e.*, treatment effect) are readily captured on an individual level by MSI-CE-MS when using a cross-over study design. **Figure 4.6** also summarizes concentration profiles measured for the three plasma

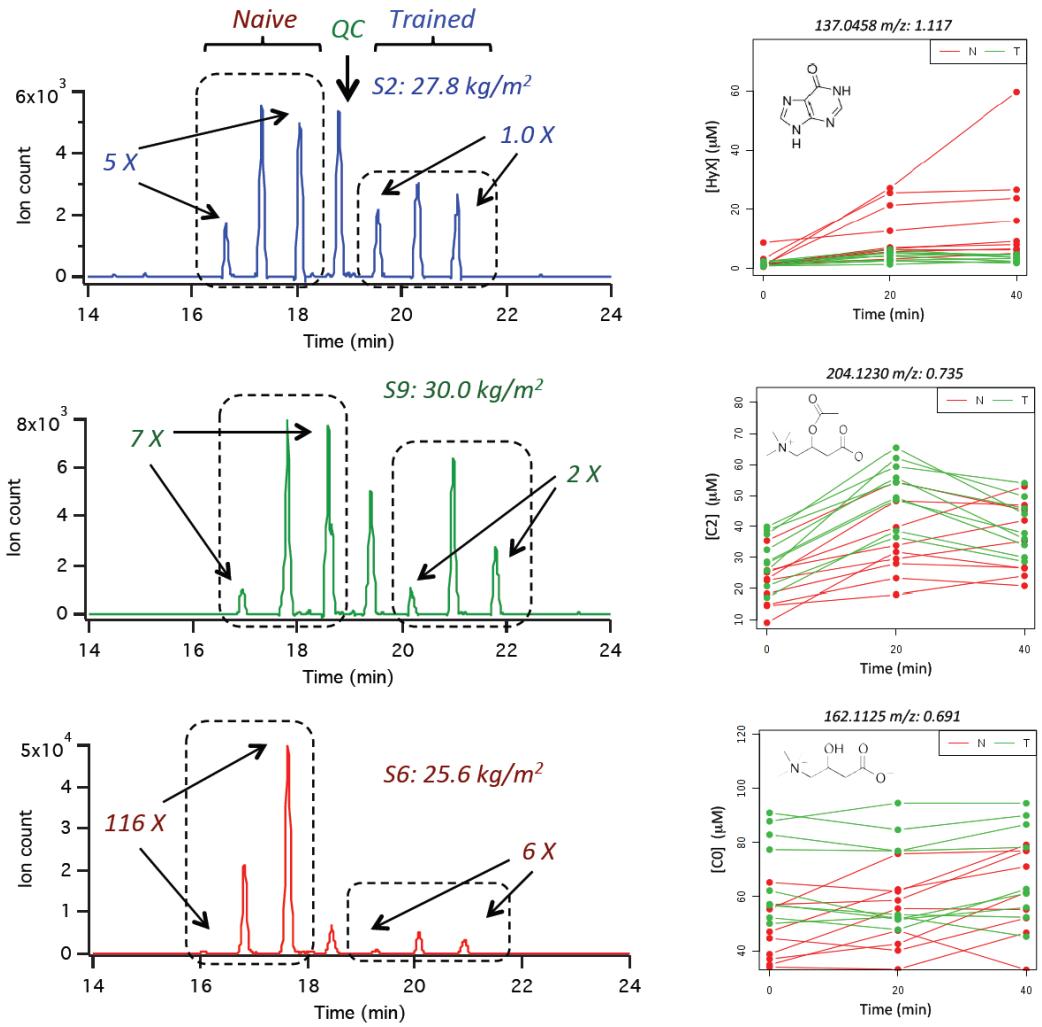


Figure 4.6. Extracted ion electropherograms for HyX highlighting the differential adaptive responses of three different participants to HIT and their plasma concentration levels that reflect attenuated adenine degradation with improved ATP recycling capacity. There is a pronounced attenuation in plasma HyX accumulation post-exercise (20 min) for trained subjects, whereas C2 undergoes a subtle attenuation only during recovery (40 min). Also, adaptive exercise training promotes improved muscle tissue mitochondrial function as related to fatty acid transport/beta-oxidation as demonstrated by upregulation in circulating levels of C0 at rest (0 min) that undergoes no significant change post-exercise after HIT.

markers of adaptive exercise training from individual subjects, namely C0, C2 and HyX. Overall, there was an average 1.5-fold (p = 1.58 E-3) increase in plasma C0 concentrations (ranging from 1.3 to 1.7-fold) from baseline while at rest, whereas there was no significant time-dependent change in C0 during exercise notably for trained subjects. Plasma C2 concentration levels also increased

significantly after HIT while at rest (average 1.4-fold increase, p = 1.63 E-4). In addition, time-dependent changes in C2 concentrations post-exercise were quite distinct depending on training status. For example, there was a significant increase in plasma C2 for naïve subjects after strenuous exercise with an average 1.8-fold increase post-HIT at 20 min (ranging from 1.3 to 3-fold) relative to baseline that subsequently plateaus upon recovery. In contrast, trained subjects exhibited a much larger increase in plasma C2 post-HIT (20 min) from baseline, which subsequently return back to baseline upon recovery. Overall, HyX had the most significant time-dependent changes over a wide dynamic range especially for naïve subjects, which was drastically attenuated after 6 week HIT intervention as reflected by an overall 9-fold reduction in plasma concentrations during recovery (p = 1.48 E-2). This data is consistent with 2-way ANOVA results, which was used to originally identify C0, C2 and HyX as significant plasma markers associated with training status and/or high intensity exercise time course.

4.5 Discussion

Strenuous exercise triggers catabolic energy demands for contracting skeletal muscle that leads to metabolic acidosis when the rates of phosphagen, glycolytic and notably mitochondrial respiration are inadequate for cellular ATP regeneration.[39] Exhaustive exercise without training is associated with exercise-induced oxidative stress, tissue injury and deleterious health outcomes.[40] In contrast, HIT utilizes short yet intermittent bursts of high-intensity exercise that serves as a potent stimulus for beneficial metabolic adaptations during physical training with a modest systemic inflammatory response.[41] Since lack of time remains a major barrier to exercise adherence, low-volume HIT offers a practical approach for enhancing cardiorespiratory fitness while reducing chronic disease risk as compared to endurance-based training protocols.[7] In our work, subjects were sedentary yet otherwise healthy overweight female adults ($BMI \approx 27$ kg/m^2) with a wide disparity in baseline aerobic fitness and adiposity as depicted in **Figure 4.1(b)**. An overall improvement in cardiorespiratory fitness for the cohort was achieved after HIT based on increases in VO_{2max} with a median gain of +4.8 mL kg^{-1} min^{-1} or about 16% from baseline. Indeed, low cardiorespiratory fitness is strongly correlated with cardiovascular disease mortality,[42] as well as higher risk of obesity, hypertension and diabetes.[43] Similarly, improved exercise capacity was also achieved after HIT as reflected by a median gain in W_{max} of 30 W or about 15% from baseline. Trained subjects were also found to complete the standardized

HIT session at a lower self-perceived exertion level of physical activity (*e.g.*, reduced by about 2.5 on a 10 point-scale) that is consistent with performing the same exercise workload at a lower average *HR*. Interestingly, there was no direct correlation between gains in VO_{2max} and W_{max} among subjects, as well as among other physiological (average *HR*) or body composition *(% fat, BMI*) parameters. Although there was no overall change in total body mass after exercise training,[35] a modest yet significant decrease in body fat (median ≈ -0.6%, p = 4.30 E-3 with paired t-test) was attained for all participants irrespective of feeding status or body composition with the exception of one subject (S1) in this study. Low fat loss response with exercise training can be confounded by psychosocial behavior of subjects, such as consumption of high-fat/sweet food as reward after exercise.[44] Nevertheless, large-scale intervention studies have shown that the health benefits of exercise training are derived even in the absence of significant weight loss.[10] Two subjects (S1, S2) were found to derive negligible benefits in VO_{2max} after HIT, where "non-responders" to exercise training have been reported to comprise about 20% of the population.[9] In both cases, these women were among the most aerobically fit participants with low adiposity at baseline despite being classified as overweight based on their *BMI*. Nevertheless, S1 and S2 did in fact derive salutary benefits in terms of gains in W_{max} after HIT like all other subjects in this study. These results highlight that gains in exercise endurance capacity are not only limited to physiological mechanisms related to increases in maximal cardiac output, such as VO_{2max}.[45] In this work, feeding status prior to exercise was not found to impact adaptive physiological responses to HIT,[35] however the fasted state has been reported to be more effective at improving glucose tolerance than fed-state training when subjects are consuming a hyper-caloric fat-rich diet.[46] Participants in our study were instructed to maintain a regular diet and lifestyle under free-living conditions throughout the intervention period, whereas standardized breakfasts were provided only during HIT sessions either prior (fed) to or after (fast) strenuous exercise. As a result, considerable biological variance in the plasma metabolome exists within-subjects, which has been shown to vary up to 50% even for standardized fasting morning specimens collected from healthy subjects.[47]

4.5.1 Adaptive Metabolic Responses to Exercise Training

Despite recent advances in developing genetic predictors of exercise responsiveness,[9,48] only a minor fraction of the total variance related to gains in VO_{2max} can be modeled accurately due to confounding non-genetic factors,

including diet, lifestyle and environment. Alternatively, metabolic signatures in plasma associated with glucose uptake and fatty acid metabolism in skeletal muscle provide deeper insights into exercise performance and cardiovacular disease risk.[19] In this work, 4 out of 54 plasma metabolites detected in overweight women were significantly modulated by training status and/or exercise time course (**Figure 4.3**), including HyX, C2, C0 and an unidentified cationic metabolite (*m/z* 163.1441). PLS-DA and ROC curves both confirmed that C0 was the most significant plasma metabolite upregulated over the 6-week HIT intervention period. More than 95% of total C0 is stored within skeletal muscle, where it serves as a key co-factor for carnitine palmitoyltransferase 1 to transport long-chain fatty acids into the inner mitochondria for beta-oxidation. Due to its high abundance in muscle tissues, it also acts as a buffer to regulate the mitochondrial acetyl coenzyme A pool levels from pyruvate oxidation, thereby forming a critical link between glycolytic and citric acid energetic pathways.[49] Thus, C0 is an important regulator of skeletal muscle fuel selection. For these reasons, oral C0 supplementation has long been explored as a putative ergogenic aid to enhance fat utilization notably during intense exercise provided that adequate uptake into muscle tissue is achieved.[50,51] C0 therapy is also widely used for treatment of in-born errors of metabolism as related to carnitine deficiency, fatty acid oxidation defects or organic acidemias.[52] Recently, chronic oral ingestion of C0 together with carbohydrate solution has been shown to reduce muscle anaerobic ATP production during high intensity exercise while eliciting greater exercise performance.[53] In addition, C0 supplementation may serve as a prophylaxis to enhance recovery after physical exertion as reflected by lower levels in biochemical markers of oxidative stress, purine degradation and muscle tissue injury.[54] C0 in skeletal muscle is derived from diet after transport from plasma membrane transporters, as well as *de novo* biosynthesis from the intermediate, *N*-trimethyllysine from endogenous protein breakdown and various vegetable sources.[55] Exercise training serves as a way to enhance lipid oxidation by upregulating genes associated with fatty acid uptake across the plasma membrane and transport to the mitochondria.[56] Since HIT promotes mitochondria biogenesis with increased muscle oxidative capacity,[57] enhanced C0 status measured in trained subjects in our work suggests increased biosynthesis in the absence of major dietary changes. For instance, average baseline levels of plasma C0 increased significantly (p = 1.58 E-3) from 46 μM to about 70 μM in the trained population (**Figure 4.6**), which is similar to the effects produced after high-dose oral C0 supplementation.[58] Although direct measurements of C0 were not performed in muscle tissue biopsies in this study, exercise-induced

improvements in C0 status at rest were weakly correlated (r = 0.5048, p = 0.166) with gains in aerobic endurance capacity (W_{max}). Indeed, greater C0 availability not only prevents C0 insufficiency during high-intensity exercise, but also improves fatty metabolism in muscle tissue while at rest that is relevant for treatment of chronic diseases associated with mitochondrial dysfunction, such as type 2 diabetes.[59] In addition to C0, a Wilcoxon-signed rank test revealed that four additional metabolites were also modulated with exercise training, notably GSH-Cys-SS (**Figure 4.4**). There was a significant decrease in plasma GSH-Cys-SS levels in the trained population, which is a circulating mixed oxidized disulfide derived from cell export of reduced glutathione (GSH) and thiol-disulfide exchange with plasma cysteine.[60] Aging is associated with a pro-oxidant shift in the plasma glutathione redox state reflecting declining antioxidant capacity and an increased susceptibility to age-related diseases.[61] Thus, HIT reduces plasma thiol redox status that is associated with intra-cellular glutathione homeostasis since exercise training enhances tissue-dependent glutathione antioxidant defenses.[62] Due to the low abundance and intrinsic reactivity of reduced thiols in plasma, thiol-specific maleimide labeling methods are needed for reliable determination of thiol redox status without oxidation artifacts generated during sample collection and workup.[63]

In our work, plasma HyX was the most significant plasma marker associated with strenuous exercise and training status among overweight/obese women (**Figure 4.3**), which is consistent with LC-MS-based metabolomic studies when sampling blood 60 min after exercise testing.[19] However, the magnitude and dynamics of HyX changes in plasma were highly variable between-subjects when comparing HIT trials for participants in their naïve/untrained state (**Figure 4.6**). Previous studies have shown that HyX is a marker of adenine nucleotide degradation and energetic stress during prolonged exercise that correlates with blood ammonia and lactate levels 60 min post-exercise.[64] AMP deaminase activity represents the first step in the purine nucleotide pathway when strenuous exercise depletes ATP reserves beyond capacity to recycle ADP, which has been proposed as an important mechanism to preserve muscle function.[65] Inosine monophosphate (IMP) is thus a key intermediate that can either reform AMP or be further catabolized to generate HyX, which diffuses slowly into blood with peak concentrations occurring within 20 min after exercise.[66] Accumulation of HyX in blood after strenuous exercise is associated with depletion in total adenine nucleotides in skeletal muscle with a corresponding increase in IMP, whereas time to fatigue is inversely correlated with plasma HyX concentrations post-exercise.[64] In our case, mean plasma HyX concentrations were about 2 μM at

baseline, which did not change significantly in the trained population (3 μM, $p > 0.100$) while at rest as shown in **Figure 4.6**. In contrast, circulating HyX for naïve subjects increased on average to 18 μM during recovery (20 min post-HIT) relative to baseline. Furthermore, post-exercise accumulation of HyX for trained subjects was attenuated overall by about 9-fold (or 5-fold for paired subjects) after HIT with treatment responses varying widely between-subjects ranging from only 1.4 (S1) to over 16-fold (S5, S6). Xanthine and uric acid as terminal en-products of purine degradation were not detected in plasma since they are anionic metabolites, thus require alkaline conditions for separation by CE-MS with negative ion-mode detection.[67] However, HyX has been shown to be a more sensitive purine marker of energetic stress than uric acid with obese subjects having a higher baseline concentration and a more pronounced elevation in HyX post-exercise relative to controls.[68] PLS-DA and ROC curves highlight that HyX, as well as C2 are significantly modulated post-exercise relative to baseline, whereas the HyX/C2 ratio serves as a better plasma marker of training status ($AUC = 0.975$, $p = 4.50$ E-4) following strenuous exercise (**Figure 4.4**). Oral C2 therapy is a safe and widely used nutraceutical product in human health that is currently an adjuvant for treatment of anti-retroviral induced-neuropathy in HIV-infected patients.[69] C2 also shows promising anti-depressant activity via induction of type 2 metabotropic glutamate receptors.[70] Since carnitine acetyltransferase catalyzes the interconversion of acetyl-CoA and C2 in the mitochondria, lower enzyme activity is associated with accumulation of long-chain acylcarnitines that contributes to impaired mitochondrial function[71] and obesity-related glucose intolerance.[72] The same enzyme is also involved in transport of short/medium chain acyl-CoAs, as well as some branched-chain amino acid oxidation intermediates used in energy metabolism.[73] **Figure 4.6** shows that there was a significant upregulation in resting levels of circulating C2 from 20 μM (naïve) to about 30 μM (trained) after HIT intervention ($p = 1.46$ E-2). Also, the rate of C2 increase was higher for trained subjects with a peak plasma concentration of about 52 μM post-exercise (about 16 μM higher than naïve state at 20 min) that lowers during recovery (40 min) unlike naïve participants. This data suggests exercise-induced upregulation in carnitine acetyltransferase activity. Previous studies have measured increased maximal activities in citrate synthase and β-hydroxyacyl-CoA dehydrogenase in muscle tissue after HIT confirming adaptive improvements in muscle oxidative capacity.[35] Although muscle C2 levels do not change during low-intensity exercise, strenuous exercise elicits an increase in C2 in order to maintain free CoA pools for aerobic ATP recycling that delays the onset of fatigue.[49] Thus, positive adaptive metabolic responses to HIT as measured in

plasma are associated with improved carnitine status at rest along with greater acetylation capacity and lower adenine purine degradation for trained subjects during strenuous exercise.

4.5.2 Accelerated Workflow for Biomarker Discovery

Since large-scale metabolomic studies often use separations to enhance MS performance, sample throughput is limited while requiring resource-intensive and complicated data pre-processing prior to statistical analysis.[31] Additional time is also required for quality assurance to ensure system stability throughout the analysis, such as the use of intermittent QCs analyzed after a batch of runs within 2 h or less.[36] Identification of unknown yet biologically significant metabolites remains a major challenge for biomarker discovery in MS-based metabolomics since methods rely on retention index and/or mass spectral matching with putative candidates.[74] In cases when entries do not exist in public databases or chemical standards are unavailable,[75] *de novo* identification of unknown metabolites require *in silico* approaches for modeling separation[76] and/or collision-induced dissociation processes in MS/MS.[77] One way to better streamline efforts for unknown identification is to apply a rigorous screening approach on metabolomic data to avoid false-discoveries, such as a dilution trend filter.[33] In our work, MSI-CE-MS was used as a convenient multiplexed platform for high-throughput metabolomic profiling (≈ 3-5 min/sample) based on serial injection of seven sample segments within a single fused-silica capillary.[32] **Figure 4.7** summarizes an accelerated data workflow for biomarker discovery used in this work, which highlights several stages that take advantage of the multiplexed analysis capabilities of MSI-CE-MS. A dilution trend filter experiment (stage 2) is applied on a pooled plasma sample to encode information temporally for high data fidelity, which serves as an untargeted primary screen to certify sample-derived metabolites using signal pattern recognition.[32] Due to the experimental design (stage 1), MSI-CE-MS was configured to assess metabolomic responses to strenuous exercise, as well as evaluate adaptive training responses to HIT for each subject while using an "internal" QC to confirm lack of system drift or method bias (stage 3). In this way, high sample throughput and high data fidelity are achieved without significant loss in separation performance, such as resolution of

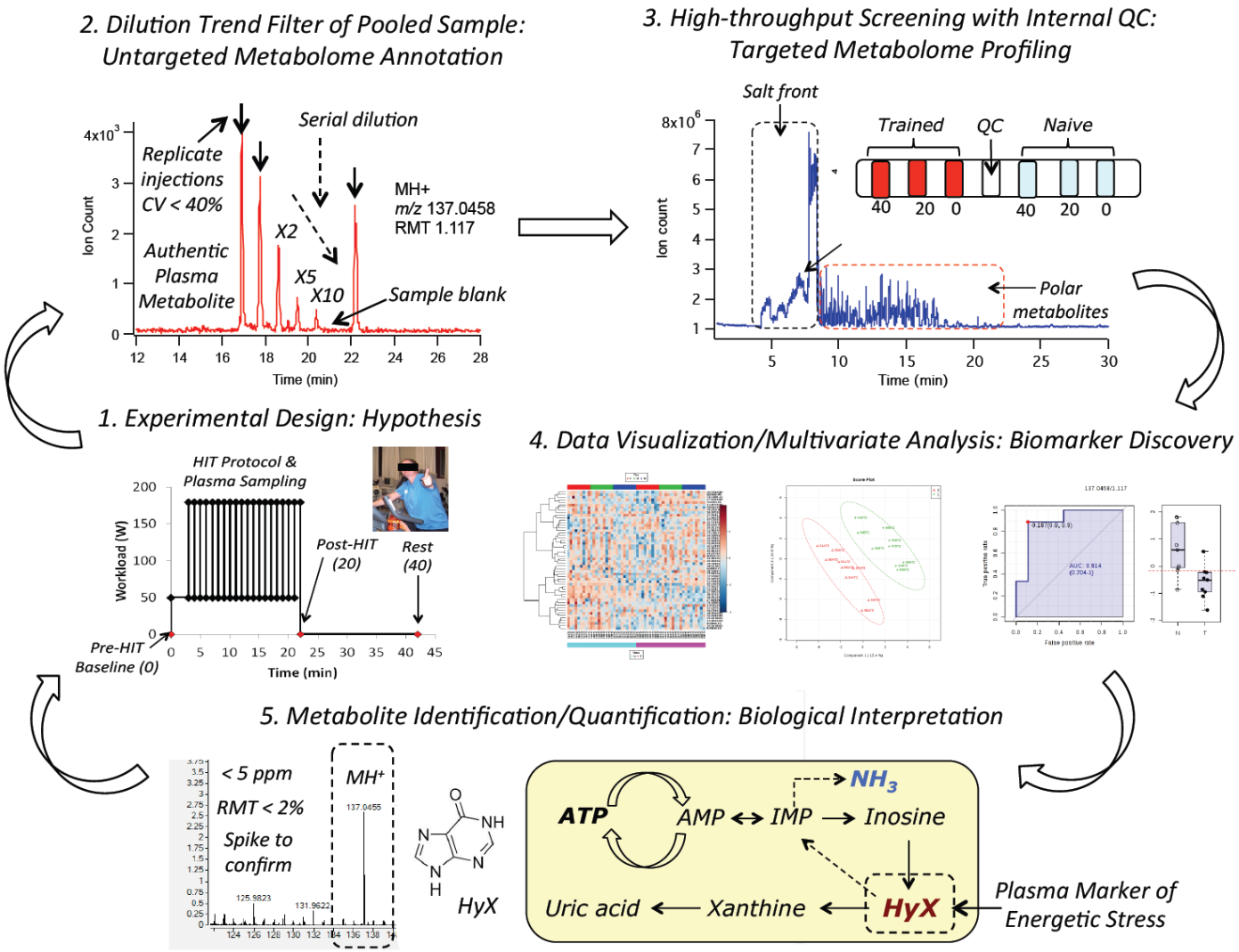


Figure 4.7. An accelerated workflow in metabolomics for biomarker discovery that takes advantage of multiplexed separations by MSI-CE-MS. In this case, a seven-segment format captures dynamic metabolomic responses to strenuous exercise for individual subjects, as well as their adaptive responses after HIT. Unlike conventional MS-based data workflows, a single dilution trend filter of a pooled QC sample is used as a primary screen to certify reproducible plasma-derived molecular features while rejecting signal artifacts based on signal pattern recognition. This allows for targeted analysis of authentic metabolites from individual samples with high data fidelity that avoids data-overfitting and false discoveries when using complementary multivariate data analysis methods. Unambiguous identification and quantification of lead plasma markers (HyX) and their associated metabolic pathway (purine degradation) provides deeper insight into exercise responsiveness that differ widely between subjects.

isobaric/isomeric interferences. Data pre-processing is also simplified since peak picking targets authentic yet reproducible features pre-screened from a pooled sample. Furthermore, confirmation of putative biomarkers of interest after statistical analysis (stage 4) is greatly simplified since individual responses to strenuous exercise and adaptive training responses are derived from a single extracted ion electropherogram (**Figure 4.6**) without complicated data conversion, pre-processing and time-alignment using bioinformatic software.[78] Alternative injection configurations can also be used in MSI-CE-MS to best suit the experimental design, such as a spike/recovery study using a standard addition calibration to evaluate method accuracy.[32] This is critical for reliable quantification of biologically significant metabolites (stage 5) that serve as markers of energetic stress as related to the adenine degradation pathway, such as HyX.

4.6 Concluding Remarks

MSI-CE-MS offers a high throughput platform in metabolomics for biomarker discovery that is achieved using a single TOF-MS system without hardware modifications or expensive infrastructure investments. In this work, several metabolites were identified as plasma markers of strenuous exercise and training status following a 6-week HIT intervention. Positive metabolic adaptations to HIT was realized despite differences in baseline fitness, adiposity and dietary status of participants under free-living conditions. Notably, exercise-induced upregulation in C0 status was indicative of enhanced muscle oxidative capacity, whereas lower plasma thiol redox status reflected greater intra-cellular antioxidant capacity. Also, adaptive changes in C2 and HyX responses post-exercise corresponded to lower energetic stress and greater acetylation capacity for trained subjects. Due to the small size of the cohort in this pilot study, further studies are warranted to investigate the relationship between adaptive metabolic responses measured in plasma with underlying changes in muscle tissue for overweight/obese subjects. Although biomarker validation was not tested on an independent cohort or different genders, the cross-over study design, rigorous data screen with quality assurance, complementary statistical methods with cross-validation along with quantitative measurements of plasma metabolites consistent with literature findings were unique strengths of the present study. However, different buffer/ionization conditions in MSI-CE-MS, as well as orthogonal instrumental methods are needed to further expand coverage of the human plasma metabolome.

