

**Fundamental Studies of Ionization Response
And Newborn Screening By CE-ESI-MS**

**FUNDAMENTAL STUDIES OF IONIZATION RESPONSE AND
NEW STRATEGIES FOR NEWBORN SCREENING OF INHERITED
METABOLIC DISORDERS BY CE-ESI-MS**

By

KENNETH ROSS CHALCRAFT, BSc

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfillment of the Requirements

For the Degree

Master of Science

McMaster University

©Copyright by Kenneth Chalcraft, September 2008

MASTER OF SCIENCE
McMaster University
(Chemistry)
Hamilton, Ontario

TITLE: Fundamental Studies of Ionization Response and New
Strategies for Newborn Screening of Inherited Metabolic Disorders
by CE-ESI-MS

AUTHOR: Kenneth Ross Chalcraft, BSc (Brandon University)

SUPERVISOR: Dr. Philip Britz-McKibbin

NUMBER OF PAGES: 100

ABSTRACT

CE-ESI-MS has become a powerful analytical tool capable of simultaneous identification and quantification of many classes of biologically relevant molecules. For studies in metabolomics, CE-ESI-MS offers a unique platform which will allow for the systematic elucidation of unknown metabolites in complex matrices without the need for complex sample preparation steps required with other techniques. In this thesis, a novel theoretical prediction model which will allow the estimation of detector response in ESI-MS is outlined. This response model will allow researchers to quantitatively predict relative ionization efficiency of compounds based on proposed two-dimensional structures without the need for a purified standard. Another feature of this model is that it can be applied to complex biological samples without the need for off-line sample pretreatment. Also in this thesis, a novel neonatal screening method will be presented which will aid clinical chemists in determining the presence of inborn metabolic disorders. This screening method which aims to compliment current protocols will allow health care professionals to further assess dried blood spot samples by providing simultaneous separation, identification, and quantification of relevant metabolites. This method also offers an alternatives to other protocols in place which are necessary to measure acid labile compounds which cannot be assessed by standard screening techniques.

ACKNOWLEDGEMENTS

First and foremost, I would like to express sincere gratitude to my extremely able supervisor Dr. Philip Britz-McKibbin for all he has done for me over my time in his research group. Through his guidance, support, and patience I have learned a great deal about not only science, but about myself. I would also like to thank Dr. Brian McCarry for his guidance and input which greatly aided my research. As well I would also like to express my gratitude towards my labmates, Richard Lee, Jennilee Gavina, Adam Ptolemy, Giselle Sigui-Lines, and Cassandra Mills for their advice, their support and above all their friendship.

Finally I would like to thank my parents, David and Betty, and my wonderful fiancé Amber for always being there for me through anything. Without your sacrifices and unwavering encouragement I would not be here today and I love you all.

TABLE OF CONTENTS

Abstract.....	III
Acknowledgements.....	IV
Table of Contents.....	V
List of Figures.....	IX
List of Tables.....	XI
List of Abbreviations.....	XII
CHAPTER 1.....	1
Background	
1.1 Chemical Separations and Detection Methods.....	1
1.2 Capillary Electrophoresis.....	3
1.2.1 History of Capillary Electrophoresis.....	4
1.2.2 Electrokinetic Phenomenon in CE.....	5
1.2.3 Electrophoretic Mobility.....	5
1.2.4 Electroosmotic Flow.....	6
1.2.5 Apparent Mobility.....	8
1.3 Chemical Detection in CE.....	9
1.3.1 A Brief History of ESI-MS.....	10
1.3.2 Mechanisms of ESI.....	11
1.3.3 History of Ion Trap Mass Analyzers.....	13
1.3.4 3D Ion Trap Mass Analyzer.....	13
1.4 Capillary Electrophoresis Electrospray Ionization Mass Spectrometry..	15
1.4.1 CE with On-Line Sample Preconcentration and Desalting.....	17
1.4.2 Dynamic pH Junction.....	19

1.5 Inherited Metabolic Disorder Screening.....	21
1.5.1 History of Newborn Screening.....	21
1.5.2 Biomarkers and Targeted Metabolic Screening.....	22
1.5.3 Bioassay and Tandem Mass Spectrometry.....	23
1.6 Research Objectives.....	24
1.7 References.....	25
CHAPTER 2.....	27
Prediction of Ionization Response for Unknown Metabolites by CE-ESI-MS.....	27
Abstract.....	27
2.1 Introduction.....	28
2.2 Experimental.....	31
2.2.1 Chemicals and Reagents.....	31
2.2.2 Apparatus and Conditions.....	32
2.2.3 Measurement of Solute Physiochemical Parameters.....	34
2.2.4 Method Calibration	36
2.2.5 Prediction of Solute Physiochemical Parameters.....	37
2.2.6 Data Processing and Multivariate Analysis.....	38
2.2.7 RBC Lysate Preparation and Recovery Experiments.....	39
2.3 Results and Discussion.....	41
2.3.1 Multivariate Model for Solute Ionization in ESI-MS.....	41
2.3.2 Relative Response Factor in ESI-MS.....	44
2.3.3 PCA Modeling and PLS Regression.....	46
2.3.4 Model Validation and Recovery Experiments.....	48
2.3.5 Virtual Quantification without Chemical Standards.....	49

2.4 Conclusions.....	51
2.5 References.....	52
 CHAPTER 3	
Inherited Metabolic Screening of Newborns by CE-ESI-MS.....	59
Abstract.....	59
3.1 Introduction.....	60
3.2 Materials and Methods.....	63
3.2.1 Chemicals and Reagents.....	63
3.2.2 Blood Spot Collection and Sample Preparation.....	63
3.2.3 Apparatus and Conditions.....	64
3.2.4 Calibration and Validation.....	65
3.3 Results and Discussion.....	67
3.3.1 Separation Optimization.....	67
3.3.2 Sample Preparation and Analysis of Dried Blood Spot Extracts.....	68
3.3.3 Method Calibration and Detection Limits.....	70
3.3.4 Method Validation.....	72
3.4 Conclusion.....	74
3.5 References.....	75
 CHAPTER 4	
Future Outlook.....	84
4.1 Research Plans.....	84
4.2 Expansion of Model Set and Predicting Anionic Compounds in Negative-Ion Mode.....	85
4.3 Impact of Various Solvents and Other Additives to Ionization Matrix...86	

4.4 Investigation of Additional Biomarkers for Metabolic Disorder Screening.....	86
4.5 Analysis of Blood Spot Extraction Protocol.....	87
4.6 Analysis of Samples Taken from Diagnosed Patients.....	87
4.7 Simulation of metabolic profiles in diseased patients.....	87
4.8 References.....	88
Appendix I Model Cationic Metabolites Used as Testing Set.....	89
Appendix II Y-matrix values used in PLS model.....	96
Appendix III X-matrix of ten validation metabolites where parameters were predicted by computer modeling.....	99

LIST OF FIGURES

- Fig. 1.1** Debye-Hückel-Stern model of the electric double layer showing (a) the rigid and diffuse layers and (b) zeta potential (ζ) that generates the EOF at the capillary wall under an external voltage.7
- Fig 1.2** Comparison of solute transport mechanisms in separations using (a) flat flow profile of EOF generated by external voltage in CE and (b) parabolic flow profile generated by a external pressure in HPLC.....7
- Fig 1.3** Taylor cone formation and aerosol formation of charged droplets in ESI.....12
- Fig 1.4** Schematic of 3D ion trap mass analyzer.....14
- Fig 1.5** Schematic of CE-ESI-ITMS instrument depicting (a) the complete system and (b). coaxial sheath liquid ESI interface16
- Fig 1.6** Schematic showing separation of a mixture of ions based on their effective charge by CE: (a) injection, (b) zonal separation at pH >2, (c) zonal separation at pH>2 with EOF and (d) sample electropherogram under neutral pH conditions (e) zonal separation at pH<2 with suppressed EOF.....18
- Fig 1.7** Time-resolved electropherograms showing the dynamics of in-capillary Trp focusing with a dynamic pH junction from experimental data and computer simulations using Simul 5.0 (insets) showing a sample plug of 12.5 cm placed at a) 7.5 cm from distal end of capillary, b) 25 cm, and c) 40 cm.....20
- Fig 2.1** Multivariate model for describing gas-phase ion formation in ESI-MS.....54
- Fig 2.2** Overlay of extracted ion electropherograms depicting a 700-fold difference in ionization response at equimolar concentrations of 30 μ M for the initial model set of metabolites by CE-ESI-MS.....55
- Fig 2.3** (a) 2D scores plot from PCA highlighting the relationship of 40 different cationic metabolites in terms of their four measured physicochemical properties (MV , $\log D$, μ_o and z_{eff}). (b) Linear correlation plot ($y = 0.792x + 0.0040$, $R^2 = 0.9144$) between measured relative response factor (RRF) by CE-ESI-MS and predicted RRF using PLS regression of the metabolite test set. (c) 2D loadings plot that highlights the relative weight of each physicochemical parameter on metabolite location in the scores plot.....56
- Fig 2.4** (a) Overlay of extracted ion electropherograms showing 5 μ M of test (red traces-selected) and validation (blue traces) metabolites spiked in RBC lysates using CE-ESI-

MS. (b) Virtual calibration curves generated *de novo* for validation metabolites from their predicted fundamental physicochemical parameters using PLS regression.....57

Fig 2.5 Recovery experiments summarizing validation studies by CE-ESI-MS of (a) 15 test metabolites and (b) 10 validation metabolites spiked in RBC lysates at low, mid and high concentrations levels using virtual calibration curves predicted by PLS regression.58

Fig 3.1 A series of extracted ion electropherograms depicting the impact of organic solvent and solute concentration on metabolite peak shape and resolution in CE-ESI-MS, where: a) long-chain carnitines without organic modifier, b) with organic modifier, c)selective leucine resolution at various ACN concentrations.....78

Fig 3.2 Extracted ion electropherograms depicting a normal profile of amino acids and acylcarnitines detected from filtered dried blood spots(a) analyzed directly after ultracentrifugation(b) the extract was evaporated and reconstituted.....79

Fig 3.3 Extracted ion electropherogram depicting recovery studies for a filtered dried blood spot extract spiked at 12.5% maximum calibration level with twenty different amino acid and acylcarnitine metabolites.....80

LIST OF TABLES

Table 1.1 Decision criteria and upper/lower cut-off limits based on normal population controls for selected IEMs targeted in newborn screening programs by ESI-MS/MS.....	24
Table 3.1 Normal concentration levels of targeted amino acids and acylcarnitines detected in dried blood spots by CE-ESI-MS without and with solvent evaporation.....	81
Table 3.2 External calibration data for 20 targeted amino acids and acylcarnitines relevant to newborn screening of IEMs using CE-ESI-MS without chemical derivatization.....	82
Table 3.3 Validation of CE-ESI-MS for accurate and reproducible quantification of amino acid and acylcarnitine metabolites derived from dried blood samples.....	83

LIST OF ABBREVIATIONS

3D	Three Dimensional
4-OH-Pro	4-hydroxyproline
5-oxo-pro	Pyroglutamic Acid
A	Adenosine
AC	Alternating Current
ACN	Acetonitrile
Allo-Leu	Allo-Leucine
APCI	Atmospheric Pressure Chemical Ionization
APPI	Atmospheric Pressure Photo Ionization
An	Adenine
Arg	Arginine
At	Atenolol
B-Ala	beta-Alanine
BGE	Background Electrolyte
C0	Carnitine
C2	Acetyl-Carnitine
C3	Propionyl-Carnitine
C4	Butyryl-Carnitine
C5	Pentoyl-Carnitine
C8	Octanoyl-Carnitine
C14	Myristoyl-Carnitine
C16	Palmitoyl-Carnitine
Carn	Carnosine
CE-ESI-MS	Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry
Cit	Citrulline
Cl-Tyr	3-Chlorotyrosine
CV	Co-efficient of Variance
Cyst	Cystathionine
CZE	Capillary Zone Electrophoresis
σ_B	Background Noise Level
DC	Direct Current
DopN	Dopamine
diAla	Alanyl-N-Alanine
ϵ	Relative Permittivity
E	Electric Field Strength
EOF	Electroosmotic Flow
FDBSE	Filtered Dried Blood Spot Extract
G	Guanosine
GABA	gamma-Aminobutyric Acid
GC	Gas Chromatography
Gln	Glutamine

Glu	Glutamic Acid
Gn	Guanine
GSSG	Glutathione (oxidized)
GSH	Glutathione (reduced)
HCy	Homocystine
His	Histidine
HisN	Histamine
HPLC	High Performance Liquid Chromatography
IEM	Inborn Errors of Metabolism
Ile	Isoleucine
IMD	Inherited Metabolic Disorder
IS	Internal Standard
IT	Ion Trap
Ka	Acidity Constant
Leu	Leucine
LIF	Laser Induced Fluorescence
LOD	Limit of Detection
LogD	Octanol/Water partition co-efficient
LOQ	Limit of Quantitation
Lys	Lysine
μ_o	Absolute Electrophoretic Mobility
μ_{ep}	Electrophoretic Mobility
μ^A_{ep}	Apparent Electrophoretic Mobility
μ_{EOF}	Electrophoretic Mobility of Electroosmotic Flow
<i>m</i>	Sensitivity
Meta	Metoprolol
Me-Asp	N-Methylaspartic Acid
Me-His	N-Methylhistidine
Me-A	Methyladenosine
MetS	Methionine Sulfone
MeOH	Methanol
Met	Methionine
MM2	Molecular Mechanics 2
MSUD	Maple Syrup Urine Disorder
m-Tyr	meta-Tyrosine
MV	Molecular Volume
η	Viscosity
NMR	Nuclear Magnetic Resonance
NAm	Nicotinamide
NA	Nicotinic Acid
NO ₂ -Tyr	3-Nitrotyrosine
Otet	Oxytetracycline
Orn	Ornithine
o-Tyr	ortho-Tyrosine

PABA	para-Aminobenzoic Acid
PCA	Principle Component Analysis
Phe	Phenylalanine
PKU	Phenylketouria
PLS	Partial Least Squares
Pro	Proline
p-Tyr	para-Tyrosine
Q_{eff}	Effective Charge
R	Hydrodynamic Radius
RBC	Red Blood Cell
RIR	Relative Ionization Response
RPA	Relative Peak Area
RMT	Relative Migration Time
RRF	Relative Response Factor
SerN	Serotonin
Thea	Theanine
Trp	Tryptophan
TrpN	Tryptamine
Tyr	Tyrosine
TyrN	Tyramine
UV	Ultraviolet
v	Velocity
Val	Valine
ζ	Zeta Potential
Z_{eff}	Effective Charge

CHAPTER 1: Background

1.1 Chemical separations and detection methods

Accurate quantification of analytes contained within complex sample mixtures often depends on high efficiency separation techniques for their resolution prior to detection. For all separation systems, there exist two distinct processes which fundamentally influence analyte migration behavior, namely separative transport and dispersive transport.¹ Dispersive processes which cause “analyte band broadening” occur due to the natural tendency of molecules to diffuse across a concentration gradient, as well as the phenomena of mass transfer resistance and convection. All of these processes additively contribute random motion that counteracts separation. In contrast, separative transport processes differentially influence the velocity or movement of an analyte, which is associated with its unique physiochemical properties under specific conditions, such as boiling point or hydrophobicity for reverse phase LC. Understanding of the various mechanisms influencing separative transport is of fundamental importance to analytical separations, which ultimately influence the resolution of complex sample mixtures. Since dispersion tends to have a negative impact on separative transport, it can be stated that the main goal of any separation method is to maximize the extent of separative transport while minimizing dispersive forces.

In addition to the need to develop an efficient separation format, a suitable method of detection must be employed which ideally is both selective for the analyte of interest yet sensitive enough to generate a reliable signal over background noise. Detector selectivity

is the ability of a detection system to discriminate analytes of interest from other interferences in a sample, such as solvent, salts and co-migrating solutes². Detector sensitivity meanwhile is a measure of the change in response intensity as a function of analyte concentration. Other important detector attributes include its compatibility to the sample type or separation conditions without matrix effects that alter detector performance, while providing adequate reliability with good precision. Indeed, off-line sample pretreatment (*e.g.*, chemical labeling, desalting, sample enrichment *etc.*) is often critical to the success of an assay by transforming the sample in order to make it more compatible with the detector. Ultimately, a fundamental understanding of the mechanism of detection is also a key feature for enhancing analytical performance. The latter feature will represent a major theme in this thesis (Ch. 2) in regards to the feasibility for accurate prediction of analyte relative ion response when using capillary electrophoresis with electrospray ionization-mass spectrometry (CE-ESI-MS).

Common separation techniques widely used today primarily include high performance liquid chromatography (HPLC) and gas chromatography (GC). In general, each separation technique has distinct limitations under different contexts, such as the incompatibility to analyze involatile or thermally labile analytes in the case of GC to overall poor separation efficiency when using conventional HPLC. Separation methods that require minimal *off-line* sample pretreatment without complicated sample handling are preferred since it reduces total analysis times as well as long-term costs, while improving assay precision and accuracy. CE represents an alternative separation technique that combines the high separation efficiency of GC with the selectivity of

HPLC that is ideal for the separation of polar metabolites. Noteworthy, on-line sample preconcentration with desalting by CE provides a convenient platform for direct analysis of low abundance metabolites in biological samples with minimal sample handling. This latter feature will represent a major theme throughout this thesis, notably for the development of rapid and inexpensive strategies for newborn screening of inherited metabolic disorders, which will be presented in Ch. 3 of the thesis.

1.2 Capillary Electrophoresis

Capillary electrophoresis (CE) is one of the most versatile separations techniques available due to its high separation efficiency, diverse separation modes and wide selectivity for separating different classes of analytes ranging from discrete metal ions to whole microbes³. These features are a direct result of the unique separation mechanisms operative in CE, where electrokinetic (mobility) and thermodynamic (equilibria) driving forces can be used to induce separation. In addition, due to the narrow bore capillary dimensions used for separations in CE, heat which is generated by ion migration can be efficiently removed in order to reduce band broadening while permitting faster separations under high electric field strengths. Moreover, since the primary transport process in CE is electroosmosis within a fused-silica capillary, axial diffusion and mass transfer resistance effects are kept to a minimum due to the flat profile of the electroosmotic flow (EOF) thereby ensuring high separation efficiencies.

1.2.1 History of Capillary Electrophoresis

The term electrophoresis is generally defined as the differential movement of charged ions in an electric field. In 1897, Kohlrausch⁴ first described his theory pertaining to electrophoretic ion migration in an electric field, however it was Tiselius⁵ in 1930 who provided the first example of an experimentally successful electrophoretic separation performed on a protein mixture. Near the end of that decade, Svenson⁶ recognized the potential applications of electrophoresis for determining the isoelectric point of proteins, which later developed to what is now known as zone electrophoresis. While many researchers at this time were quick to note the advantages of electrophoresis, a fundamental problem remained Joule heating until the introduction of anticonductive polymer gels as support media, such as polyacrylamide. In 1967 Hjertén^{7,8} developed methods of separating small molecules in free solution with the use of small diameter quartz tubes, which allowed for adequate cooling of the separation medium. Several years after Hjertén's work, Mikkers *et al.* reported using Teflon tubes⁹ which were rotated in order to dissipate heat more efficiently. While this strategy did in fact improve heat dissipation efficiency during electrophoresis, the setup was too complex in design. Later, in 1981 Jorgenson *et al.* first reported the use of CE based on the use of inexpensive open tubular fused-silica capillaries for the analysis of dansylated-amino acids using laser-induced fluorescence (LIF) detection¹⁰. In this seminal work, Jorgenson demonstrated that fused-silica capillaries could provide a format for high efficiency and rapid electrophoretic separations of charged analytes under high voltages (up to 30 kV) without significant Joule heating.

1.2.2 Electrokinetic phenomenon in CE

Separations in CE are controlled by two distinct electrokinetic processes, namely the electrophoretic mobility of an ion (μ_{ep}) and electroosmotic flow (EOF) of the bulk solution. μ_{ep} is dependent on the physiochemical properties of an ion in an electric field, whereas the EOF is a natural electrokinetic pumping mechanism, which is associated with the properties of the capillary surface and the electrolyte solution.

1.2.3 Electrophoretic mobility (μ_{ep})

μ_{ep} is defined as the ion velocity (v) per unit of electric field strength (E) in a defined solution. For most low molecular weight metabolites whose shape can be approximated by a sphere, μ_{ep} is dependent on the effective charge (Q_{eff}) to hydrodynamic radius (R) ratio, as well as the viscosity of the buffer solution (η) at a specific temperature¹¹ as defined by the Hückel equation given by:

$$\mu_{ep} = \frac{v}{E} = \frac{Q_{eff}}{6 \cdot \pi \cdot \eta \cdot R} \quad (1)$$

Overall selectivity in CE is based primarily on differences in μ_{ep} which represents a characteristic parameter of a solute that is a function of the specific electrolyte properties of a solution. In the case of weakly ionic analytes, buffer pH represents the most important property used to optimize separations in CE by altering Q_{eff} . Other electrolyte properties of a solution that can be used to modify the apparent μ_{ep} as a way to enhance resolution, include buffer type, ionic strength, organic solvent modifiers, as well as discrete additives in the buffer, such as cyclodextrins. For instance, the introduction of charged micelles as migrating pseudo-stationary phases in CE by Terabe *et. al* in 1984¹²

greatly expanded the separation of neutral solutes similar to reverse-phase HPLC. Unlike HPLC that uses columns with covalently-immobilized stationary phases, the selectivity of the separation in CE can be easily modified by changes in the solution properties, such as additive concentration and additive type. Indeed, multiple additives in free solution can be used to optimize the separation of complex samples mixtures in CE¹³, where resolution is controlled by several electrokinetic and thermodynamic parameters.

1.2.4 Electroosmotic Flow

If a capillary with an ionizable surface is filled with an electrolyte solution, an EOF may be generated upon application of an external voltage. By definition, the EOF is the bulk movement of the solution through the capillary due to the generation of a zeta potential at the capillary wall/solution interface. Since most capillaries used in CE are composed of fused-silica, weakly acidic silanol moieties can acquire a negative charge at the wall/solution interface as determined by the properties of the electrolyte solution. According to the Debye-Hückle-Stern model¹⁴, ionization at the surface can generate an electric double layer consisting of an adsorbed (Stern layer) and mobile layer (double layer) of ions (predominately counter-ions) within a short distance of the negatively charged capillary wall. as depicted in *Fig. 1.1*. When an electric field is applied, the cations of the diffuse layer are attracted to the cathode or repelled by anode in the case of CE-ESI-MS as the ESI interface is electrically grounded. Since these ions are solvated, their electromigration causes the bulk transport of the solvent that represents the EOF. However, if the pH of the BGE is sufficiently acidic (pH < 2), ionization of the silanol

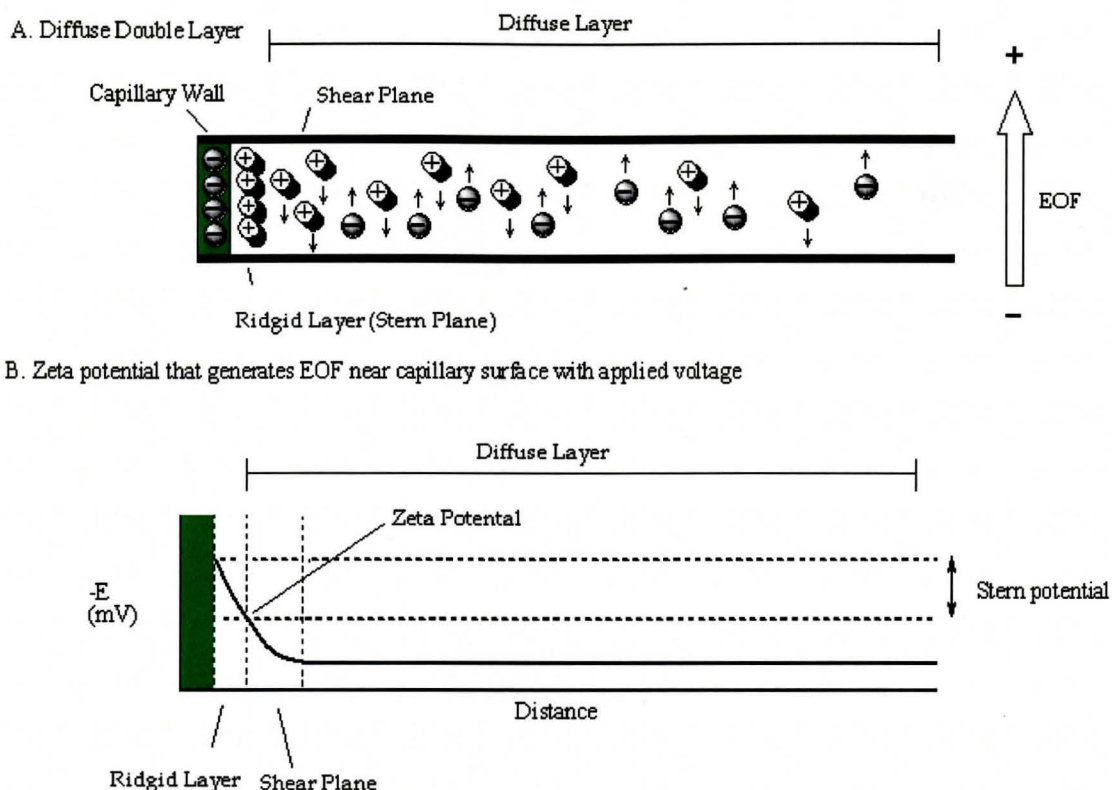


Figure 1.1. Debye-Hückel-Stern model of the electric double layer showing (a) the rigid and diffuse layers and (b) zeta potential (ζ) that generates the EOF at the capillary wall under an external voltage. The magnitude of ζ depends of nature of the capillary surface and properties of the electrolyte solution.

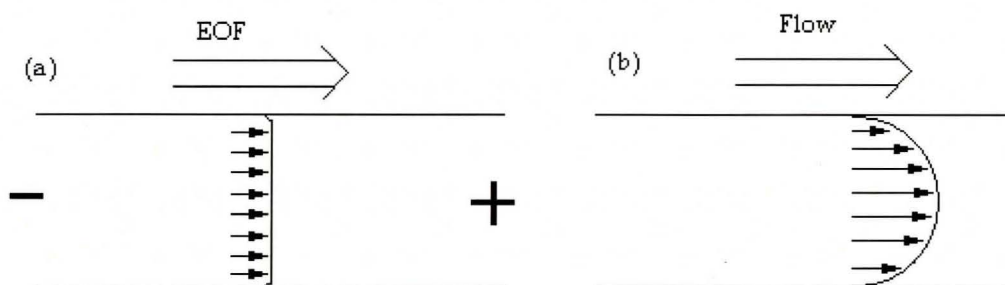


Figure 1.2. Comparison of solute transport mechanisms in separations using (a) flat flow profile of EOF generated by external voltage in CE and (b) parabolic flow profile generated by a external pressure in HPLC

moieties is effectively suppressed, thereby greatly reducing the magnitude of EOF. In general, the magnitude and direction of the EOF is dependent on the chemical properties of the capillary surface, as well as the electrolyte properties of the solution, which includes the zeta potential (ζ), the relative permittivity of the medium (ϵ) and the solution viscosity (η) as defined below:

$$\mu_{EOF} = \frac{\epsilon \cdot \zeta}{6 \cdot \pi \cdot \eta} \quad (2)$$

Unlike typical pressure-driven pumping mechanisms, EOF has a unique flat profile when occurring in capillaries that have an internal diameter less than 200 μm . This is in direct contrast to pressure-driven flow profiles normally used in HPLC, which typically exhibit increased band broadening due to axial solute diffusion due to the parabolic flow, as shown in *Figure 1.2*.

1.2.5 Apparent Mobility

In free solution CE, two electrokinetic forces determine the apparent mobility of an ion (μ_{ep}^A), which is based on the vector sum of μ_{ep} and μ_{EOF} :

$$\mu_{ep}^A = \mu_{ep} + \mu_{EOF} \quad (3)$$

In general, μ_{ep} of an ion can be readily determined experimentally [10] in an electropherogram by measuring both the apparent migration time of an analyte (t^A) and the migration time of the EOF (t_{EOF}) of a neutral marker by re-arrangement of *eq. (3)*:

$$\mu_{ep} = \frac{L_t \cdot L_d}{V} \left(\frac{1}{t^A} - \frac{1}{t_{EOF}} \right) \quad (4)$$

where, L_t is the total capillary length, L_d is the effective length to the detector and V is the separation voltage with units in cm^2/Vs . It is important to note the μ_{ep} is dependent on the specific solution properties used in the experiment, such as buffer type, ionic strength, buffer pH, solution viscosity and temperature.

1.3 Chemical detection in CE

To date, three major detection formats are commonly used in CE, namely optical, electrochemical and MS detection. UV absorbance is by far the most common detection format in CE due to its low cost, wide applicability, moderate sensitivity and overall reliability. However, it is not applicable to certain classes of analytes lacking chromophoric moieties (*e.g.*, monosaccharides) without chemical derivatization, as well as lacking selectivity when analyzing complex sample mixtures. The latter property ultimately depends on the separation efficiency and peak capacity of the separation format in order to avoid co-migrating analytes that cause spectral interferences¹⁵. CE with laser-induced fluorescence represents one of the most sensitive detection formats for quantifying trace amounts of low abundance analytes¹⁶. However, due to the lack of many intrinsically fluorescent solutes, chemical labeling with a suitable fluorophore is most often required, which can further complicate analyses due to multiple labeling and poor yield of adduct formation that is analyte-dependent. Electrochemical detection represents a sensitive and complimentary detection mode in CE for electroactive analytes that lack intrinsic chromophores. Similar to optical detection methods, electrochemical detection is limited to certain classes of analytes that suffer from poor selectivity when analyzing complex sample mixtures. An alternative to optical and electrochemical

methods in CE is MS detection, which allows for the detection of a much wider array of analytes with high selectivity. The latter feature is ultimately dependent on the type of ionization source, as well as the resolution of the mass analyzer. To date, most CE-MS methods have used electrospray ionization (ESI) as a compatible and reliable interface for solute ionization,¹⁷ although some groups have also investigated the use of alternative sources, such as atmospheric pressure chemical ionization (APCI) and photoionization (APPI).

1.3.1 A Brief History of ESI-MS

ESI is an ideal ionization technique for coupling liquid separation methods to MS due to its ability to efficiently ionize polar solutes and handle a continuous flow of liquid effluent. The first electrospray experiments reported in the literature were performed by Chapman in 1937,¹⁸ when he reported detecting gas phase ions as a result of spraying an electrified salt solution from a metal tube. Over thirty years later, Dole reported in 1968 “molecular beams of macroions” being observed from a polymer solution that had been electrosprayed using a simple ion repeller grid¹⁹. Indeed, this was the first report of a multiply charged gas phase ion. Afterwards, Iribarne and Thompson reported in 1976 evidence supporting the “ion evaporation” model as a possible mechanism for the production of gas phase ions from an electrospray source²⁰. Subsequently, Fenn *et al.* published several keynote papers describing the “free-jet” dispersion of ion clusters in a vacuum²¹⁻²³ which proposed an alternative model to ion evaporation. In the following year, Fenn *et al.* reported one of the very first practical electrospray interfaces²⁴ which could be directly coupled to liquid chromatography and this work contributed greatly to

the explosive growth of LC-MS based methods in analytical chemistry. For his work in ESI-MS, Fenn was awarded a share of the Nobel Prize in Chemistry in 2002.

1.3.2 Mechanisms of ESI

Remarkably, a fundamental understanding of the ionization mechanism operative in ESI is still unclear despite the explosive growth of ESI-MS and LC-ESI-MS in analytical laboratories world-wide. To date, two contentious models have been developed to explain how gas phase ions are generated from charged droplets in ESI-MS. The ion evaporation model was first introduced by Iribarne and Thompson²⁰ in 1974. In this theory, molecules which have been chemically ionized by addition or loss of a charged ion (typically a proton) are continuously transported through a hollow metal tube or ESI cone which is kept at a high DC potential. The intense electric field which is present at the ESI cone has two effects on ions in solution. In the case of a positive-ion MS mode, anions attracted towards the cone are electrochemically neutralized, whereas cations remain in the liquid-phase and repelled away from the ESI emitter tip. The electrohydrodynamics of this accelerating positively-charged liquid solution cause it to take the shape of a Taylor cone, until it disperses into a charged aerosol due to electrostatic repulsion forces that overcome the surface tension of the solution as depicted in *Fig. 1.3*. During the lifetime of the charged droplets, some of the ions within the surface phase desorb from the surface of the droplet into the gas-phase as the solvent evaporates under elevated temperatures. This gradual process of ion desorption from the charged droplets produces free gaseous ions, which can be

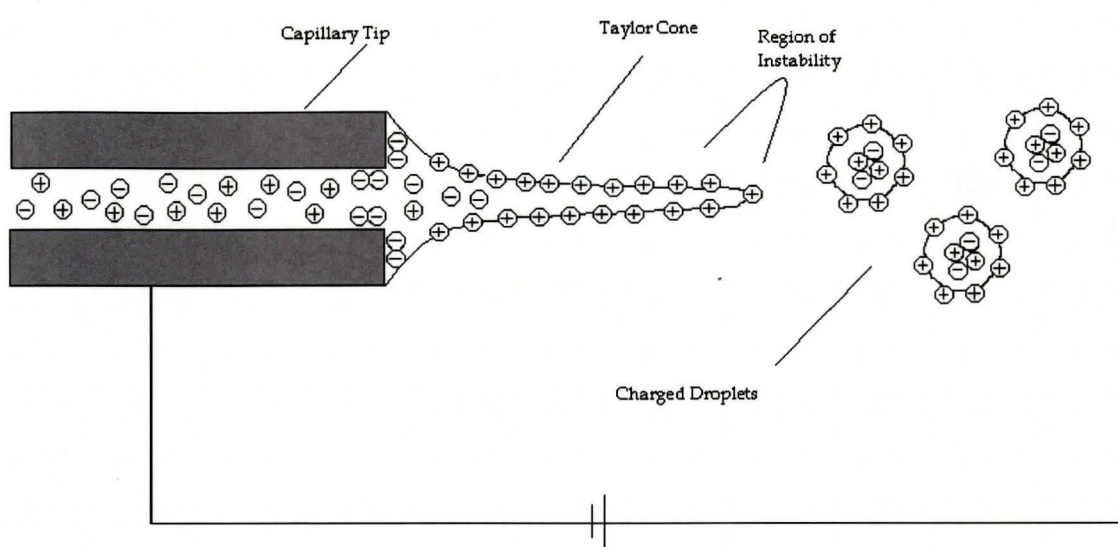


Figure 1.3. Taylor cone formation and aerosol formation of charged droplets in ESI

introduced to the mass analyzer for resolution and detection. The second mechanism for ESI proposed by Fenn²¹ is commonly known as either the droplet jet fission or Colombic fission model. In this mechanism, it is proposed that repulsive effects caused by the increasing surface charge density as a result of solvent evaporation offsets the solution surface tension resulting in droplet fission. This fission process continues to generate increasingly smaller charged droplets until a desolvated gas-phase ion is remaining, which can then be sampled by the mass analyzer. Although neither of the two mechanisms are universally accepted as being correct, in most cases both ionization processes can be operative in ESI, where the specific mechanism is dependent on the ESI conditions (*e.g.*, solution composition, flow rate, applied voltage *etc.*) and analyte physicochemical properties. However, regardless of which mechanism is dominant, neither model has been adequate for quantifying or predicting ionization responses based

on fundamental physicochemical properties of analytes, which will form the basis of the work described in *Ch. 2* of this thesis.

1.3.3 History of Ion Trap Mass Analyzers

Both varieties of the ion trap (IT) mass analyzer (3D and linear) were developed simultaneously with the standard quadrupole mass analyzer in the 1950's by MS pioneer Paul, [25] who was awarded a Nobel Prize in physics in 1989 for his contributions to this field. Shortly after his work, another MS pioneer, Dehmelt²⁶, first reported the storage of gas phase ions for radio frequency spectroscopy using an IT. This research was continued for many years using the IT to study gas-phase ion reactivity until a major design breakthrough was made by Stafford,²⁷ which permitted the first practical IT-MS instrument for use by the non-MS specialist. Stafford's IT design consisted of a damping gas to stabilize ion trajectories for improving mass resolution, as well as the use of ion scanning (ejection) via a mass instability mode, which became key features to the first commercially available 3D IT-MS instruments.

1.3.4 3D Ion Trap Mass Analyzer

A 3D quadrupole ion trap represents a low resolution, inexpensive, compact and versatile mass analyzer with excellent full-scan sensitivity that can be used to perform multi-stage MS (MS^n) experiments. These features are important for identification of low abundance yet unknown metabolites based on characteristic fragmentation losses associated with its specific molecular structure. The principles behind how the 3D IT mass analyzer operates are very similar to the standard quadrupole system where a radio

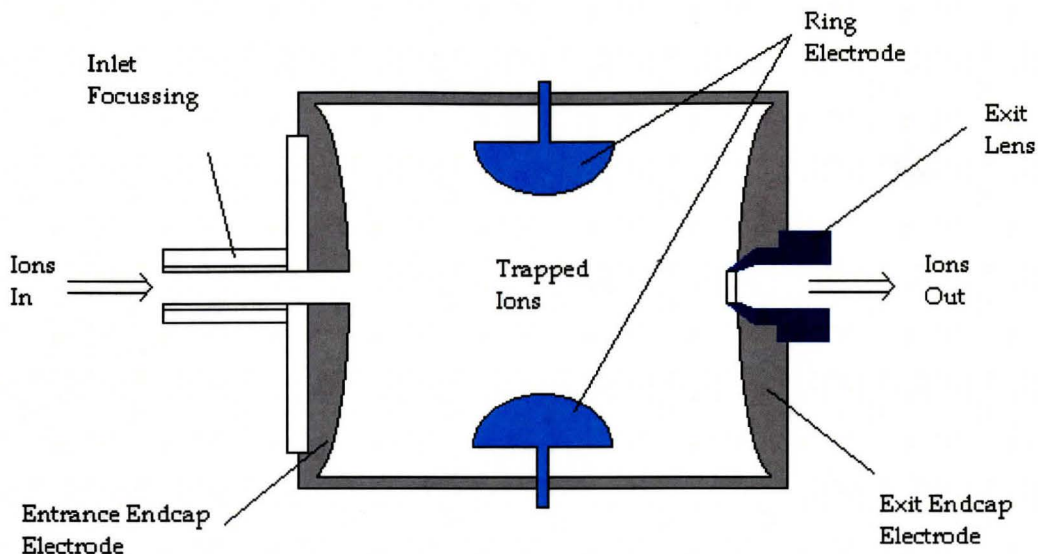


Figure 1.4 Schematic of 3D ion trap mass analyzer with a parabolic ring and two endcap electrodes, where an AC voltage is used to trap, excite/fragment, scan and eject gas-phase ions.

frequency-driven alternating current (AC) voltage is used to trap, select, excite and eject specific ions of certain m/z in a quadrupole electric field²⁸. The major difference between the two types of quadrupole mass analyzers is the geometry of the electrodes, where IT-MS instruments utilize electrodes consisting of a hyperbolic ring electrode with two end cap electrodes as depicted in *Fig. 1.4*.²⁹ Unlike quadrupole mass analyzers that function essentially as mass filters for permitting the transmission of a specific ion with a stable ion trajectory, 3D ITs first stabilize and collect all ions entering the mass analyzer simultaneously. By varying the frequency and magnitude of the AC field on the electrodes with either collisional-induced cooling or excitation (*i.e.*, fragmentation) of ions by use of a damping gas (*e.g.*, He), ions can be selectively isolated, excited and ejected prior to detection. A notable feature of IT-MS instruments is the ability to perform multi-stage MS^n experiments starting with selection of a precursor ion followed

by isolation of one or more product ions generated upon fragmentation. This permits qualitative identification of unknown ions via elucidation of their characteristic molecular fragments that comprise the chemical structure. Since IT-MS is fundamentally a low resolution mass analyzer, high efficiency separation methods are often required to expand method performance by minimizing ion-ion “space-charge” effects that can lower sensitivity and mass resolution when analyzing complex sample mixtures. Thus, optimization of separation conditions for resolving abundant co-ions, salts and other interferences in a sample prior to ESI ionization is critical to ensure accurate, sensitive and reliable analyses by IT-MS of low abundance analytes.

1.4 Capillary Electrophoresis-Electrospray Ionization-Mass spectrometry

A schematic of the instrumentation configuration of CE-ESI-MS is depicted in *Fig. 1.5*. Although the set-up required for CE is rather simple in design, modification of the ESI interface is required when coupled to MS. In CE-ESI-MS, the inlet end of a narrow bore fused-silica capillary having an internal diameter of 50 μm is immersed in a buffer reservoir after being first conditioned in acid/base, methanol and buffer solution. The distal end of the capillary extends out of the CE instrument and it is threaded through a coaxial sheath flow ESI interface³⁰, where the capillary is surrounded by two concentric stainless steel bores that stabilize ESI spray formation by flow of a sheath liquid and nebulizer gas as shown in *Fig. 1.5b*. After rinsing the capillary with a buffer solution under high pressure, the inlet end of the capillary is immersed in a sample

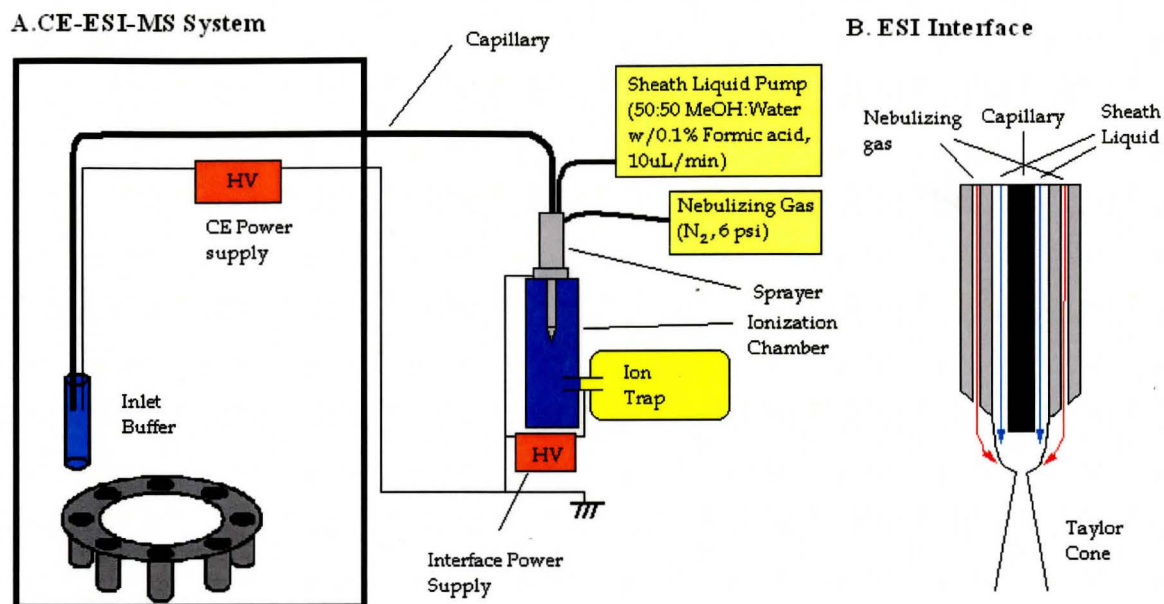


Figure 1.5 Schematic of CE-ESI-ITMS instrument depicting (a) the complete system and (b) coaxial sheath liquid ESI interface

reservoir and a sample injection is performed by either the application of a positive pressure (hydrodynamic) or the application of an electric potential (electrokinetic) for a fixed time period. Following sample injection, the inlet end of the capillary is re-immersed in a buffer reservoir containing an inert platinum electrode with application of the separation voltage (typically 20-30 kV) within the inlet buffer reservoir. Since the ESI interface is electrically grounded, the CE separation potential is isolated from the voltage applied to the ESI electrode and MS instrument. The ESI needle is contained at atmospheric pressure often at elevated temperatures to improve desolvation and solute ionization efficiency. Gaseous ion sampling into the mass analyzer is performed orthogonal to the ESI electrode with the application of a N_2 curtain gas to reduce background noise. Overall, optimization of signal response in CE-ESI-MS is a complicated process that is dependent on several operational parameters, including cone

voltage, sheath liquid flow rate, nebulizer gas flow rate, sheath liquid composition, as well as other parameters. One major advantage of the sheath liquid ESI interface is the ability to optimize solution composition to maximize ionization efficiency independent of buffer conditions used for separation. Despite the design of alternative ESI interfaces for CE-MS, such as sheathless and nanospray interfaces,³¹ coaxial sheath liquid microspray interfaces generally provide more reliable and robust long-term performance. In general, sheath liquid/gas flow rates are minimized to ensure stable ESI responses since they tend to deteriorate separation performance in terms of efficiency and sensitivity via siphoning and post-capillary dilution effects, respectively. However, these detrimental effects can be readily overcome by appropriate design of separation conditions in CE as will be discussed in the next section.

1.4.1 CE with On-line Sample Preconcentration and Desalting

The most common operating mode for CE is known as capillary zone electrophoresis (CZE), where a single continuous buffer is used for the separation. Under most conditions, the typical migration order for separations in free solution CE is cationic > neutral (unresolved) > anionic solutes, which elute at increasing migration times when detected by ESI-MS as depicted in *Figure 1.6*. The latter case assumes a normal voltage polarity and neutral/alkaline buffer conditions that induces a strong EOF, which is often an order of magnitude greater than solute μ_{ep} , thereby driving all species towards the detector. In contrast, when operating under highly acidic conditions (pH < 2), then the EOF is suppressed and most weakly acidic ions are undissociated. Under these conditions, only cations (basic or zwitter-ionic analytes) have positive μ_{ep} and thus are

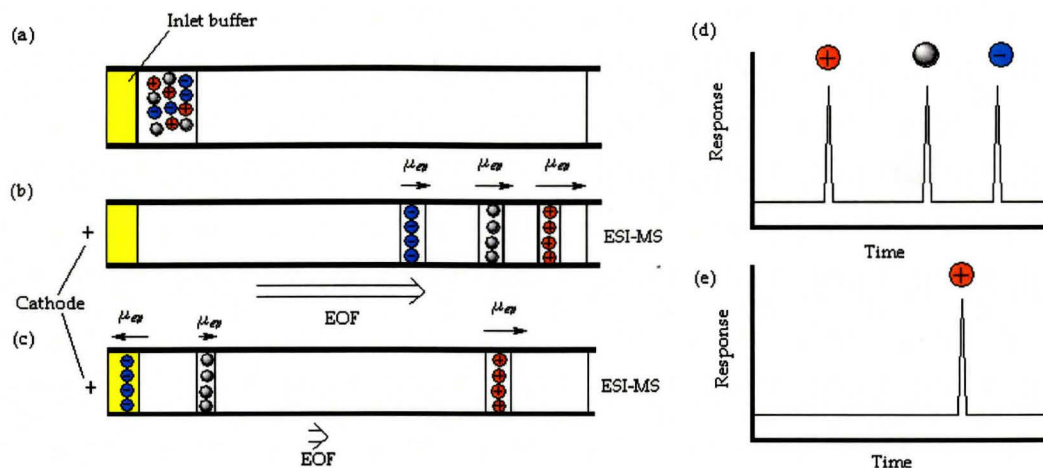


Figure 1.6 Schematic showing separation of a mixture of ions based on their effective charge by CE: (a) injection, (b) zonal separation at $\text{pH} > 2$, (c) zonal separation at $\text{pH} > 2$ with EOF and (d) sample electropherogram under neutral pH conditions (e) zonal separation at $\text{pH} < 2$ with suppressed EOF

detected by ESI-MS, whereas neutral analytes co-migrate with the slow EOF and anionic solutes (*e.g.*, strong electrolytes such as Cl^-) migrate out of the capillary inlet entirely. While front-end separation techniques are used primarily to enhance selectivity by resolving isobaric/isomeric analytes or abundant co-ions prior to ionization, CE can provide several other benefits to greatly enhance ESI-MS performance. Recently, Lee *et al.*³² demonstrated that on-line sample preconcentration with sample desalting can be performed by CE-ESI-MS for sensitive analysis of low abundance metabolites in complex biological samples. These features permit direct analysis of samples without complicated sample handling while avoiding ionization suppression effects. The latter property is also extremely important towards the development of quantitative ESI models for analyte ionization applicable to real-world samples, which will be demonstrated in *Ch. 2* of this thesis. Recently, a comprehensive review comparing different on-line preconcentration techniques in CE-MS has recently been published³³. Under optimum

conditions, enhancement in concentration sensitivity by two or three-orders of magnitude can be readily achieved by CE directly in-capillary without loss in separation performance while using conventional buffers and instrumentation. This provides a simple, inexpensive and effective format for on-line sample enrichment unlike conventional off-line sorption methods (*e.g.*, solid-phase extraction) that more than compensates for reduced sensitivity when using standard coaxial sheath liquid microspray interfaces in CE-ESI-MS.

1.4.2 Dynamic pH Junction

Dynamic pH junction represents an effective on-line preconcentration method in CE-ESI-MS since the μ_{ep} of weakly ionic analytes is strongly influenced by local buffer pH conditions when using volatile buffer pH junctions. *Fig. 1.7* depicts the mechanism of dynamic pH junction by illustrating three major steps of electrokinetic focusing after voltage application using an extremely long yet dilute sample plug of Trp. It is important to note that the sample (pH = 7) and background electrolyte (BGE, pH = 1.8) consist of a different volatile buffer system with defined pH. After hydrodynamic injection of a long sample plug that comprises up to a quarter of the capillary length, electrokinetic focusing of Trp occurs at the back-end of the BGE-sample interface since Trp has a large and positive μ_{ep} in the formic acid, pH 1.8, whereas it is electrically neutral in the sample zone. In essence, a migrating pH boundary moving across the sample zone results in preconcentration of the analyte into a narrow zone counter to diffusion. Indeed, under these conditions, about half of the capillary distance is used for sample enrichment until the buffer discontinuity is fully dissipated, when normal zonal electrophoresis and

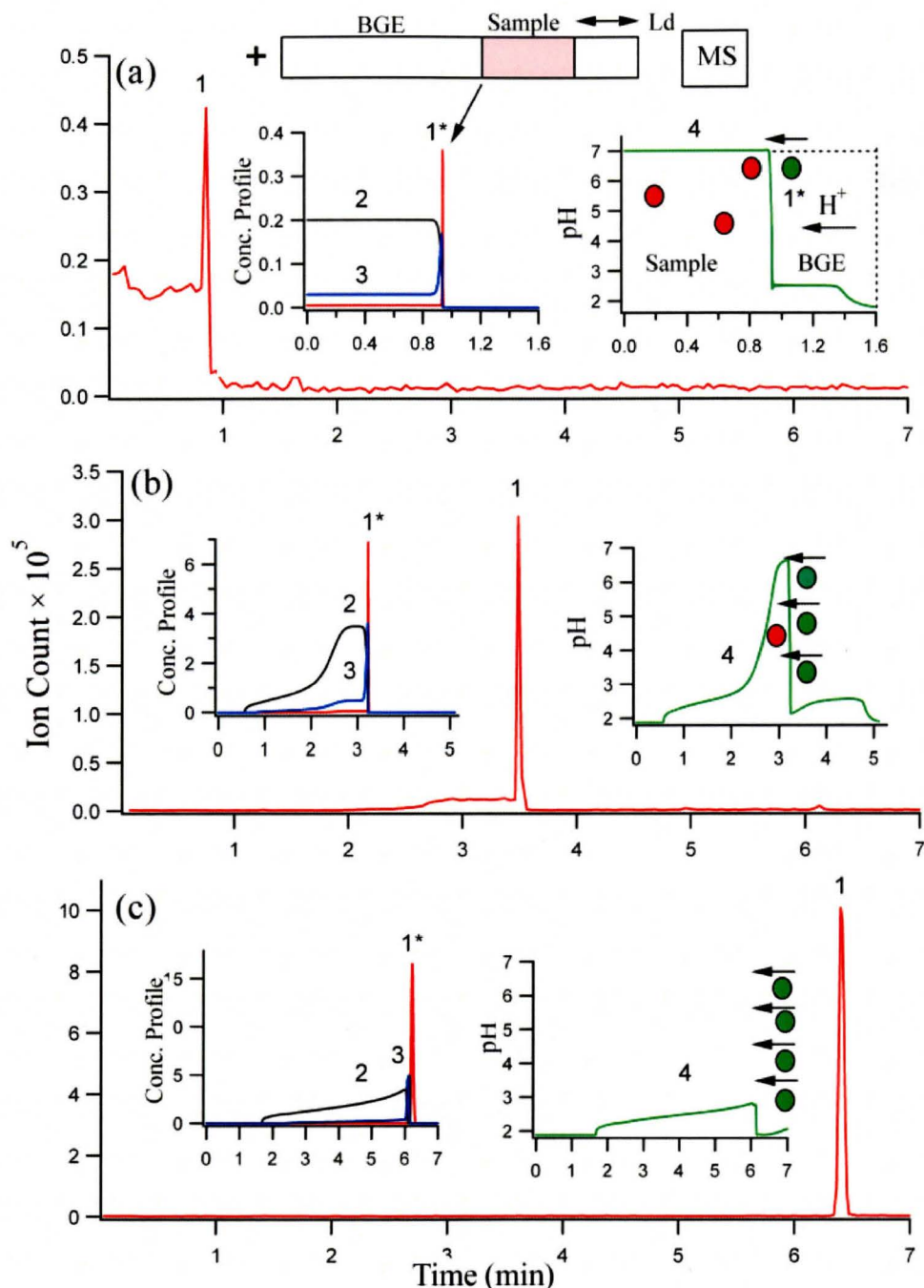


Figure 1.7 Time-resolved electropherograms showing the dynamics of in-capillary Trp focusing with a dynamic pH junction from experimental data and computer simulations using Simul 5.0 (insets) showing a sample plug of 12.5 cm placed at a) 7.5 cm from distal end of capillary, b) 25 cm, and c) 40 cm. with a BGE of formic acid at pH 1.8 and a sample matrix composed of 200mM ammonium acetate pH 7.0 and 15mM NaCl. Peak numbering corresponds to (1) Trp, (2) NH_4^+ , (3) Na^+ , and (4) pH.

diffusion occurs prior to ESI-MS detection. It is important to note that this method is applicable to high salt samples since most alkali metals (*e.g.*, Na⁺) migrate with a large positive μ_{ep} relative to most polar metabolites and their mobility is pH-independent. In other words, effective on-line sample desalting is also performed by CE-ESI-MS since involatile salts migrate well ahead (cations) or out of the capillary (anions) relative to cationic solutes as a way to avoid ionization signal suppression in ESI-MS.