Validation of lifestyle interventions that enhance the cardiorespiratory fitness and glucose tolerance responses of sedentary/obese subjects at risk for type 2 diabetes will also be examined. Cardiorespiratory fitness is a strong predictor of cardiovascular disease risk that is readily modifiable by low-volume/time-efficient exercise regimes, such as HIT. Due to large between-subject variations in treatment responses, novel markers that predict exercise responsiveness among metabolically healthy and unhealthy obese individuals[79,80] are needed to translate customized lifestyle modifications for personalized health outside a clinical setting.

4.7 Acknowledgements and Funding

We thank M. Percival for assistance during with the training phase of the project. P.B.M. and M.J.G. wish to acknowledge funding support from National Science and Engineering Research Council of Canada. P.B.M. also thanks the Canada Foundation for Innovation and the Ontario Research Fund for generous support. N.L.K. acknowledges receipt of a Queen Elizabeth II Graduate Scholarship in Science and Technology. J.B.G. held a NSERC Vanier Scholarship.

4.8 References

(1) Olshansky, S.; Passaro, D.; Hershow, R.; Layden, J.; Carnes, B.; Brody, J.; Hayflick, L.; Butler, R.; Allison, D.; Ludwig, D. *N. Engl. J. Med.* **2005**, *352*, 1138–1145.

(2) Chang, S.; Pollack, L.; Colditz, G. *PLoS One* **2013**, *8*, e66550.

(3) Dandona, P; Aljada, A.; Chaudhuri, A.; Mohanty, P.; Garg, R. *Circulation* **2005**, *111*, 1448–1454.

(4) Wild, S.; Byrne, C. *BMJ* **2006**, *333*, 1009–1011.

(5) Bruunsgaard, H. *J. Leukoc. Biol.* **2005**, *78*, 819–835.

(6) Chaput, JP; Doucet, E.; Tremblay, A. *Obes. Rev.* **2012**, *13*, 681–691.

(7) Gibala, M. J.; Little, J. P.; MacDonald, M. J.; Hawley, J. A. *J. Physiol.* **2012**, *590*, 1077–84.

(8) Hawley, J. A.; Gibala, M. J. *Diabetologia* **2012**, *55*, 535–9.

(9) Timmons, J. A.; Knudsen, S.; Rankinen, T.; Koch, L. G.; Sarzynski, M.; Jensen, T.; Keller, P.; Scheele, C.; Vollaard, N. B. J.; Nielsen, S.; Akerström, T.; MacDougald, O. A.; Jansson, E.; Greenhaff, P. L.; Tarnopolsky, M. A.; van Loon, L. J. C.; Pedersen, B. K.; Sundberg, C. J.; Wahlestedt, C.; Britton, S. L.; Bouchard, C. *J. Appl. Physiol.* **2010**, *108*, 1487–96.

(10) Boule, N. G.; Weisnagel, S. J.; Lakka, T. A.; Tremblay, A.; Bergman, R. N.; Rankinen, T.; Leon, A. S.; Skinner, S. J.; Wilmore, J. H.; Rao, D. C.; Bouchard, C. *Diabetes Care* **2005**, *28*, 108–114.

(11) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263.

(12) Langley, R.; Tsalik, E.; van Velkinburgh, J.; Glickman, S.; Rice, B.; Wang, C.; Chen, B.; Carin, L.; Suarez, A.; Mohney, R.; Freeman, D.; Wang, M.; You, J.; Wulff, J.; Thompson, J.; Moseley, M.; Reisinger, S.; Edmonds, B.; Grinnell, B.; Nelson, D.; Dinwiddie, D.; Miller, N.; Saunders, C.; Soden, S.; Rogers, A.; Gazourian, L.; Fredenburgh, L.; Massaro, A.; Baron, R.; Choi, A.; Corey, G.; Ginsburg, G.; Cairns, C.; Otero, R.; Fowler, V.; Rivers, E.; Woods, C.; Kingsmore, S. *Sci. Transl. Med.* **2013**, *5*, 195.

(13) Saito, N.; Robert, M.; Kitamura, S.; Baran, R.; Soga, T.; Mori, H.; Nishioka, T.; Tomita, M. *J. Proteome Res.* **2006**, *5*, 1979.

(14) Quanbeck, SM; Brachova, L.; Campbell, A.; Guan, X.; Perera, A.; He, K.; Rhee, S.; Bais, P.; Dickerson, J.; Dixon, P.; Wohlgemuth, G.; Fiehn, O.; Barkan, L.; Lange, I.; Lange, B.; Lee, I.; Cortes, D.; Salazar, C.; Shuman, J.; Shulaev, V.; Huhman, D.; Sumner, L.; Roth, M.; Welti, R.; Ilarslan, H.; Wurtele, E.; Nikolau, B. *Front. Plant. Sci.* **2012**, *3*, 15.

(15) Chen, Q.; Park, H.-C.; Goligorsky, M. S.; Fischer, S. M.; Gross, S. S. *PLoS One* **2012**, *7*, e37149.

(16) Southam, A.; Lange, A.; Hines, A.; Hill, E.; Katsu, Y.; Iguchi, T.; Tyler, C.; Viant, M. *Environ. Sci. Technol.* **2011**, *45*, 3759–3767.

(17) Martin, FP; Montoliu, I.; Nagy, K.; Moco, S.; Collino, S.; Guy, P.; Redeuil, K.; Scherer, M.; Rezzi, S.; Kochhar, S. *J. Proteome Res.* **2012**, *11*, 6252–6263.

(18) Westerhuis, J.; van Velzen, E.; Hoefsloot, H.; Smilde, A. *Metabolomics* **2010**, *6*, 119–128.

(19) Lewis, G. D.; Farrell, L.; Wood, M. J.; Martinovic, M.; Arany, Z.; Rowe, G. C.; Souza, A.; Cheng, S.; McCabe, E. L.; Yang, E.; Shi, X.; Deo, R.; Roth, F. P.; Asnani, A.; Rhee, E. P.; Systrom, D. M.; Semigran, M. J.; Vasan, R. S.; Carr, S. a; Wang, T. J.; Sabatine, M. S.; Clish, C. B.; Gerszten, R. E. *Sci. Transl. Med.* **2010**, *2*, 33ra37.

(20) Nieman, D. C.; Gillitt, N. D.; Henson, D. A.; Sha, W.; Shanely, R. A.; Knab, A. M.; Cialdella-Kam, L.; Jin, F. *PLoS One* **2012**, *7*, e37479.

(21) Lee, R.; West, D.; Phillips, S.; Britz-McKibbin, P. *Anal. Chem.* **2010**, *82*, 2959.

(22) Chorell, E.; Moritz, T.; Branth, S.; Antti, H.; Svensson, M. B. *J. Proteome Res.* **2009**, *8*, 2966–2977.

(23) Brugnara, L.; Vinaixa, M.; Murillo, S.; Samino, S.; Rodriguez, M. A.; Beltran, A.; Lerin, C.; Davison, G.; Correig, X.; Novials, A. *PLoS One* **2012**, *7*, e40600.

(24) Rodríguez, D. A.; Alcarraz-Vizán, G.; Díaz-Moralli, S.; Reed, M.; Gómez, F. P.; Falciani, F.; Günther, U.; Roca, J.; Cascante, M. *Metabolomics* **2012**, *8*, 508–516.

(25) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2013**, *8*, 17–32.

(26) Ishii, N.; Nakahigashi, K.; Baba, T.; Robert, M.; Soga, T.; Kanai, A.; Hirasawa, T.; Naba, M.; Hirai, K.; Hoque, A.; Ho, P.; Kakazu, Y.; Sugawara, K.; Igarashi, S.; Harada, S.; Masuda, T.; Sugiyama, N.; Togashi, T.; Hasegawa, M.; Takai, Y.; Yugi, K.; Arakawa, K.; Iwata, N.; Toya, Y.; Nakayama, Y.; Nishioka, T.; Shimizu, K.; Mori, H.; Tomita, M. *Science* **2007**, *316*, 593.

(27) Han, J.; Danell, R. M.; Patel, J. R.; Gumerov, D. R.; Scarlett, C. O.; Speir, J. P.; Parker, C. E.; Rusyn, I.; Zeisel, S.; Borchers, C. H. *Metabolomics* **2008**, *4*, 128–140.

(28) Fuhrer, T.; Heer, D.; Begemann, B.; Zamboni, N. *Anal. Chem.* **2011**, *83*, 7074–7080.

(29) Kirwan, J. A.; Broadhurst, D. I.; Davidson, R. L.; Viant, M. R. *Anal. Bioanal. Chem.* **2013**, *405*, 5147–5157.

(30) Rath, C. M.; Yang, J. Y.; Alexandrov, T.; Dorrestein, P. C. *J. Am. Soc. Mass. Spectrom.* **2013**, *24*, 1167–1176.

(31) Kuehnbaum, N. L.; Britz-McKibbin, P. *Chem. Rev.* **2013**, *113*, 2437–2468.

(32) Kuehnbaum, N. L.; Kormendi, A.; Britz-McKibbin, P. *Anal. Chem.* **2013**, *85*, 10664–9.

(33) Jankevics, A.; Merlo, M.; de Vries, M.; Vonk, R.; Takano, E.; Breitling, R. *Metabolomics* **2012**, *8*, 29–36.

(34) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84*, 5035–5039.

(35) Gillen, J. B.; Percival, M. E.; Ludzki, A.; Tarnopolsky, M. A.; Gibala, M. J. *Obesity* **2013**, *21*, 2249–2255.

(36) Kamleh, M. A.; Ebbels, T. M.; Spagou, K.; Masson, P.; Want, E. J. *Anal. Chem.* **2012**, *84*, 2670–2677.

(37) Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S. *Nucleic Acids Res.* **2012**, *40*, W127–W133.

(38) Xia, J.; Broadhurst, D.; Wilson, M.; Wishart, D. *Metabolomics* **2013**, *9*, 280–299.

(39) Robergs, R. A.; Ghiasvand, F.; Parker, D. *Am. J. Physiol. Regul. Intergr. Comp. Physiol.* **2004**, *287*, R502–R516.

(40) Gomez-Cabrera, M. C.; Domenech, E.; Vina, J. *Free Radic. Bio. Med.* **2008**, *44*, 126–31.

(41) Zwetsloot, K. A.; John, C. S.; Lawrence, M. M.; Battista, R. A.; Shanely, R. A. *J Inflamm Res* **2014**, *7*, 9–17.

(42) Lee, D.; Artero, E.; Sui, X.; Blair, S. *J Psychopharmacol.* **2010**, *24*, 27–35.

(43) Hamer, M.; O'Donovan, G. *Curr Opin Lipidol.* **2010**, *21*, 1–7.

(44) Finlayson, G.; Caudwell, P.; Gibbons, C.; Hopkins, M.; King, N.; Blundell, J. *J. Obesity* **2011**, *2011*, Article ID 615624.

(45) Bassett, D. J.; Howley, E. *Med Sci Sport. Exerc* **2000**, *32*, 70–84.

(46) Van Proeyen, K.; Szlufcik, K.; Nielens, H.; Pelgrim, K.; Deldicque, L.; Hesselink, M.; Van Veldhoven, P. P.; Hespel, P. *J. Physiol.* **2010**, *588*, 4289–4302.

(47) Crews, B.; Wikoff, W.; Patti, G.; Woo, H.; Kalisiak, E.; Heideker, J.; Siuzdak, G. *Anal. Chem.* **2009**, *81*, 8538.

(48) Bouchard, C.; Sarzynski, M.; Rice, T.; Kraus, W.; Church, T.; Sung, Y.; Rao, D.; Rankinen, T. *J Appl Physiol* **2011**, *110*, 1160–1170.

(49) Stephens, F. B.; Constantin-Teodosiu, D.; Greenhaff, P. L. *J Physiol* **2007**, *581*, 431–44.

(50) Brass, E. *Ann N Y Acad Sci* **2004**, *1033*, 67–78.

(51) Karlic, H.; Lohninger, A. *Nutrition* **2004**, *20*, 709–715.

(52) Nasser, M.; Javaheri, H.; Fedorowicz, Z.; Noorani, Z. *Cochrane Database Syst. Rev.* **2012**, *2*, Article No. CD006659.

(53) Wall, B.; Stephens, F.; Constantin-Teodosiu, D.; Marimuthu, K.; Macdonald, I.; Greenhaff, P. *J. Physiol.* **2011**, *589*, 963–973.

(54) Ho, J.; Kraemer, W.; Volek JS, Fragala MS, Thomas, G.; Dunn-Lewis, C.; Coday, M.; Häkkinen, K.; Maresh, C. *Metabolism* **2010**, *59*, 1190–1199.

(55) Servillo, L.; Giovane, A.; Cautela, D.; Castaldo, D.; Balestrieri, M. L. *PLoS One* **2014**, *9*, e84589.

(56) Tunstall, R.; Mehan, K.; Wadley, G.; Collier, G.; Bonen, A.; Hargreaves, M.; Cameron-Smith, D. *Am J Physiol Endocrinol Metab.* **2002**, *283*, E55–72.

(57) Little, J.; Safdar, A.; Wilkin, G.; Tarnopolsky, M.; Gibala, M. *J Physiol* **2010**, *588*, 1011–1022.

(58) Nüesch, R.; Rossetto, M.; Martina, B. *Drugs Exp Clin Res.* **1999**, *25*, 167–171.

(59) Stephens, F.; Constantin-Teodosiu, D.; Laithwaite, D.; Simpson, E.; Greenhaff, P. *J Clin Endocrinol Metab* **2006**, *91*, 5013–5018.

(60) Yoshiba-Suzuki, S.; Sagara, J.; Bannai, S.; Makino, N. *J Biochem.* **2011**, *150*.

(61) Jones, DP; Mody, V. J.; Carlson, J.; Lynn, M.; Sternberg, P. J. *Free Radic. Bio. Med.* **2002**, *23*, 1290–1300.

(62) Sen, C. *Mol. Cell. Biochem.* **1999**, *196*, 31–42.

(63) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(64) Sahlin, K.; Tonkonogi, M.; Söderlund, K. *Eur J Appl Physiol Occup Physiol* **1999**, *80*, 417–422.

(65) Hancock, C.; Brault, J.; Terjung, R. *J Physiol Pharmacol* **2006**, *57*, 17–29.

(66) Tullson, P.; Bangsbo, J.; Hellsten, Y.; Richter, E. *J Appl Physiol* **1995**, *78*, 146–52.

(67) Kuehnbaum, N. L.; Britz-McKibbin, P. *Anal. Chem.* **2011**, *83*, 8063.

(68) Saiki, S.; Sato, T.; Kohzuki, M.; Kamimoto, M.; Yosida, T. *Metabolism* **2001**, *50*, 627–630.

(69) Youle, M. *CNS Drugs* **2007**, *21*, 25–30.

(70) Nasca, C.; Xenos, D.; Barone, Y.; Caruso, A.; Scaccianoce, S.; Matrisciano, F.; Battaglia, G.; Mathé, A.; Pittaluga, A.; Lionetto, L.; Simmaco, M.; Nicoletti, F. *Proc Natl Acad Sci USA* **2013**, *110*, 4804–4809.

(71) Noland, R. C.; Koves, T. R.; Seiler, S. E.; Lum, H.; Lust, R. M.; Ilkayeva, O.; Stevens, R. D.; Hegardt, F. G.; Muoio, D. M. *J Biol Chem* **2009**, *284*, 22840–52.

(72) Seiler, S.; Martin, O.; Noland, R.; Slentz, D.; Debalsi, K.; Ilkayeva, O.; An, J.; Newgard, C.; Koves, T.; Muoio, D. *J Lipid Res* **2014**, *55*, 635–44.

(73) Violante, S.; Ijlst, L.; Ruiter, J.; Koster, J.; van Lenthe, H.; Duran, M.; de Almeida, I.; Wanders, R.; Houten, S.; FV, V. *Biochim Biophys Acta* **2013**, *1832*, 773–779.

(74) Kumari, S.; Stevens, D.; Kind, T.; Denkert, C.; Fiehn, O. *Anal. Chem.* **2011**, *83*, 5895–5902.

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

(76) Sugimoto, M.; Hirayama, A.; Robert, M.; Abe, S.; Soga, T.; Tomita, M. *Electrophoresis* **2010**, *31*, 2311–2318.

(77) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. *BMC Bioinformatics* **2010**, *22*, 148.

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

(79) Karelis, A. D. *Lancet* **2008**, *372*, 1281–3.

(80) Perrault, M.; Zulyniak, M. A.; Badoud, F.; Stephenson, S.; Badawi, A.; Buchholz, A.; Mutch, D. M. *PLoS One* **2014**, *9*, e88539.

4.9 Appendix A

Table A4.1. Summary of 54 cationic metabolites after dilution trend filtering of a pooled plasma sample when using MSI-CE-MS, which were consistently detected (average CV is 13% for 9 QCs) in all individual subjects across all time intervals irrespective of training/feeding status. Each metabolite is annotated by its characteristic *m/z*:RMT as a paired variable, whereas identification was confirmed by spiking plasma samples with authentic chemical standards.

Annotated Ion m/z:RMT	***Molecular Formula***	***Plasma Metabolite***	***Precision (QCs) CV% (n=9)***
76.0393:0.691	C2 H5 N O2	Glycine (Gly)	6.6
84.0444:0.905	C4 H5 N O	Fragment of Gln	7.1
84.0808:0.571	C5 H9 N	Fragment of Lys	14
86.0964:0.832	C5 H11 N	Fragment of Ile	17
86.0964:0.843	C5 H11 N	Fragment of Leu	9.0
90.0550:0.743	C3 H7 N O2	Alanine (Ala)	7.6
104.0706:0.789	C4 H9 N O2	--	7.6
106.0499:0.836	C3 H7 N O3	Serine (Ser)	4.2
114.0662:0.585	C4 H7 N3 O	Creatinine (Cretn)	8.4
116.0706:0.864	C5 H9 N O2	Proline (Pro)	8.9
118.0863:0.818	C5 H11 N O2	Valine (Val)	20
118.0863:0.954	C5 H11 N O2	Betaine (Bet)	13
119.0161:0.970	C4 H6 O2 S	--	9.6
120.0655:0.868	C4 H9 N O3	Threonine (Thr)	20
130.0499:0.908	C5 H7 N O3	Fragment of Gln	1.3
130.0863:0.562	C4 H9 N O3	Fragment of Lys	11
132.0768:0.723	C4 H9 N3 O2	Creatine (Cret)	8.9
132.1019:0.832	C6 H13 N O2	Isoleucine (Ile)	7.2
132.1019:0.843	C6 H13 N O2	Leucine (Leu)	3.7
133.0608:0.882	C4 H8 N2 O3	Asparagine (Asn)	4.9
133.0972:0.567	C5 H12 N2 O2	Ornithine (Orn)	5.9
134.0448:0.980	C4 H7 N O4	Aspartic acid (Asp)	7.5
134.0448:1.343	C4 H7 N O4	--	33
136.0757:0.951	C8 H9 O	Fragment of Tyr	5.9
137.0457:1.133	C5 H4 N4 O	Hypoxanthine (HyX)	8.0
137.0709:0.598	C7 H8 N2 O	Fragment of Cretn	15
144.1019:0.970	C7 H13 N O2	--	19
147.0764:0.905	C5 H10 N2 O3	Glutamine (Gln)	8.4
147.1128:0.562	C6 H14 N2 O2	Lysine (Lys)	8.5
148.0604:0.920	C5 H9 N O4	Glutamic acid (Glu)	6.6
150.0583:0.889	C5 H11 N O2 S	Methionine (Met)	4.7
156.0768:0.606	C6 H9 N3 O2	Histidine (His)	6.7
162.1125:0.691	C7 H15 N O3	Carnitine (C0)	18
163.1441:0.815	C7 H18 N2 O2	--	20
166.0863:0.918	C9 H11 N O2	Phenylalanine (Phe)	7.3
173.0921:0.745	C7 H12 N2 O3	--	13

Annotated Ion m/z:RMT	***Molecular Formula***	***Plasma Metabolite***	***Precision (QCs) CV% (n=9)***
175.1190:0.579	C6 H14 N4 O2	Arginine (Arg)	4.9
176.1030:0.930	C6 H13 N3 O3	Citrulline (Cit)	6.0
177.1603: 0.842	C8 H20 N2 O2	Fragment of Leu	30
182.0812:0.951	C9 H11 N O3	Tyrosine (Tyr)	2.7
192.1594:0.764	C9 H21 N O3	--	9.7
204.1230:0.735	C9 H17 N O4	Acetylcarnitine (C2)	15
205.0972:0.917	C11 H12 N2 O2	Tryptophan (Trp)	5.9
209.0921:0.872	C10 H12 N2 O3	Kynurenine (Kyn)	6.9
214.0571:0.930	C6 H7 N5 O4	Adduct of Cit	32
218.1387:0.756	C10 H19 N O4	Propionylcarnitine (C3)	31
241.0311:0.935	C6 H12 N2 O4 S2	Cystine (Cys-SS)	4.9
276.1199:1.122	C10 H17 N3 O6	--	10
286.2013:0.840	C15 H27 N O4	--	30
298.0526:0.788	C18 H15 N3 O5 S2	CysGly-Cys-SS (mixed disulfide)	9.5
337.1102:0.905	C9 H16 N6 O8	Adduct of Gln	42
427.0952:0.997	C13 H22 N4 O8 S2	GSH-Cys-SS (mixed disulfide)	57
477.1603:0.905	C18 H16 N14 O3	Adduct of Gln	30

Chapter V

Personalized Metabolomics for Assessment of Differential Glucose Tolerance Responses to High-Intensity Interval Training: Who Benefits from Exercise and Why?

V. Personalized Metabolomics for Assessment of Differential Glucose Tolerance Responses to High-Intensity Interval Training: Who Benefits from Exercise and Why?

5.1 Abstract

High intensity interval training (HIT) offers a practical approach for enhancing cardiorespiratory fitness, however its role in improving glucose regulation among sedentary normoglycemic women remains unclear. Herein, multi-segment injection capillary electrophoresis-mass spectrometry is used as a rapid screening tool in metabolomics to assess adaptive responses of overweight/obese women (*BMI* > 25, *n*=11) to an oral glucose tolerance test (OGTT) following a 6-week HIT intervention. Various statistical methods were used to classify plasma metabolic signatures associated with postprandial glucose and/or training status when using a cross-over study design. Branched-chain/aromatic amino acids and other intermediates of urea cycle and carnitine metabolism decreased over time after glucose loading, whereas adaptive changes to plasma thiol redox and ornithine status reflected increased antioxidant and detoxification capacity for trained subjects. A multi-linear regression model was also developed to predict changes in glucose tolerance based on a panel of plasma metabolites measured for subjects in their naïve/untrained state with good overall accuracy and robustness. Modest improvements in glucose tolerance were achieved with HIT, however sub-sets of women were found to exhibit negligible changes or in some cases adverse responses despite gains in aerobic endurance capacity. Since treatment outcomes to physical activity are variable between-subjects, prognostic markers are needed for designing lifestyle modifications that maximize the salutary benefits of exercise for personalized health and diabetes prevention.

5.2 Introduction

Physical inactivity is a modifiable risk factor associated with an alarming increase in obesity and chronic metabolic disorders worldwide, including type 2 diabetes.[1] Regular physical activity also improves mental health and cognitive function for healthy aging.[2] Thus, exercise represents a cost-effective lifestyle intervention for disease prevention while reducing the socioeconomic burden of late-stage treatment of diabetes complications.[3] For instance, a 10 year follow-up of intensive lifestyle modifications consisting of moderate sub-maximal exercise training lowered the incidence of diabetes by 34% in high risk adults that was

about two-fold more effective than metformin therapy.[4] Sustained long-term benefits have also been shown to reduce diabetes progression over 13 years, where greater risk reduction was associated with adherence to lifestyle changes during the intervention period.[5] However, lack of time remains a major barrier of exercise training to public health that satisfies recommended minimal guidelines with less than 10% of children reported to be involved in 60 min of moderate intensity exercise daily.[6]

High-intensity interval training (HIT) represents a practical exercise protocol that is characterized by short yet intense bursts of activity interspersed with periods of rest or low-intensity. HIT is equivalent and in many cases superior to sub-maximal continuous training while requiring less time commitments and lower total exercise volume.[7] The type of exercise, number and duration of intervals, and optimum intensity level can be adapted to meet the desired therapeutic goals and/or target populations.[8,9] In adults with coronary artery disease,[10–13] metabolic syndrome,[14] and obesity,[15,16] HIT enhances cardiorespiratory fitness and other measures of cardiovascular health, such as resting blood pressure, endothelial function, and left ventricular morphology. Extremely short duration HIT protocols have also been used to improve insulin sensitivity in sedentary populations,[17,18] whereas two-week HIT interventions reduce hyperglycemia[19] and postprandial glucose responses of patients with type 2 diabetes.[20] Although HIT can reduce abdominal adiposity in young women,[21] the insulin-sensitizing benefits associated with exercise training are independent from weight loss or body fat reduction.[22,23] However, adaptive physiological changes to HIT remain poorly understood with treatment responses variable between-subjects.[24] Exercise training can also elicit unanticipated adverse health outcomes in susceptible individuals that is difficult to predict based on body composition.[25]

Metabolomics offers a systemic approach for elucidating complex mechanisms associated with nutritional and/or exercise interventions relevant to human health.[26,27] Challenge tests in particular provide deeper insight into changes in human metabolism that is often not captured from baseline/single-point measurements alone.[28] For instance, oral glucose tolerance tests (OGTT) evaluate subtle perturbations in metabolism and diabetes risk based on dynamic responses after high-dose oral glucose loading.[29–31] Glucose, insulin, and glycated haemoglobin A1c are widely used indicators of glucose homeostasis for diagnosis of diabetes or pre-diabetes; however, other classes of metabolites are implicated in impaired glucose tolerance as related to four major axes of insulin action,

including proteolysis, lipolysis, ketogenesis and glycolysis.[28,29,32–34] To the best of our knowledge, this is the first metabolomics study to examine adaptive changes in glucose tolerance after exercise training. A cohort of overweight/obese women were recruited to participate in a 6-week HIT intervention, and standardized OGTTs were performed for each subject prior to and after exercise training using a cross-over study design. Multi-segment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) was used as a high-throughput platform in metabolomics[35] for classification of plasma metabolites modulated by glucose loading and/or training status. Multiplexed separations based on MSI-CE-MS also provide an accelerated workflow for biomarker discovery in metabolomics without complicated data pre-processing.[36] The main objective of this work was to identify plasma metabolite signatures that predict glucose tolerance changes after HIT on an individual level despite the wide disparity in baseline aerobic fitness and adiposity.

5.3 Materials and Methods

5.3.1 Subject Cohort and Study Design

Eleven overweight/obese yet healthy women (age: 18-45 years, *BMI*: 25-35 kg/m^2) participated in the 6-week supervised exercise intervention as previously described. The subjects represent a subset of a group of 16 individuals who took part in a study that evaluated the impact of pre-exercise nutritional state on adaptations to HIT.[23] In this work, dynamic metabolomic studies were performed using blood samples collected at all time intervals during two standardized OGTTs for 11 of the 16 subjects from the original trial. Participants were recruited through poster advertisement and the study protocol was approved by the McMaster University Research Ethics Board. Subjects were classified as "sedentary" based on self-reporting of physical activity $\leq$ 2 sessions/week of exercise with a duration $\leq$ 30 min. Subjects were classified as "sedentary" based on self-reporting of physical activity $\leq$ 2 sessions/week of exercise with a duration $\leq$ 30 min. Participants completed a general health questionnaire to assess overall health and verify that they did not meet any of the exclusion criteria, including pre-diabetes, diabetes, hypertension or cardiovascular diseases.

5.3.2 Body Composition, Fitness Testing and Exercise Intervention

Body composition analysis to determine percentages of bone, fat, and lean muscle tissue were measured by dual x-ray absorptiometry (DXA) scanning (Lunar Prodigy Advance, Madison, WI). Maximal oxygen uptake (VO_{2max}), maximal workload (W_{max}), and average heart rate (*HR*) during exercise trials were determined for each participant using standardized protocols. All of these parameters were measured prior to the start of the exercise training, as well as following the 6-week HIT intervention for each subject in their naïve and trained states, respectively as previously described.[23] Briefly, exercise was performed on a stationary cycle ergometer (LifeCycle C1, Life Fitness, Schiller Park, IL) while maintaining a pedal cadence of 80-100 rpm. The exercise intervention consisted of 18 HIT sessions, performed three times a week (with a 1-2 day recovery in between) over a total of 6 weeks. The HIT protocol involved intermittent cycling and consisted of 10 cycles of 60 s duration at an individualized workload designed to elicit 90% maximal *HR*. Recovery intervals that involved light cycling at 50 W were performed for 1 min between high intensity bursts for a total exercise time of 25 min for each HIT trial, which also included a 3 min warm-up and a 2 min cool down period. Standardized breakfasts were only provided for days when the subjects performed HIT training. All other meals as well as breakfasts on non-training days were under free-living conditions. However, subjects were instructed to maintain their pre-training dietary habits throughout the intervention, which was confirmed by dietary records.[23]

5.3.3 Oral Glucose Tolerance Tests

Subjects underwent an OGTT prior to the start of exercise training, as well as 72 h after the completion of the 6-week HIT intervention. Subjects fasted for ≥ 10 h overnight prior to the morning test. An indwelling catheter was inserted into the forearm vein and a resting blood sample was collected and stored on ice before processing. Participants ingested a Glucodex drink containing 75 g glucose and five more blood samples were collected at 0, 20, 30, 60, 90 and 120 min after ingestion. Fasting glucose (FG, 0 min) and 2-h post-glucose loading (2hPG, 120 min) were two key time intervals reflecting glucose homeostasis and glucose tolerance responses, respectively. Vacutainers (6 mL) with EDTA (10.8 mg) as anticoagulant were used for blood collection (Becton Dickinson, Franklin Lakes, NJ, USA). Immediately after collection, blood samples were placed on ice and subsequently centrifuged at 2,500 *g* at 4°C for 5 min to fractionate plasma from

blood cells, which were then subsequently frozen at -80°C until analysis. Frozen plasma aliquots were thawed slowly on ice, then vortexed for 30 s to mix gently. Plasma was diluted 4-fold with 200 mM ammonium acetate (pH 5) with 25 μM 3-chloro-L-tyrosine (Cl-Tyr) as internal standard. Plasma proteins were then filtered by ultrafiltration using a 3 kDa MWCO Nanosep centrifugal device (Pall Life Sciences, Washington, NY, USA) at 13,000 *g* for 15 min. A 20 μL aliquot of the diluted plasma filtrate was used for analysis by MSI-CE-MS.

5.3.4 Chemicals and Reagents

Ammonium acetate, acetic acid, formic acid, Cl-Tyr and all other metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO, USA). A 10 mM stock solution of Cl-Tyr was prepared in water and stored at 4°C. HPLC-grade acetonitrile (Honeywell, Muskegon, MI, USA) and methanol (Caledon, Georgetown, ON, Canada) were used for preparation of background electrolyte (BGE) and sheath liquid, respectively. All aqueous buffers and stock solutions were prepared with deionized water purified using a Thermo Scientific Barnstead EasyPure II LF ultrapure water system (Cole Parmer, Vernon Hills, IL, USA).

5.3.5 Instrumentation and Method Configuration

All CE-TOF-MS experiments were performed using an Agilent G7100A CE system (Mississauga, ON, Canada) interfaced with a coaxial sheath liquid Jet Stream electrospray ion source with heated gas to an Agilent 6230 time-of-flight mass spectrometer (TOF-MS). Nitrogen gas was used as the nebulizer gas in the ESI source and as the drying gas in the MS. An uncoated fused silica capillary (Polymicro Technologies, AZ, USA) with 50 μm ID and 110 cm length maintained at 25°C was used for all experiments. The background electrolyte (BGE) was 1 M formic acid containing 15% *v* acetonitrile (pH 1.8). The applied voltage was 30 kV with a total analysis time of 35 min. In MSI-CE-MS, samples and spacers containing BGE were alternately injected hydrodynamically at 100 mbar as described previously,[35] namely (1) 5 s sample, 40 s spacer; (2) 5 s sample, 40 s spacer; (3) 5 s sample, 40 s spacer; (4) 5 s sample, 40 s spacer; (5) 5 s sample, 40 s spacer; (6) 5 s sample, 40 s spacer; (7) 5 s sample, 5 s spacer. The total injection time for sample and BGE spacer segments was 280 s that is equivalent to about 30% of the total capillary length. Plasma filtrate samples were diluted 4-fold in 200 mM NH_4Ac (pH 5) with 25 μM Cl-Tyr as internal standard, which was used for determination of relative migration time (RMT) and relative

peak response of metabolites. Between runs the capillary was flushed for 10 min with BGE. The sheath liquid was 60:40 MeOH:H_2O containing 0.1% formic acid at a flow rate of 10 µL/min via 100:1 splitter. Purine and hexakis(2,2,3,3-tetrafluoropropoxy)phosphazine (HP-921) were spiked into the sheath liquid at a concentration of 0.02% v to produce corresponding reference ions at *m/z* 121.05087 and *m/z* 922.00978 for real time internal mass correction. The TOF-MS was operated in positive-ion mode for detection of cationic/zwitter-ionic metabolites over a range of *m/z* 50-1700 at an acquisition rate of 2 Hz and acquisition time of 500 ms with fixed voltage settings for following: fragmentor = 120 V, skimmer = 65 V and Oct 1 RF = 750 V. In the ESI source the settings were V_{cap} = 2000 V, nozzle voltage = 2000 V, nebulizer gas = 10 psi, sheath gas = 3.5 L/min at 195°C, and drying gas = 8 L/min at 300°C.

5.3.6 Sample Injection Configuration and Data Workflow

A seven-segment injection format was used in MSI-CE-MS for all analyses in order to enhance sample throughput without loss of information content due to ion suppression.[35] Different injection configurations can be used in order to encode information temporally according to the experimental design, however two specific formats were used in this work. First, a dilution trend filter based on a serial injection of a pooled plasma filtrate sample as quality control (QC) was performed at different dilutions, including a blank (*i.e.*, buffer filtrate).[35] Untargeted feature picking was performed using Molecular Feature Extractor (MFE, MassHunter Qualitative Analysis, Agilent Technologies Inc.) in order to detect reproducible yet authentic ions (*i.e.*, $[M+H]^+$ or $[M+Na]^+$) with peak heights over 300 counts while excluding chemical and biochemical noise. MFE was conducted over a time range (≈ 8-30 min) of the separation while excluding the salt front and the electroosmotic flow (EOF) regions of the separation, where ion co-migration and signal suppression are prevalent. An authentic plasma metabolite from the dilution trend filter was annotated by its characteristic *m/z*:RMT provided the signal was detected with adequate precision (CV < 40%, n=3) and linearity (R^2 > 0.90) in at least three dilution levels with no signal in the blank sample. This process was applied to derive a compound list of reproducible yet authentic molecular features while filtering spurious signals and background ions that comprise the majority of ion signals detected in ESI-MS.[35] A second injection configuration was also applied in MSI-CE-MS when analyzing dynamic metabolome responses to an OGTT at six time intervals (0, 20, 30, 60, 90, 120 min) for each subject either in their naïve or trained state, including a pooled

plasma QC. Thus, the same metabolite from seven different samples migrate into the ion source over a short time interval (≈ 3-5 min) under stable ionization conditions with good quantitative performance. In this way, unambiguous assignment of plasma metabolites associated with glucose loading was realized using an accelerated metabolomics workflow since it allows for targeted analysis of authentic yet reproducible metabolites pre-screened from a dilution trend filter when processing individual plasma samples.[35]

5.3.7 Quality Assurance and Statistical Analysis

Quality assurance was achieved by inclusion of a pooled QC sample as the seventh sample segment injected during each run performed by MSI-CE-MS (n = 22). The QC serves as an internal reference sample included within every separation in order to monitor long-term instrument bias and system drift that is similar to intermittent QCs used in large-scale metabolomic workflows.[37] Moreover, plasma metabolites that were initially identified and validated in the dilution trend filter were only included into the final data matrix if they were detected in a minimum of 75% of all samples with acceptable precision (CV < 40%). This rigorous approach to a metabolomics workflow results in higher data fidelity for biomarker discovery while reducing bias caused by data over-fitting.[38] Multivariate analyses of log-transformed and auto-scaled plasma metabolome data were performed using MetaboAnalyst 2.0,[39] including Wilcoxon signed rank tests, principal component analysis (PCA), hierarchical cluster analysis (HCA) and partial least-squares-discriminant analysis (PLS-DA). In addition, two-way ANOVA with Bonferroni correction was performed to classify plasma metabolites associated with training status and postprandial glucose response over a 2 h time frame. Receiver operating characteristic (ROC) curves were also performed using ROC Curve Explorer & Tester (ROCCET)[40] after *log*-transformation of data for identification of single or ratiometric plasma metabolites associated with training status (naïve; trained) and postprandial glucose (0; 120 min). Multiple linear regression (MLR) using Excel (Microsoft Inc., Redmond, WA) was used for predicting changes in glucose tolerance (2hPG) for individual subjects after HIT training, which was based on a panel of plasma metabolites measured for naïve subjects at baseline (0 min) and 120 min during the initial OGTT. Plasma metabolites were first selected as putative candidates based on their performance in PLS-DA, Wilcoxon signed rank tests and ROC curves, whereas model refining was performed to assess the optimum number of

Table 5.1. The salutary benefits of a 6 week HIT intervention involving a cohort of healthy yet overweight/obese women. Differential responses to exercise training are reflected by changes in initial aerobic endurance capacity, as well as glucose homeostasis (FG) or oral glucose tolerance (2hPG) relative to baseline (in brackets).

Subject	*Age (yrs)*	*BMI (kg/m^2)*	*Fat (%)*	*VO_2 max (mL/kg/min)*	*Workload max (W)*	*FG (mM)*	*2hPG (mM)*
S1	24	25.5	33.8	39.3 (-0.90)	248 (+18)	3.8 (+1.1)	4.8 (+0.17)
S2	25	26.3	40.0	31.1 (+4.4)	195 (+30)	5.1 (+0.88)	5.1 (+3.20)
S3	23	25.7	31.7	26.8 (+9.2)	266 (+20)	5.0 (-0.33)	4.1 (-0.54)
S4	41	27.8	40.1	29.3 (+3.1)	205 (+41)	5.0 (-0.27)	5.5 (-0.88)
S5	23	27.4	47.1	29.8 (+11)	173 (+54)	4.2 (+0.09)	5.9 (-0.88)
S6	33	25.6	35.1	33.9 (+4.8)	228 (+17)	5.4 (-1.80)	4.2 (+0.86)
S7	41	36.1	50.1	19.4 (+3.0)	190 (+40)	5.1 (-1.00)	5.2 (-1.10)
S8	21	25.1	37.1	30.4 (+7.6)	194 (+12)	4.2 (-0.17)	4.9 (-1.20)
S9	22	27.5	46.6	33.5 (+1.1)	200 (+0)	4.0 (+0.68)	7.5 (-0.71)
S10	20	32.8	48.2	21.2 (+3.8)	166 (+20)	6.2 (-1.50)	6.3 (+0.09)
S11	23	30.0	50.5	24.1 (+5.3)	170 (+30)	4.0 (-0.28)	4.2 (+0.39)
Median	*27*	*27.4*	*40.1*	*29.8 (+4.4)*	*195 (+20)*	*5.00 (-0.27)*	*5.10 (-0.54)*

metabolites to retain in the training set by a stepwise regression process in order to maximize predictive accuracy (R^2) and robustness (Q^2) when using leave-one-out at a time cross-validation. All data processing involving electropherograms was performed using Igor Pro 5.0 (Wavemetrics Inc., Lake Oswego, OR).

5.4 Results

Eleven female participants completed all trials, including the 6-week HIT intervention together with pre-/post-OGTTs in order to evaluate adaptive changes in glucose tolerance associated with exercise training. Differential metabolomic responses to HIT were analyzed by MSI-CE-MS, which were compared with physiological outcomes measured for individual subjects as summarized in **Table 5.1**, including changes in aerobic endurance capacity and glucose homeostasis/tolerance. Due to the cross-over design, both within-subject and between-subject changes in VO_{2max}, W_{max}, FG levels and 2hPG responses were measured with median/interquartile ranges for paired subjects of (4.4 ± 3.4) mL/kg/min, (20 ± 18) W, (-0.3 ±1.1) mM and (-0.5 ± 1.2) mM, respectively. Despite overall gains in aerobic endurance capacity after HIT notably in terms of W_{max} (p = 0.0272; p = 3.00 E-5 for paired t-test), only modest improvements in glucose tolerance (p > 0.05) were evident since they are obscured by large

between-subject variations. This was primarily due to the unexpectedly poor glucose tolerance response for S2 after HIT, which when excluded from the data as an outlier ($G = 0.260 > G_{crit}$; $p = 0.05$) results in a significant change in P2hG from baseline ($p = 0.0410$ for paired t-test, n=10) of about -0.62 mM. Also, there were no correlations ($R^2 < 0.10$) associated with subject body composition (*BMI, % fat*) and aerobic fitness or glucose response changes. For example, better gains in W_{max} after HIT does not necessarily correspond to better glucose tolerance, nor is *BMI* a good indicator of changes in glucose homeostasis. Moreover, FG outcomes did not correlate with 2hPG responses with some subjects having significantly lower FG levels, but with unchanged (*e.g.,* S10) or even reduced glucose tolerance (*e.g.,* S6) after exercise training. As noted previously, one subject (*e.g.,* S2) in particular was found to have negative responses to HIT in terms of adaptive changes in both FG and P2hG relative to baseline measurements before training. This is in agreement with recent studies that have reported variable responses to exercise training, including susceptible individuals denoted as "adverse responders", where the observed health effect is opposite to that desired.[25,41]

5.4.1 Individual Metabolomic Responses to OGTT

MSI-CE-MS was used to explore metabolomic responses to HIT that underlie subtle or confounding treatment effects (*i.e.,* glucose tolerance) observed among overweight yet non-diabetic women. **Figure 5.1** depicts the overall study design, including representative OGTT curves for a "positive responder" (*e.g.,* S7) highlighting significant decreases in both FG and 2hPG after HIT intervention. The seven-sample segment injection format used in MSI-CE-MS is convenient as it allows for unambiguous evaluation of metabolomic responses to postprandial glucose for each subject over six time intervals (relative to baseline) together with a pooled QC in a single run. Hence, MSI-CE-MS enables high-throughput metabolomic profiling (< 5 min/sample) of a large set of plasma filtrates (n = 11 x 6 x 2 = 132), including a series of internal QCs (n = 22) that are incorporated with each separation in order to evaluate system stability. A dilution trend filter was first used as an untargeted primary screen to filter out chemical and biochemical noise using a pooled specimen from the study while annotating authentic plasma molecular features with acceptable reproducibility and linearity.[35] In this way, conventional targeted profiling can be performed on individual plasma specimens when using a validated list of known or unknown metabolites denoted by their

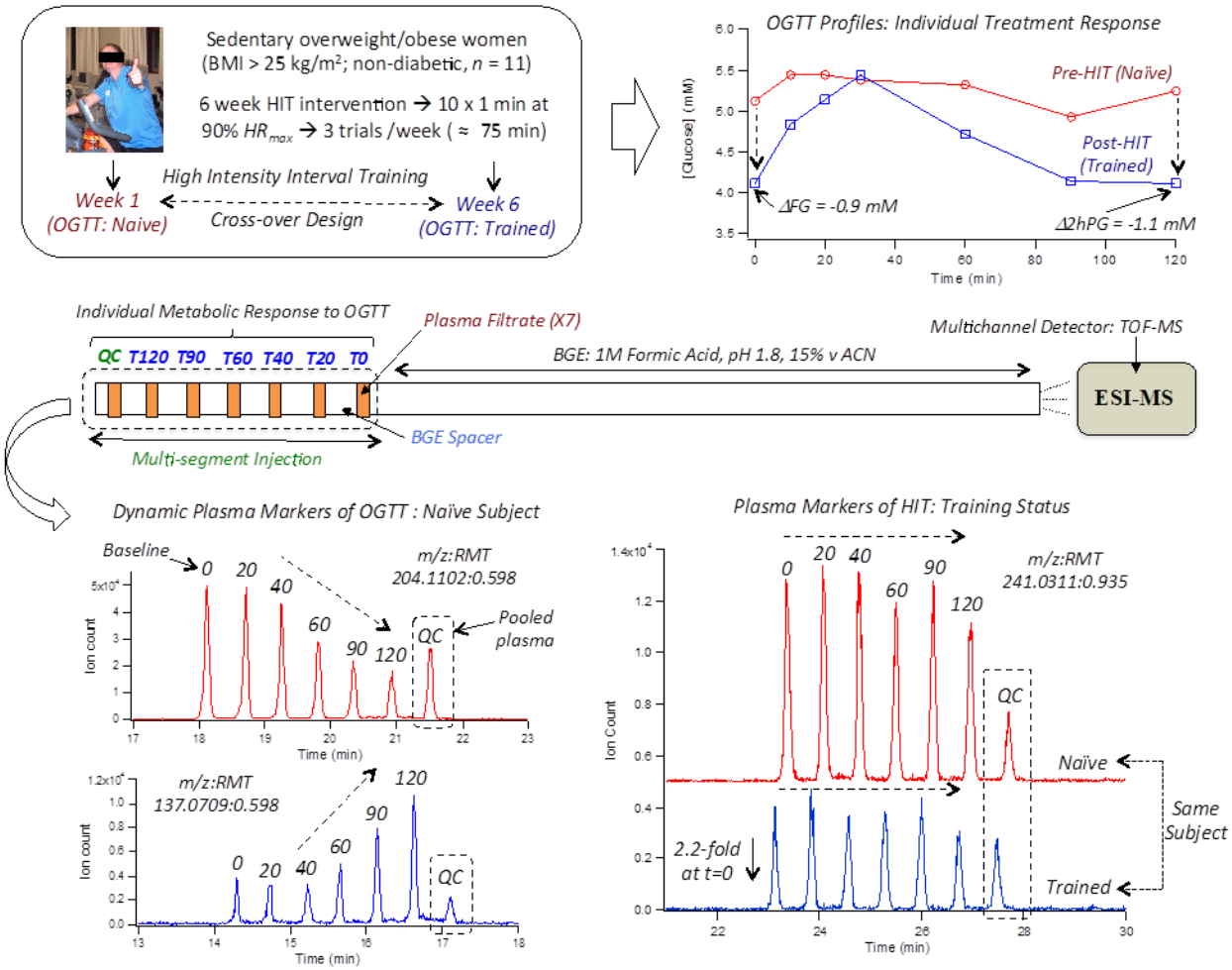


Figure 5.1. Experimental design and injection configuration used in MSI-CE-MS for elucidation of the impact of high intensity interval training (HIT) on modulating glucose tolerance in a cohort of overweight/obese females. Dynamic metabolomic responses to oral glucose loading is measured at six time intervals for each subject within a single analysis by MSI-CE-MS, including a pooled plasma that serves as quality control. In addition to changes in plasma glucose levels before (naïve) and after (trained) HIT intervention, other classes of polar metabolites undergo changes as a function of time and/or training status, which can serve as a prognostic markers of differential responses to exercise training, such as lower fasting glucose (FG) and/or plasma glucose levels at 2 h (2hPG).

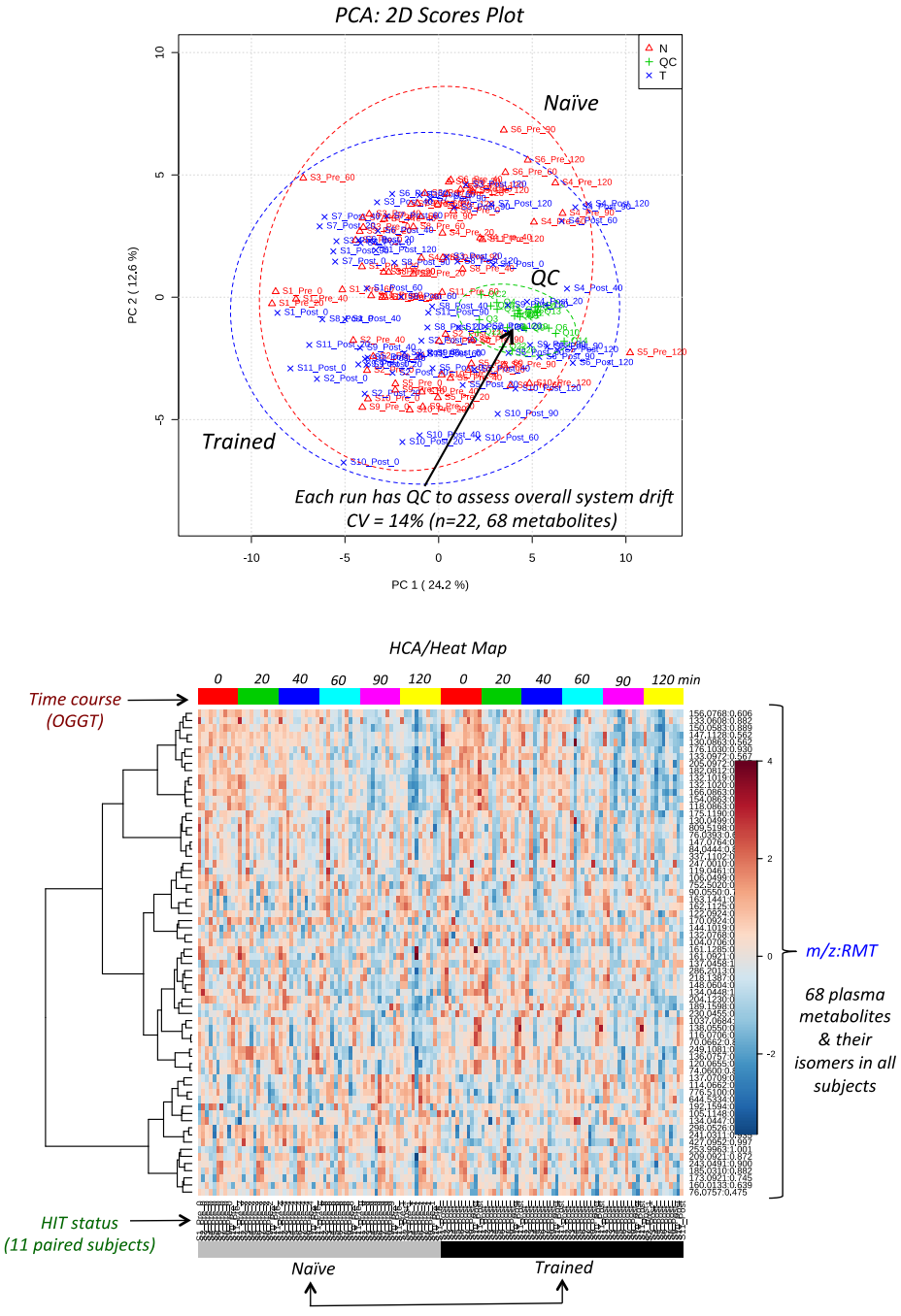


Figure 5.2. Data overview and quality assurance when using principle component analysis (PCA) that demonstrates good method reproducibility over two days as reflected by the tight clustering of QCs (*n*=22) in a 2D scores plot relative to biological variance with an average *CV* of 14% for 68 plasma metabolites. Hierarchical cluster analysis (HCA) of auto-scaled/log-transformed metabolomic data depicting the data structure of the cross-over study that is dependent on training status and glucose tolerance responses (time course) for individual subjects.

characteristic *m/z*:*RMT*.[36] **Figure 5.1** highlights three distinct metabolomic signal patterns detected by MSI-CE-MS after glucose loading as reflected by a series of extracted ion electropherograms for representative plasma metabolites. Unlike conventional metabolomic workflows that require complicated time alignment and peak picking algorithms, dynamic metabolomic responses associated with postprandial glucose are readily apparent for individual subjects by direct comparison of their normalized ion responses relative to the internal standard used in each sample segment. Overall, plasma metabolites were found to undergo either a significant decrease or increase in response over time after high dose glucose intake, whereas other metabolites remain largely unchanged, but in some cases were modulated by training status. **Figure 5.2** illustrates two methods used for assessment of overall data quality, including a 2D scores plot by PCA that highlights tight clustering of pooled QCs with an average $CV \approx 14\%$ ($n = 22$) over two consecutive days of analysis. The reproducibility of MSI-CE-MS is thus assessed when using built-in QCs incorporated into every separation, which increases sample throughput while ensuring high data fidelity. Also, a HCA heat map depicts the overall data structure of the study in terms of two factors (*i.e.*, postprandial glucose time course and training status) involving 11 paired subjects based on 68 cationic/zwitter-ionic metabolites denoted by their *m/z:RMT*, which were consistently measured in all plasma samples.