1.5 Inherited Metabolic Disorder Screening

Currently one of the most exciting developments in analytical chemistry is the emerging field of metabolomics, which was first introduced by Fiehn in 2001 as the comprehensive analysis of low molecular weight metabolites in a biological system³⁴. Metabolomic studies promise to provide an improved understanding regarding the health and well-being of an organism since metabolites represent “real-world” endpoint indicators of gene expression that is strongly associated with phenotype. One important example of the significance of metabolite analysis is in the area of newborn screening for inborn errors of metabolism (IEM). Although newborn screening is still commonly performed via analysis of targeted biomarkers associated with IEMs, metabolomics may provide an unbiased strategy to better distinguish IEM variants with lower false-positive rates by identification of multiple classes of biomarkers.

1.5.1 History of newborn screening

The term “inborn errors of metabolism” was first used by Garrod³⁵ to describe hereditary alterations of normal enzymatic reactions. Prior to the deciphering of

metabolic pathways, clinical diagnosis of these IEMs was often made by observation of gross physical symptoms such as darkened urine, musty odour or mental aberrations in which case the onset of clinical symptoms were often irreversible. With the identification and understanding of the pathways associated with metabolic disorders, researchers have been able to identify characteristic biochemical “roadblocks” in normal enzyme function which typically lead to the accumulation of substrate or unusual by-products that are not normally found in healthy individuals. At approximately the same time as Garrod introduced the concept of IEMs, the basis of many fundamental genetic theories were being developed including the “one-gene, one-enzyme” theory introduced by Beadle and Tatum³⁶. This theory stated that a gene would be responsible for the alteration or omission of an enzyme and this would lead to an “all or none” effect on metabolism. Unfortunately, this simplistic view of genetic function allowed diagnosis of only the most severe phenotypes of IEMs. As knowledge of genetics and enzyme function progressed over the years however, more specific and sensitive methods have been developed which permit early diagnosis of IEMs prior to onset of clinical symptoms. This is significant since it permits early-stage intervention and treatment of newborns that reduces infant mortality and morbidity, which can often be managed by specific dietary adjustments or restrictions.

1.5.2 Biomarkers and Targeted Metabolic Screening

As the understanding of the metabolic pathways effected by inherited metabolic disorders increased, researchers began to identify certain characteristic compounds found in biological fluids which could be used as diagnostic markers for identifying diseases

and these became known as biomarkers. *Table 1.1.* lists many metabolic disorders which are currently screening for in North America. Of note, many disorders are not characterized by one biomarker but by ratios of biomarkers or the presence/absence of a number of biomarkers simultaneously. This is a reflection of the nature of many metabolic disorders which arise because of no or partial loss of enzyme activity, which in turn creates large excesses of substrate and little to no product for a given pathway.

1.5.3 Bioassay and Tandem MS Screening

In the early 1960's, Robert Guthrie introduced the first procedure suitable for universal newborn screening to diagnose phenylketouria (PKU), which is associated with elevated levels of phenylalanine (Phe) in dried blood spots collected on filter papers³⁷. This procedure which was based on a bacterial inhibition assay became the first cost-effective and widely used newborn screening method. Following Guthrie's efforts, a number of different bioassay and enzymatic assay methods were developed for various targeted biomarkers for expanded newborn screening of other IEMs. However, because of the cost-effectiveness, low positive predictive outcome and low frequency of certain IEMs, many of these tests were not implemented by most healthcare programs. Since mid-1990's, clinical chemists realized the potential of tandem MS (ESI-MS/MS) as a new platform for high-throughout screening of multiple IEMs simultaneously in contrast to traditional assays targeting specific biomarkers. *Ch. 3* of this thesis will outline and discuss a new strategy to further improve the performance of current newborn screening practices for IEMs when using CE-ESI-MS.

Table 1.1 Decision criteria and upper/lower cut-off limits based on normal population controls for selected IEMs targeted in newborn screening programs by ESI-MS/MS.

IEM Disorder	“Positive” decision criteria (μM or conc. ratio)
Arginemia ⁴³	> 30 arg
Citrullinemia type I ⁴³	> 30 citrulline, normal Met
Citrullinemia type II ⁴³	> 30 citrulline, > 40 Met
Homocystinuria ⁴³	> 40 Met
Hypermethioninemia ⁴²	> 60 Met
Maple syrup urine disorder ³⁹	> 200 (leu +ile), >2 allo-leu, >150 Val, Val/Phe >3
Hydroxyprolinemia ³⁹	> 250 (leu +ile), <2 allo-leu, 4-OH-Pro present
5-oxo-Prolinuria ⁴⁰	5-oxo-Pro/Phe > 1.22
Phenylketouria ⁴³	> 120 Phe and Phe/Tyr >2.0
Tyrosinaemia type I,II, and III ⁴³	> 100 Tyr and > 40 Met
Hyperprolinemia ⁴⁰	> 5.62 Pro
Ornithine transcarbamylase deficiency ⁴³	< 6.0 Citrulline
Carnitine uptake disorder ⁴¹	Low total carnitines
Carnitine palmitoyltransferase def. I ⁴³	> 60 free carnitine
Carnitine palmitoyltransferase def. II ⁴³	> 0.3 C14, > 2.5 C16, >0.2 C16:1, >1 C18, > 1.8 C18:1
Short chain acyl-CoA dehydrogenase def. ⁴³	>4.0 C3, >0.6 C4
Carnitine transporter defect ³⁸	< 10 free carnitine, < 5 total carnitines
Methylmalonic acidemia ⁴³	> 4 C3, and C3/free carnitine > 0.2
Propionic acidemia ⁴³	> 6.8 C3

1.6 Research Objectives

The main goal of this thesis is to develop fundamental and applied analytical strategies based on CE-ESI-MS required for future progress in metabolomics research. The first section of this work examines fundamental ionization processes in ESI in order to develop a quantitative model for accurate prediction of relative ionization response for unknown metabolites when commercial standards are unavailable. The second section of the thesis optimizes and validates a novel CE-ESI-MS method applicable for newborn

screening of inherited metabolic disorders when using dried blood spot extracts. Three specific research objectives/hypotheses will be addressed in this thesis:

1. Develop a fundamental understanding of the factors influencing differential ionization responses for metabolites in ESI-MS
2. Validate a multivariate quantitative model for accurate prediction of the relative ionization response for unknown metabolites by CE-ESI-MS
3. Develop an alternative method for newborn screening of inherited metabolic disorders by CE-ESI-MS with improved specificity and sensitivity.

1.7 References

1. Giddings J.C.; *Unified separation science*, John Wiley & sons, New York, NY, **1991**
2. Skoog D.; West, D.; West, . Holler, F.; *Analytical Chemistry: An Introduction 5th ed.*, Saunders College Publishing, Philadelphia PA. USA, **1990**
3. Armstrong, F. D.W.; Schulte, G.; Schneiderheinze J.M.; and Westenberg, D.J.; *Anal. Chem.*, **1999**, 71, 5465
4. Kohlrausch, F.; *Wiedemanns Ann.* **1897**, 62, 209
5. Tiselius, A.; *Inaugural Dissertation Upsala*, **1930**
6. Svensson, H.; *Kolloid.Z.* **1939**, 87, 180
7. Hjerten, S.; *Arkiv. Kemi*, **1958**, 13, 151
8. Hjerten, S.; *Chromatogr. Rev.*, **1967**, 9, 122
9. Mikkers, F.E.P.; Everaerts, F.M.; Verheggen, T.P.; *J. Chromatogr.*, **1979**, 169, 11
10. Jorgenson, J.; Lukacs, K.D.; *Anal. Chem.*, **1981**, 53, 1298
11. Landers, J.P.; ed., *Handbook of capillary electrophoresis*, CRC Press, Boca Raton, FL USA, **1997**
12. Terabe, S.; Otsuka, K.; Ichikawa, K.; Tsuchiya, A.; Ando, T.; *Anal. Chem.*, **1984**, 56, 111
13. Kranack, A.R.; Bowser, M.T.; Britz-McKibbin, P.; Chen, D.D.Y.; *Electrophoresis*, **1998**, 19, 338
14. Jandik, P.; Bonn, G.; *Capillary electrophoresis of small molecules and ions*, VCH Publishers Inc., New York, N.Y. USA, **1993**
15. Pavia, D.; Lampman, G.; Kriz, G.; *Introduction to Spectroscopy 3rd ed.*, Thomson Learning Inc., Toronto On., **2001**
16. Dovichi, N.J.; Martin, J.C.; Jett, J.H.; Trkula, M.; Keller, R.A.; *Anal. Chem.*, **1984**, 56, 348
17. Maxwell, E.J.; Chen, D.D.Y.; *Analytica Chimica Acta*, 2008, in press
18. Chapman, S.; *Phys. Rev.*, **1937**, 52, 184

19. Dole, M.; Mack, L.; Hines, R.L.; Mobley, R.C.; Ferguson, D.; Alice, M. *J. Chem. Phys.*, **1968**, 49, 2240
20. Iribarne, J.V.; Thompson, B.A.; *J. Chem. Phys.* **1976**, 64, 2287
21. Yamashita, M.; Fenn, J.B.; *J. Phys. Chem.*, **1984**, 88, 4451
22. Yamashita, M.; Fenn, J.B.; *J. Phys. Chem.*, **1984**, 88, 4671
23. Wong, S.; Meng, L.; Fenn, J.B.; *J. Phys. Chem.*, **1988**, 92, 546
24. C. Whitehouse, R. Dreyer, M. Yamashita, J.B. Fenn, *Anal. Chem.*, **1985**, 57, 675
25. Paul, W.; Steinwedel, H.; *Zeitschrift für Naturforschung*, **1953**, 8A-448
26. Dehmelt, H.; *Adv. At. Mol. Phys.*, **1967**, 3, 53
27. Stafford, G.C.; Kelley, P.E.; Syka, J.E.; Reynolds, W.E.; Todd, J.F.; *Int. J. Mass Spectrom. Ion processes*, **1984**, 60, 85
28. DeHoffman, E.; Charette, J.; Stroobant, V.; *Mass spectrometry : principles and practice*, John Wiley & Sons, Chichester UK, **1996**
29. March, R.; Todd, J.; ed., *Practical aspects of ion trap mass spectrometry*, CRC Press, Boca Raton Fl. USA, **1995**
30. Smith, R. D.; Udseth, H.R.; Barinaga, C.J.; Edmonds, C.G.; *J. Chromatogr.*, 1991, 559, 197
31. Zamfir, A.D.; *J. Chromatogr. A.*, 2007, 1159, 2
32. Lee, R.; Ptolemy, A.S.; Niewczas, L.; Britz-McKibbin P.; *Anal. Chem.*, 2007, 79, 403
33. Imamik, K.; Monton, M.R.N.; Ishihama, Y.; Terabe, S.; *J. Chromatogr. A.*, 2007, 1148, 250
34. Fiehn, O.; *Comp. Funct. Genom.*, **2001**, 2, 115
35. Garrod, A.E., *Lancet*, **1902**, 13, 1616
36. Stent, G.S.; Calender, R.; “*One gene-One enzyme theory, molecular genetics: an introductory narrative*” 2nd ed. Freeman Press, San Francisco CA, **1978**
37. Guthrie, R.; Susi, A.; *Pediatrics*, **1963**, 32, 338
38. Wilcken, B.; Wiley, V.; Hammond, J.; Carpenter, K. *N. Engl. J. Med.* **2003**, 348, 2304-2312.
39. Wilcken, B. *J. Inherit. Metab. Dis.* **2007**, 30, 129-133
40. Shigematsu, Y.; Hata, I.; Kikawa, Y.; Mayumi, M.; Tanaka, Y.; Sudo, M.; Kado, N. *J. Chromatogr. B* **1999**, 731, 97-103.
41. Zoppa, M.; Gallo, L.; Zacchello, F.; Giordano, G. *J. Chromatogr. B* **2006**, 831, 267-273.
42. Ghoshla, A. K.; Guo, T.; Soukhova, N.; Soldin, S. J. *Clin. Chim. Acta* **2005**, 358, 104-112.
43. Han, L.S.; Ye, J.; Qiu, W.J.; Gao, X.L.; Wang, Y.; Gu, X.F.; *J. Inherit. Metab. Dis.*, **2007**, 30, 507

CHAPTER 2

Virtual Quantification of Metabolites Without Chemical Standards by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry

Abstract

Despite the widespread use of electrospray ionization-mass spectrometry (ESI-MS), a quantitative model for predicting the ionization efficiency of diverse classes of analytes in real-world samples is lacking. This is particularly relevant in the context of metabolomics when the quantification of unknown metabolites identified by ESI-MS remains unfeasible when purified commercial standards are unavailable. Herein we demonstrate a multivariate strategy for *de novo* quantification of metabolites using capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS) based on fundamental electrokinetic, thermodynamic and molecular properties of an ion. A model set of cationic metabolites was used to accurately predict the ionization efficiency of various classes of amino acids, amines, peptides, nucleosides and acylcarnitines, where solute relative ion response is dependent on additive thermodynamic and electrokinetic processes. Virtual calibration curves were generated *in silico* for accurate quantification of micromolar levels of metabolites in filtered red blood cell (RBC) lysates without ionization suppression. This strategy allows for direct quantification of recently discovered metabolites based on their putative chemical structure, as well as providing deeper insight into the selectivity of solute ionization efficiency in ESI-MS.

2.1 Introduction

There is growing interest in metabolomics for revealing the functional impact of gene expression on the phenotype of an organism that is required for new advances in drug development,¹ disease diagnosis,² environmental toxicology³ and agriculture.⁴ Two instrumental platforms widely used in metabolomic studies include NMR and MS, which provide quantitative and qualitative information suitable for comprehensive metabolite analyses.^{5, 6} ESI-MS is the method of choice for direct analysis of polar metabolites due to its high sensitivity and direct compatibility with separation techniques, such as liquid chromatography (LC)⁷ and capillary electrophoresis (CE)⁸. A major challenge in metabolomics remains the identification of a large fraction of unknown yet biologically relevant metabolites that do not correspond to known candidates within conserved metabolic pathways. In cases when no match is found within public databases,⁹ several empirical candidate structures can be deduced from accurate mass, isotopic composition and fragmentation information.^{10,11} However, reliable quantification of novel metabolites remains elusive while they are not commercially available, difficult to synthesize or costly to purify. This dilemma is considerable when less than 10% of total metabolite peaks detected in biological samples can be reliably quantified due to limited access to chemical standards.¹²⁻¹⁴ Thus, new strategies that permit rapid yet accurate quantification of recently identified metabolites (*e.g.*, drug metabolites, disease biomarkers, xenobiotics *etc.*) based on their putative chemical structure are needed in ESI-MS when purified standards are unavailable.

The development of ESI-MS^{15, 16} as an efficient means for ionizing intact biological molecules has revolutionized modern instrumental analysis. Recent evidence¹⁷ supports an ion evaporation mechanism for describing gas-phase ion formation involving the desolvation of low molecular weight metabolites from charged droplets. Despite the explosive growth in ESI-MS applications, a fully quantitative model for predicting solute ionization response is still lacking^{18, 19}. This is relevant given the wide disparity in analyte ionization efficiency which can result in responses that differ by over three-orders of magnitude despite equimolar concentration levels in solution. Clearly this phenomenon is a reflection of distinct physicochemical properties of solutes that impact gas-phase ionization processes under an electric field. Iribane *et al.*²⁰ first postulated that the selectivity of ion response in ESI-MS can be related to differences in the relative affinities of solutes for the charged droplet surface. Kebarle *et al.* later attributed differences in analyte sensitivity based on the rates of ion evaporation from ESI droplets due to factors associated with solute solvation energy and surface activity.^{21, 22} An equilibrium partitioning model for ESI-MS was later introduced by Enke²³ to describe solute ionization response via competitive displacement of analytes with other electrolyte co-ions for the surface of a charged droplet where ion desorption is favored. To date, several reports have examined the relationship of various solute thermodynamic properties on ionization efficiency in ESI-MS, including non-polar surface area, free energy of solvation, octanol-water partition coefficient, reverse-phase HPLC retention time, and gas-phase proton affinities.²⁴⁻²⁸ However, most ESI models have had limited success in predicting solute ionization responses in complex sample matrices for two

explicit reasons: 1. univariate correlations of single physicochemical parameters to describe gas-phase ionization are inadequate to model the behavior of diverse classes of analytes that differ significantly in terms of their charge, polarity or size (*e.g.*, metabolites) and 2. real-world samples often induce analyte ionization suppression due to background matrix effects (*e.g.*, involatile salts, abundant co-ions) that is highly variable and sample-dependent. The latter issue can be addressed if appropriate *off-line* sample pretreatment is performed (*e.g.*, desalting) prior to direct-infusion ESI-MS studies, or preferably if a high efficiency separation technique is coupled to ESI-MS to resolve low abundance analytes from major co-ions present in a sample mixture, including isomeric and isobaric ions.

Recently, Caetano *et al.* explored several multivariate models for predicting the ionization response of over 160 drug standards using computationally derived molecular descriptors by LC-MS with ESI and atmospheric pressure chemical ionization (APCI).²⁹ After model refining, six molecular descriptors reflective of 2D chemical structure were found to be significant in predicting responses in LC-ESI-MS, however the physical significance of these low-order theoretical descriptors for influencing ESI response was unclear.²⁹ In addition, the model was not validated for drug quantification in real-world samples. In this report, we introduce a multivariate strategy for accurate prediction of the relative ion response (RIR) of fifty different metabolites using CE-ESI-MS based on four intrinsic electrokinetic, thermodynamic and molecular properties associated with an ion. Accurate quantification of diverse classes of cationic metabolites (*e.g.*, amines, amino acids, peptides, nucleosides, acylcarnitines *etc.*) in red blood cell (RBC) lysates without

significant bias caused by matrix-induced signal suppression was demonstrated based on virtual calibration curves derived from partial-least squares (PLS) regression. The use of RIR was a critical feature in this work for normalizing changes in apparent ionization response that minimized long-term instrumental variations. Method validation was demonstrated by a series of recovery experiments involving twenty-five different metabolites spiked at three different concentration levels in RBC lysates by CE-ESI-MS. To the best of our knowledge, this is the first validated strategy that permits *in silico* quantification of diverse classes of metabolites in complex biological samples using ESI-MS. Noteworthy, our study demonstrates the feasibility for *de novo* quantification of metabolites based on their putative chemical structure, where multiple physicochemical parameters can be determined experimentally or estimated via computer molecular modeling. CE-ESI-MS offers a powerful format for rapid quantification of low abundance metabolites that is relevant to rapidly expanding metabolomic initiatives, while providing improved understanding of the underlying factors influencing the selectivity of analyte ionization efficiency in ESI-MS.

2.2 Experimental

2.2.1 Chemicals and reagents

De-ionized water used for buffer and sample preparations was obtained using a Barnstead EASY-pure II LF ultrapure water system (Dubuque, IA, USA). Formic acid was used as the background electrolyte (BGE), which was prepared by dilution of concentrated formic acid (Sigma-Aldrich Inc., St. Louis, MO) with de-ionized water. All other chemicals used including all sample metabolites and organic solvents were

purchased from Sigma-Aldrich, including *O*-acetyl-*L*-carnitine (C2), adenine (An), adenosine (A), *L*-alanine (Ala), β -alanine (β -Ala), γ -aminobutyric acid (GABA), *p*-aminobenzoic acid (PABA), *L*-arginine (Arg), *L*-asparagine (Asn), *L*-aspartic acid (Asp), atenolol (At), *O*-butyryl-*L*-carnitine (C4), *L*-carnitine (C0), *L*-carnosine (Carn), creatinine (Cre), *p*-chloro-*L*-tyrosine (Cl-Tyr), *L*-citrulline (Cit), *L*-cystathionine (Cyst), 2,3-dihydroxy-*L*-phenylalanine (DOPA), dopamine (DopN), *L*-glutamic acid (Glu), *L*-glutamine (Gln), guanine (Gn), guanosine (G), histamine (HisN), *L*-histidine (His), *L*-homocysteine (HCy), 5-hydroxyl-*L*-tryptophan (OH-Trp), *L*-leucine (Leu), *L*-isoleucine (Ile), *L*-lysine (Lys), metoprolol (Meta), methanol (MeOH), *N*-methyl-*L*-aspartic acid (Me-Asp), 3-methyl-*L*-histidine (Me-His), 1-methyl-adenosine (Me-A), *L*-methionine (Met), *L*-methionine sulfone (MetS), nicotinamide (NA), nicotinic acid (NA), *p*-nitro-*L*-tyrosine (NO₂-Tyr), *n*-octanol, *O*-octoyl-*L*-carnitine (C8), *L*-ornithine (Orn), oxytetracycline (Otet), oxidized glutathione (GSSG), *L*-phenylalanine (Phe), *L*-threonine (Thr), tryptamine (TrpN), tyramine (TyrN), *L*-tryptophan (Trp), *O*-propionyl-*L*-carnitine (C3), *L*-tyrosine (Tyr), *m*-*L*-tyrosine (*m*-Tyr), *o*-*L*-tyrosine (*o*-Tyr), reduced glutathione (GSH), serotonin (SerN), *L*-theanine (Thea) and *L*-valine (Val). Stock solutions of analytes were prepared in de-ionized water or 1:1 MeOH:H₂O and subsequently diluted in 200 mM ammonium acetate, pH 7.0 for calibration studies or diluted standard solutions were spiked into filtered RBC lysates for recovery and validation experiments.

2.2.2 Apparatus and conditions

Separation and detection studies were performed on an Agilent CE system equipped with an XCT 3D ion trap mass spectrometer, an Agilent 1100 series isocratic pump, and a

G16107 CE-ESI-MS coaxial sheath-liquid sprayer interface (Agilent Technologies Inc., Waldbronn Germany). All separations were performed on uncoated fused-silica capillaries (Polymicro Technologies Inc., Phoenix, USA) with 50 μm internal diameter and 80 cm total length. Samples were injected hydrodynamically for 75 s under low pressure (50 mbarr) followed by a 10 s injection of background electrolyte (BGE) which consisted of 1.4 M formic acid, pH 1.8. The sample injection plug length used in this work was equivalent to 6.2 cm or about 8% of the total capillary length, which provided a simple strategy for *on-line* sample preconcentration of dilute solutes via electrokinetic focusing in a discontinuous electrolyte system.³⁰ CE separations were performed at 20°C with an applied voltage of 25 kV. The sheath liquid consisting of 1:1 MeOH:H₂O with 0.1% formic acid was supplied by the 1100 series isocratic pump at a flow rate of 10 $\mu\text{L}/\text{min}$. Nitrogen was used as both a nebulizing and a drying gas supplied at 6 psi and 10 L/min respectively, whereas helium at 6×10^{-6} mbarr was used as a damping gas for the ion trap. All MS analyses were performed using a 5 kV cone voltage in positive-ion mode (+ESI) at 300° C. MS data was recorded within a range of 50-750 m/z using an ultrascan mode of 26,000 m/z per second. All calibration and validation studies were performed using fixed ESI-MS conditions described above, whereas different batches of BGE/sample and sheath liquid solutions were prepared frequently throughout the duration of the study (\approx 1 year). Fused-silica capillaries used for CE separations which also serve as the sample outlet in ESI had average lifespans of about 1 month after initial conditioning for 20 min each using MeOH, 0.1M HCl, de-ionized H₂O and BGE. The distal end of the capillary was precisely cut with a Shortixtm diamond capillary column

cutter (SGT Middelburg, Holland) followed by burning a short segment (≈ 1 cm) section of the polyimide coating in order to expose the bare fused-silica capillary outlet end that is then inserted into the coaxial sheath liquid interface. In general, the capillary was flushed with BGE for 10 min to ensure adequate rinsing prior to every sample injection. Evidence of capillary deterioration based on unstable/variable currents after repeated buffer rinsing and conditioning signaled preparation of a new capillary.

2.2.3 Measurement of Solute Physicochemical Parameters

In this study, four fundamental physicochemical parameters were selected as input parameters (X-matrix) to describe solute ionization response in ESI-MS, namely molecular volume (MV), octanol-water distribution coefficient ($\log D$), absolute mobility (μ_o) and effective charge (z_{eff}). The latter parameter was derived from experimentally determined pK_a values ($pK_a \approx 1.6-3.5$, ionic strength ≈ 150 mM) for model metabolites containing an acidic functional group in their structure (*e.g.*, α -COOH), which impacted their z_{eff} under the acidic BGE conditions (pH = 1.8) used for separation. In cases when metabolites had weakly acidic functional groups with $pK_a > 4$ or for amines with $pK_a > 8$, z_{eff} was considered nominally as +1 or +2 for monovalent (*e.g.*, dopamine) and divalent (*e.g.*, histamine) amines, respectively. CE-ESI-MS was used to measure apparent pK_a values corrected for ionic strength³¹ (≈ 150 mM) via non-linear regression of measured changes in apparent solute mobility (μ_{ep}^A) as a function of buffer pH (pH $\approx 1.6-4.0$) as reported previously.³⁰ Determination of μ_o was also performed by CE-ESI-MS based on linear extrapolation to zero ionic strength of measured μ_{ep}^A for model metabolites ($I \approx 25$ -

150 mM, pH 1.8) using Pitt's equation,³² where the ionic strength of the BGE was adjusted using ammonium acetate. Direct measurements of $\log D$ were also performed by CE-ESI-MS by *off-line* equilibration of a metabolite standard prepared in de-ionized water (250 μL), which was added to an equal volume of *n*-octanol. This mixture was then vortexed vigorously for 30 min and centrifuged at 10,000 rpm for 3 min. An aliquot was taken from each immiscible phase, which was then diluted in a methanol solution containing saturated ammonium acetate with Trp as an internal standard. Each sample was then analyzed directly by CE-ESI-MS using a 3 s injection (50 mbarr) with a 1.4 M formic acid as BGE. $\log D$ was determined based on the average relative peak area ratio for each metabolite measured in the *n*-octanol and de-ionized water phases. Due to the immiscibility and high resistance of *n*-octanol in the sample plug relative to the aqueous BGE, the use of a saturated ammonium acetate solution in *n*-octanol was found to improve stabilization of the current during CE-ESI-MS analyses. Because of weak partitioning for some hydrophilic metabolites into the *n*-octanol phase, high concentrations of standard (\approx mM) were needed in order to quantify the low levels of distributed solute, which often had weak ionization responses. All experimentally determined parameters (*i.e.*, $\log D$, μ_o and z_{eff} or pK_a) were measured in triplicate to ensure good precision with errors reported as $\pm 1\sigma$. Also, computer molecular modeling was used to determine MV based on the Connolly solvent-excluded volume for model metabolites, which was performed using Chem3D Ultra software, version 8.0 (CambridgeSoft Inc., Cambridge, USA). All chemical structures were energy-minimized using an iterative molecular mechanics 2 (MM2) algorithm with molecular dynamics as a

way to determine a stable molecular configuration prior to computing relevant parameters. Overall, experimentally determined parameters were performed for over forty different model metabolites as the training/test set, which encompassed a diverse class of biologically relevant cationic metabolites, including amines, amino acids, peptides, nucleosides and acylcarnitines, as well as their modified analogues. *Appendix 1* summarizes the chemical structures and physicochemical parameters measured for the forty metabolites used in this study after model refining.