5.4.2 Dynamic Metabolomic Responses to Glucose Challenge and Exercise Training

Complementary statistical methods were used to classify plasma metabolites associated with glucose challenge tests and/or training status. **Figure 5.3** summarizes between-subject variance using 2-way ANOVA with Bonferroni correction ($p < 0.05$) to correct for multiple hypothesis testing, where 21 and 3 metabolites were associated with postprandial glucose time course and HIT training, respectively. Metabolites were identified by spiking plasma samples with authentic standards whenever available after searching public databases for putative molecular structures based on their accurate mass/empirical formula and electromigration behaviour (*i.e.*, cations), whereas unknown metabolites remain annotated by their *m/z:RMT*. Oxidized disulphides were the only metabolite class (3 out of 68) significantly modulated after exercise training as reflected by lower circulating levels of cysteinylglycine-cysteine disulfide (GlyCys-Cys-SS), glutathionylcysteine-cysteine disulfide (GSH-Cys-SS) and cystine (Cys-SS).

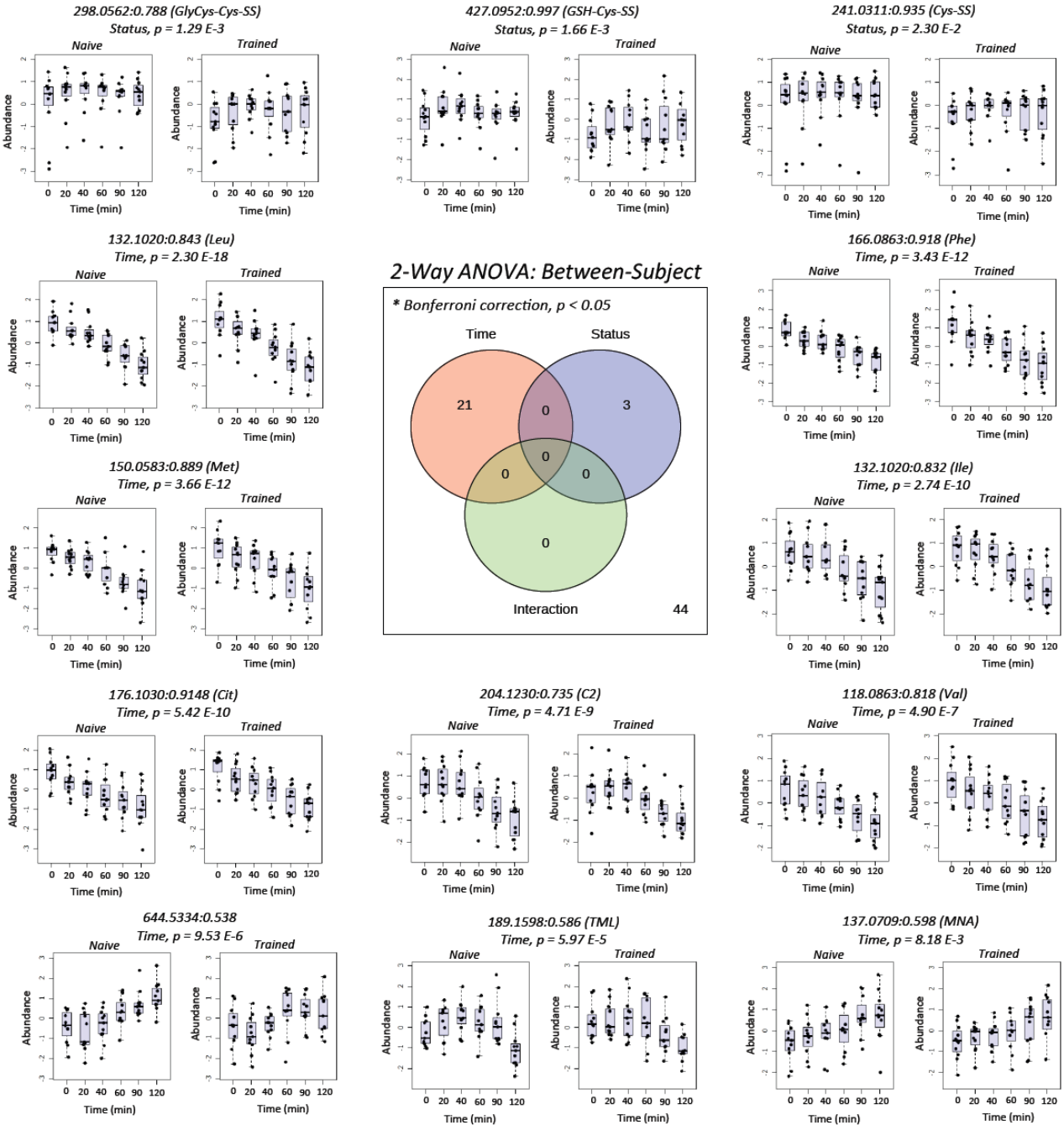


Figure 5.3. 2-way ANOVA (between-subject) for evaluation of dynamic plasma metabolomic changes associated with time (*i.e.*, postprandial oral glucose) and/or training status (*i.e.*, HIT) when using MSI-CE-MS after a dilution trend filter for unambiguous identification of authentic molecular features. Overall, plasma thiol redox status was significantly lower for plasma oxidized disulfides upon exercise training (CysGly-CysSS, GSH-CysSS and CysSS), whereas specific amino acids were lower in plasma after glucose loading as highlighted by representative metabolic profiles, notably Leu, Phe, Met, Ile and Cit ($p < 1 \times 10^{-10}$). Other classes of metabolites were also modulated with unique time-dependent profiles, including several unknown molecular features. No metabolites were significantly impacted by both training status and postprandial glucose loading (*i.e.*, no interaction terms).

However, oxidized disulfides and the majority of other plasma metabolites (44 out of 68) were independent of postprandial glucose time course. In contrast, a large number of plasma metabolites were found to undergo significant time-dependent decreases after glucose loading (21 out of 69) as summarized in **Table 5.2**, including branched-chain (Leu, Ile, Val), aromatic (Phe, Trp, Tyr), and sulfur amino acids (Met), as well as urea cycle intermediates (Cit) and carnitine precursors/metabolites (C2, Lys). In contrast, two specific plasma metabolites were found to increase in plasma upon postprandial glucose, including *N*-methylnicotinamide (MNA) and an unknown plasma metabolite. The mean responses for all plasma markers of postprandial glucose were not significantly modulated after HIT for trained subjects at rest with the exception of Ornithine (Orn). **Figure 5.4** confirms that thiol redox status as related to circulating oxidized disulfides (GSH-Cys-SS, GlyCys-CysSS, CysSS) was shifted lower following HIT intervention, whereas both Orn and trimethyllysine (TML) were upregulated after exercise training, which comprise the top five candidates in the variable importance in the projection (VIP) when using PLS-DA (R^2 = 0.986). A Wilcoxon-signed rank test that takes advantage of the cross-over design further supports that plasma thiol redox (GSH-Cys-SS) and Orn status were significantly modulated ($p < 0.05$) after HIT, whereas ROC curves demonstrate good classification of training status when using a single (GlyCys-Cys-SS, *AUC* = 0.769, p = 2.86 E-2) or ratiometric (TML/GlyCys-Cys-SS, *AUC* = 0.876), p = 5.81 E-3) plasma marker.

Figure 5.5 shows excellent group classification (0 min vs. 2hPG) when using PLS-DA (R^2 = 0.995) as determined by major decreases of plasma amino acids in the VIP ranking, notably Leu and Cit. Moreover, a Wilcoxon-signed rank test highlights that circulating levels of branched-chain amino acids (Leu/Ile), a carnitine precursor (TML) and a urea cycle intermediate (Cit) were significantly depleted in plasma after P2hG intake. Furthermore, ROC curves confirm that Leu is the most significant plasma marker of glucose loading after 2 h relative to baseline (*AUC* = 0.992, p = 1.94 E-6), whereas the Leu/carnitine (Leu/C0) ratio provides unambiguous classification of the distinct metabotype (*AUC* = 1.00, p = 2.80 E-8) of women who transition from a fasting/catabolic to fed/anabolic state. Similar results were found when comparing glucose tolerance responses for untrained women prior to HIT (data not shown). As expected, glucose challenge tests performed while fasting generate more pronounced changes in plasma metabolism when comparing baseline and P2hG postprandial glucose responses relative to long-term exercise training under free living conditions without explicit diet regulation (**Figure 5.4**).

Table 5.1. Summary of 24 significant plasma metabolites associated with postprandial oral glucose time course and HIT training status as classified by 2-way ANOVA with Bonferroni correction ($p < 0.05$).

Plasma Metabolite	*m/z:RMT*	*2-Way ANOVA (p-value)*	*Significance (Metabolic Pathway)*
L-Leucine (Leu)	132.1020: 0.843	2.30 E-18	Branched-chain amino acid; decrease with time
L-Phenylalanine (Phe)	166.0863:0.918	3.43 E-12	Aromatic amino acid; decrease with time
L-Methionine (Met)	150.0583:0.889	3.66 E-12	Sulfur amino acid; decrease with time
L-Isoleucine (Ile)	132.1020:0.832	2.74 E-10	Branched-chain amino acid; decrease with time
L-Citrulline (Cit)	176.1030:0.915	5.42 E-10	Urea cycle intermediate; decrease with time
O-Acetyl-*L*-carnitine (C2)	204.1230:0.735	4.71 E-9	Antioxidant/fatty acid oxidation; decrease with time
Adduct of Creatinine ($C_9H_{16}N_2O_6$)	249.1081:0.538	2.06 E-8	--
L-Valine (Val)	118.0863:0.818	4.90 E-7	Branched-chain amino acid; decrease with time
L-Lysine (Lys)	147.1128:0.562	5.40 E-6	Cationic amino acid; decrease with time
C_5H_9NO	136.0757:0.591	2.41 E-5	--
N^6-Trimethyl-*L*-lysine (TML)*	189.1598:0.586	5.97 E-5	C0 precursor/Fatty acid oxidation; increase with time
L-Tryptophan (Trp)	205.0972:0.917	1.13 E-4	Aromatic amino acid; decrease with time
L-Tyrosine (Tyr)	182.0812:0.951	2.34 E-4	Aromatic amino acid; decrease with time
Dopamine (Dpn)*	154.0863:0.801	3.42 E-4	Biogenic amine; decrease with time
L-Histidine (His)	156.0768:0.606	4.34 E-4	Cationic amino acid; decrease with time
Pipecolic acid (PA)*	130.0863:0.562	8.22 E-4	Degradation by-product of Lys; increase with time
Cystinylglycine-cysteine disulfide (CysGly-CysSS)	298.0526:0.788	1.29 E-3	Antioxidant/thiol redox status; decrease with training
Glutathione-cysteine disulfide (GSH-Cys-SS)	427.0952:0.997	1.66 E-3	Antioxidant/thiol redox status; decrease with training
N-methylnicotinamide (MNA)*	137.0709:0.598	8.18 E-3	Vitamin B3 metabolite; increase with training
L-Ornithine (Orn)	133.0972:0.567	1.05 E-2	Urea cycle intermediate; decrease with time and increase with training
L-Cystine (Cys-SS)	241.0311:0.935	2.30 E-2	Antioxidant/thiol redox status; decrease with training
L-Arginine	175.1190:0.579	3.31 E-2	Urea cycle/NO biosynthesis; decrease with time

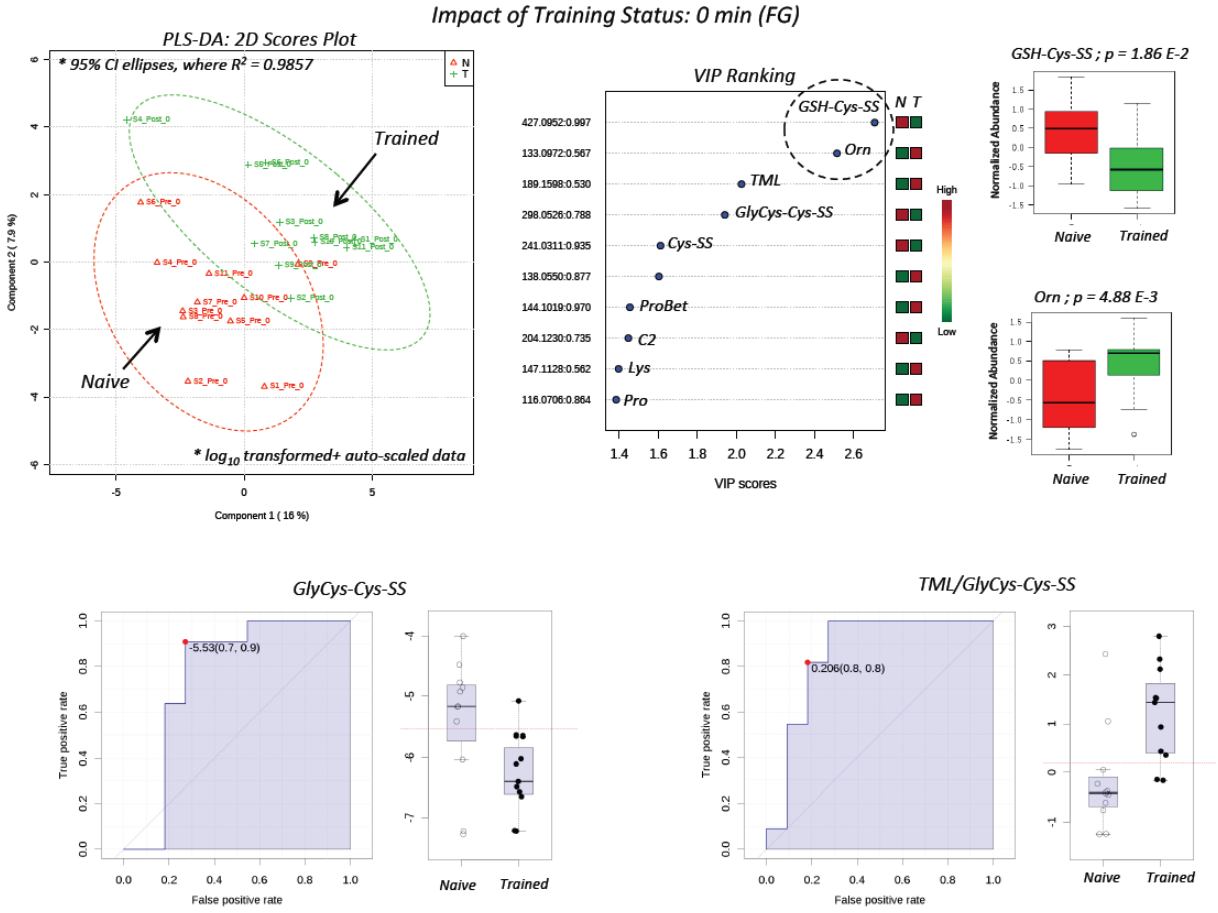


Figure 5.4. Identification of plasma markers of training status when using PLS-DA, where lower oxidized disulfides (GSH-Cys-SS, GlyCys-Cys-SS, Cys-SS) and higher ornithine (Orn) and trimethyllysine (TML) levels were measured in trained subjects after 6 week HIT intervention at rest (0 min). A paired Wilcoxon rank test confirms that both GSH-Cys-SS and Orn are significant features ($p < 0.05$) modulated by HIT status. Similarly, receiver operating characteristic (ROC) curves also demonstrates that plasma GlyCys-CysSS (*AUC* = 0.769, p = 2.86 E-2) and notably the TML/GlyCys-Cys-SS ratio (*AUC* = 0.876, p = 5.81 E-3) are among the most significant plasma metabolites used to classify training status among a heterogeneous cohort of overweight women.

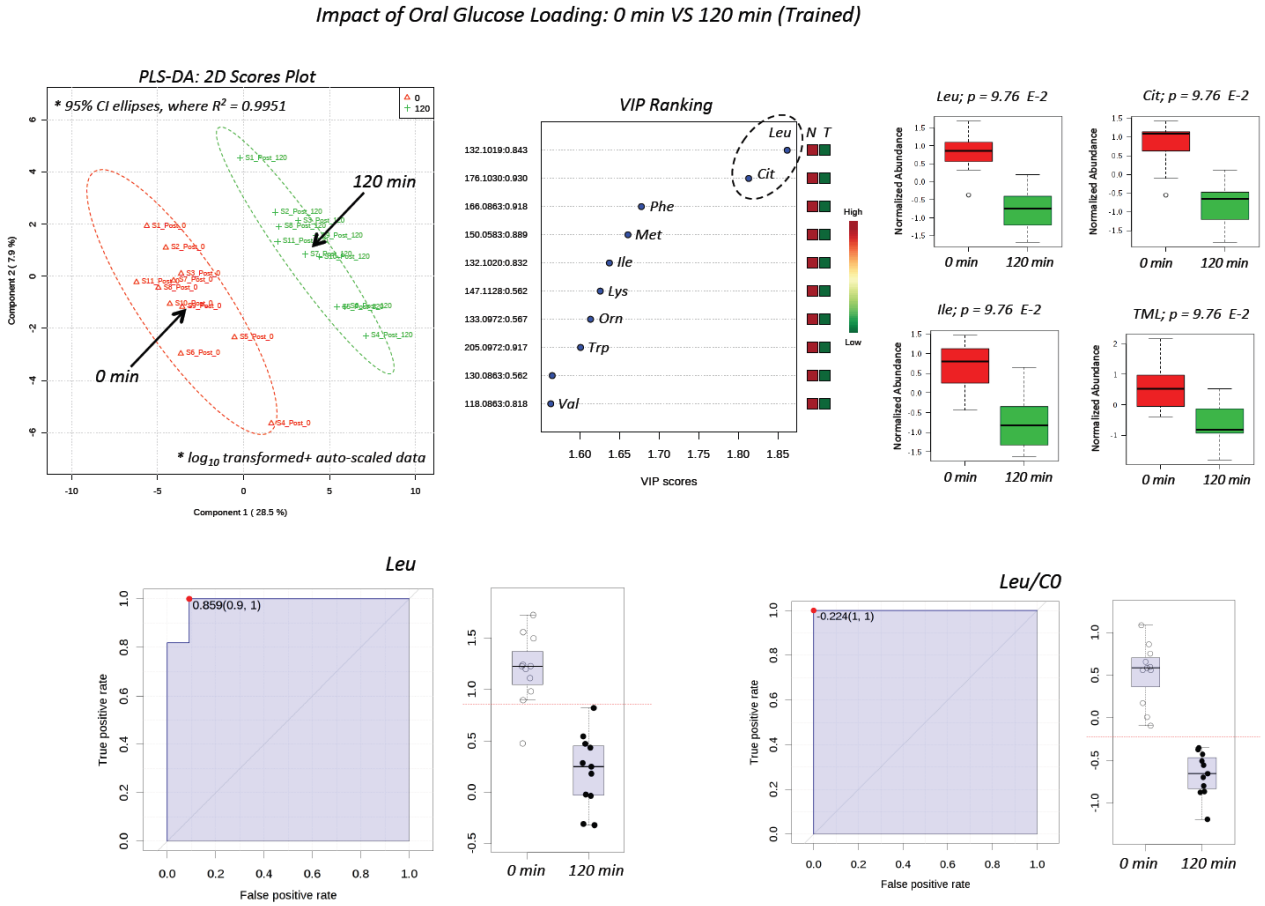


Figure 5.5. Identification of dynamic plasma markers in response to a standardized oral glucose tolerance test (OGTT) when using PLS-DA, where lower plasma amino acids (Leu/Ile, Orn, Phe, Met) were measured over 2 h following a glucose load. A paired Wilcoxon rank test confirms that five amino acids and their by-products were most significantly decreased ($p < 0.001$) at 2 h relative to baseline levels. Similarly, receiver operating characteristic (ROC) curves also demonstrates that plasma Leu (*AUC* = 0.992, $p = 1.94$ E-6) and notably Leu/C0 ratio (*AUC* =1.00, $p = 2.80$ E-8) are among the most significant plasma metabolites modulated by postprandial glucose for trained subjects.

5.4.3 Prognostic Markers of Glucose Tolerance Responses to HIT

Due to the large between-subject variability in physiological outcomes with HIT that is uncorrelated with body composition (**Table 5.1**), a multiple linear regression (MLR) model was developed to predict changes in glucose tolerance (*i.e.,* Δ2hPG) based on a panel of plasma metabolites measured for untrained subjects at 0 min and 120 min. Metabolites were selected based on their importance as markers of training status (*e.g.,* CysSS, TML/GlyCysSS) or glucose loading (*e.g.,* Leu, Leu/C0) from top candidates identified in PLS-DA, Wilcoxon-signed rank tests and ROC curves. The MLR model was optimized iteratively by minimizing the number of plasma metabolites used as variables while maintaining good accuracy (R^2 > 0.900). Permutation testing of the training set was assessed using leave-one-out cross-validation, which demonstrated good model accuracy (R^2 = 0.990) with adequate robustness (Q^2 = 0.641) when predicting changes in 2hPG responses for individual subjects withheld from the original model with the exception of one subject (S5). Putative prognostic markers of glucose tolerance included compounds associated with amino acid, thiol redox, urea cycle and carnitine metabolism; however, there was no direct correlation between gains in aerobic endurance capacity and changes in glucose tolerance for trained subjects relative to their naïve state. **Figure 5.6** shows that a panel of eight plasma metabolites as their single or ratiometric variables (six of eight metabolites with p < 0.05) was able to predict changes in 2hPG after a 6-week HIT intervention despite the considerable disparity in glucose tolerance responses. Although HIT elicited an improvement in glucose tolerance for a major fraction of the cohort (≈ 54% of subjects with Δ2hPG of 0.5 mM or lower than baseline), there was also a sub-group of non-responders (≈ 27% of subjects with negligible changes in 2hPG) and adverse responders (≈ 18% of subjects with Δ2hPG of 0.5 mM or higher than baseline), such as S2. Overall, Orn status was the most significant variable (p = 0.0085) for predicting glucose tolerance changes in women after exercise training.