2.2.4 Method Calibration

Calibration curves were constructed for more than forty different metabolites using CE-ESI-MS by measuring their average relative ion response (RIR) at six different concentrations (2, 8, 16, 30, 50, 100 μM) in triplicate ($n = 3$) relative to 50 μM diAla as the internal standard (IS). The use of an IS was critical for normalizing significant differences in apparent ion responses caused by day-to-day variations in ESI-MS, as well as long-term spray stability, buffer/sheath liquid composition and capillary alignment/distal end cutting. The concentration range for calibration was selected such that the lowest response at 2 μM was above the limit of quantification (LOQ, $S/N > 10$) for the method to ensure adequate precision ($CV < 20\%$). All calibration studies were performed via *on-line* sample preconcentration by CE-ESI-MS using 1.4 M formic acid, pH 1.8 as the BGE with metabolite standards prepared in 200 mM ammonium acetate, pH 7 as the sample solution. All samples were injected hydrodynamically using a 75 s (50 mbar) sample injection plug followed by a 10 s injection of BGE. The total relative peak area ratio for ions as derived from extracted ion electropherograms were measured

for the molecular ion (MH^+), as well as any significant fragment ions (*e.g.*, amines, MH^+ -17) or isotope contributions (*e.g.*, ^{35}Cl and ^{37}Cl) provided that the latter contributions comprised more than about 5% of MH^+ . This method of integration was performed so that all major gas-phase species associated with an ion were considered in order to accurately quantify total apparent ion response as a function of analyte concentration in solution, which is summarized in *Appendix 2*.

2.2.5 Prediction of Solute Physicochemical Parameters

In order to demonstrate the feasibility for *de novo* quantification of metabolites based on their putative chemical structure using ESI-MS, *in silico* estimation of solute physicochemical parameters was performed for ten different metabolites as a validation set not used in the original training set. Computer molecular modeling was used to determine MV similar to the method described above, whereas μ_o was predicted using the Hubbard-Onsager hydrodynamic model for ion migration,³³ where the valence charge (z_o) and MV of metabolites were used as input parameters as described in a previous report.³⁰ The latter model was developed by performing non-linear regression fitting³⁰ to experimentally determined μ_o values associated with the forty-four metabolites resulting in coefficient terms of $a = (0.00345 \pm 0.00087)$, $b = (5.9 \pm 13)$, and $c = (0.484 \pm 0.051)$ with a χ^2 of 3.76×10^{-8} . Good linear correlation between predicted and measured μ_o values was observed as reflected by a *slope* = (1.016 ± 0.041) and $R^2 = 0.9665$. Prediction of pK_a of validation metabolites was performed using ACD/Lab pK_a DB (Advanced Chemical Development Inc., Toronto, Canada), which was then converted

into z_{eff} based on the pH of the BGE used for CE separations. A reasonably good predictive accuracy for pK_a using ACD/Lab pKa DB was confirmed by correlation of experimental and predicted pK_a among thirty-two weakly acidic metabolites with a $slope = (0.780 \pm 0.049)$ and $R^2 = 0.9440$. In this work, $logD$ for most metabolites were predicted using ALOGPS 2.1 (Virtual Computational Chemistry Laboratory, <http://www.vcclab.org>) unless otherwise stated. Good correlation between experimentally measured and predicted $logD$ using ALOGPS 2.1 was demonstrated for thirty-six different metabolites with a $slope = (0.817 \pm 0.048)$ and $R^2 = 0.9453$, however poor correlation was noticed for amines and peptides. It was found that $logD$ for the latter sub-classes of metabolites were predicted with better accuracy when using ACD/Lab pKa DB software. *Appendix 3* summarizes the chemical structures and predicted physicochemical parameters for the ten validation metabolites used in this study.

2.2.6 Data Processing and Multivariate Analysis

All data processing and linear/non-linear regression was performed using Igor Pro 5.0 (Wavemetrics Inc., Lake Oswego, OR, USA). Principal component analysis (PCA) and partial-least squares (PLS) regression was performed using a multivariate analysis add-in for Excel (Microsoft Inc., Redmond, WA, USA) developed by Dr. Richard Brereton at Bristol University that is freely available for download at <http://www.chm.bris.ac.uk/org/chemometrics/addins/index.html>. PCA was performed as an unsupervised dimensionality reduction method for correlating inter-analyte variations in terms of their four physicochemical parameters (X-matrix: 40 samples X 4 variables)

that was useful for identifying sample groupings, outliers and qualitative trends. The X-matrix was pre-processed using standardized data (MV , $\log D$, μ_o , z_{eff}) involving over forty different cationic metabolites as the initial model set, where the first two PCs comprised about 76% of the total sample variance (PC1 = 50.4% and PC2 = 25.4%). PLS regression was then used as a supervised multivariate calibration method where a quantitative relationship between the X- and Y-matrices was determined using a set of latent variables that maximize X-Y co-variance.²⁹ In this case, the Y-matrix (40 samples X 6 variables) was comprised of the average relative ion response (RIR) for model metabolites as normalized to an internal standard, which was measured at six different concentration levels (2, 8, 16, 30, 50, 100 μM) in triplicate ($n = 3$) by CE-ESI-MS. PLS regression was then applied to accurately predict RIRs for model metabolites as derived from their four physicochemical parameters, which permits the generation of virtual calibration curves for each solute over a 50-fold concentration range. The predicted relative response factor (RSF) for each metabolite was then determined based on the slope of the virtual calibration curve. Indeed, the major criterion in this work was good predictive accuracy for metabolite quantification while retaining significant chemical diversity, thus model refining involved the rejection of model analytes as outliers only when high residual values were observed. This was the case for several metabolites used in the original training set, including low molecular weight and highly hydrophilic metabolites (*e.g.*, Ala, β -Ala, GABA), high molecular weight and hydrophobic metabolites (*e.g.*, metoprolol, Otet), as well as labile metabolites (*e.g.*, GSH, GSSG) that had very weak, strong or variable RIRs, respectively. This inherent bias is a reflection of the specific

composition of metabolites selected as the training set, which tended to emphasize moderately-sized and relatively hydrophilic ions (e.g., amines/amino acids) with mean $MV \approx 135 \text{ \AA}^3$, $\log D \approx -1.95$, $\mu_o \approx -3.00 \times 10^{-4} \text{ cm}^2/\text{Vs}$ with $z_{\text{eff}} \approx +0.91$. Overall, accurate quantification of micromolar levels of metabolites in RBC lysates was achieved in this work (mean absolute error < 50%) provided that metabolites possessed physicochemical properties within the range of the training set. Overall, test and validation metabolites examined in this study possessed over a 100-fold difference in ionization response in ESI-MS ranging from creatinine to *O*-octoyl-*L*-carnitine, which had the lowest and highest RSF, respectively.

2.2.7 RBC Lysate Preparation and Recovery Experiments

Fingerprick micro-sampling by a consenting healthy adult volunteer was performed by first lancing a finger tip with a disposable BD Genie Lancet (BD Vacutainer Systems Inc., NJ, USA). A blood sample ($\approx 50\text{-}100 \mu\text{L}$) was then collected by placing a SAFE-T-FILL capillary collection tube (Ram Scientific Inc., Kabe Germany) at a 90° angle to the droplet such that the blood was directly transferred into an EDTA-filled microtube. The blood sample was then inverted several times to ensure proper mixing. Equal portions of the blood sample was transferred into two separate microtubes using a disposable pipette tip with SUPERSILK surface and wide orifice (VWR International, Mississauga, ON) to prevent hemolysis. The samples were then centrifuged at 4°C for 10 min at 500 g for blood fractionation. The plasma layer was then removed and RBCs were washed with phosphate buffer solution (10mM NaH_2PO_4 , 150 mM NaCl, pH 7.4), vortexed for 1 min

and then centrifuged at 7,000 g for 1 min, which was repeated three-times. Washed RBCs were then lysed by the addition of a 3-fold volume excess of ice-cold de-ionized water, which was vortexed for 10 s then centrifuged at 7,000 g for 1 min at 4°C to pellet cell debris. Ultracentrifugation of the RBC lysate was performed using a 3 kDa filter (PALL Corporation, Ann Arbor, Michigan) at 14,000 rpm for 20 min at 4°C to remove protein (*e.g.*, hemoglobin). The filtered RBC lysate was diluted 1:1 with 400 mM ammonium acetate, pH 7.0 prior to CE-ESI-MS analyses. Recovery experiments were performed to validate model accuracy for both test (25) and validation (10) metabolites by spiking in specific concentrations of metabolite standards prepared in 400 mM ammonium acetate, pH 7.0 into filtered RBC lysate samples with a 1:1 dilution. Spiking studies were performed in triplicate ($n = 3$) at low (5 μM), mid (25 μM) and high (60 μM) concentration levels not originally used in the Y-matrix. Metabolites used for spiking studies were selected to cover a wide range of solute size, polarity and charge provided they were not detected in original non-spiked filtered RBC sample by CE-ESI-MS.

2.3 Results and Discussion

2.3.1 Multivariate Model for Solute Ionization in ESI-MS

Although several tenets are well-established regarding ionization processes in ESI-MS,¹⁸ a fully quantitative model for predicting the ionization response of metabolites in real-world samples is still lacking. Researchers have long considered ESI as an electrolytic cell where droplet charging is mediated via an electrophoretic process under an intense electric field, which induces charge separation of electrolytes within

solution.^{18, 34} Excess charge build-up on the surface of the solution ultimately leads to instability that overcomes liquid surface tension resulting in the formation of a fine mist of charged droplets, where solute ionization can occur via an ion evaporation mechanism.¹⁷ This highlights the similarity underlying CE and ESI processes since ion electromigration are both driven by an external DC voltage. Indeed, CE and ESI-MS represent orthogonal separation techniques where ions are first resolved on the basis of their effective charge to size ratio in free solution and then detected in the gas-phase as a function of their mass to charge ratio. However, since ESI processes also involve desorption of ions from the surface of charged droplets at a gas-liquid interface, additional factors that impact solute surface activity are most likely important.¹⁹ In this work, four fundamental physicochemical properties were examined to model the relative ionization efficiency of solutes in ESI-MS. The modeling of several parameters for describing ionization responses using a multivariate approach is critical when comparing a diverse series of analytes, which can often show anomalous behaviour when using univariate correlations based on $\log P$ and/or pK_a .²⁴ For instance, μ_o and z_{eff} (or pK_a) represent fundamental electrokinetic/thermodynamic parameters that can be used to simulate analyte electromigration behavior in CE,³⁵ which is useful for qualitative identification of putative metabolite candidates based on their characteristic relative migration time (RMT).³⁰ In addition, $\log D$ and MV were selected as important thermodynamic/molecular parameters reflective of ion surface activity and solvation energy. Additional parameters including Onsager slope,³³ solvent-accessible surface area

and various other molecular descriptors were also examined, but were found not to be significant in improving the predictive accuracy of the model used in this study.

Fig. 1 depicts a multivariate model for describing solute ionization in CE-ESI-MS,²³ where mixtures of ions are first separated within a narrow fused-silica capillary under an applied electric field. CE also serves other functions to enhance ESI-MS performance, including the resolution of isobaric and isomeric ions, as well as on-line sample preconcentration with in-capillary sample desalting for improved concentration sensitivity.³⁰ These features are particularly relevant since background matrix effects (*e.g.*, involatile salts, abundant co-migrating ions) can suppress analyte ionization, which has limited the applicability of ESI-MS models to predict solute responses in biological samples.³⁶ Similar to ion spray configurations widely used in LC-ESI-MS, a coaxial sheath-flow interface was used in this work to stabilize droplet formation,³⁷ where a make-up solvent flow and nebulizer gas mix with the effluent from the capillary tip, which is positioned orthogonal to the orifice entrance of the MS. *Fig. 2* depicts a series of extracted ion electropherograms in CE-ESI-MS for a variety of cationic metabolites used as our initial model set, where there was over a 300-fold difference in RIR for an equimolar 30 μ M sample mixture. For example, β -Ala (9) and Meta (19) represent the least and most responsive metabolite detected as their protonated molecular ion (MH^+) in *Fig 2*, respectively, whereas differences in RIR were also noted among the positional isomers, such as *o*-Tyr and Tyr. These observations emphasize the intrinsic selectivity of ESI, where major and subtle structural changes in chemical structure can significantly impact solute ionization efficiency.

2.3.2 Relative Response Factor in ESI-MS

The ionization efficiency (IE) of an analyte (A) ion in ESI-MS can be described by the additive contributions of equilibrium partitioning²³ and rate of ion evaporation³⁴ that impact gas-phase ion formation from charged droplets using the following equation:

$$IE_A = P \cdot f \left(\frac{K_A C_A}{K_A C_A + K_E C_E} + \frac{k_A C_A}{k_A C_A + k_E C_E} \right) [Q] \quad (1)$$

where, P is the sampling efficiency of the mass spectrometer, f is the fraction of the droplet charge converted to gas-phase ions, K_A , k_A and C_A are the partition coefficient of analyte, rate constant for analyte desorption and analyte concentration, respectively, K_E , k_E and C_E are the partition coefficient, rate constant for co-ion desorption and concentration of all other co-ions (*i.e.*, BGE, co-migrating ions) present in the droplet, respectively, and Q is the concentration of excess charge on the droplet. *Eq (1)* takes into account both thermodynamic and electrokinetic factors that influence solute ionization efficiency that is important when comparing the selectivity in ESI-MS among diverse classes of metabolites. Because of day-to-day variations in spray stability, solution composition and capillary alignment which can often result in significant changes to apparent ion signals,³⁸ solute ionization efficiency was measured in terms of RIR, where the ion response of the solute at a specific concentration was normalized to a fixed internal standard (IS, 50 μ M diAla). Assuming that there are no changes to ESI-MS operating conditions during CE separation, then P , f , Q , K_E , k_E and C_E parameters

described in eq. (1) are constant, where RIR of an ion can be expressed by the following equation:

$$RIR_A(MV, \log D, \mu_o, z_{eff}) = \left(\frac{K_A}{K_{IS} C_{IS}} + \frac{k_A}{k_{IS} C_{IS}} \right) C_A + b \quad (2)$$

where, b is a constant and the coefficient term, $\left(\frac{K_A}{K_{IS} C_{IS}} + \frac{k_A}{k_{IS} C_{IS}} \right)$ represents the RRF of an ion that is equivalent to the slope of its calibration curve (*i.e.*, sensitivity). The RRF of solute can be experimentally determined by measuring the average RIR as a function of its concentration. When considering a homologous series of analytes of similar intrinsic charge where the rates of analyte evaporation are equivalent ($k_A \approx k_B$), then RRF of an ion is primarily determined by the magnitude of its relative partition coefficient (K_A/K_{IS}) for the surface of the charged droplet. In such cases, physicochemical properties that enhance the stability of the ion for residing near the droplet surface result in a larger RRF, such as greater hydrophobicity or high $\log D$. In contrast, a highly hydrophilic ion that has a low relative partition coefficient (*i.e.*, K_A/K_{IS}) may still generate a high RRF due to electrokinetic factors that favor faster rates of desorption, such as high charge density or large μ_o . The major hypothesis in this work is that the RRF of an ion can be predicted based on four intrinsic physicochemical parameters when using multivariate modeling provided that there are not significant changes in co-ion electrolyte composition (C_E) during spray formation. The latter assumption will be demonstrated to be valid since CE can be used to separate low abundance metabolites from major co-ion interferences present in complex sample mixtures.

2.3.3 PCA Modeling and PLS Regression

Fig. 3(a) depicts a 2D scores plot from PCA analysis of forty different cationic metabolites using the four experimentally measured physicochemical parameters listed in *Appendix 1*. The scores plot provides a simple way to qualitatively distinguish different classes of metabolites by comparison with the variables in the loadings plot as shown as an inset in *Fig. 3(b)*. It is apparent that three major groupings of metabolites can be readily classified in *Fig. 3(a)*, namely cationic amino acids (*e.g.*, His), amines (*e.g.*, DopN) and a large subset of neutral amino acids, peptides and nucleosides/purines. In addition, there are several notable outliers with distinct chemical properties, such as polyamines (*e.g.*, HisN), acidic amino acids (*e.g.*, Asp) and hydrophobic amines (*e.g.*, At). Overall, the significance of PC1 is inferred as being associated with the effective charge (z_{eff}) or mobility (μ_o) of metabolites, which generally increases from left (NO₂-Tyr) to right (HisN), whereas PC2 is related to the polarity ($\log D$) and size (MV) of the ion, which increased from bottom (Asp) to top (At) of the 2D scores plot. Although both MV and $\log D$, as well as μ_o and z_{eff} represent pairs of variables that have significant collinearity, neglecting of any one these physicochemical parameters was found to significantly reduce the predictive accuracy of the model. *Fig. 3(c)* depicts a correlation plot between experimentally measured (*i.e.*, slope of calibration curve) and predicted RRFs for 40 different cationic metabolites when performing PLS regression, which showed good linearity as reflected by a $slope = 0.792$ and $R^2 = 0.9144$. The two outliers

in the correlation plot represent the ions with the highest ionization efficiency in the model set, namely At and C2; however, omission of these two apparent leverage data points did not significantly impact either slope or linearity of the correlation plot. Overall, *MV* was the single most important variable that was positively correlated with solute ionization efficiency in this study. In general, metabolites with the highest RRF in this study tended to be high *MV*/hydrophobic ions, such as the synthetic β -blocker drug, At. However, several high *MV*/hydrophilic ions ($\log D < -4$) with high intrinsic charge/mobility also had high RRF, such as Arg and Cyst, which are a cationic amino acid and dipeptide, respectively. These two opposing trends emphasize that both thermodynamic and electrokinetic factors can influence apparent solute ionization efficiency in ESI-MS, which is highly dependent on the intrinsic chemical properties of an analyte. For instance, although hydrophobic ions can more readily accumulate near the surface of charged droplets due to their higher surface activity and lower solvation energy,²³ hydrophilic ions with higher mobility can be more efficiently desorbed at a faster rate due to their higher mobility/charge density. Noteworthy, the impact of covalent modifications to the ionization efficiency of metabolite analogues can also be deduced from this work, such as common metabolic processes involving decarboxylation, methylation, acetylation, hydroxylation, dechlorination and nitration. For example, the decarboxylation of His (to HisN) was observed to induce over a 10-fold reduction in solute ion response, whereas nitration of Tyr (to NO₂-Tyr) generated less than a 2-fold increase in sensitivity. Future studies will better examine the impact of specific metabolic transformations on ESI ion response, which is relevant for rapid quantification

of newly discovered drug metabolites and xenobiotics when purified commercial standards are not readily available.

2.3.4 Model Validation and Recovery Experiments

Two subsequent recovery studies were next performed to better validate the model in terms of allowing for accurate quantification of micromolar levels of cationic metabolites in complex biological samples when using CE-ESI-MS. Fifteen different test metabolites used in the original model were spiked at three different concentration levels (5, 25 and 60 μM) in filtered RBC lysates. In this case, CE was crucial for direct quantification of metabolites in the presence of high levels of involatile salts and abundant co-ions in the sample matrix as shown in *Fig. 4(a)*. It is apparent that excess Na^+ and GSH are resolved in the original sample plug, where minor solutes (spiked at 5 μM level) migrate within about a 7 min separation window that permits their detection by ESI-MS without ionization suppression. This is essential for successful application of quantitative PLS regression methods that were calibrated with neat standard solutions in the absence of matrix effects. *Fig. 5(a)* demonstrates validation of the spike and recovery study for fifteen metabolites in the test set at three different concentration levels. Overall, good accuracy was achieved in this study as reflected by a mean absolute error about 30%. In fact, the recovery was in some cases approaching experimental error notably when quantifying metabolites with weak RRFs at the 5 μM concentration level (near LOQ), where method precision was poor ($\text{CV} \approx 20\%$). It is important to note the use of experimentally derived calibration curves for direct quantification of test metabolites in this study did not significantly improve accuracy relative to virtual calibration curves

predicted from PLS regression. In addition, *de novo* quantification of ten different cationic metabolites not used in the original training set was also performed using a spike and recovery study in filtered RBC lysates, where their physicochemical parameters were estimated using computer modeling. PLS regression was then used to predict the RIR for these cationic metabolites over a 50-fold concentration range. *Fig. 4(b)* depicts virtual calibration curves for the validation set, which predicts significantly different RRF due to differences in their physicochemical parameters, where Orn and C8 represent ions with the lowest and highest relative ionization efficiency, respectively. *Fig. 5(b)* summarizes the results for the recovery experiment for the ten validation metabolites, which provided reasonably accurate predictions with a mean absolute error under 50%. Indeed, this is rather remarkable given the context that the initial calibration and validation studies by CE-ESI-MS were performed over several months apart under variable system/sample conditions with input parameters estimated by computer modeling.

2.3.5 Virtual Quantification Without Chemical Standards?

To the best of our knowledge, this work demonstrates the *proof-of-concept* for virtual quantification of micromolar levels of cationic metabolites in biological samples without ionization suppression effects when using CE-ESI-MS. However, there are several caveats to consider when modeling relative ionization efficiency in ESI-MS using multivariate calibration methods. Firstly, *de novo* quantification of metabolites based on their putative chemical structures ultimately relies on accurate estimation of physicochemical parameters using commercial or public computer software that needs to

be adequately validated, notably in the case of pK_a and $\log D$. For instance, a bias in input $\log D$ values > 0.5 units for the validation set resulted in up to a 30% overestimate in metabolite quantification (ranged from ≈ 5 -30%) notably for weakly responsive analytes (*e.g.*, *Me-Asp*, *Orn*) at low concentrations. Secondly, the robustness of the model is also dependent on appropriate selection of a diverse class of model solutes (test set) that span an appropriate range of physicochemical properties to ensure good performance. In our case, model metabolites were predominately moderately-sized and hydrophilic ions that represented several classes of biologically relevant cationic metabolites; however, the model was not successful in predicting the responses of small/highly hydrophilic metabolites (*e.g.*, *Ala*, *GABA*) and large/highly hydrophobic metabolites (*e.g.*, *Met*, *Otet*) outside of its training range. Moreover, solutes that are pH-sensitive or thermally-labile are also not appropriate for predictive modeling due to their variable responses (*e.g.*, *GSH*). In addition, amphiphilic metabolites that tend to form micelles in aqueous solutions can result in bias, such as phospholipids and long-chain acylcarnitines. This may explain the high error ($\approx 130\%$) observed for *C8* at higher solute concentrations during spike and recovery studies in *Fig. 5(b)*. Nevertheless, we envision that researchers can successfully apply this strategy by selection of a series of model solutes that is guided by their target class of metabolite required for quantification. Although re-calibration is required whenever using different types MS instruments due to changes in ion sampling efficiency (ESI interface/mass analyzer), the applicability of the physicochemical parameters for describing analyte ionization efficiency in ESI should remain universal. In this study, accurate quantification of low abundance metabolites in complex biological

samples was only feasible due to the high efficiency separation of major co-ion interferences (*e.g.*, Na⁺, GSH) in the sample matrix by CE prior to ESI-MS detection. Indeed, ionization suppression due to matrix effects remains the most significant hindrance to the development of quantitative models in ESI-MS when applied to real-world samples.

2.4 Conclusions

Accurate quantification of a diverse class of metabolites as derived from their fundamental physicochemical properties (MV , $\log D$, μ_o and z_{eff} or pK_a) was realized when using CE-ESI-MS in conjunction with multivariate analyses. Both thermodynamic and electrokinetic factors can contribute to the selectivity in solute ionization efficiency in ESI-MS that can vary over three-orders of magnitude. Forty different cationic metabolites were used as a training set for predicting the relative ion response (RIR) of analytes over a 50-fold concentration range when using PLS regression. Method validation was demonstrated by recovery experiments of twenty-five different metabolites spiked in filtered RBC lysates at three different concentration levels (5, 25 and 60 μM) with good accuracy. The feasibility for *de novo* quantification of metabolites based on their chemical structure was also demonstrated in this study when using physicochemical parameters estimated from computer modeling. Further work is needed to expand this strategy to other classes of metabolites under both positive-ion and negative-ion modes in ESI-MS. The potential for applying a similar strategy to quantify metabolites when using LC-ESI-MS would also be of significant interest. It is hoped that

this work can contribute to a deeper understanding of fundamental ESI processes while providing a reliable yet simple strategy for quantification of newly discovered metabolites when purified commercial standards are unavailable.

2.5. References

- (1) Lindon, J. C.; Holmes, E.; Nicholson, J. K. *FEBS J.* **2007**, *274*, 1140-1151.
- (2) Dunn, W. B.; Broadhurst, D. I.; Deepak, S. M.; Buch, M. H.; McDowell, G.; Spasic, I.; Ellis, D. I.; Brooks, N.; Kell, D. B.; Neyses, L. *Metabolomics* **2007**, *3*, 413-426.
- (3) Lee, S. H.; Woo, H. M.; Jung, B. H.; Lee, J. G.; Kwon, O. S.; Pyo, H. S.; Choi, M. H.; Chung, B. C. *Anal. Chem.* **2007**, *79*, 6102-6110.
- (4) Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Ferniw, A. R. *Nat. Protoc.* **2006**, *1*, 387-396.
- (5) Moco, S.; Bino, R. J.; Vos, R. C. H. D.; Vervoort, J. *Trends in Anal. Chem.* **2007**, *26*, 855-866.
- (6) Lenz, E. M.; Wilson, I. D. *J. Proteome Res.* **2007**, *6*, 443-458.
- (7) Wilson, I. D.; Plumb, R.; Granger, J.; Major, H.; Williams, R.; Lenz, E. A. *J. Chromatogr. B* **2005**, *817*, 67-76.
- (8) Monton, M. R. N.; Soga, T. *J. Chromatogr. A* **2007**, *1168*, 237-246.
- (9) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R. E.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S. *Nucleic Acids Res.* **2007**, *35*, D521-D526.
- (10) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2006**, *7*, 234-244.
- (11) Lim, H. K.; Chen, J.; Sensenhauser, C.; Cook, L.; Subrahmanyam, V. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1821-1832.
- (12) Styczynski, M. P.; Moxley, J. F.; Tong, L. V.; Walther, J. L.; Jensen, K. L.; Stephanopoulos, G. N. *Anal. Chem.* **2007**, *79*, 966-973.
- (13) Kind, T.; Tolstikov, V.; Fiehn, O.; Weiss, R. H. *Anal. Biochem.* **2007**, *363*, 185-195.
- (14) Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. *J. Proteome Res.* **2003**, *2*, 488-494.
- (15) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
- (16) Dole, M.; Hines, R. L.; Mack, L. L.; Mobley, R. C.; Alice, M. B.; Ferguson, L. D. *J. Chem. Phys.* **1968**, *49*, 2240-2249.
- (17) Nguyen, S.; Fenn, J. B. *PNAS* **2007**, *104*, 1111-1117.
- (18) Cole, R. B. *J. Mass Spectrom.* **2000**, *35*, 763-772.
- (19) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804-817.
- (20) Iribane, J. V.; Thomson, B. A. *J. Chem. Phys.* **1975**, *64*, 2287-2294.
- (21) Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 3654-3668.

- (22) Sumner, J.; Nicol, G.; Kebarle, P. *Anal. Chem.* **1988**, *60*, 1300-1307.
- (23) Enke, C. G. *ANal. Chem.* **1997**, *69*, 4885-4893.
- (24) Henriksen, T.; Juhler, R. K.; Svensmark, B.; Cech, N. B. *J. Amer. Soc. Mass. Spectrom.* **2005**, *16*, 446-455.
- (25) Null, A. P.; Nepomuceno, A. I.; Muddiman, D. C. *Anal. Chem.* **2003**, *75*, 1331-1339.
- (26) Cech, N. B.; Krone, J. R.; Enke, C. G. *Anal. Chem.* **2001**, *73*, 208-213.
- (27) Cech, N. B.; Enke, C. G. *Anal. Chem.* **2000**, *72*, 2717-2723.
- (28) Amad, M. H.; Cech, N. B.; Jackson, G. S.; Enke, C. G. *J. Mass Spectrom.* **2000**, *35*, 784-789.
- (29) Caetano, S.; Decaestecker, T.; Put, R.; Daszykowski, M.; VanBocxlaer, J.; Heyen, Y. V. *Anal. Chimica Acta* **2005**, *550*, 92-106.
- (30) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403-415.
- (31) Vcelakova, K.; Zuskova, I.; Kenndler, E.; Gas, B. *Electrophoresis* **2004**, *25*, 309-317.
- (32) Roy, K. I.; Lucy, C. A. *Electrophoresis* **2003**, *24*, 370-379.
- (33) Li, D.; Lucy, C. A. *Anal. Chem.* **2001**, *73*, 1324-1329.
- (34) Tang, L.; Kebarle, P. *Anal. Chem.* **1991**, *63*, 2709-2715.
- (35) Hruska, V.; Jaros, M.; Gas, B. *Electrophoresis* **2006**, *27*, 984-991.
- (36) Constantopoulos, T. L.; Jackson, G. S.; Enke, C. G. *J. Amer. Soc. Mass. Spectrom.* **1999**, *10*, 625-634.
- (37) Manisali, I.; Chen, D. D. Y.; Schneider, B. B. *Trends in Anal. Chem.* **2006**, *25*, 243-256.
- (38) Leito, I.; Herodes, K.; Huopolaenen, M.; Virro, K.; Kummapas, A.; Kruve, A.; Tanner, R. *Rapid Commun. Mass. Spectrom.* **2008**, *22*, 379-384.

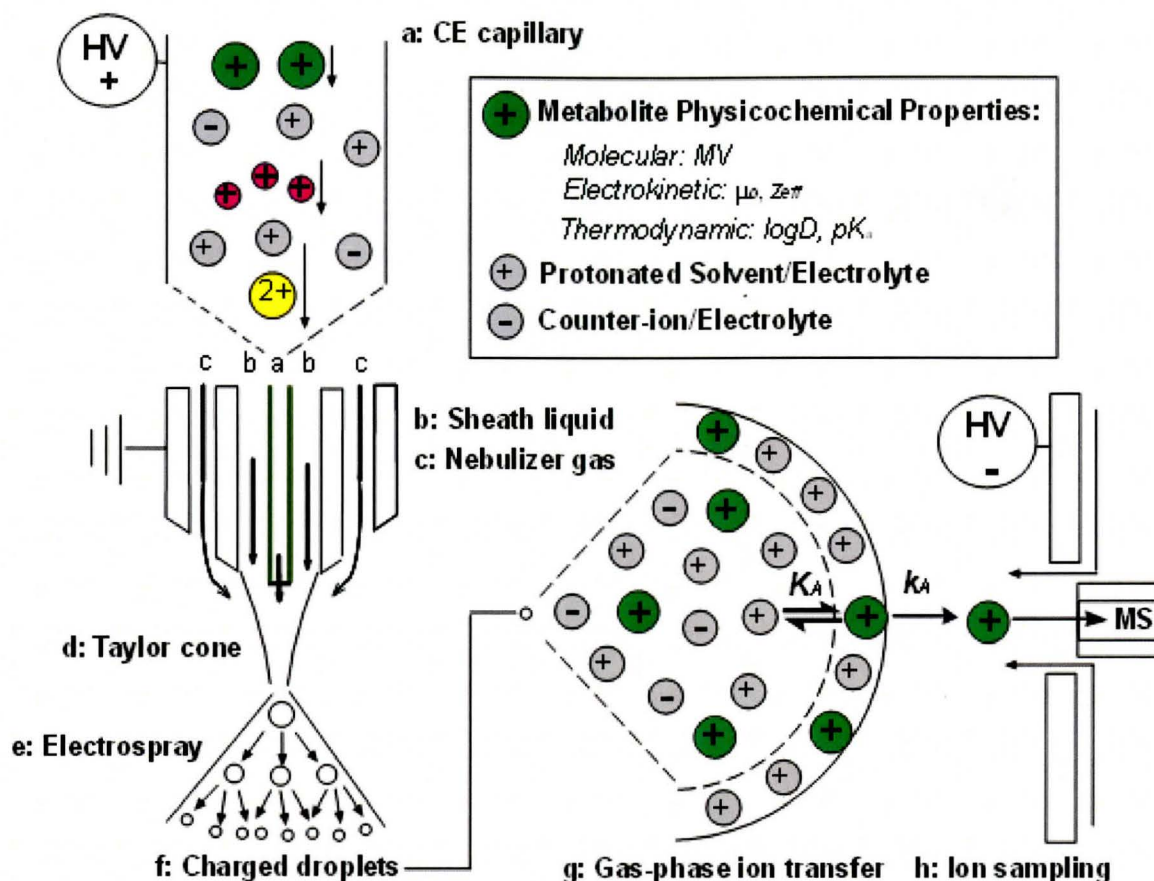


Figure 1 Multivariate model for describing gas-phase ion formation in ESI-MS, where thermodynamic (K_A) and electrokinetic factors (k_A) influence solute ionization efficiency. The relative response factor (RRF) of metabolites can be predicted based on four fundamental physicochemical properties, including MV , $\log D$, μ_0 and z_{eff} (pK_a). Note that CE separation prior to ESI-MS is important for resolving isobaric/isomeric ions, as well as minimizing ionization suppression effects, which is critical to robust predictive models in ESI-MS when applied to complex biological samples.

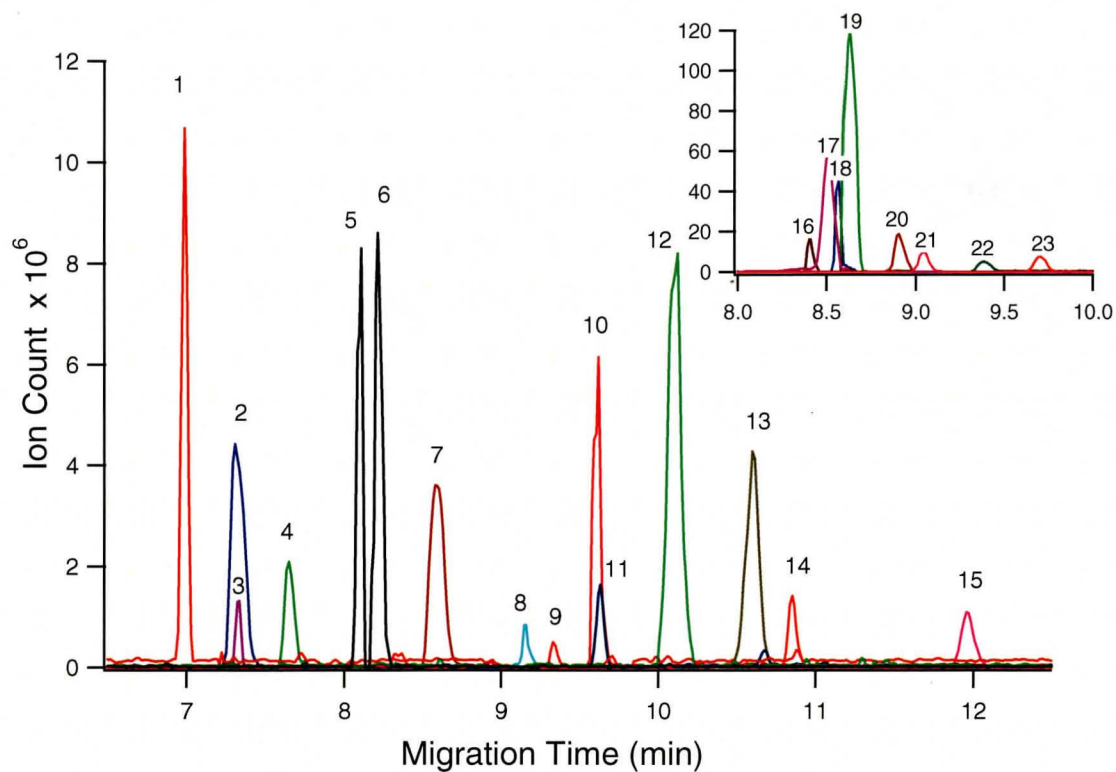


Figure 2 Overlay of extracted ion electropherograms depicting a 700-fold difference in ionization response at equimolar concentrations of 30 μM for the initial model set of metabolites by CE-ESI-MS. Analytes correspond to: 1-His, 2-TrpN, 3-Crea, 4-Sero, 5-Ile 6-Leu, 7-Cyst, 8-GABA, 9- β -Ala, 10-Glu, 11-Gn, 12-G, 13-GSSG, 14-NA, 15-Otet, 16-Carn, 17-At, 18-C2, 19-Meta, 20-Trp, 21-Phe, 22-OH-Trp, 23-NO₂-Tyr.

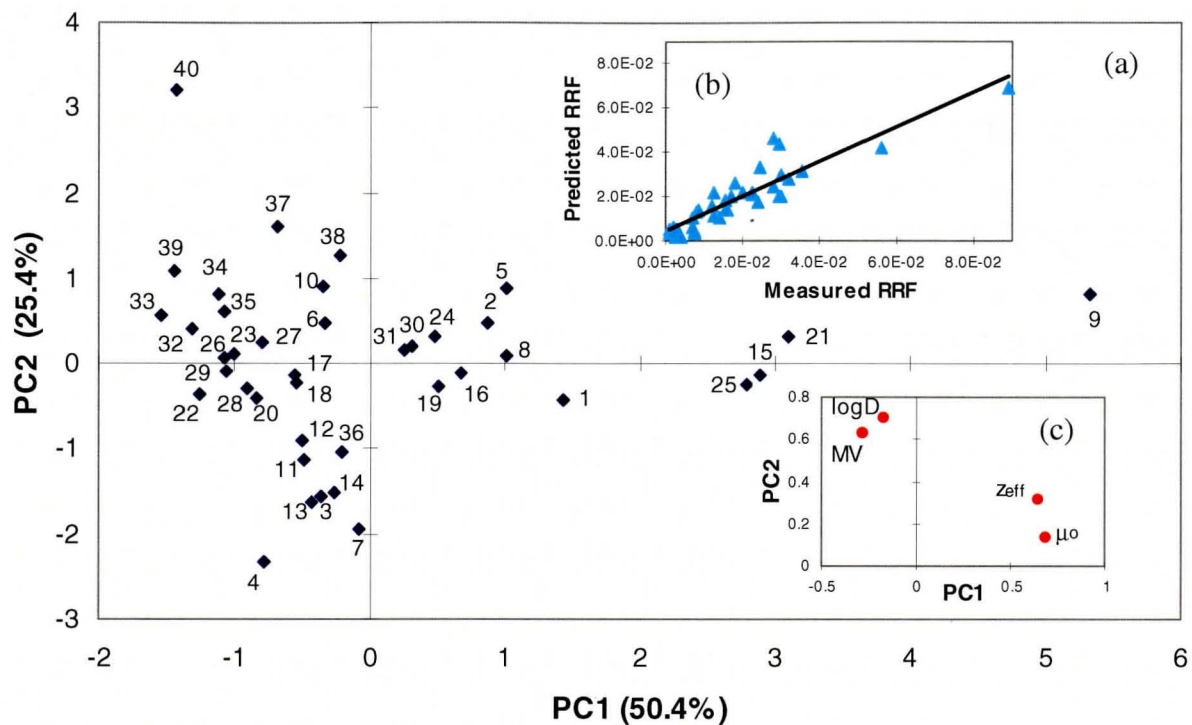


Figure 3 (a) 2D scores plot from PCA highlighting the relationship of 40 different cationic metabolites in terms of their four measured physicochemical properties (MV , $\log D$, μ_o and z_{eff}). (b) Linear correlation plot ($y = 0.792x + 0.0040$, $R^2 = 0.9144$) between measured relative response factor (RRF) by CE-ESI-MS and predicted RRF using PLS regression of the metabolite test set. (c) 2D loadings plot that highlights the relative weight of each physicochemical parameter on metabolite location in the scores plot. Model metabolites correspond to: 1-Creat, 2-Ad, 3-Thr, 4-Asp, 5-NAm, 6-NA, 7-Asn, 8-Gn, 9-His, 10-PABA, 11-HCy, 12-Val, 13-Glu, 14-Gln, 15-His, 16-TyrN, 17-Leu, 18-Ile, 19-DopN, 20-Met, 21-Lys, 22-MetS, 23-Phe, 24-TrpN, 25-Arg, 26-*m*-Tyr, 7-*o*-Tyr, 28-Tyr, 29-DOPA, 30-SerN, 31-C0, 32-Cl-Tyr, 33-NO₂-Tyr, 34-Trp, 35-OH-Trp, 36-Cyst, 37-A, 38-C2, 39-G, 40-At.

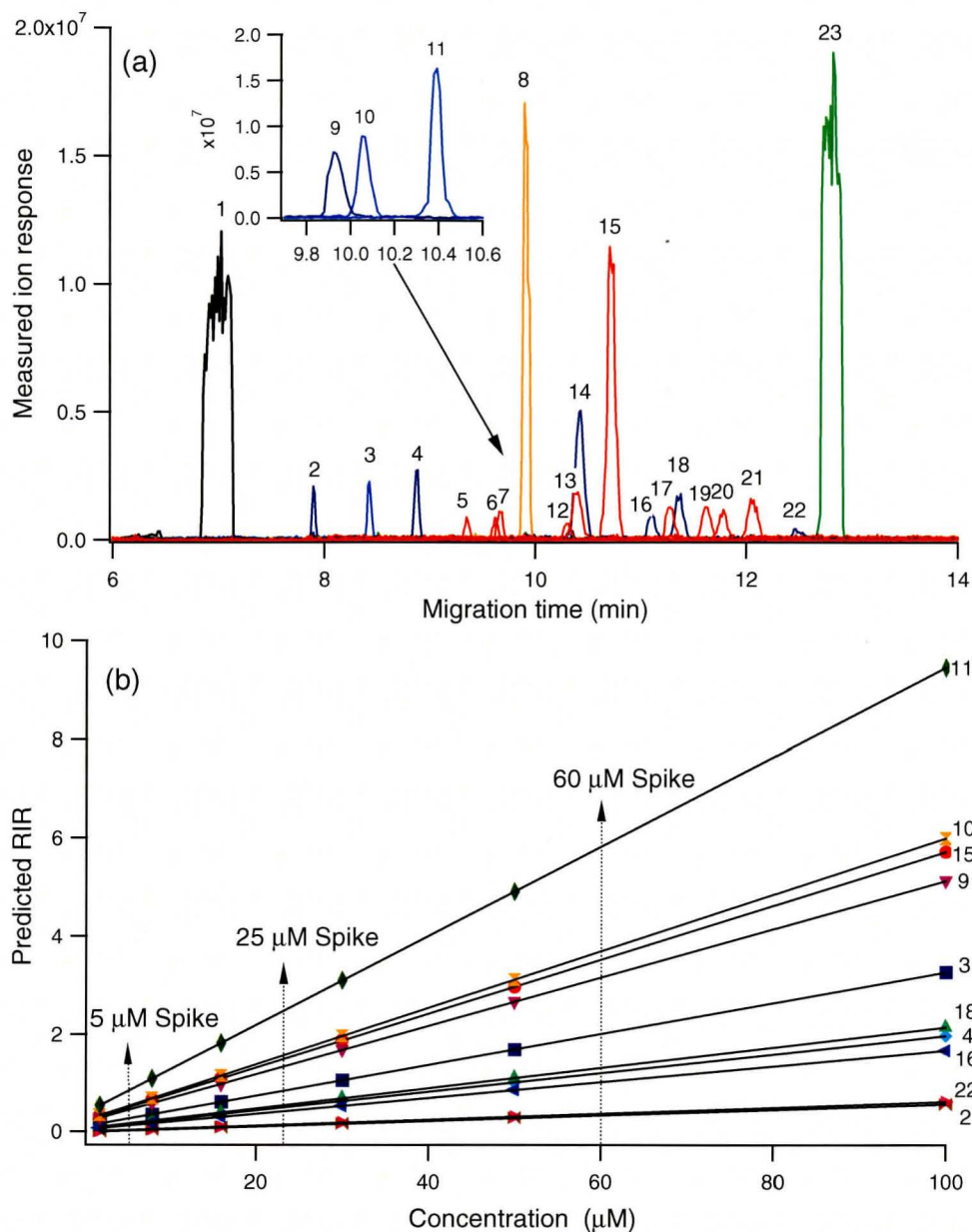


Figure 4 (a) Overlay of extracted ion electropherograms showing 5 μM of test (red traces-selected) and validation (blue traces) metabolites spiked in RBC lysates using CE-ESI-MS. (b) Virtual calibration curves generated *de novo* for validation metabolites from their predicted fundamental physicochemical parameters using PLS regression, where method accuracy was validated from recovery experiments at 5, 25 and 60 μM . Analytes correspond to: 1-Salt front (Na^+), 2-Orn, 3-Carn, 4-Me-His, 5-TyrN, 6-DopN, 7-SerN, 8-DiAla (IS), 9-C3, 10-C4, 11-C8, 12-PABA, 13-Cysta, 14-Me-A, 15-At, 16-Cit, 17-OH-Trp, 18-Thea, 19-Cl-Tyr, 20- NO_2 -Tyr, 21-G, 22-Me-Asp and 23-GSH. Note that both excess salt (1) and GSH (23) represent major co-ion interferences present in RBC lysate samples.

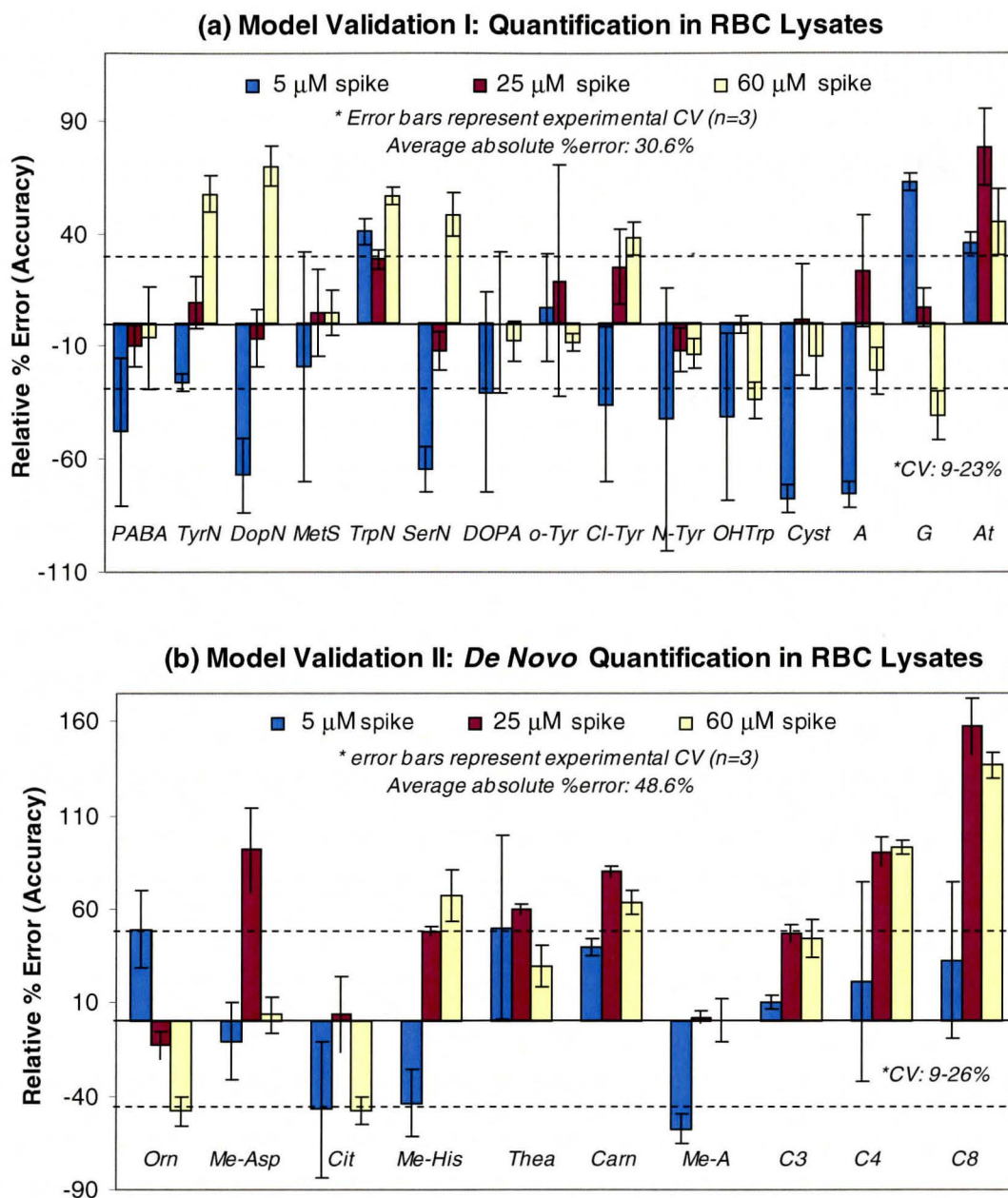


Figure 5 Recovery experiments summarizing validation studies by CE-ESI-MS of (a) 15 test metabolites and (b) 10 validation metabolites spiked in RBC lysates at low, mid and high concentrations levels using virtual calibration curves predicted by PLS regression. For validation metabolites, four physicochemical parameters were estimated *de novo* from their chemical structure via computer modeling. Overall good recovery without significant ionization suppression was achieved for cationic metabolites in +ESI with an average absolute error under 30 and 50% as indicated by the dotted lines for test and validation metabolites, respectively. The relative %error or accuracy for the recovery studies from filtered RBC lysates was determined by $(\text{predicted concentration} - \text{actual spiked concentration}) * 100 / \text{actual spiked concentration}$.

CHAPTER 3

Newborn Screening of Inborn Errors of Metabolism by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: A Second-tier Platform with Improved Specificity and Sensitivity

Abstract

The advent of electrospray-ionization tandem MS (ESI-MS/MS) has given rise to expanded newborn screening programs for early detection of multiple classes of inborn errors of metabolism (IEM) prior to the onset of clinical symptoms. However, conventional ESI-MS/MS methods are limited by poor specificity, complicated sample handling and low positive predictive outcome that can contribute to a high rate of false-positives. Herein, we report a convenient second-tiered method based on capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS) that provides a stereoselective platform for the direct analysis of amino acids and acylcarnitines from dried blood spot extracts without chemical derivatization. On-line sample preconcentration with desalting by CE-ESI-MS allowed for improved concentration sensitivity when detecting low abundance metabolites in complex biological samples without ionization suppression or isomeric/isobaric interferences. A rigorous method validation demonstrated that accurate yet precise analyses can be achieved when using a single non-deuterated internal standard. CE-ESI-MS represents a promising yet cost-effective second-tier method for newborn screening programs that is compatible with ESI-MS/MS technology in cases when improved specificity and sensitivity is warranted.

3.1 Introduction

Universal neonatal screening for inborn errors of metabolism (IEMs) was first pioneered by the work of Guthrie and co-workers in 1961, who introduced a low-cost bacterial inhibition assay for early detection of phenylketonuria (PKU) from dried blood spots collected on filter paper.¹ Until recently most public health agencies have limited mass screening to only a few classes of IEMs despite the potential benefits of early diagnosis/intervention as a way to reduce infant mortality and morbidity prior to the onset of clinical symptoms.² The cost-effectiveness of newborn screening for rare disorders has long been hampered by conventional assays that target a single biomarker without providing sufficient positive predictive outcome.^{3, 4} The advent of robust tandem MS (MS/MS) technology has given rise to expanded newborn screening programs⁵⁻¹¹ since it offers a selective, sensitive and high-throughput platform for the quantitative analysis of multiple biomarkers simultaneously, such that additional IEMs can be screened with minimal incremental costs. To date over 20 different amino acid, organic acid and fatty acid oxidation IEMs can be routinely screened using electrospray ionization-mass spectrometry¹² (ESI-MS) as documented by one of the longest running newborn screening programs in North America.¹³

Despite the remarkable success of ESI-MS in expanding neonatal screening programs, a number of significant analytical challenges still remain. Sample pretreatment of dried blood spots typically involves methanol (MeOH) extraction followed by a butanol-HCl esterification step.¹⁰ However, this procedure compromises the analysis of acid-labile metabolites (*e.g.*, Asn, Gln) relevant to known¹⁴ and recently

discovered IEMs,¹⁵ whereas butylated sulfur-containing amino acids (*e.g.*, Met, Hcy) are often not reliably quantified due to inadequate sensitivity.¹⁶ Alternative chemical derivatization strategies have been reported to further enhance ionization efficiency of amino acids¹⁷ and acylcarnitines,¹⁸ but they are time-consuming while requiring expensive stable-isotope labeled internal standards to ensure accurate recovery. A second major limitation to expanded newborn screening programs using ESI-MS is the lack of specificity for isomeric metabolites, which contributes to higher incidences of false-positives for IEMs associated with branched-chain amino acids¹⁹ and C4/C5-acylcarnitines.⁵ In addition several isobaric ions have been identified from dried blood samples that can contribute to bias, such as pivaloylcarnitine,²⁰ propionyl glycine²¹ and 4-hydroxyproline.⁹ Indeed, the nutritional status and use of drug intervention on infants and/or nursing mothers can generate unanticipated interferences which complicate mass screening initiatives.⁵ As a result, several research groups have explored second-tier diagnostic tests based on LC-ESI-MS^{16, 19, 21-23} as a way to improve assay specificity for isomeric/isobaric metabolites without chemical derivatization. However, adequate resolution of all targeted metabolites by HPLC is challenging using a single column type or elution condition due to the diverse polarity of targeted metabolites ranging from hydrophilic cationic amino acids to lipophilic long-chain acylcarnitines. For example, various modes of separation have been reported to resolve sub-classes of underivatized amino acids or acylcarnitines using reverse-phase,²² volatile ion-pair reverse-phase¹⁶ or hydrophilic interaction²⁴ chromatography. Thus, there is urgent need for the development of a simple, rapid yet selective separation platform for simultaneous resolution of

underivatized amino acid and acylcarnitine isomers that is complementary with existing ESI-MS infrastructure.