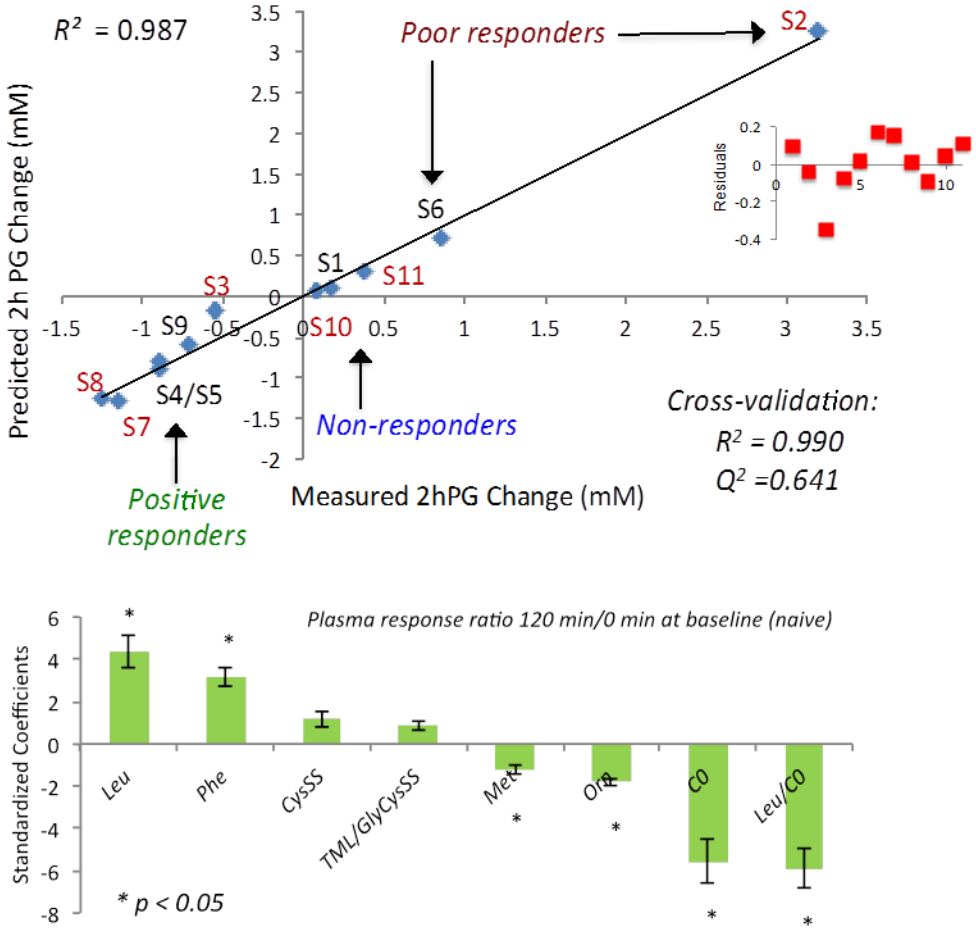


Figure 5.6. A multiple linear regression (MLR) model (R^2 = 0.987) for predicting changes in 2 h post-glucose (2hPG) response after 6 week HIT intervention based on a panel of eight plasma metabolites measured at baseline (120 min/0 min) for naïve female subjects. Good overall model accuracy (R^2 = 0.990) and adequate robustness (Q^2 = 0.641, S5 was excluded) was evaluated when using leave one out cross-validation as a permutation test on the original training set. Despite the large between-subject variability in glucose tolerance outcomes, positive responses to HIT intervention was realized in a major fraction of subjects (54%, 6 out of 11) with significant decreases in 2hPG levels (< 0.50 mM) relative to baseline that were not directly associated with body composition or gains in aerobic capacity with exercise training. However, there were also sub-groups of overweight women who had no significant change in 2hPG levels after HIT (27% as non-responders or 3 out of 11) with two participants denoted as poor responders (18% or 2 out of 11) with adverse glucose tolerance outcomes, notably in the case of S2.

5.5 Discussion

Type 2 diabetes is a chronic metabolic disorder associated with impaired glucose tolerance and insulin resistance that represents a major health crisis worldwide.[42] Glucose challenge tests provide a more sensitive approach for detecting asymptomatic diabetes than FG levels or glycated haemoglobin A1C,[43] whereas postprandial hyperglycemia is an independent risk factor for cardiovascular disease in diabetics,[44] as well as acute cardiovascular events in patients without a known diabetes history.[45] Although OGTTs suffer from significant within-subject variability with cut-off thresholds dependent on age and ethnicity,[46] results can be improved when evaluating multiple time points after glucose loading or using standardized meal tests that are also more palatable.[47] In our work, two standardized 75 g oral glucose challenge tests were performed on healthy non-diabetic women in the morning while fasting prior to and after a 6 week HIT intervention.[23] Participants in the study were primarily overweight (*BMI* ≈ 25-30 kg/m^2) young female adults who have a wide disparity in adiposity[48] ranging from 32-50% body fat indicating a heterogeneous cohort in terms of body composition and baseline aerobic fitness as summarized in **Table 5.1**. Overall, there were large differences in OGTT response curves between subjects; however, participants did not have impaired fasting glucose levels (*i.e.,* FG < 5.6 mM) or impaired glucose tolerance (*i.e.,* 2hPG < 7.8 mM) with the exception of a subject (S10) who had a slightly elevated FG (6.2 mM) and another subject (S9) with a borderline 2hPG result (7.5 mM) at baseline. Interestingly, in both cases, their glucose responses were restored to within normal ranges after HIT. The lack of correlation between FG and 2hPG outcomes in this work highlights that these two parameters of abnormal glucose metabolism do not measure the same population while likely underlying different pathophysiological processes in diabetes progression.[49]

5.5.1 Dynamic Metabolic Responses to Postprandial Glucose

During periods of fasting, such as an overnight fast prior to an OGTT, glucose homeostasis is maintained via glycogenolysis and gluconeogenesis, as well as hepatic ketogenesis which produces β-hydroxybutyrate and acetoacetate from metabolism of fatty acids released by adipose tissue.[50] Glucose ingestion causes the pancreas to release insulin and this promotes the uptake of glucose into metabolically active peripheral tissues, such as skeletal muscle.[32,33] Glucose acts as a primary source of energy in the glycolysis pathway to form pyruvate under aerobic conditions that is then converted to acetyl coenzyme A, which is the entry

point to the citric acid cycle. Concomitantly, there is a transition from catabolic (*e.g.*, proteolysis) to anabolic pathways (*e.g.*, protein biosynthesis) when switching from a fasting to a feeding state as triggered by an oral glucose challenge.[32] In this work, amino acids were the major class of metabolite (**Table 5.2)** that undergo a decrease in plasma over a 2 h time frame after glucose intake as shown in **Figure 5.3**. Previous studies have shown that insulin levels typically peak around 30 min after glucose loading, whereas plasma amino acids do not return to baseline until after 6 h.[51] Proteolysis is a pathway that is inhibited with glucose loading, resulting in selective uptake of (conditionally) essential amino acids by tissue for protein biosynthesis, notably branched-chain (Leu, Ile, Val) and aromatic amino acids (Phe, Trp, Tyr), as well as Met and Lys.[52] This is in agreement with previous metabolomic studies that have noted specific decreases in plasma amino acids after glucose ingestion reflecting decreased protein catabolism due to insulin action.[31,33,53,54] However, two previous reports were unable to resolve Leu and Ile when using LC-MS,[29,32] thus measurements were based on changes in total Leu. In our case, leucine isomers are baseline resolved prior to ionization when using MSI-CE-MS with Leu (p = 2.30 E-18) having the most significant time-dependent decrease in plasma among all other polar metabolites, including Ile (p = 2.74 E-10), whereas Val (p = 4.90 E-7) is taken up by peripheral tissue to a lower extent (**Table 5.2**).[32,51] For instance, recent studies have shown that plasma Ile levels are blunted after glucose loading in insulin resistant subjects relative to controls.[31] In this work, Phe (p = 3.43 E-12) and Met (p = 3.66 E-12) were among the top three amino acids in addition to Leu that decrease in plasma with postprandial glucose. Indeed, branched-chain and aromatic amino acids represent predictors of insulin resistance in women after a 6 year follow-up, namely circulating levels of Leu, Val and Phe.[55] Similarly, elevated branched-chain amino acids in children is associated with obesity that may independently predict insulin resistance.[56] Thus, there is growing interest in understanding branched-chain amino acid metabolism in hyperglycemia and obesity-related insulin resistance[57,58] that is relevant to cardiovascular health outcomes[59] and vascular deterioration.[60] Thus, the specific composition of dietary amino acids has a profound impact on human health not only as essential macronutrients, but also as modulators of hormonal signalling. For instance, Met restriction extends lifespan while also reducing insulin resistance and body weight gain in mice fed on a high fat diet,[61] as well as reducing cancer risk with aging.[62]

Other plasma metabolites that respond to a glucose challenge include cationic amino acids and their metabolites derived from Lys and Arg. For instance, Orn (p = 1.05 E-2) and notably Cit (p = 5.42 E-10) are two non-

proteinogenic amino acid intermediates of Arg that comprise the urea cycle that decrease with postprandial glucose as shown in **Figure 5.3**. This is in agreement with previous work that has demonstrated that circulating levels of Cit and/or Orn[29,31–33] along with other urea cycle intermediates (*e.g.*, orotate)[31] decrease in response to a glucose challenge. Dysregulation in Arg metabolism is associated with insulin resistance and impaired nitric oxide production in diabetic patients with elevated plasma arginase activity[63] that contributes to vascular aging.[64] Amino acids associated with C0 biosynthesis, including Lys, Met and the non-proteinogenic amino acid, TML,[65] were also found to undergo changes in plasma with postprandial glucose. C0 is a cofactor for β-oxidation of fatty acids that mediates their transport as acylcarnitine esters that also prevents deleterious accumulation of acyl-CoA derivatives in the mitochondria.[66] C0 insufficiency due to aging and overnutrition is implicated in insulin resistance and impaired mitochondrial function,[67] whereas C0 therapy has been proposed for the treatment and prevention of diabetes.[68] Our recent work has shown that exercise training upregulates circulating levels of C0 while improving the cardiorespiratory fitness of sedentary women as reflected by attenuation of plasma hypoxanthine post-exercise, which is a marker of energetic stress due to irreversible ATP depletion.[36] The ROC curve in **Figure 5.5** highlights that Leu/C0 is a promising marker (*AUC* = 1.00; p = 2.80 E-8) for differentiation of glucose loading (relative to Leu alone) that may prove useful as a metabolic signature for early detection of diabetes.[31] Acylcarnitines are also associated with glucose loading that indicate a transition from a fasting to a fed state since lipolysis and fatty acid beta-oxidation pathways are inhibited by insulin with glucose serving as the primary source of energy. Although several medium and long-chain acylcarnitines have been shown to decrease from baseline levels in an OGTT,[28,34] *O*-acetyl-*L*-carnitine (C2) was the only acylcarnitine to decrease in plasma over 2 h (p = 4.71 E-9) as shown in **Figure 5.3**. Due to the small sample volumes (≈ 10 nL) injected in-capillary, detection of sub-micromolar levels of medium or long-chain acylcarnitines requires on-line sample preconcentration methods to boost concentration sensitivity in CE-ESI-MS[69] that precludes use of the multiplexed injection format in this work.[35] An increase in plasma C2 due to mitochondrial dysfunction is implicated in pro-inflammatory pathways related to insulin resistance[70] while also serving as a serum marker for predicting impaired glucose tolerance together with glycine and lysophosphotidylcholine.[71] Also, increased plasma C2 responses post-exercise was previously shown to reflect increased acetylation capacity for trained subjects during strenuous exercise after HIT.[36] As shown in **Figure 5.**3, MNA was one of only two metabolites detected with a significant increase in plasma

over 2 h (p = 8.18 E-3) after glucose loading that was not modulated by training status. B vitamins were recently reported to decrease in plasma after glucose challenge, however secondary metabolites of niacin were not measured.[31] MNA is a metabolite of nicotinamide and nicotinic acid that has been shown to increase methyl depletion and oxidative stress with high-dose/chronic dietary intake.[72] Nicotinamide supplementation has been investigated in several large-scale clinical trials for preventing the onset of type 1 diabetes;[73] however, nicotinamide overload may unintentionally contribute to the development of type 2 diabetes, insulin resistance, and childhood obesity[74] that coincides with public health initiatives involving niacin-fortification of cereal grains.[75] Further work is needed to better characterize nicotinamide metabolism and dietary patterns in relation to changes in glucose tolerance and exercise training.

5.5.2 Adaptive Metabolic Responses to Exercise Training

One of the major goals of this pilot study is to evaluate the impact of HIT for enhancing oral glucose tolerance responses on an individual level. To date, there have been few studies examining the efficacy of HIT to improve body composition and/or insulin sensitivity for overweight yet normoglycemic women as a target population.[21,23,76] There are important hormonal factors that contribute to greater biological variability in women (*e.g.,* menstrual cycle, oral contraceptive usage) notably when assessing their metabolic responses to an oral glucose challenge.[77] In this work, only 3 out of 68 plasma metabolites were significantly modulated by training status when comparing glucose tolerance responses before and after HIT when using 2-way ANOVA with Bonferroni correction ($p < 0.05$) as shown in **Figure 5.3**. All three metabolites represent circulating oxidized thiols as their intact symmetric (Cys-SS) or mixed cysteine disulfides (GlyCys-Cys-SS, GSH-Cys-SS) that decrease from baseline after exercise training. Plasma thiols have previously been reported to be subtly perturbed in response to glucose loading since GSH biosynthesis is inhibited by glucose and activated by insulin.[29] In our case, there was no significant time-dependent change in oxidized disulfides with postprandial glucose due to large between-subject variations. Cys-SS is the major circulating thiol that plays a key role in regulating the extra-cellular redox state of plasma protein, whereas reduced glutathione (GSH) represents the major intra-cellular antioxidant to maintain redox homeostasis that mediates detoxification and cell functions, such as apopotosis.[78] GSH biosynthesis is limited by the availability of Cys that is derived by the influx and reduction of Cys-SS into cells, whereas export of GSH with

thiol-disulfide exchange with Cys-SS generates GSH-Cys-SS.[79] Also, recycling of GSH and glutathione *S*-conjugates by γ-glutamyltransferase results in release of GlyCys that forms GlyCys-Cys-SS and other mixed disulfides in plasma.[80] Thus, measurement of plasma thiols provides deeper insight of intra-cellular glutathione homeostasis since exercise training enhances tissue-dependent glutathione antioxidant defenses.[81] Reduced thiols were not detected by MSI-CE-MS due to their lower concentration levels in plasma, as well as their susceptibility to auto-oxidation and protein thiol-disulfide exchange reactions. Thiol-specific chemical derivatization is needed during sample workup to form stable adducts with greater ionization efficiency, which is critical for sensitive yet artefact-free determination of plasma thiol redox status in metabolomics,[82] such as the GSH/GSSG and Cys/CysSS redox couples.[83] Redox analysis of human plasma shows a linear increase in oxidative events with age as reflected by a pro-oxidant shift in Cys/CysSS, whereas loss in antioxidant capacity corresponds to a more oxidized GSH/GSSG environment.[84] The pathological effects of elevated plasma oxidized disulfides is also implicated in renal or liver failure,[85] early stage atherosclerosis,[86] mitochondrial diseases in children,[87] and diabetes complications due to prolonged hyperglycemia.[88] Physical activity functions as an antioxidant by upregulating antioxidant genes that promote beneficial cell adaptations to exercise-induced oxidative stress.[89] In contrast, exhaustive prolonged exercise without training is deleterious as it can trigger high levels of reactive oxygen species and glutathione depletion that promotes contractile muscle dysfunction and tissue injury.[90] In this case, HIT provides a sufficient oxidative stimulus via transient bursts of high intensity exercise followed by periods of recovery/low activity that elicit a low systemic inflammatory response.[91] Overall, the TML/GlyCys-Cys-SS ratio (*AUC* = 0.876, p = 5.81 E-3) was found to be the most significant plasma marker for classifying training status as shown in **Figure 5.4**. Overall 95% of C0 is sequestered in muscle tissue, where it comes from meat/dairy product intake or it is synthesized endogenously using TML as a precursor from vegetables or protein turnover.[65] C0 insufficiency is associated with aging that compromises mitochondrial function[67] that can also become limiting during strenuous exercise.[92] Thus, a reduction of plasma thiol redox status together with upregulation in C0 biosynthetic pathways is associated with positive adaptive metabolic responses to exercise training.

Orn levels were also found to be upregulated (p = 4.88 E -3) after exercise training when using a Wilcoxon-signed rank test, as well as PLS-DA as shown in **Figure 5.4**. Orn is a key substrate in the urea cycle that is required for ammonia detoxification in the mitochondria. Hyperammonemia occurs in liver diseases, as

well as strenuous exercise due to deamination of adenosine monophosphate (*i.e.*, purine nucleotide cycle) and branched-chain amino acids (*e.g.*, Leu) as energy sources used by contracting muscle.[93] Toxic ammonia accumulation is associated with muscle fatigue and organ injury that can lead to deleterious effects on brain function, including encephalopathy. For instance, ornithine transcarbamoylase deficiency is the most common urea cycle disorder that results in seizures and developmental delays in children if left untreated by protein restriction and/or Cit therapy.[94] Indeed, oral supplementation of Orn has been shown to attenuate physical fatigue in subjects performing ergometer cycling by promoting lipid metabolism and mitochondrial function.[95] Analogous ergogenic effects have also been reported with dietary Cit supplementation that is more effective than Arg while also improving insulin responses by reducing nitric-oxide mediated pancreatic insulin secretion.[96] The much higher rate of peripheral tissue uptake of Cit (p = 5.42 E-10, **Table 5.2**) measured with postprandial glucose relative to Orn or Arg may reflect its greater bioavailability that is consistent with recent metabolomic studies.[31] Overall, Orn was the only plasma metabolite in this study to be independently modulated by exercise training (**Figure 5.4**) and glucose challenge (p = 0.0105, **Table 5.2**), although the interaction term was not significant when using 2-way ANOVA (**Figure 5.3**). In fact, Orn was the second most significant feature for classifying P2hG responses from baseline (AUC = 0.917, p = 5.73 E-6, data not shown) after Leu, as well as among the top four plasma markers for distinguishing training status (AUC = 0.769, p = 4.44 E-2, data not shown) in addition to oxidized disulfides when using ROC curves.

5.5.3 Prediction of Individual Glucose Tolerance Responses

Since low cardiorespiratory fitness and physical inactivity are associated with metabolic syndrome,[97] practical exercise interventions are needed for chronic disease prevention while promoting healthy aging. For instance, a supervised home-based exercise program has been shown to improve glucose tolerance and aerobic fitness levels in obese women at risk for gestational diabetes during pregnancy.[98] However, differential treatment responses between-subjects hampers translation of novel exercise protocols to population health, such as low-volume HIT.[7] Our work has demonstrated that there is a poor correlation between gains in aerobic endurance capacity (VO_{2max}, W_{max}) with changes in glucose tolerance (FG, 2hPG) among normoglycemic women as summarized in **Table 5.1**. Indeed, pre-diabetic subjects have been reported to derive a greater training-induced change in glucose tolerance than age and *BMI*-matched healthy controls.[99] Despite

comprising an overweight cohort (*BMI* ≈ 27) of young women with a large disparity in adiposity, there was no significant correlation of body composition with changes in glucose tolerance after HIT. Body fat distribution plays a key role in obesity-related insulin resistance, however, the functional differences of adipose tissue and their roles in diabetes progression remains unclear, such as visceral and subcutaneous fat.[100] Recent studies have examined microRNA expression in skeletal muscle to predict differential gains in maximum aerobic capacity with exercise training.[101] Also, plasma metabolites of exercise were found to be correlated with fitness and resting heart rate in healthy individuals reflecting underlying glucose utilization and lipid metabolism, such as glycerol, niacinamide, glucose-6-phosphate, pantothenate and succinate.[102] To the best our knowledge, this is the first metabolomics study to model changes in glucose tolerance modulated by exercise training on individual subjects. Since there is considerable within-subject/biological variation to OGTTs that vary by about 15% and 46% for FG and 2hPG measurements, respectively,[103] mean decreases in FG (-0.27 mM or relative decrease in mean by 5.4%) and notably 2hPG (-0.54 mM or relative decrease in mean by 11%) after HIT are nonetheless significant when considering paired subjects in the cross-over study design. Thus, there is about a two-fold greater treatment effect in glucose tolerance as compared to glucose homeostasis for the cohort although not all the same subjects responded in the same way (**Table 5.1**). For instance, S6 had the largest decrease in FG levels yet had a modest reduction in glucose tolerance relative to baseline following HIT. The modest treatment effects are obscured by much larger between-subject variations in glucose tolerance with Δ2hPG ranging from -1.2 mM (*i.e.*, positive responder, S7) to +3.2 mM (*i.e.*, poor responder, S2) with three subjects (S1, S10 and S11) having non-significant changes (< ± 0.4 mM, non-responder) as summarized in **Table 5.1**. Despite the small population size and gender-specific cohort examined in this work, the disparity in treatment responses to HIT as related to Δ2hPG is consistent with large-scale endurance training interventions, where 25% of individuals are reported to have no measurable improvements in insulin sensitivity despite a mean increase of 10% ($p < 0.001$) in the absence of substantial weight loss.[41] Notably, two subjects with either borderline impaired FG (S10) or glucose tolerance (S9) measurements at baseline responded positively after exercise training.

A MLR model comprising a panel of eight plasma metabolites (single or ratiometric markers) was developed to predict changes in oral glucose tolerance (Δ2hPG), including key compounds associated with branched-chain/aromatic amino acids, urea cycle, thiol redox status and carnitine metabolism. For instance,

Orn (p = 0.0085), Phe (p = 0.017) and Leu/C0 (p = 0.025) were among the most significant plasma markers associated with Δ2hPG responses as shown in **Figure 5.6**. The feeding status of subjects on the days of their HIT sessions was previously shown not to impact gains in aerobic endurance capacity or changes in glucose tolerance.[23] Although there is growing recognition of the important roles of branched-chain/aromatic/sulfur amino acids in glucose homeostasis,[52] dysregulation in Arg metabolism also contributes insulin resistance, hypertension and cardiovascular disease risk.[63] Although Orn was not among the most significant plasma metabolites uptaken after glucose loading (**Table 5.2**, p = 0.015), untrained subjects with attenuated changes in plasma Orn (Δ2hPG) prior to intervention were positively correlated *(*r = 0.613, exception of S5) with improved glucose tolerance outcomes; this was also consistent with the greater training-induced increases to Orn status measured after HIT intervention (**Figure 5.4**). Cross-validation of the training set demonstrated adequate model robustness, however further studies on an independent cohort are warranted to better validate the prognostic value of plasma markers to accurately predict glucose tolerance changes after exercise training. Expanded coverage of anionic metabolites is also needed when using CE-MS under alkaline conditions and negative ion mode detection,[104] as well as other orthogonal analytical platforms due to the chemical diversity and wide dynamic range of the human metabolome.[105] Indeed, a confounding factor in this study is that participants were involved in a 6-week HIT intervention study under free living conditions without explicit dietary control with the exception of exercise trial sessions and OGTTs performed while fasting. For instance, a single day of dietary standardization has been shown to provide adequate normalization of the human metabolome as measured by NMR with biological variance largely determined by genetics, lifestyle and age/gender.[106] This strategy avoids the burden for prolonged dietary control while minimizing biological variance needed for detecting subtle treatment effects in real-world settings. Although there is continued debate on the optimal level of exercise intensity and duration needed to enhance glucose tolerance when using exercise interval training,[107] undesirable physiological outcomes may occur for susceptible individuals due to deleterious oxidative stress and tissue injury.[108] In this context, we anticipate that prognostic markers of exercise-induced glucose tolerance changes developed in this work may be used as a novel screening tool to identify adverse responders (*e.g.*, S2/S6, **Figure 5.6**) based on their characteristic metabotype measured during baseline testing. This is required for the translation of personalized lifestyle interventions that support recovery following bouts of

high-intensity exercise while enhancing treatment efficacy, such as concurrent nutritional supplementation together with HIT.[109]

5.6 Concluding Remarks

Safe yet efficacious lifestyle modifications based on physical activity are needed for primary prevention of diabetes notably for high-risk populations, such as overweight/obese women. However, body composition alone provides poor insight into health status, whereas treatment responses to exercise are highly variable between-subjects that can lead in some cases to unanticipated adverse health outcomes. In this work, MSI-CE-MS was used as a high-throughput screening format for characterizing dynamic metabolomic responses to a glucose challenge on individual subjects both prior to and after a 6-week HIT intervention. An accelerated metabolomics workflow allowed for unambiguous identification of biomarkers in plasma associated with postprandial glucose loading and training status when using complementary statistical methods. Plasma thiol redox status was reduced as reflected by lower circulating levels of oxidized disulfides, whereas plasma ornithine status was upregulated indicating greater intra-cellular antioxidant and detoxification capacity for trained subjects, respectively. In addition, specific plasma metabolites were associated with time-dependent peripheral tissue uptake following glucose loading, notably branched-chain/aromatic/sulfur amino acids, as well as urea cycle and carnitine intermediates. The Leu/C0 ratio was the most significant plasma marker of postprandial glucose reflecting a transition from a fasting to fed metabolic state. Overall, there was a modest improvement in glucose tolerance for participants after HIT, although treatment effects were obscured by large between-subject variations, which were not correlated with body composition nor gains in aerobic endurance capacity. For the first time, a panel of plasma metabolites was explored as prognostic markers for predicting changes in glucose tolerance for sedentary women, which allows for potential screening of adverse responders to exercise training. Future work will examine a more diverse cohort in order to develop personalized lifestyle interventions that maximize the many salutary benefits of exercise, including improved skeletal muscle oxidative capacity, cardiorespiratory fitness and mental health.

5.7 References

(1) Woolf, K.; Reese, C.; Mason, M.; Beaird, L.; Tudor-Locke, C.; Vaughan, L. *J. Am. Diet. Assoc.* **2008**, *108*, 948–959.