Capillary electrophoresis-mass spectrometry (CE-MS) represents a maturing high efficiency microseparation technique ideal for the resolution and detection of polar metabolites with growing applications in drug analysis, clinical diagnosis and metabolomic studies.²⁵⁻²⁷ To date most commercial CE-ESI-MS instruments use a coaxial sheath liquid interface that provides a robust format for gas-phase ionization of polar metabolites that is compatible with complex biological samples.²⁸ Recently a comparison of CE-MS and GC-MS performance for amino acid profiling in plant cell cultures²⁹ highlighted the unique benefits of CE-MS in a laboratory setting, including ease of sample preparation, higher sample throughput, lower operating costs with analogous reproducibility. In addition, on-line sample preconcentration with desalting can be readily integrated with CE-ESI-MS via the use of discontinuous electrolyte systems for boosting concentration sensitivity,³⁰ which also minimizes ionization suppression due to sample matrix effects. Despite the potential benefits of CE as a separation platform for neonatal screening,³¹⁻³⁴ to date there have been only two previous reports on the analysis of acylcarnitines using CE-ESI-MS.^{35, 36} To the best of our knowledge, this work represents the first study of CE-ESI-MS for the simultaneous analysis of twenty different targeted amino acid and acylcarnitine metabolites directly from dried blood spots with minimal sample handling. A rigorous validation of the optimized method demonstrated a simple, sensitive yet stereoselective strategy for accurate metabolite quantification without chemical derivatization or deuterated internal

standards, which serves as a powerful adjunct to MS/MS technology widely used in existing newborn screening programs.

3.2 Materials and Methods

3.2.1 Chemicals and reagents.

De-ionized water used for buffer and sample preparations was obtained using a Barnstead EASY-pure II LF ultrapure water system (Dubuque, IA). All chemicals including buffers, organic solvents, analyte standards were obtained from Sigma-Aldrich (St. Louis, MO, USA), including *L*-arginine (Arg), *L*-carnitine (C0), *O*-acetyl-*L*-carnitine (C2), *O*-myristoyl-*L*-carnitine (C14) and *O*-palmitoyl-*L*-carnitine (C16), *L*-citrulline (Cit), *L*-alanyl-*L*-alanine (diAla), *L*-glutamine (Gln), *L*-gluamic acid (Glu), reduced glutathione (GSH), *L*-histidine (His), *L*-leucine (Leu), *L*-isoleucine (Ile), *L*-allo-isoleucine (*allo*-Ile), *L*-lysine (Lys), *L*-methionine (Met), *L*-ornithine (Orn), *L*-phenylalanine (Phe), *L*-proline (Pro), *trans*-4-hydroxy-*L*-proline (4-OH-Pro), 5-oxo-*L*-proline or pyroglutamic acid (5-oxo-Pro), *L*-tryptophan (Trp), *L*-tyrosine (Tyr), and *L*-valine (Val). Other acylcarnitines standards were purchased from Larodan Fine Chemicals Inc. (Malmö, Sweden), including propionyl-*L*-carnitine HCl (C3), butyryl-*L*-carnitine HCl (C4), and octanoyl-*L*-carnitine HCl (C8).

3.2.2 Blood Spot Collection and Sample Preparation.

Blood samples from a consenting healthy adult volunteer were collected using a finger-prick method via disposable lancets (Unistik 3, Owen Munford Ltd., Georgia,

USA) and spotted on a Grade 903 Protein Saver Card (Whatman Inc., NJ, USA), which was dried overnight. A 3.2 mm (1/8 inch) disk ($\approx 3.4 \mu\text{L}$)¹² was punched out manually with a hole puncher (Whatman Inc.) from each dried blood spot into a 0.5 mL centrifuge tube that contained 100 μL of ice-cold 50:50 MeOH:H₂O with the internal standard dialanine (DiAla, 100 μM), which was extracted under sonication for 10 min. The resulting extract solution was then filtered through a 3 kDa Nanosep® centrifugal filter (Pall Life Sciences, MI, USA) at 15,000 g for 10 min prior to analysis. The resulting filtrate was then diluted with an aqueous ammonium acetate solution (400 mM, pH = 7.0) 1:1 to produce the final sample solution used for analysis (200 mM ammonium acetate, 25% MeOH, 50 μM DiAla). In most cases, dried blood samples were analyzed directly by CE-ESI-MS without additional sample pretreatment steps involving chemical derivatization or solvent/reagent evaporation. In order to demonstrate the feasibility for trace detection of low abundance metabolites in healthy controls using volume-restricted samples, the solvent of the extract was evaporated under a gentle stream of N₂ and reconstituted in 20 μL of sample solution, which provided an additional 10-fold sample enrichment prior to CE-ESI-MS analysis.

3.2.3 Apparatus and Conditions.

Separation and detection studies were performed on an Agilent CE system equipped with an XCT 3D ion trap mass spectrometer, an Agilent 1100 series isocratic pump, and a G16107 CE-ESI-MS coaxial sheath-liquid sprayer interface (Agilent Technologies, Waldbronn Germany). All separations were performed on uncoated fused-silica

capillaries with 50 μm internal diameter and 80 cm total length. Filtered dried blood spot extracts (FDBSE) were injected hydrodynamically for 75 s under low pressure (50 mbarr) followed by a 10 s injection of background electrolyte (BGE) which consisted of 1.4 M formic acid, pH 1.8 with various amounts of organic solvent. The sample injection plug length used in this work was equivalent to 6.2 cm or about 8% of the total capillary length.³⁰ All BGE solutions with organic solvent modifier were pH-adjusted to an apparent pH 1.8 with concentrated formic acid. CE separations were performed at 20°C with an applied voltage of 30 kV. The sheath liquid consisting of 1:1 MeOH:H₂O with 0.1% formic acid was supplied by the 1100 series isocratic pump at a flow rate of 10 $\mu\text{L}/\text{min}$. Nitrogen was used as both a nebulizing and a drying gas supplied at 6 psi and 10 L/min respectively, whereas helium at 6×10^{-6} mbarr was used as a damping gas for the ion trap. All MS analyses were performed using a 5 kV cone voltage in positive-ion mode at 300° C. MS data was recorded within a range of 50-750 m/z using an ultrascan mode of 26 000 m/z per second.

3.2.4 Calibration and Method Validation.

The linearity of the measured ion response for twenty targeted metabolites was examined using five replicates at five different concentrations over a 90-fold concentration range for most amino acids, whereas a 200-fold concentration range was used for Arg, 4-OH-Pro and all acylcarnitines. The concentration range selected for each metabolite was based on their ionization efficiency and biologically relevant concentration levels. The limit of detection (LOD) and limit of quantitation (LOQ) were

determined by the average background noise level (σ_B) and measured sensitivity (m) for each metabolite as derived from their calibration curve as defined by $(3*\sigma_B)/m$ and $(10*\sigma_B)/m$, respectively. Recovery studies were performed by triplicate analysis of filtered blood spot extracts spiked with all twenty metabolites at 12.5, 50, and 87.5% of the maximum concentration used for calibration. Percent recovery was calculated as $(\text{measured concentration} - \text{endogenous concentration}) * 100 / \text{added concentration}$, where the endogenous concentration was the detectable concentration of metabolite present in a normal dried blood spot extract (non-spiked sample) without solvent evaporation. Method precision was measured in terms of relative peak area (RPA) or relative migration times (RMT) ratios as normalized to an internal standard for assessment of method reliability. Intraday precision was determined by examination of five replicates for each calibration point using standard solutions (25 replicates total), five replicates of each recovery concentration spiked in filtered dried blood spot extracts (15 replicates total) or 10 replicates of a 50% recovery sample filtered dried blood spot extracts, which were performed in a single day. Interday precision was assessed by performing ten replicate measurements over three consecutive days (30 replicates total) of a filtered blood spot extract sample stored at 4°C containing metabolites spiked at 50% of the maximum calibration concentration.

3.3 Results and Discussion

3.3.1 Separation Optimization.

One of the main advantages of CE is the ability to seamlessly integrate several sample pretreatment functions directly in-capillary during electromigration as a way to enhance the performance of ESI-MS in terms of specificity, sensitivity and robustness. In this work, on-line sample preconcentration with desalting was performed using a discontinuous electrolyte system for electrokinetic focusing of weakly ionic amino acids and acylcarnitines, whereas strong electrolyte salts (*e.g.*, Na⁺) present in the original sample migrate ahead of analytes due to their high intrinsic mobility that is pH-independent.³⁰ This feature is important for improving concentration sensitivity in ESI-MS for the detection of low abundance metabolites without chemical derivatization while minimizing ionization suppression effects. *Fig. 1* highlights the main challenge of simultaneous resolution of polar amino acid isomers and lipophilic long-chain acylcarnitines by CE-ESI-MS when using an acidic aqueous BGE. Significant peak tailing and band broadening was observed for myristoyl carnitine (C14) and palmitoyl carnitine (C16) that was severe at concentrations above 10 μ M. Since the critical micelle concentration (cmc) of C14 and C16 is relatively low in aqueous solutions (cmc \approx 50 μ M for C14 in water),³⁷ self-aggregation and surface adsorption of these surface-active metabolites onto the capillary wall can result in significant peak skewing in CE-ESI-MS, as reported previously by Heinig and Henion.³⁵ The addition of increasing amounts of acetonitrile (ACN, 0-25% volume) as an organic modifier to the BGE was found to improve peak shape for the long-chain acylcarnitines, however it compromised the

resolution of Leu isomers. The use of ACN exceeding 25% to the BGE resulted in loss in Leu isomer resolution. In this study, optimum BGE conditions was determined to be using formic acid, 15% ACN, pH 1.8, since it maintained baseline resolution ($R_s > 1.1$ at 100 μM) of the three major Leu isomers (Leu, Ile, *allo*-Ile), while ensuring good peak shape and linearity in the measured responses for C14 and C16 up to biologically relevant levels of 40 μM as shown in *Fig. 1*. Stereoselective analysis of branched-chain amino acids is important for improved specificity in the diagnosis of variant forms of maple syrup urine disease (MSUD), hydroxyprolinemia, as well as newborns receiving total parenteral nutrition, which can confound conventional screening methods based on ESI-MS/MS that quantify total Leu and OH-Pro content.³⁸ Indeed, *allo*-Ile has been validated as a stereoselective pathognomonic marker of MSUD with lower cut-offs set at around 2 μM to prevent false-negative results.¹⁹ To the best of our knowledge, this is the first reported method that provides stereoselective resolution of branched-chain amino acids, while enabling the separation of long-chain acylcarnitines under a single separation format.

3.3.2 Sample Preparation and Analysis of Dried Blood Spot Extracts.

As first described by Millington *et al.*,¹¹ sample preparation for newborn screening by MS/MS has typically involved MeOH extraction of dried blood spots on Guthrie cards using either 3.2 mm or 4.75 mm hole-punched diameters prior to butyl ester derivatization. Despite the lower extraction recoveries for polar amino acids by this method,³⁹ ionization suppression effects are reduced due to lower levels of involatile salts

(*e.g.*, Na⁺) present in the MeOH extract. Alternatively, 1:1 MeOH:H₂O can be used as a solvent to improve amino acid extraction efficiency thereby reducing the need for chemical derivatization, provided that samples are adequately desalted via anion-exchange solid-phase extraction and filtered to remove protein prior to ESI-MS analysis.³⁹ In our case, a simple two-step extraction and ultracentrifugation procedure using 100 µL of 1:1 MeOH:H₂O on a 3.2 mm dried blood spot was performed in about 20 min without off-line sample desalting. Due to the lack of chemical derivatization and ionization suppression effects when using CE-ESI-MS, the reliance on expensive deuterated internal standards for metabolite quantification is not required unlike conventional analyses by direct-infusion ESI-MS. This represents an important advantage in terms of reducing long-term operating costs, while enabling accurate quantification of a wider class of metabolites in expanded newborn screening programs when deuterium-labeled analogues are commercially unavailable, such as glutarylcarnitine for the early detection of glutaric aciduria type I.⁴⁰

Fig. 2(a) depicts the direct analysis of underivatized amino acids and acylcarnitines by CE-ESI-MS from filtered dried blood spot extracts of a healthy adult volunteer, which reflect lower metabolite concentration levels as compared to normal or abnormal metabolite profiles found in newborns.⁸ Seven different targeted metabolites associated with IEMs were detected in this case, as well as several other biologically relevant metabolites, including Gln, Trp, Lys, His and reduced glutathione (GSH). Since butanolic-HCl is not used for chemical derivatization, endogenous Gln can be reliably quantified by CE-ESI-MS without bias due to a hydrolysis artifacts (*i.e.*, Glu), whereas

the isobaric ion, Lys is easily resolved prior to ionization. *Fig. 2(b)* shows the analysis of the same filtered dried blood spot extract after solvent evaporation and reconstitution in a 20 μL sample solution, which provided about a 10-fold sample enrichment relative to *Fig. 2(a)*. Unlike sample introduction procedures used in laboratories for universal newborn screening by ESI-MS/MS which often require $> 100 \mu\text{L}$ of sample for sufficient tube washing, loading and data acquisition,⁸ CE is compatible to volume-restricted samples even when using long sample injection plugs with on-line sample preconcentration with sample volumes not exceeding $\approx 100 \text{ nL}$ per injection. Due to the overall sensitivity enhancement contributed by both on-line and off-line sample preconcentration steps, more reliable quantification of low abundance metabolites was realized in dried blood extracts as shown in *Fig. 2(b)*, including additional minor amino acids, such as Phe, Cit and Glu. Note that carnitine (C0) and *O*-acetyl-*L*-carnitine (C2) were the only acylcarnitines detected due to the low nanomolar concentrations levels of medium and long-chain acylcarnitines normally present in dried blood spot extracts of healthy newborns.⁷ *Table 1* summarizes the results of direct and reconstituted filtered dried blood spot extract analyses by CE-ESI-MS which generated consistent results for the determination of low micromolar levels of amino acids and nanomolar levels of acylcarnitines without chemical derivatization and/or off-line sample desalting.

3.3.3 Method Calibration and Detection Limits.

Table 2 summarizes external calibration data for twenty different amino acid and acylcarnitine metabolites analyzed by CE-ESI-MS, where there was over a 3000-fold

difference in ionization efficiency between C8 (most responsive) and 5-oxo-Pro (least responsive) due to their widely different physicochemical properties, varying in terms of molecular volume, polarity and effective charge. Lower sensitivities were found for the long-chain acylcarnitines C14 and C16 relative to C8 as a result of their longer migration times and broader peaks. Method linearity was excellent ($R^2 > 0.996$) over a 90-200-fold concentration range for all metabolites examined with an average CV for quantification of about 9.7% based on normalized analyte peak areas relative to the internal standard, diAla. The LOD and LOQ of this method ranged from low micromolar to low nanomolar concentration levels that was metabolite-dependent as reflected by their widely different calibration sensitivity. Improved detectability was achieved for most underivatized amino acids in this work relative to LC-ESI-MS/MS with off-line desalting,³⁹ as exemplified by significantly lower LODs for Arg (\approx 500-fold), Phe/Orn (\approx 20-fold) and Tyr/Met (\approx 5-fold). Similarly, over a 7-fold enhancement in sensitivity for C2 and C8 acylcarnitines was realized without butyl esterification relative to conventional ESI-MS/MS.⁷ Although direct comparison of assay sensitivity for newborn screening is challenging due to the considerable disparity in sample pretreatment, ESI conditions and MS/MS scanning modes, our work demonstrates that CE-ESI-MS using a low resolution 3D ion-trap mass analyzer can provide superior sensitivity together with ease of sample handling as a result of the combined benefits of on-line sample preconcentration with desalting by CE. Further improvements in LOD can be achieved by multiple reaction monitoring of targeted metabolites using CE-ESI-MS/MS in order to reduce background

noise, which was demonstrated for the direct analysis of amino acids in diluted urine samples.⁴¹

3.3.4 Method Validation.

Fig. 3 depicts a recovery study performed on all twenty amino acids and acylcarnitines spiked at 12.5% of maximum calibration level during extraction of a dried blood spot. It is apparent that resolution of major isomeric and isobaric ions was achieved within a short 7 min separation window by CE-ESI-MS, ranging from the cationic amino acid Arg to the long-chain acylcarnitine C16, that is not readily feasible by conventional HPLC elution methods. In general, the migration order for amino acids is based on their effective charge to size ratio with migration times increasing from basic < neutral < aromatic \approx acidic amino acids. The significant difference in both ion response and migration time of 5-oxo-Pro relative to Pro is due to the stronger acidity (lower pK_a) of the α -carboxylic acid for 5-oxo-Pro that results in its 20-fold lower ionization efficiency in positive-ion mode ESI-MS, as well as its longer migration times in CE. In contrast, due to the common carnitine chemical/charge motif (*i.e.*, quaternary amine and β -carboxylic acid) shared by all acylcarnitines, their migration order was observed to increase linearly as a function of acyl chain length or molecular weight (MWt) ranging from C0 to C16 as reflected by a $R^2 = 0.968$ when correlating their measured RMT versus MWt. Since the electromigration behavior of metabolites can be quantitatively predicted in CE based on their fundamental physicochemical properties,

RMT represents a useful parameter for *de novo* identification of unknown metabolites that provides qualitative information complementary to MS.³⁰

Table 3 summarizes CE-ESI-MS performance in terms of accuracy and reproducibility that is relevant to newborn screening applications. Overall, excellent method accuracy was demonstrated by CE-ESI-MS based on five replicate spike and recovery experiments performed at 12.5, 50 and 87.5% maximum calibration level, which resulted in average recoveries for all twenty metabolites of 112, 101 and 96%, respectively. Overall precision for quantification was also good with average CV of 8.4% and 10.5% based on inter-day ($n = 10$) and inter-day ($n = 30$) recovery studies performed at 50% maximum calibration level. These results are consistent with published recovery studies by ESI-MS/MS⁷ and LC-ESI-MS/MS²³ when using a deuterated internal standard for each metabolite, which highlights the fact that ionization suppression effects are not significant within the separation window of cationic metabolites in CE-ESI-MS. Indeed, due to the widespread use of dried blood spot filters for sample collection in expanded newborn screening programs, it has been recognized that isotope-labeled internal standards spiked into the extract are not as accurate or precise as traditional isotope-dilution MS methods within homogenous solutions due to differences in extraction efficiency and true volume of blood.¹² Although CE typically has poor migration time reproducibility caused by run-to-run differences in the magnitude of the electroosmotic flow as compared to HPLC, inter-day relative migration times (RMT) was excellent when apparent migration times were normalized relative to an internal standard as reflected by an average CV $\approx 1.5\%$ for all targeted metabolites.

Thus, our work highlights the fact that a single non-deuterated internal standard can be used to effectively normalize variations in both ion responses and apparent migration times in CE-ESI-MS, while providing comparable assay accuracy and precision relative to conventional ESI-MS/MS or LC-ESI-MS/MS with multiple isotope-labeled internal standards. Although ESI-MS/MS is still the method of choice for high-throughput screening due to its rapid analysis times < 3 min/sample, CE-ESI-MS represents a viable yet unrecognized second-tiered method for expanded newborn screening programs that can greatly improve assay specificity and sensitivity. This is particularly relevant in the context of the low positive predictive value of routine ESI-MS/MS methods for the diagnosis of IEMs due to the high rate of false-positives,¹² which contribute to higher healthcare costs and unnecessary parental stress if preliminary ESI-MS/MS results are not adequately validated.

3.4 Conclusion.

In this study, a simple, sensitive yet stereoselective method was developed for the analysis of amino acid and acylcarnitine metabolites from filtered dried blood spot extracts by CE-ESI-MS without chemical derivatization, sample desalting or complicated sample handling. Enhanced detectability of twenty underivatized metabolites associated with IEMs was achieved when using on-line sample preconcentration with desalting by CE-ESI-MS without sample matrix effects or interferences caused by co-migrating isobaric/isomeric metabolites. Analogous method accuracy and precision was realized in spiked dried blood spot extracts by CE-ESI-MS when using a single non-deuterated

internal standard as compared to conventional methods based on ESI-MS/MS or LC-ESI-MS/MS with multiple stable-isotope labeled internal standards. CE-ESI-MS offers significant advantages as a second-tier method for expanded newborn screening programs relative to LC-ESI-MS in terms of lower operating costs due to inexpensive reagents (e.g., open tubular fused-silica capillaries, aqueous buffers), shorter total analysis times (e.g., no off-line solid-phase extraction), and improved selectivity for low-abundance metabolites and their isomers of widely different polarity under a single platform. Future work will further validate CE-ESI-MS as a complementary screening method to ESI-MS/MS for improving positive predictive outcome in expanded newborn screening programs, notably for better discrimination of premature or low birth weight newborns undergoing nutritional modification or medication without underlying IEM conditions. New advances in non-targeted metabolomic screening of IEMs⁴² also offer exciting strategies for improving existing targeted approaches for newborn screening by ESI-MS/MS for identification of multiple biomarkers associated with classical and variant forms of a disorder.

3.5 References

- (1) Guthrie, R.; Susi, A. *Pediatrics* **1963**, *32*, 338-343.
- (2) Wilcken, B.; Wiley, V.; Hammond, J.; Carpenter, K. *N. Engl. J. Med.* **2003**, *348*, 2304-2312.
- (3) Cipriano, L. E.; Rupa, A.; Zaric, G. S. *Value In Health* **2007**, *10*, 83-97.
- (4) Wilcken, B. *J. Inherit. Metab. Dis.* **2007**, *30*, 129-133.
- (5) Garg, U.; Dasouki, M. *Clin. Biochem.* **2006**, *39*, 315-332.
- (6) Schulze, A.; Lindner, M.; Kohlmuller, D.; Olgemoller, K.; Mayatepek, E.; Hoffmann, G. F. *Pediatrics* **2003**, *111*, 1399-1406.
- (7) Chace, D. H.; Hillman, S. L.; Hove, J. L. K. V.; Naylor, D. W. *Clin. Chem.* **1997**, *43*, 2106-2113.

- (8) Rashed, M. S.; Bucknall, M. P.; Little, D.; Awad, A.; Jacob, M.; Alamoudi, M.; Alwattar, M.; Ozand, O. T. *Clin. Chem.* **1997**, *43*, 1129-1141.
- (9) Chace, D. H.; Hillman, S. L.; Millington, D. S.; Kahler, S. G.; Roe, C. R.; Naylor, E. W. *Clin. Chem.* **1995**, *41*, 62-68.
- (10) Chace, D. H.; Millington, D. S.; Terada, N.; Kahler, S. G.; Roe, C. R.; Hofman, L. F. *Clin. Chem.* **1993**, *39*, 66-71.
- (11) Millington, D. S.; Kodo, N.; Norwood, D. L.; Roe, C. R. *J. Inher. Metabol. Dis.* **1990**, *13*, 321-324.
- (12) Chace, D. H.; Kalas, T. A.; Naylor, E. W. *Clin. Chem.* **2003**, *49*, 1797-1817.
- (13) Frazier, D. M.; Millington, D. S.; McCandless, E. E.; Koeberl, D. D.; Weavil, S. D.; Chaing, S. H.; Muenzer, J. *J. Inher. Metab. Dis.* **2006**, *29*, 76-85.
- (14) Trinh, M.-U.; Blake, J.; Harrison, J. R.; Gerace, R.; Ranieri, E.; Fletcher, J. M.; Johnson, D. W. *Clin. Chem.* **2003**, *49*, 681-684.
- (15) Haberle, J.; Gorg, B.; Toutain, A.; Rutsch, F.; Benoist, J.-F.; Gelot, A.; Suc, A.-L.; Koch, H. G.; Schliess, F.; Haussinger, D. *J. Inher. Metabol. Dis.* **2006**, *29*, 352-358.
- (16) Piraud, M.; Vianey-Saban, C.; Bourdin, C.; Acquaviva-Bourdain, C.; Boyer, S.; Elfakir, C.; Bouchu, D. *Rapid Commun. Mass. Spectrom.* **2005**, *19*, 3287-3297.
- (17) Johnson, D. W. *Rapid Commun. Mass. Spectrom.* **2001**, *15*, 2198-2205.
- (18) Minkler, P. E.; Ingalls, S. T.; Hoppel, C. L. *Anal. Chem.* **2005**, *77*, 1448-1457.
- (19) Oglesbee, D.; Sanders, K. A.; Lacey, J. M.; Magera, M. J.; Casetta, B.; Strauss, K. A.; Totorelli, S.; Rinaldo, P.; Matern, D. *Clin. Chem.* **2008**, *54*, 542-549.
- (20) Shigematsu, Y.; Hata, I.; Kikawa, Y.; Mayumi, M.; Tanaka, Y.; Sudo, M.; Kado, N. *J. Chromatogr. B* **1999**, *731*, 97-103.
- (21) Zoppa, M.; Gallo, L.; Zacchello, F.; Giordano, G. *J. Chromatogr. B* **2006**, *831*, 267-273.
- (22) Ghoshla, A. K.; Guo, T.; Soukhova, N.; Soldin, S. J. *Clin. Chim. Acta* **2005**, *358*, 104-112.
- (23) Qu, J.; Wang, Y.; Luo, G.; Wu, Z.; Yang, C. *Anal. Chem.* **2002**, *74*, 2034-2040.
- (24) Jauregui, O.; Sierra, A. Y.; Carrasco, P.; Gratacos, E.; Hegardt, F. G.; Casals, N. *Anal. Chim. Acta* **2007**, *599*, 1-6.
- (25) Gaspar, A.; Englmann, M.; Fekete, A.; Harir, M.; Schmitt-Kopplin, P. *Electrophoresis* **2008**, *29*, 66-79.
- (26) Monton, M. R. N.; Soga, T. *J. Chromatogr. A* **2007**, *1168*, 237-246.
- (27) Servais, A. C.; Crommen, J.; Fillet, M. *Electrophoresis* **2006**, *27*, 2616-2629.
- (28) Manisali, I.; Chen, D. D. Y.; Schneider, B. B. *Trends in Anal. Chem.* **2006**, *25*, 243-256.
- (29) Williams, B. J.; Cameron, C. J.; Workman, R.; Broeckling, C. D.; Sumner, L. W.; Smith, J. T. *Electrophoresis* **2007**, *28*, 1371-1379.
- (30) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. *Anal. Chem.* **2007**, *79*, 403-415.
- (31) Senk, P.; Kozak, L.; Foret, F. *Electrophoresis* **2004**, *25*, 1447-1456.
- (32) Friedecky, D.; Tomkova, J.; Maier, V.; Janostakova, A.; Prochazka, M.; Adam, T. *Electrophoresis* **2007**, *28*, 373-380.