(2) Mata, J.; Hogan, C.; Joormann, J.; Waugh, C.; Gotlib, I. *J. Abnorm. Psychol.* **2013**, *122*, 45–50.

(3) American Diabetes Association. *Diabetes Care* **2013**, *36*, 1033–1046.

(4) Knowler, W. C.; Fowler, S. E.; Hamman, R. F.; Christophi, C. a; Hoffman, H. J.; Brenneman, A. T.; Brown-Friday, J. O.; Goldberg, R.; Venditti, E.; Nathan, D. M. *Lancet* **2009**, *374*, 1677–86.

(5) Lindström, J.; Peltonen, M.; Eriksson, J. G.; Ilanne-Parikka, P.; Aunola, S.; Keinänen-Kiukaanniemi, S.; Uusitupa, M.; Tuomilehto, J.; Finnish Diabetes Prevention Study (DPS) *Diabetologia* **2013**, *56*, 284–293.

(6) Colley, R.; Garriguet, D.; Janssen, I.; Craig, C.; Clarke, J.; Tremblay, M. *Heal. Rep.* **2011**, *22*, 15–23.

(7) Gibala, M. J.; Little, J. P.; MacDonald, M. J.; Hawley, J. A. *J. Physiol.* **2012**, *590*, 1077–84.

(8) Hawley, J. A.; Gibala, M. J. *Diabetologia* **2009**, *52*, 1709–13.

(9) Gosselin, L.; Kozlowski, K.; DeVinney-Boymel, L.; Hambridge, C. *J. Strength Cond. Res.* **2012**, *26*, 2866–2871.

(10) Warburton, D. E. R.; McKenzie, D. C.; Haykowsky, M. J.; Taylor, A.; Shoemaker, P.; Ignaszewski, A. P.; Chan, S. Y. *Am. J. Cardiol.* **2005**, *95*, 1080–4.

(11) Wisløff, U.; Støylen, A.; Loennechen, J. P.; Bruvold, M.; Rognmo, Ø.; Haram, P. M.; Tjønna, A. E.; Helgerud, J.; Slørdahl, S. A.; Lee, S. J.; Videm, V.; Bye, A.; Smith, G. L.; Najjar, S. M.; Ellingsen, Ø.; Skjaerpe, T. *Circulation* **2007**, *115*, 3086–94.

(12) Moholdt, T. T.; Amundsen, B. H.; Rustad, L. A.; Wahba, A.; Løvø, K. T.; Gullikstad, L. R.; Bye, A.; Skogvoll, E.; Wisløff, U.; Slørdahl, S. A. *Am. Heart J.* **2009**, *158*, 1031–7.

(13) Munk, P. S.; Staal, E. M.; Butt, N.; Isaksen, K.; Larsen, A. I. *Am. Heart J.* **2009**, *158*, 734–41.

(14) Tjønna, A. E.; Lee, S. J.; Rognmo, Ø.; Stølen, T. O.; Bye, A.; Haram, P. M.; Loennechen, J. P.; Al-Share, Q. Y.; Skogvoll, E.; Slørdahl, S. A.; Kemi, O. J.; Najjar, S. M.; Wisløff, U. *Circulation* **2008**, *118*, 346–54.

(15) Tjønna, A. E.; Stølen, T. O.; Bye, A.; Volden, M.; Slørdahl, S. A.; Odegård, R.; Skogvoll, E.; Wisløff, U. *Clin. Sci.* **2009**, *116*, 317–26.

(16) Whyte, L. J.; Gill, J. M. R.; Cathcart, A. J. *Metabolism* **2010**, *59*, 1421–8.

(17) Babraj, J. a; Vollaard, N. B. J.; Keast, C.; Guppy, F. M.; Cottrell, G.; Timmons, J. a *BMC Endocr. Disord.* **2009**, *9*, 3.

(18) Richards, J. C.; Johnson, T. K.; Kuzma, J. N.; Lonac, M. C.; Schweder, M. M.; Voyles, W. F.; Bell, C. *J. Physiol.* **2010**, *588*, 2961–72.

(19) Little, J. P.; Gillen, J. B.; Percival, M. E.; Safdar, A.; Tarnopolsky, M. A.; Punthakee, Z.; Jung, M. E.; Gibala, M. J. *J. Appl. Physiol.* **2011**, *111*, 1554–60.

(20) Gillen, J. B.; Little, J. P.; Punthakee, Z.; Tarnopolsky, M. A.; Riddell, M. C.; Gibala, M. J. *Diabetes. Obes. Metab.* **2012**, *14*, 575–7.

(21) Trapp, E. G.; Chisholm, D. J.; Freund, J.; Boutcher, S. H. *Int. J. Obes.* **2008**, *32*, 684–91.

(22) Duncan, G.; Perri, M.; Theriaque, D.; Hutson, A.; Eckel, R.; Stacpoole, P. *Diabetes Care* **2003**, 557–562.

(23) Gillen, J. B.; Percival, M. E.; Ludzki, A.; Tarnopolsky, M. A.; Gibala, M. J. *Obesity* **2013**, *21*, 2249–2255.

(24) Timmons, J. A. *J. Appl. Physiol.* **2011**, *110*, 846–53.

(25) Bouchard, C.; Blair, S. N.; Church, T. S.; Earnest, C. P.; Hagberg, J. M.; Häkkinen, K.; Jenkins, N. T.; Karavirta, L.; Kraus, W. E.; Leon, A. S.; Rao, D. C.; Sarzynski, M. A.; Skinner, J. S.; Slentz, C. A.; Rankinen, T. *PLoS One* **2012**, *7*, e37887.

(26) Lee, R.; West, D.; Phillips, S.; Britz-McKibbin, P. *Anal. Chem.* **2010**, *82*, 2959.

(27) Chorell, E.; Moritz, T.; Branth, S.; Antti, H.; Svensson, M. B. *J. Proteome Res.* **2009**, *8*, 2966–2977.

(28) Krug, S.; Kastenmüller, G.; Stückler, F.; Rist, M. J. M.; Skurk, T.; Sailer, M.; Raffler, J.; Römisch-Margl, W.; Adamski, J.; Prehn, C.; Frank, T.; Engel, K. K.-H.; Hofmann, T.; Luy, B.; Zimmermann, R.; Moritz, F.; Schmitt-Kopplin, P.; Krumsiek, J.; Kremer, W.; Huber, F.; Oeh, U.; Theis, F. F. J.; Szymczak, W.; Hauner, H.; Suhre, K.; Daniel, H. *FASEB J.* **2012**, *26*, 2607–2619.

(29) Wopereis, S.; Rubingh, C. M.; van Erk, M. J.; Verheij, E. R.; van Vliet, T.; Cnubben, N. H. P.; Smilde, A. K.; van der Greef, J.; van Ommen, B.; Hendriks, H. F. J. *PLoS One* **2009**, *4*, e4525.

(30) Bain, J. R.; Muehlbauer, M. J. *Diabetes* **2013**, *62*, 2651–2653.

(31) Ho, J. E.; Larson, M. G.; Vasan, R. S.; Ghorbani, A.; Cheng, S.; Rhee, E. P.; Florez, J. C.; Clish, C. B.; Gerszten, R. E.; Wang, T. J. *Diabetes* **2013**, *62*, 2689–98.

(32) Shaham, O.; Wei, R.; Wang, T. J.; Ricciardi, C.; Lewis, G. D.; Vasan, R. S.; Carr, S. a; Thadhani, R.; Gerszten, R. E.; Mootha, V. K. *Mol. Syst. Biol.* **2008**, *4*, 214.

(33) Spégel, P.; Danielsson, A. P. H.; Bacos, K.; Nagorny, C. L. F.; Moritz, T.; Mulder, H.; Filipsson, K. *Metabolomics* **2009**, *6*, 56–66.

(34) Zhao, X.; Peter, A.; Fritsche, J.; Elcnerova, M.; Fritsche, A.; Häring, H.-U.; Schleicher, E. D.; Xu, G.; Lehmann, R. *Am. J. Physiol. Endocrinol. Metab.* **2009**, *296*, E384–93.

(35) Kuehnbaum, N. L.; Kormendi, A.; Britz-McKibbin, P. *Anal. Chem.* **2013**, *85*, 10664–9.

(36) Kuehnbaum, N. L.; Britz-McKibbin, P.; Gillen, J. B.; Gibala, M. J. *Submitted to J. Proteome Res.* **2014**.

(37) Kamleh, M. A.; Ebbels, T. M.; Spagou, K.; Masson, P.; Want, E. J. *Anal. Chem.* **2012**, *84*, 2670–2677.

(38) Broadhurst, D. I.; Kell, D. B. *Metabolomics* **2007**, *2*, 171–196.

(39) Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S. *Nucleic Acids Res.* **2012**, *40*, W127–W133.

(40) Xia, J.; Broadhurst, D.; Wilson, M.; Wishart, D. *Metabolomics* **2013**, *9*, 280–299.

(41) Boule, N. G.; Weisnagel, S. J.; Lakka, T. A.; Tremblay, A.; Bergman, R. N.; Rankinen, T.; Leon, A. S.; Skinner, S. J.; Wilmore, J. H.; Rao, D. C.; Bouchard, C. *Diabetes Care* **2005**, *28*, 108–114.

(42) Ginter, E.; Simko, V. *v Exp Med Biol.* **2012**, *771*, 42–50.

(43) Bartoli, E.; Fra, G. P.; Carnevale Schianca, G. P. *Eur J Intern Med* **2011**, *22*, 8–12.

(44) Ceriello, A.; Davidson, J.; Hanefeld, M.; Leiter, L.; Monnier, L.; Owens, D.; Tajima, N.; Tuomilehto, J.; International Prandial Glucose Regulation Study Group *Nutr Metab Cardiovasc Dis* **2006**, *16*, 453–456.

(45) Henareh, L.; Agewall, S. *Cardiovasc Diabetol* **2012**, *11*, 93.

(46) Utzschneider, K. M.; Prigeon, R. L.; Tong, J.; Gerchman, F.; Carr, D. B.; Zraika, S.; Udayasankar, J.; Montgomery, B.; Mari, A.; Kahn, S. E. *Diabetologia* **2007**, *50*, 2516–25.

(47) Wolever, T. M.; Chiasson, J. L.; Csima, A.; Hunt, J. A.; Palmason, C.; Ross, S. A.; Ryan, E. A. *Diabetes Care* **1998**, *21*, 336–40.

(48) Shah, N. R.; Braverman, E. R. *PLoS One* **2012**, *7*, e33308.

(49) Unwin, N.; Shaw, J.; Zimmet, P.; Alberti, K. G. *Diabet. Med.* **2002**, *19*, 708–23.

(50) Cahill, G. F. *Ann. Rev. Nutr.* **2006**, *26*, 1–22.

(51) Hijikata, Y.; Shiozaki, Y.; Sameshima, Y. *J. Clin. Chem. Clin. Biochem.* **1985**, *23*, 259–64.

(52) Adams, S. H. *Adv Nutr* **2011**, *2*, 445–56.

(53) Pan, R. M.; Mauron, C.; Glaeser, B.; Wurtman, R. J. *Metabolism.* **1982**, *31*, 937–43.

(54) Felig, P.; Wahren, J.; Hendler, R. *Diabetes* **1975**, *24*, 468–75.

(55) Würtz, P.; Soininen, P.; Kangas, A. J.; Rönnemaa, T.; Lehtimäki, T.; Kähönen, M.; Viikari, J. S.; Raitakari, O. T.; Ala-Korpela, M. *Diabetes Care* **2013**, *36*, 648–55.

(56) 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.; Fleischman, a *Pediatr. Obes.* **2013**, *8*, 52–61.

(57) Lu, J.; Xie, G.; Jia, W.; Jia, W. *Front Med* **2013**, *7*, 53–9.

(58) 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.; Svetkey, L. P. *Cell Metab.* **2009**, *9*, 311–26.

(59) Batch, B. C.; Hyland, K.; Svetkey, L. P. *Curr. Opin. Clin. Nutr. Metab. Care* **2014**, *17*, 86–9.

(60) Mels, C. M.; Schutte, A. E.; Schutte, R.; Huisman, H. W.; Smith, W.; Fourie, C. M.; Kruger, R.; van Rooyen, J. M.; Malan, N. T.; Malan, L. *Amino Acids* **2013**, *45*, 1405–13.

(61) Ables, G. P.; Perrone, C. E.; Orentreich, D.; Orentreich, N. *PLoS One* **2012**, *7*, e51357.

(62) Cavuoto, P.; Fenech, M. *Cancer Treat. Rev.* **2012**, *38*, 726–736.

(63) Kashyap, S. R.; Lara, A.; Zhang, R.; Park, Y. M.; DeFronzo, R. A. *Diabetes Care* **2008**, *31*, 134–9.

(64) Pernow, J.; Jung, C. *Cardiovasc Res* **2013**, *98*, 334–43.

(65) Servillo, L.; Giovane, A.; Cautela, D.; Castaldo, D.; Balestrieri, M. L. *PLoS One* **2014**, *9*, e84589.

(66) Mingrone, G. *Ann NY Acad Sci* **2004**, *1033*, 99–107.

(67) Noland, R. C.; Koves, T. R.; Seiler, S. E.; Lum, H.; Lust, R. M.; Ilkayeva, O.; Stevens, R. D.; Hegardt, F. G.; Muoio, D. M. *J Biol Chem* **2009**, *284*, 22840–52.

(68) Ringseis, R.; Keller, J.; Eder, K. *Eur J Nutr* **2012**, *51*, 1–18.

(69) Chalcraft, K. R.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 307–314.

(70) Adams, S. H.; Hoppel, C. L.; Lok, K. H.; Zhao, L.; Wong, S. W.; Minkler, P. E.; Hwang, D. H.; Newman, J. W.; Garvey, W. T. *J Nutr* **2009**, *139*, 1073–81.

(71) 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.; Doring, 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.; Illig, T. *Mol Syst Biol* **2012**, *8*, 615.

(72) Li, D.; Luo, N.; Ma, Q.; Li, S. Z.; Shi, Q.; Cao, Y.; Zhou, S. S. *Pharm Biol* **2013**, *51*, 8–12.

(73) Cabrera-Rode, E.; Molina, G.; Arranz, C.; Vera, M.; Gonzalez, P.; Suarez, R.; Prieto, M.; Padron, S.; Leon, R.; Tillan, J.; Garcia, I.; Tiberti, C.; Rodriguez, O. M.; Gutiérrez, A.; Fernández, T.; Govea, A.; Hernández, J.; Chiong, D.; Dominguez, E.; Di Mario, U.; Diaz-Diaz, O.; Diaz-Horta, O. *Autoimmunity* **2006**, *39*, 333–40.

(74) Zhou, S.-S.; Li, D.; Sun, W.-P.; Guo, M.; Lun, Y.-Z.; Zhou, Y.-M.; Xiao, F.-C.; Jing, L.-X.; Sun, S.-X.; Zhang, L.-B.; Luo, N.; Bian, F.-N.; Zou, W.; Dong, L.-B.; Zhao, Z.-G.; Li, S.-F.; Gong, X.-J.; Yu, Z.-G.; Sun, C.-B.; Zheng, C.-L.; Jiang, D.-J.; Li, Z.-N. *World J. Gastroenterol.* **2009**, *15*, 5674–5684.

(75) Li, D.; Sun, W. P.; Zhou, Y. M.; Liu, Q. G.; Zhou, S. S.; Luo, N.; Bian, F. N.; Zhao, Z. G.; Guo, M. *World J. Gastroenterol.* **2010**, *16*, 2378–87.

(76) Talanian, J.; Galloway, S.; Heigenhauser, G.; Bonen, A.; Spriet, L. *J Appl. Physiol.* **2007**, *102*, 1439–1447.

(77) Huffman, K.; Shah, S.; Stevens, R.; Bain, J.; Muehlbauer, M.; Slentz, C.; Tanner, C.; Kuchibhatla, M.; Houmard, J.; Newgard, C.; Kraus, W. *Diabetes Care* **2009**, *32*, 1678–1683.

(78) Jones, D. P. *Rejuvenation Res.* **2006**, *9*, 169.

(79) Yoshiba-Suzuki, S.; Sagara, J.; Bannai, S.; Makino, N. *J Biochem.* **2011**, *150*.

(80) Zhang, H.; Forman, H.; Choi, J. *Methods Enzym.* **2005**, *401*, 468–483.

(81) Sen, C. *Mol. Cell. Biochem.* **1999**, *196*, 31–42.

(82) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(83) Blanco, RA; Ziegler, T.; Carlson, B.; Cheng, P.; Park, Y.; Cotsonis, G.; Accardi, C.; Jones, D. *Am. J. Clin. Nutr.* **2007**, *86*, 1016–1023.

(84) Jones, DP; Mody, V. J.; Carlson, J.; Lynn, M.; Sternberg, P. J. *Free Radic. Bio. Med.* **2002**, *23*, 1290–1300.

(85) Andersson, A.; Lindgren, A.; Arnadottir, M.; Prytz, H.; Hultberg, B. *Clin. Chem* **1999**, *45*, 1084–1086.

(86) Ashfaq, S.; Abramson, J.; Jones, D.; Rhodes, S.; Weintraub, W.; Hooper, W.; Vaccarino, V.; Harrison, D.; Quyyumi, A. *J. Am. Coll. Cardiol.* **2006**, *47*, 1005–1011.

(87) Salmi, H; Leonard, J.; Rahman, S.; Lapatto, R. *Scand. J. Clin. Lab. Invest.* **2012**, *72*, 152–157.

(88) Sekhar, R.; McKay, S.; Patel, S.; Guthikonda, A.; Reddy, V.; Balasubramanyam, A Jahoor, F. *Diabetes Care* **2011**, *34*, 162–167.

(89) Gomez-Cabrera, M. C.; Domenech, E.; Vina, J. *Free Radic. Bio. Med.* **2008**, *44*, 126–31.

(90) Powers, S. K.; Jackson, M. J. *Physiol. Rev.* **2008**, *88*, 1243–1276.

(91) Zwetsloot, K. A.; John, C. S.; Lawrence, M. M.; Battista, R. A.; Shanely, R. A. *J Inflamm Res* **2014**, *7*, 9–17.

(92) Stephens, F. B.; Constantin-Teodosiu, D.; Greenhaff, P. L. *J Physiol* **2007**, *581*, 431–44.

(93) Wilkinson, D. J.; Smeeton, N. J.; Watt, P. W. *Prog Neurobiol* **2010**, *91*, 200–19.

(94) Tuchman, M.; Lee, B.; Lichter-Konecki, U.; Summar, M. L.; Yudkoff, M.; Cederbaum, S. D.; Kerr, D. S.; Diaz, G. A.; Seashore, M. R.; Lee, H. S.; McCarter, R. J.; Krischer, J. P.; Batshaw, M. L.; Urea Cycle Disorders Consortium of the Rare Diseases Clinical Research Network *Mol Genet Metab* **2008**, *94*, 397–402.

(95) Sugino, T.; Shirai, T.; Kajimoto, Y.; Kajimoto, O. *Nutr Res* **2008**, *28*, 738–43.

(96) Hickner, R. C. C.; Tanner, C. J. J.; Evans, C. A. A.; Clark, P. D. D.; Haddock, A.; Fortune, C.; Geddis, H.; Waugh, W.; McCammon, M. *Med Sci Sport. Exerc* **2006**, *38*, 660–6.

(97) Franks, P. W.; Ekelund, U.; Brage, S.; Wong, M. Y.; Wareham, N. J. *Diabetes Care* **2004**, *27*, 1187–93.

(98) Ong, M. J.; Guelfi, K. J.; Hunter, T.; Wallman, K. E.; Fournier, P. A.; Newnham, J. P. *Diabetes Metab* **2009**, *35*, 418–21.

(99) Jenkins, N. T.; Hagberg, J. M. *Med Sci Sport. Exerc* **2011**, *43*, 2231–40.

(100) Hamdy, O.; Porramatikul, S.; Al-Ozairi, E. *Curr Diabetes Rev* **2006**, *2*, 367–73.

(101) Davidsen, P. K.; Gallagher, I. J.; Hartman, J. W.; Tarnopolsky, M. A.; Dela, F.; Helge, J. W.; Timmons, J. A.; Phillips, S. M. *J. Appl. Physiol.* **2011**, *110*, 309–17.

(102) Lewis, G. D.; Farrell, L.; Wood, M. J.; Martinovic, M.; Arany, Z.; Rowe, G. C.; Souza, A.; Cheng, S.; McCabe, E. L.; Yang, E.; Shi, X.; Deo, R.; Roth, F. P.; Asnani, A.; Rhee, E. P.; Systrom, D. M.; Semigran, M. J.; Vasan, R. S.; Carr, S. a; Wang, T. J.; Sabatine, M. S.; Clish, C. B.; Gerszten, R. E. *Sci. Transl. Med.* **2010**, *2*, 33ra37.

(103) Mooy, J. M.; Grootenhuis, P. A.; de Vries, H.; Kostense, P. J.; Popp-Snijders, C.; Bouter, L. M.; Heine, R. J. *Diabetologia* **1996**, *39*, 298–305.

(104) Kuehnbaum, N. L.; Britz-McKibbin, P. *Anal. Chem.* **2011**, *83*, 8063.

(105) Psychogios, N.; Hau, D.; Peng, J.; Guo, A.; 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.; Smith, S.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.; Goodfriend, T.; Wishart, D. *PLoS One* **2011**, *6*, e16957.

(106) Winnike, J. H.; Busby, M. G.; Watkins, P. B.; O'Connel, T. M. *Am J Clin Nutr* **2009**, *90*, 1496–501.

(107) Earnest, C. P. *Med. Hypotheses* **2008**, *71*, 752–61.

(108) Wright, D. C.; Swan, P. D. *Diabetes Spectr.* **2001**, *14*, 93–97.

(109) Smith, A. E.; Walter, A. A.; Graef, J. L.; Kendall, K. L.; Moon, J. R.; Lockwood, C. M.; Fukuda, D. H.; Beck, T. W.; Cramer, J. T.; Stout, J. R. *J Inter Soc Sport. Nutr* **2009**, *6*, 5.

Chapter VI

Future Directions of MSI-CE-MS in Metabolomics: A High-Throughput Platform for Biomarker Discovery

VI. Future Directions of MSI-CE-MS in Metabolomics: A High-Throughput Platform for Biomarker Discovery

6.1 Overview of Major Thesis Contributions

The work presented in this thesis has contributed new analytical strategies for biomarker discovery in metabolomics when using CE-MS that enhances sample throughput with high data fidelity. Time-resolved metabolomic studies were performed to evaluate differential adaptive responses to high-intensity interval training and oral glucose tolerance tests among overweight/obese female subjects at risk for type 2 diabetes. The motivation for this work stems from the urgent need to validate simple yet practical lifestyle interventions that promote human health yet enable primary prevention of chronic metabolic diseases. If poor nutrition/diet and sedentary lifestyles are contributing factors to chronic disease development, then they are also part of the solution for their prevention. This is critical for reducing the socioeconomic burden of public healthcare in Canada that is no longer sustainable when relying of late-stage treatment options for disease management. Similar to other therapeutic interventions however, treatment responses between subjects are highly variable for reasons that remain poorly understood; indeed, in some cases, strenuous exercise regimes can elicit potential adverse effects among sub-groups of individuals. One of the major goals of this thesis was to develop efficient metabolomic tools for discovery of a panel of plasma biomarkers that may predict differential treatment responses to exercise training, as well as explore the putative benefits of low volume/high intensity exercise training to health status beyond glucose tolerance.

Chapter I provided an in-depth overview of the important roles that separation science plays notably in MS-based metabolomics research. CE-MS is ideal for high efficiency separation of ionic metabolites (*i.e.,* ionic metabolome) and their isomers in complex sample mixtures, however separation and ionization conditions need to be optimized to ensure adequate resolution and sensitivity. *Chapter II* introduced a new assay for selective analysis of weakly acidic estrogens, as well as their anionic conjugates when using an alkaline buffer system under negative ion mode ESI-MS. Ionization conditions when using a coaxial sheath liquid interface in CE-MS was optimized via surface response modeling. This method enabled selective analysis of intact sulfate and glucuronide conjugates of estrogen and their positional isomers in human urine, which is needed for deeper insight into the underlying mechanisms associated with estrogen dysregulation. Unlike conventional reversed-phase LC-MS methods

that have poor retention of strongly ionic metabolites, CE-MS allows for resolution of intact anionic steroid conjugates without enzyme deconjugation and complicated sample workup. However, concentration sensitivity in CE-MS was inadequate to detect sub-micromolar detection of low abundance metabolites of biological relevance, such as catechol estrogen isomers that are associated with breast cancer risk.[1] The low concentration sensitivity is associated with the small injection volumes (10 nL) used in CE-MS, which is about 100-fold lower than conventional LC-MS methods. Furthermore, post-capillary dilution effects[2] from the make-up solvent flow used in the coaxial sheath liquid interface in CE-MS, as well as ion suppression from anionic electrolytes used in the background electrolyte further reduces sensitivity, notably under negative ion mode conditions. New approaches to boost sensitivity in CE-MS are needed for deeper coverage of the human urine metabolome[3] based on pre-column derivatization,[4] on-line sample preconcentration[5] and/or low flow/sheathless interfaces.[6,7]

Although sensitivity remains a key limitation in CE-MS technology, *Chapter III* developed an innovative approach for high-throughput metabolome profiling based on multi-segment injection (MSI)-CE-MS that offers high data fidelity when performing untargeted metabolite profiling. Large-scale metabolomic studies require major investment in infrastructure and complicated data pre-processing procedures that are resource-intensive and time consuming. Alternatively, sample throughput can be enhanced by serial injection of multiple segments of sample within a single capillary since ions migrate and undergo ionization under steady-state conditions when using an isocratic buffer system and homogenous sheath liquid solution, respectively. In this case, optimization of the sample injection configuration in MSI-CE-MS was performed together with rigorous method validation of its analytical merits, such that seven or more plasma filtrate samples were analyzed simultaneously without loss in data integrity.

A major advantage of MSI-CE-MS is the ability to enhance sample throughput up to 10-fold while being able to encode information temporally based on signal pattern recognition that better streamlines data processing. For instance, a dilution trend filter was used in MSI-CE-MS to unambiguously reject spurious or background peaks that comprise the majority of detectable features when performing untargeted metabolome profiling since they have a characteristic pattern comprising 1, 2 or 7 signals in MSI-CE-MS unlike authentic sample-derived features. One of the major breakthroughs of this work was the development of a new data workflow for biomarker discovery in metabolomics

that avoids the complicated and time-consuming data pre-processing procedures. For instance, development of an "authentic peak list" comprising definitive molecular features after rigorous screening of a pooled plasma sample using a dilution trend filter (*i.e.,* "untargeted" metabolome profiling) allows for conventional "targeted" metabolome profiling when subsequently analyzing individual samples. This significantly reduces data pre-processing procedures while avoiding data over-fitting without tedious time alignment and peak picking algorithms that are subject to bias. However, one disadvantage of this data workflow is that it restricts analysis to common metabolites present in most samples, whereas unique metabolites expressed in individual subjects may be not captured during primary untargeted screening of the pooled sample due to its low abundance upon dilution.

Large-scale metabolomic initiatives require intermittent quality controls (*e.g.*, a pooled reference sample analyzed every 5-10 runs) in order to demonstrate system stability/robustness over the duration of a study to avoid false discoveries. A key advantage of MSI-CE-MS is the ability to incorporate a quality control sample within each injection sequence together with a series of different time-course samples derived from the same subject to clearly evaluate responses to stimulus and/or treatment efficacy on an individual level. In this case, rapid screening is realized when using a single replicate analysis provided that QC samples demonstrate adequate intermediate precision ($CV < 30\%$). Moreover, reliable quantification is ensured since relative changes in metabolite responses can be compared with less variation since solutes enter the ion source from six different samples within a short time interval (2-5 min) when spray is stable. *Chapter IV* and *Chapter V* of the thesis represent two-stages of a human intervention trial[8] utilizing MSI-CE-MS to investigate the potential health benefits of HIT on a cohort of healthy (*i.e.*, non-diabetic) yet sedentary/overweight women classified by a $BMI > 25$ who are at risk for type 2 diabetes. To date, the majority of HIT interventions have examined male subjects or diabetic patients in small-scale yet well-controlled supervised exercise training studies, where known metabolite (*e.g.*, glucose) or protein (*e.g.*, insulin) markers of outcome response are measured; it is not clear whether HIT has benefits for prevention of diabetes on otherwise healthy women and whether there are other unanticipated impacts on metabolism after exercise training.

In *Chapter IV*, plasma samples taken during HIT exercise sessions at the start and end of the 6 weeks intervention (naive and trained groups) were analyzed to assess the impact of HIT on response to and short-term recovery from

acute exercise. Several metabolites were found to be markers of exercise impact as well as correlated to physiological/fitness response to the exercise protocol. Hypoxanthine was identified as a marker of energetic stress in plasma and strongly correlated to subject response HIT, where it can be used to differentiate good responders from weak responders, which is important since response to exercise training is quite variable. Hypoxanthine is a key intermediate of adenine nucleotide (ATP) degradation in contracting skeletal muscle that accumulates irreversibly in plasma during prolonged high intensity exercise, notably for unfit subjects. *L*-carnitine was identified as a marker of exercise training status in resting plasma samples, differentiating naive from trained subjects and is key cofactor involved in fatty acid metabolism and their transport inside the mitochondria. Also, *O*-acetyl-*L*-carnitine was also identified as a distinct plasma marker that increased post-exercise for trained subjects indicative of improved acetylation capacity that is needed to avoid mitochondrial stress during strenuous exercise. However, further validation of these metabolites is required in an independent cohort together with qualitative identification of unknown metabolites of biological significance. In this work, each subject's adaptive metabolic response to strenuous exercise after HIT was captured within a single run when using MSI-CE-MS, with 3 pre- and 3 post-intervention samples along with a QC sample analyzed simultaneously. These dynamic metabolite responses to strenuous exercise and short-term recovery allow for deeper insight into responses to exercise training that elicit definitive gains in cardiorespiratory fitness for most subjects despite underlying between-subject variations in treatment responses.

In *Chapter V*, MSI-CE-MS was applied to determine time-dependent responses to OGTTs that may be modulated after 6-week HIT intervention, where differential metabolic changes to post-prandial glucose was evaluated for each subject in their naïve/untrained and trained state. Typically, glucose and insulin are quantified during an OGTT to assess glucose resistance and insulin sensitivity, but in this work MSI-CE-MS was used for plasma metabolome profiling to correlate a subject's individual response to HIT based on their characteristic metabolic phenotype. Overall, there was a mean benefit to glucose tolerance within individuals, although inter-subject variation was high. Most subjects did show improvements in aerobic fitness and endurance capacity outcomes as reflected by a decrease in average maximum heart rate during HIT, as well as increase in maximum workload and VO_{2max}, although these fitness benefits were not correlated with changes in body composition or glucose tolerance. Changes in the plasma metabolome were determined to be caused by both the long-term HIT

exercise intervention as well as acute glucose consumption. Circulating oxidized disulfides were decreased overall in trained subjects suggesting a reduced plasma thiol redox status and increased intra-cellular antioxidant capacity after HIT. Also, decreases in plasma branched-chain, aromatic, and sulfur-containing amino acids, as well as urea cycle and carnitine cycle intermediates, were also measured after glucose loading due to their specific uptake into peripheral tissues. A unique aspect of this work was the utilization of a panel of plasma metabolites for prediction of changes in glucose tolerance induced by exercise training. Stratification of participants was successfully accomplished using prognostic plasma markers to screen for poor and notably adverse responders to HIT.

In this work, MSI-CE-MS was used for high throughput analysis of a large number (n = 154) of samples over two days, where each OGTT (six time points) was analyzed in a single run along with an internal QC that served as a pooled reference sample. MSI-CE-MS was again used to streamline data processing via the use of a dilution filter for targeted metabolomics and to reject spurious peaks, background ions and irreproducible features. The metabolomics applications in *Chapter IV* and *Chapter V* have identified several metabolites that are related to high intensity interval training and oral glucose tolerance tests, providing insight into groups of people who have differential responses to exercise intervention and can help lead to personalized interventions for various health issues such as obesity, diabetes, and diabetes-related disorders.

6.2 Long-Term Robustness and Ruggedness of MSI-CE-MS

Although MSI-CE-MS was validated in *Chapter III* for quantitative analysis of cationic metabolites in plasma filtrate samples with good linearity, accuracy and repeatability (intra-day), further assessment of its long-term analytical performance is warranted. The robustness of MSI-CE-MS should be assessed by systematically varying different analytical parameters of the separation, specifically BGE composition (*e.g.,* pH, ionic strength, percent ACN), sample composition (*e.g.,* pH of ammonium acetate, dilution factor, plasma from different subjects), capillary characteristics (*e.g.,* length, terminus geometry, temperature), and injection conditions (*e.g.,* sample injection length, spacer injection length). This robustness study can be accomplished efficiently using an experimental design to systematically vary multiple factors at once and monitor the resulting performance, such as a Plackett-Burman design.[9] This is advantageous over varying each factor individually and separately monitoring the

change in response, which is laborious and time consuming. A Plackett-Burman design is well suited to robustness studies because it uses a reduced number of runs compared to full factorial or fractional factorial designs and is valuable for determining whether a method is robust to many changes, although the effect of each factor cannot be determined.[9] The number of factors (X) that can be evaluated at once is $X = N - 1$, where N is the number of runs.[10] The ranges evaluated should be representative of the expected, typical variability when appropriate personnel with proper glassware/equipment and training and not designed to intentionally make the experimental design fail. Two levels ($L = 2$) permits the two extremes of the range to be included with no centre points analyzed. In this way, the robustness of the method under potential extreme (but practical) conditions can be assessed to determine whether MSI-CE-MS is well suited to slight method variations.

There are two key sources of variation that have limited reproducibility in CE-MS, namely random changes in the electroosmotic flow (EOF) and reproducible coupling of the capillary to the ion source. The latter constraint is related to the importance of capillary installation into the coaxial sheath liquid sprayer assembly that depends on the quality of the capillary cut (smooth, flat edge with diamond blade cutter) and alignment of the capillary terminus in the ion source. In addition, changes in EOF result in larger variations ($CV \approx$ 1-2%) in the apparent migration time of metabolites in CE-MS relative to reversed-phase LC, which relies on the use of a pressure-driven pump for fluid transport. Although separations are performed under strongly acidic conditions results in EOF suppression, changes in capillary surface (between capillaries), temperature changes (non-temperature controlled sections of the capillary) and sample matrix effects (high salt samples) result in variations in the EOF. This can affect resolution and notably cause co-migration of fast migrating ions (*e.g.*, ornithine) with the salt front or closely migrating isomers (*e.g.*, isoleucine/leucine) when using MSI-CE-MS. Future work can explore the use of narrow fused-silica capillaries (25 μm) in order to reduce Joule heating in conjunction with temperature control of the capillary section between CE instrument and ion source using a circulating water bath. Also, recent work has demonstrated that dynamic coated capillary surfaces using layered ionic polymers can contribute to improved EOF stability and migration time reproducibility even when analyzing diluted urine samples.[11]

In order for MSI-CE-MS to be appropriate for large-scale metabolomic studies involving multiple analysts and instruments in different laboratories,

operating conditions need to be rigorously optimized and tested for ruggedness under over several weeks/months. The ruggedness of should be assessed by varying the experimental conditions, specifically the user, reagents, capillary, ESI sprayer, and CE-MS instrumentation. In addition, a thorough investigation of the inter-day reproducibility is required to determine the degree of instrumental variability and set appropriate thresholds for significant biological variability. This comprehensive assessment of the method's reproducibility is critical to its success in applications to real samples, in order to ascertain which measured variations are true metabolic changes.

6.3 Anionic Metabolome for Comprehensive Metabolomics via MSI-CE-MS

In this work, MSI-CE-MS has only been applied to the cationic metabolome when using strongly acidic conditions and positive ion mode as described in *Chapters III*, *IV*, and *V*. However, in order to expand metabolome coverage, the suitability for MSI-CE-MS as a multiplexed separation platform for the anionic metabolome needs to be investigated as well. The high buffer pH assay described in *Chapter II* is well suited for coverage of the anionic metabolome, including primary and secondary metabolites such as steroid conjugates. A limitation of the ammonium bicarbonate (pH 9.5) buffer system used in this assay is that its high conductivity restricts the voltages that can be applied and the pH that can be obtained. Analytes with high pK_a (*e.g.,* sugars, free estrogens and steroids, pK_a >12) will not be ionized and therefore not resolved in CE under these conditions. Alternatively, a low-conductivity amine buffer system can be used to provide high buffer capacity, reduce ion suppression and minimize Joule heating effects under strongly alkaline conditions (pH 12.6) as shown in **Figure 6.1.** Resolution of structural (*e.g.,* glucose, galactose, fructose, mannose) and positional (*e.g.,* glucose 1-phosphate, glucose 6-phosphate) isomers is achieved due to small differences in their pK_a using diethylamine as a volatile and low conductivity buffer system at pH 12.6. In this case, comprehensive profiling of typically neutral carbohydrates and reduced polyols, anionic sugar phosphates, cationic aminosugars is achieved simultaneously under a single condition. The major limitation of this technique is the poor sensitivity of CE-MS due to dilution at the coaxial sheath interface and poor response of many compounds in ESI in negative mode. These two high pH conditions under negative ion mode can be readily combined with the high throughput MSI-CE-MS strategy described in *Chapter III*

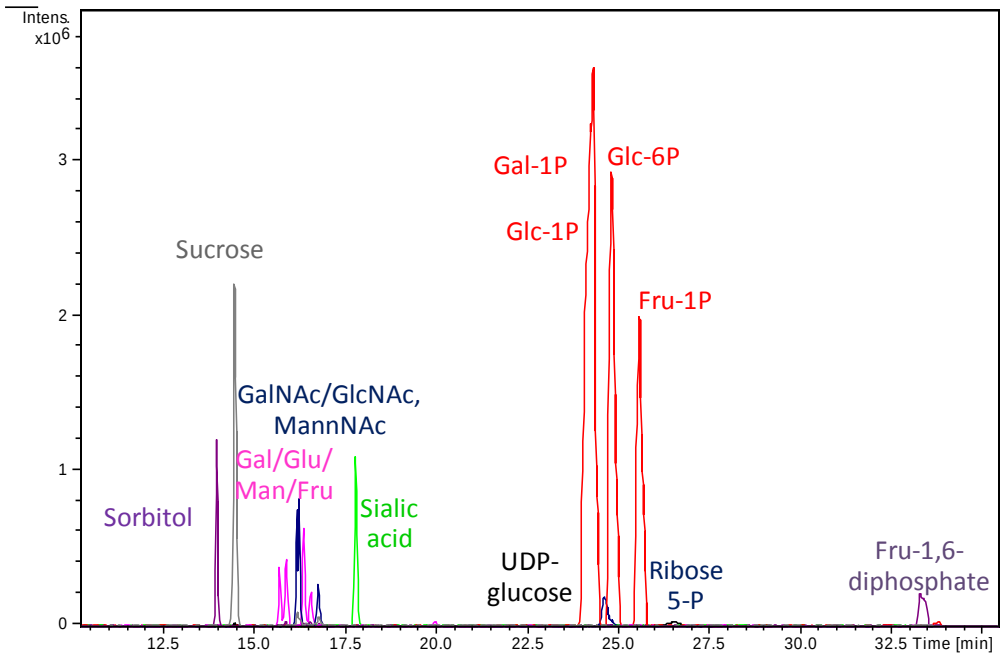


Figure 6.1. Overlay of EIEs showing resolution of various classes of compounds with a wide range of physicochemical properties and polarity, including three sets of isomers (BGE = 1 M diethylamine, pH 12.6. Structural isomers galactose (Gal), glucose (Glu) , mannose (Man), and fructose (Fru) are fully resolved, while n-acetyl galactosamine (GalNAc)and n-acetyl glucosamine (GlcNAc) are resolved from n-acetyl mannosamine (MannNAc). Positional isomers galactose 1-phosphate (Gal-1P) and glucose 1-phosphate (Glc-1P) comigrate and are resolved from their isomers glucose 6-phosphate (Glc-1P) and fructose 1-phosphate (Fru-1P).

to reduce analysis times and increase throughput for anionic metabolites in complex biological samples. When using high pH BGE the EOF is much faster, and the end of the effective analysis window will no longer be defined the late migrating EOF, resulting in a much longer separation window compared to low pH buffer conditions. The potential for late migrating chloride in biological fluids to cause ion suppression needs to be assessed to determine where it migrates relative to anions of interest and define the end of the effective analysis window.

Although a wide separation window increases peak capacity when analyzing multiple segments of samples simultaneously in a single run, major anionic interferences from matrix may result in ion suppression, such as plasma chloride and EDTA used as an anticoagulant during collection. EDTA is particularly problematic in anionic metabolite screening as it is negatively charged under the separation conditions (pH 9.5) and causes interference with the

analysis of anionic metabolites. Serum can be used as an alternative biofluid offering similar overall metabolome composition to plasma without the anticoagulant additive.[12] In addition, higher metabolite concentrations in serum compared to plasma may enable increased sensitivity and better detection of low abundance metabolites that may be of interest.[12]

For comprehensive metabolomics, information from both positive and negative modes is required to analyze cationic (under acidic BGE conditions) and anionic (under BGE alkaline) species by MSI-CE-MS, respectively. Data processing must be performed carefully to ensure that zwitterions (detected as $[M+H]^+$ and $[M-H]^-$) are identified and denoted as being the same compound to avoid redundancy that can affect statistical analyses further down the pipeline. It is common in metabolomics to combine multiple ionization modes, orthogonal separation mechanisms, and/or sampling methods to maximize metabolome coverage. By combining these two modes of MSI-CE-MS, comprehensive screening of the ionic metabolome will be achieved while maintaining high sample throughput with high data fidelity for biomarker discovery.

6.4 MSI-CE-MS for Multiple Biofluids for Biomarker Validation

Work in *Chapters IV* and *V* demonstrated the feasibility of MSI-CE-MS to be used to analyze multiple plasma samples from a single individual, including a baseline and a time course after stimulus/recovery (*i.e.* HIT trial; high glucose intake) that was repeated for each subject based on their training status (naïve/trained). The versatility of MSI-CE-MS enables metabolome profiling to be performed using various injection configurations, such as replicates, time-dependent samples, and QCs. One interesting avenue that remains to be explored with MSI-CE-MS is its applicability to analyze multiple biofluids simultaneously from the same subject within a single run. This would enable more information to be obtained about metabolism within an organism as a whole, while permitting more and/or better validated biomarkers to be identified when using multiple biofluids. In traditional multi-sample analyses, one analysis is required for each sample as has been done in metabolite profiling of plasma, hair, kidney, and liver samples in diabetic mice.[13] This enables identification of unique biomarkers that may be present in only one biofluid/tissue, confirmation of common metabolites, and corroborating information from various sources to be pooled together to form a more complete picture of the underlying metabolic processes. MSI-CE-MS can be applied, for example, to plasma, urine, and muscle tissue extract from a single

individual pre- and post-intervention (along with a pooled QC) in a single 7-segment sequence. In *Chapter IV*, hypoxanthine was identified as a key plasma marker of energetic stress notably for untrained/unfit subjects during HIT trials, as well as 20 min post-exercise when at rest. The analysis of muscle tissue extracts could confirm acute ATP consumption and irreversible IMP degradation during strenuous exercise, while urine analysis could also confirm accumulation of excreted xanthine and uric acid as end-products of nucleotide triphosphate hydrolysis. Thus, a more consistent mechanistic interpretation for the study is derived when considering the overlap in measured intra-cellular, circulating and excretory metabolic patterns involving a common pathway. It is important to note, in the development of methods using MSI-CE-MS for multiple biofluids/samples, that differences in sample composition (*e.g.*, ionic strength, pH) will affect the migration of the analytes and the overall separation. It is crucial to properly validate the reproducibility MSI-CE-MS and assess ion suppression when using multiple biofluids to ensure adequate resolution of fast migrating species, isomers, and reproducible migration times.

6.5 MSI-CE-MS with Applied Voltage Segment Spacers

The MSI-CE-MS method presented in *Chapter III* for high-throughput metabolome profiling utilized a hydrodynamic injection of a series of background buffer as physical spacers between sample injection plugs in order to displace them in space within the capillary prior to voltage application. Alternatively, voltage segments can be applied between plugs prior to the application of the final separation voltage, as previously demonstrated in multiple-injection CE with UV detection for the simultaneous analysis of up to 21 sample plugs "in series."[14,15] Although this method is capable of analyzing large numbers of samples in a single run, the major limitation is that only a single highly abundant compound can be analyzed in a simple matrix along with an internal standard for peak normalization. However, voltage segments to induce transient electrokinetic transport instead of hydrodynamic injections of buffer will allow for an increased effective capillary length without loss in peak capacity because the sample plugs will have the entire length of the capillary in which to separate under an applied voltage. This is a benefit over hydrodynamic injections, where the number of injections is limited because as the number of spacers is increased, the remaining effective capillary length is correspondingly reduced. In this case, the effective

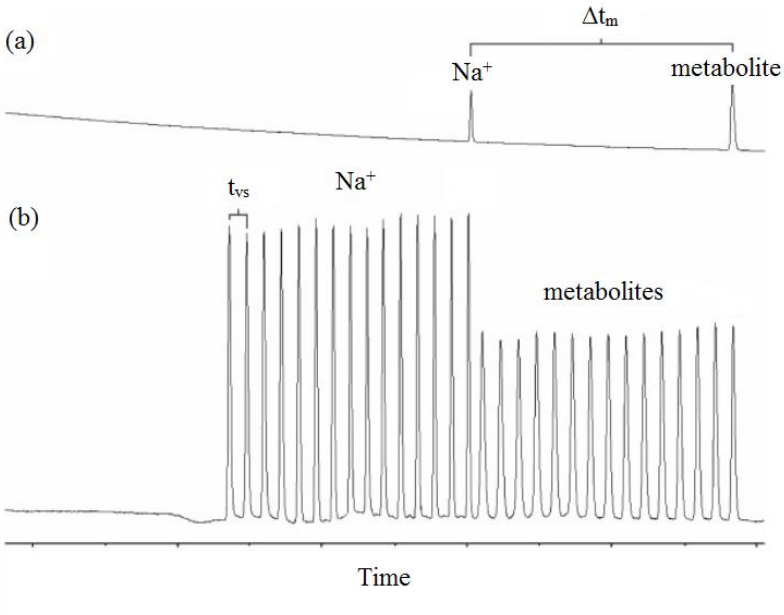


Figure 6.2. (a) Theoretical total ion electropherogram showing resolution of salt front (Na^+) and a metabolite in a single injection**, (b)** MICZE using voltage segments to partially resolve 15 sample segments from each other prior to final voltage application. The difference in migration time (Δt_m) of two analytes (*e.g.*, salt front and early migrating metabolite) in a single run can be used to optimize the number of sample injection segments and the time for each voltage segment (t_{vs}). Figure modified from Amini et al.[15]

separation window is decreased only by the addition of more samples, which is comparable to performing on-line sample preconcentration with extended sample injection times.[16] The salt front in biological/high salt samples remains a sample throughput limitation in MSI-CE-MS with voltage spacers because it will need to pass through the sample plugs and be resolved in order to prevent ion suppression. However, the number of injections can readily be maximized and optimized by first performing a single injection to determine the electrophoretic mobility of both the salt front and the earliest migrating compounds of interest. The difference in migration time (Δt_m) information can be used to methodically design the number of injections and voltage spacer time (t_{vs}) while maximizing the number of metabolites resolved from the salt front to increase metabolome coverage without ion suppression as depicted in **Figure 6.2**. Further work will examine the benefits of coupling hydrodynamic sample injections with transient electrokinetic separations in order to further boost sample throughput beyond 7 or

10 sample segments by MSI-CE-MS while still retaining high data fidelity and reproducibility.

6.6 Sheathless ESI for Increased Sensitivity

An ongoing limitation of CE-MS is its lack of sensitivity compared to LC-MS methods. As discussed in *Chapter I,* sheathless CE-MS interfaces can provide increased sensitivity due to reduced dilution via elimination of sheath liquid/nebulizer gas during spray formation. Increased sensitivity via a robust sheathless interface would permit low abundance metabolites to be reliably detected and increase the metabolome coverage and therefore increase the number of potential biomarkers that can be identified. Recent developments in sheathless CE-MS have led to more robust interfaces over past designs and will be a valuable addition to CE-MS and metabolomic profiling if they are demonstrated to have long-term stability. Several recent advances in CE-MS interface development have addressed underlying concerns regarding method sensitivity and robustness. For instance, the integration of on-line sample preconcentration together with a commercial low flow/sheathless interface for CE-MS provides an effective strategy to enhance concentration sensitivity by over two orders of magnitude while maintaining high separation efficiency relative to a coaxial sheath liquid interface as shown in **Figure 6.3**.[7] In this interface, the capillary terminus is etched with hydrofluoric acid to make the capillary wall thin and porous, and electrical contact is achieved by immersing the porous zone in a conductive buffer reservoir. However, this process produces fragile capillary tips due to the narrowed, etched capillary walls that are prone to failure and suffer from instability.[7,17] Nevertheless, this low flow ion source allows for deeper coverage of the urinary metabolome with nanomolar detection limits for most polar metabolites with over a 4-fold increase in detected molecular features (50-450 *m/z*), including low abundance ions above *m/z* 300, such as small peptides.[7]

As discussed in *Chapter I,* current development in low flow/sheathless interfaces to improve concentration sensitivity has focused on optimizing emitter tip geometry and the mechanism of electrical contact. Stable electrical contact at the capillary terminus is necessary to act as the terminal electrode and for electrospray, but remains a challenge in sheathless interfaces. One proposed CE-MS interface is a “junction-at-the-tip” design where the CE capillary is connected to a tapered, beveled metal sprayer via a flow-through micro-vial which provides electrical contact for the terminal electrode and provides a stable flow of solution

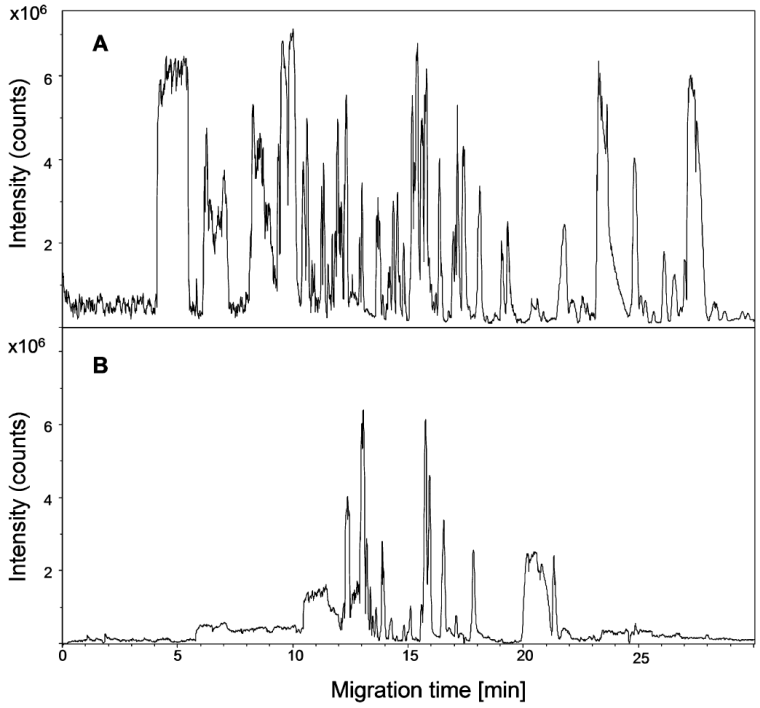


Figure 6.3. (A) Base peak electropherogram (m/z 50−450) of human urine obtained with sheathless CE-MS using a porous tip sprayer for expanded metabolome coverage. Conditions: BGE, 10% acetic acid (pH 2.2); sample injection, 2.0 psi for 30 s (~1% of capillary volume). (B) Base peak electropherogram (*m/z* 50−450) of human urine obtained with CE-MS using a coaxial sheath liquid interface. Conditions: BGE, 10% acetic acid (pH 2.2); sample injection, 0.5 psi for 30 s (~1% of capillary volume).[7]

for ESI as shown in **Figure 6.4**. This beveled tip sheathless interface is promising because it has been demonstrated to provide 2- to 13-fold increase in sensitivity (*S/N*) while being physically robust.[18] Unlike the coaxial sheath liquid interface which requires fine adjustment of the capillary position, the terminal end of the fused-silica capillary is inserted flush into the metal emitter that should improve installation. However small variations in its installation cause changes in the micro-vial volume and affect resolution due to band broadening caused by mixing. Additionally, this emitter design can only tolerate low currents in the capillary (< 15 μA), which requires less conductive buffers, narrower capillary diameters and/or lower applied voltages.