- (33) Boulat, O.; McLaren, D. G.; Arriaga, E. A.; Chen, D. D. Y. *J. Chromatogr. A* **2001**, *754*, 217-228.
- (34) Garcia, A.; Barbas, C.; Aguilar, R.; Castro, M. *Clin. Chem.* **1998**, *44*, 1905-1911.
- (35) Heinig, K.; Henion, J. *J. Chromatogr. B* **1999**, *735*, 171-188.
- (36) Elgstoen, K. B. P.; Zhao, J. Y.; Anacleto, H. F.; Jellum, E. *J. Chromatogr. A* **2001**, *914*, 265-275.
- (37) Ho, J. K.; Duclos, R. I.; Hamilton, J. A. *J. Lipid Res.* **2002**, *43*, 1429-1439.
- (38) Schadewaldt, P.; Bodner-Leidecker, A.; Hammen, H.-W.; Wendel, U. *Clin. Chem.* **1999**, *45*, 1734-1740.
- (39) Nagy, K.; Takats, Z.; Pollreisz, F.; Szabo, T.; Vekey, K. *Rapid Commun. Mass. Spectrom.* **2003**, *17*, 983-990.
- (40) Lindner, M.; Ho, S.; Fang-Hoffmann, J.; Hoffmann, G. F.; Kolker, S. *J. Inherit. Metab. Dis.* **2006**, *29*, 378-382.
- (41) Soga, T.; Kakazu, Y.; Robert, M.; Tomita, M.; Nishioka, N. *Electrophoresis* **2004**, *25*, 1964-1972.
- (42) Wirkoff, W. R.; Gangaiti, K. A.; Barshop, B. A.; Siuzdak, G. *Clin. Chem.* **2007**, *53*, 2169-2176.

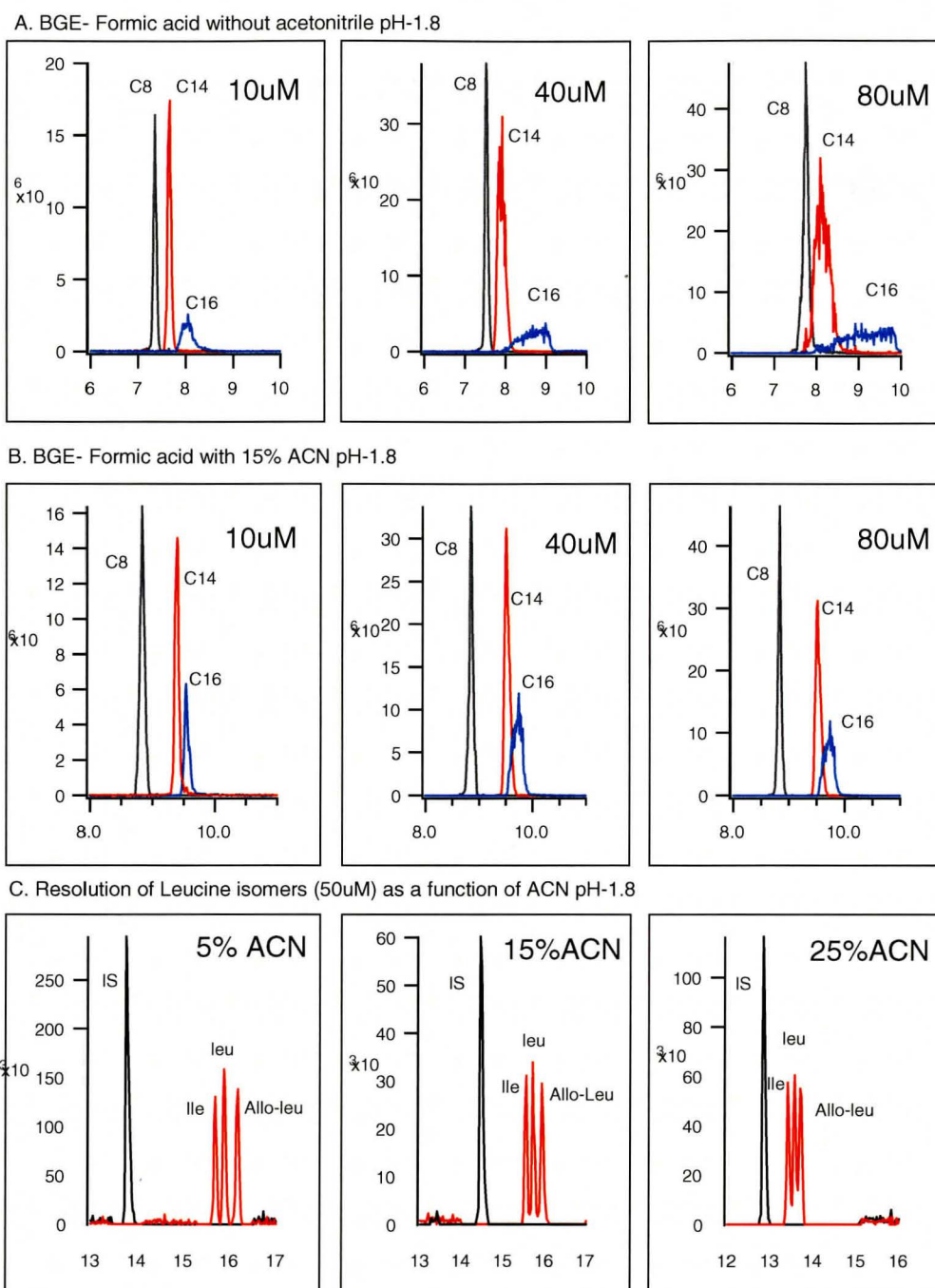


Figure 1 A series of extracted ion electropherograms depicting the impact of organic solvent and solute concentration on analyte peak shape and resolution in CE-ESI-MS, where: a) long-chain acylcarnitines without organic modifier, b) with organic modifier, c) stereoselective leucine isomer resolution at various ACN concentrations.

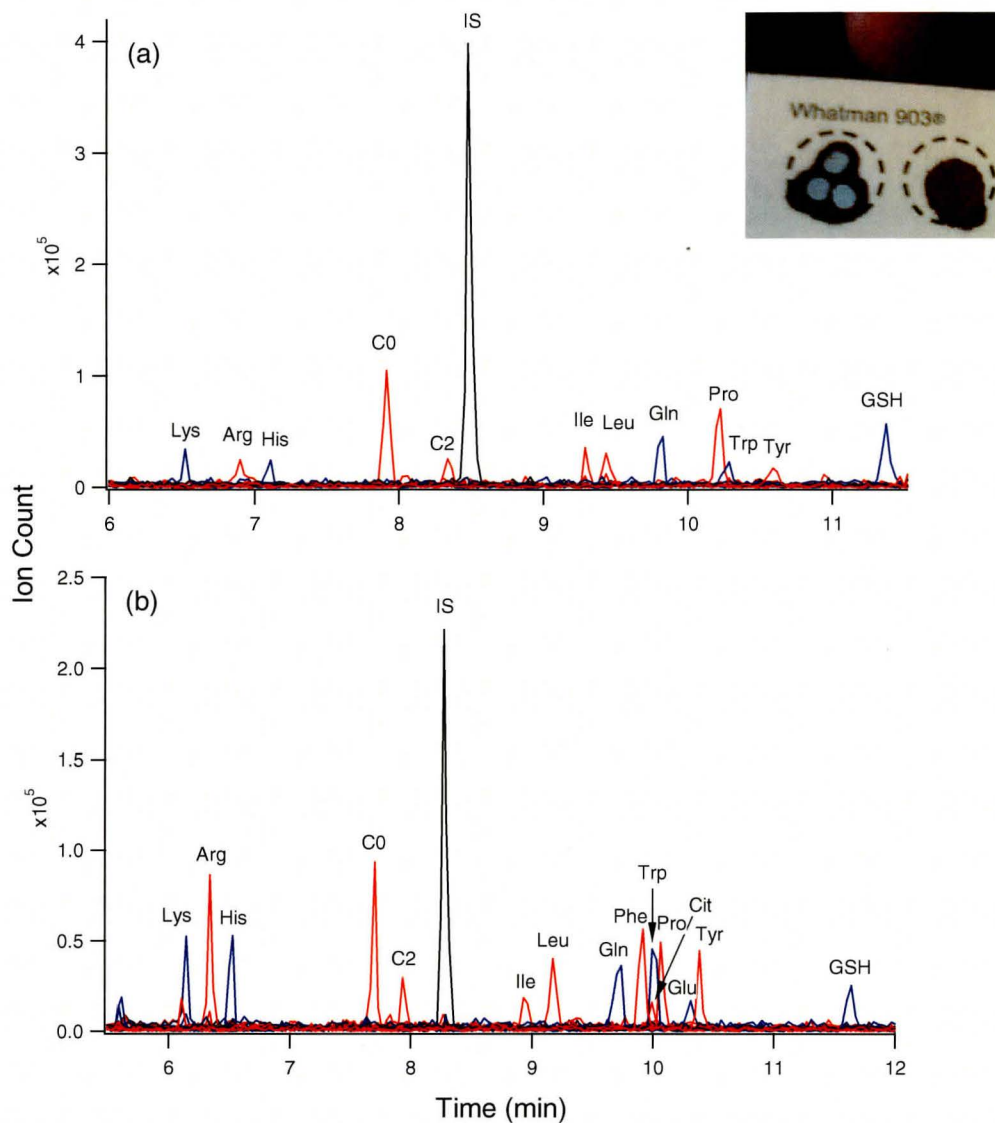


Figure 2 Extracted ion electropherograms depicting a normal profile of amino acids and acylcarnitines detected from filtered dried blood spots (3.2 mm) extracted in 100 μL of 1:1 MeOH:H₂O from a healthy adult volunteer by CE-ESI-MS. Samples were either (a) analyzed directly after ultracentrifugation and a 2-fold dilution in sample solution or (b) the extract was evaporated and reconstituted in 20 μL of sample solution to provide a 10-fold off-line sample enrichment prior to analysis. BGE: Formic acid, 15% ACN, pH 1.8; Sample: 200 mM ammonium acetate, 25% MeOH, pH 7.0. All other conditions are described in the Apparatus and Conditions section. Analyte traces depicted in red represent metabolites targeted in current expanded newborn screening, whereas blue traces are other detectable yet biologically relevant metabolites.

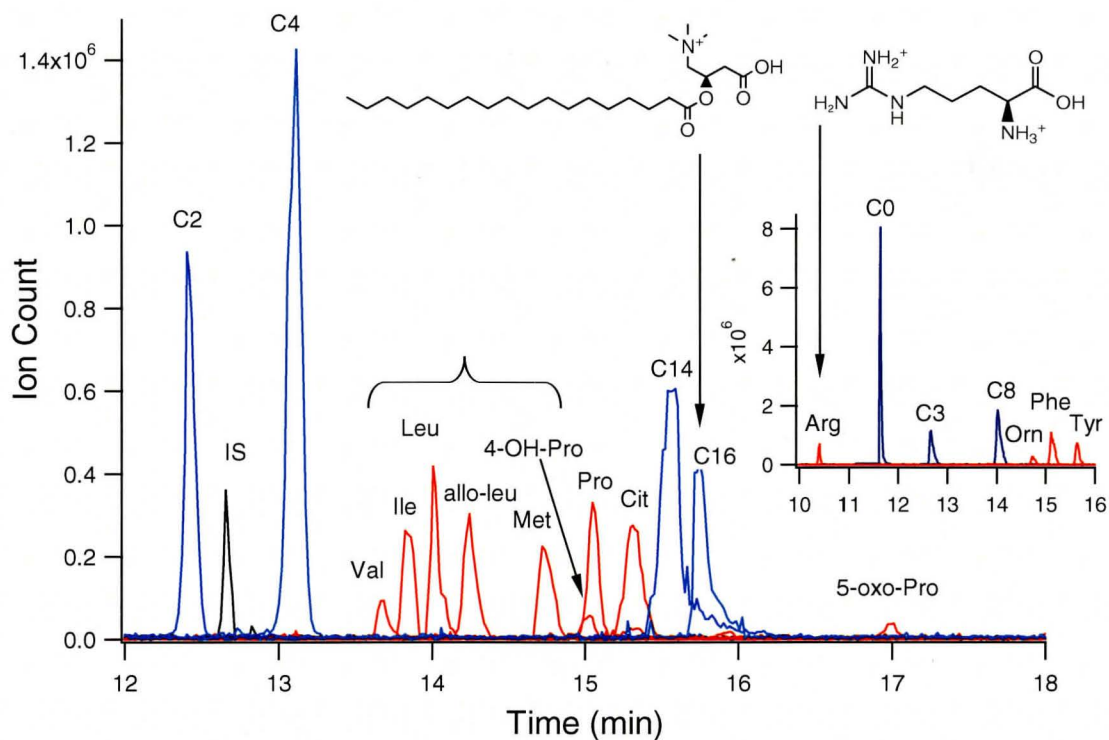


Figure 3. Extracted ion electropherogram depicting recovery studies for a filtered dried blood spot extract spiked at 12.5% maximum calibration level with twenty different amino acid and acylcarnitine metabolites commonly targeted in expanded newborn screening programs. Note that metabolites of widely different polarity can be analyzed simultaneously by CE-ESI-MS (*e.g.*, Arg and C16) while providing resolution of isomeric and isobaric analytes (*e.g.*, Leu, Ile, *allo*-Ile and 4-OH-Pro). Conditions are similar to described in Fig.2.

Table 1. Normal concentration levels of targeted amino acids and acylcarnitines detected in dried blood spots by CE-ESI-MS without and with solvent evaporation.

Metabolite	Extract Conc. ^a (μM)	Blood Conc. ^b (μM)	Extract Conc. ^c (μM)	Blood Conc. ^b (μM)
Arg	(0.20 \pm 0.01)	(12 \pm 1)	(2.1 \pm 0.1)	(12 \pm 1)
Cit	n.d.	--	trace	--
Phe	n.d.	--	(2.4 \pm 0.5)	(14 \pm 3)
Tyr	(0.32 \pm 0.05)*	(19 \pm 3)	(2.8 \pm 0.1)	(16 \pm 0.6)
Ile	(0.21 \pm 0.01)*	(12 \pm 1)	(2.1 \pm 0.1)	(12 \pm 0.6)
Leu	(0.23 \pm 0.4)*	(13 \pm 1)	(1.8 \pm 0.1)	(11 \pm 0.3)
Pro	(1.6 \pm 0.1)	(94 \pm 6)	(16 \pm 1.0)	(91 \pm 6)
C0	(0.12 \pm 0.01)	(7.0 \pm 0.6)	(1.2 \pm 0.1)	(7.1 \pm 0.2)
C2	(0.03 \pm 0.01)*	(1.7 \pm 0.6)	(0.12 \pm 0.02)	(0.7 \pm 0.1)

^a Direct analysis of dried blood extracts by CE-ESI-MS without solvent evaporation (n=3)

^b Estimated concentration of metabolites in whole blood assuming $\approx 3.4 \mu\text{L}$ in a 3.2 mm dried blood spot volume¹²

^c Analysis of 10-fold preconcentrated dried blood extract by CE-ESI-MS with solvent evaporation (n=3)

* Detectable metabolite concentration below LOQ

Table 2. External calibration data for 20 targeted amino acids and acylcarnitines relevant to newborn screening of IEMs using CE-ESI-MS without chemical derivatization.

Metabolite	Sensitivity ^a (μM) ⁻¹	Linear range ^b (μM)	Linearity (R^2)	Precision ^c (CV %)	LOD ^d (μM)	LOQ ^d (μM)
C8	1.332	0.2-40	0.9993	10.3	0.008	0.02
C14	0.9322	0.2-40	0.9986	11.9	0.08	0.2
C4	0.8144	0.2-40	0.9991	9.0	0.01	0.03
C3	0.5179	0.2-40	0.9998	8.3	0.01	0.03
C2	0.3699	0.2-40	0.9999	8.1	0.03	0.10
C16	0.2466	1-40	0.9961	15.0	0.3	1.0
Arg	0.1849	0.2-40	0.9995	7.4	0.06	0.20
C0	0.1688	5-450	0.9994	10.9	0.03	0.13
Phe	0.0515	5-450	0.9999	9.8	0.15	0.5
Tyr	0.0434	5-450	0.9991	9.3	0.5	1.5
4-OH-Pro	0.0336	0.2-40	0.9993	9.5	0.06	0.20
Cit	0.0165	5-450	0.9992	6.8	2.0	5.0
allo-Ile	0.0143	5-450	0.9999	10.7	2.0	5.0
Orn	0.0138	5-450	0.9999	11.6	2.0	5.0
Leu	0.0125	5-450	0.9991	10.0	2.0	5.0
Pro	0.0125	5-450	0.9995	7.3	0.9	2.5
Met	0.0126	5-450	0.9999	8.0	2.0	5.0
Ile	0.0102	5-450	0.9990	9.1	2.0	5.0
Val	0.00277	15-450	0.9953	11.7	5.0	15.0
5-oxo-Pro	0.00162	15-450	0.9986	11.1	5	15

^a Sensitivity represents the slope of a linear calibration plot based on average relative peak areas of analyte to internal standard (50 μM DiAla).

^b Linear range based on a five-point calibration plot performed in triplicate ($n=3$).

^c Precision based on the average CV derived from relative peak areas over entire range of calibration.

^d LOD ($S/N \approx 3$) and LOQ ($S/N \approx 10$) were defined as minimal signals detectable above background noise for serially diluted standards.

Table 3. Validation of CE-ESI-MS for accurate and reproducible quantification of amino acid and acylcarnitine metabolites derived from dried blood samples without solvent evaporation or deuterated internal standards.

Metabolites	MH ⁺ (m/z)	Recovery at 12% level ^a (CV %)	Recovery at 50% level ^a (CV %)	Recovery at 88% level ^a (CV %)	Intra-day precision ^b (CV %)	Inter-day precision ^b (CV %)	RMT ^c (Mean)	Inter-day precision ^c (CV %)
Pro	116	123	96	92	8.2	10.7	1.208	1.9
Val	118	127	126	116	9.2	12.5	1.079	1.1
5-oxo-Pro	130	131	82	74	10.1	11.5	2.152	5.8
4-OH-Pro	132	102	85	94	9.4	10.9	1.364	2.5
Ile	132	139	127	106	5.7	9.6	1.092	1.1
Leu*	132	99	126	107	6.6	9.3	1.107	1.2
allo-Ile	132	105	97	98	9.3	10.2	1.125	1.3
Orn	133	110	100	82	10.2	11.0	1.171	1.6
Met	150	113	99	90	6.7	9.7	1.172	1.5
C0*	162	96	93	79	7.2	10.9	0.917	0.50
Phe	166	126	93	89	9.1	9.6	1.204	1.6
Arg*	175	94	92	90	8.7	9.3	0.817	0.71
Cit	176	123	116	103	7.2	10.3	1.216	1.4
Tyr*	182	109	96	91	9.4	9.3	1.248	1.7
C2*	204	126	104	94	8.7	10.8	0.973	0.27
C3	218	113	96	95	8.0	9.1	0.999	0.13
C4	232	109	92	95	9.4	9.3	1.035	0.22
C8	288	119	133	107	7.5	12.5	1.099	0.48
C14	372	91	98	127	10.2	12.9	1.210	0.97
C16	400	95	72	94	7.7	10.4	1.343	3.7
Average	--	112	101	96	8.4	10.5	--	1.5

^a Mean recovery after spiking analyte standards at 12%, 50% and 88% maximum calibration level based on five replicates (n=5) for assessing method accuracy. Endogenous metabolites detected in normal adult dried blood extracts denoted by *.

^b Reproducibility data based on inter-day precision using ten replicate analyses at 50% maximum calibration level (n=10), whereas inter-day precision were

based on grand mean of ten replicate analyses performed over three consecutive days (n=30).

^c Average RMTs normalized to diAla and RMT precision based on ten replicate measurements performed over three consecutive days (n = 30).

CHAPTER 4: Future Outlook

4.1 Research Plans

For the work presented in this thesis, there are a number of areas which warrant further investigation. Firstly, an extension of the initial model set of compounds used in the electrospray response prediction model should be investigated in order to further improve its predictive capabilities for metabolites which possess properties that lie on the farthest extremes of the ranges examined with the model test set. Also, another area regarding this work which requires further investigation is the application of this model to anionic compounds in negative-ion MS mode, which has been observed to have significant differences in mechanism and performance relative to positive-ion mode MS. Yet another aspect of this work which can be studied further is the impact of various solvents, buffers and additives can have on the response of compounds in the ESI-MS process and how will these variables impact response prediction.

In regards to the work done on creating a new inherited metabolic disorder screening protocol a number of issues of this research also warrant further investigation. One area which needs to be demonstrated is can this model accurately determine the concentrations of compounds which were not available during the initial study. Another area of interest regarding this work is to apply the method to infants who have been diagnosed with an IEM in order to create a library of metabolite profiles corresponding to the various phenotypes of the diagnosed disorder. Also, a virtual library of expected metabolite profiles can and should be generated using predictive models which can be used to

simulate relative migration times in CE and their relative responses in ESI-MS based on work presented in this thesis.

4.2 Expansion of Model Set and Predicting Anionic Compounds in Negative-ion Mode

Since the initial model was limited to the relatively small number of compounds which were available at the time of the study, the model does not encompass the full range of analytes which may be of interest to researchers. This is an especially important limitation for any interested researchers in the pharmaceutical, polymer, and natural products industries since it is common that the products they are studying can have properties which lie well beyond the initial ranges used in the model. Examples of this are polymers with high molecular volumes and multiple charges, as well as natural products derived from oils and fat which can be exceptionally hydrophobic.

Also since this model only examined cationic metabolites, it has only studied half of the compounds which can ultimately be detected by ESI-MS. While it would be reasonable that many physicochemical parameters which impact ionization in positive-ion mode would also impact ionization in negative-ion mode, there has been previous research which suggests that the ionization mechanism in negative-ion mode is distinct [1,2]. Therefore, since the mechanism may differ between the two modes, further investigation must be done in order to determine whether the developed model can also be applied to anions (*e.g.*, nucleotides, organic acids, sugar phosphates etc.) in negative-ion MS mode or if the model will need to be refined for these metabolites.

4.3 Impact of Various Solvents and Other Additives to Ionization Matrix

Since it has been determined that various molecular parameters have a significant impact on ionization response in ESI it would be worth investigating the impact that experimental variables would have on response prediction since these variables would have a direct impact on the properties used in the model. Experimental variables such as organic solvents which may be required in the BGE (as used in Ch. 3 of this thesis) could impact how compounds partition in charged droplets, which in turn would impact ionization efficiency. While any impact from changing the matrix should be compensated for by the use of an internal standard, it is unlikely that all compounds will be impacted with the same magnitude which may alter the relative response factor.

4.4 Investigation of Additional Biomarkers for Metabolic Disorder Screening

It has been shown in chapter 3 that some relevant biomarkers of inherited metabolic disorders require that the conditions of the analysis be altered in order to obtain results which are both reproducible and accurate. For the twenty targeted metabolites associated with IEMs examined in this thesis, these experimental conditions are acceptable however there are a number of metabolites which were unavailable at the time of the study and it would be prudent to ensure that further modifications to the method are not required in order to quantify them from dried blood spots. This is especially true of very long chain acylcarnitines such as sterylacarnitine, as well as acylcarnitine isomers (e.g., butyryl and isobutyryl acylcarnitine), which will require stereoselective resolution while minimizing peak tailing due to surface adsorption.

4.5 Analysis of Blood Spot Extraction Protocol

For the procedure outlined in *Ch. 3*, it was found that a 50:50 mixture of cold methanol:water produced acceptable extraction of amino acids from dried blood spots. However, since no long chained hydrophobic acylcarnitines were expected to be found in the samples studied, it is unclear whether these conditions would be adequate for the extraction of this class of metabolite. Further investigation of the extraction efficiency of the method outlined should be carried out systematically to determine if this is indeed the optimal protocol or if further modification in solvent composition is warranted.

4.6 Analysis of Samples from Diagnosed Patients

To definitively validate the diagnostic capability of the developed method for metabolic screening real samples from patients which have been previously diagnosed will need to be obtained and analyzed. At the time of method development, such samples were not available. Also the analysis of samples taken from patients diagnosed with a metabolic disorder will aid in the creation of a library of metabolic profiles which will aid physicians in diagnosis of not only the primary disorder but in the assessment of phenotype which will determine what treatment should be administered.

4.7 Simulation of Metabolic Profiles in Diseased Patients

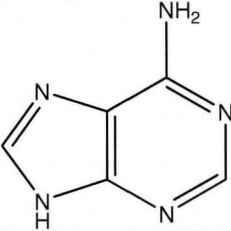
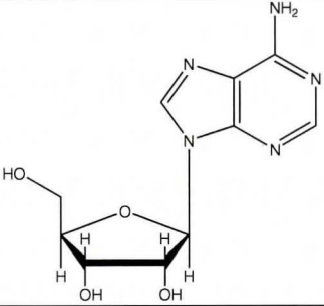
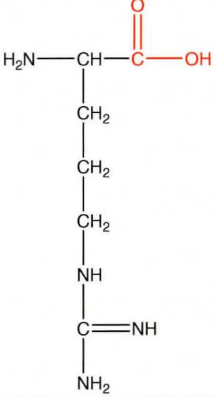
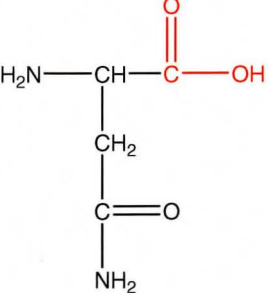
By combining the response prediction model presented in chapter 2 of this thesis in conjunction with a model developed by Lee *et. al.* [3] which is capable of predicting metabolite relative migration times (RMTs) by CE, virtual electrophoretic profiles of

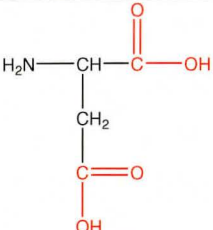
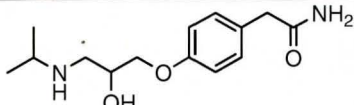
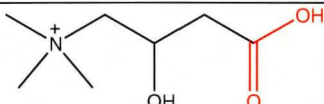
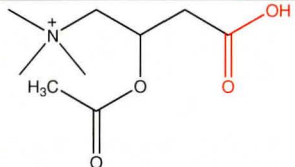
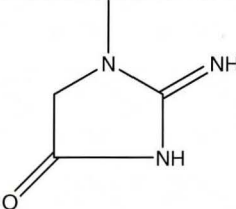
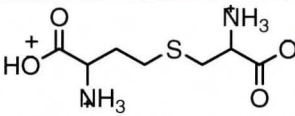
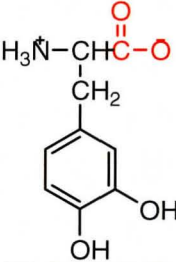
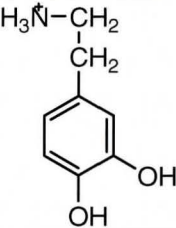
various IEMs can be created which can act as a template for professionals to use when attempting to assess disorder types and phenotypes. By having such a library available, comparisons between actual and predicted results can be done quickly thereby allowing rapid response by the physician in implementing a treatment plan.