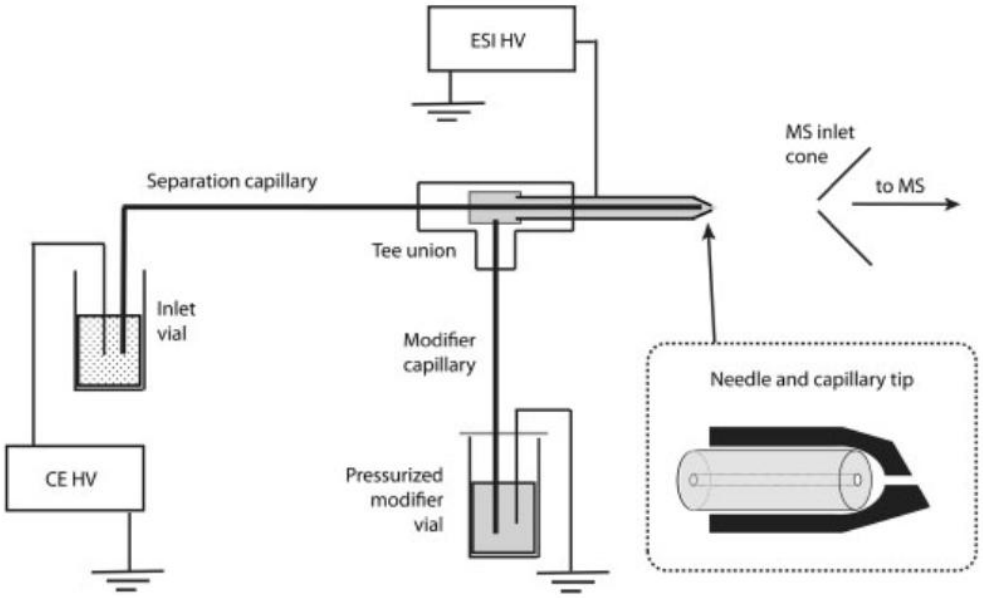


Figure 6.4. "Junction-at-the-tip" sheathless interface with a beveled tip sprayer and low flow for decreased dilution and increased sensitivity for CE-MS.[18]

In the case of plasma metabolite profiling in *Chapters IV* and *V*, the filtered list of authentic cationic metabolites comprised only about 60 unique compounds (*CV* < 40%) in pooled plasma filtrate specimens. With increased sensitivity and lowered detection limits together with anionic metabolites with negative ion mode detection, this number could be increased by 5 to 10-fold, which would greatly enhance overall metabolome coverage by CE-MS.[7] Prior to utilizing this sheathless interface for metabolomics applications, it will need to be rigorously optimized (*e.g.*, buffer composition, voltages, flow rates) for maximizing sensitivity while maintaining adequate reproducibility and robustness. If this interface can be shown to have long-term reliability between analysts and capillaries, it will be a critical factor in further establishing CE-MS as a valuable high-throughput metabolomics screening tool for biomarker discovery.

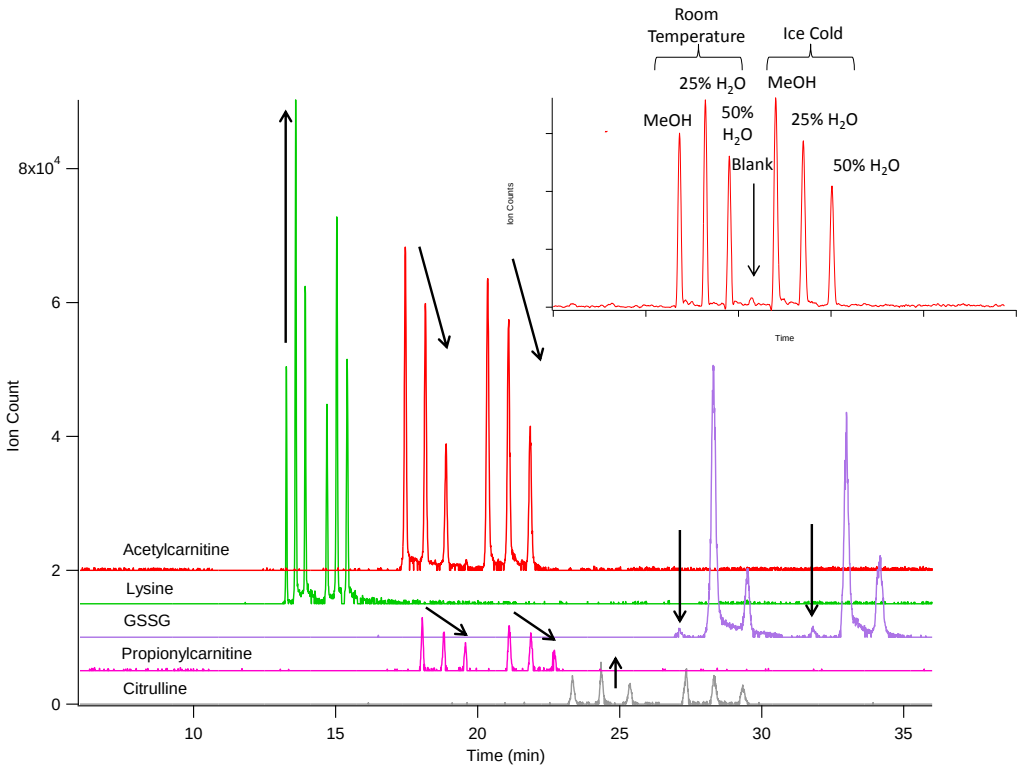


Figure 6.5. Optimization of three extraction solvents (MeOH, 25% H_2O in MeOH, 50% H_2O in MeOH), under two temperature conditions (room temperature, ice cold) for dried blood spot extracts in a single run with MSI-CE-MS. Extraction efficiency of polar metabolites (*i.e,* GSSG, lysine, citrulline) is maximized with water, whereas extraction of less polar metabolites (*i.e.* acetylcarnitine, propionylcarnitine) is maximized with higher MeOH content (unpublished data).

6.7 MSI-CE-MS for Optimization of Extraction Efficiency

An ongoing source of bias in metabolomics arises from sample preparation, particularly when extracting metabolites from a complex matrix where the extraction solvent and conditions directly affects the metabolites that will be analyzed. Optimization of sample preparation conditions is critical to both maximize extraction efficiency and reduce bias. MSI-CE-MS offers a unique way to optimize sample preparation and extraction protocols, where differences in extraction efficiency are readily observed when analyzing multiple samples with different experimental conditions in a single run. This decreases the analysis time required for optimization, while increasing data quality and decreasing variability in separation conditions. In a seven-segment injection sequence, six different experimental conditions can be monitored along with a blank as shown in

Figure 6.5. Extraction efficiency depends on the physicochemical properties of the analytes and a suitable sample preparation must be carefully designed to reduce bias and maximize metabolome coverage. Extraction of polar metabolites, such as amino acids, is improved with water content in the solvent, particularly for GSSG which shows poor recovery with pure MeOH as an extraction solvent. Extraction of less polar compounds, such as carnitines, is maximized with high MeOH content, and therefore a compromise must be made in solvent in order to maintain metabolome coverage of metabolites with varying physicochemical properties and polarities.

6.8 Personalized Health via Lifestyle Interventions for Disease Prevention

The HIT intervention study discussed in *Chapters IV* and *V* was small, consisting of 16 female participants from whom not all plasma samples were able to be collected for the HIT exercise (n = 9) and/or the OGTTs (n = 11) due to various problems occurring during blood collection. The 6 week intervention was short and long-term effects after HIT training on metabolite profile and physiological parameters were not followed up in this investigation. In particular, there was no significant difference in body composition, VO_{2max}, or metabolic profile found between the fasted and fed groups, and it would be valuable to confirm this finding in a larger study with more participants over a longer intervention period. In this case, blood samples (and other biological specimens, such as muscle tissue and 24 hr urine) can be taken at various points throughout intervention to monitor adaptive responses to training and determine if there is a point at which metabolic changes are dampened or attenuated as the subject adjusts to and becomes accustomed to the strenuous exercise routine. In addition, it would be beneficial to determine whether these physiological and metabolic changes persists after the trial is completed in order to determine whether HIT training needs to be continued to maintain health benefits. Besides exercise training, a simpler lifestyle intervention for high risk subjects that would ensure good compliance would only comprise changes in diet/nutrition.

The metabolomics of HIT has not yet been investigated in diverse populations. The work presented in this thesis focused on moderately overweight/obese female subjects. Expanding a future study to include men, a broader age group, and subjects recently diagnosed with pre-diabetes (yet not receiving medication, such as metformin) would introduce more variation in the population. This subject heterogeneity will provide more factors that can aid in

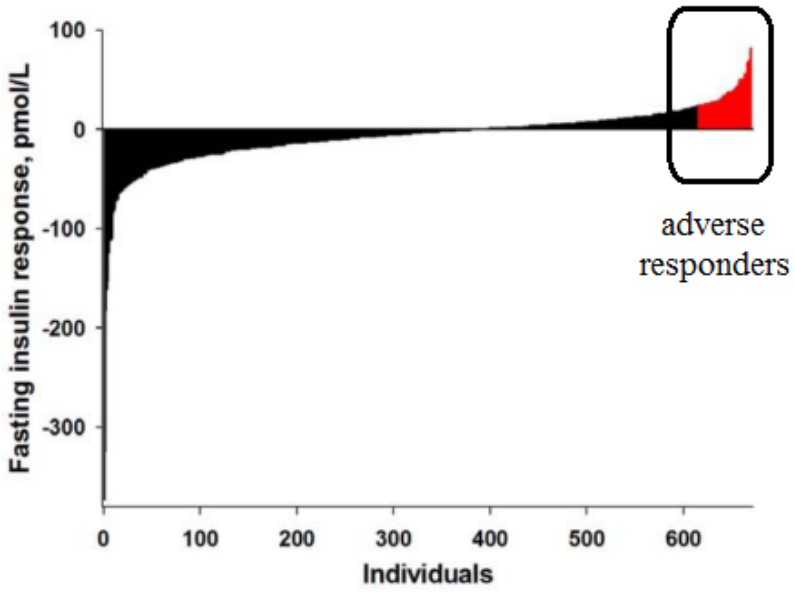


Figure 6.6. Distribution of fasting insulin responses in the Heritage study where adverse responders (8.3%) to exercise showed increased fasting insulin response after a 20-week endurance training program.[19]

classification of differential responses to HIT training to aid in the development of customized lifestyle interventions for obesity and diabetes-related disorders. For personalized metabolomics to be effective, both the factors that affect an individual's response (*e.g.*, metabolic profile, physiological measurements) and the desired health/clinical outcome (*e.g.*, adipose tissue loss, insulin sensitivity) need to be identified and considered. To tailor HIT as a personalized intervention, the exercise must be optimized for different individuals based on their characteristic blood metabolic phenotype (*i.e.*, metabotype), where different training variables are adjusted to best suit the individual, such as maximum heart rate level, interval duration and/or nutritional supplementation.

Metabolomics can be used to probe the responses of individual subjects to strenuous exercise and/or nutrition intervention and lead to personalized protocols that promote human health and ultimately prevent chronic disease development. This is of particular interest because several recent studies have indicated that a small group of the population (8-14%) constitute adverse responders to various

types of exercise training (**Figure 6.6**) and in this case, exercise may have the opposite health effect than that which is desired.[19,20] It is critical to tailor interventions to account for these adverse responders, for whom the prescribed exercise may actually do more harm than good. The discovery of a panel of metabolites that can serve as prognostic biomarkers of responses to HIT intervention would allow for patient stratification in order to guide treatment decision-making processes that ultimately avoids potential adverse effects while enhancing overall treatment efficacy.

6.8.1 Protein Supplementation for Retention of Muscle Mass During Periods of Energy Restriction

There are many strategies for the therapeutic treatment of obesity including behaviour modification via dietary and physical activity interventions, and pharmacologic intervention.[21] The behavioural intervention for obesity involves a caloric deficit over extended periods of time which can be obtained through decreasing the ingested calories and/or increasing calories burned during exercise. Energy restriction often leads to altered whole-body protein metabolism and loss of skeletal muscle due to insufficient protein intake and a negative muscle protein balance.[22] Increasing or maintaining lean muscle mass helps to offset the declines in metabolic rate that are observed during periods of caloric restriction.[22] Increased protein consumption has been shown to reduce the mass of muscle lost during these periods of energy restriction.[23,24] Recent studies have demonstrated that whey to be more effective than casein or soy in muscle protein synthesis and the maintenance of skeletal muscle due to its specific amino acid composition.[24,25] The high-throughput MSI-CE-MS platform using an accelerated workflow for biomarker discovery can be used in a current obesity intervention study to analyze multiple biofluids simultaneously. This work is currently in progress together with Dr. Phillips' laboratory in the Department of Kinesiology, where plasma and urine samples have already been collected, processed, and stored, awaiting analysis for metabolic profiling. In the 2-week nutritional intervention, overweight/obese adults (n = 42) were grouped into three groups: whey protein supplement, soy protein supplement, and maltodextrin control. The objective of the study was to assess the effect of protein supplementation (50 g/day) during periods of caloric restriction (750 kcal/day) on changes in body composition (fat and lean muscle mass) and identify metabolites associated with these changes. The efficacy of the intervention is assessed using intra- and inter-subject variation

by comparing pre- and post-intervention plasma and urine samples and correlating subjects' metabolic profiles to physiological measurements (*e.g.*, BMI, body composition). It is hypothesized that metabolites related to protein metabolism (*e.g.*, amino acids, creatinine, urea) will be elevated during protein supplementation and metabolites related to obesity (*e.g.*, glucose, lactate, free fatty acids, choline, amino acids, creatine) will be modulated as fat mass is lost. Metabolites that are positively or negatively correlated with muscle or fat mass retention can be used to classify participants based on their response to the nutritional intervention. Specifically, the two protein supplements (whey, soy) may be more effective at maintaining muscle mass in different populations, aiding in the development of personalized interventions for obesity.

6.8.2 Biomarker Discovery for Stroke Classification and Intervention

MSI-CE-MS can be applied for stroke biomarker discovery using a modified 7-segment injection approach for reliable detection of metabolites that are not consistently detected across all samples. In this design, three samples are injected in 3 different dilution patterns as shown in **Figure 6.7**. Two samples are injected at two levels (typical preparation ("2"), 2-fold diluted preparation ("1") and one sample is injected twice without dilution ("2") along with a blank in characteristic dilution patterns. These three characteristic dilution patterns can be used to identify which segment a certain metabolite originated from in cases where compounds identities and corresponding relative migration times are unknown. This approach, using paired samples with definitive dilution pattern for pattern recognition is of particular importance when a metabolite is not present in all samples and it is challenging to identify its source. In the stroke study, 2 subjects under different conditions (stroke patient at onset of stroke and seven days after hospitalization, along with health age-matched control) can be analyzed in a single run as a quick primary screen to detect differentially expressed metabolites in three different samples. Stroke patients presenting with different types of stroke (*i.e.* ischemic, hemorrhagic) can be classified based on their metabolite profiles which can lead to personalized interventions for stroke. Classification of stroke type is critical to determining the appropriate treatment but diagnosis is often ambiguous when using current diagnostic techniques (*e.g.*, CT scan, MRI), resulting in incorrect treatments and adverse side effects.[26,27] The determination of stroke biomarkers would be an invaluable tool in stroke

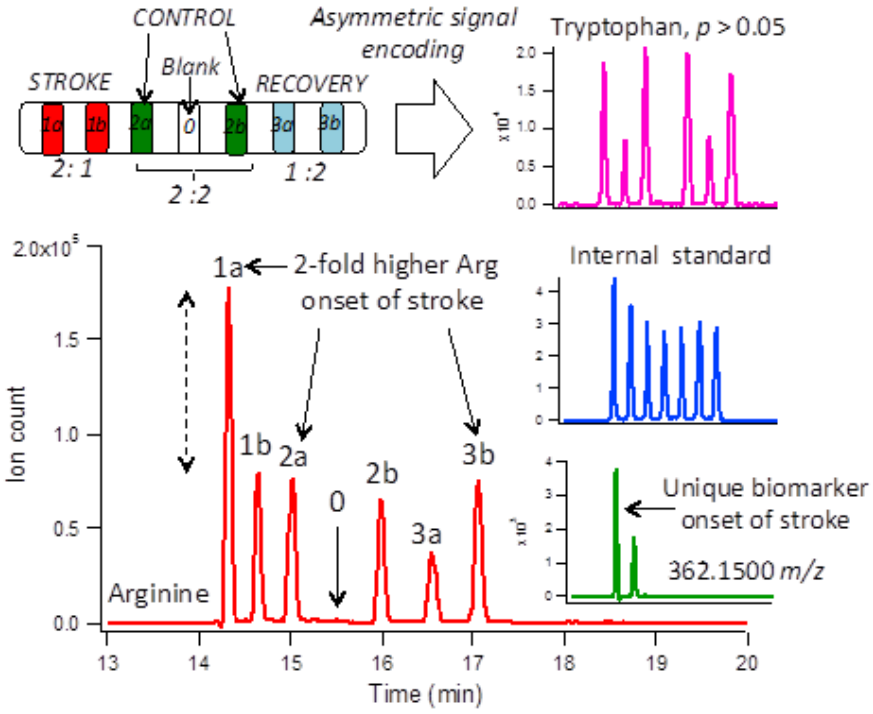


Figure 6.7. Three samples (1, 2, 3) are injected in different characteristic dilution patterns (2:1, 2:2, 1:2) along with a blank sample (0). Pattern recognition based on these asymmetric dilution profiles is used to identify metabolites that are not present across all samples and/or whose identity and relative migration are unknown (unpublished data).

intervention as patients' metabolite profiles could be screened upon hospital admission in order to classify their stroke and establish an appropriate course of treatment.

In summary, multiplexed separations based on MSI-CE-MS offers a revolutionary approach for biomarker discovery that provides high sample throughput and high data fidelity without added infrastructure/operating costs. To the best our knowledge, no other instrumental technique offers such a diverse array of benefits in terms of analytical and bioinformatic merits within a cost-effective format that we anticipate will transform metabolomics research.

6.9 References

(1) Kuehnbaum, N. L.; Britz-McKibbin, P. *Anal. Chem.* **2011**, *83*, 8063.

(2) Mokaddem, M.; Gareil, P.; Belgaied, J.-E.; Varenne, A. *Electrophoresis* **2009**, *30*, 1692.

(3) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. *PLoS One* **2013**, *8*, e73076.

(4) D'Agostino, L. A.; Lam, K. P.; Lee, R.; Britz-McKibbin, P. *J. Proteome Res.* **2011**, *10*, 592–603.

(5) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403–415.

(6) Maxwell, E. J.; Zhong, X.; Zhang, H.; van Zeijl, N.; Chen, D. D. Y. *Electrophoresis* **2010**, *31*, 1130.

(7) Ramautar, R.; Busnel, J.-M.; Deelder, A. M.; Mayboroda, O. A. *Anal. Chem.* **2012**, *84*, 885–892.

(8) Gillen, J. B.; Percival, M. E.; Ludzki, A.; Tarnopolsky, M. A.; Gibala, M. J. *Obesity* **2013**, *21*, 2249–2255.

(9) Swartz, M.; Krull, I. *LCGC North Am.* **2006**, *24*, 480–490.

(10) Plackett, R. L.; Burman, J. P. *Biometrika Trust* **1946**, *33*, 305–325.

(11) Ramautar, R.; Nevedomskaya, E.; Mayboroda, O.; Deelder, A.; Wilson, I.; Gika, H.; Theodoridis, G.; Somsen, G.; de Jong, G. *Mol. BioSyst.* **2011**, *7*, 194–199.

(12) 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.; Wang-Sattler, R. *PLoS One* **2011**, *6*, e21230.

(13) Tsutsui, H.; Maeda, T.; Min, J. Z.; Inagaki, S.; Higashi, T.; Kagawa, Y.; Toyo'oka, T. *Clin. Chim. Acta* **2011**, *412*, 861.

(14) Lodén, H.; Pettersson, C.; Arvidsson, T.; Amini, A. *J. Chromatogr. A* **2008**, *1207*, 181–185.

(15) Amini, A.; Lodén, H.; Pettersson, C.; Arvidsson, T. *Electrophoresis* **2008**, *29*, 3952–3958.

(16) Lee, R.; Britz-McKibbin, P. *Anal. Chem.* **2009**, *81*, 7047.

(17) Settlage, R. E.; Russo, P. S.; Shabanowitz, J.; Hunt, D. F. *J. Microcolumn Sep.* **1998**, *10*, 281–285.

(18) Maxwell, E. J.; Zhong, X.; Zhang, H.; van Zeijl, N.; Chen, D. D. Y. *Electrophoresis* **2010**, *31*, 1130.

(19) Bouchard, C.; Blair, S. N.; Church, T. S.; Earnest, C. P.; Hagberg, J. M.; Häkkinen, K.; Jenkins, N. T.; Karavirta, L.; Kraus, W. E.; Leon, A. S.; Rao, D. C.; Sarzynski, M. A.; Skinner, J. S.; Slentz, C. A.; Rankinen, T. *PLoS One* **2012**, *7*, e37887.

(20) Boule, N. G.; Weisnagel, S. J.; Lakka, T. A.; Tremblay, A.; Bergman, R. N.; Rankinen, T.; Leon, A. S.; Skinner, S. J.; Wilmore, J. H.; Rao, D. C.; Bouchard, C. *Diabetes Care* **2005**, *28*, 108–114.

(21) Shea, J.; Diamandis, E. P.; Sharma, A. M.; Despres, J.-P.; Ezzat, S.; Greenway, F. *Clin. Chem.* **2012**, *58*, 968–973.

(22) Pasiakos, S. M.; Vislocky, L. M.; Carbone, J. W.; Altieri, N.; Konopelski, K.; Freake, H. C.; Anderson, J. M.; Ferrando, A. A.; Wolfe, R. R.; Rodriguez, N. R. *J. Nutr.* **2010**, *140*, 745–751.

(23) Deibert, P.; König, D.; Schmidt-Trucksaess, A.; Zaenker, K. S.; Frey, I.; Landmann, U.; Berg, A. *Int. J. Obes. Relat. Metab. Disord.* **2004**, *28*, 1349–1352.

(24) Westerterp-Plantenga, M. S.; Nieuwenhuizen, A.; Tomé, D.; Soenen, S.; Westerterp, K. R. *Annu. Rev. Nutr.* **2009**, *29*, 21–41.

(25) Tang, J. E.; Moore, D. R.; Kujbida, G. W.; Tarnopolsky, M. A.; Phillips, S. M. *J. Appl. Physiol.* **2009**, *107*, 987–992.

(26) Furie, K. L.; Maas, M. B. *Biomark. Med.* **2009**, *3*, 363.

(27) Saenger, A. K.; Christenson, R. H. *Clin. Chem.* **2010**, *56*, 21–33.