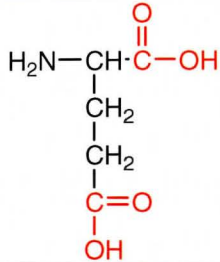
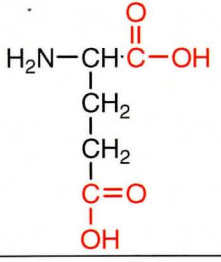
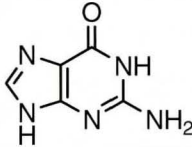
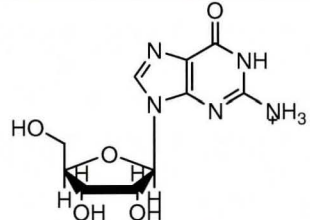
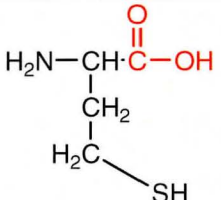
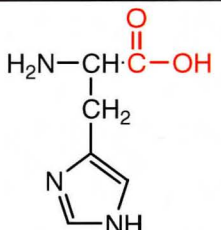
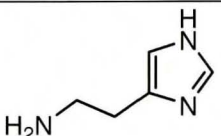
4.8 References

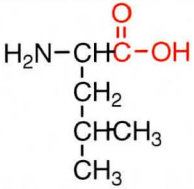
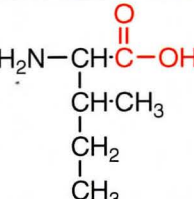
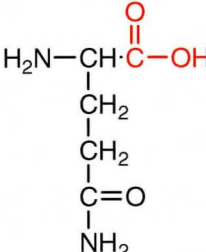
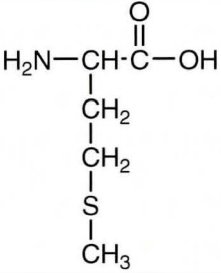
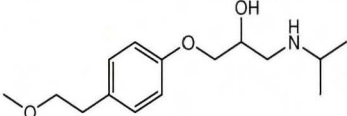
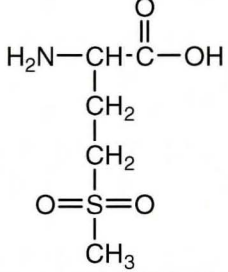
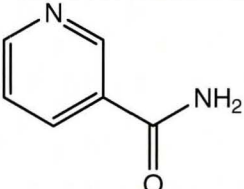
1. R.B. Cole, A. Kamel-Harrata, *J. Am. Soc. Mass Spectro.*, **1993**, 4, 546
2. T. Henriksen, R.K. Juhler, B. Svensmark, N.B. Cech, *J. Am. Soc. Mass Spectro.* **2005**, 16, 446
3. R. Lee, A.S. Ptolemy, L. Niewczas, P. Britz-McKibbin, *Anal. Chem.*, **2007**, 79, 403

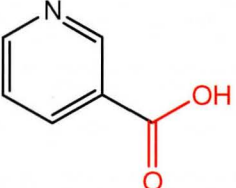
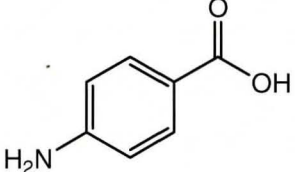
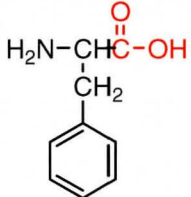
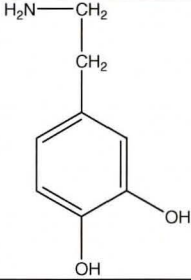
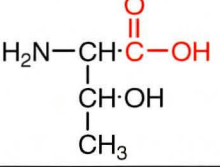
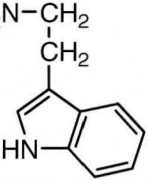
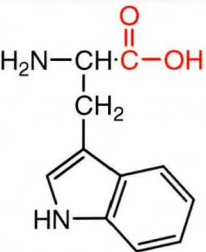
Appendix I 40 different model cationic metabolites and their four experimentally measured physicochemical parameters used in PCA modeling and PLS regression

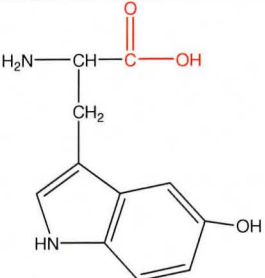
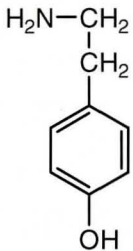
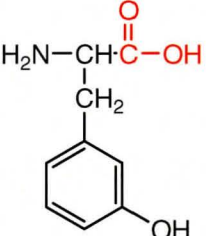
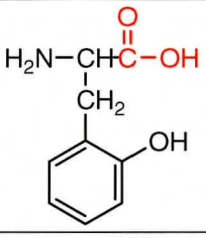
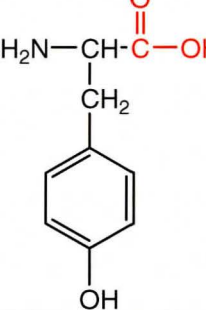
<i>Metabolite</i>	<i>MV</i> (\AA^3)	<i>LogD</i> ($\pm 1\sigma$)	μ_o ($\text{cm}^2/\text{V}\cdot\text{s}$) ($\pm 1\sigma$)	<i>pKa</i> ($\pm 1\sigma$)	<i>Chemical Structure</i>	<i>MH⁺</i> <i>fragments</i> (<i>relative</i> <i>intensity</i> <i>ratio</i>)
Adenine	89.4	-0.08 (0.01)	3.93e-4 (4.92e-6)	n/a		136
Adenosine	195	-1.01 (0.05)	2.50e-4 (3.3e-6)	n/a		268
Arginine	142	-4.20 (0.20)	4.97e-4 (2.0e-5)	1.78 (0.08)		175
Asparagine	95.1	-3.82 (0.05)	2.47e-4 (3.5e-6)	2.31 (0.04)		133


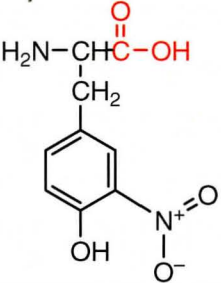
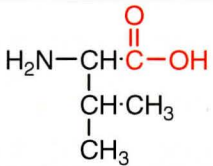
Aspartic acid	93.8	-3.89 (0.10)	2.33e-4 (3.6e-6)	1.89 (0.12)		134
Atenolol	253	0.10 (0.01)	2.25e-4 (3.4e-6)	n/a		267.3
Carnitine	155	-2.49 (0.10)	3.23e-4 (3.3e-6)	n/a		162
Carnitine, Acetyl-(C2)	197	-1.81 (0.10)	3.11e-4 (1.6e-5)	n/a		204
Creatinine	88.3	-1.76 (0.10)	4.49e-4 (2.5e-5)	n/a		114
Cystathionine	174	-4.80 (est)	2.84e-4 (3.26e-6)	2.47 (0.13)		223
L-Dopa	147	-1.96 (0.05)	2.05e-4 (6.1e-6)	2.22 (0.12)		198
Dopamine	125	-2.37 (0.21)	3.30e-4 (1.8e-5)	n/a		154,137 (1:1)

Glutamine	114	-3.64 (0.06)	2.35e-4 (4.4e-6)	2.37 (0.08)		147,130 (5:2)
Glutamic acid	113	-3.69 (0.07)	2.34e-4 (5.4e-6)	2.19 (0.09)		148,130 (9:2)
Guanine	95.6	-0.95 (0.05)	4.095e-4 (3.5e-6)	n/a		152
Guanosine	207	-1.85 (0.10)	1.79e-4 (3.6e-6)	n/a		284,152 (1:1)
Homocystine	105	-2.56 (0.50)	2.35e-4 (6.6e-6)	2.25 (0.13)		136
Histidine	118	-3.32 (0.05)	4.74e-4 (1.7e-5)	1.99 (0.02)		156
Histamine	98.5	-2.31 (0.10)	7.36e-4 (9.6e-6)	n/a		112

Leucine	124	-1.52 (0.01)	2.49e-4 (6.9e-6)	2.35 (0.01)		132,86 (5:2)
Isoleucine	125	-1.71 (0.01)	2.49e-4 (6.3e-6)	2.32 (0.02)		132,86 (5:2)
Lysine	131	-3.05 (0.05)	5.41e-4 (3.1e-5)	1.70 (0.08)		147,130 (7:2)
Methionine	127	-1.88 (0.09)	2.32e-4 (4.9e-6)	2.14 (0.09)		150
Metoprolol	279	1.55 (0.10)	2.23e-4 (4.7e-5)	n/a		268
Methionine sulfone	138	-1.96 (0.10)	1.95e-4 (1.9e-6)	2.03 (0.07)		182
Nicotinamide	94.6	0.40 (0.05)	4.39e-4 (1.2e-5)	n/a		123

Nicotinic acid	94.4	0.39 (0.10)	2.98e-4 (5.1e-5)	4.75 (0.01)		124
p-benzoic acid	102	0.80 (0.10)	2.792e-4 (2.9e-6)	2.39 (0.10)		138
Phenylalanine	140	-1.38 (0.11)	2.22e-4 (5.5e-6)	2.19 (0.08)		166,120 (7:2)
Serotonin	147	-2.18 (0.21)	3.29e-4 (1.8e-5)	n/a		177,160 (2:5)
Threonine	93.2	-2.92 (0.06)	4.49e-4 (5.2e-6)	2.15 (0.11)		120
Tryptamine	142	-1.91 (0.05)	3.57e-4 (1.9e-5)	n/a		161,144 (2:5)
Tryptophan	166	-1.05 (0.05)	2.25e-4 (6.9e-6)	2.30 (0.12)		205,188 (3:2)

5-OH-trp	173	-1.66 (0.06)	2.21e-4 (6.9e-6)	2.36 (0.11)		221,204(5:2)
Tyramine	120	-2.00 (0.09)	3.60e-4 (1.9e-5)	n/a		138,121 (1:2)
Tyrosine, meta	143	-1.55 (0.02)	2.11e-4 (4.4e-6)	2.18 (0.08)		182
Tyrosine, ortho	144	-1.44 (0.05)	2.33e-4 (7.1e-6)	2.37 (0.08)		182
Tyrosine, para	145	-2.30 (0.05)	2.16e-4 (5.9e-6)	2.23 (0.08)		182

Tyrosine, 3-Cl	160	-1.42 (0.09)	1.99e-4 (5.3e-6)	2.17 (0.08)		216,218 (3:1)
Tyrosine, 3-NO ₂	163	-1.12 (0.10)	1.93e-4 (3.3e-6)	2.06 (0.08)		227
Valine	107	-2.30 (0.12)	2.32e-4 (8.8e-6)	2.31 (0.07)		118

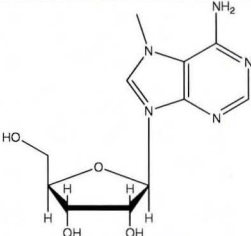
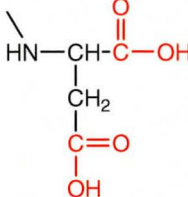
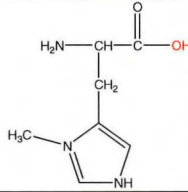
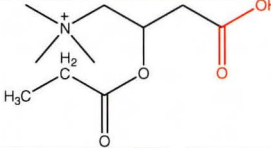
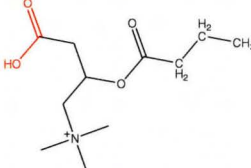
Appendix II Y-matrix values summarizing relative peak area measurements at six different concentration levels used in PLS regression for prediction of RIRs and virtual calibration curves

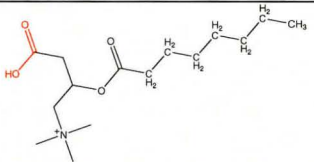
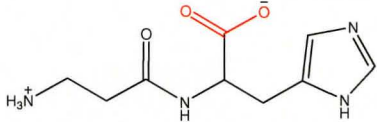
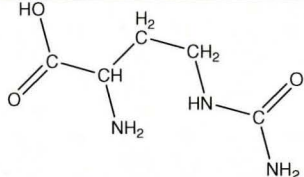
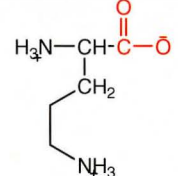
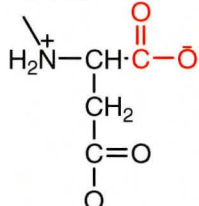
<i>Metabolite (peak mass)[relative peak area ratio]</i>	<i>2μM</i>	<i>8μM</i>	<i>16μM</i>	<i>30μM</i>	<i>50μM</i>	<i>100μM</i>
Adenine (136)	0.0321(0.0041)	0.0781(0.0092)	0.139(0.0040)	0.247(0.019)	0.399(0.078)	0.783(0.072)
Adenosine(268)	0.0767(0.010)	0.253(0.064)	0.488(0.066)	0.899(0.051)	1.487(0.076)	2.955(0.16)
Arginine(175)	0.112(0.016)	0.254(0.019)	0.444(0.064)	0.777(0.065)	1.252(0.11)	2.440(0.15)
Asparagine(175)	0.0208(0.0031)	0.0452(0.0032)	0.0777(0.0064)	0.134(0.013)	0.216(0.023)	0.419(0.036)
Aspartic acid(133)	0.0282(0.0035)	0.0455(0.0051)	0.0684(0.0061)	0.109(0.0090)	0.166(0.012)	0.310(0.029)
Atenolol(267)	0.658(0.039)	1.190(0.14)	1.901(0.345)	3.144(0.29)	4.920(0.18)	9.360(0.38)
Carnitine(162)	0.202(0.014)	0.333(0.034)	0.508(0.039)	0.814(0.11)	1.251(0.078)	2.343(0.072)
Carnitine,Acetyl-(204)	0.307(0.029)	0.641(0.081)	1.088(0.091)	1.869(0.10)	2.985(0.13)	5.775(0.45)
Creatinine(114)	0.0132 (0.0080)	0.0270(0.0036)	0.0453(0.0023)	0.0773(0.0093)	0.123(0.0057)	0.237(0.019)
Cystathionine(223)	0.109(0.0091)	0.217(0.019)	0.360(0.029)	0.611(0.065)	0.970(0.070)	1.867(0.12)
L-Dopa(198)	0.0625(0.0012)	0.181(0.010)	0.339(0.021)	0.616(0.021)	1.011(0.093)	1.999(0.31)
Dopamine(154)	0.0201(0.0040)	0.0431(0.0024)	0.080(0.0031)	0.129(0.0091)	0.209(0.029)	0.401(0.064)
Dopamine(137)	0.0216(0.0019)	0.0455(0.0060)	0.071(0.0021)	0.131(0.010)	0.208(0.012)	0.406(0.051)
Dopamine(total)[1:1]	0.0417(0.0044)	0.0886(0.0064)	0.151(0.0037)	0.260(0.014)	0.417(0.031)	0.807(0.082)
Glutamine(147)	0.0406(0.010)	0.0707(0.0065)	0.116(0.010)	0.183(0.031)	0.283(0.026)	0.530(0.049)
Glutamine(130)	0.0163(0.0020)	0.0283(0.0023)	0.0390(0.0051)	0.071(0.0081)	0.111(0.030)	0.216(0.063)
Glutamine(total)[5:2]	0.0569(0.010)	0.0990(0.0064)	0.155(0.011)	0.254(0.032)	0.394(0.040)	0.746(0.080)
Glutamic acid(148)	0.0429(0.0036)	0.0790(0.0056)	0.129(0.023)	0.214(0.032)	0.339(0.021)	0.645(0.094)
Glutamic acid(130)	0.0093(0.00010)	0.0081(0.0010)	0.028(0.0015)	0.048(0.0051)	0.073(0.0036)	0.141(0.010)
Glutamic acid(total)[9:2]	0.0522(0.0036)	0.0971(0.0057)	0.157(0.023)	0.262(0.032)	0.412(0.021)	0.786(0.095)
Guanine(152)	0.0428(0.0039)	0.0495(0.0046)	0.0584(0.0067)	0.0741(0.0075)	0.0964(0.012)	0.152(0.0020)
Guanosine(284)	0.066(0.0038)	0.141(0.032)	0.259(0.053)	0.452(0.16)	0.735(0.023)	1.440(0.21)
Guanosine(152)	0.056(0.0052)	0.115(0.015)	0.258(0.046)	0.460(0.069)	0.742(0.034)	1.447(0.16)
Guanosine(total)[1:1]	0.122(0.0064)	0.292(0.035)	0.517(0.070)	0.912(0.17)	1.477(0.041)	2.887(0.26)
Homocystine(136)	0.0100(0.00090)	0.0201(0.0036)	0.0335(0.0054)	0.0570(0.014)	0.0906(0.012)	0.174(0.019)
Histidine(156)	0.0550(0.0060)	0.139(0.016)	0.251(0.023)	0.447(0.034)	0.727(0.063)	1.427(0.10)
Histamine(112)	0.00650(0.0010)	0.0132(0.0032)	0.0220(0.0041)	0.0375(0.0034)	0.0597(0.0033)	0.115(0.025)
Leucine(132)	0.084(0.0090)	0.152(0.0029)	0.244(0.011)	0.405(0.025)	0.634(0.012)	1.208(0.090)
Leucine(86)	0.033(0.0020)	0.0610(0.079)	0.098(0.00099)	0.162(0.011)	0.254(0.026)	0.483(0.062)

Leucine(total)[5:2]	0.117(0.0092)	0.213(0.079)	0.342(0.011)	0.567(0.027)	0.888(0.029)	1.691(0.012)
Isoleucine(132)	0.0588(0.0094)	0.126(0.012)	0.216(0.031)	0.372(0.038)	0.596(0.014)	1.155(0.13)
Isoleucine(86)	0.0235(0.0030)	0.0500(0.00091)	0.0860(0.0064)	0.149(0.011)	0.238(0.0099)	0.462(0.035)
Isoleucine(total)[5:2]	0.0823(0.0099)	0.176(0.012)	0.302(0.032)	0.521(0.040)	0.834(0.017)	1.617(0.13)
Lysine(147)	0.0529(0.0031)	0.108(0.015)	0.184(0.020)	0.314(0.014)	0.500(0.037)	0.966(0.076)
Lysine(130)	0.0151(0.00065)	0.0320(0.0031)	0.052(0.0049)	0.089(0.0036)	0.143(0.011)	0.276(0.019)
Lysine(total)[7:2]	0.0680(0.0031)	0.140(0.015)	0.236(0.020)	0.403(0.014)	0.643(0.039)	1.242(0.078)
Methionine(150)	0.0560(0.0063)	0.110(0.010)	0.176(0.02)	0.293(0.035)	0.459(0.039)	0.875(0.010)
Metoprolol(268)	1.566(0.13)	3.744(0.31)	6.648(0.59)	11.730(1.01)	18.990(1.91)	37.140(2.09)
Methionine Sulfone(182)	0.0882(0.0091)	0.182(0.023)	0.307(0.029)	0.525(0.049)	0.837(0.079)	1.617(0.11)
Nicotinamide(123)	0.0283(0.00058)	0.0477(0.0023)	0.0734(0.0091)	0.119(0.015)	0.183(0.014)	0.344(0.0080)
Nicotinic acid(124)	0.00771(0.00036)	0.0227(0.0028)	0.0428(0.012)	0.0778(0.0053)	0.128(0.027)	0.253(0.016)
p-benzoic acid(138)	0.0360(0.00075)	0.0790(0.0041)	0.136(0.024)	0.237(0.029)	0.381(0.025)	0.740(0.057)
Phenylalanine(166)	0.129(0.0091)	0.268(0.019)	0.453(0.043)	0.777(0.081)	1.251(0.096)	2.396(0.22)
Phenylalanine(120)	0.036(0.0040)	0.075(0.0091)	0.128(0.031)	0.221(0.039)	0.342(0.029)	0.685(0.54)
Phenylalanine(total)[7: 2]	0.165(0.0099)	0.343(0.021)	0.581(0.053)	0.998(0.090)	1.593(0.10)	3.081(0.58)
Serotonin(177)	0.0140(0.0019)	0.0350(0.0061)	0.0666(0.0035)	0.113(0.016)	0.185(0.0091)	0.362(0.031)
Serotonin(160)	0.0350(0.0061)	0.0880(0.0091)	0.156(0.012)	0.284(0.034)	0.461(0.039)	0.906(0.078)
Serotonin(total)[2:5]	0.0490(0.0064)	0.124(0.010)	0.223(0.013)	0.397(0.038)	0.646(0.039)	1.268(0.084)
Threonine(120)	0.0247(0.0034)	0.0405(0.0039)	0.0616(0.0060)	0.0986(0.0010)	0.151(0.0092)	0.283(0.019)
Tryptamine(161)	0.0373(0.0030)	0.0960(0.0089)	0.174(0.012)	0.310(0.039)	0.505(0.049)	0.993(0.091)
Tryptamine(144)	0.0747(0.0095)	0.191(0.010)	0.347(0.032)	0.621(0.061)	1.011(0.12)	1.986(0.12)
Tryptamine(total)[1:2]	0.112(0.0096)	0.287(0.013)	0.521(0.034)	0.931(0.072)	1.516(0.13)	2.979(0.15)
Tryptophan(205)	0.142(0.013)	0.270(0.023)	0.440(0.041)	0.739(0.037)	1.165(0.096)	2.229(0.07)
Tryptophan(188)	0.095(0.0063)	0.180(0.0091)	0.294(0.022)	0.492(0.038)	0.776(0.063)	1.487(0.099)
Tryptophan(total)[3:2]	0.237(0.014)	0.450(0.023)	0.734(0.047)	1.231(0.053)	1.941(0.11)	3.716(0.12)
5-OH-trp(221)	0.149(0.024)	0.255(0.032)	0.396(0.049)	0.642(0.032)	0.994(0.087)	1.873(0.098)
5-OH-trp(204)	0.0600(0.0099)	0.102(0.032)	0.158(0.011)	0.257(0.038)	0.397(0.023)	0.749(0.035)
5-OH-trp(total)[5:2]	0.209(0.026)	0.357(0.045)	0.554(0.050)	0.899(0.050)	1.391(0.089)	2.622(0.10)
Tyramine(138)	0.0218(0.0061)	0.0470(0.0050)	0.0810(0.00097)	0.140(0.036)	0.224(0.0019)	0.435(0.041)
Tyramine(121)	0.0437(0.0065)	0.0940(0.010)	0.163(0.023)	0.280(0.026)	0.449(0.039)	0.870(0.065)
Tyramine(total)[1:2]	0.0655(0.0089)	0.141(0.011)	0.243(0.023)	0.420(0.044)	0.673(0.039)	1.305(0.077)
Tyrosine, meta	0.112(0.011)	0.254(0.0021)	0.444(0.045)	0.777(0.081)	1.252(0.096)	2.440(0.19)
Tyrosine, ortho	0.112(0.013)	0.247(0.019)	0.428(0.065)	0.744(0.095)	1.195(0.10)	2.323(0.26)

Tyrosine, para	0.166(0.0090)	0.301(0.029)	0.481(0.036)	0.796(0.063)	1.247(0.093)	2.372(0.24)
Tyrosine, 3-Cl (216)	0.117(0.0099)	0.260(0.019)	0.451(0.0096)	0.785(0.064)	1.263(0.11)	2.456(0.19)
Tyrosine, 3-Cl (218)	0.0390(0.0066)	0.087(0.0094)	0.150(0.010)	0.265(0.034)	1.421(0.10)	0.819(0.079)
Tyrosine,3-Cl (total)[3:1]	0.156(0.012)	0.347(0.021)	0.601(0.014)	1.047(0.72)	1.684(0.15)	3.275(0.21)
Tyrosine, 3-NO ₂ (227)	0.144(0.054)	0.317(0.029)	0.549(0.049)	0.953(0.083)	1.532(0.11)	2.977(0.28)
Valine(118)	0.00886(0.00010)	0.0201(0.0061)	0.0352(0.0049)	0.0615(0.0071)	0.0992(0.0084)	0.193(0.013)

Appendix III 10 different validation metabolites where all parameters were predicted *in silico* by computer modeling

<i>Metabolite</i>	<i>Structure</i>	<i>Physicochemical properties</i>	<i>Virtual Calibration Curve (Equation of Line)</i>
1-Me-adenosine		MV: 228 Å ³ LogD: -1.82 μ_0 : 2.48e-4 cm ² /Vs Z_{eff} (pH1.8): +1.00	$y=0.055x+0.223$
N-Methyl-aspartic acid		MV: 114 Å ³ LogD: -2.58 μ_0 : 3.48e-4 cm ² /Vs Z_{eff} (pH1.8): +0.71	$y=0.006x+0.0219$
3-Me-histidine		MV: 144 Å ³ LogD: -2.94 μ_0 : 6.20e-4 cm ² /Vs Z_{eff} (pH1.8): +1.49	$y=0.019x+0.071$
Carnitine, Propyl-		MV: 214 Å ³ LogD: -1.92 μ_0 : 2.56e-4 cm ² /Vs Z_{eff} (pH1.8): +1.00	$y=0.049x+0.019$
Carnitine, Butryl-		MV: 234 Å ³ LogD: -1.73 μ_0 : 2.46e-4 cm ² /Vs Z_{eff} (pH1.8): +1.00	$y=0.0576x+0.223$

Carnitine, Octanoyl-		MV: 313 Å ³ LogD: -0.64 μ_o : 2.14e-4 cm ² /Vs Z _{eff} (pH1.8): +1.00	y=0.091x+0.374
Carnosine		MV: 182 Å ³ LogD: -5.94 μ_o : 5.50e-4 cm ² /Vs Z _{eff} (pH1.8): +1.92	y=0.055x+0.223
Citrulline		MV: 140 Å ³ LogD: -3.19 μ_o : 3.14e-4 cm ² /Vs Z _{eff} (pH1.8): +0.83	y=0.016x+0.062
Ornithine		MV: 113 Å ³ LogD: -4.33 μ_o : 6.86e-4 cm ² /Vs Z _{eff} (pH1.8): +1.82	y=0.005x+0.011
Theanine		MV: 150 Å ³ LogD: -2.56 μ_o : 3.04e-4 cm ² /Vs Z _{eff} (pH1.8): +0.73	y=0.020x+0.082